Reviewer #1 (https://doi.org/10.5194/egusphere-2025-6364-RC1)

Initial Comments

I recommend publication after revision, as detailed below

Scientific Significance – Good (2)

This paper addresses an important issue, the need to provide a sound scientific basis to evaluate the performance of the Sentinel-3 Next Generation Topography Mission, in terms of the cross-calibration with the fore-running Sentinel-3 mission during a tandem phase mission. The approach described is new, developed from previous approaches applied to previous satellite altimeter tandem phase validation studies.

Thus the prime objective of the study is to support understanding of the impact of a proposed orbit selection on measurements from the S3 and S3NG-T missions, rather than to present new scientific results on the state and behaviour of the ocean. I therefore agree the paper is most appropriately included as a technical note.

Scientific Quality – Fair (3)

The scientific approach and applied methods are valid and developed from relevant previous work. Appropriate references are provided throughout.

The discussions in Sections 4 (Results) and 5 (Assessment) do not provide sufficient detail to enable a good understanding on how figures have been generated, what can be seen in these figures, and to justify the conclusions that have been drawn.

Presentation Quality – Good (2)

The text is well written and structured, although as mentioned above, more detail is needed in the description and discussion of the figures.

Thus I would recommend some revisions primarily to sections 4 and 5 to provide more detailed presentation and discussions of the results, to provide a more solid justification of the main findings.

We completely agree that more detailed discussions were needed to solidify the justification of our main findings. We have significantly expanded the physical interpretations of our results in both Sections 4 and 5. For example, we now

provide explanations of the spatial patterns observed in the regional SLA offsets. Furthermore, we moved the detailed explanation of the extrapolation model (and its associated figure) out of the Conclusions and directly into Section 5.1. We also added comprehensive discussions explaining exactly what conclusions can be drawn when comparing our two uncertainty estimation approaches.

Two of the figures are listed as A1 and A2, and the contents of these figures provide useful and important information on the geographical variability in the processed and gridded data. I cannot find any guidance regarding constraints on the number of figures within a paper. As there are only 4 other figures, I would recommend including these figures in the main body of the paper.

We agree that the geographical variability maps provide crucial information that belongs in the main text. We have followed your recommendation and moved these figures out of the appendix. To further improve the reader's understanding of the high-frequency variability observed by SWOT, we also paired figure A2 with a complementary new figure showing the temporal evolution of this variance.

The labelling of axes on the figures and the content of the captions should be improved to allow readers to understand what is presented. For all Figures 1-4 the y axes are labelled “Uncertainty”, and the captions do not explain what the uncertainty is in.

We have corrected this across all relevant figures (Figures 1-4); all y-axes have been strictly renamed to "Regional SLA offset uncertainty [mm]". Furthermore, we have updated the figure captions and the accompanying text to explicitly state both the physical parameter being measured and the specific methodology used to compute the plotted values.

Detailed Comments

Abstract

The abstract is well written and accurately summarises the contents of the paper.

10 – “...uncertainties of approximately 2mm” – in what?

We have corrected the text to clarify that this 2 mm precision refers specifically to the uncertainty of the regional Sea Level Anomaly (SLA) offsets.

13 -Extending tandem phase to one year enables the detection of errors of ±3.5mm amplitude – see later comments around Figure 4 (Line 274 onwards).

This comment has been detailed below the comment on Figure 4, and we have added a paragraph at the end of Section 5.1 (L.243) to describe the analysis.

Introduction

The background and context for the study are in general well described.

41 – I understand that during the tandem phases the interval between successive satellites for the Jason series reference missions was around 60 seconds and for Sentinel 3A/3B it was 30 seconds. However that for ERS-2 and ENVISAT it was longer - 30 minutes. The text gives the impression the delay was under one minute for all tandem phases.

Please rewrite this section to more accurately describe the tandem missions referred to.

The tandem phases have indeed different time intervals between successive satellites:

- **ERS-2 & ENVISAT: 30 minutes**
- **TOPEX/Poseidon & Jason-1: 70 seconds**
- **Jason-1 & Jason-2: 55 seconds**
- **Jason-2 & Jason-3: 80 seconds**
- **Jason-3 & Sentinel-6 MF: 30 seconds**
- **Sentinel-3A & Sentinel-3B: 30 seconds**

The manuscript has been corrected consequently (L.37-41) with track change:

"During a tandem phase, two successive missions follow an identical ground track separated by a strictly controlled time interval. Tandem phases have been systematically implemented following the launch of new reference altimetry missions, including TOPEX-Poseidon and Jason-1 (2002; 70-second gap), Jason-1 and Jason-2 (2008; 55-second gap), Jason-2 and Jason-3 (2016; 80-second gap), and Jason-3 and Sentinel-6 Michael Freilich (2021-2022; 30-second gap). They have also been used between non-reference missions, such as ERS-2 / Envisat (2007-2008, 2010-2011; ~30-minute gap) and Sentinel-3A and Sentinel-3B (2018; 30-second gap)."

44- LEULIETTE et al - correct link and reference to correct capitalisation.

Corrected

The approach only considers using recorded satellite altimeter data sets. Was using ocean model hindcast data, sampled along the exact ground tracks, considered? It would be useful to include a short discussion on the potential to use model data for these types of study, and why it was not considered appropriate for this study.

We chose not to use traditional global Ocean General Circulation Models (OGCMs) or operational model hindcasts because they struggle to faithfully represent true high-frequency ocean variability. The variability introduced over a short 4-hour delay is heavily dominated by high-frequency, sub-daily ocean dynamics, specifically internal gravity waves, internal tides, and sub-mesoscale turbulence. While state-of-the-art global ocean models successfully reproduce large-scale mesoscale circulation, accurately resolving the exact global phase and amplitude of these specific high-frequency signals remains a significant modeling challenge.

New ultra-high-resolution experimental models, such as the MITgcm LLC4320 global simulation, are now capable of capturing internal tides and high-frequency submesoscale structures. However, these remain advanced research tools rather than standard operational global hindcasts. Therefore, relying on actual recorded satellite altimeter data sets was considered the most appropriate and robust approach to capture the true physical noise floor for this study. SWOT resolves kilometer-scale SLA structures across a 120-km swath, enabling a fair representation of the submesoscale structure.

We have added a dedicated paragraph summarizing these modeling limitations and justifying our empirical approach to the revised manuscript.

Data

83 Suggest to note here that S3A and S3B were only 30 s apart on the same ground track.

Corrected, a sentence has been added L.87 and 91 for S3A/S3B and J3/S6A tandem phases.

"The S3A/S3B tandem phase lasted from 7 June to 16 October 2018, spanning four complete 27-day cycles or ten 12-day sub-cycles, with a 30-second time interval between the satellites (Clerc et al., 2020)."

100 – Link to Jason CS Sentinel 6 product handbook

Corrected

Methods

3.1 Comparison between two missions

I find the description in this subsection confusing. Has the same gridding process been applied to S3A and S3B data and also to J3 and S6A data, covering periods in and out of the tandem phase, so that there are effectively four data sets? A table listing the gridded data sets as generated might help to clarify the situation.

Yes, the exact same gridding process (a 3°x3° spatial resolution and a 12-day temporal window) is systematically applied to all individual missions and all pairs, across all time periods (in, out, and simulated 4-hour tandem phases). To clarify this in the manuscript, we have added the suggested summary table in the appendix, detailing the individual SLA grids and the paired ΔSLA grids generated for each scenario.

	SLA grid [3° x 3°, 12 days] (Individual missions)	ΔSLA grid [3° x 3°, 12 days] (Mission pairs)
Tandem phase	S3A, S3B, J3, S6A	S3A/S3B and J3/S6A
Outside tandem phase	S3A, S3B, J3, S6A	S3A/S3B and J3/S6A
4-hour tandem phase	S3A, S3B, J3, S6A	S3A/S3B and J3/S6A

127 “For each cycle” - Does this refer to the S3A, S3B 27 day cycle, or the earlier defined 12 day “sub-cycle” (itself made up of 3 4-day subcycles)

'For each cycle' referred to the 12-day sub-cycle, but we agree the original formulation was confusing. This section has been clarified in the revised manuscript (L123-126):

“To ensure a consistent comparison between missions with differing orbital characteristics (e.g., the 10-day cycle of J3/S6A versus the 27-day cycle of S3A/S3B), we adopted a fixed 12-day temporal window for all configurations, corresponding to three 4-day sub-cycles of the S3 orbit. These gridded fields are computed independently for each 12-day period.”

If these data are used to generate Figure A1 provide an explanation and a discussion. (Note the recommendation to include this figure in the main body of the paper.)

We have moved the former Figure A1 to the main body of the paper (now Figure 3 in the Results section) and added a detailed explanation of the data and physical patterns observed.

To clarify the methodology:

- **Panel (a) is generated using the time-averaged ΔSLA grids from the S3A/S3B classical tandem phase (~30s time interval).**
- **Panel (b) is generated by adding the 4-hour oceanic variability, derived from SWOT and S3 dual-crossovers, to the S3A/S3B ΔSLA data, simulating the '4-hour tandem phase' scenario.**

We have added the following discussion regarding the interpretation of these patterns to the Results section (Lines 201-210):

"The classical tandem phase enables highly effective detection of systematic differences mainly due to instrumental errors by minimising the effect of differences in oceanic variability. These spatially correlated systematic differences are clearly visible on the regional SLA offset map in panel (a) of Fig. 2. Notably, a distinctive positive offset (up to 7 mm) dominates the high southern latitudes (south of 30°S). This can likely be attributed to differences in how the two altimeters' retracking algorithms process Sea State Bias (SSB), as the Southern Ocean is characterized by high Significant Wave Heights (SWH) and long ocean waves. These zonal patterns are significantly harder to detect in panel (b), demonstrating how the oceanic variability differences introduced by a 4-hour delay masks underlying instrumental errors."

3.2 Uncertainty Computation

Two approaches are given for the calculation of the uncertainty in the estimates of regional offsets of sea surface height anomaly. Results from the two approaches are compared in Figure 3, but it is not clear which approach was used to generate the values presented in Figures 1 and 2. Please clarify.

We have clarified in the text and the figure captions that the uncertainties presented in Figures 1 and 2 were calculated using the local temporal method.

Corrected, a sentence has been added L.195-196:

"The uncertainty presented in these results is computed using the temporal method, which is based on analysing the temporal variability of the ΔSLA grids within each grid cell, as described in Section 3.2.2."

Additionally, the captions for Figures 1 and 2 have been updated to explicitly state the methodology used.

3.3 Accounting for 4-hour variability

185 Figure A2 – is the first panel for zero time offset? – Please provide a more detailed discussion on what is in the figure, and of the important features in the figure..

To clarify the time intervals in the original appendix figures: Panel (a) of the original Figure A1 represented the classical tandem phase (which has a ~30-second time interval), whereas the original Figure A2 specifically represented the oceanic variability for a 4-hour interval.

We have expanded our analysis of the features within the 4-hour variability map in the main text. Furthermore, to improve the reader's understanding of the high-frequency oceanic variability observed by SWOT and to assess the limitations of our approach, we have added a complementary new figure showing the temporal evolution of this variance.

These two elements are now combined and discussed in detail in the Section 3.3 of the revised manuscript.

193 -remind readers that "Classic tandem phase" implies near zero delay (less than one minute).

Corrected, merged with next comment.

Results

This is a key section, as it presents the results of the estimated uncertainty in regional SSHA for the 4 hour delay tandem phase. As identified in the initial remarks, more detailed discussion is required in this section to provide the reader with a better understanding of these key findings

We have improved this section by including the former Figure A1 to the main body of the paper (now Figure 3 in the Results section) and added a detailed explanation of the data and physical patterns observed.

194 How exactly are the values for the "non-tandem" scenario calculated – from data outside tandem phase.. (duh).

We have clarified the text to explicitly state both the near zero delay of the classical tandem phase, and the data used to calculate the non-tandem scenario.

We have updated the introduction of Section 4 (L.200) to read:

"In this section, we compare the continuity performance of three configurations: the classical S3A/S3B tandem phase with a 30-second time interval, a 4-hour delayed tandem configuration, and a non-tandem scenario exploiting S3A and S3B data acquired after their tandem phase."

195 Provide a more detailed discussion of Figure 1.

Which version of Regional Inter Mission Bias Uncertainty is being plotted?

Corrected, a sentence has been added L.203-205 to explicitly state the methodology used:

"The uncertainty presented in these results is computed using the temporal method, which is based on analysing the temporal variability of the ΔSLA within each grid cell, as described in Section 3.2.2."

Figure 1 (and 2, 3 4)– Include parameter for uncertainty (on y-axis) in caption – i.e. "Regional intermission bias uncertainty in SSHA"

Corrected, all the axes have been renamed as "Regional SLA offset uncertainty [mm]"

Assessment

Comparison with J3/S6A missions

Again – provide a more detailed description and discussion of Figure 2. What measure of uncertainty has been calculated (3.2.1 or 3.2.2)?

Corrected, a sentence has been added L.232:

"The uncertainty presented in this comparison is computed using the temporal method described in Section 3.2.2."

Additionally, the caption for Figure 2 has been updated to explicitly state the methodology used.

Comparison with Cross-Over Based Approach.

231 - The text does not describe how uncertainty values are derived from the SLA difference values and the regional offset.

We have updated the text to clarify that the uncertainty is derived using the exact same local temporal method applied to the along-track configurations. A sentence has been added L.262:

"The differences are spatially averaged into 3°x3° grid cells, resulting in gridded SLA difference fields analogous to those derived from along-track comparisons. The regional SLA offset is then estimated by averaging the time series of each cell across multiple cycles. The associated uncertainty is calculated using the temporal method (described in Section 3.2.2), ensuring a homogeneous calculation with the other configurations presented in Fig. 4."

Comparison of two uncertainty estimation methods

258 Again provide a more detailed discussion of Figure 3.

What conclusions can be drawn from Figure 3? Which calculation approach is better?

We have expanded the text to directly answer both of your questions. To summarize our conclusions: Figure 3 (now Figure 6) demonstrates that both calculation approaches yield highly consistent uncertainty estimates as the number of observation cycles increases. This strong agreement acts as a cross-validation, proving the robustness of our overall uncertainty estimation framework.

Regarding which approach is "better": we now explicitly state in the text that we recommend the temporal method (Method 2) whenever sufficient data cycles are available. This is because the temporal approach is mathematically consistent with the global mean uncertainty assessment, accounts for temporal correlation and provides geographically resolved insights. However, the spatial method remains a valuable, robust alternative for very short datasets where temporal statistics cannot be reliably computed. We have restructured the corresponding paragraphs to make these conclusions and recommendations immediately clear to the reader.

Conclusions

261 – "Traditional tandem phase*(s)*, where missions fly in close formation, are crucial for mission continuity." *- add the (s)*

Corrected

273 ...through *(an)* extended calibration period. *Add the (an)*

Corrected

276 – Insufficient description in the text on how the curves in figure 4 were calculated, and also insufficient discussion of the details in Figure 4. How were the results from the J3/S6a comparison used to fit a curve to the S3A S3B results?

What are the 1, 2, 3 year thresholds?

Without this information it is not possible to assess whether the key findings of the paper are justified :

"A 4 hour tandem phase would require approximately two years of continuous observations to reach a similar level of calibration precision (Fig. 4).

"However, the demonstrated ability to detect systematic differences of ±3.5 mm within one year highlights the feasibility of this approach....)"

We agree that detailing the extrapolation method is crucial for justifying the paper's key findings. To improve the logical flow of the manuscript, we have moved Figure 4 (now Figure 5) and the detailed explanation out of the Conclusions

and into Section 5.1 (Comparison with J3/S6A missions), where the J3/S6A baseline dataset is first introduced.

To answer the specific questions:

- **Statistically, the uncertainty of a temporal mean decreases proportionally to $1/\sqrt{n}$, where n is the number of observation cycles. Both the S3A/S3B and J3/S6A classical tandem phases shared an identical operational configuration with a 30-second time delay, and their initial uncertainty results show excellent consistency. Because the S3A/S3B tandem phase was relatively short (~120 days), we used the longer J3/S6A tandem (~200 days) to achieve a more robust fit.**
- **The horizontal lines are simply visual markers for 1 year, 2 years, and 3 years so the reader can easily read the logarithmic y-axis.**

We have added a paragraph at the end of Section 5.1 (L.243) to describe the analysis.

References

Caps on Surnames of authors and co-authors. Leulliette et al (2004)

Zhao Z (2024) – non standard characters (${\mathrm{M}}_{2}$)

Corrected