Response to reviewer #2

We thank the reviewer for their efforts in reviewing our manuscript. Reviewer comments are shown below in black with the author response in red.

Suggestions for revision or reasons for rejection

The authors have primarily modified the descriptions in the manuscript, but unfortunately, these modifications are not substantial. This manuscript cites Benkiran et al. (2024; https://egusphere.copernicus.org/preprints/2024/egusphere-2024-420/), who conducted almost the same experiments and obtained similar results when comparing the impacts of assimilating SSH observations from 12 nadir satellites versus 2 swath satellites. However, their study was rejected due to a lack of novelty and unreasonable results. The same issues apply to this manuscript, as detailed below.

Benkiran et al. (2022; https://os.copernicus.org/articles/18/609/2022/) conducted experiments investigating the impacts of 2 swath satellites. This manuscript and Benkiran et al. (2024) are almost identical to Benkiran et al. (2022) except for the addition of an experiment assimilating 12 nadir satellites.

This paper differs from Benkiran et al. (2022) due to the inclusion of a comparison with the 12-nadir constellation, and the different data assimilation system used. It addresses a different question to Benkiran (2022) since it is asking the question which of two specific future constellations, proposed by a space agency, would give most impact.

The Benkiran et al. (2024) paper did indeed perform **similar** experiments to those presented here, but in fact found opposite results, i.e., the 12 nadir constellation was superior in the Met Office system, while the 2 WiSA constellation was superior in the Mercator system. As described in the introduction, these experiments were performed with different operational systems with (amongst other differences) very different data assimilation schemes. As we also described in the introduction, results from OSSEs can be very system-dependent and the aim of coordinating similar experiments with different operational systems was to allow comparison and a better understanding of the results.

The need for results from multiple systems is acknowledged by the community (see Oke & O'Kane 2011 and Fujii et al. 2019) as important in order to provide a more robust assessment of impact for future observing system design. The introduction has been updated to emphasise and clarify these points.

As mentioned in previous comments, OSSEs are useful for evaluating various observation networks that have not yet been constructed. To establish the novelty of this paper, a wide variety of experiments are needed to comprehensively investigate how the number and types of satellites contribute to analysis accuracy. However, the authors did not incorporate the comments from the first review round.

We addressed a similar comment from the reviewer in the initial review, but perhaps did not make clear enough the large computational resource required to perform these experiments. Furthermore, we addressed each of the comments raised in the initial review and updated the manuscript to include further detail of the experimental design, and further analysis of the results.

We agree that OSSEs are useful to evaluate proposed observing systems. However, due to the computational expense it was not possible for us to perform a "wide variety of experiments". Our aim here was to investigate the potential impact of two specific proposed observing networks to inform the planning of ESA as they explored options for the Sentinel-3 Next-Generation Topography mission.

The novelty here was to determine how effectively our system would be able to assimilate observations from these specific proposed networks. This allowed us to identify issues that will affect the assimilation of real wide-swath altimeter observations and so prepare to make best use of real observations as they become available.

The introduction has been updated to further clarify our aim of addressing s specific network design question and to emphasise the limitations on the number of experiments that could be run due to their computational expense.

The OSSEs showed that SSH RMSEs are 0.05-0.07 m, which is about 5-10 times larger than the prescribed SSH observation errors of 0.014 and 0.005 m for nadir and swath satellites, respectively. Since observations from 12 nadir satellites cover almost the entire global ocean, at least the RMSEs should be smaller than the observation errors. Therefore, these results are inconsistent with established data assimilation theory.

We have updated the text to clarify that the errors added to the simulated altimeter observations were intended to replicate the instrumental errors. Additionally, representation errors are introduced by the differences between the nature run and OSSE model fields. As described in section 2.3, we used the same prescribed observation errors as in our operational system which include representation errors. These vary spatially and temporally with values of 3-7cm globally.

In Section 2.4 we also describe a comparison of RMSE values for each observation type between the OSSE control and our operational system to illustrate that the results we see are similar to what we might expect to see with real observations in our operational system.

Additionally, there are still many inappropriate descriptions and terminologies, as well as insufficient details in the experimental settings and results, because the authors did not address most of the previous comments.

We addressed each of the comments raised in the initial review and updated the manuscript to include further detail of the experimental design, and further analysis of the results. Without further detail, we cannot comment here on what the reviewer still finds to be inappropriate.