[General comments]

This paper investigates the impacts of two swath and 12 nadir altimeters on the analysis and forecast accuracies conducting the OSSEs. The topic appears to be interesting, however, there are critical issues described below.

Firstly, the contents are largely similar to those of a previous paper by Benkiran et al. (2022), but with only the addition of the results assimilating 12 nadir altimeters. To establish the novelty of this paper, I strongly suggest conducting a wider variety of experiments to comprehensively investigate how many nadir altimeters are beyond 1- and 2-swath altimeters and cannot substantially increase the accuracy.

Secondly, in the OSSE, data assimilation experiments should not employ true values from the Nature Run. However, this study uses true values for the initial conditions of assimilation experiments, which is inconsistent with the OSSE protocols. Therefore, it is necessary to modify the setting of the assimilation experiments.

Thirdly, the analysis SSH RMSEs in all assimilation experiments are larger than the prescribed observation errors. According to the data assimilation theory, analysis RMSEs should be smaller than both background and observation errors. Consequently, the results appear to be inconsistent with established data assimilation theory.

Fourthly, this study aims to compare the impacts of 2 swath and 12 nadir altimeters. However, most spatial maps mainly compare the accuracy of the 2 swath and 12 nadir altimeters with 3 nadir altimeters, which is not aligned with the stated motivation.

Fifthly, this study does not conduct any statistical tests to compare the accuracy of the assimilation experiments.

Finally, this paper lacks the necessary information, particularly regarding model configuration and setup of data assimilation experiments. This deficiency is likely to be attributed to the similarities of the previous authors' work. In addition, numerous colloquial and unclear descriptions as well as typos are found. Therefore, I highly recommend using an English editing service.

Due to these reasons, significant modifications are required before publication. I recommend "major revision" or "reject with an option for resubmission" if the authors require additional time to revise the manuscript.

[Specific comments]

#1) “performance”, “quality”, and “error”: The first two terms are used in a variety of contexts, such as “performance” for computation and “quality” for water quality and statistics. “error” is defined as an instantaneous difference between state variable and true value (e.g., forecast and analysis error: $\boldsymbol{\epsilon^f = x^f - x^t}$ and $\boldsymbol{\epsilon^a = x^a - x^t}$ respectively) and it does not indicate statistic expectations such as RMSE. Therefore, to ensure precision in communication and avoid readers’ misunderstanding, it would be more appropriate to use the statistic terminologies “accuracy” and “RMSE” throughout the manuscript instead of these three words.

#2) Line 29 (L29) in Page 1 (P1): Please clarify the meaning of “blue/white/green ocean states”.

#3) The short introduction is not inadequate, but it lacks the depth needed to fully motivate the comparison of the impacts between the constellation of 2 wide-swath and 12 nadir altimeters. In particular, this paper closely resembles the previous work by the authors (Benkiran et al. 2022), except for adding an experiment assimilating 12 nadir altimeters. While the OSSE allows for the evaluation of observation networks in virtual space by changing the number and type of satellites before their real-world establishment, this paper lacks clear differences in terms of novelty from the previous paper.

To enhance the novelty of this paper, additional assimilation experiments would be beneficial to understand how many nadir altimeters can surpass the accuracy achieved by 1 or 2 wide-swath altimeters. Furthermore, considering discussions on the costs associated with the development and launch of these satellites would greatly strengthen the paper. By incorporating such discussions, this paper could offer valuable insights and guidance for future satellite missions.

#4) L15 in P2: Please clarify the meaning of “Phase A study”.

#5) The authors frequently rely on references to previous works (e.g., Benkiran et al. 2021, 2022) without providing sufficient details in this paper. It is crucial to include essential information within the context of this study to avoid readers having to refer to multiple papers. Specifically, details on the WiSA concept should be thoroughly described in Section 2, while the model configuration and assimilation settings should be specified in Section 3.

Please also consider integrating Section 2 into Section 1 to provide a smoother flow of information.

In Section 3, please provide detailed information on the model configuration, including vertical resolution, initial conditions, spin-up period, boundary forcing, and treatment of sea ice if the system is global, differences in configurations and forcings between NEMO versions 3.1 and 3.6, and the settings of the data assimilation experiments.

#6) L21 in P2: Please specify the “WISA #A orbit”.

#7) L31in P2: Please remove “free” because the free run refers to a simulation independent of the Nature Run. It would be better to add descriptions of the free run in the first paragraph of subsection 3.1 rather than subsection 3.4 to clarify the differences between the free and Nature runs.

#8) The first paragraph in subsection 3.2: Please add the information on observation errors for each variable.

#9) The second paragraph in subsection 3.2: It is essential to show the SSH observation errors of swath altimeters. If the errors undergo spatiotemporal variations, please present the spatial pattern and average over the global ocean.

#10) L14 and others: Please remove the space between “Figure 1” and “A” and use a consistent format to indicate figures (e.g. Figure 1a) rather than ambiguous descriptions such as “right figure”.

#11) The last paragraph in subsection 3.2: Please indicate the coverage ratios of each assimilated observation.

#12) L20 in P3: Please add the information on what data assimilation scheme the SAM2 is based on. If the variational method is used in this study, please specify the details of the prescribed background errors.

#13) L22 in P3: Please specify “adaptivity scheme” and “observation residuals”.

#14) L31 in P3: Please specify “uncontrolled temporal frequencies”.

#15) L34 in P3: We do not have access to true values in real world and cannot initiate simulations from initial conditions based on true values. In the OSSE, both the free run and data assimilation experiments cannot employ true values from the Nature Run. However, this study uses true values from the Nature Run as the initial conditions for all assimilation experiments. Therefore, the procedure in this study deviates from the typical OSSE setup.

#16) Subsection 3.4: Please specify the reasons why only the Sentinel-6 is assimilated for all data assimilation experiments.

#17) L9-L18 in P4: The detailed information on validation methods should not be included in the result section. Please describe it by inserting the method section between Sections 3 and 4.

#18) L15–17 in P4: Please clarify the reasons why the cutoff scale is defined as 200 and 500 km in this study. The scale of ocean eddies is about 100 km. If the authors intend to separate the meso- and large-scale, the inclusion of the latter one (i.e., 500 km) might not be necessary. Please specify what kind of filters are applied to what values (state vector, error, and RMSE).

#19) L24-25 in P4: Please specify how the Nature Run and free run reproduce the SSH variance compared with the observations, although the authors might have described the detail in the previous paper. This information is important to demonstrate the reproducibility of the different versions of NEMO.

#20) L27 in P4 and others: “temporal evolution” indicates time differentiation, d/dt, and is not appropriate to use here and in similar descriptions. “time series” would be a better expression to be consistent with the authors’ intent.

#21) Section 4: To validate the results, this study mainly used two statistics: RMSEs and variance

errors. For consistency with the dimension of the prescribed background and observation errors (cf. Farchi and Bocquet 2018), it would be more appropriate to use RMSEs rather than variance errors.

#22) L31 in P4 and others: Please summarize the names of data assimilation experiments in subsection 3.4.

#23) Section 4: As described in subsection 3.2, the prescribed observation errors of nadir altimeters are 2 cm (i.e., 4 cm^2 in variance), which are smaller than the best analysis variance errors of 9.6 cm^2. Therefore, the results of the assimilation experiments are inconsistent with the data assimilation theory. Please specify the reasons why the variance errors are larger than the prescribed observation errors.

#24) Section 4: Although the authors showed the time series of variance errors, it is necessary to apply statistical tests, such as paired samples t-test, to detect the significant RMSE differences among all experiments, especially between the 2 swath and 12 Nadir experiments.

#25) Section 4: Most of the spatial pattern figures showed the statistics differences between 3 Nadirs and 12 Nadirs and between 3 Nadirs and 2 Swath, with the reference defined as 3 Nadirs. However, the main motivation of this paper is to compare the impacts of 2 swath and 12 nadir altimeters. Therefore it is cruicial to show the RMSE differences between 2 swath and 12 nadir altimeters. To detect significant differences between 2 swaths and 12 nadir altimeters, it would be necessary to apply the statistical test to each grid cell.

#26) L42 in P4: It would be better to calculate areas than points. Please calculate degradation areas as well.

#27) L24 in P5: Please specify “a variance preserving form”.

#28) Subsection 4.2: It would be expected that 2 swath altimeters would have larger impacts, especially on mesoscale than 12 Nadir altimeters. However, the accuracy differences are substantial at spatial and time scales longer than 200 km and 10 days, respectively (Figs. 8 and 10). It is essential

to understand the role of altimeters by investigating the causes of why the impacts are substantial at the longer scale only.

#29) L30-32 in P5: Please specify the method used to calculate time spectral coherence (i.e., correlation coefficient). Is it a spatially averaged correlation coefficient? Furthermore, please clearly define “effective temporal resolution”.

#30) The description style is more colloquial throughout the manuscript, especially after the fourth paragraph in subsection 4.2. For example, the sentence “On the right-hand figure, we have the difference between 12 Nadirs and 3 Nadirs” in L35-36 in P5 might be understood by the readers. However, for scientific journals, this colloquial and subjective style is not appropriate, and it is necessary to make objective descriptions. Therefore, I highly recommend using an English editing service.

#31) The fifth paragraph in subsection 4.2: The Kuroshio and Kuroshio Extension regions would have almost the same order of dynamic energy (i.e., the sum of mean and eddy kinetic energy) as the Gulf Stream region. Since the energy spectra would exclude the mean kinetic energy, the description of the “less energetic Kuroshio region” in L47 in P5 is not reasonable.

To clarify the different impacts of 2 swath and 12 nadir altimeters, it is essential to investigate what phenomena are reproduced with different accuracies.

#32) L2 in P6 and others: “mean of … error” is a bias? Please use appropriate terminologies.

#33) L3 in P6: “error” is not “innovation”.

#34) The second paragraph in subsection 4.3, subsection 4.4: The detailed descriptions of validation of the velocities in the second paragraph in subsection 4.3 and particle tracking in subsection 4.4 would not be necessary because the results are largely consistent with the SSH.

#35) L18 in P6: “significant” can be used only when statistical tests are performed.

#36) L24-28 in P7: The meaning of the descriptions is unclear. Please carefully revise and clarify the descriptions.

#37) Section 5: The results from the OSSE in virtual may not necessarily align with the OSE in real. It would be important to inform readers that the validation results might not be consistent between OSSE and OSE due to various factors such as model biases.

#38) The OSSE indicates a method used to evaluate the impacts of observation networks and does not involve data assimilation experiments. Consequently, “NatRun-OSSEs” in Fig. 3 and similar expressions may not appropriate.

#39) Figure 8: The correlation coefficient between free run and nature run appears to be completely zero. Even in the chaotic Gulf Stream regions, the coefficient would not be zero. Please carefully review the calculation.

#40) There is no label on y-axis in Fig. 8a, y-axis of Fig. 10a, and color scales of Fig. 11.

[Technical corrections]

Line 1 (L1) in Page 1 (P1): “for” in the Title might be “on”.

L9 in P1: “Surface Water Ocean Topography” might be “Surface Water Ocean Topography (SWOT) mission”.

L15 in P1: “has” should be “have”. Hereafter, we use right arrows → for similar replacement.

L18 in P1: Insert “The results showed that” before “These two configurations” to clarify that the results performed by the authors are shown hereafter.

L35 in P1: “plays” → “play”.

L15 in P2: Please spell out “CNES”.

L31-32 in P2: “to represent … synthetic observations” might be “to generate true values and synthetic observations”. It would be better to use “generate observations” rather than “simulate observations” in the OSSE.

L39 in P2: Remove “same”.

L40-43 in P2: Please specify satellite and in-situ observations for each variable.

L44 in P2: “simulated” should be “distributed”. Remove “along” before “the swath”.

L45 in P2: Please clarify “(S1 and S2)”.

L1 in P3: “(S1, S2)” → “(S1 and S2)”.

L6 in P3: “but” should be inserted before “separated”.

L28 in P3: “injected” → “inserted”.

L39 in P3-L2 in P4 and others: The caption of tables should be located at the top of the tables.

L40 in P3: “Column” → “column”

L41 in P3: “Column 2” → “the second column”, and “standard” → “assimilated”.

L15 in P4: “of high variability”→ “with large variations”.

L18 in P4: “depth” → “ocean interior”

L18 in P4: Please clarify “system mass”.

L31 in P4: “smaller” should be inserted before “SSH error”.

L6-7 in P5: Please clarify what experiments are compared.

L21-22 in P5: Please check the descriptions.

L28 in P5: Which scale “onwards” indicates, longer or shorter?

L32 in P5: “wavelengths” → “timescale”

L27-28 in P6: Please specify which way the particle tracking is conducted, offline or online.

L3 in P10: “Truth Run” → “ Nature Run”

Figure 2: The color scale should be modified because there are no minus values. Units with a color scale “Cm^2” should be “cm^2”, and this is the same for others.

Figure 3: The color of forecast error variance is difficult to see. Please modify the line type and color.

L9 in P12: The descriptions of the fourth and fifth columns are not consistent with the Table 2.

Label of Fig. 7: Please clarify which RMSE is shown in Fig. 7, forecast or analysis RMSEs. Please clarify which data assimilation experiments are subtracted from which experiments.

Figure 8b: The unit in the label of x-axis “day” should be “day^{-1}”.

L8 in P14: “cm2” → “cm^2”

Labels of Figs. 9 and 10: These are almost same as Figs. 7 and 8. Please use the expression “As in … but for …”.

===Reference

Farchi and Bocquet (2018), **25**, 4, 765-807, Nonlinear Processes in Geophysics