Anonymous Referee #2 - Comment on essd-2022-438

We would like to express our gratitude to Referee #2 for meaningful comments and suggestion. We have carefully reviewed your comments and have made necessary updates to our manuscript. We provide point-to-point response to the referee comments shown in blue whereas the revision made to the main text in shown *blue Italics*.

Kind regards

This study presents a method called an automated altimetry mapping procedure (AltiMaP) that allocates altimetry virtual stations (VS) to the Multi-Error Removed Improved Terrain Hydrography (MERIT Hydro) river network. Although this study shows an improvement over the traditional method of allocating VS to the coarse-resolution river network, I have some questions and suggestions to the authors.

We are grateful to the referee #2 for valuables comments and suggestions.

1. First of all, the authors are using the simulated WSEs from the CaMa-Flood model to be compared with the WSEs from altimetry with AltiMaP. I am not sure how this approach (i.e., using simulated WSEs as a reference to evaluate observed WSEs) can be convincing, especially considering the fact that the CaMa-Flood is a global hydrodynamic model calibrated/validated with available in-situ network (I assume so) which are sparse in large river basins, such as Amazon, Congo and Mekong. In addition, how AltiMaP assigned altimetry-observed WSEs can be used in the future for better calibration/validation of the CaMa-Flood model (this could be added in the summary section or in a separate discussion section)?
Can authors elaborate on this?

We extend our gratitude to referee #2 for providing insightful comments. We wish to clarify that our objective in comparing with CaMa-Flood results is to demonstrate how utilizing the AltiMaP method (developed in this manuscript) can enhance the comparison of simulated water surface elevation (WSE) with satellite altimetry. Our intention is not to evaluate the CaMa-Flood model itself. Rather, we employ CaMa-Flood as an illustrative example of model simulations to discuss how a systematic allocation of VSs can lead to improved model-satellite altimetry comparisons.

It is crucial to emphasize that we do not intend to use CaMa-Flood simulations as a reference; instead, we aim to evaluate the AltiMaP and ordinary allocation methods using the same dataset for both observations and simulations. In this study, we did not undertake an extensive calibration process for CaMa-Flood; instead, we utilized the standard parameters. In response to referee #2's first part of the comment, we have revised the manuscript, incorporating the additional text shown in purple.

"We forced the CaMa-Flood hydrodynamic model using the runoff simulated by the Variable Infiltration Capacity (VIC) LSM (Liang et al., 1994) with bias correction (VIC BC) (Lin et al., 2019). The standard model parameters were used in this simulation including parameters such as river bathymetry, river width, and Manning's coefficient. For comparison with WSEs simulated by CaMa-Flood, we mapped VSs to a 6′-resolution global river network after allocating VSs to the

MERIT Hydro network at 3″-resolution using AltiMaP, because the CaMa-Flood river map was derived by upscaling the MERIT Hydro flow direction map using FLOW algorithm (Yamazaki et al., 2009). Then we compared the resulting simulated WSEs with observed WSEs mapped onto the river network based on the MERIT Hydro using the AltiMaP algorithm and the ordinary allocation method, i.e., converting longitude and latitude to the CaMa-Flood grid. In this evaluation, our primary objective is to assess the potential improvement brought about by the AltiMaP method when comparing simulated WSE with the ordinary allocation method. For a fair and unbiased evaluation, we employ the same dataset for both observations (i.e., satellite altimetry) and simulations. By doing so, we create a consistent and controlled environment to assess the performance of the AltiMaP method in comparison to the ordinary allocation method. We would like to emphasize that our intention is not to treat the CaMa-Flood simulation results as an absolute reference. Rather, we utilize them as a basis for evaluating the allocation methods concerning satellite altimetry data. Our aim is to investigate whether the AltiMaP method offers any notable advancements in the accuracy of simulated WSEs when compared to satellite-derived measurements."

Secondly, as the referee #2 correctly mentioned one of the objective of the mapping of the VS into MERIT Hydro is to use them for calibration of model parameter (e.g., Zhou et al., 2022), correct state variables using assimilation (e.g., Revel et al., 2023), and model evaluation (Modi et al., 2022). The river bottom elevation parameter was corrected using a rating curve method by using both satellite altimetry and in-situ river discharge data (Zhou et al., 2022). The model can evaluate and calibrate using in-situ discharge, satellite altimetry, and inundation extent (Modi et al., 2022). In addition, the satellite altimetry data can assimilated into the model to correct the state variable (Revel et al., 2023). All the above were implemented to CaMa-Flood model and used the AltiMaP for them. We revised the text as below as per the referee #2's suggestion.

"Because we used river network-related variables in the AltiMaP VS allocation algorithm, we were able to calculate distances and elevation differences between each VS and the unit-catchment river mouth. These parameters are particularly important for comparing WSEs simulated by coarse-resolution, large-scale river routing models such as CaMa-Flood, which are based on discretized river reaches with a representative elevation for each pixel. Minimizing the distance and elevation difference between the VS and unit-catchment river mouth is critical for improving the accuracy of WSE simulations. Thus, this elevation difference may be used as a proxy to interpret bias between simulated and observed WSEs (Fassoni-Andrade et al., 2021). Satellite altimetry data are also extremely useful for evaluating and calibrating hydrodynamic models (e..g., Zhou et al., 2022) and correcting variables through data assimilation (e.g., Revel et al., 2023b), which requires correct VS allocation to a river network map. The river bathymetry parameter can be calibrated using rating curve method developed using satellite altimetry and in-situ river discharge data (Zhou et al., 2022). Furthermore, the model can be evaluated using multi-variables (i.e., river discharge, WSE, and inundation extent) (Modi et al., 2022)."

2. Line 141: how is this assumption valid? It is not certain whether the observed WSE time series available from the HydroWeb is really the one over the floodplain (which could be flag 20 or 30), or over the open river channel. If the HydroWeb time series are indeed for the floodplain, AltiMaP may be erroneously assigning the VS to the open river channel. Can authors elaborate on this?

We would like to express our gratitude to the referee #2. Since the satellite altimetry can be successfully retrieved river wider than 0.8 km (Birkett and Beckley, 2010), we think the above assumption is valid in the case of the retrackers used in HydroWeb (ICE-1, ICE-2, etc) or other similar datasets even though some other methods can be useful for deriving WSE in narrow rivers (e.g., Sulistioadi et al., 2015).

Moreover, we provided the details about the secondary location (the small channel) as kx2 and ky2. Therefore, the users of the dataset can have flexibility to text both locations. This assumption is mostly used for the flag 30 and the amount of VS categorized as flag 30 is only 1.34% of the total VSs.

We analyzed the root mean squared error (RMSE) and bias using original and secondary location as shown in the Figure S2. We found that our assumption is mostly correct as the RMSE and biases of original location (kx1, ky1) is lower than those from the secondary location (kx2, ky2).

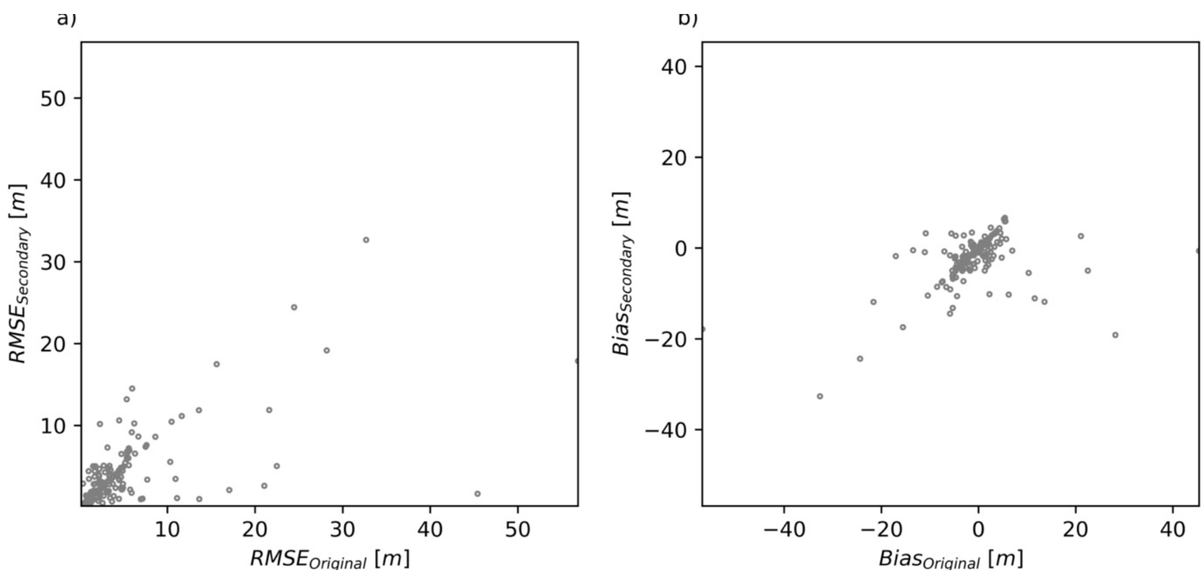

Figure S2: Comparison of root mean squared error (RMSE: a) and bias (b)for Original and Secondary allocations.

We believe the referee is referring to the floodplain lakes in "HydroWeb is indeed for floodplain" because radar echo will be contaminated from the ground (backscatter is useful from a smooth surface). Usually, HydroWeb defined the satellite altimetry over rivers and lakes separately. Therefore, if HydroWeb assign one VSs as river and provide a timeseries of floodplain lake means the observations were taken at non-nadir direction and those VS needed to correct or remove from the datasets. Checking the biased VS against MERIT DEM elevation can one way to identify such erroneous VSs. But to understand error is due non-nadir view or other issues need further investigation. In addition, MERIT Hydro incorporate waterbody map derived from Landsat and river flow direction delineated from MERIT DEM. Hence, most of the river permanent water areas were considered in the MEIRT Hydro.

Considering the referee #2's comment we added following discussion to the manuscript.

"4.4 Limitations and Future Perspectives
Even though AltiMaP is suitable in mapping the VSs into the given river network with D8 connection, the method is not capable of identifying non-nadir observations (such as floodplain lakes near the river channel). One of the major problems in the conventional altimeters in low-resolution mode (LRM) such as ENVISAT was correcting the observations from the non-nadir view was treated as nadir observations (Calmant et al., 2008; Frappart et al., 2006; da Silva et al., 2012). The dual antenna configuration of the CryoSat-2 allows precise position of reflecting point in the radar footprint and solve the signal location along-track and across-track directions (Cretaux, 2022). Moreover, ICESat-1/2 data can also be a great source of importance over terrestrial waters, but the longer revisit time limit the applications in hydrology. Satellites such as CroySat-2 and ICESat-2 provide an addition challenge in using them in river monitoring. CryoSat-2 with its' drifting orbit ~7.5km makes it challenging to define VSs as in repeat orbits (Schneider et al., 2017). With the complex ground track configuration of ICESat-2 makes it complex to use in river monitoring because the assigning method would differ depend on the satellite track orientation with respect to the river centerline (Scherer et al., 2023). However, with slight modification to the AltiMaP, We would be able to map such data into the MERIT Hydro."

3. It is mentioned that mean observed WSEs are used to be compared with MERIT DEM elevation. But it is not explained how "mean" has been obtained. Did the authors simply take the mean of the entire WSE time series? Or did the authors consider the water cycle of the basins? If the entire time series has been simply used to compute the means, that will lead to an inherent bias due to the seasonality of WSE changes. Please clarify.

Thank you very much for the comment. We used simply used all the observation for simplicity for calculating the mean. As we compare the mean WSE with the maximum variation of WSE (30m), the seasonal bias can be smaller than the amplitude of the WSE pulse. Therefore, we believe the bias due the seasonality can be ignored. But we firmly believe it can be considered in the future.

Moreover, the users of the dataset can easily modify this condition as the filtering is separated from the allocation/mapping process.

4. Figure 7: Even using AltiMaP, I see majority of the VSs have high RMSEs over the world. This demonstrates that basically HydroWeb WSEs and CaMa-Flood WSEs are not comparable. There are many factors behind this (as authors mentioned them), but I think the authors should not simply use the time series from HydroWeb without quality check. I'm not saying HydroWeb data is inaccurate, but I'm saying some of their time series may be inaccurate because of the inherent limitation of altimetry over land.

We firmly agree with the referee #2 that the quality checks must be implemented before using satellite altimetry before using them for any calibration/validation/assimilation purposes. Here, we introduce a simple quality check by comparing with the MERIT DEM elevation. However, more sophisticated quality control methods can be considered in the future studies.

In addition, we can get a good understanding which of the VS can be used with large-scale hydrodynamic model by the data provided in the AltiMaP dataset such as distance to the unit-catchment mouth (dist_to_mouth). By considering the only the VSs near to the unit-catchment mouth one make fare comparison with WSE simulated by CaMa-Flood.

Minor comment:

1. Abstract: "much lower (10.6%)" is a bit of exaggeration in my opinion. I would say "a meaningful improvement" or something like that.

Thanking the referee #2 we revised it.

Reference:

1. Modi, P., Revel, M. and Yamazaki, D.: Multivariable Integrated Evaluation of Hydrodynamic Modeling: A Comparison of Performance Considering Different Baseline Topography Data, Water Resour. Res., 58(8), 1–20, doi:10.1029/2021WR031819, 2022.
2. Revel, M., Zhou, X., Yamazaki, D. and Kanae, S.: Assimilation of transformed water surface elevation to improve river discharge estimation in a continental-scale river, Hydrol. Earth Syst. Sci., 27(3), 647–671, doi:10.5194/hess-27-647-2023, 2023.
3. Zhou, X., Revel, M., Modi, P., Shiozawa, T. and Yamazaki, D.: Correction of River Bathymetry Parameters Using the Stage–Discharge Rating Curve, Water Resour. Res., 58(4), 1–26, doi:10.1029/2021WR031226, 2022.