Anonymous Referee #1 - Comment on essd-2022-438

We would like to express our gratitude to Referee #1 for insightful comments and suggestion. We have carefully reviewed your comments and have made necessary updates to our manuscript. We provided 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

I have reviewed the paper "AltiMaP: Altimetry Mapping Procedure for Hydrography Data" by Revel et al. (2023), which introduces a method for allocating virtual stations (VSs) from satellite altimetry data to a river network. This method can improve the accuracy of comparing simulated water surface elevations (WSEs) using satellite altimetry data with those obtained from hydrodynamic models. I find this paper to be well written and recommend it for publication after addressing the following comments and suggestions.

We are thankful to Referee #1 for agreeing to review our manuscript. We appreciate the positive feedbacks from Referee #1.

1) Line 14: The authors have used numbers 10, 20, 30, and 40 to represent different flags, but it would be helpful if they could provide a more detailed explanation of their reasoning. Why not 1, 2, 3, and 4.

We appreciate the comment by Referee #1. We used flags as 10, 20, 30, and 40 for represent the different altimetry mapping procedures. We used those number rather than 1, 2, 3, and 4 to give some flexibility for adding some subdivision inside each flag. For example, we have given the following table for further assign the altimetry mapping.

Table S1: Secondary Flags used in the AltiMaP. Here large and small river are with respective to each river section. The upstream catchment area was used to define the small and large rivers.

Main Flags	*Secondry Flag*	*Description*
10	*10*	*VS was found on the river centerline*
	11	*VS was found on the river channel but not in the centerline and assigned to the nearest centerline*
	12	*VS was found in the unit-catchment mouth*
20	*20*	*VS was found in the ground and assinged to the nearest single channel centerline*
	21	*VS was found in the ground near large river channel in in mult-channel river and assinged to the larger river centerline*
30	*30*	*VS was found in the ground near small river channel in mult-channel river and assinged to the large river centerline*

	31	*VS was found in bifuricating channel and assinged to the large river centerline*
40	*40*	*VS was found in the ocean and assinged to nearest river channel*

For simplicity we have given only the main flags in the manuscript. We will attach the above table as a supplementary.

2) Line 37: The authors should provide statistical measures such as mean absolute error or any to define what they consider to be reasonable accuracy. Please avoid using qualitative words.

Thanking the referee #1, we revised the text as follows:

"Satellite altimetry has facilitated direct and reasonably accurate measurements of terrestrial water levels over the past 30 years, with uncertainties ranging from a few centimeters to a few decimeters depending on the environment and altimeter employed. (Cretaux, 2022; Papa et al., 2022)."

3) Lines 42-45: To improve the readability of the paper, it would be helpful if the authors provide separate references for each different radar altimetry mission instead of listing all the references together. This will make it easier for readers to identify and access the specific sources of information relevant to each mission.

The references shown here is the studies who used the mention satellites for observing lakes and rivers. They have used multiple satellite in each respective study. Hence, it may be better to provid them as common citations. Moreover, we provided the details about the retrackers used, agency, data sources in addition to previous Table 1. Updated Table 1 is attached below.

"Table 1: Satellites altimetry missions which are commonly used for water surface elevation observations. Some characteristics are outlined such as nominal orbit period, temporal resolution, intertrack difference, orbit height, inclination, retracker, agency and data source."

Satellite	*Norminal Orbit Period*	*Temporal Resolution (days)*	*Inter-track distacne at Equater (km)*	*Orbit Height (km)*	*Inclination (°)*	*Retracker*	*Agency*	*Data Source*
T/P	*1992-2006*	*10*	*315*	*1336*	*66*	*onboard*	*NASA - CNES*	*PODAAC*
ERS-1	*1991-2000*	*35*	*80*	*785*	*98.52*	*ICE-1, ICE-2*	*ESA*	*ESA*
ERS-2	*1995-2011*	*35*	*80*	*785*	*98.52*	*ICE-1, ICE-2*	*ESA*	*ESA*
GFO	*1998-2008*	*17*	*165*	*784*	*108*	*Ocean*	*US Navy / NOAA*	*NOAA*
ENVISAT	*2002-2012*	*35*	*80*	*800*	*98.55*	*ICE-1*	*ESA*	*ESA*
Jason-1	*2001-2013*	*10*	*315*	*1336*	*66*	*ICE*	*NASA - CNES*	*AVISO*
Jason-2	*2008-2016**	*10*	*315*	*1336*	*66*	*ICE-3*	*NASA - CNES - EUMESTAT - NOAA*	*AVISO*

Jason-3	*2016-2022**	*10*	*315*	*1336*	*66*	*ICE*	*NASA - CNES - EUMESTAT - NOAA*	*AVISO*
SARAL/Alti Ka	*2013-2016**	*35*	*75*	*800*	*98.5*	*ICE-1*	*ISRO - CNES*	*AVISO*
Sentinel-3A	*2016-Current*	*27*	*104*	*814.5*	*98.65*	*OCOG*	*ESA*	*COPERNICUS*
Sentinel-3B	*2018-Current*	*27*	*52*	*814.5*	*98.65*	*OCOG*	*ESA*	*COPERNICUS*
Sentinel-6MF	*2022-Current*	*10*	*315*	*1336*	*66*	*OCOG*	*ESA*	*COPERNICUS*

4) Table 1: The authors could add an additional column discussing how these different data sources differ from each other in their data generation algorithms.

We would like to thank the referee #1 for the valuable comments. We added few columns to the previous Table 1 as explained above. We added inclination angle, retracker used, agency, and data source. We hope this additional column would give readers information about data accessibility and the differences in the data generation algorithms.

5) Line 46: The authors should provide a brief explanation of how the temporal resolution and inter-track distance of satellite altimetry data affect the temporal and spatial resolution of the data.

We would like to express our gratitude to referee #1 for the valuable suggestions. The frequency of revisits in a repeating satellite orbit determines its temporal resolution. A shorter revisit time means more frequent observations for a single location. Additionally, a smaller inter-track distance results in closer observation locations, with VSs being situated nearer to one another.

We have revised and added the following text to the Introduction section:

"Higher temporal resolution, achieved through frequent passes or shorter revisit times, captures temporal changes with finer granularity, while a smaller inter-track distance provides a higher spatial resolution by offering closely spaced measurements. Consequently, a combination of higher temporal and spatial resolutions in satellite altimetry data enhances the ability to monitor the dynamic processes in the terrestrial surface waters."

6) Lines 70-77: The authors should provide a more detailed discussion of existing studies that have attempted to accurately locate VSs, rather than only discussing the problem. By doing so, they can highlight the research gap that their work aims to address and demonstrate how their method contributes to the current state of research on this topic. Improving the research gap in detail will help readers better understand the significance of the authors' contribution and appreciate the originality of their approach.

We would like to thank referee #1 for the suggestions. To our knowledge, this is the first attempt to systematically allocate satellite altimetry locations to the global river network. A few studies have used satellite altimetry for calibration and validation of model outputs but most of them have either compared WSE anomalies or used VSs near the simulation locations (e.g., Meyer Oliveira et al., 2021; De Paiva et al., 2013). Some others compared satellite altimetry with WSE simulated by the small-scale models (e.g., Domeneghetti et al., 2021; Jiang et al., 2019, 2021). In addition, Schneider et al., (2017) projected the CryoSat-2 observed into the river centerline of Brahmaputra river. In such instances the VS may allocated manually to the river centerline. But in the cases where absolute WSE is needed on large-scale calibration approaches such as in Zhou et al., (2022), an automated robust allocation method is essential. Therefore, we developed an automated mapping method for VSs for large-scale hydrodynamic models.

We have added a paragraph to Introduction Section highlighting the importance of our study as follows (new additions were shown in *purple*) considering the referee #1's suggestions:

"Apart from other model limitations such as uncertainty in model parameters, simplified physics, and bias in forcing, the discrepancy in the virtual station location in the river network is a considerable contributor to the bias in simulated water surface elevation when compared to satellite altimetry observations. Large-scale model calibration studies have utilized WSE anomalies for comparison with simulations, where the rough allocation of VSs in the river proves suitable (e.g., Meyer Oliveira et al., 2021; De Paiva et al., 2013). Conversely, small-scale studies have manually allocated VSs along the river centerline (e.g., Domeneghetti et al., 2021; Jiang et al., 2019; Schneider et al., 2017). Calibrations requiring absolute WSE observations, such as calibration of river bottom elevation using rating curves, demand meticulous allocation of virtual stations (VSs) within the river pixels (Zhou et al., 2022). To effectively utilize satellite altimetry observations for supporting large-scale hydrodynamic model development, a method is required to map representative locations of VSs to relevant river pixels. Moreover, an automated mapping approach becomes essential to facilitate the global-scale model evaluations. Therefore, the development of an automated method for mapping VSs into the river network is paramount to the evaluation of hydrodynamic models on a global scale."

7) Line 117: The authors should provide a more detailed explanation of their data selection criteria, such as period and temporal resolution, and explain why they chose to use satellite altimetry data from HydroWeb instead of other sources.

Thank you, for the selection of VSs, we simply select all the VSs listed in the HydroWeb as potential VSs in this study. However, we compared the satellite altimetry observations with the CaMa-Flood simulated WSE from 2002 to 2019 considering the data availability. The methods we presented here can be applied to any pre-processed data set such as HydroWeb, DAHITTI, etc. However, we used HydroWeb in this study because its' global availability and easy access.

We have revised the following text considering the referee #1's suggestions. Revised or added text have been shown in *purple*:

"2.1 Satellite altimetry data

Satellite altimetry observes water surface heights by measuring the time it takes for radar/laser pulses to bounce back from smooth surfaces. Although satellite altimetry missions were developed for ocean surface observations, they have increasingly been applied to observe lakes and rivers (Abdalla et al., 2021; Calmant et al., 2008; Calmant and Seyler, 2006; Yang et al., 2022). Several agencies have already processed their original satellite altimetry data and produced data archives *for studying WSEs, including the HydroWeb (Crétaux et al., 2011; Santos da Silva et al., 2010), Hydrosat (Tourian et al., 2016, 2022), Database for Hydrological Time Series of Inland Waters (DAHITTI; Schwatke et al., 2015), Global Reservoirs and Lakes Monitor (G-REALM; Birkett and Beckley, 2010), Copernicus Global Land Service (CGLS; Calmant et al., 2013; Crétaux et al., 2011), River & Lake (Birkett et al., 2002), Hidrosat (Santos da Silva et al., 2010; da Silva et al., 2012), and Global River Radar Altimetry Time Series (GRRATS; Coss et al., 2020) archives. In this study, we utilized satellite altimetry data obtained from HydroWeb (https://hydroweb.theia-land.fr, last accessed on 2 February 2023), which offered 12523 VSs at the time of data acquisition. For the study, we considered all available VSs from HydroWeb due to its convenient data retrieval process and global coverage. Initially, we identified all the VSs listed in HydroWeb as potential candidates for inclusion in this research."*

8) Lines 118-119: The authors could improve the readability by moving the discussion about identifying and removing biased VSs to Section 2.4 and providing a more detailed explanation of their criteria for identifying VSs.

We apricate the comment from the referee #1. We moved the explanation about filtering biased VSs to the Section 2.4.

9) Line 120: The authors could include a brief discussion of each data source used in the study and how the data were derived.

We appreciate the suggestion and add some information about MERIT Hydro to Section 2.2. The added text is highlighted in purple.

"An accurate flow direction map is essential for simulating realistic surface water dynamics at the global scale. The river network used in this study is a 3" flow direction map derived from the MERIT DEM (Yamazaki et al., 2017) and water body datasets including the Global 1" Water Body Map (G1WBM; Yamazaki et al., 2015), Global Surface Water Occurrence (GSWO; Pekel et al., 2016), and OpenStreetMap, which are referred to as MERIT Hydro (Yamazaki et al., 2019). The MERIT Hydro generation involved following steps. Initially a "conditioned DEM" was created by lowering the elevation of water pixels in MERIT DEM based on G1WBM, GSWO, and OpenStreetMap. Subsequently, an initial flow direction was determined based on topographic slope using "Steepest Slope Method". Some adjustments were made to ensure the flow continuity. Finally, endorheic basins were detected using Global 3" Water Body Map and Landsat tree density maps (Yamazaki et al., 2019). The MERIT Hydro include an adjusted DEM, river width, height over the nearest drainage, flow accumulation area, and flow direction data. The 3" MERIT Hydro was used to determine whether VSs were located on land, river, or ocean pixels. The allocation procedure for the higher-resolution flow direction map is described in Section 2.3."

10) Lines 144-145: The authors should elaborate on how river bathymetry and river bank height cause deviations.

We thank referee #1 for the question. In CaMa-Flood hydrodynamic model the WSE is diagnosed using river bathymetry $(z - b)$ and riverbank elevations (z).

$$WSE = d + z - b$$

Where d is the river water depth, z is the riverbank height, and b is the river channel depth. Because of limited data availability for river bathymetry, a power-law relationship was employed to estimate the river channel depth. Additionally, riverbank elevations were derived from spaceborne DEMs that may have inherent errors. Consequently, any inaccuracies in the river bathymetry or riverbank elevation can significantly impact the calculations of WSE.

11) Section 3.1: The section lacks sufficient discussion.

We would like to express our gratitude for the suggestion. We have revised and updated the Section 3.1. Updated text were highlighted in *purple*.

"3.1 Allocation of VSs to the river network

Figure 4a shows the global distribution of flags 10, 20, 30, and 40, which VSs initially located on river channel river, land with a single-channel river nearby, land with a multi-channel river nearby, and ocean pixels, respectively. Flag 10 was the most common, accounting for 71.74% of all VSs, followed by flags 20 (26.88%), 30 (1.34%), and 40 (0.04%). Flags 10 and 20 were evenly distributed worldwide. Mostly, large rivers such as Amazon, Congo, Nile, Ob, etc. consist of flags 10 or 20 which indicate the low inconsistencies between VS locations and the river network. Flag 40 is distributed near the ocean in Congo River, Santee River in United States, Lumi Semanit River in Albania, Mahavavy River in Madagascar, and Luni River in India. In addition, flag 30 can be seen mostly in mid-streams where multi-channel rivers exist. Hence, different flags shows different geographical characteristics.

The log probability distributions of upstream catchment areas for different flag values are also shown in Figure 4b. The median upstream catchment areas were 2.73×10^4, 9.95×10^3, 2.16×10^4, and 3.95×10^4 km^2 for flags 10, 20, 30, and 40, respectively. Flag 40 represented the largest median upstream catchment area because those are closer to the ocean and have large upstream catchment area. The distribution of flag 40 was strongly right skewed, influenced by the larger upstream catchment area of downstream Congo River. Flag 20 had the smallest median upstream catchment area, which indicates that most flag 20 VSs were in upstream reaches.

Figure 4c depicts the probability distribution of riverbank elevation for each flag. Lines represent the probability distributions of elevation for flags 10 to 30, with median values of 112.9m, 147.0m, and 141.2m for flag 10, flag 20, and flag 30, respectively. Notably, flag 40 was not visible in Figure 4c due to its very low elevation, with a median of 0.0m (mean=0.54m and std=1.21m). Flags 10 to 30 were distributed from mean sea level to 4790.0m, and there was no significant difference in elevation observed among flags 10 to 30.

The river width distribution for each flag is demonstrated in Figure 4d. Flag 20 exhibited the smallest median river width at 41.4m, with a relatively low standard deviation of 193.3m. On the other hand, Flag 40 displayed the largest median river width of 224.0m, but its variation was substantial (std=1336.6m) due to the wider Congo downstream, which measures around 3170.0m. Flag 10 showed a median river width value of 222.0m, comparable to Flag 40, but with a lower variation (std=683.6m). Meanwhile, Flag 30 exhibited a median river width of 77.7m, falling between the median river widths of Flag 10 and Flag 40. The large variation in river width observed for Flag 10 was due to its widespread distribution across the rivers, while the substantial variation of Flag 40 was influenced by the VSs' location in the Congo River."

12) Figure 4: The authors should improve the overall quality of the figures, and in the density distribution plot, they should change the color code to make the density distribution line for Flag 10 visible.

Thanking the referee #1, we have improved the figures as shown below.

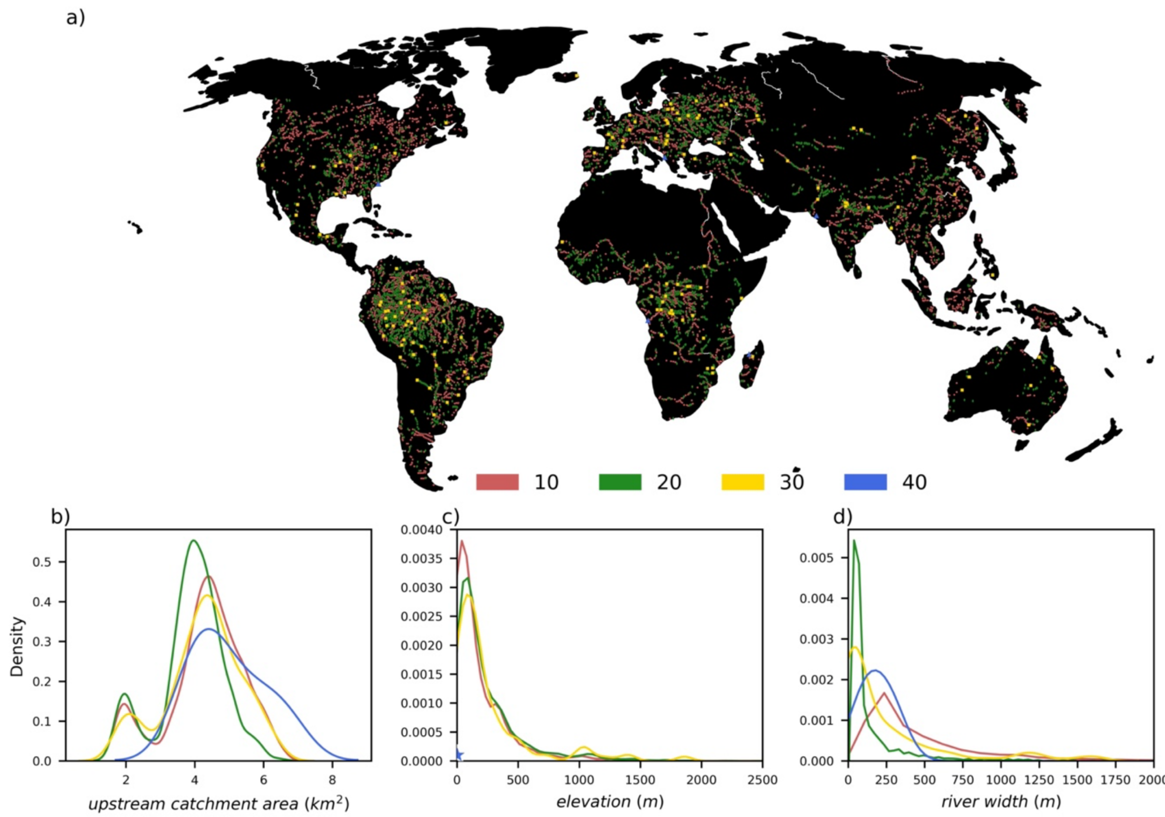

Figure 4: Global map of allocation flags. Panel at lower left corner shows probability distribution of the upstream catchment area in log scale for different flags. Flags 10, 20, 30, and 40 are indicated by light blue, medium blue, dark blue, and red colors, respectively.

13) Figure 5: The authors should improve the overall quality of the figures. The y-axis level is missing, and there is overlap between Figures 5a and 5b.

We would like to thank the referee #1. We have improved quality of Figure5, and a common y-axis label was added to Figures 5b for 5b, 5c, and 5d.

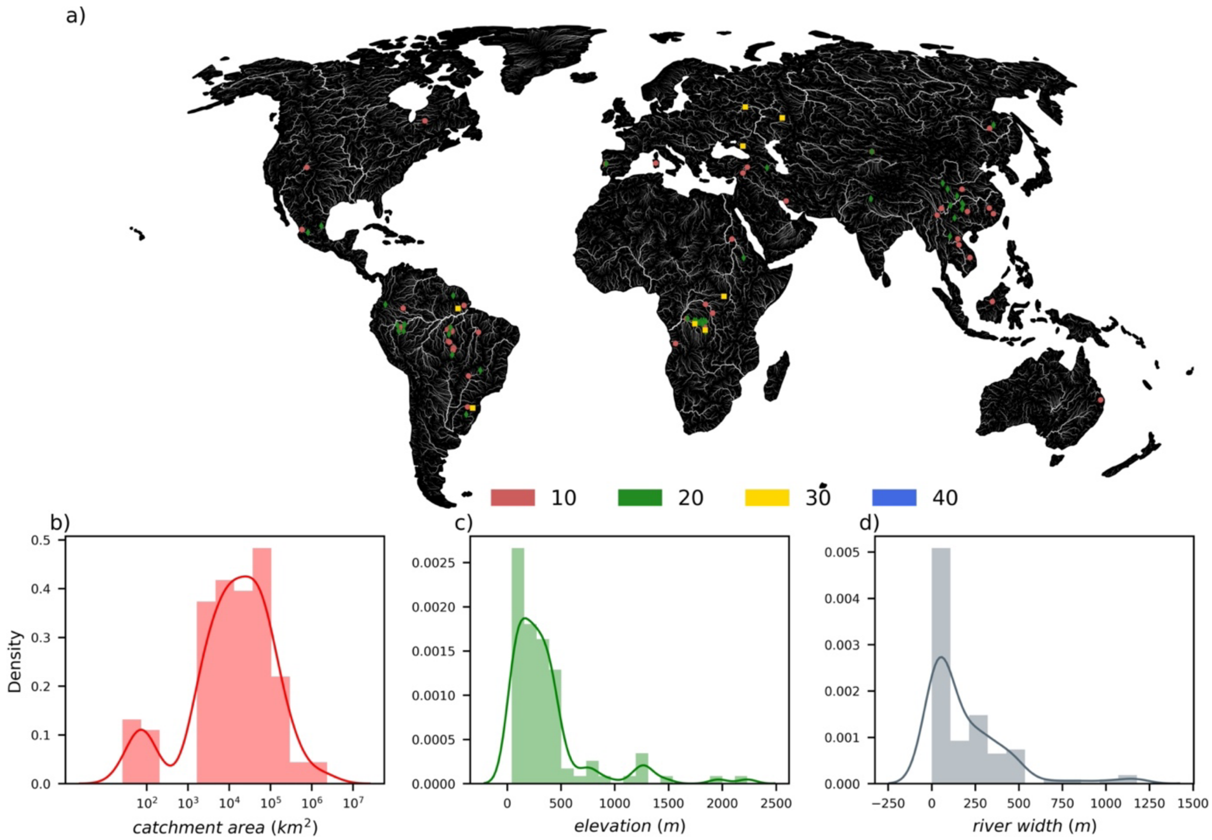

Figure 5: a) Global distribution, b) histogram of catchment area (km2), c) histogram of elevation (m), and d) histogram of river width (m) of biased VSs. Light blue circles, medium blue diamonds, dark blue squares, and red triangles for flags 10, 20, 30, and 40, respectively in panel a.

14) Line 228: The authors should explain why they compared the evaluated results in terms of RMSE? please consider showing the correlation coefficient and bias as well.

We express our gratitude for referred #1 for the valuable suggestions. We use RMSE because it represents overall nature of the errors and useful in evaluating the WSE against satellite altimetry. We have added correlation coefficient and bias to the Table 3

"Table 3: Median statistics of the error of simulated WSE using CaMa-Flood hydrodynamic model. RMSE (root mean squared error), Bias, and CC (correlation coefficient) were presented. The simulated WSE is compared with HydroWeb satellite altimetry data where the VS s were allocated using AltiMaP or the ordinary allocation method."

	AltiMaP			*Ordinary*		
	RMSE	*bias*	*CC*	*RMSE*	*bias*	*CC*
All	*2.68*	*-0.01*	*0.67*	*2.98*	*-0.99*	*0.67*
Flag 10	*2.65*	*-0.43*	*0.67*	*2.94*	*-1.87*	*0.68*
Flag 20	*2.71*	*-0.17*	*0.66*	*3.06*	*-2.46*	*0.66*
Flag 30	*2.72*	*-0.60*	*0.64*	*2.85*	*-1.97*	*0.61*
Flag 40	*0.85*	*-0.37*	*0.02*	*0.94*	*-0.30*	*0.02*

15) Line 229: The authors should explain why the elevation causes an increase in RMSE.

Thank you for the question. In rivers at higher altitudes, the internal slope of the unit-catchment is more pronounced, resulting in greater height variation within the model grid compared to rivers at lower elevations. As a consequence, even VSs situated closer to the unit-catchment mouth can exhibit larger elevation biases compared to the river grid in lower elevation areas.

16) Line 230: The authors should rewrite the sentence to clarify that there is no change in RMSE before a certain threshold (<200 m/km) and that the medium of RMSE increases from 2 to 4 m, not just RMSE.

Thanking the referee #1, we revised the text as follows. The revisions are shown in purple:

"As the distance from the VS location to the unit catchment mouth increased, the median RMSE of simulated WSE increased (Figure 6), mainly due to the difference in elevation between these points. Thus, large errors may be associated with simulated WSE when the VS is located far from the unit catchment mouth. Similarly, the median RMSE of simulated WSE increased slightly as the slope within the unit-catchment increased until slope < 200 m/km, with larger slopes (> 200 m/km) showing an increase in median RMSE from 2 to 4 m. This variation may have been caused by the non-uniformity of slopes within unit catchments of the CaMa-Flood model; however, it was well within the range of variation within unit-catchment slope bins, which reached up to 8 m."

17) Line 265: The authors should explain why?

We would like to thank the referee 1# for asking for the clarification. Firstly, the RMSE was similar in global scale because flag 10 have more than 70% which may not contribute for the large error due to allocation method. Most Flag 20-40 would account for the differences in RMSE between AltiMaP and ordinary method. Secondly, there were VSs with lower RMSE in ordinary method than AltiMaP may be due to compensating for error due to other reasons such parameter errors in the model.

We have revised the text as follows:

"RMSEs were calculated for WSEs simulated by CaMa-Flood and forced by VIC BC runoff (Lin et al., 2019). The spatial distributions of WSE RMSEs for VS allocations obtained using AltiMaP and the traditional method of allocating VSs to the CaMa-Flood grid are shown in Figure 7. Traditional VS allocation was performed using directly converting longitude and latitude information to coarse-resolution (i.e., 0.1°) grids. At the global scale, RMSEs were generally similar between both VS allocation methods. However, the satellite altimetry was better represented by AltiMaP for 17.52% of VSs (negative ΔRMSE) and by the traditional method for only 12.85% of VSs (positive ΔRMSE). The lower ΔRMSE of ordinary method may be due to the fact allocation to a nearby grid by ordinary method compensate for the errors in the model such as river bathymetry error (Modi et al., 2022)."

18) Figure 7: The authors should discuss the expert method and ordinary method in the text to help readers better understand the differences between them. Alternatively, they could use constant terms to avoid confusion between the traditional method and ordinary method terms.

Here expert method simply refers to the method developed in this manuscript (AltiMaP). Thanking the referee, we revised all the instance of expert method to AltiMaP.

19) Figure 8: The authors should address the same comment as Figure 7. Additionally, they should move Figure 8 from outside of Table 3 to improve the organization of the paper.

We appreciate the referee #1 for pointing out unvolunteered error. We have revised the expert method to AltiMaP and move Figure 8 outside the Table 3.