Dear Handling Editor and Referees,

We want to convey our appreciation for the valuable feedback and constructive comments received. Your insightful input and suggestions have enriched the quality of our manuscript. We are grateful for the time and effort you invested in reviewing our work. 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*. The revisions are highlighted in purple.

Kind regards

Referee #4 - João Paulo Brêda

I want to congratulate the authors for this work. This study describes the development of a dataset given by the allocation of virtual stations to the MERIT Hydro river centerline. In addition, the authors cross-validated the WSE simulated by a global hydraulic model (CaMa-Flood) built from the MERIT-Hydro dataset and the VS observations. To be honest, the authors did not present any great novelty with this approach for VS allocation, however, working with such big data on a global scale and the posterior analysis deserves a publication. Also, the dataset is going to be useful, especially for the next applications of global models built from the same MERIT-Hydro dataset.

We would like to thank the Referee #4 for the valuable comments and suggestions.

I just have a few comments that the authors could consider before publishing:

1. Are both datasets (MERIT and VS) referenced to the same geoid?

   We would like to thank the Referee #4 for question. Water surface elevations (WSEs) from satellite altimetry were converted to EGM96 before comparing MERIT and VS.

   Considering the Referee #4's comments, we revised text is shown in purple as follows:

   "RMSEs were calculated for WSEs simulated by CaMa-Flood and forced by VIC BC runoff (Lin et al., 2019). Both simulations and observations were converted to the same geoid before calculating RMSE (i.e., EGM96). 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)."

2. How does the algorithm automatically classify nearest multi-channel and nearest single-channel? It is not specified in the manuscript. (I assume that the rivers -flagged 10- are indicated on the MERIT-Hydro itself).

   We would like to thank the Referee #4 for the valuable comment. The AltiMaP algorithm identifies river sections perpendicular to a given river, assessing their downstream connectivity. In addition, we defined a distance threshold in perpendicular direction to identify only multi-channel rivers.

   Considering the referee's comments we modified the flowing (updated text in purple)

   "VSs must be assigned to river network pixels of the hydrodynamic model for accurate comparison of simulated and observed WSEs. The DEM-based river network can deviate from the cause of the actual river due to errors in DEM and low representability of the coarse-resolution of the river network (Amatulli et al., 2022; Paz et al., 2006; Yamazaki et al., 2009). Moreover, the reported location of the VS provided in HydroWeb can be further away from the actual river because HydroWeb provides the center of the search region, within a range of a few kilometers (e.g., 5 km × 5 km). Therefore, an important step in allocating VSs to large-scale hydrodynamic models is to assign each VS to a river centerline on a higher-resolution flow direction map (e.g., MERIT Hydro, at 3"). A schematic diagram of this allocation process is shown in Figure 1. Initially, the satellite altimetry auxiliary data (e.g., longitude and latitude) for each VS were converted into 3" pixels. Then we flagged each VS according to the land cover of the initial allocation of the pixel, with 10, 20, 30, and 40 representing river channel, land with the nearest single-channel river, land with the nearest multi-channel river, and ocean pixels, respectively (Figure 1). The secondary flags also defined to represents more special cases as defined supplementary Table S1. Finally, we searched for the centerline of the nearest river according to geometric distance and allocated the VS to that location. VSs initially located on land pixels with the nearest multi-channel rivers were allocated to the nearest largest channel of the multi-channel river (considering the upstream catchment area). The AltiMaP identifies multi-channel river by searching in a direction perpendicular to the specified river considering their downstream connectivity. We assume the observation is from the largest river when there are multiple river (Supplementary Figure S1) channels near the VS location because backscatter from the narrow river can be highly influenced by non-water features and mostly successful retrievals of WSE can be seen on larger rivers than ~0.8 km. (Birkett et al., 2002)."

   In the AltiMaP algorithm, if the initial allocation of VSs coincides with a MERIT Hydro-defined river, the algorithm will define it as flag 10.

3. Line 194. Just an observation: The criteria for the removal of a VS should be also related to the standard deviation of the respective VS (or the maximum water level difference). If the water level doesn't vary more than a few meters (1 or 2) annually it doesn't make sense to keep a VS that its mean is more than 10 m higher or lower than the actual DEM.

   We would be grateful for the Referee #4 for the valuable suggestions. We also think that the standard deviation should also be considered in some instances depending on the application

of the satellite altimetry data. But in this study, we try to keep the comparison simple as possible by using MERIT elevation only as reference on the standard deviation for WSEs are hard to find.

Secondly, the removal of biased VS is a post-processing of the AltiMaP. Hence, we keep room for the users to tailor their requirements in removing biased VSs. Depending on application the users have the flexibility to change the criteria for filtering biased VS.

In addition, this kind of low variation (e.g., 1-2m) of WSE can be mostly due to non-nadir direction observations. Correcting such kind of errors are beyond the scope of our study but we have discussed this in the section "4.4 Limitations and Future Perspectives".

4. Line 309. Verb missing: "may be very small"?

   Thanking the referee, we corrected it.

5. Line 413. Is this "min_val, max_val" supposed to be there?

   Thank you very much pointing out this mistake. "min_val, max_val" should be removed from this text.

Anonymous Referee #2

We express our gratitude for Referee #2 for the time and effort dedicated to our manuscript.