Dear reviewer,

Thank you for reviewing our paper, your constructive comments and your attention to details. We appreciate your voluntary effort and we revised the manuscript according to your comments. The comments have been addressed as following:

How easily can the user apply the Python code to allocate GRDC gauges to their own mesh? I note there are a lot of large-scale global river network meshes that used by different models. Even at the same spatial resolution, they can have different representations of river network because different algorithms were used. Without the authors' experience, can we successfully use the Python code?

We already have some feedback from users on implementing our Python code. Some Python expertise is needed to adapt the code to your needs. We have a documentation on GitHub https://github.com/iiasa/CWATM_grdc_calibration_stations and zenodo https://doi.org/10.5281/zenodo.6906577 where each module, the input and output files are described. Even if models use a different representation of river network there is a kind of standard format to describe the eight-*direction* (D8) *flow* model network. We are using the D8 direction codings also used by the MERIT Hydro Dataset (Yamazaki et al., 2019), which is also used in ArcGIS https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/how-flow-direction-works.htm. In case you use another direction coding like the LDD network https://pcraster.geo.uu.nl/pcraster/4.4.0/documentation/pcraster_manual/sphinx/secdatbase.html you can adapt our code quite easily or using the pyflwdir library https://pypi.org/project/pyflwdir/0.4.3/ to convert formats from Eilander et al. 2021, which has a good documentation on flow direction. In addition, we are not limited to 5 or 30 arcmin. Changing our script to other resolutions and non lat/lon projections is also possible.

We added to the paper in 5. Code and data availability Line 432ff:
We used input data from MERIT Hydro with a resolution of 3'' and an eight-*direction* (D8) *flow* model network format, but the code can be changed to use any resolution and non-geographical projections as input/output format.

Theoretically, the proposed method should be more accurate than previous method. But I failed to see it from the presented results, for example, some clarifications are needed in Figure 2 and Figure 7 (see my detailed comments below). In addition, only two basins were shown to demonstrate the improvement by using the proposed method. I wonder how many gauges in total will be improved (similar to the basin in Figure 2)? This will be a critical metric to report. It will be a significant contribution If the proposed method finds improved drainage area as the previous method for a large fraction of the selected gauges.

Thank you, for this question. This number would indeed show the advantage of the proposed method.

We changed line 374ff:

For the 2741 selected station resolution of 30', we found 68 cases (2%) where the station location would account for the wrong basins, which the UPA and distance method could not detect. For 684 stations (25%), we chose basin representations of the stations that fit better to similarity and UPA than to UPA and distance. For the 6414 selected stations for 5' resolution, we had 23 cases of station mismatch (0.7%) and 680 (11%) where we chose another basin representation than with UPA and distance.

Line 21: … such as Nash-Sutcliffe and Kling-Gupta for calibrating global hydrological models.
We changed this

Line 31: GSIM database (cited by the author) provides over 30,000 stations for streamflow measurement, which is more than GRDC. Also, GSIM provides the shapefile of the drainage area for each station.
We added this sentence above.

Line 157: [0,1]

Figure 2: Subplot (a) seems to plot the contributing area of cell No 14 instead of cell No 12.
Thank you for looking so carefully at this. Yes indeed, it is cell No 14.

Line 189-Line192: This be described earlier, probably before the presentation of Figure 1 and 2.
We put it before Figure 1

Line 237: Do you mean use stations with UPA larger than 10km^2?
We changed to: an UPA larger than or equal to 10 km^2

Figure 5: increase the font size. I can barely read the station number from the subplot (a). Add legend for the red and green circles.
Changed figures and increased the font size

Figure 7: What does the red circle mean? It will be helpful to plot the gauge location and allocated cell centers (both the right and wrong ones) on the map too. I don't think Figure 7b is a good example to show the benefit of the proposed method. The dark blue area is very close to the light blue area, though the light blue area is closer to UPA derived from high resolution DEM. But the dark blue area outlet is closer to the gauge location. So, the previous method (compare UPA and distance to original gauge location) should give us the right contributing area on the coarse resolution mesh.
Thank you for pointing this out. We didn't check with the UPA and distance for comparison.
Even for 7a you might find the right basin with the UPA and distance method, because the right basin is 107 km from the station point and the wrong one 120 km.
Therefore, we choose better examples where only the similarity method gets to the right result.
We changed line 344ff:
The station "Above Babine River" of the Skeena River in Canada, GRDC No. 4245920, is the station shortly before the junction with the Babine River. If we take the location of the station GRDC No. 4245920 on 30', we get the Skeena and the Babine River joined together. We have to move the station to allocate it to the correct basin. The reported UPA of the station is 12,400 km^2. If we had selected only by upstream area or by weighted upstream area and distance, we would have chosen the Babine River (UPA of 30': 12,495 km^2) in preference to the Skeena River (UPA of 30': 11,937 km^2). Figure 7a shows that the selected 30' basin in darker blue (Skeena River) with the lower UPA fits better with the high-resolution basin even if the distance to the cell center of the Skeena basin is 59 km (distance between green dot and dark blue square) compared to the distance to the Barbine River of 28 km (distance between green dot and red square).
Figure 7b shows a station mismatch selected by the UPA at 5'. The river Khudan in Russia, GRDC No. 2907025, has a reported UPA of 7,800 km^2. We only shifted the station by 0.8 km to fit the 3' high-resolution network. If we select by UPA, the Uda River, with an UPA on 5' of 7,901 km^2, fits better than the Khudana River, with an UPA on 5' of 7,673 km^2. Also, the cell center of the Uda River is closer to the station (4.4 km) than the cell center of the Khudan River (8.2 km). Selecting by area and shape similarity points to the correct basin, shown in dark blue in Figure 7b.

Line 358: shown in dark blue in Figure 7b?
Corrected

In addition, we added in acknowledgment:
The project has received funding from European Union’s Horizon EUROPE Research and Innovation Programme under Grant Agreement N° 101059264 (SOS-WATER).