Earth Syst. Sci. Data Discuss., referee comment RC1
https://doi.org/10.5194/essd-2022-231-RC1, 2022

Comment on essd-2022-231

Anonymous Referee #1

Referee comment on "The use of GRDC gauging stations for calibrating large-scale hydrological models" by Peter Burek and Mikhail Smilovic, Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2022-231-RC1, 2022

General comments

This manuscript describes the new location data of GRDC stations for calibrating large-scale hydrological models. The calibration of these models relies on the accuracy of stations allocated on the gridded river network of different spatial resolutions. The data is developed based on the idea of the 'intersection over union ratio approach', which quantifies the similarity of the watershed shapes between low- and high-resolution gridded watersheds. The new dataset's accuracy was reasonably good compared to the previous versions; thus, the dataset attracts interest in hydrological communities and is worth publishing. Although the estimated precision is promising, however, I think the manuscript does not contain an adequate description as a data paper. Given the location of GRDC stations is a widely used essential information in hydrology and earth system science studies, I think the manuscript is worth publishing on ESSD, after corrections on a few ambiguous parts.

Specific comments:

(1) L 89. Is "Upstream area error" used in the subsequent manuscript?

(2) L93-94. The unit of the distance was not described here. The values of OC calculated with the described equation differed depending on the unit of the distance used. Because the values of (1-'area accordance') are always less than one, the OC values are predominantly determined by the distance if the unit is in kilometers or meters (usually larger than 1). At least, the second term of the equation should be normalized to the range from zero to one. Also, the authors should explain how the weighting of the second term, 2, was determined.

(3) L 141=143: The choice of the weighting factor for calculating ED should be explained.

(4) L 144: Figures 1 (and Figure 2 as well) are not appropriate for the examples describing the automatic upscaling process based on the similarity of shape, because it seems that only the 'area accordance' suffice for the selection of station 7. I would suggest the authors select a more appropriate example. I think Figure 7 worth explains the problem of mismatch in the upscaled stations and how they are solved with the proposed procedures. Upon revising, please consider the necessity of Figure 2 as it was not referred to in the manuscript.

Technical corrections:

(1) L18-21: I think this part is not relevant to the main context. Suggest deleting.

(2) L 92: Delete “)”.

(3) L 133: The number of cells for each resolution does not consistent with those described in line 303. I think this line should be corrected.

(4) Figure 3 was not referred to in the manuscript.