Dear Dr. Do,

Please find below our responses to your comments.

This manuscript introduces the African Database of Hydrometric Indices (ADHI), an unprecedented collection of streamflow signatures for Africa. I believe the data product will be greatly appreciated by the regional and global hydrology communities (myself included) as it can potentially fill a significant gap in in-situ records of streamflow and thus can advance hydrologic research over the tropics.

Thank you for your interest in this dataset.

Although I fully support the publication of this dataset, I have the feeling that the dataset (and associated manuscript) has been developed in a rush, and thus has missed an opportunity to become a great product that benefits a broader range of users. For example, only time series of annual mean, annual max, and annual 7-d min streamflow indices are published although the authors have done an excellent job in synthesizing and quality-controlling that much daily streamflow data.

The main goal of submitting this data paper in ESSD is to open the discussion with potential end-users of this dataset, about potential improvements on the short term but also on the long term. Contrary to the Global Streamflow Indices and Metadata Archive (GSIM), this ADHI database will be updated over time, by including recent discharge measurements but also ancillary data to document the catchment characteristics. To provide these hydrometric indices is a first step, the next step would be to provide much more attributes such as those available in the CAMELS dataset (Addor et al. 2017).

According to your recommendation and those of the other reviewer, but also the feedback received from some users who already downloaded the database (240 as in January 2021), we improved the database to better fulfill the end-user needs. The new upgraded version should appear in the online repository in a couple of days.

I'm recommending a major revision to encourage the authors to improve and make the dataset more attractive to the international community. Below, I listed three major improvements that I strongly suggest the authors consider.

1. Expand the streamflow indices that could be accessed publicly.

- I appreciate the challenges in data restriction that the authors could have faced, but I think that it is defensible to increase the number of published time-series indices from three (in the current phase) to that described in Gudmundsson et al. (2018) - which I believe the authors have mentioned in their manuscript.

We agree, following also the recommendations of Reviewer 1, we added more hydrological signatures, using the TOSSH toolbox recently released (https://sebastiangnann.github.io/TOSSH_development/p2_signatures.html), to ensure the homogeneity of the calculation procedures across different datasets.

- I do think that at the minimum, time series of monthly indices (mean, max, min) would be highly appreciated by the global community to support a wide range of hydro-climatological research.

We added these time series in the database.

- I feel that the static percentiles that currently published in the summary text file could be re-processed using the block-window approach (e.g. yearly) to derive time-series that are useful to assess hydrological changes in Africa.

We added the time series of the percentiles.

2. Providing catchment shapefiles

Figure 3 shows that the authors have also compiled/generated a great collection of catchment boundaries. This is another great asset that could benefit a broad range of end-users. I think publishing this information will not inflict any troubles regarding data policies.

Indeed, the catchment boundaries were delineated using a public dataset (HydroSheds DEM). As indicated in section 2.4 we were able to compare for some basins the areas calculated with the available metadata. But uncertainties remain both on these metadata from very different organizations, but also on the basin delimitation procedure due to the uncertainty on the coordinates of some stations (as indicated also in the GSIM paper). The automatic relocation procedure of stations applied for the GSIM database is probably not optimal in regions with complex orography or areas of low relief and we experienced the exact same issue.

This is why in the last months we tried to: 1/ collect ancient metadata mostly from field campaigns about catchment maps and stations coordinates (in particular many scanned documents from this portal https://horizon.documentation.ird.fr) 2/ delineate the catchment boundaries within a GIS for all catchments where the automatic delineation did not work. Now the catchment boundaries will be provided in the updated database (in shapefile format).

3. Although I have not provided any specific comments on the manuscript (as I expect a major revision to make the manuscript become stronger), I have some general comments on the writing that may help better highlight the contribution of this dataset:

- The title: please consider some assertion titles such as "The production of seventy-year long streamflow indices for 1500 stations across Africa." This type of title reflects better the usefulness of the AHDI and thus will be more attractive to prospective users.

This title would be misleading to the readers, since not all the 1500 stations have time series over the last seventy years. We modified the title to mention the time period considered, 1950-2018.

- Some figures were not associated with an insightful discussion (Figure 1 is not exactly what described in Section 2.1; Figure 5 was completely left out in the discussion). Please expand your discussion regarding any "lesson-learned"

working with this dataset. For instance, some discussion about the relationship between the annual precipitation (shown as the background of Figure 1) and annual streamflow (generated by the authors) could be useful; section 4 contains effectively only two lists of bullet points- but could be expanded to include examples of "spurious patterns", substantial local changes, or improvement relative to the GRDC (see below.)

We re-organized the figures and produced new ones, in particular showing the stations from the different data sources and the link between annual precipitation and mean runoff.

The section 4 is about the content of the database, where we describe the file contents.

Following your recommendation, we expanded the sections 3.1 and 3.2 to include more details (and figures) on the indices computed: the links between runoff and catchment area, mean precipitation, the spatial variability of the different indices in relation to the climatic zone and catchment properties etc. However, we think that a deepened analysis of African hydrology is not in the scope of this data paper, as the suggestion of showing results about local changes (trends). This work would require a deeper analysis of the provided data (and this analysis would not fit in a single paper), in order to study long term trends, water balance components, using different precipitation and evapotranspiration datasets, the relationships with land use and geology - among other possible topics of interest.

- I also think a map showing improvement of ADHI relative to the GRDC database (perhaps in Section 4) could be useful for end-users. For instance, the authors can classify stations into three categories (i) new stations (relative to GRDC), (ii) extended record stations, and (iii) no improvement. The efforts of the authors to publish this data are greatly commended, and I am very excited about the release of the updated AHDI.

This is a good idea; the information was already present in the metadata but we included in the revised manuscript a map of GRDC/non-GRDC stations.

It should be noted that we did not merge the station data from SIEREM and GRDC: if the same station exists in the two databases, we kept only the one with the longest records.