

Reviewer 4 – Franziska Clerc-Schwarzenbach

General comments

Many thanks to the authors for their efforts that they put into the creation of this new dataset version. Indeed, this dataset offers exciting new opportunities for hydrological modelling, as stated in the conclusions! I enjoyed reading the manuscript and exploring the dataset a lot and I think that both the manuscript and the dataset have been very carefully and well designed.

Many thanks Franziska for the kind comments and the helpful feedback. Please see our detailed responses below.

My main suggestion for improvement is that the authors could give some guidance on which of the data should be used when there are several options available for the same variable (especially for the hydrometeorological time series), if possible. While the users of the dataset can still use a different option, this would lead to less random decisions and thus more comparability among different studies using the dataset. As the meteorological data used in CAMELS-GB v1 are not updated anymore, I assume that users should use HadUK whenever possible. Maybe this (or some other guidance) can be stated clearly in the text and / or in a read me file.

Thanks. In response to several reviewer comments, we have now added a new section to the paper (Section 3.3) detailing guidance on the use and selection of the hydro-meteorological datasets.

Furthermore, I think that it is great that in many instances in the manuscript, users are made aware of potential issues and what to take care of when working with the data. For an easier access to all these ‘warnings’, I would suggest to include one table, list, or graphic that summarizes them. I can imagine that users tend to be more aware of all these important points when they are well visible and all listed in one place.

Thanks – we believe the new section 3.3 now addresses this comment.

Minor comments regarding the dataset

- When clicking the link to the dataset via ‘Assets’, ‘Datasets’, I ended up at the CAMELS-GB (v1) archive. The link given in the abstract is fine, but currently not clickable (see technical corrections).

Thanks for spotting this – we will make sure this is updated on the ESSD website and that the link is clickable in the final manuscript.

- I think that a simpler bulk download option (other than via WGET) would be helpful for users that quickly want to get all the data, without any detours. Maybe, the different directories in which the single files are currently accessible could be made available as zip files. Especially for the timeseries files, this would make things easier.

We appreciate that different users will have different expectations for data access. We have received feedback on this dataset and the previous CAMELS-GB release that many users just want data for one or a few stations (and therefore would not want to download the entire dataset). We need to balance the needs of these users with users that do want to download the entire dataset. Currently, it is not possible to enable both these options in the data centre, although they are considering future developments to enable this.

Our expectation is that most users wanting to analyse the entire dataset will have programming skills and an example wget statement is provided on the download page, as well as further guidance to help:

<https://eidc.ac.uk/help/getdata/downloadData#programmaticAccess>. Furthermore, we are developing online notebook services to allow direct access to and analysis of CAMELS-GB and other datasets, which would not be possible with a large zip file.

- For the groundwater wells, I would suggest to also add the latitude and longitude in addition to the easting and northing currently given (as done for the gauges).

Thanks – we will make this change in the revised dataset.

Minor comments regarding the manuscript

The manuscript is easy to understand and clearly structured. Below, I list some minor suggestions for improvement in chronological order:

- L47: I would split this sentence, as I think it contains two points. On the one hand, an increasing number of LSH datasets, and on the other hand large LSH projects such as Caravan that would also not be possible without the FAIR principles.

We have revised this sentence to make clear this is a separate point:

“There is now a growing community of LSH datasets, such as the CARAVAN dataset, which included data for 6830 catchments when published in 2023 (Kratzert et al., 2023) but now contains data for more than 20,000 catchments.”

- L57: The catchments of CAMELS-GB are not only part of Caravan, but also of EStreams. This could be added here, alternatively, I would take out the plural form ‘datasets’.

Thanks, we have revised the sentence to:

“It has also been incorporated into continental (do Nascimento et al., 2024) and global catchment datasets (Kratzert et al., 2023).”

- Figure 1: I would suggest to add the full words to the titles of b) and c) and I think it would be easier to grasp the figure if the variable itself was noted again on the axes of the plots (instead of just the dataset with the units).

Thanks, we have made these changes.

- Caption of Figure 1 as well as L116: I would stick to ‘precipitation’ instead of using ‘rainfall’.

Thanks, we have updated this.

- L169: How did you handle the shifts from and to daylight saving time? This could be added here.

There is no shift from and to daylight saving. All data is stored against the time in UTC / GMT. We have added a sentence to clarify that the timeseries data is reported in UTC with no daylight saving time adjustment. In addition, all references to specific times (e.g., 09:00) throughout the text, as well as in Tables 3 and 5, have been updated to explicitly indicate UTC (e.g., 09:00 UTC).

“Hourly hydro-meteorological time series of rainfall and river flow are provided for the 671 catchments from 1st October 1990 09:00 UTC to 1st October 2022 08:00 UTC (Table 3). All timeseries data in CAMELS-GB are reported in UTC with no daylight saving time adjustment.”

- Figure 2: In b), it could be helpful to add to the axes labels that this is the fraction of wet hours, so one does not have to search for it in the caption first. In c), ‘Elevation’ should probably not be capitalized, and I would suggest to give the unit of the elevation as m a.s.l. instead of just as m.

Thanks, we have made these changes.

- L294: I appreciate that you clearly state that the PET data without interception loss were used for the calculations. I think this paragraph would benefit from an explanation why you decided to do so.

Thanks, we have revised this sentence to:

“(without correction for interception as it has more consistency with other global and national PET products)”

- L323: I think it would be better to write something like ‘a trend for increasing urban land cover’ instead of just ‘increasing land cover’ here, as the latter may be confusing when read without the preceding sentence.

Thanks, this sentence has been revised to:

“While a trend of increasing urban land cover over time...”

Technical corrections

Thanks these have been revised as suggested. We will ensure all links are clickable in the final manuscript.

- L29, L157, L448: These links are currently not clickable (I think because parts of the link are not linked).

- L37+: Since ‘Caravan’ is not an abbreviation, the name of this dataset is not capitalized.

- L45: ‘finable’ instead of ‘findable’.

- L85: I think this ‘is’ should be a ‘was’.

- L89: Subtitle should not be in title-case, and you write ‘timeseries’ as one word elsewhere.

- L130+: space missing before ‘km’. Happens also for other units, e.g., L144 before ‘mm’.

- L300: ‘half low date’ instead of ‘half flow date’.

- L399: The hyphen after ‘high’ should be a dash.

- L418: Should this then be ‘abstraction and discharges data’?