

GRDC-Caravan: extending Caravan with data from the Global Runoff Data Centre

Claudia Färber, Henning Plessow, Simon Mischel, Frederik Kratzert, Nans Addor, Guy Shalev, and Ulrich Looser

Reply to Justification (Minor Revision)

Dear Dr. Hassler,

Thank you very much for your valuable comments and questions. Please see our response below.

Comments:

- 1) Data description, Table 1 and the respective dataset: It takes quite a while to find the column names in the csv in the table due to the different orders and the names including numbers for different months, classes, etc. that are not sorted together. It would help if you had a) the same order in both the csv and the Table 1, b) put the logically related tables in columns together (e.g. _1 to _12 of the same attribute), or everything that is part of the "hydrology" part9, and c) if you put the attribute name also in a separate column in Table 1, similar to Tables 2-4. And it seems the column name for "Land surface runoff" is missing?

We have resorted Table 1 and the respective dataset. Both are now sorted in alphabetical order, but we kept the classification into thematic groups in the data description file to provide the users a better overview what is available. The attribute name is now in a separate column. The column name for "Land surface runoff" has been corrected.

- 2) Data description tables in general: It's quicker to relate the tables in the description to the csv files, if you also put the file name in the table header. Also: for completeness you could add the gauge_id to all tables - at the moment it is a bit random that you have a separate paragraph on it that is meant for all attribute tables, but nevertheless it shows up in Table 4. Simply adding it to all tables as first column would simplify things.

We have fully addressed this comment and have added the file path into the table header and the gauge_id as first row to all tables.

- 3) Data description Table 2: The names probably refer to an older version of the dataset. As you're describing the one including the Penman-Monteith estimates in the ESSD Paper and refer to that version, the column names in Table 2 need to be updated.

We apologize for this issue. The table has been updated accordingly.

- 4) side note: The header of Table 3 could be more descriptive, as "metadata" pretty much refers to all information in the whole data description. Maybe "location attributes" or something like that would be more specific?

The table header was changed to "gauging station attributes".

Questions:

- 1) In the license file there is one license from Belgium that says "no sharing with third parties allowed". Can this data still be shared publicly as part of the dataset?

We apologize for this issue. We have contacted the Hydrological Service of Wallonia again to clarify the distribution of the data in GRDC and the present dataset. They confirmed the permission to share the data. The license file has been updated accordingly.

- 2) In the netCDF time series: was there a specific reason why the files are not cf conform? (as sticking to the standard names etc. usually helps with using, understanding and plotting the netCDFs...)

You are right that following the conventions would have been better. However, the used NetCDF format has been defined in the original Caravan dataset and our dataset is an extension to it. Changing it now, would complicate the compatibility between the different Caravan extensions. Therefore, we followed the naming conventions and structure of the Caravan dataset.