

## Major comments

1.

- Comment: L 106. “A Minimum of five years (1826 days) of observations required. We considered this sufficient to calculate streamflow signatures.” Many CAMELS data sets, including the first one for the USA, require about 20 or 30 years of Q observations for inclusion. For long-term studies, WMO recommends at least 30 years of hydrometeorological observations. The reason for these stricter requirements is that hydrometeorological variables often have long-term persistence (e.g. O’Connell et al., 2022) and temporal clustering (e.g., Chagas et al., 2024; Lun et al., 2020), thus, 5 years of data may not be representative of long-term hydrological behavior. In order to include the hydrological indices of catchments with fewer than about 20 years of data, it would be interesting for the manuscript to assess whether these values are representative of long-term conditions, for example by comparing with neighboring catchments that have longer data.
- Actions: Modified the corresponding line, cautioning that short timeseries may not represent long-term conditions, and that the end user needs to evaluate suitable timeseries length for their application. Added comparison between the indices of short and long timeseries to section 6.2 of the revised manuscript (section 7 of the preprint).
- Reply: We agree that five years is not long enough for representing long term state of a catchment robustly. The previous formatting was not particularly informative of our intents and could be read in a misleading way. The text now explains that we wanted to include catchments with shorter time periods as they can still be very useful for studying individual events or calibrating/training hydrological models, for example. We consider it reasonable to provide indices also for these catchments, so that the end user can get a quick overview of the hydrologic behavior of the catchment during the period of available data. The length of time series used for calculating the indices is also clearly indicated in the hydrologic indices data file, allowing for easy filtering based on user requirements. As requested, we compared indices of short timeseries to longer neighbouring catchments. The results are now included in the section 6.2 of the revised manuscript (section 7 in the preprint).

2.

- Comment: L 211. “In addition to providing the aforementioned data, we combined the data into one convenience PET attribute from 1981, which was also used for calculating climatic signatures (see section 5.2). This was done by using snow evaporation when the snow depth of the catchment was over zero, and filling the snow-free days with FMI PET if it was available; the remainder was completed by ERA5-Land potential evaporation.” Combining the time series from two different sources can be problematic because each source is usually built from completely different data and models, particularly evapotranspiration variables which often have high uncertainties. In order to keep the combined PET time series, it would be interesting to include a more in-depth investigation of whether that is appropriate, including comparisons with other data sources, where possible, and assessing if it does not introduce step changes in PET and other spurious artifacts.
- Action: Investigated step changes, modified text at sections 3.2 and 6.3 of the revised manuscript (4.2 and 7 in the preprint)
- Reply: Combining different data sources for evaporation is indeed not ideal, and investigating the step changes gives users important information to decide whether this combination suits their needs. We want to emphasize, that all presented evaporation datasets are available separately in CAMELS-FI, so that users are free to select a dataset that suits their requirements, and are not forced to use the combined pet. Comparison to a reference dataset would have been interesting, but unfortunately the regular evaporation measurements conducted in Finland have been conducted as pan evaporation, and only for the summer months. This prevents direct comparisons to the pet methods utilizing Penman-Monteith, and also seasonal dynamics outside of summer. According to our analysis, all the other changes are typically stepless, except changes between fmi and snow evaporation.

These are almost fully limited to springtime. There, the median change is  $1.8 \text{ mm d}^{-1}$  (3 x typical variation), and can be even larger. However, it should be noted that even though the change is too quick (one day), this sort of increase spread over slightly longer time, such as a week, could be somewhat realistic, since melting and evaporating snow 1) has high albedo 2) Prevents temperature from rising too much above zero and 3) increases the relative humidity of air above. After the snow has melted, temperature is free to rise and relative humidity is often quite low, leading to rapid increase in potential evaporation (Betts et al. 2001).

3.

- Comment: Section 4.1 Hydrologic time series. The manuscript could benefit from a more in-depth description of how the streamflow measurements are conducted by the data providers (SYKE and ELY), including how the measurements have changed over time (if such information is available), how the data were quality checked, and whether data quality flags are available and included. Some description is already present in Section 7, but I believe that it could be expanded. If the data providers have documentation describing how the measurements are or were conducted, it may be worth citing it in the manuscript.
- Action: Added data quality flags to the data. Revised text appropriately at section 3.1 and 6.1 of the revised article (4.1 and 7 in the preprint) by adding more detailed information on how the measurements were conducted and expanded quality descriptions.
- Reply: Thanks for suggesting the addition of quality flags! We added them to the dataset, as well as related remarks associated with some of the quality flags. We added information on the equipment used for the measurements to the text, along with the changes to those. We also added information on the quality check process. We were unable to find detailed citable documentation that was peer-reviewed or in English, but cited a relevant publication by SYKE, in Finnish. Large portions of the metadata provided are based on internal knowledge from SYKE, provided by Dr Jari Uusikivi (one of the authors), whose team is responsible for the observations.

4.

- Comment: I could not find geological indices. The manuscript mentions that hydrogeological data are not available. However, the manuscript could include other variables such as geological type, to align the data set more closely with the CAMELS data sets.
- Action: Added geological rock type to the dataset
- Reply: The original decision to exclude lithology was made based on discussions with two Finnish geology professors (Esa Heilamo and Antti Ojala), who estimated that the fracture density would be more relevant than the rock type, as most of Finland is covered in either metamorphic or plutonic rocks, which conduct water poorly outside of cracks. Unfortunately, no high quality crack density data covering the entire Finland exists. However, including the rock type data was relatively simple, so we decided to add that as suggested, in case somebody finds it useful. We would like to note that available geological data already exhibits large variations between different CAMELS datasets, so it is basically impossible to align CAMELS-FI to all of them. After deliberation, we decided that CAMELS-CH includes a rock type classification, that is similar to available rock type data for Finland, so we used that as a basis of inspiration.

### Minor comments

5.

- Comment: Fig. 1 and Section 5.6 (Human influence). Does the regulation referred to here indicate regulation through artificial dams? What are the criteria for defining the regulation classes (not regulated, minor and major regulation)? It is not clear in either the figure caption or the section.
- Action: clarified section 5.6 (4.7 in the revised manuscript)

- Reply: The active regulation is mostly conducted through dams, but there are also exceptions, such as adjustable flood levees, extractions, or permanent lowering of lake levels. It should be noted that catchments from class 1 (no active regulation) can still have dams, but they often only have the role of preventing the level of the body of water behind them from going below a certain level. Some small hydropower stations without a reservoir are also included in class 1 (after consultation with topic experts from SYKE), because their discharge is not dependent on regulation limits or electricity prices.

6.

- Comment: Table 1. The difference between gauge\_id and basin\_id is not clear. The gauge\_id is described as “catchment identifier”.
- Action: Added definition of a basin to line 154 of the revised manuscript (line 142 of the preprint)
- Reply: Basins are defined as all the locations that share an outlet to sea. Each gauge in CAMELS-FI has a corresponding catchment, and since their relationship is one-to-one, the gauge\_id can be used to refer to both the gauge and its catchment.

7.

- comment: Table 3. How was the attribute “regulation\_level” calculated?
- Action: The description was already revised based on the other referee’s comment
- Reply: It was classified based on the criteria laid out at lines 315 to 320 in the revised manuscript (287 to 291 in the preprint).

8.

- comment L128: “Kammonen, Luro (1358)”. Is 1358 the gauge\_id? It may be worth clarifying.
- Action: clarified
- Reply: yes, clarified.

## References

Betts, A. K., Ball, J. H., and McCaughey, J. H.: Near-surface climate in the boreal forest, *Journal of Geophysical Research: Atmospheres*, 106, 33529–33541, <https://doi.org/10.1029/2001JD900047>, 2001.