ESSD Submission by Coxon et al

CAMELS-GB: Hydrometeorological time series and landscape attributes for 671 catchments in Great Britain

General Response

We thank the reviewers for taking the time to review the paper and their comments, which have helped to improve this manuscript and the quality and clarity of the research.

The main comments from the reviewers focused on (1) improved accuracy of the meta-data descriptions, and (2) the addition of new datasets in CAMELS-GB (e.g. a river network, additional catchment attributes).

In response to these comments, we have revised the text to ensure all meta-data descriptions are accurate and added new clarifications on the limitations of some attributes/timeseries (for example, the representativeness of grassland PET for other land use types). We have also revised Figure 2 and added new Figures into the Supplementary Information following reviewers suggestions. We have also changed the method of download so users can access the data more easily.

While we welcome the suggestion of new datasets/attributes for CAMELS-GB, the additional datasets/attributes suggested were either not open access (so would be difficult to release as part of an open access dataset), or would be inconsistent with other attributes provided and require additional analysis (which is outside the scope of this study). The process of uploading the dataset to the EIDC took several months, hence we will wait until significant updates are necessary to add additional attributes/datasets.

Detailed responses to all comments are provided below. Author responses are in **bold** and any modifications to the manuscript are in *italic* below each of the reviewer's comments.

Kind Regards,

Gemma Coxon (on behalf of all co-authors), July 2020

Reviewer 1

The creation of the paper and dataset was motivated by the lack of having one consistent and comprehensive large sample hydro-meteorological dataset for Great Britain. As is outlined in the objectives of the paper, such a large sample dataset would be of great value for many different research purposes. The paper then describes how an impressive amount of data and meta data were combined into one single data set: CAMELS-GB. Furthermore, limitations of the different data and meta data sources were mentioned. The authors have put in some great effort to produce a very comprehensive hydrometeorological data and meta data set that will be a valuable resource for many hydrological studies and more.

The meta-data descriptions are elaborate but could sometimes be more accurate. The mentioning of the limitations of several meta-data sets at the end of different sections is useful. Below my suggestions and comments, which are presented in order of appearance in the paper as I do not have any major criticisms.

We thank Reviewer 1 for their positive assessment of the paper and their helpful comments. Please see our detailed responses below.

Line 31: You could add here that these discharge uncertainty estimates are made for different flow quantiles, which is a nice thing to have.

Agreed. We have modified the sentence to:

"Importantly, we also derive human management attributes (including attributes summarising abstractions, returns and reservoir capacity in each catchment), as well as attributes describing the quality of the flow data including the first set of discharge uncertainty estimates (provided at multiple flow quantiles) for Great Britain"

Line 71: are not allowed to instead of "cannot"

This has been changed in the manuscript

Line 85; I agree that "subject to change" negatively affects the repeatability of analyses. However, "subject to change" can also mean subject to improvement, for example in the quality of the streamflow records. The latter might speak in favor of sometimes directly using the most up to date data from the NRFA. A comment on that, and on whether it is planned to occasionally create a new version of the CAMELS-GB dataset, might be useful.

Flow timeseries are occasionally reprocessed when a rating curve has been revised (for example). We agree that this could mean an improvement in the quality of the flow data, however, these changes aren't documented online so it would be difficult to track this and would require the user to re-calculate all the hydrologic signatures (as these could change with new flow timeseries). For reproducibility purposes, we would suggest that users use the flow data contained in CAMELS-GB if they are to use the dataset, rather than re-downloading the dataset from the NRFA. We have made this clearer in Section 5.2 (see response to the next comment with the full changes to this section).

There are currently no plans to regularly update CAMELS-GB (simply compiling the dataset took over three years of work), however, we are keen to expand it in the future and will likely update the rest of the attributes/timeseries at this point. We have now made this clear in the conclusions.

"While a wealth of data is provided in CAMELS-GB, there are many opportunities to expand the dataset that were outside the scope of this study. Currently there are no plans to regularly update

CAMELS-GB, however, future work will concentrate on 1) expanding the dataset to include higher resolution data (such as hourly rainfall e.g. Lewis et al., 2018, and flow timeseries) and datasets for the analysis of trends (such as changes in land cover over time)"

Related to the latter comment; The dataset is fixed for a certain time period (1970-2015) for good reasons; however, some researchers might want to include some of the more recent events. In that case, they might either extent the CAMELS-GB data set with NRFA data or directly download complete time series of the NRFA. Provide a remark somewhere what the preferred option would be.

If users wanted to use an extended time period then they would likely also need to extend the meteorological timeseries and recalculate the climatic indices and hydrological signatures. In this case we would suggest that they directly download the complete time series from the NRFA. We have added a sentence to Section 5.2 to make this clear.

"If users wish to extend the timeseries beyond that available in CAMELS-GB, we suggest downloading and using the extended flow timeseries available from the NRFA website and re-calculating the hydrological signatures using the code we have archived."

Line 154: Add a note why it was not possible to derive suitable surface area for these catchments.

We have modified this sentence to include the reason for this.

"All gauges from the UK SLA network are included in CAMELS-GB except catchments from Northern Ireland (due to a lack of consistent meteorological datasets across the UK) and two gauges where no suitable surface area catchment could be derived (e.g. a groundwater spring for which surface catchment area is not hydrologically relevant)."

Line 177 (out of personal interest): Why were shapefiles transformed to ASCII grids? You could also overlay shapefiles and gridded data to derive (weighted) catchments averages. Or is this less accurate / consistent?

The code we implemented to derive the time series used catchment ascii grids to derive weighted catchment averages. Give the high resolution of the grids, we imagine this wouldn't have a significant impact on the timeseries but in future we would certainly look at overlaying the shapefiles on gridded data to avoid transforming it from shapefile to ascii.

Line 205: You could add a note here on human-induced non-stationarities, as you specifically included human influenced time series in the CAMELS-GB dataset.

Agreed. We have modified the sentence to:

"From previous analyses, it is important to note that there are key known non-stationarities over this period in hydro-meteorological data and human activity (see for example Hannaford and Marsh, 2006) for GB."

Lines 232-252: Great that both PET and PETI are included. Both products are derived for grassland, which is of course perfectly fine. However, a note on how representative grassland PET is for some of the catchments where e.g. forestry or agriculture are the dominant land use class (as shown in Figure 2) would be nice.

We have added the following to Section 5.1 to highlight this:

"It is important to note that the effect of seasonal land cover is not accounted for in the CHESS-PE products– this means that for arable agriculture which may have bare soil for part of the year, or deciduous trees which lose leaves and thus reduce both transpiration and interception, the potential evapotranspiration could be lower during winter than is estimated here. This leads to a varying

difference between the PET and PETI of grass and other land cover types throughout the year (Beven, 1979).”

Lines 232-252: How does the PET estimation method for CAMELS-GB compare to PET estimation methods of the other CAMELS datasets? Is it the aim here to use the best possible method that the data allows or to use a method that is consistent across all CAMELS datasets (which favors a comparison across datasets)?

As discussed in Section 2 of the paper, our priority for CAMELS-GB is to provide the best possible PET estimates for GB. We acknowledge that this may reduce the comparability with other CAMELS datasets (Addor et al., 2017; Alvarez-Garreton et al., 2018; Newman et al., 2015), which use different PET formulations relying on different atmospheric variables. CAMELS-US generated PET using the Priestly-Taylor method (Priestley and Taylor, 1972) and CAMELS-CL used temperature data and a formulation proposed by Hargreaves and Samani, (1985).

We are striving to increase the consistency among the CAMELS datasets (in terms of time series, catchment attributes, naming conventions and data format, see Addor et al., 2019), and to create a dataset that is globally consistent. We anticipate that this will happen as part of a second phase, which will build upon the current first phase, focussed on the release of national products, such as CAMELS-GB. We have added this to the conclusions to make this clear to the reader:

“Currently there are no plans to regularly update CAMELS-GB, however, future work will concentrate on 1) expanding the dataset to include higher resolution data (such as hourly rainfall e.g. Lewis et al., 2018, and flow timeseries) and datasets for the analysis of trends (such as changes in land cover over time), and 2) refining the characterisation of uncertainties in catchment attributes and forcing (particularly for rainfall data). We are also striving to increase the consistency among the CAMELS datasets (in terms of time series, catchment attributes, naming conventions and data format, see Addor et al., 2019), and to create a dataset that is globally consistent. We anticipate that this will happen as part of a second phase, which will build upon the current first phase that is focussed on the release of national products, such as CAMELS-GB.”

Line 264: “97% (654) of the gauges have at least 20 years” this cannot be seen in Figure 1a.

Sorry for the confusion, this was referring to results in Figure 1b not 1a. We have rewritten the sentence to make this clear.

“Figure 1a shows the flow data availability for all gauges contained in the CAMELS-GB dataset covering different time periods. Over the 46 year time period (1970 – 2015), 60% (401) of the gauges have 5% missing flow data or less and 81% (542) of the gauges have 20% missing flow data or less. Figure 1b shows the number of years of available flow data for each CAMELS-GB gauge across Great Britain. 97% (654) of the gauges have at least 20 years of data and 70% (468) of the gauges have at least 40 years of data.”

The comments below refer to either the section or the dataset that is described in this section:

Section 6.1: Provide a reason for some rare but substantial differences between gauge elevation and minimum elevation.

Gauge elevation is based on information from the originating measuring authorities (EA, SEPA, etc.), this may relate to either the gauge local datum or another point at the station, and may have been derived from various methods including contour maps and GPS devices (at different points in time, so some definitely not accurate to 10m). The minimum elevation is the minimum height of the IHDTM 50m grid cells used to define the catchment boundary. There may be

differences between the gauge elevation and catchment minimum elevations due to accuracies in the originating elevations sources and the accuracy of the 50m gridded representation of surface elevation. We have modified section 6.1 to make this clear.

"Catchment elevation is extracted from CEH's 50m Integrated Hydrological Digital Terrain Model and the minimum and maximum catchment elevation within the catchment mask is provided alongside different percentiles (10th, 50th and 90th). On occasion, minimum elevation may differ slightly from the gauge elevation attribute. The latter are as reported to the NRFA by the measuring authorities and derived in a variety of ways with different levels of accuracy. Furthermore some may refer to the bank top, the gauge minimum, or a local datum. The minimum elevation attribute provides a more consistent metric (though itself limited in accuracy due to the 50m grid representation)."

Section 6.1: Why do two catchments have NA values in their mean elevation, but do have values for e.g., min and max elevation. Please check.

For two catchments (18011 and 26006) where automatic derivation of the catchment boundary from the IHDTM for the gauge location was not possible and catchment masks were manually derived, no appropriate pre-computed values for the mean elevation or mean drainage path slope was available. We have added this clarification to Section 6.1 and specified in Section 4 that some of the catchment masks were manually derived.

"For two catchments (18011 and 26006) where automatic derivation of the catchment boundary from the IHDTM for the gauge location was not possible and catchment masks were manually derived, no appropriate pre-computed values for the mean elevation or mean drainage path slope was available."

Section 6.2: High and low prec timing; Instead of providing NAs for tied values, you could provide both seasons.

This would require making changes to the data hosted on the Environmental Information Data Centre server. The process of uploading the dataset to the EIDC took several months, hence we will wait until significant updates are necessary to make the changes suggested above.

Section 6.2: Definition of seasonality unclear. Provide a reference to the exact method.

We added a reference to the exact equation in Table 2.

"seasonality and timing of precipitation (estimated using sine curves to represent the annual temperature and precipitation cycles; positive (negative) values indicate that precipitation peaks in summer (winter) and values close to zero indicate uniform precipitation throughout the year). See equation 14 in (Woods, 2009)"

Section 6.2: Why an absolute definition for low precipitation frequency and a relative definition for high precipitation frequency (and why these thresholds)? Figure S4f makes sense according to the relative definition but is a bit counter intuitive.

The rationale is that the lower (absolute) threshold defines when a day is considered "dry" and is assumed to apply to all catchments (i.e. it is not location dependent). The higher (relative) threshold was selected to categorise "high precipitation events", a relative threshold was selected to account for the differences in precipitation regimes from one catchment to the next.

Section 6.2: The provided meta data could be extended with annual averages of the other meteorological variables, at least with average annual temperature.

We agree that we could have added a wide variety of additional meteorological attributes but we wanted to maintain consistency with the previous CAMELS datasets for the meteorological and hydrological attributes so will not be extending the attributes included.

Section 6.3: Provide a reference to the method used to calculate streamflow elasticity.

We added a reference to the exact equation in Table 2

"streamflow precipitation elasticity (sensitivity of streamflow to changes in precipitation at the annual timescale, using the mean daily discharge as reference). See equation 7 in (Sankarasubramanian et al., 2001), with the last element being $\bar{P}/\bar{Q}$ *not* $\bar{Q}/\bar{P}$*"*

Section 6.3: Good that two base flow indices are provided. I personally liked using the BFIHOST, but the latter is not directly derived from the streamflow record and therefore might not fit in this sub dataset. Might it fit somewhere else? Or was there another reason that it was excluded?

We also very much like using BFIHOST. However, as you rightly point out, it doesn't fit in the hydrological attributes (as it is derived primarily from soils data rather than streamflow data) and we decided not to include it elsewhere as the source data for BFIHOST are not open access.

Section 6.3: I think zero_q_freq is the fraction and not the percentage of time with zero flow. Please check. This might also explain why you have some NAs in the slope of the FDC. Please check as well.

You are correct – this is the fraction, not the percentage. We have changed the text in Table 2 to reflect this.

"fraction of days with Q = 0"

This is also the reason for some NAs in the slope of the FDC (as you suggested) and we updated Table 2 to make the user aware.

"slope of the flow duration curve (between the log-transformed 33rd and 66th streamflow percentiles). There can be NAs in this metric when over a third of the flow time series are zeros (see zero_q_freq)"

Section 6.3: Any reasoning / reference on why you chose 9 times median flow or 0.2 times mean flow as thresholds for high flow / low flow events?

We followed the definitions adopted by the previous CAMELS datasets. These thresholds were originally suggested by Clausen and Biggs (2000) and Westerberg and McMillan (2015). We added these two references to Table 2.

Section 6.5, line 332: Mention the depth range of the top soil layers.

Added.

"As this dataset only characterises the top soil layers (up to 1.3m)"

Section 6.5: Nice that you provide ranges of e.g. the tawc! Clarify in table 2 that tawc is calculated over the soil depth available for roots (if that is the case).

This is the case and has been added.

Section 6.7: Nice that discharge uncertainty estimates for different quantiles are provided!

Thank you!

Section 6.8: Weren't UKBN catchments also labeled as suitable for low- medium and high flow assessments? Why isn't this information included in the current data set?

UKBN catchments were labelled as suitable for low, medium and high flows. These data are available as open access here (https://nrfa.ceh.ac.uk/benchmark-network) so can be easily included as part of any analyses in conjunction with CAMELS-GB. As discussed above, we will be waiting for a significant update to CAMELS-GB before including additional attributes.

Section 6.8: I completely understand the uncertainties with regard to the human interventions, which are nicely outlined in the limitations. However, what should be commented on is that some of the benchmark catchments seem to be relatively heavily impacted by a human intervention of some sort, e.g., the occurrence of a significant amounts of abstractions or the presence of several reservoirs. As a user of the dataset, does this mean that I should interpret some of the benchmark catchments with care? Or that I should be extra careful when interpreting the abstraction and reservoir information?

Both the UKBN classifications and the abstractions information should be treated with care, as there are limitations in both.

As noted by Harrigan et al., (2018), the UKBN sought to be a 'best available' classification of human disturbances based on available information and expert judgment input from the gauging authorities. However, inevitably it is not perfect, and some compromises had to be made (i.e. some impacts tolerated) especially in the heavily impacted south and east of the UK, to ensure coverage in these regions (especially because good hydrometric data quality was also a key criteria in the UKBN selections so the pool of potential stations was smaller than in CAMELS-GB). In such otherwise sparsely covered areas, abstractions, discharges etc were sometimes tolerated on a case-by-case basis if (i) there was a pressing need for a catchment to fill a gap (either geographically or in terms of representativeness) and available information suggested they (ii) had a modest overall influence on flows or (iii) were known to be stable over time.

Similarly, the Artificial Influences dataset generated for CAMELS-GB is also not without limitations as noted in the paper (6.8.2 and 6.8.3). The dataset provides gross totals that point to the possible influence of abstractions (or reservoirs etc) but does not actually quantify the net influence of these impacts on the actual flow regime (unlike other artificial influence schemes, as discussed). It would be possible to have high potential influences in a catchment without them manifesting themselves as a major detectable influence on the streamflow regime. Moreover, all UK artificial influence datasets are subject to quality limitations, as outlined in these sections.

We have clarified this in Section 6.8.1 and 6.8.2.

Section 6.8: It might be useful to additional add the Factors affecting runoff codes, which are presented in the UK hydrometric register, as indicative information on the type of human influence that might have been present at some point in time in the catchment? Factors affecting runoff codes might also be highly uncertain, but together with the already presented data on abstractions and reservoirs, they might provide some additional clues on possible human influences.

The FAR classification was originally intended as a way of highlighting the presence of impacts that would affect the water balance in a catchment. While it can provide a crude guide to the presence of impacts, it does not give information on the extent of these impacts – nor does its absence indicate a lack of impacts. It was also created a long time ago and has not been routinely (nor systematically) updated. While it is another source of information, it is not one we would want to include in CAMELS-GB given the focus on the new impact datasets which are quantitative. However, we should have signposted this dataset in the manuscript and have now done so at the beginning of Section 6.8.

"We focused on providing quantitative data of human impacts in CAMELS-GB, however it is important to note that additional datasets are available that qualitatively characterise human impacts

in GB including the Factors Affecting Runoff (FAR) codes available from the National River Flow Archive."

Line 514-518: State that this is annual average precipitation. For me, it would be enough to just mention mm / year (and delete mm / day).

We have added it is an annual average but want to keep both figures as the numbers provided in the dataset are mm/day.

"The wettest areas of the UK are in mountainous regions with a maximum of 9.6 mm day-1 (annual average of 3500 mm year-1) in the north-west."

Line 524: Add space between number and unit (mm)

Changed.

Figure 1a: As you already have a second y-axis on the right, you might also consider adding a second x-axis on the top that indicates the accumulated amount of years with missing data.

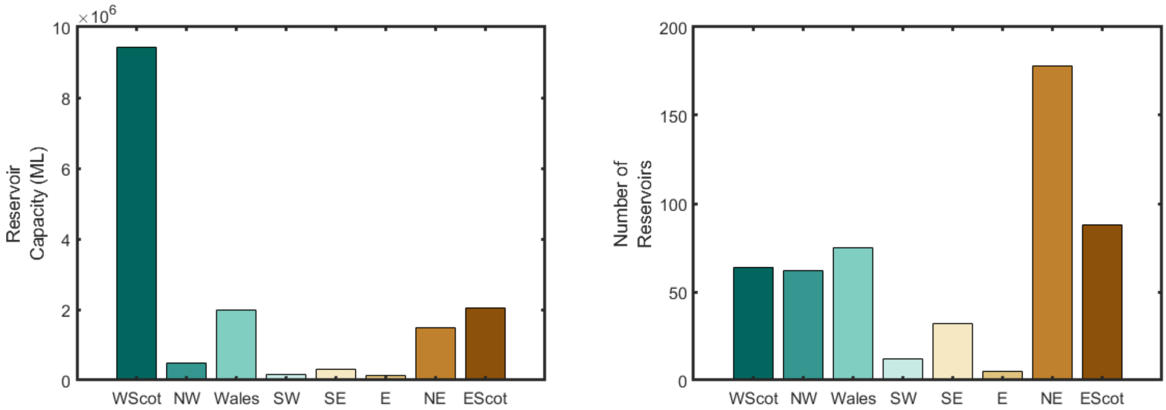
Thanks for the suggestion but we think this will make the plot too busy.

Figure 2e: The distribution of total reservoir volume might be more informative (although prob. not very different from the number of reservoirs).

Thanks, we have altered Figure 2e to display total reservoir capacity in each region and have altered the text in section 6.9.

"Large reservoir capacity is concentrated in the more mountainous northern and western regions of GB, particularly in Western Scotland (Figure 2e)."

For interest, the figure below shows the difference between total reservoir capacity per region and number of reservoirs. While the pattern is broadly the same (i.e. the northern and western regions have the largest number of reservoirs and total reservoir capacity), Western Scotland interestingly has fewer but larger reservoirs compared to the North-East.



Supplement (S2): Nice that all these maps are added. It would be helpful if they had titles, so you do not always need to read the caption. . .

Agreed. We will make this change in the revised manuscript.

Dataset: Clear and easy to process (in my case with R, but I am sure that this will be the same for other software). It would be nice to have one .zip file in the parent directory, which allows you to directly download all the time series data at once.

Agreed. The method of download has been changed so you can simply download a .zip file.

Reviewer 2

This study presents the first large-scale comprehensive hydrometeorological dataset for Great Britain. Authors synthetize the range of different data type (time and space support) from allied science fields into single, ready for use database in the well-known CAMEL format. The sources and structure of the data are well described; data aggregation procedures within the selected watersheds are specified. Comments on the possible limitations (quality and uncertainties) for the some variables are given. The format of the database is simple and self-describing. Manuscript is well structured and easy to read.

We thank Reviewer 2 for their positive assessment of the paper and their helpful comments. Please see our detailed responses below.

However, some critical comments must be noted:

1) In comparison with, CUAHSI ODM for instance, the CAMEL metadata schema is poorly developed. The database schema does not provide interoperation with data sources and feedback from community. There is no version control for observations, derived values and data series.

We have used the principal data centre for UK Freshwater research data (the Environmental Information Data Centre) to host the data. The EIDC provides DOIs but does not currently provide mechanisms for versioning or for community feedback. The data within CAMELS-GB are from primary data sources which are themselves versioned at the dataset level but do not provide information on versions or changes at the observation level.

It is important to note that CAMELS-GB (and the other CAMELS datasets) are the result of grass-root efforts led by individual hydrologists. Other initiatives supported by larger institutions and sustained funding rely on a more developed data management scheme, which we recognise the value of and find inspiring for future development stages. However, we would like to stress that no budget was available to produce and release CAMELS-GB. We anticipate that after this first phase, focussed on the release of national CAMELS datasets such as CAMELS-GB, we will be able to focus our efforts on increasing the consistency among the CAMELS datasets, as well as their interoperability and their data standards (see Addor et al., 2019). We have added a sentence to the conclusions to make this clear.

"We are also striving to increase the consistency among the CAMELS datasets (in terms of time series, catchment attributes, naming conventions and data format, see Addor et al., 2019), and to create a dataset that is globally consistent. We anticipate that this will happen as part of a second phase, which will build upon the current first phase that is focussed on the release of national products, such as CAMELS-GB."

2) Addition of a drainage network layer would facilitate navigation through the data and trace the hierarchy of nesting watersheds.

We agree that this would be a useful dataset to include. Currently, however, there are no open access, high-resolution river networks for Great Britain to share as part of the CAMELS-GB dataset so this is not possible.

3) Please give an explicit comment on data spatial aggregation - are nested watersheds areas been included or not.

We have now made this clear in Section 3 and added a new figure to the Supplementary Info showing the catchment areas.

“This results in a total of 671 catchments (includes nested catchments – see Supplement Fig S1) covering a wide range of climatic and hydrologic diversity across GB that is representative of the wider gauging network”

4) Since the database is essentially a set of text files, the use of the version control system (the GitHub, etc.) will minimizes efforts for support local copy DBs consistence in the future.

The EIDC, as described above, does not make use of a file-level version control system, but provides versioning at the dataset / DOI level only. We agree this would be useful for future versions of CAMELS-GB and stress that users can use GitHub on their systems to track changes in the dataset, which could be communicated and documented with new releases.

References

Addor, N., Newman, A. J., Mizukami, N. and Clark, M. P.: The CAMELS data set: catchment attributes and meteorology for large-sample studies, Hydrology and Earth System Sciences, 21(10), 5293–5313, doi:https://doi.org/10.5194/hess-21-5293-2017, 2017.

Addor, N., Do, H. X., Alvarez-Garreton, C., Coxon, G., Fowler, K. and Mendoza, P. A.: Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges, Hydrological Sciences Journal, 0(0), 1–14, doi:10.1080/02626667.2019.1683182, 2019.

Alvarez-Garreton, C., Mendoza, P. A., Boisier, J. P., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., Lara, A., Puelma, C., Cortes, G., Garreaud, R., McPhee, J. and Ayala, A.: The CAMELS-CL dataset: catchment attributes and meteorology for large sample studies – Chile dataset, Hydrology and Earth System Sciences, 22(11), 5817–5846, doi:https://doi.org/10.5194/hess-22-5817-2018, 2018.

Beven, K.: A sensitivity analysis of the Penman-Monteith actual evapotranspiration estimates, Journal of Hydrology, 44(3), 169–190, doi:10.1016/0022-1694(79)90130-6, 1979.

Hannaford, J. and Marsh, T.: An assessment of trends in UK runoff and low flows using a network of undisturbed catchments, International Journal of Climatology, 26(9), 1237–1253, doi:10.1002/joc.1303, 2006.

Hargreaves, G. and Samani, Z.: Reference Crop Evapotranspiration from Temperature, [online] Available from: https://doi.org/10.13031/2013.26773 (Accessed 14 July 2020), 1985.

Harrigan, S., Hannaford, J., Muchan, K. and Marsh, T. J.: Designation and trend analysis of the updated UK Benchmark Network of river flow stations: the UKBN2 dataset, Hydrology Research, 49(2), 552–567, doi:10.2166/nh.2017.058, 2018.

Lewis, E., Quinn, N., Blenkinsop, S., Fowler, H. J., Freer, J., Tanguy, M., Hitt, O., Coxon, G., Bates, P. and Woods, R.: A rule based quality control method for hourly rainfall data and a 1 km resolution gridded hourly rainfall dataset for Great Britain: CEH-GEAR1hr, Journal of Hydrology, 564, 930–943, doi:10.1016/j.jhydrol.2018.07.034, 2018.

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., Arnold, J. R., Hopson, T. and Duan, Q.: Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance,

Hydrology and Earth System Sciences, 19(1), 209–223, doi:https://doi.org/10.5194/hess-19-209-2015, 2015.

Priestley, C. H. B. and Taylor, R. J.: On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters, Mon. Wea. Rev., 100(2), 81–92, doi:10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2, 1972.

Sankarasubramanian, A., Vogel, R. M. and Limbrunner, J. F.: Climate elasticity of streamflow in the United States, Water Resources Research, 37(6), 1771–1781, doi:10.1029/2000WR900330, 2001.