



## **CAMELS-GB: A large sample, open-source, hydro-meteorological dataset for Great Britain**

Gemma Coxon (1), Nans Addor (2), John Bloomfield (3), Jim Freer (1), Matt Fry (4), Jamie Hannaford (4), Nicholas Howden (5), Chris Jackson (6), Emma Robinson (4), Thorsten Wagener (5), and Ross Woods (5)

(1) University of Bristol, Geographical Sciences, Bristol, United Kingdom (gemma.coxon@bristol.ac.uk), (2) Climatic Research Unit, University of East Anglia, Norwich, United Kingdom, (3) British Geological Survey, Wallingford, Oxfordshire, United Kingdom, (4) Centre for Ecology and Hydrology, Wallingford, United Kingdom, (5) Civil Engineering, University of Bristol, Bristol, United Kingdom, (6) British Geological Survey, Keyworth, Nottingham, United Kingdom

Data underpin our knowledge and understanding of the hydrological system; they are used to drive, test and evaluate hydrological models and advance our understanding of hydrological processes and dynamics. With the increasing availability of observational datasets, the integration of information from many catchments for data and modelling analyses is becoming increasingly common. The production of new, open source, datasets for large samples of catchments is vital to advance knowledge on hydrological processes, to underpin common frameworks for model evaluation across complex domains, and to ensure hydrological research is reusable and reproducible through the use of common datasets and code. However, the availability of open source, large-sample catchment datasets is notably sparse.

In this study, we present the first large sample, open-source, hydro-meteorological catchment dataset for Great Britain, CAMELS-GB (Catchment Attributes and MEteorology for Large-sample Studies). CAMELS-GB integrates a wealth of different spatial and time-varying datasets formulated from national, continental and global products based on observational, satellite and derived products. The dataset consists of hydro-meteorological timeseries, catchment attributes and catchment boundaries for >800 catchments that cover a wide range of climatic, hydrological, landscape and human management characteristics across Great Britain. Daily timeseries covering 1970-2015 are provided for a range of hydro-meteorological data (including rainfall, potential evapotranspiration, temperature, radiation, humidity and flow) and cover several major hydrological episodes (i.e. nationally important drought and flood periods). A comprehensive set of summary attributes are quantified describing a range of catchment characteristics including topography, climate, hydrology, land cover, soils and (hydro)-geology. Importantly, we also derive human impact attributes (including abstraction returns, percentage urban and gauge distance from reservoir), as well as attributes describing the quality of the flow data (including discharge uncertainty estimates and out of bank flow). This ensures users can identify, for example, human impacted catchments when evaluating models or analysing data. The dataset and code used to derive the data will be made open source and provided with comprehensive metadata to allow its use in a wide range of hydro-meteorological data and environmental modelling analyses. These data are part of a wider international initiative to generate multi-national, large-sample datasets to facilitate reproducible and reusable hydrological research.