Supporting information for "River flow simulations and derived metrics from a model intercomparison project covering 628+ British catchments, 1980-2019 (The GB Hydro-MIP dataset)."

Rosanna Lane, Emma Robinson, Helen Baron, Elizabeth Cooper, Victoria Bell, Stephen Birkinshaw, Jamie Brown, Burak Bulut, Steven Cole, Gemma Coxon, Brendan Dalmijn, Laura Devitt, Emma Ford, Nicole Forstenhäusler, Melike Kiraz-Safari, Eugene Magee, Robert Moore, Simon Moulds, Sopan Patil, Ponnambalam Rameshwaran, Ben Smith, John Wallbank, Albrecht Weerts, Yanchen Zheng

1 Summary

This dataset provides output from a community model intercomparison project (the GB Hydro-MIP) which evaluated modelled daily river flows from a range of machine-learning, hydrological, land-surface and machine-learning models across over 628 catchments in Great Britain. The core dataset includes: 1) river flow time-series from 31 model simulations, 5 lower benchmarks and 1 observational dataset, over the period 1980-2019 for up to 673 catchments, 2) performance metrics calculated over the calibration period (1980-1999), evaluation period (2000-2019) and full period for each river flow time-series for a core subset of 628 catchments with good data availability, and 3) hydrological signatures calculated over the calibration period, evaluation period and full period for each river flow time-series for the core 628 catchments. The models were run following the same protocol with consistent meteorological inputs. A further nine modelled river flow time-series are given for a subset of models covering 1980-2022 using a different set of driving data. The data are all provided as .csv files.

A research paper (Lane et al., The GB Hydro-MIP Dataset: River Flow Simulations and Evaluation Metrics for a Diverse Selection of Hydrological, Land-Surface, and Machine Learning Models across 628 Catchments in GB) describing the dataset will also be made available, which will be referred to as "the GB Hydro-MIP data paper" throughout this document. We strongly recommend that users also read the paper before using the dataset.

This document outlines the contents of the dataset, detailing the file structure and contents of each file, and briefly outlines specifics of the model applications. For more information about the participating models and limitations of this dataset please see the GB Hydro-MIP data paper.

2 Context

The GB Hydro-MIP arose from the recognised need for a community model intercomparison project (MIP) for UK hydrology, aimed at improving understanding of current modelling

capabilities and identifying the relative strengths and weaknesses of different modelling approaches. The project was developed through extensive community engagement, including workshops, protocol co-development and review, an open call for model submissions, webinars, conference presentations, a hackathon, and several collaborative research papers. Coordination of the MIP was carried out as part of the Hydro-JULES research programme under its cross-cutting theme on model benchmarking and uncertainty. The success of the GB Hydro-MIP reflects the contributions of the many researchers and organisations who provided model simulations, expertise, and feedback throughout the project.

For more information about the development of the MIP and the modelling protocol please see the GB Hydro-MIP data paper.

3 Contents of the dataset

3.1 Folders and file contents

The data are stored in 4 folders: streamflow_timeseries, streamflow_timeseries_extra, derived_metrics, derived_signatures. An additional file GBHMIP_gauge_list.csv listing basic catchment information about all the included river flow gauges is given in the main folder. A full list of filenames within each folder is provided in Table 1.

The folder streamflow_timeseries includes 37 .csv files: 1 file providing the observed streamflow time-series (GBHMIP_obs.csv), 5 files providing streamflow time-series for each benchmark (GBHMIP_bm_b*x*.csv), and 31 files providing streamflow time-series for each model (GBHMIP_mod_ *ModelName[_RunName]*.csv). Here, *x* is the number of the benchmark run (1-5), *ModelName* is a shortened model name, and the optional *_RunName* further describes the unique element of this run where the same model has been applied more than once.

The folder streamflow_timeseries_extra includes 10 files: 1 extended observed streamflow time-series and 9 additional streamflow time-series for models using a different forcing dataset and extended time periods (GBHMIPextra_mod_ *ModelName[_RunName]*.csv).

The folder derived_metrics includes 108 files containing the calculated metrics for each benchmark/model time-series (metrics*_period_ModelName[_RunName]*.csv). Here, *period* refers to the period over which the metrics were calculated, which is either cal (the calibration period, 1980-1999), eval (the evaluation period, 2000-2019), or wholets (the whole time-series, 1980-2019).

The folder derived_signatures includes 111 files: 108 files containing the calculated signatures for each benchmark/model time-series for each of the three periods (sigs*_period_ModelName[_RunName]*.csv), and 3 files containing the calculated signatures calculated for the observed time-series over each of the three periods (sigs*_period*_obs.csv).

There are also three supporting documents that may be used to understand and process the dataset alongside this supporting information. These include one file listing information about each model run (model_list.csv), one file listing information about each of the calculated

metrics (metric_list.csv), and one file listing information about each of the calculated signatures (signature_list.csv).

Table 1. Files included in the GB Hydro-MIP dataset. Words in grey italics within the filename(s) indicate the naming convention – please see the full text for details.

Folder	Filename(s)	File count	Description
streamflow_timeseries	GBHMIP_bm_b*x*.csv	5	Benchmark flow time-series across all 673 catchments
	GBHMIP_mod_*ModelName[_RunName]*.csv	31	Modelled river flow time-series across all 673 catchments, separate file for each model run submitted to the MIP
	GBHMIP_obs.csv	1	Observed river flow time-series
streamflow_timeseries_extra	GBHMIPextra_mod_*ModelName[_RunName]*.csv	9	Modelled river flow time-series that were provided in addition to the main MIP using different driving data
	GBHMIPextra_obs.csv	1	Observed river flow time-series covering an extended time-period
derived_metrics	metrics_*period_ModelName[_RunName]*.csv	108	Calculated performance metrics across 628 core catchments, separate file for each model run and time-period
derived_signatures	sigs_*period_ModelName[_RunName]*.csv	108	Calculated streamflow signatures across 628 core catchments, separate file for each model run and time-period
	sigs_*period*_obs.csv	3	Calculated streamflow signatures across 628 core catchments using observed data, separate file for each time-period

	GBHMIP_gauge_list.csv	1	A list of all MIP gauges with basic information (location, area, whether they are core catchments)

3.2 File formatting and variables

3.2.1 streamflow_timeseries and streamflow_timeseries_extra folders

Files containing streamflow time-series (filenames beginning GBHMIP_bm, GBHMIP_mod or GBHMIPextra_mod) have the following format. A table of simulated flows (m^3/s) for each of the 673 gauges, with -1 values indicating missing data. Different gauges are provided on each column, different days on each row. Column headings are YYYY, MM, DD, providing the date of each row, followed by each of the 673 gauging station numbers (https://nrfa.ceh.ac.uk/data/about-data/gauging-station-information/station-number). The run information provided alongside these files is given in light grey in section 6.3.

3.2.2 derived_metrics folder

This folder contains performance metrics calculated from simulated and observed river flows. The code to calculate these metrics has been made available on Zenodo (https://doi.org/10.5281/zenodo.23591728 (Lane, 2026)). Many are standard performance metrics used for hydrological modelling. Errors in hydrological signatures have been calculated as Sim_sig_value - Obs_sig_value and percentage errors in hydrological signatures have been calculated as ((Sim_sig_value - Obs_sig_value)/Obs_sig_value)*100. For these calculations, the observed and simulated signature values were calculated using only days when both observed and simulated flows were available for each specific model.

The metric_list.csv file includes a table of information about the metrics, with a row for each metric and nine columns (headername, short_name, long_name, units, lower_lim, upper_lim, optimal_value, Description, Reference).

Files containing derived metrics (filenames beginning metrics_) are formatted as tabular data, with a column for each metric and row for each gauge. The columns are fully described in Table 2. Some metrics could not be calculated for all models/gauges, and these fields have been left blank.

Table 2. Description of variables in metric files (metrics_period_ModelName[_RunName].csv).

Column header	Description	Units
gauge	Gauging station number	-
kge	Kling-Gupta Efficiency (KGE)	-
kge_invflows	KGE calculated using inverse flows	-
kge_np	KGE with non-parametric components	-
nse	Nash Sutcliffe Efficiency (NSE)	-
mae(mm/day)	Mean Absolute Error (MAE)	mm/day

mae(%)	Mean Absolute Percentage Error (MAE)	%
bias(mm/day)	Mean Bias	mm/day
bias(%)	Mean Percentage Bias	%
sd_err(%)	Error in flow Standard Deviation	%
hfv_err(%)	High Flow Volume Error (Q1-Q5)	%
mfv_err(%)	Median Flow Volume Error (Q33-Q66)	%
lfv_err(%)	Low Flow Volume Error (Q70-Q95)	%
pearson_r	Pearson Correlation Coefficient	-
spearman_r	Spearman's Rank Correlation Coefficient	-
sigerr_q5(mm/day)	High Flow (Q5) Bias	mm/day
sigerr_q95(mm/day)	Low Flow (Q95) Bias	mm/day
sigerr_freq_high(days)	High Flow Frequency Error	days
sigerr_freq_zero(days)	Zero Flow Frequency Error	days
sigerr_freq_low(days)	Low Flow Frequency Error	days
sigerr_dur_high(days)	High Flow Duration Error	days
sigerr_dur_low(days)	Low Flow Duration Error	days
sigerr_hfd(days)	Half Flow Date Error	days
sigerr_slope_fdc	Slope of the Flow Duration Curve Error	-
sigerr_q5(%)	High Flow (Q5) Percentage Bias	%
sigerr_q95(%)	Low Flow (Q95) Percentage Bias	%
sigerr_q_mean_djf(%)	Winter (DJF) Flow Bias	%
sigerr_q_mean_mam(%)	Spring (MAM) Flow Bias	%
sigerr_q_mean_jja(%)	Summer (JJA) Flow Bias	%
sigerr_q_mean_son(%)	Autumn (SON) Flow Bias	%

3.2.3 derived_signatures folder

This folder contains hydrological signatures calculated from simulated and observed river flows. These were largely based on code produced as part of the TOSSH toolbox (Gnann et al. (2021) https://doi.org/10.1016/j.envsoft.2021.104983) and code produced for CAMELS (Addor et al. https://github.com/naddor/camels). The GB Hydro-MIP code has been made available at https://doi.org/10.5281/zenodo.23591728.

The signature_list.csv file includes a table of information about the signatures, with a row for each signature and four columns (signature_headername, signature_name, units, Description).

Files containing derived streamflow signatures (filenames beginning sigs_) are formatted as tabular data, with a column for each signature and row for each gauge. The columns are fully described in Table 3.

Where signatures are presented in mm/day, the catchment areas given in GBHMIP_gauge_list.csv were used to convert volumetric discharge into specific discharge.

*Table 3. Description of variables in streamflow signature files (sigs_period_ModelName[_RunName].csv). *wetness_index is only provided in the observation-based file sigs_[period]_obs.csv.*

Column header	Description	Units
gauge	Gauging station number	-

q_mean	Mean flow	mm/day
q5	Q5 (High flow exceeded 5% of the time)	mm/day
q95	Q95 (Low flow exceeded 95% of the time)	mm/day
q_mean_djf	Mean winter flow (DJF)	mm/day
q_mean_mam	Mean spring flow (MAM)	mm/day
q_mean_jja	Mean summer flow (JJA)	mm/day
q_mean_son	Mean autumn flow (SON)	mm/day
q_7daymin	Minimum 7-day mean river flow	mm/day
cov	Coefficient of variation (CoV) of flow (i.e. standard deviation normalised by the mean)	-
freq_high	High flow frequency (number of high flow days per year) where high flow is defined as 9×q50	days
freq_zero	No flow frequency (number of no flow days per year)	days
freq_low	Low flow frequency (number of low flow days per year) where low flow is defined as 0.2×mean_flow	days
dur_zero	No flow duration (mean number of consecutive low flow days)	days
dur_high	High flow duration (mean number of consecutive high flow days) where high flow is defined as 9×q50	days
dur_low	Low flow duration (mean number of consecutive low flow days) where low flow is defined as 0.2×mean_flow	days
hfd	Mean half flow date (days since start of the (water) year on which the cumulative discharge reaches half of the annual discharge)	days
slope_fdc	Slope of the Flow duration curve between the log-transformed Q33 and Q66	-
runoff_ratio	Runoff Ratio (ratio of mean daily discharge to mean daily precip)	-
stream_elasticity	Streamflow precipitation elasticity (sensitivity of river flow to change in precipitation at the annual timescale)	-
bfi	BFI (baseflow index). BFI is the ratio between baseflow (volume) and total river flow (volume). We used the UKIH separation method.	-
wetness_index*	Wetness Index is the total precipitation divided by the total potential evapotranspiration.	-

4 Spatial coverage

Flow data are provided for 673 gauges across Great Britain, and derived metrics/signatures are provided for a ‘core’ subset of 628 of these catchments. These core gauges all had at least ten years of river flow observations in both the calibration and evaluation periods. Locations of all gauges are provided in GBHMIP_gauge_list.csv and shown in Figure 1.

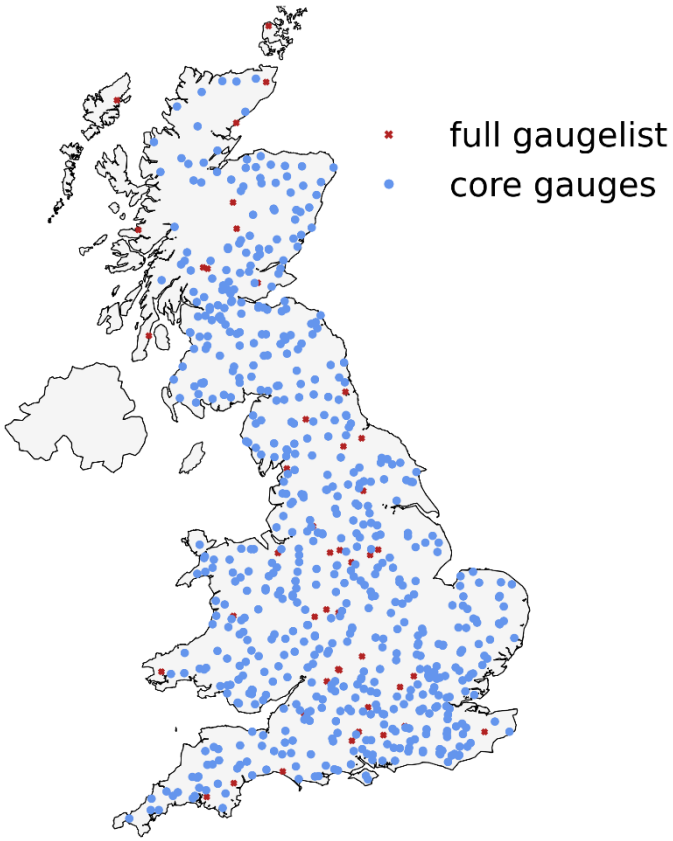

Figure 1. Spatial coverage of the GB Hydro-MIP dataset.

5 Temporal coverage and resolution

River flow time-series are provided at daily resolution over the period 1st January 1980 to 31st December 2019 for the main model intercomparison (streamflow_timeseries folder), and for an extended period of 1st January 1980 to 31st December 2022 for the additional simulations (streamflow_timeseries_extra folder). This time period was split equally into calibration and evaluation periods, with model calibration occurring within the period 1st January 1980 to 31st December 1999 and model evaluation being carried out over 1st January 2000 to 31st December 2019 (or 31st December 2022 in the case of the extra extended runs).

The derived metrics and derived signatures have been calculated over three separate time-periods: the calibration period, evaluation period and whole time-series.

6 Modelling

The models were run by different research groups, all following the GB Hydro-MIP modelling protocol. These all used the same driving data (see section 6.1), with the same temporal coverage (see section 4), but there were no restrictions placed on model spatial and temporal resolution, ancillary datasets, or calibration approach. For more details about the modelling, including descriptions of each model, please see the GB Hydro-MIP data paper. For more in depth descriptions of each modelling framework, model descriptions have been provided at https://doi.org/10.5281/zenodo.23945456 (Lane et al., 2026).

6.1 Input data

The models used consistent input datasets. For the main MIP, these were HadUK-Grid precipitation, CHESS-PE and CHESS-met (Met Office et al., 2024; Robinson, Blyth, et al., 2023b, 2023a). These datasets were available on a 1km grid or as catchment-average data produced as part of CAMELS-GB v2 (Coxon et al., 2026). For the optional extra simulations, input datasets were HadUK-Grid precipitation, Hydro-PE HadUK-Grid and Had-UK Grid temperature, also available as 1km gridded datasets or catchment-average time-series produced as part of CAMELS-GB v2 (Coxon et al., 2026; Met Office et al., 2024; Robinson, Brown, et al., 2023).

Other ancillary input datasets differed between models.

6.2 Models

*Table 4. Model simulations included in the GB Hydro-MIP dataset. * indicates model runs that have been provided with the optional extra inputs as well as the main MIP.*

Run name (in filename)	Model detail	Reference
bm_b1	Benchmark #1. Mean flow.	Knoben (2024)
bm_b2	Benchmark #2. Mean flow for each day of the year.	Knoben (2024)
bm_b3	Benchmark #3. Rainfall runoff ratio applied monthly.	Knoben (2024)
bm_b4	Benchmark #4. Adjusted precipitation benchmark (APB) model.	Knoben (2024)
bm_b5	Benchmark #5. Adjusted smoothed precipitation benchmark (ASPB) model.	Knoben (2024)
mod_LSTM	Long Short-Term Memory (LSTM) machine learning model trained with 22 static attributes and a standard 365-day lookback window. Trained on 1980-1994, validated on 1995-1999.	(Kratzert et al., 2018)
mod_LSTM_30daylookback	As mod_LSTM but with a shorter 30-day lookback window.	(Kratzert et al., 2018)
*mod_LSTM_ SpaceTimeSplit22attr	LSTM machine learning model trained with 22 static attributes using k-fold spatial cross validation. Trained on 1987-1999 and validated on 1980-1986.	(Kratzert et al., 2018)
*mod_LSTM_ SpaceTimeSplit45attr	as LSTM_SpaceTimeSplit22attr but an extended set of 45 static attributes.	(Kratzert et al., 2018)

*mod_LSTM_ TimeOnlySplit22attr	as LSTM_SpaceTimeSplit22attr but no spatial cross validation.	(Kratzert et al., 2018)
mod_EXPHYDRO	Exponential Bucket Hydrological model (EXP-HYDRO) lumped hydrological model.	(Patil & Stieglitz, 2014)
*mod_GR6J	GR6J lumped hydrological model.	(Pushpalatha et al., 2011)
*mod_HBV	HBV lumped hydrological model.	(Bergström, 1992)
*mod_HBV_NoSnow	HBV lumped hydrological model with the snow module switched off.	(Bergström, 1992)
mod_PDM	Probability Distributed Model (PDM) lumped hydrological model using the Grid-to-Grid snow module.	(Bell et al., 2016; Moore, 2007)
mod_PDM_NoSnow	Probability Distributed Model (PDM) lumped hydrological model.	(Moore, 2007)
mod_PycatchmodHUK	Pycathmod-HUK semi-distributed hydrological model (four hydrological response units).	(Wilby et al., 1994)
mod_PDM2PAR_RRMT	Rainfall-Runoff Modelling Toolbox (RRMT) lumped hydrological model combination (PDM+2PAR).	(Wagener et al., 2002)
mod_PDMCRES_RRMT	RRMT lumped hydrological model combination (PDM+CRES).	(Wagener et al., 2002)
mod_PDMLEAK_RRMT	RRMT lumped hydrological model combination (PDM+LEAK).	(Wagener et al., 2002)
mod_PEN2PAR_RRMT	RRMT lumped hydrological model combination (PEN+2PAR).	(Wagener et al., 2002)
mod_PENCRES_RRMT	RRMT lumped hydrological model combination (PEN+CRES).	(Wagener et al., 2002)
mod_PENLEAK_RRMT	RRMT lumped hydrological model combination (PEN+LEAK).	(Wagener et al., 2002)
mod_DECIPHeR_cat	DECIPHeR/ DECIPHeR-GW semi-distributed hydrological model (best calibration for each catchment).	(Coxon et al., 2019; Zheng et al., 2025)

mod_DECIPHeR_reg	DECIPHeR/ DECIPHeR-GW semi-distributed hydrological model (best calibration applied across each region).	(Coxon et al., 2019; Zheng et al., 2025)
*mod_G2G	Grid-to-Grid (G2G) distributed hydrological model with a temperature-based snow module.	(Bell et al., 2009, 2016)
*mod_G2G_Abs	Grid-to-Grid (G2G) – Anthropogenically influenced distributed hydrological model	(Rameshwaran et al., 2022)
*mod_G2G_NoSnow	Grid-to-Grid (G2G) distributed hydrological model without the optional snow module.	(Bell et al., 2009)
mod_HBV-TYN	HBV-TYN conceptual rainfall-runoff model, adapted to be fully distributed.	(He et al., 2022)
mod_SHETRAN_cal	SHETRAN physically-based distributed hydrological model, calibrated to each catchment.	(Lewis et al., 2018; Smith et al., 2024)
mod_SHETRAN_uncal	SHETRAN physically-based distributed hydrological model, uncalibrated.	(Lewis et al., 2018)
mod_WFLOW-SBM	wflow_sbm spatially distributed hydrological model.	(Ali et al., 2025; van Verseveld et al., 2024)
mod_WFLOW-SBM_BRT	wflow_sbm spatially distributed hydrological model.	(Ali et al., 2025; van Verseveld et al., 2024)
mod_WFLOW-SBM_wateruse	wflow_sbm spatially distributed hydrological model with an added water use component.	(Purnamasari et al., 2025; van Verseveld et al., 2024)
mod_JULES	Joint UK Land Environment Simulator (JULES) land-surface model with TOPMODEL runoff generation option.	(Best et al., 2011)
mod_JULESPDM	Joint UK Land Environment Simulator (JULES) land-surface model with PDM runoff generation option.	(Best et al., 2011)

6.3 Model run information – calibration, setup and missing values

Full descriptions of all the models are given in the GB Hydro-MIP data paper. This section provides specific details on the runs that were submitted to the MIP to expand upon Table 4, covering calibration approach, details of model setup and explanations for missing values. The modelling teams submitted metadata alongside each model run under the following 17 categories:

1. title
2. creator_name
3. creator_institution
4. creator_email
5. date_created
6. input_PE
7. input_precip
8. input_temperature
9. time_period_start
10. time_period_end
11. model_name
12. model_reference
13. model_spatial_resolution
14. model_timestep_minutes
15. model_calibration
16. model_setup
17. model_MIP_notes

This metadata is also provided below.

bm_bx.csv: provides benchmark flow simulations. The benchmarks are 1) mean flow, 2) daily mean flow, 3) flows estimated using monthly rainfall-runoff ratios, 4) the adjusted precipitation benchmark (APB), and 5) the adjusted smoothed precipitation benchmark (ASPB). All benchmarks were created following Knoben et al. (2024) https://doi.org/10.1002/hyp.15288 using the HydroBM Python package (https://github.com/wknoben/hydrobm). Benchmarks were defined using data over the calibration period (1980-1999) and used to predict over the evaluation period (2000-2019). They are all intended to be lower benchmarks, with b1-b2 demonstrating what can be achieved with river flow only and no knowledge about precipitation, b3 demonstrating what can be achieved given the additional knowledge of precipitation, and b4-b5 indicating what can be achieved with very parsimonious models. 6 catchments had insufficient data to calculate benchmarks b1,b3,b4 and b5, and 22 catchments had insufficient data to calculate b2, so these values have been set as nodata (-1). Note, some catchments have nodata values at the start of the simulation for b4 and b5 due to the precipitation time lag parameter.

```
================================================================
FILE: GBHMIP_bm_b1.csv
================================================================
title: mean flow benchmark
```

creator_name: Rosanna Lane
creator_institution: UK Centre for Ecology and Hydrology (UKCEH)
creator_email: roslan@ceh.ac.uk
date_created: 28/03/2026
input_PE: None
input_precip: None.
input_temperature: None.
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: b1_mean_flow
model_reference: Knoben et al. (2024) https://doi.org/10.1002/hyp.15288
model_spatial_resolution: lumped catchment.
model_timestep_minutes: N/A.
model_calibration: the benchmarks are defined using data over the calibration period (1980-1999) and used to predict over the evaluation period (2000-2019).
model_setup: benchmarks were created following Knoben et al. (2024) https://doi.org/10.1002/hyp.15288 using the HydroBM Python package (https://github.com/wknoben/hydrobm).
model_MIP_notes: This is intended as a lower benchmark for the UK Hydro-MIP dataset, as an indication of what can be achieved with streamflow only and no knowledge about precipitation. 6 catchments had insufficient data for this benchmark to be calculated and so values have been set as nodata (-1).

==========

FILE: GBHMIP_bm_b2.csv

==========

title: daily mean flow benchmark
creator_name: Rosanna Lane
creator_institution: UK Centre for Ecology and Hydrology (UKCEH)
creator_email: roslan@ceh.ac.uk
date_created: 28/03/2026
input_PE: None
input_precip: None.
input_temperature: None.
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: b2_daily_mean_flow
model_reference: Knoben et al. (2024) https://doi.org/10.1002/hyp.15288
model_spatial_resolution: lumped catchment.
model_timestep_minutes: N/A.
model_calibration: the benchmarks are defined using data over the calibration period (1980-1999) and used to predict over the evaluation period (2000-2019).
model_setup: benchmarks were created following Knoben et al. (2024) https://doi.org/10.1002/hyp.15288 using the HydroBM Python package (https://github.com/wknoben/hydrobm).
model_MIP_notes: This is intended as a lower benchmark for the UK Hydro-MIP dataset, as an indication of what can be achieved with streamflow only and no knowledge about

precipitation. 22 catchments had insufficient data for this benchmark to be calculated and so values have been set as nodata (-1).

==========

FILE: GBHMIP_bm_b3.csv

==========

title: rainfall runoff ratio to monthly benchmark
creator_name: Rosanna Lane
creator_institution: UK Centre for Ecology and Hydrology (UKCEH)
creator_email: roslan@ceh.ac.uk
date_created: 07/04/2026
input_PE: None
input_precip: HadUK-Grid.
input_temperature: CHESS-Met.
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: b3_rainfall-runoff_ratio
model_reference: Knoben et al. (2024) https://doi.org/10.1002/hyp.15288
model_spatial_resolution: lumped catchment.
model_timestep_minutes: 1440
model_calibration: the benchmarks are defined using data over the calibration period (1980-1999) and used to predict over the evaluation period (2000-2019).
model_setup: benchmarks were created following Knoben et al. (2024) https://doi.org/10.1002/hyp.15288 using the HydroBM Python package (https://github.com/wknoben/hydrobm).
model_MIP_notes: This is intended as a lower benchmark for the UK Hydro-MIP dataset, as an indication of what can be achieved with very parsimonious models. 6 catchments had insufficient data for this benchmark to be calculated and so values have been set as nodata (-1).

==========

FILE: GBHMIP_bm_b4.csv

==========

title: adjusted precipitation benchmark benchmark
creator_name: Rosanna Lane
creator_institution: UK Centre for Ecology and Hydrology (UKCEH)
creator_email: roslan@ceh.ac.uk
date_created: 28/03/2026
input_PE: None
input_precip: HadUK-Grid.
input_temperature: CHESS-Met.
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: b4_APB
model_reference: Knoben et al. (2024) https://doi.org/10.1002/hyp.15288
model_spatial_resolution: lumped catchment.
model_timestep_minutes: 1440

model_calibration: the benchmarks are defined using data over the calibration period (1980-1999) and used to predict over the evaluation period (2000-2019).
model_setup: benchmarks were created following Knoben et al. (2024) https://doi.org/10.1002/hyp.15288 using the HydroBM Python package (https://github.com/wknoben/hydrobm).
model_MIP_notes: This is intended as a lower benchmark for the UK Hydro-MIP dataset, as an indication of what can be achieved with very parsimonious models. 6 catchments had insufficient data for this benchmark to be calculated and so values have been set as nodata (-1). Note, some catchments have nodata values at the start of the simulation due to the precipitation time lag parameter.

==========
FILE: GBHMIP_bm_b5.csv
==========

title: adjusted smoothed precipitation benchmark benchmark
creator_name: Rosanna Lane
creator_institution: UK Centre for Ecology and Hydrology (UKCEH)
creator_email: roslan@ceh.ac.uk
date_created: 28/03/2026
input_PE: None
input_precip: HadUK-Grid.
input_temperature: CHESS-Met.
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: b5_ASPB
model_reference: Knoben et al. (2024) https://doi.org/10.1002/hyp.15288
model_spatial_resolution: lumped catchment.
model_timestep_minutes: 1440
model_calibration: the benchmarks are defined using data over the calibration period (1980-1999) and used to predict over the evaluation period (2000-2019).
model_setup: benchmarks were created following Knoben et al. (2024) https://doi.org/10.1002/hyp.15288 using the HydroBM Python package (https://github.com/wknoben/hydrobm).
model_MIP_notes: This is intended as a lower benchmark for the UK Hydro-MIP dataset, as an indication of what can be achieved with very parsimonious models. 6 catchments had insufficient data for this benchmark to be calculated and so values have been set as nodata (-1). Note, some catchments have nodata values at the start of the simulation due to the precipitation time lag parameter.

mod_DECIPHeR_cat.csv: provides flows from the DECIPHeR & DECIPHeR-GW models (Coxon et al., 2019; Zheng et al., 2025). DECIPHeR & DECIPHeR-GW was calibrated through running each catchment using 5000 different parameter sets. The best simulation was selected based on an equally weighted combination of KGE, NSE and logNSE for the calibration periods (1980-1999 inclusive). DECIPHeR and DECIPHeR-GW used the same Digital Terrain Analysis (DTA) and HRU delineation settings. The DECIPHeR model was set up for all 287 most downstream catchments to include all the 673 CAMELS-GB catchments. The DECIPHeR model does not account for groundwater processes, whereas the DECIPHeR-

GW model includes surface-groundwater interactions. Due to the lack of geological data for Scotland, DECIPHeR-GW was only constructed for England and Wales, covering a total of 240 most downstream catchments. Results provided here represent the best performance that DECIPHeR or DECIPHeR-GW can achieve for each catchment. Specifically, if DECIPHeR outperforms DECIPHeR-GW in a catchment, the DECIPHeR simulation is provided; otherwise, the DECIPHeR-GW simulation is provided. Neither model considers human activities, so the simulation results may be poor in catchments with heavily modified flow regimes. A total of 93 catchments were simulated using the DECIPHeR-GW model (full list provided in output file). Five catchments (25029,40022,42026,42027, 46014) have no observed data during the pre-2000 calibration period, so observation from the evaluation period were used to select the best simulation. Gauge 39125 was removed due to issues with the digital terrain analysis.

==========
FILE: GBHMIP_mod_DECIPHeR_cat.csv
==========
title: DECIPHeR & DECIPHeR-GW daily flow (m3 s-1) covering 1980-2019, generated for the UK Hydro-MIP
creator_name: Yanchen Zheng,Laura Devitt,Gemma Coxon
creator_institution: University of Bristol
creator_email: yanchen.zheng@bristol.ac.uk,laura.devitt@bristol.ac.uk,gemma.coxon@bristol.ac.uk
date_created: 16/12/2025
input_PE: CHESS PET
input_precip: HadUK-Grid
input_temperature: None
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: DECIPHeR & DECIPHeR-GW
model_reference: Coxon et al. (2019) DECIPHeR v1
model_spatial_resolution: catchment, HRU-based
model_timestep_minutes: 1440
model_calibration: DECIPHeR & DECIPHeR-GW is calibrated through running each catchment 5000 different parameter sets.We select our best simulation based on a combined consideration of KGE, NSE and logNSE for the calibration periods (1980-1999 inclusive).
model_setup: DECIPHeR and DECIPHeR-GW use the same Digital Terrain Analysis (DTA) and HRU delineation settings. The DECIPHeR model was set up for all 287 most downstream catchments to include all the 673 CAMELS-GB catchments. The DECIPHeR model does not account for groundwater processes, whereas the DECIPHeR-GW model includes surface-groundwater interactions.Due to the lack of geological data for Scotland, DECIPHeR-GW was only constructed for England and Wales, covering a total of 240 most downstream catchments.
model_MIP_notes: Our simulation results provide the best-calibrated outputs at both the catchment-by-catchment and river basin scales. The results we provide represent the best performance that DECIPHeR or DECIPHeR-GW can achieve for each catchment.Specifically, if DECIPHeR outperforms DECIPHeR-GW in a catchment, we provide the DECIPHeR simulation; otherwise, we provide the DECIPHeR-GW

simulation.Neither model considers human activities, so the simulation results may be poor in catchments with heavily modified flow regimes.A total of 93 catchments were simulated using the DECIPHeR-GW model, and the list is as follows.[17001,20007,26005,26006,27032,27038,27073,27080,28003,28039,28080,28115,28116,31025,32008,33022,33028,33033,34002,34003,34005,34008,34011,36003,37003,37005,37008,37009,37011,37018,37019,37031,37033,38003,38007,38012,38017,38021,38029,38030,39004,39005,39007,39011,39012,39014,39022,39027,39028,39029,39035,39037,39049,39056,39061,39078,39087,39095,39101,39114,39120,39125,39127,40012,40013,40016,40018,40021,41023,42001,42004,42008,42010,42016,42024,42027,44011,45003,45005,52004,54003,54004,54044,54052,54063,54096,55032,68020,69012,69032,73015,101002, 101005]. Five catchments (25029,40022,42026,42027, 46014) have no observed data during the pre-2000 calibration period, so observation from the evaluation period were used to select the best simulation. Gauge 39125 was removed due to issues with the digital terrain analysis.

mod_DECIPHeR_reg.csv: as above, but providing the best calibrated outputs at the river basin scale (rather than per catchment). Catchments 31025, 67010 and 67018 have zero values in the river-basin best simulation; the catchment-by-catchment results should be used for these gauges instead.

==========

FILE: GBHMIP_mod_DECIPHeR_reg.csv

==========

title: DECIPHeR & DECIPHeR-GW daily flow (m3 s-1) covering 1980-2019, generated for the UK Hydro-MIP
creator_name: Yanchen Zheng,Laura Devitt,Gemma Coxon
creator_institution: University of Bristol
creator_email: yanchen.zheng@bristol.ac.uk,laura.devitt@bristol.ac.uk,gemma.coxon@bristol.ac.uk
date_created: 16/12/2025
input_PE: CHESS PET
input_precip: HadUK-Grid
input_temperature: None
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: DECIPHeR & DECIPHeR-GW
model_reference: Coxon et al. (2019) DECIPHeR v1
model_spatial_resolution: catchment, HRU-based
model_timestep_minutes: 1440
model_calibration: DECIPHeR & DECIPHeR-GW is calibrated through running each catchment 5000 different parameter sets.We select our best simulation based on a combined consideration of KGE, NSE and logNSE for the calibration periods (1980-1999 inclusive).
model_setup: DECIPHeR and DECIPHeR-GW use the same Digital Terrain Analysis (DTA) and HRU delineation settings. The DECIPHeR model was set up for all 287 most downstream catchments to include all the 671 CAMELS-GB catchments and 2 FDRI catchments. The DECIPHeR model does not account for groundwater processes, whereas the DECIPHeR-GW model includes surface-groundwater interactions.Due to the lack of geological data for Scotland, DECIPHeR-GW was only constructed for England and Wales, covering a total of 240 most downstream catchments.

model_MIP_notes: Our simulation results provide the best-calibrated outputs at both the catchment-by-catchment and river basin scales. The results we provide represent the best performance that DECIPHeR or DECIPHeR-GW can achieve for each catchment.Specifically, if DECIPHeR outperforms DECIPHeR-GW in a catchment, we provide the DECIPHeR simulation; otherwise, we provide the DECIPHeR-GW simulation.Neither model considers human activities, so the simulation results may be poor in catchments with heavily modified flow regimes.A total of 84 catchments were simulated using the DECIPHeR-GW model, and the list is as follows.[17001,20007, 25020 Downsteam (25021 internal catchments),26005,31025, 32006 (32008),34002,34003,34005,34008,36003,37003, 37005. 37009,37018,37019,37031,37033,38021,39004,39005,39049,39056,39095,40016,40021, 41009 (41011,41022, 41027), 41014 (41019, 41025),41023,42001, 43007 (43009, 43018),44011,45003,45005,52004, 54057 (54001,54002,54003,54004,54005,54007,54008,54011,54012,54014,54015,54016,54018,54019,54020,54022,54024,54025,54028,54029,54032,54034,54036,54038,54044,54048,54049,54052,54060,54063,54080,54095, 54096),68020,69012,69032,73015,101002, 101005]. Catchments 31025, 67010 and 67018 have zero values in the river-basin best simulation; the catchment-by-catchment results should be used for these gauges instead. Gauge 39125 was removed due to issues with the digital terrain analysis.

mod_EXPHYDRO.csv: provides flows from the Exponential Bucket Hydrological model (EXP-HYDRO) (Patil & Stieglitz, 2014). The Particle Swarm Optimisation (PSO) algorithm was used with Kling-Gupta Efficiency (KGE) as the objective function over the period 1 January 1980 to 31 December 1999. Any days with missing data within the data range are ignored for KGE calculations. EXP-HYDRO was run at a spatially lumped resolution using catchment-averaged inputs. After calibration, the optimised model was run for the full period until 31 December 2019. EXP-HYDRO does not explicitly account for human modifications in a catchment. Therefore, it is not expected to perform well for catchments with heavily modified flow regimes.

==========

FILE: GBHMIP_mod_EXPHYDRO.csv

==========

title: Exponential Bucket Hydrological model (EXP-HYDRO) daily mean flows (m3 s-1) covering 1980-2019, generated for the UK Hydro-MIP.
creator_name: Sopan Patil
creator_institution: Bangor University
creator_email: s.d.patil@bangor.ac.uk
date_created: 08/09/2025
input_PE: CHESS-PET
input_precip: HadUK-Grid
input_temperature: CHESS-MET
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: Exponential Bucket Hydrological model (EXP-HYDRO)
model_reference: Patil and Stieglitz (2014) Modelling daily streamflow at ungauged catchments

model_spatial_resolution: Catchment (spatially lumped)
model_timestep_minutes: 1440
model_calibration: The Particle Swarm Optimisation (PSO) algorithm is used with Kling-Gupta Efficiency (KGE) as the objective function, and the data range from 01/01/1980 to 31/12/1999. Any days with missing data within the data range are ignored for KGE calculations.
model_setup: EXP-HYDRO is run at a spatially lumped resolution using catchment-averaged inputs. After calibration, the optimised model is run all the way until 31/12/2019.
model_MIP_notes: EXP-HYDRO does not explicitly account for human modifications in a catchment. Therefore, it is not expected to perform well for catchments with heavily modified flow regimes.

mod_G2G.csv: provides flows from the Grid-to-Grid (G2G) model (Bell et al., 2009) coupled with the temperature-based snow module (Bell et al., 2016). G2G is an uncalibrated model mainly parameterised using spatial datasets of landscape properties (e.g. HOST soil types). The model was run for each 1km grid square across Great Britain, and flows were extracted for grid squares which best relate to the locations of gauging stations. The snow module was included for these simulations but using average daily temperatures (rather than the standard daily min/max temperatures). G2G simulates naturalised flows on a 1km grid and is therefore not expected to match observations for catchments with heavily modified flow regimes or small areas.

==========
FILE: GBHMIP_mod_G2G.csv
==========
title: Grid-to-Grid daily mean flows (m3 s-1) using HadUK-Grid precip, CHESS PETi and CHESS-Met temperature inputs (produced for UK Hydro-MIP)
creator_name: Rosanna Lane
creator_institution: UK Centre for Ecology and Hydrology (UKCEH)
creator_email: roslan@ceh.ac.uk
date_created: 7/08/2025
input_PE: CHESS PETi
input_precip: HadUK-Grid
input_temperature: CHESS-Met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: Grid-to-Grid (G2G)
model_reference: Bell et al. (2009) Use of soil data in a grid-based hydrological model to estimate spatial variation in changing flood risk across the UK. JoH
model_spatial_resolution: 1km grid
model_timestep_minutes: 15
model_calibration: G2G is an uncalibrated model mainly parameterised using spatial datasets of landscape properties (e.g. HOST soil types).
model_setup: The model has been run for each 1km grid square across Great Britain, and flows have been extracted for grid squares which best relate to the locations of gauging stations. The snow module was included for these simulations, but using average daily temperatures (rather than the standard daily min/max temperatures).

model_MIP_notes: G2G simulates naturalised flows on a 1km grid and is therefore not expected to match observations for catchments with heavily modified flow regimes or small areas.

mod_G2G_NoSnow.csv: as for mod_G2G.csv, but without the inclusion of the optional snow module.

==========
FILE: GBHMIP_mod_G2G_NoSnow.csv
==========
title: Grid-to-Grid daily mean flows (m3 s-1) using HadUK-Grid precip and CHESS PETi inputs and no snow module (produced for UK Hydro-MIP)
creator_name: Rosanna Lane
creator_institution: UK Centre for Ecology and Hydrology (UKCEH)
creator_email: roslan@ceh.ac.uk
date_created: 7/08/2025
input_PE: CHESS PETi
input_precip: HadUK-Grid
input_temperature: None
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: Grid-to-Grid (G2G)
model_reference: Bell et al. (2009) Use of soil data in a grid-based hydrological model to estimate spatial variation in changing flood risk across the UK. JoH
model_spatial_resolution: 1km grid
model_timestep_minutes: 15
model_calibration: G2G is an uncalibrated model mainly parameterised using spatial datasets of landscape properties (e.g. HOST soil types).
model_setup: The model has been run for each 1km grid square across Great Britain, and flows have been extracted for grid squares which best relate to the locations of gauging stations. The snow module has not been included for these simulations.
model_MIP_notes: G2G simulates naturalised flows on a 1km grid and is therefore not expected to match observations for catchments with heavily modified flow regimes or small areas. The snow module was switched off for this simulation to demonstrate the impact of snow processes.

mod_G2G_Abs.csv: provides flows from Grid-to-Grid (G2G) - Anthropogenically influenced (Abstractions with Hands-Off Flow (HoF) conditions and discharges) (Rameshwaran et al., 2022). The model application otherwise followed mod_G2G.csv.

==========
FILE: GBHMIP_mod_G2G_Abs.csv
==========
title: Grid-to-Grid anthropogenically influenced daily mean flows (m3 s-1) using HadUK-Grid precip and CHESS PETi inputs (produced for UK Hydro-MIP)
creator_name: Ponnambalam Rameshwaran
creator_institution: UK Centre for Ecology and Hydrology (UKCEH)
creator_email: ponr@ceh.ac.uk
date_created: 22/01/2026

input_PE: CHESS PETi
input_precip: HadUK-Grid
input_temperature: CHESS-Met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: Grid-to-Grid (G2G) - Anthropogenically influenced (Abstractions and discharges with Hands-off-Flow (HoF))
model_reference: Rameshwaran, et al. (2022) Use of abstraction and discharge data to improve the performance of a national-scale hydrological model. https
model_spatial_resolution: 1km grid
model_timestep_minutes: 15
model_calibration: G2G is an uncalibrated model mainly parameterised using spatial datasets of landscape properties (e.g. HOST soil types).
model_setup: The model has been run for each 1km grid square across Great Britain, and flows have been extracted for grid squares which best relate to the locations of gauging stations. These simulations include abstractions/discharge data for England as well as the optional snow module.
model_MIP_notes: G2G simulates naturalised flows on a 1km grid and is therefore not expected to match observations for catchments with heavily modified flow regimes or small areas.

mod_GR6J.csv: provides flows from the GR6J (Genie Rural a 6 parametres Journalier) model (Pushpalatha et al., 2011). Calibration algorithm that optimises the error criterion (kge2) selected as objective function using the Irstea procedure described by C. Michel. Missing data for some stations due to unavailability of observed flows (less than 10% available in the calibration period 1980-1999).

==========
FILE: GBHMIP_mod_GR6J.csv
==========

title: GR6J daily mean flows (m3 s-1) generated for the Hydro-JULES MIP
creator_name: Burak Bulut & Eugene Magee
creator_institution: UK Centre for Ecology & Hydrology
creator_email: BurBul@ceh.ac.uk & EugMag@ceh.ac.uk
date_created: 15/10/2025
input_PE: CHESS-PETI
input_precip: HADUK
input_temperature: None
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: GR6J Genie Rural a 6 parametres Journalier
model_reference: Pushpalatha et al. (2011). A downward structural sensitivity analysis of hydrological models to improve low-flow simulation. Journal of hydrology 411(1-2) 66-76.
model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: Calibration algorithm that optimises the error criterion (kge2) selected as objective function using the Irstea procedure described by C. Michel
model_setup: catchment

model_MIP_notes: Missing data for stations due to unavailability of observed flows (less than 10% available in the calibration period 1980-1999).

mod_HBV.csv: provides flow from the HBV (Hydrologiska Byrans Vattenbalansavdelning) model (Bergström, 1992). Particle Swarm Optimization (PSO) was used with up to 50 iterations and 2000 function evaluations including a hybrid local refinement - the error criterion (kge2). Missing data for some stations due to unavailability of observed flows (less than 10% available in the calibration period 1980-1999).

==========
FILE: GBHMIP_mod_HBV.csv
==========
title: HBV daily mean flows (m3 s-1) generated for the Hydro-JULES MIP
creator_name: Burak Bulut
creator_institution: UK Centre for Ecology & Hydrology
creator_email: BurBul@ceh.ac.uk
date_created: 15/10/2025
input_PE: CHESS-PETI
input_precip: HADUK
input_temperature: CHESS
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: HBV (Hydrologiska Byrans Vattenbalansavdelning)
model_reference: Bergstrom S. (1995) The HBV Model. In
model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: Particle Swarm Optimization (PSO) was used with up to 50 iterations and 2000 function evaluations including a hybrid local refinement - the error criterion (kge2)
model_setup: catchment - snow module ON
model_MIP_notes: Missing data for stations due to unavailability of observed flows (less than 10% available in the calibration period 1980-1999).

HBV_NoSnow.csv: as above, but without the snow module.

==========
FILE: GBHMIP_mod_HBV_NoSnow.csv
==========
title: HBV daily mean flows (m3 s-1) generated for the Hydro-JULES MIP
creator_name: Burak Bulut
creator_institution: UK Centre for Ecology & Hydrology
creator_email: BurBul@ceh.ac.uk
date_created: 15/10/2025
input_PE: CHESS-PETI
input_precip: HADUK
input_temperature: CHESS
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: HBV (Hydrologiska Byrans Vattenbalansavdelning)
model_reference: Bergstrom S. (1995) The HBV Model. In

model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: Particle Swarm Optimization (PSO) was used with up to 50 iterations and 2000 function evaluations including a hybrid local refinement - the error criterion (kge2)
model_setup: catchment - snow module OFF
model_MIP_notes: Missing data for stations due to unavailability of observed flows (less than 10% available in the calibration period 1980-1999).

HBV-TYN.csv: provides flows from the HBV-TYN model (He et al., 2022). Each catchment was calibrated independently using a Monte Carlo approach with 50,000 uniformly sampled parameter sets. Simulations included a one-year warm-up (1979). Calibration optimised daily-flow Nash–Sutcliffe Efficiency (NSE) over the 1980–1999 period, using predefined plausible parameter ranges. The optimal parameter set for each catchment was selected based on the highest NSE. The model was implemented at 1km spatial resolution across the full catchment area. Flow time-series were derived from the grid cells that best represent the positions of the respective gauging stations. The model simulates naturalised flows on a 1km grid and is therefore not expected to reproduce observed discharge in catchments with heavily modified flow regimes or in small catchments where local interventions strongly influence flows. Simulated discharge for catchments 025029, 040022, 042026, 042027, and 046014 is not provided because calibration could not be performed due to insufficient observed discharge data within the calibration period. Simulated discharge for catchments 021014 and 054014 is not included because these catchments have not yet been incorporated into the model domain.

==========
FILE: GBHMIP_mod_HBV-TYN.csv
==========
title: HBV-TYN daily mean flows (m3 s-1) covering 1980-2019, generated for the UK HydroMIP.
creator_name: Nicole Forstenhaeusler
creator_institution: Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia
creator_email: n.forstenhaeusler@uea.ac.uk
date_created: 29/01/2026
input_PE: CHESS PET
input_precip: HadUK-Grid
input_temperature: CHESS-met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: HBV-TYN
model_reference: He, Y., Manful, D., Warren, R., Forstenhäusler, N., Osborn, T.J., Price, J., Jenkins, R., Wallace, C. and Yamazaki, D., 2022. Quantification of impacts between 1.5 and 4 C of global warming on flooding risks in six countries. Climatic Change, 170(1), p.15.
model_spatial_resolution: 1km grid
model_timestep_minutes: 1440
model_calibration: Each catchment was calibrated independently using a Monte Carlo approach with 50,000 uniformly sampled parameter sets. Simulations included a one-year

warm-up (1979). Calibration optimised daily-flow Nash–Sutcliffe Efficiency (NSE) over the 1980–1999 period, using predefined plausible parameter ranges. The optimal parameter set for each catchment was selected based on the highest NSE
model_setup: The model was implemented at 1km spatial resolution across the full catchment area. Flow time series were derived from the grid cells that best represent the positions of the respective gauging stations
model_MIP_notes: The model simulates naturalised flows on a 1km grid and is therefore not expected to reproduce observed discharge in catchments with heavily modified flow regimes or in small catchments where local interventions strongly influence flows. Simulated discharge for catchments 025029, 040022, 042026, 042027, and 046014 is not provided because calibration could not be performed due to insufficient observed discharge data within the calibration period. Simulated discharge for catchments 021014 and 054014 is not included because these catchments have not yet been incorporated into the model domain.

mod_JULES.csv: provides flows for the Joint UK Land Environment Simulator (JULES) model (Best et al., 2011). JULES is an uncalibrated model. The configuration used was based on RAL3 (Bush et al., 2025), which includes the TOPMODEL runoff generation option. The model has been run for each 1km grid square across Great Britain, and flows have been extracted for grid squares which best relate to the locations of gauging stations. The river code is called every JULES model timestep (i.e. 30 mins). The rose suite used is u-dy491at revision r357855, this is stored in the Met Office Science Repository Service. JULES simulates naturalised flows on a 1km grid and is therefore not expected to match observations for catchments with heavily modified flow regimes or small areas. Gauge 31002 doesn't match a sensible location and was excluded from the analysis.

==========
FILE: GBHMIP_mod_JULES.csv
==========
title: JULES daily mean flows (m3 s-1) using HadUK_Grid precip and CHESS for all other meteorological inputs
creator_name: Emma Robinson
creator_institution: UK Centre for Ecology and Hydrology (UKCEH)
creator_email: emrobi@ceh.ac.uk
date_created: 16/07/2026
input_PE: calculated within JULES
input_precip: HadUK-Grid
input_temperature: CHESS-Met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: JULES
model_reference: https://gmd.copernicus.org/articles/4/677/2011/gmd-4-677-2011-discussion.html
model_spatial_resolution: 1km grid
model_timestep_minutes: 30
model_calibration: JULES is an uncalibrated model
model_setup: The model has been run for each 1km grid square across Great Britain, and flows have been extracted for grid squares which best relate to the locations of gauging

stations. The river code is called every JULES model timestep (i.e. 30 mins) The rose suite used is u-dy491 at revision r357855. The difference between this and the RAL3PDM run is that this version uses TOPMODEL rather than the PDM saturation excess option
model_MIP_notes: JULES simulates naturalised flows on a 1km grid and is therefore not expected to match observations for catchments with heavily modified flow regimes or small areas. Gauge 31002 doesn't match a sensible location so has been excluded from the analysis.

mod_JULESPDM.csv: as above, but this version uses the PDM runoff generation option, rather than TOPMODEL. The rose suite used is u-dy491 at revision r356416 for 1980-1984 inclusive, and r357608 for the rest of the run.

==========
FILE: GBHMIP_mod_JULES+PDM.csv
==========

title: JULES daily mean flows (m3 s-1) using HadUK_Grid precip and CHESS for all other meteorological inputs
creator_name: Emma Robinson
creator_institution: UK Centre for Ecology and Hydrology (UKCEH)
creator_email: emrobi@ceh.ac.uk
date_created: 16/07/2026
input_PE: calculated within JULES
input_precip: HadUK-Grid
input_temperature: CHESS-Met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: JULES
model_reference: https://gmd.copernicus.org/articles/4/677/2011/gmd-4-677-2011-discussion.html
model_spatial_resolution: 1km grid
model_timestep_minutes: 30
model_calibration: JULES is an uncalibrated model
model_setup: The model has been run for each 1km grid square across Great Britain, and flows have been extracted for grid squares which best relate to the locations of gauging stations. The river code is called every JULES model timestep (i.e. 30 mins) The rose suite used is u-dy491 at revision r357608. The difference between this and the RAL3 run is that this version uses the PDM saturation excess option, rather than TOPMODEL
model_MIP_notes: JULES simulates naturalised flows on a 1km grid and is therefore not expected to match observations for catchments with heavily modified flow regimes or small areas. Gauge 31002 doesn't match a sensible location so has been excluded from the analysis.

mod_LSTM.csv: provides flows from the LSTM model using the NeuralHydrology framework (CudaLSTM). Global training with NSE loss; best-validation-NSE checkpoint exported. 365-day lookback LSTM; 22 static CAMELS-GB attributes concatenated with dynamic inputs. Daily discharge_vol simulations for CAMELS-GB basins over the defined export period. CAMELS-GB data did not include 21014 and 54014, so flows could not be given for these gauges. Training with fixed temporal splits (no spatial split). The model output some negative flow values - these have been set to 0.

==========
FILE: GBHMIP_mod_LSTM.csv
==========

title: NeuralHydrology LSTM daily simulated flows (m3 s-1)
creator_name: Emma Ford
creator_institution: University of Oxford
creator_email: emma.ford@hertford.ox.ac.uk
date_created: 07/10/2026
input_PE: CHESS-PE
input_precip: HadUK-Grid
input_temperature: CHESS-Met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: NeuralHydrology LSTM
model_reference: Kratzert et al. (2018; 2019) LSTM rainfall-runoff modelling; NeuralHydrology framework (CudaLSTM); configuration follows Lees et al. (2021).
model_spatial_resolution: Catchment-average forcings per CAMELS-GB basin
model_timestep_minutes: 1440
model_calibration: Global training with NSE loss; 30 epochs fixed in advance following Lees et al. (2021); final epoch exported with no checkpoint selection. Temporal splitting
model_setup: 365-day lookback LSTM; hidden size 64; 22 static CAMELS-GB attributes concatenated with dynamic inputs.
model_MIP_notes: Daily simulations for CAMELS-GB basins over the defined export period; trained on specific discharge (mm/day) and converted to m3 s-1 using CAMELS-GB catchment area. Values before 01/01/2000 are in-sample (training and validation periods) and should not be used for evaluation. CAMELS-GB data did not include 21014 and 54014 so flows could not be given for these gauges. Training with fixed temporal splits (no spatial split). 17 gauges had no observed flow in the training period and were not trained on.

mod_LSTM_30daylookback.csv: as above but with a shorter 30-day lookback period. This is expected to lead to reduced performance but provides an interesting example of LSTM sensitivity to lookback-period length.

==========
FILE: GBHMIP_mod_LSTM_30daylookback.csv
==========

title: NeuralHydrology LSTM daily simulated flows (m3 s-1)
creator_name: Emma Ford
creator_institution: University of Oxford
creator_email: emma.ford@hertford.ox.ac.uk
date_created: 18/03/2026
input_PE: CHESS PETi
input_precip: HadUK-Grid
input_temperature: CHESS-Met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: NeuralHydrology LSTM

model_reference: Kratzert et al. (2018, 2019) LSTM rainfall–runoff modelling; NeuralHydrology framework (CudaLSTM).
model_spatial_resolution: Catchment-average forcings per CAMELS-GB basin
model_timestep_minutes: 1440
model_calibration: Global training with NSE loss; best-validation-NSE checkpoint exported (epoch 15). Temporal splitting
model_setup: 30-day lookback LSTM; 22 static CAMELS-GB attributes concatenated with dynamic inputs.
model_MIP_notes: Daily discharge_vol simulations for CAMELS-GB basins over the defined export period. CAMELS-GB data did not include 21014 and 54014, so flows could not be given for these gauges. Training with fixed temporal splits (no spatial split). The model output some negative flow values - these have been set to 0.

mod_LSTM_SpaceTimeSplit22attr.csv: provides flows from the Long Short-Term Memory (LSTM) model (Kratzert et al., 2018; Lees et al., 2021). It was calibrated using k-fold (k=6) spatial cross validation between 1987-1999. A single LSTM was trained on all catchments in k-1 folds over the calibration period and used to predict discharge in the left-out fold over the test period. This LSTM model used 22 catchment attributes equivalent to those used by Lees et al. (2021). The model outputs some negative flow values - these have been set to 0. CAMELS-GB data did not include 21014 and 54014, so flows could not be given for these gauges.

==========
FILE: GBHMIP_mod_LSTM_SpaceTimeSplit22attr.csv
==========

title: Long Short-Term Memory (LSTM) daily mean flows (m3 s-1) covering 1980-2019, generated for the UK Hydro-MIP
creator_name: Simon Moulds
creator_institution: School of GeoSciences, University of Edinburgh, UK
creator_email: simon.moulds@ed.ac.uk
date_created: 02/12/2025
input_PE: CHESS-PET
input_precip: HadUK-Grid
input_temperature: CHESS-met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: LSTM
model_reference: Kratzert et al. (2018, 2019, 2022); Lees et al. (2021)
model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: Calibrated using k-fold (k=6) spatial cross validation between 1987-1999
model_setup: A single LSTM was trained on all catchments in k-1 folds over the calibration period and used to predict discharge in the left out fold over the test period. This LSTM model used 22 catchment attributes equivalent to those used by Lees et al. (2021).
model_MIP_notes: The model output some negative flow values - these have been set to 0.

mod_LSTM_ SpaceTimeSplit45attr.csv: as above, but using 45 catchment attributes.

==========
FILE: GBHMIP_mod_LSTM_SpaceTimeSplit45attr.csv
==========
title: Long Short-Term Memory (LSTM) daily mean flows (m3 s-1) covering 1980-2019, generated for the UK Hydro-MIP
creator_name: Simon Moulds
creator_institution: School of GeoSciences, University of Edinburgh, UK
creator_email: simon.moulds@ed.ac.uk
date_created: 02/12/2025
input_PE: CHESS-PET
input_precip: HadUK-Grid
input_temperature: CHESS-met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: LSTM
model_reference: Kratzert et al. (2018, 2019, 2022)
model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: Calibrated using k-fold (k=6) spatial cross validation between 1987-1999
model_setup: A single LSTM was trained on all catchments in k-1 folds over the calibration period and used to predict discharge in the left out fold over the test period. This LSTM model used 45 catchment attributes.
model_MIP_notes: The model output some negative flow values - these have been set to 0.

mod_LSTM_ TimeOnlySplit22attr.csv: as above, but trained using temporal split only (no spatial cross validation).

==========
FILE: GBHMIP_mod_LSTM_TimeOnlySplit22attr.csv
==========
title: Long Short-Term Memory (LSTM) daily mean flows (m3 s-1) covering 1980-2019, generated for the UK Hydro-MIP
creator_name: Simon Moulds
creator_institution: School of GeoSciences, University of Edinburgh, UK
creator_email: simon.moulds@ed.ac.uk
date_created: 06/10/2026
input_PE: CHESS-PET
input_precip: HadUK-Grid
input_temperature: CHESS-met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: LSTM
model_reference: Kratzert et al. (2018, 2022); Lees et al. (2021)
model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: Calibrated using k-fold (k=6) spatial cross validation between 1987-1999

model_setup: A single LSTM was trained on all catchments in k-1 folds over the calibration period and used to predict discharge in the left out fold over the test period. This LSTM model used 45 catchment attributes.
model_MIP_notes: The model output some negative flow values - these have been set to 0.

mod_PDM.csv: provides flows from the Probability Distributed Model (PDM) (Moore, 2007) using the Grid-to-Grid snow module (Bell et al., 2016). A variety of model structures are calibrated for each catchment for the period 1980–1999 inclusive to optimise the modified Kling-Gupta Efficiency (KGE') applied to square-root transformed flows (KGE'[sqrt]) and achieve zero bias. Final selection among model structures primarily optimises 0.8×KGE'[sqrt] + 0.2×NSE[log] but with a preference given to simpler model structures and those deemed more hydrologically plausible for the particular catchment. HadUK-Grid precipitation was pre-processed at a 1km spatial resolution using the G2G snow module. It was catchment-averaged for use in all stages of PDM modelling (including calibration). Flows are provided for the 628 core gauges only.

==========
FILE: GBHMIP_mod_PDM.csv
==========

title: Probability Distributed Model (PDM) daily mean flows (m^3/s) with snow modelling covering 1980-2019 and generated for the UK HydroMIP.
creator_name: John Wallbank, Robert Moore, Steven Cole
creator_institution: UK Centre for Ecology & Hydrology (UKCEH)
creator_email: johwal@ceh.ac.uk
date_created: 31/12/2025
input_PE: CHESS-PETI
input_precip: HadUK-Grid
input_temperature: CHESS-met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: Probability Distributed Model (PDM)
model_reference: Moore, R. J. (2007) The PDM rainfall-runoff model. Hydrol. Earth Syst. Sci., 11, 483–499. https
model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: A variety of model structures are calibrated for each catchment for the period 1980 – 1999 inclusive to optimise the modified Kling-Gupta Efficiency (KGE') applied to square-root transformed flows (KGE'[sqrt]) and achieve zero bias. Final selection among model structures primarily optimises 0.8*KGE'[sqrt] + 0.2*NSE[log] but with a preference given to simpler model structures and those deemed more hydrologically plausible for the particular catchment.
model_setup: Snow processes included. HadUK-Grid precipitation was pre-processed at a 1km spatial resolution using the Grid-to-Grid (G2G) snow module. It was catchment-averaged for use in all stages of PDM modelling (including calibration).
model_MIP_notes: valid flows at 628 core gauges only

mod_PDM_NoSnow.csv: as above, but without inclusion of snow processes. Catchment-average HadUK precipitation from the CAMELS-GB v2 dataset was used directly for all stages of the modelling (including calibration).

==========

FILE: GBHMIP_mod_PDM_NoSnow.csv

==========

title: Probability Distributed Model (PDM) daily mean flows (m^3/s) without snow modelling covering 1980-2019 and generated for the UK HydroMIP.
creator_name: John Wallbank, Robert Moore, Steven Cole
creator_institution: UK Centre for Ecology & Hydrology (UKCEH)
creator_email: johwal@ceh.ac.uk
date_created: 31/11/2025
input_PE: CHESS-PETI
input_precip: HadUK-Grid
input_temperature: None
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: Probability Distributed Model (PDM)
model_reference: Moore, R. J. (2007) The PDM rainfall-runoff model. Hydrol. Earth Syst. Sci., 11, 483–499. https
model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: A variety of model structures are calibrated for each catchment for the period 1980 – 1999 inclusive to optimise the modified Kling-Gupta Efficiency (KGE') applied to square-root transformed flows (KGE'[sqrt]) and achieve zero bias. Final selection among model structures primarily optimises 0.8*KGE'[sqrt] + 0.2*NSE[log] but with a preference given to simpler model structures and those deemed more hydrologically plausible for the particular catchment.
model_setup: Snow processes not included. Catchment average HadUK precipitation from the CAMELS-GB v2 dataset was used directly for all stages of the modelling (including calibration).
model_MIP_notes: valid flows at 628 core gauges only.

mod_PDM2PAR_RRMT.csv: provides flows from a model combination within the Rainfall-Runoff Modelling Toolbox (RRMT). Soil moisture accounting with probability distributed model (PDM) + flow routing model with two conceptual reservoirs in parallel (2PAR) (Wagener et al., 2002). The model was calibrated using a uniform random Monte Carlo sampling approach. Each of the model parameters were assigned upper and lower bounds based on feasible parameter ranges. For each of the 673 catchments, 10000 parameter sets were generated through uniform random sampling within these bounds. Kling-Gupta Efficiency (KGE) metric is chosen as the calibration objective function to obtain the best parameter set. Five years of data (1974-1979) were used as a warm-up period.

==========

FILE: GBHMIP_mod_PDM+2PAR_RRMT.csv

==========

title: PDM+2PAR model structure of RRMT simulating daily mean flows (m3 s-1) using HadUK-Grid precip, CHESS PET inputs (produced for UK Hydro-MIP)
creator_name: Melike Kiraz-Safari
creator_institution: Boğaziçi University - Institute of Environmental Sciences
creator_email: melike.kiraz@bogazici.edu.tr
date_created: 19/12/2025
input_PE: CHESS-PET
input_precip: HadUK-Grid
input_temperature: None
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: Soil moisture accounting with probability distributed model (PDM) + Flow routing model with two conceptual reservoirs in parallel (2PAR) from Rainfall-Runoff Modelling Toolbox (RRMT)
model_reference: Wagener, T., Lees, M.J., and Wheater, H.S. (2001). A toolkit for the development and application of parsimonious hydrological models. Mathematical Models of Small Watershed Hydrology, 2, 1-34. Additional reference
model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: The model was calibrated using a uniform random Monte Carlo sampling approach. Each of model parameters was assigned upper and lower bounds based on feasible parameter ranges. For each of the 673 catchments, 10000 parameter sets were generated through uniform random sampling within these bounds. Kling-Gupta efficiency metric (KGE) is chosen as the calibration objective function to obtain the best parameter set.
model_setup: Five years of data (1974-1979) were used as a warm-up period.
model_MIP_notes: The model structure is expected to perform well in the baselow-dominated catchments.

mod_PDMCRES_RRMT.csv: as above, but soil moisture accounting with probability distributed model (PDM) + flow routing model with conceptual reservoir (CRES) from Rainfall-Runoff Modelling Toolbox (RRMT).

==========
FILE: GBHMIP_mod_PDM+CRES_RRMT.csv
==========
title: PDM+CRES model structure of RRMT simulating daily mean flows (m3 s-1) using HadUK-Grid precip, CHESS PET inputs (produced for UK Hydro-MIP)
creator_name: Melike Kiraz-Safari
creator_institution: Boğaziçi University - Institute of Environmental Sciences
creator_email: melike.kiraz@bogazici.edu.tr
date_created: 19/12/2025
input_PE: CHESS-PET
input_precip: HadUK-Grid
input_temperature: None
time_period_start: 01/01/1980
time_period_end: 31/12/2019

model_name: Soil moisture accounting with probability distributed model (PDM) + Flow routing model with conceptual reservoir (CRES) from Rainfall-Runoff Modelling Toolbox (RRMT)
model_reference: Wagener, T., Lees, M.J., and Wheater, H.S. (2001). A toolkit for the development and application of parsimonious hydrological models. Mathematical Models of Small Watershed Hydrology, 2, 1-34. Additional reference
model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: The model was calibrated using a uniform random Monte Carlo sampling approach. Each of model parameters was assigned upper and lower bounds based on feasible parameter ranges. For each of the 673 catchments, 10000 parameter sets were generated through uniform random sampling within these bounds. Kling-Gupta efficiency metric (KGE) is chosen as the calibration objective function to obtain the best parameter set.
model_setup: Five years of data (1974-1979) were used as a warm-up period.
model_MIP_notes: The model structure is expected to perform well in the wet catchments without heavily human-modified flow regimes.

mod_PDMLEAK_RRMT.csv: as above, but soil moisture accounting with probability distributed model (PDM) + flow routing model with leaky aquifer model structure (LEAK) from Rainfall-Runoff Modelling Toolbox (RRMT).

==========
FILE: GBHMIP_mod_PDM+LEAK_RRMT.csv
==========

title: PDM+LEAK model structure of RRMT simulating daily mean flows (m3 s-1) using HadUK-Grid precip, CHESS PET inputs (produced for UK Hydro-MIP)
creator_name: Melike Kiraz-Safari
creator_institution: Boğaziçi University - Institute of Environmental Sciences
creator_email: melike.kiraz@bogazici.edu.tr
date_created: 19/12/2025
input_PE: CHESS-PET
input_precip: HadUK-Grid
input_temperature: None
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: Soil moisture accounting with probability distributed model (PDM) + Flow routing model with leaky aquifer model structure (LEAK) from Rainfall-Runoff Modelling Toolbox (RRMT)
model_reference: Wagener, T., Lees, M.J., and Wheater, H.S. (2001). A toolkit for the development and application of parsimonious hydrological models. Mathematical Models of Small Watershed Hydrology, 2, 1-34. Additional reference
model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: The model was calibrated using a uniform random Monte Carlo sampling approach. Each of model parameters was assigned upper and lower bounds based on feasible parameter ranges. For each of the 673 catchments, 10000 parameter sets were generated through uniform random sampling within these bounds. Kling-Gupta efficiency metric (KGE) is chosen as the calibration objective function to obtain the best parameter set.

model_setup: Five years of data (1974-1979) were used as a warm-up period.
model_MIP_notes: The model structure is expected to perform well in leaky-aquifer catchments (e.g., chalk-dominated catchments), where significant subsurface losses prevent a fully closed water balance.

mod_PEN2PAR_RRMT.csv: as above, but soil moisture accounting model based on Penman Drying Curve (PEN) + flow routing model with two conceptual reservoirs in parallel (2PAR) from Rainfall-Runoff Modelling Toolbox (RRMT).

==========
FILE: GBHMIP_mod_PEN+2PAR_RRMT.csv
==========

title: PEN+2PAR model structure of RRMT simulating daily mean flows (m3 s-1) using HadUK-Grid precip, CHESS PET inputs (produced for UK Hydro-MIP)
creator_name: Melike Kiraz-Safari
creator_institution: Boğaziçi University - Institute of Environmental Sciences
creator_email: melike.kiraz@bogazici.edu.tr
date_created: 19/12/2025
input_PE: CHESS-PET
input_precip: HadUK-Grid
input_temperature: None
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: Soil moisture accounting model based on Penman Drying Curve (PEN) + Flow routing model with two conceptual reservoirs in parallel (2PAR) from Rainfall-Runoff Modelling Toolbox (RRMT)
model_reference: Wagener, T., Lees, M.J., and Wheater, H.S. (2001). A toolkit for the development and application of parsimonious hydrological models. Mathematical Models of Small Watershed Hydrology, 2, 1-34. Additional reference
model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: The model was calibrated using a uniform random Monte Carlo sampling approach. Each of model parameters was assigned upper and lower bounds based on feasible parameter ranges. For each of the 673 catchments, 10000 parameter sets were generated through uniform random sampling within these bounds. Kling-Gupta efficiency metric (KGE) is chosen as the calibration objective function to obtain the best parameter set.
model_setup: Five years of data (1974-1979) were used as a warm-up period.
model_MIP_notes: The model structure is expected to perform well in the baselow-dominated catchments.

mod_PENCRES_RRMT.csv: as above, but soil moisture accounting model based on Penman Drying Curve (PEN) + flow routing model with conceptual reservoir (CRES) from Rainfall-Runoff Modelling Toolbox (RRMT).

==========
FILE: GBHMIP_mod_PEN+CRES_RRMT.csv
==========

title: PEN+CRES model structure of RRMT simulating daily mean flows (m3 s-1) using HadUK-Grid precip, CHESS PET inputs (produced for UK Hydro-MIP)

creator_name: Melike Kiraz-Safari
creator_institution: Boğaziçi University - Institute of Environmental Sciences
creator_email: melike.kiraz@bogazici.edu.tr
date_created: 19/12/2025
input_PE: CHESS-PET
input_precip: HadUK-Grid
input_temperature: None
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: Soil moisture accounting model based on Penman Drying Curve (PEN) + Flow routing model with conceptual reservoir (CRES) from Rainfall-Runoff Modelling Toolbox (RRMT)
model_reference: Wagener, T., Lees, M.J., and Wheater, H.S. (2001). A toolkit for the development and application of parsimonious hydrological models. Mathematical Models of Small Watershed Hydrology, 2, 1-34. Additional reference
model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: The model was calibrated using a uniform random Monte Carlo sampling approach. Each of model parameters was assigned upper and lower bounds based on feasible parameter ranges. For each of the 673 catchments, 10000 parameter sets were generated through uniform random sampling within these bounds. Kling-Gupta efficiency metric (KGE) is chosen as the calibration objective function to obtain the best parameter set.
model_setup: Five years of data (1974-1979) were used as a warm-up period.
model_MIP_notes: The model structure is expected to perform well in the wet catchments without heavily human-modified flow regimes.

mod_PENLEAK_RRMT.csv: as above, but soil moisture accounting model based on Penman Drying Curve (PEN) + flow routing model with leaky aquifer model structure (LEAK) from Rainfall-Runoff Modelling Toolbox (RRMT).

==========
FILE: GBHMIP_mod_PEN+LEAK_RRMT.csv
==========
title: PEN+LEAK model structure of RRMT simulating daily mean flows (m3 s-1) using HadUK-Grid precip, CHESS PET inputs (produced for UK Hydro-MIP)
creator_name: Melike Kiraz-Safari
creator_institution: Boğaziçi University - Institute of Environmental Sciences
creator_email: melike.kiraz@bogazici.edu.tr
date_created: 19/12/2025
input_PE: CHESS-PET
input_precip: HadUK-Grid
input_temperature: None
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: Soil moisture accounting model based on Penman Drying Curve (PEN) + Flow routing model with leaky aquifer model structure (LEAK) from Rainfall-Runoff Modelling Toolbox (RRMT)

model_reference: Wagener, T., Lees, M.J., and Wheater, H.S. (2001). A toolkit for the development and application of parsimonious hydrological models. Mathematical Models of Small Watershed Hydrology, 2, 1-34. Additional reference
model_spatial_resolution: catchment
model_timestep_minutes: 1440
model_calibration: The model was calibrated using a uniform random Monte Carlo sampling approach. Each of model parameters was assigned upper and lower bounds based on feasible parameter ranges. For each of the 673 catchments, 10000 parameter sets were generated through uniform random sampling within these bounds. Kling-Gupta efficiency metric (KGE) is chosen as the calibration objective function to obtain the best parameter set.
model_setup: Five years of data (1974-1979) were used as a warm-up period.
model_MIP_notes: The model structure is expected to perform well in leaky-aquifer catchments (e.g., chalk-dominated catchments), where significant subsurface losses prevent a fully closed water balance.

mod_PycatchmodHUK.csv: provides flows from pycatchmod-HUK (Wilby et al., 1994). The model was calibrated 1980-1999 (inc.) optimised using a differential evolution algorithm with a multi-weighted objective function predominantly weighted towards the relative residual error (low flows) but including NSE and quartiles (Q50 Q70 Q80 Q95). Pycatchmod simulates naturalised flows for a lumped output by combining four individual Hydrological Response Units (HRUs). Uncertainty is increased the more modified the flow regimes of the catchment are.

==========
FILE: GBHMIP_mod_Pycatchmod+HUK.csv
==========
title: pycatchmod-HUK daily mean flows (m3 s-1) covering 1980-2019 generated for the UK Hydro-MIP.
creator_name: Jamie Brown
creator_institution: Hydrology.UK
creator_email: jamie.brown@hydrology.uk.com
date_created: 15/01/2026
input_PE: CHESS PETI
input_precip: HadUK-Grid
input_temperature: None
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: pycatchmod-HUK
model_reference: Wilby R. Greenfield B. Glenny C. 1994. A coupled synoptic-hydrological model for climate change impact assessment. Journal of Hydrology. 153. p265-290
model_spatial_resolution: Lumped with seperated HRUs.
model_timestep_minutes: Daily.
model_calibration: 1980-1999 (inc.) optimised using a differential evolution algorithm with a multi weighted objective function predominantly weighted towards the relative residual error (low flows) but including NSE and quartiles (Q50 Q70 Q80 Q95)
model_setup: 4 HRUs set up

model_MIP_notes: pycatchmod simulates naturalised flows for a lumped output by combining four individual HRUs. Uncertainty is increased the more modified the flow regimes of the catchment are.

mod_SHETRAN_cal.csv: provides flows from SHETRAN, set up nationally (Lewis et al., 2018) and calibrated for each catchment using the method described in Smith et al. (2024). It is a gridded model with default resolution of 1km. Three catchment simulations are run at 200m as they are deemed too small for a 1km resolution simulation (39061, 76011, 80005). 16 catchment simulations were run at 5km resolution as these were too large to calibrate without reducing the calibration/validation periods (8006, 15006, 21009, 27009, 28009, 28022, 33035, 39001, 39002, 39072, 54001, 54032, 54057, 54095, 55023, 76007). Models used a two-layer soil-aquifer subsurface with the aquifer layers extending to 20-25m depth. Soils and land covers were simplified to either rural or urban cell types and parameterised uniformly across the catchment. Models are similar to those in the references above with the only changes being the swapping out of CEH-GEAR rainfall data to HadUK-Grid precipitation. The model was calibrated over the period 1990-1999, with spin-up occurring 1980-1989, in contrast to the other models. Consequently, the first two simulation years (1980 and 1981) have been removed from this submission.

==========

FILE: GBHMIP_mod_SHETRAN_cal.csv

==========

title: Calibrated SHETRAN UK Daily Flows
creator_name: Ben Smith
creator_institution: Newcastle University Water Group
creator_email: Ben.Smith4@ncl.ac.uk
date_created: 20/05/2026
input_PE: CHESS-PET
input_precip: HadUK-Grid
input_temperature: CHESS-MET
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: SHETRAN
model_reference: "Lewis et al. (2018) ""Development of a system for automated setup of a physically-based spatially-distributed hydrological model for catchments in Great Britain."""" Environmental modelling & software 108
model_spatial_resolution: 1km (plus 3x200m and 16x5km)
model_timestep_minutes: Variable 0-60
model_calibration: The model spin-up period was 1980-1998, and the calibration period was 1990-1999. "Calibrated using the method described in Smith at al. (2024) ""Physically-based modelling of UK river flows under climate change"". Frontiers in Water 6 1468855. 2016"
model_setup: Gridded models across the UK. Default resolution is 1km. Three simulations are run at 200m as they deemed too small for a 1km resolution simulation (39061 76011 80005). 16 simulations were run at 5km resolution as these were to large to calibrate without reducing the calibration/validation periods (8006 15006 21009 27009 28009 28022 33035 39001 39002 39072 54001 54032 54057 54095 55023 76007). Two layer model as per references above. Soil/subsoil with aquifer layer down to 20-25m depth. Soils and land cover

are simplified to either rural or urban cell types and parameterised uniformly across the catchment.
model_MIP_notes: Models are similar to those in the references above with the only changes being the swapping out of CEH-GEAR rainfall data to HadUK.

mod_SHETRAN_uncal.csv: provides flows from the uncalibrated SHETRAN model (Lewis et al., 2018). Spatially parameterised using the method described in Lewis (2016). All models are run at 1km except for the 5 largest catchments that are run at 5km (15006, 28022, 39001, 54032, 54057) and the three smallest that are run at 200m (39061, 76011, 80005). SHETRAN is a physically-based spatially-distributed model of surface/subsurface/and channel flows. Models used a two-layer soil-aquifer subsurface as per references above but with variable soils, geologies, and land uses based on dominant class within each grid cell. The model was calibrated over the period 1990-1999, with spin-up occurring 1980-1989, in contrast to the other models. Consequently, the first two simulation years (1980 and 1981) have been removed from this submission.

==========
FILE: GBHMIP_mod_SHETRAN_uncal.csv
==========

title: Uncalibrated SHETRAN UK Daily Flows
creator_name: Ben Smith
creator_institution: Newcastle University Water Group
creator_email: Ben.Smith4@ncl.ac.uk
date_created: 01/12/2025
input_PE: CHESS-PET
input_precip: HadUK-Grid
input_temperature: CHESS-MET
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: SHETRAN
model_reference: Elizabeth Lewis et al. "Development of a system for automated setup of a physically-based spatially-distributed hydrological model for catchments in Great Britain." Environmental modelling & software 108 (2018)
model_spatial_resolution: 1km
model_timestep_minutes: 0-60 (variable)
model_calibration: Uncalibrated. Spatially parameterised using Elizabeth Lewis "A robust multi-purpose hydrological model for Great Britain". Thesis. Newcastle University, 2016.
model_setup: 200m 1km and 5km gridded models across the UK. All models are run at 1km apart from the 5 largest that are run at 5km (15006 28022 39001 54032 54057) and the three smallest that are run at 200m (39061 76011 80005). Physically-based spatially distributed model of surface/subsurface/and channel flows. Two layer model as per references above. Soil/subsoil with aquifer layer down to 20-25m depth. Variable soils, geologies, and land uses based on dominant class within each grid cell. Parameterised using expert judgement and Aquifer Properties Manual.
model_MIP_notes: Models are similar to those in the references above with the only changes being the swapping out of CEH-GEAR rainfall data to HadUK.

mod_WFLOW-SBM.csv: provides flows from the wflow_sbm model (van Verseveld et al., 2024). This is an uncalibrated model where parameter values are induced from globally available gridded data, translated to model input based on literature. The exception is the vertical-to-horizontal saturated hydraulic conductivity ratio, which was created with a pedotransfer function estimated by the Random Forest (RF) ML algorithm (Ali et al., 2025). The wflow_sbm model is a distributed (gridded) model, set up covering the entirety of Great Britain on a 1km resolution. River flow was extracted at the cells corresponding to the gauge locations.

```
================================================================
FILE: GBHMIP_mod_WFLOW-SBM.csv
================================================================
title: Daily mean flows (m3 s-1) using HadUK-Grid precip, CHESS-PE pet and CHESS-Met temperature inputs.
creator_name: Brendan Dalmijn, Albrecht Weerts
creator_institution: Deltares
creator_email: albrecht.weerts@deltares.nl
date_created: 2025-12-01
input_PE: CHESS-PE
input_precip: HadUK-Grid
input_temperature: CHESS-Met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: wflow_sbm
model_reference: van Verseveld et al (2024) Wflow_sbm v0.7.3, a spatially distributed hydrological model
model_spatial_resolution: 1km grid
model_timestep_minutes: 1440
model_calibration: Uncalibrated model where parameter values are induced from globally available gridded data, translated to model input based on literature. The exception is the vertical-to-horizontal saturated hydraulic conductivity ratio, which was created with a pedotransfer function estimated by the Random Forest (RF) ML algorithm (Ali et al (2024) Machine learning for predicting spatially variable lateral hydraulic conductivity
model_setup: The wflow_sbm model is a distributed (gridded) model, set up covering the entirety of Great Britain on a 1km resolution. The streamflow was extracted at the cells corresponding to the gauges locations.
model_MIP_notes: None.
```

mod_WFLOW-SBM_BRT.csv: as above, except the vertical-to-horizontal saturated hydraulic conductivity ratio was created with a pedotransfer function estimated by the Boosted Regression Trees (BRT) ML algorithm (Ali et al., 2025).

```
================================================================
FILE: GBHMIP_mod_WFLOW-SBM_BRT.csv
================================================================
title: Daily mean flows (m3 s-1) using HadUK-Grid precip, CHESS-PE pet and CHESS-Met temperature inputs.
creator_name: Brendan Dalmijn, Albrecht Weerts
```

creator_institution: Deltares
creator_email: albrecht.weerts@deltares.nl
date_created: 2025-12-01
input_PE: CHESS-PE
input_precip: HadUK-Grid
input_temperature: CHESS-Met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: wflow_sbm
model_reference: van Verseveld et al (2024) Wflow_sbm v0.7.3, a spatially distributed hydrological model
model_spatial_resolution: 1km grid
model_timestep_minutes: 1440
model_calibration: Uncalibrated model where parameter values are induced from globally available gridded data, translated to model input based on literature. The exception is the vertical-to-horizontal saturated hydraulic conductivity ratio, which was created with a pedotransfer function estimated by the Boosted Regression Trees (BRT) ML algorithm (Ali et al (2024) Machine learning for predicting spatially variable lateral hydraulic conductivity
model_setup: The wflow_sbm model is a distributed (gridded) model, set up covering the entirety of Great Britain on a 1km resolution. The streamflow was extracted at the cells corresponding to the gauges locations.
model_MIP_notes: None.

mod_WFLOW-SBM_wateruse.csv: provides flows from the wflow_sbm model with an added water use component (Purnamasari et al., 2025; van Verseveld et al., 2024). Parts of the GB are rerun including anthropogenic water use and adjusted soil depth (5m instead of 2m) and use a different way to fit the exponential decline (f parameter).

==========
FILE: GBHMIP_mod_WFLOW-SBM_wateruse.csv
==========

title: Daily mean flows (m3 s-1) using HadUK-Grid precip, CHESS-PE pet and CHESS-Met temperature inputs.
creator_name: Brendan Dalmijn, Albrecht Weerts
creator_institution: Deltares
creator_email: albrecht.weerts@deltares.nl
date_created: 2025-12-01
input_PE: CHESS-PE
input_precip: HadUK-Grid
input_temperature: CHESS-Met
time_period_start: 01/01/1980
time_period_end: 31/12/2019
model_name: wflow_sbm
model_reference: van Verseveld et al (2024) Wflow_sbm v0.7.3, a spatially distributed hydrological model. Purnamasari et al. (2025) Improving realism of high-resolution hydrological modeling with anthropogenic water use
model_spatial_resolution: 1km grid
model_timestep_minutes: 1440

model_calibration: Uncalibrated model where parameter values are induced from globally available gridded data, translated to model input based on literature. The exception is the vertical-to-horizontal saturated hydraulic conductivity ratio, which was created with a pedotransfer function estimated by the Boosted Regression Trees (BRT) and Random Forest (RF) ML algorithms (Ali et al (2024) Machine learning for predicting spatially variable lateral hydraulic conductivity
model_setup: The wflow_sbm model is a distributed (gridded) model, set up covering the entirety of Great Britain on a 1km resolution. The streamflow was extracted at the cells corresponding to the gauges locations. We rerun parts of the GB with including anthropogenic water use and adjusted the soil depth (5m instead of 2m) and used a different way to fit the exponential decline (f parameter).
model_MIP_notes: None.

6.4 Output data

Output data for all model runs were converted into the same format – daily mean flows (m^3/s) at up to 673 gauging stations. All missing data values were set to -1 (these were mainly where a gauge could not be modelled due to a lack of input data). Where models simulated negative flows, values were set to 0 (this was only an issue for machine learning models).

7 Quality control

The river flow simulations were carried out by several independent research groups. They then went through a central quality-checking process, to make sure all output files were consistently formatted and flows appeared reasonable. These included checks that ensured: 1) consistent metadata had been provided for each model, 2) columns for all 673 gauges were included in each output file, 3) gauge names were consistently formatted and ordered, 4) there were no negative flow values, 5) all nodata values were recorded as -1, 6) flow values appeared reasonable (i.e. no obvious unit errors), 7) simulations following the main MIP protocol and those using the optional extra set of input data were kept separate and clearly distinguished. Details of who carried out each simulation, the model setup, calibration approach and any further MIP-specific considerations (e.g. missing catchments) are documented in the header lines of the output files alongside the simulated flows.

8 Limitations and user guidance

For a comprehensive discussion of dataset limitations and guidance on its use, please refer to the GB Hydro-MIP data paper.

Acknowledgements

The GB Hydro-MIP was made possible through the contributions of the many researchers and organisations who provided model simulations, expertise, and feedback throughout the project. We thank all modelling teams for participating in the intercomparison and sharing model outputs, metadata, and model descriptions. We are also grateful to the wider UK hydrological modelling community for their valuable input on the design and development of

the GB Hydro-MIP dataset. Coordination of the model intercomparison was supported by the Natural Environment Research Council award number NE/X019063/1 (Hydro-JULES).

References

Ali, A. N. A. M., Imhoff, R. O., & Weerts, A. H. (2025). *Machine learning for predicting spatially variable lateral hydraulic conductivity: a step towards efficient hydrological model calibration and global applicability*. https://doi.org/10.22541/essoar.174016989.99005964/v1

Bell, V. A., Kay, A. L., Davies, H. N., & Jones, R. G. (2016). An assessment of the possible impacts of climate change on snow and peak river flows across Britain. *Climatic Change*, *136*(3–4), 539–553. https://doi.org/10.1007/s10584-016-1637-x

Bell, V. A., Kay, A. L., Jones, R. G., Moore, R. J., & Reynard, N. S. (2009a). Use of soil data in a grid-based hydrological model to estimate spatial variation in changing flood risk across the UK. *Journal of Hydrology*, *377*(3–4), 335–350. https://doi.org/10.1016/j.jhydrol.2009.08.031

Bell, V. A., Kay, A. L., Jones, R. G., Moore, R. J., & Reynard, N. S. (2009b). Use of soil data in a grid-based hydrological model to estimate spatial variation in changing flood risk across the UK. *Journal of Hydrology*, *377*(3–4), 335–350. https://doi.org/10.1016/j.jhydrol.2009.08.031

Bergström, S. (1992). The HBV model - its structure and applications. In *Swedish Meteorological and Hydrological Institute, Norrköping*.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. . L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E. M., Boucher, O., Cox, P. M., Grimmond, C. S. B., & Harding, R. J. (2011). The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes. *Geoscientific Model Development*, *4*(3), 677–699. https://doi.org/10.5194/gmd-4-677-2011

Bush, M., Flack, D. L. A., Lewis, H. W., Bohnenstengel, S. I., Short, C. J., Franklin, C., Lock, A. P., Best, M., Field, P., McCabe, A., Van Weverberg, K., Berthou, S., Boutle, I., Brooke, J. K., Cole, S., Cooper, S., Dow, G., Edwards, J., Finnenkoetter, A., … Zhang, H. (2025). The third Met Office Unified Model–JULES Regional Atmosphere and Land Configuration, RAL3. *Geoscientific Model Development*, *18*(12), 3819–3855. https://doi.org/10.5194/gmd-18-3819-2025

Coxon, G., Freer, J., Lane, R., Dunne, T., Knoben, W. J. M., Howden, N. J. K., Quinn, N., Wagener, T., & Woods, R. (2019). DECIPHeR v1: Dynamic fluxEs and ConnectIvity for Predictions of HydRology. *Geoscientific Model Development*, *12*(6), 2285–2306. https://doi.org/10.5194/gmd-12-2285-2019

Coxon, G., Zheng, Y., Barbedo, R., Cooper, H., Fileni, F., Fowler, H. J., Fry, M., Green, A., Gribbin, T., Harfoot, H., Lewis, E., Neto, G. G. R., Qiu, X., Salwey, S., & Wendt, D. E. (2026). CAMELS-GB v2: hydrometeorological time series and landscape attributes for 671 catchments in Great Britain. *Earth System Science Data*, *18*(6), 4345–4371. https://doi.org/10.5194/essd-18-4345-2026

He, Y., Manful, D., Warren, R., Forstenhäusler, N., Osborn, T. J., Price, J., Jenkins, R., Wallace, C., & Yamazaki, D. (2022). Quantification of impacts between 1.5 and 4 °C of global warming on flooding risks in six countries. *Climatic Change*, *170*(1–2), 15. https://doi.org/10.1007/s10584-021-03289-5

Knoben, W. J. M. (2024). Setting expectations for hydrologic model performance with an ensemble of simple benchmarks. *Hydrological Processes*, *38*(10). https://doi.org/10.1002/hyp.15288

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*, *22*(11), 6005–6022. https://doi.org/10.5194/hess-22-6005-2018

Lane, R. (2026). *Python code to calculate hydrological signatures and model evaluation metrics (Version 1.0.0) [Computer software]*. Zenodo. https://doi.org/https://doi.org/10.5281/zenodo.23591728

Lane, R., Ford, E., Birkinshaw, S., Zheng, Y., Moore, R., Kiraz, M., Patil, S., Brown, J. R. C., Magee, E., Bulut, B., Forstenhäusler, N., Weerts, A., Robinson, E., & Harfoot, H. (2026). *Model descriptions for models participating in the GB Hydro-MIP*. Zenodo.

Lees, T., Buechel, M., Anderson, B., Slater, L., Reece, S., Coxon, G., & Dadson, S. J. (2021). Benchmarking data-driven rainfall–runoff models in Great Britain: a comparison of long short-term memory (LSTM)-based models with four lumped conceptual models. *Hydrology and Earth System Sciences*, *25*(10), 5517–5534. https://doi.org/10.5194/hess-25-5517-2021

Lewis, E. (2016). *A robust multi-purpose hydrological model for Great Britain*. Newcastle University.

Lewis, E., Birkinshaw, S., Kilsby, C., & Fowler, H. J. (2018). Development of a system for automated setup of a physically-based, spatially-distributed hydrological model for catchments in Great Britain. *Environmental Modelling and Software*. https://doi.org/10.1016/j.envsoft.2018.07.006

Met Office, Hollis, D., Carlisle, E., Kendon, M., Packman, S., & Doherty, A. (2024). HadUK-Grid Gridded Climate Observations on a 1km grid over the UK, v1.3.0.ceda (1836-2023). In *NERC EDS Centre for Environmental Data Analysis*.

Moore, R. J. (2007). The PDM rainfall-runoff model. *Hydrology and Earth System Sciences*. https://doi.org/10.5194/hess-11-483-2007

Patil, S., & Stieglitz, M. (2014). Modelling daily streamflow at ungauged catchments: what information is necessary? *Hydrological Processes*, *28*(3), 1159–1169. https://doi.org/10.1002/hyp.9660

Purnamasari, D., Verseveld, W. J. van, Buitink, J., Weiland, F. S., Dalmijn, B., Teuling, A. J., & Weerts, A. H. (2025). *Improving realism of high-resolution hydrological modeling with anthropogenic water use: a study on the Rhine basin*. https://doi.org/10.22541/essoar.175336966.62234347/v1

Pushpalatha, R., Perrin, C., Le Moine, N., Mathevet, T., & Andréassian, V. (2011). A downward structural sensitivity analysis of hydrological models to improve low-flow simulation. *Journal of Hydrology*, *411*(1–2), 66–76. https://doi.org/10.1016/j.jhydrol.2011.09.034

Rameshwaran, P., Bell, V. A., Brown, M. J., Davies, H. N., Kay, A. L., Rudd, A. C., & Sefton, C. (2022). Use of Abstraction and Discharge Data to Improve the Performance of a National-Scale Hydrological Model. *Water Resources Research*, *58*(1). https://doi.org/10.1029/2021WR029787

Robinson, E. L., Blyth, E. M., Clark, D. B., Comyn-Platt, E., Rudd, A. C., & Wiggins, M. (2023a). Climate hydrology and ecology research support system meteorology dataset for Great Britain (1961-2019) [CHESS-met]. In *Environmental Information Data Centre*.

Robinson, E. L., Blyth, E. M., Clark, D. B., Comyn-Platt, E., Rudd, A. C., & Wiggins, M. (2023b). Climate hydrology and ecology research support system potential evapotranspiration dataset for Great Britain (1961-2019) [CHESS-PE]. In *Environmental Information Data Centre*.

Robinson, E. L., Brown, M. J., Kay, A. L., Lane, R. A., Chapman, R., Bell, V. A., & Blyth, E. M. (2023). Hydro-PE: gridded datasets of historical and future Penman–Monteith potential evaporation for the United Kingdom. *Earth System Science Data*, *15*(10), 4433–4461. https://doi.org/10.5194/essd-15-4433-2023

Smith, B. A., Birkinshaw, S. J., Lewis, E., McGrady, E., & Sayers, P. (2024). Physically-based modelling of UK river flows under climate change. *Frontiers in Water*, *6*. https://doi.org/10.3389/frwa.2024.1468855

van Verseveld, W. J., Weerts, A. H., Visser, M., Buitink, J., Imhoff, R. O., Boisgontier, H., Bouaziz, L., Eilander, D., Hegnauer, M., ten Velden, C., & Russell, B. (2024). Wflow_sbm v0.7.3, a spatially distributed hydrological model: from global data to local applications. *Geoscientific Model Development*, *17*(8), 3199–3234. https://doi.org/10.5194/gmd-17-3199-2024

Wagener, T., Lees, M., & Wheater, H. (2002). A toolkit for the development and application of parsimonious hydrological models. In *Mathematical models of large watershed hydrology* (Vol. 1, pp. 87–126).

Wilby, R., Greenfield, B., & Glenny, C. (1994). A coupled synoptic-hydrological model for climate change impact assessment. *Journal of Hydrology*, *153*(1–4), 265–290. https://doi.org/10.1016/0022-1694(94)90195-3

Zheng, Y., Coxon, G., Rahman, M., Woods, R., Salwey, S., Rong, Y., & Wendt, D. E. (2025). DECIPHeR-GW v1: a coupled hydrological model with improved representation of surface–groundwater interactions. *Geoscientific Model Development*, *18*(13), 4247–4271. https://doi.org/10.5194/gmd-18-4247-2025