UK Hydro-MIP Protocol, version 2

July 2025

Elizabeth Cooper[1], Helen Baron[1], Emma Robinson[1], Rosanna Lane[1], Gabriele Arduini[2], Steven Birkinshaw[8], Segolene Berthou[3], Burak Bulut[1], Eleanor Blyth[1], Amulya Chevuturi[1], Douglas Clark[1], Gemma Coxon[4], Simon Dadson[1,5], Gurpreet Das[2],Richard Davis[6], Rich Ellis[1], Katie Facer-Childs[1], Emma Ford[5], Matt Fry[1], Helen Harfoot[6], Liz Lewis[7], MikeVaughan[6], Christel Prudhomme[2], Heather Rumbold[3], Ben Smith[8], Gianni Vesuviano[1]

[1]UKCEH, [2]ECMWF, [3]Met Office, [4]University of Bristol, [5]University of Oxford, [6]Environment Agency, [7]University of Manchester, [8]University of Newcastle

Introduction

Hydro-JULES is a NERC funded research programme bringing together a wide community of modellers to inform a new generation of hydrological models. As part of the Hydro-JULES cross-cutting benchmarking and model intercomparison theme, we have been consulting the community about running a hydrological Model Intercomparison Project (MIP). Following a series of meetings with stakeholders and further community consultation, we are now ready to circulate our community-agreed MIP protocol. We would be very pleased to have more models included in the MIP; if you run a hydrological model and would like to be included, please read the protocol and contact us at hydrojules@ceh.ac.uk.

The purpose of the MIP is to gather together modelled river flow outputs spanning a historical 40-year period (20 years calibration, 20 years evaluation) in order to help explore and summarise the current strengths of hydrological models used by the UK community. The MIP will include a range of types of hydrological models; crucially all of the models will use consistent driving data to generate outputs for a large number of GB catchments. This will allow us to explore the relative strengths of different approaches to hydrological modelling, enhancing our understanding of hydrological process representation in models. We aim to answer questions about the importance of different methods of process representation across flow regimes and catchment types. We also wish to investigate the relative strengths of taking a multi-model ensemble approach to hydrological modelling. We anticipate that this work will lead to at least one publication, the author list for which will include all MIP participants. We will also make all model outputs open and accessible to the hydrological community, enabling further research and analysis beyond the scope of the hydro-JULES project.

About this protocol

In this protocol we describe the experiments, input datasets, and output variables which will be included in the MIP. We also outline how to provide model results and some of the metrics that will be used to analyse them. This is the protocol for phase 1 of the MIP; future phases may have a different or wider scope, including different driving data, output variables, and/or spatial and temporal extents. Please also see our future roadmap for more details on future plans.

1. Experiments

a. **Spatial extent.** This MIP will compare modelled river flow for 673 catchments across the UK. This will comprise the 671 catchments covered by the CAMELS-GB dataset (https://doi.org/10.5194/essd-12-2459-2020) as well as the Tweed at Kingledores (https://nrfa.ceh.ac.uk/data/station/info/21014) and the Severn at Abermule (https://nrfa.ceh.ac.uk/data/station/info/54014). Discharge data availability is limited for some of these catchments, and therefore we have identified a list of 628 core gauges which have at least 10 years of flow data in both the calibration and evaluation periods. We require flow submissions for all core gauges, but encourage flow submissions for the full gauge list if possible (e.g. for models that do not require individual catchment calibration). The MIP catchment list is available to download from our website (https://hydro-jules.org/uk-hydro-mip) and is also available on JASMIN (/gws/nopw/j04/hydro_jules/data/uk_hydro_mip/supporting_data/ukhydromip_gauge_list).

b. **Temporal resolution.** We will be focusing on daily river flows. Specifically, the mean river flow in each water day (9am to 8:59am) for consistency with gauged daily flow records (https://nrfa.ceh.ac.uk/data/about-data/daily-flow-data/gauged-daily-flow-data).

c. **Time periods.** We invite modelers to submit daily simulated river flows covering the period 1980 to 2019 inclusive.
    - Data from any years pre-1980 can be used for a model spin-up. We have no restrictions on how model stores are initalised or the length of spin-up period.
    - The period 1980 – 1999 inclusive should be used for model calibration/ training.

- The period 2000 – 2019 inclusive will be used for model evaluation.

If re-calibration is not possible or these periods are unsuitable for your model, please email the MIP team on hydrojules@ceh.ac.uk. We can allow some flexibility in the calibration/evaluation periods for models where recalibration is not feasible, as long as it is clearly documented in the model output files (see section 3 below).

d. **Calibration.** We leave the choice of calibration strategy to modellers. This MIP will evaluate model performance over the whole hydrograph, including simulation of annual average flows, seasonal flows, high/low flow quantiles and flow timing; calibration objective functions should be chosen with this in mind. We recognise that the choice of calibration metric may have a large impact on simulated flows, so we ask everyone to make sure any objective function(s) used are clearly documented in the output files. We welcome multiple submissions from the same model structure with different behavioural parameter sets, including simulations employing different calibration approaches. In this case, run differences should be reflected in the output filenames (see the naming convention below) and fully described within the model_calibration section of each output file.

e. **Model setup.** Spatial resolution, process representation and ancillary datasets may be chosen by the user to suit their model. Differences between models will be documented through the Environment Agency's scientific model description questionnaire (to be circulated once flow simulations have been submitted). Further comments regarding the model setup and calibration should be added to the output files. We welcome submission of flows from different versions/ setups of your model; please make it clear how these differ in the filenames and model setup section of the output files.

2. Driving data

This MIP requires common meteorological driving data for all physical models. Gridded meteorological drivers from, or based on, the CHESS-met dataset should be used, apart from precipitation, which should be from the HadUK-Grid dataset. See Table 1 for a summary of required input datasets and where these can be accessed. For models requiring catchment-average inputs, these will be available as part of CAMELS-GB version 2. Please email Gemma Coxon (gemma.coxon@bristol.ac.uk) for early access to these data. CAMELS-GB v2 does not include catchment average data for all CHESS-met variables, please email the MIP team if these are required by your catchment model.

Modelers may otherwise aggregate or disaggregate the gridded driving data in time or space to suit their model (e.g. to create subdaily meteorological forcing) but we ask that details of any such procedures are provided along with output data.

Table 1. Summary of hydrometeorological input variables to be used within the MIP and where they can be accessed.

<table>
<tr><th rowspan="2">Variable</th><th rowspan="2">Data to use for the MIP</th><th colspan="2">Data access</th></tr>
<tr><th>For models requiring catchment-average inputs</th><th>For models requiring gridded inputs</th></tr>
<tr><td>Precipitation</td><td>HadUK-Grid precipitation</td><td rowspan="4">Available as part of CAMELS-GB v2, email Gemma Coxon for early access (gemma.coxon@bristol.ac.uk)</td><td>Download from CEDA archive: https://dx.doi.org/10.5285/b963ead70580451aa7455782224479d5 or access on JASMIN: /badc/ukmo-hadobs/data/insitu/MOHC/HadOBS/HadUK-Grid/v1.3.0.ceda/</td></tr>
<tr><td>Potential evapotranspiration (with or without interception)</td><td>CHESS-PE

Optional extra runs: Hydro-PE HadUK-Grid</td><td>Download from EIDC: https://doi.org/10.5285/8651771d-aa6d-4d0f-8bcd-b3be1f733852 or access on JASMIN: /gws/nopw/j04/hydro_jules/data/uk/driving_data/chess/chess-pe/ (available to all registered users)

For optional extra runs, Hydro-PE HadUK Grid can be downloaded from: https://doi.org/10.5285/bebb62085-ba81-480c-9ed0-2d31c27ff196</td></tr>
<tr><td>Air temperature</td><td rowspan="3">CHESS-met</td><td rowspan="3">Download from EIDC: https://doi.org/10.5285/835a50df-e74f-4bfb-b593-804fd61d5eab</td></tr>
<tr><td>Specific humidity</td><td rowspan="2">-</td></tr>
<tr><td>Downward shortwave radiation</td></tr>
</table>

Downward longwave radiation			or access on JASMIN: /gws/nopw/j04/hydro_jules/data/uk/driving_data/chess/chess-met/ (available to all registered users)
Surface air pressure			
Wind speed			
Daily temperature range			

CHESS-met comprises meteorological data for the period 1961-2019 at a daily timestep, 1km resolution at a measurement height of 1.2m. The variables available are: air temperature at 1.2m (K), specific humidity at 1.2m (kg kg-1), downward shortwave radiation (W m-2), downward longwave radiation (W m-2), surface air pressure (Pa), wind speed at 10m (m s-1), daily temperature range (K). The CHESS-met dataset also contains precipitation. However, in this MIP we will use precipitation from the HadUK-Grid dataset.

For models that require PE as an input, we require modellers to use CHESS-PE. This is the PE product most consistent with the CHESS-met drivers and the JULES model and comprises two related datasets; modellers can select either the PETI option which includes a correction for interception, or the PET version which has no correction for interception. We recognise that a limitation of the CHESS-PE product is that it only extends until 2019, and therefore runs cannot include the 2021/2022 drought. We invite any modellers submitting flows to the main MIP to also consider producing a second set of flows using Hydro-PE HadUK-Grid as the PE input. Details of these optional extra runs are given in section 7.

While we require consistent meteorological driving data, there will be no constraints on other inputs. We want to encourage as many modellers to contribute to the MIP as possible, and recognise that it would not be practical to request that everyone uses the same ancillary datasets. We also understand that there are vast differences in data requirements between models. For models requiring catchment characteristic data, this is available from the CAMELS-GB dataset: Catchment attributes and hydro-meteorological timeseries for 671 catchments across Great Britain (CAMELS-GB) - EIDC.

3. Output variables

Please provide daily mean timeseries of river flow in m^3s^{-1} for the catchments and time periods detailed above. River flow outputs can be supplied either in a netcdf format or as a csv file with 17 header lines containing all the all the required model run information. Please format the filename in lowercase as: [model]_[start-date]_[end-date]_[PE

input]_[any additional required run identifier]. For example g2g_1980_2019_chesspeti.csv. Table 2 details information that should be supplied as global attributes (if supplying flows in netcdf format) or header lines (if supplying flows as a csv).

Please supply model outputs as an array of river flows. Columns should be the year, month and day followed by flows for each of the MIP catchments (preferably with gauging station numbers in ascending order). Please supply any missing values as -1, with an explanation for why these could not be simulated given in the model_MIP_notes section. If only providing simulations for the 628 core catchments, please still include all gauges in the output file, with a note in the model_MIP_notes section saying that only core gauges were modelled.

Template model output files have been provided on JASMIN (/gws/nopw/j04/hydro_jules/data/uk_hydro_mip), with an example of what this looks like in Figure 1.

Table 2. Information that should be supplied as global attributes / as a header alongside simulated flows.

csv line no.	Name	Description	Example
1	title	Short title describing model runs	Grid-to-Grid (G2G) daily mean flows (m3 s-1) covering 1980-2019, generated for the UK Hydro-MIP.
2	creator_name	Name of person carrying out the model runs	Rosanna Lane
3	creator_institution	Primary institution of person carrying out the model runs	UK Centre for Ecology and Hydrology (UKCEH)
4	creator_email	Email address of person carrying out the model runs	roslan@ceh.ac.uk
5	date_created	dd/mm/yyyy	01/04/2025
6	input_PE	CHESS-PETI or CHESS-PET (Hydro-PE HadUK-Grid for optional extra runs)	CHESS-PETI
7	input_precip	This should be HadUK-Grid	HadUK-Grid
8	input_temperature	This should be CHESS-MET or None	CHESS-met
9	time_period_start	Start date of model outputs, given as DD/MM/YYYY	01/01/1980
10	time_period_end	End date of model outputs, given as DD/MM/YYYY	31/12/2019
11	model_name	Full name of model	Grid-to-Grid (G2G)

12	model_reference	Reference to use when discussing model	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, 377(3-4), 335-350.
13	model_spatial_resolution	Resolution (specify if catchment, HRU-based or gridded model)	1km grid
14	model_timestep_minutes	Model time step in minutes (i.e. 1440 if run at a daily resolution)	15
15	model_calibration	details of model calibration	uncalibrated. G2G is mainly parameterised using spatial datasets of landscape properties (e.g. HOST soil types) rather than via catchment calibration.
16	model_setup	details of model setup	Here, 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.
17	model_MIP_notes	Any model specific notes that are relevant for the MIP	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.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	title: Grid-to-Grid (G2G) daily mean flows (m3 s-1) covering 1980-2019, generated for the UK Hydro-MIP.																	
2	creator name: Rosanna Lane																	
3	creator institution: UK Centre for Ecology and Hydrology (UKCEH)																	
4	creator_email: roslan@ceh.ac.uk																	
5	date_created: 15/05/2025																	
6	input_PE: CHESS PETi																	
7	input_precip: HadUK-Grid																	
8	input_temperature: None																	
9	time_period_start: 01/01/1980																	
10	time_period_end: 30/12/2019																	
11	model_name: Grid-to-Grid (G2G)																	
12	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																	
13	model_spatial_resolution: 1km grid																	
14	model_timestep_minutes: 15																	
15	model_calibration: G2G is an uncalibrated model mainly parameterised using spatial datasets of landscape properties (e.g. HOST soil types).																	
16	model_se 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 as it requires i																	
17	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.																	
18	YYYY	MM	DD	01001	02001	02002	03003	04001	04003	04005	04006	05003	06007	06008	07001	07002	07003	07005
19	1980	1	1	0.956	7.798	5.2	6.072	23.376	3.531	3.064	2.794	15.313	50.24	1.981	5.662	11.095	1.605	1.441
20	1980	1	2	0.994	7.304	4.784	5.709	21.611	3.217	3.073	2.713	15.064	45.325	1.869	5.091	9.681	1.545	1.328
21	1980	1	3	0.914	6.988	4.499	5.521	20.565	3.003	2.891	2.605	14.867	42.208	1.77	4.813	8.7	1.474	1.255
22	1980	1	4	0.796	6.696	4.28	5.213	19.711	2.84	2.695	2.437	13.882	41.034	1.65	4.634	8.188	1.432	1.256
23	1980	1	5	0.714	6.432	4.239	5.169	19.029	2.841	2.796	2.488	13.261	39.938	1.617	4.79	7.917	1.424	1.255
24	1980	1	6	0.659	6.25	4.193	5.772	19.886	2.842	3.901	3.184	14.698	38.773	1.958	4.625	7.814	1.404	1.208
25	1980	1	7	0.622	6.303	4.198	7.363	25.266	3.067	4.979	4.435	23.087	45.059	2.587	5.153	7.71	1.361	1.157
26	1980	1	8	0.63	7.083	4.936	8.796	29.879	3.578	4.727	4.651	25.325	65.142	2.719	6.582	7.58	1.312	1.136
27	1980	1	9	0.714	8.512	6.598	10.546	35.779	5.621	8.057	7.077	34.814	84.535	4.452	10.848	9.106	1.254	1.317
28	1980	1	10	0.82	10.263	8.15	16.448	59.644	10.017	11.4	10.614	55.317	130.378	6.207	19.735	14.015	1.256	1.84
29	1980	1	11	0.834	10.871	8.82	20.857	70.294	8.358	9.462	9.888	42.62	171.96	4.494	16.028	24.135	1.306	1.712
30	1980	1	12	0.747	10.657	8.198	17.516	62.61	7.688	7.68	8.155	33.938	140.259	3.868	11.782	20.323	1.303	1.385
31	1980	1	13	0.693	10.136	7.842	15.593	54.372	7.944	6.746	6.951	29.693	127.974	3.598	10.96	15.561	1.208	1.206

Figure 1. Example of a model output file in .csv format.

4. Where to put results and how they will be analysed

JASMIN storage for model outputs is available, for the main runs at: /gws/nopw/j04/hydro_jules/data/uk_hydro_mip/main/ and for the optional extra runs at /gws/nopw/j04/hydro_jules/data/uk_hydro_mip/optional_extra/. To access this space users will need to be members of the Hydro-JULES group workspace. JASMIN users can apply to join if necessary (follow the guidance here: https://help.jasmin.ac.uk/docs/short-term-project-storage/apply-for-access-to-a-gws/, and please mention the MIP). If participants are not JASMIN users, please contact the team to make alternative arrangements for uploading model outputs (hydrojules@ceh.ac.uk).

We will run analysis against NRFA daily flow data and make the analysis and analysis tools available to all MIP contributors.

We will evaluate flows using a range of metrics and hydrological signatures, over the 20-year evaluation period. These will include summary metrics such as the Kling Gupta efficiency and Nash-Sutcliffe efficiency, decomposed KGE metrics of bias, variance and correlation, and analysis of a range of hydrological signatures such as errors in annual/seasonal flows, flow quantiles, and peak flows. We will endeavor to consider observational uncertainties in our analysis where this information is available. A formal list

of performance metrics is still being devised in consultation with the community, and all simulated streamflow timeseries will be made openly available for anyone wishing to calculate additional performance metrics.

5. Getting involved

This MIP will pull together river flow outputs from a wide range of UK hydrological models, generating a valuable dataset. Our analysis of these outputs will help to inform understanding of hydrological process representation in models across flow regimes and catchment types. We will gather model outputs from participants through 2025 and begin analysis in 2026. If you have a model which is suitable for inclusion in the MIP, or any other queries, please contact us on hydrojules@ceh.ac.uk.

6. Guidelines

Participants involved in the MIP are committing to the following guidelines:

- To understand and follow the modelling protocol, including the guidance on submitting model data and meta-data (see section 0).
- To ensure the suitability of their model and to conduct required internal quality checks before submission.
- To submit model output according to the prescribed timelines.
- All submitted data will be made openly available (with the expectation of appropriate citation in the MIP outputs and any future use).
- To be available to provide additional information on their model and setup as required for the MIP output publication.

7. Optional extras: alternative PE

A key component of the Hydro-JULES MIP is that we require all models to be run with the same driving data. This will allow us to explore the relative strengths of different modelling approaches, focusing on how differences in process representation and model resolution impact modelled flows. We have selected CHESS-PE as the potential evapotranspiration (PE) product to be used in the MIP, as it is accompanied by the CHESS-met dataset which contains the variables required to run the JULES model.

However, we recognise that the choice of driving data will impact flows and other models may have been calibrated based on other PE products. The impact of driving data on

modelled flows is also an interesting question. We therefore invite modellers who have submitted outputs following the main MIP protocol to submit a second set of runs using Hydro-PE. The protocol for this second set of model outputs is outlined below.

Temporal extent: The Hydro-PE HadUK-Grid product is available until 2022. We therefore propose running models from 1990 to 2022 inclusive when using Hydro-PE HadUK Grid. This extended period will still allow comparison with the main MIP over 1990-2019, while also covering the 2021/2022 drought.

Spatial extent, calibration and model setup: the model should be run over the same catchments and following the same guidelines for model setup as outlined in the main MIP protocol.