Received: 9 April 2018 | Revised: 24 July 2018 | Accepted: 29 August 2018

DOI: 10.1002/gdj3.55

DATA PAPER

WILEY Geoscience Data Journal RMetS

The MaRIUS-G2G datasets: Grid-to-Grid model estimates of flow and soil moisture for Great Britain using observed and climate model driving data

Victoria A. Bell | **Alison L. Kay** | **Alison C. Rudd** | **Helen N. Davies**

Centre for Ecology & Hydrology, Wallingford, UK

Correspondence
Victoria A. Bell, Centre for Ecology & Hydrology, Wallingford, Oxfordshire OX10 8BB, UK.
Email: vib@ceh.ac.uk

Funding information
This article was funded by a grant from the UK Natural Environment Research Council's Drought and Water Scarcity programme (NE/L010208/1).

Abstract

The MaRIUS-G2G datasets were produced for the MaRIUS (Managing the Risks, Impacts and Uncertainties of drought and water Scarcity) project, using the Grid-to-Grid (G2G) national-scale hydrological model for Great Britain. There are six separate datasets, with each of three combinations of meteorological driving data (two observation-based and one from climate model ensembles) used to produce two types of outputs (daily time-series of natural river flow for 260 sites, and monthly 1×1 km grids of natural flow and soil moisture in the unsaturated zone). The driving data required by G2G are rainfall and potential evaporation (PE). Two of the datasets from observation-based driving data use rainfall from CEH-GEAR (CEH-Gridded Estimates of Areal Rainfall) with PE from MORECS

Dataset

MaRIUS-G2G consists of 6 separate datasets, with each of 3 combinations of driving data used to produce 2 types of outputs, flow and soil-moisture:
Identifier: doi:10.5285/5f3c1a02-d5c4-4faa-9353-e8b68ce2ace2
Title: Grid-to-Grid model estimates of daily mean river flow for gauged catchments in Great Britain (1960 to 2015): observed driving data [MaRIUS-G2G-MORECS-daily]
Identifier: doi:10.5285/e911196a-b371-47b1-968c-661eb600d83b
Title: Grid-to-Grid model estimates of monthly mean flow and soil moisture for Great Britain (1960 to 2015): observed driving data [MaRIUS-G2G-MORECS-monthly]
Identifier: doi:10.5285/0ceb4f85-0bbf-49f0-ab70-cfc137ab7d4d
Title: Grid-to-Grid model estimates of daily mean river flow for gauged catchments in Great Britain (1891 to 2015): observed driving data [MaRIUS-G2G-Oudin-daily]
Identifier: doi:10.5285/f52f012d-9f2e-42cc-b628-9cdea4fa3ba0
Title: Grid-to-Grid model estimates of monthly mean flow and soil moisture for Great Britain (1891 to 2015): observed driving data [MaRIUS-G2G-Oudin-monthly]
Identifier: doi:10.5285/f6cac471-7d92-4e6d-be8a-9f7887143058
Title: Grid-to-Grid model estimates of daily mean river flow for gauged catchments in Great Britain: weather@home2 (climate model) driving data [MaRIUS-G2G-WAH2-daily]
Identifier: doi:10.5285/3b90962e-6fc8-4251-853e-b9683e37f790
Title: Grid-to-Grid model estimates of monthly mean flow and soil moisture for Great Britain: weather@home2 (climate model) driving data [MaRIUS-G2G-WAH2-monthly]
Creator: Centre for Ecology and Hydrology (CEH) [Bell, V.A., Rudd, A.C., Kay, A.L. and Davies, H.N.].
Publisher: NERC Environmental Information Data Centre (EIDC), UK
Publication year: 2018
Resource type: Metadata document
Version 1.0

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2018 The Authors. *Geoscience Data Journal* published by Royal Meteorological Society and John Wiley & Sons Ltd.

(Met Office Rainfall and Evaporation Calculation System), and cover the period 1960–2015, while the other two use CEH-GEAR rainfall with PE derived from 5 km temperature data using the Oudin method, and cover 1891–2015. The two datasets based on driving data (rainfall and PE) from the climate model ensembles cover three periods: Historical Baseline (1900–2006), Near-Future (2020–2049), and Far-Future (2070–2099). Data for a 30-year Baseline period (1975–2004), against which the Near-Future and Far-Future periods should be compared, are also available directly. There are 100 members in each ensemble, and the future periods use the RCP8.5 emissions scenario. This paper provides details of the G2G model and the different sets of meteorological driving data, as well as the availability and formatting of the output datasets. It also describes some recent and potential applications of the datasets, which have already been used to support historical and future analyses of low flow and drought characteristics across Britain, and provides some guidance on how the climate model-driven datasets should (and should not) be used.

KEYWORDS
climate change, drought, hydrological modelling, river flow, soil moisture

1 | INTRODUCTION

MaRIUS (Managing the Risks, Impacts and Uncertainties of drought and water Scarcity) was a UK NERC-funded research project (2014–2017) which developed a risk-based approach to the management of droughts and water scarcity (www.mariusdroughtproject.org). As part of this overall aim, MaRIUS had several scientific and applied objectives, including a wish to promote the uptake of the research through the analysis of real and synthetic droughts, at catchment and national scales, under historical and future climatic conditions. This analysis was undertaken within a probabilistic framework supported by the generation of synthetic time-series of hydrometeorological conditions for historical and future time-periods using a regional climate model. These time-series were propagated through the MaRIUS suite of models (of hydrology, agriculture, ecology, and socio-economic impacts) to analyse the impacts of water scarcity and drought risk on a range of sectors.

The hydrological datasets presented here (Bell et al., 2018a, 2018b, 2018c, 2018d, 2018e, 2018f) consist of output from one of the MaRIUS suite of models: a national-scale grid-based hydrological model (Grid-to-Grid [G2G]). The model output has already been used in MaRIUS to support a range of drought analyses, but could also be used to support other hydrological research.

MaRIUS was one of five projects funded under the NERC UK Droughts & Water Scarcity Programme. The fifth project, ENDOWS (aboutdrought.info), aims to maximize the impact of the Programme for a diverse range of stakeholders. It also provides easier access to the wide range of datasets produced by the Programme, including the hydrological model output.

2 | DATA PRODUCTION METHODS

The G2G grid-based hydrological model is used with three combinations of meteorological driving data (two observation-based and one using climate model ensembles) to produce two types of outputs for Great Britain (GB; daily time-series for 260 sites and monthly grids). The result is six datasets, collectively termed MaRIUS-G2G (Figure 1).

2.1 | The G2G model

The G2G is a national-scale hydrological model for GB that runs on a 1×1 km grid (aligned with the GB national grid), at a 15-min time-step, and is parameterized using digital datasets (e.g., soil types, land-cover) (Bell *et al.*, 2016). The effect of urban and suburban land-cover on runoff and downstream flows is accounted for in the model. G2G has been shown to perform well for a wide range of catchments across Britain (Bell *et al.*, 2016; 2009; Formetta *et al.*, 2017) particularly those with more natural flow regimes as it currently does not include the effect of artificial influences such as abstractions and discharges on river flows. G2G is routinely used for high flow applications, for example operational fluvial flood forecasting (Cole and Moore, 2009), pluvial flood forecasting (Speight *et al.*, 2016) and assessments of the

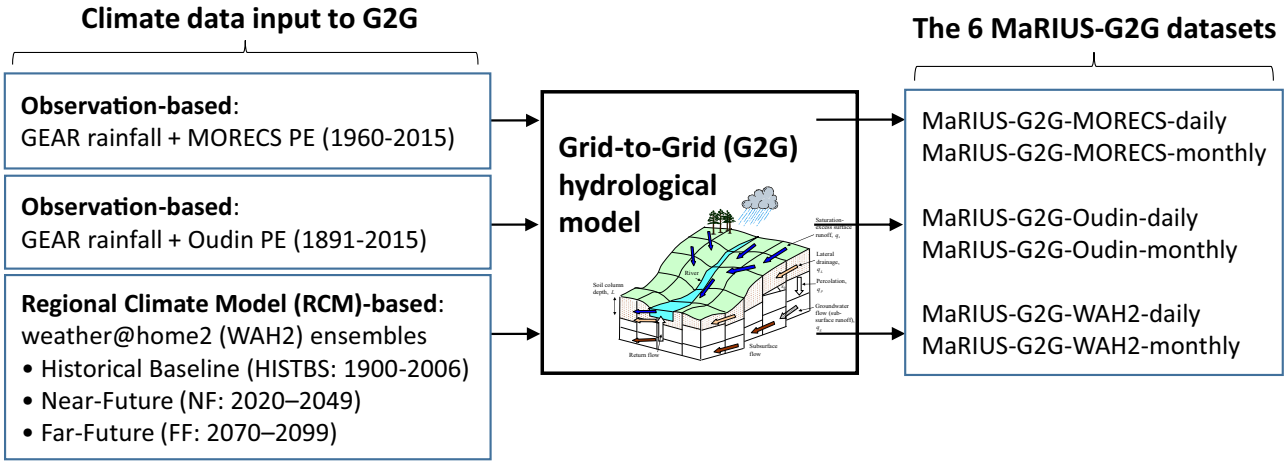

FIGURE 1 Schematic summarising the 6 MaRIUS-G2G datasets

effect of projected climate change on peak river flows (Bell *et al*., 2016; 2009). It has recently been shown to perform well for low flows and for drought identification (Rudd *et al*., 2017). The G2G generally uses spatial datasets in preference to parameter identification via calibration, and where model parameters are required (such as the kinematic wave speeds used in lateral routing), nationally applicable values are used. Thus, calibration has not been used to identify separate model parameters for individual catchments.

G2G requires input time-series of precipitation and potential evaporation (PE). The optional snow module (Bell *et al*., 2009) is not used here, thus precipitation input to G2G is assumed to be rain. The spatial data, such as topography and soil data, used to configure G2G are as in Bell *et al*. (2016).

2.2 | G2G driving data: observation-based

The first combination of observation-based meteorological driving data consists of.

- 1 km x 1 km grids of daily rainfall (CEH-GEAR: Keller *et al*., 2017; Tanguy *et al*., 2016),
- Monthly PE data on a 40 km grid (MORECS: Hough and Jones, 2008).

MORECS (Met Office Rainfall and Evaporation Calculation system) provides observation-based monthly estimates of PE from well-watered short grass using the Penman-Monteith equation (Monteith, 2005). Although the CEH-GEAR rainfall data are available for the period 1890–2015, MORECS PE is not available pre-1960, so the G2G runs with these data only cover the period 1960–2015.

The second combination of observation-based meteorological driving data covers a longer historical period, and consists of

- 1 km x 1 km grids of daily rainfall (CEH-GEAR),
- Monthly PE estimates derived from temperature data on a 5 × 5 km grid.

As neither MORECS PE nor all of the climate data required by the Penman–Monteith equation for PE (e.g., long and short wave radiation) are available for earlier historical periods, the temperature-based method of Oudin *et al*. (2005) has been used to estimate monthly PE from gridded (5 × 5 km) monthly temperature observations available from 1891 (Perry and Hollis, 2011). A set of monthly spatial correction factors has been applied to fit the long-term mean values of Oudin PE to those of MORECS (Rudd *et al*., 2017), as there are some differences in seasonal patterns, although annual means are similar (Kay and Davies, 2018). This method of estimating PE allows G2G runs covering the period 1891–2015. Note that rainfall measurements were particularly sparse for periods in the first half of the 20th century, and sometimes identified as missing in CEH-GEAR (figure 6 of Keller *et al*., 2017); these were spatially infilled (Rudd *et al*., 2017).

For both combinations of observation-based driving data, the PE are copied to each of the corresponding 1 × 1 km boxes of the hydrological model grid, and both PE and rainfall are divided equally down to the 15-min model time-step (Bell *et al*., 2016; 2009). Despite differences in the resolution of the base datasets (40 km for MORECS PE and 5 km for Oudin PE), the spatial variability in the 1 km downscaled PE data is very similar, probably as a consequence of the monthly spatial correction factors used to fit Oudin PE to MORECS PE, together with

the relatively low spatial variability in PE estimates across Britain (Robinson *et al*., 2017, figure 6).

2.3 | G2G driving data: climate model ensembles

The required rainfall and PE are from the MaRIUS weather@home2 (WAH2) regional climate model (RCM) dataset (Guillod *et al*., 2018). The WAH2 system uses volunteer computing time to do large numbers of runs of the HadRM3P RCM nested in the HadAM3P atmospheric global climate model (Guillod *et al*., 1997). Data are available for 100-member ensembles covering three periods:

- Historical baseline (HISTBS: 1900–2006),
- Near-future (NF: 2020–2049),
- Far-future (FF: 2070–2099).

Data for a 30-year Baseline (BS) period (1975–2004), against which the Near-Future (NF) and Far-Future (FF) periods should be compared, are also available (as a subset of the longer Historical BS period). Five alternative sets of NF and FF ensembles were produced, with varying sea surface temperature (SST) warming patterns and magnitudes; only the NF and FF sets based on the median patterns and warming are used here. The future periods use the RCP8.5 emissions scenario (Riahi *et al*., 2017).

A simple bias-correction scheme, based on monthly multiplicative factors, was applied to precipitation (Guillod *et al*., 1997). PE was derived from other WAH2 meteorological variables (Guillod *et al*., 1997) using the Penman–Monteith scheme (Monteith, 2005).

- For the historical periods, the PE calculation uses monthly stomatal resistance (r_s) values from MORECS (Hough and Jones, 2008).
- For the future periods, two alternative versions of PE are available: one using the same r_s values as the historical periods and one using values adjusted to allow for closure of stomata under increased carbon dioxide concentrations. Adjusting r_s decreases the projected changes in PE in the future (Rudd and Kay, 2016; Guillod *et al*., 1997), and moderates projected future decreases in low flows (Kay *et al*., 2015). The datasets provided here only use the adjusted r_s PE.
- Unlike precipitation, PE is not bias-corrected (see Guillod *et al*., 1997 for more details).

For use by G2G, the WAH2 precipitation and PE data are re-projected from the 0.22° (~25 km) rotated lat-lon RCM grid to the 1 × 1 km G2G grid. Following re-projection, spatially distributed weights based on standard average annual rainfall patterns are used to provide a non-uniform distribution of precipitation within each RCM box (Bell *et al*., 2007). Note that the WAH2 RCM assumes 360-day years (twelve 30-day months).

For the first 2 years of the HISTBS, NF and FF simulations the G2G was being "spun up," thus flow estimates from the first 2 years should be ignored in analyses, or used only for follow-on model spin-up.

2.4 | G2G outputs

Two types of outputs are produced:

- daily time-series of river flow for 260 sites across Great Britain,
- monthly 1 × 1 km grids of flow and soil moisture for Great Britain.

For daily outputs, G2G flow estimates are provided as daily mean natural flows (m^3/s) for locations corresponding to 260 National River Flow Archive (NRFA; nrfa.ceh.ac.uk) gauging stations (Figure 2). The majority of the 260 sites were selected to achieve a wide spatial coverage across Britain and are typically the furthest downstream gauge on a river. Other sites were chosen as they were part of a set of gauges used to assess the G2G for low flow events (Rudd *et al*., 2017). The gauged sites represent catchments covering approximately 65% of mainland GB, and span a very broad range of soils, relief, climate conditions, and anthropogenic influences.

For monthly outputs, G2G flow estimates are provided as monthly averages of daily mean natural flows (m^3/s), and G2G soil moisture estimates are provided as monthly averages of daily mean soil moisture in the unsaturated zone (mm water/m soil). The latter can also be interpreted as 1,000 θ, where θ has units of m water/m soil ($0 \leq \theta \leq 1$). The G2G model assumes that soil properties, including soil depth, vary spatially across Britain. Soil depth can vary from a few centimetres to several metres, and G2G model soil moisture estimates should be interpreted as depth-integrated values for the whole soil column. Both flow and soil moisture estimates are provided for every non-sea 1 × 1 km grid box, and flows are provided for every 1 km land grid box whether there is a river located in the grid box or not. Figure 3 presents example of 1 × 1 km G2G flow and soil moisture output over Britain.

3 | DATASET LOCATION AND FORMAT

The six MaRIUS-G2G datasets are available from the Environmental Information Data Centre (EIDC;

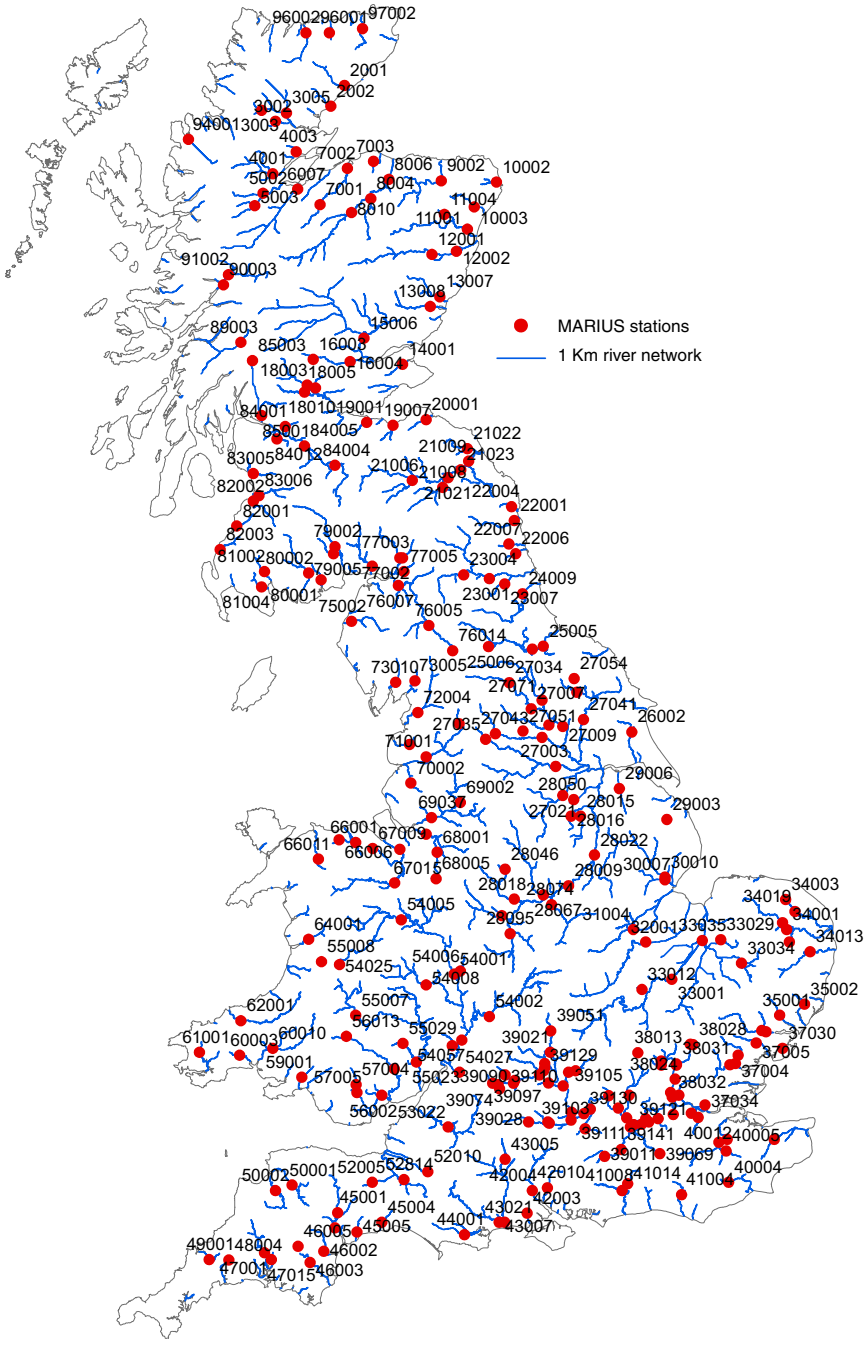

FIGURE 2 Locations of 260 river gauging (MaRIUS) stations across Britain for which daily G2G model flow estimates (m^3/s) are available

eidc.ceh.ac.uk), for non-commercial license as well as internal business use. The formats of the daily and monthly data are described below.

3.1 | Daily data

The MaRIUS-G2G-MORECS-daily (Bell et al., 2018a) and MaRIUS-G2G-Oudin-daily (Bell et al., 2018c) flow data are stored in csv format files (Table 1), with a single header line. The first column is the date, followed by a column for each catchment. The data follow the standard (365- or 366-day) Gregorian calendar. The time is recorded as the calendar date and the flows are mean values from 09:00 to 09:00 on the following day.

The MaRIUS-G2G-WAH2-daily flow data (Bell et al., 2018e) are stored in csv files, one for each catchment (Table 1), with a single header line. The first column is the date, followed by a column for each ensemble member (1–100). The data have 30-day months due to the "360_day" calendar of the climate model data. The time is recorded as

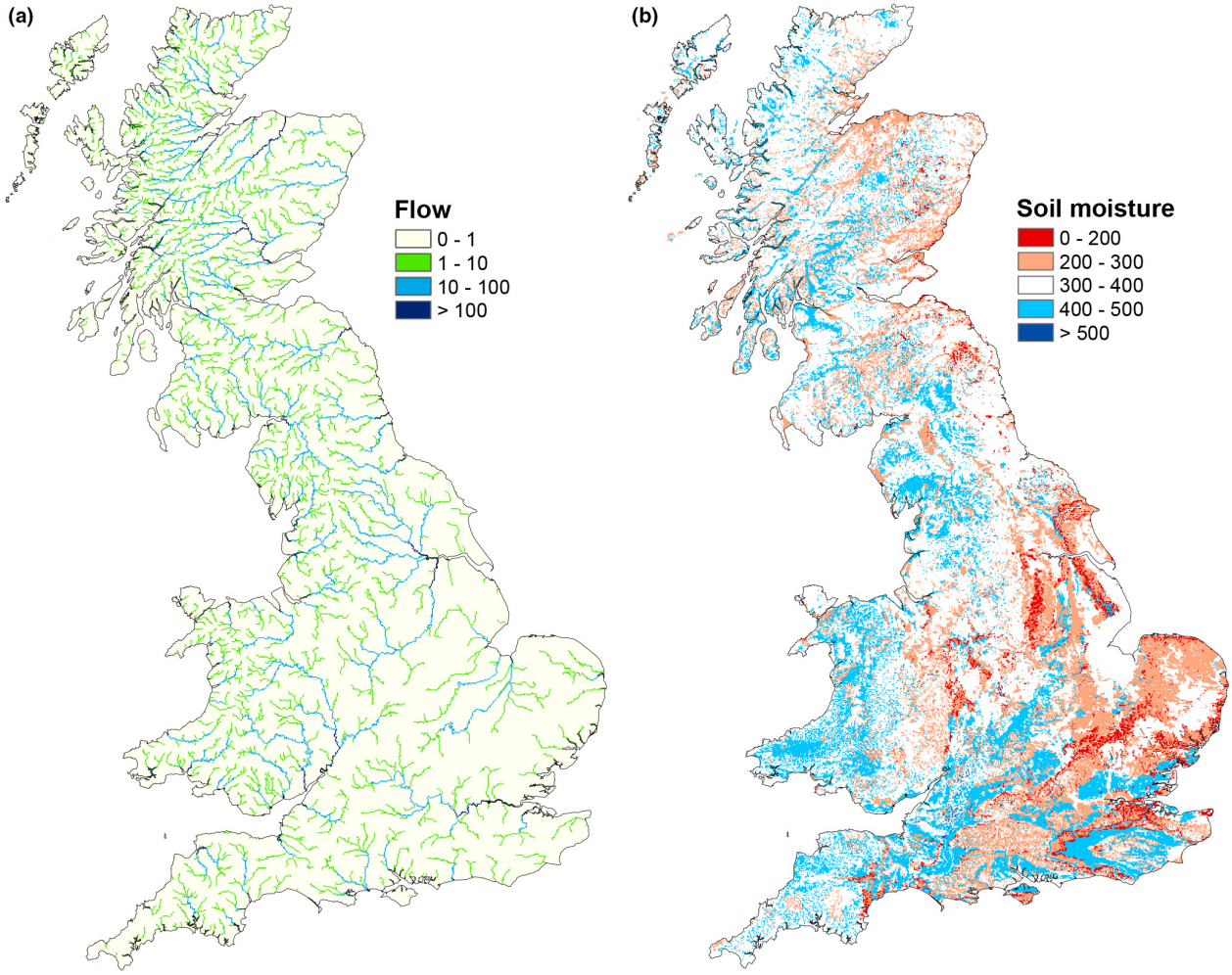

FIGURE 3 Example grids of (a) flow (m^3/s) and (b) soil moisture (mm water/m soil) for Great Britain

TABLE 1 File naming convention for daily outputs

Period	Name of .csv file	Years available
Historical	G2G_MORECS_flow_1960_2015.csv	1960–2015
Historical	G2G_Oudin_flow_1891_2015.csv	1891–2015
Historical baseline	G2G_WAH2_flow_HISTBS_[a].csv	1900–2006 (1900 and 1901 are spin-up)
Baseline	G2G_WAH2_flow_BS_[a].csv	1975–2004
Near-future	G2G_WAH2_flow_NF_[a].csv	2020–2049 (2020 and 2021 are spin-up)
Future	G2G_WAH2_flow_FF_[a].csv	2070–2099 (2070 and 2071 are spin-up)

[a]Is the NRFA catchment number (https://nrfa.ceh.ac.uk/).

the calendar date and the flows are mean values for that day (midnight to midnight).

A related file provides details of the 260 NRFA gauging stations for which a corresponding 1 × 1 km G2G grid box has been selected, including the station number, river name, location, G2G easting, G2G northing, G2G catchment area, and any data issues (e.g., station closed). All G2G catchment areas are within 8% of the NRFA catchment areas for these catchments.

3.2 | Monthly data

MaRIUS-G2G-MORECS-monthly (Bell et al., 2018b) and MaRIUS-G2G-Oudin-monthly (Bell et al., 2018d) 1 × 1 km gridded data are each stored in a single NetCDF4 file (Table 2), following CEH gridded dataset conventions. The data follow the standard (365- or 366-day) Gregorian calendar. The time stamp in the NetCDF files is "days since 1900-01-01" (for MaRIUS-G2G-

TABLE 2 File naming convention for monthly outputs

Period	Name of NetCDF file	Years available
Historical	G2G_MORECS_var_1960_2015.nc	1960–2015
Historical	G2G_Oudin_var_1891_2015.nc	1891–2015
Historical baseline	G2G_WAH2_var_HISTBS[a].nc	1900–2006 (1900 and 1901 are spin-up)
Baseline	G2G_WAH2_var_BS[a].nc	1975–2004
Near-future	G2G_WAH2_var_NF[a].nc	2020–2049 (2020 and 2021 are spin-up)
Future	G2G_WAH2_var_FF[a].nc	2070–2099 (2070 and 2071 are spin-up)

var is either "flow" or "soil".
[a]Is the ensemble member reference number (1–100).

MORECS-monthly) or "days since 1891-01-01" (for MaRIUS-G2G-Oudin-monthly).

The MaRIUS-G2G-WAH2-monthly 1 × 1 km gridded data are stored in NetCDF4 files, as one file for each period and each ensemble member (Table 2). The data have 30-day months due to the "360_day" calendar of the climate model data. The time stamp in the NetCDF files is "days since 1900-01-01".

The monthly data are provided for a 700 × 1,000 km spatial domain on the GB National Grid, from lower left corner (0, 0) to top right (700,000, 1,000,000) (in m). Values for each 1 × 1 km grid box represent the centre of the grid box (i.e., the lower left corner pixel is [500, 500]). G2G values are only provided for land grid boxes and set to missing

(−9,999) in the sea. The monthly values are nominally assigned to the first day of the month. Flow, soil moisture, and time are referenced in the files as "flow," "soil," and "time".

To aid the use of the monthly data, two further datasets are provided:

1. Digitally derived catchment area (km^2) draining to every 1 × 1 km grid box (Davies and Bell, 2018): MaRIUS_G2G_CatchmentAreaGrid.nc.
2. Estimated locations of NRFA gauging stations on the 1 × 1 km grid and as a csv file.

   i The 1 × 1 km grid (MaRIUS_G2G_NRFAStationIDGrid.nc) provides the best locations corresponding to 1,285 gauging stations, referenced by their integer NRFA station number (including the 260 stations for which daily river flow time-series are also provided). At these locations, the G2G flow estimates can be compared to observed (gauged) river flows. The (integer) file format sets ID to 0 for land, −9,999 for sea, and the NRFA station number at gauging station locations.

   ii A file (MaRIUS_NRFAStationIDs.csv) provides details of the 1285 NRFA gauging stations for which a corresponding 1 × 1 km G2G grid box can be selected.

The most appropriate G2G grid cell is identified as the one that is closest in terms of geographical location and catchment area, and additional checks have been undertaken to ensure that the G2G flows are for the correct river tributary, and not for a nearby river with a similar catchment area. Despite these checks, in some cases the derived catchment area draining to the 1 × 1 km river grid cell will be different to the "observed" NRFA catchment area. This problem can particularly affect small catchments for which discretization to a 1 × 1 km grid leads to proportionally larger errors.

4 | DATASET USE AND REUSE

4.1 | Recent and potential uses of the datasets

The observation-driven G2G flow estimates for recent historical periods (e.g., from MaRIUS-G2G-MORECS-daily) can be compared to observed (gauged) river flows (e.g., from the NRFA). Such performance assessments (e.g., Rudd *et al.*, 2017) show that G2G simulates river flows reasonably well, performing best for catchments with a natural flow regime (little anthropogenic influence) and a flow record that is considered accurate, but less well where the regime is influenced by artificial abstractions/discharges and where the subsurface hydrology is unusually complex. Long-term soil moisture observations are less plentiful than observed flow data, and an evaluation of G2G estimates of soil moisture against observations has not yet been undertaken. However, soil moisture is estimated by G2G as a precursor to estimating runoff and flow, and river flows have been subject to evaluation at hundreds of GB river locations (see above). The increasing availability of remotely sensed soil moisture products, such as COSMOS-UK (cosmos.ceh.ac.uk) and ESA CCI (www.esa-soilmoisture-cci.org/node/137), should enable an evaluation of G2G soil moisture estimates soon.

The observation-driven G2G flow and soil moisture estimates for longer historical periods (e.g., MaRIUS-G2G-Oudin-monthly) can be used to identify droughts or floods, and investigate their characteristics. For example, Rudd *et al.* (2017) use the MaRIUS-G2G-Oudin-monthly dataset to show that the threshold level method can be used to identify historic droughts in Britain (1891–2015). They then show that there is substantial spatial and temporal variability in drought characteristics, with groundwater-

dependent areas typically experiencing more severe droughts, but that there are no consistent changes through time for four 30-year time-slices covering the period 1891–2010.

The climate model-driven G2G flow and soil moisture estimates for historical and future periods can be used to investigate potential changes in drought characteristics. For example, AC Rudd (unpublished data) use the MaRIUS-G2G-WAH2-monthly datasets, and show that the severity and intensity of river flow and soil moisture droughts is projected to increase in the future. Droughts in southern and eastern regions are projected to increase in length, and droughts with the largest spatial extent across Britain are projected to increase in area.

The climate model-driven G2G daily river flow estimates for historical and future periods can be used to investigate potential changes in low flow frequency. For example, Kay *et al*. (2015) use the MaRIUS-G2G-WAH2-daily data for four catchments, and show future reductions in low flows which are generally larger in the south of the country and for the later (FF) time-period.

The climate model-driven flow and soil moisture could in theory also be used to investigate potential changes in high flow characteristics. Guillod *et al*. (1997) show that the linear monthly factors used to correct for monthly biases in WAH2 precipitation estimates lead to improved estimates of rainfall during periods of low rainfall, but they note a "small overestimation of dryness at rare frequencies" (>20 years return periods) for long rainfall accumulation times (2–4 years). Their analysis also indicated that bias-corrected WAH2 overestimates high rainfall extremes for 5- to 50-year return periods across GB, with the magnitude of the overestimate varying with location and return period. As MaRIUS-G2G-WAH2 flow and soil moisture datasets are based on bias-corrected WAH2 climate data, it follows that they are also likely to over-estimate high flow and soil moisture extremes when compared to observations, however an analysis of relative change between current and future periods may be unaffected.

The G2G data can also be used in other modelling, for example, to investigate impacts related to agriculture, ecology, or economics.

The MaRIUS-G2G-WAH2-daily data have been used for a catchment-explicit national risk assessment of possible future economic impacts of abstraction restrictions on irrigated agriculture in England and Wales due to drought management decisions (G Salmoral unpublished manuscript). The study evaluates the frequency, severity, and duration of abstraction restrictions following a risk-based analysis of economic losses using rainfall and river flow for the BS, NF, and FF. It shows how, for a set of rainfall and river flow triggers in line with those applied by current environmental regulators, different ranges of economic losses are obtained due to the related changing climate and crop specific conditions.

To assess the ecological impacts of drought on river habitat availability, observation-driven and climate model-driven G2G flow estimates (MaRIUS-G2G-MORECS-monthly and MaRIUS-G2G-WAH2-monthly) have been used in combination with hydraulic geometry estimates (e.g., mean river depth) for several hundred sites across England and Wales to assess the effect of low flows on loss of river habitat and longitudinal connectivity (Laize *et al*., 1965). There is also the potential to use the flow estimates to evaluate the resilience of wetlands to droughts. For example, river-fed wetland ecosystem models could use climate model-driven G2G flow estimates (MaRIUS-G2G-WAH2-daily) to evaluate potential hydro-ecological impacts under climate change. Similar work is being undertaken in MaRIUS using WAH2 RCM data as an input to rain-fed wetland ecosystem models (www.mariusdrought project.org).

The observation-driven G2G soil moisture estimates for recent historical periods (MaRIUS-G2G-MORECS-monthly) have been used to support studies of the effect of recent climatic changes on the phenology of arboviruses (viruses transmitted by arthropods), which have a preference for moist or semi-aquatic habitats (C Sanders unpublished data).

4.2 | How (and how not) to use the climate model-driven G2G datasets

MaRIUS-G2G-WAH2-monthly or MaRIUS-G2G-WAH2-daily data for the BS time slices (HISTBS and BS) can be compared to estimates from G2G driven by observational input data (e.g., from MaRIUS-G2G-Oudin-monthly or MaRIUS-G2G-Oudin-daily), or to observed data (e.g., river flows from the NRFA). However, comparisons in either case should only be made statistically, not by time-series equivalence. For example, WAH2 BS river flows for 1976 will not directly resemble observed reality in 1976; only statistics over long (multi-decadal) periods should be compared (e.g., mean monthly flows, or flow duration curves). Comparison of climate model-driven G2G simulations to an observation-based G2G run will indicate how biases in the WAH2 data affect the results for the BS periods; comparison to observational data themselves will be additionally affected by the accuracy of the G2G model simulations.

G2G outputs for future time slices (NF and FF) can be compared to the BS time slice estimates, NOT to observed time series or G2G simulations with inputs of observed precipitation and PE. Similarly, results from impacts models (e.g., economic, ecological, agricultural) based on G2G outputs for future time slices (NF and FF) should be

compared to those from the BS time slice, NOT to observed "real world" impacts.

Each of the 100 historical (HISTBS and BS) and future (NF and FF) ensemble members are plausible realizations of the climate of these periods, and analyses of projected future changes should look at differences between historical and future statistical distributions, rather than between individual ensemble members.

Although each of the historical (HISTBS and BS) and future (NF and FF) ensemble members is numbered from 1 to 100, historical and future ensemble members with the same ensemble number bear no relation to each other and should not be directly compared. Thus, flows from BS1 (Baseline ensemble member 1) should not be directly compared to NF1 (Near-Future ensemble member 1), any more than they should to NF2 or NF35.

OPEN PRACTICES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at DOI 10.5285/5f3c1a02-d5c4-4faa-9353-e8b68ce2ace2. Learn more about the Open Practices badges from the Center for Open Science: https://osf.io/tvyxz/wiki.

ACKNOWLEDGEMENTS

This data is an outcome of MaRIUS: Managing the Risks, Impacts and Uncertainties of droughts and water Scarcity, funded by the UK Natural Environment Research Council's Drought and Water Scarcity programme (NE/L010208/1). We thank the Met Office National Climate Information Centre for the 5 km temperature data, and Nuria Bachiller-Jareno, Matt Fry, Maliko Tanguy and Oliver Robertson for advice.

ORCID

Victoria A. Bell http://orcid.org/0000-0002-0792-5650
Alison L. Kay http://orcid.org/0000-0002-5526-1756
Alison C.Rudd http://orcid.org/0000-0001-5996-6115
Helen N.Davies http://orcid.org/0000-0001-7783-4853

REFERENCES

Bell, V.A., Kay, A.L., Jones, R.G. and Moore, R.J. (2007) Development of a high resolution grid-based river flow model for use with regional climate model output. *Hydrology and Earth System Sciences*, *11*(1), 532–549.

Bell, V.A., Rudd, A.C., Kay, A.L., Davies, H.N. (2018a). Grid-to-Grid model estimates of daily mean river flow for gauged catchments in Great Britain (1960 to 2015): observed driving data [MaRIUS-G2G-MORECS-daily]. *NERC Environmental Information Data Centre*. https://doi.org/10.5285/5f3c1a02-d5c4-4faa-9353-e8b68ce2ace2

Bell, V.A., Rudd, A.C., Kay, A.L., Davies, H.N. (2018b). Grid-to-Grid model estimates of monthly mean flow and soil moisture for Great Britain (1960 to 2015): observed driving data [MaRIUS-G2G-MORECS-monthly]. *NERC Environmental Information Data Centre*.

Bell, V.A., Rudd, A.C., Kay, A.L., Davies, H.N. (2018c). Grid-to-Grid model estimates of daily mean river flow for gauged catchments in Great Britain (1891 to 2015): observed driving data [MaRIUS-G2G-Oudin-daily]. *NERC Environmental Information Data Centre*. https://doi.org/10.5285/0ceb4f85-0bbf-49f0-ab70-cfc137ab7d4d

Bell, V.A., Rudd, A.C., Kay, A.L., Davies, H.N. (2018d). Grid-to-Grid model estimates of monthly mean flow and soil moisture for Great Britain (1891 to 2015): observed driving data [MaRIUS-G2G-Oudin-monthly]. *NERC Environmental Information Data Centre*. https://doi.org/10.5285/f52f012d-9f2e-42cc-b628-9cdea4fa3ba0.

Bell, V.A., Rudd, A.C., Kay, A.L., Davies, H.N. (2018e). Grid-to-Grid model estimates of daily mean river flow for gauged catchments in Great Britain: weather@home2 (climate model) driving data [MaRIUS-G2G-WAH2-daily]. *NERC Environmental Information Data Centre*. https://doi.org/10.5285/f6cac471-7d92-4e6d-be8a-9f7887143058.

Bell, V.A., Rudd, A.C., Kay, A.L., Davies, H.N. (2018f). Grid-to-Grid model estimates of monthly mean flow and soil moisture for Great Britain: weather@home2 (climate model) driving data [MaRIUS-G2G-WAH2-monthly]. *NERC Environmental Information Data Centre*. https://doi.org/10.5285/3b90962e-6fc8-4251-853e-b9683e37f790

Bell, V.A., Kay, A.L., Jones, R.G., Moore, R.J. and Reynard, N.S. (2009) 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.

Bell, V.A., Kay, A.L., Davies, H.N. and Jones, R.G. (2016) An assessment of the possible impacts of climate change on snow and peak river flows across Britain. *Climate Change*, *136*(3), 539–553.

Cole, S.J. and Moore, R.J. (2009) Distributed hydrological modelling using weather radar in gauged and ungauged basins. *Advances in Water Resources*, *32*(7), 1107–1120.

Davies, H.N. and Bell, V.A. (2009) Assessment of methods for extracting low-resolution river networks from high-resolution digital data. *Hydrological Sciences Journal*, *54*(1), 17–28.

Formetta, G., Prosdocimi, I., Stewart, E. and Bell, V. (2018) Estimating the index flood with continuous hydrological models: an application in Great Britain. *Hydrology Research*, *49*(1), 123–133.

Guillod, B.P., Jones, R.G., Kay, A.L., Massey, N., Sparrow, S., Wallom, D , Coxon, G., Bussi G., Freer, J., Kay, A.L., Massey, N.R., Sparrow, S.N., Wallom, D.C.H., Allen, M.R. and Hall, J.W. (2017) Managing the Risks, Impacts and Uncertainties of drought and water Scarcity (MaRIUS) project: Large set of potential past and future climate time series for the UK from the weather@home2 model. Didcot, UK: Centre for Environmental Data Analysis, https://doi.org/10.5285/0cea8d7aca57427fae92241348ae9b03.

Guillod, B.P., Jones, R.G., Dadson, S.J., Coxon, G., Bussi, G., Freer, J, Massey, N.R., Sparrow, S., Wallom, D.C.H. and Wilson, S.S. (2018) A large set of potential past, present and future hydro-

meteorological time series for the UK. *Hydrology and Earth System Sciences*, *22*(1), 611–634.

Hough, M. and Jones, R.J.A. (1997) The United Kingdom Meteorological Office rainfall and evaporation calculation system: MORECS version 2.0 – an overview. *Hydrology and Earth System Sciences*, *1*(2), 227–239.

Kay, A.L. and Davies, H.N. (2008) Calculating potential evaporation from climate model data: a source of uncertainty for hydrological climate change impacts. *Journal of Hydrology*, *358*, 221–239.

Kay, A.L., Bell, V.A., Guillod, B.P., Jones, R.G., Rudd, A.C. (2018) National-scale analysis of low flow frequency: historical trends and potential future changes. *Climate Change*, *147*(3–4), 585–599.

Keller, V.D.J., Tanguy, M., Prosdocimi, I., Terry, J.A., Hitt, O., Cole, SJ, Terry, J.A., Hitt, O., Cole, S.J., Fry, M., Morris, D.G. and Dixon, H. (2015) CEH-GEAR: 1 km resolution daily and monthly areal rainfall estimates for the UK for hydrological and other applications. *Earth System Science Data*, *7*, 143–155.

Laize, C., Nineham, N., Dunbar, M., Hannah, D., Bell, V. (2017) Loss of habitat and connectivity during drought. In: Krause, S (Ed) HydroEco 2017: 6th International Multidisciplinary Conference on Hydrology and Ecology, University of Birmingham, Birmingham, UK, 18–23 June 2017 (pp. 148-149). Available at: https://www.birmingham.ac.uk/generic/hydroeco2017/attending/Hydroeco2017-conference-programme.pdf. [Accessed 09 October 2018]

Monteith, J.L. (1965) Evaporation and environment. *Symposia of the Society for Experimental Biology*, *19*, 205–234.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F and Loumagne, C.. (2005) Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2 – towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling. *Journal of Hydrology*, *303*, 290–306.

Perry, M. and Hollis, D. (2005) The generation of monthly gridded datasets for a range of climatic variables over the UK. *International Journal of Climatology*, *25*, 1041–1054.

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann G., Nakicenovic N. and Rafaj P. (2011) RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climate Change*, *109*, 33–57.

Robinson, E.L., Blyth, E., Clark, D.B., Finch, J. and Rudd, A.C. (2017) Trends in evaporative demand in Great Britain using high-resolution meteorological data. *Hydrology and Earth System Sciences*, *21*(2), 1189–1224.

Rudd, A.C., Bell, V.A. and Kay, A.L. (2017) National-scale analysis of simulated hydrological droughts (1891–2015). *Journal of Hydrology*, *550*, 368–385.

Rudd, A.C. and Kay, A.L. (2016) Use of very high resolution climate model data for hydrological modelling: estimation of potential evaporation. *Hydrology Research*, *47*(3), 660–670.

Speight, L., Cole, S.J., Moore, R.J., Pierce, C., Wright, B., Golding, B, Cranston M., Tavendale A., Ghimire S. and Dhondia J. (2016) Developing surface water flood forecasting capabilities in Scotland: an operational pilot for the 2014 Commonwealth Games in Glasgow. *J. Flood Risk Management*, *11*, 884–901.

Tanguy, M., Dixon, H., Prosdocimi, I., Morris, D.G., Keller, V.D.J. (2016) Gridded estimates of daily and monthly areal rainfall for the United Kingdom (1890–2015) [CEH-GEAR]. Wallingford, UK: NERC Environmental Information Data Centre, https://doi.org/10.5285/33604ea0-c238-4488-813d-0ad9ab7c51ca.

How to cite this article: Bell VA, Kay AL, Rudd AC, Davies HN. The MaRIUS-G2G datasets: Grid-to-Grid model estimates of flow and soil moisture for Great Britain using observed and climate model driving data. *Geosci Data J*. 2018;5:63–72. https://doi.org/10.1002/gdj3.55