Improved Hydrology for Regional Environmental Prediction

Elizabeth Cooper, Alberto Martínez-de la Torre, Toby Marthews,
Rich Ellis, Alison Kay, Matthew Wiggins, Simon Dadson,
Ponnambalam Rameshwaran, Nick Reynard and Douglas Clark

UK Centre for Ecology and Hydrology,
Wallingford.

May 2022

Summary

In this report we document work aimed at improving the quality of the representation of land and rivers in multi-year, coupled atmosphere-land (UM-JULES) simulations over the British Isles. The approach taken was to use standalone (uncoupled) simulations of JULES to investigate the potential to improve the coupled system.

The use of alternative soil ancillary information, generated by using a data assimilation framework and observations of soil moisture from the COSMOS-UK network to optimise the constants in a pedotransfer function, was found to result in improved simulations by JULES of river flow in a diverse sample of British rivers (as measured by standard statistics). The revised soil parameters tended to increase the variability of the simulated river on short timescales, and reduce variability on the annual timescale. In catchments with a large influence of slow baseflow the revised parameters tended to give poorer simulations, with too much variability on short timescales.

A new representation of groundwater processes was implemented in JULES and applied, for the first time, across Great Britain. This was shown to allow an influence of groundwater on near-surface hydrology and fluxes over large parts of the country. Modelled river flows were more realistic in many cases, though much of the improvement was due to differences in the representation of runoff generation rather than the introduction of groundwater.

A more physically-complete parameterisation of river physics, using the local inertial equation, was applied in a range of catchments, again for the first time. Simulated river flows were improved in most cases.

The final section of this report offers a perspective on how terrestrial hydrology and its impacts could be considered in the next generation of UK Climate Projections.

The work reported here has significantly improved our capability in several areas of land surface modelling, to the extent that new developments could be tested in nationwide simulations of JULES. In each area the results are encouraging and further development will continue in the Hydro-JULES project.

UK Centre for
Ecology & Hydrology

Contents

1 Introduction

The Regional Environmental Prediction Project is a Met Office R&I Strategy flagship activity, under the Seamless Environmental Prediction and Capturing Environmental Complexity themes, to develop regional coupled prediction capabilities with the aim of improving predictions across kilometre-scale atmosphere, land and marine components through more explicitly representing the feedbacks and interactions in the environment.

The existing implementation of the regional coupled system over the British Isles includes the coupling between hydrology and ocean in season-length runs. Although successful, this first demonstrator needs further improvements in the quality of the river outflow into the ocean. Developments underway will also lead to prototype climate runs with a regional atmosphere-land-ocean coupled system, which do not currently include river hydrology.

The Hydro-JULES project is a NERC-funded research programme with a wide range of activities across the terrestrial water cycle, to underpin hydrological research in the UK. Hydro-JULES is delivered by UKCEH in partnership with BGS and NCAS. Included in the project are several developments to the JULES land surface model.

This report summarises work at UKCEH aimed at improving the quality of the representation of land and rivers in multi-year, coupled atmosphere-land (UM-JULES) simulations over the British Isles. Key developments to JULES from Hydro-JULES are considered. The approach taken was to use standalone (uncoupled) simulations of JULES to investigate the potential to improve the coupled REP system. Ultimately this work can inform the design of the next generation of products from UK Climate Projections (hereinafter UKCPnext).

Specifically in this project we:

- Generated alternative soil ancillary (characteristic) data and assessed the impact on modelled hydrology
- Introduced a representation of groundwater processes
- Investigated the use of a more physically-complete representation of river flow and inundation
- Considered how hydrological impacts should be included in UKCPnext.

2 JULES configuration, data, and codes

2.1 JULES configurations

The Joint UK Land Environment Simulator (JULES) is described in Best et al. (2011) and Clark et al. (2011), while descriptions of many recent developments can be found on the JULES website.

In this work we created JULES configurations and suites based on RAL3M (as implemented in suite mi-bc095). The RAL3M configuration is a recent update on the Regional Atmosphere and Land configuration RAL1 (Bush et al., 2020). The grid used is a variable-resolution rotated pole grid; the JULES simulations described below considered the central constant-resolution part of the domain in which the resolution is approximately 2.2km. This constant-resolution grid has size 532 x 654, and is referred to as the REP grid below. In terms of land hydrology, RAL3M uses a TOPMODEL-based parameterisation of runoff generation.

The branches of JULES that were used and suites are referenced in each section below, but a common starting point was the JULES branch vn6.2_daily and the JULES suite u-ci969. This JULES branch is based on vn6.2 with a development from this project to make it easier to specify daily input files (daily templating, which was subsequently adopted for vn6.3).

Arguably the largest differences between the UM and JULES runs concern the timestep length and the, inevitably, different ways in which near-surface meteorology are represented. The UM used a timestep of 60 s, whereas JULES was run with a timestep of 1800 s. This longer JULES timestep length was used largely because of the limitation already present through the use of hourly precipitation data; there would likely be little benefit in using a shorter timestep with the same input data. By contrast, there is likely to be significant variability of precipitation in the UM on timescales shorter than one hour, which will affect the simulation of surface fluxes (including runoff). The influence of such differences in setup between the UM and JULES are assessed in Section 2.5 below.

2.2 UM data

We were provided with ancillary files for RAL3M that were used with UM suite mi-bc095, an initial state file derived from the UM, and meteorological and hydrological outputs for 2012–2017 were available via MASS.

Table 1 lists the variables retrieved and the available frequencies. PP files were converted to netCDF files suitable for use with JULES using the iris python library. Two sets of driving data were created: meteorological to drive the full JULES model and hydrological to drive a river routing only version of JULES. To create a driving dataset with hourly values, where a variable was only available as 3-hourly, data the values were repeated for each of the 3 hours.

Variable	Frequency	STASH code	Stream
Shortwave	3hr	1235	apd.pp
Longwave	3hr	2207	apd.pp
Rainfall	1hr	4203	apq.pp
Snowfall	1hr	4204	apq.pp
Temperature at 1.5m	3hr	3236	apd.pp
Wind components at 10m	1hr	3225 (u), 3226 (v)	apd.pp
Surface pressure	3hr	0409	apd.pp
Specific humidity at 1.5m	3hr	3237	apd.pp
Surface runoff	3hr	8234	apd.pp
Subsurface runoff	3hr	8235	apd.pp

Table 1: UM variables retrieved from MASS.

River routing ancillary data and code

The river ancillary data used in this work are essentially those of Davies et al. (2022), with minor modifications and reformatting to allow their use with JULES. Those are 1km data on a grid that is essentially that of the CHESS OSGB dataset (albeit over a larger domain). JULES cannot currently regrid runoff from the REP grid to the 1km OSGB grid, hence we had to identify an alternative approach. Initially we tried to use UniFHy (Hallouin et al. (2021); this is the modelling framework being developed in the Hydro-JULES project) but that is also currently unable to support this combination of grids. In the end we used a preprocessing stage to regrid the runoff fields (produced by JULES at 2.2km) and then used a 'rivers only' branch of JULES to perform the routing. Surface and sub-surface runoff outputs from the UM or JULES were regridded to this 1km using IRIS functionality.

The rivers-only branch (r21857_camaflood) is the latest in a line of branches that allow simulations in which only river routing is active, driven by runoff fields. (This is similar to the rivers executable recently introduced to JULES by ticket 1084 development.) This branch was used to route flows using RFM; a separate branch was used for CaMa-Flood (see later).

The River Flow Model (RFM) implements a finite difference solution to the 1-D kinematic wave equation - see Appendix B of Lewis et al. (2018). The parameter values used with RFM were chosen to match those used in previous work by the Met Office (see Table C3 of Lewis et al. (2018)) and are listed in Table 2.

Parameter	Value
a_thresh	13.0
cland	0.4
criver	0.5
cbland	0.05
cbriver	0.05
retr	0.005
retl	0.005

Table 2: Parameter values for RFM.

2.3 Measured flow amounts

We used daily flow data from the National River Flow Archive (NRFA). The bulk of our analysis was based on the 191 stations in Great Britain that were selected by the eFLaG project (Hannaford et al., 2022) on the basis of offering national-scale representivity (of climate and catchment characteristics such as soils), spatial distribution, and a balance of 'near-natural' and disturbed catchments (80 of the catchments are also included in the near-natural Benchmark compilation of Harrigan et al. (2017). The locations covered are shown in Fig.3a of Hannaford et al. (2022). For each station the most appropriate location on the model river grid for comparison was chosen by starting at the declared station location and considering neighbouring locations that best matched the declared catchment area.

Where a smaller number of stations is used for illustration, and for compatibility with previous work, we also make use of the 13 catchments of Martínez-de la Torre et al. (2019), some of which are not in the eFLaG dataset. These catchments and some basic characteristics are listed in Table 3. On occasions a subset of 11 was used, excluding stations 27304 and 54001 (which are upstream of 27009 and 54057 respectively).

Gauge number	Gauge name	BFI	Area km^2
12002	Dee at Park	0.53	1844
15006	Tay at Ballathie	0.65	4587
27009	Ouse at Skelton	0.45	3315
27034	Ure at Kilgram Bridge	0.32	510
27041	Derwent at Buttercrambe	0.70	1586
39001	Thames at Kingston	0.63	9948
39081	Ock at Abingdon	0.65	234
43012	Avon at Knap Mill	0.86	1706
47001	Tamar at Gunnislake	0.46	917
54001	Severn at Bewdley	0.53	4325
54057	Severn at Haw Bridge	0.57	9895
71001	Ribble at Samlesbury	0.34	1145
84013	Clyde at Daldowie	0.46	1903.1

Table 3: Gauge numbers, names, Baseflow Index and catchment area reported in the NRFA for the 13 catchments selected by Martínez-de la Torre et al. (2019).

2.4 Calculation of river flow metrics

We considered various existing codes for the comparison of observed and measured river flows before concluding that we would develop a new program. This new evaluation code uses the hydroeval API for Python (Hallouin, 2021) which gives access to a suite of evaluation metrics, including Nash-Sutcliffe Efficiency and the Original Kling-Gupta Efficiency. The Base Flow Index (BFI) was also generated using a methodology from Gustard et al. (1992). Times with missing data within the observation set were removed from the modelled data before metrics were calculated. In addition to calculating metrics for each station, the code calculates and plots summary statistics for the dataset as a whole, and plots hydrographs for each station. This tool has a good level of functionality and has improved our capability to analyse modelled river flows (in particular with NRFA data), but it is our expectation that it will continue to be developed (e.g. through the Hydro-JULES project) to provide enhanced capability in future.

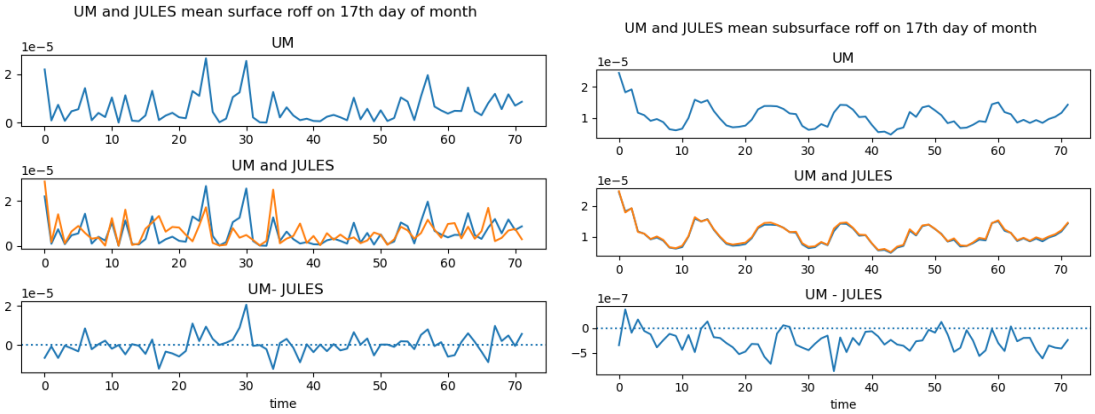
2.5 Comparison of results from JULES and the UM

In this section we test the assumption that runs of standalone JULES offer insight into the performance of JULES in coupled mode. We compare diagnostics from the UM (suite mi-bc095) with corresponding outputs from standalone JULES (suite u-ck812).

Figure 1 shows surface and subsurface runoff from the UM and JULES for a day in the middle of each month. There is no systematic increase in the difference between these as the simulation time increases, i.e. no observable drift through the simulation period.

Figure 1b shows that the differences in subsurface runoff are very small; Fig. 1a shows relatively larger differences between the modelled surface runoff fields. This may be a consequence of the different timesteps used in the models, and in particular the differences in rainfall intensity that might be expected between the 60 s timestep of the UM and the hourly meteorological data that

are used to drive JULES. This level of similarity seen in domain-average fluxes was generally also seen in spatial patterns and time series at selected points – UM and JULES results tended to be similar but not identical.



(a) surface runoff

(b) subsurface runoff

Figure 1: Comparison of domain-averaged surface and sub-surface runoffs for UM and JULES standalone runs. Values on the x-axis correspond to number of months through the simulation.

We also compared modelled river flows, which provide a measure integrated over the catchment. River routing was not active in the UM run and to calculate river flow we used UM daily runoff fields as input to a rivers-only configuration of JULES that used the RFM parameterisation (see Section 2.2). Figure 2 shows the calculated daily flows for JULES against the equivalent UM outputs for the catchments used in Martínez-de la Torre et al. (2019) and described in table 3. In general there is a reasonable level of agreement between the models.

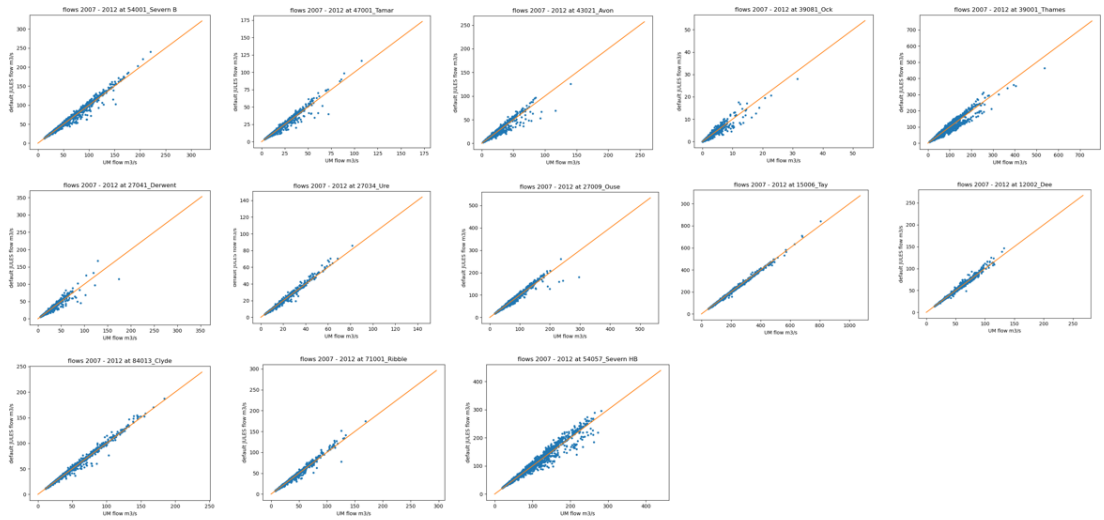


Figure 2: Daily flow values for default JULES runs plotted against daily flow values for routed UM flow outputs.

Figure 3 shows the correlation coefficient (r) for UM and JULES riverflow outputs compared to observations in 191 catchments. The correlation coefficient for the UM is lower in almost all catchments - possibly reflecting a timing mismatch. Note also that NRFA daily flows generally refer to the 24-hour period ending at 09H GMT each day (or 00H GMT at a minority of stations), whereas both UM and JULES outputs are averages ending at 00H GMT. This will affect the calculated statistics, more so in smaller catchments. Overall JULES and the UM show similar values for r in many catchments, though JULES does seem to perform better than the UM in catchments with large BFI – possibly relating to increased surface runoff generated by higher rainfall intensities in the UM.

In this section we have established that the runoff fields from simulations of JULES, forced by UM meteorology, look broadly similar to those from the UM when similar configurations and

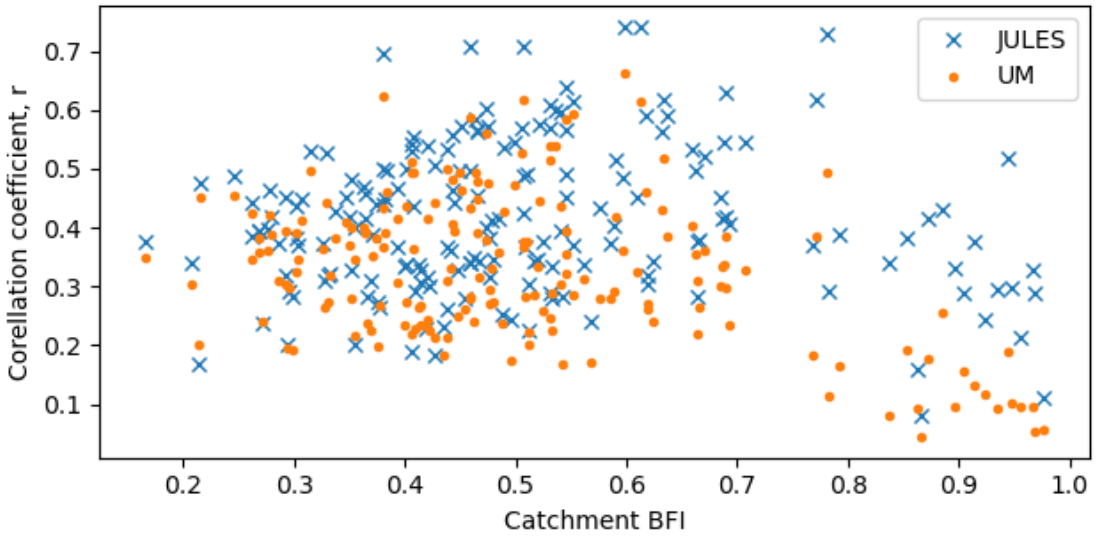


Figure 3: Correlation coefficients between JULES (blue) and UM (orange) streamflow measurements vs NRFA observations for 191 catchments.

ancillary files are used. We conclude that the standalone configuration would appear to be correct, and that standalone JULES can indeed provide insight into the behaviour of the land surface in the coupled model.

3 Alternative soil ancillary data

3.1 Introduction to soil ancillary work

Work within the Met Office (e.g. presentation by Chris Short, Unified Model User Workshop, June 2021) and in Hydro-JULES (Pinnington et al., 2021; Cooper et al., 2021) has shown that JULES soil ancillary data can have an important impact on atmospheric and land states. Here we describe three different soil ancillary files and compare the impacts of using them in standalone JULES runs which closely match the RAL3 MetOffice configuration. Soil ancillaries contain the soil physics parameters necessary for JULES to represent soil hydraulics processes. These soil physics parameters are generated for each model grid cell using an underlying soil texture database and a set of pedotransfer functions (PTF). In this work we use the Brooks and Corey option in JULES for soil hydraulic processes, and we use the Cosby PTF (Cosby et al., 1984), although there are a large number of alternative options as in, e.g. Marthews et al. (2014) and Tóth et al. (2015). The three soil ancillary files we compare are:

- Default (as provided by UKMO); uses Harmonised World Soil Database (HWSD) soil texture data (Database) and standard Cosby PTF)
- SoilGrids; uses SoilGrids topsoil texture data (Hengl et al., 2017) and standard Cosby PTF
- Optimised uses HWSD soil texture data and a modified Cosby PTF in which the standard PTF constants have been optimised using COSMOS-UK soil moisture data (Stanley et al., 2021) and techniques from Pinnington et al. (2021) and Cooper et al. (2021).

The soil ancillary options are described further in the next section, while in Section 3.3 we compare the impact of the different ancillaries on various aspects of JULES.

3.2 Alternative soil ancillary data

Default soil ancillary file

This file is the soil data ancillary provided by the Met Office. It is based on HWSD soil textures and the standard Cosby PTF.

SoilGrids soil ancillary file

This file uses the SoilGrids 250m soil texture data (Hengl et al., 2017) regridded to the REP rotated pole grid (UKCP grid). This texture data is used in a standard Cobsy PTF to generate the soil physics values in the ancillary file. The SoilGrids textures are more smoothly varying than the HWSD equivalent over the GB domain; Fig. 4 shows the percentages of sand per grid in the HWSD database (LH column), SoilGrids database (middle column) and the difference between the two (RH column).

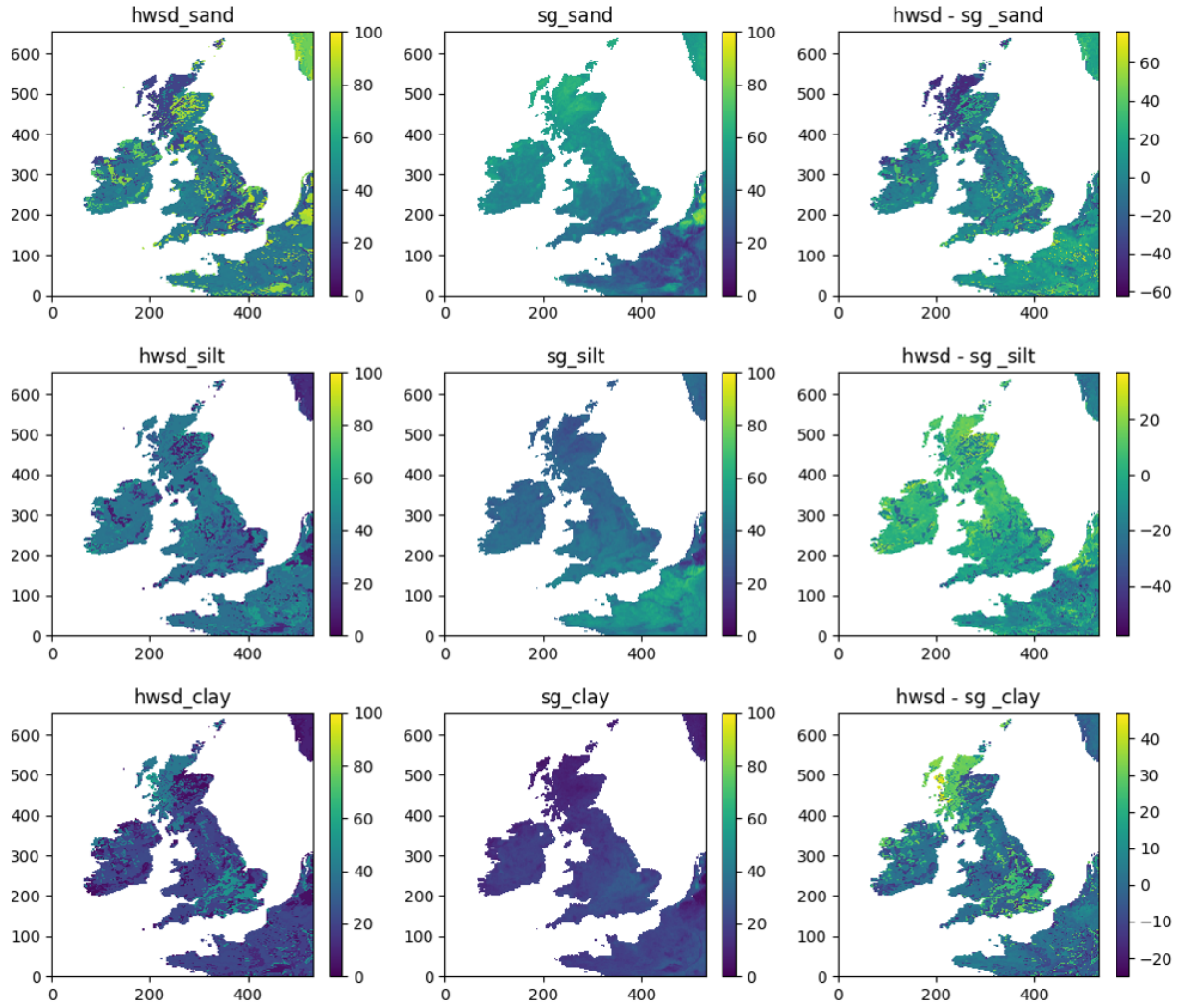


Figure 4: Topsoil sand fractions for hwsd and soilgrid databases.

Optimised HWSD ancillary file

For this ancillary file we used HWSD soil textures. We used the lavendar data assimilation framework (Pinnington, 2021; Pinnington et al., 2020) and soil moisture observations from the COSMOS-UK network to optimise the constants in the Cosby PTF, following the approach in Cooper et al. (2021). We used daily soil moisture observations from 16 COSMOS-UK sites throughout 2017 in the data assimilation algorithm. The corresponding 2017 JULES runs were driven with observed meteorology from each COSMOS-UK met station. It would have been preferable to do the optimisation runs over the period of interest (2007 - 2012) but the COSMOS-UK network only began operating in 2014.

Figure 5 shows soil moisture time series at the 16 sites used in this study for 2017 and 2018. The plots show that at all sites RAL-JULES is able to better match the soil moisture observations (pink dots) after data assimilation (posterior runs; yellow line) than before data assimilation (prior runs; blue line). Note that all of the observations in 2017 are used to optimise the PTF constants; these PTF constant values are then used at all sites.

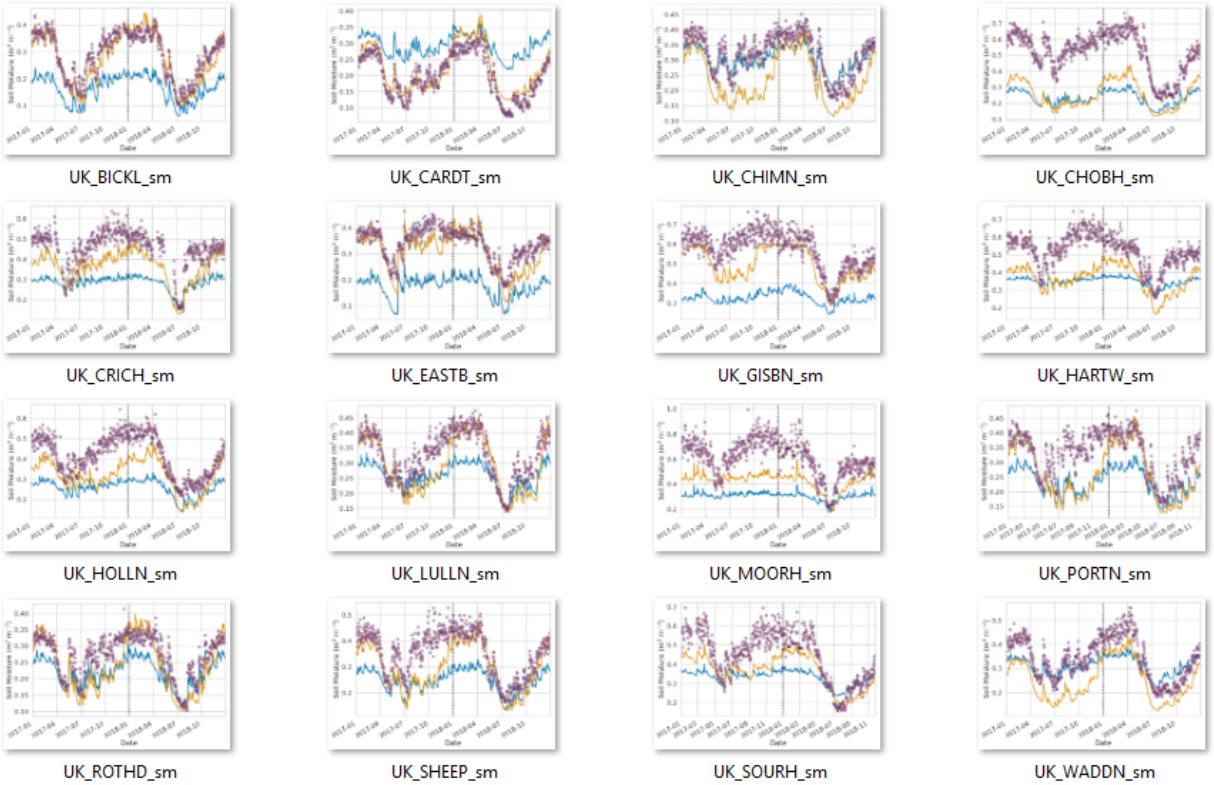


Figure 5: Soil moisture time series at 16 COSMOS-UK sites used in study. Pink dots are COSMOS observations, blue lines are initial RAL-JULES soil moisture outputs, yellow lines are RAL-JULES outputs after data assimilation

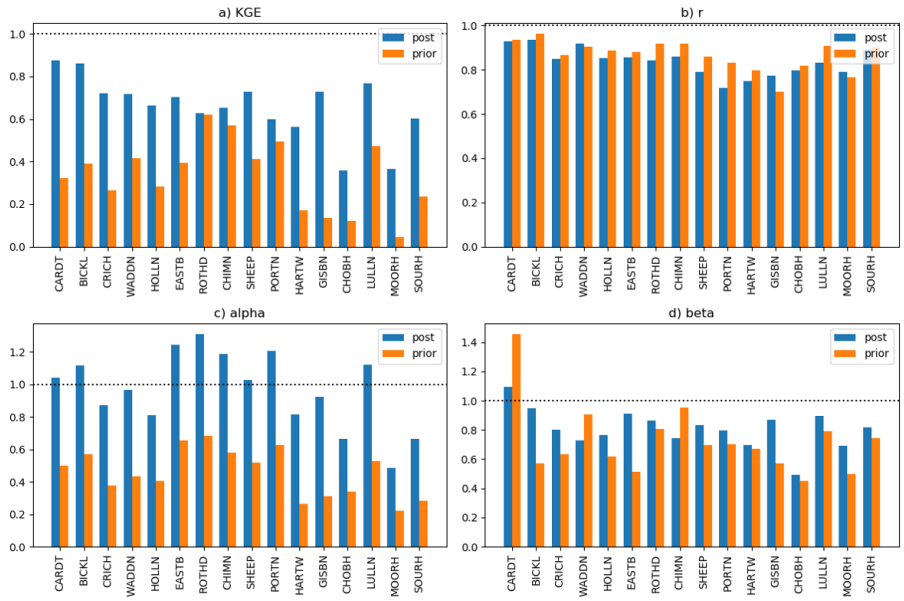


Figure 6: Kling-Gupta efficiency metric and components for the timeseries shown in Fig. 5. The ideal value of KGE, R, alpha and beta is 1; this is shown on each plot with a dotted line.

Figure 6 shows Kling-Gupta Efficiency (KGE) metrics for the times series shown in Fig. 5. As in previous work, optimisation of the constants in the Cosby PTF leads to larger KGE values, representing a better fit to the soil moisture observations at all 16 sites. Figure 6 also shows that in agreement with previous work, the correlation coefficient, r, is not greatly impacted by the optimisation. The posterior value of beta is closer to the ideal value of 1 than the prior value at the majority of sites; this represents a decrease in bias between the model and the observations. The largest improvement following optimisation is in the alpha component of the KGE. The alpha component represents how well the spread in the model matches the spread in the observations and the optimisation acts to increase the spread in the JULES soil moisture outputs to better match the observations.

3.2.1 Comparison of ancillary files

Default (HWSD) and SoilGrids files

Figure 7 shows the b soil physics parameter plotted over the REP domain for the default soil ancillary file (HWSD) and the SoilGrids ancillary (SG). Figure 7 illustrates the fact that the smoother soil textures in the SoilGrids database result in more smoothly varying soil physics parameter files over the domain. Figure 7 shows that the SoilGrids values for the b parameter have a narrower range than those generated using the HWSD soil textures, with no b values at the higher end of the HWSD range.

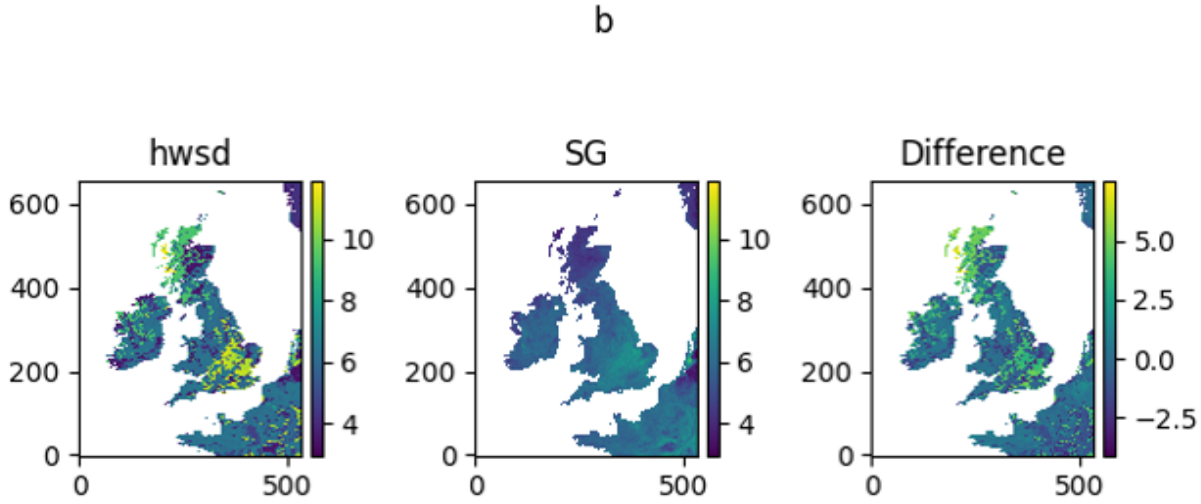


Figure 7: Soil physics b parameter plotted over the REP domain for the default (HWSD) runs, the SoilGrids (SG) runs, and the difference between them. The spatially smoother SoilGrids texture fields lead to smoother soil physics parameter variations.

Figure 8 shows histogram plots of the soil physics parameters used in the default runs (blue) and the SoilGrids runs (orange).

Default and optimised HWSD ancils

Figure 9 shows histogram plots of the soil physics parameters used in the default runs (blue) and the optimised runs (orange). Figure 9 shows that the range of values for b and satcon is very much narrower in the optimised ancils; we also see that the satcon values are generally smaller and the sm_{sat} values are much bigger. These changes acts to allow the soil to become wetter at saturation, and slow down the movement of https://www.overleaf.com/project/62345be93929e26f74a7b937water through the soil.

3.3 Impact of alternative soil ancillary data

3.3.1 Heat fluxes

The different soil ancillary files produce different values of soil moisture, and this influences the pattern of latent and sensible heat fluxes. There is very little difference in heat fluxes between the

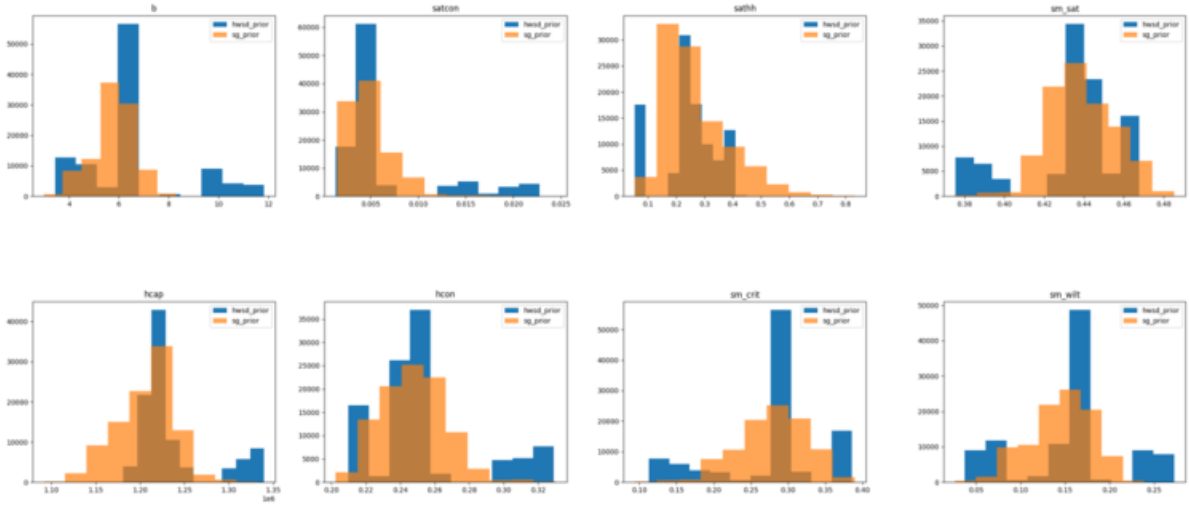


Figure 8: The eight soil physics parameter values showing the difference in the range of values generated for the default runs (blue) and the SoilGrids runs (orange).

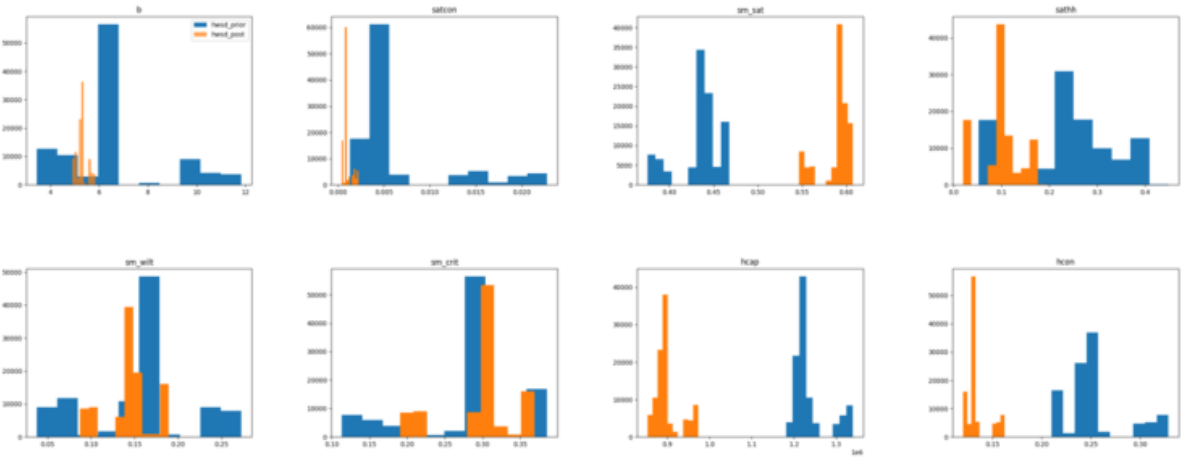


Figure 9: The eight soil physics parameter values showing the difference in the range of values generated for the default runs (blue) and the optimised ancillary runs (orange).

default and SoilGrids runs (not shown). The optimised soil ancillary run gives slightly reduced latent heat flux and slightly increased sensible heat flux in the summer months. This leads to a reduced value of evaporative fraction in the summer months for the optimised ancillary runs, as shown in figure 10.

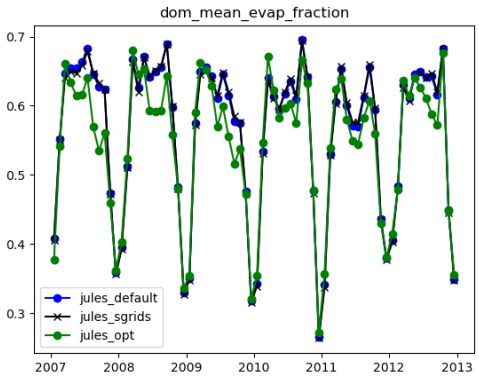


Figure 10: Evaporative fraction from runs of JULES using different soil ancillary data.

Figure 11 shows the evaporative fraction across the domain for at a time when the optimised ancillary JULES run predicts lower evaporative fraction than the other runs (July 2008). The difference in evaporative fraction is greatest around the Thames Valley region, parts of eastern England, and in the part of mainland Europe contained in the REP domain, where differences in soil physics parameters tend to be large.

3.3.2 River flows

Surface and subsurface runoff outputs were routed into river flows on a 1km (CHESS) grid, as described in Section 2.2, and compared with NRFA river flow observations at a number of gauges.

Figure 12 shows the KGE metric (as described in Knoben et al. 2019) for the SoilGrids and optimised ancillary runs plotted against the KGE for the default run. Each point represents one of 191 GB catchments. The right hand panel in Fig. 12 shows a subset of the catchments, where the KGE score is greater than -0.41 and the forecast streamflow is therefore regarded as having some skill (Knoben et al., 2019). Figure 12 shows that the SoilGrids KGE scores are fairly similar to the default, with the red points close to the dotted 1:1 correspondence line. There are 43 catchments (23%) in which the SoilGrids ancillary gives an improvement in KGE score (closer to the ideal maximum value of 1) and 148 (77%) where the score becomes slightly worse. For the optimised ancillary file, the LH panel of Fig. 12 shows that for catchments in which the default KGE score is very low, the soil ancillary degrades the forecast even further, with even lower KGE scores. However, the right hand panel shows that in the skilful KGE score range, the optimised ancillary gives rise to many better KGE scores than the default, with the majority of black points lying above the dotted line. In total there are 138 catchments (72%) in which the KGE score is improved for the optimised ancillary and 53 (28%) where the KGE score for the optimised ancillary runs is worse than for the default runs.

The KGE is made up of three components (correlation, variability bias and mean bias; Knoben et al. (2019)); values were calculated for all catchments, with those for 13 catchments shown for default and optimised ancillary runs in Fig. 13. The correlation coefficient, r is generally worse (further from the ideal value of 1) for the optimised ancillary run than for the default run, with an improvement in only 8 out of 191 catchments (4%). The alpha component (related to variability bias) of the KGE is larger in the optimised ancillary runs than the default run in every catchment. This reflects the fact that the spread of the flows is increased for the optimised ancillary run. When the default alpha is too low (<1; not enough variance in the flow) an increase in alpha is an improvement, but in catchments where alpha is already too large the optimised ancillary runs amplify this. The beta component (related to mean bias) also increases in the majority of catchments for the optimised ancillary runs. The alpha and beta parameters are improved (closer to 1) for the optimised ancillary runs in 141 and 151 catchments respectively.

Nash-Sutcliffe Efficiency (NSE) scores for the 191 catchments are shown plotted against baseflow index (BFI) in Fig. 14. The left hand panel demonstrates that the optimised ancillary runs produce very bad NSE scores in catchments with the highest BFI; this is also true for the KGE score (not

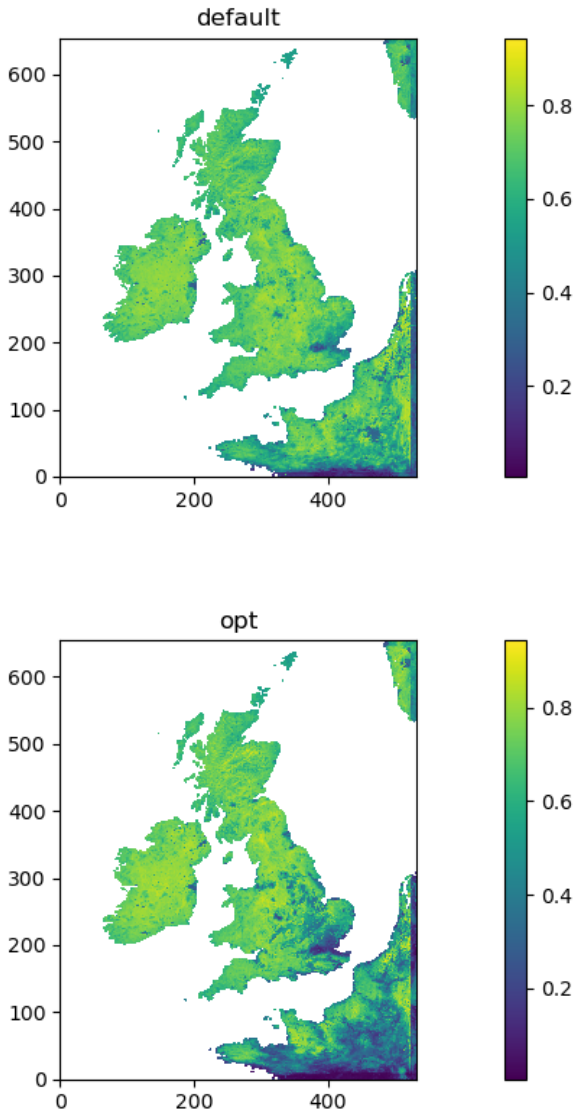


Figure 11: Mean evaporative fraction over July 2008 for default (top) and optimised (bottom) ancillary JULES runs

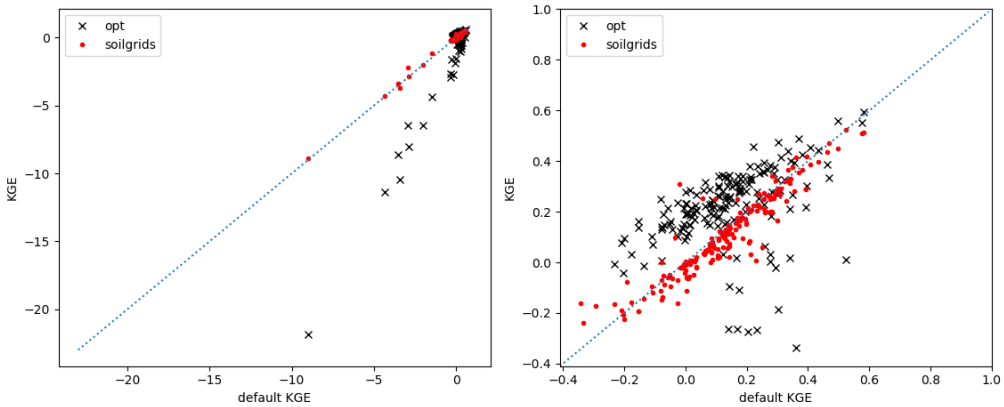


Figure 12: Kling-Gupta Efficiency metric for 191 catchments, with SoilGrids (red) and optimised ancillary (black) run values plotted against default run values. Left hand panel shows all the calculated KGE scores; right hand panel shows only scores in the range of forecast skill

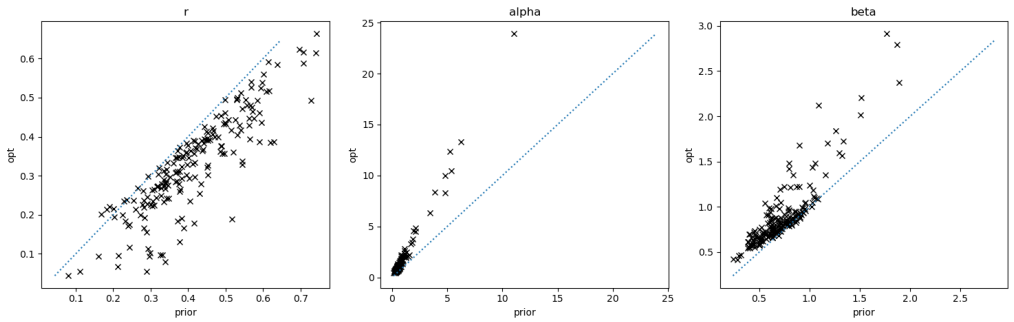


Figure 13: Components of the KGE for 191 catchments; optimised values plotted vs default values. Ideal value of each of the components is 1.

shown). The right hand panel shows only NSE scores above zero, as this is the threshold at which a forecast is often assumed to show some skill 6. The default runs give rise to better NSE scores than the optimised ancillary runs in 179 (94%) of the catchments. The default and SoilGrids NSE scores are more similar, with the default run scoring better in 98 (51 %) catchments.

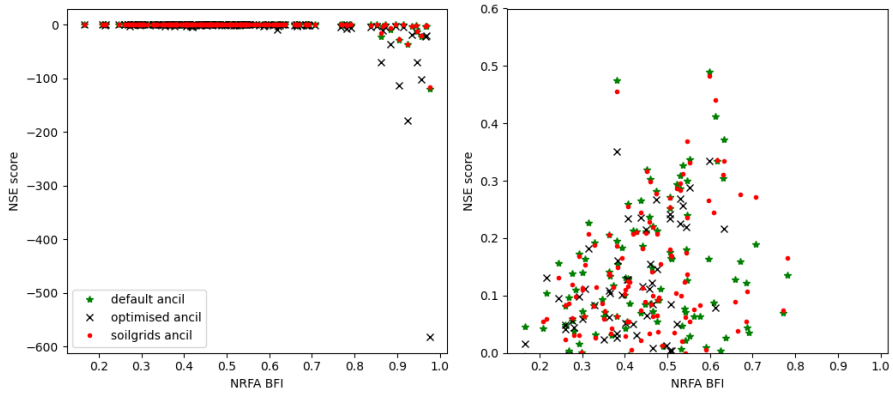


Figure 14: NSE scores with BFI for all 191 catchments on left panel; same data with cutoff of NSE > 0 in right panel.

Illustration for selected catchments

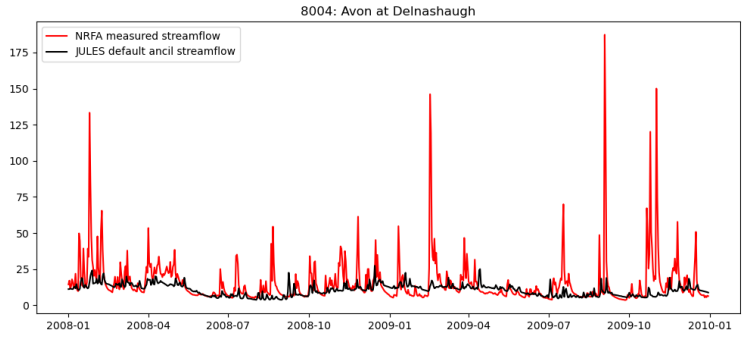
Figures 15 and 16 show two years of hydrograph data for two catchments with similar observed BFIs. Red lines show NRFA daily flow measurements and black lines show JULES streamflow estimates. Gauge 8004 is the Avon at Delnashaugh, which has a BFI of 0.55 (NRFA) and a catchment area of 542.8 km^2; gauge 54057 is the Severn at Haw Bridge, with a BFI of 0.57 (NRFA) and area 9895 km^2. These catchments are very different sizes, but both have BFI values close to the mean of 0.50 for the 191 catchments studied here. In Fig. 15a we can see that while the baseflow appears to be captured well by the default soil ancillary runs, the modelled streamflow is not as flashy as the observations; the highest flows are missed by the model. Figure 15b shows the streamflow resulting from the optimised soil ancil. Now more of the higher flows are captured, but the modelled low flows are too low. The default ancillary predicts the measured BFI very accurately (0.51 over the 2007 - 2012 time series); the optimised ancillary underestimates the BFI (0.39) but scores better on the KGE metric (-0.01 for the optimised ancillary vs. -0.14 for the default), with slightly reduced bias and a better match in the variance of the modelled flow. Results in catchment 54057 show a similar picture in Fig. 16, with the optimised soil ancillary providing better capture of the higher flows and improved spread of the modelled flow leading to an increase in the KGE score (0.38 optimised vs 0.28 default). However, the NSE score is worse for the optimised ancil, giving a score of 0.23, vs 0.31 for the default run. The BFI is also underestimated using the optimised soil ancillary (0.35 vs 0.54 for the default ancil).

Figures 17 and 18 show example results for catchments with more extreme observed BFI values. In Fig. 17a we see daily flow time series for catchment 39019, the Lambourn at Shaw, with BFI = 0.97 (NRFA) and area 234.1 km^2. Figure 17a shows the default soil ancillary produces streamflows which are too flashy compared to observations, and underestimates the low flows in this catchment.

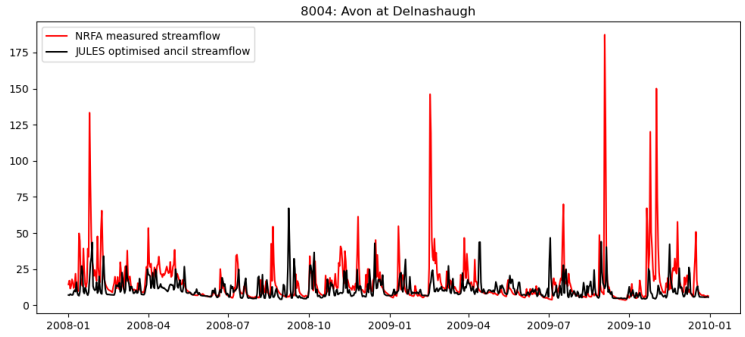
Figure 17b shows that these deficiencies are enhanced by the optimised ancillary (note the different y axis labels in figs17a and 17b), with the high flows becoming higher still and the low flows even lower. The optimised ancillary therefore scores lower in both NSE and KGE metrics for the optimised ancillary (Default NSE = -2.8, optimised NSE = -20.6; default KGE = -0.19, optimised KGE = -2.72).

Results for catchment 21023, are shown in Fig. 18. This gauge is Leet Water, Coldstream which has a low BFI of 0.32 (NRFA) and catchment area of 113 km^2. In this catchment we see again the pattern of flashier flows and reduced baseflow for the optimised ancillary compared to the default. The optimised ancillary again significantly underestimates the BFI (0.06 cf 0.17 for the default run) and makes the NSE worse relative to the default runs. The optimised ancillary runs are able to better capture higher observed flows, giving a slight improvement in the KGE score here.

The catchments highlighted here have illustrated features found in many of the other catchments, and also drawn attention to the fact that whether a change represents an improvement or not will often depend on the chosen metric.



(a) default soil ancil



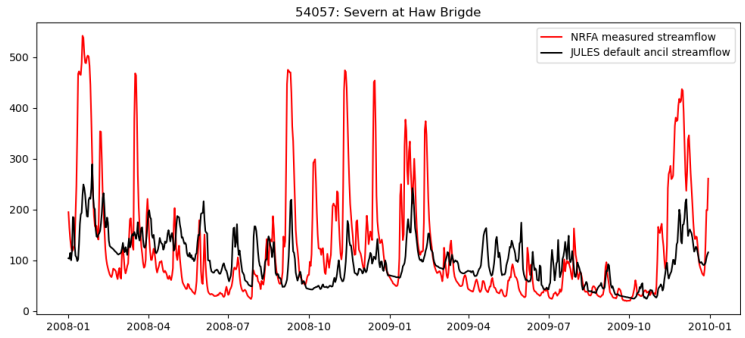
(b) optimised soil ancil

Figure 15: Hydrographs for gauge 8004: Avon at Delnashaugh. Red lines show NRFA observations; black lines show JULES streamflow.

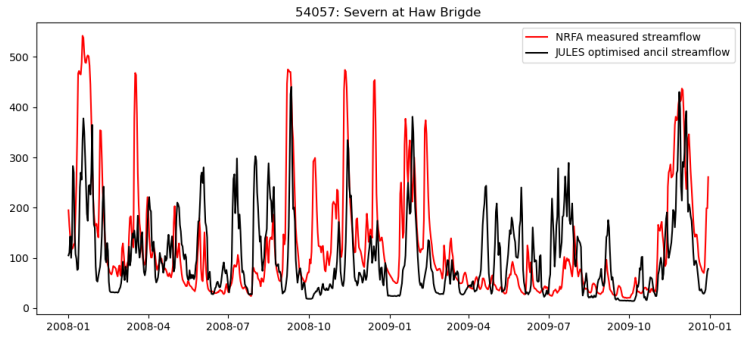
Cross spectral analysis of river flow

We compared modelled and observed streamflow for 13 catchments in terms of cross spectral analysis; the catchments used are those of Martínez-de la Torre et al. (2019) and listed in table 3. Note that we inadvertently used a value of a_thresh = 1 in the routing configuration for these results (meaning that all points were considered to be rivers and hence used the wave speeds for river rather than land points); elsewhere in this report a value of a_thresh = 13 was used (see Section 2.2). This is expected to have a small impact on the spectral results, but is unlikely to affect the main conclusions.

Spectral analysis allows assessment of the average amplitude of discharge on different timescales (Martínez-de la Torre et al., 2019). In Fig. 19 we show amplitude ratios of the modelled to observed flows at the annual (top) and several day (bottom) timescales. The ideal value for the amplitude ratios is 1, shown on the plots with a dotted line. The data shows that at the annual scale all of the JULES runs underestimate the variability of the flows, with the optimised ancillary runs

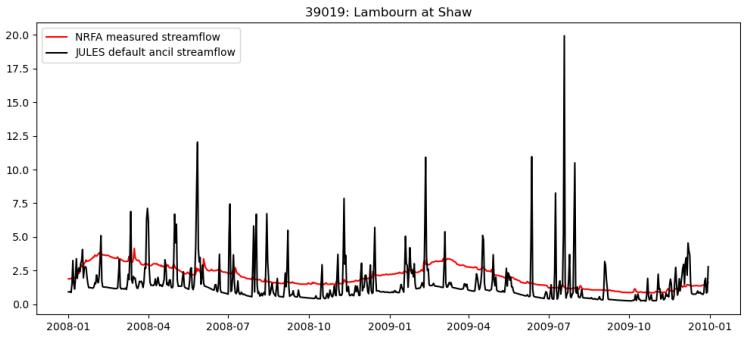


(a) default soil ancil

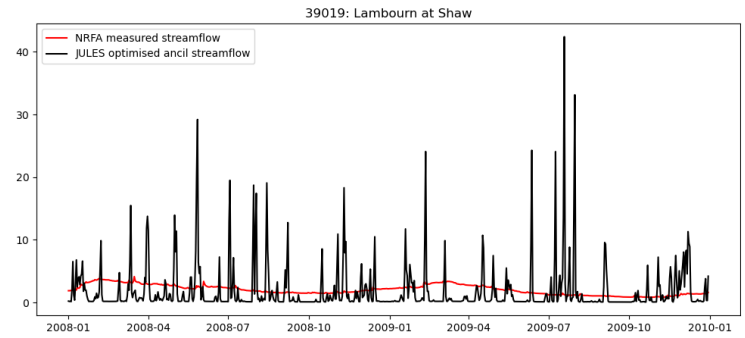


(b) optimised soil ancil

Figure 16: Hydrographs for gauge 54057: Severn at Haw Bridge. Red lines show NRFA observations; black lines show JULES streamflow.

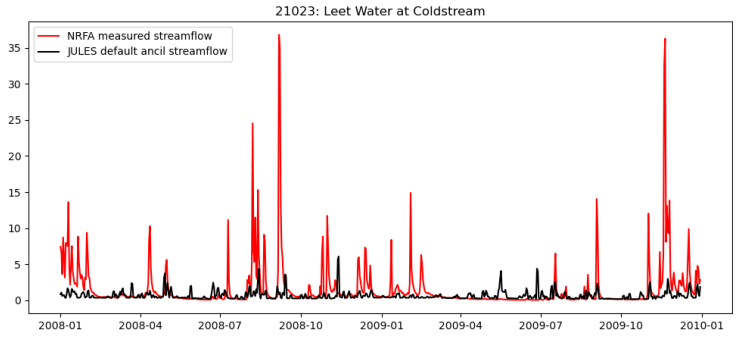


(a) default soil ancil

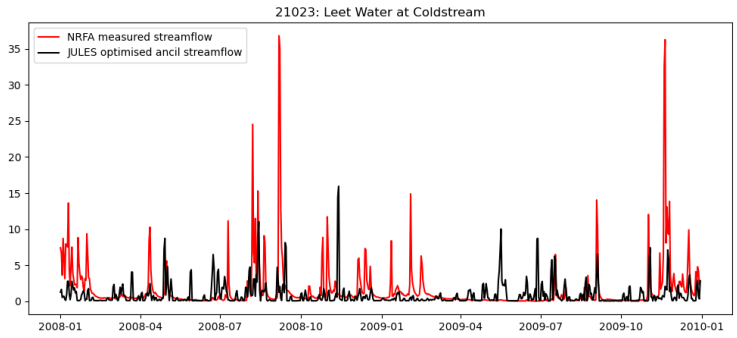


(b) optimised soil ancil

Figure 17: Hydrograph for gauge 39019: Lambourn at Shaw. Red lines show NRFA observations; black lines show JULES streamflow.



(a) default soil ancil



(b) optimised soil ancil

Figure 18: Hydrograph for gauge 21023: Leet Water, Coldstream. Red lines show NRFA observations; black lines show JULES streamflow.

underestimating to the greatest degree. At shorter timescales, the optimised ancillary run gives rise to larger amplitude ratios than the default run at all catchments. In the 9 (of 13) catchments where the default run underestimates the short term flow variability, the optimised ancillary run therefore improves this measure. In catchments where short term flow variability is already overestimated in the default run, the optimised ancillary makes the situation worse.

Optimising the soil ancillary based on COSMOS-UK soil moisture measurements reduces the saturated conductivity and absolute value of the matric potential at saturation (parameters satcon and sathh) and increases the moisture content at saturation (sm_sat). This allows the soil to hold more water in the winter and to better match the seasonal cycle of the COSMOS observations. Applied at a catchment scale, the optimised soil ancillary has the effect of decreasing baseflow and increasing flashier flows over catchments, relative to the default soil ancil. This increased variation in the flow values matches the increase in variation of the soil moisture values. We find that the optimised ancillary can improve streamflow by some measures in catchments where the default ancillary underestimates the flashiness, but performs particularly badly in catchments with the largest measured BFI.

3.4 Summary and future work on soil ancillary data

We have found that:

- the differences between the default and soil grids ancillary fields are small at the catchment scale, with the SoilGrids JULES runs producing similar river flows to the default runs.
- the optimised ancillary allows soils to hold on to more water. This leads to reduced evaporative fraction in the summer months compared to the default ancillary file.
- the optimised ancillary gives rise to flashier river flows, enhancing flows at short timescales to better match the observations.
- the optimised ancillary substantially decreases the baseflow compared to the default ancillary, and this significantly degrades the match to river flow observations, especially in catchments with larger baseflow indices.

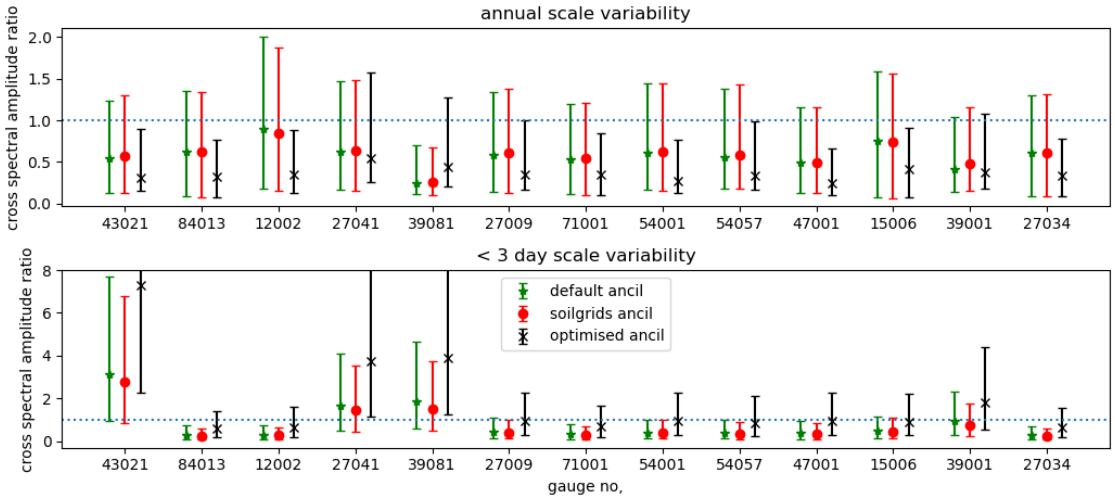


Figure 19: Cross spectral amplitude ratios, comparing modelled streamflow from runs with alternative soil ancillary files, to measured NRFA streamflow. Gauge numbers and locations are described in Table 3.

Possible avenues of further exploration include:

- Validation (or otherwise) of the new soil ancillary file using flux tower measurements of latent and sensible heat fluxes. This would allow us to work out whether the optimised ancillary fields are having a positive or negative impact on JULES flux estimates as well as soil moisture.
- Using additional observations in the data assimilation framework. There are many possible options for this, such as using flux tower or streamflow measurements alongside soil moisture observations to further constrain the PTF constants. It would also be possible to look at optimising parameters controlling different processes such as those in RFM and/or the runoff generation mechanisms in JULES. It would be important here to investigate which processes are sensitive to which observations, and to guard against overfitting of any particular parameter or process.
- Rerunning this analysis using the new groundwater parameterisation described in the next section. It is likely that this optimisation process produces parameters which are implicitly correcting for process not yet represented in JULES, such as those concerning groundwater. It would therefore be interesting to examine the extent to which the parameters change once groundwater process are explicitly considered.
- Investigation of a selection of alternative PTFs, many of which include additional inputs as well as sand, silt, clay. It would also be particularly interesting to use PTFs specifically designed for particular types of soil (e.g. highly organic soils).

4 Groundwater

In this project we have finished the implementation of a new Dynamic GroundWater scheme for JULES that has been developed in a collaboration between UKCEH and the British Geological Survey (BGS). We have also applied the new scheme across GB.

4.1 Overview of the parameterisation

The Dynamic Groundwater (hereafter DGW) parameterisation is based on the groundwater scheme from LEAFHYDRO (Walko et al., 2000; Miguez-Macho et al., 2007; Fan et al., 2007; Martínez-de la Torre and Miguez-Macho, 2019) and is illustrated by Figures 20–22. The model introduces a prognostic water table depth that fluctuates in the model as a result of three main interactions keeping the groundwater balance:

- Groundwater recharge (R): Water flux between the groundwater reservoir and the soil. Depending on the soil wetness and evaporative demands, the recharge can be downwards, causing the water table to rise, or upwards, causing the water table to deepen.
- Lateral groundwater flow (Qn): Water flux to or from neighbour cells within the saturated groundwater reservoir. This flux is governed by topography and the water table head elevation in the cells.
- Groundwater-river flow (Qr): This can occur as groundwater discharge (subsurface runoff) into the streams when the water table head is above the river bed, maintaining stream baseflow (Fig. 22), or as river infiltration to the groundwater reservoir when the water table head is below the river bed.

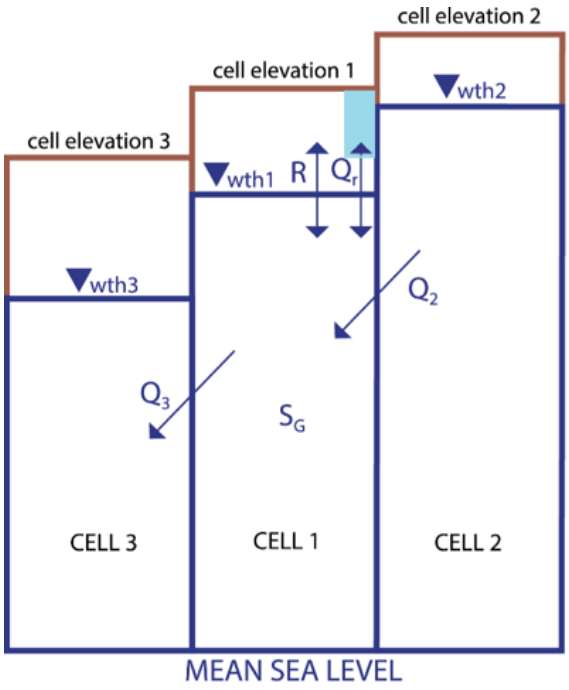
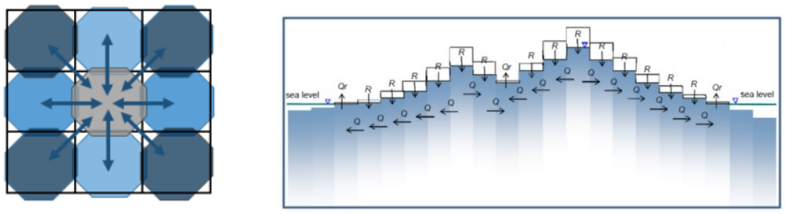


Figure 20: Schematic view of the DGW groundwater balance in a cell.

We have added two new developments in JULES-DGW that were not in the previous work. For the calculation of recharge when the water table is below the resolved soil layers JULES-DGW uses a new formulation based on the saturated conductivity at depth and the water table position (Niu et al., 2007; Batelis et al., 2020). In addition, for the calculation of lateral flows, the original versions of the model used an exponential decay of saturated conductivity with depth and applied an anisotropy ratio to calculate the lateral conductivity from the vertical conductivity. Now the DGW included in JULES includes an option to use ancillary fields for the depth to the bottom of the aquifer and the saturated conductivity at depth. Through this the user can choose to provide ancillary layers if the area of interest is well characterized and such parameters at depth are available.

The JULES-DGW code can be found in the branch vn5.2_dgw_leafhydro_bgs.



(a) View from top (b) lateral flow cross section

Figure 21: Schematics of lateral flow, from Fan et al. (2013).

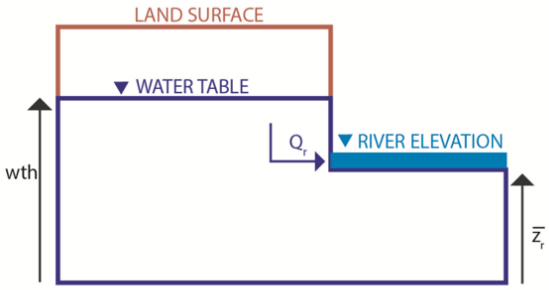


Figure 22: River-groundwater flow for gaining streams.

4.2 Simulations and groundwater ancillary data

We developed suite u-ck085 for use with JULES-DGW in conjunction with UM meteorology for 2007–2012, similar to the suites used in earlier sections outlined in Sec. 2.1. The JULES configuration is as close to RAL3M as possible, with minor changes to allow the use of the vn5.2-based DGW code.

We carried out a set of simulations to study model performance and the effect of introducing DGW:

- simulation WT (water table) with the DGW scheme on and PDM used to calculate surface runoff, with parameter settings as recommended in Martínez-de la Torre et al. (2019);
- simulation FD (free-drain) with the DGW scheme off and again PDM to calculate surface runoff;
- simulation TOP using TOPMODEL to calculate runoff as in RAL3M.

Comparison of WT and FD shows the effect of DGW, while comparison of WT and TOP shows the impact of different ways of dealing with both the surface runoff and subsurface hydrology. Note that the DGW scheme cannot be used with TOPMODEL as these include contrasting descriptions of groundwater processes. The WT run included two spin up cycles over 2007–2012 as the initial water table depth (from external data, see below) needs to adjust to the model.

JULES-DGW requires additional ancillary fields which were prepared in this project. Here we briefly summarize what this new dataset contains, the main data sources used, and how the parameters were calculated for this work.

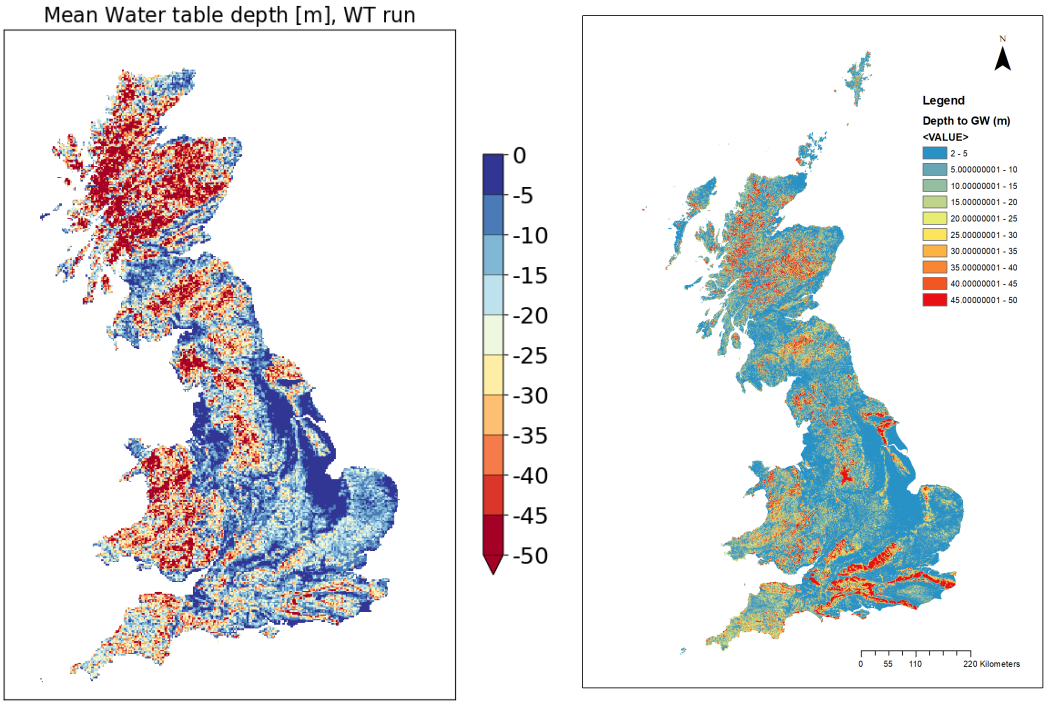
- *zw_eq*: Initial water table depth, calculated using an iterative 2D groundwater model that finds a balance between long-term recharge driven by the atmosphere and lateral flows driven by the topography (Fan et al., 2007, 2013). We extracted this field from a global 1km dataset available through the eartH2Observe project (https://wci.earth2observe.eu) which was produced using an updated version of the methodology presented in Fan et al. (2013)).
- *klat*: Deep saturated conductivity. This parameter has been provided by BGS for this project in the form of a dataset covering Great Britain and drawing on many years of work and experience with geological surveys, hydrogeological interpretation and borehole pumping tests. We regridded the data from the original 1km OSGB grid.

- *elev*: Cell elevation. Calculated as mean elevation from high resolution DEM data (100m) provided by the MetOffice.
- *z_river*: Riverbed depth from surface. Calculated as the difference between the cell elevation and the lower high resolution elevation found within the 2.2km cell.
- *rc_eq*: River conductance at equilibrium. Calculated from a long-term balance between R, Q_{n} and Q_{r} (Miguez-Macho et al., 2007). For this balance we used long-term recharge from a JULES run over 1961-2015 (Blyth et al., 2019) driven by the CHESS forcing dataset (https://eip.ceh.ac.uk/chess).
- *rc_a*: River conductance parameter for deviation from rc_eq due to water table position, depending on the cell slope (Martínez-de la Torre and Miguez-Macho, 2019). The calculation of this involves the cell slope, which was calculated from the high resolution DEM data provided by the Met Office.
- *rc_in*: Parameter for fixed river-groundwater flow in losing stream. Not used for this work.
- *sy*: Specific yield for groundwater. We use values provided by BGS. This parameter has been provided by BGS for this project in the form of a dataset covering Great Britain and drawing on many years of work and experience with geological surveys, hydrogeological interpretation and borehole pumping tests. We regridded the data from the original 1km OSGB grid.
- *zbot*: depth to the bottom of the aquifer. We use the value recommended by BGS of 100 metres over Great Britain.
- *fdepth*: Parameter for exponential decay of saturated conductivity with depth below the resolved soil layers (Fan et al., 2007). Not used for this work.

4.3 Impact of representing groundwater

4.3.1 Mean hydrological patterns

The first result we show is the mean water table depth for the 6-year simulation, shown in Fig. 23.



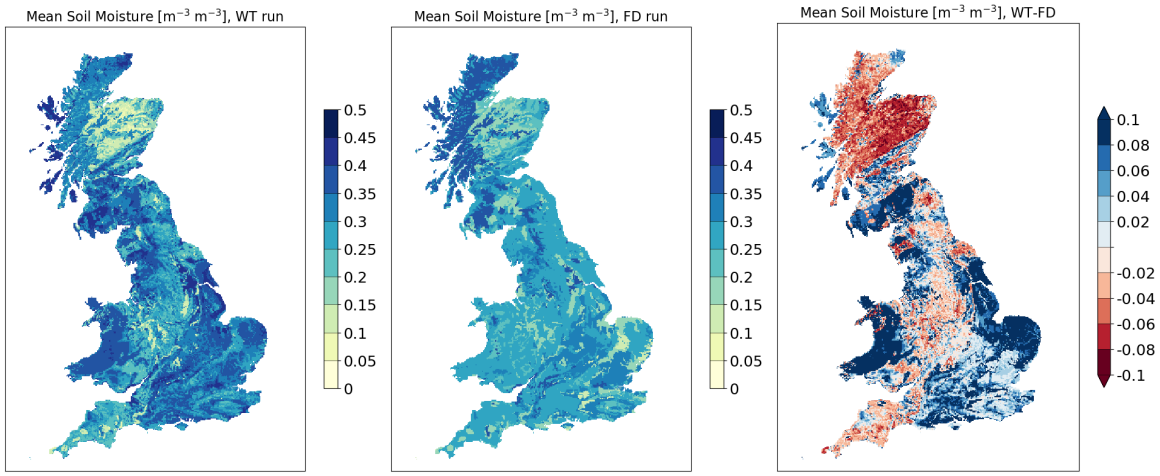
(a) JULES-DGW

(b) BGS modelled

Figure 23: Water table depth.

The main patterns in water table depth from JULES-DGW in Fig. 23a resemble those in the separate estimate produced by BGS (the modelled Depth to Groundwater dataset) shown in Fig. 23b. Both products show very deep groundwater levels in mountainous regions in the north and west, and shallower regions in the east. Although the main areas of shallow water table depths are found in the southeast of the island, there are extensive areas with depths below 15-20 m scattered across the domain, showing a strong potential connection and influence of groundwater dynamics on JULES soil moisture and water flows.

Typically the large-scale soil moisture pattern should be dominated by precipitation input and soil texture. The influence of groundwater becomes an additional factor as shown in Fig. 24. The difference between the mean top-3 m soil moisture fields from the WT and FD runs (Fig. 24c) shows patterns from the water table depth distribution (Fig. 23a), reaching generally higher values (wetter WT run) where the water table is shallower and negative values (drier WT run) in regions with a deeper water table (even though there are some exceptions such as over Wales and South Scotland where the WT run presents wetter patterns even with deep water table). The similarity between the patterns of soil moisture WT-FD differences and mean water table depth illustrates the controlling role of groundwater dynamics in soil moisture spatial variability.



(a) (b) (c)

Figure 24: Mean soil moisture.

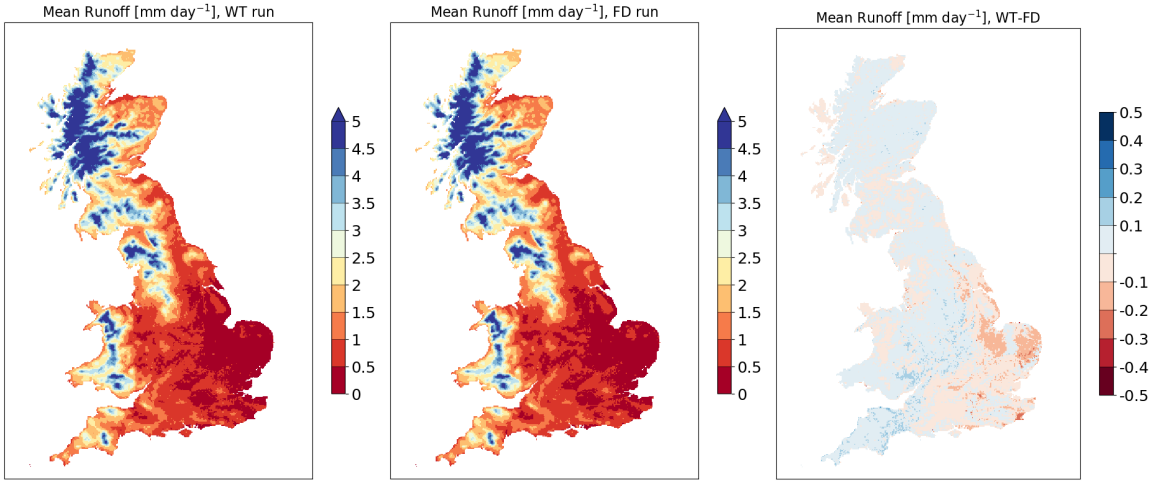
The differences in soil moisture suggest that DGW will introduce differences in the runoff and evapotranspiration (ET hereafter) water fluxes that complete the JULES water budget. We find that even if the total runoff differences are not large (mostly below 0.1 mm day^{-1}, Fig. 25c), spatial patterns are introduced in the runoff partition between surface and subsurface runoff (Fig. 26).

The areas highlighted in Figures 24c and 23a with drier soil and deeper water table in the WT run result in lower surface runoff (PDM surface runoff production is based on the wetness of the top soil layers) and higher subsurface runoff (the water drains quickly over those hilly areas to the subsurface). More interestingly, wetter areas of shallow water table over the southwest and other regions that show higher surface runoff (again as a consequence of the PDM behaviour with the wetter soil) and lower subsurface runoff in general, indicating that the water table presence might act as a buffer in the subsurface, keeping the soil wetter and not draining water to the river network as much as the free-drain approach does.

4.3.2 Impacts in different groundwater regimes

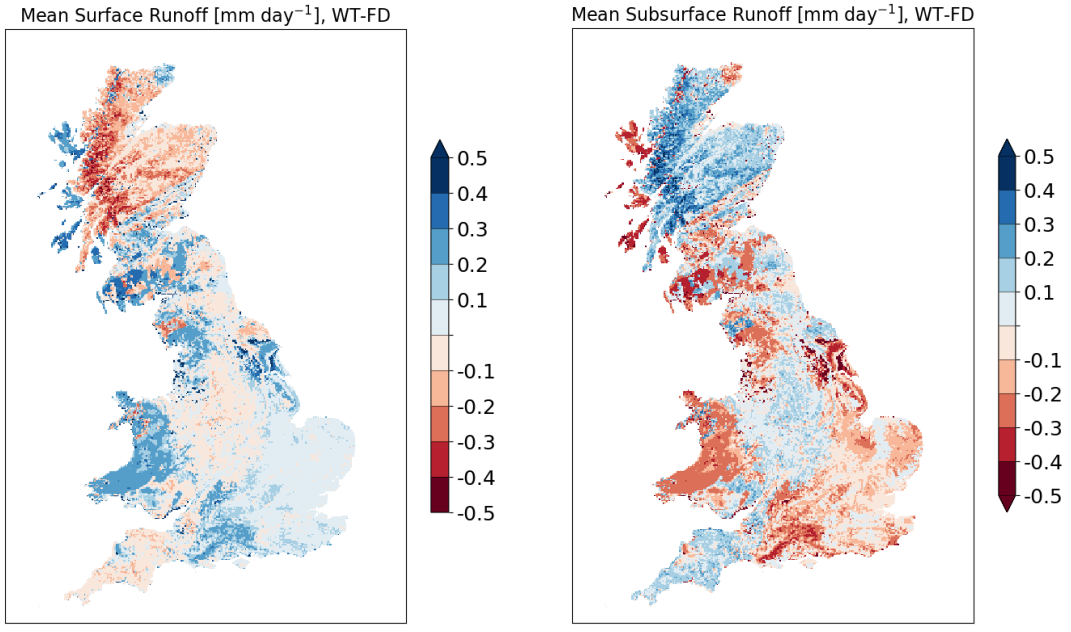
The differences described in the preceding section tell us that there will be a clear influence of groundwater dynamics on JULES hydrology; here we try to understand how those differences apply during the simulation period over different regions where the water table is found closer to or further from the surface (more or less connected to the surface fluxes). The next set of figures show the monthly evolution of the main JULES water variables, dividing the domain in three main areas depending on the local depth to water table: (1) mean water table depth is below 20 m, (2) between 20 and 3m, and (3) shallower than 3 m.

Deep water table regions cover 46 % of the domain (Fig. 27a) and DGW is not expected to make a big difference to the simulation in these areas as the groundwater is disconnected from the



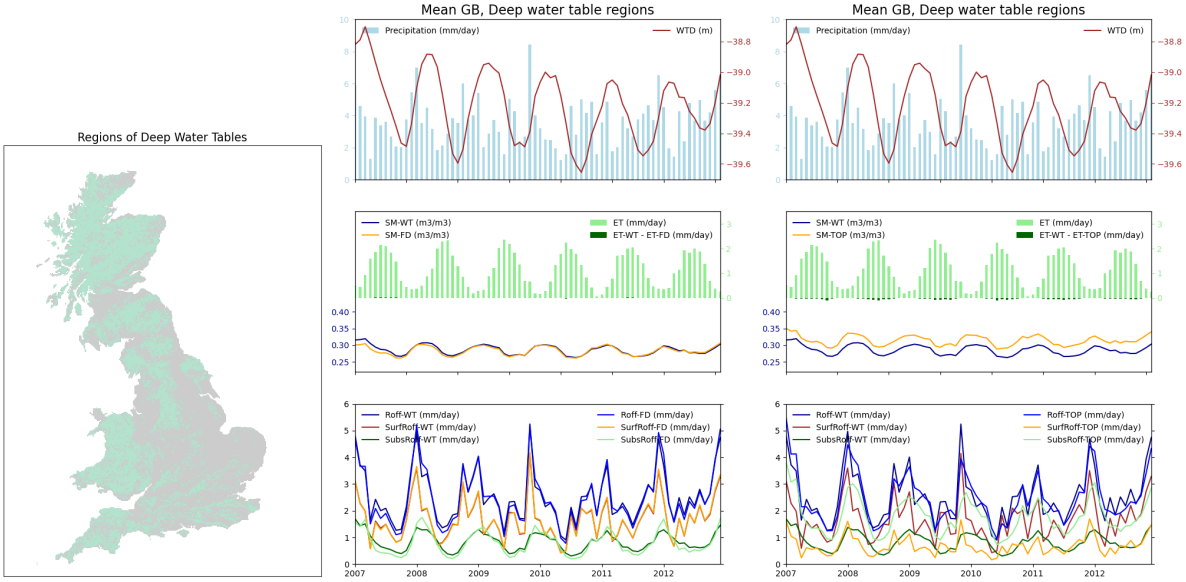
(a) (b) (c)

Figure 25: Mean runoff.



(a) (b)

Figure 26: Mean runoff partition differences.



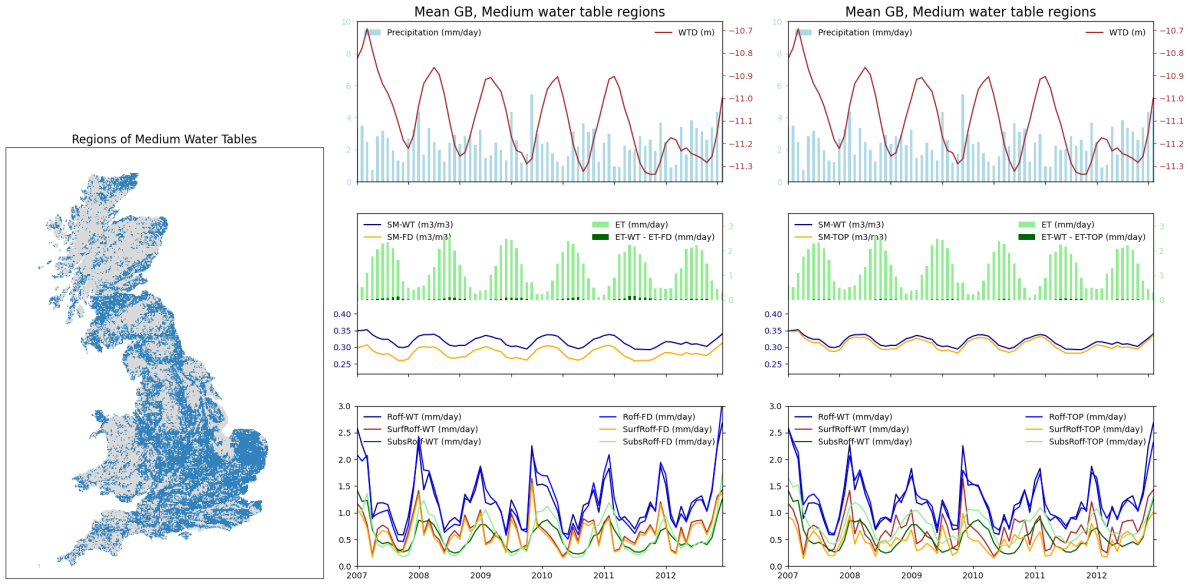
(a) $WTD < 20m$, 46% of land (b) WT-FD (c) WT-TOP

Figure 27: Monthly evolution averaged over the deep ($wtd > 20m$) regions. (b) and (c): Precipitation and water table depth in the top plots; ET and ET differences (WT-FD, WT-TOP in (c)) in the top of the middle plots; mean soil moisture in the resolved column for WT and FD (TOP in (c)) runs in the bottom of the middle plots; and total runoff, surface runoff and subsurface runoff for WT and FD (TOP in (c)) runs in the bottom plots.

surface. Even though the water table is found very deep into the soil (close to 40 m in average), its seasonal evolution is clear in Fig. 27b and we can appreciate how it responds with a lag of 2–3 months to the precipitation input at the surface. The changes in soil moisture are very small or negligible (except in the first year when the water table is still adjusting after the spin up phase) and therefore there is no real influence on ET. The bottom plot in Fig. 27b does not show significant differences between WT and FD runs, as expected. These hilly regions are largely dominated by surface runoff, and the small differences in total runoff come from the timing of the subsurface runoff (green lines) which, in the WT run, does not show such high peaks in winter and keeps the baseflow slightly higher than the FD run during the drier months. Looking at Fig. 27c we point out that DGW has a larger impact over these deep water table regions when we compare it to the TOP run, showing drier soil and a very different partition of the runoff, even if the total runoff differences are not as large. Effectively the WT run treats these regions as surface runoff dominated, whereas looking at the runoff partition of the TOP run these regions could be characterized as baseflow dominated areas. Higher runoff peaks during wet months in the WT run in these regions will have an impact, however, on the river flow (next section).

In Fig. 28 we analyse another large portion of the land (46 %) where the water table is found at shallower depths (10-11 m in average) and the impact from DGW is expected to be more significant. The soil moisture is now higher in the WT run, and this shows an impact in the ET fluxes to the atmosphere (Fig. 28b, middle plot) which is larger than over deeper water table regions, as here the precipitation input is lower and the ET demands are better satisfied with a wetter soil. Also there are 2 periods (Feb-Mar 2011 and Feb-Mar 2012) that show consecutive months of very low precipitation, and this have an impact on the water table which does not rise to its usual level during the winter 2011-2012, thus causing lower differences in soil moisture during the last year and lower differences in ET during the 2012 ET season. This might be showing the kind of soil moisture memory that the water table introduces in the soil and other studies have pointed out with former versions of DGW (Miguez-Macho et al., 2007; Martínez-de la Torre and Miguez-Macho, 2019).

The DGW runoff impact in Fig. 28b is similar to that found over deeper water table regions, with lower winter subsurface runoff peaks and slightly higher baseflow during the drier season, but such differences in subsurface runoff have a larger impact on total runoff over these areas where the water table is more connected to the system, as now the surface and subsurface flows are of similar magnitude. In Fig. 28c we point out that again the TOP run characterizes more clearly all these regions as subsurface flow dominated (as seen over the deeper water table regions), and the differences in the resulting total runoff are not as large as they were over deeper areas but still



(a) $20 < WTD < 3m$, 46% of land (b) WT-FD (c) WT-TOP

Figure 28: Monthly evolution averaged over the medium ($20 < wtd < 3m$) regions. (b) and (c): as in Fig. 27

.

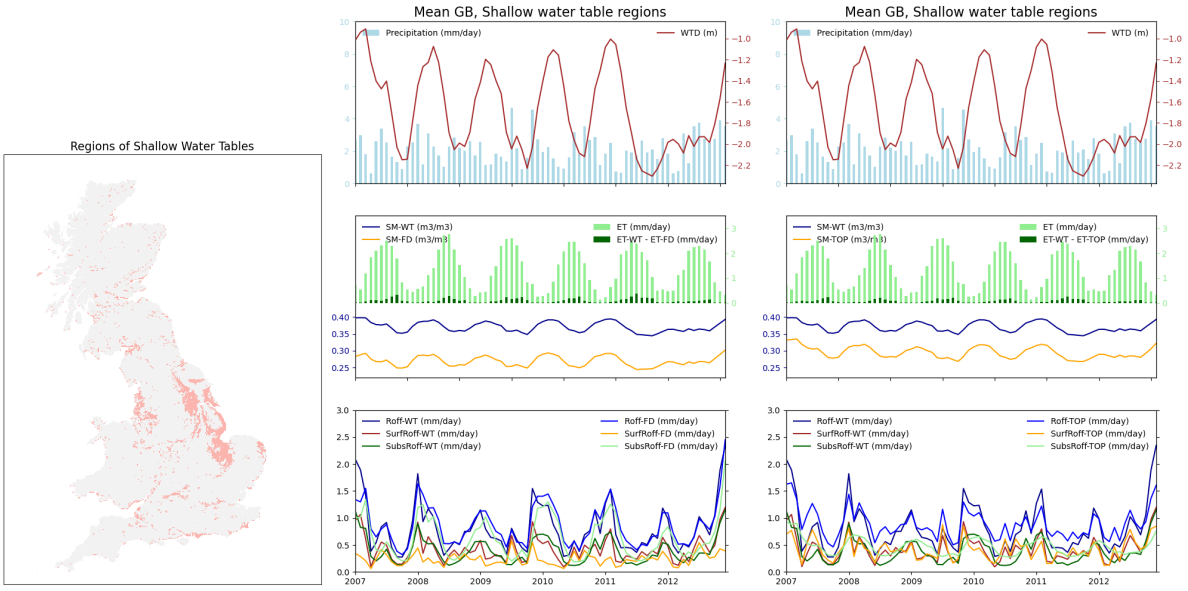
appear to come from higher winter peaks with the WT run from surface flow (from PDM very likely producing higher surface runoff with similar or slightly higher soil moisture).

The areas of the domain where the water tables oscillates around very shallow values (mean water tables less than 3 m deep) are found over many regions, particularly in the east, but they only cover 8% of the total domain (Fig. 29a). Over these areas the soil moisture is much wetter in the WT run (Fig. 29b) as the water table is actually within the resolved soil layers (we used the typical JULES soil layer distribution for this work, down to 3 m), and the impact on ET is larger still that it was over the medium water table depth areas. The surface runoff is, probably as a consequence of such wet soil, much higher in the WT run, causing the subsurface flow to be lower than in the FD run. We were expecting the groundwater to dominate the runoff over these regions (this is an issue for further study). The expected behaviour that does occur, even if it might not be too significant over such wet areas, is that the subsurface runoff continues to drop, for 1-2 months more in the FD run than in the WT run, to lower values every year during the dry season, while the WT run starts to recover earlier (even if it reaches lower values in this case). It is over these shallower water table areas where the differences in total runoff (and not only the runoff partition) are found to be larger between the WT and the TOP run (Fig. 29c). Again as seen when comparing WT and FD, the subsurface runoff recovers during or after summer earlier and quicker in the WT run, but the values are higher for the TOP run and result in higher total runoff during the low flow seasons.

4.3.3 River flows

We evaluated river flow from JULES-DGW using the NRFA daily flow observations. In Fig. 30 we have selected the stations where the model in the WT run shows skill (defined as in Sec. 3.3.2). We found skill in 50% of catchments in terms of NSE and 93% (a total of 153 catchments) in terms of KGE.

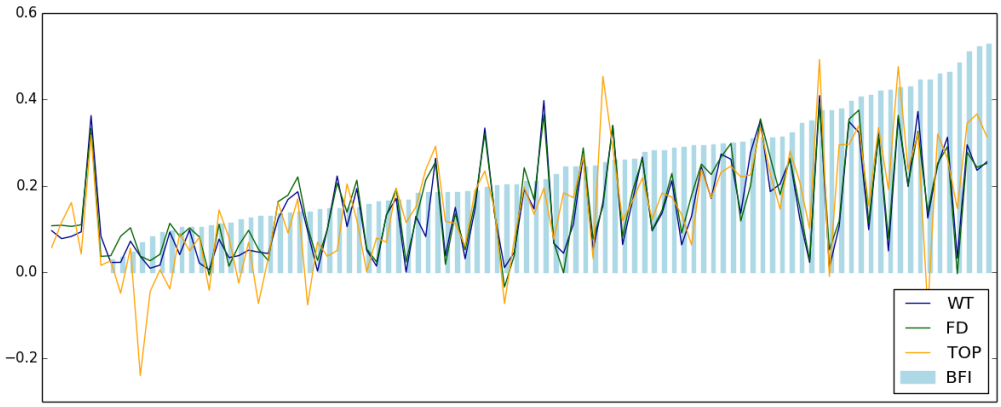
The differences in NSE metrics (Fig. 30a) between the WT and FD runs are generally small, however WT tends to improve performance where there was already good skill in the FD run, and where FD performs better (lower BFI catchments) generally the skill is anyway low ($NSE < 0.2$). The TOP run is more different (consistent with it having very different representations of runoff processes). The WT run NSE score improves on the scores from TOP in 48% of the catchments, with the comparative performance not showing a clear relation with BFI values (blue columns). The KGE metric (Fig. 30b) shows again that the improvement in skill from FD to WT is small although it occurs in 54% of the shown catchments, but clearly the WT run outperforms the TOP run in most stations (86%, a total of 153), more clearly where the BFI is small but also with high



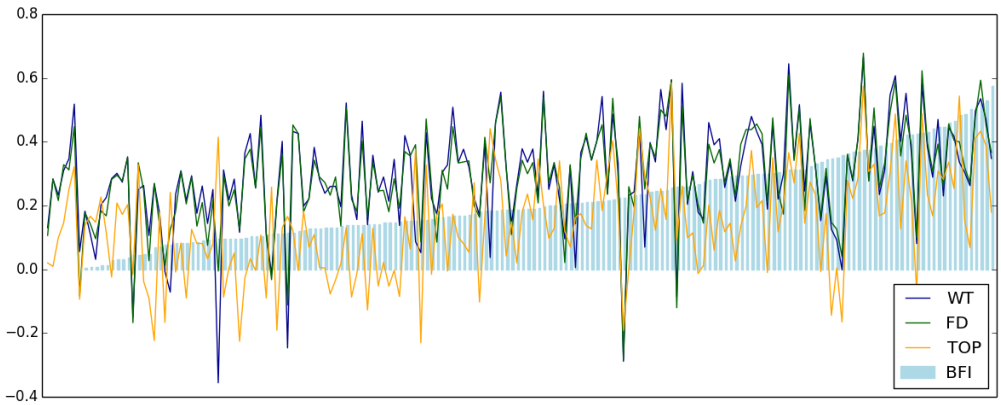
(a) $WTD < 3m$, 8% of land (b) WT-FD (c) WT-TOP

Figure 29: Monthly evolution averaged over the shallow ($wtd < 3m$) regions. (b) and (c): as in Fig. 27

.



(a) NSE



(b) KGE

Figure 30: Performance metrics at gauge stations with skill for daily river flow of WT, FD and TOP runs evaluated against NRFA observations. BFI in light blue columns.

BFIs.

In Fig. 31 we present hydrographs at selected stations to illustrate aspects of the behaviour of JULES-DGW. We use some of the same gauge stations as were used earlier (Sec. 3.3.2) with a range of values of BFI, but here we replace station 21023 (a small catchment with low BFI and poor performance from all configurations) with station 39001 (Thames at Kingston, 9945 km^2) which covers a large catchment with regions of connected water tables.

For the Avon at Delnashaugh (Fig. 31a) no run shows skill in terms of NSE, but KGE is clearly improved by both WT and FD with improved peaks from precipitation events, albeit it often smaller than observed peaks. In contrast the TOP run is not flashy enough and also overestimates the baseflow. However, over this catchment the improvement largely comes from switching to FD, and WT and FD perform very similarly.

For the larger catchment of the Severn at Haw Bridge (Fig. 31b) WT outperforms both FD and TOP. The winter high flows are better captured by WT and FD (for example at the end of 2009). At times spring-summer baseflows tend to be overestimated by the TOP run (e.g. June-July 2008) whereas WT simulates these better. These results are particularly encouraging as this is a station integrating the water balance of a large (9869 km^2) region with extensive areas of shallow water table (Fig. 23a).

The Thames at Kingston (Fig. 31c) is another large catchment covering areas of connected water table. The WT run slightly improves on the performance of the FD run in terms of KGE but has lower NSE, with the main difference seeming to be related to lower baseflow with WT. However this baseflow is oftentimes underestimated with WT (summer 2008, spring 2009). The TOP run shows similar NSE score but lower KGE than the other runs; as in other catchments it has a tendency to overestimate baseflow and underestimate high flows.

The final catchment highlighted here is the 238 km^2 Lambourn (Fig. 31d) for which a very high BFI (0.97) is indicative of the importance of groundwater in the chalk geology. All simulations were too flashy (as was also the case for the soil ancillary work shown in Fig. 17b) and the WT run produced the lowest baseflow. This catchment illustrates the unexpected behaviour discussed in the previous section in regions with a very shallow water table region, in which run WT produced low baseflow and high surface flows from a very wet soil. The reasons for this need to be understood through further work.

This small selection of catchments has shown a range of responses to model configuration and has suggested that the best configuration is currently different in different areas. The introduction of DGW has improved results in some catchments, including large areas with important groundwater systems. There is little impact of DGW in other catchments, consistent with lower importance of groundwater in those areas. However, even in catchments with major groundwater influence, the current configuration of JULES-DGW is not guaranteed to improve the simulation.

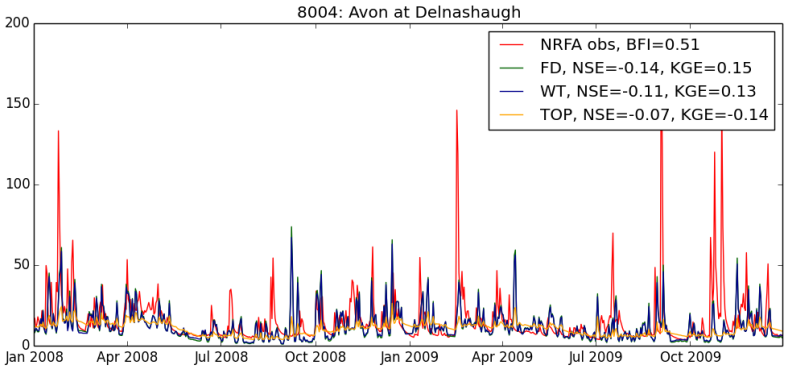
Cross spectral analysis of river flow

We compared modelled and observed streamflow for 13 selected catchments (Table 3) in terms of cross spectral analysis, as was done previously for soil ancillaries (Sec. 3.3.2). In Fig. 32 we show amplitude ratios of the modelled to observed flows at the annual (top) and s̃everal day (bottom) timescales. The data show that at the annual scale all of the runs (with the exception of the Dee catchment, 12002, in WT and FD runs) underestimate annual variability, with the FD run consistently showing the amplitude closest to observations. The WT run shows reduced amplitude (compared with FD) which could be a result of the buffering effect on streamflows of groundwater, with WT amplitude also being lower than that from TOP in most catchments.

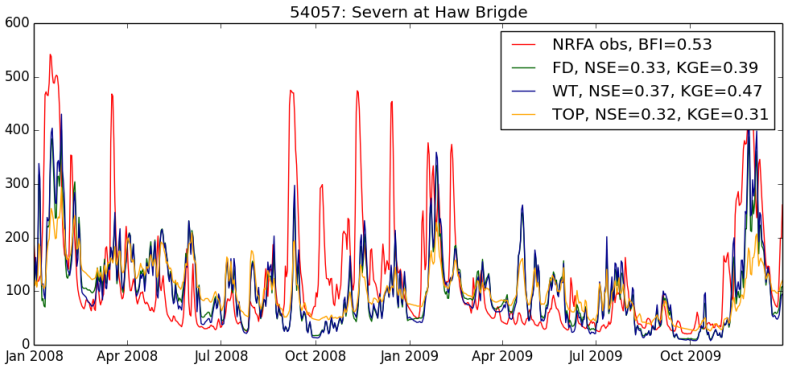
At shorter timescales, the opposite result is generally found, with the WT runs having higher amplitude ratios and therefore better results in terms of < 3 day scale variability in most catchments. At this timescale the introduction of groundwater has little impact in many catchments, except that in catchments where short term flow variability is already overestimated in the TOP and FD runs (higher BFI catchments where all simulations overestimate the flashiness: 27041, 39081, 43021), the WT run further increased the variability, giving poorer results.

4.4 Implementing groundwater in JULES vn6.3

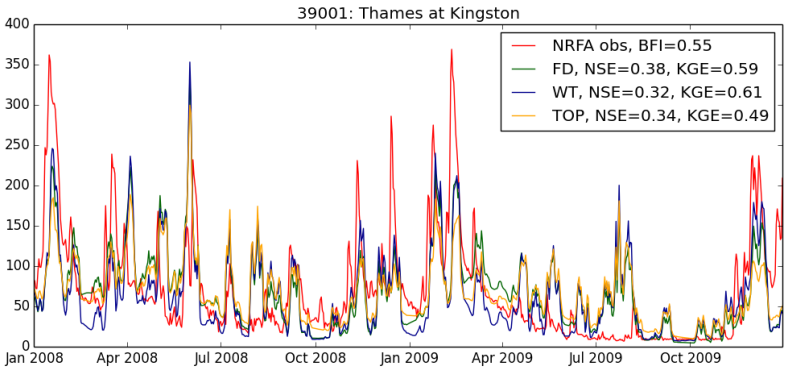
During the course of this project we have started to implement the groundwater model (originally added on a branch from vn5.2) in the latest version of JULES (vn6.3) using the branch vn6.3_groundwater. This is work in progress; at present the new code can successfully replicate results in a short test case. Further work is required in various areas, including adding provision for parallel runs using MPI (which was already catered for in the vn5.2_dgw_leafhydro_bgs branch



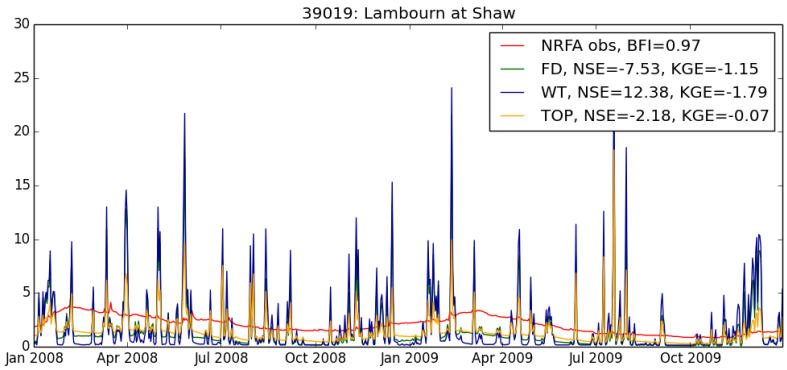
(a)



(b)



(c)



(d)

Figure 31: River flow (m^3 s^{-1}) and metrics for gauges 8004 (a, Avon at Delnashaugh), 54057 (b, Severn at Haw Bridge), 39001 (c, Thames at Kingston) and 39019 (d, Lambourn at Shaw) [2008-2009 shown, the metrics refer to the 6-year simulation]. Red lines show NRFA observations; blue lines show WT run, green lines show FD run and orange lines show TOP run.

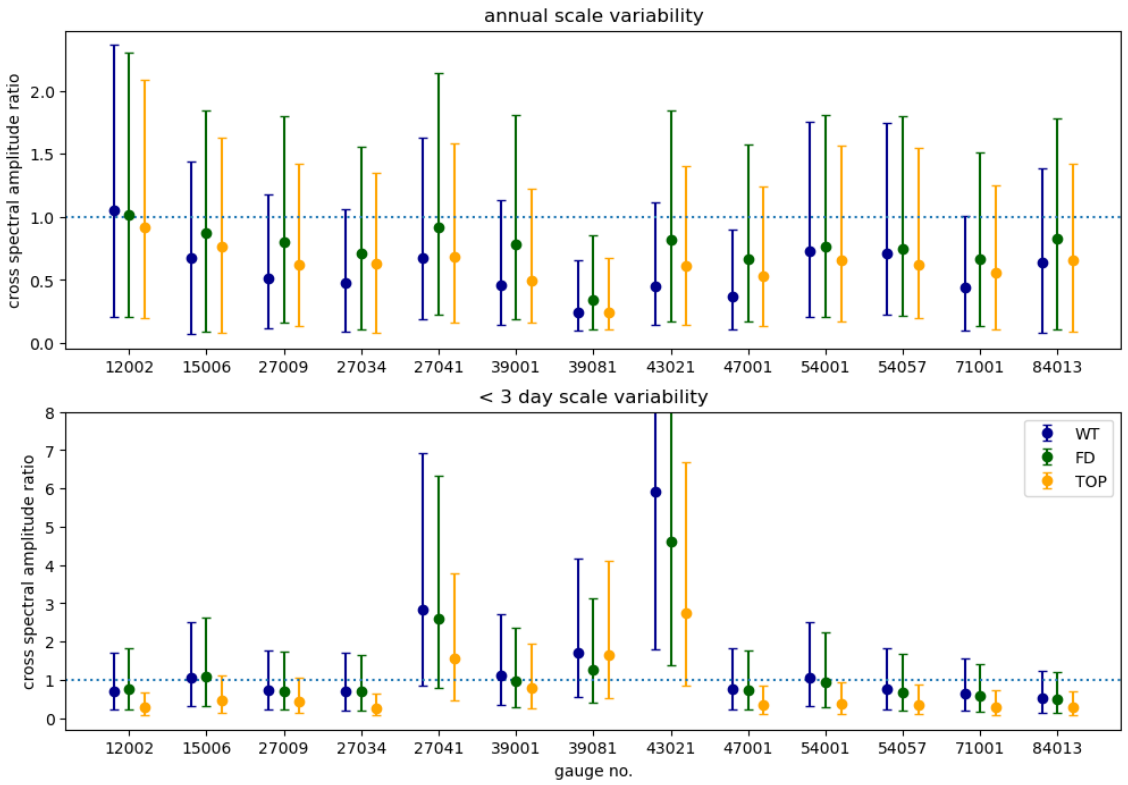


Figure 32: Cross spectral amplitude ratios, comparing modelled streamflow from runs with alternative parameterisations of runoff and groundwater, to measured NRFA streamflow. Gauge numbers and locations are described in Table 3.

used in earlier sections). It is anticipated that this development will be completed over the next couple of months.

4.5 Summary and future work on groundwater

This project produced ancillary data for JULES-DGW for a GB domain and successfully ran the model for several years. As the first nation-wide simulations using JULES-DGW these represent a significant step forward in our modelling capability.

The initial results showed here are promising, with the DGW scheme acting as a buffer in the subsurface and connecting groundwater to the JULES surface hydrology over a large part of the domain (roughly half of Great Britain). Groundwater is contributing to baseflow during drier seasons and wetter soil results in altered ET fluxes. The river flow analysis was also encouraging, demonstrating improved performance relative to the RAL3M-like TOP run at most NRFA stations – though much of that improvement cannot be attributed to DGW (due to the different representation of surface runoff in the WT and TOP runs). The cross spectral analysis of variability at the annual scale confirmed this picture for variability, but results were more mixed at a shorter timescale.

We are strongly encouraged by the results shown here to carry on with the work started with this project and with further analysis of these runs. Some areas for future work are:

- Validation of the position and time evolution of the modelled water table depth using observations from borehole measurements (BGS data in the National Groundwater Level Archive.
- Further study of seasonal variations in different catchments, building on the work in Sec. 4.3.2 based on large areas with similar water table depths.
- Study the impact and behaviour of the lateral flows of groundwater.
- Investigate the impact of assuming an exponential decay with depth of saturation conductivity instead of the data provided by BGS, which is not be available for regions outside Great Britain.

- Assess the impact of different parameter values for river-groundwater flow, in order to better understand unexpected behaviour when the water table is very shallow.
- Validation of changes in evaporation shown here, using in situ measurements from flux towers.
- Investigate the best model configuration, particularly considering the parameterisation of runoff generation. This might include considering whether aspects of existing schemes can be combined in new ways.

5 An alternative parameterisation of river flow and flooding: CaMa-Flood

5.1 JULES-CaMaFlood

The global flood simulation model Catchment-based Macro-scale Floodplain (CaMa-Flood) has been used widely for predictions of inundation extents because of its sophisticated approach (e.g. Hoch and Trigg 2019; Marthews et al. 2021). CaMa-Flood is the only open-source global river routing model that is based on the local inertial approximation of the Saint Venant equations, which takes into account the backwater effects of downstream elements including lakes, tributaries and estuaries. By including these effects, CaMa-Flood is able to produce a much better characterisation of many wetlands and other areas whose dynamics are dominated by surface water inundation.

The river routing and overbank inundation aspects of CaMa-Flood have been added as options in development branches of JULES; this model is described as JULES-CaMaFlood. The runs of JULES-CaMaFlood reported in this report used the branch vn6.3_cmf. The results below represent the first assessment of this new capability for the rivers of Great Britain.

Ancillary data for JULES-CaMaFlood

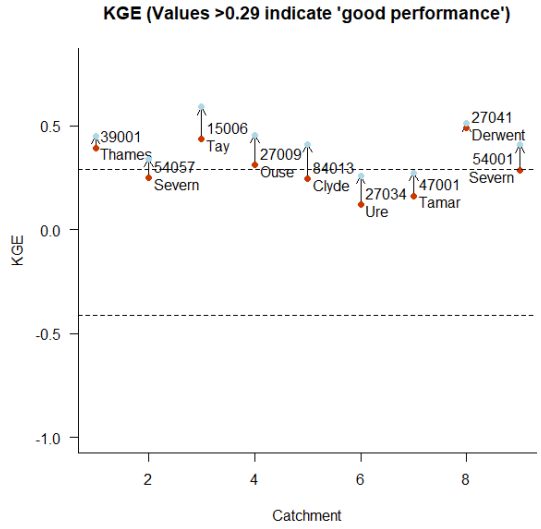
JULES-CaMaFlood was run using the same river-related ancillary data as were used for the simulations with RFM (Davies et al., 2022). An additional field, the river elevation, was prepared for JULES-CaMaFlood from the same data sources by Helen Davies at UKCEH. During the course of this using these elevation data in simulations with JULES-CaMaFlood it was found that at a relatively small number of locations the downstream point was at a higher elevation, which results in isolated locations with considerably deeper rivers. These anomalous elevations arose because of the *ad hoc* approach used to derive the elevation field, on top of the existing ancillary layers which focus on drainage areas and flow directions, not elevation. The river elevation field must be considered as provisional until future work can identify an improved algorithm that does no introduce these jumps.

5.2 Results from JULES-CaMaFlood

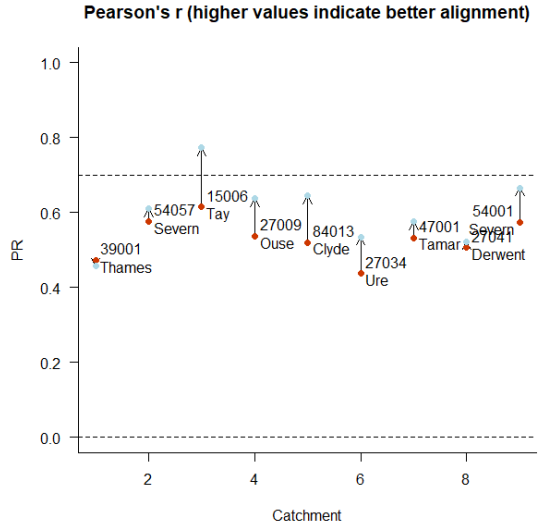
JULES-CaMaFlood was driven by UM runoff fields, as outlined in Sec. 2.2. The occasional jumps in the river elevation field at a small number of points, discussed above, made GB-wide simulations impractical and instead we focused on simulating nine of the catchments from Table 3.

The river flows simulated by JULES-CaMaFlood were compared with NRFA flow observations, and also with the corresponding simulations of JULES with RFM (referred to as JULES-RFM here) that were described in Sec. 2.2 and 3.3.2. The results for KGE, BFI and correlation coefficient are shown in Fig. 33, while the flows for selected catchments are shown in Fig. 34.

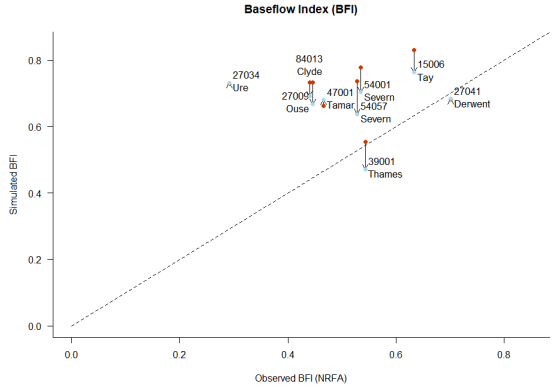
Figure 33a shows that JULES-RFM generally did a reasonable job in these catchments, as measured by KGE, but JULES-CaMaFlood has improved results in every catchment. Looking at the simulated hydrographs in Fig. 34 we can see that peak flows are captured better by JULES-CaMaFlood, particularly in the Tay and Ure (Figs. 34a and 34c), while low flows are most obviously improved in the Tay. Similarly, the correlation between model and NRFA flows improved in most catchments (Fig. 33b), with only the Thames showing a minor deterioration with JULES-CaMaFlood, and NSE also increased in most cases (not shown). The story for BFI was similarly positive, as shown in Fig. 33c which compares the modelled and observed values of BFI. For most catchments the use of JULES-CaMaFlood resulted in a lower BFI, closer to the value calculated from the observed flows, with the Thames again being an exception. The poorer performance in terms of BFI for the Thames looks to be consistent with the more flashy nature of the flows from JULES-CaMaFlood when compared with those from JULES-RFM (Fig. 34b), which often results in excessively low flows from JULES-CaMaFlood. In reality the recession of flows over these long



(a) KGE

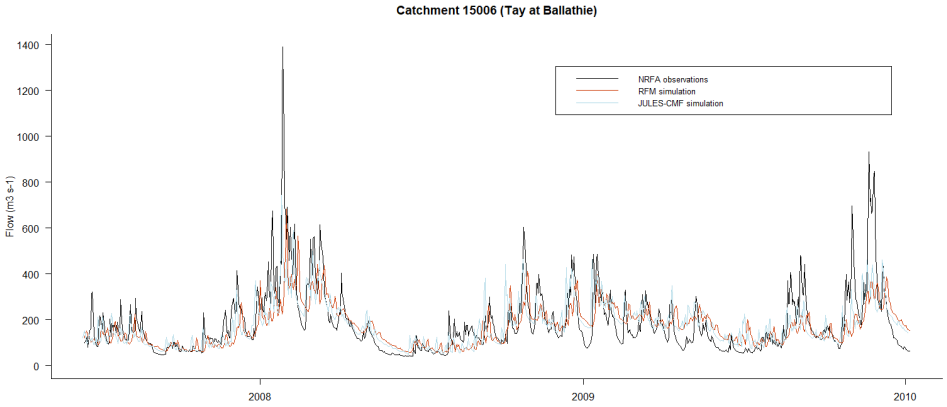


(b) Correlation

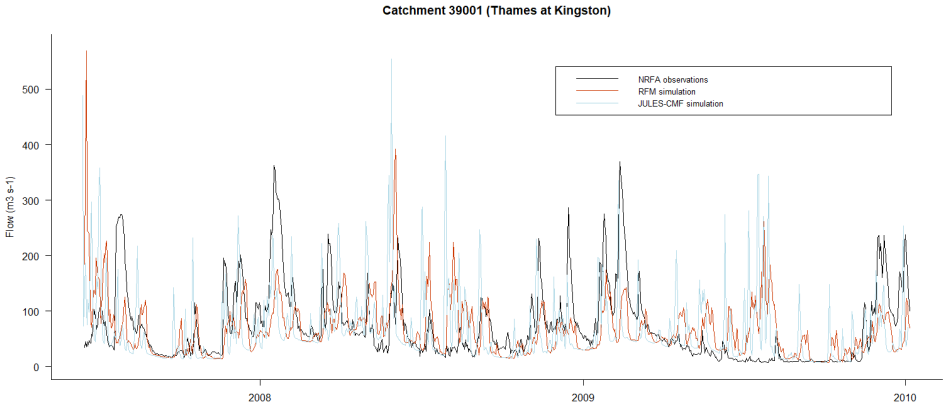


(c) BFI

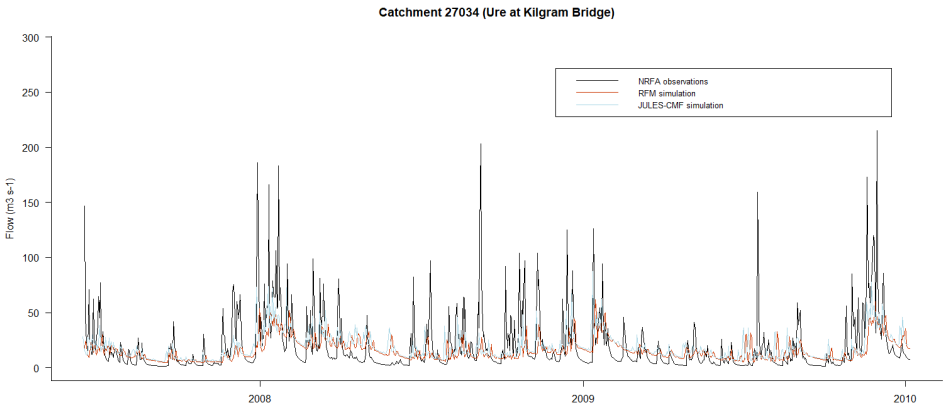
Figure 33: Metrics of daily river flow, comparing observed values from NRFA with simulated values. Arrows indicate the change from JULES-RFM to JULES-CaMaFlood. (a) KGE; horizontal lines indicate values of -0.41 and 0.29 which indicate that a simulation has some skill and can be considered good, respectively (Knoben et al., 2019) (b) correlation coefficient (Pearson's r) (c) Baseflow Index.



(a) Tay



(b) Thames



(c) Ure

Figure 34: Daily river flow (m^3 s^{-1}) for (a) Tay at Ballathie) (b) Thames at Kingston (c) Ure at Kilgram Bridge. Black lines show NRFA observations; light blue lines show JULES-CaMaFlood; orange lines show JULES-RFM.

timescales is controlled at least as much by geology and groundwater processes as by in-stream routing, and the poorer performance of JULES-CaMaFlood in this respect may in fact be indicative of deficiencies in the simulated runoff production (by TOPMODEL) in these simulations.

The pattern of generally improved results for a diverse subset of large catchments in GB is encouraging, particularly given that these runs represent the first time JULES-CaMaFlood has been used to simulate rivers at this scale, and using new ancillary fields.

5.3 Summary and future work on JULES-CaMaFlood

This project has seen the first use of JULES-CaMaFlood to simulate river flows and flooding in major British rivers. The existing code has been improved, and various issues in both the code and the ancillary data have been addressed. This represents a major advance in our capability.

CaMa-Flood uses several parameters that have been optimised for other configurations and ancillary data, often at coarser resolution. We are confident that further development of the code, and optimisation of parameter values, will improve the fit to measured flow and inundation data.

Here we have shown that the use of JULES-CaMaFlood has generally improved the simulation of river flow for the nine catchments studied; future work will extend the simulations and analysis to many more catchments, including many smaller rivers. A perceived advantage of CaMa-Flood is that it includes a more complete description of the physics of river flow than is available in the kinematic wave description of RFM. Although this already appears to have paid dividends, with improved results in large catchments, it might be that an even larger advantage will be found for smaller rivers in which backwater effects can be more important, particularly close to the confluence with larger streams. The analysis to date has also focussed on river flow and future work should also consider river flooding; another feature of CaMa-Flood is its ability to simulate river level (i.e. the depth of water in the river) which could result in improved simulations of flooding.

However, JULES-CaMaFlood generally requires a much shorter timestep for numerical stability than is used with JULES-RFM. Any improvement in the skill of simulations with JULES-CaMaFlood will need to be balanced against the increased computational burden. Careful analysis, likely including studies of individual flow and flood events, will be required so that we can gain a proper understanding of when and where JULES-CaMaFlood can be expected to produce better results. If the CaMa-Flood based parameterisation is to be used more widely within JULES it may be necessary to consider allowing the river model to run in parallel – at present all river parameterisations in JULES run in serial on a single processor.

In comparison to existing river options within JULES, JULES-CaMaFlood is more demanding in terms of ancillary information, requiring further ancillary variables to describe channel characteristics including depth, width, roughness and elevation. Any future development and uptake of JULES-CaMaFlood will need to be accompanied by corresponding development of our ability to generate the required ancillary files.

6 Towards UKCPnext

A configuration for JULES

This report has described recent developments in the code and data used with JULES that have the potential to improve the representation of terrestrial hydrology in the coupled REP model. Substantial progress has been made, and there are clear indications that improvements can be delivered, but clearly this is not the end of the process. In each area investigated there is further work to be done, as summarised at the end of previous sections of this report. Equally important is the need to identify the best combination of parameterisations, parameter values, and ancillary data – that is to integrate developments in different areas into the final package (configuration). In any such exercise there is a need to balance often conflicting demands. Aspects of each of the developments described in earlier sections (soil ancillaries, groundwater, and flow routing) appear to offer potential benefits. Although they have not yet been fully validated, we believe that all of the developments have shown sufficient promise to merit further investigation and consideration for inclusion in a configuration of the coupled model.

The importance of information on impacts

Publishing data and reports on potential future changes in climate, globally and for the UK, is clearly important, but it is only the first step. Many users require information on the consequent impacts, which are experienced more directly by people and by the environment. Examples might

include information on changes in heat waves rather than just changes in temperatures, or information on changes in water availability and flood risk rather than just changes in precipitation. Such information is necessary in order to develop appropriate policies and adaptation strategies in response to climate change.

A particular issue is that, on release of new climate projections, users often want information on how impacts are likely to differ from those using previous generations of climate projections. But the availability of such information is, at least initially, very limited. This causes a problem as:

> End users are often unable to process climate change information; paradoxically, the availability of new projections—intended to improve the response to climate change—may slow or halt adaptation action, with users reluctant to follow old plans in case they are wrong, but unable to formulate new approaches. (Kay et al., 2020)

Furthermore, the availability of

> Consistent, salient indicators, for example of changes in river flow, would help users to evaluate the impact of new climate science on current plans (Kay et al., 2020)

Having directly-simulated river flows available from UKCPnext would likely be a welcome addition, as it would reduce the time required to provide some information on potential hydrological impacts. An added advantage would be complete consistency between the climate and hydrological simulations, and the inclusion of land-surface feedbacks. Independent ('offline') hydrological simulations (i.e. driving separate hydrological models with data from the climate projections) derive their own soil moisture levels for example, which could be inconsistent with equivalent variables within the land-surface model used by the climate modelling system.

However, having no directly comparable simulated flows from previous generations of climate projections would complicate any required impact comparisons. There may still be a requirement for offline simulations of river flows, to fill this gap. There are also a number of other reasons that offline simulations may be required or desired.

Reasons for offline hydrological impact modelling

One perceived issue may be the performance of river flows simulated within the climate modelling system. For such river flow simulations to be relied upon for adaptation planning, substantial evaluation of performance would be required, both when driven with observation-based data, and by climate model data. This would need to look at not just the performance for average flows but for high/low flows, as it is often the extremes that are of the most interest in terms of impacts on people and the environment.

A possible related issue could be any perceived bias in the climate model data. Impacts modellers often apply 'bias-correction' techniques to adjust climate model data before use to drive impact simulations, although such methods are not as straightforward as many may believe since they generally lack a sound physical basis, do not satisfy conservation laws, and can even introduce 'artefacts' into the climate data (e.g. Ehret et al. 2012; Teng et al. 2015; Maraun et al. 2017). Nevertheless, bias-correction is often considered necessary, and the presence of non-linearity and thresholds in hydrological responses may mean that its use, or not, has a significant effect on simulated impacts (e.g. Willkofer et al. 2018). An additional complication is that bias-correction may need to be tailored to the impact under consideration; for example, a particular method may work well for one aspect of precipitation but not for another (e.g. Guillod et al. 2018).

There may also be a desire to use climate projections from other sources alongside UKCPnext, each applied in the same way. This is particular likely if the range of projections derived from the Hadley Centre global/regional climate models is perceived to be different to that derived from alternative climate models, as is the case in UKCP18 (Murphy et al. 2018, Fig 5.1).

In terms of hydrological impact assessment, there is an increasing desire to include potential hydrological model uncertainty (structure and/or parameterisation), as stakeholders gain greater understanding of uncertainty and how it can be dealt with in decision-making processes. There may also be a need to incorporate factors that may not be (fully) included in direct UKCPnext river flow simulations (e.g. abstractions/discharges, river regulation etc.), or to model smaller areas in greater detail (e.g. urban areas subject to flood risk). Related to this, there may be a desire to explore the effect of different potential adaptation options (e.g. future changes in abstractions/discharges, land-cover change, natural flood management etc.). In addition, there may be a need to simulate further variables than those related to water quantity (e.g. water quality, erosion), which may not be directly available within UKCPnext. Furthermore, there may be a desire to understand and

analyse the drivers of projected changes in river flows in different catchments, which may be easier to do with offline simulations.

Summary

Information on potential future changes in climate is crucial, but many users require information on impacts, experienced more directly by people and the environment. UKCPnext directly providing simulated river flows would be a welcome addition, to enable a faster assessment of potential hydrological impacts. However, there may still be a requirement for 'offline' hydrological simulations, for greater comparability to hydrological impacts derived from previous generations of climate projections. There are also a number of other reasons that offline simulations may be required or desired, including possible perceived issues with model performance or bias, consistent application of a broader range of climate projections, or the inclusion of a wide range of other factors affecting river flows.

7 Summary and final thoughts

This project was designed to investigate the relevance of new developments for the RAL3M configuration, using standalone JULES but driven by UM meteorology at the resolution of the coupled model. Although this addresses the coupled environment, it did not allow us to quantify the impacts on land hydrology of deficiencies in the simulated meteorology. In particular, any errors in the simulated precipitation fields – for example a poor simulation of the timing, location or intensity of a rainfall event – can potentially compromise the modelled riverflow. There are indications that this had some impact on results shown in this project, e.g. in Fig. 16 there is no sign of a modelled peak to match that observed early in October 2008, suggesting the UM did not capture this event. In future this could be addressed by adding new JULES runs that use best-estimates of meteorology, for example from CHESS (though these can come with their own limitations, for example the daily CHESS data needs to be disagrregated in time). Initially these runs could be done on the REP grid, allowing existing ancillary data to be reused. The analysis of modelled river flow should also incorporate statistics that are less sensitive to mistiming of the input precipitation – for example flow-duration curves look at the distribution of flows rather than seeking to match the flow on a given day.

The inclusion of further processes and alternative data sources within the JULES system adds to the challenge of identifying a best configuration. Previous work by Martínez-de la Torre et al. (2019) developed the use of PDM within a UK-focussed configuration, but more recently the RAL3M configuration has shifted to using the TOPMODEL-based parameterisation of runoff. Almost inevitably each of these alternative parameterisations and configurations will give better results for some locations/events/variables but will be poorer for others. Thus there is a need to better understand the controls on model performance and to quantify the trade offs as we optimise some aspects of the modelled system in preference to others. The optimisation might largely be achieved by modifying parameter values and fields, but it is possible that aspects of the JULES code also need to be refactored – for example to combine positive aspects of the TOPMODEL simulation of runoff with the representation of major aquifers via DGW.

The work undertaken in this project has significantly advanced our modelling capability and knowledge. In terms of developments to the JULES code (and supporting ancillary data), both JULES-DGW and JULES-CaMaFlood have been taken to the point where it has been possible to apply the models for GB-wide, multi-year simulations. This focus on developing technical capacity has meant that some of the analysis of results is more preliminary in nature – clarifying that the new parameterisations appear to be functioning reasonably, but with more detailed examination still to be completed. Although this particular project is at an end, many of these future analyses and items of research will be taken up in the Hydro-JULES project.

Acknowledgements

This work was funded by the Met Office under Contract P109933. The majority of model runs and analysis was performed using the JASMIN scientific data analysis environment .

We acknowledge and appreciate the assistance provided by the following people:

- Graham Weedon (Met Office) for performing cross-spectral analysis.

- Colleagues at BGS for providing groundwater ancillary data and for ongoing collaboration on the JULES-DGW code.
- Helen Davies (UKCEH) for providing additional river ancillary data for use with JULES-CaMaFlood.

References

Stamatis-Christos Batelis, Mostaquimur Rahman, Stephan Kollet, Ross Woods, and Rafael Rosolem. Towards the representation of groundwater in the joint uk land environment simulator. *Hydrological Processes*, 34(13), 2020. doi: 10.1002/hyp.13767.

M. J. Best, M. Pryor, D. B. Clark, G. G. Rooney, R. L. H. Essery, C. B. Ménard, J. M. Edwards, M. A. Hendry, A. Porson, N. Gedney, L. M. Mercado, S. Sitch, E. Blyth, O. Boucher, P. M. Cox, C. S. B. Grimmond, and R. J. Harding. The joint uk land environment simulator (jules), model description – part 1: Energy and water fluxes. *Geoscientific Model Development*, 4(3): 677–699, 2011. doi: 10.5194/gmd-4-677-2011. URL https://gmd.copernicus.org/articles/4/677/2011/.

E. M. Blyth, A. Martínez-de la Torre, , and E. L. Robinson. Trends in evapotranspiration and its drivers in great britain: 1961 to 2015. *Progress in Physical Geography: Earth and Environment*, 43(5), 2019. doi: 10.1177/0309133319841891. URL https://doi.org/10.1177/0309133319841891/.

M. Bush, T. Allen, C. Bain, I. Boutle, J. Edwards, A. Finnenkoetter, C. Franklin, K. Hanley, H. Lean, A. Lock, J. Manners, M. Mittermaier, C. Morcrette, R. North, J. Petch, C. Short, S. Vosper, D. Walters, S. Webster, M. Weeks, J. Wilkinson, N. Wood, and M. Zerroukat. The first met office unified model–jules regional atmosphere and land configuration, ral1. *Geoscientific Model Development*, 13(4):1999–2029, 2020. doi: 10.5194/gmd-13-1999-2020. URL https://gmd.copernicus.org/articles/13/1999/2020/.

D. B. Clark, L. M. Mercado, S. Sitch, C. D. Jones, N. Gedney, M. J. Best, M. Pryor, G. G. Rooney, R. L. H. Essery, E. Blyth, O. Boucher, R. J. Harding, C. Huntingford, and P. M. Cox. The joint uk land environment simulator (jules), model description – part 2: Carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, 4(3):701–722, 2011. doi: 10.5194/gmd-4-701-2011. URL https://gmd.copernicus.org/articles/4/701/2011/.

E. Cooper, E. Blyth, H. Cooper, R. Ellis, E. Pinnington, and S. J. Dadson. Using data assimilation to optimize pedotransfer functions using field-scale in situ soil moisture observations. *Hydrology and Earth System Sciences*, 25(5):2445–2458, 2021. doi: 10.5194/hess-25-2445-2021. URL https://hess.copernicus.org/articles/25/2445/2021/.

B J Cosby, G M Hornberger, R B Clapp, and T R Ginn. A Statistical Exploration of the Relationships of Soil Moisture Characteristics to the Physical Properties of Soils. *Water Resources Research*, 20(6):682–690, 1984. doi: 10.1029/WR020i006p00682.

Harmonized World Soil Database. Harmonized World Soil Database (version 1.2). URL http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/.

H.N. Davies, P. Rameshwaran, and V.A. Bell. Gridded (1km) physical river characteristics for the uk, 2022. URL https://doi.org/10.5285/6da95899-f3b8-4089-b621-560818aa78ba.

U. Ehret, E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, and J. Liebert. Hess opinions "should we apply bias correction to global and regional climate model data?". *Hydrology and Earth System Sciences*, 16(9):3391–3404, 2012. doi: 10.5194/hess-16-3391-2012. URL https://hess.copernicus.org/articles/16/3391/2012/.

Y. Fan, H. Li, and G. Miguez-Macho. Global patterns of groundwater table depth. *Science*, 339 (6122):940–943, 2013. ISSN 0036-8075. doi: 10.1126/science.1229881.

Ying Fan, Gonzalo Miguez-Macho, Christopher P. Weaver, Robert Walko, and Alan Robock. Incorporating water table dynamics in climate modeling: 1. water table observations and equilibrium water table simulations. *Journal of Geophysical Research: Atmospheres*, 112(D10), 2007. ISSN 2156-2202. doi: 10.1029/2006JD008111. D10125.

B. P. Guillod, R. G. Jones, S. J. Dadson, G. Coxon, G. Bussi, J. Freer, A. L. Kay, N. R. Massey, S. N. Sparrow, D. C. H. Wallom, M. R. Allen, and J. W. Hall. 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, 2018. doi: 10.5194/hess-22-611-2018. URL https://hess.copernicus.org/articles/22/611/2018/.

A. Gustard, A. Bullock, and J. M. Dixon. Low flow estimation in the united kingdom. (ih report no.108), 1992. URL http://nora.nerc.ac.uk/id/eprint/6050/1/IH_108.pdf.

T. Hallouin, R. J. Ellis, D. B. Clark, S. J. Dadson, A. G. Hughes, B. N. Lawrence, G. M. S. Lister, and J. Polcher. Unifhy v0.1: A community framework for the terrestrial water cycle in python. *Geoscientific Model Development Discussions*, 2021:1–30, 2021. doi: 10.5194/gmd-2021-419. URL https://gmd.copernicus.org/preprints/gmd-2021-419/.

Thibault Hallouin. hydroeval: an evaluator for streamflow time series in Python, April 2021. URL https://doi.org/10.5281/zenodo.4709652.

J. Hannaford, J. Mackay, M. Ascott, V. Bell, T. Chitson, S. Cole, C. Counsell, M. Durant, C. R. Jackson, A. Kay, R. Lane, M. Mansour, R. Moore, S. Parry, A. Rudd, M. Simpson, K. Facer-Childs, S. Turner, J. Wallbank, S. Wells, and A. Wilcox. eflag: enhanced future flows and groundwater. a national dataset of hydrological projections based on ukcp18. *Earth System Science Data Discussions*, 2022:1–40, 2022. doi: 10.5194/essd-2022-40. URL https://essd.copernicus.org/preprints/essd-2022-40/.

Shaun Harrigan, Jamie Hannaford, Katie Muchan, and Terry J. Marsh. Designation and trend analysis of the updated UK Benchmark Network of river flow stations: the UKBN2 dataset. *Hydrology Research*, 49(2):552–567, 10 2017. ISSN 0029-1277. doi: 10.2166/nh.2017.058. URL https://doi.org/10.2166/nh.2017.058.

Tomislav Hengl, Jorge Mendes de Jesus, Gerard B. M. Heuvelink, Maria Ruiperez Gonzalez, Milan Kilibarda, Aleksandar Blagotić, Wei Shangguan, Marvin N. Wright, Xiaoyuan Geng, Bernhard Bauer-Marschallinger, Mario Antonio Guevara, Rodrigo Vargas, Robert A. MacMillan, Niels H. Batjes, Johan G. B. Leenaars, Eloi Ribeiro, Ichsani Wheeler, Stephan Mantel, and Bas Kempen. Soilgrids250m: Global gridded soil information based on machine learning. *PLOS ONE*, 12(2): 1–40, 02 2017. doi: 10.1371/journal.pone.0169748.

Jannis M Hoch and Mark A Trigg. Advancing global flood hazard simulations by improving comparability, benchmarking, and integration of global flood models. *Environmental Research Letters*, 14(3):034001, feb 2019. doi: 10.1088/1748-9326/aaf3d3. URL https://doi.org/10.1088/1748-9326/aaf3d3.

Alison Lindsey Kay, Glenn Watts, Steven C. Wells, and Stuart Allen. The impact of climate change on u. k. river flows: A preliminary comparison of two generations of probabilistic climate projections. *Hydrological Processes*, 34(4):1081–1088, 2020. doi: https://doi.org/10.1002/hyp.13644. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.13644.

W. J. M. Knoben, J. E. Freer, and R. A. Woods. Technical note: Inherent benchmark or not? comparing nash–sutcliffe and kling–gupta efficiency scores. *Hydrology and Earth System Sciences*, 23 (10):4323–4331, 2019. doi: 10.5194/hess-23-4323-2019. URL https://hess.copernicus.org/articles/23/4323/2019/.

H. W. Lewis, J. M. Castillo Sanchez, J. Graham, A. Saulter, J. Bornemann, A. Arnold, J. Fallmann, C. Harris, D. Pearson, S. Ramsdale, A. Martínez-de la Torre, L. Bricheno, E. Blyth, V. A. Bell, H. Davies, T. R. Marthews, C. O'Neill, H. Rumbold, E. O'Dea, A. Brereton, K. Guihou, A. Hines, M. Butenschon, S. J. Dadson, T. Palmer, J. Holt, N. Reynard, M. Best, J. Edwards, and J. Siddorn. The ukc2 regional coupled environmental prediction system. *Geoscientific Model Development*, 11(1):1–42, 2018. doi: 10.5194/gmd-11-1-2018. URL https://gmd.copernicus.org/articles/11/1/2018/.

D. Maraun, T. Shepherd, and M. Widmann. Towards process-informed bias correction of climate change simulations. *Nature Climate Change*, 7:764–773, 2017. doi: https://doi.org/10.1038/nclimate3418.

T. R. Marthews, C. A. Quesada, D. R. Galbraith, Y. Malhi, C. E. Mullins, M. G. Hodnett, and I. Dharssi. High-resolution hydraulic parameter maps for surface soils in tropical south america. *Geoscientific Model Development*, 7(3):711–723, 2014. doi: 10.5194/gmd-7-711-2014. URL https://gmd.copernicus.org/articles/7/711/2014/.

T. R. Marthews, S. J. Dadson, D. B. Clark, E. M. Blyth, G. Hayman, D. Yamazaki, O. R. E. Becher, A. Martínez-de la Torre, C. Prigent, and C. Jiménez. Inundation prediction in tropical wetlands from *JULES-CaMa-Flood* global land surface simulations. *Hydrology and Earth System Sciences Discussions*, 2021:1–31, 2021. doi: 10.5194/hess-2021-109. URL https://hess.copernicus.org/preprints/hess-2021-109/.

A. Martínez-de la Torre, E. M. Blyth, and G. P. Weedon. Using observed river flow data to improve the hydrological functioning of the jules land surface model (vn4.3) used for regional coupled modelling in great britain (ukc2). *Geoscientific Model Development*, 12(2):765–784, 2019. doi: 10.5194/gmd-12-765-2019. URL https://gmd.copernicus.org/articles/12/765/2019/.

Alberto Martínez-de la Torre and Gonzalo Miguez-Macho. Groundwater influence on soil moisture memory and land–atmosphere fluxes in the iberian peninsula. *Hydrology and Earth System Sciences*, 23(12), 2019. ISSN 4909-4932. doi: 10.5194/hess-23-4909-2019.

Gonzalo Miguez-Macho, Ying Fan, Christopher P. Weaver, Robert Walko, and Alan Robock. Incorporating water table dynamics in climate modeling: 2. formulation, validation, and soil moisture simulation. *Journal of Geophysical Research: Atmospheres*, 112(D13), 2007. ISSN 2156-2202. doi: 10.1029/2006JD008112. D13108.

J.M. Murphy, G.R. Harris, and D.M.H. Sexton. Ukcp18 land projections: Science report. met office hadley centre, exeter, uk. 2018.

Guo-Yue Niu, Zong-Liang Yang, Robert E. Dickinson, Lindsey E. Gulden, and Hua Su. Development of a simple groundwater model for use in climate models and evaluation with gravity recovery and climate experiment data. *Journal of Geophysical Research: Atmospheres*, 112(D7), 2007. ISSN 2156-2202. doi: 10.1029/2006JD007522. D07103.

E Pinnington. LaVEnDAR vn 1.0.0. *https://github.com/pyearthsci/lavendar*, 2021. doi: doi: 10.5281/zenodo.2654852.

E. Pinnington, T. Quaife, A. Lawless, K. Williams, T. Arkebauer, and D. Scoby. The land variational ensemble data assimilation framework: Lavendar v1.0.0. *Geoscientific Model Development*, 13(1):55–69, 2020. doi: 10.5194/gmd-13-55-2020.

E. Pinnington, J. Amezcua, E. Cooper, S. Dadson, R. Ellis, J. Peng, E. Robinson, R. Morrison, S. Osborne, and T. Quaife. Improving soil moisture prediction of a high-resolution land surface model by parameterising pedotransfer functions through assimilation of smap satellite data. *Hydrology and Earth System Sciences*, 25(3):1617–1641, 2021. doi: 10.5194/hess-25-1617-2021. URL https://hess.copernicus.org/articles/25/1617/2021/.

V. Stanley, S.and Antoniou, A. Askquith-Ellis, L.A. Ball, E.S. Bennett, J.R. Blake, D.B. Boorman, M. Brooks, M. Clarke, H.M. Cooper, N. Cowan, A. Cumming, J.G. Evans, P. Farrand, M. Fry, O.E. Hitt, W.D. Lord, R. Morrison, G.V. Nash, D. Rylett, P.M. Scarlett, O.D. Swain, M. Szczykulska, J.L. Thornton, E.J. Trill, A.C. Warwick, and B. Winterbourn. Daily and sub-daily hydrometeorological and soil data (2013-2019) [cosmos-uk], 2021. URL https://doi.org/10.5285/b5c190e4-e35d-40ea-8fbe-598da03a1185.

J. Teng, N. J. Potter, F. H. S. Chiew, L. Zhang, B. Wang, J. Vaze, and J. P. Evans. How does bias correction of regional climate model precipitation affect modelled runoff? *Hydrology and Earth System Sciences*, 19(2):711–728, 2015. doi: 10.5194/hess-19-711-2015. URL https://hess.copernicus.org/articles/19/711/2015/.

B. Tóth, M. Weynants, A. Nemes, A. Makó, G. Bilas, and G. Tóth. New generation of hydraulic pedotransfer functions for europe. *European Journal of Soil Science*, 66(1):226–238, 2015. doi: https://doi.org/10.1111/ejss.12192. URL https://bsssjournals.onlinelibrary.wiley.com/doi/abs/10.1111/ejss.12192.

Robert L. Walko, Larry E. Band, Jill Baron, Timothy G. F. Kittel, Richard Lammers, Tsengdar J. Lee, Dennis Ojima, Roger A. Pielke Sr., Chris Taylor, Christina Tague, Craig J. Tremback, and Pier Luigi Vidale. Coupled atmosphere–biophysics–hydrology models for environmental modeling. *Journal of Applied Meteorology*, 39(6):931–944, 2000. doi: 10.1175/1520-0450(2000)039⟨0931:CABHMF⟩2.0.CO;2.

Florian Willkofer, Franz-Josef Schmid, Holger Komischke, Jane Korck, Marco Braun, and Ralf Ludwig. The impact of bias correcting regional climate model results on hydrological indicators for bavarian catchments. *Journal of Hydrology: Regional Studies*, 19:25–41, 2018. ISSN 2214-5818. doi: https://doi.org/10.1016/j.ejrh.2018.06.010. URL https://www.sciencedirect.com/science/article/pii/S2214581817303324.