Developing water supply reservoir operating rules for large-scale hydrological modelling

Saskia Salwey[1], Gemma Coxon[1,2], Francesca Pianosi[2,3], Rosanna Lane[4], Chris Hutton[5], Michael Bliss Singer[6,7,8], Hilary McMillan[9], Jim Freer[1,10]

[1] School of Geographical Sciences, University of Bristol, Bristol, United Kingdom.
[2] Cabot Institute, University of Bristol, Bristol, United Kingdom.
[3] School of Civil, Aerospace and Design Engineering, University of Bristol, Bristol United Kingdom.
[4] UK Centre for Ecology & Hydrology, Wallingford, United Kingdom.
[5] Wessex Water Services Ltd, Bath, United Kingdom.
[6] School of Earth and Environmental Sciences, Cardiff University, Cardiff, United Kingdom.
[7] Earth Research Institute, University of California Santa Barbara, Santa Barbara, USA.
[8] Water Research Institute, Cardiff University, Cardiff, United Kingdom.
[9] Department of Geography, San Diego State University, San Diego, California, USA.
[10] University of Saskatchewan, Centre for Hydrology, 116-1151 Sidney Street, Canmore, Alberta, Canada.

Correspondence to: Saskia Salwey (saskia.salwey@bristol.ac.uk)

Abstract. Reservoirs are ubiquitous water infrastructure~~key components of many water supply systems~~, providing functional capability to manage, and often mitigate, hydrological variability across space and time. The presence and operation of a reservoir controls the downstream flow regime, such that in many locations understanding reservoir operations is crucial to understanding the hydrological functioning of a catchment. ~~Although s Substantial progress has been made~~Despite many advances in modelling reservoir operations, inclusion of reservoirs in large-scale hydrological modelling remains challenging, particularly when the number of reservoirs is large and data access is limited. ~~but several key challenges remain, particularly for large-scale applications including hundreds of reservoirs. In these cases, generic and uncalibrated reservoir operating rules are often applied. However, these rules were developed from global reservoir databases that consist mostly of large irrigation reservoirs and thus are not transferable to smaller reservoirs or those fulfilling other purposes, such as water supply. An alternative option is to use a calibrated, data-driven approach but such techniques require reservoir inflows, outflows and storage data which are rarely available across hundreds of reservoirs. To overcome these problems, h~~Here we design a set of simple reservoir operating rules (with only two calibrated parameters) focused on simulating small water supply reservoirs across large scales with various types of open access data (i.e. catchment attributes and flows at downstream gauges~~general catchment attributes such as surface area or reservoir capacity, and flows at downstream gauges~~). ~~Using Great Britain as a case study, we~~We integrate our rules into a national-scale hydrological model of Great Britain and compare hydrological simulations ~~from two modelling scenarios,~~ with and without the new reservoir component. Our simple reservoir operating rules significantly increase model performance in reservoir-impacted catchments, particularly when the rules are calibrated individually at each downstream gauge. We also test the feasibility of using transfer functions (which transform reservoir and catchment attributes into operating rule parameters) to identify a nationally-consistent ~~parameterisation~~calibration. This works well in ~50% of the catchments, while nuances in individual reservoir operations limit performance in others. We suggest that

our approach should provide a lower benchmark for simulations in catchments containing water supply reservoirs, and that more complex methods should only be considered where they outperform our simple approach.

1 Introduction

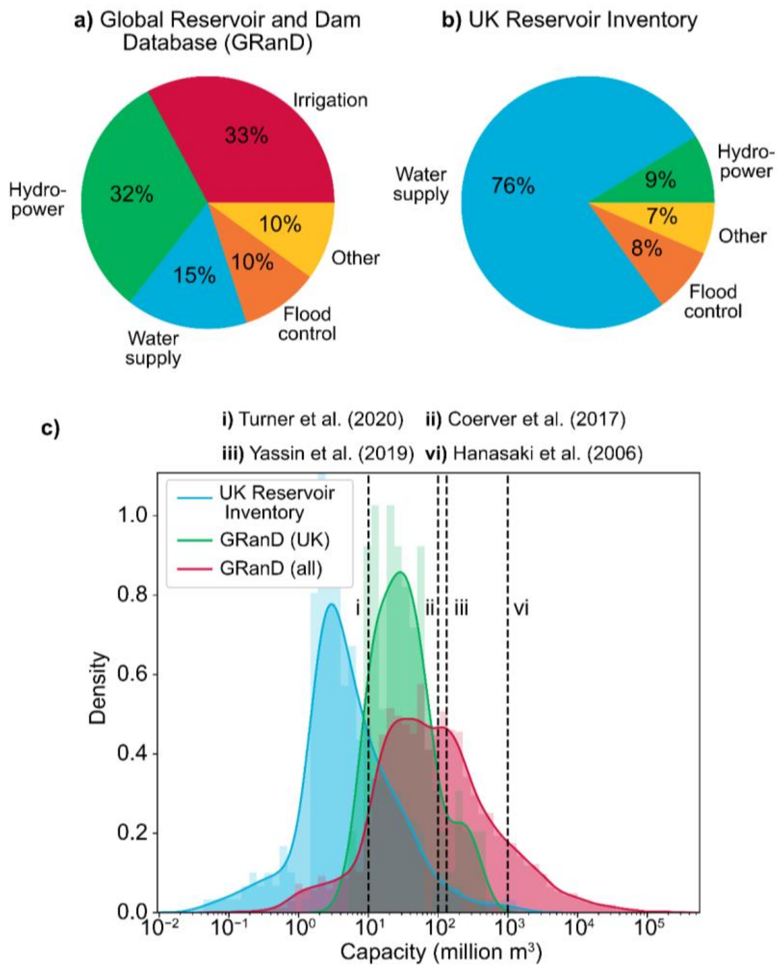
Effective and reliable water resource management is essential for food, water, and energy security (Sardo et al., 2023; Carrillo and Frei, 2009; Brown et al., 2015). To cope with increasing hydrologic variability and to ensure a reliable supply of water, national to continental- scale solutions are needed (Mcmillan et al., 2016). This requires more integrated and resilient water resource systems which can manage, and often mitigate, hydrological variability across space and time (Dobson et al., 2020; Wendt et al., 2021; Gaupp et al., 2015). A key part of these inter-connected water management systems is reservoirs. Reservoirs play a vital role in the supply and management of water resources and their operations significantly alter downstream flow (Döll et al., 2009; Tebakari et al., 2012; Vörösmarty et al., 2003; Adam et al., 2007; Salwey et al., 2023). As a result, appropriately representing reservoirs and their operating rules in hydrological modelling frameworks is a key area of research (Brown et al., 2015).

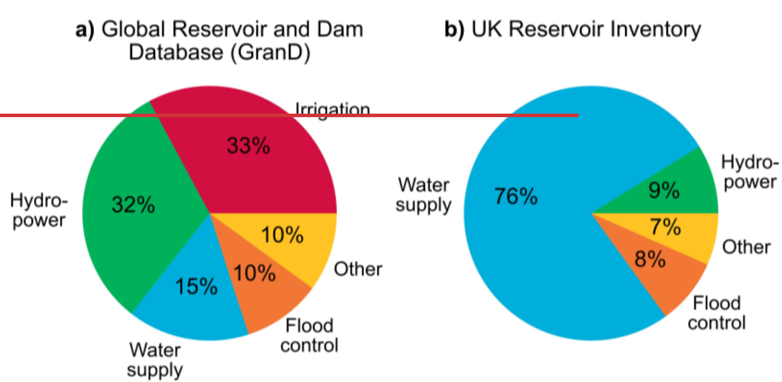
To model reservoir operations at the largest scale, global reservoir databases and uncalibrated operating rules are available (Hanasaki et al., 2006; Wisser et al., 2010; Lehner et al., 2011b). By simulating how much water is released from a reservoir at each timestep, uncalibrated reservoir operating rules integrated into global hydrological models have been shown to yield significant improvements in streamflow simulations (Abeshu et al., 2023; Hanasaki et al., 2006). However, global reservoir rules and datasets are often not suitable for application over national/continental scales. Using Great Britain (GB) as an example, the distribution of both reservoir type and size is markedly different when comparing data from global (Global Reservoir and Lakes Database, G~~r~~RanD) and national (UK Reservoir Inventory) reservoir databases (Figure 1). Over three-quarters of the reservoirs in GB are designed for water supply, whereas globally, reservoirs are primarily designed for irrigation (33%) and/or hydropower (31%). Furthermore, reservoirs in global databases (GR~~r~~anD) tend to be much larger than in the UK reservoir inventory. Consequently, reservoir operating rules developed from these global databases, for global-scale application, are often unsuitable for applications in national-scale models.

One option for developing more tailored reservoir operating rules at the national scale is to use a calibrated, data-driven approach. ResOpsUS (Steyaert et al., 2022) is a national US dataset providing historical timeseries of reservoirs storage, outflow and inflow for ~~thousands of~~over 600 US reservoirs. This dataset has enabled the development of a national-scale inventory of tailored, empirically derived, operating rules for each reservoir (Turner et al., 2021). When forced with observed inflow data, these data-driven rules reproduce downstream flow observations significantly better than uncalibrated, generic operating rules (Turner et al., 2020). However, these data-driven operating rules no longer outperform the generic alternatives when integrated into a hydrological model, i.e. when forced with simulated inflows ("online testing"), instead of observed ones ("offline testing") (Turner et al., 2021). Furthermore, extensive datasets such as ResOpsUS are seldom available at national-scale and consequently the approach is challenging to apply elsewhere.

In this paper, we develop a set of simple reservoir operating rules tailored towards water supply reservoirs that can be implemented across local, national or global scales. We focus on water supply reservoirs as there is a lack of generic operating rules for this type of reservoir, despite their importance for water supply and management in many countries, including our application domain (Great Britain). Although offline testing of operating rules is common in the literature (Zhao et al., 2016; Yassin et al., 2019), here we integrate and test reservoir representation in a hydrological model from the start.

Our simple operating rules have parameters which are linked to catchment and reservoir attributes via transfer functions. Parameter regionalisation, where transfer function parameters are calibrated by assuming prior relationships between model parameters and catchment attributes, is common in hydrological modelling (Samaniego et al., 2010), but has not previously been applied to modelling reservoir operating rules. We present the results from two methods of calibration. The first method uses common bounds for the transfer function parameters but within these bounds finds an "optimal" parameter set for each catchment (we call this a catchment-by-catchment calibration). The second method identifies one set of "optimal" transfer function parameters that can be estimated and applied across all reservoirs (we call this a nationally-consistent calibration). This latter method facilitates the simulation of operating rules in ungauged or data-poor basins. The simplicity of our rules and use of transfer functions allows us to simulate reservoir operations over hundreds of reservoirs using only open-source data.





Formatted: Normal (Web), Justified, Line spacing: 1.5 lin

Figure ~~1~~**1**: a, b) Pie charts showing the distribution of reservoir types across the a) G~~r~~RanD database and b) UK Reservoir Inventory. The ‘other’ category groups together reservoirs designed for uses such as recreation, navigation and fishing which make up a small proportion of the database. c) Histogram showing the distribution of reservoir capacities across the UK ~~R~~reservoir Inventory (blue), UK Reservoirs in G~~r~~RanD (green) and full GR~~r~~anD database (red). Dashed lines and associated labels i-vi represent the smallest reservoirs considered by some key papers discussed in the Introduction.

Formatted: Font: 10 pt, Bold

Formatted: Font: 10 pt

2 Developing large-scale reservoir operating rules

The following section describes the generic reservoir operating rules introduced in this paper for the large-scale simulation of water supply reservoirs. We discuss their specific application to Great Britain in Section 3.

2.1 Operating rules

As is common in modelling reservoirs, we consider reservoirs to be zero dimensional points, where their dynamics are controlled by a mass balance equation. The reservoir mass balance is updated at every timestep and represented with the following equation:

$$\frac{\Delta S}{\Delta t} = I_t - CF_t - ABS_t - spill_t \qquad (1)$$

Where S represents the reservoir storage and t is time. I_t is the inflow simulated by the hydrological model per unit time, CF_t is the volume per unit time of water released into the downstream river to fulfil environmental flow requirements (known as the compensation flow), ABS_t is the volume per unit time of water abstracted from the reservoir for public water supply and $spill_t$ is the volume of water remaining above the reservoir capacity per unit time which must be released downstream (this is calculated after all other fluxes have been calculated). Equation 1 does not include evaporation as this is not a big component of the mass balance for reservoirs in Great Britain (see section 3.4) (Dobson et al., 2020), however, evaporation could easily be included into the mass balance for reservoirs where this is important.

We use transfer functions to determine the relationships between catchment attributes (e.g. catchment size or mean annual rainfall), reservoir attributes (e.g. capacity or use) and the rates of compensation flow (CF) and abstraction (ABS) as follows:

$$ABS = f_1(catchment\ attributes, reservoir\ attributes, parameter_{1\ldots n}) \qquad (2)$$

$$CF = f_2(catchment\ attributes, parameter_{1\ldots n}) \qquad (3)$$

The catchment and reservoir attributes used within these functions will vary depending on what data are available and a selection of attributes may have to be tested before a sensible relationship is established. In some cases attributes may be combined (for example by normalising reservoir storage by catchment area). In this study we calibrate the parameters in the transfer functions above both in a catchment-by-catchment manner, and nationally, identifying one parameterisation for the entire sample of catchments. The development of the transfer functions for our study area (Great Britain) is described in more detail in section 3.5.1. The compensation flow and abstraction fluxes at each timestep CF_t and ABS_t are then calculated (in m^3~~Ml~~/day) based on the current reservoir storage as follows:

$$CF_t = \begin{cases} CF & \boldsymbol{if}\ S_t > S_{min} + CF \cdot \Delta t \\ (S_t - S_{min})/\Delta t & \boldsymbol{if}\ S_{min} < S_t < S_{min} + CF \cdot \Delta t \\ 0 & \boldsymbol{if}\ S_t \le S_{min} \end{cases} \cancel{\begin{cases} CF & \boldsymbol{if}\ S_t > S_{min} + CF \cdot \Delta t \\ (S_t - S_{min})/\Delta t & \boldsymbol{if}\ S_{min} < S_t < S_{min} + CF \cdot \Delta t \\ 0 & \boldsymbol{if}\ S_t \le S_{min} \end{cases}} \tag{4}$$

$$ABS_t = \begin{cases} ABS & \boldsymbol{if}\ S_t > S_{min} + CF \cdot \Delta t + ABS \cdot \Delta t \\ (S_t - S_{min})/\Delta t - CF & \boldsymbol{if}\ S_{min} < S_t < S_{min} + CF \cdot \Delta t + ABS \cdot \Delta t \\ 0 & \boldsymbol{if}\ S_t \le S_{min} + CF \cdot \Delta t \end{cases} \cancel{\begin{cases} ABS & \boldsymbol{if}\ S_t > S_{min} + CF \cdot \Delta t + ABS \cdot \Delta t \\ (S_t - S_{min})/\Delta t - CF & \boldsymbol{if}\ S_{min} < S_t < S_{min} + CF \cdot \Delta t + ABS \cdot \Delta t \\ 0 & \boldsymbol{if}\ S_t \le S_{min} + CF \cdot \Delta t \end{cases}} \tag{5}$$

In this instance CF_t is given priority and removed before ABS_t hence the calculation of ABS_t must ensure there is enough storage for the CF_t to be removed first. This step ensures there is sufficient storage for these fluxes to be removed and prevents the reservoir from being drained below its minimum capacity ~~(~~S_{min} (which can either be specified using site-specific data or estimated as a percentage of total reservoir capacity). Note that whilst in this study we use a fixed value for CF and ABS over time, seasonal, or sub-seasonal transfer functions could be developed to vary these parameters over time if appropriate.

To implement these operating rules into a hydrological model, the user will need data describing reservoir use (in this case the reservoir ought to be designed for water supply), capacity (to represent storage) and location (to locate the reservoir on the river network). These data can be obtained nationally from datasets such as the UK Reservoir Inventory (Durant and Counsell, 2018), Inventory of Dams in Germany (Speckhann et al., 2021), or National Inventory of Dams in the US (Usace, 2018), or globally from datasets such as GR~~r~~anD (Lehner et al., 2011b) and GeoDAR (Wang et al., 2022). In order to define the transfer functions used in the operating rules above, a small sample of observed compensation flow and abstraction data is needed (ideally for at least 10 reservoirs). These data can be found in documents published by water companies (e.g. Water Resource Management Plans (WRMP), or Drought Plans), academic literature, or, where a gauge is located close to a reservoir outflow, can be inferred from the downstream flow timeseries.

3 Application to national-scale hydrological modelling in Great Britain

The following section describes the application of the simple operating rules introduced above to the national-scale hydrological modelling of Great Britain (GB). Like many other countries, GB faces increasing water scarcity, where changing patterns of rainfall and evapotranspiration could add to the increasing pressures of future demand (Watts et al., 2015; Dobson et al., 2020). At present, water management is carried out mostly by local water companies, but to ensure water supply remains resilient to change, GB is considering several more regional or national strategic solutions (Murgatroyd and Hall, 2020). Reservoirs make up a large component of the domestic water supply system in GB and have a significant influence on river flows (16% of all river basins contain one or more reservoirs) (Salwey et al., 2023; Tijdeman et al., 2018). Due to the size and type of reservoirs found across GB (mostly small water supply reservoirs), global-scale approaches to reservoir representation are not applicable. This serves as a good case study for somewhere where existing reservoir operating rules are not suitable, and the national context of water management is particularly important.

To demonstrate the application of our reservoir operating rules across GB, they are implemented into the DECIPHeR hydrological model (section 3.1). We use hydrometeorological data from 1970-2020 (section 3.2) to run model simulations in two samples of catchments: reservoir catchments, i.e. all those catchments draining into a gauge that lies downstream of one or more water supply reservoirs, and near-natural catchments, which have no upstream reservoirs (Figure 2). The near-natural simulations use Multiscale Parameter Regionalisation (MPR) (Mizukami et al., 2017; Samaniego et al., 2017; Lane et al., 2021) to estimate DECIPHeR's natural model parameters. When using the term 'natural model parameters' we refer to the seven standard DECIPHeR parameters which are designed to simulate hillslope hydrology unimpacted by humans (section 3.3). In the reservoir catchments, DECIPHeR is run both with, and without, reservoir representation (section 3.4 and 3.5) to compare the difference in model performance before and after incorporating our new reservoir operating rules (section 3.6). Since most national-scale models of GB do not include reservoir representation (e.g. G2G or GR4J; (Smith et al., 2019; Rudd et al., 2019)) we consider this to be a suitable benchmark. Finally, the model is evaluated against a suite of model performance metrics (section 3.6) to better understand where and when our simple reservoir operating rules result in better (or worse) model performance and to act as a benchmark for future model improvements.

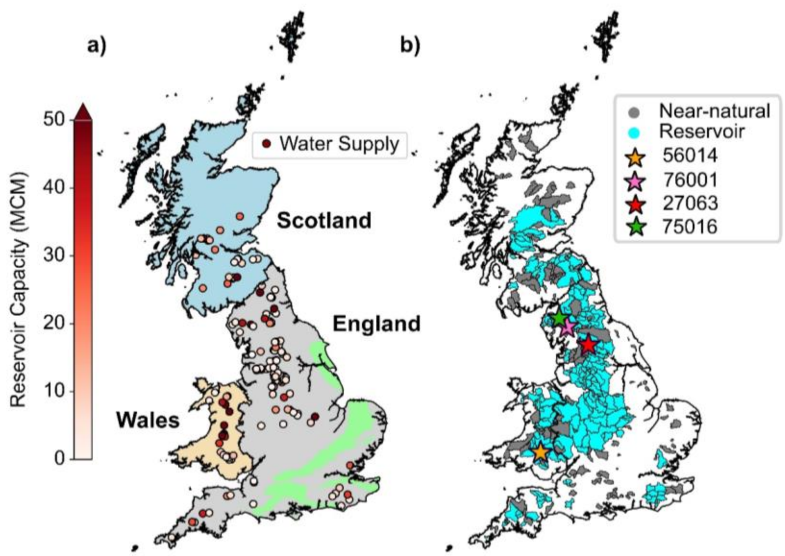


Figure 2. Distribution of (a) water supply reservoirs across GB and (b) near-natural and reservoir catchments used in this study. Reservoirs are coloured by their storage capacity and the four catchments featured in Figures 5-8 are highlighted with stars on subplot b.

Formatted: Font: Not Bold

Formatted: Font: Not Bold

Formatted: Font: Not Bold

3.1 DECIPHeR

DECIPHeR (Dynamic fluxEs and Connectivity for Predictions of HydRology) is a flexible, semi-distributed hydrological modelling framework which has previously been implemented across a range of scales (e.g. catchment to national scale) and locations (e.g. Europe, Asia, Africa) and has both been coupled to other models and had additional modules incorporated (Shannon et al., 2023; Dobson et al., 2020; Devitt, 2019; Fadhliani et al., 2021). The model has been applied nationally in Great Britain and demonstrated good performance (Lane et al., 2021; Coxon et al., 2019), with generally better model performance in wetter catchments in the North and West of GB. However, since the model has no reservoir representation, performance is usually poor in catchments downstream of reservoirs. At present, in these locations the model has no knowledge of reservoir locations, and flow in these catchments is simulated as if they were natural.

DECIPHeR uses hydrological response units (HRUs) to split up the landscape into non-contiguous spatial elements that share similar characteristics in landscape attributes (e.g., soil, topography or geology) and spatially varying inputs (e.g., rainfall). Each HRU then acts as a separate model store capable of having different spatial inputs, model parameter values and/or model structures to represent different and localized processes. In this study, HRUs were classified using a 2.2-km input grid (consistent with national climate projection data), which were further sub-divided by gauged sub-catchments (which include those defined by reservoir nodes) and percentiles of slope and upslope accumulated area (i.e. the area of land draining to a particular point in the landscape). This ensures that HRU's cascade downslope to the bottom of the valley and the spatial variability of the climatic inputs is appropriately represented. ~~In this study, HRUs were classified using a 2.2-km input grid (consistent with national climate projection data), which were further sub-divided by gauged sub-catchments (which include those defined by reservoir nodes) and three equal classes of slope and accumulated area.~~

3.2 Hydrometeorological Data

To drive the hydrological model, precipitation and potential evapotranspiration (PET) timeseries are needed. In this study, we use observation-based gridded daily precipitation and PET data derived from the HadUK-grid dataset which provides a number of climate variables on a 1km x 1km grid across the UK (Hollis et al., 2019). Daily precipitation data from HadUK-grid is available from 1891-present and derived from the Met Office UK rain gauge network. The observed precipitation data from the rain gauge network are quality controlled, and then inverse-distance weighted interpolation is used to generate the daily rainfall grids (Hollis et al., 2019). Daily PET was calculated using the Penman- Monteith equation applied to climate variables available from HadUK-grid (Robinson et al., 2023). These data are available from 1969- 2021. While the climatic variables are available on a 1km x 1km grid, these were upscaled to a 2.2-km grid for use in the hydrological modelling. This was chosen to align with the existing model setup and the grid used for national climate projections (Robinson et al., 2021; Lane and Kay, 2022).

In this study, we run the model from 1970- 2020 since it encompasses a variety of climatic conditions. The first 5 years of the time window were used as a spin-up period where no model evaluation is carried out. Simulations are evaluated from 1975 onwards (or from the date the reservoir construction was completed) using daily streamflow timeseries from the UK National River Flow Archive (NRFA) (https://nrfa.ceh.ac.uk/). Since 96% of reservoirs in GB were built by 1980, we can evaluate the model performance across most of the simulated period (where the flow data are available at the relevant gauge).

3.3 Calibration in near-natural catchments

In this study we calibrate the parameters in the reservoir operating rules independently from the natural model parameters. This avoids unrealistic parameterisations or equifinality, where natural parameters might mimic reservoir processes (Dang et al., 2020).

DECIPHeR has seven natural model parameters which describe how much water the soils can store and how permeable they are, the river channel velocity and the transmissivity of the sub-surface (Coxon et al., 2019). To generate nationally-consistent parameter fields for DECIPHeR's natural model parameters, we use multiscale parameter regionalisation (MPR), following the method introduced by Samaniego et al. (2010) and applied in DECIPHeR by Lane et al. (2021). High resolution parameter fields are generated by linking model parameters to spatial catchment characteristics via transfer functions and

~~then~~subsequently using MPR to upscale~~s~~ the~~se~~ parameter fields to the model resolution. ~~This method links parameters to spatial catchment characteristics via transfer functions, resulting in a hydrologically meaningful parameterisation~~ ~~calibration~~ ~~where the same rules are used to determine model parameters across the country.~~ Transfer functions were defined for each natural model parameter (see Lane et al. 2021), and the transfer function parameters were calibrated simultaneously across all non-reservoir (or near-natural) catchments. Catchments with reservoirs were excluded from this calibration, as the purpose was to find parameter fields which resulted in good model performance for natural catchments before the addition of any reservoir component.

We calibrated the transfer function parameters using a set of simulations in near-natural catchments selected from the UK benchmark network (Figure 2b). The UK benchmark network (Harrigan et al., 2018) consists of 137 catchments chosen for their lack of human influence and near-natural flow regime. In each catchment, we ran 5,000 simulations sampling the transfer function parameters between set bounds. The top 10 natural transfer function parameter combinations were then chosen by calculating the non-parametric KGE ~~(assessed using the non-parametric KGE~~ (Pool et al., 2018) (~~,~~ see section 3.6) in all near-natural catchments. ~~. We then~~ The 10 combinations with the highest average non-parametric KGE across all the near-natural catchments were subsequently used to determine ~~unbiased values for~~ the natural model parameters in reservoir catchments.

3.4 Integrating reservoirs into the river network

To integrate new reservoir representation into DECIPHeR, we modified the river routing and represent each reservoir as a zero dimensional point on the river network. Channel flow routing in DECIPHeR is modelled using a set of time delay histograms for the points on the river network where river flow timeseries is required. A fixed channel wave velocity is applied throughout the network to account for delay and attenuation in the simulated flows. The reservoir points are placed at their outflow locations as nodes on the river network. These nodes break up the river reach such that during a simulation, incoming river flow is manipulated according to the operating rules described in Section 2.1 before it continues downstream. Reservoir storage is also simulated, and the timeseries can be obtained as an output. In this study we do not consider evaporation from the reservoirs, this is partly because the flux is small in GB, and partly because we model reservoirs as zero-dimensional points and so we already simulate evaporation from the underlying area. We do not have evaporation relationships nor surface area data, and we note that other studies also opted to exclude evaporation from reservoirs across Great Britain, where even the largest reservoir (Kielder) only has evaporation equal to 3% of its inflow (Dobson et al., 2020).

We use a 50-m gridded digital elevation model (Intermap Technologies, 2009) to generate the river network in DECIPHeR, extracting headwater cells from an open-access river network which maps the rivers across GB, generated by the Ordinance Survey ~~OS rivers layer~~ (Ordnance Survey, 2023). These cells are then routed downstream to generate a river network. Once the river network has been generated, reservoir locations and capacities were extracted for water supply reservoirs from the UK Reservoir Inventory (Durant and Counsell, 2018) which contains data on UK reservoirs with storage exceeding 1.6 million cubic metres (MCM) and a selection of smaller ones. After cross-referencing the UK Reservoir Inventory with the Global Reservoir and Dam Database (Lehner et al., 2011a), we found that some of the Scottish reservoirs in GR~~r~~anD were not included in the UK Reservoir Inventory, and in several locations the capacities were significantly different. Consequently, in Scotland, the UK reservoir inventory has been supplemented with data from the Scottish Environment Protection Agency (SEPA). This provided an additional 4 water supply reservoirs and where mismatches in capacities were identified, the UK Reservoir Inventory has been updated using the supplementary SEPA data.

In total 207 reservoirs from the UK reservoir inventory are classified as water supply. We excluded 47 of the UK Reservoir Inventory water supply reservoirs from this analysis, either because there was no gauge downstream of the reservoir and thus results could not be evaluated (11), they were outside of Great Britain (2) or because they could not be placed on the river network (34), which was usually because the reservoir appeared to be disconnected from the river channel.

3.5 Simulations in reservoir catchments

Simulations in reservoir catchments are carried out at all gauges located downstream of one or more water supply reservoirs. This is a total of 264 catchments (Figure 2b). In each catchment the model is run both with and without reservoir representation. The no-reservoir scenario runs ~~Each catchment was simulated~~~~Te~~10~~n~~ simulations ~~per reservoir catchment were run~~in each reservoir catchment using the top 10 natural transfer function parameter combinations (section 3.3). In the reservoir scenario, for each of the same 10 parameter combinations, we sample the reservoir parameters 500 times, resulting in ~~each catchment was instead simulated~~ 5000 simulations per catchment.~~s were run consisting of~~~~Each catchment was simulated 10 times in the no-reservoir scenario using the top 10 natural transfer function parameter combinations and 5000 times in the reservoir scenario sampling reservoir parameters (500 times for each of the top 10 natural transfer function parameter combinations).~~ The minimum capacity of each reservoir (Smin) is set to 10% of the maximum capacity (which is obtained from relevant databases). At the very start of a simulation, S_t is set to 90% of the reservoir's maximum capacity (since simulations begin in winter when reservoirs are usually full).

3.5.1 Defining reservoir transfer functions

In order to define the reservoir transfer functions, we used a small sample of catchments for which compensation flow data and abstractions estimates are available to determine which catchment and reservoir attributes (e.g. catchment area, rainfall, reservoir capacity) exhibit the strongest relationships with the fluxes (see results section 4.2 for the chosen attributes). The small sample consists of 9 catchments with compensation flow data and 16 with abstraction estimates. Although data for these fluxes are not available on a large-scale, in some cases compensation flow is recorded in Water Resource Management Plans and Drought Plans, and where there is a suitable downstream gauge, hydrological signatures can be used to infer abstraction volume (by looking at changes to the water balance) and compensation flow (from plateaus in the flow duration curve) (see Section 1 in the Supplementary Material). In this study abstraction and compensation flow remain constant throughout the simulation (i.e. the same volume is released or abstracted at every timestep), but where appropriate they could be varied throughout the year.

Since they are based on only a limited number of observed data points, the transfer functions and the data they use contain significant uncertainty. To account for this, we define upper and lower bounds for each transfer function parameter (p) and sample within the chosen parameter space. The upper and lower bounds are determined after assessing the relationship between the non-parametric KGE and the two parameters (ABS and CF) in a selection of catchments (see Section 2 in the Supplementary Material). By using transfer functions which account for local information, we avoid sampling unrealistic parameter space and can begin to understand how these fluxes might be estimated without calibration.

3.5.2 Calibration of reservoir parameters (catchment-by-catchment and nationally-consistent calibration)

After running 5000 simulations in each catchment, we consider two types of calibration. The first is a catchment-by-catchment calibration for which we identify the best performing simulation and set of reservoir parameters in each catchment (this could leave us with a different optimal reservoir transfer function parameter combination in each catchment). The second looks for the best nationally-consistent calibration, where each catchment uses the same set of reservoir transfer function parameters. The best nationally-consistent simulation is chosen by first calculating the difference in non-parametric KGE between each of the 5000 reservoir simulations and the best no-reservoir simulation, and then identifying the median difference in KGE for each of the 5000 simulations across all of the catchments with a contributing area of more than 25% (i.e. more than 25% of the catchment is drained through a reservoir). We chose to use only gauges draining a high proportion of the catchment since these are the most impacted by the reservoir representation but note that the results are very similar if we include more or fewer gauges in this sample (see Section 5 in the Supplementary Material). The simulation with the highest median KGE difference ~~(this was +0.10)~~ was then chosen as the best nationally-consistent calibration.

3.6 Model evaluation

To evaluate the flow simulations in the both the benchmark and reservoir catchments, we use the non-parametric KGE (Pool et al., 2018; Gupta et al., 2009). The non-parametric KGE metric is comprised of three diagnostically meaningful components considering the errors in mean flow, flow variability and the correlation between observed and simulated flow. The non-parametric KGE uses the flow duration curve to investigate flow variability instead of the standard deviation, and the Spearman rank correlation instead of the Pearson correlation coefficient. Since previous work (Ferrazzi and Botter, 2019; Salwey et al., 2023) has shown that reservoirs can have a significant impact on the flow duration curve and water balance, we considered this a suitable metric to investigate the flow components in both benchmark and reservoir catchments. We also calculated the normalised Mean Absolute Error (nMAE) to complement the Spearmans rank correlation, which was more informative in catchments with little variability in river flows (see section 4.3).

Flow timeseries are only evaluated at gauges where there is more than 20 years of observed data between 1975 (or the reservoir construction date if this is later) and 2020. After these criteria have been enforced, we are able to evaluate the model in 205 out of 264 catchments. The model is evaluated for both the catchment-by-catchment calibration and the nationally-consistent calibration ~~for~~ by comparing simulations with~~-~~ –and –without reservoir representation.

For completeness, in a few selected catchments we also compared our operating rules to the widely used non-irrigation reservoir rule introduced by Hanasaki et al. (2006). However, since the Hanasaki rule assumes that no abstractions are taken directly from the reservoir, these rules are not well suited to water supply reservoirs in our domain (see section 4.3 and section 8 in the Supplementary Material). We could not compare our operating rules to any of the data-driven approaches in the literature (e.g. Turner et al. (2020)) since their high data requirements could not be fulfilled at the national scale in GB. Comparing the simulations which use our new operating rules to the simulations of the pre-existing hydrological model without reservoir representation thus remains the most feasible and relevant way to evaluate the new proposed model.

4 Results

4.1 Calibration in near-natural catchments

~~The results~~Model performance ~~of~~~~from the~~from simulations in near-natural catchments ~~simulations for~~ with the ~~best~~ top (highest median non-parametric KGE across 137 catchments) nationally-consistent ~~set of transfer function parameters~~calibration are displayed in Figure 3~~2~~. When considering the top 10 ~~transfer function parameter sets~~near-natural simulations across all 137 catchments, the median KGE score ranges from 0.83-0.84. While the model generally captures the mean flow, flow variability and correlation well, there are ~~a few~~ some catchments ~~from the 137~~ which have poor performance. For example, the Aldbourne at Ramsbury (39101) and the Ewelme at Ewelme Brook (39065) (which are both chalk catchments) have non-parametric KGE scores of -0.69 and -0.11 in the best performing simulation respectively. In general, ~~T~~the poorer performing catchments are largely chalk catchments, ~~where~~ since here the model is not able to capture flow losses from inter-catchment groundwater flows, which has also been noted in previous studies (Coxon et al., 2019; Lane et al., 2021; Lane et al., 2019). While this is an area of model improvement for future studies (see for example Oldham et al. (2023)), it is less significant for this study as reservoirs are typically not constructed in groundwater dominated catchments in GB.

~~The top 10 transfer function parameter combinations (or simulations, decided based on those with the highest median non-parametric KGE across all 137 catchments) are carried forward to be used in simulations in reservoir catchments.~~

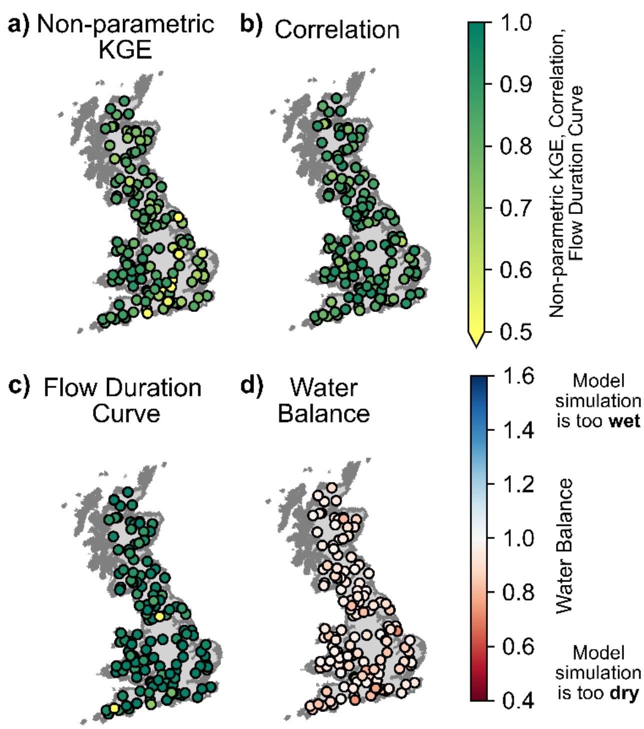
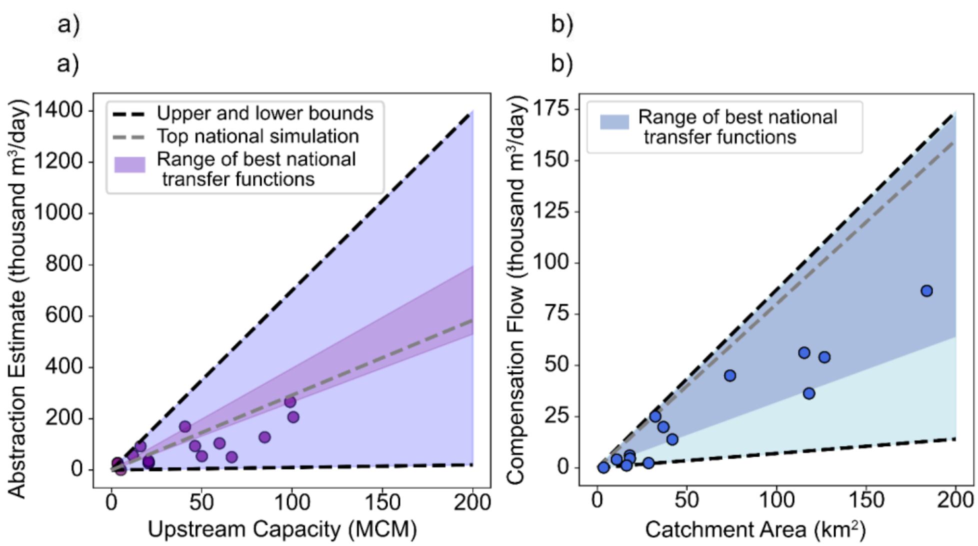


Figure ~~2~~3. (a) Non-Parametric KGE and its components **(b, c, d)** for the transfer function parameter combination with the highest median KGE (0.84) across 137 near-natural catchments.



4.2 Reservoir ~~T~~transfer function definition

Figure ~~3~~4. Relationship between (a) reservoir capacity upstream of a gauge and the reservoir abstraction volume (ABS) and (b) catchment area upstream of a gauge and compensation flow (CF). Dots represent data from a sample of catchments where abstractions could be estimated using a water balance hydrological signature and compensation flows could be extracted from Drought Plans, WRMPs and observed downstream flow duration curves (see Section 3.5.1 and Section 1 in Supplementary Material). Grey dashed lines represent the linear transfer functions associated with the ~~best~~ top performing simulation from the nationally-consistent calibration, the darker shading represents the spread of the top 5% of nationally-consistent simulations and black dashed lines represent the limits of the transfer function parameters based on the sensitivity of model performance to these parameters (see Table S1 and Section 2 in the Supplementary Material).

We tested a number of catchment/reservoir attributes to define the reservoir transfer functions used in this study (see Section 1 in the Supplementary Material) relying on data from a small sample of catchments (see Section 3.5.1). We found that catchment area was the most appropriate attribute to identify the compensation flow (CF), and the upstream reservoir capacity was best for identifying the abstraction volume (ABS) (see Figure ~~3~~ 4 and Equations 6 and 7 below). Since the observations (Figure 4) do not show any evidence of non-linearity, we chose to use a linear (and hence more parsimonious) relationship for both transfer functions: ~~We also found that a linear relationship was adequate for both transfer functions:~~

$$ABS = ResCapacity * p1 \tag{6}$$

$$CF = CatchArea * p2 \tag{7}$$

The top nationally-consistent calibration associated with the ABS parameter (marked on Figure 4a with a grey dashed line) generates parameters which are similar to those observed in the literature. However, the top nationally-consistent transfer function selected for estimating the CF parameter (marked on Figure 4b with a grey dashed line) lies close to the upper end of the sampling limits (Table 1) and does not match the observations. To investigate the sensitivity of the model to each of the reservoir parameters (CF and ABS) Figure 4 also shows the variability in the transfer functions associated with the top 5% of nationally-consistent simulations (this is displayed on Figure 4 with darker shading). The top 5% of simulations are those with the highest average non-parametric KGE (calculated across the full sample of reservoir catchments). This shows that the model's predictive performance is more sensitive to ABS (p1) than CF (p2). The regional differences in the transfer function parameters selected by the catchment-by-catchment calibration can be seen in section 9 of the supplementary material. ~~The best nationally-consistent transfer function associated with ABS (marked on Figure 3a with a grey dashed line) is in line with the observed data. However, the nationally-consistent transfer function used to estimate CF (marked on Figure 3b with a grey dashed line) lies close to the upper end of the sampling limits. The range of variability of the transfer functions associated with the top 5% of national-consistent simulations (displayed on Figure 3 by darker shadings) also shows that transfer function parameter ABS (p1) is much more identifiable than that of CF (p2), or in other words, the model's predictive performance is much more sensitive to ABS than CF.~~

Table 1: Range of variability of the transfer function parameters (p) for use across GB. Upper and lower bounds have been determined to prevent parameter values from becoming unrealistic, whilst being as wide as possible to enable the feasible parameter space to be fully sampled.

Transfer function parameter	Lower Bound	Upper Bound
p_1	0.0001	0.007
p_2	0.07	0.87

4.3 Model evaluation (reservoir catchments)

After running DECIPHeR both with and without reservoir representation across GB, we produced 5000 flow simulations *with* reservoir representation and 10 simulations *without* reservoir representation in 205 reservoir catchments.

The following results have been split into two sections. The first (4.3.1) presents the results from a catchment-by-catchment calibration, identifying the optimum set of transfer function parameters in each catchment (considering there are 2 calibrated transfer function parameters and 205 catchments this approach identifies 410 parameters). The second section (4.3.2) presents the results from a nationally-consistent ~~parameterisation~~ calibration (a total of 2 transfer function parameters assuming the same relationship between catchment and reservoir attributes in every catchment).

4.3.1 ~~Best~~ Top individual simulations (catchment-by-catchment calibration)

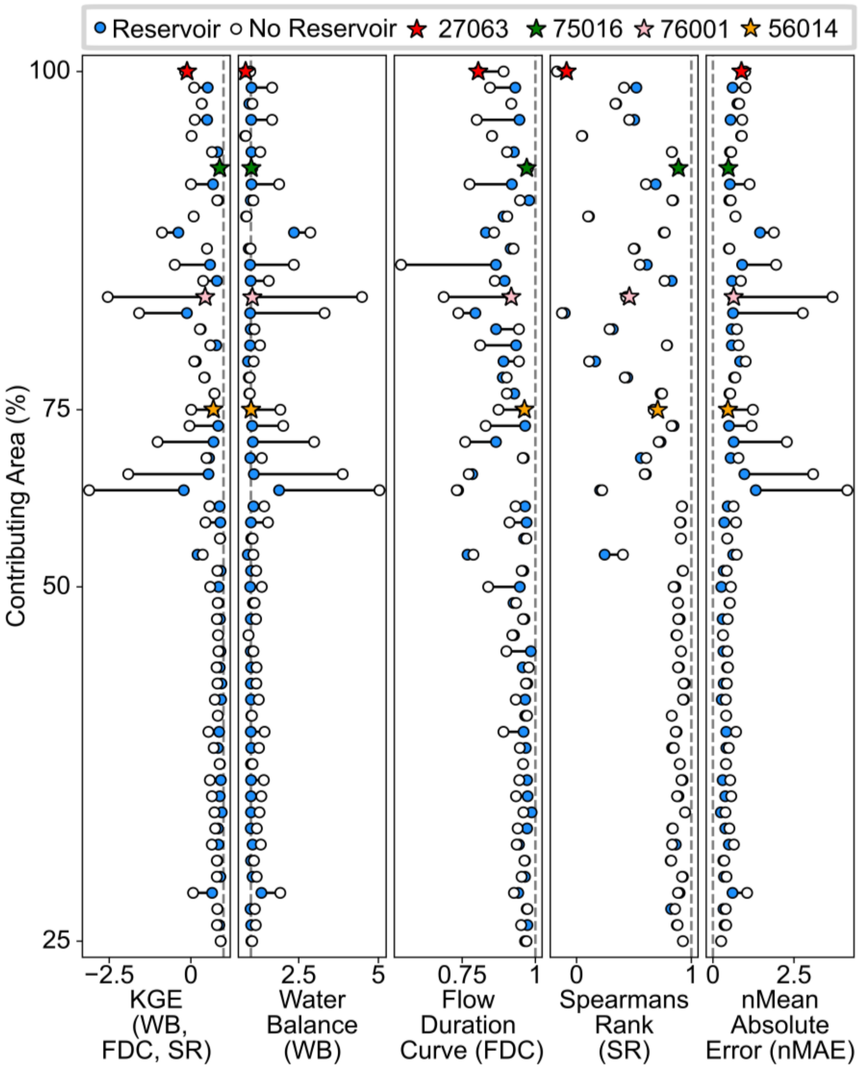


Figure ~~4~~5. Difference in performance between the top reservoir simulation in each catchment and the top no-reservoir simulation. Results are presented for the non-parametric KGE metric and its relative components as well as the normalised Mean Absolute Error (nMAE). Catchments are ordered based on their contributing area (proportion of the catchment that is drained through a reservoir). Grey dashed lines represent the optimum value for each metric, points falling closest to these lines have the best performance. Four catchments are highlighted using star markers and investigated in more detail in section 4.4.

Figure ~~4~~ 5 shows the maximum improvement in non-parametric KGE and its respective components for the ~~best~~ top performing simulation at all gauges downstream of a reservoir with a contributing area higher than 25% (for full results see Section 3 in the Supplementary Material). Figure 5 also highlights four catchments with star markers which are designed to demonstrate where the operating rules are working well and where improvements are needed (see Figure 2 for the location of these catchments across GB). Catchments 76001 and 56014 (pink and yellow stars) show large improvements in the KGE where the operating rules are working well. Comparatively, changes in the KGE at catchments 27063 and 75016 are minimal. This is discussed in more detail in section 4.4. The plot also displays an alternative to the Spearman's Rank correlation metric, the Normalised Mean Absolute Error. We find that for gauges with a contributing area below 25% (of which there are 157, not displayed in Figure ~~4~~5), only 9 (or 5%) of the gauges show a non-parametric KGE improvement of more than 0.1. Only one gauge shows a decrease of more than 0.1 which suggests that the reservoir representation is not worsening model performance in catchments where reservoirs have a minimal impact. Since such a small percentage of each of these catchments is controlled by (or drained through) a reservoir, we do not expect reservoir representation to make a large difference here and exclude them from the analysis and plots below.

Of the 55 gauges with a contributing area higher than 25%, 51 have a higher non-parametric KGE when the model includes reservoir representation compared to a model without it. 28 (of 55) gauges have a non-parametric KGE increase of more than 0.1, 18 have a non-parametric KGE increase of more than 0.3, 11 have a non-parametric KGE increase of more than 0.5 and 6 have a non-parametric KGE increase of more than 1. The median change in KGE is + 0.11 and the mean is + 0.38. The largest improvement in KGE is 2.99 which is seen at the Haweswater Beck at Burnbanks (76001) (denoted by a pink star on Figures 2 and 5-8) where the metric increases from -2.55 to 0.44, largely driven by the water balance component which decreases from 4.49 to 1.04. The largest decrease in KGE is -0.16 at the St Neot at Craigshill Wood (48009) which is largely driven by a decrease in the correlation component of 0.16. The median KGE across all gauges with a contributing area exceeding 25% rises to 0.82 from 0.58 after the inclusion of reservoir representation. When you consider gauges with a contributing area higher than 50% and 75% respectively, the median KGE is slightly lower but sees a larger improvement, rising to 0.55 from 0.20 pre-reservoir representation and 0.5 from 0.11 pre-reservoir representation. All gauges with a KGE improvement of more than 0.6 have a contributing area exceeding 65%.

In general, the largest improvements in KGE tend to come from the water balance and flow duration components of the metric and the smallest from the correlation. This component appears to be very insensitive to the inclusion of reservoir representation. We find that where compensation flow dominates a hydrograph Spearman's rank cannot appropriately rank so many similar data points, and these flow plateaus contain very similar data points with large differences in ranks (see Section 6 in the Supplementary Material and the discussion in Section 5.3). As a result, we calculated several other correlation-based metrics. Of these we chose the normalised mean absolute error (nMAE) to be displayed in the results section. Compared to the RMSE or Pearson's correlation this metric does not put as much emphasis on the high flows (which in many reservoir catchments do not dominate much of the flow regime) and unlike the Spearman's rank, this metric can process many data points of a similar value, suitably evaluating the ability of a model to recreate the compensation flow. Reductions in the nMAE appear to be correlated with reductions in the water balance.

4.3.2 ~~Best~~ Top overall simulation (nationally-consistent calibration)

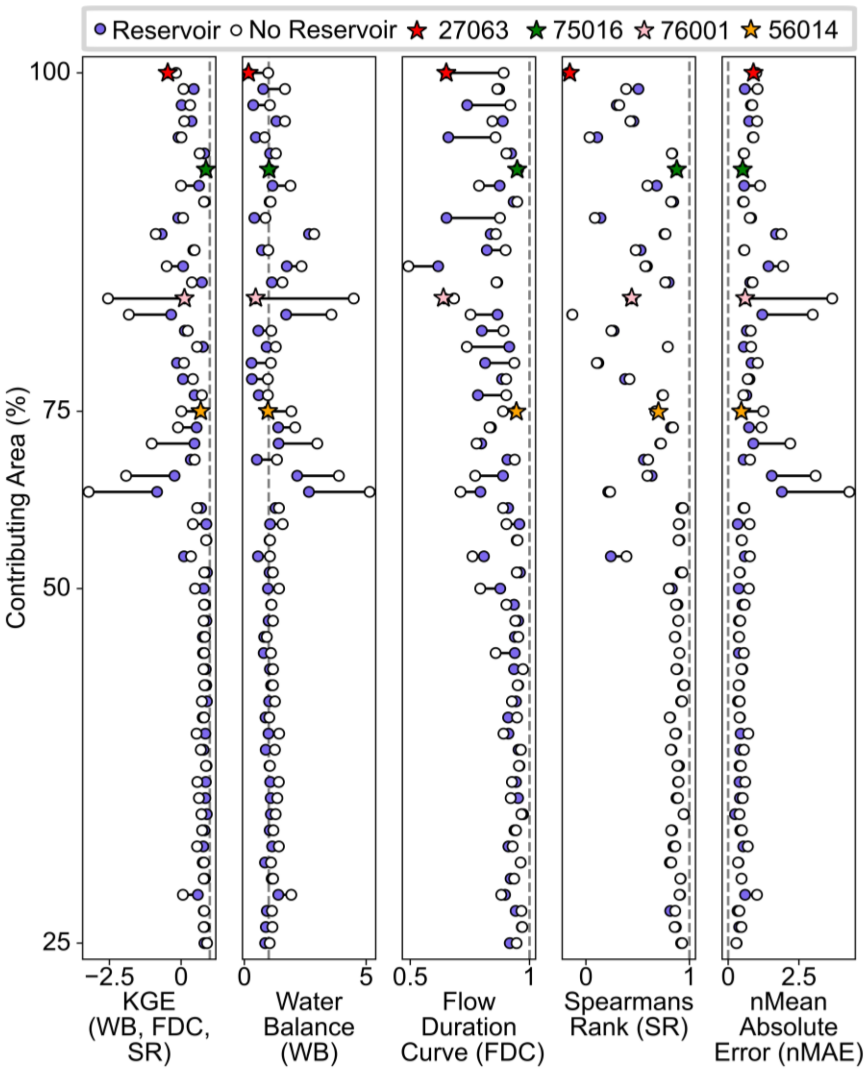


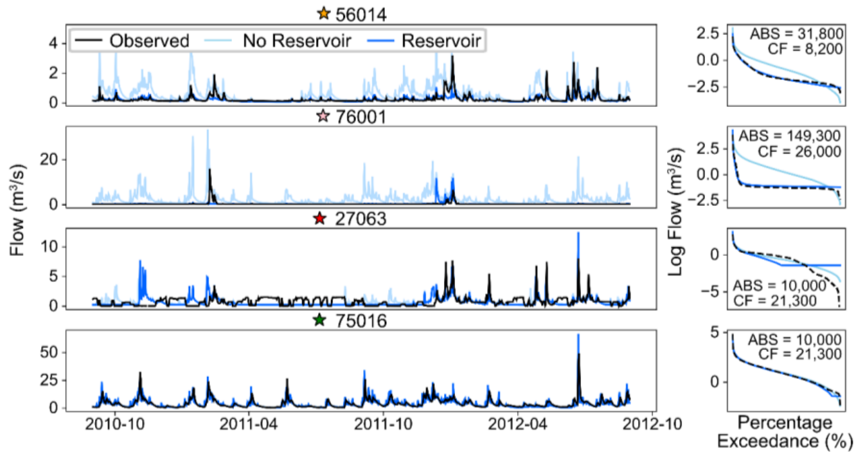
Figure ~~5~~6. Difference in performance between the nationally ~~best~~ top reservoir simulation and the nationally ~~best~~ top no-reservoir simulation. Results are presented for the non-parametric KGE metric and its relative components as well as the normalised Mean Absolute Error (nMAE). Catchments are ordered based on their contributing area (or proportion of the catchment that is drained through a reservoir). Grey dashed lines represent the optimum value for each metric, points falling closest to these lines have the best performance. Four catchments are highlighted using star markers and investigated in more detail in section 4.4.

Figure ~~5~~ 6 shows the improvement in non-parametric KGE and its respective components for a nationally-consistent ~~parameterisation~~ calibration at all gauges downstream of a reservoir with a contributing area higher than 25% (for full results see Section 4 in the Supplementary Material). Of the 55 gauges with a contributing area higher than 25%, 38 have a higher non-parametric KGE when the model includes reservoir representation compared to a model without it. 27 (of 55) gauges have a non-parametric KGE increase of more than 0.1, 12 have a non-parametric KGE increase of more than 0.3 and 9 have a non-parametric KGE increase of more than 0.5. The largest improvement in KGE is 2.78 which is seen at gauge 76001 (denoted by a pink star on Figures 2 and 5-8) where the metric increases from -2.55 to 0.23. The largest decrease in KGE is -0.35 at gauge 54081. The median KGE across all gauges with a contributing area exceeding 25% is 0.73, an increase from 0.56 without reservoir representation. When you consider gauges with a contributing area higher than 50% and 75% the median KGE with reservoir representation drops to 0.37 (increasing from 0.17 pre-reservoir representation) and 0.29 (increasing from 0.10 pre-reservoir representation).

When using a nationally-consistent ~~parameterisation~~ calibration there are 17 catchments where model performance decreases after including reservoir representation. Of these, 8 have a decrease in KGE exceeding 0.1. In general, these are catchments where the model with no reservoir representation captures the water balance well, but when the reservoir representation forces an abstraction, this component of the KGE significantly decreases, and usually decreases the flow duration curve metric too. These catchments appear to function differently from the rest of the sample, where abstractions are not taken directly from the reservoir and releases are controlled by a different set of rules. Overall, the correlation component of the KGE shows very minimal change between reservoir and no-reservoir simulations (which is contrary to visual changes in the correlation of hydrographs).

4.4 Example reservoir catchment simulations

4.4.1 ~~Best~~ Top individual simulation (catchment-by-catchment)



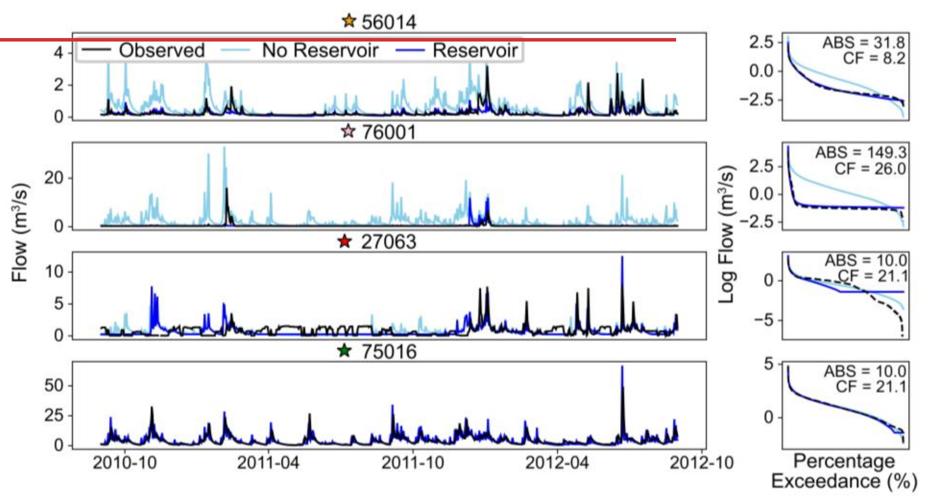


Figure ~~6~~7. Hydrographs and flow duration curves from the best individual simulations (catchment-by-catchment) for selected reservoir catchments. CF and ABS are recorded on each catchments flow duration curve in m^3 ~~Ml~~/day.

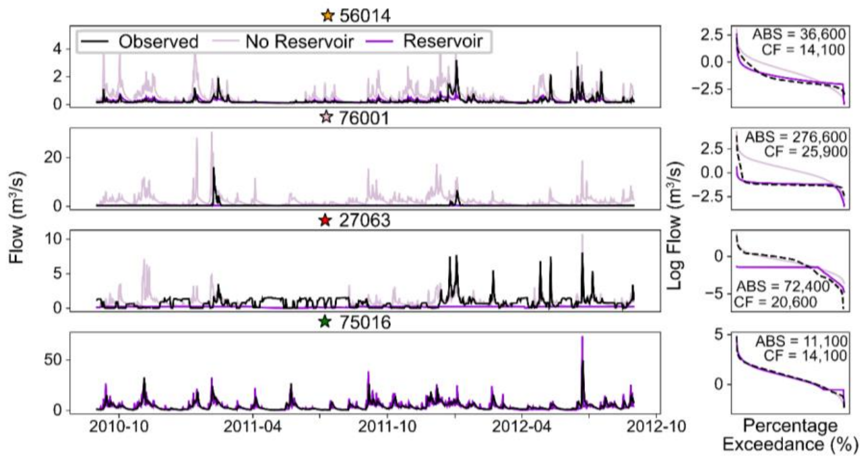
Formatted: Normal (Web), Line spacing: 1.5 lines

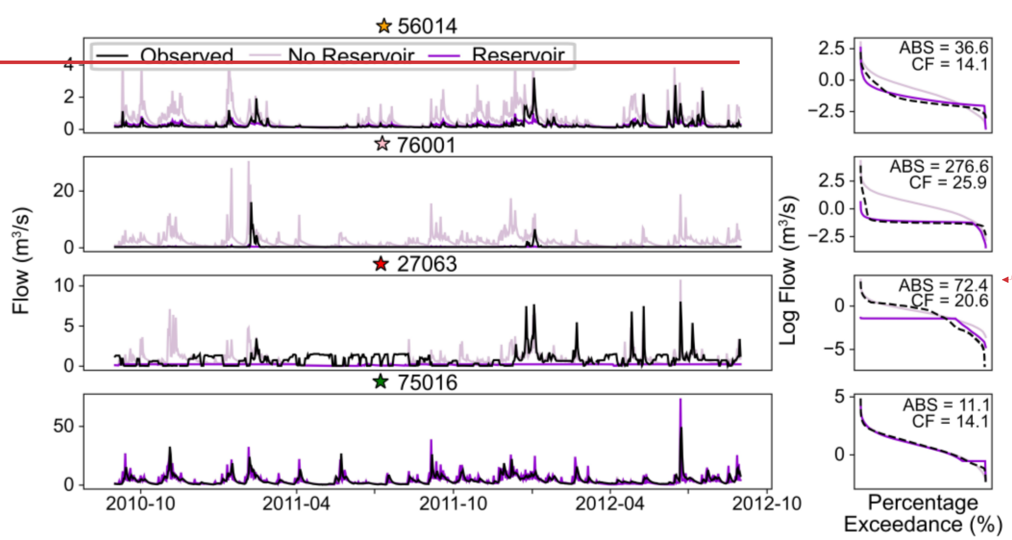
Formatted: Font: 10 pt

Simulation results for ~~T~~ the Usk at Usk Reservoir (56014) (yellow star) and the Haweswater Beck at Burnbanks (76001) (pink star) in Figure 7 demonstrate some of the central improvements made by the new reservoir operating rules. Peaks seen in the no-reservoir model (without reservoir representation) are not seen (or are decreased) in the model with reservoir representation (where reservoirs are absorbing peaks in inflow by increasing storage) allowing for compensation flow to dominate the flow duration curve and hydrograph. Both gauges 56014 and 76001 see large improvements in the KGE (0.01 to 0.69 and -2.55 to 0.44 respectively) which are largely facilitated by the improvements in the water balance and FDC components. The correlation (spearman's rank) component of the metric has only a very small increase at both locations (56014 sees an increase of 0.67 to 0.69 and 76001 from 0.43 to 0.45) despite visually having a much more representative hydrograph. This highlights some of the problems with calculating spearman's rank on data with little variability. The storage simulations in these two catchments follow a broadly yearly pattern of drawdown and refill. By comparing storage timeseries simulated at Haweswater (the reservoir upstream of 76001) to local level data (from the Hydrology Data Explorer; https://environment.data.gov.uk/hydrology/explore) we can see that the broad patterns in the simulated storage match the observed data well (see Section 7 in the Supplementary Material).

Unlike the first two examples, the newly included reservoir representation does not substantially improve the KGE at the Dibb at Grimwith Reservoir (27063) (red star) (KGE increases from -0.18 to -0.11). The reservoir located in this catchment plays a central role in regulating downstream flow which is not anticipated by our simple rules. The routine releases can be seen in the observed hydrograph, but since these play a different role to the compensation flow and are instead pulses of water intended to maintain downstream flow, they are not recreated by our simple rules. The ABS parameter here is very low to account for the fact that there are no abstractions but even this small abstraction decreases the water balance component of the non-parametric KGE from 0.98 to 0.83. Finally, the Cocker at Scalehill (75016) (green star) provides an example of a location where the reservoir outflow is generally unregulated. The reservoir in this catchment (Crummock Water) is very small and is full for most of the simulation (see Section 7 in the Supplementary Material), meaning the outflow is largely unimpacted and thus can be well recreated by both the simulation without the reservoir and the simulation with the reservoir.

4.~~2~~4.2 ~~Best~~ Top national simulation (nationally-consistent)





Formatted: Normal (Web), Line spacing: 1.5 lines

Figure ~~7~~8. Hydrographs and flow duration curves from best median simulation (nationally-consistent calibration) for selected reservoir catchments. CF and ABS are recorded on each catchments flow duration curve in m^3~~Ml~~/day.

Formatted: Font: 10 pt

Results with the nationally-consistent calibration (Figure 8) show~~s~~ similar ~~Many of the same distinctions can be made~~ differences between simulations with and without reservoirs in catchments 56014 (yellow star) and 76001 (pink star) ~~between the nationally-consistent and catchment-by-catchment calibration.~~ Peaks in the no-reservoir simulations are absorbed by the reservoirs and compensation flow dominates much of the hydrograph. At gauge 76001, the ABS parameter has increased from 149.3 in the catchment-by-catchment simulation to 247.7 with a nationally-consistent calibration. This abstraction is likely to be much higher than reality and explains the decrease in the water balance to 0.44 (compared to 1.04 in the catchment-by-catchment results). However, this still brings the metric much closer to 1 than the no-reservoir simulation which achieves a value of 4.49. Both of these catchments are relatively insensitive to changes in the CF parameter. The increase in the ABS parameter in both of these catchments is reflected in the simulation of reservoir storage (see Section 7 in the Supplementary Material for reservoir storage simulations). These reservoirs (particularly Haweswater reservoir upstream of gauge 76001) are much more consistently drawn down in the nationally-consistent simulations.

A similar over abstraction is also seen in catchment 27063 (red star), where the nationally-consistent ~~parameterisation~~ calibration enforces a daily abstraction of 64~~.~~8,000 m^3~~Ml~~/day, meaning that since the reservoir is never full (and never spills), the compensation flow dominates the hydrograph. Here enforcing the nationally-consistent transfer function parameters reduces the non-parametric KGE from -0.18 to -0.42. Most of the performance loss here comes from the water balance component, followed by the flow duration curve. Finally, the model performance remains very constant at gauge 75016 (green star). This is because despite the enforced abstraction, the reservoir still remains full for the majority of the simulation. In this catchment, the KGE remains at 0.86 across the reservoir and no-reservoir simulations from the best catchment-by-catchment and nationally-consistent calibrations, and ~~. It~~ is ~~very~~ insensitive to the reservoir parameters.

Finally, Section 8 of the Supplementary Material reports a comparison of the top~~/best~~ nationally-consistent simulation with the widely used Hanasaki rule (Hanasaki et al. 2006) in this selection of catchments. Although the Hanasaki rule has no calibrated parameters and is therefore arguably simpler than ours, we found that it delivers a much poorer performance. This is largely because the Hanasaki rule does not allow for abstraction from the reservoirs which is a key component of reservoir operation in most GB reservoirs.

5 Discussion

5.1 Can we improve model performance with simple operating rules?

After integrating a set of simple operating rules into a national-scale hydrological model, we found that large gains in model performance are possible with only two additional calibrated parameters. The best results were produced when these parameters were calibrated at each downstream gauge, but amongst reservoirs with a single purpose (water supply), a nationally-consistent ~~parameterisation~~ calibration can also make significant improvements.

The improvements we have achieved in simulating streamflow with reservoir representation are similar to others seen in the literature (Turner et al., 2020; Yassin et al., 2019; Coerver et al., 2018). However, what makes this study unique is twofold: firstly, the simplicity of our operating rules. Many of the alternative sets of calibrated reservoir operating rules introduced in the literature have far more calibrated parameters than the two we have introduced here (e.g. Yassin et al. (2019) recommend 6 parameters that are determined for every month of the year leaving 72 total parameters, Turner et al. (2021) introduce a data-driven scheme with 19 parameters). Most of these rules are, to some extent, attempting to recreate specific operating policies or rule curves and by extension introduce significant additional complexity compared to our flux-based approach. A second key advantage of our operating rules is their minimal data requirement. It is not uncommon for a set of reservoir operating rules to require storage, inflow and release data to calibrate its parameters, which is rarely available over large scales (Yassin et al., 2019; Turner et al., 2020; Ehsani et al., 2017). Contrastingly, our rules (which only require a reservoir's location, capacity and catchment area) ought to be more transferable to large-scale modelling, particularly in regions where inflow, outflow and storage timeseries are unavailable (such as GB).

To our knowledge, this is also the first time water supply reservoirs have been the focus of a large-scale study. Unlike hydropower (Abeshu et al., 2023) or irrigation (Hanasaki et al., 2006) reservoirs, water supply reservoirs are rarely the focus of large-scale studies, despite the fact that 22% of reservoirs globally (according to the GR~~r~~anD database) play a role in water supply. Instead, in many uncalibrated models, this type of reservoir is often collated into one 'non-irrigation' category (Hanasaki et al., 2006; Wisser et al., 2010). In this case reservoir rules usually aim to (where possible) release mean flow at all times of the year or reduce intra-annual variability. Since these rules facilitate no abstractions or compensation flow requirement, we consider them unsuitable for most of the reservoirs in our sample. Although we have only tested our approach at water supply reservoirs, a similar set of transfer functions and simple rules could be designed to suit reservoirs of other purposes. While we do not expect our rules to outperform more complex approaches, our rules provide a simple and practical starting point as a benchmark for incorporating reservoir representation into hydrological modelling where, due to data limitations, none of the pre-existing approaches could be applied.

5. 2 Can we identify a nationally-consistent ~~parameterisation~~calibration?

Overall, a nationally-consistent ~~parameterisation~~ calibration across most of the reservoirs in our sample worked well where 49% of gauges (with a contributing area higher than 25%) saw the non-parametric KGE increase by more than 0.1 after the inclusion of the nationally-consistent operating rules. This is promising given this approach uses only 2 parameters (compared with 410 in the catchment-by-catchment approach) and open source catchment and reservoir attributes, thus reducing computational requirements (where a model no longer needs to be calibrated in every catchment) and facilitating the application of our operating rules to ungauged basins, or to reservoirs located in countries with less data available for calibration. We find that within our sample of reservoirs, catchment area and reservoir capacity are reasonable predictors of the compensation flow and abstraction volume across most water supply reservoirs.

There are very few examples of calibrated operating rules which undertake a similar nationally-consistent ~~parameterisation~~calibration. Yassin et al. (2019) introduce rules which may be applied in a similar, nationally-consistent manner, but the parameters are extracted from inflow, storage and release data which are not available in GB (or many other locations). Turner et al. (2021) extrapolate their rules to data-scarce reservoirs, but the ~~parameterisation~~ calibration varies from location-to-location where rules are fitted to observed data. We suggest that our approach can act as an informative lower benchmark (Seibert et al, 2018) to compare to more complex approaches that involve more detailed calibration, more parameters or higher data requirements.

However, while the nationally-consistent calibration worked well for many of the reservoirs in our sample, there were some catchments where a nationally-consistent calibration did not work well, particularly those which contained reservoirs fulfilling multiple purposes or regulating downstream flow. Although we included only reservoirs classified as water supply reservoirs (from the UK Reservoir Inventory) in our sample, in practise some of these reservoirs fulfil multiple objectives (e.g. Cow Green reservoir plays a role in flood management, and Kielder water is used for hydropower). Furthermore, ~7 of the reservoir catchments in our sample contained upstream reservoirs which play a different role in the water supply system than the rest of the sample. In these locations reservoirs focus on facilitating downstream abstractions (rather than those taken directly from the reservoir). It is no surprise that our rules do not work well here, where they are likely to miss some crucial coordination with the downstream river and misrepresent the purpose of the reservoir (Rougé et al., 2019). However, future work might consider defining new transfer functions to describe the operating rules at reservoirs in this sample (see section 5.4 for more detail).

5.3 Metrics to evaluate reservoir-impacted timeseries

Although much of the literature assessing reservoir operating rules evaluates their success with metrics such as RMSE (or nRMSE) (Turner et al., 2020), KGE (both parametric and non-parametric) (Yassin et al., 2019) and NSE (Voisin et al., 2013), we advise that this is interpreted and carried out with caution.

Standard metrics such as the non-parametric KGE worked well in our near-natural catchments, however, when used to evaluate reservoir-impacted hydrographs, their shape and distribution meant that the correlation component of the metric was not informative. The Spearman's Rank was not able to characterise correlations between two timeseries with low variability (i.e. when the compensation flow dominates the regime), which is often the case in reservoir-impacted timeseries. Although the Spearman's Rank was chosen for the non-parametric KGE over the Pearson correlation for its lower sensitivity to extreme values and focus on mean and low flows (Pool et al. 2018), in many reservoir-impacted catchments it was this portion of the hydrograph which the metric could not evaluate properly.

Although several other metrics were tested to look at the timing, or correlation, of our simulated flow, these were often very influenced by the high flows. Whilst matching the timing of these high flows is an important component of simulating reservoir-impacted flows, we were interested in where a reservoir absorbed a peak in inflow (releasing only the compensation flow) or spilled in broadly the right week/ month rather than the exact day. The Pearson correlation and RMSE put too much emphasis on the daily peaks, giving more weight to larger errors. Comparatively, the nMAE was less influenced by the peaks in flow and large errors, providing a better evaluation of a timeseries dominated by the compensation flow. We suggest that future studies should seek to develop new signatures which replace the correlation component of the KGE evaluation metric and can better capture behaviour in human-influenced catchments ~~We suggest that future studies should consider replacing the correlation component of the KGE evaluation metric, and instead focus on identifying signatures or tests which can better capture behaviour in human-influenced catchments~~ (Kiraz et al., 2023). Standard metrics like the KGE should be calculated on impacted timeseries with caution, where their ability to evaluate natural timeseries does not always translate.

5.4 Limitations and future work

A limitation of this study was our inability to capture reservoir operations at gauges where upstream reservoirs fulfil multiple purposes as well as facilitating water supply. Future work might investigate whether this second cluster of multi-purpose or river regulating reservoirs could be represented by a similar set of simple rules. By extension, national-scale inventories could benefit from sub-categories for reservoir purpose, including a multi-purpose category. Furthermore, although these rules have only been tested at water supply reservoirs in Great Britain, they may be useful for simulating reservoirs in other locations. Whilst operations will vary country-by-country, this simple approach could be used to design rules and transfer functions for application elsewhere. Where a nationally-consistent approach is not appropriate (perhaps due to multi-purpose reservoirs or more complex coordination), transfer functions could be useful in defining the parameter bounds for calibration and establishing relationships between reservoir and catchment attributes and model parameters.

6 Conclusions

This study presents a set of new, simple operating rules designed to simulate operations at water supply reservoirs across large scales. We demonstrate their application across GB, where national-scale hydrological modelling has not previously included reservoir representation. Our approach performs well across a large-sample of reservoirs, with the largest performance gains established from a catchment-by-catchment calibration. Although it performs less well, our nationally-consistent calibration should act as an informative lower benchmark for simulating operations at water resource reservoirs before more complex rules are considered. The results of this study should encourage the inclusion of reservoirs in national-scale hydrological modelling across GB, since we have identified large gains in performance with minimal data and added complexity.~~The results of this study should mandate the inclusion of reservoirs in large-scale hydrological modelling across GB, since we have identified large gains in performance with minimal data and added complexity.~~

7 Code and data availability

Formatted: Heading 1

The DECIPHeR model code is available at https://github.com/uob-hydrology/DECIPHeR. The UK Reservoir Inventory database (Durant and Counsell, 2018) and PET data (Robinson et al., 2023) are available from the CEH Environmental Data Centre (https://eidc.ac.uk/). Rainfall data (Hollis et al., 2019) is available from the CEDA archive (https://archive.ceda.ac.uk/) and flow timeseries are available from the NRFA (https://nrfa.ceh.ac.uk/). Flow outputs, parameter sets and performance metrics from the best performing model simulations (associated with both a catchment-by-catchment and nationally-consistent calibration) are available from the University of Bristol data repository, data.bris, at https://doi.org/10.5523/bris.3elcv1fhj0cxl2u45mmkb8y8op.

8 Supplement link

9 Author contributions

With guidance from GC and FP, SS was responsible for the development of the reservoir representation and implementation of the operating rules, model simulations and output analysis. SS wrote the initial manuscript with substantial contributions from GC and FP. RL helped with the model calibration as well as providing feedback and edits to this manuscript. CH, MS and HM helped guide the research design and JF developed the schema for integrating the reservoir rules into DECIPHeR. All co-authors edited and contributed to the manuscript.

10 Competing interests

At least one of the (co-)authors is a member of the editorial board of Hydrology and Earth System Sciences.

11 Special issue statement

12 Acknowledgements

We gratefully acknowledge the discussions with Yanchen Zheng about the development of DECIPHeR and integration of Multiscale Parameter Regionalisation. We would also like to thank Toby Dunne for his guidance on the implementation of reservoir representation within the DECIPHeR routing scheme and Jan Seibert for his useful discussions and advice.

13 Financial support

This work is funded by a NERC GW4+ Doctoral Training Partnership studentship from the Natural Environmental Research Council (NE/S007504/1) and Wessex Water Ltd. GC was supported by a UKRI Future Leaders Fellowship [MR/V022857/1]. Rosanna Lane was supported by the Natural Environment Research Council award number NE/X019063/1 as part of the Hydro-JULES programme delivering National Capability.

References

Abeshu, G. W., Tian, F., Wild, T., Zhao, M., Turner, S., Chowdhury, A., Vernon, C. R., Hu, H., Zhuang, Y., and Hejazi, M.: Enhancing the representation of water management in global hydrological models, Geoscientific Model Development Discussions, 1-41, 2023.

Adam, J. C., Haddeland, I., Su, F., and Lettenmaier, D. P.: Simulation of reservoir influences on annual and seasonal streamflow changes for the Lena, Yenisei, and Ob' rivers, J Geophys Res-Atmos, 112, 10.1029/2007jd008525, 2007.

Brown, C. M., Lund, J. R., Cai, X. M., Reed, P. M., Zagona, E. A., Ostfeld, A., Hall, J., Characklis, G. W., Yu, W., and Brekke, L.: The future of water resources systems analysis: Toward a scientific framework for sustainable water management, Water Resources Research, 51, 6110-6124, 10.1002/2015wr017114, 2015.

Carrillo, A. M. R. and Frei, C.: Water: A key resource in energy production, Energy Policy, 37, 4303-4312, 2009.

Coerver, H. M., Rutten, M. M., and van de Giesen, N. C.: Deduction of reservoir operating rules for application in global hydrological models, Hydrology & Earth System Sciences, 22, 2018.

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

Dang, T. D., Chowdhury, A. F. M. K., and Galelli, S.: On the representation of water reservoir storage and operations in large-scale hydrological models: implications on model parameterization and climate change impact assessments, Hydrology and Earth System Sciences, 24, 397-416, 10.5194/hess-24-397-2020, 2020.

Devitt, L.: Evaluation of a Hydrological Modelling Framework, DECIPHeR, for Use in Large and Data Scarce River Basins: Upper Niger Case Study, University of Bristol, 2019.

Dobson, B., Coxon, G., Freer, J., Gavin, H., Mortazavi-Naeini, M., and Hall, J. W.: The Spatial Dynamics of Droughts and Water Scarcity in England and Wales, Water Resources Research, 56, e2020WR027187, 10.1029/2020WR027187, 2020.

Döll, P., Fiedler, K., and Zhang, J.: Global-scale analysis of river flow alterations due to water withdrawals and reservoirs, Hydrology and Earth System Sciences, 13, 2413, 10.5194/hess-13-2413-2009, 2009.

Durant, M. and Counsell, C.: Inventory of reservoirs amounting to 90% of total UK storage, NERC Environmental Information Data Centre, Wallingford, 10.5285/f5a7d56c-cea0-4f00-b159-c3788a3b2b38, 2018.

Ehsani, N., Vörösmarty, C. J., Fekete, B. M., and Stakhiv, E. Z.: Reservoir operations under climate change: Storage capacity options to mitigate risk, Journal of Hydrology, 555, 435-446, 2017.

Fadhliani, Zulkafli, Z., Yusuf, B., and Nurhidayu, S.: Assessment of streamflow simulation for a tropical forested catchment using dynamic TOPMODEL—Dynamic fluxEs and ConnectIvity for Predictions of HydRology (DECIPHeR) Framework and Generalized Likelihood Uncertainty Estimation (GLUE), Water, 13, 317, 2021.

Ferrazzi, M. and Botter, G.: Contrasting signatures of distinct human water uses in regulated flow regimes, Environmental Research Communications, 1, 071003, 10.1088/2515-7620/ab3324, 2019.

Gaupp, F., Hall, J., and Dadson, S.: The role of storage capacity in coping with intra-and inter-annual water variability in large river basins, Environmental Research Letters, 10, 125001, 2015.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of hydrology, 377, 80-91, 2009.

Hanasaki, N., Kanae, S., and Oki, T.: A reservoir operation scheme for global river routing models, Journal of Hydrology, 327, 22-41, 10.1016/j.jhydrol.2005.11.011, 2006.

Harrigan, S., Hannaford, J., Muchan, K., and Marsh, T. J.: Designation and trend analysis of the updated UK Benchmark Network of river flow stations: the UKBN2 dataset, Hydrology Research, 49, 552-567, 10.2166/nh.2017.058, 2018.

Hollis, D., McCarthy, M., Kendon, M., Legg, T., and Simpson, I.: HadUK-Grid—A new UK dataset of gridded climate observations, Geoscience Data Journal, 6, 151-159, 2019.

Kiraz, M., Coxon, G., and Wagener, T.: A Signature-Based Hydrologic Efficiency Metric for Model Calibration and Evaluation in Gauged and Ungauged Catchments, Water Resources Research, 59, e2023WR035321, 2023.

Lane, R. and Kay, A.: Gridded simulations of available precipitation (rainfall+ snowmelt) for Great Britain, developed from observed data (1961–2018) and climate projections (1980–2080), UK CEH [data set], UK CEH [data set], 10, 2022.

Lane, R. A., Freer, J. E., Coxon, G., and Wagener, T.: Incorporating uncertainty into multiscale parameter regionalization to evaluate the performance of nationally consistent parameter fields for a hydrological model, Water Resources Research, 57, e2020WR028393, 2021.

Lane, R. A., Coxon, G., Freer, J. E., Wagener, T., Johnes, P. J., Bloomfield, J. P., Greene, S., Macleod, C. J., and Reaney, S. M.: Benchmarking the predictive capability of hydrological models for river flow and flood peak predictions across over 1000 catchments in Great Britain, Hydrology and Earth System Sciences, 23, 4011-4032, 2019.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., and Magome, J.: High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management, Frontiers in Ecology and the Environment, 9, 494-502, 2011a.

Lehner, B., Liermann, C. R., Revenga, C., Vorosmarty, C., Fekete, B., Crouzet, P., Doll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rodel, R., Sindorf, N., and Wisser, D.: High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management, Frontiers in Ecology and the Environment, 9, 494-502, 10.1890/100125, 2011b.

McMillan, H., Booker, D., and Cattoën, C.: Validation of a national hydrological model, Journal of Hydrology, 541, 800-815, 2016.

Mizukami, N., Clark, M. P., Newman, A. J., Wood, A. W., Gutmann, E. D., Nijssen, B., Rakovec, O., and Samaniego, L.: Towards seamless large-domain parameter estimation for hydrologic models, Water Resources Research, 53, 8020-8040, 2017.

Murgatroyd, A. and Hall, J. W.: The Resilience of Inter-basin Transfers to Severe Droughts With Changing Spatial Characteristics, Frontiers in Environmental Science, 8, 571647, 10.3389/fenvs.2020.571647, 2020.

Oldham, L. D., Freer, J., Coxon, G., Howden, N., Bloomfield, J. P., and Jackson, C.: Evidence-based requirements for perceptualising intercatchment groundwater flow in hydrological models, Hydrology and Earth System Sciences, 27, 761-781, 2023.

Pool, S., Vis, M., and Seibert, J.: Evaluating model performance: towards a non-parametric variant of the Kling-Gupta efficiency, Hydrological Sciences Journal, 63, 1941-1953, 2018.

Robinson, E., Kay, A., Brown, M., Chapman, R., Bell, V., and Blyth, E.: Potential evapotranspiration derived from the UK Climate Projections 2018 Regional Climate Model ensemble 1980-2080 (Hydro-PE UKCP18 RCM), 2021.

Robinson, E. L., Brown, M. J., Kay, A. L., Lane, R. A., Chapman, R., Bell, V. A., and Blyth, E. M.: Hydro-PE: gridded datasets of historical and future Penman–Monteith potential evaporation for the United Kingdom, Earth System Science Data, 15, 4433-4461, 2023.

Rougé, C., Reed, P. M., Grogan, D. S., Zuidema, S., Prusevich, A., Glidden, S., Lamontagne, J. R., and Lammers, R. B.: Coordination and Control: Limits in Standard Representations of Multi-Reservoir Operations in Hydrological Modeling, Hydrology and Earth System Sciences Discussions, 1-37, 10.5194/hess-25-1365-2021, 2019.

Rudd, A. C., Kay, A., and Bell, V.: National-scale analysis of future river flow and soil moisture droughts: potential changes in drought characteristics, Climatic Change, 156, 323-340, 2019.

Salwey, S., Coxon, G., Pianosi, F., Singer, M., and Hutton, C.: National-Scale Detection of Reservoir Impacts Through Hydrological Signatures, Water Resources Research, e2022WR033893, 2023.

Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resources Research, 46, 2010.

Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Zink, M., Wanders, N., Eisner, S., Müller Schmied, H., Sutanudjaja, E. H., and Warrach-Sagi, K.: Toward seamless hydrologic predictions across spatial scales, Hydrology and Earth System Sciences, 21, 4323-4346, 2017.

Sardo, M., Epifani, I., D'Odorico, P., Galli, N., and Rulli, M. C.: Exploring the water–food nexus reveals the interlinkages with urban human conflicts in Central America, Nature Water, 1, 348-358, 2023.

Shannon, S., Payne, A., Freer, J., Coxon, G., Kauzlaric, M., Kriegel, D., and Harrison, S.: A snow and glacier hydrological model for large catchments–case study for the Naryn River, central Asia, Hydrology and Earth System Sciences, 27, 453-480, 2023.

Smith, K. A., Barker, L. J., Tanguy, M., Parry, S., Harrigan, S., Legg, T. P., Prudhomme, C., and Hannaford, J.: A multi-objective ensemble approach to hydrological modelling in the UK: an application to historic drought reconstruction, Hydrology and Earth System Sciences, 23, 3247-3268, 2019.

Speckhann, G. A., Kreibich, H., and Merz, B.: Inventory of dams in Germany, Earth System Science Data, 13, 731-740, 2021.

Steyaert, J. C., Condon, L. E., Turner, S. W. D., and Voisin, N.: ResOpsUS, a dataset of historical reservoir operations in the contiguous United States, Scientific Data, 9, 1-8, 10.1038/s41597-022-01134-7, 2022.

Tebakari, T., Yoshitani, J., and Suvanpimol, P.: Impact of large-scale reservoir operation on flow regime in the Chao Phraya River basin, Thailand, Hydrological Processes, 26, 2411-2420, 10.1002/hyp.9345, 2012.

Tijdeman, E., Hannaford, J., and Stahl, K.: Human influences on streamflow drought characteristics in England and Wales, Hydrology and Earth System Sciences, 22, 1051-1064, 10.5194/hess-22-1051-2018, 2018.

Turner, S. W. D., Doering, K., and Voisin, N.: Data-Driven Reservoir Simulation in a Large-Scale Hydrological and Water Resource Model, Water Resources Research, 56, e2020WR027902, 10.1029/2020WR027902, 2020.

Turner, S. W. D., Steyaert, J. C., Condon, L., and Voisin, N.: Water storage and release policies for all large reservoirs of conterminous United States, Journal of Hydrology, 603, 126843, 10.1016/j.jhydrol.2021.126843, 2021.

USACE, U. S. A. C. o. E.: National Inventory of Dams [dataset], 2018.

Voisin, N., Li, H., Ward, D., Huang, M., Wigmosta, M., and Leung, L.: On an improved sub-regional water resources management representation for integration into earth system models, Hydrology and Earth System Sciences, 17, 3605-3622, 2013.

Vörösmarty, C. J., Meybeck, M., Fekete, B., Sharma, K., Green, P., and Syvitski, J. P.: Anthropogenic sediment retention: major global impact from registered river impoundments, Global and planetary change, 39, 169-190, 10.1016/S0921-8181(03)00023-7, 2003.

Wang, J., Walter, B. A., Yao, F., Song, C., Ding, M., Maroof, A. S., Zhu, J., Fan, C., McAlister, J. M., and Sikder, S.: GeoDAR: georeferenced global dams and reservoirs dataset for bridging attributes and geolocations, Earth System Science Data, 14, 1869-1899, 2022.

Watts, G., Battarbee, R. W., Bloomfield, J. P., Crossman, J., Daccache, A., Durance, I., Elliott, J. A., Garner, G., Hannaford, J., Hannah, D. M., Hess, T., Jackson, C. R., Kay, A. L., Kernan, M., Knox, J., Mackay, J., Monteith, D. T., Ormerod, S. J., Rance, J., Stuart, M. E., Wade, A. J., Wade, S. D., Weatherhead, K., Whitehead, P. G., and Wilby, R. L.: Climate change and water in the UK - past changes and future prospects, Progress in Physical Geography-Earth and Environment, 39, 6-28, 10.1177/0309133314542957, 2015.

Wendt, D. E., Bloomfield, J. P., Van Loon, A. F., Garcia, M., Heudorfer, B., Larsen, J., and Hannah, D. M.: Evaluating integrated water management strategies to inform hydrological drought mitigation, Natural Hazards and Earth System Sciences, 21, 3113-3139, 2021.

Wisser, D., Fekete, B. M., Vörösmarty, C., and Schumann, A.: Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network-Hydrology (GTN-H), Hydrology & Earth System Sciences, 14, 2010.

Yassin, F., Razavi, S., Elshamy, M., Davison, B., Sapriza-Azuri, G., and Wheater, H.: Representation and improved parameterization of reservoir operation in hydrological and land-surface models, Hydrology and Earth System Sciences, 23, 3735-3764, 10.5194/hess-23-3735-2019, 2019.

Zhao, G., Gao, H., Naz, B. S., Kao, S.-C., and Voisin, N.: Integrating a reservoir regulation scheme into a spatially distributed hydrological model, Advances in Water Resources, 98, 16-31, 2016.