Supplement of Hydrol. Earth Syst. Sci., 28, 4203–4218, 2024
https://doi.org/10.5194/hess-28-4203-2024-supplement

Hydrology and Earth System Sciences
Open Access
EGU

Supplement of

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

Saskia Salwey et al.

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

Supplementary material

S1. Transfer function definition

The data used to define the transfer functions for estimating CF and ABS has all been extracted from open access Water Resource Management Plans (WRMPs) and Drought Plans (e.g. United Utilities Final Drought Plan 2022 contains the compensation flows at 11 reservoirs) or, where a gauge is sufficiently close to the reservoir outflow, inferred from downstream timeseries. To infer CF from a downstream flow timeseries, plateaus can be identified in the low flow portion of a flow duration curve. In this study, the data for the CF parameter came from a mixture of sources (the downstream gauge record, WRMPs, Drought Plans), whereas the ABS estimates were all inferred from downstream gauges. To infer ABS from a downstream gauge precipitation and PET data can be used to calculate the water balance. For this method to be successful, a gauge must be sufficiently close to a reservoir outflow, and there is an underlying assumption that all deviations from a closed water balance can be attributed to reservoir-related abstractions. In some cases this method will be inferring the total abstractions across multiple reservoirs. For our sample of gauges this was done by using flow, precipitation and PET data from the National River Flow Archive, CEH_GEAR and CHESS-PE datasets. For more detail on inferring ABS or CF from a downstream flow timeseries see Salwey et al. (2023).

Our intention was for these transfer functions to be as simple as possible and therefore we started by testing linear relationships. These worked reasonably well so we did not test any other functions. We narrowed our search down by only using open access attributes. In some cases we tried combining these. The attributes tested in the transfer functions are listed in Tables S1 and S2 below. The Normalized Upstream Capacity (NUC) compares the capacity of upstream reservoirs to the average volume of precipitation received by the catchment in a year (this metric is defined further in Salwey et al. (2023)), the Average Precipitation comes from the CAMELS-GB dataset, Reservoir Capacity can be defined as the sum of total reservoir capacity found upstream of a gauge (data on reservoir capacity comes from the UK Reservoir Inventory and SEPA datasets) and Catchment Area comes from the National River Flow Archive. To choose the final function we compared the R^2 values which are presented in Tables S1 and S2 below.

Table S1. Information on the attributes tested in the compensation flow (CF) transfer function and their associated r^2 values and plot numbers.

Parameter	Attribute(s)	Plot	R^2
CF	Normalized Upstream Capacity (NUC)	A	0.08
CF	Average Precipitation	B	0.03
CF	Reservoir Capacity	C	0.35
CF	Reservoir Capacity / Area	D	0.03
CF	Catchment Area	E	0.88

Table S2. Information on the attributes tested in the abstraction volume (ABS) transfer function and their associated r^2 values and plot numbers.

Parameter	Attribute(s)	Plot	R^2
ABS	Normalized Upstream Capacity (NUC)	A	0.09
ABS	Average Precipitation	B	0.27
ABS	Reservoir Capacity	C	0.66
ABS	Reservoir Capacity / Area	D	0.16
ABS	Gauge Area	E	0.09

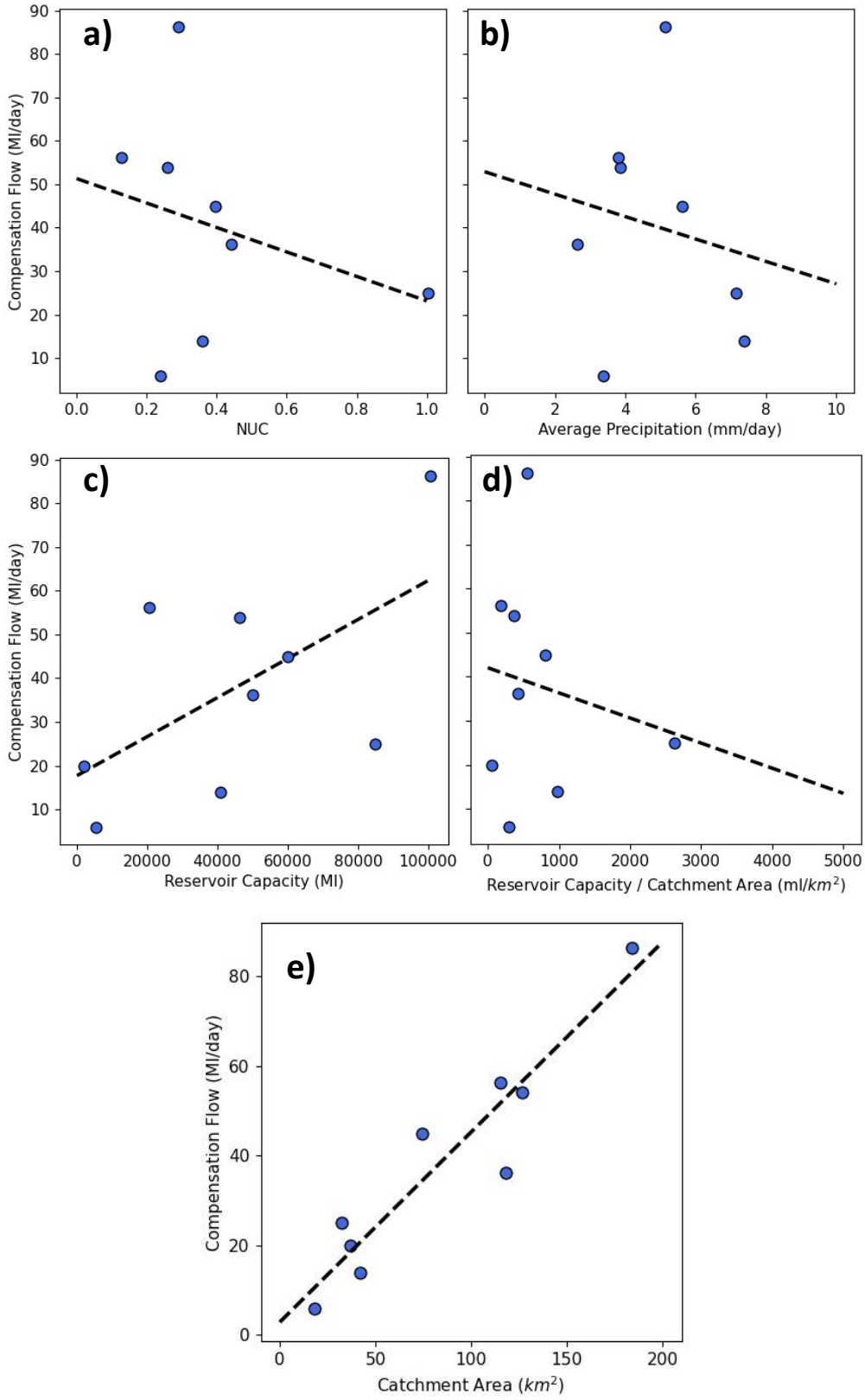

Figure S1. Relationships between a selection of catchment and reservoir attributes and the compensation flow volume.

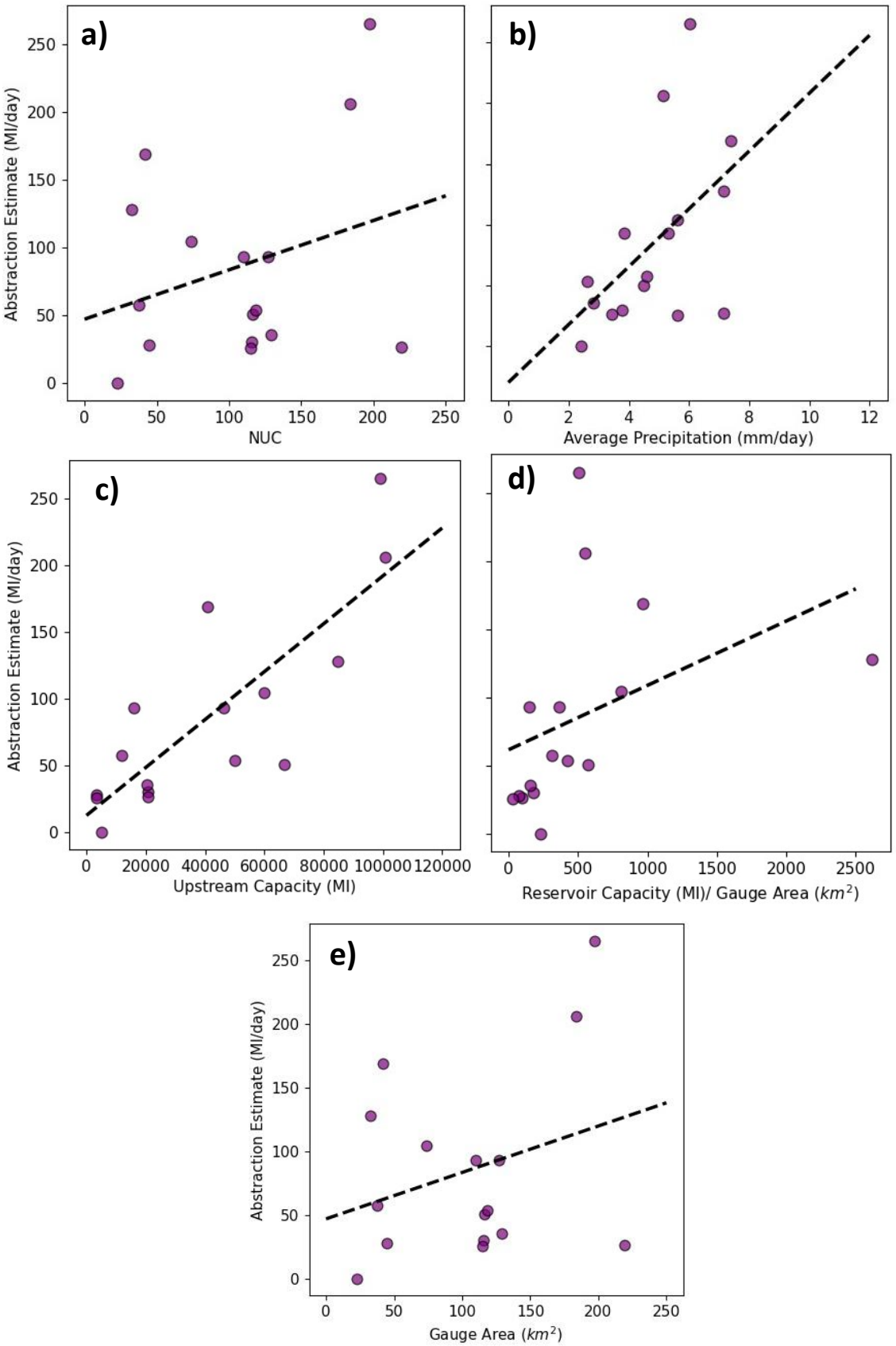

Figure S2. Relationships between a selection of catchment and reservoir attributes and the abstraction volume.

S2. Parameter sensitivity

When defining the bounds for sampling global parameters in the transfer functions we did some initial sensitivity checks, to make sure that our parameter space was representative. Below are some examples of the ABS and CF parameter sensitivity at a selection of gauges. We found that our parameter ranges were sufficient (i.e. the maximum KGE is captured by the parameters) and that in general results were very insensitive to the CF parameters and very sensitive to ABS (though as you move further from a reservoir the results become very insensitive to both parameters).

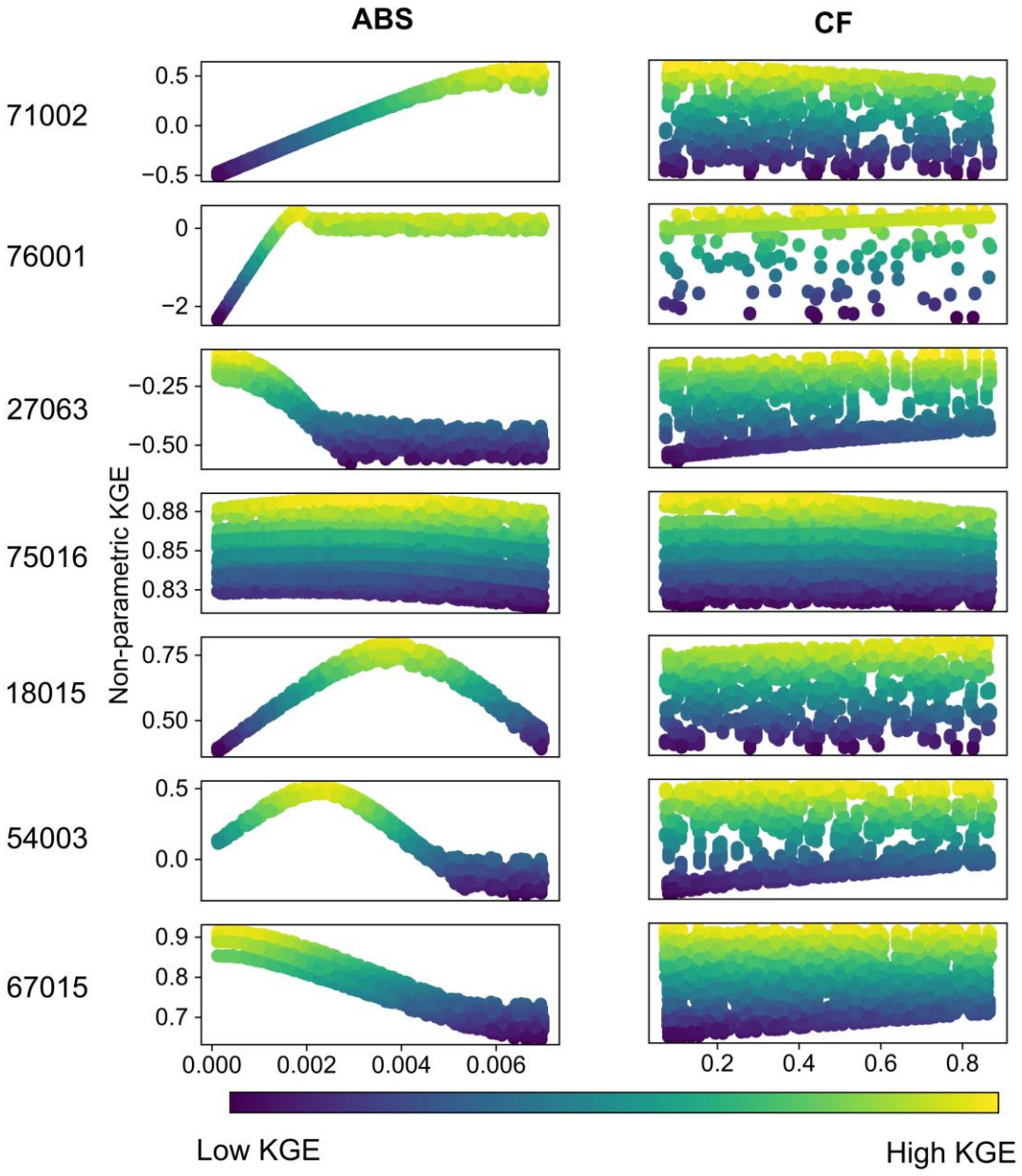

Figure S3. Parameter sensitivity at a selection of 7 gauges. Points are colored by their non-parametric KGE.

S3. Full results all gauges (catchment-by-catchment)

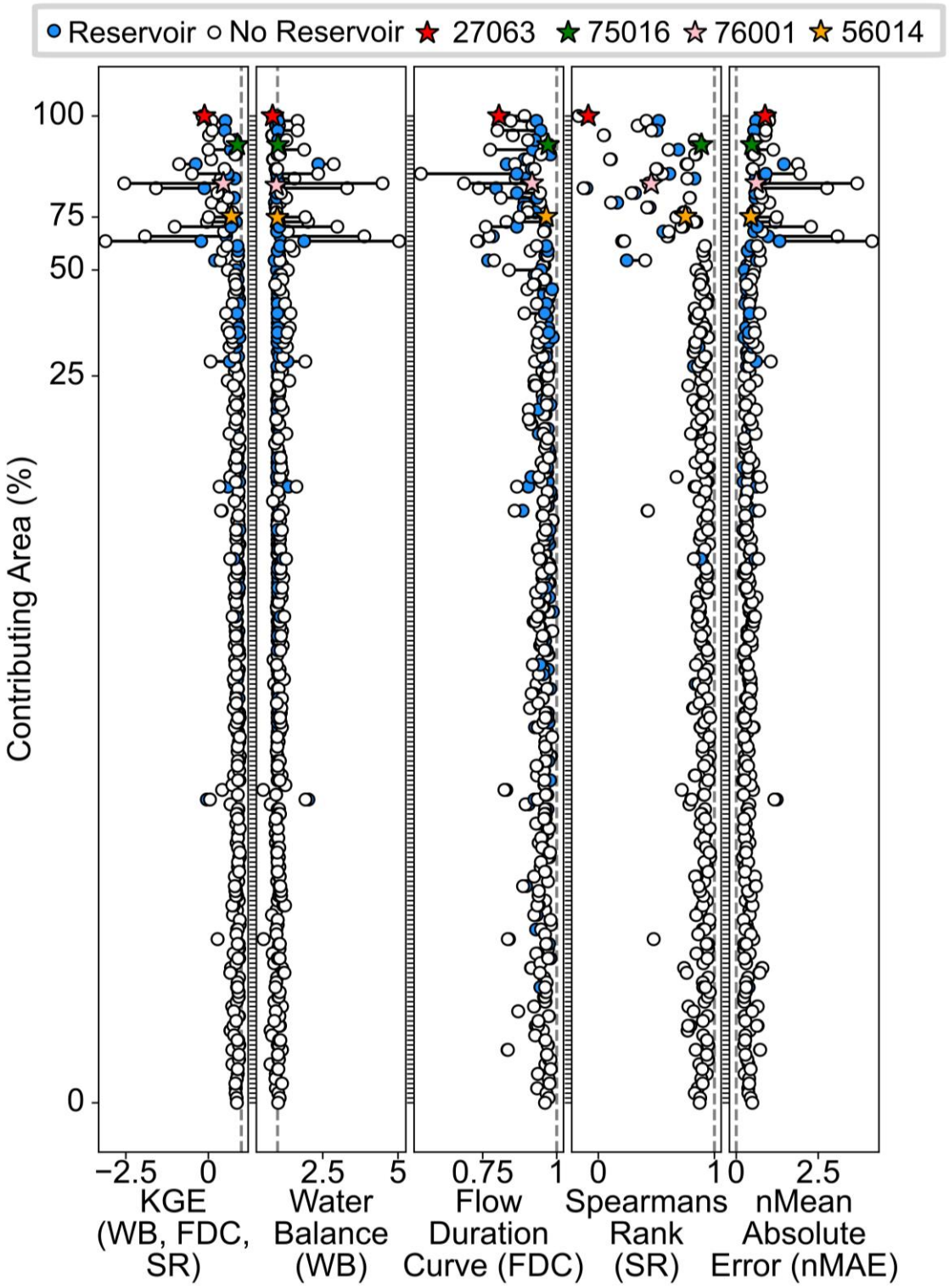

Figure S4. Difference in performance between the top reservoir simulation in each catchment and the best performing no-reservoir simulation. Results are presented for the non-parametric KGE metric and its relative components as well as the normalised Mean Absolute Error. Catchments are ordered based on their contributing area. 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 of the main paper.

S4. Full results all gauges (nationally consistent calibration)

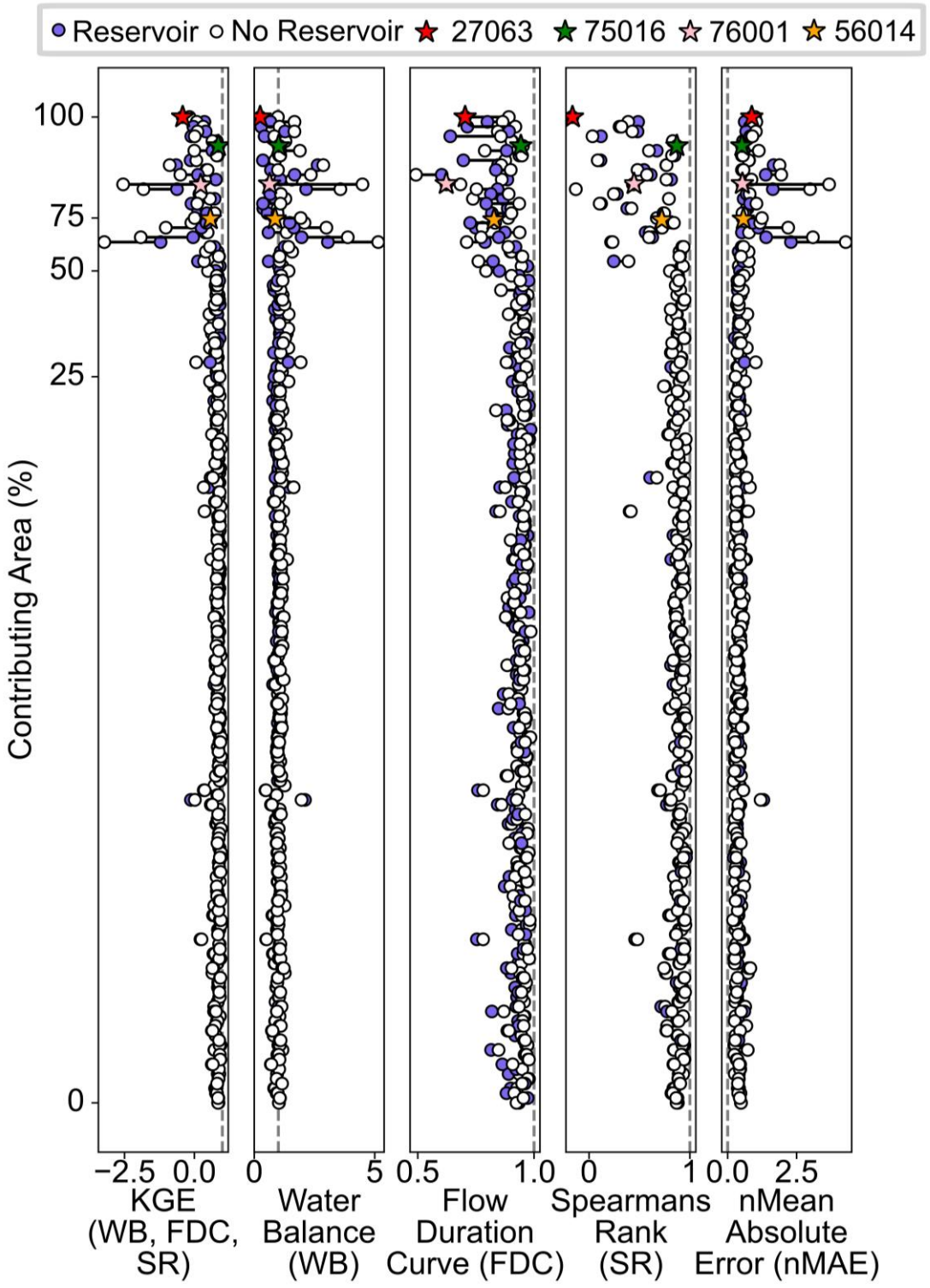

Figure S5. Difference in performance between the nationally best reservoir simulation and the nationally best no-reservoir simulation. Results are presented for the non-parametric KGE metric and its relative components as well as the normalised Mean Absolute Error. Catchments are ordered based on their contributing area. 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 of the main paper.

S5. Alternative subsets for identifying the best nationally consistent simulation

To identify the best nationally consistent parameterization we find the median difference in KGE between reservoir and no-reservoir simulations across a subset of gauges. The results in the main paper present results from the simulation with the best median KGE difference taken from only the gauges with a contributing area exceeding 25%. We chose this because these are the gauges where we see most impact, and if we are to consider the median, we do not want this to exclude the most impacted catchments (since those with a contributing area exceeding 25% only account for 25% of the total sample of reservoir catchments).

However, we found that our results were reasonably insensitive to this choice, and the results where the subset contains all gauges (Figure S6), gauges with a contributing area over 50% (Figure S7) and gauges with a contributing area over 75% (Figure S8) are presented below to demonstrate this and supplement the plot in the main paper.

The median KGE (for all gauges with a contributing area higher than 25%) is 0.72-0.76 across all of the subsets. When using all gauges to select the best nationally consistent simulation only 33 of the 58 gauges see an improvement in KGE, whereas when you subset by contributing area (25, 50 or 75%) 38 gauges see an improvement in KGE. 25-27 gauges have a KGE improvement of more than 0.1 no matter which subset you choose, 12-15 have an increase of 0.3 and 9-10 an increase of 0.5.

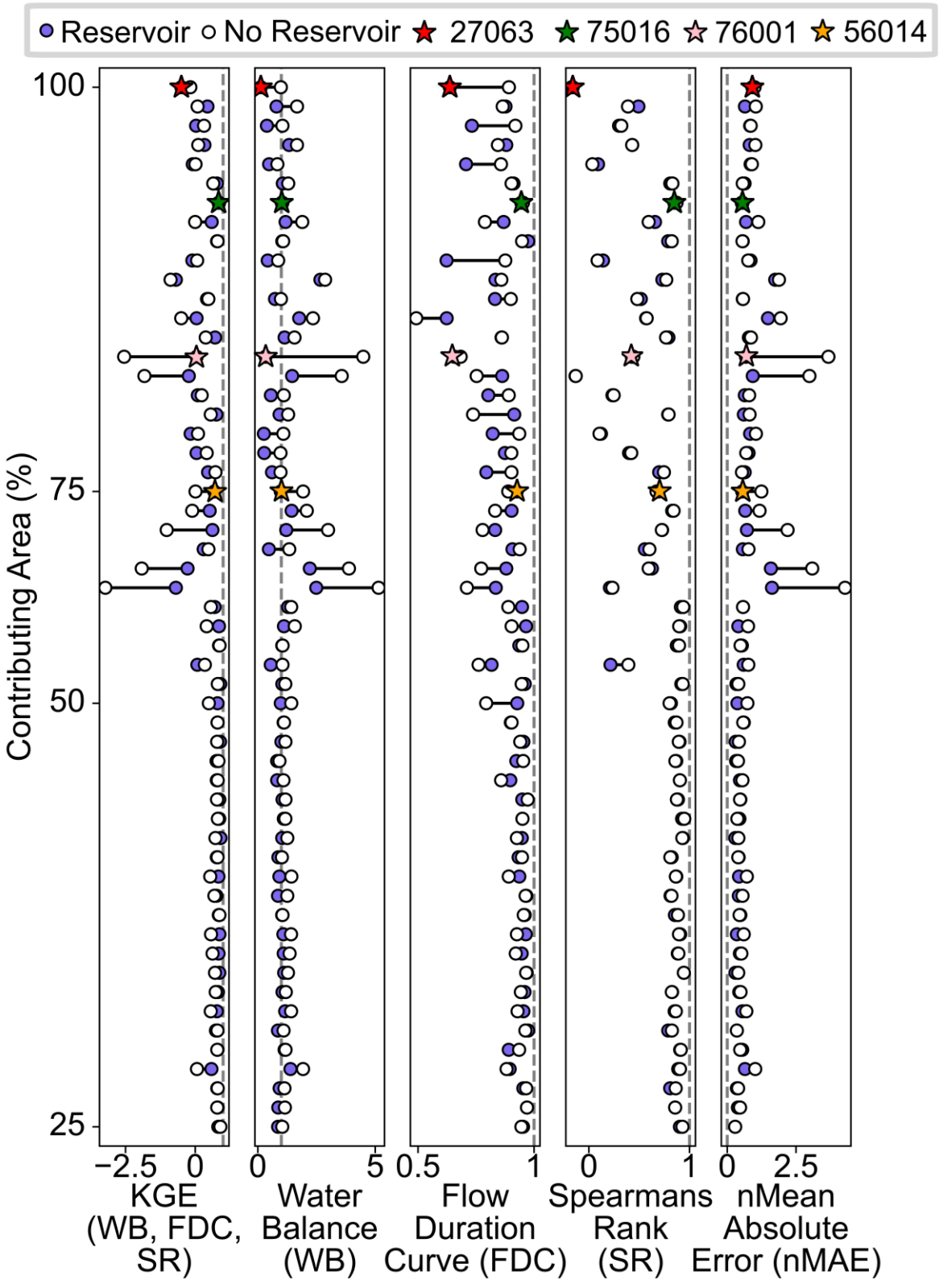

Figure S6. Difference in performance between the nationally best reservoir simulation and the nationally best no-reservoir simulation. Here the nationally best simulations are chosen using all reservoir gauges.

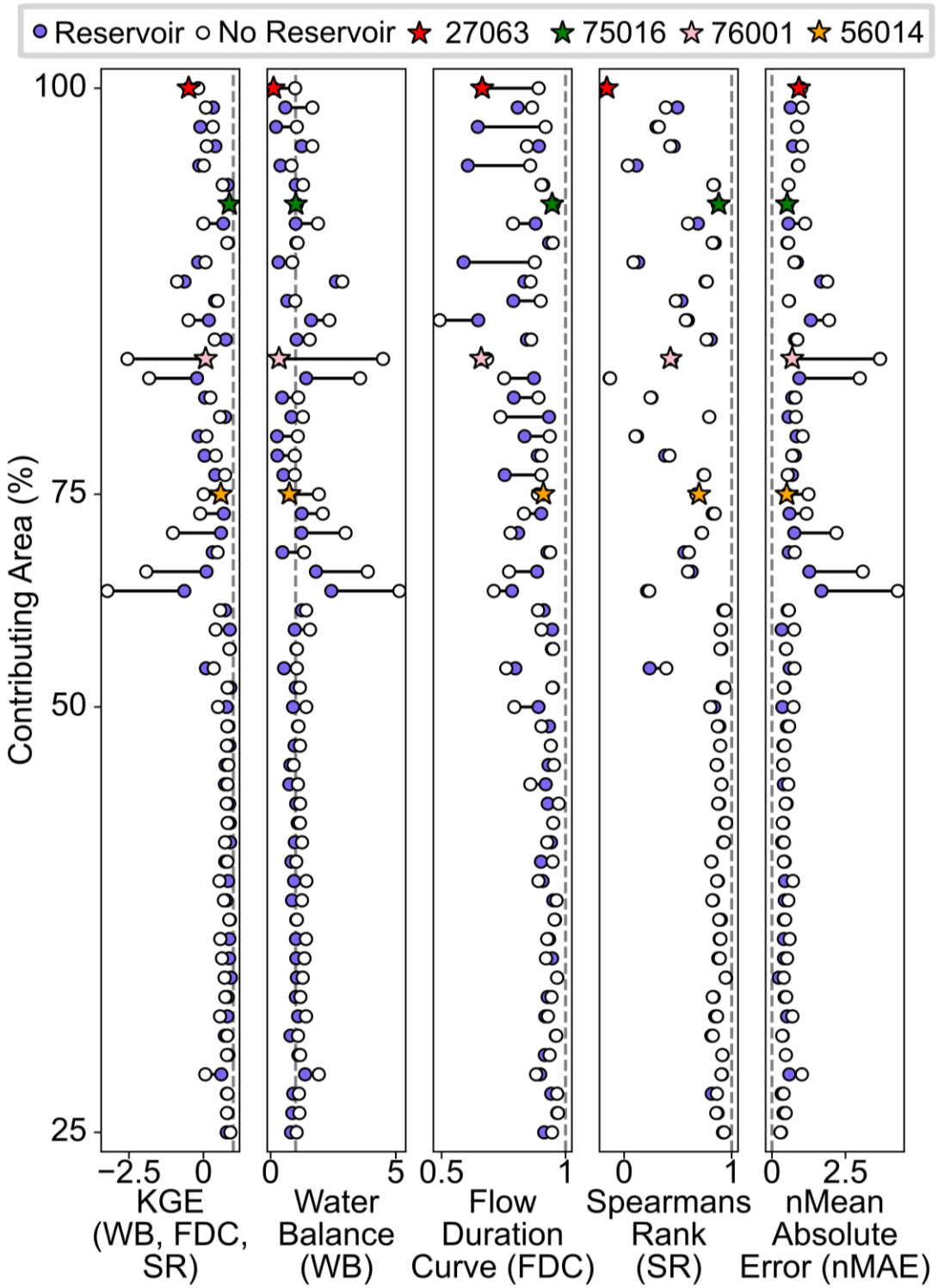

Figure S7. Difference in performance between the nationally best reservoir simulation and the nationally best no-reservoir simulation. Here the nationally best simulations are chosen using only reservoir gauges with a contributing area exceeding 50%.

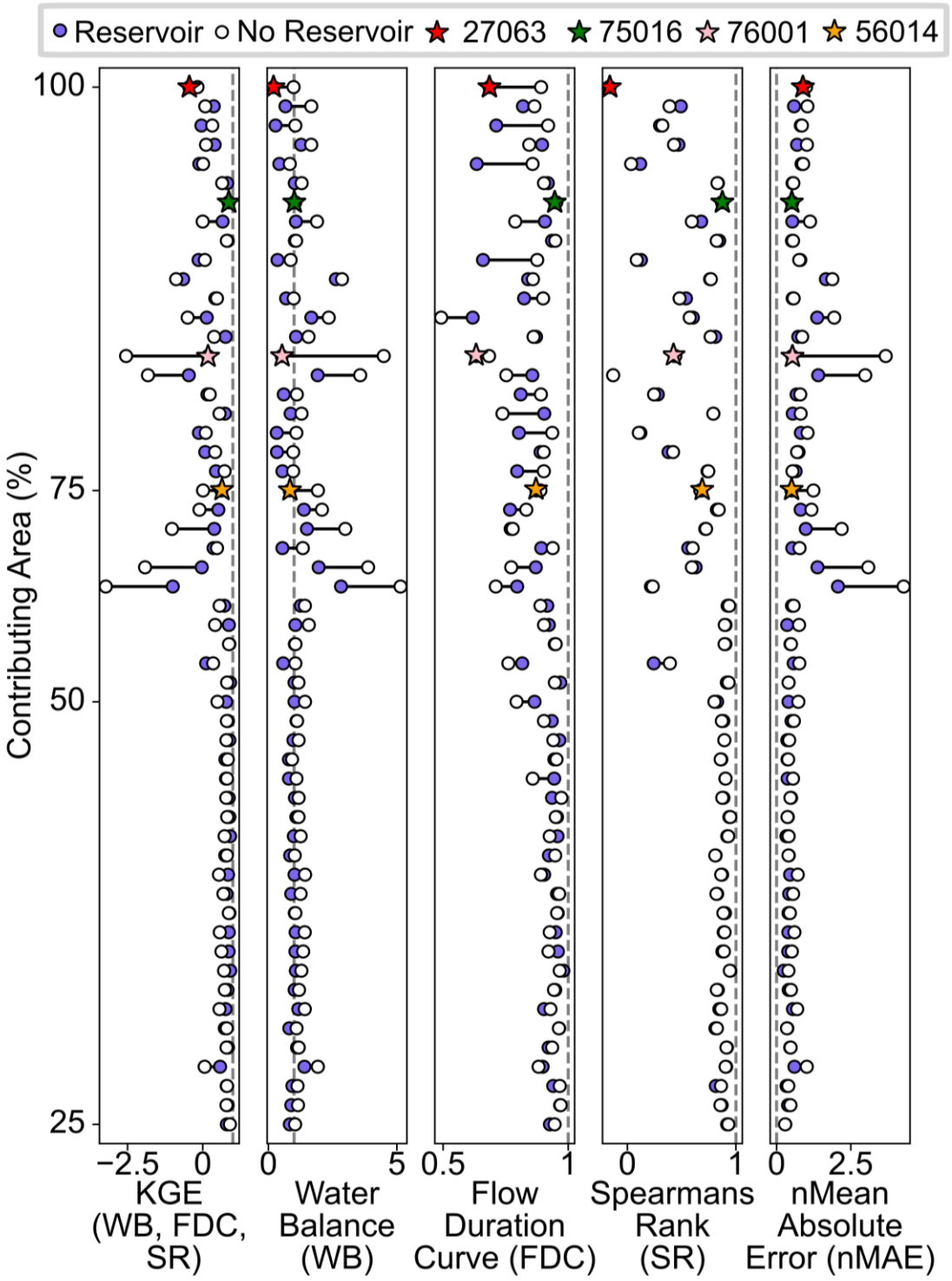

Figure S8. Difference in performance between the nationally best reservoir simulation and the nationally best no-reservoir simulation. Here the nationally best simulations are chosen using only reservoir gauges with a contributing area exceeding 75%.

S6. Problems with the Spearman's rank metric

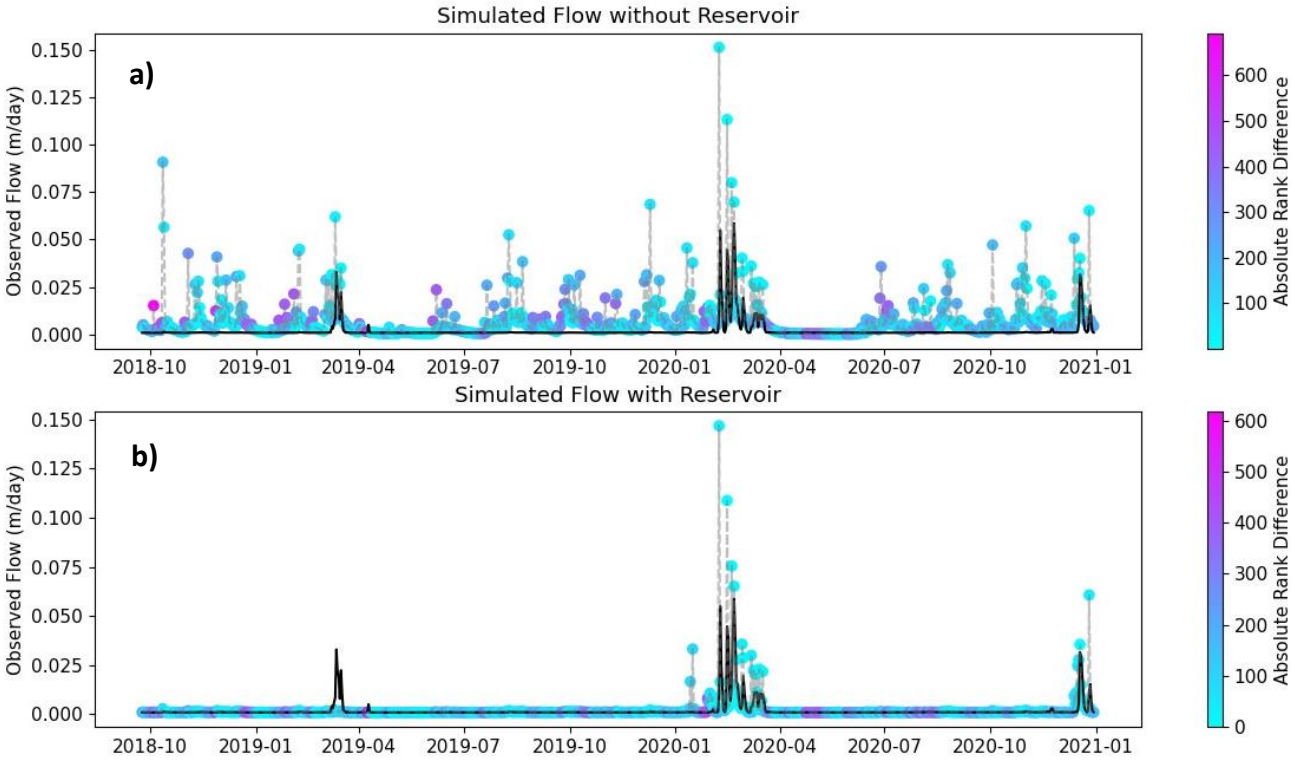

Figure S9. Observed flow at gauge 76001 and **(a)** flow simulated without reservoir representation and **(b)** flow simulated with reservoir representation with points colored based on the rank difference between observed and simulated flow.

Figure S9 visualizes some of the problems we encountered with using the Spearman's Rank metric on impacted timeseries. The metric is able to rank points in the no-reservoir (naturalized) timeseries reasonably well (see Figure S9a), where the differences in rank between observed and simulated timeseries are higher (pink dots) when there is an unmatched peak in flow. However, even in the no-reservoir scenario there is still a plateau of low flow in April/ May/ June of 2020 where you can begin to see some of the problems arising. Here there are data points where although the flow magnitude is similar, the ranks are vastly different. This is same problem that we see in Figure S9b where the simulated timeseries includes reservoir representation and is dominated by compensation flow. We would expect the largest rank differences (between observed and simulated data, represented by pink dots) to be found where the peaks in flow are over or underestimated, but instead these are all found in the flow plateaus. We believe this problem stems from having so many data points with similar values which is common in reservoir-impacted timeseries.

S7. Reservoir storage simulations

As part of the reservoir representation integrated into DECIPHeR in this study, the model is able to output reservoir storage timeseries. Although we do not have access to any large-sample datasets to evaluate or calibrate the model using storage data across GB, the simulated storage is displayed below in Figure S10 for the four catchments featured in Figures 6 and 7 in the main paper. Results are presented in each catchment for both the best nationally consistent parameterization and the

catchment-by-catchment parameterization to contextualize some of the changes in the hydrograph and KGE components seen between the two calibrations.

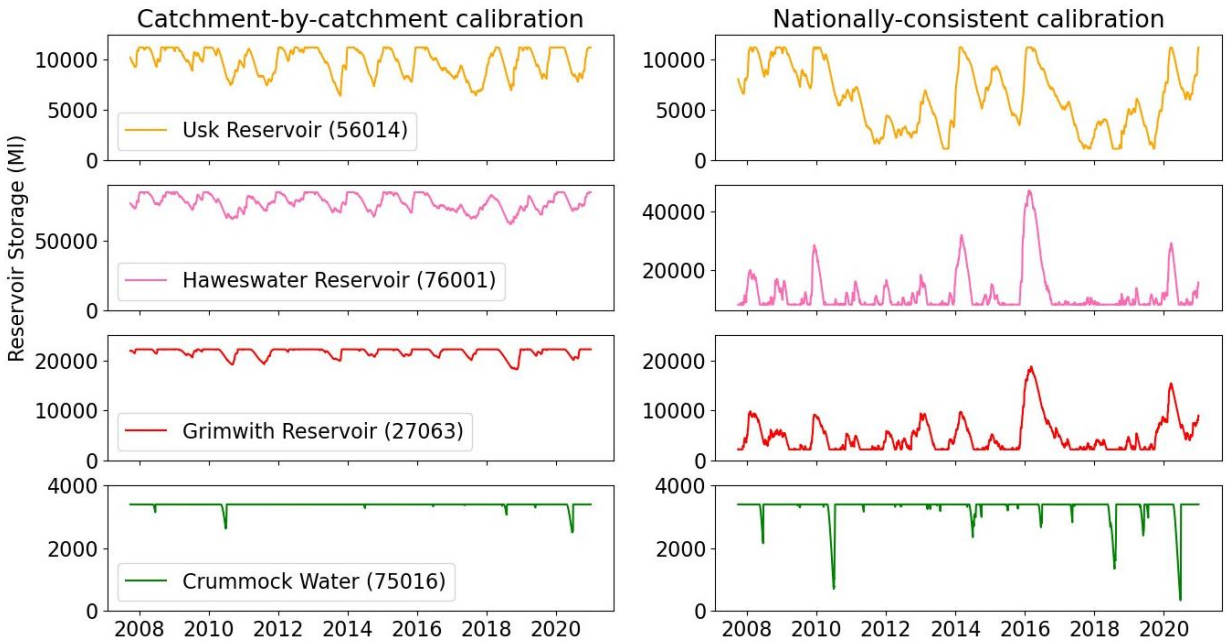

Figure S10. Storage simulations for four reservoirs associated with best catchment-by-catchment (column one) and nationally-consistent (column two) simulations. Chosen reservoirs match those featured in Figures 6 and 7 in the main paper.

In some locations across GB, river level data is available from the Hydrology Data Explorer (https://environment.data.gov.uk/hydrology/explore). In Figure S11 we have used this observed data to compare the storage patterns simulated at Haweswater reservoir in the catchment-by-catchment calibration to river level data. Although this river level data cannot be directly compared to changes in reservoir volume (because the bathymetry of the reservoir is not known), we note that the cycles of drawdown and refill are similar, suggesting that the models simulation of storage is reasonable.

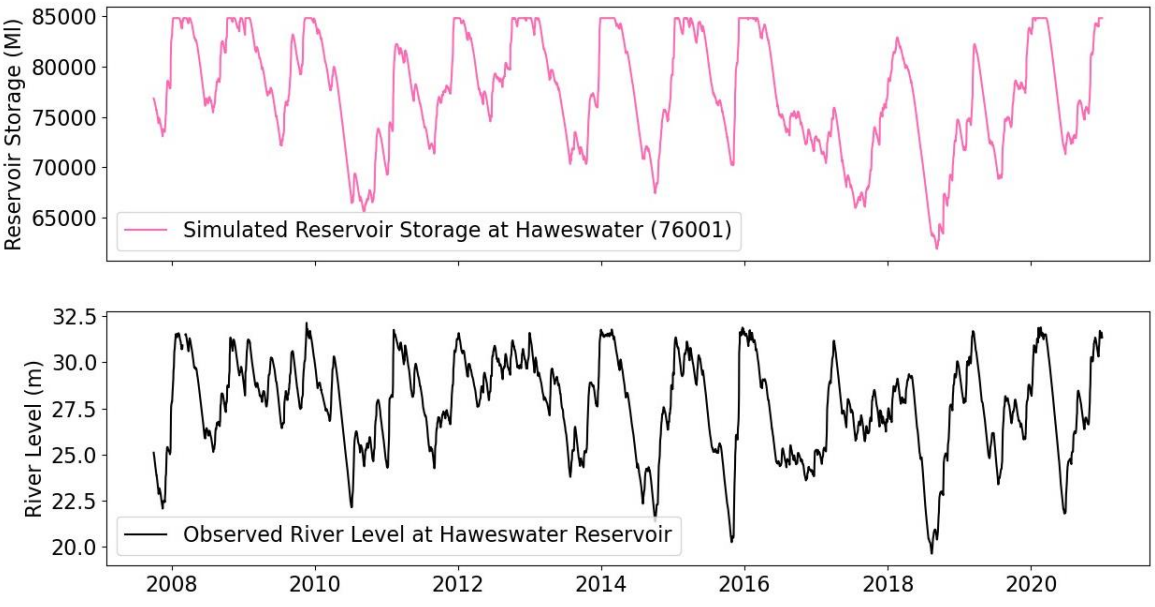

Figure S11. Simulated storage from the best performing catchment-by-catchment calibration at Haweswater reservoir (top) and observed river level data from Haweswater reservoir (bottom).

S8. Hanasaki rule comparison

To compare the new water supply reservoir operating rules introduced in this paper to an alternative set of rules, Figures S12 and S13 show the hydrographs associated with simulations using both our new rules and the rules defined by Hanasaki et al. (2006).We find that the Hanasaki rules are not well suited to simulating flow downstream of water supply reservoirs, largely because they have no abstraction component which is a key part of operations at water supply reservoirs in GB. By not accounting for the abstractions taken from the reservoir, reservoirs simulated by the Hanasaki rules are full much more often and therefore often spill in periods where in reality only the compensation flow is released. Table S3 compares the non-parametric KGE scores associated with both the Hanasaki rules and the rules introduced in this paper. It is clear from this table that the Hanasaki rules are not able to well represent the water balance and at three of the four featured gauges where the non-parametric KGE achieved with the Hanasaki rules is lower than what is achieved by a model with no reservoir representation.

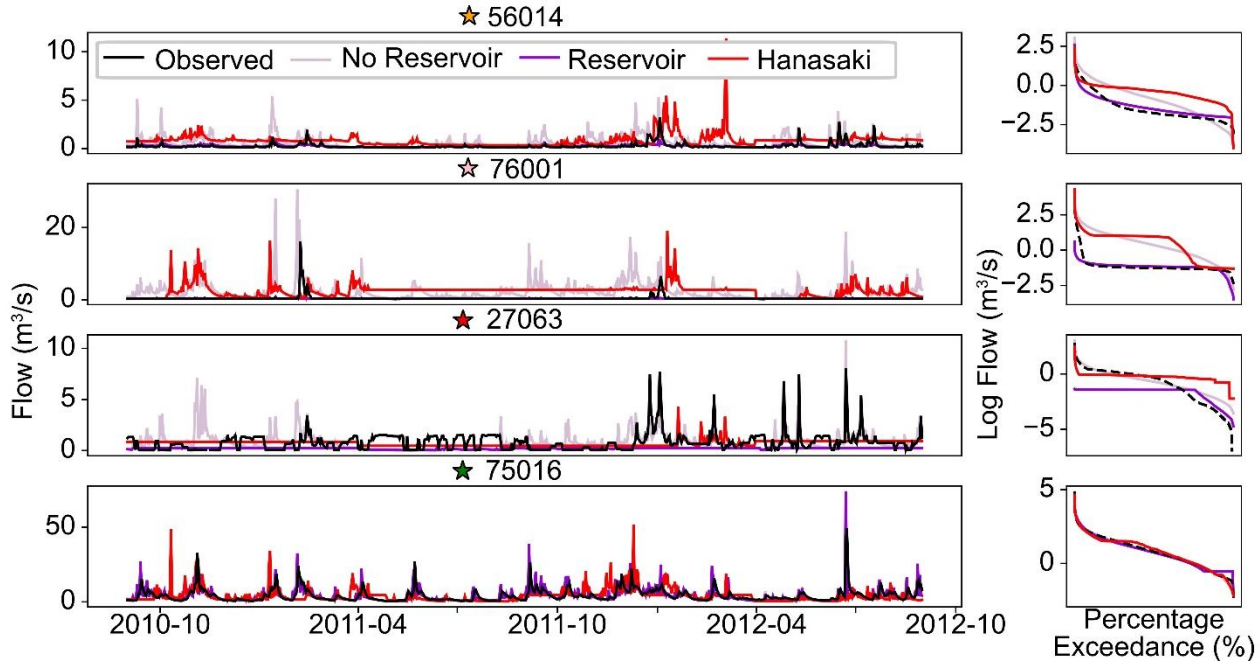

Figure S12. Hydrographs and flow duration curves top median simulation (nationally-consistent calibration) for selected reservoir catchments compared to simulations using the Hanasaki non-irrigation reservoir rules (Hanasaki et al. 2006).

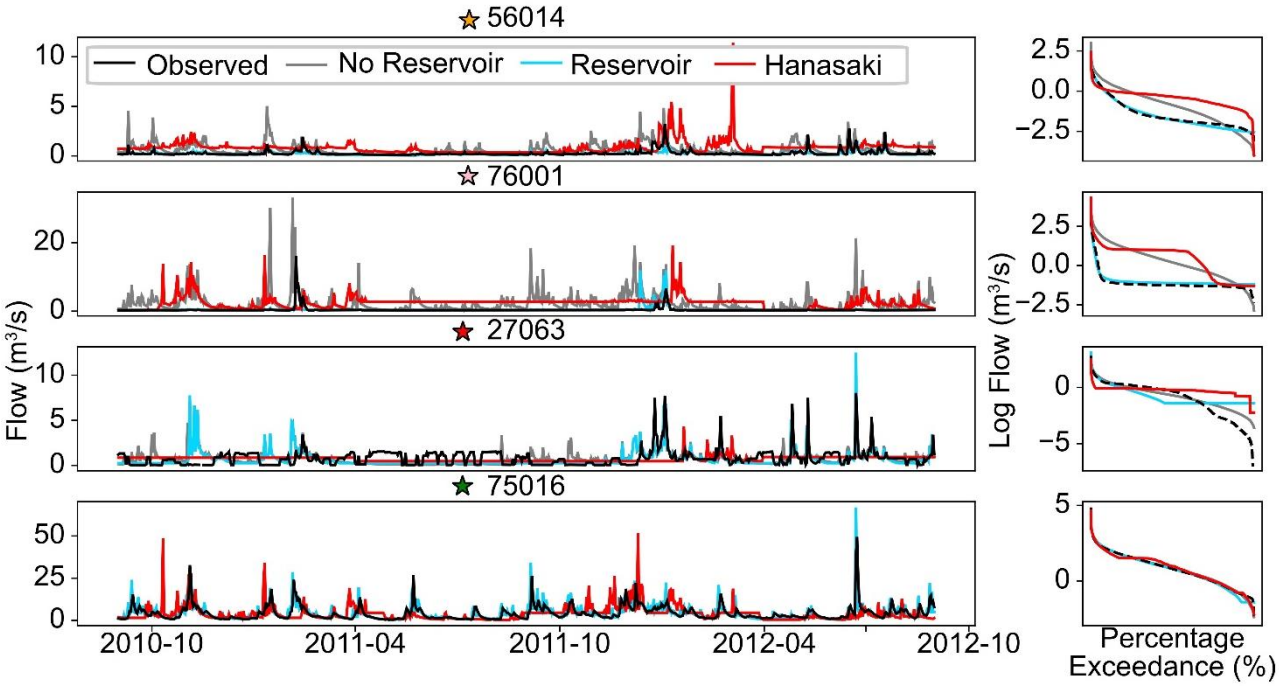

Figure S13. Hydrographs and flow duration curves from the top individual simulations (catchment-by-catchment) for selected reservoir catchments compared to simulations using the Hanasaki non-irrigation reservoir rules (Hanasaki et al. 2006).

Table S3. Non-parametric KGE and associated components for simulations at four featured gauges using water supply operating rules (with both catchment-by-catchment (CBC) calibration and nationally-consistent (NC) calibration), no reservoir representation and Hanasaki et al. (2006) non-irrigation operating rules.

Catchment	Operating rules	Water balance	FDC	Spearman's rank	Non-parametric KGE
56014	Water supply rules (CBC)	1.00	0.96	0.69	0.69
	Water supply rules (NC)	0.85	0.83	0.68	0.60
	No reservoir representation (NC)	1.93	0.89	0.67	0.01
	Hanasaki non-irrigation rules	1.93	0.69	0.40	-0.15
76001	Water supply rules (CBC)	1.04	0.92	0.45	0.44
	Water supply rules (NC)	0.63	0.62	0.43	0.23
	No reservoir representation (NC)	4.49	0.68	0.44	-2.54
	Hanasaki non-irrigation rules	4.49	0.63	0.35	-2.61
27063	Water supply rules (CBC)	0.83	0.80	-0.08	-0.11
	Water supply rules (NC)	0.24	0.70	-0.16	-0.42
	No reservoir representation (NC)	0.98	0.89	-0.17	-0.17
	Hanasaki non-irrigation rules	0.98	0.73	0.11	0.06
75016	Water supply rules (CBC)	1.01	0.97	0.89	0.88
	Water supply rules (NC)	1.01	0.94	0.87	0.86
	No reservoir representation (NC)	1.04	0.95	0.87	0.85
	Hanasaki non-irrigation rules	1.04	0.93	0.83	0.81

S9. Spatial distribution of reservoir transfer function parameters

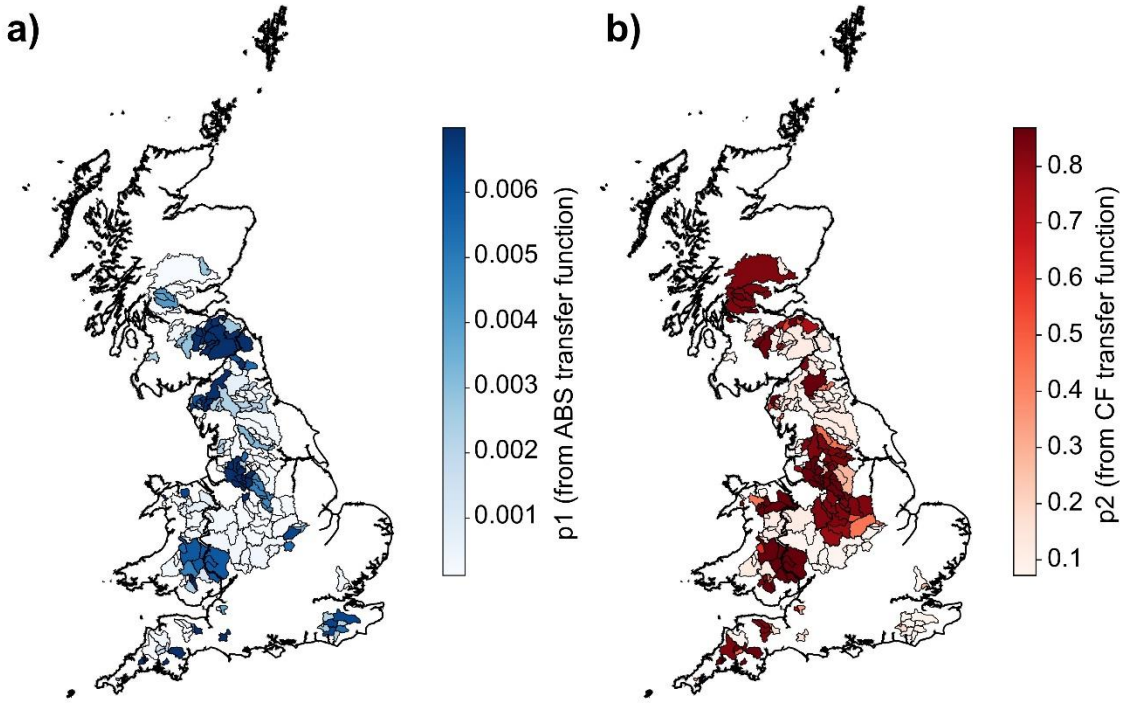

Figure S14. Spatial distribution of transfer function parameters associated with both CF and ABS. Catchments are colored by the transfer function parameters identified in the catchment-by-catchment calibration.

References

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.

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.