Response to reviewers

General comments

This paper reports on the development of a generic algorithm for water supply dams and its application to a hydrological model for the United Kingdom. The development of hydrological models that include reservoir operations is essential for estimating flow rates and water resources. The paper argues that a relatively simple model can represent the operation of water supply dams in the UK. The results reported are generally considered reasonable, but some of the descriptions were difficult to understand and should be refined.

Thank you for your review and helpful comments, we will address each of them in the response below.

Specific comments

Line 185, ”three equal classes of slope and accumulated area”: This part is hard to read. What do you mean by “equal classes”? What is “accumulated area”?

Thanks, we have clarified this in the manuscript as follows::

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 makes sure that the spatial variability of the climatic inputs is appropriately represented.

Line 218, “to determine unbiased values for the natural model parameters in reservoir catchment”: What do you mean by “unbiased”?

By using the term ‘unbiased’ we refer to the fact that DECIPHER’s seven standard model parameters (which have been designed to simulate the hillslope hydrology and are referred to as ‘natural’ model parameters) should not be overcompensating for any reservoir processes. We agree that the term ‘unbiased’ may be confusing so we have removed it such that the sentence now reads:

The top 10 natural transfer function parameter combinations (assessed using the non-parametric KGE (Pool et al., 2018), see section 3.6) are subsequently used to determine the natural model parameters in reservoir catchments.

Line 234, “the OS river layer”: What is this?

We have clarified this in the manuscript:

We use a 50-m gridded digital elevation model (Intermap Technologies, 2009) to generate the river network in DECIPHeR, extracting headwater cells from the an open-access river network which maps the rivers across GB generated by the Ordnance Survey (Ordnance Survey, 2023).

Line 252, “5000 times in the reservoir scenario sampling reservoir parameters”: I couldn’t understand these runs and how they were implemented. What do you mean by “sampling”?

‘Sampling’ refers to the random selection of the 5000 parameter combinations chosen from within the set bounds. We will change this sentence to instead read:

Each catchment was simulated 10 times in the no-reservoir scenario using the top 10 natural transfer function parameter combinations. In the reservoir scenario each catchment was instead simulated 5000 times using 500 reservoir parameter combinations for each of the top 10 natural transfer function parameter combinations.

Line 264, "In this study abstraction and compensation flow remain constant throughout the simulation." This is an interesting point. It implies that the summation of the abstraction and the compensation flow must be equal to the mean annual inflow (otherwise, the long-term water balance does not close). The key difference between this study and Hanasaki et al. (2006) is whether to separate the outflow of reservoirs into abstraction and compensation flow. This point would be mentioned somewhere.

We agree that the one key difference between our operating rules and the Hanasaki operating rules is that our rules include a public water abstraction which is removed directly from the reservoir (and therefore the water balance is not closed within the catchment but is exported for external water supply). This is a very important component of the reservoir operations in GB. We will distinguish this important difference between our rules and the Hanasaki rules in a new paragraph which we will add to the Model Evaluation section (to avoid repetition, see our response to your final comment for a copy of this paragraph).

Line 304, "The results from the near-natural simulations for the best nationally-consistent set of transfer function parameters": It is a bit awkward to read this. Why don't you put a simulation name? I think the combination must be [NATural or REServoir], [INDividual or national-CONsisnt]; hence, this run might be called "NAT-CON."

Thank you for this suggestion but we would prefer to avoid introducing more acronyms.

Line 315 "(or simulations, decided based on those with the highest median non-parametric KGE across all 137 catchments)": Hard to read. Better to clearly explain/define this in the Methods section.

We will add the following sentence to the section describing these runs in the methods section to make this clearer (Line 219):

We identify the top 10 natural transfer function parameter combinations by calculating the non-parametric KGE (Pool et al., 2018, see section 3.6) in all near-natural catchments and selecting the 10 combinations with the highest average non-parametric KGE.

Line 333, "to define the transfer functions used in this study": It is a bit confusing. I believe at least two transfer functions are used in this study: one for hydrological parameters and the other for reservoirs. For better readability, these should be clearly distinguished.

We agree this is difficult to distinguish. We will alter the manuscript to always specify whether we are referring to the natural transfer functions or reservoir transfer functions.

Figure 3: Please confirm whether the label "Top national simulation" is correct. I expected "Best national simulation". Anyway, the mixture of top and best is confusing for me.

Thank you for pointing this out. We will be more consistent with our language here and use only 'Top'.

Line 338-339 (Equations 6-7): Again, this is a very interesting point. The compensation flow is a function of the catchment area. Because the inflow to a reservoir is basically proportional to the catchment area, it can be said that this equation is similar to Hanasaki et al. (2006; CF ~ mean annual inflow). The abstraction is a function of reservoir capacity. Because the reservoir capacity is only weakly correlated with the inflow usually, it can be said that this formulation is different from Hanasaki et al. (2006). Then, let me come back to my previous point. The equations 6-7 do not seem to guarantee the water balance shown in Equation 1. From this perspective, it is doubtful that the overall modeling framework is hydrologically reasonable. This point should be additionally discussed in the manuscript.

The water balance shown in equation 1 is guaranteed by equations 4 and 5 which ensure that the abstraction and compensation flows are reduced to zero if the storage is too low. We agree that the CF should be proportional to the mean annual inflow and we also agree that by including an abstraction flux our rules are different from those introduced by Hanasaki. We will make this clearer by distinguishing the key differences between our rules and the Hanasaki rules in a new paragraph which we will add to the Model Evaluation section (to avoid repetition, see our response to your final comment for a copy of this paragraph).

Line 341 “is in line with the observed data”: What does it mean by “in line with”? The gray dotted line is basically above the blue dots.

To check that our natural and reservoir functions were generating physically realistic parameter values for the CF and abstraction parameters (and to define the transfer functions in the first place) we compared the parameters identified by our transfer functions to the CF and abstraction values recorded in the literature (where this information was available). By saying that our chosen transfer functions identify ABS parameters “in line with the observed data” we mean that the parameters chosen by the best nationally-consistent transfer functions are similar to (or in line with) those recorded in the literature.

That said, this is only true for the ABS parameter and we agree that the gray dotted line is above the blue dots. We found that the model was very insensitive to the choice of CF parameter and so although the best nationally-consistent transfer functions identify ABS parameters which are in line with the observed data the same is not true for the CF and we will point this out here.

We will rephrase this section of text below to account for your following two comments as well as this.

Line 343 “at the upper end of the sampling limits”: What are the “sampling limits?”

These are introduced in Table 1. These are the bounds within which the transfer function parameters are sampled. We will reference this table in the text which we will rephrase below.

Line 343 “the range of variability of the transfer functions associated with the top 5% of national-consistent simulations”: Top 5% in terms of what? What do you mean by “the range of variability”?

Here we are referring to the 5% of simulations with the highest average KGE value (across the whole sample of catchments). The range of variability refers to the amount of variability in the shape of the transfer functions within this top 5%. We will make this clearer in the text and agree it could be much better described.

The text will address all three of the previous comments by being rephrased to:

The top nationally-consistent transfer function associated with the ABS parameter (marked on Figure 3a 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 3b 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 3 also shows the variability in the transfer functions associated with the top 5% of nationally-consistent simulations (this is displayed on Figure 3 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).*

Line 376 (Figure 4 caption) “top reservoir simulation”: Again, what is the difference between “top” and “best”?

As above, we will be more consistent with our language here and use ‘top’.

Line 502- “the simplicity of our operating rules”: Ideally, the author should also demonstrate that their model outperforms the even simpler model (Hanasaki et al. (2006), which requires no calibration parameter).

We have now demonstrated for four key reservoir catchments the differences in model simulations between our reservoir rules and Hanasaki’s rules. This is demonstrated by two figures and a table (see below) which will be added to the supplementary information. Our results reveal some inherent problems associated with using the Hanasaki rules to simulate flow downstream of water supply reservoirs and show that in three of the four examples the Hanasaki rules are outperformed by not just our operating rules but also the model with no reservoir representation. We will add the following paragraph to the model evaluation section:

Our model lies within a spectrum of complexity regarding approaches to simulating reservoir operations. The operating rules implemented in this study are designed to simulate the key components of water supply reservoir operation using a physically realistic representation of the abstraction and compensation flow fluxes, whilst also minimizing the complexity and number of parameters. Although rules such as those introduced by Hanasaki et al. (2006) have no calibrated parameters and are therefore arguably more simple than ours, we found that they demonstrated poor performance at water supply reservoirs. This was largely because the Hanasaki rules do not allow for abstraction from the reservoirs which is a key component of reservoir operation across GB (see section 8 in the supporting information). On the other hand, the high data requirements associated with alternative data-driven approaches to defining operating rules (e.g. Turner et al. (2020)) meant that this type of rule could not be implemented at the national-scale across GB. In this paper we instead compare simulations using our new operating rules to a model without reservoir representation. This comparison allows us to evaluate our approach against the pre-existing methodology included in DECIPHeR (where reservoirs are not represented).

The following will be added to the supplementary material:

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.*

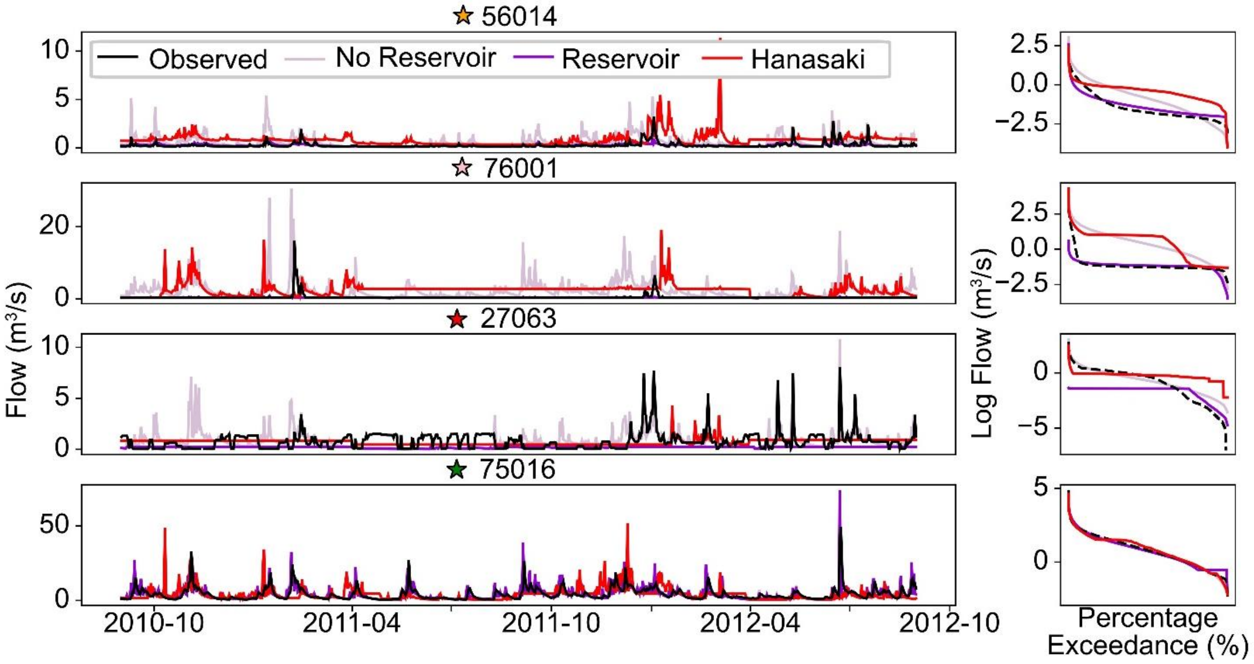

Figure S12. *Hydrographs and flow duration curves top median simulation (nationally-consistent calibration) for selected reservoir catchments compared to simulations using the Hanasaki non-irrigation reservoir rules (Hanasaki et al. 2006).*

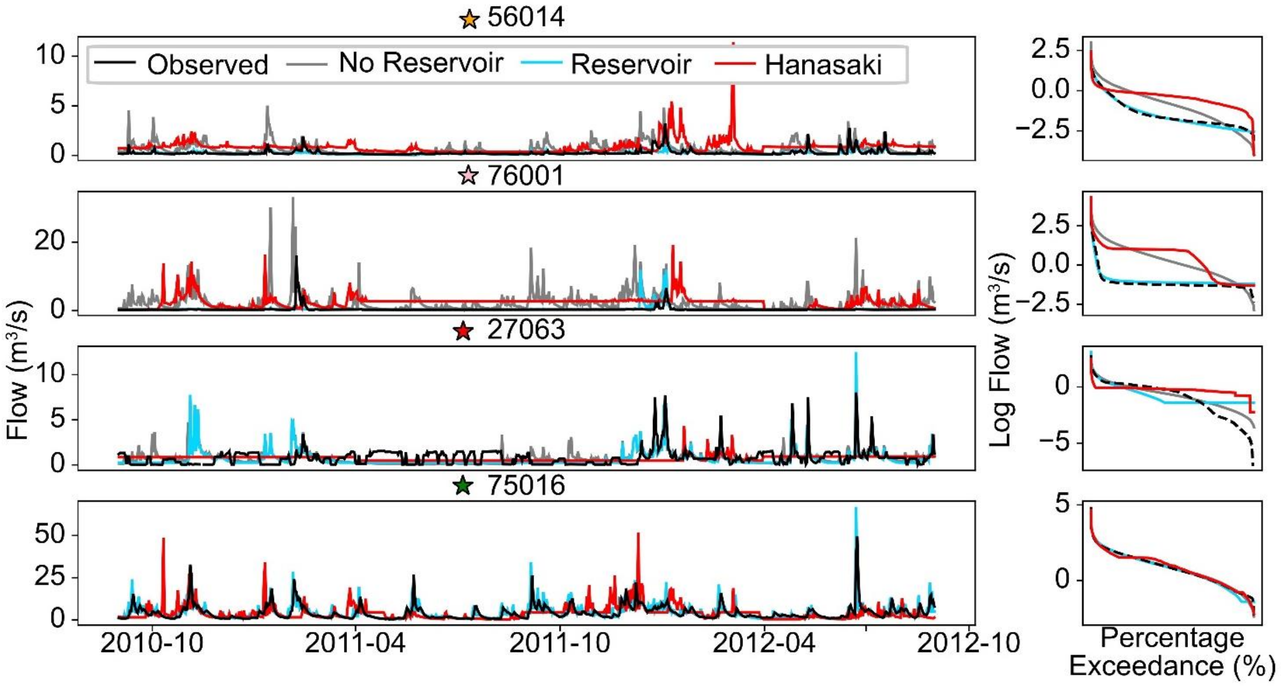

Figure S13. *Hydrographs and flow duration curves from the top individual simulations (catchment-by-catchment) for selected reservoir catchments compared to simulations using the Hanasaki non-irrigation reservoir rules (Hanasaki et al. 2006).*

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