Review Comments:

This paper describes a simplified reservoir operating scheme that is distinctly oriented toward smaller non-irrigation reservoirs that only uses two parameters. This reservoir scheme uses general catchment and reservoir parameters that are more readily available than reservoir time series data. To test the sensitivity and performance of this reservoir scheme, the authors derived parameters for water supply reservoirs throughout Great Britain that encompass reservoirs of multiple sizes and uses. They also analyzed the feasibility of using transfer functions to create a nationally consistent parameterization. In their work, they evaluate the impact of two calibration methods one within each catchment and the other across the entirety of Great Britain. Results show that the methodology works well in catchments with primarily water supply uses and is limited in catchments with multi-purpose reservoirs. This article does a really good job of creating simplified curves and evaluating metrics for evaluating reservoir dynamics with a unique focus on water supply reservoirs.

Thank you for your useful review, we will address each of your comments individually below.

Major Edits:

Equation 4 and 5 have Smin, which I assume refer to dead storage, however, this variable is not defined in the text. Please clarify this in the text so the reader is better able to follow what the equations are referring to.

We will define this in the text, thank you.

Section 3.5.1: As readers may be unfamiliar with the study domain, it could be useful to have a figure that shows the general makeup of the reservoirs used in this study colored by use, size and a second panel with the naturalized and non-naturalized catchments and general characteristics. This would allow a reader who is unfamiliar with the catchments in GB have a better idea of the characteristics of the catchments mentioned in the paper.

Thank you for pointing this out. We will the figure below to section 3 which shows the size and distribution of water supply reservoirs across GB and the reservoir and natural/ near-natural catchments.

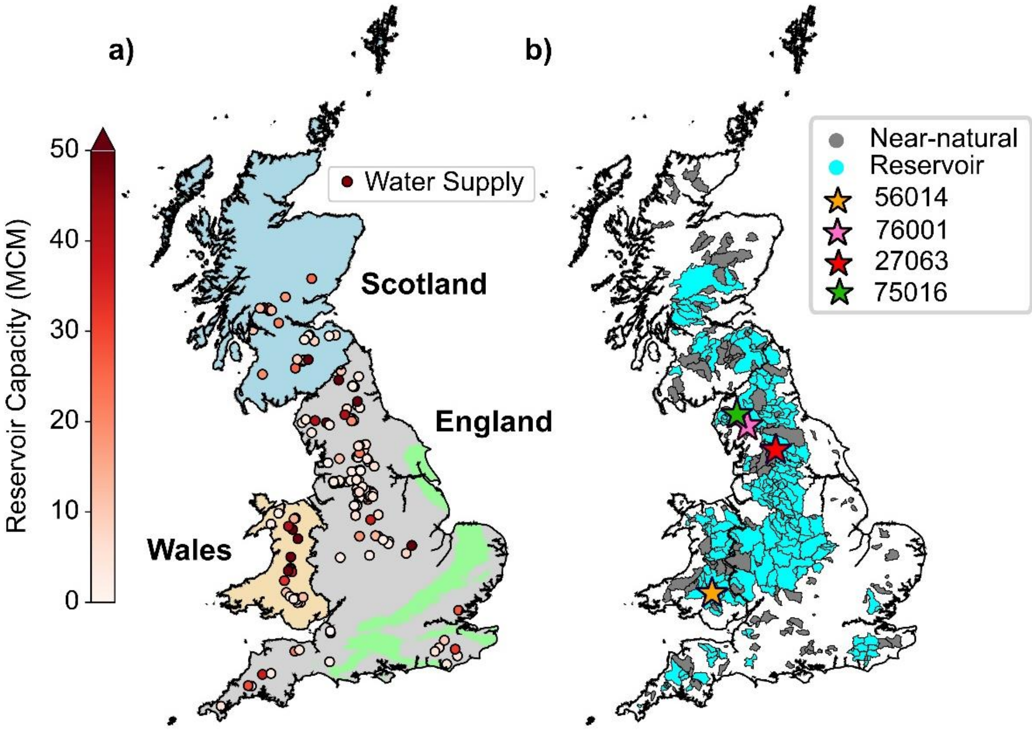

Figure 2. *Distribution of (a) water supply reservoirs across GB and (b) near-natural and reservoir catchments used in this study. Reservoirs are coloured by their storage capacity and the four catchments featured in Figures 6 and 7 are highlighted with stars on subplot b.*

Line 308: I am personally not familiar with the catchments in GB. To increase clarity, I would include a catchment map either with numbers or highlight these catchments in Figure 2 so that the reader knows which area you are referring to. Alternatively, you could leave out the reference to these specific catchments and describe them more generally. The same also goes for Line 388, 390, and 460. I would make the distinction as to why you picked these catchments and give an overview of them prior to mentioning them so the reader is not confused.

Thank you for this suggestion. We will describe these catchments more generally as you have suggested, in this instance we agree that the catchment characteristics are more important than their specific location. For example, at line 308 we will alter the text to read:

For example, the Aldbourne at Ramsbury (39101) and the Ewelme at Ewelme Brook (39065) which are both chalk catchments…

We will also make sure that all references to the four featured catchments (denoted by stars on Figures 4-7) are highlighted in the text and we have added the locations of these catchments to the new Figure 2 introduced above. For example, on line 388 where we reference the Haweswater Beck at Burnbanks (76001) we will add (denoted by a pink star on Figures 2 and 4-7).

Line 313: This is an interesting note about reservoir construction in GB and it's impacts on your work. What would be the larger impacts on other regions that are more groundwater regions dominated such as the southwestern US or for dams built on limestone or more porous rock such as the Mosul Dam in Iraq?

We are currently coupling the surface water components of DECIPHeR to a groundwater model (Rahman et al, 2023) so it would be important to include this model development if the model was applied to reservoirs in catchments with significant amounts of groundwater flow.

Line 337: The rational for using a linear relationship is not clear from the text. Why was a linear relationship better than a non-linear one?

Thank you for pointing this out, we will alter the text to read:

Since the observations (Figure 3) do not show any evidence of non-linearity, we chose to use a linear (and hence more parsimonious) relationship for both transfer functions.

Line 350: Including Table 1 with the upper and lower bounds of the transfer function parameters is nice, however, the Table is not referenced at all in the text. Instead of a table, it might be nice to include a figure of the parameter ranges, average parameter value, or another metric to show the regional differences in transfer function parameters across different catchments in GB.

Thanks – we have now added a reference to this table into the text. The upper and lower bounds of the parameter ranges are displayed on Figure 3 but we will add the figure below to the supplementary material which shows the regional differences in transfer function parameters by plotting the parameters chosen in each catchment by the catchment-by-catchment calibration spatially across GB.

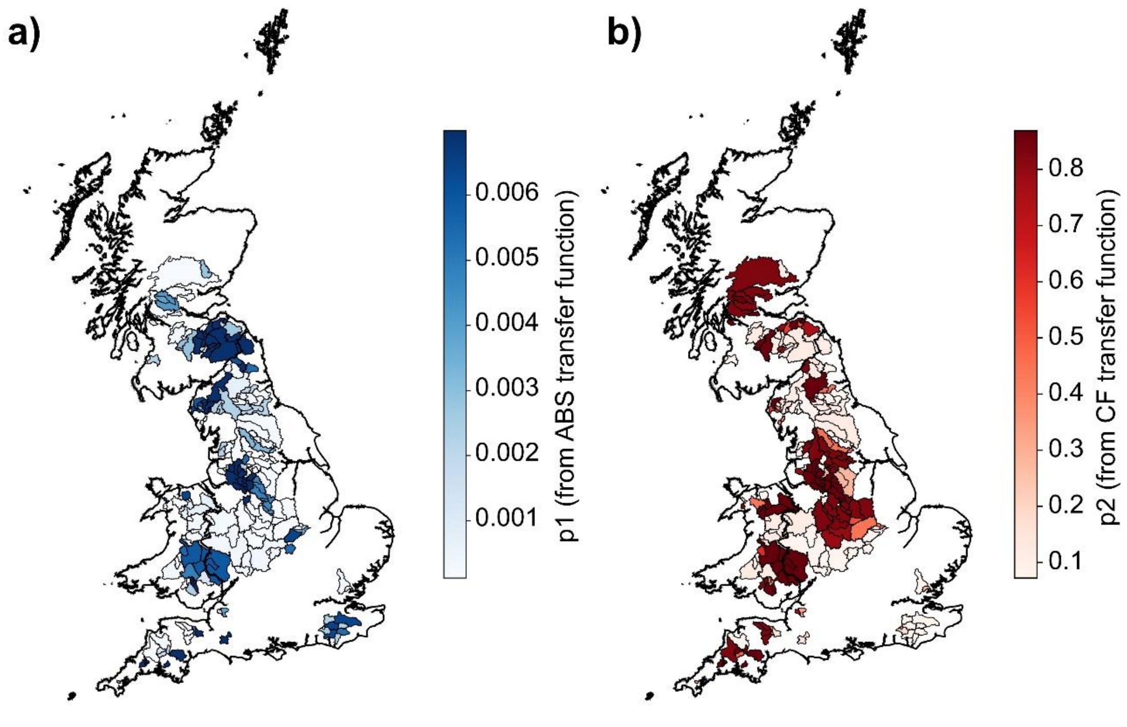

Figure S14. *Spatial distribution of transfer function parameters associated with both CF and ABS. Catchments are colored by the transfer function parameters identified in the catchment-by-catchment calibration.*

Figure 4 and Figure 5: Why did you pick the four catchments denoted by stars in these figures? What characteristics cause those to be chosen for the case studies and do these directly relate to the ones mentioned in Line 308, 388 and 390? If so, I would make the connection between these clearer.

We chose these four catchments because they show the successes and areas for improvement for our operating rules. This description is included in Section 4.4 where these catchments are investigated in more detail, but we will add the following description into the text after the 4 featured catchments and their results are first introduced in Figure 4 to make this clear earlier on:

Figure 4 highlights four catchments with star markers which are designed to demonstrate the central improvements and improvements needed in the new operating rules. 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.

Only one of the catchments referenced in lines 308, 388 and 390 directly relates to those featured using the stars, but we will point this connection out in the manuscript by referencing the star color. Line 388 will become:

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 4-7) 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.

Line 549: Can you expand on what you would suggest other do in cases where your methodology is limited (i.e. in the case of multipurpose reservoirs)? Should a combination of methodologies be used or a more generalized approach such as Hanasaki et al., 2006? Additionally, how would the authors envision upscaling this approach to other hydrologic models?

Thank you for your comment. We discuss this in the limitations and future work section of the discussion but will make this clearer here by rephrasing the text to say:

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).

Section 4: The authors do a really nice job of explaining their methodology and results. It would be interesting to see a comparison between this simpler model and one or the more generalized approaches or perhaps the approach that is already in DECIPHeR as another point of evaluation.

Thank you for your comment. At present there is no alternative approach for representing reservoirs in DECIPHeR, so by comparing our results to the model with no reservoir representation we are comparing our results to the current approach. We will make this clearer in the manuscript.

Regarding a comparison to a more generalized approach, this point has been raised by both yourself and the second reviewer so we will add a comparison to the Hanasaki rules into the supporting information (see Figures S12 and 13 and Table S3 below) and add the following paragraph to the Model Evaluation section to address this:

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.

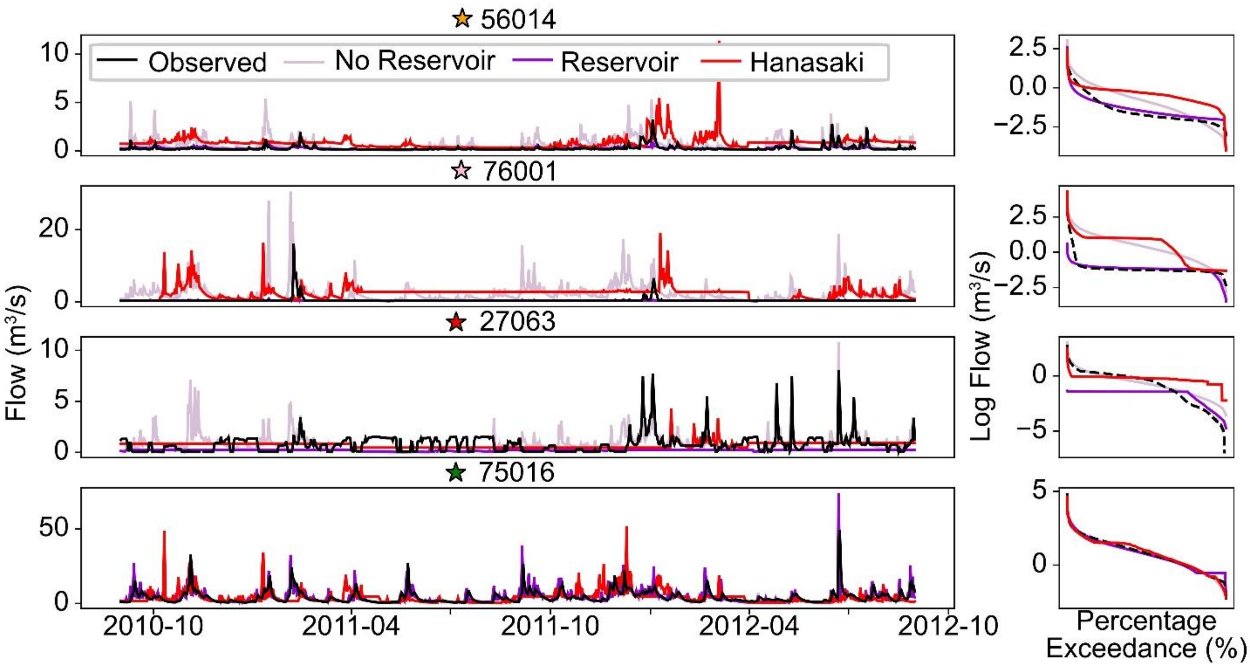

Figure S12. *Hydrographs and flow duration curves top median simulation (nationally-consistent calibration) for selected reservoir catchments compared to simulations using the Hanasaki non-irrigation reservoir rules (Hanasaki et al. 2006).*

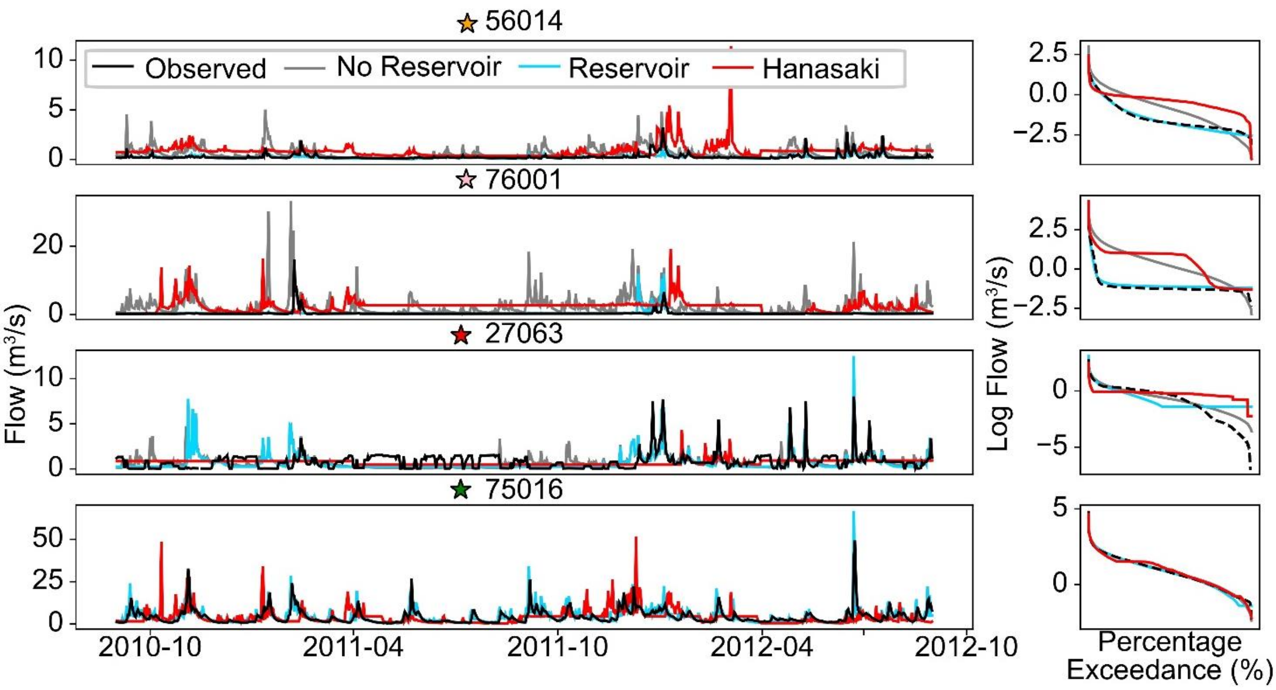

Figure S13. *Hydrographs and flow duration curves from the top individual simulations (catchment-by-catchment) for selected reservoir catchments compared to simulations using the Hanasaki non-irrigation reservoir rules (Hanasaki et al. 2006).*

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

Minor Edits:

Line 53 and 55: I believe the abbreviation for GRanD should have a capital R as well.

Thank you for pointing this out, we will correct this across the manuscript.

Line 60: ResOpsUS contains over 600 dams. I would edit this line accordingly.

Thank you, we will correct this.

Line 125 – 128: I would suggest adding parenthesis to equation 4 and 5 to denote the right side of the comparison.

Thank you, we will add this.

Figure 3: I would switch units in this figure from Ml/day to cubic meters per day since most audiences are more familiar with cubic meters.

We will change these units.

Figure 6: The color difference between reservoir and observed is hard to distinguish. I would suggest altering the colors to be a bit more clear. I would also change the units of CF and ABS to be cubic meters /sec or per day so the units across all graphs are consistent. I do like that you kept the colors of the stars the same throughout Figure 4, 5 and 6.

We will alter the colors to make this clearer, thank you.

I like the colors on Figure 7 and the difference between the three curves is much easier to see. I would also change the units on CF and ABS to cubic meters per second or day for consistency.

We will change these units.

Line 570: Could you give an example of a signature or test that you would suggest for other studies?

We will reword this section to make it clearer that these new metrics still need to be developed. The text will now read:

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 (Kiraz et al., 2023).

Line 589 – 590: Perhaps I missed a section, but I am not sure where the first half of this sentence comes from since the large scale hydrologic models that I am aware of all contain reservoirs in Great Britain. If the authors are referring to water supply reservoirs specifically, I would suggest rephrasing.

Here we are referring to national-scale hydrological models, we will alter this sentence to read:

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.

Although we acknowledge that there are several global-scale hydrological models which include reservoirs in GB, the resolution of such models is too coarse to inform water resources planning.