# Response to reviewers

We thank the editor and three reviewers for the high-quality, constructive comments. They have really helped to improve the paper.

We addressed all reviewers' comments as completely as we could. As a result, most sections have been amended. The main changes to the paper are listed as follows:

- ➔ A new Section 3.1 explains in detail the methodology for the analysis, and why we used the Method of Morris.
- ➔ In the results, a new paragraph (Section 4.1.2 and Figure 6) justifies why we argue coordination between reservoirs is observable in historical operations, both to avoid flooding during the Spring 2011 and to avoid drought in 2012-13.
- ➔ Offline water balance models have been setup for both the drought and flood events where we are comparing simulation results to historical operations. They show that the coordination measures observed in the historical record would have been sufficient to avoid flood and drought in the simulation, if they had been implemented. These models are discussed in Sections 4.3 and 4.4 (and a new Figure 11)
- ➔ In Section 4.4, we decided to focus exclusively on the two reservoirs in the basin with a flood control role: Palisades and American Falls.

In what follows, page and line numbers given in "Response" parts refer to the original submission, whereas page and line numbers given in "Manuscript changes" parts refer to the revised submission.

## <u>Reviewer 1:</u>

<u>Comment 1:</u> This study deploys the Method of Morris to evaluate the sensitivity of release and storage time series to adjustments in the parameters of generic reservoir operating rules. The analysis is performed using a high resolution hydrological model (WBM) and focuses on a cascade of reservoirs located in the Upper Snake River Basin. The paper is well written, with clearly defined methodology and easy-to-follow results section. The experiments conducted are ill-suited to the aims of the study, leaving too many confounding factors. Specifically, I disagree with the authors key claim that the approach exposes coordination between reservoirs missing from the generic operating schemes (see specific comments below). While I do not doubt that coordination does occur in this particular cascade system, I cannot see how this sensitivity analysis exposes that coordination unequivocally.

<u>Response 1:</u> We want to thank the reviewer for their thoughtful and constructive comments. We believe they provide a basis for us to clarify and be more explicit with the aims and assumptions of our work, both with textual explanations and supplementary data. The

concern that our work does not distinguish when non-coordination is the most likely cause of the simulated drought and flood errors is also well-taken. In a revised version, we will show how coordinated operations would have avoided some (most) of the water extreme related impacts simulated in Sections 4.3 and 4.4, even in a model where (as in any hydrological model) there are other sources of error.

Our detailed response to the key concern expressed by the reviewer (that the errors seen in the paper may not be due to the absence of the coordination in the reservoir release rule) considers multiple aspects in this order:

1) Coordination is not represented in reservoir operation rules as presented in the current literature on large-scale hydrological models. This is not something this paper tries to demonstrate but a fact supported by the literature review made in the introduction.

2) Instead, this paper tries to diagnose some possible consequences of this non representation of coordination. Thus, our study is not focused on identifying the dominant parameters in the rule (for a reservoir in isolation, those can be deduced from the equations). Our Method of Morris analysis is a quantitative diagnostic for better understanding how the influence of the reservoir rules' parameters evolves both through time and through the multi-reservoir cascade. This diagnosis is helpful for addressing two key questions (i) does the assumption that reservoirs can be parameterized separately (made in studies that propose new and more sophisticated release rules) hold? And (ii) what are the potential consequences of this assumption for flood and drought assessments by increasingly detailed hydrological models in a changing world?

3) Whereas question (i) can be addressed by looking at simulation results alone (see Section 4.2), question (ii) is better examined by using a benchmark where we know for a fact that efficient multi-reservoir coordination has historically minimized flood and drought risk: the USRB. We provided evidence of that coordination when comparing historical and simulated storage and release trajectories. This is apparent when narrating the drought event in Section 4.3: p19 lines 2-4, 8-11 and 15-18, and when narrating the flood event in Section 4.4: p21 lines 25-34 and p23 lines 16-20. We however do agree with the reviewer that we did not provide direct evidence that coordination led to release behaviors that could not be accounted for with improved parameterization of the reservoir release rules. Please refer to point (A) below for a detailed explanation of how we propose to address this in the revision.

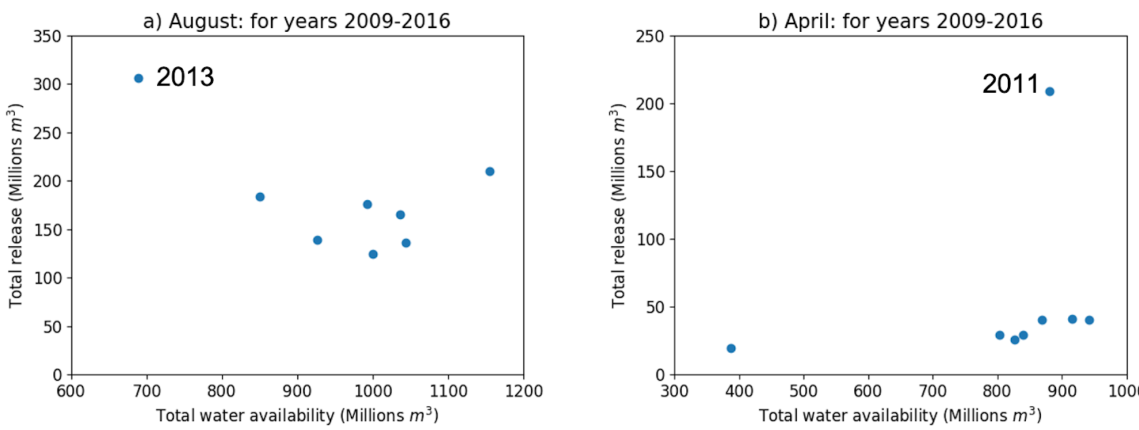
4) More specifically, our comparative analyses between historical and simulated release decisions in Sections 4.3 and 4.4 show qualitatively different behavior when contrasting historical data with simulation of the same period. We agree with the reviewer's concern that due to potential other errors in the model, our time-variant diagnosis approach does not in itself rule out that the observed differences may have little to do with the coordination issue. Please refer to point (B) below for a detailed explanation of how we propose to address this in the revision.

5) Please also note that as this is a diagnostic study, our goal is not to fully explain the errors or to correct them. There is to our knowledge no study of a real operational context using a

hydrological modeling experiment where one would absolutely know every source of error or potential confounding factor. That being said, we are grateful to the reviewer for remarks that underlined that we need to clarify the presence or absence of confounding factors.

Manuscript changes 1:

A) We have provided a more quantitative classification and proof of coordination in the historical record, for both the high flow and low flow event. In support of this, we now include the figure below as Figure 6 and added a paragraph in Section 4.1 (Page 17 lines 5-18) to discuss it. We included this Figure because both panels show unusual historical release patterns concurrent with both the flood and drought events. For the drought in August 2013 (Section 4.3; left panel) as for the flood in April 2011 (Section 4.4; right panel), the release response (y-axis) to storage + monthly inflows (that's total water availability, x-axis), is unusual. In the 2013 drought, the reservoir has already been emptied to rescue American Falls levels downstream but releases are still the highest in 8 years, whereas in the 2011 flood, releases are around 8 times as high whereas water availability is similar for 6 of the other 7 years. In both cases the release difference with other years cannot be explained by immediately available variables.




As noted in our methodology (page 12 lines 7-11), we aim to "find qualitative behaviors in historical (i.e., observed) operations that cannot be accounted for by a reservoir's release rule simply by changing parameter values or integrating a near-term (less than a month) inflow forecast such as in some existing rules (e.g., Biemans et al, 2011). Other variables would be necessary to incorporate these behaviors within the hydrological models. They will be interpreted as preliminary indications of cooperation, and the events they highlight will be investigated further."

B) In Sections 4.3 and 4.4, we built simple offline water balance model of the reservoir cascade to investigate where observed historical coordination would have been sufficient to avert the simulated water extremes, notwithstanding other sources of error in the model (which will always exist). These simple offline water balance models of the reservoir cascade, detailed below for (i) the drought event, and (ii) the flood event, were used to conduct 'what-if' experiments where we look at the consequences of storage dynamics based on combining simulated inflows and storages (incorporating other sources of error in the model) with a releases that are modified from the simulated ensemble to introduce coordination. In detail, we have for each experiment conducted with an offline water balance model:

(i) for the drought event (Section 4.3, page 21 line 30 to p 22 line 3):

"To test whether the observed differences between the historical record and simulation are an artefact of other errors in the hydrological model, we built a simple offine water balance model over 2012-2013, taking the three reservoirs' simulated water balance offline. For each of the 400 ensemble members, we then simulated operations obtained by releasing 8 million m3 more each day in July and August than planned by release rules -- a total extra release of 496 million m3 [matching the difference between historical and simulated trajectories observed in Figure 8]. The extra release concerns only Palisades in 2012, and both Palisades and Jackson Lake in 2013. To make sure that underestimating routing times between reservoirs does not falsely cause the reservoirs to store water, we choose a conservatively high (for the area) routing time of 7 days between each reservoir and the one downstream. Results (Supplementary Information Figures 1 and 2) show that with this simple coordination water balance measure, American Falls would not have emptied -- and reservoirs upstream of it would not have lost their capacity to supply downstream agriculture either."

Note that the 7-day routing delay is a conservative assumption: if higher releases with long delays are enough for the water to reach the downstream reservoir in time, then the shorter actual delays would be enough as well.

(ii) for the flood event (Section 4.4) we focus exclusively on the two reservoirs that provide flood control in the USRB: Jackson Lake and Palisades (and we took American Falls out of the analysis). This improved focus enabled us to better describe the coordination between these two reservoirs, and to highlight the real-world consequences of peak releases as high as those simulated out of Palisades (page 25 lines 18-21). We also describe an additional offline reservoir water balance experiment (Section 4.4.3 starting p25, line 26), that quantitatively verifies that a simple early release / two-reservoir coordination policy avoids catastrophic flooding within this version of the model.

C) we have added additional text in a new Section 3.1 (p 10 line 28 to p 12 line 27) to clarify the changes in our methodology and go beyond solely explaining the method of Morris, moving current Sections 3.1 and 3.2 to 3.2 and 3.3 respectively. In particular, section 3.1 now details:

1. that the analysis is seeks to "diagnose the potential consequences of non-coordination within large-scale models' reservoir release rule, our dual objectives are to (A) understand how these consequences play out through space and time in a large-scale hydrological model and (B) compare and contrast the implications of the reservoir rule representation for capturing real-world coordinated

reservoir operations", (page 10 lines 29-32).

2. the method of Morris is used a evaluative model diagnostic tool, where "the focus is not solely on which parameters in the release rule are most influential in regulating flows, but importantly clarifying how the set of dominant controls on water flows and reservoir storage levels evolves along a complex multi-reservoir cascade through time. Parametric sensitivities are then used alongside storage and release trajectories for the simulated ensemble to assess how sets of dominant controls in a point in time and at a given reservoir can be associated to high- or low-flow conditions." (page 12 lines 1-5)

3. that "the comparison will be supplemented by an analysis exploiting offline reservoir water balance models that implement basic coordination on the simulated ensembles at times where there are indications of cooperation. These offline reservoir water balance models help to show that coordination alone could explain the qualitative differences between simulated and historical operations. Note that we cannot fully quantify and explain all sources of errors: there is to our knowledge no study of a real operational context using a hydrological modeling experiment where one would absolutely know every source of error or potential confounding factor." (page 12 lines 11-19)

D) update the abstract to reflect the above points and clarify the methods and goals of the study. We inserted (page 1 lines 9-12):

"We employ a time-varying sensitivity analysis that utilizes Method of Morris factor screening to track how the release rule parameters that control reservoir storage and release evolve 1) along the cascade, and 2) in time according to seasonal high- and low-flow events. We combine this with a comparative analysis of historical operation and targeted experiments with simple offline reservoir water balance models."

<u>Comment 2:</u> I also have some concerns with the study design, and in particular the use of the high-resolution hydrological model, which introduces a severe and unnecessary computational constraint while also supplying the reservoirs with (likely) inaccurate historical inflows. The latter leaves the reader unsure as to whether the difference between simulated and observed storage/release is more a function of erroneous inflow than inaccurate operating decisions. This is particularly important when demonstrating the inadequacy of the reservoirs rules in representing flood and drought (it could simply be that the upstream hydrology from the model is delivering the wrong volumes of water). Since the actual observed inflows are available for this system, it would seem far more prudent to develop a simple, offline cascade model. I suspect such a radical change to the experimentation at this stage would be unrealistic, so I would instead encourage the authors to change the aims and storyline offered here. The simulations are sound and there is a great need for the community to learn more about the nature and performance of generic reservoir schemes. I

would support a revision if either (a) the authors can convincingly rebut my concerns listed below, or (b) a new angle is developed with a more defensible conclusion.

Response 2: Our introduction highlights that large scale hydrological models are being employed with increasingly ambitious assessment goals (p3 lines 2-14). These large-scale assessments of hydrological risks are not carried out with surrogates (offline simpler versions) of full-scale models. In this regard, WBM and the reservoir rule system represent a state-of-the-art representative of this class of assessment model that is distinguished in its representation of human infrastructure systems. Our study has a clear and direct design: evaluating model outcomes in the same conditions as large-scale hydrological vulnerability assessments are carried out.

Moreover, it is non-trivial to connect complex large regional infrastructures to the broader natural components of the water balance. For instance, inflows to other reservoirs need to reflect the consequences of simulated upstream releases (including what that means for flow routing along the river course where there are lateral inflows). There is also the question of which other parts of the model should be included: for instance there are potential consequences for water withdrawals in case of low-flow conditions. This would have required testing and arbitrarily excluding components of the water balance from the online surrogate to avoid making it overly complex.

This being said, we acknowledge that supplementary experiments involving simple offline reservoir water balance models would complement the main experiment (conducted with a full-scale model) to help clarify the contribution of coordination in averting impacts from water extremes. Therefore, we therefore thank the reviewer for their suggestion here.

Manuscript changes 2: We explicitly clarify that our study aims “to evaluate model outcomes in the same conditions as large-scale hydrological vulnerability assessments are carried out” (page 12 lines 24-25).

We also used complementary experiments using offline simple reservoir water balance models (see Manuscript changes 1).

Comment 3: Section 3.2. The justification for the 10% decision is unconvincing. Suppose a dam has an average inflow of 100 cumecs. An Rmin of 0.1 would be 10 cumecs and would vary by a maximum of 1 cumec. So the left-hand side of the operating curve hardly moves. In contrast, if your Rmax is 5 (=500 cumecs) then you’d vary this parameter by plus or minus 50 cumecs. So the right hand side of the curve will shift wildly in comparison. This is surely why Rmin appears to be unimportant in your sensitivity analysis; you’ve barely moved it. I appreciate that there are physical reasons why Rmin would be expected to be somewhat less variable than Rmax in absolute terms (although both are highly uncertain), but the uniform 10% assumption is not ideal either, and it invalidates your later statements about which parameters constitute the signature of parametric influence (page 15).

Response 3: This criticism of the chosen multipliers would be exactly on point if the goal of this sensitivity analysis was to ascertain which parameters in the reservoir rule are dominant. Instead (and as noted above), this analysis diagnoses how the parameter signature (i.e., the set of parameters that dominate and in which direction they influence outputs) evolves throughout the cascade and with time. The core contribution is not what single parameter controls are dominant, but that the dynamics in sensitivities are complex, highly dynamic, and non-separable. This is the core concern expressed in our Introduction, that these parametric behaviors are in direct contradiction to the current reservoir rule forms representation of reservoirs as being individually separable and independent parameterization problems without any information on coordination..

Manuscript changes 3: We acknowledge that the rationale for using the Method of Morris has not been explained with the requisite precision in the previous submission, so it is now explicit in Section 3, page 12 lines 1-5 in particular (see also the Manuscript changes #1 in this document).

Comment 4: Section 4.2 through 4.4. The results described here show that the operations of upstream dams can affect the decisions made at downstream dams. This occurs because any change upstream has an effect on the inflows into a downstream dam, thus affecting its storage levels and therefore its releases (which are a function of storage). Sensitivity analysis exposes how the decisions taken at one dam are affected by the rules deployed at other dams (often with some intriguing complexities). This is insufficient, however, to claim that the models must be missing *coordination* (which presumably means dam operator A looking at the storage levels of dams B and C to inform his or her decision). Similarly, the results that describe inadequate representation of flood and drought mitigation *could* be caused by uncoordinated reservoir schemes. But inadequate mitigation could also be simply because the operating schemes at individual dams are insufficient. It's possible that more realistic operating schemes at the individual dams (i.e., not necessarily incorporating coordination, but with more realistic structure and/or parameterization) would provide the correct mitigation responses. It's also possible that the failure of these models to represent flood and drought mitigation is partly caused by bias in your inflow data (which is not shown or compared to observed at the upstream dam). Given these possibilities, the proposed framework fails to demonstrate unequivocally that a missing piece of the reservoir model is coordination.

Response 4: This comment echoes the reviewer's main concern above and further illustrates it. We thank the reviewer for taking the time to formulate it. It has been helpful in formulating our response to the reviewer's main concern (please see Response 1 and Changes to Manuscript 1).

Comment 5: Figure 1. Difficult to interpret due to color scheme used.

Response 5: We agree.

Manuscript changes 5: We amended the figure to improve the color scheme, and clarified the legend.

Comment 6: Table 1. Please provide some additional justification for these parameters. Has there been any validation done?

Response 6: the reviewer is right to point out that readers should know where the figures in Table 1 come from.

Manuscript changes 6: We added this in Section 2.4 (page 6 lines 2-14):

The general form of the reservoir rule was first presented by Proussevitch et al. (2013) and validated using the GRanD database (Lehner et al., 2011). Variants of this rule have been used with a daily time step on the Niger river basin (Oyerinde et al., 2016), and with large-scale assessments using WBM (Grogan et al, 2015; 2017; Zaveri et al, 2016; Liu et al., 2017). The fine-tuning of the parameters when establishing this version of the rule was made using a set of 22 large North-American and Eurasian reservoirs in offline mode, including the two largest reservoirs in the USRB (Palisades and American Falls, daily release NSE 0.70 and 0.60 respectively). Similar to what happens when a reservoir rule that classifies reservoirs by purpose is used in a large-scale model, we did not fine-tune the rule to each reservoir. This allows us to use the reservoir rule in conditions that are similar to what is done in most state-of-the-art hydrological models.

Comment 7: Why would the Rref for a water supply reservoir be 0.1? This parameter should vary widely across water supply reservoirs as a function of demand relative to inflow (and 0.1 is very low).

Response 7: Thanks for this question. According to the convention followed in WBM's release rule, water is withdrawn directly from water supply reservoirs, therefore the goal is to keep water in the reservoir unless it is close to being completely full.

Manuscript changes 7: We amended the sentence on the "water supply" use (page 10 lines 12-14)

"The  primary function for the``Water supply" use  is to keep releases minimal and storage maximal in order to maximize the quantity of water that can be drawn directly from the reservoir -- except for circumstances that require flood control at near-full storage."

Comment 8: Generally I don't see why the purpose of the reservoir would control these parameters. Rmin (and indeed the whole left-hand section of the operating curve) would likely be strongly determined by environmental flow requirements, which are independent of reservoir purpose. I would also suggest discussing the omission of seasonality in the operational parameters (which could apply to non-irrigation reservoirs). If there is no good justification for these rules then this is ok (it's not the purpose of your experiment to defend the status quo)â˘A˘ Tin this instance just make a statement to inform the reader.

Response 8: The idea that reservoir purpose controls release rule parameters is reproduced here because it is ingrained in the literature (see p2 lines 23-35 and p7 lines 11-14). We agree with the reviewer that no parametrization is perfect, be it in our paper or in the wider literature. However, and as the reviewer is right to point out, our goal is not to improve existing parameterizations. Instead, it is to diagnose them using a reservoir rule with several state-of-the-art characteristics.

Comment 9:

P6 line 1 - reservoir data *were* derived…

P6 line 1 - not clear who performed these manual updates

Manuscript changes 9: We corrected the typo and specified that authors made the manual updates (using the "we" page 6 line 5)

Comment 10: P6 line 7 - How accurate are these demand data compared to USGS estimates, which are US specific.

Response 10: The GPW is based on US census data, and is processed to a gridded resolution on par with our model grid, whereas USGS estimates are county level. We verified the difference between the two demand estimates and found it to be below 50%, in a basin where industrial and domestic demands are two orders of magnitude smaller than irrigation demands. Therefore, we deemed the GPW data to be acceptable for this study.

Comment 11: Section 2.1. I hadn't studied the Method of Morris previously, but this is an excellent description that educates the reader.

Response 11: We appreciate this kind comment.

Comment 12: P16 line 7 - The high variability in release (relative to observed) is surely just caused by the right-hand side slope of the operating curve being too steep, or the Sref being too low. Why should this point to incorrect representation of coordination?

Response 12: We are not sure whose reservoir's Sref or right-hand slope the reviewer is referring to. We are also not sure what they mean by "incorrect": in Section 4.2 we do not assess whether the representation is correct. Instead, as stated clearly at the top of the Section (p 15 lines 32-33), “Our analysis in Figure 6 focuses on flows from a single year to clarify the complex interactions between upstream releases and Minidoka operations”

In other words, in Section 4.2 and Figure 6 we use results from the method of Morris to demonstrate that release rule parameters upstream influence inflows, storage and release decision at Minidoka. In other words, we show that the assumption that this reservoir can be parameterized separately from the others does not hold here.

This means that the reservoirs should be parameterized jointly. And that would account for coordination.

Manuscript changes 12: We amended the last sentence of Section 4.2 to read: “Joint parameterization would explicitly account for coordination within reservoir operations, but would also require searching a parameter space of very high dimensionality.”

Comment 13: P16 line 27 - this may have been a low flow event, but your figure for Jackson shows clearly that the drawdown was in large part caused by sustained high release through 2013.

Response 13: There is no contradiction between the two statements. In fact, as can be shown in the Figure in this document (left panel), there was little water and high releases. This is explained by coordination with downstream reservoirs.

Manuscript changes 13: We inserted that figure (as Figure 6 page 17) to clarify this.

Comment 14: It would be helpful to see inflow, storage, and release graphs for the full time series for all reservoirs (perhaps in supplemental material).

Manuscript changes 14: We presented this data in supplementary material, using similar graphs as in the paper.

## Reviewer 2:

Comment 1: This article provides an evaluation of the consequences of the lack of a representation of reservoir coordination within a multi-reservoir system when simulating flood and drought events in large-scale hydrological models. The model Water Balance Model simulates a multi-reservoir system in USA. The model includes the representation of each reservoir operation policy (using predefined parameters according to each reservoir purpose) but it does not represent the coordination between reservoirs. The global sensitivity analysis Method of Morris is used to assess the effect of the parameterization to the model outputs. Authors conclude that the representation of reservoir policies independently is not enough and that, in addition, we need to capture reservoir coordination in large-scale models to properly simulate flood and drought effects.

Response 1: We thank the reviewer for their clear understanding of our paper, and more broadly for their thoughtful and comprehensive review.

Comment 2: The article is well written and structured. The Introduction makes a good review of the hydrological impacts of multi-reservoir systems and previous attempts in representing reservoir systems in hydrological models.

Response 2: We thank the reviewer for their kind words.

Comment 3: The methodology is well defined with the exception a few aspects that need further explanation. The article does not explicitly say where the parameters of the reservoir rules (Table 1) come from. Moreover, the authors do not specify the parameter ranges. If the values in Table 1 were obtained by calibration in a previous work, the authors could show the ranges applied in that calibration or just reference that work. If there is no previous calibration, how the predefined values of the parameters produce a good agreement between observed and simulated storages and releases (e.g. Figure 5) in normal climate conditions?

Response 3: Thanks for pointing out the need to clarify how the parameters were specified for the reservoir rules .

As for parameter ranges for parameter sampling, the methodology does specify and explain the rationale for this (p 12 lines 17-20), even though we acknowledge precisions should be given. Please note that parameters ranges given in the methodology are not informed from release rule calibration but by the necessity to design a diagnostic experiment whose results are easy to interpret. A major reason for this is that calibration ranges are usually not considered in studies using already-parameterized release rules.

Manuscript changes 3: we inserted in Section 2.4 a paragraph explaining the origin of the rule and of its current parameterizations, as well as the associated assumptions (p 8 lines 2-14; the paragraph can also be in response to Reviewer #1, under “Manuscript changes 6”).

We also explained why we import previously-calibrated parameter values and apply a sampling range that has more to do with the design of experiment than with parameter calibration. On one hand, (page 12 lines 24-26) we:

“aim to evaluate model outcomes in the same conditions as large-scale hydrological vulnerability assessments are carried out. For instance, we do not fine-tune reservoir rule parameters to individual basins, and instead use purpose-based parameterizations typical in large-scale studies (e.g., Biemans et al., 2011; Voisin et al., 2013a; Yoshikawa et al., 2014).”

On the other hand (page 13 lines 25-30):

“This analysis uses a range of plus / minus 10% around base values for all parameters in Table 1. These modest 10% ranges would be conservative if our focus were calibration and not diagnosis. Results (Section 4) will demonstrate that our narrow sampling yields quite substantial effects when compounded across the reservoir cascade in periods where coordinated operations are significant. Besides, there are two reasons for choosing the same range across all parameters: 1) it accounts for the fact that each parameter does not have the same base value across all reservoirs, and 2) it facilitates comparisons between different parameters' sensitivity indices.”

Comment 4: The results and discussion are also clear and well structured but there is a lack of discussion of how the methodology applied here can be used by others. I was wondering if this could be done using a different model where the parameters of the reservoir rules are unknown and need to be obtained by calibration.

Response 4: The goal of this paper is not to propose a calibration or parametric model identification framework on what constitutes a “good enough” representation of reservoir operations, but to shed light on what the unintended consequences of not representing coordination can be. This study contributes a diagnostic evaluation of the effects that can emerge with existing reservoir rule representations, if used within large-scale hydrological for flood and drought assessments.

Manuscript changes 4: The revised version has a subsection (3.1, start page 10 lines 28) in the methodology that clarifies our rationale in the design of the diagnostic analysis, and how the Method of Morris is only part of the analysis. It also clarifies that we are using the example of the USRB to draw “lessons can then be applied to understand the potential consequences of not considering coordination between reservoirs over larger study areas” (page 12 lines 22-23). This study is not intended to focus on more general aspects of parameter calibration or model identification..

Comment 5: Lastly, I think that the paper needs further and clearer discussion on why the lack of representation of reservoir coordination is most likely to be the main reason of this failure to simulate flood and drought events.

Response 5: We agree with this comment. We have written a detailed answer to the same comment by Reviewer 1 (please refer toResponse 1 to Reviewer 1)

Comment 6: In conclusion, this paper makes a relevant contribution to the growing discussion around the representation of reservoir systems in hydrological models and it has clear practical implications. The authors provide practical recommendations and possible solutions. While the representation of reservoir coordination is still very difficult to implement in models, this study highlights its importance and the need to, at least, consider this limitation when modelling catchments containing reservoir systems under extreme conditions.

Response 6: We would like to thank the reviewer for their kind words and accurate assessment.

Comment 7: While the authors provide a justification for the 10% range used for the sensitivity analysis, in my opinion, it would be interesting to also show what range of variation around base values should be applied to properly simulate any of the drought and flood events.

Response 7: As we discuss above, the intent of this study is not to correct the errors observed but rather, to diagnose them as an artifact of reservoir rule representations using a state-of-the-art assessment model example in an institutionally complex reservoir cascade. For this diagnosis, the 10% range is adequate in showing how the parametric effects across the reservoir cascade in the USRB are highly complex, interdependent, and dynamic.

Manuscript changes 7: where the ranges are introduced in the manuscript, we added explanations (page 13 lines 25-30) which we also quote at the end of “Manuscript changes 3” of this document.

Comment 8: Page 15, Lines 17-19: What if the releases were represented as cumulative releases, the sensitivity would be as consistent as for the storage?

Response 8: This is correct. However, if we used cumulative releases then at Jackson Lake we would have the exact same and opposite sensitivity indexes compared to storage. This information is captured in the current results.

Comment 9: Page 21, Lines 2-4: Could you please provide with a possible explanation to this unexpected result?

Response 9: Lines 5-12 provide this explanation.

Manuscript changes 9: We replaced “Instead” by the much more appropriate “This is because” at p 22 line 22 (formerly p 21 line 4)

## Reviewer 3:

Comment 1: The manuscript presents a new large scale reservoir operations model with generic operating rules associated with the reservoir main operational purpose such as flood control or irrigation, or default. The reservoir model stands out from equivalent models in that the releases are decided daily based on the daily storage level, shapes with combined log and exponential curves that accelerate the release in times of floods when close to full capacity and slows down the release in times of droughts with differences in the thresholds and propensity to release and store based on the purpose of the reservoir. The overall release is scaled by the long term mean annual flow. The model is implemented at high resolution ( .6 km, daily time step) over the Upper Snake River Basin, which is a snowmelt driven basin. The method of Morris is used to identify the reservoir release parameters that tend to be most influential in the reservoir release and storage variations throughout an 8 year period. Upstream reservoirs are used to evaluate the approach while downstream reservoirs are used to evaluate the impact of upstream reservoirs. A flood and drought events are evaluated with respect to observed operations to categorize the error associated with the lack of representation of reservoir coordination. Authors conclude that reservoir coordination is needed to represent flood and drought in typical reservoir models, and that optimization of rules with foresight would help in this endeavor. All simulations were performed on very high performing computational resources taking 2 days for 8 year simulation over the Upper Snake River Basin.

Response 1: Thanks for this detailed and accurate account of our work. One (very minor) clarification being that each simulation does not take 2 days (see line 9 on page 13).

Comment 2: The subject is very interesting for the HESS community and the manuscript is well written but there are a number of concerns that would need to be addressed before consideration for publication. The main concerns are about the two (great) highlights of the paper : the new model and the time sensitive analytics; i) the manuscript presents a new large scale model, with a very interesting concept for the releases that is however not enough evaluated and discussed, and ii) the approach to quantify the contribution of reservoir coordination to better represent floods and droughts needs to be improved – it is based on inference statements and the model could be modified to include information about upstream reservoir release to demonstrate the point about coordination.

Response 2: Thanks for this comment.

(i) Regarding the general concern on the model, we would like to point out that the hydrological model is not new, and neither is the rationale for the release rule. In fact, WBM is a well established representative of the broader class of large scale hydrological assessment models that are being used in regional to global applications, as highlighted in p2 lines 19-22 and p5 lines 18-29. As for the release rule itself, it also has a published record of application in WBM (Grogan et al, 2015; 2017; Zaveri et al, 2016; Liu et al., 2017).

What is more, the release rule we use is not put forward for its novelty but for being a state-of-the-art representative of the emerging types of rules that are being employed as reviewed in our Introduction. We highlight this when we detail the rule in Section 2.4 (see p. 11 lines 8-15)

(ii) It is not our intent to precisely quantify the exact contribution of reservoir coordination to overall prediction errors, but rather to use an institutionally complex multi-reservoir cascade that is known to exploit high levels of coordination to illustrate qualitative artifactual behaviors that can emerge from the absence of coordination in how large-scale hydrological models abstract major storages.

This being said, we agree that a new kind of release rule for large-scale hydrological models could take into account coordination. The goal of the paper is not to propose such a rule and demonstrate it, but to diagnose the issue to motivate future efforts in the research community to develop and evaluate new reservoir operation representations that better account for coordination.

Manuscript changes 2:

1. We give details on the provenance and previous uses of the release rule in Section 2.4 ( page 8 lines 2-14)

2. We clarified this goal of our paper in a revised version, in the methodology (bottom page 10, first sentence Section 3.1).

Comment 3: Minor feedbacks are that the reference to typical reservoir model is misleading and the analytics with the method of Morris is very hard to follow.

Response 3: We address the reviewer's concern within their point 1) below. As for the analytics, we appreciate the feedback from the 3 reviewers.

Manuscript changes 3: We made clarifications to interpretability as discussed in our responses here and for Reviewers 1 and 2.

Comment 4: Reference to typical reservoir operations model seems misleading. At the scale of the Upper Snake River Basin, typical reservoir operations models have a nodal architecture and represent accurate reservoir operating rules that can be revisited in optimization mode and especially in forecast mode to mitigate reservoir and drought events. The manuscript here refers to very large scale spatially distributed reservoir models that have been developed initially to be fully coupled with hydrology model and research land-surface-atmosphere interactions. Those models are typically applied over multiple independent large river basins. I would suggest to not refer to typical reservoir model where most of the community understand reservoir models where rules can be optimized and are applied to one basin at a time. Please refer to large scale distributed reservoir model or equivalent differentiation from nodal operational reservoir models.

Response 4:

We agree that had the goal been to propose the most accurate operating rules possible for this basin, we would have chosen very different rule forms. However, the goal of our study is to diagnose the implications for how state-of-the-art large scale hydrological assessment models represent reservoir operation rules. We understand that this point needs to be made even more explicit, and the revision will add a subsection to the methodology that will point this out.

As for the reviewer's concern with the geographical scale of our study, we would like to point out that release rules such as they are encoded in large-scale hydrological models must not lead to large qualitative errors in mid-size basins such as the USRB. Indeed, larger-scale studies would contain a set of basins that are not hydrologically connected with 1) a number of mid-size basins, and 2) large basins containing several headwater basins such as the USRB. Therefore, insights from a diagnostic approach on the USRB will be relevant to large-scale assessments if we apply the release rule in the same way (i.e., without over-parameterizing it).

For these reasons, the release rule is applied here as it would be within a large-scale hydrological model, that is, without fine-tuning the parameters to each individual reservoir.

Manuscript changes 4: We added a subsection in the methodology (3.1, bottom page 10) to detail our rationale and assumptions explicitly in a single location. Among other things, it explains why we use a single mid-sized basin, and why we do not tailor the rule to this basin (page 12 lines 20-26)

"The geographical extent of the USRB is small compared with that of areas traditionally considered in large-scale hydrological models. This is because a focus on a smaller area makes drawing quantitative and qualitative lessons from our diagnostic analysis easier. These lessons can then be applied to understand the potential consequences of not considering coordination between reservoirs over larger study areas. This is also why we do not fine-tune reservoir rule parameters to individual basins, and instead use purpose-based parameterizations typical in large-scale studies (e.g., Biemans et al., 2011; Voisin et al., 2013; Yoshikawa et al., 2014)."

Comment 5: A new large scale reservoir operations model : please provide more details

Response 5: We would like to clarify that our paper is not focused on a "large scale reservoir operation model". In fact our point is quite the opposite: pointing out the limitations of standard rule-based representations in hydrological models that consider and parameterize reservoirs separately as reviewed in detail in our Introduction.

Manuscript changes 5: The revision explicitly reminds the reader that the reservoirs are not coordinated in the model, but only in real-life operations, when introducing the methodology (page 10 lines 29-32):

"To diagnose the potential consequences of non-coordination within large-scale models' reservoir release rule, our dual objective is to (A) understand how these consequences play out through space and time in a large-scale hydrological model and (B) compare and contrast with the impacts of real-world coordinated reservoir operations."

We also provide evidence of cooperation within Section 4.1, with Figure 6 and Section 4.1.2, both on page 17.

Comment 6: what is the river routing process for this high spatial resolution and daily time step? A recommendation in the introduction is not to aggregate reservoir storage but many reservoirs have less than 2 days in travel time. How does the reservoir model decision release algorithm adjust stability?

Response 6: Thanks for this. We reply to each sentence separately and in order:

1. Similar to the above, we would like to point out that the release rule presented here does not consider what happens downstream, where the river routing is done by the hydrological model (WBM).

2. Likewise, in the introduction's discussion of Shin et al (2019) (p 2 lines 5-7 original submission) we did not make any prescriptive recommendation related to aggregating reservoir storages. It is part of the second paragraph in the introduction whose role is to set context, which is the evolution of hydrological models towards higher resolution applications. Our work looks at what happens with a model that is relatively highly-resolved, but far from hyperresolution. At the resolution we use though, it makes little sense to aggregate storages because they are less than two days apart: this would negate the sought advantages of a higher resolution..

3. We apologize but we are not sure what you mean by "adjust stability" here. This said, the reservoir rule mechanisms are completely specified in Section 2.4.

Comment 7: How are the 6 parameters initialized? Are the necessary data widely available? What are the assumptions?

Response 7: Thanks for pointing out our lack of explanation of where the parameters of the reservoir rules come from. As for the assumptions, we would like to point out that the key one, common to release rules for large-scale hydrological models is that by construction, each reservoir gets separate parameters that do not depend explicitly on the behavior of reservoirs upstream.

Manuscript changes 7: In a revised version, we inserted in Section 2.4 a paragraph explaining the origin of the rule and of its current parameterizations, as well as the associated assumptions (p 8 lines 2-14; the paragraph can also be in response to Reviewer #1, under "Manuscript changes 6").

The key assumption that this rule (similar to other rules) assumes non-coordinated reservoirs has also been made explicit in Section 2.4 where the release rule is introduced page 8 lines 30-32):

"Note that this parameterization, similar to those of other state-of-the-art rules in large-scale hydrological models, do not account for possible coordination mechanisms in multi-actor, multi-reservoir systems."

Comment 8: Evaluation of the smoother release curves with other models. In other equivalent models that are cited (Hanasaki, Doell, Biemans, Voisin, etc) , releases are decided daily based on reservoirs minimum and maximum capacities, minimum environmental flow and tend to follow monthly storage and releases targets with no foresight, but using long term mean monthly inflow, which also tends to be regulated or natural flow depending on the models. What is the improvement for those rules? The obvious features are the changes in release rates - how does it improve the flow representation in general?

Response 8: The reviewer cites four release rules from four papers (Hanasaki et al (2006), Doll et al (2003), Biemans et al (2011), Voisin et al (2013)) that have release rules that decide a monthly release target in order to analyse outputs with a monthly time step. Given this, it makes sense for the rules to be based on monthly parameters.

Yet, as detailed in our introduction (p3 lines 15-35) the transition of applications to include flood concerns means that shorter timescales have to be considered. The smooth monthly curves are not appropriate for finer time scaled extremes (floods) and for large storages those extremes carryover in the effects on water operations for droughts. So in short, there is tension and difficulty in resolving floods and droughts in complex cascades like USRB using standard rule forms and assumptions of independence between reservoirs.

Please note that we do not try to use a "best" rule but a rule that shares some key characteristics with state-of-the-art available rules (and we would urge the reviewer to refer to newer representations, e.g. see references in p3, lines 30-35, rather than those they cite here).

Manuscript changes 8: The revision clarifies that the version of the Hansaki et al rule that is used for flooding is modified by Mateo et al (2014) at p3 line 16 to be used with daily time step (See p 3 lines 18-19)

Comment 9: reservoir coordination. Note that the use of a rolling past 20-year of mean monthly regulated inflow provides a minimum of reservoir coordination mostly during extreme events. "Some" coordination is represented through the use of mean monthly regulated flow and also the allocation of water demand to a number of reservoirs based on how full they are. This feature is not present in this model representation, and would likely not drastically change extreme events. Yet it does represent "coordination" around releases and other water management performance metrics than flow and storage, rather coordination on meeting basin-scale water demand. There were statement throughout the paper saying that there was no coordination at all, which seemed then inaccurate and should be clarified.

Response 9: We thank the reviewer for this comment. It underlines that there are different interpretations of what coordination means. In this paper, we focus on active forms of coordination by human operators, where the dynamics across space and time in release decisions occur in a way that cannot be explained by the immediate or short-term hydrological or climate conditions. A key feature of coordination as understood in this paper is that upstream reservoirs react to and anticipate downstream water issues.

By contrast, the reviewer seems to focus in this comment on two forms of passive coordination, in which reservoirs adjust to varying on-site conditions that are typically imposed on downstream reservoirs by upstream operations. These conditions unfold concurrently to the release decision (inflows influenced by upstream water management, or at-site withdrawals).

Side note on inflows as a means of coordination: the reservoir model uses long-term (5 year) mean (regulated) inflow (see page 8 lines 2; Figure 3 on page 9), and the paper illustrates that this may not be sufficient to represent coordination

Manuscript changes 9: We clarified the difference between our paper's focus (active coordination) and what is common in release rules in large-scale hydrological models (passive coordination) in the introduction by defining more precisely what we mean by non-coordination (page 4 lines 4-5): "consequences of a release decision on downstream reservoir levels (and management objectives) are not considered"

Comment 10: evaluation of the model and transfer to other regions: âA ˘ c whether the coordination ´ between reservoirs was represented or not, how does it affect the vulnerability metrics at the scale of the basin, which is what those models were initially developed for?

Response 10: We thank the reviewer for keeping their eye on the end-goal here. We agree that large-scale hydrological models are increasingly used to assess water vulnerability. Vulnerability metrics are useful when the difference between a situation and the other is quantitative, but this work is striving to highlight qualitative differences: an adverse event vs. its absence.

This being said, we agree that it is important to put the results into context. For instance, we do explain the consequences of losing control of regulating downstream releases at American Falls in a low-flow event (p 16 lines 29-32).

Manuscript changes 10: We clarified that emptying American Falls instantly disrupts irrigation schedules (and forces farmers to watch their crops wither), by amending the sentence (page 20 lines 5-7) that now reads:

“Therefore, it is key that American Falls, the main reservoir in the Snake River plain located just upstream of Minidoka, is not empty so that it can keep regulating water levels in downstream reservoirs, ensuring irrigation needs are met.”

As for the 2011 flood event, we added that “The simulated peak is almost 40% higher than the highest observed daily discharge over the past 40 years. Maximal Palisades release over this period -- 1140m3/s on the 20 June 1997 -- corresponds to a flooding event which led to six counties declaring a state of disaster, leading to over USD 11 million in relief by the federal U.S. government.”

Comment 11: Most of those models have been developed for application to a wide range of ´ climatology conditions. The model here is applied to a relatively very small basin for its kind. If this manuscript will be used as reference for this large scale reservoir models, it should be either evaluated with respect to other generic rules, or the applicability to larger regions and very different regions should be presented.

Response 11: Thanks for this comment. Indeed, we are using an example of generic rule that shares key common characteristics with other generic rules; these rules are usually applied to larger hydrological areas.

As explained in a previous comment, it is important to point out what the absence of coordination may lead to if using these rules in headwater basins that would be part of larger-scale studies with large-scale hydrological model.

Manuscript changes 11: We clarified this reason for zooming in to a midsize basin in the methodology section (Section 3.1, page 12 lines 20-26)

Comment 12: Evaluation of the contribution of reservoir coordination – artifact of the model? - the main assumption is that the daily releases are based on storage only. All other equivalent models used an estimate of the expected monthly inflow. The main conclusion of the paper is that the coordination between reservoirs should be represented. While I do believe in this conclusion, it seems that the reference to “typical reservoir model” is not justified if the monthly inflow (proxi for foresight without forward running all the models involved) is not represented at all like in other models. My recommendation would be to modify the experiment to evaluate perhaps incremental and simple levels of coordination (

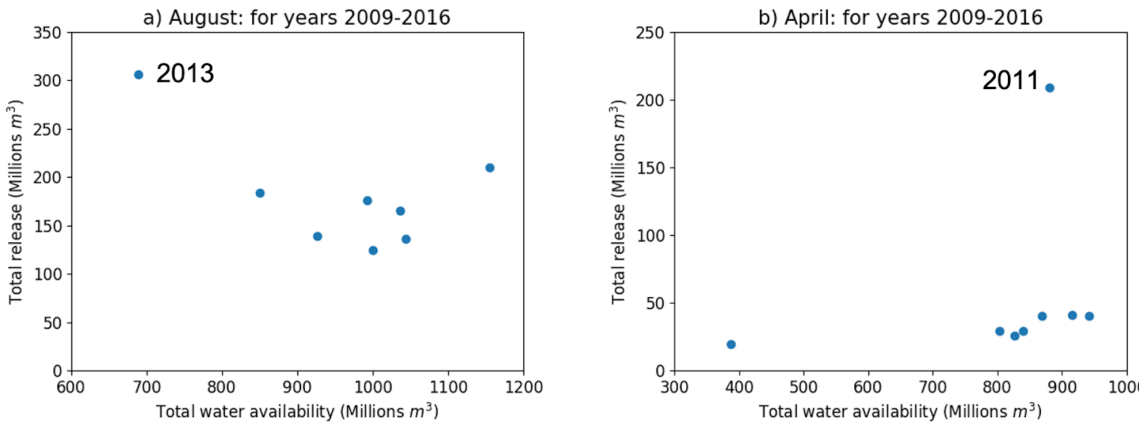
aka adding inflow as parameter for the decision release, or a proxi for inflow) to complement the interpretation of the results and provide more quantifiable statements.

Response 12: We do agree that models that schedule release over a monthly time step also use inflows over a monthly time step. Then, release is naturally a function of available water (beginning-of-month storage plus monthly inflow) makes more sense than determining release as a function of storage alone. Likewise, daily release decisions should be a function of beginning-of-day storage and daily inflows. In WBM, this is implicitly the case because releases are actually calculated on an hourly time-step, enabling the model to account for day-to-day inflow rate variations as the day progresses; this produces a daily release total.

Note that this rationale of taking the same time step for inflows and releases is common to most release rules for large-scale hydrological models, up to the most recent rules (e.g. Yassin et al, 2019, in this journal).

Beyond this theoretical reasoning, we conducted another check for added confidence that the results are not an artefact of using monthly expected inflows. In the two events examined in this paper, we produced the Figure 1 below, which plots historical monthly release (y-axis) as a function of available water (storage + monthly inflows) for all years of the modeled period 2009-2016 at the most upstream reservoir (Jackson Lake). Results show that historical operations could not be replicated simply by incorporating inflows (even monthly inflows) to the release rule.

For the drought in August 2013 (Section 4.3; left panel) as for the flood in April 2011 (Section 4.4; right panel), the release response is unusual. In the 2013 drought, the reservoir has already been emptied to rescue American Falls levels downstream but releases are still the highest in 8 years, whereas in the 2011 flood, releases are around 8 times as high whereas water availability is similar for 6 of the other 7 years.




Manuscript changes 12 : We clarify in Section 2.4 that the rule is separately implemented at the hourly time step in the model, to assimilate inflow into outflow calculations, in order to produce a total release at the daily time step in the model (page 8 lines 5-6):

“WBM's reservoir module operates on a hourly time step to closely follow storage variations and yield a daily release total.”

We inserted this Figure in the revised version (as Figure 6 on page 17), assorted with an explanation that highlights how the anomalous releases (in August 2013 and April 2011 respectively) correspond to coordination that could not be replicated by simply considering expected monthly inflows (Section 4.1.2 on page 17)

<u>Comment 13:</u> Evaluation of the contribution of reservoir coordination during extreme events I found it extremely hard to follow the text and interpretation of the drivers of the release (annual flow versus objective of this reservoir or upstream reservoirs, and shape of release) by just looking at the figures. Most of the text describes the observed operations and coordination and how the model does not capture it. It is unclear how the method of Morris helps with the interpretation during extreme events. While the visualization is very nice to show the data, it seems that those figures could go in the supplemental material and another figure that compiles those time series and support the text would help.

<u>Response 13:</u> Thanks for this comment that will help us clarify our manuscript. We believe that we were not explicit enough in explaining the rationale for the method of Morris.

<u>Manuscript changes 13:</u> In the revised manuscript, we explain in a new Section 3.1 (methods) that when we use the method of Morris (page 12 lines 1-5):

“The focus is not solely  on which parameters in the release rule are most influential in regulating flows, but importantly clarifying how the set of dominant controls on water flows and reservoir storage levels evolves along a complex multi-reservoir cascade through time. Parametric sensitivities are then used alongside storage and release trajectories for the simulated ensemble to assess how sets of dominant controls in a point in time and at a given reservoir can be associated to high- or low-flow conditions.”

<u>Comment 14:</u> Overall discussion and recommendation versus computational resources needs The authors conclude that foresight should be represented, which is also very sound. Yet the computational resources brought forward for such a relatively small basin are huge which decrease the feasibility at a continental or global scale. Optimization also bring other uncertainties and more computational needs. While the authors seem to indicate that this is what we should do, those were actually drawbacks and motivation for developing those large scale generic models. The recommendation is confusing and perhaps the authors could provide a clarification on new model performances to make it possible now? Please also note that nodal models that typically support reservoir operation optimizations do not provide the spatially distributed feedback into the hydrology model to represent hydrology-land-surface-atmosphere interactions. Maybe the authors meant that we need different types of large scale reservoir models? This would be very sound – just need to be clearer about recommendations then.

Response 14: We fully agree with the reviewer that there is a research challenge ahead in terms of incorporating human complexity into large-scale hydrological models. We recognize that a first step in addressing this challenge is to signal the consequences of ignoring it, and that is what our paper is trying to do. Clearly, it does not try to provide an easy fix, and while the discussion outlines some ways forward, we fully acknowledge that solutions are not going to be easy to design and implement.

# Coordination and Control: Limits in Standard Representations of Multi-Reservoir Operations in Hydrological Modeling

Charles Rougé[1,2], Patrick M. Reed[2], Danielle S. Grogan[3], Shan Zuidema[3], Alexander Prusevich[3], Stanley Glidden[3], Jonathan R. Lamontagne[4], and Richard B. Lammers[3]

[1]Department of Civil and Structural Engineering, University of Sheffield, Sheffield, United Kingdom
[2]Department of Civil and Environmental Engineering, Cornell University, Ithaca, NY, United States
[3]Water Systems Analysis Group, University of New Hampshire, Durham, NH, United states
[4]Department of Civil and Environmental Engineering, Tufts University, Medford, MA, United States

**Correspondence:** Charles Rougé (c.rouge@sheffield.ac.uk)

**Abstract.** Model-based risk assessment of hydrological extremes needs to consider the interactions between the many stakeholders in a river basin as well as the institutions and regulations that mediate them. Unfortunately, commonly employed representations of human-operated structures in hydrological models are limited in their ability to capture human-mediated coordination and control actions in complex river basin systems. This study contributes a detailed diagnostic analysis of the parametric controls and their effects in standard reservoir representations in flood and drought modeling. Our diagnostic analysis uses the Water Balance Model (WBM), which features detailed representations of the human infrastructure coupled to the natural processes that shape water balance dynamics. Our analysis focuses on challenges posed by human-mediated coordination and control actions using the multi-reservoir cascade of the Upper Snake River Basin (USRB) in the Western U.S. We employ a time-varying sensitivity analysis that utilizes Method of Morris factor screening to ~~quantify how the parametrizations of the reservoir release rules impact modeled flows throughout the USRB~~track how the release rule parameters that control reservoir storage and release evolve 1) along the cascade, and 2) in time according to seasonal high- and low-flow events. We combine this with a comparative analysis of historical operation and targeted experiments with simple offline reservoir water balance models. Our results demonstrate the importance of understanding the state-space context in which reservoir releases occur and where operational coordination plays a crucial role in avoiding or mitigating water-related extremes. Understanding how major infrastructure is coordinated and controlled in major river basins is essential to properly assessing future flood and drought hazards in a changing world. This implies that the validation of hydrological models for this purpose should move beyond the usual goodness-of-fit checks of outlet flows to incorporate an assessment of the actual emergency response operations used to mitigate hydrological extremes.

## 1 Introduction

The cumulative impacts of reservoir cascades on river flows has been recognized and demonstrated worldwide by early global hydrological models (Dynesius and Nilsson, 1994; Vörösmarty et al., 1997). Since then, these findings, frequently corroborated in the literature (e.g., Nilsson et al., 2005; Adam et al., 2007; Döll et al., 2009; Biemans et al., 2011; Grill et al., 2019), have taken a new significance with the planned or ongoing construction of more than 3,700 major dams, most of them in the global South (Zarfl et al., 2015). This new wave of dam construction cements the role of man-made reservoirs as key actors on the global hydrological cycle. A striking illustration of this fact is the cumulative consequences of building multiple dams on river flow regimes, ecosystem benefits, or sediment transport in previously relatively undammed major river basins such as the Amazon (Latrubesse et al., 2017; Timpe and Kaplan, 2017) or the Mekong (Schmitt et al., 2018).

In parallel, and as a response to evolving flood and drought risks in a changing world, communities involved in large-scale hydrological modeling aim to address the challenges posed by representing, monitoring, and forecasting these risks at fine resolutions in both space and time (Wood et al., 2011; Bierkens, 2015). For high-resolution modeling of multiple reservoir systems, reservoirs should not be lumped together, but rather, their individual impacts on system dynamics should be carefully accounted for (Shin et al., 2019). In this context, better representations of how human societies (mis-)manage their water resources needs to be integrated in these models (Wada et al., 2017), especially since currently state-of-the-art models yield mixed results for the modeling of monthly extremes (Zaherpour et al., 2018). There remains opportunities for research to determine which aspects of human management are most urgent to integrate in standard reservoir representations. One such aspect is coordination between reservoirs, long-recognized as a key aspect of water management (e.g., Lund and Guzman, 1999). A recent intermodel comparison by Masaki et al. (2017) found discrepancies between models when representing flows across large reservoir cascades. This echoes an earlier study that found deteriorating goodness-of-fit of monthly releases along such cascades (Adam et al., 2007).

The present study uses a diagnostic analysis of commonly-employed parametrizations for reservoir release decisions to evaluate their implications within multi-reservoir cascades that are critical for managing floods and droughts. We focus on a highly resolved model of the Upper Snake River Basin (USRB) – 30 arc seconds spatial resolution for an average grid cell of about 0.6 square kilometer, and a daily time step – that encompasses a total of 128 reservoirs in the Western U.S. Our model-based representation of the USRB exploits the Water Balance Model (e.g. Wisser et al., 2010), which is well-suited for regional or global scale hydrological assessments (e.g. Wisser et al., 2008; Grogan et al., 2017) and includes a representation of human impacts on the water cycle. The remainder of this introduction reviews reservoir representations in hydrological models, including their use for flood and drought modeling, and the key goals of our contributed diagnostic assessment.

Early attempts at representing man-made reservoirs modeled them as natural reservoirs (i.e., lakes; Meigh et al., 1999; Coe, 2000; Döll et al., 2003). In 2006, representations proposed separately by Haddeland et al. and Hanasaki et al. introduce the idea that man-made reservoirs should have a distinct parametrization that reflect the reservoir's purpose, leading to two different kinds of reservoir representations (Nazemi and Wheater, 2015). Haddeland et al. (2006) optimize release for the upcoming year assuming future inflows are known, and following a management objective in line with the reservoir's primary purpose. This optimization-based scheme has been extended in several studies, most notably van Beek et al. (2011) who replaced perfect foresight of the next year's inflows with an uncertain forecast (for other improvements, see also Adam et al., 2007; Wada et al., 2014). Alternatively, Hanasaki et al. (2006) propose a parametrization that simulates releases based on a set of site-specific parameters such as long-term average inflow, reservoir capacity and beginning-of-year storage, and downstream water demands. There exist several refinements of this rule, by changing the definition of what constitutes downstream demand (Döll et al., 2009), by considering more reservoir purposes (Biemans et al., 2011; Yoshikawa et al., 2014), by allowing the reservoir's primary purpose to vary seasonally (Voisin et al., 2013a), or by proposing a general rule differentiating refill and drawdown seasons for large multipurpose reservoirs (Wisser et al., 2010).

This first generation of reservoir representation has led to improved simulations of historically observed discharge at the monthly time scale (Pokhrel et al., 2012; Li et al., 2015; Veldkamp et al., 2018). It has been integrated into increasingly complex modeling frameworks. For instance, the rule first proposed for a global flow routing model by Hanasaki et al. (2006) has been integrated as part of the global hydrological model H08 (Hanasaki et al., 2008), which then has been integrated into a land surface model that models the carbon, energy and water cycles (Pokhrel et al., 2012). Similarly, the rule proposed by Voisin et al. (2013a) has been incorporated into increasingly complex modeling frameworks accounting for regional-scale feedbacks between climate, socio-economic systems and heavily managed water, energy and food systems (Voisin et al., 2013b; Kraucunas et al., 2015). As the models including these reservoir representations have grown more complex, so have the questions asked of them. Applications typically include assessments of past and present water withdrawals, human impacts on hydrology, water stress and scarcity ~~(e.g., Biemans et al., 2011; Wada et al., 2011, 2014; Yoshikawa et al., 2014; Hanasaki et al., 2018; ?; Meza et al., 2019)~~ (e.g., Biemans et al., 2011; Wada et al., 2011, 2014; Yoshikawa et al., 2014; Hanasaki et al., 2018; Liu et al., 2019; Meza et al., 2019) . Recently, modelling frameworks have been extended to include water quality (Wanders et al., 2019) or economic appraisals of the consequences of scarcity (Bierkens et al., 2019). These models are also increasingly being used for appraisals of future water scarcity under integrated socio-economics and climatic scenarios (e.g., Hanasaki et al., 2013; Hejazi et al., 2015; Jägermeyr et al., 2016; Herbert and Döll, 2019).

Reservoir management is also critical for understanding flooding, where simulations must resolve much finer timescales (i.e., daily or shorter). Reservoir rules like those of Hanasaki et al. (2006) can be modified to be applied with a daily time step to investigate the potential of reservoir management to alleviate flooding (Mateo et al., 2014), or be modified to better model the periods when reservoirs are nearly full (Shin et al., 2019). Large-scale or global flooding assessments are made more complex by the fact that hydrologic routing by itself is insufficient for floodplain modeling (e.g., Sampson et al., 2015; Schumann et al., 2016). In this context, a good first approximation to account for reservoirs is to allocate flood storage capacity following an extreme precipitation event, especially since this alone can dramatically alter a flood's outcome (Metin et al., 2018). Yet, subtle changes in flood management by reservoirs can have decisive impacts, both in theory (Najibi et al., 2017), and in observed catastrophic flooding events like in Kerala (Southern India) in 2018 (Mishra et al., 2018). A finer assessment of the capacity of reservoirs for flood management involves explicit consideration of the multipurpose nature of reservoirs, as they often are assigned flood control duties on top of other uses. To achieve this, the representation proposed in LISFLOOD by Burek et al. (2013) partitions storage into different compartments; Zajac et al. (2017) demonstrated the merits of this formulation for flood impact assessment at the global scale.

Similar to Burek et al. (2013), several representations of varying complexity have been proposed to divide active storage capacity into several compartments, both to obtain sensible operations at submonthly time steps and to account for the fact that most reservoirs are inherently multipurpose installations (Wu and Chen, 2012; Zhao et al., 2016; Wang et al., 2019; Yassin et al., 2019). Another way to account for the complex nature of operations at a daily time step has been to directly emulate observed operations using machine learning techniques (Ehsani et al., 2016; Coerver et al., 2018). Both types of approaches have also been implemented in search for representations that can adapt to evolving climate conditions (Ehsani et al., 2016; Zhao et al., 2016). Thus, Ehsani et al. (2017) demonstrated the role of reservoir storage in alleviating the impacts of both floods and droughts under a changing climate in the Northeastern U.S..

It is worth noting, however, that all of the reservoir representations discussed above do not account for coordination within multi-reservoir systems. In other words, consequences of a release decision on downstream reservoir levels (and management objectives) are not considered. To date, there has not been a carefully designed diagnostic model evaluation of the implications of errors in representing actual human coordination and controls in high-impact, complex river basin contexts. This study links observed operations for recent high- and low-flow events in the USRB's reservoir cascade to clarify how standard representations of release rules capture the underlying coupled human-natural processes that are critical to model-based assessments of our vulnerabilities to extremes. The diagnostic model evaluation approach used in this work employs time-varying sensitivity analysis (e.g., Reusser and Zehe, 2011; Herman et al., 2013b; Guse et al., 2014; Pianosi and Wagener, 2016; Lamontagne et al., 2019; Quinn et al., 2019). We explicitly map how reservoir rule parameterizations relate to the qualitative as well as quantitative impacts of model behavior across the successive reservoirs within the USRB cascade at a daily time-scale. Building on prior successful diagnostic model evaluation studies, our sensitivity analysis is based on the factor screening capabilities of the Method of Morris (Morris, 1991; Campolongo et al., 2007), which requires significantly less computation time than other methods while providing high-fidelity measures of model controls (Herman et al., 2013a; Iooss and Lemaître, 2015).

This diagnostic assessment exploits the Water Balance Model with a simulation-based parametric release rule ~~that is a variant of that proposed by~~ ~~Oyerinde et al.~~ ~~(2016) for a key reservoir in the Niger river basin~~introduced by Proussevitch et al. (2013) and incorporated to WBM in several recent large-scale assessments (Grogan et al., 2015, 2017; Zaveri et al., 2016; Liu et al., 2017) . This representation is state-of-the-art in its ability to reproduce the climatological daily water balance of a single reservoir over the year with high accuracy. The possibility to use different parametrizations depending on the reservoir's perceived use and behavior, and the fact that release behavior qualitatively depends on storage level, are all features that capture the advanced reservoir representations currently in use in other models. ~~It is also representative of the typical~~ Note that we do not seek to validate this release rule, but rather, to use it as a typical example of release rule abstractions in large-scale hydrological models, in that it does not feature direct coordination between reservoirs' release decisions.

The rest of this work is structured as follows. Section 2 presents the study area and model used for the analysis, including a detailed explanation of the reservoir rule. Section 3 introduces and justifies the methodological aspects of the analysis. Section 4 presents the results from the diagnostic approach. Section 5 and 6 discuss the implications of our findings as well as our overall concluding remarks.

## 2 Study area and model

### 2.1 The Upper Snake River Basin

The Snake River originates east of the Teton Range in western Wyoming, then crosses the mountains into the Snake River Plain in southern Idaho. After flowing west through the entirety of that plain, it flows north to join the Columbia River. This work focuses on the Upper Snake River Basin (USRB; Figure 1), which has a drainage area of about 92,000 sq. km and is characterized by a snow-dominated, semi-arid climate. To ensure water availability for the whole agricultural season, the U.S. Bureau of Reclamation has built and operated a network of reservoirs, canals and lateral distribution ditches since the early twentieth century (U.S. Bureau of Reclamations, 2012). Since then, a diverse array of demands, including hydropower, irrigation, ecological conservation, and downstream water allocation, has increasingly required the USRB to be extensively managed with a network of dams of a broad range of sizes: 128 reservoirs of over 10 $hm^3$ (10 million $m^3$) throughout the basin, for a total volume of 6.93 $km^3$. Its waters are over-allocated across the USRB's competing demands (McGuire et al., 2006). The over-allocation is at least partially the result of historical perceptions of water availability where the twentieth century was wet compared with previous centuries (Wise, 2010). In fact, water availability is decreasing (Ahmadalipour and Moradkhani, 2017), forcing farmers to adapt to drier conditions (Hoekema and Sridhar, 2011). These drying trends are expected to worsen with climate change, especially as this will be accompanied by an increasing mismatch between seasonal patterns of water availability and use (Hamlet and Lettenmaier, 1999; Rauscher et al., 2008; Wise, 2012).

The USRB is also vulnerable to rain on snow events that can lead to extreme flooding. These are a common occurrence in the wider U.S. Northwest and are expected to get worse in the future (Musselman et al., 2018). In the USRB, a historically significant flood that caused widespread damage occurred in February 1962, with rainfall on frozen ground following a particularly cold spell (Thomas and Lamke, 1962). This was despite the recent completion of the Palisades Dam giving the Minidoka Project a significant ability to coordinate storage capacity for both water supply and flood control. Following this event, the USRB was also the site of the Teton Dam failure in 1976 (Independent Panel To Review Cause of Teton Dam Failure, 1976). All of these characteristics – heavy reliance on institutionally coordinated reservoir management in a drought- and flood-prone area that has experienced the consequences of dam failure, and where water extremes are expected to get worse with climate change – make the USRB a particularly relevant basin to study the representation of reservoirs within large-scale hydrological models.

### 2.2 The Water Balance Model

The University of New Hampshire Water Balance Model (WBM) is a process-based, modular, gridded hydrologic model that simulates spatially and temporally varying water volume and water quality across a wide range of spatial domains from global half-degree grid cell resolution (e.g., Grogan et al., 2017) to local 120 m grid cells (Stewart et al., 2011). WBM represents all major land surface components of the hydrological cycle, and tracks fluxes and balances between the atmosphere, above-ground water storage (e.g., snowpack, glaciers), soil, vegetation, groundwater, and local runoff. A digitized river network connects grid cells enabling simulation of flow through the river and groundwater systems and for simulating water temperature (Stewart

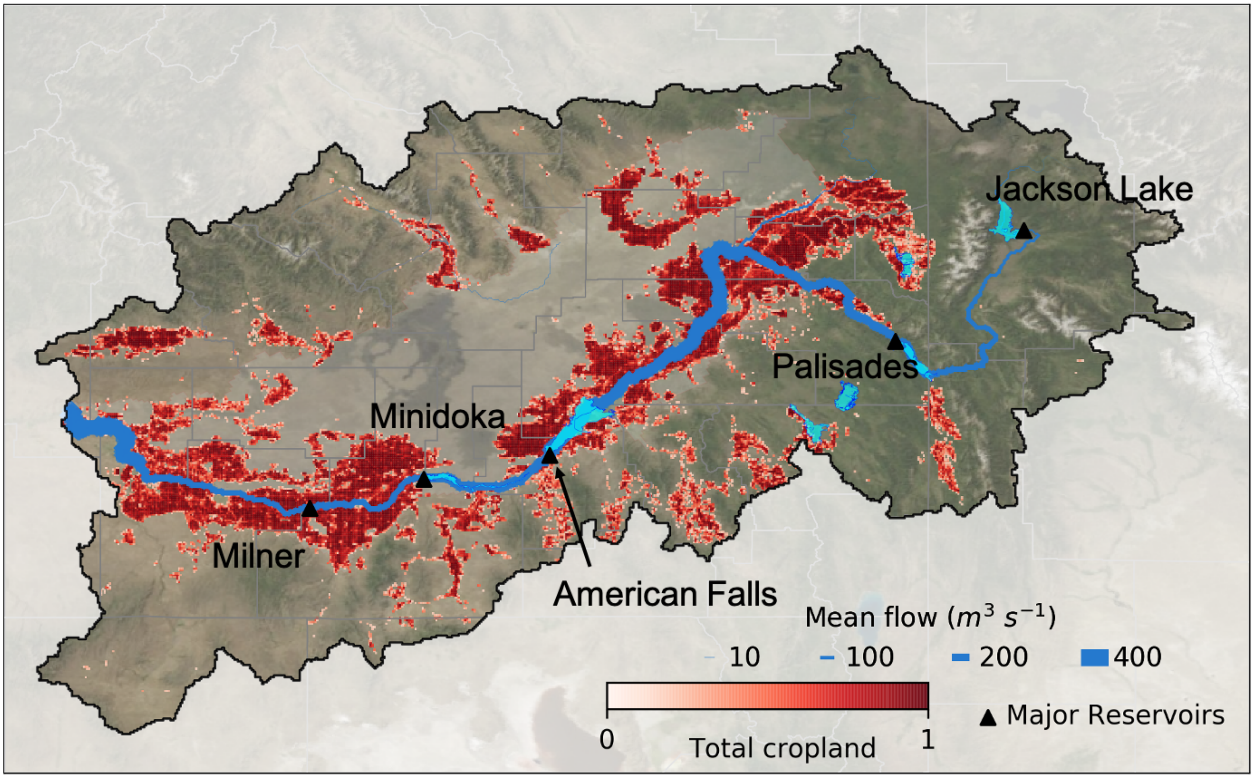



**Figure 1.** Upper Snake River basin (USRB) with the five reservoirs on the main stem of the Snake River.

et al., 2013). Direct human influences on the water cycle include domestic, industrial, and agricultural (irrigation and livestock) water demand and use, and the impacts of impervious surfaces. WBM accounts for the operation of dams and reservoirs (Wisser et al., 2010), inter-basin hydrological transfers (Zaveri et al. 2016), and agricultural water use from irrigation (Wisser et al., 2010; Grogan et al., 2015; Grogan, 2016; Zaveri et al., 2016; Grogan et al., 2017). Additionally, new WBM modules have been developed recently to include the use of sub-grid elevation band distributions derived from a high-resolution elevation dataset to improve handling of snowpack in mountainous regions.

### 2.3 WBM representation of the USRB

A drainage network of the USRB that covered an area of 92,900 $km^2$ (compared to USGS's estimate of 92,700 $km^2$) was developed at a spatial resolution of 30-arcseconds (approximately 780 $m$) based on HydroSHEDS (Lehner et al., 2008) corrected to better represent drainage as mapped by the US Geological Survey's National Hydrography Data (nhd.usgs.gov). Reservoir data ~~was~~ were derived from the National Inventory of Dams (nid.usace.army.mil)~~following manual updates that included dam additions, and updates to~~. We manually added dams and updated reservoir capacities, locations, and upstream drainage areas. WBM simulations used gridMET (Abatzoglou, 2013) for contemporary precipitation and temperature, and Modern Era

Retrospective-Analysis for Research and Applications, version 2 (MERRA2; Gelaro et al., 2017) for open water evaporation, windspeed, relative humidity, leaf area index, and albedo to calculate potential evapotranspiration following Monteith (1965). Snow accumulation and melt followed the temperature-index based formulation of Willmott et al. (1985). Human population density, which controls both domestic and industrial water demand, came from the Gridded Population of the World (GPW) dataset (Center For International Earth Science Information Network-CIESIN-Columbia University, 2016). WBM used Food and Agricultural Organization (FAO) estimates of livestock density for cattle (Steinfeld et al., 2006) at 5 minute resolution following Wisser et al. (2010). These data compared favourably with the U.S. Department of Agriculture's (USDA) National Agricultural Summary Statistics (NASS) for 2005, but exhibit more realistic spatial variability than county-level averages. USDA Soil SURvey GeOgraphic (SSURGO) data parameterized available water capacity for the USRB.

WBM uses an adaptation of FAO's Irrigation and Drainage paper (Allen et al., 1998) to estimate crop water requirements based on potential evapotranspiration, soil moisture, and a crop coefficient ($k_c$) defining a particular crops' water use efficiency. Details regarding the crop water demand calculations are provided in previous works (Grogan et al., 2017; Wisser et al., 2010). This study used the US Department of Agriculture's Crop Data Layer (CDL) estimates of crop types (and land cover) at 30 m resolution (Han et al., 2012), aggregated by surface area averaging and remapped to a consistent group of crops as monthly irrigated and rainfed crop areas (MIRCA) crops (Portmann et al., 2010). We applied $k_c$, planting dates, and crop depletion factors from MIRCA to the CDL crop fractions. Open water, impervious area, and wetland data also came from CDL data. Process based representation of irrigation technology was recently introduced to WBM following key aspects of the formulation of Jägermeyr et al. (2015). Irrigation technologies used in the USRB varied by county following Maupin et al. (2014) and Dieter et al. (2018). Additional details regarding the specific implementation of irrigation technologies will be reported in a separate paper.

### 2.4 Reservoir representation within WBM

WBM's release rule for managed reservoirs expresses daily release as a fraction of long-term (five years or more) mean release at the reservoir as illustrated in Figure 3. This is a refined convention for release rules within hydrological models (Hanasaki et al., 2006; Wisser et al., 2010) to be primarily controlled by instantaneous reservoir storage and purpose rather than statistics on the probability distribution of inflow rates. ~~In~~ WBM's reservoir module operates on a hourly time step to closely follow storage variations and yield a daily release total. The general form of the reservoir rule was first presented by Proussevitch et al. (2013) and validated using the GRanD database (Lehner et al., 2011). Variants of this rule have been used with a daily time step on the Niger river basin (Oyerinde et al., 2016), and with large-scale assessments using WBM (Grogan et al., 2015, 2017; Zaveri et al., 2016; Liu et al., 2017). The fine-tuning of the parameters when establishing this version of the rule was made using a set of 22 large North-American and Eurasian reservoirs in offline mode, including the two largest reservoirs in the USRB (Palisades and American Falls, daily release Nash-Sutcliffe eficiency (NSE) coefficient 0.70 and 0.60 respectively). Similar to what happens when a reservoir rule that classifies reservoirs by purpose is used in a large-scale model, we did not fine-tune the rule to each reservoir. This allows us to use the reservoir rule in conditions that are similar to what is done in most state-of-the-art hydrological models.

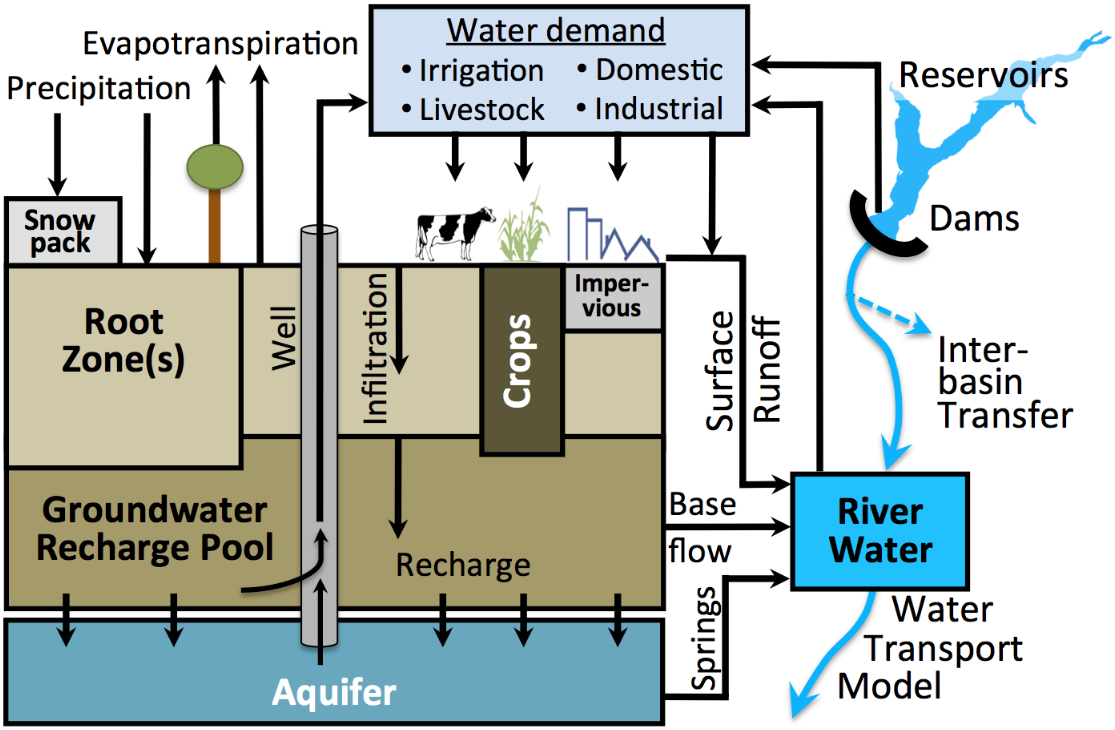



**Figure 2.** Conceptual representation of the Water Balance Model as used in this study.

WBM's release rule, there are qualitatively different behaviors delimited by a reference storage $S_{ref}$, below which the priority is to refill the dam and above which release levels increase rapidly as the reservoir gets nearly full. We call $R_{ref}$ the release at reference storage. For storage $S < S_{ref}$, the rule designed to favor filling the reservoir expresses release $R$ a logarithmic function of storage $S$:

$$R = R_{\min} + \frac{\ln(1+P_R S)}{\ln(1+P_R S_{ref})}(R_{ref} - R_{\min}) \tag{1}$$

where $R_{\min}$ is release at minimal storage, and $P_R$ is a shape parameter for the logarithmic part of the rule that controls the propensity for release. Indeed, $P_R$ close to zero leads to an almost linear rule whereas the higher $P_R$, the more release gets close to $R_{ref}$ even for near-empty storage. For $S \geq S_{ref}$, release $R$ varies exponentially with storage $S$:

$$R = R_{ref} + \frac{(S - S_{ref} + \Delta S)^{P_S} - \Delta S^{P_S}}{(S_{\max} - S_{ref} + \Delta S)^{P_S} - \Delta S^{P_S}}(R_{\max} - R_{ref}) \tag{2}$$

where $R_{\max}$ is release at full storage and $\Delta S$ is computed from the other parameters to ensure that the transition between the logarithmic and exponential parts of the release rule is "smooth" (continuously differentiable). The exponential shape parameter $P_S$ is the propensity for storage, since it minimizes releases until storage is close to its maximal level.

Thus, there are six parameters for the reservoir rule in Figure 3: shape parameters $P_R$ and $P_S$ which represent respectively the propensity for release at low storage (getting releases closer to $R_{ref}$ faster) and for storage in near-full reservoirs (delaying

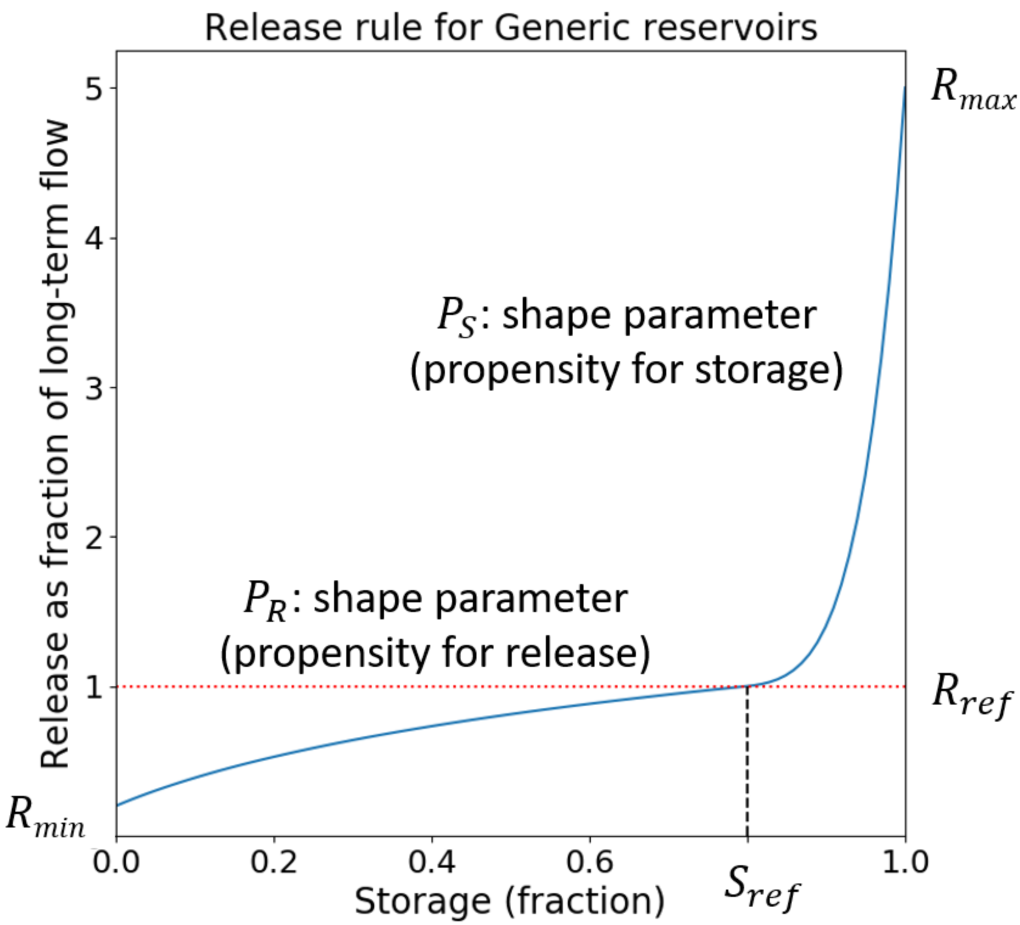



**Figure 3.** 6-parameter reservoir rule. Release is scaled by long-term annual inflow while storage is scaled by the active capacity: it is 0 at dead storage and 1 at full storage.

| Purpose | | $P_R$ | $P_S$ | $R_{\min}$ | $R_{\max}$ | $S_{ref}$ | $R_{ref}$ | $irrFreq$ | Number of reservoirs |
|---|---|---|---|---|---|---|---|---|---|
| Irrigation | (low) | 1 | 3 | 0.01 | 2 | 0.8 | 0.1 | range $[0,1]$ | 73 |
| | (high) | 297 | 3 | 0.292 | 0.885 | 0.949 | 1.44 | | |
| Generic | | 4 | 6 | 0.2 | 5 | 0.8 | 1 | N/A | 33 |
| Hydroelectric | | 200 | 3 | 0.2 | 1.25 | 0.9 | 1 | N/A | 12 |
| Water Supply | | 1 | 6 | 0.1 | 5 | 0.7 | 0.1 | N/A | 10 |

**Table 1.** Parameters for reservoirs in the USRB. The last column classifies the basin's 128 reservoirs by primary purpose.

releases for as long as possible); releases $R_{\min}$ and $R_{\max}$ at minimum and maximum storage; and the coordinates $S_{ref}$ and $R_{ref}$ of the (reference) inflection point. Note that this parameterization, similar to those of other state-of-the-art rules in large-scale hydrological models, do not account for possible coordination mechanisms in multi-actor, multi-reservoir systems. These parameters depend on the reservoir's primary purpose, as shown in Table 1.

"Irrigation" represents the dominant primary use: taken together, irrigation reservoirs represent a storage capacity of 6.27 $km^3$, or just over 90% of the USRB's total storage capacity. Note that "Irrigation" reservoirs require a seventh parameter to model the need to refill to store water for the irrigation season, and release it with the appropriate timing. This parameter, noted $irrFreq$, represents the relative frequency of water demand for irrigation throughout the year. It affects the release rule through each of the other parameters $p_i$ with $1 \leq i \leq 6$, according to:

$$p_i = p_i^{low} + irrFreq \cdot (p_i^{high} - p_i^{low}) \tag{3}$$

with $irrFreq$ between 0 and 1 and the low and high values of the parameters defined in Table 1. This results in three distinct release rules depending on time of year, as shown in Figure 4. Winter features a refill phase ($irrFreq = 0$) with low releases except for keeping a flood control compartment available, whereas peak irrigation season is a drawdown phase ($irrFreq = 1$) with high releases no matter the storage level. A shoulder season ($irrFreq = 0.5$) smooths out the transition between the two.

The reservoir rule for "Hydroelectric" primary use shows a near constant release except at very low storage levels, thanks to a very high propensity for release when $S < S_{ref}$. ~~A primary use for~~ The primary function for the"Water supply" ~~keeps release minimal except for a flood control behavior at near-full storage ,~~ use is to keep releases minimal and storage maximal in order to maximize the quantity of water that can be drawn directly from the reservoir – except for circumstances that require flood control at near-full storage. Other reservoirs mix different uses, and they are represented by the "Generic" rule form that corresponds to Figure 3, and represents an implicit trade-off between uses that prioritize release and those that prioritize storage.

Despite its relative simplicity, the release rule proposed here shares several important characteristics with other rules proposed in the literature. The logarithmic and exponential portions mirror the intuition that the release behavior is qualitatively different depending on storage levels, a trait emphasized by some recent release rules (Wu and Chen, 2012; Zhao et al., 2016; Wang et al., 2019; Yassin et al., 2019). Besides, the representation of reservoirs based on their primary purpose has been a recurring theme since the seminal release rules by Hanasaki et al. (2006) and Haddeland et al. (2006); the time-varying $irrFreq$

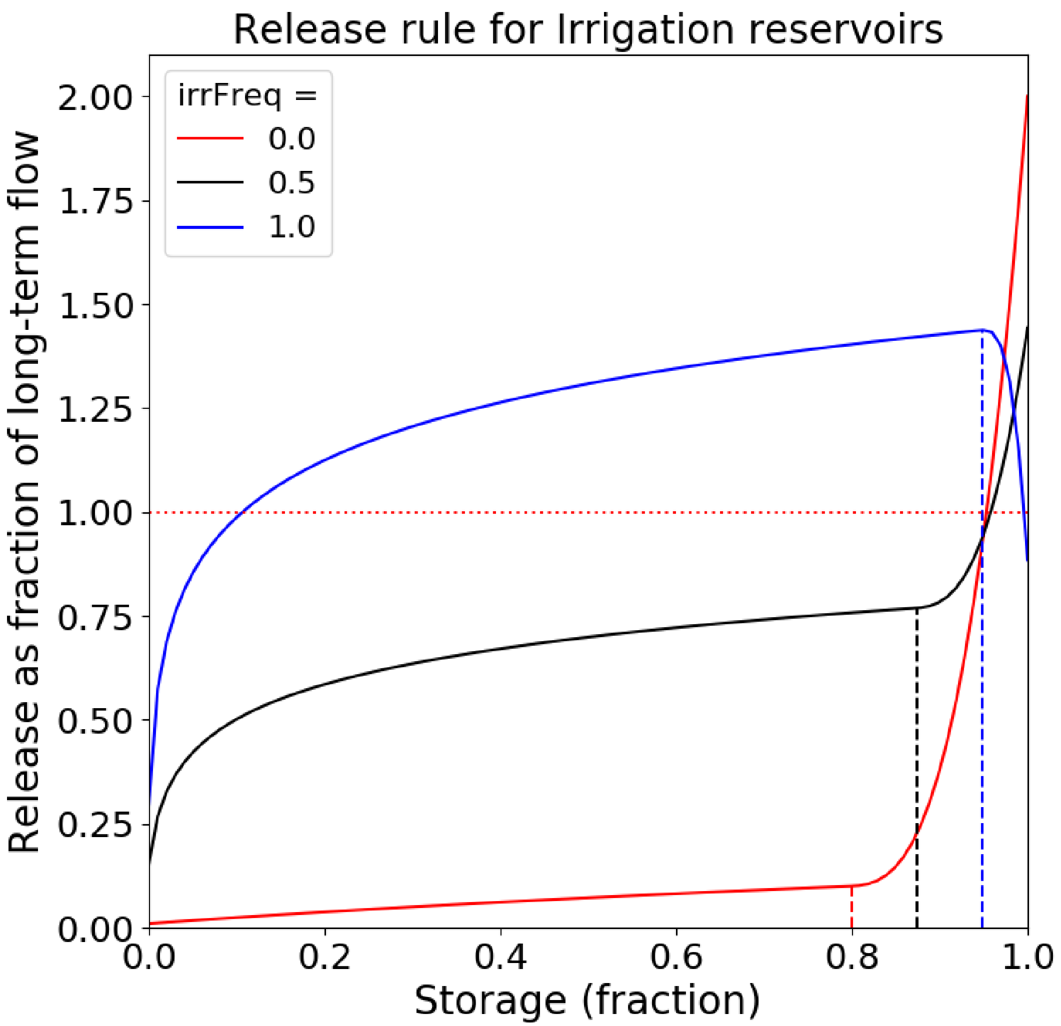



**Figure 4.** Impact of the seasonal shape parameter for reservoirs whose primary purpose is irrigation. Dashed line indicate the inflection point $S^*$. We have $irrFreq = 1$ between July 18 and August 30 included, 0 between October 12 and April 23 included, and 0.5 during the shoulder seasons.

parameter also enables irrigation reservoir to have a flood control behavior in winter, similar to the improvement proposed by Voisin et al. (2013a). Finally, there is the option to fine-tune individual reservoir's release rule parameters to better represent actual (generally multi-purpose) operations. Yet, adjusting parameters implies the assumption that the release rule is able to qualitatively capture the main processes at play in operations of a multireservoir multi-reservoir system. The following section introduces the experimental setup to diagnose this.

## 3 Methodology

### 3.1 General approach

To diagnose the potential consequences of non-coordination within large-scale models' reservoir release rule, our dual objectives are to (A) understand how these consequences play out through space and time in a large-scale hydrological model and (B) compare and contrast the implications of the reservoir rule representation for capturing real-world coordinated reservoir operations. To fulfill objective (A) with WBM and its release rule, we design a sensitivity analysis experiment that uses the Method of Morris on the parameters of this release rule. The focus is not solely on which parameters in the release rule are most influential in regulating flows, but importantly clarifying how the set of dominant controls on water flows and reservoir storage levels evolves along a complex multi-reservoir cascade through time. Parametric sensitivities are then used alongside storage and release trajectories for the simulated ensemble to assess how sets of dominant controls in a point in time and at a given reservoir can be associated to high- or low-flow conditions. We detail the Method of Morris in Section 3.2 and the experiment we design with it in Section 3.1.

A prerequisite to fulfilling objective (B) is to find qualitative behaviors in historical (i.e., observed) operations that cannot be accounted for by a reservoir's release rule simply by changing parameter values or integrating a near-term (less than a month) inflow forecast such as in some existing rules (e.g., Biemans et al., 2011). Other variables would be necessary to incorporate these behaviors within the hydrological models. They will be interpreted as preliminary indications of cooperation, and the events they highlight will be investigated further. Then, the comparison of historical and simulated operations will be used to understand how reservoirs coordinated in the historical record, and contrast the impact of this with reservoir non-coordination in simulations. Differences observed in this comparison could be due to all sorts of error in the model, and not just to non representation of coordination. Therefore, the comparison will be supplemented by an analysis exploiting offline reservoir water balance models that implement basic coordination on the simulated ensembles at times where there are indications of cooperation. These offline reservoir water balance models help to show that coordination alone could explain the qualitative differences between simulated and historical operations. Note that we cannot fully quantify and explain all sources of errors: there is to our knowledge no study of a real operational context using a hydrological modeling experiment where one would absolutely know every source of error or potential confounding factor.

Both (A) and (B) are carried out by focusing on the USRB in 2009-2016. The geographical extent of the USRB is small compared with that of areas traditionally considered in large-scale hydrological models. This is because a focus on a smaller area makes drawing quantitative and qualitative lessons from our diagnostic analysis easier. These lessons can then be applied to understand the potential consequences of not considering coordination between reservoirs over larger study areas. For this reason, we also aim to evaluate model outcomes in the same conditions as large-scale hydrological vulnerability assessments are carried out. For instance, we do not fine-tune reservoir rule parameters to individual basins, and instead use purpose-based parameterizations typical in large-scale studies (e.g., Biemans et al., 2011; Voisin et al., 2013a; Yoshikawa et al., 2014).

### 3.2 Time-varying sensitivity analysis: Method of Morris

The Method of Morris (Morris, 1991; Campolongo et al., 2007) has proven to be a successful tool for detailed diagnostic evaluations of large and complex hydrological models ~~(e.g., Herman et al., 2013b; Zajac et al., 2017; ?)~~(e.g., Herman et al., 2013b; Zajac et al., 2 . This section presents the sampling technique used, the basic Morris sensitivity indices as well as a time-varying version of these indices. All sensitivity analyses were performed using the SALib toolbox written in Python by Herman and Usher (2017).

The Method of Morris samples points within the parametric spaces of interest by following so-called "trajectories". Two consecutive points of a trajectory share the same input values except for parameter input $i$ where they are separated by a distance $\Delta_i$. The value of input dimension $i$ is changed exactly once along a trajectory, and the order in which input dimensions are changed is random. If $D$ is the dimension of the parametric input space being sampled, then each trajectory comprises $D+1$ points. To ensure that the Morris sensitivity measures are as accurate as possible, sampling must cover the parametric input space as well as possible. This paper implements the method proposed by Ruano et al. (2012), which first generates a large number of trajectories, then selects a subset that provides near-optimal input space coverage using a computationally efficient optimization technique (as implemented here, $M = 50$ trajectories were selected out of a thousand).

To compute Morris indices from a set of $M$ input trajectories, one must run the model whose parametric input space is being sampled at each point $x$ of each trajectory. Therefore, there are $M \times (D+1)$ model runs. For each trajectory $j$ ($1 \le j \le M$), model runs yield the so-called elementary effect along input dimension $i$ for each date $t$:

$$EE_i^j(t) = \frac{f(x_1, \ldots, x_i + \Delta_i, \ldots, x_M) - f(x)}{\Delta_i} \quad (4)$$

With $M$ trajectories being sampled, sensitivity index $\mu_i(t)$ for input dimension $i$ at date $t$ is the average over the elementary effects:

$$\overline{\mu_i}(t) = \frac{1}{M} \sum_{j=1}^{M} EE_i^j(t) \quad (5)$$

In this work, we are concerned with relative contributions to sensitivity across reservoir rule parameters ($1 \le i \le D$) and over a given time period ($t \in [t_1, t_2]$). Therefore, we compute the following normalized values for the Morris sensitivity index:

$$\mu_i(t) = \frac{\overline{\mu_i}(t)}{\mu_{\max}} \quad (6)$$

where $\mu_{\max}$ is the maximal value of $|\overline{\mu_i}(t)|$ over the input space and time frame of interest:

$$\mu_{\max} = \max_{i \in [1,n], t \in [t_1, t_2]} |\overline{\mu_i}(t)| \quad (7)$$

As a result, each $\mu_i(t)$ values will be between -1 and 1. Absolute values close to 1 representing inputs that have a dominant influence on outputs, not only compared with other inputs at that date, but also compared with inputs' impacts on outputs at other dates within $[t_1, t_2]$. Positive values denote that outputs values increase with input values, whereas the contrary holds for negative values.

### 3.3 Reservoir Parameters and Ranges

We conduct this diagnostic analysis with 7 groups of parameters. Each group contains one of the 7 parameters of the release rule for all 128 reservoirs in the USRB. This analysis uses a range of $\pm 10\%$ around base values for all parameters in Table 1. ~~There~~ These modest 10% ranges would be conservative if our focus were calibration and not diagnosis. Results (Section 4) will demonstrate that our narrow sampling yields quite substantial effects when compounded across the reservoir cascade in periods where coordinated operations are significant. Besides, there are two reasons for ~~this~~choosing the same range across all parameters: 1) it accounts for the fact that each parameter does not have the same base value across all reservoirs, and 2) it facilitates comparisons between different parameters' sensitivity indices.

Our choice to explore 7 groups of parameters serves to reduce the computational burden of our diagnostic analyses, while facilitating a clear experimental mechanism to investigate the core parameterization assumptions used to capture multi-reservoir release and storage dynamics. It also meets the core objective of this study, which is to clarify the importance of multi-reservoir coordination and control to our model-based assessments of flood and drought vulnerabilities in complex river basin systems. Indeed, the chosen parameter set is necessary and sufficient to answer two key intermediary questions. First, we must understand how release rule parameters from a given reservoir influence its water balance (release and storage) through time. This makes it necessary to consider all 7 parameters of the release rule. Second, we must understand how the release rule parameters from upstream reservoirs influence subsequent at-site reservoir controls. This is key to understanding how the non-coordinated release rule affects the time-varying response to high- and low-flow extremes as we move down a reservoir cascade. Our experimental design highlights when parametric controls on reservoir releases are modified by upstream interferences. Indeed, the same parameters that increase release at a given reservoir also increase upstream releases. Both effects have opposite consequences for a reservoir's storage. Our analysis will track the instances in which upstream controls dominate at-site controls, and clarify the consequences of this for the reservoir cascade's response to hydrological extremes.

The $D = 7$ parameters and ranges thus defined are used to set up a method of Morris experiment with $M = 50$ trajectories. The ensemble size is $M \times (D+1) = 400$. We ran this experiment on The Cube cluster at the Cornell Center for Advanced Computing Results. The Cube has 32 compute nodes with Dual 8-core E5-2680 CPUs at 2.7 GHz, with 128GB of RAM. A single run of the USRB WBM takes close to seven hours on average for the USRB, with an eight-year simulation period (2009-2016) preceded by a five-year spinup period. The ensemble of 400 members took almost 3,000 hours of compute time to get and analyze, using parallel runs exploiting Open Message Passing Interface version 1.6.5.

| Reservoir name | WBM primary usage | Capacity ($hm^3$) |
|---|---|---|
| Jackson Lake | Irrigation | 1,078 |
| Palisades | Irrigation | 1,503 |
| American Falls | Irrigation | 2,145 |
| Minidoka | Water supply | 123 |
| Milner | Water supply | 62 |

**Table 2.** Reservoir cascade on the main stem of the Upper Snake River, ordered from upstream to downstream.

## 4 Results

Our results focus on the reservoir cascade on the main stem of the Upper Snake River (Table 2). The three upstream reservoirs in the table are the three largest reservoirs in the basin, and their capacity to store water for the irrigation season is crucial to the agricultural sector in the USRB. Consequently they are classified as "Irrigation" reservoirs. The two downstream reservoirs are smaller and must be maintained at high storage levels during the irrigation season so that canals can draw directly from them, leading to their classification as "Water supply" reservoirs. All but the most downstream reservoir (Milner) are part of, or associated to, the Minidoka Project, therefore their operations for water supply and flood protection are largely coordinated when deemed necessary. Using an ensemble of WBM simulations computed as specified in Section 3, we carry out a diagnostic evaluation of the parametric controls of the release rules in three steps. Initially, we focus on the upstream reservoir, Jackson Lake, where there are no interferences from other reservoirs upstream (Section 4.1). This is where imulation results enable us to quantify the main controls on a reservoir's release rulewithout interferences from other reservoirs upstream (Section 4.1), a prerequisite to studying how these controls evolve with hydroclimatic events and along the reservoir cascade. This is also where we can find indications of coordination within the historical record, as defined in Section 3.1. Next, we quantify the upstream interference with downstream releases in the USRB's cascade (Section 4.2). Finally, Sections 4.3 and 4.4 contrast actual observed operations with those from the simulated ensembles for recent USRB low and high flow events, respectively.

### 4.1 Upstreamparametric : controls on release and storage

#### 4.1.1 Dominant parametric controls in simulations

First, let us examine WBM's parametric reservoir rule's dynamic sensitivities through Jackson Lake, an upstream reservoir that is not influenced by inflows from any other reservoir in the USRB. Figure 5 provides a visualization of time-varying sensitivities at the daily time step. Simulated and observed hydrological time series overlay the sensitivities represented by the blue-to-red color scale. The left vertical axis represents the plotted reservoir state (release or storage). In Figure 5, shades of blue, red and white report the normalized Morris sensitivity index in a given time period (the horizontal axis); they are organized over seven lines, each corresponding to one of the seven reservoir rule input variables evaluated in our analysis and listed on the right vertical axis. As indicated by the colorbar right of the figure, shades of reds correspond to normalized Morris

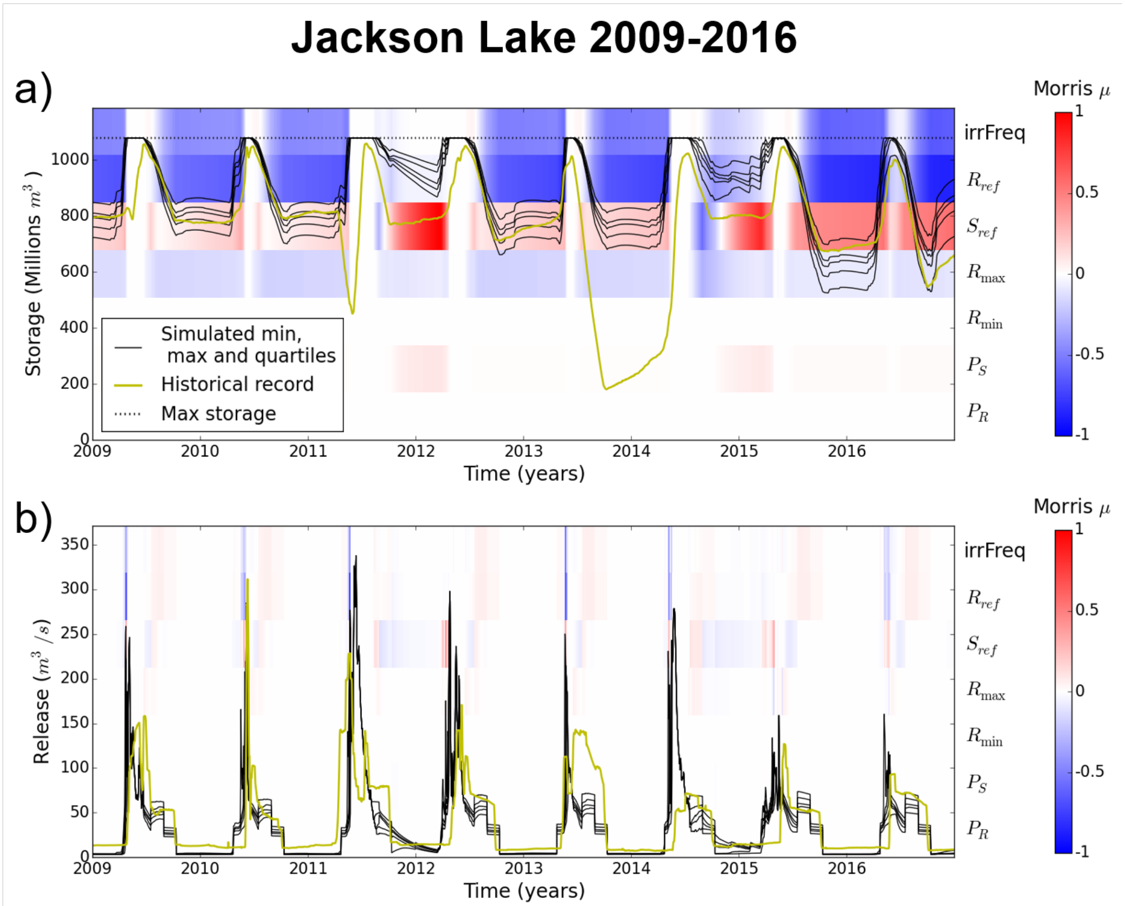



**Figure 5.** Foreground: comparison of the observed (gold line) and simulated (black lines) trajectories at Jackson Lake, for a) reservoir storage and b) release. Background: daily sensitivities to reservoir rule parameters on left y-axis.

sensitivity values close to 1, indicating that the associated variable is dominant and that higher values of it correlate with higher values of the reservoir state. Conversely, shades of blue correspond to normalized Morris sensitivity values close to -1 and indicate that the associated variable is dominant but that higher values of it correlate with lower values of the reservoir state. Finally, whites indicate sensitivity is weak compared with that or other variables and / or dates.

In panel (a) of Figure 5, storage differences that emerge across the sampled parametrizations of the evaluated WBM ensemble members are the time integral of daily release differences, therefore they indicate the cumulative effects over time of how the parameters influence the release rule. Storage sensitivities present clear annual patterns for Jackson Lake, and are more broadly representative of the dominant controls for an "irrigation" reservoir within WBM that has an absence of interactions from other reservoirs. Panel (a) shows that the Jackson Lake storage sensitivities go to zero in periods where maximum storage is attained; this is expected because then, there is no variation in storage across the ensemble. In all other periods, there are three dominant parameters influencing storage: $S_{ref}$, $R_{ref}$ and $irrFreq$, with a remarkably consistent influence over the eight-year simulation period. The direction in which these parameters influence the release rule is consistent with the release rules of Figures 3 and 4. Indeed, higher values of $R_{ref}$ and $irrFreq$ directly increase release, therefore decreasing storage over time, whereas increasing $S_{ref}$ delays the transition between the logarithmic and exponential parts of the rule, and has the opposite effect – except if storage is very high during peak irrigation season, due to the dip observable for $irrFreq = 1$ in Figure 4. Although somewhat reduced in effect relative to top three parameters controlling storage dynamics in the Jackson Lake reservoir, the maximum storage releases $R_{\max}$ are predictably inversely correlated with storage (panel (b) on Figure 5). Yet, the three dominant parameters $S_{ref}$, $R_{ref}$ and $irrFreq$ yield what can be interpreted as the "signature" of the parametric influence of the release rule governing storage over time.

Transitioning to Jackson Lake's parametric sensitivities for releases, the overall magnitude of the normalized Morris indices are substantially reduced and less consistent relative to those for storage (Figure 5 panel (b)). A potential reason for this is a dampening effect, as parameters that increase current releases also decrease future storage, and consequently limit future releases. Another potential reason for the diminished sensitivities overall in panel (b) is that contrary to storage that registers the cumulative effects of parametric differences, release sensitivities peak on particular days, making other time periods less sensitive in comparison. The parametric sensitivities for the Jackson Lake release have an opposite "signature" as that of storage during large stretches in summer and fall over 2009-2016. Indeed the three dominant parameters are $S_{ref}$, $R_{ref}$ and $irrFreq$, with higher values of $S_{ref}$ correlating with lower release whereas higher values of $R(S^*)$ and $irrFreq$ correlate with higher releases. Both "signatures" are consistent because parameters that have a sustained impact on release are expected to have an opposite effect on storage.

#### 4.1.2 Comparison with historical operations

Overall a comparison of historical (i.e., observed) versus simulated storage and releases in Figure 5 shows a broad agreement during the eight-year study period, despite two major departures, especially apparent for storage. In 2011, early-spring release in the historical record created flood control storage and enabled peak flows to be lower than in the simulated ensemble. Observations of large drawdown in the summer of 2013, with the reservoir replenished only in 2014, are not matched by the simulations. These correspond respectively to high- and In both situations, we compared release over a whole month to total water availability (initial storage plus total inflows over the month) for each of the 8 years covered by our analysis. In both cases, results from Figure 6 shows that both events correspond to a major departure with the expectation that release should be indexed on water availability. In 2013 (panel (a)), releases are much higher than any other year despite water availability that August being the lowest of the 8 years. This corresponds to a low-flow periods, both of which were managed by historical operations. We contrast these observations period during which extra water is released to help downstream reservoirs meet demand; we contrast this coordinated historical response with simulation results in detail in Sections 4.3for the low-flow event and 4.4 for the high-flow eventSection 4.3. In 2011 (panel (b)), releases are over four times higher than normal despite water availability being comparable to conditions for 6 of 7 other years. This is a prelude to an intense snowmelt season, requiring anticipation and coordination from the two main reservoirs tasked with flood control in the USRB: Jackson Lake and Palisades (U.S. Bureau of Reclamations, 2012). We constrast this response with simulation results in Section 4.4.

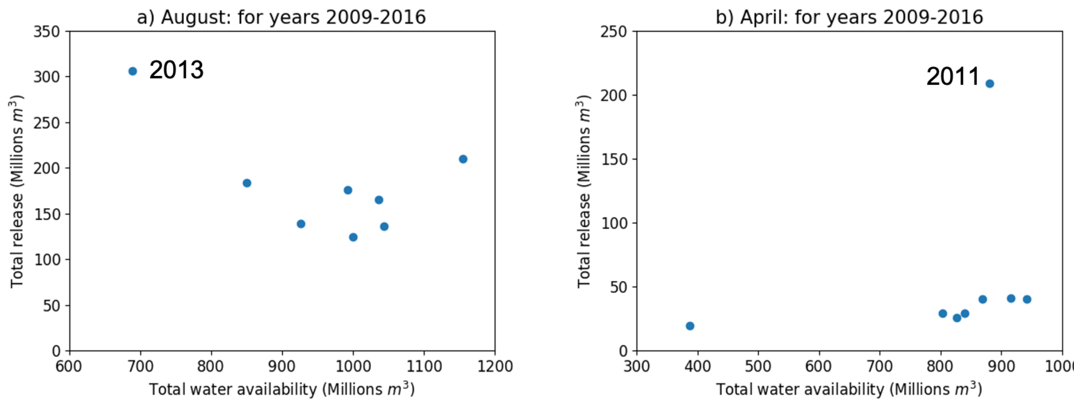



**Figure 6.** Total monthly historical release vs. water availability (beginning-of-month storage plus inflows) for each year between 2009 and 2013 at Jackson Lake. a) left: August and b) right: April

### 4.2 Absence of downstream coordination in controls

We now transition our focus to the fourth reservoir on the USRB reservoir cascade, Minidoka, which is considerably smaller than the first three. Our analysis in Figure 7 focuses on flows from a single year to clarify the complex interactions between upstream releases and Minidoka operations. Starting with inflows to the reservoir shown in panel (a) on Figure 7, the dominant parameters controlling inflows are $R_{ref}$ and $irrFreq$ (and to a lesser extent $S_{ref}$). The time-varying pattern of dominant release sensitivities across year 2013 (panel (b)) mirrors that of inflows, as dominant parameters tend to positively and negatively correlate with inflows and releases alike at the same time of year. Moreover, the strong release sensitivities to the seasonal $irrFreq$ parameter from May to October can only be due to interactions with upstream reservoirs, because $irrFreq$ only influences irrigation reservoirs' release rule whereas Minidoka is classified as a "Water supply" reservoir. These results suggest that the upstream reservoirs' rules are a dominating factor in this downstream reservoir's release decisions.

However closely variations in simulated releases in panel (b) of Figure 7 tend to follow simulated inflows in panel (a), these releases show unexpected high frequency fluctuations that are artifacts that are not meant to occur in the reservoir's release rule. This shows the unintended consequences of interactions with upstream reservoirs. In other words, it would arguably be very difficult to calibrate the parameters of Minidoka's release rule without accounting for the complex upstream interactions. Mathematically, this is termed non-separability.

All of these insights from comparing inflow and release sensitivities are confirmed by looking at Minidoka's 2013 storage sensitivities in panel (c). Similar to the release sensitivities in panel (b), the influence of $irrFreq$ on storage is a direct signature of interactions with upstream releases. In fact, the dominant storage sensitivities for the whole year are end-of-April sensitivities to $irrFreq$ and $R_{ref}$ (dark red on panel (c)). The former parameter is not defined for the Minidoka release rule, whereas the latter should be associated with negative sensitivity (with the color blue) in absence of upstream interactions. The simulated reservoir filling for Minidoka is strongly influenced by parametric artifacts outside of its own parameterization. Beyond that,

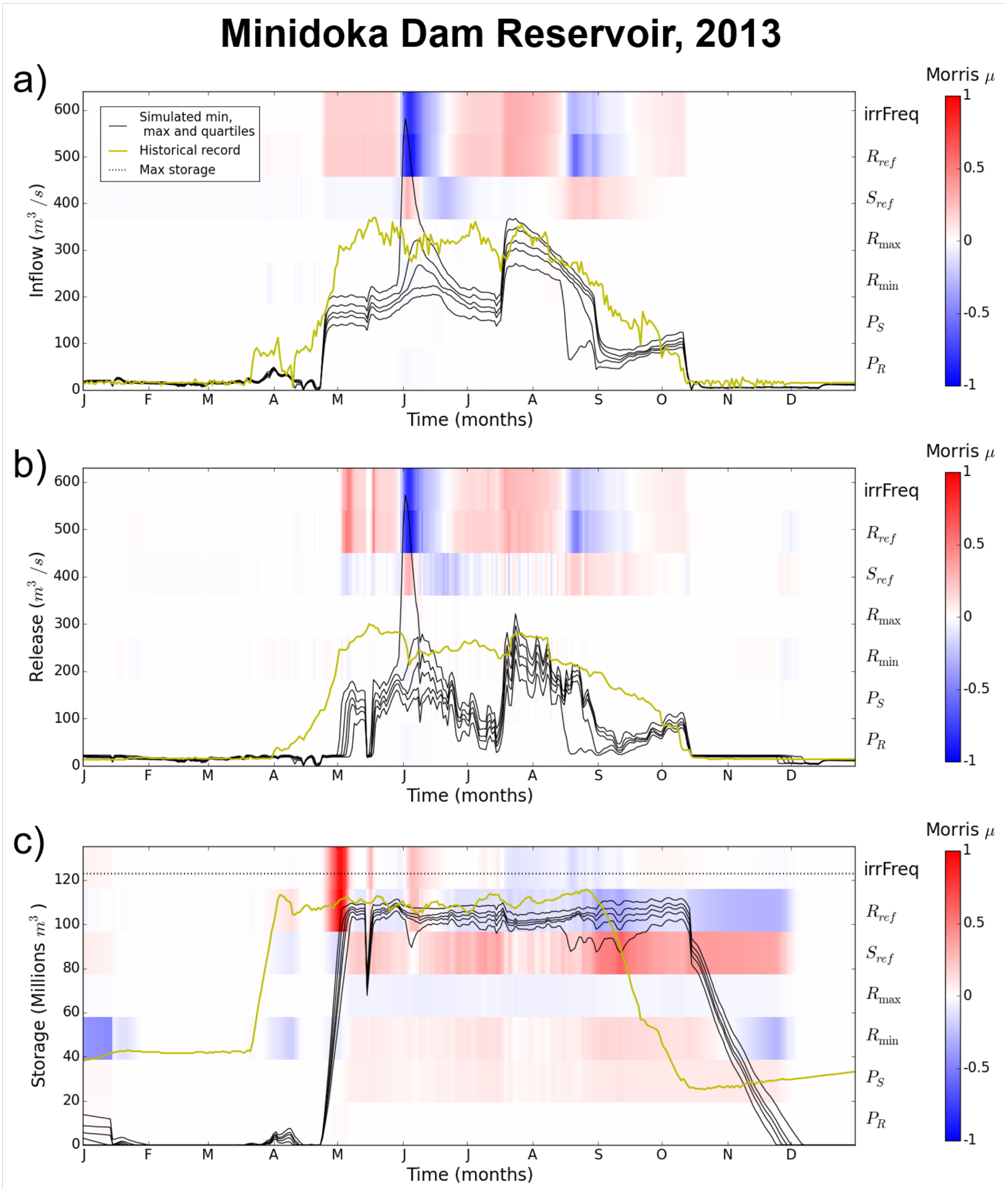



**Figure 7.** Simulated values (max, min and quartiles, shown with black lines) with historical values (gold line), and sensitivity to input variables (background), for (top to bottom): Minidoka reservoir's inflow, release and storage.

the picture of time-varying storage sensitivities is extremely complex. For instance, the direction of storage sensitivity to $irrFreq$ (i.e., positive or negative correlation with storage) does not always appear to be clear and consistent with that same parameter's sensitivities for inflows and releases (compare panels on Figure 7). This apparent complexity cannot be dissociated from upstream interactions, again reinforcing that parameterizing Minidoka's release rule cannot be done separately from the parameterizations of the upstream reservoirs. This would Joint parameterization would explicitly account for coordination within reservoir operations, but would also require searching a parameter space of very high dimensionality.

### 4.3 Drought risk

We now transition to the reservoir operations along the USRB's reservoir cascade for the consecutive dry years of 2012 and 2013. We contrast coordinated historical operations, illustrated here by storage levels in the basin's three main reservoirs, with the simulations results from our ensemble of hydrological model runs – which we term *simulated* storage in this Section and in the next (Section 4.4). We also analyze the sensitivity of simulated storage to WBM's parametric controls.

The 2012-2013 low-flow event led to a significant simulated drawdown at upstream Jackson Lake in 2013, previously observed in Figure 5. The strong deviations in the dynamics of historical (gold lines) and simulated (black lines) reservoir operations for both years 2012 and 2013 are apparent in Figure 8. Recall that the two most downstream reservoirs in the Snake River reservoir cascade — Minidoka and Milner – are smaller reservoirs that must stay full during the irrigation season so farmers can draw water through gravity irrigation. Therefore, it is key that American Falls, the main reservoir in the Snake River plain located just upstream of Minidoka, is not empty so that it can keep regulating water levels in downstream reservoirs, ensuring irrigation needs are met. For this reason our analysis will start with American Falls (panel (c) on Figure 8) and work its way upstream to shed light on the historical observed coordination, and lack thereof in the simulations, during the 2012-2013 low-flow period. The pace and magnitude of the drawdown are the defining differences between historical and simulated operations at American Falls. For both years, historical operations show reservoir levels decreasing at a near-constant rate from nearly full in early May to about $5-10\%$ by the end of summer. The drawdown season spans 4-5 months and the reservoir never loses its capacity to regulate downstream reservoir levels. Alternatively, simulated drawdown seasons are much shorter – two and half months from mid-June to the beginning of September – and the reservoir swings from full (in 2012) or nearly so (in 2013) to completely empty either for the whole ensemble (in 2012) or nearly half of it (in 2013). In other words, American Falls loses its capacity to regulate irrigation delivery or is simulated to be dangerously close to doing so.

The reason for this contrasting behavior can be found with upstream operations. For instance, the historical storage trajectory at Palisades (panel (b)) shows a marked drawdown from early July to late October 2012. On average, the reservoir released over $0.5\ km^3$ more towards American Falls in the observed record than it does in the simulations, and this enabled American Falls to keep its capacity to regulate irrigation withdrawals on the Snake River plain. Simulated storage sensitivities however, reflect the lack of coordination across the ensemble of simulations, as both the main controls on storage from April 2012 onwards are the same as for large reservoirs for which at-site controls dominate upstream interferences (see section 4.1). These controls, and the simulated storage trajectories, fully ignore any connection with the simulated events unfolding downstream.

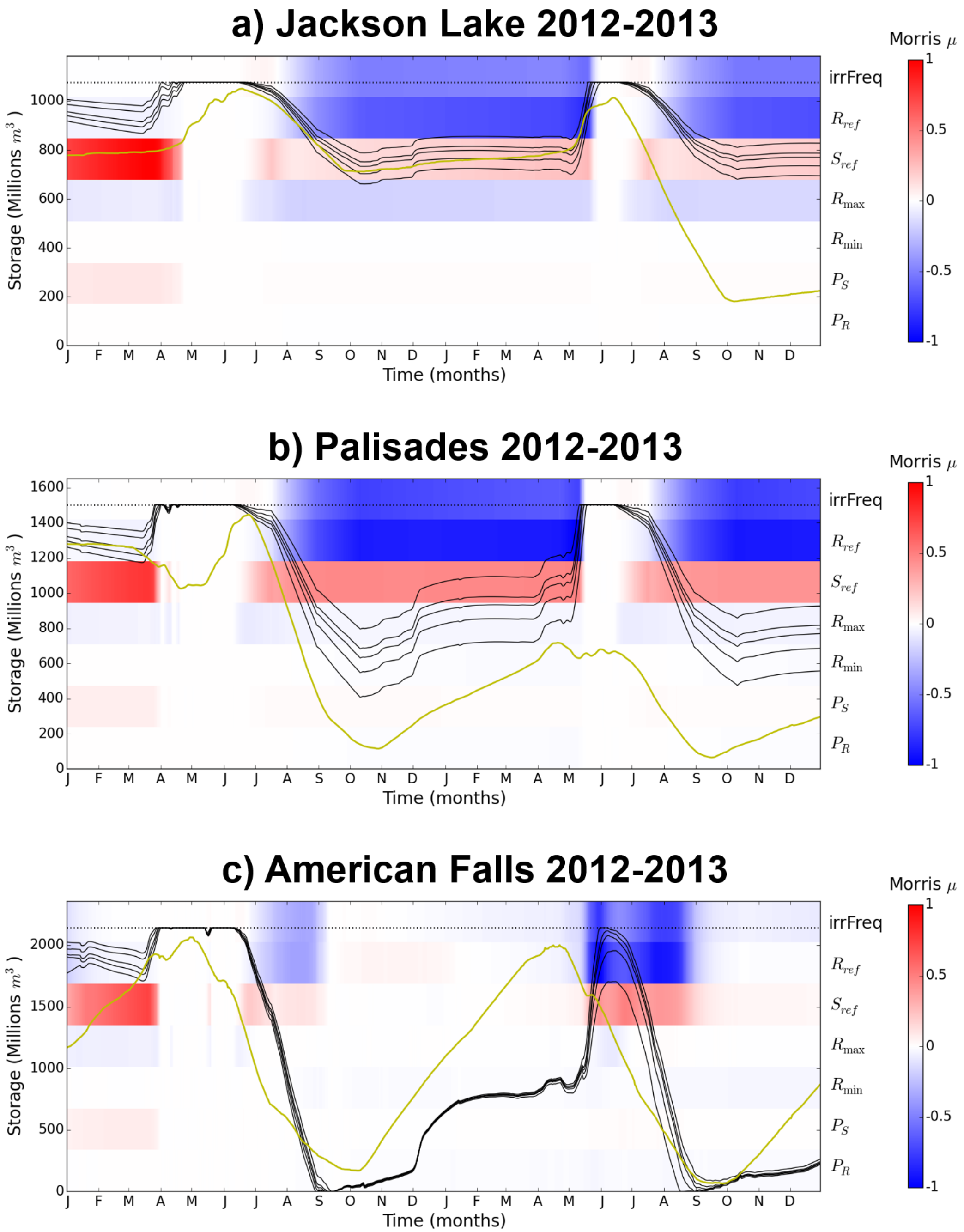



**Figure 8.** Recorded and historical storage during the dry years 2012 and 2013 with sensitivities of the simulated variables illustrated in the background. From top to bottom: storage at three largest USRB reservoirs from upstream to downstream.

Yet, in 2013 historical storage levels at Palisades had not recovered from the exceptional 2012 drawdown, so that the reservoir could no longer supply extra water the Snake River Plain. Instead, exceptional historical Jackson Lake drawdown in the summer of 2013 (panel (a)) supplied over 0.5 $km^3$ extra water to the Snake River Plain compared with what the simulations record. Thus, complex multi-year and multi-reservoir coordination was needed to avert adverse drought impacts on agriculture. The simulations do not account for this coordination, as demonstrated by both simulated storage and by the consistent parametric controls at both Jackson and Lake and Palisades for both years. The ensemble of simulations let American Falls empty whereas the two largest reservoirs upstream of it remain close to full.

To test whether the observed differences between the historical record and simulation are an artefact of other errors in the hydrological model, we built a simple offine water balance model over 2012-2013, taking the three reservoirs' simulated water balance offline. For each of the 400 ensemble members, we then simulated operations obtained by releasing an extra $50m^3/s$ during the summer than planned according to release rules, as long as the reservoir is over 20% full. This is consistent with the difference between historical and simulated releases. The extra release concerns only Palisades in 2012, and both Palisades and Jackson Lake in 2013. To make sure that underestimating routing times between reservoirs does not falsely cause the reservoirs to store water, we choose a conservatively high (for the area) routing time of 7 days between each reservoir and the one downstream. Results (Supplementary Information Figures 1 and 2) show that with this simple coordination water balance measure, American Falls would not have emptied – and reservoirs upstream of it would not have lost their capacity to supply downstream agriculture either.

### 4.4 Flood risk

We next evaluate if these representational deficits in simulating coordinated operations also yield consequential errors in the Spring of 2011, where the observed operations averted a flood by exploiting forecast-based anticipatory releases in the two upstream large reservoirs at Jackson Lake and Palisades. Following the flow from upstream ~~to downstream along the USRB's three largest reservoirs~~ Jackson Lake to downstream Palisades (Table 2), we contrast coordinated historical storage and discharge levels observed in the Spring of 2011, with the simulations results from our ensemble of hydrological model runs and the associated release and storage sensitivities to WBM's parametric controls.

#### 4.4.1 Jackson Lake

Starting upstream, we focus on the storage and release dynamics, both simulated (black lines) and historical (gold lines), at Jackson Lake (Figure 9). All simulation results fill the reservoir entirely between May 14 at the earliest and May 26 at the latest (panel (a)); this period coincides with maximal release sensitivity (panel (b)). Note that the dominant controls on simulated release during May 14-26 (panel (b)) are the same as the dominant controls on simulated storage prior to that period (panel(a)), with strong negative sensitivities to $R_{ref}$ and $irrFreq$ and strong positive sensitivity to $S_{ref}$. ~~These controls correspond~~ The dominance of these three parameters corresponds well with our prior results~~for the most common dominant parameters~~, as detailed in Section 4.1. These results however run contrary to the real-world expectation that for reservoirs with no upstream interactions, release rule parameters should influence release and storage in opposite directions.

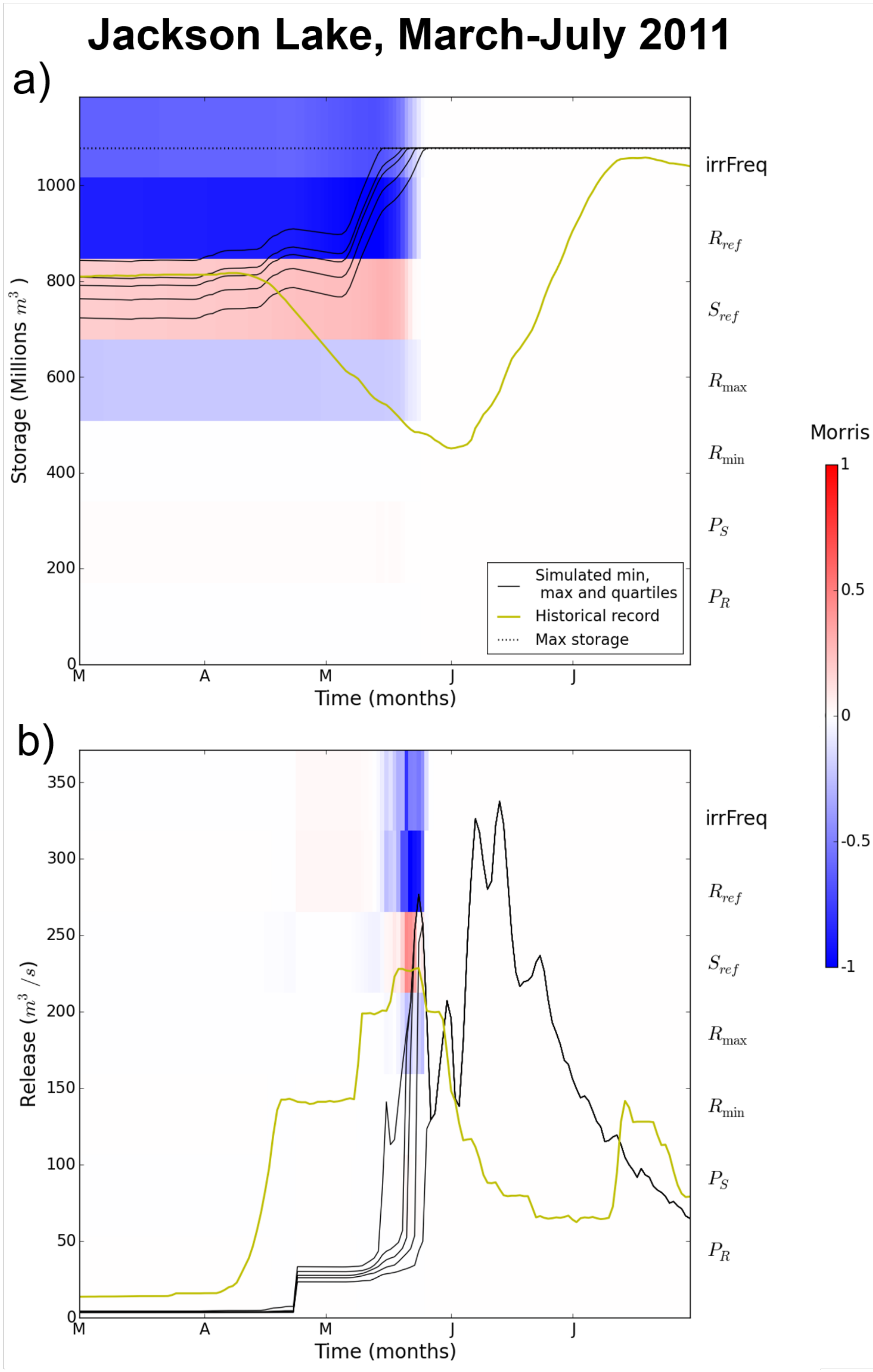



**Figure 9.** Simulated values (max, min and quartiles, shown with black lines) with historical values (gold line), and sensitivity to input variables (background), for Jackson Lake's release and storage.

Instead This is because during the snowmelt-driven peak flow season, higher simulated storage leads to quicker reservoir filling which takes away the reservoir's capacity to regulate peak flows. Once the reservoir is full, simulated peak releases out of Jackson Lake are much higher than the historical observed releases. This is essentially due to the fact that These have been mitigated by real-world reservoir operators who started releasing water in early April to decrease reservoir storage by almost half between then and early June. This created enough storage space to absorb runoff from peak snowmelt season in June, while simultaneously reducing releases to limit the reservoir's contribution to downstream high flows. By contrast, all simulated releases only increase gradually when the reservoir gets close to full capacity. Due to this lack of foresight-driven preventive releases in the simulations, Jackson Lake is full by the end of May and unable to absorb peak flow in June. This represents a large and consequential structural error in model's representation of flooding operations and vulnerabilities.

#### 4.4.2 Palisades

Moving to the next reservoir downstream, Figure 10 illustrates the simulated (black lines) and historical (gold lines) storage and release dynamics for the Palisades reservoir in March-July 2011. All simulation results fill the reservoir entirely between May 5 at the earliest and May 9 at the latest (panel (a)). Similar to Jackson Lake, this coincides with a period of maximal release sensitivity, and the dominant controls on both storage and release are identical: same parameters ($irrFreq$, $R_{ref}$ and $S_{ref}$) with the same directional effects. Put simply, parameters that favor reservoir filling in simulations diminish Palisades reservoir's capacity to store water and to absorb peak snowmelt season flows, leading to heightened simulated releases. Since Palisades is downstream of Jackson Lake and snowmelt occurs earlier at lower altitudes, simulated filling occurs earlier, and consequently the WBM abstraction of the reservoir is subsequently unable to absorb snowmelt peaks, including the one event occurring May 24-26 as the result of Jackson Lake filling. This is evidenced by parametric release sensitivities and the concurrent simulated release peak (panel(b)) around these dates that necessarily come from upstream – there is no on-site release sensitivity when Palisades is full.

By contrast, historical operations favored preventive releases as early as the end of March at Palisades, to free up almost 1.3 $km^3$ of storage space by early May – precisely at the time when the onset of snowmelt fills the reservoir up in simulations. This leaves over 1.1 $km^3$ of storage space by early June, and the comparison of Figures 9 and 10 shows that both Jackson Lake and Palisades filled at a near-constant pace throughout June, nearing being completely full around July 10. This controlled and coordinated filling of both reservoirs ensured that releases well below 700 $m^3/s$ at Palisades, a full 900 $m^3/s$ lower than the simulated peak across virtually all of the simulated ensemble. The simulated peak is almost 40% higher than the highest observed daily discharge over the past 40 years. Maximal Palisades release over this period – 1140$m^3/s$ on the 20 June 1997 – corresponds to a flooding event which led to six counties declaring a state of disaster, leading to over USD 11 million in relief by the federal U.S. government (wea). Coordination is mediated by seasonal forecasts based on snowpack height, and is apparent through the dual fact that reduction in Jackson Lake release (Figure 9 panel (b)) when Palisades starts filling back up. As a result, neither reservoir ever loses its capacity to regulate streamflow by filling completely, and that downstream releases are capped. The simulation is strongly inconsistent with the institutional flood control objectives of the reservoirs (U.S. Bureau of Reclamations, 2012).

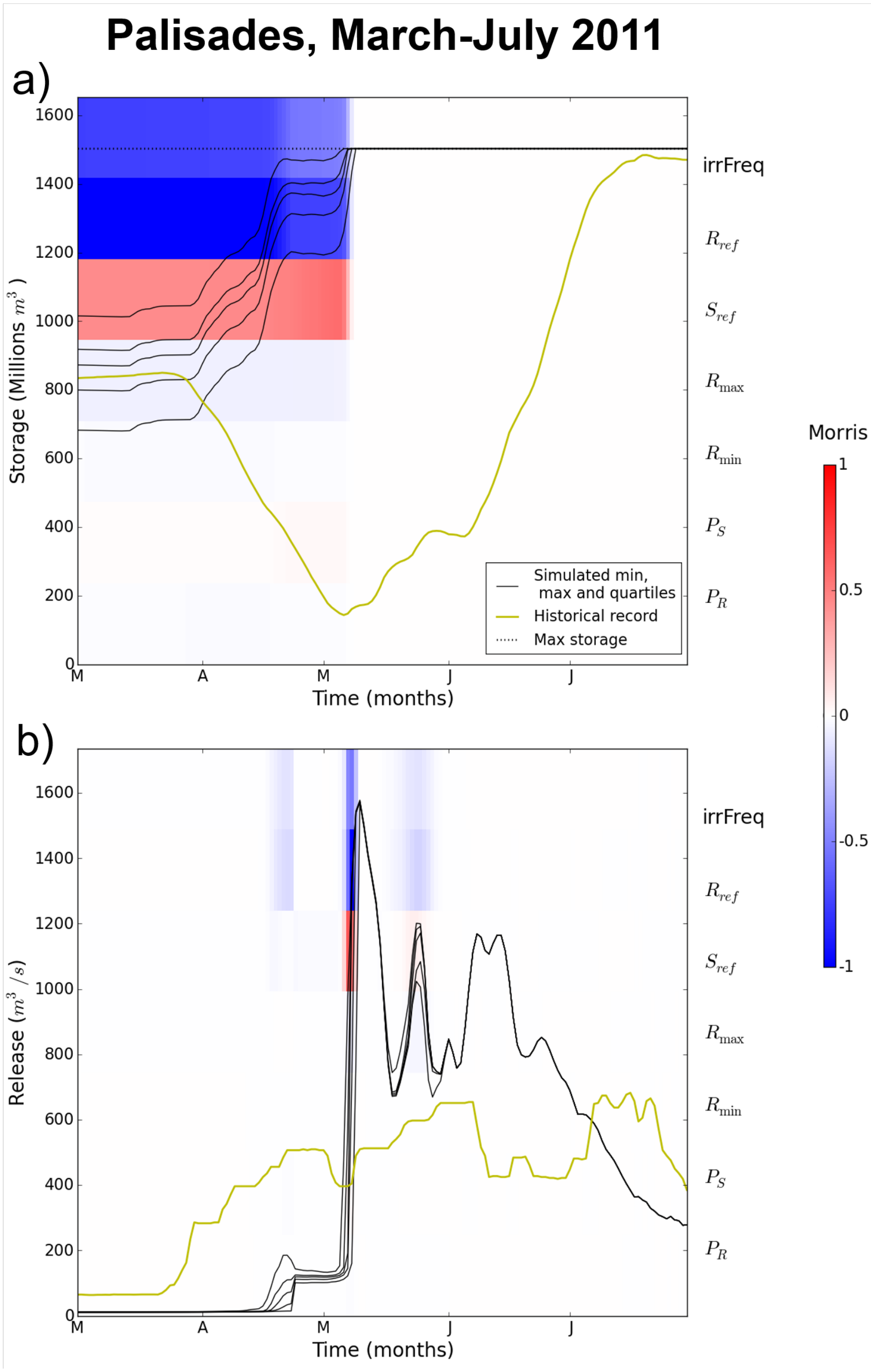



**Figure 10.** Simulated values (max, min and quartiles, shown with black lines) with historical values (gold line), and sensitivity to input variables (background), for Palisades reservoir's release and storage.

### 4.4.3 American FallsOffline water balance experiment

Simulated values (max, min and quartiles, shown with black lines) with historical values (gold line), and sensitivity to input variables (background), for American Falls reservoir's release and storage.

Transitioning to American Falls (Figure ??), the third reservoir on the cascade and the largest reservoir in the USRB, enables us to compare the compounding errors from WBM-simulated upstream release policies on high flows with historical operations. All simulated ensemble members have the reservoir entirely full by May 9, which coincides with the date where the Palisades reservoir is also full across the fullensemble. Accounting for the routing delay due to the distance between Palisades and American Falls (over 200 km), Total water availability at both Palisades and Jackson Lake over the 5-month period between March and July 2011 in the simulated ensemble is within 10% of historical water availability for all ensemble members. Yet, temperature-sensitive snowmelt processes are notoriously difficult to resolve in mountain ranges (**??**), especially at fine space and time scale. In these simulations, snowmelt yields very high modeled inflows to Palisades over a short period of time. These inflows are mirrored by the timing at which the two biggest reservoirs in the USRB are simulated to be full results in large release sensitivities in the days before and after May 9. As in Figures 9 and 10, these sensitivities reflect reservoirs losing their capacity to regulate high flows as they get full. When American Falls is full across the whole simulated ensemble, release sensitivity can only be the result of sensitivity of upstream releases to reservoir rules parameters releases observed in Figure 10 panel (b) – most notably, the recall that releases mirror inflows when a reservoir is full. In contrast, daily historical inflows to Palisades releases. never reach $1,000 m^3/s$ over the 2011 snowmelt season. Therefore, it is important to evaluate the extent to which the large qualitative and quantitative differences between simulated and historical flows observed in the previous paragraphs could have been mitigated with early releases and coordinated operations in the model.

The consequence of both largest USRB reservoirs not coordinating and losing their capacity to regulate high flows at the same time is that the simulated flood peak from Palisades getting full (Figure 10 panel (b) To answer this, we developed a simple offline water balance model for the two flood control reservoirs, taking offline inflow, release and storage trajectories for every simulated ensemble members. Then we replaced releases with a simple policy starting the last week of March – matching the timing at which operators started emptying Palisades. Palisades release is set a full $100 m^3/s$ lower than the maximum historical daily release of $682 m^3/s$, and a policy is set at Jackson Lake to match observations from Figure 9). Releases are set to be 1) combines with high releases that cause American Falls to fill completely. These compounding structural errors in representing reservoir operations yield a $2,500\ m^3/s$ simulated flood peak downstream of American Falls around May 13. The simulated flood peak then propagates downstream where the smaller reservoirs do not have the capacity to attenuate it. After filling , the reservoirs are forced to let all subsequent flow peaks pass without any simulated management or control. By contrast, coordinated operations between upstream reservoirs ensure that historical American Falls releases in 2011 only peak at $820\ m^3/s$ on average on an eight-day period starting June 2. All the while, American Falls reservoir filled very gradually from early March to be nearly full in July, never losing its capacity to regulate streamflow. This early July historical storage peak across all three main reservoirsin the USRB also maximizes water supply for irrigation purposes.

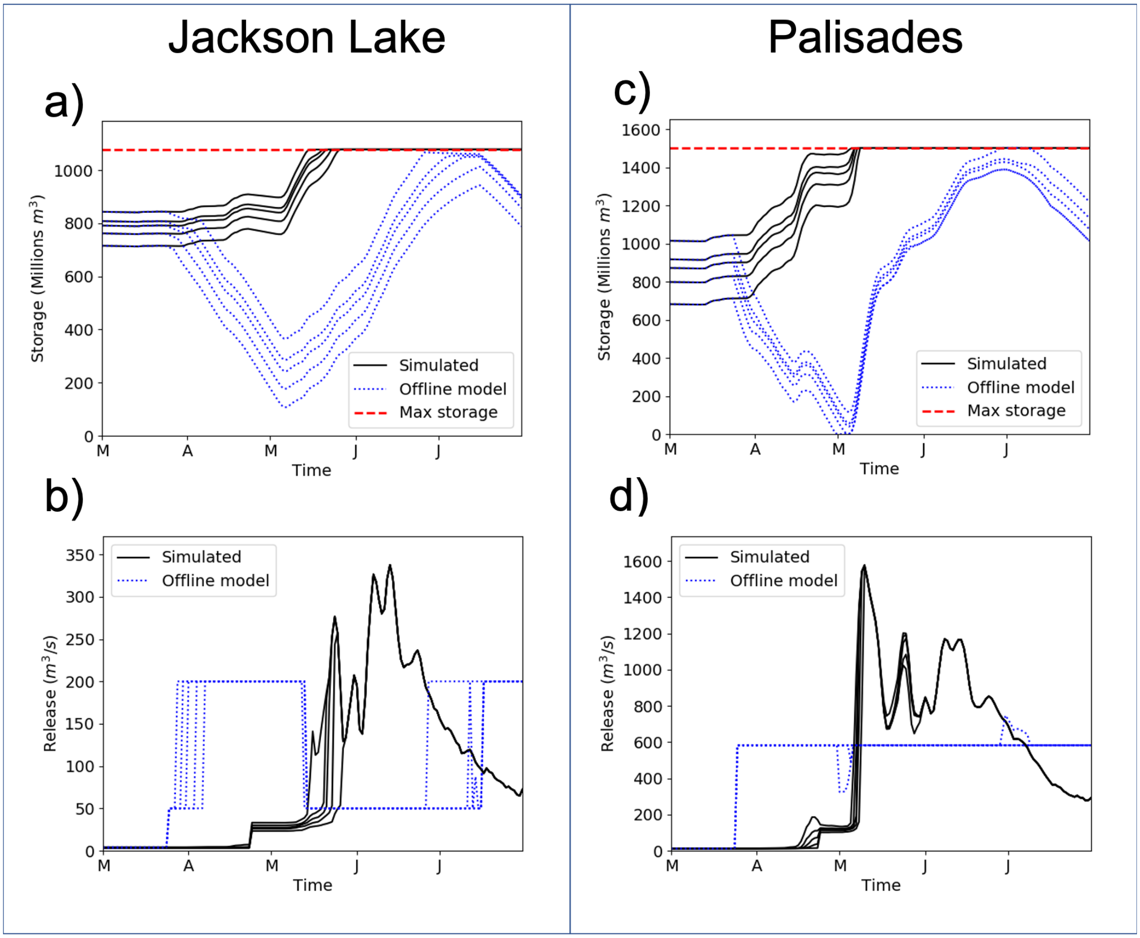



**Figure 11.** Results from the offline reservoir water balance model (dotted blue lines), compared with simulation results (continuous black lines): min, max values and quartiles for both ensembles. Differences in storage (panels (a) and (c) for Jackson Lake and Palisades respectively) are due to a simple coordinated released policy starting the last week of March (panels (b) and (d) for Jackson Lake and Palisades respectively).

$200m^3/s$ when Palisades is empty enough (less than 40% full) or Jackson Lake is nearly (over 98%) full, and 2) cut release back to $50m^3/s$ otherwise. The routing delay was fixed at one day, a conservative assumption making any excessive release from Jackson Lake immediately consequential for Palisades reservoir levels. Results from Figure 11 show that even with simulated inflows, a simple flood control policy enables to avoid filling Jackson Lake across the whole ensembles, and avoids filling Palisades in most cases. The only ensemble members for which Palisades gets filled are the ones that start with much higher initial storage at both reservoirs; even then, filling only happens in late June and peak flows are less than half those simulated without coordination.

## 5 Discussion

This work analyzes a state-of-the-art release rule from a large-scale, high-resolution hydrological model to understand the potential consequences of not capturing real-world operational coordination across reservoirs when simulating flood and drought events. It focuses on the USRB, a Western U.S. basin featuring a reservoir cascade managed with a high level of coordination to avoid both floods and water shortages, two risks made prominent by the area's geography and climate. An ensemble simulated and analyzed using the screening method known as the method of Morris shows some of the consequences of not capturing coordination, leading to 1) quick and complete drawdown of reservoirs in irrigation hotspots during hot, dry summers, and 2) simulating potentially catastrophic floods with untimely filling across the cascade. The historical record~~demonstrates~~ , and experiments based on offline water balance models of the reservoir cascade, demonstrate that in both instances, coordinated reservoir management avoided the occurrence of these events.

In both the high-flow and the low-flow events, coordination and control decisions are mediated both by other reservoirs' operations and by other decision-relevant variables. This is obvious for the averted flood of 2011 where snowpack monitoring led to forecasts of large snowmelt with enough lead time to make space in ~~key reservoirs~~two key reservoirs and coordinate their response. Similarly, the 2012-2013 decisions are mediated by water demands in the Snake River Plain. In both cases, the mix of institutional communication – between reservoirs and farmer representatives – as well as monitoring of key water supply and demand predictors are instrumental to implementing successful coordination actions in the face of adverse climatic events. Recent research on the water management institutions of the Upper Snake River basin suggests that they are well-equipped to show resilience in the face of expected climate change (Kliskey et al., 2019; Gilmore, 2019).

There is a growing body of literature highlighting the potentially highly interdependent nature of state-aware reservoir operations and institutional coordination in large multi-purpose reservoir cascades (Quinn et al., 2019). The importance of institutional context as well as location specific nature of selecting key variables for informing forecasts is a significant challenge to large-scale hydrological modeling. Poor abstractions of forecast informed reservoir operations and basin specific institutions that support coordinated emergency responses limit the value of hydrological modeling in understanding vulnerabilities to extremes. In a context where high-resolution modeling (Wood et al., 2011; Bierkens et al., 2015) is framed as a key element for informing, monitoring and forecasting these risks at exquisitely fine spatial and temporal resolutions, it is urgent to move beyond validation based exclusively on goodness-of-fit. Model evaluations need to 1) identify key human and natural processes leading to flow extremes, and 2) validate that these processes are present in the hydrological model. As recent developments in the literature on reservoir representations in hydrological models illustrate, there has been a growing sophistication in representations of release rules without addressing the key concern of capturing the key variables managers use to address unusual flow conditions in complex coupled human and natural systems. Parametrizations that are "good" in the sense that they score well with respect to one or more goodness-of-fit indicators may not necessarily represent the underlying processes correctly (e.g., Legates and McCabe Jr., 1999; Gupta et al., 2009). This point has recently been illustrated for reservoir representations in large-scale hydrological models through the flawed structural behavior of an upper Mekong (Lancang) basin model where reservoirs had been omitted ~~(?)~~(Dang et al., 2020). This is why we did not attempt to calibrate reservoir rule parameters in this work. Besides becoming an increasingly difficult task going downstream, it would only have served to mask a portion of structural model errors without actually addressing them (a.k.a "right for the wrong reasons"). To the contrary, this paper takes the view that the unintended consequences from these errors need to be exposed before "well-calibrated" but structurally deficient representations are used to assess out-of-sample flood and drought risks with future flow conditions that are often very different from those used for model calibration and evaluation. In this case, we exposed the need to refine representations of human-mediated coordination and controls in hydrological models, so they do not flag false vulnerabilities in a world where rapidly-developing global crises are expected to yield large capital investments.

The most straightforward way to represent complex human coordination processes and the key variables they rely on is to integrate actual management rules directly into hydrological models . This has been done in water resource management models such as RIVERWARE (Zagona et al., 2001) or WEAP (Yates et al., 2005)(Zagona et al., 2001; Yates et al., 2005). Such rule systems demonstrably improve hydrological models (Qiu et al., 2019), but they necessitate a direct knowledge of operations that is unavailable in most cases. Alternatively, machine learning techniques have been developed to infer reservoir operator behavior from historical observations, but often assume that decisions are taken as a function of a set of standard hydrologic variables on a reservoir-by-reservoir basis (Hejazi et al., 2008; Ehsani et al., 2016; Coerver et al., 2018)(Hejazi et al., 2008; Ehsani et al., 2016; Coer

. Recently though, applications to multi-reservoir systems in California have seen these techniques extended to consider impacts of forecast variables such as snowpack depth on operations (Yang et al., 2016), and to infer drought vulnerability from monthly operations (Giuliani and Herman, 2018). Our work demonstrates that further research is needed in this direction to fully account for complex feedbacks between climate variables, water supply and flood control objectives, and release decisions. Emerging techniques enabling storage level monitoring even in inaccessible areas including war zones (Müller et al., 2016; Avisse et al., 2017) could then make it possible to generalize machine learning-based approaches.

An alternative to reproducing historical operations is to improve operations through optimization instead. Such optimization needs to consider the distinct and sometimes conflicting management objectives, including but not limited to protection against water shortages and floods. Using the example of a single reservoir with multiple commitments in terms of flood control, water supply and hydropower production, Giuliani et al. (2014) showed that the multiple vulnerabilities associated with historical operations could be mitigated using multiobjective heuristics. This emerging approach, called evolutionary multi-objective direct policy search (EMODPS Giuliani et al., 2016), proposes reservoir rules that trade-off flood and drought vulnerabilities with other reservoir management objectives. It has been successfully applied to a flood- and drought- prone multireservoir multi-reservoir system (Quinn et al., 2017). Its results could be integrated to large-scale hydrological models to provide a best-case scenario whose merit is to highlight water resource vulnerabilities that are likely to not be a model artefact (Rougé et al., 2018), but instead stress the need for structural adaptation measures.

## 6 Conclusions

The interactions between the multiple stakeholders in major river basin systems are complex, as is the interplay between the key variables they use to monitor and manage flood and drought risks. Although large-scale hydrological models have generally sought to abstract this complexity in their representation of human processes, they at present struggle to capture coordination and control processes in ~~multireservoir~~ multi-reservoir systems. The current standard practice treats each reservoir's release independently from other reservoirs' storage levels. This paper demonstrates the unintended consequences this can have for flood and drought assessment, using a well-established hydrological model with advanced representations of multi-reservoir operations in the Upper Snake River Basin. Our diagnostic assessment of a state-of-the-art release rule abstractions in large-scale hydrological models exploits time-varying sensitivity analysis based on the Method of Morris to show how the behavioral controls in parameterized reservoir representations can inadvertently lead to amplifying errors. The diagnostic methodology used can be replicated with other hydrological models, release rule representations, and river basins. It provides insights of cumulative reservoir rules impacts across the basin at a daily time step. Application to the reservoir cascade on the Upper Snake River Basin, with its complex institutions and careful monitoring of frequent flood and water risks, showed how failure to represent the appropriate monitoring and reservoir coordination processes could lead a hydrological model to simulate flood and drought events that actual basin operators would unequivocally avoid.

This finding is consequential at a time where reservoir rules of increasing sophistication are being proposed to come to a better agreement between observed and simulated releases, and where the monitoring and forecasting of water-related risks at extremely high resolutions is hailed at the future of hydrology. It demonstrates the necessity to complement goodness-of-fit testing by devising validation techniques for the qualitative behavior of human-operated structures in the face of water emergencies. This is not a task for hydrologists alone, as developments across water resources management, operations research, machine learning and assimilation of remotely sensed data, among others, all have a role to play in tackling the challenge whose urgency this work highlights.

## Acknowledgments

This work was supported by the U.S. Department of Energy, Office of Science, Biological and Environmental Research Program, Earth and Environmental Systems Modeling, MultiSector Dynamics, Contract No. DE-SC0016162 and the material based upon work supported by the National Science Foundation under Grant No. 1639524 in the Innovations at the Nexus of Food, Energy and Water Systems (INFEWS) program.

*Code and data availability.* Core WBM code is available from the authors by request, and so is the code for this sensitivity analysis. Result data and code necessary to draw the figures are available online at https://github.com/charlesrouge/UpperSnakeRiver_reservoirs_WBM.

## References


Flooding in Idaho, https://www.weather.gov/safety/flood-states-id, retrieved 13 April 2020.

Abatzoglou, J. T.: Development of gridded surface meteorological data for ecological applications and modelling, International Journal of Climatology, 33, 121–131, https://doi.org/10.1002/joc.3413, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.3413, 2013.

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, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2007JD008525, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JD008525, 2007.

Ahmadalipour, A. and Moradkhani, H.: Analyzing the uncertainty of ensemble-based gridded observations in land surface simulations and drought assessment, Journal of Hydrology, 555, 557 – 568, https://doi.org/https://doi.org/10.1016/j.jhydrol.2017.10.059, http://www.sciencedirect.com/science/article/pii/S0022169417307369, 2017.

Allen, R. G., Pereira, L. S., Raes, D., Smith, M., et al.: Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, FAO, Rome, 300, 1998.

Avisse, N., Tilmant, A., Müller, M. F., and Zhang, H.: Monitoring small reservoirs' storage with satellite remote sensing in inaccessible areas, Hydrology and Earth System Sciences, 21, 6445–6459, https://doi.org/10.5194/hess-21-6445-2017, https://www.hydrol-earth-syst-sci.net/21/6445/2017/, 2017.

Biemans, H., Haddeland, I., Kabat, P., Ludwig, F., Hutjes, R. W. A., Heinke, J., von Bloh, W., and Gerten, D.: Impact of reservoirs on river discharge and irrigation water supply during the 20th century, Water Resources Research, 47, https://doi.org/10.1029/2009WR008929, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009WR008929, 2011.

Bierkens, M. F. P.: Global hydrology 2015: State, trends, and directions, Water Resources Research, 51, 4923–4947, https://doi.org/10.1002/2015WR017173, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015WR017173, 2015.

Bierkens, M. F. P., Bell, V. A., Burek, P., Chaney, N., Condon, L. E., David, C. H., de Roo, A., Döll, P., Drost, N., Famiglietti, J. S., Flörke, M., Gochis, D. J., Houser, P., Hut, R., Keune, J., Kollet, S., Maxwell, R. M., Reager, J. T., Samaniego, L., Sudicky, E., Sutanudjaja, E. H., van de Giesen, N., Winsemius, H., and Wood, E. F.: Hyper-resolution global hydrological modelling: what is next?, Hydrological Processes, 29, 310–320, https://doi.org/10.1002/hyp.10391, https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.10391, 2015.

Bierkens, M. F. P., Reinhard, S., de Bruijn, J. A., Veninga, W., and Wada, Y.: The Shadow Price of Irrigation Water in Major Groundwater-Depleting Countries, Water Resources Research, 0, https://doi.org/10.1029/2018WR023086, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018WR023086, 2019.

Burek, P. A., Roo, A. D., and van der Knijff, J.: LISFLOOD - Distributed Water Balance and Flood Simulation Model - Revised User Manual, Tech. rep., Publications Office of the European Union. Directorate-General Joint Research Centre, Institute for Environment and Sustainability, 2013.

Campolongo, F., Cariboni, J., and Saltelli, A.: An effective screening design for sensitivity analysis of large models, Environmental Modelling & Software, 22, 1509 – 1518, https://doi.org/https://doi.org/10.1016/j.envsoft.2006.10.004, http://www.sciencedirect.com/science/article/pii/S1364815206002805, modelling, computer-assisted simulations, and mapping of dangerous phenomena for hazard assessment, 2007.

Center For International Earth Science Information Network-CIESIN-Columbia University: Gridded Population of the World, Version 4 (GPWv4): Population Density,[object Object], 2016.

Coe, M. T.: Modeling Terrestrial Hydrological Systems at the Continental Scale: Testing the Accuracy of an Atmospheric GCM, Journal of Climate, 13, 686–704, https://doi.org/10.1175/1520-0442(2000)013<0686:MTHSAT>2.0.CO;2, https://doi.org/10.1175/1520-0442(2000)013<0686:MTHSAT>2.0.CO;2, 2000.

Coerver, H. M., Rutten, M. M., and van de Giesen, N. C.: Deduction of reservoir operating rules for application in global hydrological models, Hydrology and Earth System Sciences, 22, 831–851, https://doi.org/10.5194/hess-22-831-2018, https://www.hydrol-earth-syst-sci.net/22/831/2018/, 2018.

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, https://doi.org/10.5194/hess-24-397-2020, https://www.hydrol-earth-syst-sci.net/24/397/2020/, 2020.

Dieter, C. A., Maupin, M. A., Caldwell, R. R., Harris, M. A., Ivahnenko, T. I., Lovelace, J. K., Barber, N. L., and Linsey, K. S.: Estimated use of water in the United States in 2015, USGS Numbered Series 1441, U.S. Geological Survey, Reston, VA, http://pubs.er.usgs.gov/publication/cir1441, iP-090439, 2018.

Döll, P., Kaspar, F., and Lehner, B.: A global hydrological model for deriving water availability indicators: model tuning and validation, Journal of Hydrology, 270, 105 – 134, https://doi.org/https://doi.org/10.1016/S0022-1694(02)00283-4, http://www.sciencedirect.com/science/article/pii/S0022169402002834, 2003.

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, 2413–2432, https://doi.org/10.5194/hess-13-2413-2009, https://www.hydrol-earth-syst-sci.net/13/2413/2009/, 2009.

Dynesius, M. and Nilsson, C.: Fragmentation and Flow Regulation of River Systems in the Northern Third of the World, Science, 266, 753–762, https://doi.org/10.1126/science.266.5186.753, http://science.sciencemag.org/content/266/5186/753, 1994.

Ehsani, N., Fekete, B. M., Vörösmarty, C. J., and Tessler, Z. D.: A neural network based general reservoir operation scheme, Stochastic Environmental Research and Risk Assessment, 30, 1151–1166, https://doi.org/10.1007/s00477-015-1147-9, https://doi.org/10.1007/s00477-015-1147-9, 2016.

Ehsani, N., VÃ¶rÃ¶smarty, 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, https://doi.org/https://doi.org/10.1016/j.jhydrol.2017.09.008, http://www.sciencedirect.com/science/article/pii/S0022169417305991, 2017.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), Journal of Climate, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Gilmore, S.: Assessing the Adaptive Capacity of Idaho's Magic Valley As a Complex Social-Ecological System, Master's thesis, University of Idaho, Moscow, ID, 2019.

Giuliani, M. and Herman, J. D.: Modeling the behavior of water reservoir operators via eigenbehavior analysis, Advances in Water Resources, 122, 228 – 237, https://doi.org/https://doi.org/10.1016/j.advwatres.2018.10.021, http://www.sciencedirect.com/science/article/pii/S0309170817311193, 2018.

Giuliani, M., Herman, J. D., Castelletti, A., and Reed, P.: Many-objective reservoir policy identification and refinement to reduce policy inertia and myopia in water management, Water Resources Research, 50, 3355–3377, https://doi.org/10.1002/2013WR014700, http://dx.doi.org/10.1002/2013WR014700, 2014.

Giuliani, M., Castelletti, A., Pianosi, F., Mason, E., and Reed, P. M.: Curses, Tradeoffs, and Scalable Management: Advancing Evolutionary Multiobjective Direct Policy Search to Improve Water Reservoir Operations, Journal of Water Resources Planning and Management, 142, 04015 050, https://doi.org/10.1061/(ASCE)WR.1943-5452.0000570, http://ascelibrary.org/doi/abs/10.1061/%28ASCE%29WR.1943-5452.0000570, 2016.

Grill, G., lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., Babu, S., Borrelli, P., l. Cheng, Crochetiere, H., Macedo, H. E., Filgueiras, R., Goichot, M., Higgins, J., Hogan, Z., Lip, B., Mcclain, M. E., Meng, J., Mulligan, M., Nilsson, C., Olden, J. D., Opperman, J. J., Petry, P., Liermann, C. R., l. S/'aenz, Salinas-Rodr/'iguez, S., Schelle, P., Schmitt, R. J. P., Snider, J., anf K. Tockner, F. T., Valdujo, P. H., van Soesbergen, A., and Zarfl, C.: Mapping the world's free-flowing rivers, Nature, 569, https://doi.org/10.1038/s41586-019-1111-9, 2019.

Grogan, D. S.: Global and regional assessments of unsustainable groundwater use in irrigated agriculture, phdthesis, University of New Hampshire, http://scholars.unh.edu/dissertation/2, accessed 2019-05-21, 2016.

Grogan, D. S., Zhang, F., Prusevich, A., Lammers, R. B., Wisser, D., Glidden, S., Li, C., and Frolking, S.: Quantifying the link between crop production and mined groundwater irrigation in China, Science of The Total Environment, 511, 161 – 175, https://doi.org/https://doi.org/10.1016/j.scitotenv.2014.11.076, http://www.sciencedirect.com/science/article/pii/S0048969714016702, 2015.

Grogan, D. S., Wisser, D., Prusevich, A., Lammers, R. B., and Frolking, S.: The use and re-use of unsustainable groundwater for irrigation: a global budget, Environmental Research Letters, 12, 034 017, http://stacks.iop.org/1748-9326/12/i=3/a=034017, 2017.

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, https://doi.org/https://doi.org/10.1016/j.jhydrol.2009.08.003, http://www.sciencedirect.com/science/article/pii/S0022169409004843, 2009.

Guse, B., Reusser, D. E., and Fohrer, N.: How to improve the representation of hydrological processes in SWAT for a lowland catchment – temporal analysis of parameter sensitivity and model performance, Hydrological Processes, 28, 2651–2670, https://doi.org/10.1002/hyp.9777, https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.9777, 2014.

Haddeland, I., Skaugen, T., and Lettenmaier, D. P.: Anthropogenic impacts on continental surface water fluxes, Geophysical Research Letters, 33, https://doi.org/10.1029/2006GL026047, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006GL026047, 2006.

Hamlet, A. F. and Lettenmaier, D. P.: EFFECTS OF CLIMATE CHANGE ON HYDROLOGY AND WATER RESOURCES IN THE COLUMBIA RIVER BASIN, JAWRA Journal of the American Water Resources Association, 35, 1597–1623, https://doi.org/10.1111/j.1752-1688.1999.tb04240.x, https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1752-1688.1999.tb04240.x, 1999.

Han, W., Yang, Z., Di, L., and Mueller, R.: CropScape: A Web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support, Computers and Electronics in Agriculture, 84, 111–123, https://doi.org/10.1016/j.compag.2012.03.005, http://www.sciencedirect.com/science/article/pii/S0168169912000798, 2012.

Hanasaki, N., Kanae, S., and Oki, T.: A reservoir operation scheme for global river routing models, Journal of Hydrology, 327, 22 – 41, https://doi.org/https://doi.org/10.1016/j.jhydrol.2005.11.011, http://www.sciencedirect.com/science/article/pii/S0022169405005962, 2006.

Hanasaki, N., Kanae, S., Oki, T., Masuda, K., Motoya, K., Shirakawa, N., Shen, Y., and Tanaka, K.: An integrated model for the assessment of global water resources – Part 1: Model description and input meteorological forcing, Hydrology and Earth System Sciences, 12, 1007–1025, https://doi.org/10.5194/hess-12-1007-2008, https://www.hydrol-earth-syst-sci.net/12/1007/2008/, 2008.

Hanasaki, N., Fujimori, S., Yamamoto, T., Yoshikawa, S., Masaki, Y., Hijioka, Y., Kainuma, M., Kanamori, Y., Masui, T., Takahashi, K., and Kanae, S.: A global water scarcity assessment under Shared Socio-economic Pathways – Part 2: Water availability and scarcity, Hydrology and Earth System Sciences, 17, 2393–2413, https://doi.org/10.5194/hess-17-2393-2013, https://www.hydrol-earth-syst-sci.net/17/2393/2013/, 2013.

Hanasaki, N., Yoshikawa, S., Pokhrel, Y., and Kanae, S.: A Quantitative Investigation of the Thresholds for Two Conventional Water Scarcity Indicators Using a State-of-the-Art Global Hydrological Model With Human Activities, Water Resources Research, 54, 8279–8294, https://doi.org/10.1029/2018WR022931, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018WR022931, 2018.

Hejazi, M. I., Cai, X., and Ruddell, B. L.: The role of hydrologic information in reservoir operation - Learning from historical releases, Advances in Water Resources, 31, 1636–1650, 2008.

Hejazi, M. I., Voisin, N., Liu, L., Bramer, L. M., Fortin, D. C., Hathaway, J. E., Huang, M., Kyle, P., Leung, L. R., Li, H.-Y., Liu, Y., Patel, P. L., Pulsipher, T. C., Rice, J. S., Tesfa, T. K., Vernon, C. R., and Zhou, Y.: 21st century United States emissions mitigation could increase water stress more than the climate change it is mitigating, Proceedings of the National Academy of Sciences, 112, 10 635–10 640, https://doi.org/10.1073/pnas.1421675112, https://www.pnas.org/content/112/34/10635, 2015.

Herbert, C. and Döll, P.: Global assessment of current and future groundwater stress with a focus on transboundary aquifers, Water Resources Research, 0, https://doi.org/10.1029/2018WR023321, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018WR023321, 2019.

Herman, J. and Usher, W.: SALib: An open-source Python library for Sensitivity Analysis, The Journal of Open Source Software, 2, https://doi.org/10.21105/joss.00097, https://doi.org/10.21105/joss.00097, 2017.

Herman, J. D., Kollat, J. B., Reed, P. M., and Wagener, T.: Technical Note: Method of Morris effectively reduces the computational demands of global sensitivity analysis for distributed watershed models, Hydrology and Earth System Sciences, 17, 2893–2903, https://doi.org/10.5194/hess-17-2893-2013, https://www.hydrol-earth-syst-sci.net/17/2893/2013/, 2013a.

Herman, J. D., Kollat, J. B., Reed, P. M., and Wagener, T.: From maps to movies: high-resolution time-varying sensitivity analysis for spatially distributed watershed models, Hydrology and Earth System Sciences, 17, 5109–5125, https://doi.org/10.5194/hess-17-5109-2013, https://www.hydrol-earth-syst-sci.net/17/5109/2013/, 2013b.

Hoekema, D. J. and Sridhar, V.: Relating climatic attributes and water resources allocation: A study using surface water supply and soil moisture indices in the Snake River basin, Idaho, Water Resources Research, 47, https://doi.org/10.1029/2010WR009697, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010WR009697, 2011.

Independent Panel To Review Cause of Teton Dam Failure: Report to U.S. Department of Interior and State of Idaho on Failure of Teton Dam, Tech. rep., United States Bureau of Reclamations, https://www.usbr.gov/pn/snakeriver/dams/uppersnake/teton/1976failure.pdf, 1976.

Iooss, B. and Lemaître, P.: A Review on Global Sensitivity Analysis Methods, pp. 101–122, Springer US, Boston, MA, https://doi.org/10.1007/978-1-4899-7547-8_5, https://doi.org/10.1007/978-1-4899-7547-8_5, 2015.

Jägermeyr, J., Gerten, D., Heinke, J., Schaphoff, S., Kummu, M., and Lucht, W.: Water savings potentials of irrigation systems: global simulation of processes and linkages, Hydrol. Earth Syst. Sci., 19, 3073–3091, https://doi.org/10.5194/hess-19-3073-2015, https://www.hydrol-earth-syst-sci.net/19/3073/2015/, 2015.

Jägermeyr, J., Gerten, D., Schaphoff, S., Heinke, J., Lucht, W., and Rockström, J.: Integrated crop water management might sustainably halve the global food gap, Environmental Research Letters, 11, 025 002, http://stacks.iop.org/1748-9326/11/i=2/a=025002, 2016.

Kliskey, A., Abatzoglou, J., Alessa, L., Kolden, C., Hoekema, D., Moore, B., Gilmore, S., and Austin, G.: Planning for Idaho's waterscapes: A review of historical drivers and outlook for the next 50 years, Environmental Science & Policy, 94, 191 – 201, https://doi.org/https://doi.org/10.1016/j.envsci.2019.01.009, http://www.sciencedirect.com/science/article/pii/S146290111830042X, 2019.

Kraucunas, I., Clarke, L., Dirks, J., Hathaway, J., Hejazi, M., Hibbard, K., Huang, M., Jin, C., Kintner-Meyer, M., van Dam, K. K., Leung, R., Li, H.-Y., Moss, R., Peterson, M., Rice, J., Scott, M., Thomson, A., Voisin, N., and West, T.: Investigating the nexus of climate, energy, water, and land at decision-relevant scales: the Platform for Regional Integrated Modeling and Analysis (PRIMA), Climatic Change, 129, 573–588, https://doi.org/10.1007/s10584-014-1064-9, https://doi.org/10.1007/s10584-014-1064-9, 2015.

Lamontagne, J. R., Reed, P. M., Marangoni, G., Keller, K., and Garner, G. G.: Robust abatement pathways to tolerable climate futures require immediate global action, Nature Climate Change, https://doi.org/10.1038/s41558-019-0426-8, 2019.

Latrubesse, E. M., Arima, E. Y., Dunne, T., Park, E., Baker, V. R., d'Horta, F. M., Wight, C., Wittmann, F., Zuanon, J., Baker, P. A., Ribas, C. C., Norgaard, R. B., Filizola, N., Ansar, A., Flyvbjerg, B., and Stevaux, J. C.: Damming the rivers of the Amazon basin, Nature, 546, 363–369, https://doi.org/10.1038/nature22333, 2017.

Legates, D. R. and McCabe Jr., G. J.: Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation, Water Resources Research, 35, 233–241, https://doi.org/10.1029/1998WR900018, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1998WR900018, 1999.

Lehner, B., Verdin, K., and Jarvis, A.: New Global Hydrography Derived From Spaceborne Elevation Data, Eos, Transactions American Geophysical Union, 89, 93–94, https://doi.org/10.1029/2008EO100001, http://onlinelibrary.wiley.com/doi/10.1029/2008EO100001/abstract, 2008.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, 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, https://doi.org/10.1890/100125, https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/100125, 2011.

Li, H.-Y., Leung, L. R., Getirana, A., Huang, M., Wu, H., Xu, Y., Guo, J., and Voisin, N.: Evaluating Global Streamflow Simulations by a Physically Based Routing Model Coupled with the Community Land Model, Journal of Hydrometeorology, 16, 948–971, https://doi.org/10.1175/JHM-D-14-0079.1, https://doi.org/10.1175/JHM-D-14-0079.1, 2015.

Liu, J., Hertel, T. W., Lammers, R. B., Prusevich, A., Baldos, U. L. C., Grogan, D. S., and Frolking, S.: Achieving sustainable irrigation water withdrawals: global impacts on food security and land use, Environmental Research Letters, 12, 104 009, http://stacks.iop.org/1748-9326/12/i=10/a=104009, 2017.

Liu, X., Liu, W., Yang, H., Tang, Q., Flörke, M., Masaki, Y., Müller Schmied, H., Ostberg, S., Pokhrel, Y., Satoh, Y., and Wada, Y.: Multimodel assessments of human and climate impacts on mean annual streamflow in China, Hydrology and Earth System Sciences, 23, 1245–1261, https://doi.org/10.5194/hess-23-1245-2019, https://www.hydrol-earth-syst-sci.net/23/1245/2019/, 2019.

Lund, J. R. and Guzman, J.: Derived operating rule for reservoir in series or in parallel, Journal of Water Resources Planning and Management, 125, 143–153, 1999.

Masaki, Y., Hanasaki, N., Biemans, H., Schmied, H. M., Tang, Q., Wada, Y., Gosling, S. N., Takahashi, K., and Hijioka, Y.: Intercomparison of global river discharge simulations focusing on dam operation – multiple models analysis in two case-study river basins, Missouri-Mississippi and Green-Colorado, Environmental Research Letters, 12, 055 002, http://stacks.iop.org/1748-9326/12/i=5/a=055002, 2017.

Mateo, C. M., Hanasaki, N., Komori, D., Tanaka, K., Kiguchi, M., Champathong, A., Sukhapunnaphan, T., Yamazaki, D., and Oki, T.: Assessing the impacts of reservoir operation to floodplain inundation by combining hydrological, reservoir management, and hydrodynamic models, Water Resources Research, 50, 7245–7266, https://doi.org/10.1002/2013WR014845, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013WR014845, 2014.

Maupin, M. A., Kenny, J. F., Hutson, S. S., Lovelace, J. K., Barber, N. L., and Linsey, K. S.: Estimated use of water in the United States in 2010, no. 1405 in Circular, U.S. Geological Survey, Reston, Virginia, 2014.

McGuire, M., Wood, A. W., Hamlet, A. F., and Lettenmaier, D. P.: Use of Satellite Data for Streamflow and Reservoir Storage Forecasts in the Snake River Basin, Journal of Water Resources Planning and Management, 132, 97–110, https://doi.org/10.1061/(ASCE)0733-9496(2006)132:2(97), https://ascelibrary.org/doi/abs/10.1061/%28ASCE%290733-9496%282006%29132%3A2%2897%29, 2006.

Meigh, J. R., McKenzie, A. A., and Sene, K. J.: A Grid-Based Approach to Water Scarcity Estimates for Eastern and Southern Africa, Water Resources Management, 13, 85–115, https://doi.org/10.1023/A:1008025703712, https://doi.org/10.1023/A:1008025703712, 1999.

Metin, A. D., Dung, N. V., Schröter, K., Guse, B., Apel, H., Kreibich, H., Vorogushyn, S., and Merz, B.: How do changes along the risk chain affect flood risk?, Natural Hazards and Earth System Sciences, 18, 3089–3108, https://doi.org/10.5194/nhess-18-3089-2018, https://www.nat-hazards-earth-syst-sci.net/18/3089/2018/, 2018.

Meza, I., Siebert, S., Döll, P., Kusche, J., Herbert, C., Eyshi Rezaei, E., Nouri, H., Gerdener, H., Popat, E., Frischen, J., Naumann, G., Vogt, J. V., Walz, Y., Sebesvari, Z., and Hagenlocher, M.: Global-scale drought risk assessment for agricultural systems, Natural Hazards and Earth System Sciences Discussions, 2019, 1–26, https://doi.org/10.5194/nhess-2019-255, https://www.nat-hazards-earth-syst-sci-discuss.net/nhess-2019-255/, 2019.

Mishra, V., Aaadhar, S., Shah, H., Kumar, R., Pattanaik, D. R., and Tiwari, A. D.: The Kerala flood of 2018: combined impact of extreme rainfall and reservoir storage, Hydrology and Earth System Sciences Discussions, 2018, 1–13, https://doi.org/10.5194/hess-2018-480, https://www.hydrol-earth-syst-sci-discuss.net/hess-2018-480/, 2018.

Monteith, J.: Evaporation and environment: the state and movement of water in living organisms, Symposium of the Society for Experimental Biology, 19, 205–224, 1965.

Morris, M. D.: Factorial Sampling Plans for Preliminary Computational Experiments, Technometrics, 33, 161–174, https://doi.org/10.1080/00401706.1991.10484804, http://www.tandfonline.com/doi/abs/10.1080/00401706.1991.10484804, 1991.

Müller, M. F., Yoon, J., Gorelick, S. M., Avisse, N., and Tilmant, A.: Impact of the Syrian refugee crisis on land use and transboundary freshwater resources, Proceedings of the National Academy of Sciences, 113, 14 932–14 937, https://doi.org/10.1073/pnas.1614342113, http://www.pnas.org/content/113/52/14932.abstract, 2016.

Musselman, K. N., Lehner, F., Ikeda, K., Clark, M. P., Prein, A. F., Liu, C., Barlage, M., and Rasmussen, R.: Projected increases and shifts in rain-on-snow flood risk over western North America, Nature Climate Change, 8, 808–812, https://doi.org/10.1038/s41558-018-0236-4, 2018.

Najibi, N., Devineni, N., and Lu, M.: Hydroclimate drivers and atmospheric teleconnections of long duration floods: An application to large reservoirs in the Missouri River Basin, Advances in Water Resources, 100, 153 – 167, https://doi.org/https://doi.org/10.1016/j.advwatres.2016.12.004, http://www.sciencedirect.com/science/article/pii/S0309170816302895, 2017.

Nazemi, A. and Wheater, H. S.: On inclusion of water resource management in Earth system models – Part 2: Representation of water supply and allocation and opportunities for improved modeling, Hydrology and Earth System Sciences, 19, 63–90, https://doi.org/10.5194/hess-19-63-2015, https://www.hydrol-earth-syst-sci.net/19/63/2015/, 2015.

Nilsson, C., Reidy, C. A., Dynesius, M., and Revenga, C.: Fragmentation and Flow Regulation of the World's Large River Systems, Science, 308, 405–408, https://doi.org/10.1126/science.1107887, http://science.sciencemag.org/content/308/5720/405, 2005.

Oyerinde, G. T., Wisser, D., Hountondji, F. C., Odofin, A. J., Lawin, A. E., Afouda, A., and Diekkrüger, B.: Quantifying Uncertainties in Modeling Climate Change Impacts on Hydropower Production, Climate, 4, https://doi.org/10.3390/cli4030034, http://www.mdpi.com/2225-1154/4/3/34, 2016.

Pianosi, F. and Wagener, T.: Understanding the time-varying importance of different uncertainty sources in hydrological modelling using global sensitivity analysis, Hydrological Processes, 30, 3991–4003, https://doi.org/10.1002/hyp.10968, https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.10968, 2016.

Pokhrel, Y., Hanasaki, N., Koirala, S., Cho, J., Yeh, P. J.-F., Kim, H., Kanae, S., and Oki, T.: Incorporating Anthropogenic Water Regulation Modules into a Land Surface Model, Journal of Hydrometeorology, 13, 255–269, https://doi.org/10.1175/JHM-D-11-013.1, https://doi.org/10.1175/JHM-D-11-013.1, 2012.

Portmann, F. T., Siebert Stefan, and Döll Petra: MIRCA2000–Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling, Global Biogeochemical Cycles, 24, https://doi.org/10.1029/2008GB003435, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008GB003435, 2010.

Proussevitch, A., Shiklomanov, A., Frolking, S., Glidden, S., Lammers, R., and Wisser, D.: Log-Exponential Reservoir Operating Rules for Global And Regional Hydrological Modeling, American Geophysical Union, Fall Meeting 2013, abstract #GC21B-0827, 2013.

Qiu, J., Yang, Q., Zhang, X., Huang, M., Adam, J. C., and Malek, K.: Implications of water management representations for watershed hydrologic modeling in the Yakima River basin, Hydrology and Earth System Sciences, 23, 35–49, https://doi.org/10.5194/hess-23-35-2019, https://www.hydrol-earth-syst-sci.net/23/35/2019/, 2019.

Quinn, J., Reed, P., Giuliani, M., and Castelletti, A.: What is controlling our control rules? Opening the black box of multi-reservoir operating policies using time-varying sensitivity analysis, Water Resources Research, 55, 5962–5984, https://doi.org/10.1029/2018WR024177, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018WR024177, 2019.

Quinn, J. D., Reed, P. M., Giuliani, M., and Castelletti, A.: Rival framings: A framework for discovering how problem formulation uncertainties shape risk management trade-offs in water resources systems, Water Resources Research, 53, 7208–7233, https://doi.org/10.1002/2017WR020524, http://dx.doi.org/10.1002/2017WR020524, 2017.

Rauscher, S. A., S., P. J., S., D. N., and M., B. M.: Future changes in snowmelt-driven runoff timing over the western US, Geophysical Research Letters, 35, https://doi.org/10.1029/2008GL034424, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008GL034424, 2008.

Reinecke, R., Foglia, L., Mehl, S., Herman, J. D., Wachholz, A., Trautmann, T., and Döll, P.: Spatially distributed sensitivity of simulated global groundwater heads and flows to hydraulic conductivity, groundwater recharge, and surface water body parameterization, Hydrology and Earth System Sciences, 23, 4561–4582, https://doi.org/10.5194/hess-23-4561-2019, https://www.hydrol-earth-syst-sci.net/23/4561/2019/, 2019.

Reusser, D. E. and Zehe, E.: Inferring model structural deficits by analyzing temporal dynamics of model performance and parameter sensitivity, Water Resources Research, 47, https://doi.org/10.1029/2010WR009946, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010WR009946, 2011.

Rougé, C., Tilmant, A., Zaitchik, B. F., Dezfuli, A. K., and Salman, M.: Identifying key water resource vulnerabilities in data-scarce transboundary river basins, Water Resources Research, 0, https://doi.org/10.1029/2017WR021489, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2017WR021489, 2018.

Ruano, M., Ribes, J., Seco, A., and Ferrer, J.: An improved sampling strategy based on trajectory design for application of the Morris method to systems with many input factors, Environmental Modelling & Software, 37, 103 – 109, https://doi.org/https://doi.org/10.1016/j.envsoft.2012.03.008, http://www.sciencedirect.com/science/article/pii/S1364815212000904, 2012.

Sampson, C. C., Smith, A. M., Bates, P. D., Neal, J. C., Alfieri, L., and Freer, J. E.: A high-resolution global flood hazard model, Water Resources Research, 51, 7358–7381, https://doi.org/10.1002/2015WR016954, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015WR016954, 2015.

Schmitt, R. J. P., Bizzi, S., Castelletti, A., and Kondolf, G. M.: Improved trade-offs of hydropower and sand connectivity by strategic dam planning in the Mekong, Nature Sustainability, 1, 96–104, https://doi.org/10.1038/s41893-018-0022-3, 2018.

Schumann, G. J.-P., Stampoulis, D., Smith, A. M., Sampson, C. C., Andreadis, K. M., Neal, J. C., and Bates, P. D.: Rethinking flood hazard at the global scale, Geophysical Research Letters, 43, 10,249–10,256, https://doi.org/10.1002/2016GL070260, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL070260, 2016.

Shin, S., Pokhrel, Y., and Miguez-Macho, G.: High-Resolution Modeling of Reservoir Release and Storage Dynamics at the Continental Scale, Water Resources Research, 55, 787–810, https://doi.org/10.1029/2018WR023025, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018WR023025, 2019.

Steinfeld, H., Gerber, P., Wassenaar, T., Castel, V., Rosales, M., and de Haan, C.: Livestock's long shadow, Tech. rep., Food and Agriculture Organisation of the United Nations (UN-FAO), http://www.fao.org/docrep/010/a0701e/a0701e00.HTM, 2006.

Stewart, R., Wollheim, W., Miara, A., Vorosmarty, C., Fekete, B., Lammers, R., and Rosenzweig, B.: Horizontal Cooling Towers: Riverine Ecosystem Services and the Fate of Thermoelectric Heat in the Contemporary Northeast U.S., Environmental Research Letters, 8, 025 010, https://doi.org/10.1088/1748-9326/8/2/025010, 2013.

Stewart, R. J., Wollheim, W. M., Gooseff, M. N., Briggs, M. A., Jacobs, J. M., Peterson, B. J., and Hopkinson, C. S.: Separation of river network-scale nitrogen removal among the main channel and two transient storage compartments, Water Resources Research, 47, https://doi.org/10.1029/2010WR009896, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010WR009896, 2011.

Thomas, C. A. and Lamke, R. D.: Floods of February 1962 in Southern Idaho and Northeastern Nevada, Tech. rep., United States Geological Survey, 1962.

Timpe, K. and Kaplan, D.: The changing hydrology of a dammed Amazon, Science Advances, 3, https://doi.org/10.1126/sciadv.1700611, http://advances.sciencemag.org/content/3/11/e1700611, 2017.

Turner, S. W. D., Xu, W., and Voisin, N.: Inferred inflow forecast horizons guiding reservoir release decisions across the United States, Hydrology and Earth System Sciences, 24, 1275–1291, https://doi.org/10.5194/hess-24-1275-2020, https://www.hydrol-earth-syst-sci.net/24/1275/2020/, 2020.

U.S. Bureau of Reclamations: Project details - Minidoka Project, https://web.archive.org/web/20120925152815/http://www.usbr.gov/projects/Project.jsp?proj_Name=Minidoka+Project, accessed: 2019-07-19, 2012.

van Beek, L. P. H., Yoshihide, W., and P., B. M. F.: Global monthly water stress: 1. Water balance and water availability, Water Resources Research, 47, https://doi.org/10.1029/2010WR009791, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010WR009791, 2011.

Veldkamp, T. I. E., Zhao, F., Ward, P. J., de Moel, H., Aerts, J. C. J. H., Schmied, H. M., Portmann, F. T., Masaki, Y., Pokhrel, Y., Liu, X., Satoh, Y., Gerten, D., Gosling, S. N., Zaherpour, J., and Wada, Y.: Human impact parameterizations in global hydrological models improve estimates of monthly discharges and hydrological extremes: a multi-model validation study, Environmental Research Letters, 13, 055 008, http://stacks.iop.org/1748-9326/13/i=5/a=055008, 2018.

Voisin, N., Li, H., Ward, D., Huang, M., Wigmosta, M., and Leung, L. R.: On an improved sub-regional water resources management representation for integration into earth system models, Hydrology and Earth System Sciences, 17, 3605–3622, https://doi.org/10.5194/hess-17-3605-2013, https://www.hydrol-earth-syst-sci.net/17/3605/2013/, 2013a.

Voisin, N., Liu, L., Hejazi, M., Tesfa, T., Li, H., Huang, M., Liu, Y., and Leung, L. R.: One-way coupling of an integrated assessment model and a water resources model: evaluation and implications of future changes over the US Midwest, Hydrology and Earth System Sciences, 17, 4555–4575, https://doi.org/10.5194/hess-17-4555-2013, https://www.hydrol-earth-syst-sci.net/17/4555/2013/, 2013b.

Vörösmarty, C. J., Sharma, K. P., Fekete, B. M., Copeland, A. H., Holden, J., Marble, J., and Lough, J. A.: The Storage and Aging of Continental Runoff in Large Reservoir Systems of the World, Ambio, 26, 210–219, http://www.jstor.org/stable/4314590, 1997.

Wada, Y., van Beek, L. P. H., Viviroli, D., Dürr, H. H., Weingartner, R., and Bierkens, M. F. P.: Global monthly water stress: 2. Water demand and severity of water stress, Water Resources Research, 47, https://doi.org/10.1029/2010WR009792, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010WR009792, 2011.

Wada, Y., Wisser, D., and Bierkens, M. F. P.: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, Earth System Dynamics, 5, 15–40, https://doi.org/10.5194/esd-5-15-2014, https://www.earth-syst-dynam.net/5/15/2014/, 2014.

Wada, Y., Bierkens, M. F. P., de Roo, A., Dirmeyer, P. A., Famiglietti, J. S., Hanasaki, N., Konar, M., Liu, J., Müller Schmied, H., Oki, T., Pokhrel, Y., Sivapalan, M., Troy, T. J., van Dijk, A. I. J. M., van Emmerik, T., Van Huijgevoort, M. H. J., Van Lanen, H. A. J., Vörösmarty, C. J., Wanders, N., and Wheater, H.: Human–water interface in hydrological modelling: current status and future directions, Hydrology and Earth System Sciences, 21, 4169–4193, https://doi.org/10.5194/hess-21-4169-2017, https://www.hydrol-earth-syst-sci.net/21/4169/2017/, 2017.

Wanders, N., van Vliet, M. T. H., Wada, Y., Bierkens, M. F. P., and van Beek, L. P. H. R.: High-Resolution Global Water Temperature Modeling, Water Resources Research, 0, https://doi.org/10.1029/2018WR023250, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018WR023250, 2019.

Wang, K., Shi, H., Chen, J., and Li, T.: An improved operation-based reservoir scheme integrated with Variable Infiltration Capacity model for multiyear and multipurpose reservoirs, Journal of Hydrology, 571, 365 – 375, https://doi.org/https://doi.org/10.1016/j.jhydrol.2019.02.006, http://www.sciencedirect.com/science/article/pii/S0022169419301416, 2019.

Willmott, C. J., Rowe, C. M., and Mintz, Y.: Climatology of the terrestrial seasonal water cycle, Journal of Climatology, 5, 589–606, https://doi.org/10.1002/joc.3370050602, http://libra.msra.cn/Publication/2839197/climatology-of-the-terrestrial-seasonal-water-cycle, 1985.

Wise, E. K.: Tree ring record of streamflow and drought in the upper Snake River, Water Resources Research, 46, https://doi.org/10.1029/2010WR009282, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010WR009282, 2010.

Wise, E. K.: Hydroclimatology of the US Intermountain West, Progress in Physical Geography: Earth and Environment, 36, 458–479, https://doi.org/10.1177/0309133312446538, https://doi.org/10.1177/0309133312446538, 2012.

Wisser, D., Frolking, S., Douglas, E. M., Fekete, B. M., Vörösmarty, C. J., and Schumann, A. H.: Global irrigation water demand: Variability and uncertainties arising from agricultural and climate data sets, Geophysical Research Letters, 35, n/a–n/a, https://doi.org/10.1029/2008GL035296, http://dx.doi.org/10.1029/2008GL035296, l24408, 2008.

Wisser, D., Fekete, B. M., Vörösmarty, C. J., and Schumann, A. H.: Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network- Hydrology (GTN-H), Hydrology and Earth System Sciences, 14, 1–24, https://doi.org/10.5194/hess-14-1-2010, https://www.hydrol-earth-syst-sci.net/14/1/2010/, 2010.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P.: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, Water Resources Research, 47, https://doi.org/10.1029/2010WR010090, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010WR010090, 2011.

Wu, Y. and Chen, J.: An Operation-Based Scheme for a Multiyear and Multipurpose Reservoir to Enhance Macroscale Hydrologic Models, Journal of Hydrometeorology, 13, 270–283, https://doi.org/10.1175/JHM-D-10-05028.1, https://doi.org/10.1175/JHM-D-10-05028.1, 2012.

Yang, T., Gao, X., Sorooshian, S., and Li, X.: Simulating California reservoir operation using the classification and regression-tree algorithm combined with a shuffled cross-validation scheme, Water Resources Research, 52, 1626–1651, https://doi.org/10.1002/2015WR017394, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015WR017394, 2016.

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, https://doi.org/10.5194/hess-23-3735-2019, https://www.hydrol-earth-syst-sci.net/23/3735/2019/, 2019.

Yates, D., Sieber, J., Purkey, D., and Huber-Lee, A.: WEAP21 – A Demand-, Priority-, and Preference-Driven Water Planning Model, Water International, 30, 487–500, https://doi.org/10.1080/02508060508691893, https://doi.org/10.1080/02508060508691893, 2005.

Yoshikawa, S., Cho, J., Yamada, H. G., Hanasaki, N., and Kanae, S.: An assessment of global net irrigation water requirements from various water supply sources to sustain irrigation: rivers and reservoirs (1960–2050), Hydrology and Earth System Sciences, 18, 4289–4310, https://doi.org/10.5194/hess-18-4289-2014, https://www.hydrol-earth-syst-sci.net/18/4289/2014/, 2014.

Zagona, E. A., Fulp, T. J., Shane, R., Magee, T., and Goranflo, H. M.: RIVERWARE: A GENERALIZED TOOL FOR COMPLEX RESERVOIR SYSTEM MODELING, JAWRA Journal of the American Water Resources Association, 37, 913–929, https://doi.org/10.1111/j.1752-1688.2001.tb05522.x, https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1752-1688.2001.tb05522.x, 2001.

Zaherpour, J., Gosling, S. N., Mount, N., Schmied, H. M., Veldkamp, T. I. E., Dankers, R., Eisner, S., Gerten, D., Gudmundsson, L., Haddeland, I., Hanasaki, N., Kim, H., Leng, G., Liu, J., Masaki, Y., Oki, T., Pokhrel, Y., Satoh, Y., Schewe, J., and Wada, Y.: Worldwide evaluation of mean and extreme runoff from six global-scale hydrological models that account for human impacts, Environmental Research Letters, 13, 065 015, http://stacks.iop.org/1748-9326/13/i=6/a=065015, 2018.

Zajac, Z., Revilla-Romero, B., Salamon, P., Burek, P., Hirpa, F. A., and Beck, H.: The impact of lake and reservoir parameterization on global streamflow simulation, Journal of Hydrology, 548, 552 – 568, https://doi.org/https://doi.org/10.1016/j.jhydrol.2017.03.022, http://www.sciencedirect.com/science/article/pii/S0022169417301671, 2017.

Zarfl, C., Lumsdon, A. E., Berlekamp, J., Tydecks, L., and Tockner, K.: A global boom in hydropower dam construction, Aquatic Sciences, 77, 161–170, https://doi.org/10.1007/s00027-014-0377-0, https://doi.org/10.1007/s00027-014-0377-0, 2015.

Zaveri, E., Grogan, D. S., Fisher-Vanden, K., Frolking, S., Lammers, R. B., Wrenn, D. H., Prusevich, A., and Nicholas, R. E.: Invisible water, visible impact: groundwater use and Indian agriculture under climate change, Environmental Research Letters, 11, 084 005, https://doi.org/10.1088/1748-9326/11/8/084005, https://doi.org/10.1088%2F1748-9326%2F11%2F8%2F084005, 2016.

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, https://doi.org/https://doi.org/10.1016/j.advwatres.2016.10.014, http://www.sciencedirect.com/science/article/pii/S0309170816305504, 2016.