Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., author comment AC3
https://doi.org/10.5194/gmd-2022-59-AC3, 2022

Author Response on RC2

Danielle S. Grogan et al.

Author comment on "Water balance model (WBM) v.1.0.0: a scalable gridded global hydrologic model with water-tracking functionality" by Danielle S. Grogan et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-59-AC3, 2022

Reviewer #2

This manuscript presents the UNH Water Balance Model, a hydrology model that has been developed over several decades, but released publicly (open source code) for the first time.

The authors provide a literature review of the model history and evaluate the current model performance against river discharge observations and irrigation water supply requirements. The WBM performs better across North America and Europe in terms of discharge and irrigation water supply, but relatively poorly across Asia and portions of South America.

The manuscript then provides examples of regional simulations including the Indus River watershed as well as the Wyoming headwaters contributing to river flow to the Colorado, Colombia and Mississippi River. This provides an opportunity to demonstrate the novel water component tracking functionality that enables identification of source regions and source stocks for river discharge and irrigation water supply.

The tracking feature appears to be a significant advance in river flow diagnostics and is capable of determining source spatial regions, and source water components (precip, agriculture, groundwater). This should be a valuable tool for land management and government policy makers. Finally, the authors provide an overview of WBM run-time instructions describing the necessary input data and setup scripts to perform a simulation.

This reviewer was impressed with the breadth of the manuscript included 1) a literature performance review, 2) multi-domain simulations with emphasis on the new diagnostic tracking functionality and 3) an overview of a model setup.

We would like to thank Reviewer #2 for their kind words and thorough commentary on the paper. It will significantly improve the paper.

This reviewer would have liked much more discussion devoted to WBM performance.

We will generate a more detailed assessment of river discharge performance and how this model compares to other similar models. In response to this comment and others, we will be adding a section in the discussion regarding WBM

performance with respect to observations and other global hydrologic models.

The WBM had a strong high bias in simulating global irrigation water supply as compared to other studies (Table 6). This apparently was caused by a systematic underestimation of discharge for the China/Asia region, but very little discussion (only a mention regarding better parameters are needed) was devoted to this topic. Whereas the authors provided a comparison against similar hydrology models in terms of simulated irrigation supply, no comparison was provided for discharge rates against other models for better context and perhaps lead to a discussion of what model components or parameters are most in need of improvements. This reviewer would have liked more justification or explanation to describe the skill of WBM such that a new user could avoid certain regions or pay special attention to parameters which are poorly constrained. Also an inclusion/reference of a model tutorial would be helpful for new users to begin interacting with the WBM.

The reviewer, along with others, identified a greater need to focus on WBM model performance in our manuscript, which we plan to accommodate in any revision of the manuscript. We have compiled a table of prior GHM estimates for total exorheic and endorheic river discharge. Though this table is still in draft form, we note that WBM's estimate of global discharge is in line with other's estimate over the same time period: about 40,000 km^3 yr^{-1} exorheic discharge and 2,000 km^3 yr^{-1} endorheic discharge. These estimates are consistent with those of Sutanudjaja et al. (2018) over the same time period (2000 to 2010), and a bit higher than a variety of studies that modeled epochs between 1960 and 2000, which generally coalesced around 36,000 to 39,000 km^3 yr^{-1}. We will also build out our comparison of model fit statistics across space presented in Figure 6 to include discussion about common causes for model regional misfit by GHMs.

The reviewer also suggested the development of a full tutorial. While we realize that this may indeed be helpful for some users, we would like to see if the instruction manual we prepared and released through our GitHub repository is found to be lacking by the community before we endeavour to invest in the development of a full tutorial, which we understand to be quite time intensive.

Line 66: Very nice explanation of the value of this source component tracking feature in this paragraph.

We thank the reviewer for this kind assessment.

Line 85: Global models come an out-of-the-box setup of preferred sub-model structures and parameterizations. There is a table devoted to the key default parameters (Table 2), but no discussion of the optimization process that is required for regional simulations, or representation of the range of parameters to make these regional simulations perform well. The authors do present some discussion and results based upon the contribution of uncertainty due to the climate forcing (Figure 3), but missed an opportunity to discuss the contribution of parameter uncertainty in the manuscript.

Several reviewers expressed an interest in seeing greater detail presented regarding calibration and parameterization strategies, and we would like to provide this general response. Any revision to our manuscript will provide additional detail needed to form a baseline understanding of the parameterization strategy for the model; however, we acknowledge at the outset of this response that there is to date untapped potential for more rigorous evaluation of uncertainty quantification using WBM.

First, following comments from Reviewers 2 and 5, we think that building out Table 2 to include a greater cross-section of parameters commonly adjusted in

regional studies is appropriate. We note that this table will be redundant to the WBM_Usage_and_Input_Reference.xlsx spreadsheet on the WBM GitHub page; however, we agree that providing a subset of this reference within the manuscript will improve the readability. Any revision to this table will include default values, reasonable ranges, parameter description, and important citations where applicable. We note that to the extent possible, we have relied on structuring the model consistent with empirically meaningful parameters. As such, values presented in Table 2 will often reflect syntheses of field observations and uncertainty as characterized therein. Other model parameters are more synthetic and have less direct connection to field observation. Many of these parameters have been evaluated through calibration exercises over the years in studies summarized in Section (3.1). The reasonable ranges to be included in any revision to Table 2 will be based on what the authors consider as appropriate starting points for parametric uncertainty analyses based on a combination of prior experience and physical meaning.

Previous work to calibrate WBM has generally leveraged manual calibration, with several instances of more rigorous calibration attempts. Parameterization of core WBM components evolved through iterative attempts to capture response in both global and regional contexts. Generally, it has been found that parameterization schemes as represented by the default parameter values in WBM_Usage_and_Input_Reference.xlsx reflect reasonable compromises that adequately represent discharge time-series in global simulations, as well as regional contexts focused on temperate, humid, and modestly developed watersheds. We plan to highlight that uniform parameterization can be applied to unique watersheds to capture non-default response (Samal et al. 2017, Zuidema et al. 2018), or that spatially varying parameterizations can capture more finely resolved nuance in watershed properties (Zuidema et al. 2020).

How modular is WBM in terms of testing particular hypotheses about hydrology and competing methods, etc. ? There is some brief description at the very end of turning flags on/off, but no specifics in how this influences representation of hydrology. Given the source tracking capability it would be interesting to test the impact of certain model assumption/hypotheses.

We agree with the reviewer that combined with WBM's source tracking capability, varying the processes represented by the model provide for powerful tools to address hypotheses in regional and global hydrology. Generally, this strategy underlies much of the development that has gone into recent work that leverages WBM, including Grogan et al. (2017), Zuidema et al. (2020), Rougé et al. (2021). We have added new optional functionality throughout the history of development in part because it captured important processes that improved representations. While we acknowledge that such experimentation makes for compelling analyses that show the importance of capturing anthropogenic processes in GHMs (Veldkamp et al. 2018), we consider repeating such experimentation here to be beyond the scope of this manuscript.

Model Description:

Line 140: I am assuming the representation of snow is considered single-layer, and does not include multiple layers. Things like snow properties and albedo are not explicitly taken into account. Also insolation and aspect are not considered within the melting term? Care to comment how this might influence your snow source? Has this ever been validated against gridded snow data sets?

The snow model is a simple single-layer approach with elevation bands providing

within grid cell variability (description starts on line 115 in the manuscript). Grogan et al. (2020) corroborated daily model predictions of snow water equivalent to 1034 observation points in the US Northeast. We will update the text to explicitly mention the single layer representation and estimates of the quality of fit to snow water equivalent.

Line 200: "Actual evapotranspiration (AET) from naturally vegetated land areas is a function of the PET, soil moisture, and soil properties."

So rooting depth is not taken into account? What function or purpose is setting the soil moisture pool depth to that of the rooting depth? I assume it's a single layer soil subsurface then?

In response to this question and direct requests from other reviewers, we plan to include details of the soil water balance calculation in any revision of the manuscript. AWC is a difference between soil field capacity (porosity that can hold water) and wilting point (minimal porosity that plants cannot extract water from). Because it is a relative metric (m/m), it needs to be multiplied by the rooting depth to have it in mm.

Line 210: "While AET from other land cover types (e.g., forest or grassland) can be parameterized and simulated, no published study has yet used this option of WBM. Actual evapotranspiration from other consumptive water uses are described below in Section 2.2.5."

Are the default parameters provided for forest/grassland types to calculate AET, or does the user have to provide them? If this option has never been used, how does the model treat forest/grassland if run globally – which includes forest and grassland cover?

WBM parameters that control all hydrological cycle processes have default values that a user can set and overwrite. This public version of WBM has over 250 of them (see WBM_Usage_and_Input_Reference.xlsx in the GitHub repository). Regarding this particular comment, default parameters that control AET were inherited from Federer et al. (2003). Specifically it includes a flag from the Hamon PET model, and optimized soil drying parameters for the best estimate of Global river runoff and discharge. The default AET settings do not use specific types of land cover and treat land as a generic land cover type, but the user can optionally set AET parameters for each land cover type. Using default settings in WBM is equivalent to treating the land-surface as covered by a reference vegetation (Allen et al. 1996).

Line 275: The term "Shallow groundwater pool" is not used in Figure 1. I assume this is the same thing as "groundwater recharge pool"? If so, make sure the terminology is consistent.

Yes, and thank you for pointing out the need for this correction.

Do the grid cells communicate for surface runoff and subsurface discharge or does this get routed directly to river transport model?

WBM assumes that surface runoff and "subsurface discharge", which we call baseflow, do not interact on the landscape, but both contribute to river flow as independent flow-paths. We will make this distinction clearer in the description of the model.

The term "unlimited unsustainable groundwater source" seems confusing. Is there a better way to describe this fossil ground water?

From the model perspective the water users are tapping into a pool of water that is not explicitly recharged and therefore it is unsustainable, even though this pool is not finite. The concept and the term "unsustainable" has been used in our previous published work (Grogan et al., 2017; Liu et al., 2017).

Several terms for this concept are used in the literature depending on the application, and we make particular reference to the useful table of definitions in Bierkens and Wada (2019). These different terms have been acknowledged in the text (starting on line 289). We selected our terminology as an adaptation to the "physically non-sustainable groundwater use or groundwater depletion" defined by Bierkens and Wada (2019), and our conceptualization is consistent with that use. Because we define a flux of water that satisfies demand unmet by local resources (i.e. non-sustainable groundwater use), this implies a pool of water from which this water is drawn, which we do not represent explicitly, that is best identified in our use as a tracking descriptor as "non-sustainable groundwater". We adopt "unsustainable" as equivalent to "non-sustainable" because the word "unsustainable" is defined in both the Merriam-Webster and Cambridge dictionaries and "non-sustainable" is not. While we do agree with the reviewer that it is the flux or the use of groundwater that is unsustainable, it is illogical to describe subsequent fates or flow paths abstracted unsustainably as "use". For instance, in describing the fractions of primary water components that drain to the ocean, it is awkward to describe a fraction of this discharge as "unsustainable groundwater use", when we are clearly treating it as representing a fraction of flow within the river system.

We are also unsatisfied with the alternative terminology posed (fossil groundwater) by the reviewer. Again, following definitions from Bierkens and Wada (2019), some terms imply knowledge of the age of recharge to the aquifer that we do not characterize in typical WBM simulations (e.g. recharging 12,000 years before present (Jasechko et al. 2017) in the case of fossil groundwater). Therefore, we prefer terminology that implies no assumptions of the era of recharge.

We acknowledge that the terminology that we have been using in this and our prior papers is not ideal, though it has been deliberated extensively. However, we view it as satisfactory for the time-being and for the purposes we are using it. We plan to clarify our definition of unsustainable groundwater in any revision of the manuscript.

Line 550-555: Here you mention spin up time in the context of water source tracking, but I feel this discussion could come much earlier when describing the model dynamics and features themselves.

We will be adding some further detail to the General Overview (Section 2.1), which will provide a brief introduction to typical practices in model spin-up.

Table 4: Please define 'relict' here.

We agree with the reviewer that the definition of relict water found on Line 856-857 comes too late in the manuscript and should be moved to the area around Table 4.

Because the model is becoming open source, the presumption is to allow for a wider user community. Do you provide a tutorial to familiarize the user with WBM? The authors provide a four-step description towards the end of the manuscript, plus a reference to an instruction manual, but a tutorial would be a great advance.

As stated earlier in our response, we will assess if the instruction manual released through our GitHub repository is found to be lacking by the community.

Model Validation:

Line 620: I recognize this section is devoted to a summary of WBM literature, but it is difficult to evaluate the skill of the model without the context of comparing against other similar hydrology models. There must be some model hydrology intercomparison studies to show here for global river discharge. Certainly Figure 3 might benefit for a comparison against other models.

The reviewer, along with others, identified a greater need to focus on WBM model performance in our manuscript, which we plan to accommodate in any revision of the manuscript. We have compiled a table of prior GHM estimates for total exorheic and endorheic river discharge that shows how this parameterization of WBM results are consistent with other estimates of global discharge during the same model epoch of 2000 to 2010. We will also build out our comparison of model fit statistics across space presented in Figure 6 to include discussion about common causes for model regional misfit by GHMs.

Line 624: What sort of parameter calibration is performed here? Hand tuning, or more formal DA approaches? Are these parameters available for the user?

Throughout the development history of WBM, default values for most parameters have been established through a combination of manual calibration, pre-calibration approaches such as Generalized Likelihood Uncertainty Estimation, and from literature review of applicable properties. The parameter set used here reflects those default values, and no separate calibration was performed for this manuscript. The revised manuscript will build out Table 2 to include a more complete listing of commonly adjusted parameters during regional calibration exercises, and their physically plausible ranges. For a complete listing of parameters, their units, complete descriptions of their behavior and interaction with the model, the reviewer is referred to the WBM_Usage_and_Input_Reference.xlsx spreadsheet on the WBM GitHub page.

Section 3.2.1: Table 2 is good, but a physical definition for the parameters should be stated in the table and not just in text. Perhaps a summary table of parameter values for the literature review manuscripts performed at different regions/resolutions, in addition to the default values.

In any requested revision of the manuscript, we plan to expand the table to include definitions for each parameter, reasonable ranges of parameter values based on literature review or prior experience, units, and parameter definitions. This information will be provided as a subset of the most important and common parameters adjusted, a complete listing of which can be found in the WBM_Usage_and_Input_Reference.xlsx spreadsheet available in the WBM GitHub repository.

Line 690: "Above, we reviewed previously-published WBM validations. As none of the prior versions of WBM code have been released open source, it is important to validate the exact model structure in this first open-source release. "

Was there significant mechanistic changes to these pre-release versions? Was adding the tracking capability the only significant difference from this official release version? From Table 7 it seems like you add some new functionality from previous versions: "Added rainfed agriculture, other land cover types, inter-basin transfers, domestic and livestock water demand", but you don't mention that here, and it seems to the reader that the only change is making it open source, when there are some structural changes.

We propose the following suggestions to clarify any confusion over the collection of processes represented in this release of WBM, and prior studies that we reviewed in Section 3.1. First, to address your question directly, yes, there were 'significant mechanistic changes' compared to many of the prior studies reviewed in Section 3.1. Many of these studies targeted specific research questions requiring new, unique, and narrowly applicable process representation. The open-source release presented here, represents a core trunk of process representation that we consider is widely applicable by itself, and is a useful starting point or baseline from which deviations in the code needed to accommodate future unique process representation can be described. Therefore, the most significant mechanistic changes are the removal of processes that we consider one-off or too narrowly corroborated by empirical data to distribute as having validity over global scope.

We propose that any revision to the manuscript should more clearly spell out how the core functionality included in WBM (v1.0.0) reflects this core of functionality common to preceding studies (including tracking), despite unique process representation in some of the referenced previous work. Furthermore, we also propose that WBM (v1.0.0) be added as a separate row to Table 7, with an entry that reflects the relationship between the released version and prior work.

Table 6: Appreciated how the WBM irrigation withdrawal estimate was put in the context of other studies. Would it be possible to construct something similar for global discharge? Especially since the author attributes the high irrigation withdrawal estimate (China) to relatively high discharge rates in the Asia domain, it seems like it would be worthwhile to hone in on discharge biases, and diagnose where and what location these are occurring.

Yes, we will be adding an additional table and brief discussion of global discharge estimates from this and other global modeling studies in any requested revision to the manuscript. Regarding your specific comment about a higher irrigation withdrawal prediction than other prior work in China, we would like to point the reviewer to Figures 5c and 6c, which illustrate low biases in discharge in China and southeast Asia.

Line 800: "Global discharge is dominated by rain over most of the globe, with snowmelt an important contributor at the poles, and both glacier runoff and unsustainable groundwater important regionally."

"Snowmelt is an important contributor at the poles": That seems like an oversimplification. Seems that Figure 8 shows there is significant snowmelt contributions well down into the northern mid-latitudes especially for mountainous terrain such as the Rocky Mountains and the Himalayas where a larger population rely on snow runoff for irrigation etc. This should be mentioned.

Thank you for this comment. We propose to modify the text in this section to read as follows:

Global discharge is dominated by rain over most of the globe, with snowmelt an

important contributor at high latitudes and high altitudes, with both glacier and unsustainable groundwater important regionally.

Figure 8: Glacier run-off seems to be unrealistic in the SouthWest and MidWest US where none should be occurring. Glaciers are apparently determined solely by snow input and melting algorithm and not prescribed like land cover types?

The Glacier Runoff panel in Figure 8 shows the fraction of glacier water found in river discharge. There are glaciers found in the Colorado and Mississippi Rivers and the meltwater travels downstream. That water is used for irrigation and there will be non-zero fractions of glacier water found throughout these regions.

Figure 11: Perhaps outline the watershed domain of the Indus river in this figure within the larger figure 10 for better reference and perspective.

Thank you for this suggestion, we will include the requested outline in any revision of the manuscript.

Figure 11b: There are no units in this figure. Why is there such a large discontinuity between months 12 and month 1? Why does glacier runoff have such a strong contribution upstream at point A, yet such a small contribution downstream at point B?

We will include units in the figure in any revision of the manuscript.

The plot reflects the large differences in flow between the months of December and January. We note that there are similarly large differences in flow between the months of November and December. Most of the water at the river mouth during December reflects negligible water withdrawals upstream for irrigation as there is no or little irrigation in the month of December.

Point A is located in the foothills of the Himalayas and a large portion of the basin upstream is covered by glacier. Downstream of Point A is the Indo-Gangetic Plain, which contains extensive irrigated agriculture, and large amounts of the glacier water are lost to evapotranspiration. The basin outlet (Point B) includes drainage from several eastern tributaries originating in India. These tributaries are supported predominately by rain water, and from extraction of unsustainable groundwater.

4.2 Return flow tracking

Line 867: I feel like this “relict” and “pristine” distinction should be made earlier, because it is used in an earlier table.

We agree with the reviewer that the definition of relict water found on Line 856-857 comes too late in the manuscript and should be moved to the area around Table 4.

Return water diagnostics and source water diagnostics and tracking seem to be some of the most useful components of the system. Would it be easy to port these ‘diagnostics’ modules to other hydrology models or is it baked into the system?

The code is deeply integrated into the model, however, the concepts and techniques could be adopted by other modeling groups and are very similar to the techniques used in HBV (Stahl et al. 2017, Weiler et al. 2018).

Figure 12 and 13: Care to comment on why the irrigation return flow waters are so high in the Northern India region? It might help the reader contextualize this diagnostic, and make for interesting discussion of what makes that region so unique.

We will make the following addition to the caption:

The large fraction of irrigation return flow in northern India and Pakistan results from this region being one of the most intensively irrigated regions in the world. Water is used and reused multiple times with low irrigation efficiency (Zaveri et al., 2016).

Figure 14: These plots are interesting and potentially very valuable. Couple questions, in Figure 14b could the color scale be changed to emphasize the 0 to 0.2 discharge fraction? A large part of the watershed domain seems to be the same color of green that is hard to distinguish at all.

Could you not color Wyoming such that the river network can be seen easily? It doesn't seem to be a need to color Wyoming dark blue since it is the headwater region. Also in Figure 14a is this distribution pattern mostly driven by snowmelt? It doesn't look like the distribution is picking up the summer monsoon season where a larger fraction of the water should derive from Arizona and New Mexico. I suppose this depends if the climate forcing captures the SouthWest monsoon.

We did not color Wyoming separately from the rest of the grid cells. All small rivers in Wyoming have 100% discharge originating in the state and therefore they are colored in the darkest blue. In a few cases where rivers enter into Wyoming, such as the North Platte River, they reduce the fraction of Wyoming-sourced water and lighten the color. As more Wyoming-sourced water is added to those rivers, the colors will get progressively darker. We will try to make a clearer figure in any revised manuscript.

Regarding Figure 14a, yes, the large peak in Fig. 14a is a result of snowmelt.

Model Code development section: Although I do find this section interesting it may be better suited for the appendix, such that the reader can immediately go from the results to the discussion. You have provided steps, and there is access to an instruction manual, but do you also provide a simple user tutorial for a cut down domain and provide the input files so the user can familiarize themselves with the steps?

Another reviewer expressed appreciation that this section was included. Therefore, we will consider where this section is best located, in the main text or in the supplement, and we encourage the Editor to provide an opinion. The tutorial request was discussed above.

Discussion:

The tracking capabilities of this model which can attribute discharge source regions, or impact from agriculture on discharge are a great feature and worthy of discussion. The authors refer to the modular nature of the model in order to toggle on/off different features, but they don't provide a test case of this, where for a single experiment, different model structures/assumptions/parameters are switched up to identify the model sensitivity. Just a suggestion.

We do provide a variety of simulation runs showing the breadth of the model's capabilities, and more detailed analyses can be found in the manuscripts cited that make specific use of this functionality (Grogan et al. 2017, Zuidema et al. 2020). We feel that any additional sensitivity analyses using tracking

functionality would detract from the focus of documenting the WBM v1.0.0 model, and would be best suited to separate future work allowing us to keep the manuscript to a manageable length.

I was certainly expecting at least more discussion concerning model performance related to river discharge and irrigation water withdrawal as covered in the first part of the manuscript. It was a bit concerning that the WBM was a bit of an outlier when estimating global irrigation water withdrawal, and this was partially attributed by the authors to low discharge rates across China and Asia. Very little to no discussion or explanation was provided for this. The discharge rates seemed to perform relatively well in other regions includes North America, but it was difficult to contextualize given no comparison in performance was provided from other hydrology models.

In any requested revision of the manuscript, we will be drawing out our discussion of model performance as suggested by the reviewers. We expect to provide a greater number of efficiency measures than the index of agreement, also suggested by several reviewers. We will be adding an additional table, structured similarly to Table 6, that compares WBM's estimate of global discharge to other macro-scale models, which shows WBM's estimate of global discharge to be consistent with other models.

To highlight some of the key points in the discussion of any text revision, we will note that input data dictates the model's ability to capture observed discharge response (Figure 3). The dataset that we use in this analysis (MERRA2, Gelaro et al. 2017) is a reanalysis product that draws on large observational datasets available for the global north, and better simulation performance in these regions likely reflects data available for reanalysis.

References:

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

Bierkens, M. F., and Wada, Y. (2019). Non-renewable groundwater use and groundwater depletion: a review. Environmental Research Letters, 14(6), 063002.

Federer, C. A., Vörösmarty, C., & Fekete, B. (2003). Sensitivity of annual evaporation to soil and root properties in two models of contrasting complexity. Journal of Hydrometeorology, 4(6), 1276-1290.

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., Zhao, B., 2017. 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

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, Environ. Res. Lett., 12, 034017, https://doi.org/10.1088/1748-9326/aa5fb2, 2017.

Grogan, D. S., Burakowski, E. A., and Contosta, A. R.: Snowmelt control on spring

hydrology declines as the vernal window lengthens, Environ. Res. Lett., 15, 114040, https://doi.org/10.1088/1748-9326/abbd00, 2020.

Jasechko, S., Perrone, D., Befus, K. et al. Global aquifers dominated by fossil groundwaters but wells vulnerable to modern contamination. Nature Geosci 10, 425–429 (2017). https://doi.org/10.1038/ngeo2943

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, Environ. Res. Lett., 12, 104009, https://doi.org/10.1088/1748-9326/aa88db, 2017.

Rougé, C., Reed, P., Grogan, D., Zuidema, S., Prusevich, A., Glidden, S., Lamontagne, J., and Lammers, R.: Coordination and Control: Limits in Standard Representations of Multi-Reservoir Operations in Hydrological Modeling, https://doi.org/10.5194/hess-2019-589, In review.

Samal, N. R., Wollheim, W., Zuidema, S., Stewart, R., Zhou, Z., Mineau, M., Borsuk, M., Gardner, K., Glidden, S., Huang, T., Lutz, D., Mavrommati, G., Thorn, A., Wake, C., and Huber, M.: A coupled terrestrial and aquatic biogeophysical model of the Upper Merrimack River watershed, New Hampshire, to inform ecosystem services evaluation and management under climate and land-cover change, 22, 18, https://doi.org/10.5751/ES-09662-220418, 2017.

Stahl, K., Weiler, M., Freudiger, D., Kohn, I., Seibert, J., Vis, M., Gerlinger, K., and Böhm, M.: Final report to the International Commission for the Hydrology of the Rhine basin (CHR), 146, 2017.

Sutanudjaja, E. H., van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H. C., Drost, N., van der Ent, R. J., de Graaf, I. E. M., Hoch, J. M., de Jong, K., Karssenberg, D., López López, P., Peßenteiner, S., Schmitz, O., Straatsma, M. W., Vannametee, E., Wisser, D., and Bierkens, M. F. P.: PCR-GLOBWB 2: a 5-arcmin global hydrological and water resources model, 11, 2429–2453, https://doi.org/10.5194/gmd-11-2429-2018, 2018.

Veldkamp, T. I. E., Zhao, F., Ward, P. J., de Moel, H., Aerts, J. C. J. H., Müller Schmied, H., 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, Environ. Res. Lett., 13, 055008, https://doi.org/10.1088/1748- 9326/aab96f, 2018.

Weiler, M., Seibert, J., and Stahl, K.: Magic components—why quantifying rain, snowmelt, and icemelt in river discharge is not easy, 32, 160–166, https://doi.org/10.1002/hyp.11361, 2018.

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, Environ. Res. Lett., 11, 084005, https://doi.org/10.1088/1748-9326/11/8/084005, 2016.

Zuidema, S., Wollheim, W., Mineau, M. M., Green, M. B., and Stewart, R. J.: Controls of Chloride Loading and Impairment at the River Network Scale in New England, 47, 839–847, https://doi.org/10.2134/jeq2017.11.0418, 2018.

Zuidema, S., Grogan, D., Prusevich, A., Lammers, R., Gilmore, S., and Williams,

P.: Interplay of changing irrigation technologies and water reuse: example from the upper Snake River basin, Idaho, USA, 24, 5231–5249, https://doi.org/10.5194/hess-24-5231-2020, 2020.