

Report #1 (Anonymous Referee #2):

I am pleased with the authors' response to my initial set of questions and suggestions. Just a couple exceptions listed below:

Near Line 140 the authors refer to several studies which 'evaluate' the model performance against observations of SWE (snow), but make no effort to quantify the performance skill.

Response:

We have added a summary of the metrics reported in the previous publications' SWE validations to Section 3.1 Published WBM validation and evaluation *Additional validation and evaluation metrics.*

The term 'unsustainable groundwater' makes sense, as thoroughly explained by the authors. My main issue was that the phrase unlimited, unsustainable groundwater seems like an internal contradiction.

Response:

While in reality, unsustainable groundwater must be limited by definition, we use the word "unlimited" here to clarify that the model does not have any cap on the volume of water that can be extracted from this groundwater category. This caveat is very important for understanding the limitation of the model for evaluation sustainability of groundwater resources. We've clarified this in the text in Section 2.2.2 Groundwater.

Report #2 (Anonymous Referee #6)

This paper introduces and describes the most recent version (1.0.0) of the University of New Hampshire Water Balance Model with source tracking functionality. The authors evaluates the current model performance against GRDC river discharge observations and FAO's irrigation withdrawals for the period 2000-2009. Considering the above-mentioned variables, the model performs well in regions such as North America and poorly in Asia. The source tracking functionality as demonstrated with examples in the paper distinguishes this model from other GHM's. Overall, this paper provide a quite thorough description of the model. The source tracking functionality is indeed interesting. However, a few questions should be addressed prior to publication in GMD.

Abstract

Line 11: "WBM was first published in 1989; here we describe the first fully open source WBM version"

Could the authors please add the model version number here for clarity?

Response: we have added the version number

Line 16: "Users can determine what proportion of any flux consists of each of the primary inputs of water to the surface of the terrestrial hydrologic cycle, previously extracted water for human uses, or runoff generated from any place on the Earth's surface."

I believe you want to highlight the fact that a user can know proportions of the primary source in a given flux but is it difficult to grasp such meaning with this kind of sentence formulation and length. Kindly rephrase for clarity.

Response: We have revised this sentence for clarity.

Introduction

Here I would also like to see a closing paragraph on the progression of the WBM models. What was lacking in those models and the reason for the new model? Why should user's care about this new model?

Response: The progression of WBM models is covered in detail in Section 5, including statements on what is new about each model. Given the length of this paper, we prefer not to add redundant text on this topic to the introduction.

In addition, I suggest explicitly stating your research question, which will improve comprehension.

Eg. The aim of this paper is to provide an overview of the newest model version WBM 1.0.0 by

1. Describing the new model comprehensively
2. Showing and discussing standard model output for selected domain X or the entire globe
3. Providing insights into model evaluation, and giving guidance for the users of model output (conditional if the focus is not on providing standard outputs for users)

Response: We have clarified the scope of the manuscript at the end of section 1. However, as the manuscript is a model description paper, we have not framed the manuscript around a research question and suggest that we can address the referees comment without framing the introduction around a question.

Model description

The authors really provided a thorough description of the model.

Line 104: “WBM is modular and is able to accept climate, land use/land cover, water management, and water demand inputs from other models and data sources.....”

With Modularity, you mean WBM code (or software) is written in small modules right? If so, you could state the number of modules that makes up WBM or show the software architecture, which will highlight WBM’s modularity.

Response: We added a list of the modules that can be turned off or on to this paragraph.

Line 134. “Table 1 presents a cross-section of parameters that are typically...”

Add references to Table 1 if any

Response: All parameters in Table 1 are described in the relevant text, which contains all required references for values and ranges.

Model Evaluation

Line 935: ‘Model code’: This is well suited for the appendix as it breaks the continuity of the results and dissection and decrease comprehension. You could also briefly summarize the important aspect of the model code (language, code structure, etc.). In addition, refer the readers to the appendix for full and detailed description of the model code in addition to how to use the code

Response: Other reviewers found this section useful, and we think it is helpful to readers for understanding the model as per the scope and instructions provided by GMD for model description papers. If the editor would like us to move this section to the appendix, we will do so. Otherwise, we suggest leaving it where it is.

Discussion

This study obviously has limitations. I expect to see more of them discussed here

Response: The evaluation metrics are intended to quantify the model’s limitations. We use four different evaluation metrics, and describe the results of prior studies that report additional metrics for different model fluxes and domains. Since this is a model description paper, and not a research paper, it is somewhat challenging to discuss limitations beyond quantifying the model performance under a default parameter set. As discussed in the paper, regional studies must parameterize and evaluate the model for the given study region, and must then identify the model limitations for that specific region

and research question. We have added a sentence emphasizing this regional limit to the discussion section. Further, we do describe the limitations of evaluating the tracking functionality in section 4, and have moved a sentence from the conclusion to the discussion to remind the reader that corroboration of tracking functionality is not yet possible.

Report #3 (Anonymous Referee #4):

The authors have adequately addressed all comments and I think the manuscript is in a really good shape.

The only minor detail is that (l. 28) there seems to be no reference for Gochis et a., 2020 (which should also be et "al.") and if you want to distinguish between LSMs and ESMs here, shouldn't it be CESM rather than CLM?

Response: We have added the Gochis et al. (2020) reference to the list, and fixed the typo. We categorize CLM as the land surface model, and add CESM as the earth system model.