Comment on gmd-2022-59

Anonymous Referee #5

Referee 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-RC5, 2022

Overall comments:

The authors present and describe a new open-source version of the global hydrologic model WBM, emphasizing new capabilities for tracking water sources. The paper is well written, and the overview is fairly comprehensive, including theory, examples, plentiful references to earlier literature, and a discussion of how this open-source version relates to other versions of WBM that have been used over the 3 decades since the first version was created. Not only is the model now open source, but the authors have provided a Singularity container to simplify access/usage. Overall, this is a nice contribution, and I recommend publication after minor revisions.

Below are listed specific comments keyed to particular line numbers, sections, or equations:

Introduction: I appreciate the overview of applications of GHMs, which seems like a useful entry point for those new to the topic.

Sec 1.1: the need for water tracking is well motivated here.

69-70 typo?

86-88 at some point around this section it would be useful to describe how WBM handles gridding. Here an example is given of a fixed-width (120 m) grid, and later examples are noted of lat-lon based grids. Does the user have a choice between these? Does the model

account for the variable size of fixed-longitude grid boxes? What happens at the poles? Or does the global configuration exclude very high latitude regions like Antarctica?

124 and ff: thank you for listing units of each variable.

Eq 3 & 4, should this be P^e? (also, in general, instead of introducing equations with 'defined according to', it can be helpful to say something more descriptive like 'so-and-so depends on temperature T and precipitation rate P according to')

Eq 4 and other math: if you want text-like typesetting, e.g., the word "if" in eq 4 or a sub/super-script like "max", use \text{if}, W_i^\text{max}, etc. (requires \usepackage{amsmath})

142 grammar around lapse rates

158 is Pt a user-defined param? A fixed fraction of P? Calculated in some other way?

Eq 7 and others: consider the more traditional use of a dot (\cdot) or no symbol at all to represent multiplication, as opposed to an asterisk (which I think traditionally means convolution, even though most modern programming languages use it as a multiplication operator)

212 no cap

eq 13 ff: this is somewhat confusing because the phrase 'immediately moved' suggests a discontinuity but the differential equation suggests differentiability / continuity. Please clarify (maybe via a delta operator in front of R_EXC, which is 1 if volume of retention pool exceeds the threshold and 0 otherwise?)

250ff isn't there a unit mis-match between W and R in eqs 14-18? And between R and T in 18?

286 unclear what 'stock' means here (it seems to be a standard term with WBM, so please define it before using)

Sec 2.2.3: I appreciate the references to papers that describe the routing methods, but it would still be helpful to have a bit more information on linear reservoir routing. For example, does it mean that each grid cell's river discharge output to its neighbor is calculated as a linear reservoir, that is, as a function of river water within the cell? What, briefly, is the basis for assigning a reservoir coefficient? Are these constant or do they depend, for instance, on channel geometry?

418 does 'scaler' mean 'scaling factor' or 'scalar' or something else?

429-437 Can you elaborate a bit on this treatment and why it is needed?

473 capitalization consistency

Sec 2.2.7 source tracking: this section is a bit confusing, partly (I think) because of the wide variety of sources that could be tracked. If I understand right, a user would not normally track ALL of these sources in any given model run, but rather would pick a type of source to track - is that right? Actually, reading forward in section 4, we learn that there are 3 options. It would be helpful to list these options up here in section 2. In addition, a couple of examples would potentially help a lot. They need not be very elaborate, but could be as simple as something like: 'a user interested in X might choose to track sources Y and Z'.

545 again, are these 3 mutually exclusive (i.e., one chooses from among them), or are all 3 tracked simultaneously? (later text suggests the former, but at this point in the text it is not clear)

557 daily time step: helpful to mention this much earlier

Sec 3: helpful to define what you mean by validation (I'm not personally a stickler for semantics, but some would consider the term problematic, and better described by confirmation, testing, or evaluation).

Sec 3.1: are the summarized validation studies performed in conjunction with some kind of calibration / parameter optimization? Or is calibration only used in regional applications? You kind of answer this question around 635-640 but it would be helpful to clarify near the start of this section.

Validation generally: it would be interesting to summarize some of the lessons from

testing and validation, it terms of what might be behind systematic under- or over-prediction of discharge. For example, have past validation exercises revealed certain gaps in knowledge, and/or mathematical approximations that would need to be refined in order to improve model performance? This might fit well under Results or Discussion.

Sec 3.2: thanks for differentiating between validation of different versions, and including this section devoted to the open-source version. It's a nice reminder (and demonstration) that testing of models should ideally include the specific code implementation alongside the theory and numerical algorithms.

Eq 31: the first time I read this, my mind immediately went to cancellation of errors -- but then realized that this is actually desirable for a bias metric. You might consider reversing the order of 32 and 31, and introducing the MBE with a phrase like 'in order to measure systematic bias' or something to that effect, so readers don't get hung up on it.

711 observations per year, or total?

753 tense

Sec 5: I appreciate the code history and summary of different versions

943 uniformly spaced... in geographic coords? (again, helpful to explain grid set up early in the paper)

1039 typo