

Groundwater Hysteresis Increasingly Decouples Flowing Network Length from Streamflow as Snow Shifts to Rain

Elijah N. Boardman, Mark S. Wigmosta, Nicole M. Fernandez, John A. Whiting, Adrian A. Harpold

Response to Reviewer Comment RC1

We greatly appreciate the reviewers' helpful comments on this manuscript, which we will address in a revision. Our comments are interspersed into the review in blue.

This study from Boardman et al. describes reconfiguration and application of a spatially distributed, physically based hydrologic model to simulate flowing network length and streamflow in a headwater catchment. The model shows expansion and contraction of the flowing stream network relative to discharge, and the authors compare the model outputs to a common field-based description of L-Q using a power-law relationship. An important contribution of the work is that the model reveals nonmonotonicity in the simulated L-Q relationship, which the authors state is due to the hysteretic nature of the flowing stream length and flow in the watershed driven by groundwater.

General comments:

- Both the science and the presentation of the manuscript are very high quality. I have several relatively minor comments that in my opinion the authors should address; however I suggest that the paper be published in HESS upon completion of these minor comments.
- In my opinion, this paper represents an important step forward for the headwater modeling community.

We thank the reviewer for their interest in our study, and we look forward to addressing the minor issues raised below.

Specific comments:

- L28: This definition of headwater streams doesn't align with what current literature suggests (Cf. Golden et al., 2025; <https://doi.org/10.1038/s44221-024-00351-1>). I recommend revision or inclusion of a citation where other folks have adopted this definition that you all use.

We have revised our phrasing to “first-order streams,” which is consistent with the phrasing used by the study cited in our manuscript (Downing et al. 2012).

- L176: I really appreciate the written description of the model formulation. However, it would also be useful for future studies if equations were presented here that describe how the processes were parameterized. Please consider inclusion of these.

The DHSVM routing equations are already described in Wigmosta et al. (1994, 2002), and the changes implemented here primarily affect the geometry, not the fundamental routing equations. Nevertheless, we will include several key equations in our revision to more clearly illustrate how the existing routing equations interact with our updated geometry.

- L196: It seems odd to jump from S1 and S2 to S10.

We appreciate that the Supporting Information numbering may be out of order in the main manuscript—the intention was to number the figures within a logical chronological order (from setting up the model to results), but sometimes the results (S10) are also mentioned to illustrate the functioning of the methods (S1-2).

- L260: Can reaches be less than 30-m in the formulation? If so, where is that possible and how are those lengths determined.

Yes, the reaches can be smaller when necessary to account for the network topology, i.e., at confluences. This has been clarified in the manuscript.

- L286: Didn't you state that this version of DHSVM has a new instream routing algorithm? Wouldn't prior sensitivity analyses of DHSVM therefore not be applicable here?

This has been clarified that the sensitivity analyses are “prior” in the sense of this study, but still within the scope of the revised model.

- L315: Just for clarification: this is comparing Qsim and Qobs at the outlet of the catchment, right? No other catchments upstream have flow data (at least that are used to calibrate the model)?

Correct. There are some upstream measurement locations that have been operated idiosyncratically, but these locations do not have reliable stage-discharge relationships and are not considered here. However, it is also important to note that the gauge location is not at the “outlet” of the DHSVM model, but rather, the DHSVM model is extended several kilometers downstream to reduce uncertainty associated with the outlet subsurface boundary condition.

- L324: It'd be helpful to know what the benchmarks are that you're considering here.

This has been clarified to “sub-optimal streamflow NSE and log-scale NSE (relative to the best values achieved by the model, i.e., 0.87 for both).”

- L326: It might be helpful to let the reader know that you'll talk about power law anomaly in the next section. It took me awhile to understand what L_A and EC_A were referring to.

Thank you for pointing out this missing context—we have added a sentence explaining that L_A refers to a potential hysteresis signal that will be introduced in more detail in Sect. 3.2.

- L330: It sounds like L_A and EC_A weren't incorporated into model calibration, but rather were included after calibration to constrain the realizations. Was it not possible to incorporate these directly into the calibration algorithm as well?

It would be theoretically possible to calibrate these directly, but the data volume generated is difficult to work with. Each evaluation of L_A requires generating many years of daily stream network maps, which is about 2 GB of data per parameter set. Thus, testing this for hundreds of parameter sets is computationally expensive. Also, our prior research of DHSVM calibration suggests that it might be better not to directly calibrate metrics directly related to the research question (i.e., better to calibrate NSE instead of L_A) to avoid overfitting:

<https://hess.copernicus.org/articles/29/6333/2025/>

- L393: It feels like it might be worthwhile to include some information (maybe a short paragraph) on the climate scenarios and how you ran these through DHSVM.

We added a sentence at the end of Sect. 2.2 explaining that we also test a simple uniform-warming test in order to evaluate the potential sensitivity of the model to changes in rain-snow partitioning, but this should not be interpreted as a realistic climate “scenario.”

- Fig 1: When showing the stream network, it would be helpful to also identify where the dry reaches are. E.g., with grey lines or something similar.

Although we agree in principal that this could be interesting, we feel that adding additional lines to the stream network maps could be confusing and result in visual clutter. Moreover, illustrating the dry reaches would make it harder to see the critical network-disconnection phenomenon. Finally, our diagrams are consistent with the style of Godsey and Kirchner (2014) for the same catchment, aiding in easier visual comparison.

- L485: Why not use a Mann Kendall test here and include the p-value? That seems like it would provide more statistical merit to the analysis than the sign of the Sen's slope.

We appreciate that multiple nonparametric statistical tests could be conducted to evaluate trends. We will test the Mann-Kendall test as well to determine if it provides similar results.

- L491: This is a super interesting finding. One point that could be worthy to include in your discussion: It does seem like there's quite a bit of scatter in the (simulated) L-Q plots shown in the studies from both Gao and Mahoney, which also demonstrate nonmonotonicity in L-Q... (<https://doi.org/10.1016/j.jhydrol.2021.126522>; <https://doi.org/10.1016/j.jhydrol.2023.129422>). Are these models also incorporating a degree of hysteresis (albeit perhaps not quite as explicitly as is being done here)?

We have also noticed the scatter in the Gao and Mahoney papers. It is a bit unclear whether the models used in those studies could produce a hysteretic groundwater behavior, since as far as we understand, those models use static topographically based routing schemes. It could also be possible that variations in the spatial configuration of $P - ET$ could lead to more or less streamflow being generated by higher or lower parts of the stream network. Since the authors of those studies do not appear to comment directly on the scatter observed in their L-Q plots, we will refrain from making any definitive comparison. However, we have mentioned that similar scatter is observed in other modeling studies in the discussion of our Fig. 3.

- Fig 4. Same as fig. 1 – it'd be helpful to show the dry streams on these maps.

As before, we think that adding the dry streams would make the figures more visually confusing, especially since the maps need to be easily compared to each other. However, the full geomorphic network (and hence implicitly the dry streams) is shown in Fig. S3.

- L726: could just be worthwhile to note that EC isn't always conservative and therefore may not universally be a good tracer.

Good point, and we have added this caveat.