Hydrol. Earth Syst. Sci. Discuss., 6, C2600–C2611, 2009
www.hydrol-earth-syst-sci-discuss.net/6/C2600/2009/

Hydrology and
Earth System
Sciences
Discussions

Interactive comment on "Iterative approach to modeling subsurface stormflow based on nonlinear, hillslope-scale physics" *by* J. H. Spaaks et al.

J. H. Spaaks et al.

jspaaks@uva.nl

Received and published: 13 November 2009

We would like to thank Anonymous Referee #2 for taking the effort of writing a clear and concise Interactive Comment. We feel that the manuscript has benefited greatly from the constructive feedback we received. In the sections below, we cover specific issues in more detail. The Referee's comments are verbatim in bold face; our comments are in normal face.

1 Major points

1.) There is no discussion of how the results of this study compare with other hillslope modeling studies or other studies that looked at the complexity required to represent hillslopes. There is also no comparison of the results with those from other modeling studies for the Panola hillslope (e.g. Lehmann et al, Tromp-van Meerveld and Weiler, and Hopp et al). In fact there is no mention of these previous modeling studies at all. Therefore it is not clear, how the results of this study compare to or are different from the previous studies and how this study compliments the previous studies. There should be some discussion of how the results of these model simulations compare to those of the previous studies. This could be done in the discussion sections for each model simulation.

If we understand the comments correctly, the Referee brings up 2 issues:

1. Discuss how the results compare to other studies that looked at complexity in relation to the available data; This can be interpreted in relation to 2 aspects:
    (a) the number of calibrated parameters in relation to the available data;
    (b) what processes are needed to explain the data
2. Discuss how the modeling results compare to those of other studies in terms of the description/prediction/hindcasts for Panola specifically

With regard to 1(a), the purpose of the calibration was not to determine the best parameter set as one would do when making a prediction. Rather, it was necessary for comparing one model to the other. For this, it would have been pointless to take just any parameter set; for each model, we needed the parameter set that yielded discharge curves of the right order of magnitude. We did not intend the paper to be about parameter identifiability in relation to the available data, so we think there is no need to discuss this either.

With regard to 1(b), we intentionally limited the number of processes that were included. Our study was not aiming to identify structural inadequacy of the presented models in terms of the included processes. Rather we compared different implementations of subsurface stormflow, and how the model results related to the collective field knowledge of subsurface stormflow. Therefore, we believe that including a discussion about structural adequacy would be somewhat misplaced.

With regard to 2, we did not aim to reproduce the observations of the Panola hillslope *specifically*. We used this hillslope because the fill-and-spill theory originated there, and there was a good record of precipitation and discharge available. As we mentioned in the article, we do not expect that even our most complex, calibrated model structure gives accurate predictions of, say, distributed groundwater tables. Because of these considerations, we do not think a comparison with model results from the literature is useful.

2.) The description of the models is not clear. For example on P5209L25-26 and P5217L1-4 a lumped model with spatially discontinuous groundwater is described. Is it an actual lumped model? If so, how can there be spatially discontinuous groundwater? Or is it a spatially distributed model with the same parameters for each element? Similarly it is not clear how for the lumped model described on P5213L4 water is routed from one element to another element or how for the lumped model on P5220L6 and P5227L23 there are different spatial elements or patches of saturation. The different models should be described better and the term lumped model shouldn't be used if the model consists of spatial elements.

The Referee wonders how spatially discontinuous groundwater patterns can be represented by a lumped model. The answer is that we take an implicit approach to

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

modeling the effect of spatially discontinuous groundwater distributions within the element. In this approach, we assume that a body of transient groundwater exists within the soil domain of one spatial element, i.e. it is not present everywhere. Furthermore, its exact size and shape are unknown. We assume that emergence of water at the soil-bedrock interface occurs at the same rate everywhere within the element. Emerging water has to travel to the body of transient groundwater, before it can be discharged from the element. A water parcel that happens to emerge far away from the body of transient groundwater has to travel a long way through the unsaturated zone compared to a parcel of water that emerges close to the body of transient groundwater. The travel time for all parcels within an element is modeled according to a frequency distribution of travel times (Equations 12 and 13).

For the first three experiments, we apply our concepts in a lumped setup. For experiment 4, we represent the hillslope as a transect of 8 spatial elements, each of which is governed by the mechanism of either $Model_1$ or $Model_3$.

Regarding the routing of water in a lumped setup, we assumed that $\frac{\Delta H}{\Delta s}$ from Equation 3 can be approximated by the average slope of the bedrock.

In the revised version of the manuscript, we have added some text to clarify our concepts. Also, we deleted the reference to the flow direction algorithm on page 5213, line 14–15. While the models can be set up in lumped, transect, or 2-D spatially distributed mode, it is not necessary to include this reference here, since we only use lumped and transect setups.

3.) P5217: Is this not a lumped model where water reaches the saturated zone over time (using the travel time distribution) instead of instantaneously? It is not clear to me how this variable travel time to reach the saturated zone (which I assume accounts for variable time through the unsaturated zone) is the same as a model with spatially discontinuous transient saturation. Describe this model in more detail using clearer terminology.

The Referee seems to understand the text correctly as it is now, but has a hard time placing it into context due to the confusion about the lumped *vs.* a distributed setup (see the previous point). Now that we have described the model setup more clearly, this should no longer be a problem.

4.) P5222L11: Do you see this inflection point for all events? Is the inflection point caused by macropore flow shutting off? Field and modeling studies have shown that macropore flow is important for this hillslope. Or can you not represent it because you do not include a decrease in Ksat with depth in the model? Discuss other possibilities for this inflection as well.

We selected this particular event because it has the dynamics typically observed in subsurface stormflow dominated hillslopes. We evaluated models that were capable of the type of nonlinear behavior that has been observed for different subsurface stormflow dominated hillslopes around the world. For the specific case of Panola, the inflection point is not caused by macropore flow shutting off, at least not for the macropores that were directly observed. The hydrograph for macropore flow at Panola decreases quite smoothly, and keeps decreasing past the inflection point. The inflection point is rather the result of the double peak that exists in the hydrograph of matrix flow. The first part of the double peak may itself be the result of macropore flow some distance upslope from the trench, but we do not have data to corroborate this.

Of course, the model behavior may be improved by including additional processes in the model, or by using a different concept for the processes that were included. While we could have investigated any given hypothesis, we chose to investigate whether we could reproduce the inflection point (as well as the other evaluation criteria) by using $Model_3$ in a spatially distributed transect setup. It turned out that $Model_3$ was not capable of doing so.

5.) P5226L1: Many modeling studies, including those for Panola, have shown the

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

importance of initial conditions and the importance of modeling a period longer than one storm. How did you take care of these antecedent moisture conditions and how did you check that the moisture conditions are well represented? The values given in table 3 for field capacity and initial moisture content seem really low. Are these really reasonable (for winter)? Where did you get these values from?

As the Referee points out, initial soil moisture conditions and modeling a period longer than one storm are important. Generally speaking, calibration using multiple events essentially amounts to compensating for errors in the initial and boundary conditions (as well as allowing the parameters to compensate for errors in the model structure). Doing so improves the predictions and allows for better estimation of the total uncertainty. However, we were not aiming to make an accurate prediction for Panola specifically. As long as the initial conditions were within reasonable limits, there was no real need to assess their appropriateness. Also keep in mind that our comparisons are based on calibrated models only. Calibrating each model using different initial conditions would probably have changed the model performance a little bit, but not the shape of the curve. For example, we believe that Table 2 would not have changed.

Nevertheless, it has become apparent from the reviewer's comments that some clarification was needed in the description of our experimental design's rationale. In the revised version of the paper, we therefore included a section similar to our explanation above.

6.) Table 1: What are these parameter ranges based on? Are they the same as those used by Hopp et al. and Tromp-van Meerveld and Weiler?

We chose the parameter ranges as follows: first, we performed a manual calibration of each model. After we identified which part of the parameter space yielded good fits and behavior, we limited the parameter space to a few orders of magnitude (typically 2.5–4, but less where appropriate) around that initial best point. In the revised version of the manuscript, we added a section on calibration.

2 Minor points

***) P5207L8: Replace “impermeable” by “lower permeability” as recent research has shown that the bedrock is not necessarily impermeable. This has also been shown for the Panola watershed.**

Changed.

***) P5209L10 and elsewhere in the text: Don’t use the term virtual experiments here. These model simulations are not virtual experiments as per Weiler and McDonnell 2004 but rather model simulations to test model structures and model suitability.**

The term ‘virtual experiments’ was coined by Markus Weiler and Jeff McDonnell in their 2004 paper. The authors present an approach with which first-order controls on flow and transport at the hillslope scale can be identified. They compare results from 2 model setups that use a different drainable porosity. This is somewhat similar to a sensitivity analysis, in that the effect of changing the value of one parameter is isolated. It also goes beyond a simple sensitivity analysis in that it evaluates in what way specifically the model results differ. Their paper aims to better understand why the model is doing what it is doing and how that relates to observable patterns in the real world. This is exactly what we were aiming for with our virtual experiments, in particular for the transect case (virtual experiment 4). From a philosophical perspective, the only difference is that we are comparing model structures, rather than parameterizations.

***) P5213L1: Precipitation or throughfall?**

We assumed that throughfall is equal to precipitation. This seems a reasonable approximation given that leaf-out only occurs later in the season. Also, interception losses reported in the literature are relatively small (≈ 2 mm for storms during the dormant season, Cappellato and Peters, 1995; JoH 169, p.135) compared to the relatively large event (59 mm) and the large bedrock losses.

***) P5215L23: How many model parameterizations were used in total?**

In Tables 1 and 3, we have added the total number of SCEM-UA parameterizations that were evaluated.

***) P5219L14: How much is this influenced by the objective function? If it is influenced so much by the objective function, wouldn't it be better to try a different objective function first rather than changing the model?**

The optimal parameter set changes with (1) the objective function which is used to rate the quality of the model simulation in SCEM-UA; (2) the particular subset of the data which is used to calibrate the model to the observations. As a rule of thumb, you only change the objective function if you realize that even the optimal solution for the current objective function will not reproduce the aspects of the data that you are focusing on (for example, high flows). However, our goal was to have a global fit on all of the data for this event. We could have used other objective functions such as the sum of absolute deviations or the Nash-Sutcliffe efficiency, but this would not have changed our conclusions.

***) P5220L1: Reference papers that describe this.**

Reference added as suggested.

***) P5222L12: Do you have any field evidence for that? If so include or include**

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

references.

Changed.

***) P5224L3 and P5225L1: What is the Nash-Sutcliff efficiency for these models?**

In the revised version of the manuscript, the Nash-Sutcliffe efficiencies have been included for virtual experiment 4.

***) P5526L1-5: If the model results are so dependent on these initial conditions, errors and choice of where to position the transect how can these models be applied in other cases? Please discuss or describe the importance of this issue in more detail and give suggestions for how to deal with this.**

Of course, it would be great if we can accurately predict the behavior of other, perhaps ungauged, hillslopes. However, we believe that there is still a lot of research to be done before we can hope to achieve this. Recently, various types of nonlinearities have been recognized in the hillslope hydrological system. In terms of extrapolation to other hillslopes, the implications of these nonlinearities are not very well understood at all. Therefore, we can only give some general advice. In the revised version of the paper, we dedicated a new section to describing what we believe is needed for progress in hillslope hydrology. This new section advocates a more widespread use of calibration methods that account for uncertainty about the model states (as well as model parameters; e.g. Vrugt et al. 2005, WRR 41, W01017).

***) P5226L6: Be more specific about what the ”more effective experimental design” should entail? Based on the results of this study, what should an effective experimental design look like or what should it include?**

Upon re-reading the text on page 5226, line 6, we noticed that it was ambiguous. We changed the wording to emphasize the point we are making in the rest of this paragraph. What we meant is that you cannot distinguish which model—if any—is appropriate for the hillslope shown in Fig. 6, without introducing distributed data. The reason why we do not actually do this, is that (1) we wanted to show a model that could represent the nonlinear discharge behavior of Panola and other hillslopes, while also satisfying additional ‘soft’ criteria (summarized in Table 2); (2) A comparison with distributed groundwater tables would not be useful, since we are using a transect of spatial elements. Therefore a comparison would need to incorporate aggregating transient groundwater tables along isolines, which is a less than trivial problem.

***) P5226L15-19: This sentence is not clear to me. What do you mean by “unjustified to continue the iterative research cycle”? And why is it unjustified? Rewrite this section and explain better.**

For our explanation, the reader is referred to the Referee’s previous question in this document. In the revised version of the paper, we rewrote parts of the discussion of Virtual Experiment 4 and explained our point of view more elaborately.

3 Editorial points

P5206L2: replace “soil water transport” by “runoff”?

Changed as suggested.

P5206L17: “discharge being too steep” replace by “hydrograph being too steep”

Changed as suggested.

P5207L9-10: rewrite this sentence. It isn’t right to say that these regions have

steep slopes because of the unconsolidated nature of the materials.

Changed.

P5207L24: move reference to the end of the sentence?

Changed as suggested.

P5208L9: replace “organically” by “organic”

Changed.

P5208L26: replace “with” by “for”?

Changed as suggested.

P5210L11: You should reference the WRR 2008 Panola dataset here, assuming that that is where you obtained the data from.

Changed as suggested.

P5210L14: replace “Our streamflow” by “The subsurface flow”

Changed as suggested.

P5210L22: replace “sum of squares” by “sum of squared errors”?

Changed as suggested.

P5211L8: replace “steep increase” by “sharp increase”

Changed as suggested.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

P5213L14: describe what “Vm” stands for here instead of on P5214L2.

Changed as suggested.

P5216L9: do you mean “observation times” instead of “observation points”?

Changed as suggested.

P5216L16: insert “, as observed for many field studies” at the end of the sentence?

Changed as suggested.

P5218L19-20: reword this sentence. “Discharge becomes less steep”?

Changed.

P5224L4-6: the repetition of this sentence at the end of each paragraph is a bit awkward. Can you not just mention that it is given for all (subsequent) model simulations when you first introduce table 2?

Changed as suggested.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 6, 5205, 2009.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper