Response to comments from Reviewer 2 on "*The evolution of process-based hydrologic models: Historical challenges and the collective quest for physical realism*" by Martyn P. Clark et al.

[Responses are in red font at the bottom each sub-section].

1 Summary

The manuscript is one in a series of discussion papers by the lead author over the last few years and was prompted by the Symposium in Honor of Eric Wood: Observations and Modeling across Scales held June 2-3, 2016 in Princeton, NJ. The authors use three questions posed by Freeze and Harlan [1969] to examine progress in process-based hydrologic modeling over the last fifty years and to define outstanding research challenges. The manuscript is generally well-written, but is more narrow in focus than its title and introduction suggest. The specific answers to the three questions are not based on general surveys of the field, but are provided through the lens of the symposium topics. This is probably a good thing, since it constrains the length of the manuscript, but it requires some rewriting or additional wording to clearly define how the questions from Freeze and Harlan are evaluated in the rest of the manuscript.

Thanks for these comments. We have revised the introduction to clarify the scope of the paper. Specifically, in the Introduction we state:

We do not mean to provide a comprehensive review; rather, we present possible solutions to outstanding modeling problems, focusing attention on the research sphere of Eric F. Wood.

2 Specific comments

1. For example, the first question discusses whether physically-based mathematical descriptions of hydrologic processes are available and whether the relationships between the component phenomena are well enough understood. This question is addressed in section 2 (model structure) by focusing on scaling relationships and the representation of local processes in regional models. While an important aspect of model representation and one that continues to challenge the community, it is not the sole challenge to the question posed by Freeze and Harlan (but an obvious angle given the topic of the symposium). The manuscript would be improved by a better connection between the question from Freeze and Harlan and the discussion of scaling. Explain why scaling is the main topic that is being discussed and how it relates or ranks compared to other aspects of question one. Following that, the discussion of scaling in section 2 can remain largely unchanged.

   We have modified the discussion at the start of section 2 to clarify that we focus on contributions from Eric F. Wood.

2. Same for model parameters (section 3). The problem is not only that spatial information is not always of sufficient resolution and quality or that the information does not exist, but also that some parameters are not directly observable at the scale of the application (it's not even always clear that the equations we use at certain

spatial scales are necessarily the right ones). The challenge (l.229 and following) is not only to make the best use of the information that we have, but there may also be an opportunity to change our physical descriptions to make better use of the available information at a particular scale. Section 3 focuses mostly on parameter upscaling and does not discuss the use of new data sources, the use of inverse methods, etc. I am not advocating to discuss all topics because the manuscript will lose focus, but it would be good to motivate better why the authors focus on this particular aspect of question 2 from Freeze and Harlan.

While we appreciate the limitations in scope, we do discuss the points raised by this reviewer.

Some parameters are not directly observable at the scale of the application: We discuss the challenges as "*when spatial information does exist it may have limited spatial representativeness and relevance – for example, the information on hydraulic conductivity from soil pits may only have weak relations with the transmission of water throughout catchments [Beven 1989].*"

Section 3 focuses mostly on parameter upscaling and does not discuss the use of new data sources, the use of inverse methods, etc: We already discuss new data sources "*[…] there are numerous opportunities to improve information on geophysical properties, including estimates of vegetation structure [Simard et al. 2011], soil depth [Pelletier et al. 2016], soil properties [Chaney et al. 2016b], bedrock depth and permeability [Fan et al. 2015] and the physical characteristics of rivers [Gleason and Smith 2014].*" We also discuss the use of inverse methods "*[…] there is considerable scope to improve the way that multivariate data is used to constrain model parameter values […]*".

As the reviewer notes, it is impossible to discuss everything, and we believe that the additional motivation provided at the start of the paper (i.e., where we state that we simply present examples, and that we do not intend to be comprehensive) is sufficient for the reader to appreciate the limited scope of our paper.

3. The section on model execution (section 4) requires some discussion of the purpose or end goal of our simulations. One could argue that we run at higher resolutions and finer time steps simply because we can and because we lack the scaling relationships that allow us to be more "economical" with computing resources. The statement (l.320-321) that "[...] more complex models may not have as much physical realism as computationally frugal alternatives" raises the question 'why are we doing them?'. And the final paragraph of section 4.2 leaves me again wondering what the end goal is of the model simulations.

   Good point. We included additional discussion in the section on model execution:

   A key reason for conducting such spatially resolved simulations is to understand explicit spatial controls on hydrologic processes – for example, Maxwell and Condon [2016] use high resolution continental-domain ParFlow simulations to understand the controls of groundwater flow on the partitioning of evapotranspiration into bare soil evaporation and transpiration.

4. l.389-l.399: It would be nice to call out some of the specific advances that have been made in response to the questions from Freeze and Harlan.

   We appreciate this comment. Specific modeling advances are defined throughout the paper, and we do not see the need repeat the modeling advances in the conclusions. Rather, we have expanded the conclusions to address the comment from Eric Wood to define a path forward for the community.

3 Minor comments

- l.101-102: "both sets of solutions can occur in the same model". Do you mean for the same process? This is not entirely clear and an example would be helpful.

  We have revised the text to state that many models include a mix of small-scale and large-scale flux parameterizations (e.g., VIC uses a large-scale parameterization of infiltration, yet relies on small-scale equations to simulate the storage and transmission of water through the soil matrix).

- l.103-104: "are readily shared among different modeling groups". I don't quite understand what "shared" means in this context (even given the example).

  The revision above helps address this issue. Further, we have revised the text to state "*When viewed in this way, the different solutions to the scaling/closure problem can be shared among different modeling groups that employ very different modeling approaches.*"

- l.129: "summarizes recent in developing". Word missing after "recent"

  Fixed. "*[…] recent advances […]*"

- l.141: "Another class of methods is" Suggested change "Another class of methods consists of"

  Suggestion adopted. Thanks.

- l.143: "described earlier" Where, I cannot find the earlier reference.

  Clarified. The text now reads "*Examples of this class of methods include the empirically derived storage-discharge relationships described earlier, where the large-scale transmission of water is often defined as a linear (or near-linear) function of water storage [Ambroise et al. 1996; Clark et al. 2008; Fenicia et al. 2011; Brauer et al. 2014].*"

- l.157: "However, if l and D are comparable in scale, this becomes problematic." Explain why.

  We have modified the text to state "*However, it becomes more difficult to define scale-aware parameterizations if l and D are comparable in scale.*"

- l.284: "for upper 50 m". Suggested change "for the upper 50 m"

Done.

- l.289: "on the other hand". Missing a "on the one hand"

  We now simply state: *"[…] these datasets have been globally extrapolated from locally established empirical relationships between subsurface properties and surface lithology [Hartmann and Moosdorf 2012]."*

- l.291: "guide the interaction" should read "guide to the interaction

  We have revised the text to state *"As a consequence, they provide useful information on the interaction between groundwater and evaporation, but have limited use […]"*