The paper represents a first attempt to use parflow to simulate continental scale hydrologic processes, demonstrating a methodology for using existing physical observations to generate a numerical simulation that results in a natural organization of stream networks and groundwater flow. Such an approach may also be of interest to the geomorphology and geochemistry communities. The manuscript should definitely be published in GMD.

The previous review comments reflect considerable differences of opinion regarding novelty and the conceptualization. Although the authors have done a commendable job in responding to these comments, one thing I suggest the authors consider is a more detailed statement, possibly at the end of the results/discussion regarding (a) how or could such simulations could be improved through better parameterization (for example better resolution in the depth of subsurface, fixing state boundaries, etc? There is a general statement at line 487 but thoughts on how to prioritize additional products would be useful. (b) Do the authors think the resolution (perhaps for surface water) and parameterization impacts any of the scaling results (for example around figure 10 and 11).

Does the paper address relevant scientific modelling questions within the scope of GMD? Does the paper present a model, advances in modelling science or a modelling protocol that is suitable for addressing relevant scientific questions within the scope of EGU?

Yes

1. Does the paper present novel concepts, ideas, tools, or data?

Yes

2. Does the paper represent a sufficiently substantial advance in modelling science?

Yes

3. Are the methods and assumptions valid and clearly outlined?

Mostly, but a few areas could be improved.

4. Are the results sufficient to support the interpretations and conclusions?

Yes

5. Is the description sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)? In the case of model description papers, it should in theory be possible for an independent scientist to construct a model that, while not necessarily numerically identical, will produce scientifically equivalent results. Model development papers should be similarly reproducible.

For MIP and benchmarking papers it should be possible for the protocol to be precisely reproduced for an independent model. Descriptions of numerical advances should be precisely reproducible.

Mostly, a few small details that could be expanded as noted below.

6. Do the authors give proper credit to related work and clearly indicate their own new/original contribution?

Yes

7. Does the title clearly reflect the contents of the paper? The model name and number should be included in papers that deal with only one model.

Yes

8. Does the abstract provide a concise and complete summary?

Yes

9. Is the overall presentation well structured and clear?

Mostly, a few small comments and typos.

10. Is the language fluent and precise?

Yes.

11. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?

Yes, with exceptions of a minor issue of missing spaces between values and units and some small typos.

12. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?

Only small suggestion of more detail in the methods.

13. Are the number and quality of references appropriate?

Yes

14. Is the amount and quality of supplementary material appropriate? For model description papers, authors are strongly encouraged to submit supplementary material containing the model code and a user manual. For development,

technical and benchmarking papers, the submission of code to perform calculations described in the text is strongly encouraged.

N/A

Detailed comments/suggested corrections:

Line 17: "tools for" and "processes; however, "
Line 19: need to be or can only be?
Line 22: Why give the area twice? First time seems sufficient.
Line 30: recommend removing the word novel – readers should decide for themselves.
Line 45: "bridge this gap" is rather informal. Suggest replacing this phrase.
Line 78: "groundwater-surface water"
Line 97: "extent (i.e. continental scale)" (there is more than one continent).
Line 98: missing space "(1 km)
Line 134: define Sx, Sy?
Line 166: "Physically this" – define "this" or delete.
Line 205: "functional relationships that impose"
Line 208: This somewhat repeats what is stated at Line 162.
Line 221-237: missing spaces between values and units throughout: 102 m, 1 km, 100 m, 1 x 1 km, etc. please correct. This occurs throughout manuscript and some figures and should be corrected.

Line 229: It seems to me that several things are missing from the discussion. First, how is porosity specified? Second, at some point in the manuscript a discussion or acknowledgment regarding the bottom layer of the model relative to realistic physical conditions should be included. I am not familiar with the exact details or issues associated with the Gleeson et al. (2001) map, but the current simulation includes a highly simplified treatment of the actual bottom boundary of the subsurface – namely the presence of low-porosity impermeable bedrock. I realize there is no other option than the assumption made, but if for example depth of regolith would improve the model considerably vs. not at all, that is important information to provide to the community.

Finally, it would be helpful to provide more information about the various products used here. I realize the references are included but for example, how is the soil conductivity in the two layers determined? Also, the basics of the Gleeson map construction.

Figure 1: I did not find the description very clear relative to figure 1. What is plotted in (E)? Conductivity where? Below 2 m?

Line 230: SSURGO – please define acronym.
Line 237: minor point but Sx,Sy ook lowercase in equation – are uppercase previously
Line 239: tense "reduce"
Line 245:246: This is unclear – what kind of uncertainty? "As such" does not fit here.
Line 248: Interfaces "between" property values.
Line 257: possibly add "variability using best available data". Also "impact of uncertain"

Line 260-262: This statement is rather confusing. Please clarify. Was a new product created from another product or are the authors referring to an actual product? I also found it unclear exactly how unsaturated flow is handled- prescribed recharge? Please clarify with respect to equation 1.

Line 277: "far from being trivial". What does this mean? Can authors be specific? Trivial is a relative word.

Line 281: "still better than 60% efficiency"

Line 293: "checked for plausibility against"? Do authors mean: "are compared to observations"?

Line 342: Change "while" to "although"

Line 344: "in future work"?

Line 347: What is a plausibility check? This is an evaluation against parameters that were not used to develop the simulations. Out of curiousity, are any of these datasets used to generate the P-ET or other "products" ?

Line 354: Please clarify this statement –

Line 497: what does within "the confines of the numerical experiment" mean?