Earth Surf. Dynam. Discuss.,
https://doi.org/10.5194/esurf-2020-64-RC1, 2020

Earth Surface Dynamics Discussions

Interactive comment on "How do modeling choices impact the representation of structural connectivity and the dynamics of suspended sediment fluxes in distributed soil erosion models?" *by* Magdalena Uber et al.

Anonymous Referee #1

Received and published: 26 October 2020

1 Summary

Uber et al. present a numerical modeling study that explores how modeling choices related to computational mesh generation, parameterization, and source-classification grouping influences a variety of output metrics describing hydrograph and sedigraph characteristics. The authors focus their work in two well studied mesoscale catchments and connect the results of their sensitivity analysis with basin-scale characteristics (e.g., mean slope).

The presented research represents a valuable contribution to understanding how the choices made in setting up a computational model influence model results. Placing this level of attention on model set up is rarely described in formal publications, yet is critical to understanding when and how models can be applied to understand and/or predict systems of interest. In addition, the author's well designed numerical experiments permits an assessment of how discretization and parameterization impact basin hydrograph and sedigraph dynamics.

Below I describe comments and recommendations first in narrative form and then as line-level comments. My most substantial concern is that the paper lacks an overarching introduction to the study design—a section in which the authors set up the specific questions or hypothesis that they seek to address and connect them with a conceptual description of their numerical experiment design. A related comment is that I found the explanation of the modeling choices difficult to follow. Both of these issues meant that it was difficult to connect the study design and methods with the results and discussion.

I recommend acceptance after major revisions and look forward to seeing this paper published.

2 Narrative Comments

2.1 Addition of an "Study Design" Section

The experimental design employed by the authors is valid and appropriate for the questions that they seek to pose. However, I found description clearly connecting the big picture questions ("what controls sediment flux from mesoscale watersheds") to the scenario design currently introduced by Section 3.4 and Table 2 was missing, or spread across too many sections of the paper.

I recommend that a new section be placed immediately after the introduction. In this section you would describe your experimental design and connect it to the big picture you have laid out in your introduction. Such a section would include the specific questions and hypotheses each scenario's experiment seeks to answer and an explanation of why this question was targeted.

While the reader may not know the details of the two sites or the model, your introduction should provide enough information such that this section can come before the more detailed methods section. Such a section will introduce to the reader the concrete questions your scenarios were designed to address.

Such as section should a description of the type of model analysis method used (e.g., a series of one-at-a-time sensitivity studies) and explain why this sort of method is appropriate to address the study objectives. Pianosi et al. (2016) is a good place to start for background on this topic. This will help the reader understand the type of results you will obtain.

In such a section, I would also like to see an introduction to why two catchments are used and why calculating whole-catchment connectivity metrics (described in Section 3.1); e.g., doing the same set of simulations across two catchments with different geology/land use/etc allows you to isolate how transferable your results are to catchments with different properties. This would also allow you to set up why you calculate a variety of catchment connectivity metrics (presented in Table 1) and explicitly state that you will eventually work to connect those connectivity metrics with the variability identified by the sensitivity analysis (a start at this is done at L461).

2.2 Improve explanation of modeling choices

The core of the study hinges on connecting the modeling set up described in Section 3.3 to the scenarios described in Section 3.4. However, I found it difficult to connect

these two sections, mostly because I found it hard to follow exactly what the authors varied in their modeling set up.

The most constructive form of feedback I think I can provide here is a summary of what I understood after reading the paper four times, as well as what I would recommend so that I might have understood this after the first reading.

Based on my reading, what I understand is that there Iber requires a computational mesh, and the mesh size can vary in space. Each mesh cell has a value for Manning's n and a value for α. *Choice 1*: The considered area is divided up into three conceptual domains which influence the grid cell size and Manning's n value based on the CDA (hillslope, channel, badlands). Based on the delineation of these domains the mesh is discretized.

Next the mesh is parameterized with a spatially variable for Manning's n value. You might have chosen to let Manning's n vary smoothly, or something else, but you have chosen that the domain will get two Manning's values (channel and hillslope). *Choice 2* focuses on those values.

While water can fall on and run across the entire computational mesh, sediment can only be sourced from the bare bedrock areas. In these areas, the propensity to produce sediment is parameterized by α.

I don't think the following was ever stated, but in order to produce the source proportion sedigraphs, I believe that some method of source tracking can be chosen in order to elucidate the dynamics of the basin. Different classification of these tracked sources is represented by *Choice 3* (I think).

Thus Scenarios 2a–2d focus on Choice 1, Scenarios 3a–f focus on Choice 2, and I think that different delineations of source tracking (Choice 3), along with different choices for Manning's n yield Scenario 4.

I would recommend the following to the authors:

- Revise section 3.3 to describe more clearly what the modeling choices are such that they set the reader up to understand the details of scenario design discussed in the following section.
- In Section 3.3 or in the new "study design" section proposed above, explain why these choices are important to focus on. Are they the only choices? Are they the only ones which carry uncertainty? There are many things you might have focused on (e.g., assess the sensitivity to the channel grid cell size), but you chose these elements, why? To be clear, I think the elements you've chosen are great, I just want more description of why they were chosen.
- Clarify how the source classification is represented in model specification. Does this choice not influence the model physics, but just the model output that permits a different view on the dynamics?
- Explain why sediment is only sourced from the bare bedrock.

2.3 Improve connection between study design and discussion

The structure of the discussion roughly follows the three non-base case scenarios and presents the most salient aspects of the results. However, within each of the major discussion sections, I found the text difficult to follow. I suspect that by being more explicit about the target questions and hypotheses earlier in the text the authors will be able to very lightly restructure the discussion such that the reader is easily able to connect the discussion with the study intent and numerical experiments.

In addition, the end of the discussion starts to tie together the basin-scale metrics presented in Table 1 and the numerical modeling results. It would be beneficial to introduce earlier on that you will do this and describe in more detail how this is accomplished (e.g., regression, rank correlation). Knowing that this sort of analysis is coming

will help explain why all of the basin-scale metric are calculated and discussed starting at L136.

2.4 Figures

The interactive figures provided by Uber et al. (2020) are a fantastic complement to the paper. I might consider adding catchment as a facet (e.g., facet grid with scenario catchment) because this would facilitate comparison between catchments. I'd also like to applaud your consistency in the use of color to denote geological unit across figures. This should be a standard expectation, but it isn't, and it makes comprehension much better.

My primary concern with figures relates to the maps presented in Figure 1. This figure shows us inconsistent information across the two catchments (e.g., badlands only shown in 1a) and does not show us all of the information used in the modeling study that is the focus of the work. I recommend that Figure 1 be redrafted into a series of rows that shows the reader the main elements used in model initialization for each catchment. For example, row one might show a shaded relief map with the river system and badlands areas, row 2 would show the considered geologic units used, row 3 might show the weighting factor W presented by Borselli et al. (2008), while row 4 would show the roughness based weighting factor of Cavalli et al. (2013).

2.5 Code availability

For the purposes of computational reproducibility, state the version of Iber used.

No statement has been made about model input file availability. Such files should be digitally archived for the purpose of reproducibility.

3 Line Level Comments

Bullet points in this Section indicate "<LineNumber>", "F<Figure Number>", or "T<Table Number>".

- 36 The term "Mediterranean and mountainous" is used a few times, first here. Mediterranean could be interpreted a few ways: e.g., places with a Mediterranean climate, places near the Mediterranean. Recommend being more specific about what is meant.
- 56 Recommend giving an example of your objectives and thus how structural connectivity is represented to anchor this abstract concept on a concrete example or two.
- 76 I suspect the sentence that ends in this line needs a reference.
- 87 Be more specific about which models and provide examples with associated references.
- 100 Additional subsubsection headers would have helped me understand this section more easily. For example Section 3.1 discusses both a description of the catchments and connectivity metrics calculated, and Section 3.3 discusses many different aspects of the model set up. I would split each of these subsections into multiple subsubsections.
- 129 A few lines or a paragraph summarizing the similarities and differences of the two catchments would benefit the reader here.
- 136 Some statements about why these connectivity metrics were chosen would benefit the reader. In addition, explain (here or in something like the proposed "Study Design" section) what you expect to learn from these metrics and how they are used.

- 137 The distance to the outlet metric has been called the "width function" by the landscape evolution modeling community Hancock et al. (2010, 2002). Work by this community has shown that it is not a particularly good metric for comparing catchment topography, but is a does provide a good assessment of hydrology. It may be useful to connect with this literature.
- 138 Mathematically represent the connectivity indices of Borselli et al. (2008) and Cavalli et al. (2013) here so that the reader can more clearly understand what they represent.
- 171–173 This detail of model set up should be located elsewhere. Probably is a subsubsection of Subsection 3.3 (see also the comment at L237 and 289.
- 211 Being able to connect this discussion of badlands in model set up to a consistent picture of where badlands are located is why I mentioned earlier that Figure 1 should be revised to include consistent information about each catchment.
- 215 Connect and justify the choice of a 5 m minimum grid size with relevant field observations and the numerics of the Iber model? E.g., how does this compare with the range of values for channel width in each catchment? Do the numerics of Iber benefit from a relationship between minimum grid cell size and channel width (e.g., smallest grid cell = channel width, 10 grid cells = channel width).
- 217 20 m seems like a rather large grid cell size for gullied areas. Explain and/or justify this value.
- 222 The erosion source locations should be shown in Figure 1 in addition to the subplots shown in later figures.
- 222 If I'm interpreting this correctly, I believe you are saying that sediment production can only occur in the areas of bare bedrock. This should be explained further and justified. In addition, discuss how this model set up decision impacts the

implications of this study for overall soil erosion (as these bare bedrock patches only make up a small portion of the study watershed).

227–236 It is difficult to understand if this section of text is summarizing the work of Uber et al. (2019) or if it is presenting an analysis of modeling results. Revise to clarify this point.

227 Introduce the units of α when the variable is first presented.

237 No discussion of time discretization, model run duration, or external forcing (e.g., rain) is present in the prior subsection. These elements of model set and running should be discussed.

237 Based on the results presented, it appears that Iber has the capability of tracking the source of water/sediment as it moves through the catchment and that how these source regions are grouped is what is meant by the "source classification" column of Table 2. This aspect of the model should be discussed. As best as I can tell this is a critical aspect of interpreting Scenario 4.

In addition, it is not clear whether this choice of model set up impacts the dynamics of water and sediment (or if it just impacts how they are analyzed). E.g., are simulation 1 and 4a and 4b the same simulation just analyzed/post processed differently?

260 The simulations of Scenario 3 represent two one-at-a-time sensitivity studies (Sc. 3a–3c for sensitivity to hillslope Manning's n and Sc. 3d–3f for channel). Recommend using more formal language to describe the numerical experiments as it will help the reader anticipate the type of results presented.

268 It is not clear to me how the the different options for source classification of Scenario 4 relate to changes in the parameterization of the model. Were different values of α used? Something else? Clarify.

In addition, these scenarios include two options for the Manning's n value, the base case and one in which the hillslope value is low and the channel value is high. The results of Scenarios 4c and d are discussed at L454. Formally introduce what the purpose of this sub-scenario is.

272–274 This sentence, in which you link the changes to the model set up with a hypothesis is exactly the sort of text that a "Study Design" section would benefit from. Recommend that similar sentences for each scenario exist and be present in such a section.

280 This section clearly describes what model output metrics were used, however it does not explain why these output metrics were chosen or justify why the are appropriate given the overall goals of the study. This section should be expanded to include this information.

289 This sentence describing model run details should go elsewhere in the text. Probably in a section on external forcing, along with the text currently located at L171–173 (see comment at L237).

296 Be more specific about which aspects of the model. Some aspects are sensitive and some are not.

307 Connect this statement with new text earlier in the paper describing why two catchments are used. Set the reader up for this sort of discussion by explaining why two catchments are used, and comparing/contrasting them.

313 Justify why this is a reasonable interpretation and connect with literature.

337 This statement presents a different conclusion than Table 3 and the text near L296 which states that different CDA values result in output metric variability. These three elements of results and discussion should be consistent.

344-350 The purpose and reasoning of the argument you advance here is not clear. As you highlight it in the conclusion (L487) I believe you think it is an important point. Recommend this text be revised.

352 The section of Table 2 that shows the results of Scenario 3 indicates that changing Manning's n in the hillslope has a larger impact on the results than changing the channel value. This should be discussed.

372 What is meant by "more stable"?

379 Here and elsewhere, sensitivity should be presented as a relative measure. E.g., this output was more sensitive to choice/parameter A than to choice/parameter B. Without the comparison the statement is uninterpretable.

392 Here you discuss both a contrast between the two catchments, the analysis of Scenario 4, and connecting basin-wide metrics of IC with the sensitivity results. Recommend structuring the section to help the reader anticipate this.

393 Introduce this idea in the study design.

397–399 This has already been stated.

402 Add a figure reference.

407 More specific. E.g., close = first, or something different?

421 It is not clear if Scenario 4 represents a different approach to tracking something else? Because the description of how Sc. 4 was constructed is incomplete is is nearly impossible to understand the results of Sc. 4.

423–425 Give the reader a little more context about "typical interpretations of discharge-sediment flux hysteresis" and provide a description of what a clockwise vs counterclockwise loop means.

431 Not sure what is meant by this sentence.

448 Unclear if distance to the outlet (or stream) being considered is related to the parameterization or the analysis of the results.

461 This sentence starts a new line of inquiry: which basin-wide metrics (Table 1) best predict the sensitivities documented by the numerical experiments. A more explicit discussion of the methods used here (e.g., comparing basin wide metrics to sensitivity ranking) should be added to the methods. In addition, the description of this analysis should be expanded.

465 This sentence is not clear.

468 It is not clear that your study design supports this type of analysis. To my ability to tell you have not varied the location and/or erodibility of the sediment sources within the catchment. As such, your study design does not permit assessment of how variability in location of sediment sources influences the output metrics.

469 The point you are making here is not clear, mostly because the text introduced at L344-350 is not clear.

478 Unclear how the study is about source soils when the only erodible material is the exposed bedrock. This should be addressed here and earlier in the text.

ost Figures In the many multi-panel plots I recommend use of consistent x and y axis limits and/or explicit notation of inconsistent axis limits in Figure captions.

F10–13 The panel (f) is the sort of information that would be great to have in a revised Figure 1. The background color scheme for the inset maps (distance to outlet, distance to stream) should be represented by a legend.

T2 The layout of the table makes it difficult to see the difference between the scenario 4 options.

T3 1. Why are the simulations used for Scenario 4 not in the table?

2. Recommend adding some vertical lines to help guide the viewer in separating Sc. 1, Sc. 2, and the two halves of Sc. 3.

3. Overlaying the table text on top of a tile plot is a great addition. However, the darkest blue values make reading the text impossible.

4. Not clear why some values have NA, explain.

References

Borselli, L., Cassi, P., and Torri, D.: Prolegomena to sediment and flow connectivity in the landscape: a GIS and field numerical assessment, Catena, 75, 268–277, 2008.

Cavalli, M., Trevisani, S., Comiti, F., and Marchi, L.: Geomorphometric assessment of spatial sediment connectivity in small Alpine catchments, Geomorphology, 188, 31–41, 2013.

Hancock, G., Willgoose, G., and Evans, K.: Testing of the SIBERIA landscape evolution model using the Tin Camp Creek, Northern Territory, Australia, field catchment, Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group, 27, 125–143, 2002.

Hancock, G., Lowry, J., Coulthard, T., Evans, K., and Moliere, D.: A catchment scale evaluation of the SIBERIA and CAESAR landscape evolution models, Earth Surface Processes and Landforms, 35, 863–875, 2010.

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., and Wagener, T.: Sensitivity analysis of environmental models: A systematic review with practical workflow, Environmental Modelling & Software, 79, 214–232, 2016.

Uber, M., Legout, C., Nord, G., Crouzet, C., Demory, F., and Poulenard, J.: Comparing alternative tracing measurements and mixing models to fingerprint suspended sediment sources in a mesoscale Mediterranean catchment, Journal of Soils and Sediments, 19, 3255–3273, 2019.

Uber, M., Nord, G., Legout, C., and Cea, L.: Modeled contributions of sediment sources to total suspended sediment flux in two mesoscale catchments, https://doi.org/http://dx.doi.org/10.17178/EROSION_MODEL.2020, 2020.

Interactive comment on Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2020-64, 2020.