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 Uber et al.

Answer to reviewer #1

We wish to thank the anonymous referee #1 for this very detailed and constructive review of our study and acknowledge the time spent and effort made. His/her comments helped us to substantially improve the paper and we hope that the changes made accordingly will contribute to an easier understanding of the text. In the following, the reviewer comments appear in black italic and our answers are provided in blue. When there are quotations from the text of the article, they appear in quotation marks.

As a general response, we would like to point out that, upon reading many of the reviewer's comments or questions, we realized that the second objective of the article was mis-explained and therefore misunderstood. While the first objective is achieved by performing a sensitivity analysis of the choices made during the construction of the models (modelling scenarios 1, 2 and 3 in Table 2), the second objective corresponds to an opening towards the understanding of the temporal dynamics of fine sediment fluxes as a function of the geomorphological characteristics of two different watersheds, in particular due to the location of the sources and their structural connectivity. Thus, the scenarios 4 described in Table 2 allow a better visualization and interpretation of the contributions of the different subcategories of sedimentary sources to the outlets.

In the revised version of the manuscript, we propose to add, as proposed, a study design section that links the different modeling scenarios in Table 2 to the two objectives reformulated in the table to be more explicitly linked to those announced in the introduction, also slightly reformulated.

After reading the comments of referee #1 we also became aware that the title of the manuscript only referred to our first objective while it did not refer to the second one. Thus we changed the title to the manuscript to "How do modeling choices and erosion zone locations impact the representation of connectivity and the dynamics of suspended sediments in a multi-source soil erosion model?"

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. […]

Thank you for your general acknowledgement and positive feedback on our work.

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.

We addressed your concerns by including a short section "study design" as you proposed and made changes in the description of the modeling scenarios to better understand the modeling choices (see the further points).

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

Thanks again for the constructive proposals and the recommendation for publication.

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).

We introduced a section "study design" as you proposed. However, we introduced it as an introduction of the modeling scenarios section, as it is directly linked to the description of the scenarios. While this section is short we hope that the changes made in further sections will also help to improve the understanding of the study design. Thank you also for the recommendation of introducing our approach with the paper by Pianosi et al. and the hints to be more precise on the type of sensitivity analysis conducted.

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.

Thank you for your summary from an outside perspective which helped us to be more precise on some parts, see comments below.

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.

This is correct. We try to be more precise by changing the first sentence in section 3.3 that now reads "As a distributed model, Iber requires a computational mesh which is made up by three main modeling units with different spatial discretizations and roughness coefficients, i.e. the river network, the hillslopes and the badlands."

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).

This is correct. Again, we try to be more precise by adding "Values for Manning's n and erodibility were assigned spatially distributed to each mesh element." in line 219 of the initially submitted version of the manuscript.

Choice2 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.

We reformulated the sentence starting in line 222 which now reads "While runoff is generated and routed in the entire catchment, the production of sediment was limited to the potential erosion zones. The latter include all the mesh elements in the modeling unit "badland" and the mesh elements on the hillslopes that belonged to the diffuse agricultural sources in the Claduègne catchment. The erosion zones were classified according to ..."

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.

To be more precise about that, we reformulated the sentence starting in line 225 which now reads "Sediment production (Eq. 2) was calculated in each mesh element of the potential erosion zones for each source class separately. Sediment transfer is then routed over the entire catchment. Thus, separate sedigraphs for each source class were obtained at the outlet and the contribution of each source class to total sediment flux could be calculated for every time step."

Furthermore, we thoroughly revised the description of the model in section 3.2 to be precise about this aspect. In Eq. 2 we added the subscript s to be more explicit about the fact that it was solved for each sediment class separately.

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.

As mentioned in the general answer, the last set of scenarios (Sc. 4) were designed to answer the second objective written at the end of the introduction. The aim of Sc. 4 is to better interpret the modelled temporal dynamics of sediment fluxes for various groups of sediments depending on their geology and also on their distance to the outlet or to the river network. Thus, Sc. 4a and 4b do not really correspond to choices during modeling set up as the overall sedigraphs are not modified. They just allow a better visualization of the sediment origin (in subgroups) in order to facilitate the comparison with the connectivity indicators. To go further in the discussion and the interpretation of the impact of the location of sources within the catchment, and particularly to assess to which extent the conclusions derived from Sc. 4a and 4b were dependent on changes in roughness parameters, Sc. 4c and 4d were added, but they were initially not designed to be part of the sensitivity analysis conducted for objective 1.

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.*

This suggestion was accepted and included in the new study design section where it reads "Based on preliminary studies that are not reported here, these factors were found to be the most important ones in determining sediment flux dynamics. While other factors (erodibility, rainfall intensity) crucially influence absolute values of erosion and suspended sediment concentration, their values are less important to determine arrival times and temporal dynamics of source contributions."

- *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?*

Yes, this is the case. E.g. the sedigraphs of the sources "Limestone 1" (close to the outlet) and "Limestone 2" (further) in scenario 4a sum up to the sedigraph of the source "Limestone" (which includes close and distant subsources) in the basic scenario. Thank you for pointing out that this was not clear. The following sentence was added at the end of section 3.4 "It should be stressed that this source classification does not influence model physics, i.e. total sediment yield from a source (close + distant sources) remains the same as in the basic scenario where they are not differentiated."

- *Explain why sediment is only sourced from the bare bedrock.*

We changed the sentence starting in line 126 to “The land use is dominated by forests and scrublands, which are permanently covered by vegetation and are thus assumed to be negligible as sediment sources. Agricultural zones are barely present in the catchment.”

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.

Thank you for that remark. We included this idea in the new study design section by adding “[…] indicators of structural connectivity of the two catchments are used to describe the configuration of sediment sources in the catchment. They are compared to the modeled hydro-sedimentary fluxes both qualitatively by visual analyses and quantitatively by means of the calculation of characteristic time scales of the hydrographs and sedigraphs (e.g. time of concentration, lag time)”
We prefer not to use a specific term like regression or rank correlation because we are comparing only 5 data points at a maximum (4 sources in the Galabre catchment + liquid discharge)

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.

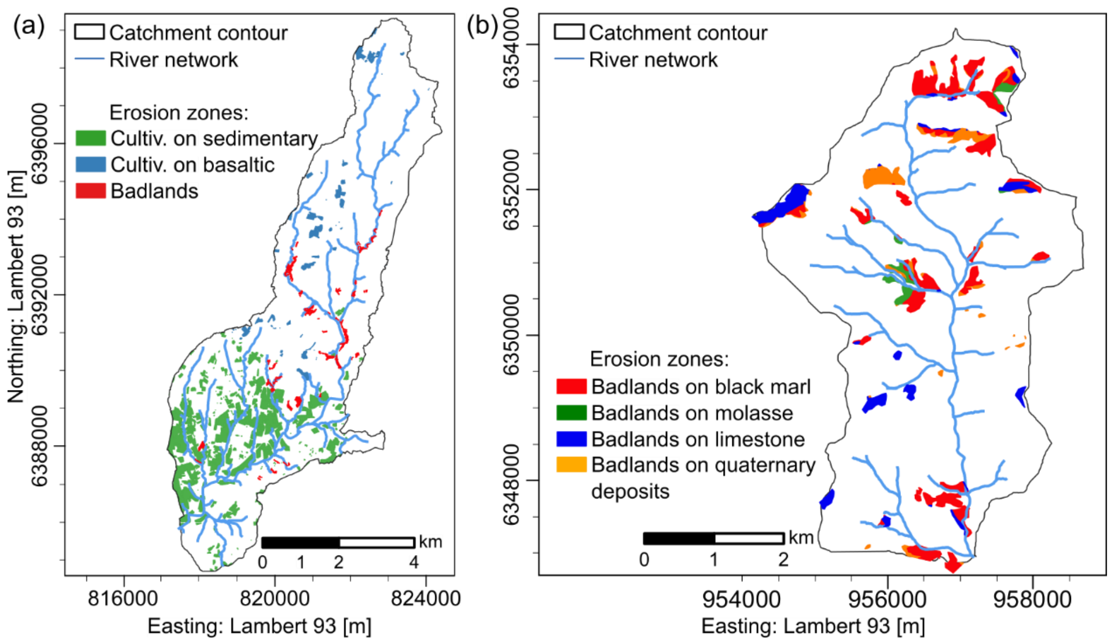
This is a nice idea, but it is not easy to implement. We prefer to keep the interactive figure as they are.

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.

Thank you very much for the positive feedback on the (interactive) figures.

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)

We revised figure 1 in a way that consistent information is shown for the two catchments. The land use information in fig 1b was omitted as it is indeed redundant for this study. Now the figure shows the erosion zones that were considered in the two catchment as colored patches.



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).

We prefer to keep figure 1 simple as the paper already has several figures that are composed of different subfigures. The information you request is contained in Magdalena Uber's PhD thesis available at https://tel.archives-ouvertes.fr/tel-02926078 and a reference was added in the caption of figure 1.

2.5 Code availability

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

We changed the introduction of the model to be more clear about the fact that we worked with a version of the model that is in development: "Surface runoff, sediment transport and soil erosion in the study catchments were modelled with an ad-hoc version of the software Iber (Bladé et al., 2014) developed in previous studies by the authors (Cea et al. 2016)". While the hydraulic model can be downloaded from the iberaula website, the erosion and sediment transport module is still a research version developed initially by Cea et al. (2016) which cannot be downloaded yet.

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

Given that the erosion and transport part of the model cannot be downloaded yet, we do not think there is any interest in dropping the input files on a repository.

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.

Thank you for pointing that out. We meant the terms as "having a Mediterranean and mountainous climate". However, this is not the case in the study by Vanmaercke et al that was cited in line 36, so we prefer to clarify it in line 42 where we replace "Mediterranean and mountainous watersheds" with "watersheds with a Mediterranean or mountainous climate".

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.

We replaced the sentence ending in this line by the sentence "In the context of soil erosion and sediment transfer studies it is of interest how active erosion zones are linked to the catchments outlet." to be more precise about the use of the concept of structural connectivity in this study.

76 I suspect the sentence that ends in this line needs a reference.

We added the reference "(Merrit et al., 2003)"

87 Be more specific about which models and provide examples with associated references.

The half-sentence "such as WEPP (Laflen et al. (1991)), Kineros (Woolhiser et al. (1990)) and Mike 11 (Hanley et al., 1998)" was added.

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.

We agree with the reviewer for section 3.1. Thus we split it in a first subsection labelled "Catchment descriptions" and a second one labelled "Connectivity indicators". However, we decided to keep section 3.3 unchanged.

129 A few lines or a paragraph summarizing the similarities and differences of the two catchments would benefit the reader here.

We agree to sum up the main differences at the end of the paragraph, by adding the sentences "In comparison, the Galabre catchment is smaller and steeper than the Claduègne catchment. The distribution of the erosion zones differs in the two catchments, with the ones in the Galabre catchment being more dispersed over the entire catchment but smaller in size due to the absence of diffuse agricultural sources." Their main similarity is the fact that they are both mesoscale catchments in a mountainous and Mediterranean context. As this is stated several times before we prefer not to repeat it again here.

136 Some statements about why these connectivity metrics were chosen would benefit the reader.

This was explained in lines 140-142 "The distance to the outlet and the distance to the stream of a given position in the catchment serve as proxies of longitudinal (upstream-downstream) and lateral (hillslope-channel) connectivity in the sense of Fryirs (2013)" and in lines 144 – 146 "However, neither of these measures takes into account surface roughness and slope. Thus, two of the most widely used indicators of connectivity, i.e. the IC proposed by Borselli et al. (2008) and the adjusted version of IC proposed by Cavalli et al. (2013), were calculated." We hope that adding the precisions in the brackets helps to better understand the explanations to the reader who might not be familiar with the work by Fryirs (2013).

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.

Ok, we included that explanation in the "Study design" section as you proposed.

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.

Thank you for the hint and the recommendation of the reference. We added the sentence "The distance to the outlet metric refers to the width function applied as a measure of network structure and catchment shape by Hancock et al. (2010).".

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.

We prefer to refer the reader to the original publication here and not go into too much detail. The two indices are not the most important metrics used in this paper. It is already a bit unusual to describe the calculations of these metrics in the study sites section but we took that decision in order to keep it short. Thus, going into further detail would be beyond the scope of this short description of the metrics.

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.

We agree and relocated the sentence in line 237.

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.

Ok, see our response to your comment on figure 1.

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).

Given that the same surface water and sediment routing equations are applied in all three units (the river network, the hillslopes and the badlands), the model presents a continuous representation of hillslopes and the river network. In order that the river flow strictly follows the slope, we had to choose a cell size that is in the order of magnitude of the resolution of the DEM (1 m). A smaller mesh size of 1 m for example would strongly increase the number of mesh elements and thus computation time so this value is a compromise between exact representation of the topography and computational efficiency. Thus, the minimum grid size of 5 m was chosen as a compromise between the representation of the flow structure in the river and computation time.

217 20 m seems like a rather large grid cell size for gullied areas. Explain and/or justify this value.

You are right that the topography on the steep badlands is not exactly reproduced by this value. Again, it presents a compromise between detail and computational efficiency. We did preliminary analyses that are not reported in this paper on the impact of the mesh size (only for hydrology and on a subcatchment) by conducting a convergence-of-the-mesh experiment: starting at a coarse mesh size and then gradually decreasing it. At some point the results converged, i.e. a smaller mesh size did not lead to significantly different results. This is how the optimal mesh size was determined. The resulting optimal mesh size of 20 m for badlands is related to the fact that the detachment of rainfall is modelled in a simple way according to the rainfall amount. This optimal mesh size would have been different if detachment by overland flow was implemented as the topography impacts the water heights and velocities.

222 The erosion source locations should be shown in Figure 1 in addition to the subplots shown in later figures.

We revised figure 1 so that it now shows the erosion zones clearly.

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).

Thank you for the remark. We remind that the erosion zones were previously defined in the sediment fingerprinting studies by Legout et al. (2013) and Uber et al. (2019) for the Galabre and Claduègne catchment respectively. In line 126 we noted that in the Galabre catchment the land use classes other than the badlands are " permanently covered by vegetation and are thus assumed to be negligible as sediment sources". In the Claduègne catchment diffuse, agricultural sources are considered. We stress that difference between the two catchments following your comment on line 129 by adding the sentence "The distribution of the erosion zones differs in the two catchments, with the ones in the

Galabre catchment being more dispersed over the entire catchment but smaller in size due to the absence of diffuse agricultural sources”. Also, we hope that with changes made to figure 1 it is now easier to see the extent and location of the erosion zones.

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.

We rephrased the two sentences ending in line 235: ”SSYs,ev is the contribution of source s to SSYev and was calculated based on the mean source contributions. They were estimated with sediment fingerprinting in the Claduègne catchment by Uber et al., 2019 and in the Galabre catchment by Legout et al., 2013.” We hope that in this way it gets clear that the reference Uber et al., 2019 refers only to the sediment fingerprinting in the Claduègne catchment. The rest of the section explains the calculations made for this study.

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

Thank you for pointing it out. We state the unit in line 230 where the formula is given now.

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.

Based on your comment above we move the description of the hyetograph (rainfall forcing) here. Further we added “The simulated time is 24 h, including 12 h of rain and 12 h for the fluxes to reach the outlet” (line 295) to be precise about the model run duration here. The description of the model (section 3.2) was revised thoroughly and now states the method of time discretization: “The solver is explicit in time, meaning that the maximum time step that can be used to evolve the equations in time is limited by the Courant-Friedrichs-Lewy (CFL) condition (Courant et al. 1967). This implies that the time step in typical applications is of the order of one second or less. The CFL condition is implemented in the solver and thus, the computational time step is automatically evaluated from the grid size, water velocity and water depth”

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?

We hope that this gets evident after our general answer at the beginning of this document and the changes we made following your earlier comments in narrative form:
In line 278 it now says: “It should be stressed that this source classification does not influence model physics, i.e. total sediment yield from a source (close + distant sources) remains the same as in the basic scenario where they are not differentiated.” Further in line 225, it now says “Sediment production (Eq. 2) was calculated and routed to the outlet in each mesh element for each source class separately. Thus, separate sedigraphs for each source class were obtained at the outlet and the

contribution of each source class to total sediment flux could be calculated for every time step." Eq. 2 was also changed to be more explicit that it was solved for each class separately.

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.

We stated that in the new Study design section you proposed earlier and repeat it in line 256: "We tested the impact of varying the CDA threshold on the modeled hydro-sedimentary response while keeping all other parameters unchanged compared to the basic scenario (one-factor-at-a-time sensitivity analysis).

268 It is not clear to me how 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.

There are no changes in the parameterization of the model. We hope our general answer and specific response to your comment on line 278 allow a better understanding about the aims of Sc. 4. For example, in the Claduegne catchment, the difference is that instead of having three source classes in the basic scenario (badland, basaltic, sedimentary) in scenario 4b and 4d there are 6 source classes (badlands-close, badlands-distant, basaltic-close, basaltic-distant, sedimentary-close, sedimentary-distant). This is visualized in figures 10 and 11. We hope the changes made as explained in our response to your narrative comments and the one on line 278 make this source classification and its implication easier to understand.

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.

Thank you for pointing that out. We added the sentences "Besides the values for Manning's n used in the basic scenario, in Sc. 4c and 4d we used values for Manning's n that were less contrasted between the hillslopes and the river network. This was done to assess whether the interpretation of Sc. 4a and 4b (i.e. the discussion on how the location of the sources in terms of their distance to stream or outlet, impacts the temporal dynamics of SS fluxes at the outlet) depended on the values of n." after the description of Sc.4.

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.

Following this comment, we made sure that for every Scenario a sentence like this explain why this scenario was created. For scenario 2 we added line 258: "As different values for Manning's n were applied in the river network units on one hand and in the hillslopes and badlands units on the other hand, the travel times of the sediments from source to sink vary depending on the length of the river network in the model. Thus, it can be assumed that modeled sediment dynamics are sensitive to this parameter.". For scenario 3 we think the explanation is already in the text: "As one of the objectives of this study is to assess the impact of choices made during model set-up on the simulated sediment

flux dynamics, the model was run with different values of Manning's n in the river network units on one hand and in the hillslopes and badlands units on the other hand".
In the study design section, it is now stated: "The underlying hypothesis is that both modeling choices (notably CDA threshold and Manning's n) and catchment characteristics (structural connectivity of the sources) determine travel times from the sources to the outlet. With the presented study design, it could be assessed whether modeling choices or actual catchment configurations were more important in generating output variability".

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 they are appropriate given the overall goals of the study. This section should be expanded to include this information.

Thanks for pointing that out. We added the sentence "We use these metrics to quantitatively assess differences in model output between the scenarios described above." At the end of this section.

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).

We moved this information in section 3.3 as you proposed.

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

We changed "the model was sensitive" to "modeled hydrographs and sedigraphs were sensitive".

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.

We hope that the introduction of the modeling scenarios with the new section presenting the study design allow to better understand the interest of studying two catchments. Particularly the following sentence was added to: "With the presented study design, it could be assessed whether modeling choices or actual catchment configurations were more important in generating output variability."

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

We do not have found any relevant study to cite for this purpose. However, analyzing all the characteristics of both catchments leads to a clear contrast of their slopes. Whatever the compartments (hillslopes, intermittent streams and main stream) the slopes are on average two to three times higher in the Galabre than in the Claduègne catchments, leading to modelled hydrological response times smaller in the Galabre than in the Claduègne catchment in accordance with measurements.

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.

The emphasis here is on “in this range”. We rephrased it so that it becomes more evident: “Overall, our results showed that the thresholds of 15, 35 and 50 ha produced very similar results. Thus, in this range, the model was not very sensitive to the CDA threshold.”

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.

Thank you for pointing out that the paragraph was not clear, we rephrased it: “This result showed that it is important to use a CDA threshold that is in the right order of magnitude. Field observations or detailed maps (i.e. topographic map at scale 1:25000) can be valuable sources of this information. The sensitivity of model output to variations of the CDA threshold was also observed by other authors (Pradhanang and Briggs, 2014). For our modeling set-up it is reassuring that model results converged in the “right” order of magnitude that can be expected from field observations.”

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.

It is true that generally changing n on the hillslopes has a larger impact than changing n in the river network. But this might not be true universally. Thus, we prefer to keep the formulation as it is (“Interestingly, in the Claduègne catchment liquid discharge was more sensitive to changes in $n_{hillsl.}$ than to n_{river} while solid discharge was more sensitive to n_{river}. This was not the case in the Galabre where both liquid and solid discharges were more sensitive to $n_{hillsl.}$”, line 360). Actually, changing n on the hillslopes had less impact on the sedigraphs than what could be expected. We discuss that in the paragraph “Our results showed that even though modeled liquid discharges were sensitive to $n_{hillsl.}$, the sedigraphs of the main sources and thus of total suspended solid discharge were much less sensitive to this parameter (Figure 8). This was due to the fact that in both catchments the main sediment sources were located close to the river (Table 1, Figure 2). Thus, only a small fraction of the trajectory of particles was located on the hillslopes.” (lines 378-381).

372 What is meant by “more stable”?

We added “more stable in time” to be more precise.

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.

We added the percent change with respect to the basic scenario as a quantitative measure of sensitivity: “Our results showed that even though modeled liquid discharges were sensitive to $n_{hillsl.}$,(e.g. maximum liquid discharge changed by 24% in the Claduègne catchment and 12% in the Galabre catchment), the sedigraphs of the main sources and thus of total suspended solid discharge were much less sensitive to this parameter (maximum solid discharge changed by 3% in the Claduègne catchment and by 1% in the Galabre catchment , Figure 8)”

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.

We hope that the clarification made on objective 2 help the reader to better anticipate what is compared and discussed in this section.

393 Introduce this idea in the study design.

As you recommended, we announced the comparison of the two catchments in the study design section : “With the presented study design, it could be assessed whether modeling choices or actual catchment configurations were more important in generating output variability.”

397–399 This has already been stated.

Thank you for pointing that out. We propose to delete the sentence “The rising limb of the hydrograph was also steeper in the Galabre than in the Claduègne catchment (shorter Tlag and Tc, Figure 5, Table 3).” The second sentence we prefer to keep however. The steeper slopes of the Galabre catchment are assumed to be the reason for several findings: the faster reaction of the catchment, the steeper hydrograph and sedigraph, the lower sensitivity to n in the river.

402 Add a figure reference.

The figure reference is given 3 lines above: “From Figures 7 and 9 a general pattern of the contribution of the different geological sources to total solid discharge can be derived: In the Claduègne catchment […]” To make it more evident that this paragraph refers to figures 7 and 9 we propose to replace the full stop with a colon in line 400.

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

We added a sentence: “In the Galabre catchment at the onset of the event (“1”), suspended sediment originated almost entirely from the black marls, i.e. the source closest to the outlet.”

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 it is nearly impossible to understand the results of Sc. 4.

Thank you for pointing out that the description of Sc. 4 was insufficient to understand it from an external perspective. We hope that this gets clearer after the changes we made in the methods section according to your comments above. But as it seems to be an important point, we added a further explanation on how results were obtained here: “To further address the respective roles of the distance to the outlet and the distance to the stream on the pattern of source contributions to total solid discharge throughout events, the geological sources were subdivided based on these measures in the scenarios 4a to 4b (Table 2). In this way, model output consisted of separate sedigraphs for the close and distant subsources of a given source class. The sum of these sedigraphs is the same as the sedigraph of that source class in the basic scenario.”

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.

We expanded the paragraph by giving a short description of the interpretations of Q-SSC flux hysteresis: "Figures 10 and 11 showed for the Galabre catchment that the limestone sources that were close to the river and the ones that were close to the outlet exhibited a clockwise discharge-sediment flux hysteresis pattern while the distant ones exhibited an anticlockwise pattern. These results confirmed typical interpretations of hysteresis loops, i.e. the assumption that clockwise loops indicate a dominance of close sources because maximum sediment flux occurs before peak discharge while anticlockwise hysteresis patterns indicate a dominance of more distant sources (Bača, 2008; Misset et al., 2019). The results further and highlighted that the sedigraphs of the different sediment sources were strongly related to their location in the catchments and their structural connectivity."

431 Not sure what is meant by this sentence.

We rephrased the sentence which now says "Thus, the mean distance to the outlet was not sufficient to determine travel times of the sources to the outlet."

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

The latter is the case. The sentence was rephrased accordingly: "When the results were analyzed in terms of the distance to the outlet, it was remarkable that […]"

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.

We understood that this comment is related to the comment above that the reviewer wished to have a more explicit statement of the method used to "correlate" basin metrics to the metrics of the sedigraph. But, as stated earlier, we wish to refrain using statistical terms such as correlation or rank analysis for the comparison of only 5 data points.

465 This sentence is not clear.

Thank you for pointing that out. The idea behind this sentence is explained in the following sentences so we deleted this unclear sentence.

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.

Indeed we cannot prove this statement with quantitative metrics of sensitivity. Nonetheless, we think that the analysis is justified. We did not vary locations of the sources but we compared different sources with different locations. Concerning erodibility, it is true that we don't report on how changes made in the erodibility coefficient effect model output. This is due to the fact that erodibility is linearly related to detachment rates in our model. Thus, changing the value of alpha changes absolute values but not the temporal dynamics of sediment fluxes. We stressed that following your earlier comments by adding "While other factors that were not considered here (erodibility, rainfall intensity) crucially influence absolute values of erosion and suspended sediment concentration, their values are less important to determine arrival times and temporal dynamics of source contributions" in the new study design section.

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

Thank you again for noticing that this point was not clear. We hope that the changes we made in the results section (former L344-350) make it easier to follow this conclusion.

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.

We changed "source soils" to "sources" here. We also revised the description of what was considered a source in section 3.3: "[…] the potential erosion zones. The latter include all the mesh elements in the modeling unit "badland" and the mesh elements on the hillslopes that belonged to the diffuse agricultural sources in the Claduègne catchment". Furthermore, figure 1 now shows clearly what was considered as a source in the two catchments (Badlands in the Galabre catchment, Badlands as well as cultivated soils in the Claduègne catchment).

Most 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.

Whenever this was possible we used consistent x and y limits. However, whenever two erosion zones were compared, it was not possible because then the dynamics in the graphs of the less erosive zone would not be visible because of the very different erodibility of the sources (e.g. the y-axis of fig. 6). Furthermore, we focus on temporal dynamics and not on absolute values in this study. Thus, we did not state this explicitly in the figure legends.

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.

As noted above, we prefer to keep figure 1 simple to stress the most important information on the location of the erosion sources and wish to keep the panel (f) in these figures where the focus is on the distance to the outlet and distance to the stream metrics.

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

We revised the column "Aim" in table 2 to better relate this table to the 2 objectives of the study. In the text we better explained why two sets of values for n were used in Sc. 4 following your comment above.

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

As the classification of the sources was different in Sc. 4 than in the other scenarios we would have to give all 3 metrics (T_lag, T_c, T_spr) for each one of 31 subsources so this would add nearly 100 lines to the table which is already quite long.

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

We prefer to keep the classic table layout without vertical lines.

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.

We changed the text color to white so that it is easier to read the text on the darkest blue shades.

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

Following your comment, we explained this in the caption of the table: "NA values indicate that the hydrograph or sedigraph did not recede to 0.1 Qmax within the simulated time."