Response to the Anonymous Referee #2 – Epple et al. In Review at SOIL

The authors want to thank the anonymous author for his/her comments on the manuscript. In the following, we would like to put forward adaptions and corrections of our manuscript in response to the referee's feedback and various comments:

R = Referee's comment

A = Authors' response

R: The hypothesis that model cannot keep up with the data is not justify in the introduction. Rather 15 lines of the introduction (out of 55) are about climate change, which is out-of the scope of the article, and there is no real demonstration that their hypothesis is plausible.

A: Thank you very much for pointing out to us, that the introduction shows a misbalance. In reaction to this criticism, we would like to take the focus away from climate change and further elaborate on our hypothesis. Nevertheless, climate change has an important influence on soil erosion worldwide, which makes the model development all the more important (Guo et al. 2019). Therefore, we would like to keep a few sentences regarding climate change and its influences on soil erosion in our introduction. We propose the following changes:

"Climate change has resulted in the increase of frequency and magnitude of weather extremes, leading to spatially differentiated changes of soil erosion (Nearing et al., 2005; Routschek et al., 2014; Li and Fang, 2016). Alongside its direct impact, climate change also triggers indirect drivers of soil erosion, such as crop management or land use changes (Li and Fang, 2016), which vary greatly from region to region (Klik and Eitzinger, 2010; Nunes et al., 2013; Hu et al., 2020b). In this context model development is necessary to improve the simulation of soil erosion processes (Guo et al., 2019)."

We would like to edit and rephrase the introduction regarding our hypothesis as follows:

"Increasing computing power is constantly pushing the limits of data collection and processing and new opportunities to assess and use soil data and understand soil erosion processes become available. The continuous development and improvement of measurement techniques leads to spatially and temporally highly resolved information at different scales (Li et al., 2017).

This raises the question which of the formerly limiting factors, both in terms of model capabilities and data availability may have become obsolete. Can new and modernized measurement techniques, novel data on soil and soil erosion, a new range on resolutions or new achievements regarding computing power help to further develop these models adequately? Most process-based models demand measured soil data (e.g. texture, water retention function, hydraulic conductivity or aggregate stability) most of which are not available in a high resolution. However, recent data assessment techniques especially regarding photogrammetric methods offer new data on soil and soil erosion, and with this a broad range of model development as well as model ensemble strategies. Furthermore, we regard cross-process as well as cross-scale understanding of soil erosion on different spatial and temporal scales as a vital step towards improving process-based soil erosion modelling. To fully understand and achieve an impact reduction by applying adapted management strategies, models need to integrate the current understanding of soil erosion processes from splash to gully erosion (Parsons, 2019; Li and Fang, 2016). For future model development we find it crucial to determine whether the present assessment techniques are contributing to the improvement of process-based soil erosion models. Most models do *not include high resolution soil or surface input data and work only on specific scales, which begs the question whether models can keep up with the data."*

R: On the contrary, the limiting factors for soil erosion modelling that are cited in the introduction (based on "many" reviews) point to problems linked to poor data availability. The hypothesis I would formulate reading the introduction is "can the data keep up with the models?"

A: Both are true. Most process-based models demand measured soil data which in are often not available in a high resolution. However, new data assessment techniques especially regarding photogrammetric methods offer a broad range of model development as well as model ensemble strategies. Nevertheless, while the assessment of soil data is constantly improving, offering new data and new resolutions, process-based soil erosion models are limited by their unchanged and partly outdated structure. In the process of restructuring our introduction, we would like to remove the limiting factors and make use of these aspects later on in chapter 2.1.

R: The authors do not really describe what they consider to be an "expert based" soil erosion model.

A: The referee's statement is not entirely clear. As we have never mentioned, "expert based" models, we do not see a relevance in describing them. Nevertheless, we would like to add a few sentences further explaining process-based soil erosion models.

"Process-based soil erosion models are more complex than empirical models when it comes to their input data, computing requirement, calibration necessity, and how user-friendly they are. However, due to physically based descriptions of soil erosion and sediment transport, they offer an understanding and reproducibility of the occurring processes (Hajigholizadeh et al., 2018). Such models allow for an separate consideration of individual components of soil erosion processes as well as a better understanding of the relationship between cause and impact within soil erosion research (Scherer, 2008). While these models are process-oriented, they still contain empirical parts based on laboratory and site-specific field experiments. Even though they can be extrapolated to other scales, this aspect must still be treated with care (Parsons, 2019)."

R: They give a list of soil erosion models that are supposed to be expert based. However, some of them are based on the USLE type soil erosion equation. Are these expert based? This list of 44 soil erosion models is very confusing. Why listing so many models and still not be exhaustive? What is the reasoning that led to the elaboration of this list? What is the use of this, almost three-page table, in the demonstration of the authors?

A: Our goal regarding the table was to show limitations most of the currently used models have in common. Based on these information we want to present solutions created through new assessment techniques to overcome some of the models' shortcomings. After receiving feedback from the two referees, we recognize that the table might be a bit confusing and that there is altogether too much focus on the table itself. Assuming that the referee is referring to process-based models, we would like to revise the table according to the comments of both referees (corrections, reducing the columns and double checking every model and their right to be included) and to relocate it to the supplement. Instead of having both the opportunities and the limitations in the table, we have decided to solely focus on the opportunities in order to make the information more accessible. In conclusion, we would assume that the limitations can be inferred from the former. The limiting factors common in many models will instead be listed in chapter 2.1:

"Next to their capabilities, the authors also list challenges and shortcomings of the models. The following difficulties, present a selection which can be found in many models and are sorted by how frequently they are mentioned:

- *High data demand*
- *Process description of gully erosion*
- *Process description of rill initiation (and rill structure)*
- *Sediment connectivity*
- *High computing demand*
- *Assumption of homogenous input parameter*
- *Site-specific calibration*
- *Temporal unchanging soil and surface input parameters*
- *The risk of equifinality"*

R: The different paragraphs that compose the article are also like a list of items that are pertinent to soil erosion modelling; however, they are only very superficially treated. For example there is a paragraph about tracing, that cite, sometimes very rapidly, some tracing techniques, again without being exhaustive, and I don't see how this paragraph contributes to an overall objective of the paper.

A: The exemplary tracing techniques discussed give insight to the area in which such approaches could be useful. We find this information of interest as it provides an initial understanding and offers information which we return to later in the third chapter. We would like to revise this section to better connect the different statements. Moreover, we will submit our manuscript for extensive proofreading in order to better connect paragraphs and statements and link them more strongly to our hypothesis.

R: The same for the paragraph on remote sensing that mostly deals with topography measurements, omitting many innovations in digital soil mapping.

A: We would like to expand on our description of the opportunities remote sensing offers today. While aspects of digital soil mapping have been mentioned briefly, we would like to further elaborate and discuss them. Furthermore, we aim to summarize our results in a more concise manner and restructure the paragraphs to provide a clearer understanding.

R: The statements of the authors are often not rigorous. For example, one of the only reference to illustrate that model cannot cope with the data is Thomsen et al (2015). The authors say that Thomson et al. point out that the possibilities offered by SfM and TLS exceed the integration possibilities of the LISEM Model. First, this is not the objective of the article, and second, the experiments are carried out on 1m² plots. LISEM was never meant to run on 1 m² plots. What would TLS and SF would bring on a 100 ha catchment? The authors should have instead reported the discussion in the paper of Thomsen that go deeper in the issue of the parameterization of random roughness, which have implications in process representation of soil erosion models.

A: With this statement, the referee brings up a valuable discussion. While they point out correctly, that LISEM was never intended to run on 1 m², the model, being physically based, claims transferability to other scales. Consequently, the model should ideally be transferable to a small scale, such as a 1m² plot. If models solely work on a large scale, the small but no less important soil erosion processes are neglected. This brings us back to the hypothesis. While the data is ready to work on a small scale, the same cannot be expected from the models. New measurement techniques allow the monitoring of processes, which until then could not be observed or "only" as a final result (e.g. headcut retreat, filling small ponds or aggregates disintegration). Such data can then be used to further develop process-based soil erosion models. However, we would like to revisit this section as well and state our discoveries and arguments more clearly including the discussion above.

R: The authors talk about data assessment on large scale for the reference Eltner et al. 2018, which is about experiments on plots from 600 to 2400 m².

A: We would like to change this to "larger scales", as we understand the criticism on the definition of "large scale". While we talk about scales from micro plot to catchments, we believe a field can already been seen as a larger scale in comparison to the micro plot.

R: The authors also never cite the research on upscaling methods, which permit to integrate small-scale features for large scale modelling.

A: While we see a lot of potential in upscaling methods and researchers have already discussed the challenges accompanied by upscaling, the method itself has not yet been established. Such cross-scale assessment and integrations of soil erosion could prove valuable for improved process understanding. This brings us back to the question whether models can keep up with the data. We would like to include further thoughts on this topic in our manuscript.

R: On the newly available data, the authors should distinguished, between the techniques that brings new information (that was not available before) and the ones that bring higher resolutions, as the implications are different, and should be treated separately.

A: Thank you for the advice. However, we believe that separating these aspects so distinctly could be problematic. Even though some techniques already exist, they become a valuable asset for soil erosion modeling due to higher resolution. Furthermore we presented examples in our manuscript that higher resolution actually offers new information (e.g. on erosional processes). Separating these "two" aspects does not seem right. However, we would like to clearly distinguish between the two main directions assessment techniques offer us today:

i) On the one side methods which can be used for a new process understanding. A process understanding researchers were not yet able to observe in such a high spatial and temporal resolution across scale.

ii) On the other side, methods enabling new data assessment, which can be used to feed models.

R: And then the questions that is interesting, is "what is the detail representation in my model that is optimum to achieve a good model performance. This question is valid both in terms of the nature of process representation (how many parameters should I integrate) and it's resolution.

A: This is an interesting and valuable research question. However, it is a question, which cannot be answered by a review, and is therefore not answered in our manuscript. The choice for the right resolution, as well as the number of parameters is dependent on the research question. Studies on a large catchment will be based on a different research question and use a different resolution than studies in a sub-catchment. We would like to add these aspects to our manuscript in chapter 3.3:

"The choice for the right scale as well as the right resolution is still dependent on the research question. Both small and large scales combined with different resolutions can bring out some processes while masking others. Even if a model is capable of handling the whole variety of scale and resolution, the appropriate choice must be made by the modeller themselves."

Furthermore, we would like to deepen our discussion on scale. Different models represent different processes that again are working on specific scales. Until recently, the process understanding was also dependent on scale, e.g. measuring splash erosion by splash cup experiments. Whereas today we have spatial and temporal high resolution data enabling the identification and mapping of processes in scales where they have not been visible before. Small-scale processes have been neglected in larger scale modelling even though these processes have an important influence on the modelling results. An example is again the splash erosion, which is measured by splash cups but not on scales dominated by interrill flow. Aggregate breakdown, preparing the soil surface, also has an important role for larger scales rather than the micro plot only. With the possibilities of today we can detect processes in a high resolution, which due to assessment techniques used to be only visible in different scales. This could lead to an overall better process understanding and consequently help improve the process based soil erosion modelling.