The sensitivity of landscape evolution models to spatial and temporal rainfall resolution:

Reviewer 1 Comments

This paper deals with a very interesting and relevant question for the scientific community working on sediment transfers in mesoscale river basins: how do the spatial and temporal resolution of the meteorological forcing impact modelled sediment yields? While this issue has already been addressed from a purely hydrological standpoint, it remains understudied in modelling approach dealing with landscape evolution and soil erosion. However it seems to me that the conclusions raised by the authors are not supported by enough simulations. My main concern is about the potential effect of changes in soil hydrological properties (spatially and temporally) as the spatiotemporal resolution of rainfall is changed. This is not at all considered by the authors in their simulations while they recognise at the end of the discussion that it may change considerably the sensitivity of landscape evolution models to rainfall resolution. As hydrological properties might be scale-dependant, changing only the spatiotemporal resolution of rainfall between runs without considering potential scale interactions between rainfall and soil behaviour may lead to erroneous conclusions on the sensitivity of landscape models. I know that adding runs in which the soil properties are randomly changed (m and K parameters) will need considerable additional computation time but the conclusions of the paper would be more supported and strengthened.

We would like to thank the reviewer for their comments and thorough review. Aside from typo's and other minor points/clarifications, the main point the reviewer asks us to address is the interaction with soil properties and the balance between precipitation (P) and infiltration (I).

We agree completely with the reviewer that soil and land use properties might influence our results. However, the focus of this study is to examine just the impact of spatial and temporal rainfall resolution. In our parameterisation, hydrological factors that will change spatially are deliberately treated globally so we can look solely at the role of rainfall resolution. The experimental set up (e.g. having different hydrological areas defined by the rainfall grid resolution) is contingent upon the deliberately limited research questions we are asking – and to look at both soil properties and rainfall resolution would, we suggest, require a completely different model set up.

We believe that it is important to consider that basin hydrology – both in terms of soil properties – and the driving precipitation – is often dealt with incredibly simplistically in LEM's, if at all! Therefore, our motivation is to explore not just the sensitivity to resolution – but to show the difference between having no representation and some representation of a distributed hydrology in LEM's.

Since this paper was first submitted – we have also submitted (and now published in early view) an article that takes a tightly constrained look at the impact of spatial changes in the TOPMODEL m value on the geomorphic outputs over longer time scales (Coulthard & Van De Wiel, 2016)*. There is certainly a place for a study looking at both together. This could be looking at a combination of the two approaches opens up the CAESAR-Lisflood model to a framework of modelling using Hydrological Response Units (HRU), a common approach in semi-distributed hydrological models, such as Dynamic-TOPMODEL and SWAT. This allows rainfall, land cover and soil properties to be represented at higher resolution than a global lumped estimate, but divided into broadly hydrologically homogenous regions.*

Whilst we have not carried out any additional research to answer the points raised by the reviewer above and below, we certainly accept their validity – and have added a section to the discussion/limitations section

Coulthard, T. J., & Van De Wiel, M. J. (2016). Modelling long term basin scale sediment connectivity, driven by spatial land use changes. Geomorphology. http://doi.org/10.1016/j.geomorph.2016.05.027

Concerning the structure of the manuscript, the result section is very short and could be expanded, particularly if additional simulations are presented. The discussion section is rather heterogeneous in answering the 3 research questions written in the introduction. Section 4.2, addressing question 2, is very short and does not fully address it, as the authors recognise that more simulations would be required. Also Section 4.3, addressing question 3, is not supported by the data (no reference to them). I would suggest focusing the results and the discussion on question 1.

We have changed the research questions – and dropped number 2.

If this question is fully addressed according to the above mentioned issues dealing with the hydrological basin properties, this would represent a substantial contribution to earth surface mass transfers. For those reasons, I do not recommend acceptance of the manuscript in its present form.

Specific remarks:

P2 L31-33: "Improved model performance" is not only "tempered by increased uncertainty surrounding precipitation data", but by the uncertainty in the budget of precipitation (P) versus infiltration (I) or storage in the soil.

P3 L8-9: I fully agree with the authors here. This also refers to the ability to simulate correctly P-I budgets, spatially and temporally. As argued in the general comments, the authors should try to address that issue in their numerical sensitivity analysis.

See first comment. There is of course a need to look at the sensitivity of models to both spatial and temporal changes in precipitation AND land use – but in this paper we have focused on just one. This is to (a) make the experimental design simpler and (b) because spatial changes in land cover is really a different research/science question that we have answered in a subsequent paper.

P4 L30: I found the description of the model spatial discretisation quite confusing (not sure to have completely understood yet), mentioning here "area lumped parameters" and later (P5 L18) "grid cell size Dx" without giving any typical size for Dx. Is it the DEM resolution (50m) mentioned

Changed to say 50m.

P7 L31? Could the authors try to be more specific on spatial discretisation and if possible limit the reference to previous papers to very specific model details that are not essential for that study?

There are no references in this paragraph, the three before or the two after - think this relates to the overall model description.

P6 L1-3: As far as I understand, the hydrological model is adapted to the rainfall grid. Thus I agree with the authors that it enables having different levels of storage and runoff in each cell, but only due to rainfall variations. Varying also m and K would also create variations (i.e. P-I budget), but the authors kept constant those parameters (P8 L3).

This is true - but the cells do not cover hydrologically homogenous areas and values of m and k would be difficult to determine. m and k are not scale dependent, therefore the use of global values here is justified, yet we acknowledge that adjusting these values to local conditions will change the outputs - this is beyond the scope of this study (which is motivated only by rainfall resolution).

P7 L29: How can the authors be confident in their conclusions with only 2 additional long term random simulations?

In each random scenario the rainfall distribution was reshuffled every ten years, producing a significant element of randomness to the full 1000 year record simulated. However, both these two random runs produced similar results to each other. Additional random scenarios could be run, yet this is computationally expensive and would yield similar results. The random runs were motivated by wanting to disrupt the spatial pattern in the ten year record, which is repeated 100 times in the long term runs - in reality, we could have been confident with just one test as this achieved this, but the only very minor variations between the two reinforce this further.

We can see that this may not be clear from the text – so this has been clarified in the methods section. Part of this is trying to explain how methods evolved in response to results – but within the methods section!

P8 L1 : Which initial grainsize distribution was used to run the 30 year model used as an initial condition? Which grain sizes are given in Table 3?

The grainsizes in Table 3 are those used to initialise the model with a global distribution, which was then spun up using a thirty year simulation. The spun-up grain size distribution was used for the tests. Text has been added to make this much clearer.

P8 L16: “considerable differences”: this is not new and references should be given to situate these data.

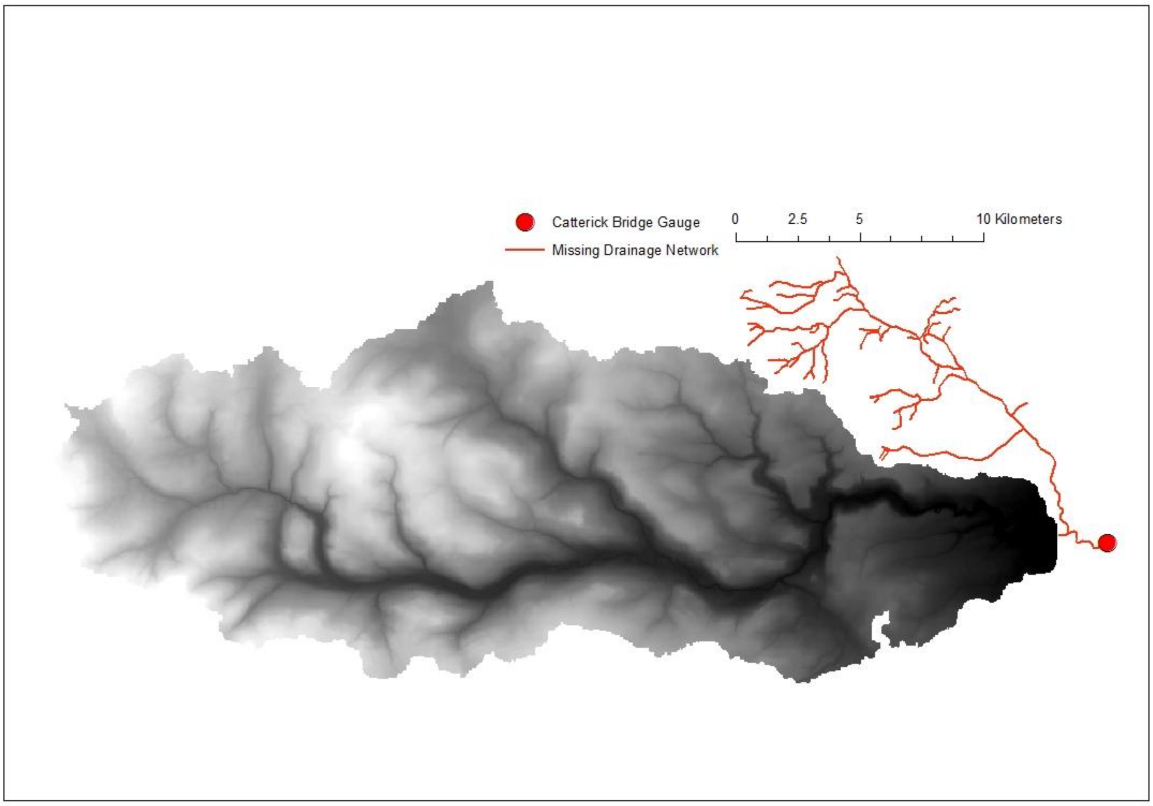
Yes - this well established and long been understood. Line 16 has been altered in the manuscript to make this clearer.

P8 L21-23: I agree that these changes are minor. Could they be significantly different if m and K parameters were also randomly changed from one run to another?

Possibly, but as previously mentioned that is not the aim of this study.

P8 L28: “also drains an additional tributary”. This may be critical. What is the drainage area of this tributary? How are the authors confident with that comparison between model and measurements?

The purpose of the hydrology tests was not to assess the model's performance and skill at estimating the hydrological conditions, but to demonstrate the effects on the hydrology when driving the model with different resolutions of input data. We believe that the data presented do this effectively.



The image above shows the full extent of the Swale catchment we have used, plus the location of the Catterick Bridge gauge station and the missing drainage network. The system is predominantly Gilling Beck , which becomes Skeeby Beck before entering the River Swale. The information for the Catterick Bridge Gauge station (http://nrfa.ceh.ac.uk/data/station/info/27090) states it has a draingage area of 499.4km^2 - of which, the Complete Swale DEM is 415km^2 - so the missing area covers 17% of the drainage area. It's overall contribution to the discharge of the River Swale is not known.

P9 L11 : Very little difference is observed between random 1 and 2. This relates to a previous comment. Are 2 random simulations enough? Why not having done more, as presented later with the jumbled runs (P10 L9) for answering another question. Could these differences be more important if additional random simulations were performed? Otherwise this result (little difference) contradicts somehow with the results in Figure 7 showing a great dependence of the sediment yield to the rainfall allocation. Overall, it seems to me that the authors tried to address too many questions in the paper without running enough simulations to address each of them.

The jumble runs involve the shuffling of the rainfall distribution just the once. Random 1 and 2 each involve 100 reshuffles of the rainfall distribution and therefore include a higher degree of randomness. They are akin to carrying out 100 10 year re-shuffled simulations. So we were happy with their results – and the simple comparison of Random 1 and 2 showed they were performing correctly. If these runs had involved a single shuffling at the beginning of the 1000 year then we would expect that each run would show much greater difference in the erosion patterns seen.

P10 L8-9: if this issue is so important, it should be introduced as an objective of the paper. As written it appears like an additional side issue. The description of these jumbled runs should be added in the method section and removed from the discussion.

This has been changed. Moved to methods/results and removed from discussion.

P10 L5-10: I fully agree with the authors that it is a major limitation to this study. Thus I recommend the authors to try to assess how m and Ks variations could impact the sensitivity analysis as it will help to generalize their findings.

We believe that Reviewer 1 is referring to Page 12 here. The CAESAR-Lisflood model does not presently allow for a variable K value. This study looks solely at the influence of the rainfall input resolution on the modelled sediment yields and geomorphology of the catchment. By allowing for spatial variation of m and K as suggested, this would add a further variable from the standard approach of using a global value for these parameters. It is our belief that, although it would be an interesting avenue of investigation (indeed, similar work by one of the authors has recently been published in this area), it would only confuse the purpose of the investigation in this manuscript.

P10 L10: Why 20 different records? Does this number has an impact on the range covered in Figure 7 (i.e. from -7 to 2,5% for X-axis and from -15 to 60% for Y-axis)?

Twenty was chose as there are 20 rain cells in this domain -and this showed a clear pattern, as would have with ten. Running for 100 tests might extend the range covered in Figure 7, but not the step difference between the temporal resolutions which was the point of the Jumble tests.

P18 Table 3: Evaporation was set to 0. How does it impact the conclusions? K is missing in this table.

K is not a variable parameter in C-L - Evaporation was set to 0, yet this was constant throughout the tests. It will make a small difference, as would including the vegetation parameters or a bedrock layer making the region transport limited. This needs to be viewed as a conceptual sensitivity test into rainfall resolution alone. Varying other factors is beyond the scope of this study.

P27 Figure 6: Total rainfall : The authors should specify over how many years.

Over the available NIMROD record (ten years).

P28 Figure 7: Why was this analysis done on the upper Swale only? The complete basin is characterised by more rainfall cells and would have probably exhibited more variations in the random redistribution of rainfall (see author's comment in section 4.2, line 5). I find this figure very interesting in addition to the results from Table 8 for example, as it clearly shows the impact of the jumbled runs. This sensitivity of the sediment yield to different spatial and temporal distribution of the rainfall raises again the question : would this sensitivity be the same if also m and K parameters were included in a wider jumbled run numerical analysis.

Computational efficiency was the main reason. This involved 160 tests - each taking roughly 8 hours for the upper Swale, and 2 - 3 days for the Complete Swale. Also consistency - the majority of the additional tests (1000 year, Random 1 and 2) were conducted using the Upper Swale. We have added a note to indicate that long model run times restricted how many simulations we could carry out.

Again, we feel the global values for m and K are justified for this study for reasons stated previously throughout the response. By varying the values in each cell, it would make it difficult to disentangle the influence of the rainfall resolution, and the influence of varying these values.

Technical correction:

P2 L15 : delete reference at the end of the sentence

Done

P3 L19-20 : can not find those three references in the reference list

Thank you – they had been manually added not in the referencing SW!

P3 L28 : Coulthard et al. (2013a)

this reference seems correct ?

P5 L4 : Add units for Qtot P6 L11-13 : n (Manning) is missing in the list

Added.

P7 L29: were then compared

Thank you.

P28 Figure 7: homogenize the colors for 6 hour (yellow and purple)

Changed.