We thank **Christopher Skinner** for his review of our manuscript and making the highly constructive comments and suggestions. We are glad to hear our effort to revise the Manuscript. The author's response is shown below in blue text.

In the second Table 2 (there are two) on page 22, the sediment conservation ability in the source area increases from 0.5 to 0.55 between the PP and EP scenarios, yet I don't think there are any additional interventions in this area (levees and vegetation are in the deposit area). I think this may be an effect of the way the authors have applied the spatially varied "m" in the model. In UP and PP, the "m" has a global value of 0.008 and in EP the vegetated areas are given a separate value of 0.02. It isn't stated by the authors but I believe **their rainfall input is catchment lumped (please could the authors confirm)**. My understanding of CL is that for a lumped input it will average all the "m" values and create a single lumped input from it, in this case making the input for the whole catchment less flashy. Alternatively, the authors could **specify two separate rainfall input areas**, one for the vegetated area and one for the rest of the catchment, in effect making two hydrological response units (HRUs) for the model, each with its own input based on the local "m" value. I don't think this needs to be done for a revised manuscript as I doubt it would change their conclusions materially, but it should at least be acknowledged.

Firstly, we are so sorry about the confusion about the tables' labels and we have corrected them. We agree the increase of the sediment conservation ability between the PP and EP scenarios is an effect of the spatially varied m-values in the model.

1. *the rainfall input*

Actually, the rainfall input in Scenario UP and Scenario PP is catchment lumped. While in Scenario EP, we divided it into two separate but identical rainfall for the regions with different m-values.

Further on "m", I concur with the comments from Jorge that where possible the value should be calibrated against gauged data. If this is not available, basing the value on land cover, as the authors have done, is reasonable. However, the authors are using downscaled hourly rainfall, not observed hourly rainfall, so **any calibration would need to account for this**.

2. *Calibration*

Admittedly, it is essential to calibrate the hydrological components before replicated work. We follow both of the two reviewers' suggestions and calibrate the parameters by replicating the flash flood event in July 2018 using C-L.
There are no huge differences in geomorphology, channel location, and landcover before 2013 and after 2018 in our catchment found from the field surveys. Based on Scenario PP (with two check dams), we changed the rainfall series to the two-week hourly precipitation in July 2018, which is recorded by the rain gauge 2.5 km away from the catchment placed in 2015. The simulation results (**错误!未找到引用源。**c and

错误!未找到引用源。d) showed the erosion and maximum flood depth deposition distributions in Scenario PP on July 15th, 2018. As shown in 错误!未找到引用源。c and 错误!未找到引用源。d, we selected three locations randomly to compare the simulation results with remotely sensed images and photos. The results (Figure 2) showed reliable results including sediment deposition and the peak flood depth, which indicate that the flash flood event was replicated successfully by the C-L.

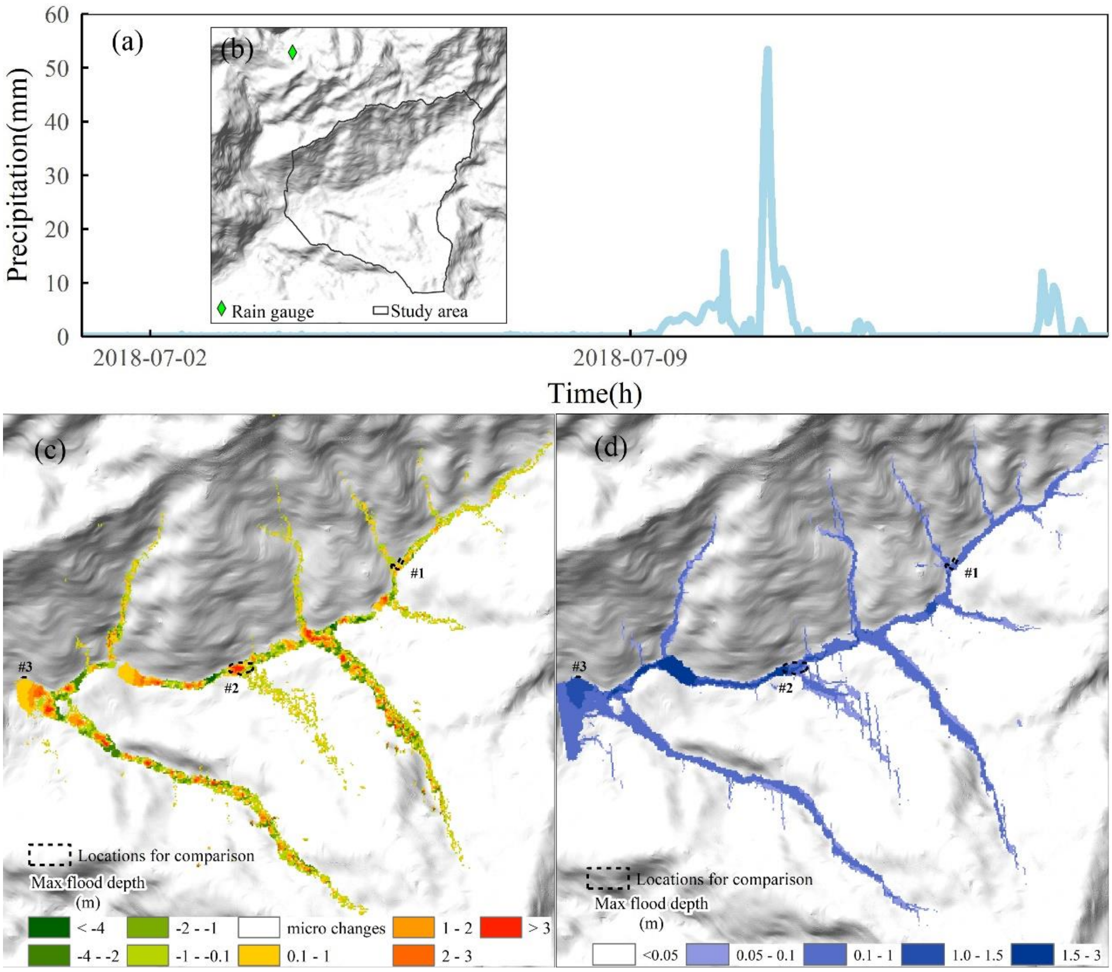

Figure 1. The input rainfall series (a and b) and simulation results of the flash flood event in July 2018 (c and d).

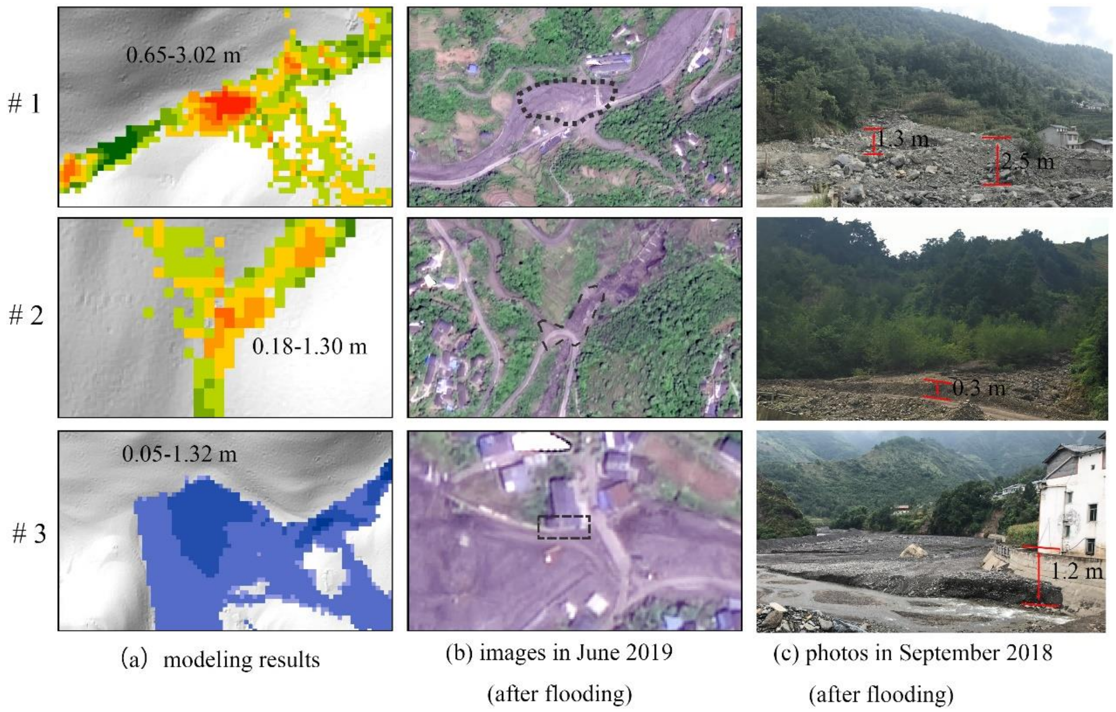

Figure 2. The comparison of the simulation results to images (GF-2 with 8-m resolution) and photos after the flash flood event in July 2018.

I also concur with Jorge's comment on **spin-up period**. There are no details in the manuscript and it would be helpful to know. In this case, where much of the eroded material is fresh and loose, it could be argued a spin-up might actually be a counterproductive in this instance.

3. *Spin-up period*

Admittedly, didn't spin up the model to mix the grain sizes. The purposes of the process are to eliminate the 'walls' and the 'depressions' in the cells and avoid the intense erosion in the hill slope in the early run time. Actually, we preprocess the DEMs by filling the sink based on Environmental Systems Research Institute's (ESRI's) ArcMap (ArcGIS, 10.8) to eliminate the problematic pixels. Moreover, for our catchment, the fine grains were distributed homogeneously both in the hill slope and the channel five years after the strong earthquake. Therefore, we think a huge difference would not exist. However, we will continue to compare the difference in future work.

The downscaling of the rainfall to hourly is really important (as shown nicely in Figure 3). Sorry to push one of my papers, but Coulthard and Skinner (2016: https://doi.org/10.5194/esurf-4-757-2016) provides some analysis of why and it would be useful to refer to this here. Unfortunately, I found the description of how this was done not clear – please could the authors **revisit this description** so it is easier to follow. It would be useful to also know the spatial resolution of the rainfall product used and the spatial resolution it applied to the model with (I assumed it was lumped).

Thanks for the recommendation. The findings from a range of simulations have revealed that using time-averaged climate inputs may be under-predicting basin sediment yields as well as introducing spatial biases through under-predicting or over-predicting erosion (Coulthard and Skinner, 2016), which helps to explain why we downscale the daily rainfall into hourly rainfall. Therefore, we would reference it in our revised manuscript.

4. *Downscaling method description*

The description of the temporal downscaling process was revised as:
In this research, we compared three scenarios using identical precipitation data during 2011 and 2013 as mentioned in section 3.1. The source daily precipitation of one station in 2011-2013 was downloaded from China Meteorological Administration (http://data.cma.cn). The rainfall intensity and the frequency of extreme events affect patterns of erosion and deposition (Coulthard and Skinner, 2016; Coulthard et al., 2012). And we used the stochastic downscaling method to generate hourly data to best capture the hydrological events in this study, which was introduced by (Li et al., 2020; Lee and Jeong, 2014). The referenced hourly precipitation was measured from the pluviometer located 20 km from the study area in 2016, with annual total precipitation of 684 mm. The rainfall in 2016 was characterized by (1) hourly precipitation from 1.1 mm to 35.4 mm and (2) the maximum and average duration of a rainfall event up to 24 h and 2.8 h. The main processes of the downscaling method are as follows.

- extracting the measured daily rainfall closest to the referenced daily rainfall in 2011-2013 through the threshold setting and producing the genetic operators from the extracted hourly rainfall;
- mixing on the genetic operators by genetic algorithm (Goldberg, 1989) composed of reproduction, crossover and mutation and repeating until the distance between the predicted daily rainfall and the measured rainfall is less than the setting threshold;
- normalizing the hourly precipitation to remain the daily rainfall value unchanged.

The input of generated hourly precipitation is catchment lumped in Scenario UP and EP and divided into two separate but identical rainfall in Scenario EP.

A **verification of the model outputs** for the PP scenario by comparing them to real-world observations would strengthen the analysis of the paper. For example, Figure 10 and related discussion could be included within the results as a form of verification for the model outputs.

5. *Verification*

Similar to the reply on the first review, we add the verification both in the discharge and erosion/deposition features. As shown in Figure 3, we compare two types of discharge recorded in published research (Feng et al., 2017; Guo et al., 2018) with those of simulation results to confirm the physical plausibility. As shown in Figure 4, we verify the erosional and deposited features by providing photographic evidence from the observed landscape and compare them to cross-sections from the simulated results.

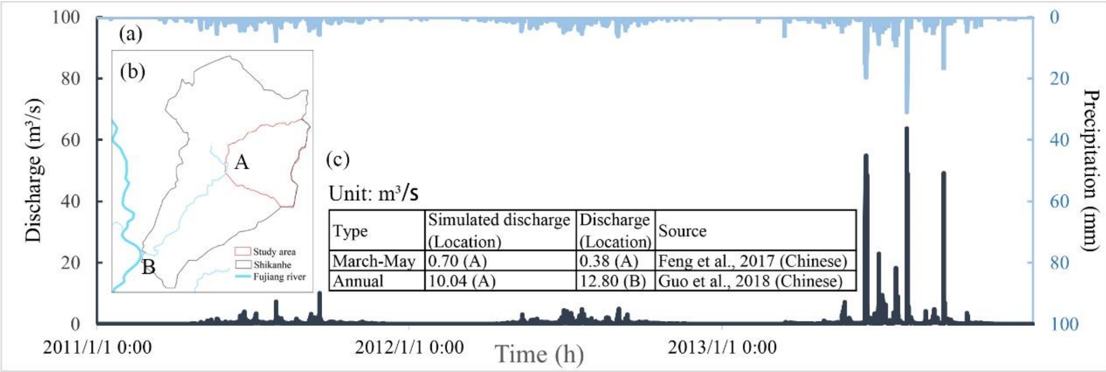

Figure 3. (a) The simulation discharge in 2011-2013 in Scenario PP; (b) the verification location; (c) the comparison of the simulated to the recorded discharge.

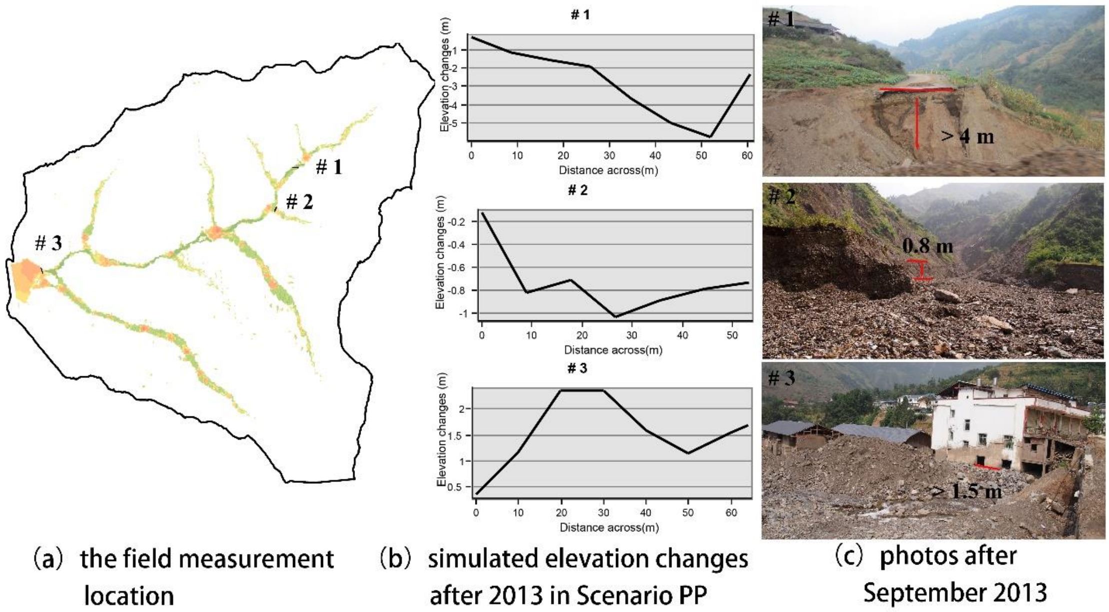

Figure 4. The comparison of cross-sections from the simulation results to the photos in the field measurement locations after 2013 in Scenario PP.

I would recommend that the authors include in the discussion notes on how the outputs of this analysis could be used – ie, why is this work useful. Is the intention that these modelling approaches will be used in the future to design debris-flow management schemes and help to inform decision making, for example?

6. *The discussion of application*

Followed by the recommendations, we enrich the first section in discussion to model uncertainty and application. The application was discussed as follows.

The methods applied in the study further demonstrate the role of C-L as a tool to understand the short-medium term or the long-term geomorphology changes (Ramirez et al., 2022; Li et al., 2020; Coulthard et al., 2012) and observe the effectiveness of natural hazard interventions measures provided different rainfall patterns . For example, the mitigation facilities in this study were effective, especially engineering measures that cooperated with vegetation revetments in the upstream area, which would help decision-makers to optimize the management strategies to control mountain disasters. Geotechnical engineering has its disadvantages event though it is a mature technology that identifies and fixes problems quickly (Peng and Yongming, 2013), such as the greater work and expense and the difficulty of maintenance. While the 'green development', the vegetation cover was effective to prevent erosion by strengthening topsoil and absorbing excess rainwater with its roots (Reichenbach et al., 2014; Stokes et al., 2014; Forbes and Broadhead, 2013; Mickovski et al., 2007). Alternatively, the methods could be used to study the tree planting patterns in different slopes.

On the language in the manuscript, I found the vast majority of the manuscript well written and easy to follow. There were a few instances where phrasing is not quite comfortable and I think some editorial guidance would be sufficient to improve these. Some in-line references contain initials and these should be corrected.

Thanks for the suggestions and we would polish the language and revise the references in the new manuscript.

Reference

Coulthard, T. J. and Skinner, C. J.: The sensitivity of landscape evolution models to spatial and temporal rainfall resolution, Earth Surf. Dyn., 4, 757–771, https://doi.org/10.5194/esurf-4-757-2016, 2016.

Coulthard, T. J., Hancock, G. R., and Lowry, J. B. C.: Modelling soil erosion with a downscaled landscape evolution model, Earth Surf. Process. Landforms, 37, 1046–1055, https://doi.org/10.1002/esp.3226, 2012.

Feng, W., He, S., Liu, Z., Yi, X., and Bai, H.: Features of Debris Flows and Their Engineering Control Effects at Xinping Gully of Pingwu County, J. Eng. Geol., 25, 2017.

Forbes, K. and Broadhead, J.: Forests and landslides: the role of trees and forests in the prevention of landslides and rehabilitation of landslide-affected areas in Asia, FAO, 14–18 pp., 2013.

Goldberg, D. E.: Genetic Algorithms in Search, Optimization and Machine Learning, 1st ed., Addison-Wesley Longman Publishing Co., Inc., USA, 1989.

Guo, Q., Xiao, J., and Guan, X.: The characteristics of debris flow activities and its optimal timing for the control in Shikan River Basin Pingwu Country, Chinese J. Geol. Hazard Control, 29, 2018.

Lee, T. and Jeong, C.: Nonparametric statistical temporal downscaling of daily precipitation to hourly precipitation and implications for climate change scenarios, J. Hydrol., 510, 182–196, https://doi.org/10.1016/j.jhydrol.2013.12.027, 2014.

Li, C., Wang, M., Liu, K., and Coulthard, T. J.: Landscape evolution of the Wenchuan earthquake-stricken area in response to future climate change, J. Hydrol., 590, 125244, https://doi.org/10.1016/j.jhydrol.2020.125244, 2020.

Mickovski, S. B., Bengough, A. G., Bransby, M. F., Davies, M. C. R., Hallett, P. D., and Sonnenberg, R.: Material stiffness, branching pattern and soil matric potential affect the pullout resistance of model root systems, Eur. J. Soil Sci., 58, 1471–1481, https://doi.org/10.1111/j.1365-2389.2007.00953.x, 2007.

Peng, C. and Yongming, L.: Debris-Flow Treatment: The Integration of Botanical and Geotechnical Methods, J. Resour. Ecol., 4, 097–104, https://doi.org/10.5814/j.issn.1674-764x.2013.02.001, 2013.

Ramirez, J. A., Mertin, M., Peleg, N., Horton, P., Skinner, C., Zimmermann, M., and Keiler, M.: Modelling the long-term geomorphic response to check dam failures in an alpine channel with CAESAR-Lisflood, Int. J. Sediment Res., 37, 687–700, https://doi.org/10.1016/j.ijsrc.2022.04.005, 2022.

Reichenbach, P., Busca, C., Mondini, A. C., and Rossi, M.: The Influence of Land Use Change on Landslide Susceptibility Zonation: The Briga Catchment Test Site (Messina, Italy), Environ. Manage., 54, 1372–1384, https://doi.org/10.1007/s00267-014-0357-0, 2014.

Stokes, A., Douglas, G. B., Fourcaud, T., Giadrossich, F., Gillies, C., Hubble, T., Kim, J. H., Loades, K. W., Mao, Z., McIvor, I. R., Mickovski, S. B., Mitchell, S., Osman, N., Phillips, C., Poesen, J., Polster, D., Preti, F., Raymond, P., Rey, F., Schwarz, M., and Walker, L. R.: Ecological mitigation of hillslope instability: Ten key issues facing researchers and practitioners, Plant Soil, 377, 1–23, https://doi.org/10.1007/s11104-014-2044-6, 2014.