

Review 2:

We would like to thank Reviewer 2 for their thoughtful and constructive comments on our manuscript. The reviewer raised several important points that have helped us improve the clarity and rigour of our analysis. In response, we have revised relevant sections to address each suggestion, with particular attention given to the clarification of our fragility index and exposure methodology.

1. 28-30 – incomplete sentence – needs copy edit.

Thank you for your comment and highlighting this oversight in the grammatical structure of our sentence. We agree in its current form has the potential to be read as incomplete, creating an element of confusion. We therefore propose a rewrite of the sentence to generate better clarity in both its meaning and structure in lines 28-30.

2. 107-108 - Please clarify this sentence – it sounds like landslides are being included as a subcategory of debris flows as written. Also – why would debris flows be constrained to the south? I can find lots of documentation about debris flows in northern provinces in China, maybe eliminate.

Thank you for your comment. You're absolutely right that the original sentence could be misinterpreted as suggesting that debris flows are a separate category from landslides and that they are geographically constrained to southern provinces, which is not accurate. Debris flows are a specific type of landslide, and while some post-2008 earthquake data did indicate a concentration of debris flows in southern Sichuan—particularly in steep, high-relief catchments, the broader national pattern shows that both landslides and debris flows occur across many regions in China, including the northern provinces. To better reflect this nuance and avoid overgeneralisation, we have rephrased the sentence in lines 111 to 114. This revision should more accurately reflect the distribution while maintaining clarity.

3. 179 – Clarify “properties”

Thank you for your comment. We agree, the use of the term “properties” was an overgeneralised term here. To remedy this, we offer a sentence rewrite which provides greater detail in explaining the features being mapped in lines 187 to 189.

4. 190 – Commas make this sentence awkward

Thank you for your comment. We agree the commas in this sentence (line 199) create awkwardness when reading the sentence. We have therefore removed the commas to improve readability and grammatical accuracy.

5. 231 – Is this assessed cell by cell?

Thank you for your comment. We can confirm that the exposure assessment was conducted at the individual building level rather than cell-by-cell. All analysis steps, including digitisation of buildings, assignment of fragility values and exposure calculations were performed within a GIS environment and supported in excel. Fragility values were manually assigned based on the spatial relationship between buildings and mapped hazard zones and then validated using literature. We have clarified this in the revised manuscript to improve transparency and replicability in lines 239 to 248.

6. 236 – Why is fragility just assessed as a binary? Would be nice to have a little clarification there.

Thank you for your comment. Fragility was assessed as a binary value primarily due to limitations in available data and the need for robust, replicable classification using remote sensing and literature sources. Given the scale of the study areas and the reliance on satellite imagery, a binary index (0 or 1) was the most reliable way to differentiate between buildings clearly exposed and/or affected by debris flows and those not. Buildings shown to be inundated or damaged, or situated in highly susceptible locations e.g., along channels or gully mouths were assigned a fragility of 1, while others were set to 0. This simplification avoids over interpretation of uncertain data and ensures consistency when validating against historical damage reports. Future work could definitely look to incorporate more nuanced fragility gradations where detailed onsite or structural data is available. We have provided greater detail in lines 248 to 254 for clarification.

7. Figure 3 – What do numbered tributaries correspond to?

Thank you for pointing this out. The numbered tributaries in the three catchments in Figure 3 correspond to the main contributing streams in each gully, identified and labelled based in their position within the hydrological network. We will add a brief note to the figure caption to clarify this in lines 330 to 333.

8. 362-363 - This is a little confusing as per earlier - isn't exposure just being assessed by whether a structure is in the path of a debris flow or not?

Thank you for your comment. You are correct that exposure in our model is primarily based on whether a building is located within the path of the debris flow, reflecting direct hazard impact. However, we also recognise that building quality and structural characteristics, particularly in our unmodified catchment Xiaojia, hold significant influence in terms of overall vulnerability and damage outcomes. Due to our limited access to detailed records on building construction and materials, our current vulnerability assessment simplifies fragility to a binary classification as previously discussed based upon observed damage and location relative to debris flow channels. We have added further clarification to the manuscript (lines 387 to 392) acknowledging that whilst exposure is defined spatially, building quality is an important factor influencing susceptibility and could therefore be incorporated into future studies with greater access to detailed structural data records.

9. **Figure 5 – In Cutou Gully, the shading of the tributary channel with the pale blue colour only implies that it wasn't there 2005-2010? Also, some of the roads (dashed purple, solid red) lack symbols in the legend. Not a huge deal though. Also, were any of the areas shaded as built environment in 2005-2010 subject to debris flows? It's a little hard to parse whether the changes in built environment from 2011-2019 represent just additions to pre-existing infrastructure or replacement to areas that were damaged. More importantly, the spatial scale is not correct. The small-scale bar seems to represent 5 km (it's pre>y small, I may be misreading) but checking out the field sites on Google Earth they're an order of magnitude smaller than that.**

We appreciate the reviewer's thorough comments and have revised Figure 5 accordingly to improve clarity and accuracy (lines 405 to 415).

- The pale blue shading of the tributary channel represents the mapped extent of the channel as observed in the 2011–2019 period. We have updated the figure caption to clarify that this feature reflects changes post-2008, not necessarily the absence of the channel in 2005–2010, and does not imply the channel did not exist earlier only that the resolution of the satellite images from this time period made it difficult to map the channel.
- The legend has been amended to include all road symbology, including the previously missing dashed purple and solid red lines, to ensure consistency and improve interpretability.
- Regarding areas mapped as built environment from 2005–2010, we believe that several of these were damaged during debris flow events, particularly in zones near the gully mouth. However, due to limitations in the resolution and frequency of available satellite imagery, it was not always possible to distinguish between post-event reconstruction of existing structures and the addition of entirely new infrastructure. This limitation is now explicitly noted in the updated figure caption.
- Finally, the spatial scale bar was incorrectly labelled. The figure has been corrected, and the scale bars now accurately reflects the true extent of the study areas. The caption has also been revised to remove any ambiguity regarding scale.

We trust these changes address the reviewer's concerns and improve the overall clarity and utility of Figure 5.

10. **383-390 - Given the questions I had about figure 5, maybe it would be helpful to include some example satellite images showing damage in these events and your methodology of assessment. This would be fine just as a supplement.**

Thank you for the helpful suggestion. We agree that including example satellite images illustrating damage from the debris flow events, would enhance the clarity and robustness of the analysis. Figure 4 already presents satellite images of the three study locations, highlighting areas of scarring from previous debris flow activity and zones of increased erosion. Dam locations were approximated and built environment areas shaded based on risk of damage derived from proximity to high erosion areas. Critical infrastructure is also annotated where appropriate.

To further support the analysis, we have prepared a supplementary document that provides additional satellite images illustrating observed damage and changes to the built environment from individual events. We believe this addition will improve transparency and aid readers in understanding the spatial and temporal dynamics discussed in the manuscript.

11. **397-398 - Is there an available volume estimate for any of the debris flows in these catchments? Forgive me if you already listed it somewhere and I just couldn't find it.**

Thank you for your question. We did previously detail volume estimates for key debris flows at the study sites between lines 132 to 145 of the manuscript. For instance, the large debris flows on August 2019 in both Cutou and Chediguan gullies measured over $50 \times 10^4 \text{ m}^3$ each. Cutou gully, notably, has experienced a cumulative $11 \times 10^6 \text{ m}^3$ of coseismic deposits following the 2008 earthquake (Yan et al., 2014),

contributing to its high frequency of post-seismic debris flows. Additionally, a major debris flow of $64 \times 10^4 \text{ m}^3$ damaged infrastructure in Chediguan in 2019 despite the presence of a check dam.

In Xiaojia, while debris flow hazard is considered moderate due to fewer historical occurrences, a notable event following heavy rainfall remobilised approximately $30,000 \text{ m}^3$ of deposits in 2010, resulting in local flooding impacts (Liu et al., 2014). These volume estimates, drawn from prior studies and event records, form an important basis for our analysis of hazard exposure and infrastructure vulnerability within the catchments.

- 12. 412-415 – Please clarify what we’re looking at in 6iii that shows the effectiveness of the check dams – I’m struggling to see what’s illustrative in the figure. Also, the confusion I had a little earlier with visualising expansion in the built environment in Figure 5 is compounded here – the text states that expansion was restrained in Xiaojia Gully but in Figure 6iii it appears that all the development was post-2005 at the very least.**

Thank you for your insightful comment. To clarify regarding Figure 6(iii), Xiaojia serves as our unengineered comparison catchment, as it does not have check dams like Cutou and Chediguan. The figure illustrates the predicted debris flow runouts under different volume scenarios, highlighting how exposure increases with runout size. The “effectiveness of check dams” mentioned in the text refers to the comparison between Xiaojia’s unmitigated exposure and the reduced exposure seen in Cutou and Chediguan at lower runout volumes. In other words, Figure 6(iii) shows that without engineered mitigation (i.e., check dams), exposure in Xiaojia increases more steadily with volume, supporting the argument that check dams in the other catchments help limit damage during smaller debris flows.

Regarding the built environment expansion, we agree that Figure 6(iii) reflects the cumulative built environment at the time of analysis, including all development post-2005. When we mention that expansion in Xiaojia is “restrained,” we refer specifically to the comparatively limited rate and spatial extent of new development between 2011 and 2019 relative to Cutou and Chediguan, as illustrated in Figure 5. We will clarify this distinction in the revised manuscript to reduce any confusion between cumulative development and post-event expansion in lines 454 to 462.

- 13. Figure 6 – By “low and high runouts” do you mean the 10^4 and 10^5 m^3 debris flow simulations in LAHARZ? Please clarify. Maybe it would be helpful to show the result of some of the smaller simulations at a smaller scale, though, to visualise the effect of the check dams.**

Thank you for your observation. Yes, the terms “low” and “high” runouts correspond to the 10^4 m^3 and 10^5 m^3 debris flow volume simulations respectively, as previously detailed in the manuscript (lines 437 to 439). To improve clarity, we have now included these volume values in brackets directly within the figure caption. Regarding the visualization of smaller-scale simulations, we initially excluded these from the figures because their limited spatial extent appeared as simple lines, which did not effectively illustrate the influence of the check dams. We believe that the accompanying textual explanation more clearly conveys the impacts of check dams on these smaller events. However, we appreciate the suggestion and have added a sentence to the text to explicitly highlight this in lines 441 to 446).

- 14. 415-417 – I agree that the levee effect might be an important consideration here and the check dams might have played a role in encouraging development, but three catchments is an extremely small sample size to rely on. I hesitate to ask the authors to expand on their analysis at this point, but it would substantially strengthen the paper to look at a few other catchments with and without check dams in place and see whether this trend holds up to a larger sample size.**

We appreciate the reviewer’s thoughtful insight regarding the limited sample size of three catchments and the potential influence this hazard on generalisability of the levee effect and check dam impacts on development patterns. We agree that expanding the analysis to include additional catchments, both with and without check dams, would further strengthen the evidence to support or potentially refute this trend. However, due to the data availability constraints and scope of this current study to explore three gullies within close geographical context, our focus was limited to these well-characterised gullies along the Minjiang where we were able to obtain a range of satellite images across the study timescale to provide detailed accounts of the built environment evolution. We have acknowledged this limitation and highlighted the need for future research to explore this hypothesis across a broader set of catchments to confirm the observed patterns in lines 474 to 470. We hope this study serves as an initial case study that encourages future work and discussion into this area.

15. **420-423 – I don't quite understand what this means. The landscape seems like a really important factor in all these simulations, but I wouldn't invoke it if you're not directly analysing it here – it just gets the reader thinking that it might be more important than the check dams.**

Thank you for your comment. We agree that directly analysing landscape factors is beyond the scope of our current study. To avoid confusion, we have revised the manuscript (lines 468 to 471) to focus solely on the observed limits of check dam effectiveness under extreme debris flow volumes, without speculating on other factors such as landscape resilience. This clarification should help maintain a clear emphasis on the role of engineered mitigation structures in our analysis.

16. **Figure 7 – What drives the extreme jump up in destruction in Cutou under the largest debris flow volume simulation? Also, I think it would be more effective to have at least a little analysis of debris flow inundation decoupled from building damage since that's dependent on a lot of factors not controlled for in your analysis. How do the check dams control the distribution of erosion and deposition?**

Thank you for your comment and suggestions regarding Figure 7. We appreciate the opportunity to clarify and expand on these points. Regarding the pronounced increase in building damage observed in Cutou under the extreme volume simulation, this is primarily driven by the combination of the debris flow magnitude exceeding the retention capacity of the check dams resulting in overtopping (most common mechanism of check dam failure observed in Sichuan gullies) and the spatial distribution of buildings within the flow path (gully mouth), resulting in disproportionately higher damage. We have now clarified this in the revised manuscript to better explain this dynamic in lines 523 to 536.

We also agree with the importance of decoupling debris flow inundation from building damage, as structural damage depends on multiple factors not controlled for in our study, such as building materials and construction quality. Due to the lack of comprehensive data on these factors across the study sites, we used exposure as a proxy for damage risk. This limitation is now explicitly acknowledged in the manuscript. Furthermore, we have included a discussion on how check dams influence erosion and deposition patterns. Specifically, these structures trap sediment upstream, reducing downstream sediment loads and runout distances which helps mitigate exposure during low to moderate debris flow events. However, during extreme events, this function becomes less effective as sediment volumes exceed check dam capacity. This nuanced role of check dams has been incorporated into the revised text to provide a more complete understanding of their effectiveness. Finally, we appreciate your suggestion to perform more detailed numerical analysis of multiple hazard events and we agree this would strengthen the findings. While beyond the scope of the current study, it is highlighted as an important direction for future research.

17. **469 – I don't think this even needs to be stated**

Thank you for your comment. Upon review we agree this sentence is not needed and it has been removed from the revised manuscript.

18. **481-482 – How did the dams fail? Collapse? Overtopping?**

Thank you for your comment. You are correct that the check dams in both Cutou and Chediguan failed primarily through overtopping during the 2019 debris flow events. We have now included this important detail in the revised manuscript to clarify the failure mechanism and its implications for physical vulnerability in these catchments in lines 552 to 554.

19. **519-520 – Declined from what? Earlier industrial presence? Or is it just less industrialised than the other catchments?**

Thank you for your question. The phrase "declined industrial presence" in reference to Chediguan is intended to indicate a reduction in industrial activity compared to historical levels, rather than a comparison to other catchments. This decline reflects changes in land use and economic activity in the area over recent decades. We have revised the manuscript to clarify this point for greater precision in lines 595 to 597.

20. **525-526 – Doesn't it go without saying that a larger debris flow is going to subject a larger area to potential damage?**

Thank you for your comment. We agree that it is intuitive for larger debris flows to impact a wider area, which naturally leads to increased potential damage. However, we included this observation to explicitly quantify and demonstrate this relationship within the context of our specific catchments and debris flow volumes using the LAHARZ model. Our aim was to establish a clear, data-supported correlation between runout volume and exposure, which underpins the subsequent analysis of how check dams influence this dynamic. We have revised the text to clarify this intent and to better emphasise the added value of this

quantified relationship in the context of engineered mitigation measures and urban development patterns in lines 602 to 605.

21. 529-530 – is it that check dams are not effective against the largest flows or that the check dams just aren't big enough to be effective?

Thank you for your comment. To clarify, the check dams in Cutou and Chediguan are, from our investigation designed to mitigate small to medium debris flow volumes which limits their effectiveness against extreme runoff volumes as shown by our simulations. In other words, we are not suggesting that check dams are inherently ineffective against large flows in general, but rather that, within the specific context of mountainous Sichuan landscape their features (size/design) are insufficient to fully contain or mitigate these extreme events. We have updated the manuscript to better reflect this in lines 607 to 612.

22. 531 – What results if you simulate large debris flows without the check dams present?

Thank you for your question. We did perform additional simulations of debris flows without the presence of check dams in Cutou and Chediguan. These showed that while check dams mitigate damage from smaller events, during extreme debris flows their failure (e.g., overtopping and damage) can exacerbate impacts by releasing stored sediment and debris downstream resulting in flows of similar size to those with check dams. In some areas of the flow path, the presence of damaged check dams led to greater building and infrastructure loss compared to scenarios without check dams, due to sudden dam breaches increasing flow volumes and velocities. We have added this clarification and the comparative simulation results to the revised manuscript to better reflect these dynamics (lines 612 to 617).

23. 532 – “raised development level” to me implies an increase in elevation – maybe reword.

Thank you for your comment. We agree and have reworded the sentence within the revised manuscript in lines 617 to 618.

24. 551 – Needs more than 3 data points for substantive support of this claim, I'd say. A lot of other factors could be encouraging or discouraging development.

Thank you for your comment. We agree that relying on only three catchments provides limited scope to conclusively support the levee effect theory. We have clarified this in the manuscript in lines 647 to 651 by presenting our findings as preliminary evidence that aligns with the levee effect concept, while acknowledging that other socio-economic and geographic factors could also influence development patterns. We recommend that future studies incorporate larger datasets and broader case studies to substantiate these observations further.

25. 566 – Could it also be that construction of check dams has been prioritised in areas with a lot of development?

Thank you for your comment. We agree that it is plausible check dam construction was prioritised in areas with existing or planned urban development, reflecting a reactive rather than solely proactive approach to hazard mitigation. In the Cutou catchment, the timing of check dam construction does indeed coincide with ongoing urban expansion, which may indicate that mitigation efforts were focused on protecting increasingly developed areas. However, we highlight that the rate of urban expansion post-construction increased significantly. We have clarified this point in the manuscript to highlight that the relationship between check dams and urban growth is complex and potentially bidirectional (lines 666 to 671). This nuance reinforces the need to consider both development patterns and structural interventions in tandem when assessing debris flow risk and planning mitigation strategies.