1. **Line 217 The facility value and related parameters selection were overlooked, and that increase the vagueness of application process in case the reader was interesting in similar application design.**

We have clarified and expanded our methodology section for assigning fragility values to buildings. To address the vagueness, we have explicitly stated that the fragility values were assigned based on a combination of literature sources and satellite images. Specifically, buildings that were inundated or damaged in previous events, or those located along the channel or gully mouth, were given a fragility value of 1, while all other buildings were assigned a value of 0. We validated these values using historical damage reports from the 2008 earthquake recovery period to ensure their applicability. Additionally, we emphasised this approach allows for replicable application designs in similar hazard-prone areas, thus addressing the concern regarding the transferability of the methodology to other contexts. We believe these revisions provide the necessary clarity for readers interested in applying this approach to similar designs. Please see additions for the revised manuscript, in italics, below:

E_b is the number of buildings damaged, and C is the fragility index of the elements at risk (Zou et al., 2019). Fragility values range from 0 to +1, with higher values indicating greater susceptibility to damage and/or failure. We assigned fragility values through using a mixture of literature and satellite images; buildings shown to be inundated or damaged in previous events or situated along the channel or gully mouth were given a value of 1, all other buildings were set a value of 0. These values were validated using historical damage reports, where available, from the 2008 earthquake recovery period to ensure applicability (Zeng et al., 2015; Wei et al., 2021; Petley et al., 2023). This approach allows for replicable application designs in similar hazard-prone areas.

2. **The justification of using -1 to +2 as units of measure to be inserted and its quantification relationship to vulnerability value is missing.**

We have added a justification for using the -1 to +2 scale as units of measure in our analysis. The scale was selected based on its ability to represent a range of vulnerability values that are meaningful for our study area. We have also included a quantification relationship between this scale and vulnerability values, explaining how each unit on the scale corresponds to specific levels of vulnerability, both physical and economic. Please see the additions below, in bold italics, to be added to the revised manuscript:

The key difference between our method and that of Zou et al (2019) is the incorporation a modification factor, M, to account for the effectiveness of engineered measures like check dams in mitigating building damage and subsequent exposure. The mitigation factor, M, quantifies the influence of engineered measures, in this study check dams, on the vulnerability and subsequent exposure of buildings to debris flow impacts. The addition of this factor brings an evaluative element to the exposure assessment, quantifying the influence of check dams and assigning values ranging from -1.0 to +2.0 to reflect a spectrum of mitigation outcomes:

- *M = -1: Effective mitigation of debris flows, resulting in a significant reduction in hazard exposure, as evidenced by a decrease in the number of buildings damaged during historical events following construction.*
- *M = 0: No mitigation present; exposure levels are entirely dependent on natural site conditions.*
- *M = +1: Ineffective mitigation; there is no reduction in the number of buildings impacted in recorded debris flow events following dam construction.*

- *M = +2: Mitigation increases exposure. Recorded events of similar volume show an increase in the number of buildings impacted following dam construction.*

The above -1 to +2 scale was selected to capture a nuanced relationship between mitigation effectiveness and vulnerability. A reduction in M (e.g., -1) lowers hazard exposure by reducing flow impacts at critical locations, thereby increasing E_{df}. Conversely, an increase in M (e.g., +2) elevates exposure, as development in hazard-prone areas amplifies the potential for damage. For example, a decrease in M by one unit (from 0 to -1) reflects an improvement in flow attenuation due to effective check dams, reducing overall exposure. Conversely, an increase in M by one unit (from 0 to +1) signifies a scenario where mitigation fails, e.g. the. 2019 debris flow event in Cutou, maintaining high exposure levels. At M = +2, exposure exceeds natural vulnerability due to increased hazard presence caused by intensified land use near mitigation structures.

This scale was developed through a combination of evaluating present hazard mitigation and analysing of historical data, particularly from the 2008 earthquake recovery. Moreover, this approach, based upon the methodology proposed by Zou et al. (2019), allows for an assessment of exposure by considering both the physical resistance of buildings and the efficacy of mitigation efforts.

3. LAHARZ simulation, data processing, assumption, and technical details were missing.

We have now provided additional details regarding the LAHARZ simulation, including the specific data processing steps, key assumptions, and technical details. This includes a description of the model setup, the sources of input data, and the assumptions underlying the simulation parameters. We have also discussed the limitations of the model, and any uncertainties associated with the assumptions made. Please see additions for the revised manuscript, in italics, below:

LAHARZ is a GIS toolkit for lahar hazard mapping and modelling, developed by the USGS to calculate the area of inundation and cross sections based on empirical scaling relationships between area and volume (Schilling., 2014; Iverson et al., 1998)*. These empirical relationships allow for the creation of realistic inundation areas without a priory knowledge of the rheological parameters. The model simulates a debris flow triggered at a source point located on a digital elevation model and with an initial source volume. The model calculates the flow path downslope of the triggering location then generates a cross-section at each point downslope that represents the depositional volume for that area (Iverson et al., 1998).*

We implemented this model using the extension in ArcGIS (USGS., 2007). We used the 30m resolution ASTER DEM as an input, as it is the most reliable of the globally available DEMs. We identified the source areas of 2019 debris flows for Chediguan and Cutou and the 2011 for Xiaojia (Cutou – 351603, 3473449; Chediguan – 350846, 3453894; Xiaojia – 356666, 3439268) from satellite imagery and used these as the triggering locations for our simulations. We then prescribed three input volumes at each of these locations ($10^4\ m^3, 10^5\ m^3$ and $10^6\ m^3$). The flow volumes simulate a range of observed post-2008 debris flows, representing low, high, and extreme debris flows documented in the Fan et al., (2019a) datasets. The volumes we selected reflects the range of similar hazard events in comparable geomorphological settings such as other parts of China and Italy (Wu et al., 2016; Bernard et al., 2019). For catchments with check dams, we added barriers at each check dam location by raising the cell count of the DEM by the height of the check dam obtained from field imagery.

The model was validated by comparing simulated runout extents with observed debris flows from post-2008 events. While a 30m resolution was the only available DEM for our study locations, we tested the sensitivity of DEM resolution on the extent of the final flow. A higher, 10m resolution DEM was available *for the Cutou gully and we ran LAHARZ for that catchment. While the 10m DEM created a more effective flow path compared to the mapped data, the flow depositional area was similar in both the 10m and 30m scenario (RMSE 18m). Given the lack of a significant difference between the two DEM resolution we ran 30m scenarios across the three catchments. We note that there is not a strong understanding currently of what controls the maximum size of debris flows within Wenchuan catchments, hence we cannot attribute a particular probability to each scenario.*

4. Fig. 7 and the amount of buildings, types, and degree of vulnerability in terms of economic or physical were missing.

We have revised Figure 7 to include the total number of buildings, along with the degree of vulnerability. Additionally, the figure caption now provides this information for clarity. The lines below will be added to explain why we did not include economic data in the analysis:

> *"However, due to the lack of available data on building materials in these three regions, we were unable to quantify their influence on structural vulnerability. As a result, exposure was determined to be the primary contributing factor to building damage."*

Revised Figure 7(i) with building numbers added and figure caption below (Figure 7(ii) will remain unchanged):

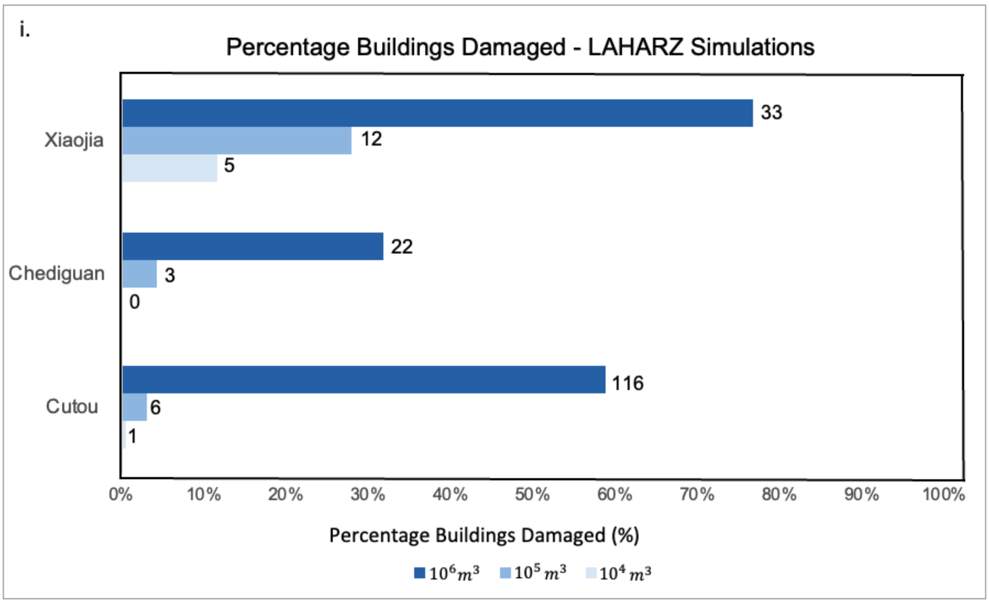

Figure 7: *Built environment impacts from three debris flow scenarios modelled using LAHARZ at Cutou, Chediguan and Xiaojia. (i). Percentage of buildings damaged as a proportion of total buildings (Cutou – 197, Chediguan – 69 and Xiaojia – 43) in each scenario. (ii) Total number of buildings damaged by each simulated debris flow.*

5. Maps and figures are very simple, and the conclusion was almost predictable, as I am still looking for scientific arguments and proofs that may increase the credibility of research contribution.

We appreciate the reviewer's feedback. In response, we have revised the introduction and research objectives to better highlight the key scientific arguments and methods used to substantiate the credibility of our findings. These revisions aim to strengthen the scientific rigor and contribution of our research.

Regarding the maps and figures, we chose to keep them simple to ensure ease of interpretation and clear visualisation of exposure changes over time (Figures 5 and 6). We believe this approach enhances the accessibility of the findings without compromising the scientific integrity of the results.