Responses to comments of Reviewer 3

(1) Line 105: Please explicitly define damage degree. Is it the vulnerability matrix? A more explicit statement of that in the beginning of 2.1 would be clarifying. Line 155: Please elaborate on the fishnet tool and what it does briefly, beyond it's purpose for your methods

***Answer**: Thank you. We have improved the definition of damage degree to let it to be more clear. Yes, it indicates the vulnerability matrix. As for the 'fishnet' tool in ArcGIS, it is used to create a grid of rectangular or square cells, forming a structure similar to net across a specified area. This grid, named 'fishnet', is usually used for spatial-related tasks, such as sampling, density calculations, and overlay analysis. Each cell in the grid is a polygon, and we can specify the cell size, and the extent of the grid. In our study, considering the difficulty in extracting single building, 'fishnet' tool was employed to generate rectangle cells from building clusters with specified cell size, 500 m^2 for a building.*

(2) Page 6, Figure 1: please explain variables and their names more explicitly, especially in the GridSearch algorithm section. Also, please indicate choice of colors and corresponding meanings

***Answer**: Thank you. We have further explained the variables presented in Fig. 1. V_p is the physical vulnerability of buildings, and P_t represents the impact pressure of a debris-flow event to buildings. HD and VD are horizontal and vertical distance of buildings to their nearest debris-flow channel. These factors in dataset I are used to develop a physical vulnerability matrix. As for the variables in dataset II, they are the depositional volume of a debris-flow event (U), area of a debris-flow catchment (A), length of the main channel for a catchment (L), the average topographic relief (R), and the average gradient of main channel (J). ρ_{df} is the mean density of the material. More details regarding the selection reason of these variables and their definitions have been provided in sections 2.1 and 2.3. Please see the 'Revised manuscript with changes marked'.*

GridSearch algorithm was used to generate the optimal hyperparameters of XGBoost, including n_estimators, learning_rate, max_depth, min_child_weight, and gamma. Explanations of these hyper-parameters are provided as below.

***n_esmators.** It determines the number of weak learners (decision tree) to be used in our model. The increased number of weak learners could improve accuracy but also* *may result in overfitting.*

learning_rate. *This parameter controls the weight assigned to each decision tree added to the model. A small learning_rate can cause slow speed of model training, but it can improve the generalisation performance by requiring more decision trees (n_estimators), ultimately leading to a more robust model.*

max_depth. *It controls the maximum number of splits that each tree can have. Deeper trees are able to describe more complex relationships, but an overfitting may not be avoided. In contrast, A smaller depth may result in underfitting because the trees are too shallow to capture data features.*

min_child_weight. *This parameter decides the minimum amount of data that is required to create a new leaf node on a decision tree. A higher value allows the algorithm to be more conservative, reducing the risk of overfitting by requiring larger samples to form a split.*

gamma. *This parameter defines the minimum improvement in the model's accuracy required to make a split. If this improvement is less than the specified gamma, the split will not occur. A higher gamma value makes the algorithm more conservative capable of preventing overfitting.*

The optimal values of these hyper-parameters have been provided in section 2.3. Please see the 'Revised manuscript with changes marked'.

As for the selection of colors in this figure, the first part with deep blue color represents 'Data collection', as shown in the improved Fig. 1. After that, 'Model development' highlighted with orange was conducted to propose a physical vulnerability matrix and develop an impact pressure estimation model. In the development of machine learning mode, the optimal hyperparameters of XGBoost were generated by GridSearch algorithm. After the determination of physical vulnerability that building may suffer when a future debris-flow event occurs, their economic loss can be estimated caused by an event (deep green). Finally, we applied this framework to the Longxihe Basin, China to conduct a case study (dark red) to suggest the potential economic loss this basin may encounter if debris flows occur in this area.

(3) Figures 7- 12: Please offer a more descriptive caption of the figures.

Answer: *Thank you. We have improved the descriptions of these figures. Please see the 'Revised manuscript with changes marked'.*