Review comments and responses

Reviewer 2

Over view

I have read the manuscript with interest. The paper develops a software for hazard assessment of coseismic landslides considering three stages after earthquake. The paper is well written, but I have some concerns regarding the innovation and the method used by the paper.

Comments

1. I agree with the authors that rapid assessment of coseismic landslides is crucial to emergency response after strong earthquakes. The paper combined logistic regression and Bayesian probability methods in Matlab for assessing the spatial probability of landslides. Even the authors emphasized that there is no specialized software for seismic landslide hazard assessment, particularly in the various needs of different stages after a major earthquake, however, the methods they use are traditional methods, nothing new about the methodology itself. There are quite many existing toolbox or packages in ArcGIS, QGIS and R, which can be used for the same analysis as the authors did in Matlab.

Authors' response: Yes. Currently, there are many toolboxes or packages for assessing the spatial probability of landslides in different programming languages. However, these toolboxes are built by the traditional hazard assessment process, which requires landslide samples for modelling. As a result, the prediction results often lag behind the actual application, which cannot satisfy the emergency assessment of earthquake-induced landslides (Ma et al 2020). Therefore, to solve this problem, we integrated a new generation of earthquake-triggered landslide hazard model (Xu_{2019} model), so that our software can serve for the emergency assessment of earthquake-induced landslides in the Sichuan-Yunnan area. Secondly, most studies on data-driven model have used the same ratio of sliding samples to non-sliding samples. Such a sampling method artificially exaggerates the proportion of sliding samples in the study area (Allstadt, et al 2018, Nowicki Jessee et al 2018); thus, the assessment results only consider the relative hazard level, but do not represent the real occurrence probability of landslides. Consequently, the resulting probability of the model overestimates the actual landslide occurrence probability (Shao et al 2020, Nowicki Jessee et al 2018). We proposed a real probability prediction method of coseismic landslides by the bayesian probability method and LR model (Shao et al 2020). The results of this model represent the occurrence probability of landslide rather than the relative hazard level (Shao et al 2020) and thus can calculate the landslide area of the quake-affected area. Thirdly, to our knowledge, although there are quite many existing toolboxes or packages in ArcGIS, QGIS and R, there is no specific software for regional landslide hazard assessment based on matlab language, so our work will also help those familiar with matlab language to carry out the earthquake-induced landslide assessment.
Allstadt, K.E. et al., 2018. Improving Near‐Real‐Time Coseismic Landslide Models: Lessons Learned

from the 2016 Kaikōura, New Zealand, Earthquake. Bulletin of the Seismological Society of America, 108, 1649-1664.

Ma, S., Xu, C., Shao, X., 2020. Spatial prediction strategy for landslides triggered by large earthquakes oriented to emergency response, mid-term resettlement and later reconstruction. International Journal of Disaster Risk Reduction, 43, 101362.

Shao, X., Ma, S., Xu, C., Zhou, Q., 2020. Effects of sampling intensity and non-slide/slide sample ratio on the occurrence probability of coseismic landslides. Geomorphology, 363, 107222.

Nowicki Jessee, M.A. et al., 2018. A global empirical model for near-real-time assessment of seismicallyinduced landslides. Journal of Geophysical Research: Earth Surface, 123, 1835-1859.

2. In addition, the three-stage methodology is just classified considering the different time window after an earthquake. The methods at all stages are the same. The only difference is adding more and more landslide data after an earthquake due to more available information with time, such as remote sensing images. At third stage, if we already know all coseismic landslide distribution by RS imagery interpretation, why we still need the hazard model? Even at this stage, you get high R2, it is because the overfitting of the model. It does not mean the model will have good prediction power for next event.

Authors' response: In different stages after a large earthquake, the demand to mitigate earthquake disasters is different (Ma et al 2020). Especially in mountainous areas, the spatial prediction of landslides is of great significance to short-term emergency response (stage 1), medium-term temporary resettlement (stage 2) and long-term rehabilitation and reconstruction (stage 3). In stage 1, in the absence of landslide data, the rapid emergency hazard mapping using near-real-time model can provide guidance for post disaster emergency rescue and the interpretation and identification of earthquake-induced landslide in stage 2. In stage 2, considering the timeliness, the earthquake-induced landslide hazard assessment was carried out based on partially available landslides data. The assessment results are beneficial for the improvement of the construction of earthquake-induced landslide inventory, and provide useful information on avoiding high landslide hazard areas for quake-affected areas.

In stage 3, we are faced with not only the problem of coseismic landslide identification, but also the weakened slope caused by the quake. As a result, it is critical to locate the landslides that are stable during the earthquake but unstable for a period of time after the earthquake. Such an essential process can be achieved by the hazard assessment of earthquake-induced landslides based on the complete landslide data. Thus, the results obtained in stage 3 will definitely be more objective than those obtained in the stage 2, because the training samples used in the model in this stage are abundant and more objective. That is to say, the prediction ability of the model in stage 3 is stronger than that in stage 2 (the prediction rate in the third stage is higher). Meanwhile, the evaluation results at this stage can effectively serve the town planning and long-term risk assessment of the subsequent quake-affected areas. To summarize, we suggested to perform earthquake-induced landslide hazard assessment at multiple stages in a large earthquake in order to better deal with the landslide disaster prevention and mitigation issues that earthquake areas face at various stages. The relevant descriptions have been added in Section 3.2.1 (line 288-296).

Ma, S., Xu, C., Shao, X., 2020. Spatial prediction strategy for landslides triggered by large

earthquakes oriented to emergency response, mid-term resettlement and later reconstruction. International Journal of Disaster Risk Reduction, 43, 101362.

Major comments:

3. Why the authors did not try CNN or other more advanced AI methods, which should have better performance than logistic regression and Bayesian probability methods?

Authors' response: Currently, the majority of the models employs one of several possible classification methods, including classical statistics (e.g., logistic regression, discriminant analysis, linear regression), index-based (e.g., weight-of evidence, heuristic analysis), machine learning (e.g., support vector machines, random forest) and neural networks (e.g., recurrent neural network, Convolutional neural network) (Reichenbach et al., 2018). Among them, the LR model is one of the most widely used models in the hazard assessment of earthquake-induced landslides by virtue of its simplicity, high efficiency, and high prediction accuracy (Reichenbach et al., 2018; Shao and Xu, 2022) (Fig.1). In addition, it is the preferred method for establishing the near-real-time prediction model of earthquake-induced landslides (Nowicki Jessee et al., 2018; Tanyas et al., 2019; Xu et al, 2019).

In recent years, deep learning methods, especially Convolutional Neural Networks (CNN), have been pervasively applied in landslide hazard assessment. Similar to other machine learning methods, the internal structure of CNN model is complex like a black box, and the models need to classify the independent variables before the evaluation modeling (Yang et al., 2022). Compared with the CNN model, the LR model can better avoid these two problems. This method can carry out different types of independent variables including continuous variables and discrete variables. The LR model can give specific regression coefficients of independent variables, with simple calculation process and definite physical meanings. At the same time, recent studies show that the LR regression model performed better in prediction   other machine learning models (Zhao et al., 2022). The relevant description have been added in the section 3.22.
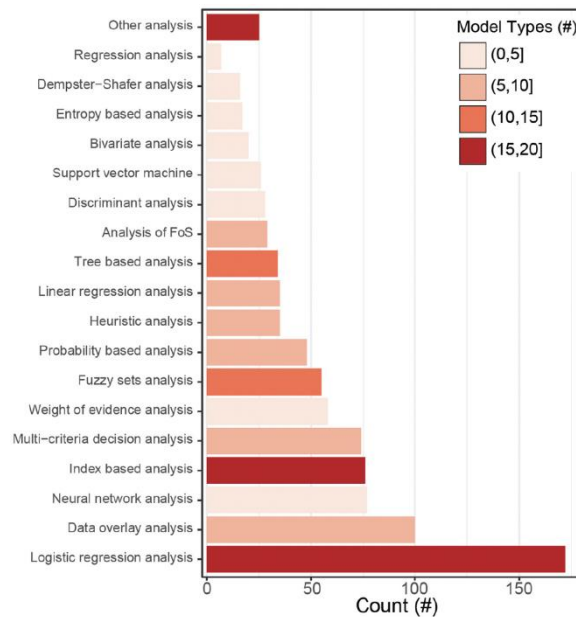
Fig.1 Horizontal bar chart shows the count of 19 model type classes used to group the 163 model names given by Reichenbach et al., 2018 in the literature databases.

Reichenbach, P., Rossi, M., Malamud, B.D., Mihir, M., Guzzetti, F., 2018. A review of statistically-based landslide hazard models. Earth-Science Reviews, 180, 60-91.

Tanyas, H., Rossi, M., Alvioli, M., van Westen, C.J., Marchesini, I., 2019. A global slope unit-based method for the near real-time prediction of earthquake-induced landslides. Geomorphology, 327, 126-146.

Xu, C., Xu, X., Zhou, B., Shen, L., 2019. Probability of coseimic landslides: A new generation of earthquake-triggered landslide hazard model. Journal of   Engineering Geology, 27, 1122.

Yang, Z., Xu, C., Shao, X., Ma, S., Li, L., 2022. Landslide hazard mapping based on CNN-3D algorithm with attention module embedded. Bulletin of Engineering Geology and the Environment, 81, 412.

Zhao, P., Masoumi, Z., Kalantari, M., Aflaki, M., Mansourian, A., 2022. A GIS-Based Landslide Hazard Mapping and Variable Importance Analysis Using Artificial Intelligent Training-Based Methods. Remote. Sens., 14, 211.

4. PGA and PGV are considered as the most important seismic factors, why the authors used intensity rather than PGA and PGV data? Besides, distance to river and distance to transportation lines are also important factors considering the river incision in mountainous regions and human work effect, why they are not considered in the model?

Authors' response: Thank the reviewers for their suggestions. Indeed, as the reviewer said, PGA and PGV are the two most important seismic factors, and these two factors can be converted to each other through specific empirical formula (Boore et al., 2014; Saffari et al., 2012). Like PGA and PGV, seismic intensity is also one of the most important seismic factors, and there are specific formulas for seismic intensity and PGA to be converted (Du et al., 2018; Xin et al., 2020). In our study, we chose the seismic intensity instead of PGA because in the official results released by the China Earthquake administration, seismic intensity is obtained by the integration of multi-source information such as instrument records, field survey and actual earthquake damages. Compared with the PGA map obtained from simple instrument records or attenuation relationship, the map of seismic intensity can better reflect the distribution of the seismic influence field. Therefore, in the three stages, we carried out the probability prediction of earthquake-induced landslides based on the rapidly obtained seismic intensity.

The reason why the distance to river and distance to transportation lines are not selected is that among all influencing factors we selected, the topographic wetness index (TWI) and land-use type can represent regional hydrological factors and human factors respectively to a certain extent. Additionally, according to the previous studies about the spatial distribution of earthquake-induced landslides in the Sichuan-Yunnan region, we found that these two influencing factors do not show strong correlation with the occurrence of earthquake-induced landslides. Furthermore, despite the fact that these two influencing factors are not taken into account in the evaluation results, the performance of the evaluation model is satisfactory. Meanwhile, the model and software we developed are adaptable and do not place rigid limits on the input of influencing factors. Peers who are interested in assessment models might add or change the corresponding independent variables during the modeling process.

Boore, D.M., Stewart, J.P., Seyhan, E., Atkinson, G.M., 2014. NGA-West2 Equations for Predicting

PGA, PGV, and 5% Damped PSA for Shallow Crustal Earthquakes. Earthquake Spectra, 30 (3), 1057-1085.

Saffari, H., Kuwata, Y., Takada, S., Mahdavian, A., 2012. Updated PGA, PGV, and Spectral Acceleration Attenuation Relations for Iran. Earthquake Spectra, 28 (1), 257-276.

Du, K., Ding, B., Luo, H., Sun, J., 2018. Relationship between Peak Ground Acceleration, Peak Ground Velocity, and Macroseismic Intensity in Western China. Bulletin of the Seismological Society of America, 109, 284-297.

Xin, D., Daniell, J.E., Wenzel, F., 2020. Review of fragility analyses for major building types in China with new implications for intensity–PGA relation development. Nat. Hazards Earth Syst. Sci., 20, 643-672.

5. It is quite obvious that from Fig.4 that the actual landslides (black polygons) are not falling in high probability zones? The model seems not satisfactory for the first stage. Many landslides in all events are failing into blue (low probability zones), while the predicted high probability zones have a few landslides. This indicates that the model has quite high false alarms from prediction perspective. In the second stage, Fig.6, it still has the mis-matching problem. In the third stage, it looks better, but this is because as I mentioned above, the overfitting of model by using a large amount of known landslides. Actually the first stage, the rapid prediction using very limited or even no available landslide information, is most important one considering the emergency response and rescue work. The model's performance at this stage is not good.

Authors' response: In the first stage, except for the Minxian earthquake, the AUC value of other earthquakes is above 0.8. But we have to admit that the evaluation results of six earthquakes based on the Xu$_{2019}$ model can be improved. We can see that landslide observations from the earthquake match well with predicted high probabilities, but the model predicts potential landsliding in a large area beyond the mapped landslide area. Especially in Minxian, Jiuzhaigou and Yushu earthquake cases, the performance of the model is not satisfactory. But most of the current near-real-time models have such problems that the model performs well when evaluated over the domain of an entire event area, but clearly, individual pixels will predict probabilities that underestimate or overestimate the landslide hazard (Nowicki Jessee et al., 2018; Allstadt, et al 2018). We propose two possible reasons for this phenomenon: (1) The resolution of the input data of the Xu$_{2019}$ model is 100m, which affects the prediction accuracy of the model to a certain extent. Therefore, there may be errors between the modeling prediction and the actual result at the regional scale. (2) Nine earthquake cases used for the establishment of the Xu$_{2019}$ model are located in China and its adjacent areas. The corresponding epicentral areas have different topographic and geological conditions, and only four cases are in the Sichuan-Yunnan area, which may weaken the applicability of the Xu$_{2019}$ model in other quake events. Therefore, in the past few years, we have been constantly supplementing the earthquake landslide database in Sichuan Yunnan region (e.g. 2014 Ms Jinggu earthquake, 2020 Ms Qiaojia earthquake, 2018 Ms 5.7 Xingwen earthquake, 2019 Changning earthquake, 2022 Ms 6.8 Luding earthquake,.etc). We suggest that with the accumulation of enough coseismic landslide inventories of earthquake cases in Sichuan-Yunnan area, we can constantly update the near-real-time earthquake-triggered landslide hazard model based on these abundant landslide data and high resolution input factors data, and further improve the accuracy of the modelling in the emergency assessment. The relevant description have been

added in the discussion section (Line 578-602).

For the second stage, the predicted landslide area (Ap) of the six events are almost the same as the observed landslide area (Ao). Except for the Jiuzhaigou earthquake, the overall error of the remaining five earthquakes is between 9% and 50%, of which the error of the results of the 2008 Wenchuan earthquake is the lowest with 9%, indicating that the assessment results of the second stage are reliable for the quantification of coseismic landslide development area in the quake-affected area. In addition, from the complete landslide inventory and prediction results of six events (Fig. 6), although some landslides are spread out on the low-hazard areas, most landslide are located in the high-hazard areas which are relatively consistent with the actual landslide distribution. Meanwhile, based on the complete landslide data, the validation results in the second stage show that the AUC values of the second stage are all above 0.85, which indicates that the model have pretty good performance.

Allstadt, K.E. et al., 2018. Improving Near-Real-Time Coseismic Landslide Models: Lessons Learned from the 2016 Kaikōura, New Zealand, Earthquake. Bulletin of the Seismological Society of America, 108, 1649-1664.

Nowicki Jessee, M.A. et al., 2018. A global empirical model for near-real-time assessment of seismicallyinduced landslides. Journal of Geophysical Research: Earth Surface, 123, 1835-1859.