Review of manuscript nhess-2024-96

General comments

In my opinion, the manuscript could be accepted after a major revision. It does not make clear what is the problem it tries to solve, and it lacks details about key elements of the methodology (i.e., the use of machine learning to calculate landslide hazard). Moreover, the discussion of the results –and essentially the manuscript itself– focuses on the seismic hazard model, while the title suggests that it is about landslide hazard too. Moreover, the documentation calculation of the landslide hazard should be improved. Moreover, given that landslide hazard modelling and the results with respect to landslide hazard are given so little coverage in the manuscript, please consider revising the title, removing from the manuscript whatever concerns landslide hazard, and focusing on seismic hazard.

Although I am neither an English native speaker nor an English language professional, I believe I have found more than a few instances, where the writing should be improved. Therefore, the manuscript does not meet editorial standards, in my opinion. Please consider having the manuscript edited by an English language professional.

As far as the figures are concerned, which have been published elsewhere and are included in the manuscript as they are or after some modification, please make sure that the reproduction rights have been secured, and inform the editor, or please consider removing them.

Abstract

Please considering stating clearly what is the main topic of the paper. It is not clear what is the problem that this paper tries to solve. It states that it presents a new probabilistic seismic hazard assessment model that accounts for multi-segment faults, and that it uses this new model to do landslide hazard assessment. As suggested in line 28-29, the new seismic hazard model makes better predictions of some ground motion intensity measures, which may lead to a better assessment of landslide hazard. Moreover, please consider finishing the abstract with a statement about the implications of the findings.

- Line 1, “Potentials”: please consider replacing with “hazard”.

Thanks! We revise it to Occurrence Probabilities.

- Lines 21-23: Please clarify why the abstract mentions this historical earthquake.

This earthquake is a typical multi-segment rupturing earthquake, as stated by many studies, such as Huang et al., 2021. We revised the words in Line 23: as demonstrated by the historical multi-segment ruptured 1515 M7.8 Yongsheng Earthquake.

- Line 24, "presented": incorrect tense. Please replace with "presents".

Revised it. Thanks!

- Line 24, "a novel seismic hazard modeling study": please replace with to "a new probabilistic seismic hazard model"

Revised it. Thanks!

- Line 25, "integrating fault slip parameters and assessing multi-segment rupturing risks": Please explain why is this being done by this paper. What is the necessity? A classical PSHA would not do?

Thanks for your question! We mean that in this region of NWYR, the rugged terrain makes it difficult to find the fault surface tracks. The climate is humid with abundant rainfall, leading to high vegetation cover, severe weathering, and significant damage to fault surface traces.

We modified the words from Line 24 to Line 28 as follows: This article presents a new seismic hazard modeling study for the NWYR, with recent fault geometrical and slip rates studies, incorporating recent findings on fault geometry and slip rates, and integrating fault slip parameters and historical seismicity rates to assess multi-segment rupturing risks.

- Line 28-29, "emerges as": Please consider replacing with "is proposed as".

Revised it. Thanks!

- Line 28-29, "most suitable": Please explain by which criteria and for which use.

We revised the words:"Among the four potential multi-segment rupture combination models examined, Model 1, characterized by multi-segment rupture combinations on single faults, particularly fracturing the Zhongdian fault, is proposed as the most suitable for the NWYR, given that the normalized misfit scores (NMS) are all below the 30%~40% threshold, supported by the alignment of modeled seismicity rates with fault slip rates.”

- Line 28-29, “supported by the alignment”: Please consider replacing with: “as suggested by the agreement”.

Thanks for your advice. We modified it to: Model 1, characterized by multi-segment rupture combinations on single faults, particularly fracturing the Zhongdian fault, is proposed as the most suitable for the NWYR, given that the normalized misfit scores (NMS) are all below the 30%~40% threshold, supported by the alignment of modeled seismicity rates with fault slip rates.

- Line 30, “demonstrated”: incorrect tense. Please replace with “demonstrates”.

Revised it. Thanks!

- Line 30, “peak ground-motion acceleration (PGA) values, calculated with a 475-year

return period from modeled seismicity rates, exhibited”: incorrect terminology. Please consider replacing with: “the peak ground acceleration for a mean return period of 475 years, which is calculated with the developed probabilistic seismic hazard model, has”

Thanks! We revised it.

- Line 32, “fault distribution”: Please clarify if the manuscript refers to the spatial distribution of the faults.

- “than the China Seismic Ground Motion Parameters Zonation Map”: please consider revising replacing with “than the PGA given by the Chinese seismic ground motion parameters zonation map.”

We revised them according to the reviewer’s comments. Thanks!

- Line 33: Please give a one-sentence description of the simulations.

Thanks! We revised it as follows:

Furthermore, we utilized PGA values with the Bayesian Probability Method and the Machine Learning Model to predict landslide occurrence probabilities, based on our peak ground motion acceleration distribution map.

- Line 34, “across”: Please consider replacing with “as a function of”.

Revised it. Thanks!

- Line 37, “highlighted”: incorrect tense. Please replace with “highlights”.

Revised it.

1. Introduction

Please justify why this study only uses the peak ground acceleration. Please state what are the ground motion intensity measures used in the literature for landslide hazard and for vulnerability to landslides.

Both peak ground acceleration (PGA) values and intensity measures can be utilized for assessing landslide hazards and vulnerabilities, as they indicate the magnitude of the seismic forces on the rock generated by earthquakes. In contrast, other parameters, such as peak ground velocity (PGV), primarily convey velocity information. PGA values can be calculated directly from probabilistic seismic hazard assessment (PSHA) studies, whereas intensity maps require further transformation from PGA values.

Line 156: The use of a machine learning model is suddenly mentioned here. Please consider mentioning it in the title, and in the abstract. Please justify the use of machine learning and consider adding comparisons of this calculation using machine learning with classical methods or cite a reference that validated this method.

Thank you for your advice. We have incorporated it into the abstract. However, we did not include it in the title since it represents only a small part of our results. Instead, we revised the title to: 'Modeling Seismic Hazard and Landslide Occurrence Probabilities in Northwestern Yunnan, China: Exploring Complex Fault Systems with Multi-Segment Rupturing in a Block Rotational Tectonic Zone,' which also implies the simulation work related to landslide occurrence.

Line 158-159: “disaster preparedness… in the area”. Indeed, this study may help in this direction, but please consider mentioning in the abstract and in the opening of the introduction that this is also part of the context of this study.

Thanks! We add it in the end of the abstract.

3. Multi-segment rupture hazard Modeling

Line 335: Please consider describing what is the state of the art in probabilistic seismic hazard assessment, and then explain why accounting for the slip rate would be an improvement.

Thank you for your advice. We added the words in the beginning of the section 3 of multi-segment rupture hazard Modeling. We emphasize the importance of fault slip rate rather the historical seismicity rate.

"Recognizing the significance of these rupture parameters in producing multi-segment rupturing, recent studies, such as those by Chatier et al. (2019), Cheng et al. (2021), Lee et al. (2022), and Chang et al. (2023), included the possibilities and probabilities of multi-segment rupturing in seismic hazard analysis. Additionally, Dutykh et al. (2013) and Rashidi et al. (2020) employed multi-segment rupturing into models of tsunami wave generation. The concept of multi-segment rupturing was also incorporated in the UCERF3 model through their complex "Grand Inversion" methodology, which integrates data on fault slip rates, historical seismicity, and paleoseismic records (Page et al., 2014; Field et al., 2014). However, for most other regional studies, collecting all the necessary input parameters remains challenging.

In seismic hazard modeling, fault slip rates can be used instead of historical seismicity data to simulate seismicity rates on faults, as slip rates span multiple seismic cycles of large-magnitude earthquakes and provide estimates of the average earthquake recurrence interval (Youngs and Coppersmith, 1985). We utilize the methodology developed by Chatier et al. (2019) to translate these fault slip rates into seismicity rates, considering both multi-segment and single-segment ruptures."

It is not uncommon to take into account the characteristic earthquake in seismic hazard models. It is not clear why this paragraph mentions this in its opening.

We appreciate the reviewer's comment regarding the mention of characteristic earthquakes in our paragraph. The Y-C model is primarily derived from the characteristic earthquake model, which provides a foundation for understanding seismic hazards. In our study, the SHERIFS code offers two options for magnitude-frequency relationships, and we specifically chose the Gutenberg-Richter (G-R) model due to its robustness and widespread acceptance in the literature.

See Line 379-383. “Therefore, in our analysis of seismicity rates for the whole seismicity rates on the regional faults, we opted to utilize the G-R relation (Gutenberg and Richter, 1944) as the Magnitude-Frequency relationship, rather than the Youngs-Coppersmith (Y-C) relation (Youngs and Coppersmith, 1985).”

Please explain why the manuscript focuses on the estimation of the PGA with a 10% probability of exceedance in 50 years. Please consider discussing the PGA for other annual probabilities of exceedance, and other intensity measures. If the national hazard map is only in terms of PGA for 475 years, please consider comparing the other intensity measures with other hazard models.

As stated in Line 107, 'Due to the high altitude, dense vegetation, and easily weathered conditions, obtaining accurate fault slip rates poses a significant challenge, often leading to considerable uncertainties.' In this region, studies on seismic hazard models are sparse, which is why we compared our results with the national hazard map in terms of PGA for a 475-year return period.

Line 581: The reader may have questions about this method, but its description is missing. The machine learning model is trained using scenarios which include the PGA as an entry parameter. However the footprint of the PGA in a scenario is different from a map of the PGA for a specific return period. Moreover, please state if the landslide hazard calculation accounts for all (or a very wide range) of annual probabilities of exceedance of the PGA (or for a very wide range of return periods), and not just the PGA for 475 years. If it does not, please explain why.

Thank you for your question. We have added text to clarify this in Section 3.4. The clarification is as follows:

Our model directly assessed the absolute probability of landslide occurrence, represented as the percentage of the landslide area within a region relative to the total area of the region (Shao et al., 2020). As a result, our hazard estimates have a true probabilistic meaning, reflecting the actual probability of landslide occurrence rather than being merely a formal expression of probability. We then calculated the probabilistic seismic susceptibility for a specific point in time within the study area, which produced a probabilistic PGA distribution map. By using this probabilistic PGA map as input for our model, we can estimate the corresponding probability of earthquake-triggered landslide occurrence. We employed these steps as the basis of our approach to calculating the probability of such landslides.

Conclusions

Line 671: Please add section title for the conclusions.

Thanks! We added it.

Line 671: The opening statement claims that the manuscript has given insights. This sentence seems out of place, because the manuscript first needs to briefly state the insights, then explain their importance, and then claim that it made valuable insights. Please consider dedicating the biggest part of the conclusions to the importance of the findings.

Thanks for your suggestion. We modified it as follows:

This study presents a comprehensive seismic hazard model for the NWYR, integrating fault slip rates and historical seismicity data to assess the risks of multi-segment ruptures and landslide occurrences. By leveraging fault slip rates and fault geometrical distributions in the NWYR, we employed the iterative method within the SHERIFS code to simulate seismicity rates for both single-segment and multi-segment ruptures. This work underscores the complexity of the fault systems within the region's block rotational tectonic environment. Our study has yielded valuable insights into the seismic hazards present in the NWYR. Through the development of fault segmentation models based on recent geological research and the application of advanced simulation techniques, we have significantly enhanced our understanding of fault activity and seismicity rates across the region. We also identified multi-segment models that best represent the observed data.

Lines 687-693: In my opinion, this is rather vague. Please consider making precise recommendations for future research.

We revised it as follows: Future seismic hazard work can be improved by utilizing geophysical data to understand fault structures where strong earthquakes are developing (Xu et al., 2017), applying geodetic data to assess energy accumulation on fault segments (e.g., Yao and Yang, 2023), using microseismicity relocation data to reveal fault asperities (Lay and Nishenko, 2022), and employing dynamic rupture simulations of single and multi-segments to enhance earthquake motion predictions (e.g., Zhang et al., 2017). These studies on fault behaviors, interactions, and multi-segment ruptures are vital for improving seismic hazard assessments. Staying vigilant and proactive in seismic risk management will better protect communities and infrastructure in the NWYR and beyond.