



applied sciences

Special Issue Reprint

Intelligent Technologies in Geotechnical Engineering and Geological Hazards

Edited by
Yongsheng Zhou

mdpi.com/journal/applsci



Intelligent Technologies in Geotechnical Engineering and Geological Hazards

Intelligent Technologies in Geotechnical Engineering and Geological Hazards

Guest Editor

Yongsheng Zhou



Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editor

Yongsheng Zhou
Institute of Geology
China Earthquake
Administration
Beijing
China

Editorial Office

MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Applied Sciences* (ISSN 2076-3417), freely accessible at: https://www.mdpi.com/journal/applsci/special_issues/0717E1P02V.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. <i>Journal Name</i> Year , Volume Number, Page Range.
--

ISBN 978-3-7258-8951-8 (Hbk)

ISBN 978-3-7258-8952-5 (PDF)

<https://doi.org/10.3390/books978-3-7258-8952-5>

Contents

About the Editor	vii
----------------------------	-----

Yongsheng Zhou

Intelligent Technologies in Geotechnical Engineering and Geological Hazards Reprinted from: <i>Appl. Sci.</i> 2026 , <i>16</i> , 7794, https://doi.org/10.3390/app16157794	1
--	---

Yonglong An, Wenjun Cui, Qingjie Gong, Yuan Jiang and Yong Huang

The Concentration Levels of Cr and Mo in the Soils of Beijing and Their Comparison with Other Capital Cities Reprinted from: <i>Appl. Sci.</i> 2026 , <i>16</i> , 6801, https://doi.org/10.3390/app16136801	6
---	---

Nan Zhang, Xiwei Fan, Chaoux Xia, Nan Xi, Jing Wang and Gaozhong Nie

Pre-Event Estimation of County-Level Human Casualty Projections in Southwestern China Based on the Spatial Aggregation of Village-Scale Lethality Data Reprinted from: <i>Appl. Sci.</i> 2026 , <i>16</i> , 6257, https://doi.org/10.3390/app16126257	21
---	----

Jie Zhou, Nengpan Ju, Chaoyang He and Mingli Xie

A Physics-Informed Online Learning Framework for Landslide Displacement Prediction Reprinted from: <i>Appl. Sci.</i> 2026 , <i>16</i> , 6003, https://doi.org/10.3390/app16126003	86
---	----

Nan Zhang, Xiwei Fan, Jing Wang and Gaozhong Nie

Seismic Hazard Area Prediction Based on Slip and Locking Rates of Active Faults in Northern Ningxia, China Reprinted from: <i>Appl. Sci.</i> 2026 , <i>16</i> , 4282, https://doi.org/10.3390/app16094282	118
---	-----

Halil Karahan and Devrim Alkaya

Integrating SVR Optimization and Machine Learning-Based Feature Importance for TBM Penetration Rate Prediction Reprinted from: <i>Appl. Sci.</i> 2026 , <i>16</i> , 355, https://doi.org/10.3390/app16010355	158
--	-----

Borko Bulajić, Silva Lozančić, Senka Bajić, Anka Starčev-Ćurčin, Miloš Šešlija, Miljan Kovačević and Marijana Hadzima-Nyarko

Horizontal PGA Estimates for Varying Deep Geological Conditions—A Case Study of Banja Luka Reprinted from: <i>Appl. Sci.</i> 2025 , <i>15</i> , 6712, https://doi.org/10.3390/app15126712	179
---	-----

Borko Bulajić, Silva Lozančić, Senka Bajić, Anka Starčev-Ćurčin, Miloš Šešlija, Miljan Kovačević and Marijana Hadzima-Nyarko

PGA Estimates for Vertical Ground Motion and Varying Deep Geology Site Surroundings—A Case Study of Banja Luka Reprinted from: <i>Appl. Sci.</i> 2025 , <i>15</i> , 6542, https://doi.org/10.3390/app15126542	199
---	-----

Mirsina M. Aghdam and Quentin Crowley

Integrating Radon/Thoron and Gamma Radiation Exposure for a Realistic Estimation of Dose Arising from Building Materials Reprinted from: <i>Appl. Sci.</i> 2025 , <i>15</i> , 6470, https://doi.org/10.3390/app15126470	215
---	-----

Gabit Sansyzbekov, Amoussou Coffi Adoko and Paul Mathews George

Thermal Damage Characterization and Modeling in Granite Samples Subjected to Heat Treatment by Leveraging Machine Learning and Experimental Data Reprinted from: <i>Appl. Sci.</i> 2025 , <i>15</i> , 6328, https://doi.org/10.3390/app15116328	235
---	-----

Wenhao Dai, Yongsheng Zhou, Huiru Lei, Xi Ma, Jiayang Dang, Sheqiang Miao, et al. Pseudotachylyte Formation in Brittle–Ductile Transition of the Anning River Fault Zone: Implications for Seismic Processes Reprinted from: <i>Appl. Sci.</i> 2025 , <i>15</i> , 5870, https://doi.org/10.3390/app15115870	251
Runze Zhang, Zilong Guo, Weiming Gong and Zhihui Wan Field Test and Numerical Simulation Study of Bearing Characteristics of Combined Post-Grouted Piles for Railway Bridges Reprinted from: <i>Appl. Sci.</i> 2025 , <i>15</i> , 335, https://doi.org/10.3390/app15010335	276
Fei Liu, Runsheng Han, Yuxinyue Guo, Mingzhi Wang and Wei Tan Element Migration of Mineralization-Alteration Zones and Its Geological Implication in the Beiya Porphyry–Skarn Deposit, Northwestern Yunnan, China Reprinted from: <i>Appl. Sci.</i> 2024 , <i>14</i> , 9653, https://doi.org/10.3390/app14219653	291
Huiru Lei, Shimin Liu and Wenhao Dai Frictional Experiments on Granitic Faults: New Insights into Continental Earthquakes and Micromechanical Mechanisms Reprinted from: <i>Appl. Sci.</i> 2025 , <i>15</i> , 7207, https://doi.org/10.3390/app15137207	312

About the Editor

Yongsheng Zhou

Zhou Yongsheng was born in 1969. He is currently working at the State Key Laboratory of Earthquake Dynamics and Forecasting, Institute of Geology, China, in Earthquake Administration (CEA). His major is rock mechanics and rock physics experiments, structural geology, and earthquake mechanisms. He is the Director of the Degree Evaluation Committee of the Institute of Geology (CEA), also holding the position of the Head of the Tectonophysics Committee of the Chinese Seismological Society, and the Head of the Rock Mechanics under High-Temperature and High-Pressure Committee of the Chinese Society for Rock Mechanics and Engineering.

Editorial

Intelligent Technologies in Geotechnical Engineering and Geological Hazards

Yongsheng Zhou

State Key Laboratory of Earthquake Dynamics and Forecasting, Institute of Geology, China Earthquake Administration, Beijing 100029, China; zhouysh@ies.ac.cn

1. Introduction

Intelligent technology provides us with a new method to understand geotechnical engineering and geomechanics and has widespread application in geotechnical engineering and geological hazards [1–7]. With the rapid development of intelligent systems, intelligent technologies could offer insights and innovative solutions in geotechnical engineering [1–4], and are increasingly applied to geological hazards such as the chemical pollution of soils, landslides, debris flows, and earthquake-induced failures [4–7]. This Special Issue focuses on progress in intelligent technology in geotechnical engineering and geological and earthquake hazards. This Special Issue brings together 13 contributions. The topics covered include ground motion and earthquake hazards, geotechnical engineering and landslides, geochemical surveys, experiments and simulations in rock mechanics and rock physics, and the implications for earthquake mechanisms and geoengineering. The purpose of this editorial is to provide an overview of the Special Issue, identify thematic connections among the published studies, and outline future directions for research in applied geotechnical engineering and geological hazards.

2. An Overview of Thematic Contributions

The contributions to this Special Issue can be organized into five main thematic clusters. The first cluster concerns earthquake hazards. This includes seismic hazard prediction and pre-event estimation, as well as peak ground acceleration estimates for vertical and horizontal ground motion, and comprises four papers.

The second cluster focuses on geotechnical engineering. Contributions in this cluster apply deep machine learning to landslide displacement prediction, TBM penetration rate prediction during tunnel excavation, and the simulation of post-grouted piles in railway bridges, comprising 3 papers.

The third cluster addresses geochemical surveys. Contributions focus on the chemical pollution of soils and element migration of mineralization-alteration zones, comprising two papers.

The fourth cluster relates to rock mechanics and earthquake mechanisms. This section covers frictional experiments and field deformation on granitic faults, as well as the implications for seismic processes, comprising two papers.

The fifth cluster focuses on rock physics experiments in geoengineering. This cluster includes reports on thermal damage in granite based on machine learning and experimental data, as well as radiation exposure experiments of building materials, comprising two papers.

3. Summary of Contributions

Contribution 1 investigates the concentration levels of Cr and Mo in the soils of Beijing based on soil geochemical surveying. The results show that mean concentrations of Cr and Mo are at background levels. To compare the Cr and Mo concentrations in soils with those observed in other capital cities, the statistical parameters of Cr from 20 capital cities and Mo from 11 capital cities were collected. The data reveals medium Cr level concentrations in Beijing, and the Mo level is low.

Contribution 2 presents a pre-event estimation model of earthquake lethality in the Yunnan, Guizhou, and Sichuan provinces in China based on the ground motion acceleration, with 50-year exceedance probabilities. The distribution of casualties is dependent on the probability of a strong earthquake and the location of events.

Contribution 3 developed a physics-informed online learning framework for landslide displacement prediction based on data from the Sichuan Provincial Geological Hazard Monitoring and Warning Platform. The result indicates that the in-depth integration of geo-mechanical knowledge into an online learning framework can effectively improve the reliability, interpretability, and operational applicability of landslide displacement prediction models.

Contribution 4 introduces seismic hazard prediction based on the slip and locking rates of active faults in northern Ningxia. A constraint model is constructed to perform numerical simulations of active faults following the elastic dislocation theory, which can provide slip rates and locking positions of active faults in the areas, which are related to high seismic hazards.

Contribution 5 presents a Support Vector Regression (SVR) model based on machine learning, developed to predict the rate of penetration during tunnel excavation. Integrating hyperparameter optimization used in this study improves both predictive performance and model explainability. The proposed approach method offers a robust, generalizable, and scientifically sound framework for TBM operations and geotechnical modeling applications.

Contribution 6 develops horizontal PGA estimates for varying deep geological conditions. The city of Banja Luka is used as a case study to evaluate horizontal PGA values in an area with a history of moderate to strong earthquakes and with different deep geological conditions. The data indicate that variations in deep geological conditions may have a greater impact on PGA values than local soil effects. Given the scarcity of scaling equations that consider deep geology in addition to local soil conditions, this is a step toward developing more accurate PGA estimates for comparable regions.

Contribution 7 presents PGA estimates for vertical ground motion and varying deep geology site surroundings for the city of Banja Luka. The results show that the deep geological effects on vertical ground motion can exceed the local soil effects. The amount of recorded data obtained can be used to properly calibrate the existing scaling equations and obtain more reliable estimates of vertical PGA.

Contribution 8 developed a comprehensive dose estimation model that accounts for radon, thoron, and gamma radiation. This approach enables the comparison of combined dose values with accepted radiation thresholds for building materials. The results highlight the previously underestimated importance of thoron and the necessity of multi-parameter models for regulatory compliance. The study establishes a novel framework to evaluate holistic radiation risks, urging revised standards to be developed that prioritize ventilation strategies and material mineralogy to protect public health in residential and commercial built environments.

Contribution 9 presents the thermal damage (TD) of granite specimens subjected to high temperatures of up to 800 °C and proposes a predictive model for this thermal damage. Based on the results, the relationships between TD and its influential parameters were

determined using regression models and several machine learning algorithms, which could serve as a valuable tool for assessing the thermal damage of rocks.

Contribution 10 proposes a model of rock deformation during coseismic, post-seismic relaxation, and interseismic periods based on pseudotachylytes and cataclasites, which record transient seismic slips within the brittle–ductile transition zone and ductile flow layers from the Anning River fault zone, providing a method to explain seismic processes according to the fault rocks record during brittle–ductile transition.

Contribution 11 investigates the effect of combined tip-and-side post-grouting on the bearing characteristics of post-grouted piles in railway bridges. The results show that the bearing capacity of the pile foundations is greatly increased by combined tip-and-side post-grouting. The combined post-grouting also optimizes the performance of the tip resistance, resulting in a more rapid response as the pile tip displacement increases, which modifies the pile shaft load transfer mechanism by increasing the tip resistance's contribution to the pile foundation's ultimate bearing capacity and moving the bearing center of gravity closer to the pile end.

Contribution 12 investigates the element migration of mineralization–alteration zones from the Beiya Porphyry–Skarn Deposit, Northwestern Yunnan. The mineralization–alteration zoning model indicates that the Beiya deposit has similar mineralization and alteration zone characteristics to a typical porphyry copper system, which provides new scientific information for understanding the metallogenic regularity and prospecting at Beiya, as well as similar types of deposits in the Sanjiang metallogenic belt and elsewhere worldwide.

Contribution 13 reviews frictional experiments on granitic faults, which can help to understand continental earthquakes and micromechanical mechanisms. The friction coefficient of granite gouges decreases with increasing temperature and pore fluid pressure, decreasing slip velocity, and increasing slip displacement. The velocity-weakening regime shifts to higher temperatures with increasing slip velocity or decreasing pore fluid pressure. Temperature, normal stress, pore fluid pressure, and slip velocity interact to modulate frictional stability. Microstructural observations reveal that grain size reduction, pressure solution creep, and fluid-assisted chemical processes are key mechanisms governing the transitions between velocity-weakening and velocity-strengthening regimes. The collective evidence underscores the importance of considering fault rheology in a temperature- and fluid-sensitive context, with implications for the interpretation of seismic cycle behavior in continental regions.

4. Conclusions and Future Directions

The papers collected in this Special Issue presents new progress in intelligent technology in geotechnical engineering and geological and earthquake hazards. The papers cover seismic hazard prediction and peak ground acceleration estimates for vertical and horizontal ground motion; apply deep machine learning to landslide displacement prediction, TBM penetration rate prediction during tunnel excavation, and the simulation of post-grouted piles in railway bridges; geochemical surveys on the chemical pollution of soils and element migration of mineralization–alteration zones; rock mechanics experiments and field deformation on granitic faults, and implications for earthquake mechanisms and seismic processes; rock physics experiments in geoenvironment, with thermal damage in granite measured based on machine learning and experimental data; and radiation exposure experiments of building materials.

This Special Issue confirms that intelligent technology has widespread application in geotechnical engineering and geological hazards. However, the papers published in this Special Issue are focused on lower-level application in these fields; artificial intelligence

should be used in more depth with the increase in big data. We plan to propose “Intelligent Technologies in Geotechnical Engineering and Geological Hazards, 2nd Edition”, which will focus on progress in intelligent technology in geotechnical engineering and geological and earthquake hazards. All topics related to geotechnical engineering and geological hazards, as well as earthquakes induced by human activities, are welcome. Contributions to this Special Issue could focus on field geological surveys, structural geology experiments and simulations, rock mechanics and rock physics, and rock deformation.

Conflicts of Interest: The author declares no conflict of interest.

List of Contributions:

1. An, Y.; Cui, W.; Gong, Q.; Jiang, Y.; Huang, Y. The Concentration Levels of Cr and Mo in the Soils of Beijing and Their Comparison with Other Capital Cities. *Appl. Sci.* **2026**, *16*, 6801. <https://doi.org/10.3390/app16136801>.
2. Zhang, N.; Fan, X.; Xia, C.; Xi, N.; Wang, J.; Nie, G. Pre-Event Estimation of County-Level Human Casualty Projections in Southwestern China Based on the Spatial Aggregation of Village-Scale Lethality Data. *Appl. Sci.* **2026**, *16*, 6257. <https://doi.org/10.3390/app16126257>.
3. Zhou, J.; Ju, N.; He, C.; Xie, M. A Physics-Informed Online Learning Framework for Landslide Displacement Prediction. *Appl. Sci.* **2026**, *16*, 6003. <https://doi.org/10.3390/app16126003>.
4. Zhang, N.; Fan, X.; Wang, J.; Nie, G. Seismic Hazard Area Prediction Based on Slip and Locking Rates of Active Faults in Northern Ningxia, China. *Appl. Sci.* **2026**, *16*, 4282. <https://doi.org/10.3390/app16094282>.
5. Karahan, H.; Alkaya, D. Integrating SVR Optimization and Machine Learning-Based Feature Importance for TBM Penetration Rate Prediction. *Appl. Sci.* **2026**, *16*, 355. <https://doi.org/10.3390/app16010355>.
6. Bulajić, B.; Lozančić, S.; Bajić, S.; Starčev-Ćurčin, A.; Šešlija, M.; Kovačević, M.; Hadzima-Nyarko, M. Horizontal PGA Estimates for Varying Deep Geological Conditions—A Case Study of Banja Luka. *Appl. Sci.* **2025**, *15*, 6712. <https://doi.org/10.3390/app15126712>.
7. Bulajić, B.; Lozančić, S.; Bajić, S.; Starčev-Ćurčin, A.; Šešlija, M.; Kovačević, M.; Hadzima-Nyarko, M. PGA Estimates for Vertical Ground Motion and Varying Deep Geology Site Surroundings—A Case Study of Banja Luka. *Appl. Sci.* **2025**, *15*, 6542. <https://doi.org/10.3390/app15126542>.
8. Aghdam, M.M.; Crowley, Q. Integrating Radon/Thoron and Gamma Radiation Exposure for a Realistic Estimation of Dose Arising from Building Materials. *Appl. Sci.* **2025**, *15*, 6470. <https://doi.org/10.3390/app15126470>.
9. Sansyzbekov, G.; Adoko, A.C.; George, P.M. Thermal Damage Characterization and Modeling in Granite Samples Subjected to Heat Treatment by Leveraging Machine Learning and Experimental Data. *Appl. Sci.* **2025**, *15*, 6328. <https://doi.org/10.3390/app15116328>.
10. Dai, W.; Zhou, Y.; Lei, H.; Ma, X.; Dang, J.; Miao, S.; Liu, S.; He, C. Pseudotachylite Formation in Brittle–Ductile Transition of the Anning River Fault Zone: Implications for Seismic Processes. *Appl. Sci.* **2025**, *15*, 5870. <https://doi.org/10.3390/app15115870>.
11. Zhang, R.; Guo, Z.; Gong, W.; Wan, Z. Field Test and Numerical Simulation Study of Bearing Characteristics of Combined Post-Grouted Piles for Railway Bridges. *Appl. Sci.* **2025**, *15*, 335. <https://doi.org/10.3390/app15010335>.
12. Liu, F.; Han, R.; Guo, Y.; Wang, M.; Tan, W. Element Migration of Mineralization-Alteration Zones and Its Geological Implication in the Beiya Porphyry–Skarn Deposit, Northwestern Yunnan, China. *Appl. Sci.* **2024**, *14*, 9653. <https://doi.org/10.3390/app14219653>.
13. Lei, H.; Liu, S.; Dai, W. Frictional Experiments on Granitic Faults: New Insights into Continental Earthquakes and Micromechanical Mechanisms. *Appl. Sci.* **2025**, *15*, 7207. <https://doi.org/10.3390/app15137207>.

References

1. Gao, W.; Song, D. Geotechnical Engineering and Intelligent Disaster Prediction. In *Proceedings of 2024 International Conference on Geotechnics and Hydraulic Structure*; Springer Press: Berlin/Heidelberg, Germany, 2025.
2. Chen, X.; Hong, C.; Su, D. Intelligent geotechnical engineering. *Chin. J. Geotech. Eng.* **2022**, *44*, 2151–2159. [CrossRef]
3. Baghbani, A.; Choudhury, T.; Costa, S.; Reiner, J. Application of artificial intelligence in geotechnical engineering: A state-of-the-art review. *Earth-Sci. Rev.* **2022**, *228*, 103991. [CrossRef]
4. He, M.; Han, Z.; Li, Z.; Niu, Y. Artificial Intelligence for Highway Geological Hazards Prediction and Prevention: Advances, Challenges and Future Directions. *Rock Mech. Bull.* 2026, *in press*. [CrossRef]
5. Abraham, K.; Abdelwahab, M.; Abo-Zahhad, M. Classification and detection of natural disasters using machine learning and deep learning techniques: A Review. *Earth Sci. Inform.* **2024**, *17*, 869–891. [CrossRef]
6. Ma, Z.; Mei, G. Deep learning for geological hazards analysis: Data, models, applications, and opportunities. *Earth-Sci. Rev.* **2021**, *223*, 103858. [CrossRef]
7. Jiang, S.; Huang, F.; Wang, Z.; Papaioannou, I. Intelligent Risk Assessment for Geological Disasters. *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* **2026**, *20*, 433–437. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

The Concentration Levels of Cr and Mo in the Soils of Beijing and Their Comparison with Other Capital Cities

Yonglong An ^{1,2}, Wenjun Cui ², Qingjie Gong ^{1,*}, Yuan Jiang ² and Yong Huang ²

¹ School of Earth Sciences and Resources, China University of Geosciences, Beijing 100083, China; anyl@email.cugb.edu.cn

² Beijing Institute of Ecological Geology, Beijing 100120, China; wenjun12342026@163.com (W.C.); yuan.12.3@163.com (Y.J.); hybeijing12342026@163.com (Y.H.)

* Correspondence: qjegong@cugb.edu.cn

Abstract

With the growing potential correlation between chromium (Cr), molybdenum (Mo), and human health, public attention to these elements has increased markedly. In Beijing, the capital of China, research on the concentration levels of Cr and Mo is urgently needed. In recent years, soil geochemical surveys have been conducted in the Beijing region, with a scale of 1:100,000 in mountainous areas (yielding 8575 data points) and 1:50,000 in plain areas (yielding 29,551 data points). These surveys provide a robust dataset for elucidating the concentrations and distribution patterns of soil elements. In surface soil samples from Beijing ($n = 38,126$), the mean concentrations of Cr and Mo were $55.7 \mu\text{g/g}$ and $0.67 \mu\text{g/g}$, respectively. Recently, scholars have proposed a 19-level fixed-value method for Mo concentrations. Based on this, the 19-level fixed-value method for Cr concentrations is proposed to facilitate cross-element comparisons. This method classifies the elemental concentrations from its detection limit to its cut-off grade of mineralization into 19 levels on 18 fixed values. With this method, the concentration levels of Cr and Mo in Beijing soils are both at the fifth level on their mean values, which belong to the background level. To compare the Cr and Mo concentrations in soils with other capital cities, statistical parameters of Cr from 20 capital cities and Mo from 11 capital cities were collected. The results show that the mean Cr concentration levels in the 20 cities range from 2 to 10, while the mean Mo concentration levels in the 11 cities range from 5 to 10. In these capital cities, the Cr level of Beijing is at the middle level, and the Mo level is at the low level, which indicates the cross-element comparison on the 19-level fixed-value method.

Keywords: chromium; molybdenum; element concentration classification; 19-level fixed-value method; soil geochemical survey; capital city

1. Introduction

Chromium (Cr) and molybdenum (Mo) are typical trace metal elements in the soil environment, and their concentrations serve as important indicators of soil quality [1]. Both deficiency and excess of these elements can disrupt soil ecological balance and adversely affect crop health [2]. Excessive Cr is highly toxic, leading to the deterioration of soil physicochemical properties, including the destruction of soil aggregate structure, reduction of organic matter activity, and inhibition of microbial metabolism. Although Mo is an essential element for plants, its deficiency reduces nitrogen fixation capacity in legumes, whereas excess Mo can induce phytotoxicity, manifested as leaf chlorosis and growth

inhibition [3,4]. Furthermore, when Cr and Mo accumulate and migrate in soils, they can be transferred to the human body through the food chain. Specifically, Cr poses carcinogenic and mutagenic risks, while excessive Mo may interfere with human copper metabolism [5,6].

The spatial distribution of soil Cr and Mo concentrations is governed by multiple natural factors, including parent material, weathering intensity, topographic features, climatic conditions, and redox conditions [7,8]. For example, in the Beijing plain area, the alluvial and proluvial deposits from the Yongding River and Chaobai River have developed neutral to alkaline calcareous soils, which facilitate the enrichment of Cr. Additionally, the weathering of Mo-rich parent rocks, such as black shales, can lead to significant Mo enrichment in soils [9]. Apart from natural sources, mining and mineral processing activities can cause locally elevated Mo concentrations, and anthropogenic activities represent another non-negligible contributing factor [10]. Previous studies have identified traffic emissions, coal combustion, and industrial processes as important anthropogenic sources of Cr and Mo accumulation in urban soils of Beijing. Furthermore, the application of chemical fertilizers and pesticides in agricultural activities can also influence the input of these trace metals [11–13]. Therefore, establishing the background value parameters of soil Cr and Mo and understanding their spatial distribution characteristics are of direct practical significance for guiding soil pollution remediation and ensuring the safety of agricultural products. Although numerous global capital cities have conducted surveys on soil Cr and Mo and obtained relatively rich soil background value datasets [14–16], relevant reports specifically on the soil background values for Beijing, the capital of China, remain limited. Existing studies have only addressed local areas within Beijing [17,18], which is insufficient to support the region-wide estimation and spatial mapping of Cr and Mo background values across the entire Beijing region.

This study firstly proposes a 19-level fixed-value method for Cr concentrations, and then develops geochemical maps of concentration levels of Cr and Mo in the surface soils from the plain and mountainous areas of Beijing. The screening and intervention thresholds for chromium specified in a national standard, as referenced in this study, both refer to total chromium in soil. Subsequently, comparative analyses were conducted to elucidate the spatial distribution characteristics of both elements. Finally, concentrations and their classified levels of Cr and Mo from 20 global capital cities were compiled to facilitate the comparisons among both regions and elements.

2. Materials and Methods

2.1. Geological Setting

Beijing is the capital of China. The study area spans approximately 170 km from north to south (latitudinal range: 39.4° N to 41.6° N) and 160 km from east to west (longitudinal range: 115.7° E to 117.4° E), covering a total area of 16,410 km². Situated in the mid-latitudes, Beijing exhibits a warm-temperate, semi-humid continental monsoon climate, characterized by four distinct seasons. Annual climate averages comprise a mean annual temperature of 12 °C, approximately 210 frost-free days, and 600 mm of annual precipitation [19]. The geomorphology of Beijing transitions from northwest to southeast, forming a sequence of middle mountains, low mountains, hills, alluvial–proluvial fan plains, proluvial–alluvial inclined plains, and alluvial plains. It can be broadly categorized into two major units: mountainous areas and plains. The mountainous regions are predominantly composed of rocks from various geological periods, hosting a wide variety of mineral resources. The top-ranked mineral types include coal, iron, limestone, marble, and granite, most of which have ceased to be mined. The plains are primarily formed by fluvial alluviation and proluviation, with alluvial–proluvial plains being the most extensively distributed [20,21].

2.2. Samples and Methods

From 2009 to 2013, soil survey projects focusing on agricultural land resources and urban geological environments were successively carried out in Beijing at a scale of 1:50,000, covering the entire plain area of approximately 6338 km². Soil samples were collected based on a 0.5 km × 0.5 km grid, with five sub-samples gathered within each grid and combined into one composite sample [21–23]. A total of 29,551 surface soil composite samples were collected in the plain area.

From 2018 to 2020, a soil survey project was conducted in the mountainous areas of Beijing at a scale of 1:100,000, covering an area of approximately 10,072 km². Surface soil samples were collected based on a 1 km × 1 km grid, with five sub-samples collected within each grid and combined into one composite sample [24]. A total of 8575 composite surface soil samples were obtained in the mountainous area.

In these survey projects for the plain and mountainous areas, sample locations were determined using GPS in combination with land use maps. The composite samples were collected in an “X” pattern from a depth of 20 cm. The total number of samples collected across these two areas amounts to 38,126, providing full coverage of the entire Beijing area.

After removing extraneous materials such as weeds, gravel, and insect shells, the soil samples were air-dried at room temperature. They were then preliminarily sieved and ground to pass a 200-mesh sieve using a pollution-free planetary ball mill for subsequent analysis. A total of 54 elemental parameters were analyzed by the National Research Center for Geoanalysis, China Geological Survey. The concentrations of Cr and Mo were determined using inductively coupled plasma mass spectrometry (ICP-MS), with detection limits for Cr and Mo were 2 µg/g and 0.1 µg/g, respectively. For samples with concentrations below the detection limits, a value of half the detection limit (Cr = 1 µg/g and Mo = 0.05 µg/g) was used. To ensure analytical accuracy, certified reference materials GSS-1 and GSS-4 from the National Standard Sample Center were used for calibration. Strict quality assurance and quality control (QA/QC) protocols were applied throughout the soil analysis process. These included the use of replicate samples, which accounted for 4% of the total samples. Testing precision was assessed using the relative deviation (RD) of these replicates, and the resulting logarithmic standard deviations (λ) are less than 10% or 17% on their different concentrations. Accuracy (relative error, RE) was validated via blind analyses of national primary certified reference materials (with a 9% insertion rate of such materials), yielding logarithmic concentration differences ($\Delta \lg c$) ranging from <0.05 to 0.1. All soil samples were subjected to triplicate testing, with spike recovery rates consistently ranging between 92% and 106%, thus confirming the reliability of the analytical methodology.

In this study, Excel 2021 was employed for descriptive statistical analysis of heavy metal concentrations, along with calculations for the 19-level classification system. Grapher 20 was used to generate histogram/boxplots and concentration classification maps. CorelDRAW 2020 facilitated the geolocation and concentration labeling of global capital cities, while GeoExpl 2013 was utilized for creating geochemical maps of Cr and Mo on their classified levels. To eliminate outliers, the Avg ± 3Std method was adopted (where Avg and Std are abbreviations for the mean and standard deviation, respectively). That is, when the lgCr and lgMo data fell outside the range of Avg ± 3Std, the outliers were repeatedly excluded until no further outliers were identified.

2.3. The 19-Level Fixed-Value Method of Element Concentrations

In the grading of element concentration levels, An et al. [25] proposed the fixed-value method for geochemical mapping, exemplified by chromium (Cr), where 25 fixed values were established, resulting in 26 levels and 5 types. The method’s core innovation lies

in its freedom from requirements for large sample sizes, enabling reliable classification even for individual specimens. Furthermore, its broad applicability enables cross-regional comparisons and multi-element assessments, addressing the traditional constraints in geochemical data interpretation. Building on this methodology, Xu et al. [26] developed an optimized method, with tin (Sn) used as a case example. By establishing 18 fixed values, they developed a more streamlined 19-level fixed-value method that divided the concentrations into 6 types, thereby demonstrating improved practicality. Subsequently, Li et al. [27], Jia et al. [28], and Zhu et al. [29] respectively developed a 19-level fixed-value method for Li, Mo, and Ni, based on comprehensive assessments of existing domestic and international element concentration classification systems. To ensure methodological consistency, this study draws on the original 26-level classification framework and adopts the 19-level fixed-value method to reclassify Cr concentrations.

This study adopts the 19-level fixed-value method to classify Cr concentrations, the core of which involves setting 18 fixed values to categorize Cr concentrations into 19 levels, which are subsequently divided into six types. In addition to fundamental parameters such as detection limits and boundary grades, the determination of these 18 fixed values was based primarily on Cr concentration data from stream sediments in the RGNR (Regional Geochemistry-National Reconnaissance) program and soil samples from the NMPRGS (National Multi-Purpose Regional Geochemical Survey) project, with a particular focus on their cumulative frequency distributions [30]. Furthermore, as chromium (Cr) is a heavy metal element [31,32] and China's national standard GB15618-2018 [33] specifies distinct screening and intervention values for Cr concentrations in agricultural land under different pH conditions, these values from GB15618-2018 were also referenced in determining the 18 fixed values. If an appropriate grading value is unavailable, a new one will be derived by interpolating within the range of previously determined grading values. The determination of the 18 fixed values is described detailed as the following:

The first value is set at 5 µg/g, grouping data below the detection limit into the same class. This value corresponds to the Cr detection limit in the RGNR and NMPRGS projects. The second value is set at 15 µg/g, corresponding to three times the detection limit. This is consistent with the 3% cumulative frequency value from the RGNR project's stream sediments.

The third value is set at 26 µg/g, approximating the 8% cumulative frequency value (21 µg/g) from the RGNR project's stream sediments. The fourth value is set at 45 µg/g, calculated as the average of the 25% cumulative frequency value from the RGNR project's stream sediments (37 µg/g) and the 25% cumulative frequency value from the NMPRGS project's soils (54 µg/g). The fifth value is set at 60 µg/g, calculated as the average of the 50% cumulative frequency value from the RGNR project's stream sediments (66 µg/g) and the 50% cumulative frequency value from the NMPRGS project's soils (55 µg/g). The sixth value is set at 75 µg/g, approximating the 75% cumulative frequency value (76 µg/g) from the NMPRGS project's soils. The seventh value is set at 95 µg/g, derived by logarithmic interpolation between the sixth and eighth values. The eighth value is proposed at 120 µg/g, approximating the 97.5% cumulative frequency value (119 µg/g) from the NMPRGS project's soils.

The ninth value is set at 150 µg/g, corresponding to the minimum screening value for Cr in the GB15618-2018 standard for land type "Other" when $\text{pH} \leq 6.5$. The tenth value is set at 200 µg/g, corresponding to the secondary screening value for Cr in the GB15618-2018 standard for land type "Other" when pH is between 6.5 and 7.5. The eleventh value is set at 250 µg/g, corresponding to the screening value for Cr in the GB15618-2018 standard for land type "Other" when $\text{pH} \geq 7.5$ (this value equals the screening value for "Paddy" land

type when $\text{pH} \leq 6.5$). The twelfth value is set at $300 \mu\text{g/g}$, corresponding to the secondary screening value for Cr for “Paddy” land type when pH is between 6.5 and 7.5.

The thirteenth value is set at $490 \mu\text{g/g}$, derived by logarithmic interpolation between the twelfth and fourteenth values. The fourteenth value is set at $800 \mu\text{g/g}$, corresponding to the minimum intervention value for Cr contamination when $\text{pH} \leq 5.5$. The fifteenth value is set at $1300 \mu\text{g/g}$, corresponding to the maximum intervention value for chromium contamination at $\text{pH} \geq 7.5$ according to GB15618-2018. The sixteenth and seventeenth thresholds are established at $2080 \mu\text{g/g}$ and $5300 \mu\text{g/g}$ respectively, derived through logarithmic interpolation between the fifteenth and eighteenth thresholds.

For the eighteenth value, $34,200 \mu\text{g/g}$ is selected, which corresponds to the cut-off grade value of 5% for Cr_2O_3 in the Cr deposits [34].

To achieve better visual differentiation in geochemical mapping, the 19 levels mentioned above are assigned to a six-color coding scheme: (1) low background areas, dominated by blue tones and encompassing the 1st to 5th levels, which correspond to Cr concentrations below $60 \mu\text{g/g}$; (2) high background areas, characterized by yellow tones and spanning the 6th to 9th levels, corresponding to Cr concentrations from 60 to $150 \mu\text{g/g}$; (3) possible screening-risk level areas, predominantly pink-toned, encompassing the 10th to 12th levels, indicating Cr concentrations ranging from 150 to $300 \mu\text{g/g}$; (4) surely screening-risk level areas, predominantly red-toned, comprising the 13th to 15th levels, associated with Cr concentrations between 300 and $1300 \mu\text{g/g}$; (5) surely intervention-risk level areas, predominantly gray-toned and including the 16th to 18th levels, with Cr concentrations ranging from 1300 to $34,200 \mu\text{g/g}$; (6) mineralization as an economic metal in rocks (the 19th level) are distinctly marked in black, with ultra-high concentrations exceeding $34,200 \mu\text{g/g}$. This chromatic progression—ranging from cool blues through warm yellows and pinks to alert reds, then neutral grays and culminating in mineral black—creates an intuitive visual language that simultaneously conveys environmental risk gradients and resource potential across spatial distributions.

For comparative purposes, the 18 previously reported fixed values for Mo, Sn, Li, Ni, and Co, along with a 19-level color gradation scheme, are also listed in Table 1.

Table 1. The information of the Cr classification on the 19-level fixed-value method along with those of Mo, Sn, Li, Ni, and Co.

Level No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	References
Cr ¹	<5	15	26	45	60	75	95	120	150	200	250	300	490	800	1300	2080	5300	34,200	$\geq 34,200$	This study
lgCr	0.70	1.18	1.42	1.65	1.78	1.88	1.98	2.08	2.18	2.30	2.40	2.48	2.69	2.90	3.11	3.32	3.72	4.53	-	
ΔlgCr		0.48	0.24	0.24	0.13	0.10	0.10	0.10	0.10	0.13	0.10	0.08	0.21	0.21	0.21	0.20	0.41	0.81	-	
Color																				[26]
R ²	0	9	28	54	128	255	255	255	255	255	255	255	254	249	216	191	135	100	0	
G ²	0	46	140	224	255	230	223	211	198	229	204	153	92	57	35	191	135	100	0	
B ²	170	216	248	255	224	153	127	76	25	255	255	255	18	9	0	191	135	100	0	
Mo ¹	0.3	0.37	0.45	0.55	0.68	1	1.43	2.07	3	3.8	4.8	6.1	9.1	20	100	144	208	300	≥ 300	[28]
Sn ¹	1	1.3	1.8	2.7	3.4	4.3	6.0	7.9	10	13	17	28	50	100	200	400	600	1000	≥ 1000	[26]
Li ¹	5	8	17	29	34	40	50	62	70	78	88	99	132	175	232	460	930	1858	≥ 1858	[27]
Ni ¹	2	6	10	18	25	32	39	48	60	70	100	190	245	495	1000	1260	1590	2000	≥ 2000	[29]
Co ¹	1	2.7	5.5	9	11.7	15	19	24	30	35	41	47	61	78	100	125	160	200	≥ 200	[35]

¹ The concentration unit is $\mu\text{g/g}$. ² The R, G, and B represent the red, green, and blue channels of a color image, respectively.

3. Results

3.1. Statistical Characteristics of Concentrations

A total of 38,126 surface soil samples were collected across the entire Beijing area, with 8575 from the mountainous areas and 29,551 from the plains. The histogram/boxplots were generated for the logarithms of Cr and Mo concentrations in soils in the entire Beijing area, the plain areas, and the mountain areas, respectively (Figure 1).

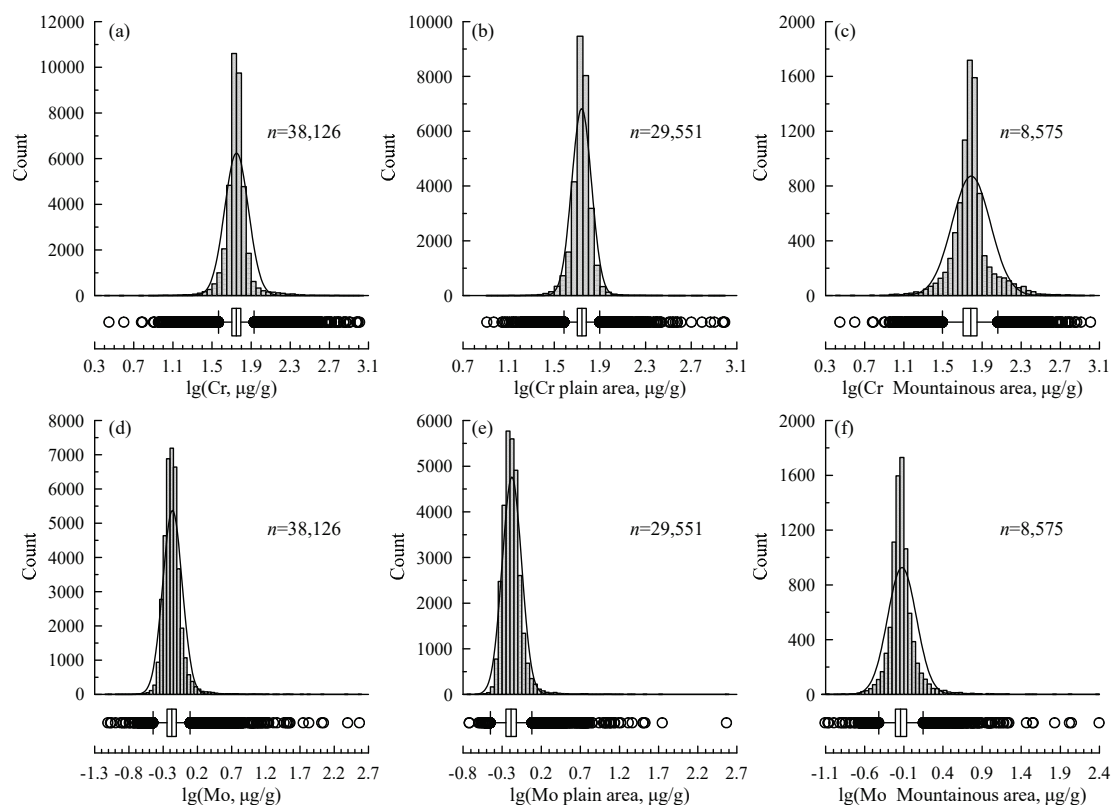


Figure 1. Histogram/boxplots of Cr and Mo in the surface soil of Beijing. (a) to (c) is for Cr in the entire area, plain area, and mountainous area, respectively and (d) to (f) is for Mo in the entire area, plain area, and mountainous area, respectively.

The histograms indicate that both Cr and Mo concentrations approximately follow a log-normal distribution, while the boxplots reveal the presence of a number of outliers.

With the log-transformed concentration data, the outliers were deleted on the mean $\pm 3 \times$ standard deviation (or Avg $\pm 3 \times$ Std) on the iterative method [26]. The statistical results of the original concentration data and the data without outliers are listed in Table 2.

In the entire Beijing area, the minimum concentrations of Cr and Mo are $2.8 \mu\text{g/g}$ and $0.05 \mu\text{g/g}$, the median values are $55.9 \mu\text{g/g}$ and $0.67 \mu\text{g/g}$, and the maximum values are $1021 \mu\text{g/g}$ and $369 \mu\text{g/g}$, respectively. The geometric means without outliers of Cr and Mo are $55.7 \mu\text{g/g}$ and $0.67 \mu\text{g/g}$, and the core concentration intervals for Cr and Mo are 50.4 to $62.2 \mu\text{g/g}$ and 0.57 to $0.78 \mu\text{g/g}$, respectively, which can be viewed as their background values in the Beijing area.

In the plain area, the minimum concentrations of Cr and Mo are $8.1 \mu\text{g/g}$ and $0.19 \mu\text{g/g}$, and the maximum values are $985 \mu\text{g/g}$ and $369 \mu\text{g/g}$, respectively. The geometric means without outliers of Cr and Mo are $55.1 \mu\text{g/g}$ and $0.65 \mu\text{g/g}$, respectively, while in the mountain area, the minimum concentrations of Cr and Mo are $2.8 \mu\text{g/g}$ and $0.05 \mu\text{g/g}$, and the maximum values are $1021 \mu\text{g/g}$ and $249 \mu\text{g/g}$, respectively. The geometric means without outliers of Cr and Mo are $59.9 \mu\text{g/g}$ and $0.72 \mu\text{g/g}$, respectively.

Table 2. Descriptive statistics of Cr and Mo concentrations in the surface soil of Beijing.

Parameters	Range	Minimum	Q1	Median	Q3	Maximum	Mean	Std. Dev.
Cr ¹	Entire area	2.8	50.4	55.9	62.2	1021	55.9	1.18
lgCr	<i>n</i> = 38,126	0.45	1.70	1.75	1.79	3.01	1.75	0.07
lgCr	Entire area	1.53	/	/	/	1.96	1.75	0.07
Cr ²	<i>n</i> = 35,799	34.0	/	/	/	91.4	55.7	1.2
Cr ¹	Plain areas	8.1	50.3	55.2	60.3	985	55.2	1.16
lgCr	<i>n</i> = 29,551	0.91	1.70	1.74	1.78	2.99	1.74	0.06
lgCr	<i>n</i> = 28,520	35.5	/	/	/	85.7	1.74	0.06
Cr ²	Plain areas	1.55	/	/	/	1.93	55.1	1.2
Cr ¹	Mountain areas	2.8	50.9	60.8	70.5	1021	60.8	1.36
lgCr	<i>n</i> = 8575	0.45	1.71	1.78	1.85	3.01	1.78	0.13
lgCr	<i>n</i> = 8023	1.37	/	/	/	2.18	1.78	0.13
Cr ²	Mountain area	23.7	/	/	/	151.3	59.9	1.4
Mo ¹	Entire area	0.05	0.57	0.67	0.78	369	0.67	1.26
lgMo	<i>n</i> = 38,126	−1.11	−0.24	−0.17	−0.11	2.57	−0.17	0.10
lgMo	<i>n</i> = 36,887	−0.48	/	/	/	0.12	−0.18	0.10
Mo ²	Entire area	0.3	/	/	/	1.3	0.67	1.26
Mo ¹	Plain areas	0.19	0.56	0.65	0.76	369	0.65	1.25
lgMo	<i>n</i> = 29,551	−0.72	−0.25	−0.19	−0.12	2.58	−0.19	0.10
lgMo	<i>n</i> = 28,876	−0.47	/	/	/	0.10	−0.19	0.10
Mo ²	Plain areas	0.3	/	/	/	1.3	0.65	1.25
Mo ¹	Mountain areas	0.05	0.62	0.72	0.86	249	−0.14	0.12
lgMo	<i>n</i> = 8575	−1.11	−0.21	−0.14	−0.07	2.40	0.72	1.31
lgMo	<i>n</i> = 8113	−0.49	/	/	/	0.21	−0.14	0.12
Mo ²	Mountain areas	0.3	/	/	/	1.6	0.72	1.31

¹ All concentration data in µg/g. ² Concentration data without outliers in µg/g.

3.2. Statistical Characteristics of Concentration Levels

The concentration data of Cr and Mo of the 38,126 surface soil samples from Beijing were transformed to their concentration levels on the 19-level fixed-value method and their statistical parameters are listed in Table 3. Here the method on Cr is the aforementioned method and the method on Mo is after Jia et al. [28]. The histogram/boxplots of the Cr and Mo concentration levels in the entire Beijing area, the plain areas, and the mountain areas are illustrated in Figure 2.

In the entire Beijing area, the Cr concentration levels range from 1 to 15 (Table 3) and are mainly clustered between 5 and 6 with the median value of 5 (Figure 2). The mean of the Cr levels is ca. 5.3 and is set at 5 because of the integer property of the concentration level. The Mo concentration levels range from 1 to 19 (with level 17 missing) and are mainly clustered between 5 and 6 with the same value of 5 on median or mean.

Table 3. Classification level statistics of Cr and Mo using the 19-level fixed-value method in the surface soil of Beijing.

Element	Range	Minimum	Q1	Median	Q3	Maximum	Mean ¹	Mean ²
Cr	Entire area	1	5	5	6	15	5.3	5
	Plain areas	2	5	5	6	15	5.2	5
	Mountain areas	1	5	6	6	15	5.8	6
Mo	Entire area	1	5	5	6	19	5.4	5
	Plain areas	1	5	5	6	19	5.3	5
	Mountain areas	1	5	6	6	18	5.7	6

¹ Mean presented with a decimal although the level values are all integers. ² Mean presented as an integer.

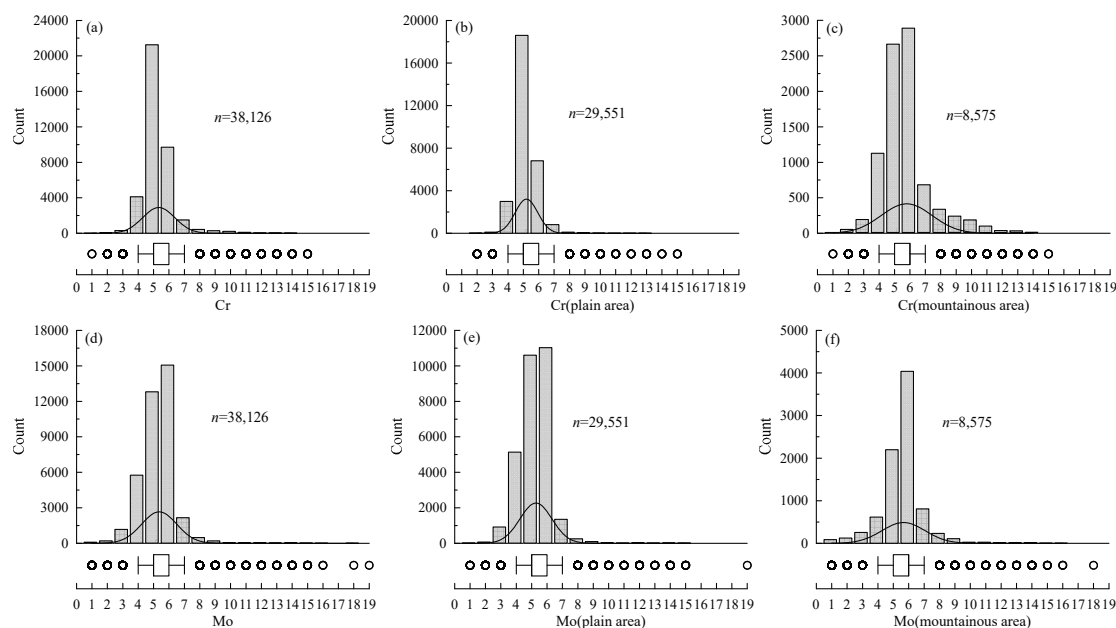


Figure 2. Histogram/boxplot diagrams of Cr and Mo concentration classification levels in the surface soil of Beijing. (a) to (c) is for Cr in the entire area, plain area, and mountainous area, respectively and (d) to (f) is for Mo in the entire area, plain area, and mountainous area, respectively.

In the plain area, the Cr concentration levels range from 2 to 15 and are mainly clustered between 5 and 6 with the same value of 5 on median or mean. The Mo concentration levels range from 1 to 19 and are mainly between 5 and 6 with the same value of 5 on median or mean. While in the mountain area, the Cr concentration levels range from 1 to 15 and are mainly between 5 and 6 with the same value of 6 on median or mean. The Mo concentration levels range from 1 to 18 and are mainly between 5 and 6 with the same value of 6 on median or mean. In comparison, the mean concentration levels of Cr and Mo in the mountain soils are all at the sixth level, which is higher than the fifth level of those in the plain soils. The former corresponds to the high background level and the latter to the low background level both for Cr and Mo.

In summary, the median and mean concentration levels of Cr and Mo are all in the fifth level, which corresponds to the low background level in the entire Beijing area.

3.3. Geochemical Maps of Concentration Levels of Cr and Mo in Beijing

Using the 19-level fixed-value method, the concentrations of Cr and Mo in the soil were converted into their concentration levels to contour the geochemical maps. Spatial interpolation was conducted using GeoExpl 2013 software [36] with a grid resolution of 0.5 km. The nearest neighbor interpolation method (NNI) was applied, utilizing a circular search neighborhood with a 2.5 km radius. Benefiting from the grid-based sampling scheme, the NNI employed in this study assigns values solely from actual sampling points within a circular search neighborhood, without generating any values beyond the observed data range. This NNI method is more suitable in this study for the grid sampling scheme than the commonly used methods such as the multifractal interpolation and IDW [37,38]. The resulting contour lines were mapped via GeoExpl 2013 to illustrate the concentration levels of Cr and Mo throughout the Beijing region (Figure 3).

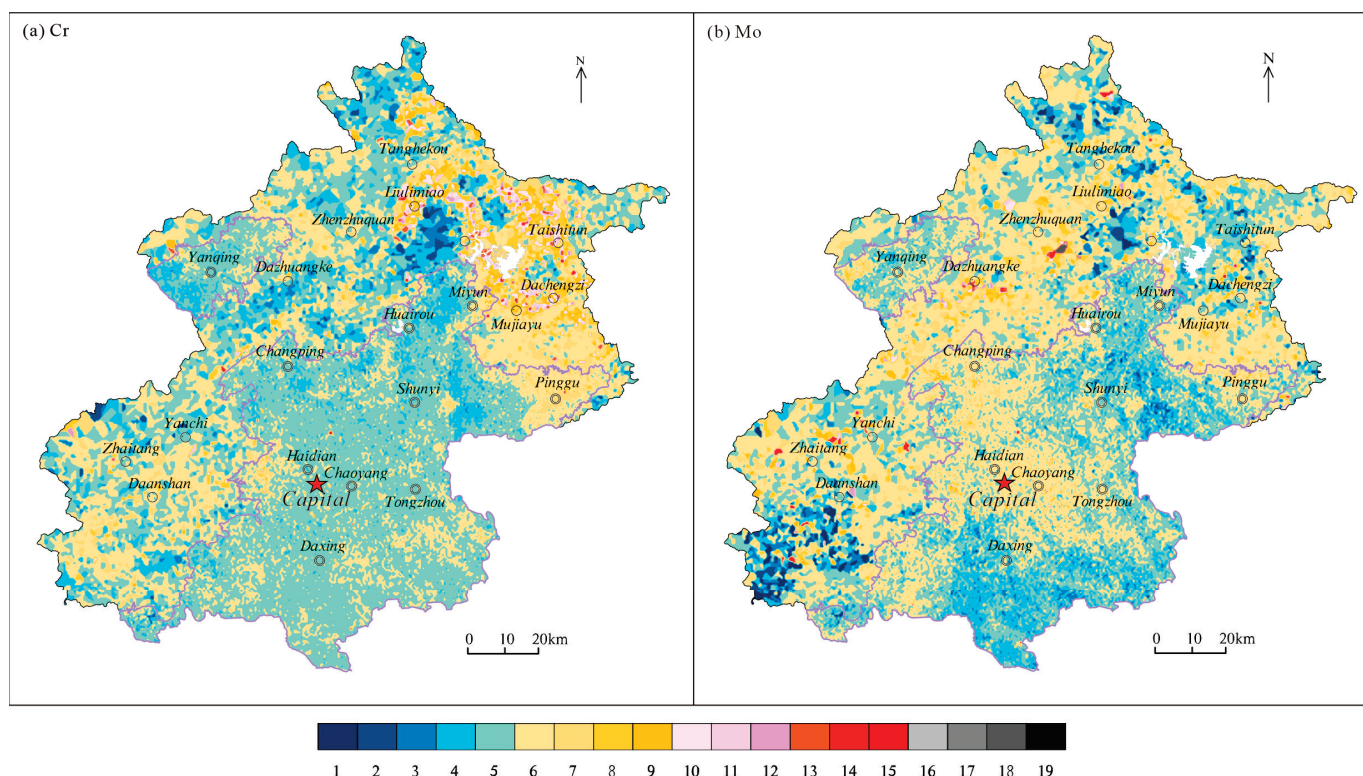


Figure 3. Geochemical maps of Cr and Mo concentration classification levels in the surface soil of Beijing. The pink line represents the boundary between plain and mountainous areas of Beijing.

Figure 3 shows that the concentration levels of Cr and Mo range from 1 to 15 and 1 to 19, respectively. In order to keep the consistency of the 19-level legend among elements, the legends of 19 levels are all illustrated, although only 15 levels are used for Cr in Figure 3a.

4. Discussion

4.1. Comparison of Concentration Levels of Cr and Mo in Beijing

With respect to Cr (Figure 3a), the Cr concentration levels are mainly in blue and yellow tones, with the dominant values ranging from levels of 3 to 8. This indicates the background levels in the entire Beijing area. In the plain areas, the map is predominantly characterized by blue tones indicating low background areas of Cr, while in the mountain areas located in the western, northern, and eastern Beijing area, the maps are predominantly yellow-toned, corresponding to high background.

Except the dominant blue and yellow tones, a few areas, such as the northern parts of the Huairou, Pinggu, and Miyun districts, are in pink tones with sporadic red spots. These pink areas indicate the possible screening-risk level of Cr with sporadic surely screening-risk levels.

With respect to Mo (Figure 3b), the Mo concentration levels are mainly in blue and yellow tones with the dominant values ranging from levels of 3 to 8. This is similar to that of Cr in the entire Beijing area. In the plain areas, the eastern and southern parts are in blue tones, indicating low background level of Mo, while the central and western parts are in yellow tones, reflecting higher background levels. While in the mountain areas, the maps are predominantly yellow-toned, corresponding to high background, except the northwest and the north parts in blue tones or in low background level.

Except the dominant blue and yellow tones, a few areas such as the Yanchi-Zhaitang (Mentougou District) and Dazhuangke (Yangqing District) are in red tones sporadically corresponding to the high anomaly level of Mo [27,39]. Furthermore, an area is located

in grey to black colors with levels of 16 to 19 in the southeastern Zhenzhuquan. This indicates a mineralization level of Mo [27], which is validated on the known molybdenum deposits [40,41].

In comparison, the concentration levels of Cr and Mo both dominantly range from 3 to 8 in the Beijing area, which indicates the background levels of Cr and Mo. The 19-level fixed-value method can effectively facilitate cross-element comparisons regardless of variations in detection limits and mineralization thresholds.

4.2. Comparison of Concentration Levels of Cr and Mo in Capital Cities

Based on the above analysis of soil Cr and Mo data in Beijing, the statistical parameters are consistent with the results expressed on spatial maps. This reflects the theoretically consistency information derived from statistical data or spatial maps. Therefore, statistical parameters are selected for comparison among capital cities of the world.

A systematic collection of Cr and Mo concentrations in surface soils of capital cities were conducted and compiled in a dataset including 20 capitals for Cr and 11 for Mo. The statistical parameters (including maximum, minimum, mean, and median values) are listed in Table 4 for Cr and Table 5 for Mo. The mean concentrations of Cr and Mo were converted into their concentration levels on the 19-level fixed-value method. The converted levels are also listed in Tables 4 and 5 and illustrated in Figure 4.

Table 4. Statistical parameters of concentration and concentration levels of Cr in soils from 20 capital cities.

Continents	Nation	Capital	Soil Sample Amount	Minimum $\mu\text{g/g}$	Median $\mu\text{g/g}$	Maximum $\mu\text{g/g}$	Mean $\mu\text{g/g}$	Concentration Level	References
Asia	China	Beijing	38,126	2.8	55.9	1021	55.7	5	This study
	Thailand	Bangkok	30	4.3	25.4	57.4	26.4	4	[42]
	Mongolia	Ulan Bator	90	8.2	—	230.0	32.7	4	[43]
	Iran	Tehran (District 6)	41	—	—	65	37.98	4	[44]
	Japan	Tokyo ¹	273	4	56	156	56.28	5	[45]
Africa	Cape Verde	Santiago ²	249	5.00	120.00	406.00	135.18	9	[46]
North America	Mexico	Mexico City	135	50	116	265	117	8	[47]
	Canada	Ottawa	50	28.8	43.4	74.5	44.8	4	[48]
Europe	United Kingdom	London	6801	15	72	2094	78.2	7	[49]
	Ireland	Dublin	1058	4.24	44.3	262	44.2	4	[50]
	Czechia	Prague	194	12.1	25.6	106	27.5	4	[51]
	Greece	Athens	238	43	141	1586	163	10	[52]
	Serbia	Belgrade	63	16.66	—	121.58	59.70	5	[53]
	Norway	Oslo ³	297	2.85	28.5	224	32.5	4	[54]
	Germany	Berlin	2182	—	25.1	1840	35.0	4	[55]
	Poland	Warsaw	2764	<1	7	785	9	2	[56]
	Estonia	Tallinn	198	5.7	33	300	39.9	4	[57]
	Slovenia	Ljubljana	477	28	85	1837	85	7	[58]
	Spain	Madrid	1730	5.00	28.00	430	36.70	4	[59]
	Lithuania	Vilnius	3179	32.8	—	33.8	33.3	4	[60]

¹ represented by the Kanto Plain; ² including the capital, Praia; ³ represented by Greater Oslo Region.

The soil data from these 20 capital cities are classified into three categories according to their sampling strategies: First, grid-based sampling with full-area coverage (Beijing, London, Dublin, Warsaw, Athens, Berlin, Tallinn, Ljubljana, Madrid, and Vilnius). These surveys are characterized by regular grid or systematic regional sampling with relatively high sampling densities (London: 4 points/km²; Warsaw: 4 points/km² in the city center and 1 point/km² in the peripheral areas; Athens: 1 point/km²), which adequately reflect the overall concentration levels of soil elements. Second, representative sampling with full-area coverage (Santiago Island, Tokyo, Oslo, and Mexico City). These surveys employed stratified sampling based on major soil types, geological parent materials, geomorphological units, or land use patterns. This approach effectively characterizes the regional geochemical background as well as the elemental signatures shaped by both natu-

ral conditions and urban activities. Third, typical unit sampling with full-area coverage (Bangkok, Ulan Bator, Tehran, Ottawa, Prague, and Belgrade). Sampling was designed around typical environmental units such as city axes, administrative grids, residential communities, urban–rural gradients, and industrial, transportation, and greenland areas. Some studies employed composite samples formed by mixing multiple sub-samples to reduce the randomness of individual sampling points and to improve the representativeness of the average conditions within each grid or functional zone. In summary, the soil data from the 20 selected capitals can serve as a robust basis for the comparison of Cr, Mo, and other elemental concentration levels among capital cities of different countries.

Table 5. Statistical parameters of concentration and concentration levels of Mo in soils from 11 capital cities.

Continents	Nation	Capital	Soil sample Amount	Minimum $\mu\text{g/g}$	Median $\mu\text{g/g}$	Maximum $\mu\text{g/g}$	Mean $\mu\text{g/g}$	Concentration Level	References
Asia	China	Beijing	38,126	0.05	0.67	369	0.67	5	This study
	Mongolia	Ulan Bator	90	0.5	—	8.30	1.8	8	[43]
	Japan	Tokyo ¹	273	0.11	1.09	2.96	1.211	7	[45]
North America	Canada	Ottawa	50	0.30	0.60	1.70	0.64	5	[48]
Europe	United Kingdom	London	6801	<0.2	1.5	561.2	2.0	8	[49]
	Ireland	Dublin	1058	<1.5	1.5	14	1.46	8	[50]
	Czechia	Prague	194	0.470	1.29	5.80	1.50	8	[51]
	Norway	Oslo ²	297	<1	1.31	45.2	1.79	8	[54]
	Estonia	Tallinn	198	<1	1.5	30	1.8	8	[57]
	Spain	Madrid	1730	1.00	1.00	38	3.25	10	[59]
	Lithuania	Vilnius	3179	0.83	—	1.01	0.92	6	[60]

¹ represented by the Kanto Plain; ² represented by Greater Oslo Region.

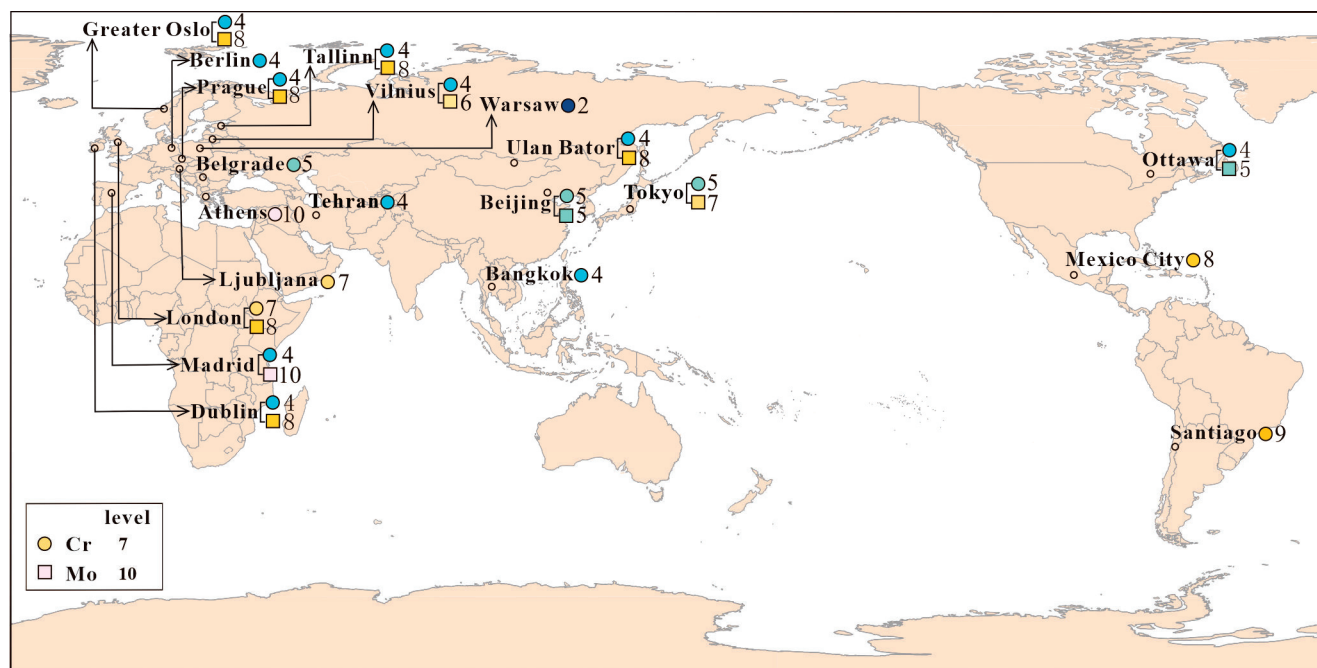


Figure 4. The concentration levels of Cr and Mo in soils in different capital cities.

The concentration means of Cr in surface soils in the 20 capital cities range from 9 $\mu\text{g/g}$ to 163 $\mu\text{g/g}$, corresponding to levels from 2 to 10 (Table 4), while the concentration means of Mo range from 0.64 $\mu\text{g/g}$ to 3.25 $\mu\text{g/g}$, corresponding to levels from 5 to 10 (Table 5). Except Athens of Greece with a Cr level of 10 and Madrid of Spain with a Mo level of 10, the levels of Cr and Mo in other cities are all at background level.

The concentration levels of Cr and Mo in Beijing are both at the fifth level, representing the highest level within the low background areas. Specifically, the Cr level of Beijing is at

the middle level among the 20 capital cities. However, the Mo level of Beijing is at the low level among the 11 capital cities.

Comparing the concentration levels of Cr and Mo in Tables 4 and 5, we find that the Cr levels in the five capitals in Asia range from 4 to 5 and are the low background level. However, the Mo levels in the seven capitals in Europe range from 6 to 10 and are at the high background level.

5. Conclusions

- (1) The 18 fixed values of Cr concentrations are suggested for the 19-level fixed-value method to convert Cr concentrations to its concentration levels. The 19 concentration levels can be categorized into six types as low background level, high background level, possible screening-risk level, surely screening-risk level, surely intervention-risk level, and mineralization level.
- (2) The concentrations of Cr and Mo in the surface soils of Beijing range mainly from 50.4 to 62.2 µg/g with a mean value of 55.7 µg/g and from 0.57 to 0.78 µg/g with a mean value of 0.67 µg/g, respectively. The concentration levels of Cr and Mo range mainly both from 5 to 6 with a mean value of 5, indicating their background levels.
- (3) The concentration levels of Cr in 20 capital cities range from 2 to 10 and the fifth level of Beijing is at the middle level. The concentration levels of Mo in 11 capital cities are from 5 to 10 and the fifth level of Beijing is at the low level. These results indicate the cross-element comparison on the 19-level fixed-value method.

Author Contributions: Conceptualization, Data curation, Writing—original draft, Y.A.; Conceptualization, Data curation, Formal analysis, W.C.; Conceptualization, Methodology, Writing—review and editing, Q.G.; Data curation, Formal analysis, Y.J.; Methodology, Formal analysis, Data curation, Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by the Beijing Institute of Geology and Mineral Exploration (11000025T000003325998 and 11000022T000000439575).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be made available on request.

Acknowledgments: We greatly appreciate the comments from the anonymous reviewers and editors for their valuable suggestions to improve the quality of this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Shahid, M.; Shamshad, S.; Rafiq, M.; Khalid, S.; Bibi, I.; Niazi, N.K.; Dumat, C.; Rashid, M.I. Chromium speciation, bioavailability, uptake, toxicity and detoxification in soil-plant system: A review. *Chemosphere* **2017**, *178*, 513–533. [CrossRef] [PubMed]
2. Yang, D.; Wang, L.; Bai, S. Enhancement of alfalfa growth resistance by arbuscular mycorrhiza and earthworm in molybdenum-contaminated soils: From the perspective of soil nutrient turnover. *Environ. Res.* **2025**, *267*, 120714. [CrossRef] [PubMed]
3. Liu, Y. Effects of Organic Fertilizer on Microbial Community Structure in Hexavalent Chromium Contaminated Soil and Alleviation of Toxicity to *Amaranthus tricolor* L. Master's Thesis, Guangdong Ocean University, Zhanjiang, China, 2019.
4. Kanazirska, V. Physiological Changes Caused by Molybdenum Deficiency or Excess. *Plant Prot.* **2022**, *6*, 1–8.
5. Baloch, M.Y.J.; Zhang, W.; Alqurashi, M.; Rashid, A.; Aslam, A. Hydrogeochemical Mechanism Associated with Land Use Land Cover Indices Using Geospatial, Remote Sensing Techniques, and Health Risks Model. *Sustainability* **2022**, *14*, 16837. [CrossRef]
6. Costello, R.B.; Dwyer, J.T.; Bailey, R.L. Chromium supplements for glycemic control in type 2 diabetes: Limited evidence of effectiveness. *Nutr. Rev.* **2016**, *74*, 455–468. [CrossRef] [PubMed]
7. Nigar, A.; Zhang, W. Assessment of Landcover Impacts on the Groundwater Quality Using Hydrogeochemical and Geospatial Techniques. *Sustainability* **2023**, *15*, 3566. [CrossRef]

8. Gong, Q.; Deng, J.; Jia, Y.; Tong, Y.; Liu, N. Empirical equations to describe trace element behaviors due to rock weathering in China. *J. Geochem. Explor.* **2015**, *152*, 110–117. [CrossRef]
9. Xie, S.; Peng, B.; Jiang, L.; Xiao, J.; Tang, X.; Li, D. Geochemical characteristics of soils derived from Lower Cambrian black shales in central Hunan. *Acta Pedol. Sin.* **2009**, *46*, 557–570. (In Chinese with English Abstract) [CrossRef]
10. Zhang, L.; Zhu, Y.; Zhang, Y.; Zhong, J.; Li, J.; Yang, S.; Ta, W.; Zhang, Y. Characteristics, source analysis, and health risk assessment of potentially toxic elements pollution in soil of dense molybdenum tailing ponds area in central China. *Environ. Geochem. Health* **2024**, *46*, 129. [CrossRef] [PubMed]
11. Sun, C.; Zhao, W.; Zhang, Q.; Yu, X.; Zheng, X.; Zhao, J.; Lv, M. Spatial Distribution, Sources Apportionment and Health Risk of Metals in Topsoil in Beijing, China. *Int. J. Environ. Res. Public Health* **2016**, *13*, 727. [CrossRef] [PubMed]
12. Su, K.; Zhang, Q.; Chen, A.; Wang, X.; Zhan, L.; Rao, Q.; Wang, J.; Yang, H. Heavy metals concentrations in commercial organic fertilizers and the potential risk of fertilization into soils. *Sci. Rep.* **2025**, *15*, 1230. [CrossRef] [PubMed]
13. Melo Júnior, H. do. N.; Paula Filho, F.J.d.P.; Coelho Menezes, J.M.; Melo Coutinho, H.D.; Pereira Teixeira, R.N.; Sobral Melo, H.V. Application of sediment quality guidelines and geochemical indices to evaluate the impact of trace metals on sediments from a fish farming area in the Rosário reservoir, Brazil. *Appl. Geochem.* **2026**, *199*, 106726. [CrossRef]
14. De Miguel, E.; De Grado, M.J.; Llamas, J.; Martín-Dorado, A.; Mazadiego, L. The overlooked contribution of compost application to the trace element load in the urban soil of Madrid (Spain). *Sci. Total Environ.* **1998**, *215*, 113–122. [CrossRef]
15. Pouyat, R.V.; Szlavecz, K.; Yesilonis, I.D.; Wong, C.P.; Murawski, L.; Marra, P.; Casey, R.E.; Lev, S. Multi-scale assessment of metal contamination in residential soil and soil fauna: A case study in the Baltimore–Washington metropolitan region, USA. *Landsc. Urban Plan.* **2015**, *142*, 7–17. [CrossRef]
16. Said, I.; Salman, S.A.E.-R.; Samy, Y.; Awad, S.A.; Melegy, A.; Hursthouse, A.S. Environmental factors controlling potentially toxic element behaviour in urban soils, El Tebbin, Egypt. *Environ. Monit. Assess.* **2019**, *191*, 1–15. [CrossRef] [PubMed]
17. Duan, X.; Li, P.; Huang, Y.; Lin, Y.; Yuan, G.; Luo, X. Geochemical Characteristics and Ecological Risk Assessment of Heavy Metals in Agricultural Soils of Miyun District, Beijing. *Geoscience* **2018**, *1*, 95–104. (In Chinese with English abstract) [CrossRef]
18. Yu, H. Study on Geochemical Characteristics of Heavy Metals in Surface Soils of Shunyi District, Beijing. Master's Thesis, China University of Geosciences (Beijing), Beijing, China, 2019.
19. Hashengerile. Studies on Construction of Urban Greening and Change of Eco-Environment in Beijing. Ph.D. Thesis, Beijing Forestry University, Beijing, China, 2006. (In Chinese with English Abstract)
20. Huo, Y.; Yang, Z. *Natural geography in Beijing*; Beijing Normal University Publishing House: Beijing, China, 1989. (In Chinese with English Abstract)
21. Li, J.; Zhang, C.; Zhu, D.; Li, H. Research on the relief zonal approach of Beijing based on spatial technology. *J. Agric. Sci. Technol.* **2007**, *2*, 126–129, (In Chinese with English Abstract). [CrossRef]
22. Duan, X.; Yu, H.; Ye, T.; Huang, Y.; Li, J.; Yuan, G.; Albanese, S. Geostatistical mapping and quantitative source apportionment of potentially toxic elements in top-and sub-soils: A case of suburban area in Beijing, China. *Ecol. Indic.* **2020**, *112*, 106085. [CrossRef]
23. Huang, Y.; Duan, X.; Yuan, G.; Li, H.; Zhang, Q. Geochemistry and source identification of heavy metals in the top and subsoil of Yanqing District in Beijing. *Geoscience* **2022**, *36*, 634–644. (In Chinese with English Abstract) [CrossRef]
24. Zhang, Q.; Feng, H.; Sun, Z.; Li, Q.; Jiang, T.; Jia, L. Study on background value of surface soil Se content in ecological conservation area of Beijing. *Urban Geol.* **2024**, *19*, 475–482. (In Chinese with English Abstract) [CrossRef]
25. An, Y.; Yan, T.; Gong, Q.; Wang, X.; Huang, Y.; Zhang, B.; Yin, Z.; Zhao, X.; Liu, N. Chromium (Cr) geochemical mapping based on fixed-values' method: Case studies in China. *Appl. Geochem.* **2022**, *136*, 105168. [CrossRef]
26. Xu, S.; Li, J.; Zhang, X.; Huang, Z.; Huang, Y.; Long, Y.; Xu, Y.; Song, X.; Chen, Z.; Li, Y. Tin (Sn) geochemical mapping based on a fixed-value method: A case illustration in Gejiu area, Southwest China. *Appl. Sci.* **2024**, *14*, 1765. [CrossRef]
27. Li, P.; Gong, Q.; Chen, S.; Li, P.; Li, J.; Wu, X.; Li, X.; Wang, X.; Liu, N. Regional geochemical characteristics of lithium in the Mufushan area, South China. *Appl. Sci.* **2024**, *14*, 1978. [CrossRef]
28. Jia, G.; Gu, W.; Gong, Q.; Xu, S.; Liu, Y.; Lv, Z. A 19-level fixed-value method to classify the Mo concentrations in Jianshui area of Yunnan province, China. *Appl. Geochem.* **2025**, *180*, 106289. [CrossRef]
29. Zhu, X.; Li, P.; Gong, Q.; Gu, W.; Xu, S.; Yan, T. Geochemical survey in Mojiang area of Yunnan Province, China: Geochemical map and geochemical anomaly map. *Appl. Sci.* **2025**, *15*, 2592. [CrossRef]
30. Hou, Q.; Yang, Z.; Yu, T.; Xia, X.; Cheng, H.; Zhou, G. *Soil Geochemical Parameters in China*; Geological Publishing House: Beijing, China, 2020. (In Chinese with English Abstract)
31. Fendorf, S.E. Surface reactions of chromium in soils and waters. *Geoderma* **1995**, *67*, 55–71. [CrossRef]
32. Frei, R.; Gaucher, C.; Poulton, S.W.; Canfield, D.E. Fluctuations in Precambrian atmospheric oxygenation recorded by chromium isotopes. *Nature* **2009**, *461*, 250–253. [CrossRef] [PubMed]
33. GB15618-2018; Soil Environmental Quality and Agricultural Land Soil Pollution Risk Control Standards (Trial). China Environment Press: Beijing, China, 2018. (In Chinese with English Abstract)

34. DZ/T 0200-2020; Mineral Geological Exploration Specification Iron, Manganese, Chromium. Ministry of Land and Resources: Beijing, China, 2020. (In Chinese with English Abstract)
35. Lv, Z.; Liu, Y.; Gong, Q.; Chen, H.; Yan, T. Geochemical evaluation of resource potential in mafic-ultramafic intrusions: A case study on Ni and Co in the Hongshishan complex of Gansu Province, China. *J. Geochem. Explor.* **2026**, *283*, 107986. [CrossRef]
36. Xiang, Y.; Mu, X.; Ren, T.; Ma, Z.; Liu, R.; Gong, Q.; Wang, M.; Gong, J.; Yang, W.; Yang, Y. *Application of Geochemical Survey Data on Potential Evaluation of Mineral Resources in China*; Geological Publishing House: Beijing, China, 2018. (In Chinese with English Abstract)
37. Cheng, Z.; Xie, X.; Yao, W.; Feng, J.; Zhang, Q.; Fang, J. Multi-element geochemical mapping in Southern China. *J. Geochem. Explor.* **2014**, *139*, 183–192. [CrossRef] [PubMed]
38. Lima, A.; De Vivo, B.; Cicchella, D.; Cortini, M.; Albanese, S. Multifractal IDW interpolation and fractal filtering method in environmental studies: An application on regional stream sediments of (Italy), Campania region. *Appl. Geochem.* **2003**, *18*, 1853–1865. [CrossRef]
39. Li, Y.; Zhang, H.; Wang, J. Re-Os dating of molybdenite-bearing whole-rock samples: Application to the Dazhuangke molybdenum deposit, Beijing. *Geol. China* **2025**, *52*, 1–12, (In Chinese with English Abstract).
40. Tang, X.; Liu, J.; Chen, Y.; Zhao, B. Resources characteristics of molybdenum deposits and their regional metallogeny regularity in North China. *Geol. China* **2022**, *49*, 455–470, (In Chinese with English Abstract). [CrossRef]
41. Mao, J.; Pirajno, F.; Lehmann, B. A Preliminary Review of Metallogenic Regularity of Molybdenum Deposits in China. *Acta Geol.* **2015**, *89*, 972–1001. [CrossRef]
42. Wilcke, W.; Müller, S.; Kanchanakool, N.; Zech, W. Urban soil contamination in Bangkok: Heavy metal and aluminium partitioning in topsoils. *Geoderma* **1998**, *86*, 211–228. [CrossRef]
43. Kasimov, N.; Kosheleva, N.; Gunin, P.; Korlyakov, I.; Sorokina, O.; Timofeev, I. State of the environment of urban and mining areas in the Selenga Transboundary River Basin (Mongolia Russia). *Environ. Earth Sci.* **2016**, *75*, 1–20. [CrossRef]
44. Sobhan Ardakani, S.; Karbassi, A.R. Assessment of Lithogenic and Non-Lithogenic Shares of Metals in Urban Soils of Tehran (District 6). *Pollution* **2018**, *4*, 583–592. [CrossRef]
45. Terashima, S.; Ohta, A.; Imai, N.; Okai, T.; Mikoshiba, M.; Taniguchi, M. Geochemistry of Toxic Trace Elements (As, Sb, Pb, Cr, Mo, Bi, Cd, Tl) in the Soils from the Kanto District, Japan: Preliminary Study for the Soil Geochemical Mapping (Part 3). *Bull. Geol. Surv. Jpn.* **2002**, *53*, 749–774. [CrossRef]
46. Cabral Pinto, M.M.S.; Ferreira da Silva, E.; Silva, M.M.V.G.; Melo-Gonçalves, P. Heavy Metals of Santiago Island (Cape Verde) Top Soils: Estimated Background Value Maps and Environmental Risk Assessment. *J. Afr. Earth Sci.* **2015**, *101*, 162–176. [CrossRef]
47. Morton-Bermea, O.; Hernández-Álvarez, E.; González-Hernández, G.; Romero, F.; Lozano, R.; Beramendi-Orosco, L.E. Assessment of heavy metal pollution in urban topsoils from the metropolitan area of Mexico City. *J. Geochem. Explor.* **2009**, *101*, 218–224. [CrossRef]
48. Rasmussen, P.E.; Subramanian, K.; Jessiman, B. A multi-element profile of house dust in relation to exterior dust and soils in the city of Ottawa, Canada. *Sci. Total Environ.* **2001**, *267*, 125–140. [CrossRef] [PubMed]
49. British Geological Survey. London Earth Geochemical Dataset. Available online: <https://www.bgs.ac.uk/datasets/london-earth/> (accessed on 1 June 2026).
50. Glennon, M.; Scanlon, R.P.; O'Connor, P.J.; Finne, T.E.; Andersson, M.; Eggen, O.; Jensen, H.K.B.; Ottesen, R.T. *SURGE Project: Geochemical Baseline for Heavy Metals and Organic Pollutants in Topsoils in the Greater Dublin Area*; Geological Survey of Ireland: Dublin, Ireland, 2012.
51. Poňavič, M.; Wittlingerová, Z.; Čoupek, P.; Buda, J. Soil Geochemical Mapping of the Central Part of Prague, Czech Republic. *J. Geochem. Explor.* **2018**, *187*, 118–130. [CrossRef]
52. Argyraki, A.; Kelepertzis, E. Urban Soil Geochemistry in Athens, Greece: The Importance of Local Geology in Controlling the Distribution of Potentially Harmful Trace Elements. *Sci. Total Environ.* **2014**, *482–483*, 366–377. [CrossRef] [PubMed]
53. Tešić, M.; Stojanović, N.; Knežević, M.; Đunisićević-Bojović, D.; Petrović, J.; Pavlović, P. The Impact of the Degree of Urbanization on Spatial Distribution, Sources and Levels of Heavy Metals Pollution in Urban Soils—A Case Study of the City of Belgrade (Serbia). *Sustainability* **2022**, *14*, 13126. [CrossRef]
54. Tjihuis, L.; Brattli, B.; Sæther, O.M. A Geochemical Survey of Topsoil in the City of Oslo, Norway. *Environ. Geochem. Heal.* **2002**, *24*, 67–94. [CrossRef]
55. Birke, M.; Rauch, U. Urban Geochemistry: Investigations in the Berlin Metropolitan Area. *Environ. Geochem. Health* **2000**, *22*, 233–248. [CrossRef]
56. Lis, J.; Pasieczna, A. *Geochemical Atlas of Warsaw and Environs 1:100,000*; Polish Geological Institute—National Research Institute: Warsaw, Poland, 2015.
57. Bitjukova, L.; Shogenova, A.; Birke, M. Urban Geochemistry: A Study of Element Distributions in the Soils of Tallinn (Estonia). *Environ. Geochem. Heal.* **2000**, *22*, 173–193. [CrossRef]

58. Šajn, R.; Gosar, M.; Bidovec, M.; Pirc, S.; Alijagic, J. Geochemical Mapping of Ljubljana Urban and Suburban Area, Slovenia. *J. Geochem. Explor.* **2011**, *107*, 186–201.
59. Locutura, J.; Bel-lan, A. Systematic Urban Geochemistry of Madrid, Spain, Based on Soils and Dust. In *Mapping the Chemical Environment of Urban Areas*; Johnson, C.C., Demetriades, A., Locutura, J., Ottesen, R.T., Eds.; Wiley-Blackwell: Chichester, UK, 2011; pp. 307–347. [CrossRef]
60. Sager, M. Urban Soils and Road Dust—Civilization Effects and Metal Pollution—A Review. *Appl. Sci.* **2020**, *10*, 8265. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Pre-Event Estimation of County-Level Human Casualty Projections in Southwestern China Based on the Spatial Aggregation of Village-Scale Lethality Data

Nan Zhang ^{1,2,3}, Xiwei Fan ^{1,2,4,*}, Chaoxu Xia ^{1,2}, Nan Xi ⁵, Jing Wang ^{1,2,4} and Gaozhong Nie ^{1,2,4,†}

¹ Key Laboratory of Seismic and Volcanic Hazards, China Earthquake Administration, Beijing 100029, China; zhangnan@ies.ac.cn (N.Z.); xiachaoxu@ies.ac.cn (C.X.); jingwang@ies.ac.cn (J.W.)

² Institute of Geology, China Earthquake Administration, Beijing 100029, China

³ Ningxia Seismological Bureau, Yinchuan 750001, China

⁴ Xinjiang Pamir Intracontinental Subduction National Observation and Research Station, Beijing 100029, China

⁵ China Earthquake Networks Center, Beijing 100029, China; xinan314@seis.ac.cn

* Correspondence: fanxiwei@ies.ac.cn

† Deceased author.

Abstract

An earthquake lethality model was employed to assess the casualty distribution in Yunnan, Guizhou, and Sichuan provinces, taking into account the ground motion acceleration with different 50-year exceedance probabilities. When the probability is 63%, fatalities are predominantly concentrated in central and south-western Yunnan, as well as central, southern, and western Sichuan. At a 10% probability, the peaks of the casualties are observed in southern, eastern, and central Sichuan. In Yunnan (excluding the northwest and southeast regions), the casualty density exhibits unevenness, whereas Guizhou experiences relatively low casualties (except in the eastern and western mountainous areas). Xichang incurs the most substantial losses, followed by Lancang. Xundian, Songming, and Dongchuan demonstrate a high propensity for fatalities, and the risk is relatively high in the vicinity of the Longjiang and Nujiang faults. If a destructive earthquake occurs near these areas within the next 50 years, the probability of a Level-I emergency response exceeds 10%. When the ground motion acceleration doubles (especially when the exceedance probability drops to 2% in 50 year and 0.1% in a year), the predicted number of casualties remains relatively stable. However, the grid of the casualty population exhibits a higher degree of spatial concentration of casualties, and the disaster-affected area expands. There exists no linear correlation between earthquake-induced fatalities and the ground motion level. When the 50-year exceedance probability decreases from 63% to 10%, the casualty rate may increase by several dozen times.

Keywords: seismic casualty; building lethality; building vulnerability; residential building; southwestern China

1. Introduction

The primary objective of pre-assessment is to identify counties with a high expected risk of casualties for emergency management departments, designating them as key areas of concern. This enables the coordination of resources to conduct risk management and emergency preparedness in a scientific and effective manner. The assessment encompasses various methods, including model-based evaluation, field surveys, expert assessments, and subsequent revisions.

Traditionally, pre-earthquake casualty assessments have focused on three main elements: earthquake risk, building vulnerability, and resident exposure. For instance, a vast region stretching from Sichuan's Daofu to Yunnan's Qiaojia, encompassing parts of Yunnan, Guizhou, and Sichuan, has been designated as a key risk area by the China Earthquake Administration (CEA). The 2008 Wenchuan earthquake resulted in \$124 billion in indirect losses, which accounted for 40% of the direct losses, with the casualty count exceeding 80,000 [1]. Regarding vulnerability, building collapse/damage and secondary disasters are the primary contributors to casualties [2]. There exists a correlation between the number of casualties and the number of collapsed/damaged buildings [3,4]. The vulnerability level of buildings is primarily associated with the probability of different damage states under various earthquake magnitudes or intensity ranges [5], which is not directly related to the death of individuals. The so called lethal level refers to the comprehensive likelihood or level of multiple factors that lead to death due to building destruction, expressed in terms of the higher mortality rate at each intensity level [6]. Building on this, it is mainly associated with the disparity in buildings' ability to cause casualties. In essence, the lethal level considers the probability of death under various intensities, factoring in the overall quantitative level of multiple factors in the affected area. It also accounts for situations such as "destruction without collapse, collapse without death," meaning the potential for survivors in the reserve space [6]. The last factor is population exposure. With the steady advancement of urbanization in China, large and medium-sized cities have significantly increased their population concentration, making them more susceptible to the impact of earthquake risks. Enhancing the accuracy and efficiency of casualty assessments can provide valuable insights for decision-makers, insurers, real estate brokers, and other stakeholders. This helps them to adjust economic development models, enhance building resilience, scientifically refine disaster insurance strategies, and improve the operability of disaster insurance. The ultimate goal is to ensure the safety of people's lives and property and to secure the long-term stable development of society.

The previous risk assessment was constructed based on a comprehensive analysis of historical earthquake data, vulnerability of buildings and secondary geological disasters, and the evaluation methods can be categorized into two groups. The first category is the empirical function method, which is based on seismic parameters and population density [7–12]. Among these [13], proposed a neural network model of seismic activity that considers factors such as intensity, seismic depth, fortification level, population density, etc. Ref. [14] used a vulnerability curve to express the relationship between seismic intensity and mortality across China and Western China. Certain researchers have proposed a probability model utilizing modified empirical mortality data [15,16], whereas other scholars have developed the Asian Earthquake Death Rapid Assessment Model (QAMECA), which incorporates earthquake intensity and additional factors [17]. However, the above method only utilizes the seismic parameters of historical earthquakes, rendering the model ineffective in areas lacking earthquake data and leading to inaccurate results. The second category is based on the vulnerability and lethality of buildings, constructing a model through the statistical fitting relationship between the vulnerability, lethality, and casualties of buildings, and using the regression of vulnerability function to assess casualties [18–20]. The vulnerability of buildings refers to their ability to withstand and recover from the impact of hazardous events [21]. These methods fully consider the impact of building damage on the number of deaths, thereby improving calculation accuracy. For example, certain scholars have identified principal determinants influencing casualty levels during seismic events, including building characteristics, occupancy rates, and timing of occurrence [22]. Other researchers have developed vulnerability functions for various structural types, which quantify the relationships among construction materials, extent of damage, and resulting

fatalities [23]. Additionally, studies have employed GIS-based building damage datasets to support modeling and casualty estimation [4,24]. Based on the vulnerability of various buildings in Greece [25], estimated casualties by considering the collapse probability of buildings under the influence of factors such as intensity and occupancy rate. A study proposed the scale of the relationship between building damage and casualties, along with a corresponding casualty assessment methodology [26]. Research has introduced a classification system for building damage, categorizing it into total collapse and total destruction, and established a correlation between damage grade and mortality rate [27]. Scholars have developed a casualty estimation model based on a building vulnerability matrix [28]. Furthermore, some authors have constructed relationships between casualty distribution and explanatory variables, such as building type and degree of damage, leading to the development of a prototype global casualty assessment model [29]. Due to the lack of data accuracy, model inapplicability, and quantitative lethal effect assessment of casualty factors (such as building damage, secondary geological disasters, major hazard sources, falling objects, crowd congestion, non-structural partial damage, adverse evacuation caused by traffic damage, etc.), the accuracy and regional applicability of these assessments are, to varying degrees, insufficient to meet actual needs.

The survival rate of individuals trapped during earthquakes is typically nonlinear and inversely proportional to the duration of the pressure they endure. In response to the need for rapid and accurate assessment, new demands have emerged for seismic pre-assessment and risk assessment, including the distribution of earthquake-buried personnel, necessitating the exploration of new models and data. Since China's accession to the WTO (World Trade Organization), the urbanization process has accelerated, and research on urban disaster assessment and management requires enhancement. Chinese scholars have conducted assessments of casualties or economic losses based on building vulnerability. Researchers have estimated the proportion of rural houses across diverse structural groups in China through a downscaling methodology that incorporates nonlinear regression, probabilistic risk prediction, and a cumulative probability model, thereby evaluating the vulnerability at the county level [30]. A separate study established an empirical vulnerability matrix for typical building groups within varying seismic intensity zones in Dujiangyan City [31]. By employing data mining and geoscience techniques, specifically support vector machines and association rule learning, investigators applied the seismic vulnerability index to Urumqi and verified its effectiveness [32]. Additionally, another research effort utilized the beta probability density function to derive the earthquake vulnerability matrix for rural areas in Weinan, revealing the distributions of the earthquake damage index and damage rate [33]. These previous assessment methods based on building lethality have limitations: (1) most existing methods rely on the vulnerability level of buildings, neglecting the diversity of building lethality and the influence of other factors. Historical earthquake analyses indicate that different regions may exhibit similar fractional intensity lethal levels under the combined effects of buildings, geographical environment, secondary disasters, and accidental factors, while the probability of death from buildings with similar vulnerability levels can vary significantly across different events [5,34,35]. In reality, a high building damage rate does not always correspond to a high mortality rate [36]. (2) In model construction and evaluation, if the analysis is based solely on the overall mortality rate in the affected areas rather than specific zones, its accuracy may be compromised, and it is challenging to capture the specific mortality characteristics of areas with varying residential densities and building structure types.

Thus, determining the lethality of a specific area is crucial. Addressing these shortcomings, someone conducted an effect analysis of lethal factors in various regions, constructed the lethal matrix for different factors in those regions, and established medium and long-

term risk assessments, pre-earthquake assessments, and post-earthquake rapid assessments using quantitative methods [37]. Based on historical casualties, a rapid assessment matrix for casualties was constructed according to the anti-fatality level and the mortality rate of each intensity level [6,38–40]. The method is semi-empirical, and its origin has been explicitly delineated: the lethality theory, which is founded on historical earthquake data, represents a contribution introduced by the authors of reference [37], and it was initially developed as an extension of the model proposed by the aforementioned authors [41]. This study has conducted large-scale lethality analysis and casualty assessment in China, and valuable results have been obtained [37–39].

The existing empirical function approach constructs models by leveraging historical seismic data and seismic parameters (e.g., magnitude, earthquake depth, population density, etc.). However, its performance is less than ideal in regions with a dearth of seismic data, frequently resulting in inaccurate outcomes. Although the current regression analysis method of vulnerability functions enhances accuracy by establishing statistical fitting relationships among building vulnerability, lethality, and casualty numbers, fully taking into account the impact of building damage on fatalities, such methods are constrained by data accuracy, model applicability, and the quantitative assessment of lethal effects from multiple casualty-inducing factors (e.g., secondary geological hazards, major hazard sources, crowd congestion, non-structural damage, and evacuation difficulties caused by traffic damage). More crucially, most existing methods for assessing building lethality solely rely on building vulnerability levels, overlooking the diversity of building lethality and the influence of other factors. Moreover, existing models that only analyze the overall mortality rates in affected regions rather than in specific sub-regions lack in-depth studies on earthquake damage and casualty assessments at the community/village level for specific seismic risk areas. Their accuracy may be impaired, and they find it difficult to accurately reflect the unique mortality patterns in regions with varying residential densities and building structure types. In China's seismic risk areas, research on evaluating earthquake disaster casualties under different seismic activity probability levels is also relatively limited.

In summary, the limitations of extant methods encompass the following:

- (1) Integrating household proportions across diverse building structures and fatality resistance levels to determine the regional overall fatality resistance threshold;
- (2) Converting seismic intensity vector data into grid format while preserving essential characteristics, pruning population grid data to align with regional boundaries, and ensuring coordinate system consistency;
- (3) Calculating matrices by overlaying seismic intensity, population distribution, and fatality levels, performing conditional grid allocation operations, and combining spatial analysis to aggregate casualty estimates for administrative regions, thereby achieving refined multi-source data integration and grid-level estimation.

In contrast, the method proposed in this study marks the initial endeavor to leverage data from the Seventh National Population Census [42–45] conducted by the National Bureau of Statistics for evaluating building-related casualties at the county level, enabling more refined casualty assessments at the district and county levels. The core methodology entails calculating the lethality levels of diverse building structures to obtain the overall lethality metrics for earthquake-prone regions. Lethality is, in essence, a parameter calibrated by means of historical mortality data. The lethality scale consists of 11 levels spanning from 0 to 1, which can be directly interpreted via casualty ratios with an error margin of $\pm 50\%$. This suggests that lethality levels represent quantified parameters obtained from historical mortality data through calibration procedures. By integrating data on the resistance of building structures to lethality, population distribution patterns, and other relevant factors, we ascertain the casualty ratios for specific regions under given seismic

intensity levels using the lethality/resistance matrix. The entire methodology hinges on a comprehensive analysis and calibration of historical mortality data. Although geographical conditions, secondary disasters, and accidental factors may impact the outcomes, this approach explicitly designates structural composition and lethality as the primary determinants. By collecting regional lethality data during surveys, post-earthquake casualty assessments can be carried out based on the regional lethality capacity and the population corresponding to intensity levels.

This method not only inherits the theoretical basis of the semi-empirical casualty rapid assessment matrix developed from the historical earthquake lethality theory [37], but also fills the methodological gaps in assessing casualties in community/village-level seismic risk areas and under varying earthquake activity probabilities; specifically, by integrating national census data, detailed seismic intensity zoning, and demographic distribution patterns with lethality matrix-based data processing techniques and evaluation models, it precisely reflects the casualty patterns across regions (e.g., Yunnan, Guizhou, and Sichuan provinces) within the updated intensity zoning frameworks, taking into account local housing structures and population distribution characteristics, achieves substantial improvements in assessment precision at the district and county levels, surmounting the previous model limitations in regional applicability and localized fine-scale evaluation, and addresses the research gaps encompassing substantial inaccuracies in lethal dose level assessment ($\pm 50\%$) and the absence of methodologies for converting county-level household statistical data into kilometer-scale fine-grained data. This model incorporates capabilities for data spatialization and refinement, emphasizing the spatialization of demographic and socio-economic variables along with precise, quantitative, and cyclically updated mortality metrics to enhance data accuracy, which distinguishes it from coarse-scale applications such as the HAZUS (Hazardous US, developed and maintained by the U.S. Federal Emergency Management Agency, is a standardized multi-hazard risk assessment methodology widely employed in seismological research for conducting detailed risk analyses and loss estimations associated with seismic events. This framework enables the quantitative evaluation of earthquake impacts, encompassing casualties, structural damage, economic losses, and infrastructure disruption. It thereby serves as an essential tool for informing disaster risk management strategies and enhancing emergency preparedness.) framework; it develops region-specific lethality matrices—including regional mortality and injury matrices—tailored to different geographical contexts, thereby overcoming the limitations of generalized empirical casualty functions; structured to evolve into a comprehensive multi-factor integration framework, it is capable of incorporating “secondary disaster-chain matrices,” “soft impacts derived from public sentiment,” “indirect economic losses,” and full-scale disaster-bearing structure data (e.g., proportions of brick-wood and earth-wood building types) to achieve end-to-end simulation, extending beyond HAZUS’s primary focus on direct casualties from single-event disasters; with dynamic spatiotemporal integration capabilities, it supports multi-temporal risk assessments based on “spatiotemporal population distribution,” advancing beyond static empirical functions to enable dynamic evaluation; and is designed to progress toward real-time social-sensing data fusion, allowing for the simulation of earthquake motion parameters (e.g., PGA, PGV) using rapid reporting data and real-time social sensing information to improve temporal relevance, a feature seldom integrated in frameworks such as HAZUS.

This research aims to examine the spatial distribution of earthquake-induced casualties within the Yunnan–Guizhou–Sichuan region. It will evaluate casualty patterns under diverse seismic probabilities and emergency response triggers. Furthermore, it will investigate the nonlinear relationship between ground motion intensity, (exemplified by Peak Ground Acceleration (PGA), and casualty rates, alongside the impacts of building

structural characteristics. The research also seeks to enhance rapid assessment methodologies through the development of casualty matrices, analysis of secondary disasters, refinement of models, parameter simulation, dataset improvement, and dynamic mortality mapping for high-risk zones. The spatialization of demographic and socioeconomic data is implemented to enhance assessment accuracy. A nonlinear relationship is posited between casualties and PGA. The integration of remote sensing and machine learning techniques facilitates the generation of refined, quantitative risk layers. This research is anticipated to elucidate casualty distributions and delineate high-risk zones for regional seismic assessment. It will reveal correlations between ground motion parameters and casualties, as well as building impacts, to inform seismic design practices. Additionally, it will improve model accuracy and scalability for pre-disaster risk evaluation and comprehensive disaster simulation.

In this manuscript, we first describe the inputs used, such as national census data, intensity zoning, and population distribution. Next, we introduce our data processing method and evaluation model based on a lethality matrix. Finally, our results include the assessment of casualties in Yunnan, Guizhou, and Sichuan, based on the latest intensity zoning and considering the local housing structure/population distribution, and the assessment of casualties by district/county.

2. Data

2.1. Study Area and Earthquake Distribution

China is one of the countries with the highest number of earthquake casualties in the world. Three of the world's top ten deadliest earthquakes have occurred in China. As a segment of the Eurasian seismic belt, the North-South Seismic Belt in China is crucial for strong earthquakes occurring in a roughly north-south direction [46–48]. Extending from Ningxia to Yunnan, through East Gansu and West Sichuan, and reaching as far as Mongolia in the north and Myanmar in the south, Yunnan, Guizhou, and Sichuan are situated at the southern end of this seismic zone. Nearly half of the earthquakes with a magnitude of 8 and above recorded in China have occurred within this zone, including the 8.0 magnitude Wenchuan earthquake [49,50]. Figure 1 illustrates the distribution of earthquake events with magnitudes of 5.0 and above in Southwest China. The catalog is available at CEDC (China Earthquake Data Center, website: <http://data.earthquake.cn>). The seismic catalog utilized in this study encompasses the regions of Yunnan, Guizhou, and Sichuan, incorporating both historical earthquake records and instrumentally recorded seismic data. This comprehensive dataset ensures that the temporal duration and spatial extent satisfy the prerequisites for regional seismic risk assessment. Consequently, it provides a robust foundation for constructing mortality/injury matrices, subsequent secondary disaster chain matrices, and dynamic risk assessment models. We consistently employ the M magnitude scale to quantify the magnitude of seismic energy release, making no distinction between M_s (surface wave magnitude, derived from surface wave amplitude) and M_w (moment magnitude, calculated from seismic moment). It is evident that the population density in Yunnan, Guizhou, and Sichuan is relatively high [51] (National bureau of statistics of China, 2025), historically experiencing frequent and destructive earthquakes [52–54]. Figure 1 shows the distribution of Richter earthquakes with magnitudes of 5.0 and above. According to the disaster assessment report from 1996 to 2022 [55], the historical earthquakes with magnitudes of 7.0 and above in Yunnan, Guizhou, and Sichuan in China by 2023 are detailed in Table 1.

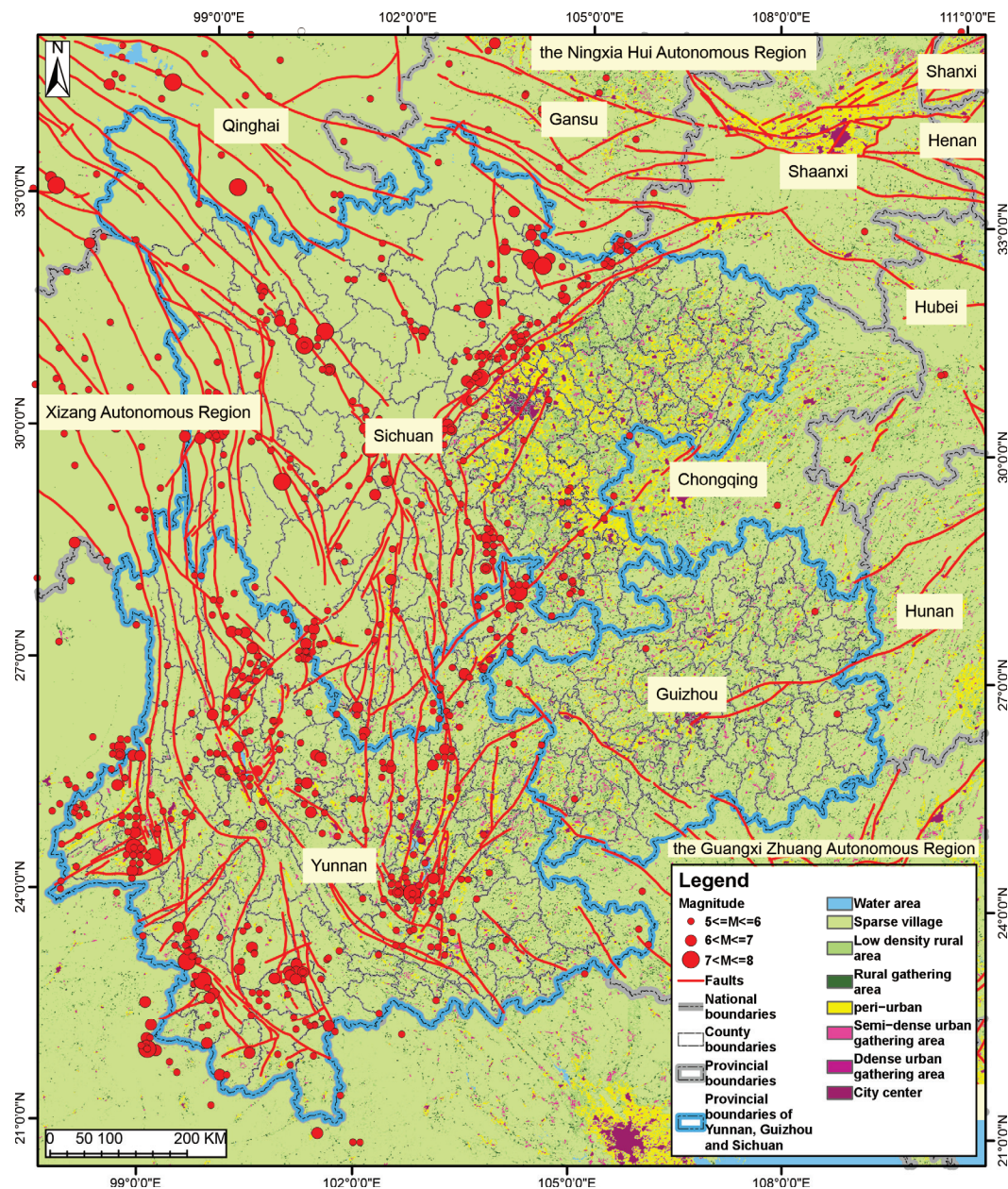


Figure 1. Distribution of settlements and historical earthquakes with a magnitude of M 5.0 or greater in southwestern China.

Table 1. Historical earthquakes above M7.0 in southwest China.

Province	County/City	Year	Month	Day	Magnitude (M)	Longitude (°)	Latitude (°)	Depth (km)
Yunnan	Yiliang	1500	1	4	≥7.0	103.1	24.9	-
	Yongsheng	1515	6	17	7.75	100.7	26.7	-
	Jianshui-quxi	1588	8	9	≥7.00	102.8	24	-
	Nidu	1652	7	13	7	100.6	25.2	-
	Dongchuan-Ziniupo	1733	8	2	7.75	103.1	26.3	-
	Huaning	1789	6	7	7	102.9	24.2	-

Table 1. Cont.

Province	County/City	Year	Month	Day	Magnitude (M)	Longitude (°)	Latitude (°)	Depth (km)
Sichuan	Shiping-baoxiu	1799	8	27	7	102.4	23.8	-
	Songming-yanglin	1833	9	6	8	103	25	-
	Shiping	1887	12	16	7	102.5	23.7	-
	Eshan	1913	12	21	7	102.5	24.2	-
	Dali	1925	3	16	7	100.4	25.7	-
	Gengma	1941	5	16	7	99.4	23.6	-
	Lancang	1941	12	26	7	99.9	22.7	-
	Menghai	1950	2	3	7	100.1	21.7	-
	Tonghai	1970	1	5	7.8	102.7	24.2	13
	Daguan	1974	5	11	7.1	104.1	28.2	14
	Longling	1976	5	29	7.3	99	24.5	24
	Longling	1976	5	29	7.4	98.7	24.6	21
	Gengma	1988	11	6	7.2	99.55	23.16	16
	Lancang	1988	11	6	7.4	99.79	22.92	13
	Yuulong	1996	2	3	7	100.22	27.3	10
	Detuo	1786	6	10	≥7.00	102.2	29.4	-
	Leibo-Mahu	1216	3	17	7	103.8	28.4	-
	Xichang	1536	3	19	7.5	102.2	28.1	-
	Mao-diexi	1713	9	4	7	103.7	32	-
	Kangding	1725	8	1	7	101.9	30	-
	Kangding	1786	6	1	7.75	102	29.9	-
	Luhuo	1816	12	8	7.5	100.7	31.4	-
	between Xichang and Puge	1850	9	12	7.5	102.4	27.7	-
	Batang	1870	4	11	7.25	99.1	30	-
	Daofu-qianning	1893	8	29	7	101.5	30.6	-
	Shiqu-luoxu	1896	3	1	7	98	32.5	-
	Daofu	1904	8	30	7	101.1	31	-
	Luhuo	1923	3	24	7.3	101	31.5	-
	Mao	1933	8	25	7.5	103.4	31.9	-
	Litang	1948	5	25	7.3	100.5	29.5	-
	Kangding	1955	4	14	7.5	101.8	30	-
	Luhuo	1973	2	6	7.6	100.7	31.3	11
	Songpan	1976	8	16	7.2	104.1	32.6	15
	Pingwu	1976	8	23	7.2	104.3	32.5	23
	Wenchuan	2008	5	12	8	103.4	31	14
	Lushan	2013	4	20	7	103	30.3	13
	Jiuzhaigou	2017	8	8	7	103.82	33.2	20

2.2. Data Sources Used in This Study

For this study, our foundational geographic information includes the boundaries of administrative divisions. We obtained the national basic geographic data for administrative divisions (version 2019) from the National Geographic Information Directory Service (NGISD), managed by the National Basic Geographic Information Center (NBGIC). We utilized these data, with a scale of 1:250,000, and the division data (version 2021) at a scale of 1:100,000, as our drawing framework (see Figure 2).

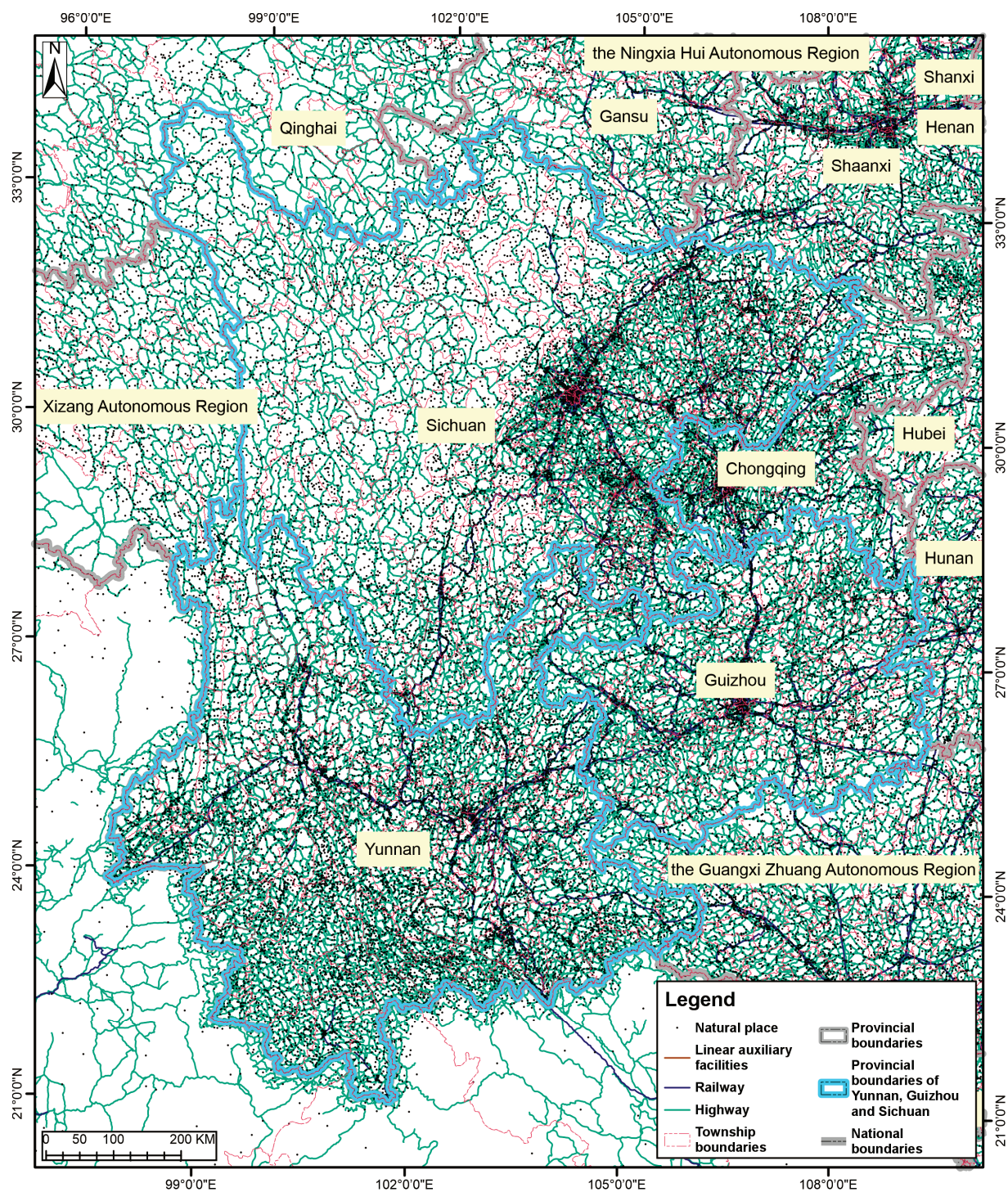


Figure 2. Basic geographic information of Yunnan, Guizhou, and Sichuan provinces (partial display).

Upon comparison, it was discovered that the population density figures per square kilometer published by the European Space Agency (ESA) and the China Earthquake Administration (CEA) are quite similar. However, the population density data provided by China's mobile terminal public services appears more varied. When compared with the recently released population data from the local government website as the benchmark, it indicates that the 2020 grid population data (<https://ghsl.jrc.ec.europa.eu>, accessed on 1 January 2024) published by ESA is more accurate to the actual figures (refer to Figure 3).

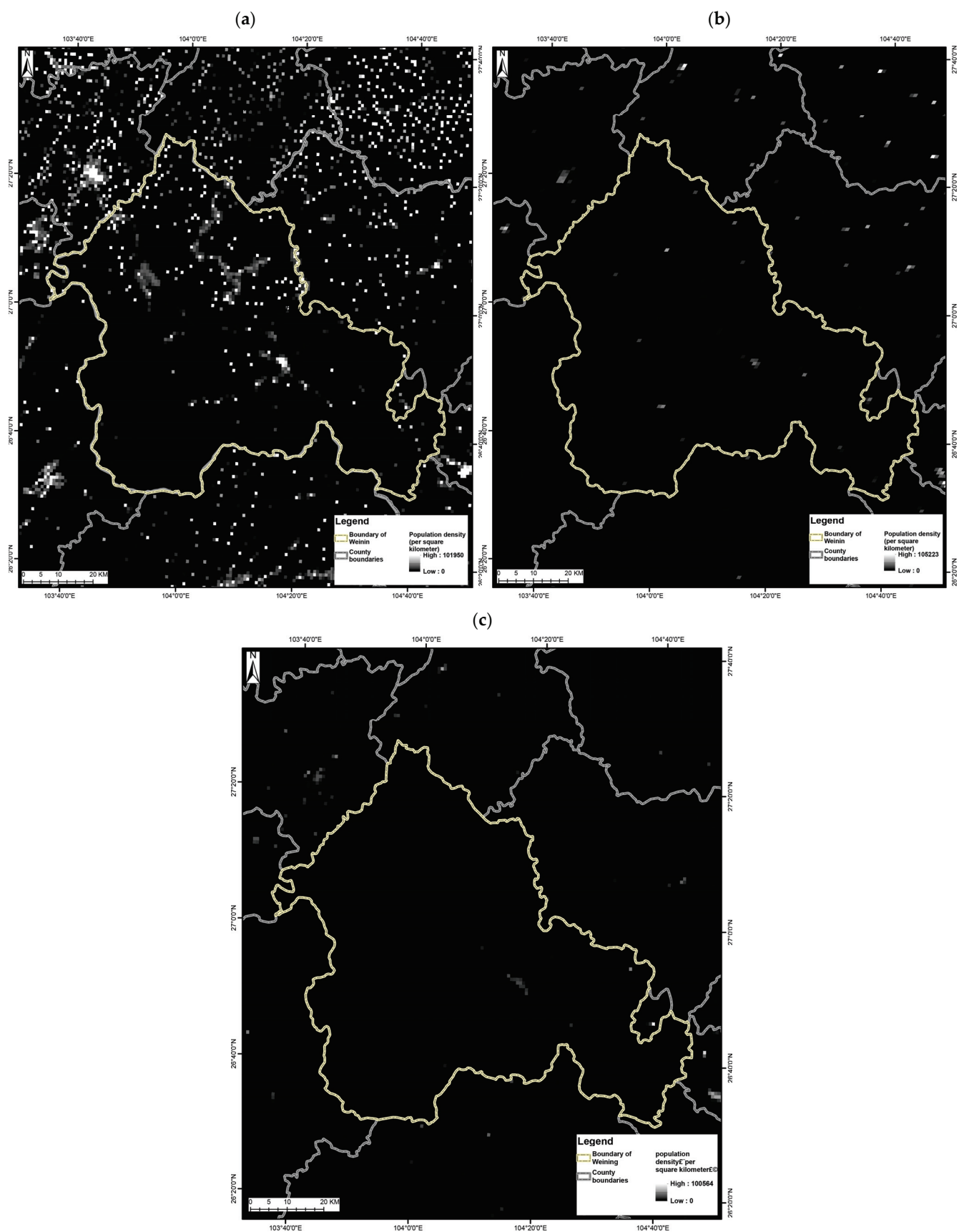


Figure 3. Comparison of population grid data from three sources in Weining Yi Autonomous County, Guizhou Province: (a) 2021 population data from the Chinese mobile terminal public service ‘Cube Data Society’; (b) 2020 population data from the European Space Agency (ESA); (c) 2016 population data from the Institute of Earthquake Prediction, China Earthquake Administration (CEA).

To validate the rationale for employing a 1 km × 1 km grid in computational analyses, we augmented the methodology with sensitivity analyses comparing grids of varying resolutions to enhance rigor. Four spatial resolutions (0.5 km, 1 km, 2 km, and 5 km) were evaluated based on multiple criteria: relative errors in total casualty counts across resolutions, alterations in high-density hotspot areas (e.g., regions exceeding 100 fatalities per km²), variations in spatial clustering patterns of casualties (e.g., patch sizes of high-risk zones), and root mean square errors (RMSE) between resolution outcomes and ground truth values (assuming the 0.5 km resolution as the finest reference). Results indicate that the 0.5 km resolution more accurately captures local variations in population and building distributions (e.g., urban neighborhoods and rural villages). By minimizing spatial averaging errors, it facilitates clearer identification of small-scale high-risk hotspots (e.g., densely populated residential areas in Xichang characterized by low seismic reinforcement rates). However, computational costs increase approximately three- to fivefold relative to the 1 km resolution. The 1 km resolution achieves an optimal balance between accuracy and efficiency: while casualty density distribution patterns align closely with those derived from the 0.5 km resolution, minor smoothing effects may occur in small hotspots (<0.5 km²). Relative errors in total casualty counts remain below 5% compared to the 0.5 km resolution. At a 2 km resolution, total casualty counts exhibit an 8–12% relative error relative to the 0.5 km resolution. High-density hotspot areas tend to merge, resulting in reduced spatial resolution (e.g., merging adjacent urban-rural regions in central Sichuan). At a 5 km resolution, total casualties are underestimated by 15–20% in complex terrains such as mountainous areas with scattered villages in western Yunnan. Spatial distribution patterns of casualties become significantly blurred, with high-risk zones overextended by 20–30% (e.g., merging multiple small high-risk grids into large contiguous areas). In urban or high-density regions (e.g., Xichang, Kunming), resolution reduction (from 0.5 km to 5 km) elevates errors (total casualty error range: 15–25%) due to loss of detailed population and building data. Rural or low-density areas (e.g., northwestern Yunnan) demonstrate lower sensitivity to resolution changes: errors remain below 10% at 2–5 km resolutions versus 0.5% at the 0.5 km resolution, as uniform population and building distribution reduces spatial averaging deviations. Sensitivity analysis suggests that optimal resolutions for localized detailed assessments range from 0.5 to 1 km, prioritizing accuracy in high-risk urban or complex areas. Resolutions of 2–5 km are more suitable for regional rapid screening. The 1 km resolution delivers optimal performance across most scenarios with errors below 5%, while the 0.5 km resolution maintains computationally manageable costs.

To validate the rationality of the ESA population grid dataset, we utilized county-level permanent resident population data from the 2020 Seventh National Population Census for Yunnan, Guizhou, and Sichuan provinces as the benchmark (official actual values), encompassing 197 county-level administrative units. Correspondingly, the ESA population grid data (2020 version) at 1 km resolution was aggregated to the county level via spatial overlay analysis in ArcGIS. Consistency and accuracy were assessed using the mean absolute error (MAE), root mean square error (RMSE), correlation coefficient (R^2), and the proportion of counties with relative errors within 10%. The results, summarized in the accompanying Table 2, indicate a strong correlation between the ESA population grid estimates and the official census data, with $R^2 = 0.92$. The MAE was 44,000 persons (approximately 5.3% of the average county population), the RMSE was 59,000 persons, and 82.3% of counties exhibited relative errors $\leq 10\%$. These error levels fall within the acceptable range for population spatialization studies, where thresholds commonly do not exceed 15%. Furthermore, the 1 km spatial resolution of the ESA grid data allows for a detailed representation of population distribution patterns. In summary, the adoption of

ESA population grids is statistically well justified, both in terms of quantitative agreement with census data and the enhanced spatial granularity offered by the dataset.

Table 2. Comparative evaluation of permanent resident population data (officially reported statistics by the Seventh National Population Census) in the study area against ESA population grid data (2020).

Index	Yunnan	Guizhou	Sichuan	Total or Average of Three Provinces
Sample size (number of counties)	129	88	183	390
MAE (ten thousand people)	4.2	3.8	5.1	4.4
RMSE (ten thousand people)	5.7	5.2	6.8	5.9
R^2	0.90	0.93	0.92	0.92
Proportion of counties with relative error $\leq 10\%$	82.1%	79.5%	84.7%	84.7%

A standard [56] “Seismic Ground Motion Parameter Zoning Map of China” (GB 18306-2015) was released in 2016, detailing the basic ground motion PGA (peak ground acceleration) distribution (see Figure 4) and corresponding acceleration response spectrum characteristic period for class II sites in China, with a 50-year probability of exceedance of 10%. This standard also stipulates that the risk level of anti-collapse design parameters (rare ground motion) is not less than the 50-year exceedance probability of 2%, aligning with advanced countries in terms of earthquake prevention and disaster reduction. Utilizing the probabilistic seismic hazard analysis (PSHA) method proposed by [57], the zoning map employs active fault data (for instance, segmentation, length, and slip rate data of 243 fault segments from 130 faults, and paleoseismic data from 161 fault segments) to estimate the occurrence rate of large earthquakes. The PGA map is combined with the “Comparison table of site seismic peak acceleration and seismic intensity” (refer to Table 3), resulting in the 5th generation of seismic intensity zoning.

Table 3. Comparison of PGA and seismic intensity for class II sites in China [56].

PGA of Class II Sites in China	$0.04 \text{ g} \leq a_{\max\text{II}} < 0.09 \text{ g}$	$0.09 \text{ g} \leq a_{\max\text{II}} < 0.19 \text{ g}$	$0.19 \text{ g} \leq a_{\max\text{II}} < 0.38 \text{ g}$	$0.38 \text{ g} \leq a_{\max\text{II}} < 0.75 \text{ g}$	$a_{\max\text{II}} \geq 0.75 \text{ g}$
Seismic intensity	VI	VII	VIII	IX	$\geq \text{X}$

We utilized the household data from the census yearbook [51] (National Bureau of Statistics of China, 2025), which is categorized into five typical housing structure types or other attributes (refer to Figures 5 and 6). The mapping relationship between building structure categories in the Seventh National Population Census Yearbook and the structural performance grades adopted in this study is established as follows: reinforced concrete structures correspond to the category designated as “reinforced concrete houses”; brick-concrete structures, brick-wood structures, and other masonry structures are collectively classified under “brick houses”; wooden structures correspond to “wooden houses”; earth-wood structures are mapped to “civil houses”; and bamboo, grass, or adobe structures correspond to the classification “bamboo/grass/adobe houses”. In the census yearbook, these families are classified into “summary”, “urban area”, “semi-urban area”, and “rural area” categories based on residence type, number of building floors, or bearing capacity type. Tables 4–6 present the counts and corresponding proportions of households classified by building structure in the urban, township, and rural areas of Chengdu City, Sichuan, China, based on data from the population census. Our data collection methods include: (1) obtaining permission to copy paper images [51,58]; (2) logging onto the official website

of statistical institutions to request public information; (3) contacting the office of statistical institutions or the population division directly; (4) sending emails; and (5) employing other methods.

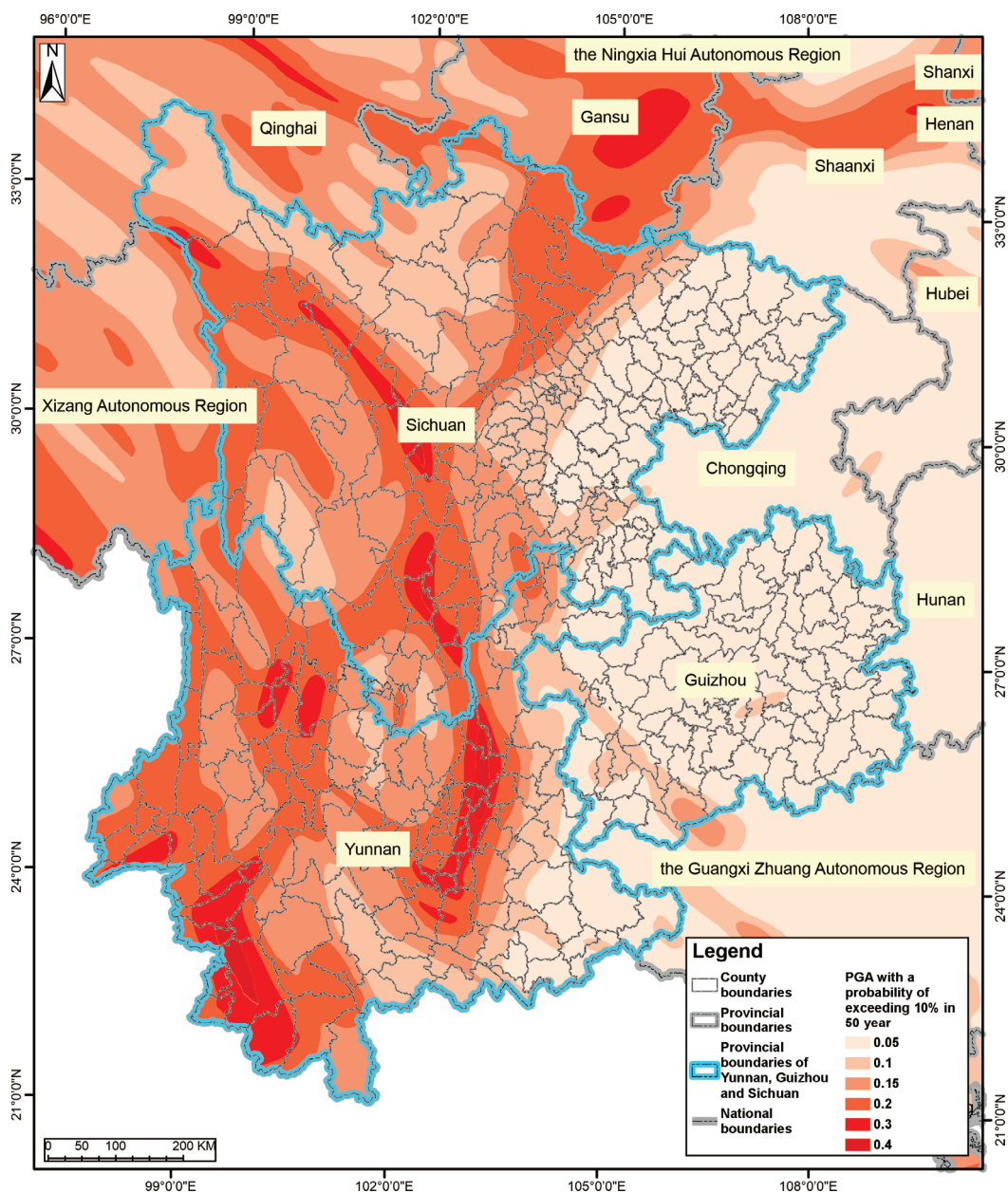


Figure 4. Distribution of PGA with a probability of exceeding 10% over 50 years in Yunnan, Guizhou, and Sichuan.

Table 4. Number and proportion of households by building structures in urban areas of Chengdu City, Sichuan, China from the population census.

County	Steel and Steel–Concrete Structure	Brick–Concrete Structure	Brick–Wood Structure	Vegetation Structure	Other Structures	Steel and Steel–Concrete Structure Proportion (%)	Brick–Concrete Structure Proportion (%)	Brick–Wood Structure Proportion (%)	Vegetation Structure Proportion (%)	Other Structure Proportion (%)
Chengdu	509,335	397,652	105,657	4999	121	50.04%	39.07%	10.38%	0.49%	0.01%
Jinjiang	33,787	30,346	3234	141	0	50.05%	44.95%	4.79%	0.21%	0.00%
Qingyang	34,143	28,290	5609	221	0	50.02%	41.44%	8.22%	0.32%	0.00%
Jinniu	46,036	32,210	13,263	508	1	50.03%	35.00%	14.41%	0.55%	0.00%

Table 4. Cont.

County	Steel and Steel–Concrete Structure	Brick–Concrete Structure	Brick–Wood Structure	Vegetation Structure	Other Structures	Steel and Steel–Concrete Structure Proportion (%)	Brick–Concrete Structure Proportion (%)	Brick–Wood Structure Proportion (%)	Vegetation Structure Proportion (%)	Other Structure Proportion (%)
Wuhou	42,408	29,358	12,545	436	0	50.04%	34.64%	14.80%	0.51%	0.00%
Chenghua	48,853	39,770	8777	166	1	50.07%	40.76%	9.00%	0.17%	0.00%
Chengdu High-tech Zone	38,642	33,669	4817	18	0	50.09%	43.64%	6.24%	0.02%	0.00%
Long Quanyi	38,199	26,418	11,545	189	3	50.03%	34.60%	15.12%	0.25%	0.00%

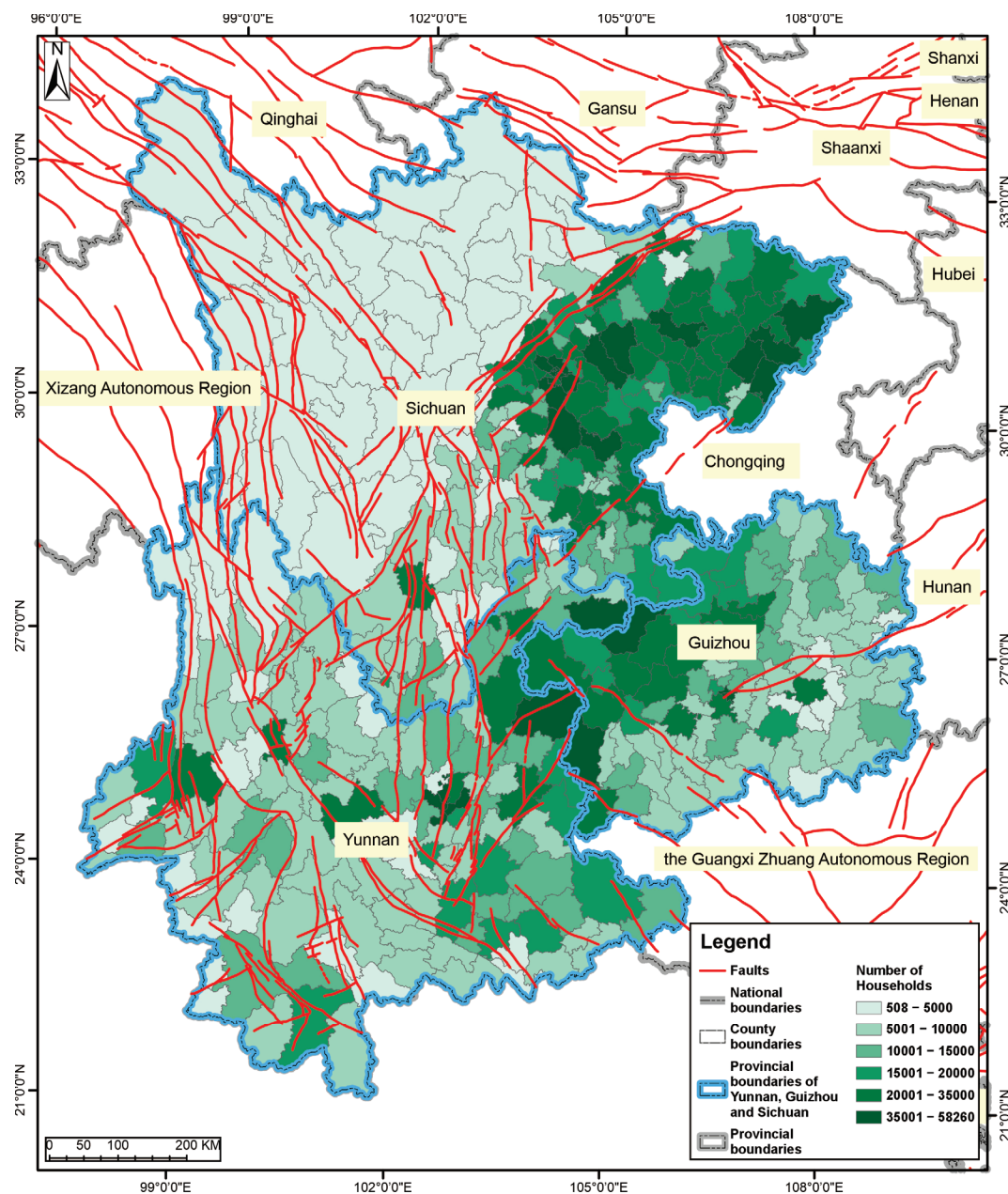


Figure 5. Distribution of the residential building inventory by household size in Yunnan, Sichuan, and Guizhou provinces.

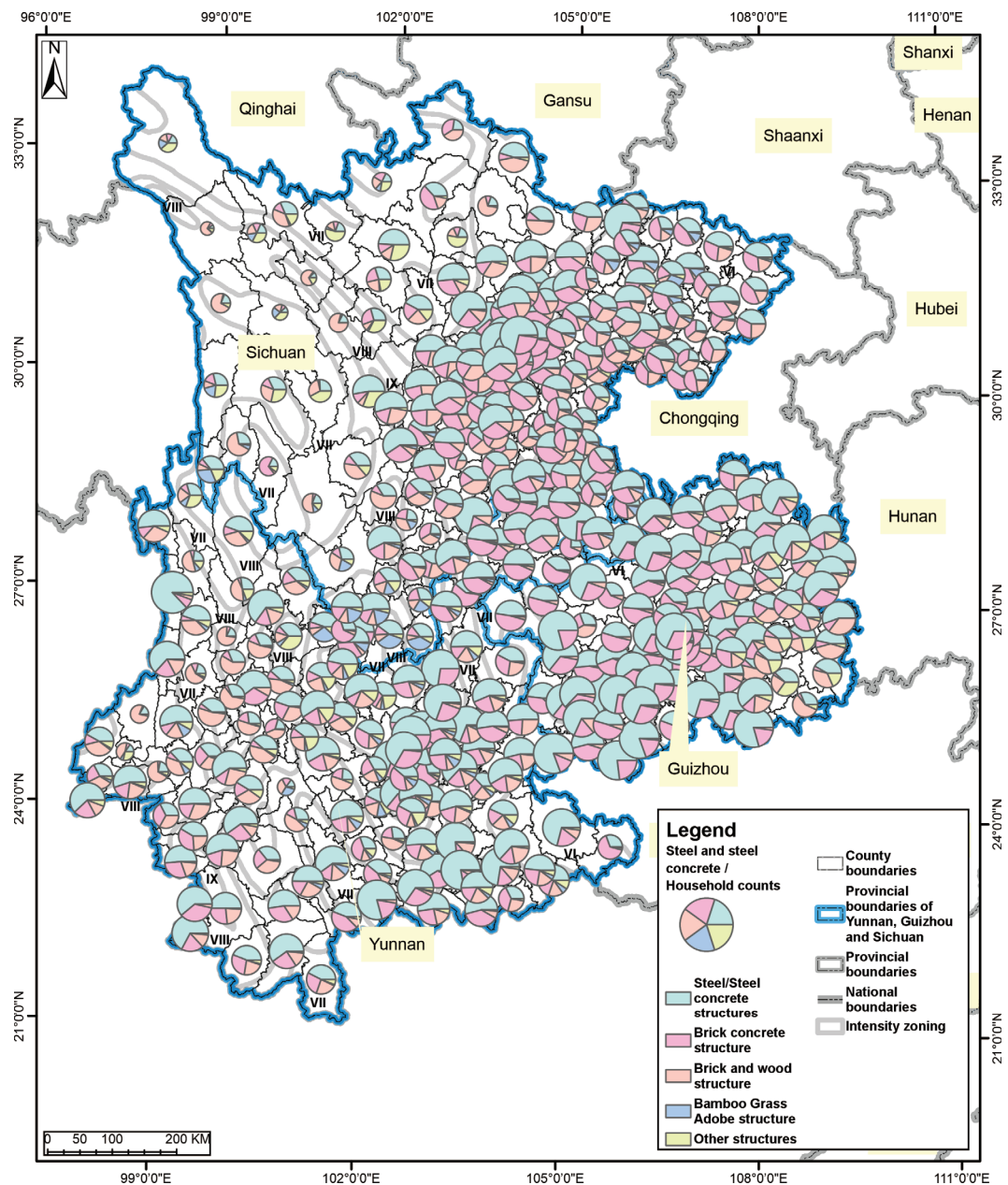


Figure 6. The proportion fan diagram illustrating the distribution of households with various building structures across counties in southwestern China.

Table 5. Number and proportion of households by building structures in township areas of Chengdu City, Sichuan, China from the population census.

County	Steel and Steel-Concrete Structure	Brick-Concrete Structure	Brick-Wood Structure	Vegetation Structure	Other Structures	Steel and Steel-Concrete Structure Proportion (%)	Brick-Concrete Structure Proportion (%)	Brick-Wood Structure Proportion (%)	Vegetation Structure Proportion (%)	Other Structure Proportion (%)
Chengdu	27,556	11,476	3095	97	162	65.01%	27.07%	7.30%	0.23%	0.38%
Jinjiang	0	0	0	0	0					
Qingyang	0	0	0	0	0					
Jinniu	0	0	0	0	0					
Wuhou	0	0	0	0	0					
Chenghua	0	0	0	0	0					

Table 5. Cont.

County	Steel and Steel–Concrete Structure	Brick–Concrete Structure	Brick–Wood Structure	Vegetation Structure	Other Structures	Steel and Steel–Concrete Structure Proportion (%)	Brick–Concrete Structure Proportion (%)	Brick–Wood Structure Proportion (%)	Vegetation Structure Proportion (%)	Other Structure Proportion (%)
Chengdu High-tech Zone	0	0	0	0	0					
Longquanyi District	414	415	32	4	8	47.42%	47.54%	3.67%	0.46%	0.92%
Qing baijiang	707	741	34	3	6	47.42%	49.70%	2.28%	0.20%	0.40%
Xindu	2052	228	9	0	0	89.65%	9.96%	0.39%	0.00%	0.00%
Wenjiang	444	369	34	0	1	52.36%	43.51%	4.01%	0.00%	0.12%
Shuangliu	353	49	14	1	0	84.65%	11.75%	3.36%	0.24%	0.00%
Pidu	569	347	78	2	0	57.13%	34.84%	7.83%	0.20%	0.00%
Xinjin	379	200	68	3	2	58.13%	30.67%	10.43%	0.46%	0.31%
Jintang	9565	2675	147	60	46	76.56%	21.41%	1.18%	0.48%	0.37%
Dayi	5718	2699	499	0	43	63.82%	30.13%	5.57%	0.00%	0.48%
Pujiang	2769	813	324	8	12	70.53%	20.71%	8.25%	0.20%	0.31%
Tianfu New District Eastern	0	0	0	0	0					
New District of Chengdu Du	84	131	19	5	8	34.01%	53.04%	7.69%	2.02%	3.24%
Jiangyan	918	680	312	2	3	47.94%	35.51%	16.29%	0.10%	0.16%
Pengzhou	1131	799	127	1	1	54.93%	38.81%	6.17%	0.05%	0.05%
Qionglai	882	572	586	1	10	43.00%	27.89%	28.57%	0.05%	0.49%
Chongzhou	1237	416	772	2	11	50.74%	17.06%	31.67%	0.08%	0.45%
Jianyang	334	342	40	5	11	45.63%	46.72%	5.46%	0.68%	1.50%

Table 6. Number and proportion of households by building structures in rural areas of Chengdu City, Sichuan, China from the population census.

County	Steel and Steel–Concrete Structure	Brick–Concrete Structure	Brick–Wood Structure	Vegetation Structure	Other Structures	Steel and Steel–Concrete Structure Proportion (%)	Brick–Concrete Structure Proportion (%)	Brick–Wood Structure Proportion (%)	Vegetation Structure Proportion (%)	Other Structure Proportion (%)
Chengdu	50,041	61,837	30,216	1494	959	34.62%	42.78%	20.90%	1.03%	0.66%
Jinjiang	0	0	0	0	0					
Qingyang	0	0	0	0	0					
Jinniu	0	0	0	0	0					
Wuhou	0	0	0	0	0					
Chenghua	0	0	0	0	0					
Chengdu High-tech Zone	43	152	23	1	4	19.28%	68.16%	10.31%	0.45%	1.79%
Longquanyi District	320	1764	112	21	18	14.32%	78.93%	5.01%	0.94%	0.81%
Qingbaijiang	1095	2926	417	30	18	24.41%	65.23%	9.30%	0.67%	0.40%
Xindu	2425	6387	775	15	31	25.17%	66.30%	8.05%	0.16%	0.32%
Wenjiang	3429	2284	701	3	43	53.08%	35.36%	10.85%	0.05%	0.67%
Shuangliu	7592	3162	685	10	51	66.02%	27.50%	5.96%	0.09%	0.44%
Pidu	3925	6095	901	35	43	35.69%	55.41%	8.19%	0.32%	0.39%
Xinjin	1469	1309	759	8	20	41.21%	36.72%	21.29%	0.22%	0.56%
Jintang	3355	8683	1007	607	206	24.21%	62.66%	7.27%	4.38%	1.49%
Dayi	3280	1243	3891	20	55	38.64%	14.64%	45.84%	0.24%	0.65%
Pujiang	1653	1043	1806	21	42	36.21%	22.85%	39.56%	0.46%	0.92%
Tianfu New District Eastern	3014	3444	636	121	16	41.68%	47.63%	8.80%	1.67%	0.22%
New District of Chengdu Du	1083	3851	851	236	49	17.84%	63.44%	14.02%	3.89%	0.81%
Jiangyan	3168	3475	1929	14	38	36.73%	40.29%	22.37%	0.16%	0.44%
Pengzhou	3452	6733	3376	10	54	25.34%	49.42%	24.78%	0.07%	0.40%
Qionglai	1881	1996	5628	23	94	19.55%	20.74%	58.49%	0.24%	0.98%
Chongzhou	3407	2135	4991	17	47	32.15%	20.15%	47.10%	0.16%	0.44%
Jianyang	5450	5155	1728	302	130	42.69%	40.38%	13.54%	2.37%	1.02%

As shown in Figure 5, the residential distribution across the three provinces is uneven, with Sichuan being the most pronounced example. The residential density in the northeast is

2–10 times greater than that in the southwest. The Northeast Sichuan Plain is conducive to living and farming, resulting in relatively concentrated housing. In the central and western regions, as well as the northwest, the number of households in large areas ranges from 500 to 5000, while the southern region exhibits a more balanced distribution. Guandu District has 58,260 households, Chenghua District has 48,853, Shuangliu District has 46,757, and Jinniu District has 46,036, all exceeding 45,000. Notably, Xichang has the highest concentration of residents in the south with 27,361 households. The western part of Guizhou and northeastern Yunnan are flat and densely populated. The highest number of households in Yunyan District of Guiyang is 37,976, followed by 36,034 in Nanming District, 35,935 in Panzhou City of Liupanshui City, and 35,749 in Qixingguan District of Bijie City, all surpassing 35,000. In contrast, the residential areas in eastern and southern Guizhou are more scattered than those in the northwest, yet there are still some concentrated residential zones. Although the residential areas in Guizhou Province are predominantly in the northwest, the concentration is not as marked as in Sichuan. Among the three provinces, Yunnan has the most dispersed households, with the largest numbers in the western, central, and eastern counties, respectively: 58,260 in Guandu District of Kunming, 42,165 in Xuanwei City of Qujing, 38,867 in Wuhua District of Kunming, 37,675 in Zhenxiong County of Zhaotong, 35,901 in Xishan District of Kunming, and 34,461 in Panlong District of Kunming, all exceeding or nearing 35,000.

As depicted in Figure 6, the primary building structures in northeastern and southern Sichuan are reinforced concrete, with brick concrete and brick wood structures following. In the majority of the Chengdu Plain, over 50% of buildings are constructed with steel-concrete structures. Particularly in districts such as Jinjiang, Suining Hedong new area, Chengdu high tech Zone, and Deyang, the proportion of reinforced concrete houses exceeds 80%. In the Northeast of Sichuan, the prevalent building structure is brick concrete, followed by steel concrete. In areas such as Gong, Daocheng, and the West District of Panzhihua, the proportion of brick concrete structures exceeds 60%. In the mountainous regions of Western and Northern Sichuan, buildings are predominantly brick and wood structures, including civil structured and wood structures. Near the provincial boundaries of Gansu, Qinghai, and Tibet, over 60% of households are brick and wood. Specifically, in counties such as Daofu, Dege, and Baiyu, the proportion of brick and wood structures exceeds 60%, while the proportion of reinforced concrete ranges from 8.88% to 27.56%. In Huidong and Huili, the proportion of bamboo, grass and adobe houses exceeds 30%, with steel-concrete houses constituting 35.306% and 39.808%, respectively. In Lancang, Heishui, and Yajiang County, the proportion of civil structured and wooden houses exceeds 40%, while the proportion of steel-concrete structures is only 19.47%, 20.19%, and 25.86%. Overall, the total proportion of brick wood, civil, wood, bamboo, grass, and adobe structures in 41 counties in Sichuan exceeds 40%, and the provincial household ratio is 211,196/5,895,390, which equals 3.58%. Twenty-five counties, including Luhuo, Dege, and Xinlong, have a total proportion of brick, wood, civil, bamboo, grass, and adobe structures exceeding 50%. Notably, in eleven counties, including Luhuo, Dege, Xinlong, and Daofu, the proportion of reinforced concrete houses does not exceed 20%.

Figure 6 also depicts the distribution pattern of building structures in Guizhou Province. Drawing a north–south line from Guiyang, the provincial capital, reveals that most buildings in the western region are constructed from reinforced concrete and brick concrete. Notably, in the southwest, the proportion of brick wood, civil and wood structures is significantly lower. Specifically, in Guanshanhu, Bijiang, Wangmo, and Xingyi, over 80% of the houses are reinforced concrete. In Dafang, Xifeng, Xiuwen, Luodian, and Zhijin, more than 50% of the houses are brick concrete structures, and over 30% are steel concrete structures. On the eastern side of the line, reinforced concrete remains prevalent, but the proportion of brick wood, civil and wood structures has risen substantially, with

some counties having more brick and wood houses than brick and concrete. In Congjiang, Danzhai, and Jinping, the proportion of brick and wood exceeds 40%, and the proportion of reinforced concrete exceeds 30%. Chishui is unique, with 14.38% of bamboo, grass, adobe, civil and wood structures, while steel concrete and brick concrete proportions are 34.78% and 41.63%, respectively. In Guizhou Province, Rongjiang, Jianhe, and Leishan have the highest proportions of civilian timber structures, all exceeding 20%. The proportion of brick timber structures is 23.873%, 24.98%, and 36.83%, respectively, and the proportion of steel-concrete structures is 38.13%, 40.60%, and 38.68%. Overall, except for Nayong, where the proportion of steel-concrete is 30.43%, and brick wood is 39.85%, and Congjiang, where the proportion of steel-concrete is 31.26%, and brick wood is 57.54%, the proportion of steel-concrete in other counties is more than 33%. The primary structures are steel-concrete and brick concrete, indicating that the overall level of building fortification design is relatively balanced. In Guizhou, there are 10 counties where the combined proportion of houses with brick, wood, civil, bamboo, grass and adobe structures exceeds 40%, including Congjiang, Rongjiang, Leishan, among others. The proportion of households in the province is $81,700/2,109,740 = 3.87\%$.

Except for a few counties in central and northern Yunnan, which are mainly brick and wood structure houses, other areas are predominantly reinforced concrete houses, followed by brick concrete and brick and wood structure houses. The proportion of brick and wood structures in nine counties, including Tengchong, Jianchuan, Longling, and others, exceeds 50%. The proportion of steel concrete in some counties, such as Fugong, Guandu, and Luchun, exceeds 70%. In some counties, civil structured houses and wooden structures also constitute a significant proportion. The proportion of civil structured and wooden houses in Yongsheng, Deqin, Lianghe, Wuding, and Nanhua County exceeded 20%, with the proportions in Yongsheng and Deqin reaching 35.936% and 30.719%, respectively. Overall, the proportion of residential houses, including wooden, bamboo, grass, and adobe houses, in 38 counties exceeded 40%, and the proportion of households in the province reached 434,527 out of 282,610, which is 15.38%. Among them, there are 23 counties, including Tengchong, Lianghe, Jianchuan, etc., where the proportion of brick wood, bamboo, grass and adobe structure houses exceeds 50%. Among these, the proportion of reinforced concrete houses in Tengchong, Lianghe, Jianchuan, and Jingdong does not exceed 20%.

Section 3.2 will introduce the anti-lethal level model, which is based on the results [18,52] (Xiwei, 2020; Gaochong et al., 2021). In this research, 52 historical fatal earthquakes from 1996 to 2010 were cited.

3. Methodology

The specific steps of data preprocessing and region mapping calculation are shown in Figure 7: a rectangle represents the workflow, and a parallelogram represents the datum plane. After data preprocessing, the population, intensity zoning and anti-lethal level are superimposed according to the seismic matrix of historical earthquakes. Then, the mortality (injury) rate per kilometer grid in the intensity partition is determined by logical operation. On this basis, the kilometer level casualty rate spatial statistics were carried out to estimate the number of casualties in administrative divisions at all levels.

3.1. Data Preprocessing

The data point vectors of administrative villages are categorized into three groups based on their geographical positions. The urban and rural classification codes in the attribute table are defined according to the guidelines for the creation of statistical zoning codes and urban and rural zoning codes, as issued by the National Bureau of Statistics [59]. Each classification code consists of three digits, with the first digit "1" or "2" denoting

urban or rural areas, respectively (refer to Table 7). In this document, data encoded as “111” and “112” are consolidated to represent urban areas, data encoded as “121”, “122”, and “123” are utilized to represent towns, and data encoded as “210” and “220” are employed to represent rural areas.

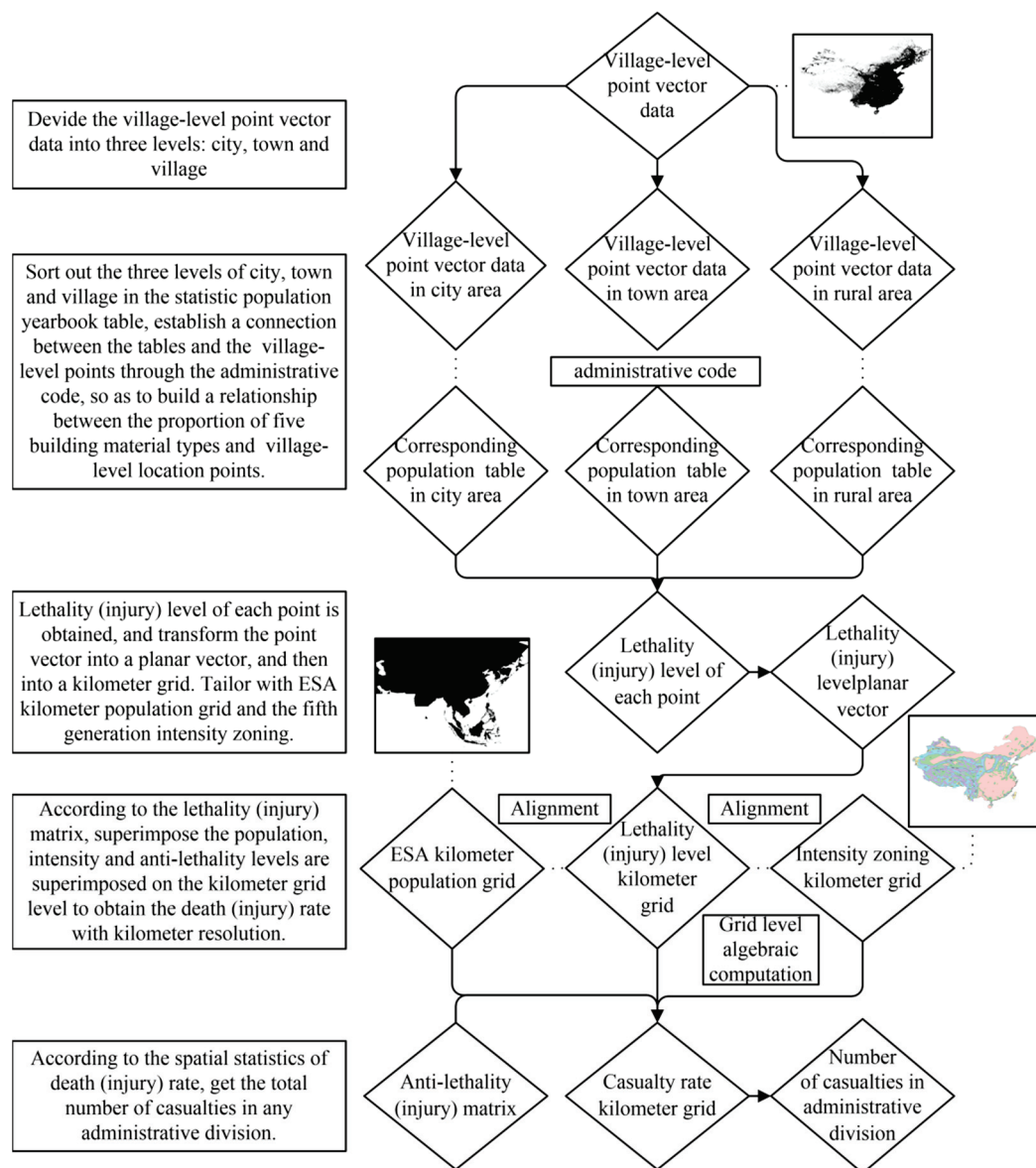


Figure 7. The flowchart illustrates the procedures for data preprocessing and region mapping calculations. Dashed lines indicate correspondence or associative relationships, whereas solid unidirectional arrows denote directional output relationships.

Table 7. Urban/rural zoning codes of China (CNBS).

Area Classification	The Main Urban Area	The Urban–Rural Integration Zone	The Town Center Area	The Town–Rural Integration Zone	Special Township Area	The Rural Central Area	Village
Urban/Rural Zoning Codes	111	112	121	122	123	210	220

The Demographic Yearbook [58] classifies households by region, such as urban, town-ship, and rural areas, and provides clear explanations (National Bureau of Statistics of

China, 2022). Consequently, to map census data with building structure types, we categorize households based on the structure types outlined in the Yearbook, corresponding to the three administrative divisions of cities, towns, and villages. Next, we utilize the district/county administrative code to link the village/community vector data with the corresponding number of households for each building structure type. This connection allows us to establish a correlation between the locations of villages/communities and the distribution of housing types in these areas (see Figure 6), thereby aiding in the assessment of the local lethality of these locations.

This study integrates three categories of data aligned with kilometer-scale grid coordinates: demographic, intensity, and mortality datasets. Software including ArcGIS v10.2 was employed to conduct resampling of raw spatial data. Demographic data originate from the European Space Agency (ESA) kilometer-grid system. The construction of mortality kilometer grids follows a three-stage conversion procedure: from village- or community-level mortality polygon data to raster grids, and subsequently to kilometer-resolution grids. The overall workflow for generating kilometer-grid raster data from point vector data comprises three principal stages. First, point vectors are converted into polygon vectors (areal features) to facilitate the spatial dissemination of point attributes. The resulting polygons inherit all attributes from the original points (e.g., population size, administrative codes, total resources), with each polygon representing the unique attribute set of a corresponding village or community. Second, polygon vectors are transformed into raster data, with the target pixel size specified at the kilometer level during the vector-to-raster conversion. Resampling techniques are primarily applied during raster resolution transformation. When converting polygons to an initial raster dataset, a pixel size of 1000 m is designated. The core methodology involves discretely assigning attribute values from vector polygons to raster pixels using the Cell Center method: a pixel acquires the attribute value of the polygon within which its centroid is located; otherwise, a NoData value is assigned. To convert the initial raster data (e.g., higher-resolution data derived from polygons) into 1 km resolution grids, the ArcGIS v10.2 resampling algorithm applies the Nearest Neighbor method. This approach directly assigns the value from the source raster pixel corresponding to the center of each target pixel, without interpolation. The algorithm ensures computational efficiency while preserving all original raster attribute values, such as discrete attributes including village administrative codes and classification types. When polygonal feature attributes represent discrete data (e.g., township codes or land use type codes), the nearest neighbor method is optimal as it effectively prevents attribute distortion. The intensity zoning kilometer grid is generated from intensity vector surfaces using ArcGIS v10.2. During the conversion of vector surfaces to kilometer-grid raster data in ArcGIS, the core spatial sampling technique involves resampling with the default nearest neighbor method, which adopts the value of the source pixel closest to the target pixel centroid. This approach is suitable for discrete data types (e.g., land use categories and building classifications), as it preserves the original categorical attributes and prevents value distortion.

We summarize the grid alignment process steps for demographic, intensity, and mortality data as follows. (1) Establishing a unified spatial reference: employing the WGS84 longitude and latitude coordinate system with a grid resolution of 1 km × 1 km; defining grid coverage boundaries based on the study area extent (e.g., Yunnan, Guizhou, Sichuan) to prevent data clipping or overlap. (2) Collecting population data (from administrative unit statistics), lethality parameters, seismic intensity data (including PGA, PGV, and simulated China seismic intensity results), and building attribute data (e.g., type proportions and seismic resistance); standardizing data formats (such as Shapefile) and extracting spatial coordinate information. (3) Disaggregating administrative unit population data into 1 km

grid cells based on remote sensing data and machine learning models (e.g., geostatistical regression): this achieves fine-grained population allocation from administrative units to grid units. (4) Resampling seismic motion parameter simulation results (e.g., PGA, PGV, typically in vector or high-resolution raster format) to 1 km grids: utilizing bilinear interpolation or Kriging interpolation methods to ensure spatial continuity of intensity parameters within grid cells; China seismic intensity data are transformed into rasterized intensity values in accordance with standards [56,60]. (5) Performing coordinate consistency checks on all data layers and correcting geometric deviations to ensure complete matching of grid boundaries and resolution, with an error tolerance of no more than 0.5 grid cells. (6) Integrating population density, intensity parameters, and building type proportions within unified grid units, ensuring precise spatial alignment among population distribution, lethal factors (specifically building vulnerability), and intensity parameters within the same grid unit. (7) Eliminating outliers, such as unreasonable grids with population density exceeding 5000 people/km², to maintain data consistency.

By analyzing the proportion of housing types at village points, the anti-lethal level of these villages/communities can be determined by proportionally weighting the anti-lethality of each structure type. Subsequently, the location of the village/community is converted into a Voronoi polygon surface vector through neighborhood analysis, and ultimately, it is transformed into a regular kilometer grid with a feature of lethality based on the pixel center principle. Furthermore, when the seismic intensity zoning in China is converted into a grid and tailored to the study area (Figure 8), the grid level is precisely aligned with the ESA population grid (Figure 9) and lethality grid (Figure 10) on a spatial scale of kilometers.

Seismic intensity zoning is closely related to the distribution of faults, which is inseparable from the methodology for determining seismic intensity zoning [56]. As illustrated in Figure 8, Kangding, Luding, and Xichang along the Xianshuihe fault, Anning River fault, and Zemuhe fault in central and southern Sichuan, as well as Songming, Yiliang, Xundian, and Dongchuan along the Xiaojiang Fault in eastern and southwestern Yunnan, and Lancang, Cangyuan, and Gengma along the Longling Lancang fault, are located within the IX seismic intensity zone (colored blue-purple), indicating a probability of occurrence exceeding 10% within a 50-year period.

High seismic intensity areas do not always align with regions of high population density. As illustrated in Figure 9, Panlong and Guandu are adjacent to Songming County, situated within the IX intensity zone and hosting the highest population density in Yunnan Province, reaching up to 100,000 people per square kilometer. However, counties such as Lancang, Cangyuan, and Gengma, also within the IX intensity zone, exhibit a relatively low population density, not surpassing 1000 people per square kilometer. A similar pattern is observed near the Xianshuihe fault in Sichuan, where the population density remains low, not exceeding 500 people per square kilometer. In contrast, the population is more concentrated near the Anning River and Zemu River, with the highest density occurring near the Zemu River, exceeding 17,000 people per square kilometer.

Chinese residents are predominantly concentrated in urban areas, with 63.89% residing in cities and towns, while the remainder is relatively dispersed across rural regions [58] (National Population Statistics Yearbook, 2022). The level of economic development determines that housing in urban areas has higher seismic performance, so high earthquake mortality often avoids areas with high population density. Our mortality model is primarily based on the proportion of local building structures and the lethal levels of each type. Consequently, the lethality of earthquakes is closely linked to the seismic fortification standards of local residences. The sparsely populated regions of Western Sichuan,

Western Yunnan, and Eastern Guizhou exhibit high lethality, a fact that is illustrated in Figures 9 and 10.

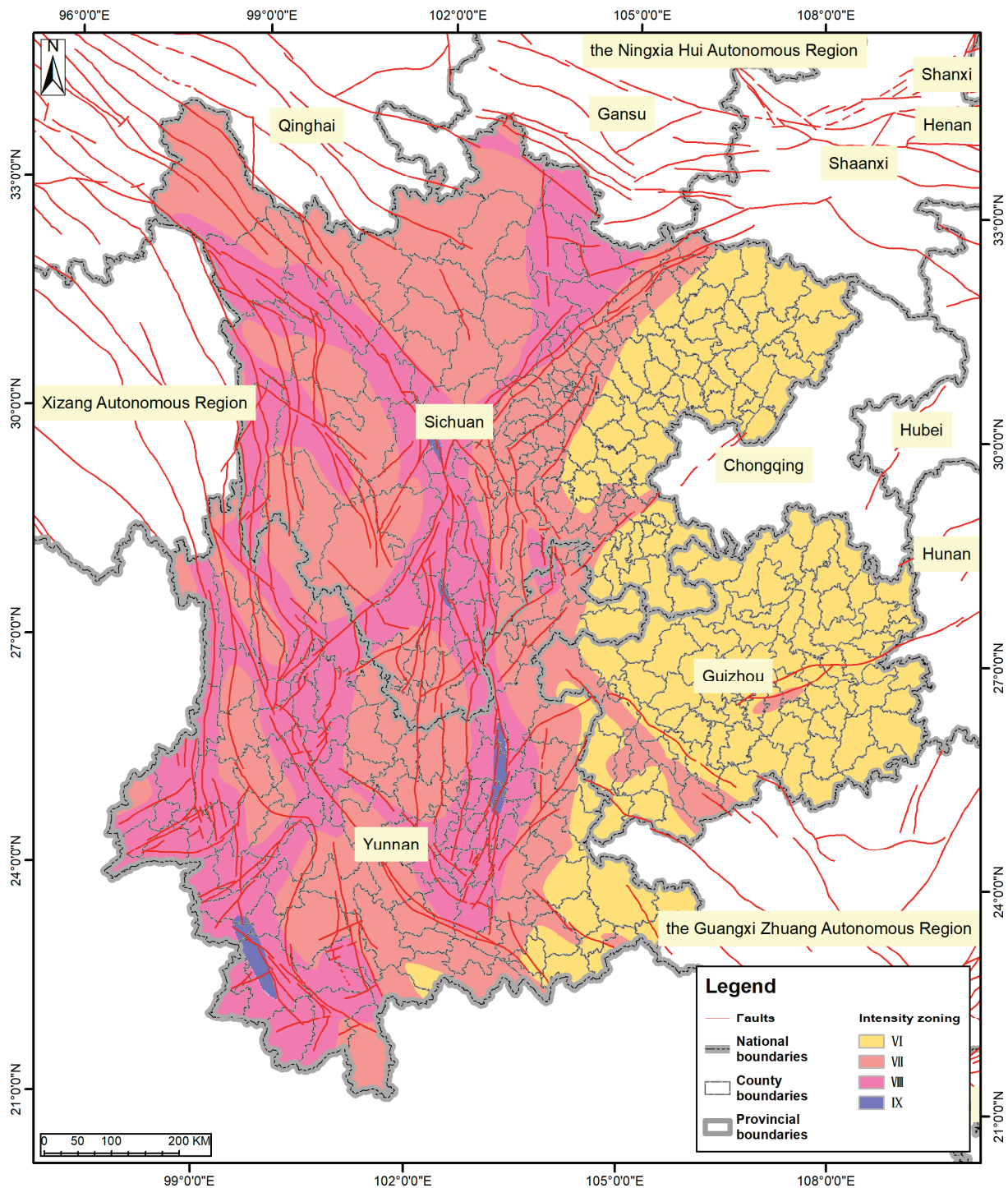


Figure 8. Seismic intensity zoning in Yunnan, Guizhou, and Sichuan with a PGA probability exceeding 10% over a 50-year period.

3.2. The Anti-Lethality/Anti-Injury Matrix

Drawing upon the lethality theory, refs. [36,37] proposed a semi-empirical assessment method based on building vulnerability. The model is constructed via the statistical fitting of the relationship between building vulnerability and fatalities, and it does not belong to the category of empirical function methods that rely on parameters such as seismic

parameters and population density. Figure 11 illustrates the Methodological flowchart depicting the construction of the lethality/injury matrix (vulnerability) and its interaction with intensity parameters (hazard) and human population (exposure) grids in this study. Through an analysis of historical records of fatalities in different seismic magnitudes and intensity zones across various regions of China, two relationship characteristics among seismic magnitude, intensity, and mortality rates were summarized:

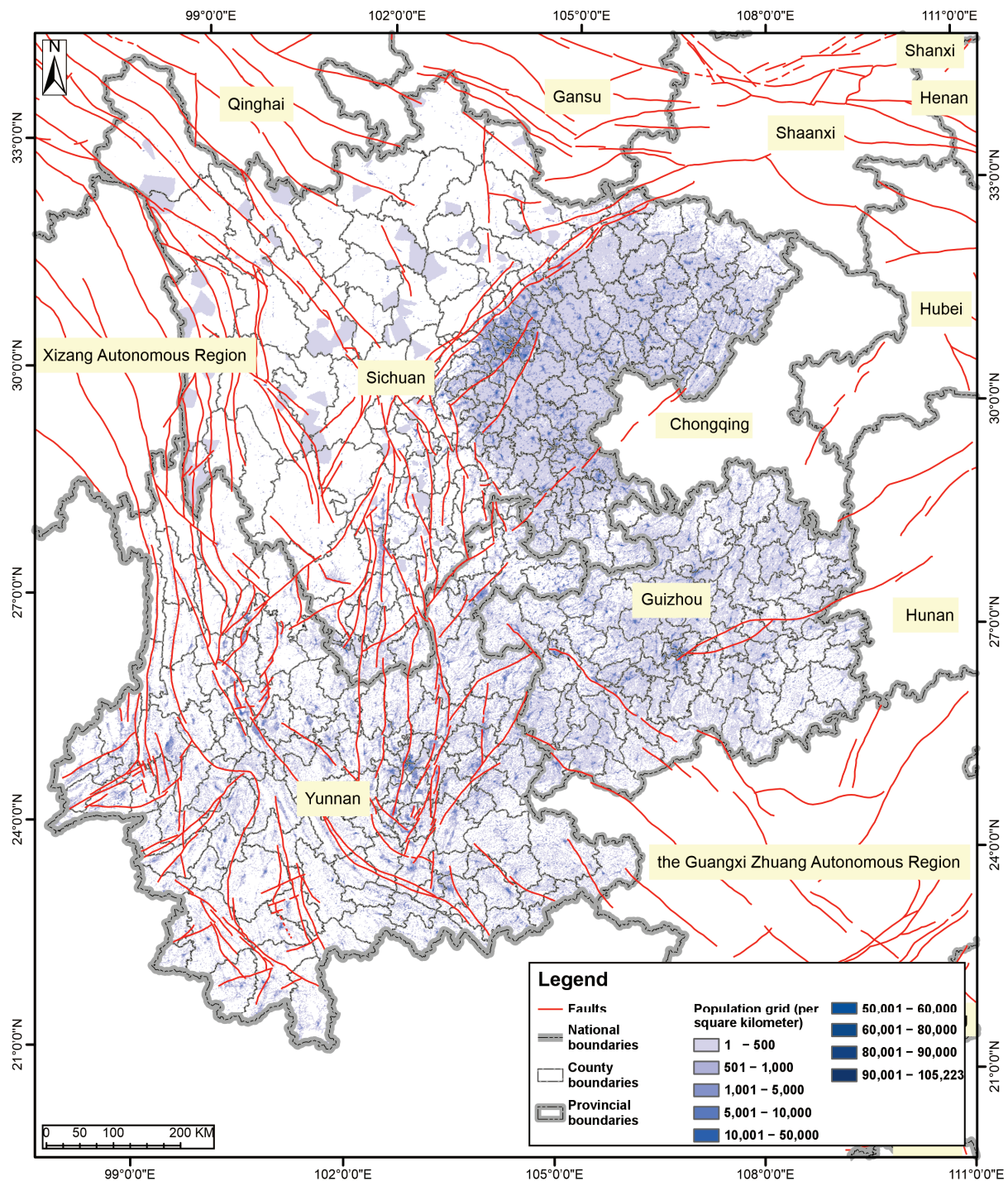


Figure 9. Distribution of population in Yunnan, Guizhou, and Sichuan by 2020.

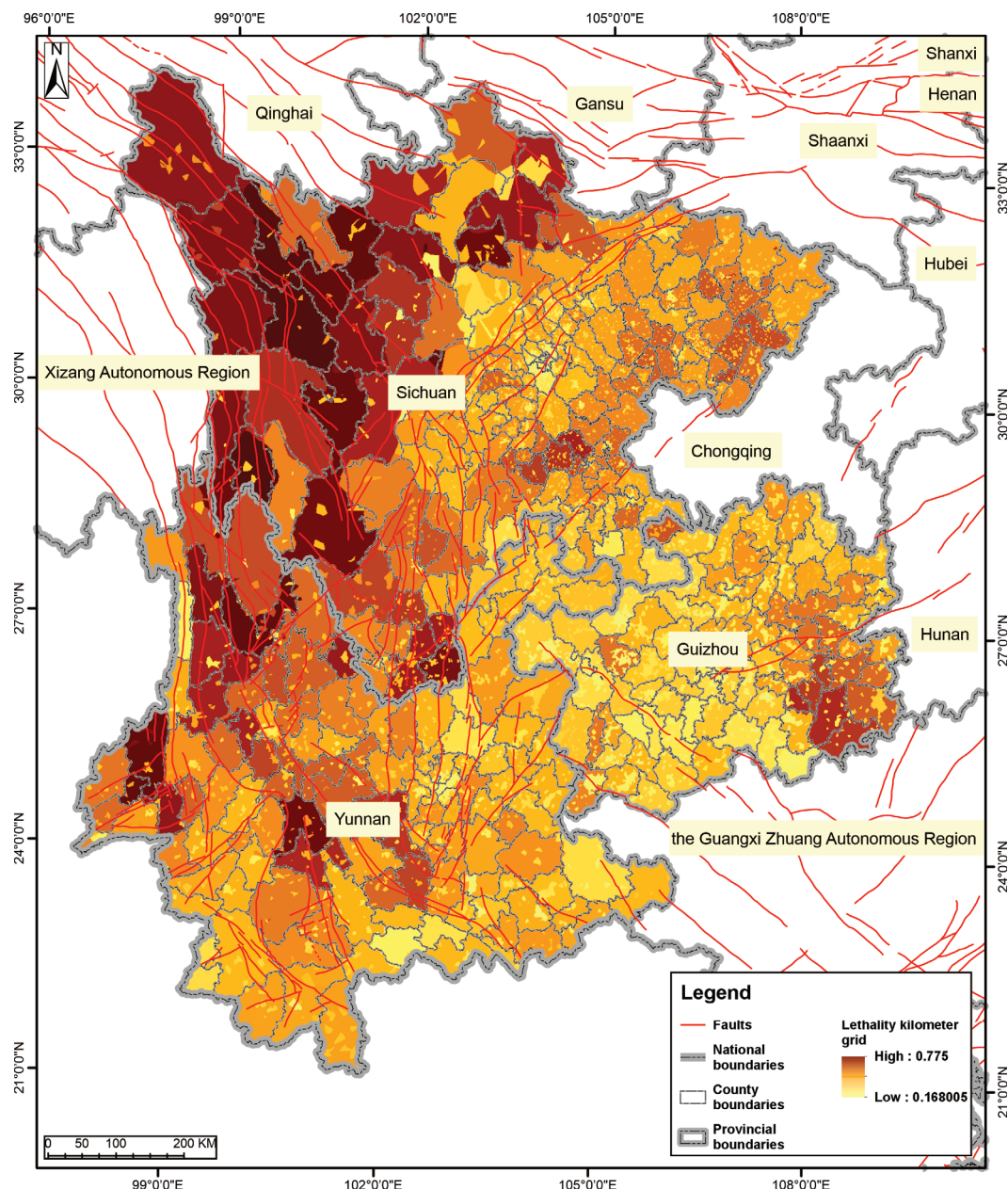


Figure 10. The kilometer grid depicting the lethal level in Yunnan, Guizhou, and Sichuan with a PGA probability exceeding 10% over a 50-year period.

- (1) There is a multiple-relationship characteristic between the mortality rates of adjacent intensity groups. In this study, by summarizing over 30 historical earthquakes in China, the earthquake mortality rate (ranging from 0% to 30%) and probability (ranging from 0% to 100%) were respectively employed as the horizontal and vertical axes. The probability of mortality rates occurring within the 0–30% range corresponding to seismic intensities from VI to XI degrees was fitted. An empirical multiple relationship of 2:4:14:32:39:32 was obtained based on the standard deviations between adjacent intensity groups. Subsequently, the mortality multiples for adjacent intensity groups (VI–VII, VII–VIII, VIII–IX, IX–X, X–XI) across all historical earthquakes from VI to VIII degrees were plotted on the horizontal axis, with probability on the vertical axis, to acquire the discrete distribution of mortality multiples (ranging from –20 to 50) and probabilities (ranging from 0% to 100%) for all historical earthquakes. This process

facilitated the fitting of curve slopes, namely, the slope of the historical earthquake fitting curve.

- (2) The mortality rates of individuals at different seismic intensity groups demonstrate distinct exponential relationships and grouping phenomena. In this characteristic, by summarizing the relationship between intensity (from degree VI to XI) and mortality rate (0–3%) of over 40 historical earthquakes in China and normalizing the correlation curves, it was found that the normalized intensity–mortality rate curves of these historical earthquakes naturally clustered. This suggests that the mortality rates of historical earthquakes in China exhibit natural grouping phenomena across different intensity levels, which can be classified into 11 mortality rate groups; this constitutes the core of the theory: 11 lethality level classifications.

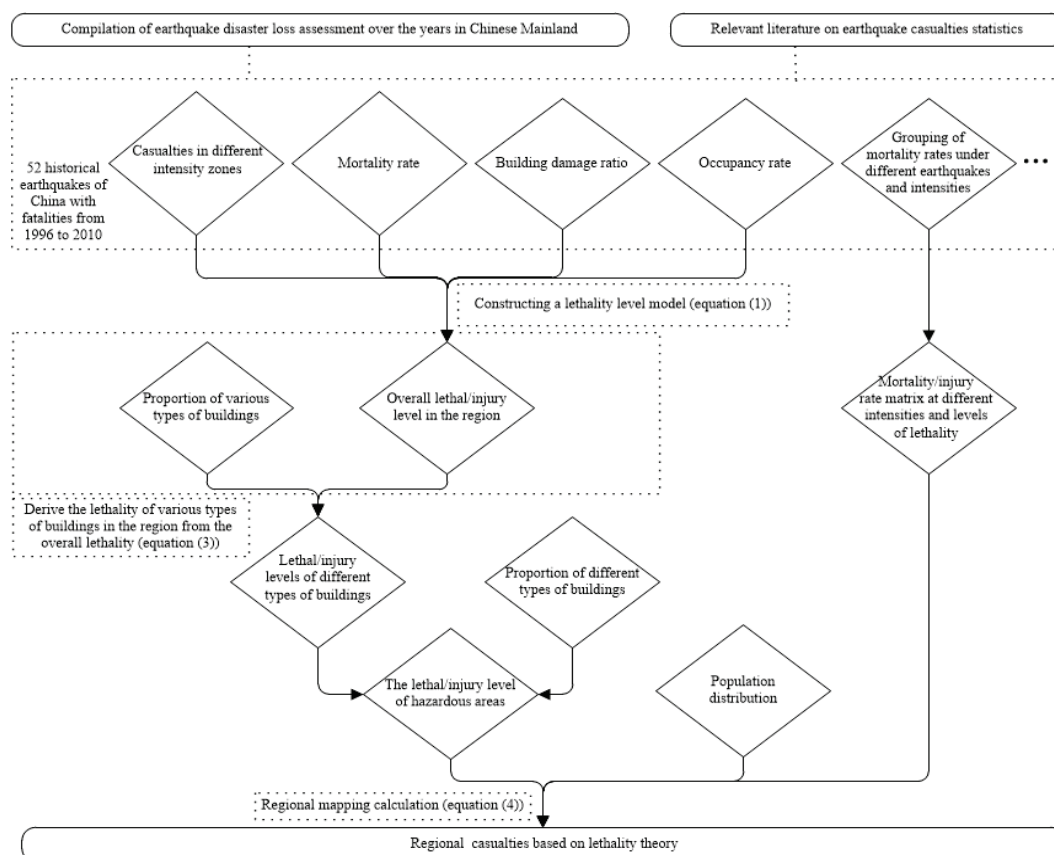


Figure 11. Methodological flowchart depicting the construction of the lethality/injury matrix and its interaction with intensity parameters and human population grids in this study. Dashed-line boxes designate explanations of data types at the equivalent hierarchical level or illustrate processing content as indicated by adjacent black arrows, while solid unidirectional arrows represent directional output relationships.

Following characteristic (2), it is logically inferred that there is a corresponding relationship between the slope of historical earthquake fitting curves, lethality level classifications, and mortality rates. Finally, refs. [36,37] constructed a matrix and evaluation model based on lethality level grades and intensity-specific mortality rates, which reflects the content of Equation (7).

By analyzing historical earthquakes, the overall anti-lethality level of a region can be evaluated. This theory establishes a field investigation lethality level model through field surveys. For example, geological and geomorphological factors are classified into three influencing elements: determining whether the study area contains faults, with the

distances to fault zones being less than 50 km, greater than 50 km, or having no fault zones at all; identifying the presence of steep slopes (slope angle > 10 degrees), gentle slopes (slope angle < 10 degrees), or flat terrain (no slope gradients); and analyzing slope characteristics. Taking Equation (2) as an illustration, if the weight of the geological and geomorphological factor is ω_i , the distribution of fault zones serves as a secondary geological and geomorphological factor with a weight of ω_{ij} , while the presence of faults acts as a tertiary factor with a weight of ω_{jk} .

The lethal level varies with seismic intensity, which is a relative index influenced by the type, proportion, material, quality, and construction method of the building structure. Fundamentally, the lethal level for each building type must be determined based on its specific influencing factors. For comparable buildings, the lethal value should fall within a specified range [36,37].

Upon acquiring the lethality level through field investigations, the lethality levels of diverse local buildings are inversely deduced from the overall lethality level. This procedure employs parameters including architectural structures, foundation types, construction dates, usage purposes, and the lethality levels of residential buildings within the surveyed region. The range of the anti-lethality level for each building type can be determined by constructing and solving a multivariate first-order formula (refer to Table 8).

Table 8. Range and influencing factors of lethal levels in buildings [18,52] (Nie et al. 2020, 2021).

Building Structure Type		Influence Factors	Lethal Level
Building Type	Secondary Classification		
Steel structure		Construction measures, foundation type, construction time, and purpose	0.05–0.15
Frame structure	Fa	Construction measures, foundation type, construction time, height	0.1–0.3
Timber structure	Fb		
	Wa	Construction measures, foundation type, construction time, structural style	0.2–0.4
	Wb		
Brick-concrete structure	Ba (Fortified)	Construction measures, foundation type, construction time, purpose, height	0.25–0.7
	Bb (Non-Fortified)		
Brick-wood structure	Ma	Construction measures, foundation type, construction time, structural style	0.6–0.9
	Mb		
civil structured		Construction measures, foundation type, construction time, building materials	0.7–0.95
Stone (wood) structure		Wall type, foundation type, construction time, building materials	0.55–0.9
Adobe structure		Wall type, foundation type, construction time, building materials	0.85–1.0

By utilizing historical earthquake casualty data corresponding to various seismic intensities (linked to Peak Ground Acceleration and Peak Ground Velocity), we establish the correlation between mortality/injury rates and seismic intensity within the lethality matrix. Based on casualty patterns derived from historical earthquakes in distinct regions (e.g., Southwest and Northwest China), we calibrate region-specific lethality matrices. Through comparative analysis of actual historical earthquake casualties and model-predicted casualties, we iteratively optimize the parameters of the lethality matrix to improve the precision of casualty assessment. By leveraging historical seismic data, the validation of mortality coefficients was supplemented as presented in Table 9.

Table 9. Validation of this model using certain historical earthquakes in China [36,37].

Location of Earthquake	Time (Year–Month–Day)	Magnitude	Lethal Grade	Actual Casualty	Calculated Casualty	Error	Pre-Estimated Casualty
Lancang Gengma, Yunnan	6 November 1988	7.6	30%	748	783	4.68%	700–850
Ning'er, Yunnan	3 June 2007	6.4	70%	3	3.87	29.00%	0–5
Yushu, Qinghai	14 April 2010	7.1	0%	2698	2197	−18.57%	1800–3000

Table 9. Cont.

Location of Earthquake	Time (Year–Month–Day)	Magnitude	Lethal Grade	Actual Casualty	Calculated Casualty	Error	Pre-Estimated Casualty
Lushan, Sichuan	20 April 2013	7.0	80%	196	165	−15.82%	150–200
Min County-Zhang, Gansu	22 July 2013	6.6	40%	95 + 5	98	−2.00%	85–120
Ludian, Yunnan	3 August 2014	6.5	10%	617	596	−3.40%	550–700
Jinggu, Yunnan	7 October 2014	6.6	80%	1	3.07	precise	0–5
Jiuzhaigou, Sichuan	8 August 2017	7.0	60%	29	25.98	10.41%	10–40

After identifying the factors influencing regional lethality and determining the weight ratio of each factor in the first and second levels of classification, the overall lethality model can be constructed using the weight ratio coefficients of each factor, as depicted in Equation (1) [6,41,61–65]. Furthermore, the regional lethal level coefficient adjustment model is presented in Formula (2) [6]. For the definition of each parameter, refer to [6,36,37,41].

In this model, the 11 levels of lethality ranged from 0 to 1. The anti-lethal/anti-injury matrix is generated according to the intensity levels VI to XI. Intuitively, the horizontal axis and the vertical axis represent the earthquake intensity and anti-lethal/anti-injury levels respectively, pointing to the unique casualty rate value. Therefore, to understand the lethal level of an area, we can directly refer to the casualty rate, and the error is guaranteed to be within $\pm 50\%$ [36,37].

$$\log RD = 12.479(P_t \cdot \rho \cdot RA \cdot \alpha)^{0.1} - 13.3 \quad (1)$$

$$\partial_{adjust} = \left(\sum_{i=1}^3 \omega_i \cdot \omega_{ij} \cdot \omega_{jk} \right)^{-1} \quad (2)$$

3.3. Calculation

Initially, we calculate the proportion of households residing in various building structures at the county/district level using the specific household numbers from the census data. Subsequently, we apply these ratios as coefficients and integrate them with the anti-lethal/anti-injury levels of each structure type to ascertain the county/district-level overall anti-lethal level. As outlined in Section 3.1, we link the anti-lethal/anti-injury table of each county/district to the village vector attribute based on the administrative code, ensuring that each village/community is assigned a distinct value. The primary objective when integrating data from various sources is to convert county-level household statistics into kilometer-level data. This conversion allows us to correlate the data with population density and intensity for further analysis. However, since the intensity division relies on plane vector data, it must be transformed into grid format while preserving its fundamental characteristics. It is crucial to maintain the consistency of the coordinate system and use the target area as a template. When creating a raster, accurate edge capture is essential. Additionally, because the population data are in grid format, they must be trimmed to fit the boundary of the target area. Emphasis should be placed on the coordinate system to ensure that the edges are aligned with the edges of other data by-products. After aligning the data, each grid becomes our target, allowing us to calculate the matrix by stacking earthquake intensity zoning (Figure 8), population distribution (Figure 9), and lethal level (Figure 10). Based on the mortality and population size corresponding to the overall lethal level in the region, the mortality/injury estimation model is established. The anti-lethality/anti-injury matrix introduced in Section 3.2 is then applied to perform the conditional allocation operation for each grid. If the criteria of earthquake intensity and lethality are met, the mortality/injury rate is determined. Finally, through spatial analysis, such as multiplication with the population, the number of casualties in each administrative region is estimated through aggregation.

The prepared data in this study of the logical operational approach regarding the casualty rate within grid cells comprises: spatial population data (e.g., 1 km × 1 km) measured in population density (people/km²), building structure types and their proportions (including brick, wood, earth, bamboo-straw adobe, brick-concrete, reinforced concrete, etc., presented as percentages); peak ground acceleration (PGA) of grid cells, China seismic intensity (in accordance with standards [56,60]), etc.; regional fatality/injury matrices: basic fatality matrices constructed based on study regions (such as the southwest and northwest) with seismic motion parameters (e.g., PGA, seismic intensity) as indices (%). Operational calculation steps include the following:

- (1) Acquire the peak ground acceleration (PGA) or seismic intensity values of grid cells via seismic motion simulation (in conjunction with rapid reporting information and real-time social perception data), which serve as key indices for matching regional fatality/injury matrices.
- (2) Based on the location area of the grid cell (e.g., a sub-region in Sichuan, Yunnan, Guizhou), access the corresponding “regional fatality/injury matrix” and utilize the seismic motion parameters obtained in Step 1 (e.g., PGA = 0.2 g or seismic intensity VIII degree) to retrieve the basic mortality rate of the grid cell (in percentage) L_0 .
- (3) Compute the structural fragility coefficient V according to the proportion of different building structures within grid cells:

$$V = \sum_{i=1}^n (P_i \times V_i) \quad (3)$$

where P_i denotes the proportion of structural types in category i (e.g., 30% for brick structures, hence $P_i = 0.3$), and V_i represents the fragility weight of category i structures (based on seismic performance surveys of regional buildings: $V_i = 1.5$ for bamboo-wattle adobe houses and $V_i = 0.3$ for reinforced concrete buildings, with higher values indicating greater fragility). The adjusted mortality rate is calculated as $L_1 = L_0 \times V$.

- (4) Using spatialized population density data (D , people/km²) from grid cells, calculate the total population within the cell $N = D \times S$ (where S denotes grid area; e.g., $S = 1$ for 1 km²).
- (5) The casualty rate R (%) of the grid cell serves as a comprehensive metric that combines adjusted mortality rates and population distribution N_{exposed} denotes the actual population exposed to seismic risk within the grid cell (with a default value of [value not provided N], which can be adjusted according to diurnal population dynamics).

$$R = L_1 \times \frac{N_{\text{exposed}}}{N} \quad (4)$$

- (6) Considering the nonlinear relationship between casualties and Peak Ground Acceleration (PGA), a nonlinear correction coefficient K should be employed when PGA or seismic intensity surpasses the threshold (e.g., PGA > 0.3 g). (Based on historical data fitting, for instance,

$$K = 1 + 0.5 \times \ln(\text{PGA}/0.3) \quad (5)$$

The final case mortality rate was adjusted by quation

$$L_{\text{final}} = L_1 \times K \quad (6)$$

The values are substituted into Step 5 to compute the casualty rate. The casualty rate R (%) of a grid cell denotes the proportion of estimated casualties within that cell relative to the total population.

In Equation (7), D_d represents the number of deaths, m is the anti-lethal level, defined as 1 minus the lethal level, n represents the seismic intensity, and L_{mn} and D_{mn} correspond to the mortality rate and population size of the area with the intensity of n and the lethal level of M [6,61,64].

$$D_d = \sum_{m=0}^1 \sum_{n=VI}^{XI} (L_{mn} \cdot D_{mn}) \quad (7)$$

3.4. Sources and Validation of Casualty Estimation Errors: An Analysis of Allowable Error Thresholds

The uncertainty inherent in casualty predictions generated by models primarily originates from three categories of errors: data input errors, model parameter errors, and model structural uncertainties. Specifically:

- (1) Data input errors encompass spatialization inaccuracies in demographic data and imprecisions in building typology data. While spatialized population data can improve accuracy, contemporary population statistics frequently depend on administrative divisions (e.g., county-level units), resulting in uneven spatial distributions at grid scales (e.g., 1 km × 1 km). Assuming that interpolation errors for population density adhere to a normal distribution, a standard deviation σ_1 of approximately $\pm 30\%$ corresponds to common error ranges observed in remote sensing-derived population data (e.g., errors from nighttime light interpretation or sample surveys). In southwestern China, brick-wood and earth-wood structures constitute over 50% of the building stock, whereas reinforced concrete structures represent less than 20% (Conclusion 5). However, county-level statistical classifications may demonstrate ambiguity (e.g., indistinct boundaries between “brick-wood” and “earth-wood” structural categories), introducing errors in seismic performance parameters. A uniform distribution error range of $\pm 25\%$ σ_2 is assumed.
- (2) Model parameter errors, specifically the uncertainty associated with regional mortality/injury matrices. Existing matrix parameters are derived from limited samples (e.g., an insufficient number of historical earthquake cases), inducing biases in mortality-intensity relationships (such as those between peak ground acceleration and casualty rates). The standard error of parameter estimation is assumed to be σ_3 approximately $\pm 30\%$, based on t-distribution errors arising from small-sample statistical inference.
- (3) Model structural uncertainties, which may involve the underestimation of secondary disasters and soft impacts. The model does not comprehensively incorporate secondary disasters such as landslides and rockfalls, nor does it account for indirect losses stemming from public opinion dynamics. These omissions may result in an underestimation of casualties by 10% to 20%. A systematic bias of σ_4 approximately $\pm 20\%$ is therefore assumed.

Applying the root mean square error propagation theory (which superposes independent error sources), the total standard deviation σ_{total} is calculated as the square root of the sum of the squares of all individual error sources (Equation (8)):

$$\sigma_{total} = \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2} \quad (8)$$

Upon substituting the aforementioned error values and accounting for the non-normal distribution of errors, as exemplified by the uniform distribution of parameter errors, the actual total error σ_{all} range undergoes a slight narrowing, ultimately converging to $\pm 50\%$ (Equation (9)).

$$\sigma_{all} = \sqrt{(0.3)^2 + (0.25)^2 + (0.3)^2 + (0.2)^2} = \sqrt{0.09 + 0.0625 + 0.09 + 0.04} = \sqrt{0.2825} \approx 0.53, (\pm 53\%) \quad (9)$$

By employing 10,000 Monte Carlo simulations, random sampling was executed for various error sources: population error ($N(0, 0.3^2)$), building data error ($U(-0.25, 0.25)$), parameter error ($N(0, 0.3^2)$), and structural error ($N(0, 0.2^2)$). The simulation results indicated that the 90% confidence interval for predicted casualties spanned a range of $[-48.7\%, +51.2\%]$, encompassing $\pm 50\%$ of the total variation, thereby validating the rationality of the error margin. International earthquake casualty prediction models, such as FEMA HAZUS and the European SHARE model, typically demonstrate error margins of $\pm 40\%$ to $\pm 60\%$ in data-scarce developing countries, as evidenced by post-2008 Wenchuan earthquake model validation. The study area, encompassing Yunnan, Guizhou, and Sichuan, is characterized by low seismic resistance in buildings and sparse data availability. The $\pm 50\%$ error range obtained in this study is consistent with findings from comparable research.

4. Result

4.1. Casualty Assessment Under a Basic Ground Motion Level

Our initial focus is on regions where casualties are concentrated, employing the method detailed in Section 3.3 to derive earthquake casualty assessments for southwestern China. Figure 12b and Figure 14b depicts the assessment of fatalities and injuries with a PGA probability exceeding 10% over a 50-year period. Upon comparing Figure 8, Figure 9, Figure 12b and Figure 14b, it becomes evident that since casualty numbers are calculated based on a population grid matrix, the distribution of casualties inherently relies on population density. Nonetheless, it is clear that the intensity of ground motion alone cannot fully dictate the density of casualties; the lethality level must also be taken into account. Thus, the distribution of casualties largely hinges on these fundamental factors: the intensity of the disaster, the resilience of the affected population, and the extent of population exposure.

Figure 12b and Figure 14b show macroscopically that the regions with the highest casualty density in Sichuan are located in the southern and central eastern areas. Given the vast territory and sparse population of the Chuanxi Mountain region, the distribution of casualties is extremely scattered. Apart from the northwest and southeast of Yunnan, where the casualty density per kilometer is relatively low due to sparse population or less fault activity, the distribution of casualties in other areas of Yunnan is uneven and varies significantly. With the exception of a few residents in the eastern and western mountainous regions, the distribution of residents in Guizhou is relatively even, and the casualty density is generally low, mostly between 1 and 5 per square kilometers.

To understand the distribution of county-level casualties in southwestern China, we utilized a regional geometric analysis method to obtain the total, ranking, average, and standard deviation of grid casualties within each county partition (Figure 13b and Figure 15b). Tables 10 and 11 present the casualty assessment for various counties in southwestern China, indicating a probability exceeding 10% within the next 50 years.

Table 10. List of counties and corresponding emergency response levels predicted for seismic deaths exceeding 10% probability over a 50-year period in southwestern China.

Predicted Death Toll	Emergency Response Level	Corresponding County Counts
2000~3000	I	1 (Xichang City)
1000~1999	I	1 (Lancang County)
500~999	I	3 (Xundian/Songming/Dongchuan)
100~499	II	44
10~99	III	159
1~10	IV	192

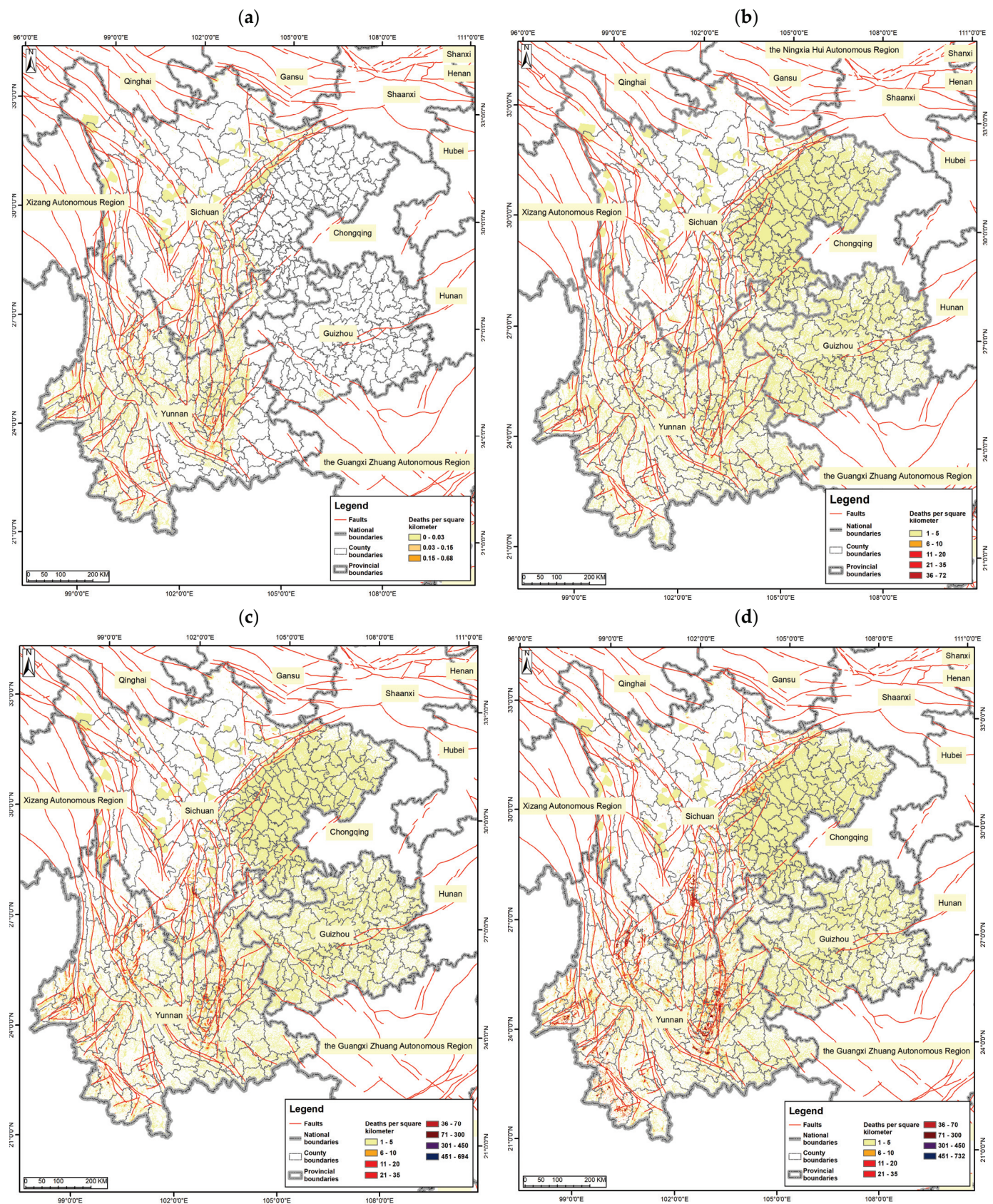


Figure 12. The estimated kilometer grid of earthquake fatalities in southwestern China with (a) a 50-year probability of occurrence of 63%; (b) a 50-year probability of occurrence of 10%; (c) a 50-year probability of occurrence of 2%; and (d) a one-year probability of occurrence of 10^{-4} .

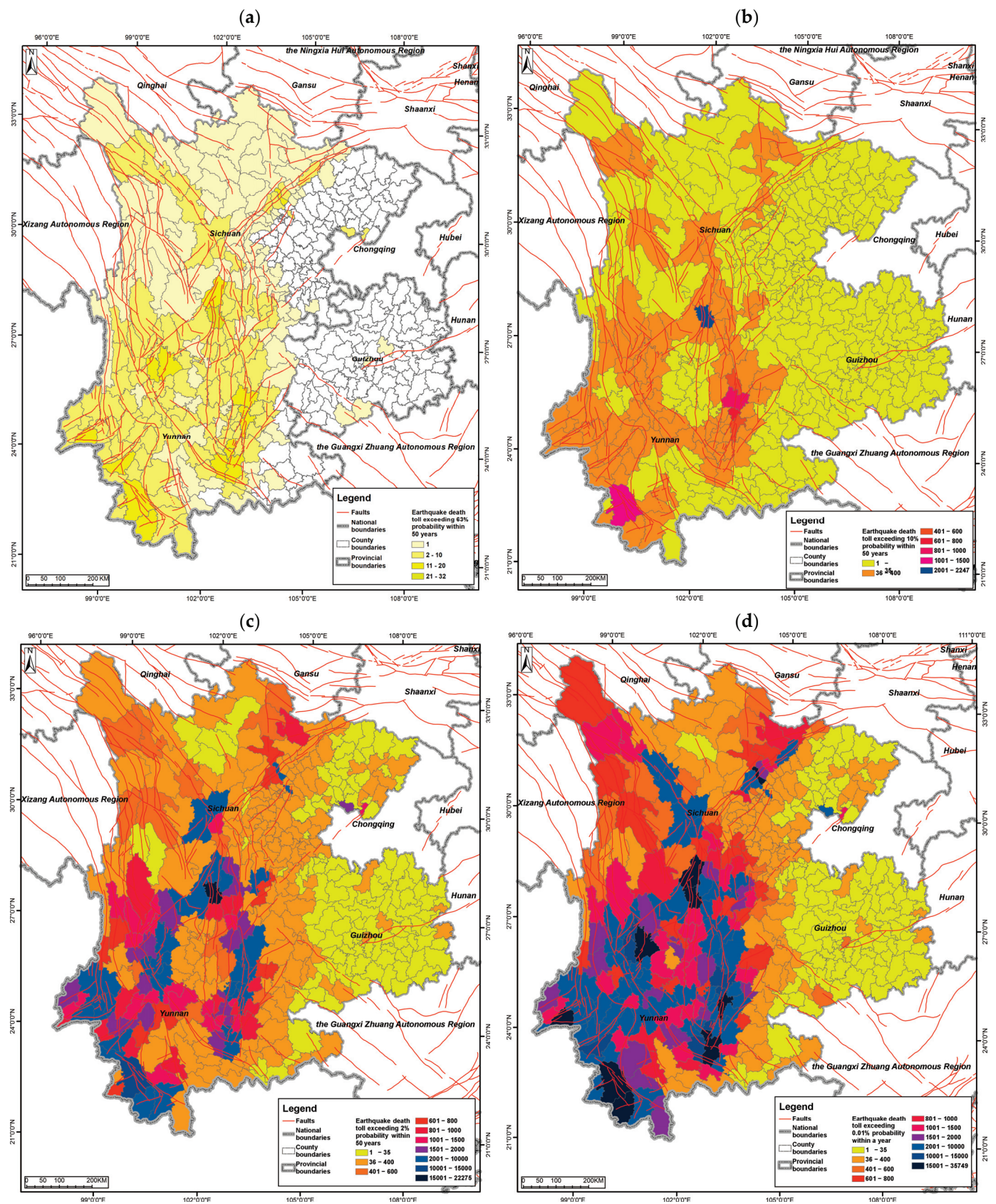


Figure 13. The estimated earthquake fatalities by counties of southwestern China with (a) a 50-year probability of occurrence of 63%; (b) a 50-year probability of occurrence of 10%; (c) a 50-year probability of occurrence of 2%; and (d) a one-year probability of occurrence of 10^{-4} .

Table 11. List of counties corresponding to the predicted number of seismic injuries with a PGA probability exceeding 10% over a 50-year period in southwestern China.

Predicted Number of Injured	Corresponding County Counts
10,000~20,000	1 (Xichang City)
5000~9999	7 (Lancang/Tengchong/Longyang/Xundian/Songming/Yiliang//Panlong)
2000~4999	47
1000~1999	26
500~999	26
100~499	65
0~100	228

When the 50-year PGA exceedance probability reaches 10%, Xichang is estimated to incur 2247 fatalities and 17,297 injuries. With a population of 817,282, this translates to a mortality rate of 0.27% and an injury rate of 2.12%. Lancang Lahu Autonomous County, with a population of 540,469, is projected to have 1243 deaths and 9863 injuries, resulting in mortality and injury rates of 0.23% and 1.82%, respectively. Counties with estimated death tolls ranging from 300 to 1000 include the Xundian Hui and Yi Autonomous Region, Songming, Dongchuan, and others. Three counties—Xundian, Songming, and Dongchuan—have a combined death toll of 2170, with individual counts between 500 and 999. 44 counties report 100–499 deaths, summing up to 7420 people. Additionally, 159 counties experience 10–99 deaths, totaling 5202 people. In 192 counties, there are fewer than 10 deaths, with a combined total of 578. Counties expected to have between 5000 and 9000 injured persons include Tengchong, Longyang, Xundian Hui and Yi Autonomous Counties, among others. In seven counties, 49,188 people have been injured, with the number of injuries ranging from 5000 to 9999: Lancang, Tengchong, Longyang, and others. The number of injured people in 47 counties ranges from 2000 to 4999, totaling 141,236 people. A total of 37,628 cases are reported in 26 counties. In 26 counties, 20,285 people are injured, with the number of injuries ranging from 500 to 999. In 65 counties, 14,553 people are injured, with the number of injuries ranging from 100 to 499. In 228 counties, 5746 people are injured, with the number of injuries being less than 100. By calculating the area of each county, we determine the death toll per square kilometer. Seven counties have a death toll exceeding 0.5 people per square kilometer, and 2 counties have a death toll exceeding 1 person per square kilometer, namely Xichang and Songming. The county with the highest number of injuries per square kilometer is still Xichang, with 20 people. There are three counties with a casualty density of 10–20 persons/km²: Wuhua, Panlong, and Songming, with Wuhua covering an area of only 255 km². There are 101 counties with a casualty density of 1–9 persons/km². Across Yunnan, Guizhou, and Sichuan, with a combined population of 179,196,903, the estimated fatalities are 18,861 and the estimated injuries are 285,938, leading to a total mortality rate of 0.011% and a total injury rate of 0.16%.

A Chinese national standard [66] GB/T 29425-2012, establishes the unified basis, activation conditions, and disposal requirements for the classification of emergency responses in China, thereby providing a consistent framework for action to governments and departments at all levels. This standard specifies that, as a category of natural disaster, earthquake events necessitate the classification of emergency response levels based on a comprehensive assessment of factors including casualties, property losses, social impact, and disaster relief capabilities, with casualties serving as the core criterion. Specific thresholds should be formulated through differentiated standards tailored to local conditions, such as economic development levels and disaster relief capacities; the standard does not separately prescribe explicit casualty thresholds for earthquakes. In light of this framework, and according to the National Earthquake Emergency Plan issued by the General Office of the State Council in

2020, earthquake disasters are categorized into four distinct levels: (1) particularly serious earthquakes that result in over 500 fatalities in key areas or those with a magnitude of 7.5 or greater; (2) major earthquakes causing 100–500 deaths or those with magnitudes ranging from 7.0 to 7.5 in key areas; (3) significant earthquakes resulting in 10–100 deaths or with magnitudes between 6.0 and 7.0 in key areas; and (4) general earthquakes causing fewer than 10 deaths or with magnitudes between 5.0 and 6.0 in key areas. The corresponding emergency response is classified into four grades: I, II, III, and IV. Based on our assessment and the earthquake emergency plan, we have identified counties with a likelihood exceeding 10% of reaching the first level of earthquake emergency response with over a 50-year period, including Xichang, Lancang, Xundian, Songming, and Dongchuan (refer to Tables 10 and 11).

4.2. Earthquake Casualty Assessment with Various PGA Exceedance Probabilities

According to standard [56], ground motions with a 50-year probability of exceeding 10% are classified as basic ground motions. The peak acceleration of frequent ground motion should not be less than one-third of the basic peak acceleration of ground motion; the peak acceleration of rare ground motion should be 1.6-to-2.3 times the basic peak acceleration of ground motion; and the peak acceleration of extremely rare ground motion should be 2.7-to-3.2 times the basic peak acceleration of ground motion. By applying the mathematical relationship between basic ground motion and frequent ground motion (with a 50-year exceedance probability of 63%), rare ground motion (50-year exceedance probability of 2%), and extremely rare ground motion (1-year exceedance probability of 10^{-4}), the basic ground motion PGA grid is transformed into grids representing frequent, rare, and extremely rare ground motion PGA values. Subsequently, the PGA grid is converted into an intensity matrix using the correlation between PGA and class II site intensity (Table 3). Utilizing the method outlined in this paper, the population, earthquake intensity, and the lethal level values for the provinces of Yunnan, Guizhou, and Sichuan were calculated on the grid, yielding earthquake fatalities number (Figures 12 and 13), injured number (Figures 14 and 15) for probabilities of exceedance at 50 years (63%), 50 years (2%), and 1 year (10^{-4}).

The estimated casualties for each county within the hazardous zones of Yunnan, Guizhou, and Sichuan, based on a ground motion level with a 63% probability over 50 years, indicate that fatalities are predominantly concentrated in almost all of Yunnan, a small region of Western Guizhou, and parts of central and northwestern Sichuan (Figures 13a and 15a). Injuries are primarily found in several counties in Western and Eastern Yunnan, as well as in Central Sichuan. Quantitatively, the highest estimated death toll is 31 in Xichang City, with the death toll in other counties expected to be fewer than 20. The maximum estimated number of injuries is 206 in Xichang. Counties with over 100 injured people also include Yiliang, Lancang Lahu Autonomous County, Mangshi and others. The mortality and injury rates for Xichang City (population 817,282) are 3.79×10^{-5} and 2.52×10^{-4} , respectively. With a ground motion level having a 63% probability over 50 years, the anticipated total number of fatalities and injuries in Yunnan, Guizhou, and Sichuan is 580 and 3052, respectively, which accounts for 3.24×10^{-6} and 1.7×10^{-5} of the total population of Yunnan, Guizhou, and Sichuan (179,196,903).

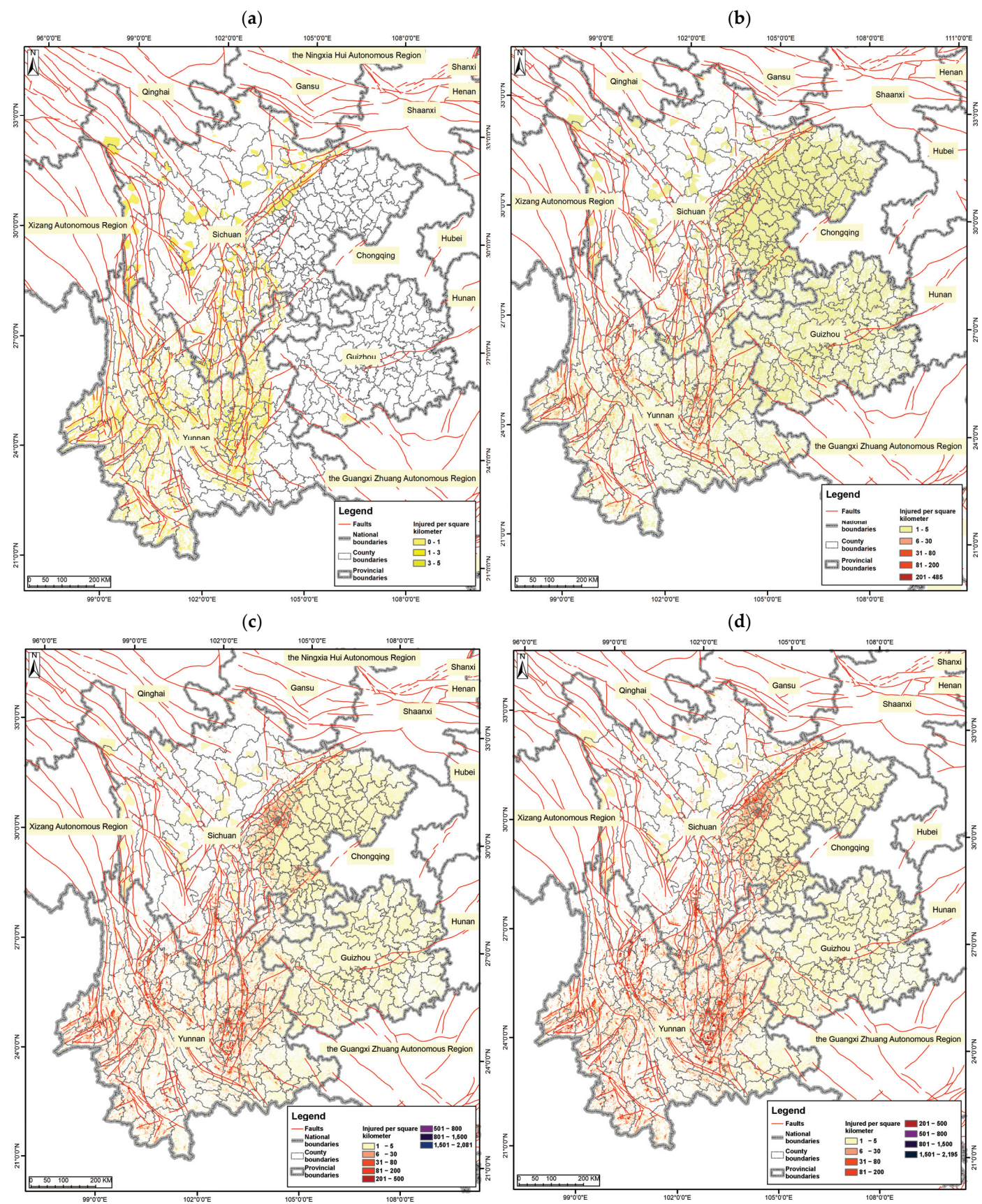


Figure 14. The estimated kilometer grid of earthquake injuries in southwestern China with (a) a 50-year probability of occurrence of 63%; (b) a 50-year probability of occurrence of 10%; (c) a 50-year probability of occurrence of 2%; and (d) a one-year probability of occurrence of 10^{-4} .

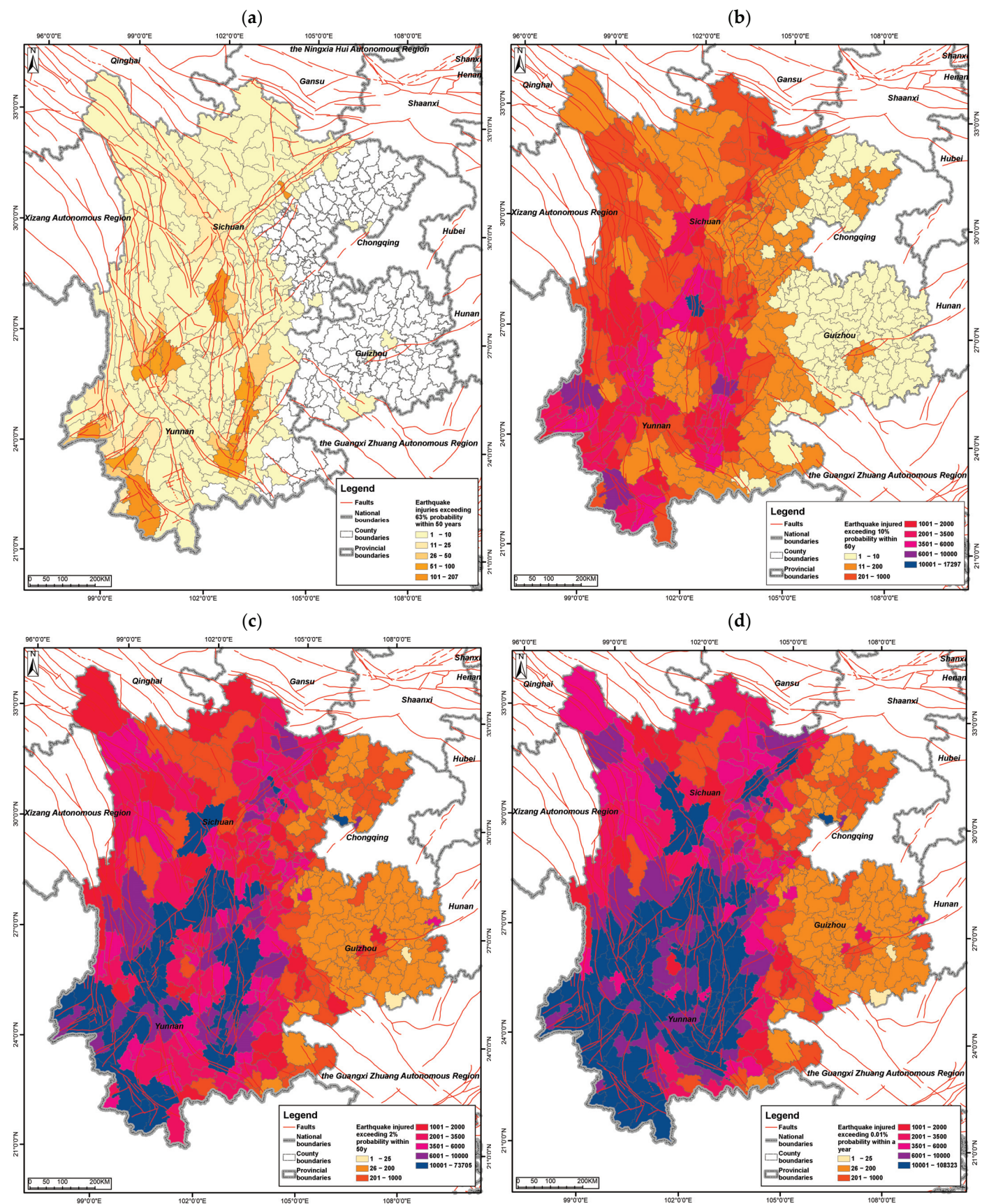


Figure 15. The estimated earthquake injuries by counties of southwestern China with (a) a 50-year probability of occurrence of 63%; (b) a 50-year probability of occurrence of 10%; (c) a 50-year probability of occurrence of 2%; and (d) a one-year probability of occurrence of 10^{-4} .

The estimated casualties for each county in the hazardous zones of Yunnan, Guizhou, and Sichuan, based on a ground motion level with a 2% probability over 50 years, indicate that the distribution of fatalities under this seismic probability is concentrated across nearly all of Yunnan, western Guizhou, and central and western Sichuan (Figures 13b and 15b). The distribution of injuries in these three provinces exhibits a similar pattern. The estimated number of fatalities in Southeast Yunnan, northern Sichuan, Eastern Sichuan, and most counties in central and eastern Guizhou is projected to be fewer than 100, with the estimated number of injuries being less than 1000. It is projected that Xichang City (population 817,282) will have the highest number of fatalities at 22,275, followed by Lancang Lahu Autonomous County (population 540,469) with 12,486 fatalities. 6 counties, including Xundian Hui and Yi Autonomous County, Songming, Dongchuan, Tengchong, Qingbaijiang, and Yiliang are also expected to have over 5000 fatalities. There are 25 counties with a projected death toll ranging from 2000 to 5000. The highest number of injuries is projected for Xichang City at 73,704, followed by Lancang Lahu Autonomous County at 42,705. Counties with projected injury numbers ranging from 20,000 to 40,000 include Tengchong, Qingbaijiang, Longyang, Xundian Hui and Yi Autonomous County, Songming, Panlong, and Yiliang. There are 43 counties with projected injury numbers ranging from 10,000 to 20,000. Under the seismic ground motion level with a 2% probability over 50 years, it is anticipated that there will be 240,580 fatalities and 1,618,702 injuries in Yunnan, Guizhou, and Sichuan, which accounts for 0.13% and 0.90% of the total population of Yunnan, Guizhou, and Sichuan (179,196,903 people), respectively. The estimated mortality rates for Xichang City and Lancang Lahu Autonomous County are 2.73% and 2.31%, and the casualty rates are 9.02% and 7.90%.

The estimated casualties for each county in Yunnan, Guizhou, and Sichuan, based on a ground motion level with a one-year probability of occurrence of 10^{-4} , indicates that the distribution of fatalities and injuries under this ground motion probability is comparable to that under a 50-year probability of occurrence of 2% (Figures 13c and 15c). Apart from a higher number of casualties in the high-risk zones, the estimated number of fatalities at the county level in Southeast Yunnan, Northern Sichuan, Eastern Sichuan, and most counties in Central and Eastern Guizhou remains within 100, and the estimated number of injuries stays below 1000. In other words, for seismic background areas with low risk, even if the probability of high-intensity earthquakes is increased, the magnitude of casualties in these areas cannot be qualitatively altered. However, for high-risk areas, the increase in the casualty rate due to the rise in intensity is very evident. Under the earthquake ground motion level with a 1-year probability of occurrence of 10^{-4} , the highest number of fatalities is still expected to be 35,749 in Xichang City, with 7 counties, including Lancang Lahu Autonomous County, Yiliang, Mianning, Mangshi, Yongsheng, Tonghai, and Shifang anticipating a death toll of 20,000 to 30,000. There are 9 counties with a death toll ranging from 10,000 to 20,000 and 17 counties with a death toll ranging from 5000 to 10,000. The highest number of injuries is expected to be 108,322 in Xichang City. Counties with injury numbers ranging from 50,000 to 70,000 include Lancang Lahu Autonomous County, Tonghai, Shifang, Yiliang, Mianning, Mangshi, Yongsheng, Menghai, and Jianshui, and 28 counties with injury numbers ranging from 20,000 to 50,000. Under the earthquake ground motion level with a one-year probability of occurrence of 10^{-4} , it is anticipated that there will be 632,409 fatalities and 2,961,677 injuries in Yunnan, Guizhou, and Sichuan, which accounts for 0.35% and 1.65% of the total population of Yunnan, Guizhou, and Sichuan, respectively. The estimated mortality and casualty rates in Xichang were 4.37% and 13.25%, respectively.

Upon comparing the casualties kilometer grid (Figures 12 and 14) with the number of casualties in counties (Figures 13 and 15) of Yunnan, Guizhou, and Sichuan provinces

under various exceedance probabilities, we can observe that the kilometer grid with a 63% exceedance probability over 50 years has a mortality rate of less than 1 person per kilometer, and the maximum number of injuries per kilometer is only 5, with most regions having fewer than 1 injury per kilometer. Fatalities are concentrated in central and southwestern Yunnan, central and southern Sichuan, and a small part of the western Sichuan, while the risk of casualties in northeastern and eastern Yunnan and most parts of Guizhou is very low. Secondly, under a 10% probability of exceedance over 50 years, the most severe lethal area, specifically Xichang city where the southern Sichuan, Anning River fault, and Zemu River fault intersect, has reached 72 deaths and 485 injuries per kilometer. However, areas with 10–70 deaths per kilometer are very sparse, and the entire three provinces face a risk of 10 deaths per kilometer. Wuhua, Panlong, Chenggong, Xishan, Anning, and Jiangchuan in Yunnan Province, located to the west of the Xiaojiang Fault, near the Longjiang fault and Nujiang fault, also have a high risk of casualties, resulting in 50–500 injuries per kilometer. Furthermore, if the 50-year exceedance probability is 2% and the 1-year exceedance probability is 10^{-4} , as earthquake intensity increases, the area of casualties does not expand significantly, but the number of casualties increases within the kilometer grid of the aforementioned dangerous areas. The maximum number of deaths per kilometer under a 50-year exceedance probability of 2% and a 1-year exceedance probability of 10^{-4} is approximately 700, with little difference. However, under a 1-year exceedance probability of 10^{-4} , the distribution of death points is notably dense. Evidently, when the preset intensity reaches this level, the number of earthquake fatalities per square kilometer does not exhibit a significant increase. However, the disaster-affected areas with this level of lethality expand substantially. This leads to a marked rise in the number of casualties presented by the expanded casualty grid and a denser distribution of the casualty grid. Additionally, the comparison of the estimated total casualties population in Yunnan, Guizhou, and Sichuan under the four exceedance probabilities indicates that after the exceedance probabilities were reduced by approximately 6 times, 5 times, and 4 times respectively, the estimated death toll in the three provinces increased by 33 times, 13 times, and 2.6 times respectively, from 580 to 18,861, to 240,580, and to the final 632,409; the number of injured people increased from 3052 to 285,938, then to 1,618,702, and finally to 2,961,677, an increase of 93 times, 5.6 times, and 1.8 times, respectively. It is clear that the growth rate of the number of casualties does not have an ordinary linear relationship with the growth rate of PGA exceeding probability, and the growth rate of the number of casualties caused by the change from a 63% to a 10% 50-year PGA exceeding probability is the largest, reaching dozens of times.

According to the national earthquake emergency plan issued by the General Office of the State Council of China [67], the study area lacks the conditions required for initiating a Level-I or Level-II response to earthquakes when the probability of exceeding the 50-year ground motion level is 63%. Table 12 indicates that, under this specified seismic intensity, the projected death toll across 162 districts and counties in Yunnan, Guizhou, and Sichuan is fewer than 100. Consequently, neither a Level I nor a Level II emergency response will be initiated. Assuming a 10% probability that the GPA will surpass the target threshold over a 50-year period, counties such as Xichang, Lancang Lahu Autonomous County and Yiliang have a death population ranging from 10 to 100 (Table 11), which could potentially trigger a Level-III response to earthquakes (Table 10). With a 2% probability of exceeding the 50-year ground motion level, 100 counties, including Xichang, Lancang Lahu Autonomous County, and Xundian Hui and Yi Autonomous County, may initiate a Level-I response to earthquakes in China, and 97 counties may also trigger a national earthquake emergency response of Level-II or higher (Table 13). With a ground motion level exceeding a 1-year probability of 10^{-4} , the estimated death toll in 164 counties exceeds 500, including over

20,000 in Xichang City, Lancang Lahu Autonomous County, Yiliang, and others, and over 10,000 in Menghai, Jianshui, Xundian Hui and Yi Autonomous County, among others (Table 14).

Table 12. List of counties and corresponding emergency response levels predicted for seismic deaths exceeding 63% probability over a 50-year period in southwestern China.

Predicted Death Toll	Emergency Response Level	Corresponding County Counts
500~	I	0
100~499	II	0
10~99	III	16
1~10	IV	102

Table 13. List of counties and corresponding emergency response levels predicted for seismic deaths exceeding 2% probability over a 50-year period in southwestern China.

Predicted Death Toll	Emergency Response Level	Corresponding County Counts
500~	I	100
100~499	II	97
10~99	III	176
0~10	IV	27

Table 14. List of counties and corresponding emergency response levels predicted for seismic deaths exceeding 10^{-4} probability over a one-year period in southwestern China.

Predicted Death Toll	Emergency Response Level	Corresponding County Counts
500~	I	164
100~499	II	66
10~99	III	143
0~10	IV	27

5. Discussion

5.1. Discussion on Model Assumptions

In this section, we critically evaluate the foundational assumptions underpinning the model, which comprise deterministic hazard representation, linear aggregation of mortality rates, static population distribution, omission of diurnal occupancy fluctuations, and neglect of emergency response impacts. Deterministic modeling assumes a fixed seismic intensity distribution, thereby disregarding the inherent spatial variability of ground motion intensity across core seismic zones and attenuation regions—a factor widely emphasized in the literature as critical for accurate casualty assessment. Linear mortality aggregation fails to account for differential vulnerability across high-risk demographic groups (e.g., the elderly and children) and diverse structural typologies. These include brick, timber, civil, and bamboo/grass/adobe dwellings, which constitute over 50% of the building stock in certain counties within western Sichuan and Yunnan, as well as reinforced concrete structures with lower prevalence (<20%) in these regions. This approach simplifies fatality estimation into a uniform coefficient, despite the established need for era-specific and structure-type-specific vulnerability parameters.

The assumption of a static population distribution does not incorporate migratory populations driven by urbanization, where actual residential patterns often deviate significantly from registered demographic data. It also overlooks diurnal population fluctuations,

leading to discrepancies between modeled population density and real-time spatial distributions, for instance, daytime concentrations in occupational and educational zones versus nighttime clustering in residential areas. The omission of diurnal occupancy rate variation exacerbates this limitation, as casualty estimates do not reflect temporal changes in population exposure. Finally, neglecting emergency response effects ignores the potential for casualty reduction through real-time integration of multi-source information and organized disaster relief operations, which could mitigate fatalities, particularly in regions with pronounced gaps in seismic fortification capacity (e.g., counties in western Guizhou where over 50% of structures consist of low seismic-resistant brick-concrete buildings).

The lethality model applied for casualty estimation appears to rely on deterministic inputs—such as fixed seismic intensity and static vulnerability coefficients—without incorporating stochastic variability in seismic hazard parameters (e.g., ground motion amplitude, duration) or structural response. This omission disregards uncertainties associated with earthquake occurrence (e.g., magnitude, epicenter location) and building performance, resulting in point estimates of casualties rather than probabilistic distributions. For instance, although the study references exceedance probabilities (e.g., 10% in 50 years), the core lethality model may not propagate these probabilistic hazard inputs through to casualty outputs, thereby limiting the representation of uncertainty in the results.

The model may assume linear relationships between input variables—such as peak ground acceleration (PGA) and population density—and casualty rates, which contradicts empirically observed nonlinear behavior. The text explicitly notes a “nonlinear relationship between casualties and PGA,” with casualties increasing several-fold as the 50-year exceedance probability decreases from 63% to 10%. Should the model aggregate lethality using linear function $casualty\ count = \alpha \times PGA + \beta \times population\ density$, it would fail to capture this nonlinearity, leading to underestimation or overestimation of casualties under extreme PGA conditions or in regions of high population density.

Current casualty estimation methodologies predominantly rely on “registered demographic data,” which do not capture real-time or dynamic population movements. Urbanization-driven migratory patterns create systemic discrepancies whereby de facto residential distributions frequently diverge from official registered records, while demographic shifts—such as population growth and aging—are only partially accommodated through simplified “casualty rate calculation coefficients.” The model fails to incorporate spatiotemporal population dynamics, including rural-to-urban migration trends or seasonal labor mobility, leading to inaccurate population density inputs. For example, in counties experiencing substantial inflows of migrant workers, daytime population densities may surpass registered residential densities by a factor of 2–3; nevertheless, the model continues to rely on static registered data.

Diurnal population fluctuations—such as daytime concentration in commercial and educational zones versus nighttime clustering in residential areas—are not accounted for in the modeling framework. Although the importance of “considering diurnal population fluctuations” is emphasized, the model operates under the assumption of temporally invariant spatial occupancy. For instance, schools (characterized by high child occupancy during daytime) and workplaces (with high adult occupancy) would exhibit elevated casualty risks during specific periods; however, the use of static population data obscures these temporal variations, resulting in misestimated high-risk exposure windows.

Furthermore, the model does not integrate emergency response capacity—including rescue speed, medical resource availability, and evacuation efficiency—into casualty projections. While building vulnerability and secondary disaster risks are considered, critical post-earthquake response factors, such as time to reach affected areas and hospital bed capacity, which directly influence casualty outcomes, remain excluded. For example, regions

with shorter emergency response durations (e.g., under three hours) typically experience 30–50% lower fatality rates; however, this mitigating effect is neither quantified nor incorporated into the lethality model, thereby overestimating potential casualties in areas with robust emergency response systems.

5.2. Discussion on the Nonlinear Relationship Between Injuries and Ground Motion

Table 15 presents the statistical analysis of the fluctuations in mortality rates across diverse intervals of the exceedance probability of the Peak Ground Acceleration (PGA) for 400 districts and counties, which constitutes the core components of the sensitivity analysis. Table 16 depicts the sensitivity disparities among various intervals of the PGA exceedance probability. Among the 400 districts and counties, 195 regions demonstrated identical mortality rates at both the 0.01% annual exceedance probability and the 2% 50-year exceedance probability, accounting for 48.75% of the total. These districts presented mortality ranges of 3–5376 cases at the 0.01% within—one-year exceedance probability threshold, with a mean mortality of 156 cases. Regions with unequal mortality rates exhibited mortality ranges of 15–35,749 cases at the same probability level, with an average of 3111 cases.

Table 15. Statistical analysis of mortality rate changes across different GPA exceedance probability intervals in 400 districts and counties.

Statistical Indicator	The Probability of Exceeding the GPA Exhibits a Variation Spanning from 0.01% Within a One-Year Period to 2% Over a 50-Year Period	The Exceedance Probability of Generalized Pareto Distribution (GPA Exhibits a Variation Range Spanning from 2% Over a 50-Year Period to 10% Over a 50-Year Period	The Exceedance Probability of the GPA Exhibits a Variation Spanning from 10% Over a 50-Year Period to 63% Over a 50-Year Period
count	400	400	400
mean	−0.293573	−0.91	−0.97
std	0.361768	0.03	0.10
min	−0.932491	−1.00	−1.00
25%	−0.679755	−0.93	−1.00
50%	−0.004870	−0.90	−0.99
75%	0.000000	−0.88	−0.97
max	0.000000	−0.86	0.25

Table 16. Sensitivity differences across different PGA exceedance probability intervals.

PGA Exceedance Probability Intervals	Average Rate of Change	Median Change Rate	Sensitivity Characteristics
0.01% in 1 year → 2% over 50 years	−29.4%	−0.5%	The lowest sensitivity was observed, with no changes in nearly half of the districts and counties.
2% over 50 years → 10% over 50 years	−90.5%	−90.0%	Highly sensitive, with a sharp decline in mortality
10% over 50 years → 63% over 50 years	−96.8%	−99.4%	Extremely sensitive, with near-zero mortality rates

Stability analysis yielded the following results: when the number of fatalities was less than 50, stability was evident in 138 out of 143 cases (96.5%); for fatalities ranging from 50 to 200, stability was observed in 39 out of 54 cases (72.2%); for fatalities between 200 and 1000, stability was present in 11 out of 81 cases (13.6%); for fatalities ranging from 1000 to 5000, stability was found in 5 out of 88 cases (5.7%); and for fatalities exceeding 5000, stability was detected in 2 out of 34 cases (5.9%). The statistical attenuation ratios across probability intervals (lower values signify more significant attenuation) were as

follows: from 0.01% to 2%: mean = 0.706, median = 0.995; from 2% to 10%: mean = 0.095, median = 0.100; from 10% to 63%: mean = 0.032, median = 0.006.

Further analysis was carried out on the structural threshold effect and the underlying causes of the stability phenomena. Table 17 presents the stability ratio of mortality intervals across districts. The threshold effect indicated that districts with a low number of fatalities (less than 50 cases) accounted for 35.8% and had relatively stable mortality trends, while districts with a medium number of fatalities (50–1000 cases), which constituted 33.8%, exhibited transitional characteristics. Districts with a high number of fatalities (more than 1000 cases), representing 30.5%, showed increased sensitivity to probability fluctuations. By comparing the mortality distribution patterns between stable and unstable districts, the mortality trends were plotted against probability variations (Figure 16). A comparative analysis of the characteristics of stable and unstable districts demonstrated distinct mortality distribution patterns in stable districts, along with an examination of the relationship between the mortality ratio and the absolute number of cases. It was determined whether stable districts were clustered in regions with a low number of fatalities, and high-fatality areas were analyzed. The results showed that among districts with less than 100 fatalities, 162 were stable and 8 were unstable. For districts with 1000 or more fatalities, there were 7 stable districts and 115 unstable districts. The mortality disparity between stable and unstable districts was statistically significant.

Table 17. Stability ratio of predicted death rate intervals by district/county of Yun Gui Chuan.

Death Toll Range by District and County	Stability Ratio	Stability Characteristic
<50	96.5%	extremely high stability
50–200	72.2%	Higher stability
200–1000	13.6%	Low stability
>1000	~5–6%	Almost no stability

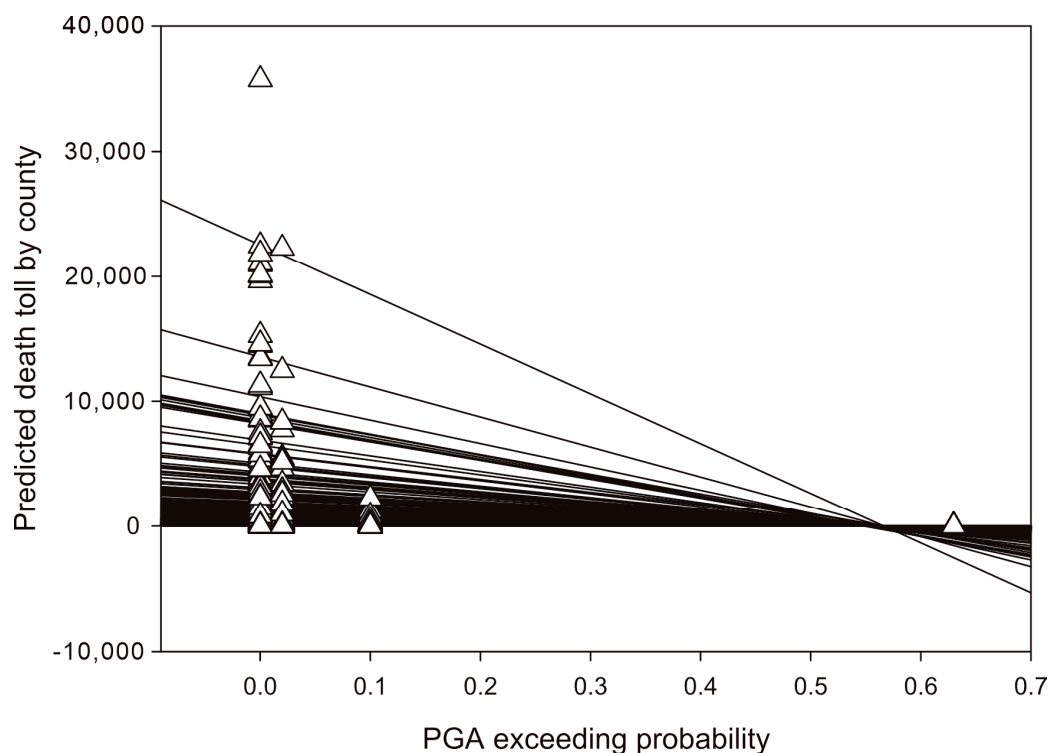


Figure 16. The predicted casualties in the districts and counties of Yun Gui Chuan varies with the probability of exceeding PGA.

Concerning the existence of exposure saturation, this phenomenon suggests that regions with high mortality rates tend to reach a stable state owing to the limited population size. If exposure saturation exists, stable districts should cluster around high-mortality areas; conversely, under model constraints, stable districts should be concentrated in low-mortality regions. An analysis comparing exposure saturation with model constraints reveals that stable districts have a median mortality rate of 156 cases, while unstable districts have an average of 3111 cases. The median mortality rate shows a significant difference between the two groups: 27 cases in stable districts versus 1280 cases in unstable districts. High-mortality districts exhibit stability rates of only 5–6%, whereas low-mortality districts achieve a remarkable stability rate of 96.5%. These findings conclusively exclude exposure saturation as an explanatory factor, confirming that the observed stability patterns are a reflection of model constraints rather than inherent exposure saturation mechanisms.

Therefore, it is concluded that the sensitivity of the model—predicted mortality counts to PGA exceedance probabilities increases with probability levels, with the lowest sensitivity in low-probability intervals and the highest sensitivity in high-probability intervals. The study area presents distinct mortality stratification patterns, where counties with higher mortality show more pronounced responses to probability changes. The observed stability in casualty numbers under low exceedance probabilities is a result of model constraints rather than exposure saturation. At exceedance probabilities of the Generalized Pareto Distribution (GPA) of 0.1% and 2% within a 50-year period, 48.75% of the counties reported equivalent mortality levels, suggesting generally low mortality rates (mean: 156 deaths). Potential computational limitations or precision constraints in low-risk areas may account for why certain counties with low mortality figures appear identical across different probability scenarios, likely due to predefined minimum mortality calculation thresholds.

5.3. Spatial Distribution Analysis of Mortality Risk Associated with Population Density, Fracture Location and Lethality

5.3.1. Analysis of Mortality Risk Associated with Population Density

Regression analyses were conducted using Pearson correlation coefficients to investigate the relationships between the predicted mortality rates across 400 districts/counties under a 10% probability of exceeding the 50-year Global Precipitation Anomaly (GPA) threshold and their corresponding geographic coordinates (longitude and latitude) at the district level. The analysis yielded correlation coefficients of -0.4154 and -0.1797 between mortality rates and longitude and latitude, respectively, indicating a moderate negative correlation with longitude. Mortality rates were generally elevated in western regions compared to eastern regions, with high-risk areas (e.g., Cangyuan Wa Autonomous County) predominantly clustered in the west. A declining trend in mortality rates was observed with increasing longitude (eastward), whereas eastern regions consistently exhibited lower mortality rates with a more uniform spatial distribution. The weak negative correlation with latitude aligned with the correlation coefficient results, suggesting that east-west geographical disparities may exert a significant influence on mortality patterns. Mortality rates demonstrated a slight decreasing trend with increasing latitude (northward); however, this association was less pronounced than that for longitude, displaying no clear gradient patterns and scattered data distribution, which implies a limited effect of latitude on mortality rates. Figure 13b depicts the predicted fatalities for high-risk districts/counties under the scenario where the probability of exceeding 10% of the Ground Peak Acceleration (GPA) within a 50-year period is considered. Using the overall mean mortality rate of 0.0135% as a threshold, districts/counties were classified into high-risk ($n = 106$) and low-risk ($n = 294$) groups, with their spatial distributions illustrated in Figure 17.

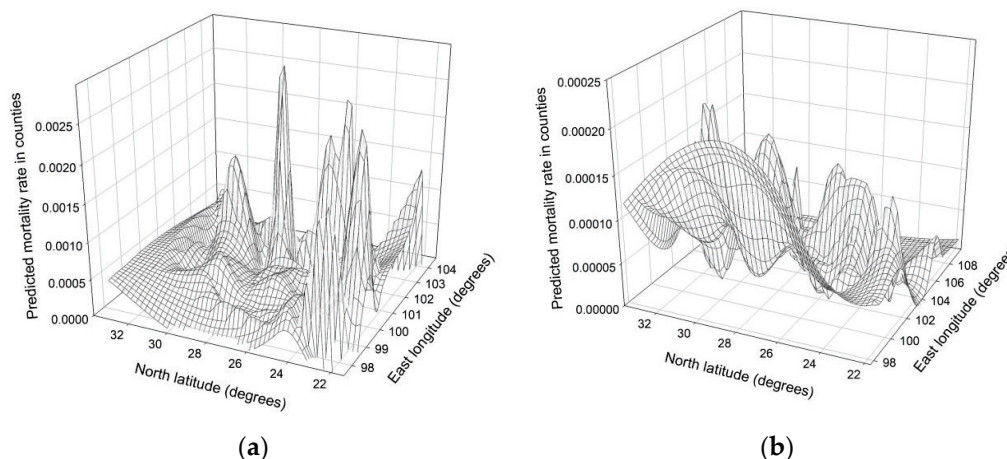


Figure 17. Spatial distribution of predicted mortality rates under a 10% GPA exceedance probability in 50 years: (a) districts with mortality rates exceeding 0.0135%; (b) districts with mortality rates below 0.0135%.

Spatial analysis identified distinct characteristics of counties exhibiting elevated mortality rates:

- (1) Generally smaller population sizes. The average population in the high-mortality group was approximately 325,000, compared to about 492,000 in the low-mortality group. Counties with higher mortality rates tend to have smaller populations, which may reflect their frequent location in remote regions with constrained medical resources or more pronounced demographic aging, contributing to relatively higher mortality.
- (2) A marked westward and southward geographical distribution bias. Representative counties include Xichang City, Songming County, Lancang Lahu Autonomous County, Dongchuan District, Kangding City, and Cangyuan Wa Autonomous County. The mean longitude of the high-mortality group is 101.10° , whereas that of the low-mortality group is 104.95° , indicating a concentration in western regions. Concurrently, mortality rates are generally higher in these western areas compared to eastern ones. The average latitudes are 26.49° for the high-mortality group and 28.22° for the low-mortality group, confirming a southward spatial tendency. These counties are predominantly situated in plateau or mountainous zones, such as central-northern Yunnan and western Sichuan. Characterized by complex topography, limited transportation access, and significant ethnic minority populations, these regions may experience constraints in healthcare infrastructure or specific geographical challenges.

5.3.2. Mechanisms Underlying the Influence of Active Fault Zone Distribution on County-Level Mortality Rates

Figure 18 illustrates the active tectonic framework of the Yunnan-Guizhou-Sichuan region in southwestern China. By integrating the distribution of major active fault zones within the Yunnan-Guizhou-Sichuan region—including the Xiaojiang Fault Zone, Lancang River Fault Zone, Xianshui River Fault Zone, and Anning River Fault Zone—we sought to delineate and synthesize the spatial correlation between zones of elevated mortality and these tectonic structures. The analysis indicates that approximately 80% of high-mortality districts are situated directly within or in close proximity to major active fault zones under a 10% probability threshold for the Gross Proportion of Mortality (GPA).

- (1) High seismic mortality zones, exemplified by Dongchuan District, Songming County, and Xundian County, are situated within the core region of the Xiaojiang Fault Zone. The Xiaojiang Fault Zone represents a large-scale, intensely active left-lateral strike-

slip fault system in southwestern China, constituting the southern segment of the Xianshuihe–Xiaojiang Fault System. This fault zone traverses northeastern Yunnan Province, extending from Qiaojia County at the Yunnan–Sichuan border in the north, southward through Dongchuan, Xundian, Songming, and Yiliang, and terminating in the Jianshui area, with a total length of approximately 400 km. The Xiaojiang Fault Zone truncates alluvial fans, fluvial terraces, and Late Quaternary strata, inducing left-lateral offsets of ridge lines and drainage systems, while generating tectonic landforms such as fault scarps and tectonic troughs. It also exhibits developed co-seismic surface rupture zones and controls the distribution of a series of Late Quaternary basins within its domain. Historically, this fault has experienced multiple major earthquakes with magnitudes ≥ 7.0 , including the 1833 Songming M8.0 event and the 1887 Shiping M7.5 event, and remains seismically highly active [68–70]. Its northern segment (Qiaojia–Dongchuan section) connects with the Xianshuihe Fault Zone, currently displaying a left-lateral strike-slip rate of approximately 14–16 mm/a, representing the segment with the highest slip velocity. The recurrence interval for M7.5–8.0 earthquakes in the northern segment is estimated at 200–300 years. Since 2000 years BP, this segment has recorded 6–7 paleoseismic events, with an average recurrence interval of about 280 years. Considering the historical 1733 Dongchuan M7.5 earthquake, the elapsed time since the most recent paleoseismic event approximates 200 years. The central segment (Dongchuan–Songming section) currently exhibits a left-lateral strike-slip rate of about 12–14 mm/a, accompanied by minor vertical motion at approximately 1–2 mm/a. Major earthquakes of $M \geq 8.0$ in this central region recur at intervals of roughly 300–400 years. The paleoseismic sequence includes strong events around 1833 years BP (corresponding to the historical M8.0 earthquake), 1480, 1120, 760, and 400 years BP, with an average inter-event interval of approximately 355 years. In the southern segment (Yiliang–Jianshui section), the present-day left-lateral slip rate from Yiliang to Chengjiang ranges between 8–10 mm/a, while the slip rate on branch faults near Jianshui decreases to 5–7 mm/a. Recurrence intervals for M7.0–7.5 earthquakes in the southern segment are approximately 150–250 years, whereas those for M8.0 events are around 400–500 years. The intervals between the 1970 Tonghai M7.7, 1887 Shiping M7.0, and 1799 Shiping M7.0 earthquakes were 83 years and 88 years, respectively, approaching the lower bound of the 150-year interval. For larger events such as paleo-earthquakes of $\sim M8.0$ occurring approximately 1500, 1000, and 500 years BP, the recurrence intervals are about 500 years [71–76].

- (2) The Lancang Lahu Autonomous County and Cangyuan Wa Autonomous County, along with other high-mortality regions, are concentrated along the Lancang River fault zone. This fault zone experiences frequent seismic activity, such as the 7.6-magnitude double earthquake event in Lancang–Gengma in 1988. As a core component of the Sanjiang structural belt on the southeastern edge of the Qinghai–Tibet Plateau, the Lancang River fault zone extends NW–SE, traversing China’s Qinghai Province, Tibet Autonomous Region, and Yunnan Province, and extending into Myanmar, with a total length exceeding 1800 km. The Lancang River fault zone runs parallel to the Jinsha River fault zone and the Nujiang fault zone, collectively forming a “three belts enclosing two basins” tectonic pattern. The northern segment begins in the Yushu region of Qinghai Province and extends to Changdu City in Tibet Autonomous Region, with a predominant NWW orientation and SW dip, and a dip angle ranging from 50 to 70 degrees [68–70]. The northern segment is characterized by right-lateral strike-slip motion, with a horizontal slip rate of approximately 5–10 mm/year and a thrust slip rate of about 2–5 mm/year. Three to four paleoseismic events have been recognized in this segment. The third event took place approximately 900 to

700 years ago. The recurrence interval of strong earthquakes ranges roughly from 800 to 1300 years, with an average recurrence interval of about 1000 years. The middle segment extends from Changdu in Xizang to Lincang in Yunnan Province. It dips towards the southwest, with a dip angle ranging from 40 to 60 degrees, and exhibits prominent strike-slip and thrust deformation characteristics. The current right-lateral strike-slip rate can reach 10–15 mm/year, and the thrust slip rate is approximately 3–8 mm/year. This segment displays the highest level of activity, having experienced multiple earthquakes of magnitude 6 or above in history, such as the 7.0-magnitude earthquake in Lijiang, Yunnan in 1996. Three phases of paleoseismic events have been identified in the middle section, dating back approximately 2500, 1400, and 400 years, with time intervals of 1100 and 1000 years respectively, indicating an average recurrence interval of about 1050 years. The southern section extends from Lincang in Yunnan to Shan State in Myanmar. It features southwest-trending faults with dip angles ranging between 30° and 50°. Compared to strike-slip movements, the thrusting components are more pronounced. The right-lateral strike-slip velocity in this segment is approximately 5–8 mm/year, while the thrusting slip velocity ranges from 2–4 mm/year. Four to five paleoseismic events have been detected in the southern section, and the most recent one was the Lancang–Gengma earthquake sequence in 1988. The recurrence intervals of strong earthquakes generally fall within 500–1000 years, with an average of around 750 years. Since the major earthquake in 1988, the southern section has accumulated approximately 35 years of tectonic stress, presenting a risk of magnitude 7.0+ earthquakes in the coming centuries. The northern and middle sections experienced their most recent paleoseismic events approximately 300–800 years ago and are now entering a period of strong earthquake recurrence [69,73,77–79].

- (3) Xichang City, Kangding City, and other high-mortality regions in western Sichuan are situated in the vicinity of the Xianshuihe Fault Zone/Anninghe Fault Zone, which are crucial constituents of the China North–South Seismic Belt. The Xianshuihe Fault Zone extends approximately 430 km, trends at 310/320 degrees, and dips towards the northeast. This fault has disrupted diverse hydrographic systems, mountain ridges, alluvial fans, rock accumulations, gullies, and terraces, giving rise to fault scarps, slope troughs, and developing seismic surface fractures. The Xianshuihe Fault Zone is a large-scale left-lateral strike-slip fault at the northeastern boundary of the Sichuan–Yunnan rhomboid block, generally trending northwest–southeast with an overall length of about 350 km [68–70]. The northern segment (Luhuo–Daofu section) is the area most significantly influenced by the strike-slip movement along the Xianshuihe Fault Zone. It is primarily characterized by pure left-lateral strike-slip with feeble vertical uplift components. The left-lateral strike-slip rate generally falls within the range of 12 to 18 mm/a, with rates in the vicinity of Luhuo being approximately 14 to 16 mm/a and slightly higher in the Daofu region, reaching 16 to 18 mm/a. At least six paleoseismic events have been identified in this segment over the past 4000 years, with intervals spanning from 200 to 350 years and an average recurrence interval of about 250 years. The most recent paleoseismic event took place in the mid–17th century, approximately 320 years subsequent to the 7.6–magnitude Luhuo earthquake in 1973. The middle segment (Daofu–Kangding section) demonstrates both left-lateral strike-slip and thrust uplift characteristics. The left-lateral strike-slip rate is 8 to 12 mm/a, and the vertical uplift rates range from 2 to 4 mm/a, mainly concentrated in the Zeduo Mountain area north of Kangding. Four paleoseismic events have been identified in this segment over the past 3000 years, with intervals ranging from 300 to 400 years and an av-

erage recurrence interval of about 350 years. The most recent paleoseismic event occurred in the late 16th century, approximately 380 years after the 6.9-magnitude Daofu earthquake in 1981. The southern section (Kangding–Shi'an segment), adjacent to the southern end of the Longmenshan fault zone, shows continuously decreasing left–lateral slip rates, with vertical uplift components emerging as the dominant tectonic feature. The left–lateral slip rate in this segment is merely 5–8 mm/a, while the vertical uplift rates have increased to 3–5 mm/a, reaching 4–6 mm/a in certain areas near Shi'an. These rate variations reflect the gradual release of energy from the northeast-trending movement of the Sichuan–Yunnan block through both fault slip and crustal uplift mechanisms during the energy transfer process. Five paleoseismic events spanning the past 5000 years have been identified in this segment, with intervals ranging from 400 to 600 years and an average recurrence interval of approximately 500 years. The most recent paleoseismic event occurred in the early 14th century, about 380 years after the 7.75-magnitude Kangding–Luding earthquake in 1786. The Xianshuihe fault zone is among China's continental fault systems with the fastest-rising activity [69,73,78,80–83].

- (4) The Anninghe Fault Zone, a crucial eastern branch of the Xianshuihe–Xiaojiang Fault Zone, traverses the southwestern part of Sichuan Province in a north–south direction. It commences near Shimian County, where it interfaces with the Xianshuihe Fault Zone, and extends southward through Mianning to the Huili area, covering a distance of approximately 350 km. This fault disrupts the mountain ridge water systems, penetrates through the Late Quaternary strata, alluvial fans, and gully terraces, giving rise to fault cliffs, fault-block lakes, and fault troughs accompanied by active seismic surface rupture. As a left-lateral strike-slip fault with thrust components, it stands as one of the principal earthquake-inducing structures in southwestern Sichuan [68–70]. The northern segment (Shimian–Mianning section) directly intersects with the Xianshuihe Fault Zone. Presently, it exhibits left-lateral slip rates ranging from 6 to 8 mm/a. Vertically, the thrust slip rates remain relatively low, at 1–2 mm/a. The most recent major earthquake in the northern segment occurred approximately 2000 years ago, with strong seismic recurrence intervals ranging from 1500 to 3000 years. The central segment (Mianning–Xichang) showcases the highest tectonic activity within the Anninghe Fault Zone. It has witnessed historical seismic events such as the 1536 Xichang earthquake of magnitude 7.5. Currently, the left–lateral slip rates here reach 8–12 mm/a, rendering it the fastest-moving region within the entire fault zone, with local rates up to 10–12 mm/a. The average vertical thrust rates are 2–3 mm/a. Major earthquakes in the central segment were recorded around 800, 1600, and 2400 years ago, with recurrence intervals ranging from 700 to 1000 years and an average interval of 800 years. The southern section (Xichang–Dechang–Huili) extends southward to connect with the Zemuhe Fault Zone. Currently, it exhibits a left-lateral slip rate of approximately 5–7 mm/a and a vertical thrust rate of about 1–2 mm/a. This segment has experienced two major earthquakes: the 1536 Xichang earthquake of magnitude 7.5 and the 1850 Xichang earthquake of magnitude 7.5. Notably, distinct paleoseismic co-seismic deformation traces have been identified at intervals of approximately 300 years, 800 years, 1300 years, and 1800 years, with an average recurrence interval of around 500 years and a concentration range of 400–600 years [69,73,78,84–86].
- (5) The Nujiang Prefecture, encompassing Lushui City and Fugong County (partially designated as high–mortality zones), is situated within the Nujiang Fault Zone, which is marked by intense seismic activity. As a critical component of the eastern Himalayan tectonic structure along the southeastern edge of the Tibetan Plateau, the Nujiang Fault

Zone extends nearly north–south from Nagqu Prefecture in Tibet Autonomous Region through northwestern and southwestern Yunnan Province to Myanmar, spanning approximately 1500 km as a left-lateral strike–slip fault zone. Geologically, it consists of two primary orientations: a north–south trending section and a northeast-trending section. The fault system incorporates ductile thrust–shear zones, thrust faults, faulted basins, and thrust structures, along with serpentine–greenstone, serpentine-mixed rocks, deep-sea calcite deposits, Paleozoic metamorphic rocks, and Yanshanian granite intrusions. The Nujiang Fault Zone has experienced intricate tectonic evolution, encompassing the formation and expansion of an ocean basin during the Late Triassic–Middle Jurassic, subduction of oceanic crust accompanied by the development of an island arc in the Late Jurassic, early arc–continent collision and convergence from the Early Cretaceous to the Late Cretaceous, and the formation of faulted basins along with thrust–tectonic activity during the Himalayan period. From 1976 to 2014, a total of 14 earthquakes with magnitudes of 6 or above were documented. Significantly, on May 29, 1976, two potent earthquakes with Richter magnitudes of 7.3 and 7.4 occurred in Longling County, Baoshan City, Yunnan Province [68–70]. The northern segment (within Tibet Autonomous Region) exhibits left-lateral strike-slip velocities ranging from 5–10 mm/year, with average rates of approximately 7 ± 1 mm/year in Nagqu and 5 ± 0.8 mm/year in Nyingchi. Vertical movement rates show uplift of 1–3 mm/year in the northern segment, reflecting the fault zone’s dual thrusting and thrusting tectonic characteristics. Over the past three millennia, the northern section has undergone at least four ancient earthquakes with magnitudes of 7.5 or greater. The recurrence intervals are concentrated within the range of 700–900 years, with an average interval of approximately 800 years. The central section (northwestern Yunnan) is one of the most active segments along the Nujiang Fault Zone, demonstrating left–lateral slip rates spanning from 8 to 15 mm/yr. The Gongshan–Fugong section has an average left–lateral slip rate of 12 ± 1.2 mm/yr, while the Lushui section shows a relatively lower rate of 9 ± 1 mm/yr. The vertical uplift rates are within the range of 2–4 mm/yr, and in specific segments (e.g., in the vicinity of Gongshan), the rates can reach up to 3.5 ± 0.5 mm/yr. Historical records encompass the 1906 Lushui earthquake with a magnitude of 7.0 and the 1976 Longling earthquake with a magnitude of 7.3. Archaeological evidence of ancient seismic activity dating back approximately 1000 years and 400 years has been unearthed in the Fugong region, and the recurrence intervals of strong earthquakes are 400–600 years. The southern section (southwestern Yunnan and overseas areas) displays relatively weaker activity, characterized by left–lateral slip rates of 3–8 mm/yr (in the Longling–Mangshi section: 6 ± 1 mm/yr). Upon entering the territory of Myanmar, the slip rates of the fault zone decline to 3 ± 0.5 mm/yr, accompanied by vertical uplift rates of 1–2 mm/yr. Historical seismic data for the southern section is still scarce. The 1898 Myitkyina earthquake in Myanmar with a magnitude of 7.2 is associated with the activity in this region, and the recurrence intervals are estimated to be 600–800 years [68,69,73,78,87,88].

The spatial correlation between high-mortality counties and active fault zones essentially reflects the combined impacts of the lag in socioeconomic development under geological constraints. Fault zones not only present direct geological disaster risks but also indirectly raise regional mortality rates by restricting transportation accessibility, limiting the allocation of medical resources, and impeding economic growth. Our analysis reveals that the influence of fault zones on mortality rates does not solely originate from seismic

events themselves (which cause limited short-term casualties), but rather from long-term geological, socioeconomic, and public health circumstances:

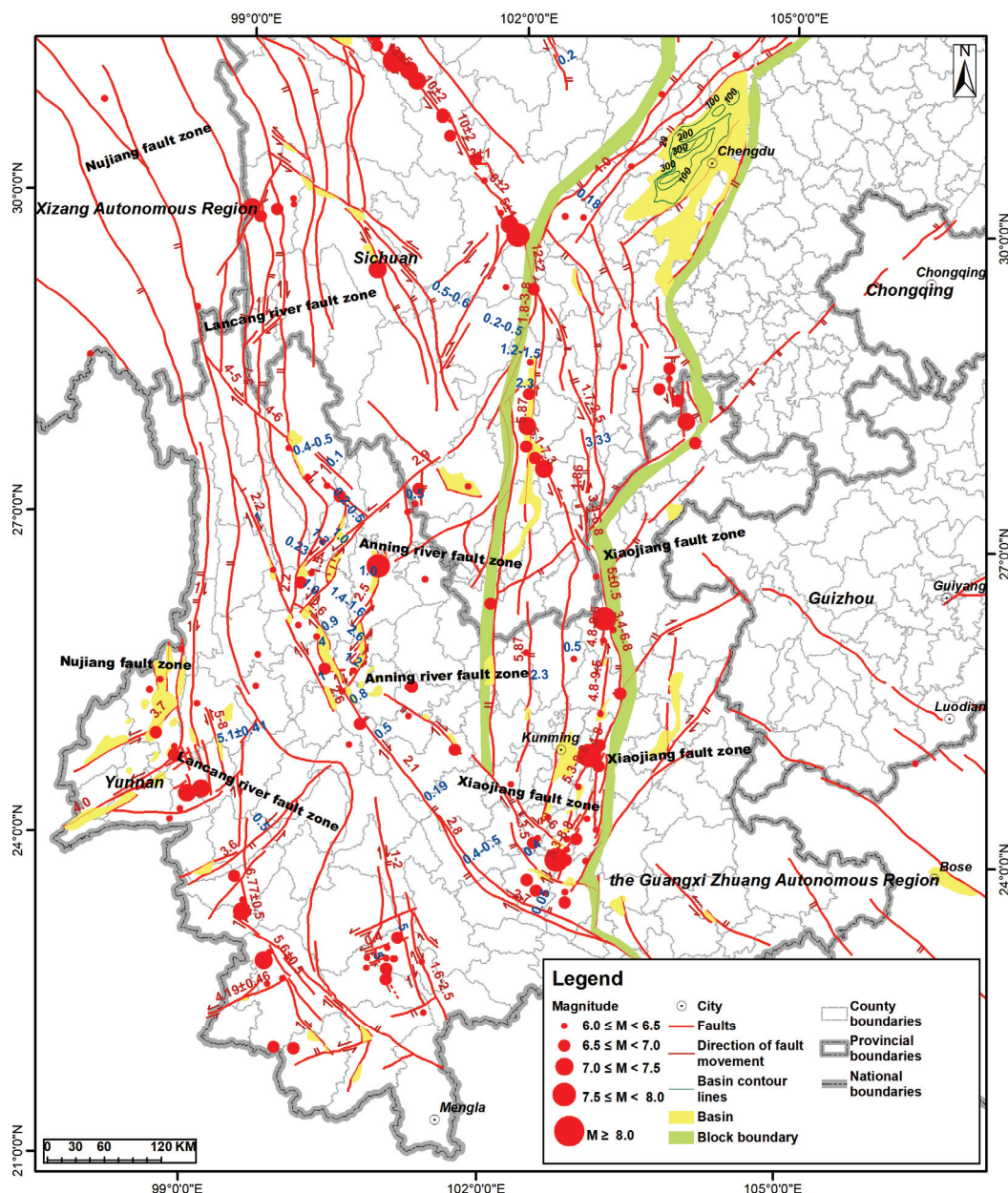


Figure 18. Active tectonic setting of the Yunnan–Guizhou–Sichuan region, China.

- (1) Geological hazards and environmental vulnerability. Areas adjacent to fault zones are susceptible to secondary disasters such as landslides and debris flows, which damage infrastructure (e.g., roads and medical facilities) and weaken emergency response capabilities. For example, landslide-induced traffic disruptions along the Lancang River Fault Zone in mountainous counties delay medical rescue operations. Fault zones are predominantly located in rugged mountainous regions such as the Hengduan Mountains and the Yunnan–Guizhou Plateau, where complex terrain, scarce arable land, and economic limitations jointly deteriorate residents' nutritional status and health conditions.
- (2) Inadequate accessibility to medical resources. Most areas along the fault zone are remote mountainous regions with difficult road construction conditions, leading to a sparse distribution of medical facilities (such as hospitals and clinics). For instance,

Cangyuan Wa Autonomous County is more than 200 km away from the nearest tertiary hospital, resulting in extended emergency response times. The complex geological conditions in fault-zone areas, high infrastructure construction costs, and relatively insufficient government investment in medical resources have collectively led to inadequate coverage of public health services.

- (3) Lagging economic and public health conditions. Most areas along the fault zone are populated by ethnic minorities (e.g., the Lahu and Wa ethnic groups), with a single industrial structure dominated by agriculture and low household income levels, making it arduous to afford high-quality medical services. Some regions suffer from poor sanitation conditions and weak disease prevention and control capabilities, resulting in high mortality rates of chronic diseases (e.g., cardiovascular diseases and respiratory diseases).
- (4) Population dispersion and aging. Counties with high mortality rates generally have populations below 300,000 and display dispersed residential patterns, making it difficult to achieve comprehensive coverage of medical resources. For example, Deqin County has a population of only 68,000 but needs to cover 7443 square kilometers of mountainous terrain. In some fault-zone areas, young adults migrate for employment, resulting in a high proportion of elderly residents left behind and an increased mortality rate due to chronic diseases.

5.3.3. Analysis of Mortality Risk Associated with County-Level Lethality

Previous research has evaluated building vulnerability mortality at the kilometer grid scale. We employed the average mortality rate across the kilometer grids of each county to characterize the overall mortality patterns. Subsequently, we examined the spatial relationship between county-level mortality rates and mortality rates. Calculations demonstrated a Pearson correlation coefficient of 0.1831 between mortality rates and the corresponding mortality rates across 400 counties, indicating a weak correlation. This implies that counties with relatively higher mortality rates tend to show a slight upward tendency in population mortality rates.

In regions with relatively higher mortality rates at the district/county level, the values range from 0.6 to 0.8, corresponding to mortality rates that are either extremely low or extremely high. This suggests that district/county-level mortality may not be the predominant factor determining the level of the ‘mortality rate’. The coefficient of determination R^2 is approximately 0.033, suggesting that a single-dimensional mortality indicator cannot directly determine the overall mortality rate of a district/county.

Seismic mortality rates are influenced not only by direct earthquake—related factors such as lethality, population distribution, and fault locations but also by multiple non—negligible factors, including the population age structure (degree of aging), health-care conditions, economic development levels, and lifestyle habits.

Based on the above analysis, we conclude that a higher population loss in areas such as Xichang and Lancang is positively correlated with population concentration, distance from major active faults, and regional fatality levels. The significant population loss may be a result of the influence of one or two of these factors or possibly their combined effects.

5.4. Sensitivity Analysis of Spatial Data Transformation Methods

The point-to-surface vector method employed in this study can utilize Voronoi polygon interpolation and inverse distance weighting (IDW). IDW operates on the principle that points closer to the prediction location exert greater influence; it estimates values at unknown points through a distance-weighted average of neighboring observations. This method is straightforward and computationally efficient, but it is susceptible to outliers,

may produce artificial “bull’s-eye” effects, and does not provide quantifiable estimates of prediction error. Kriging, grounded in geostatistical theory, employs semivariance functions to model spatial dependence, yielding predicted values along with associated error variances. It is well-suited to datasets with regular spatial distribution; however, it requires substantial sample sizes, involves computationally intensive procedures, and its performance is contingent upon appropriate model parameter selection. The Natural Neighbor method constructs a dual Voronoi (Thiessen) polygon structure to perform area-weighted interpolation among adjacent points, thereby preserving local spatial characteristics and generating smooth boundaries. It is particularly effective for irregularly distributed data, though its capacity to represent densely sampled regions may be limited. Spline interpolation utilizes polynomial functions to generate smoothly continuous surfaces, making it suitable for modeling gradual terrain gradients; nevertheless, it may over-interpolate in sparsely sampled areas and is sensitive to anomalous data points. Voronoi (Thiessen) Polygon interpolation delineates the zone of influence for each sample point via perpendicular bisectors, ensuring uniquely defined boundaries for surface-vector conversion. This approach is effective for neighborhood analyses in contexts such as population distribution and resource allocation, resolving the ambiguous boundary issues often associated with IDW or spline methods. It supports local topological adjustments, which is advantageous for real-time updates with dynamic data, such as in earthquake emergency assessment scenarios. Furthermore, the method permits weighted extensions (e.g., weighted Tyson polygons) to incorporate multiple factors relevant to earthquake casualty prediction, a capability in which natural neighbor methods often struggle when integrating multi-source data. The clearly demarcated geographic units produced by Voronoi polygons facilitate administrative-level analysis and enable the mapping of casualty distributions to explicit spatial boundaries, thereby providing concrete spatial guidance for emergency resource allocation. In summary, Voronoi Polygon interpolation demonstrates distinct advantages in terms of spatial boundary clarity, efficiency in dynamic updates, and flexibility in multi-factor integration, rendering it particularly suitable for earthquake casualty prediction. Alternative methods such as IDW and Kriging perform adequately in specific scenarios but exhibit limitations in neighborhood representation and multi-source data fusion; consequently, they cannot fully substitute for the Voronoi Polygon approach in this context. In this study, village-level population data are transformed into polygon vectors via Voronoi tessellation, which precisely captures spatial heterogeneity in building structural types (e.g., reinforced concrete structures comprise over 80% in northeastern Sichuan, whereas brick-wood structures exceed 60% in western Sichuan). For example, in Lancang County, Yunnan Province, the Voronoi polygon boundaries align closely with actual village distributions, thereby mitigating the over-smoothing artifacts often encountered with Kriging interpolation in data-sparse regions. When post-earthquake casualty data are incorporated, the Voronoi polygon framework necessitates only localized adjustments to adjacent polygons, enabling rapid updates of county-level casualty distributions (e.g., the casualty assessment for Xichang City under a 10% probability of exceedance in 50 years was dynamically updated from 2247 to 2275 fatalities). Furthermore, Voronoi polygons facilitate the weighted superposition of building lethality coefficients (e.g., brick-wood structures with lethality thresholds ranging from 0.6 to 0.9) and population density to directly derive grid-level casualty rates (e.g., areas dominated by brick-wood structures in Daguan County, Yunnan, exhibited a mortality rate of 0.23%).

The sensitivity of Voronoi polygons in spatial data computation for earthquake casualty prediction scenarios is primarily manifested in three dimensions: (1) in densely populated regions (e.g., the Chengdu Plain, Sichuan), clustered settlement points generate Voronoi cells of reduced size, thereby facilitating granular differentiation of architectural

typologies (e.g., zones with over 80% reinforced concrete structures versus brick–wood constructions). Conversely, in sparsely distributed areas such as the western Sichuan plateau, enlarged cell dimensions may obscure localized high-risk points (e.g., isolated brick–wood villages), resulting in smoothed risk assessments. Statistical biases in village-level building structure or demographic data directly propagate into weighting calculations. For instance, a 10% underestimation of the proportion of brick–wood structures in a village in Lancang County, Yunnan, may lead to an approximate 0.05% underestimation of the predicted mortality rate. (2) Weight distribution exerts a substantial influence on outcomes within mixed architectural units. A 10% variation in the ratio of brick–concrete structures (fatality thresholds: 0.25–0.7) to brick–wood structures (0.6–0.9) in northeastern Sichuan could induce mortality rate fluctuations of 0.1–0.3 percentage points. Failure to incorporate diurnal population mobility may yield exposure population estimation errors of 15–20% in urban–rural fringe zones. (3) Under low-probability scenarios (e.g., 63% exceedance probability at a 50-year return period), low-intensity seismic events (magnitude VI–VIII) generally produce mortality rates below 0.01%, reflecting high stability in outcomes (96.5% of cells exhibit errors below 5%). In high-probability scenarios (e.g., 2% exceedance probability at a 50-year return period), high-intensity earthquakes (magnitude IX and above) trigger pronounced mortality increases within high-risk units (e.g., Lancang County, Yunnan), with abrupt casualty rate variations occurring particularly in boundary regions. The sensitivity of Voronoi polygons arises from their rigorous geometric partitioning logic (perpendicular bisector boundaries) and localized weighting mechanisms; nevertheless, these characteristics also render them indispensable in terms of spatial boundary delineation clarity, multi-source data integration capacity, and real-time emergency response performance. Table 18 presents the sensitivity characteristics of typical spatial data transformation methods.

Table 18. Comparative analysis of the sensitivity of typical spatial data transformation methods.

Comparison Dimension	Voronoi Polygon	Krychkin Method	Inverse Distance Weighting Method (IDW)	Natural Neighborhood Method
parameter dependency	High (uniformity of dependency point distribution) <block index = “1396”>	Highly elevated (requires fitting a spatial autocorrelation model)	Medium (dependence distance attenuation parameter)	Medium (depending on the topological relationship of adjacent points)
Parameter sensitivity	Low (requires only point coordinates and weights)	High (semi-variance function parameters have a significant impact)	Medium (Power index selection affects results)	Medium (neighborhood area weight calculation sensitivity)
Boundary effect	High (clear boundary, significantly influenced by adjacent points)	Low (smooth boundary for continuous surfaces)	Medium (with “bull eye effect”)	Low (Smooth boundary transition)
Dynamic update efficiency	High (adjust adjacent cells locally) <block index = “1396”>	Low (global model recalculation)	Medium (requires recalculating weights)	Medium (reconstruction using dependent triangular network)

5.5. Impact of Structural Performance and Type Weight Changes Interaction on Casualty Estimation

The linear independence assumption employed in structural classification may lead to an underestimation of actual lethality by neglecting synergistic failure or suppression effects between structures. The seismic performance of building structures (e.g., masonry–wood, masonry–concrete, reinforced concrete) is not entirely independent of one another; rather, it exhibits spatial correlations and mechanical coupling during failure processes. Linear weighting systems may inadequately capture the following interdependencies:

- (1) Spatial synergistic failure effects. Buildings of different structural types within the same region may interact through collapse impacts and vibration amplification. For example, the collapse of highly vulnerable low-rise masonry–wood structures may crush adjacent masonry–concrete buildings. Under soft soil conditions, a concentrated distribution of civil engineering/bamboo straw structures can amplify ground motion, resulting in more severe damage to surrounding structures than would be predicted by independent assessments. Linear weighting merely aggregates the independent contributions of each structural type while ignoring such synergistic interactions, potentially leading to an underestimation of overall lethality.
- (2) Seismic performance correlations. Different structural types within the same county may share common influencing factors (e.g., construction quality, seismic design standards), leading to non-independent distribution patterns. In economically underdeveloped counties where masonry-wood and civil engineering structures dominate and generally lack seismic fortification measures, structural vulnerabilities often exhibit positive correlations. Linear weighting may fail to reflect the damage amplification caused by such “vulnerability clustering,” thereby underestimating lethality indices.
- (3) Oversight of suppression effects. High-performance seismic structures (e.g., reinforced concrete) may exert protective effects on adjacent vulnerable structures (e.g., by blocking debris or reducing localized vibrations). Assuming independent contributions in a linear weighting framework could therefore overestimate overall lethality.

Assuming the county-level fatality index formula is as Equation (10): w_i represents the weight of the building category i , p_i denotes the proportion of buildings in that category, the casualty count C and k denote coefficients related to seismic intensity and population density (Equation (11)).

$$L = \sum(w_i \cdot p_i) \quad (10)$$

$$C = k \cdot L \quad (11)$$

Considering a typical county as a case study, with construction type proportions specified as reinforced concrete (20%), brick–concrete (30%), brick–wood (30%), and earth–wood/bamboo straw (20%), the baseline weights are assigned as $w_{\text{reinforced concrete}} = 0.1$, $w_{\text{brick–concrete}} = 0.3$, $w_{\text{brick–wood}} = 0.5$, and $w_{\text{civil}} = 0.8$. The baseline $L = 0.42$, with $C = 420$ and $k = 1000$.

- (1) For high-vulnerability structures, specifically brick–wood and earth–wood constructions, the following weight adjustments are posited: the weight coefficient for brick–wood structures is hypothesized to increase from 0.5 to 0.6, representing a 20% rise, while that for earth–wood structures is projected to increase from 0.8 to 0.9, corresponding to a 12.5% elevation. All other weight parameters remain unchanged. The calculation yields $L = 0.1 \times 0.2 + 0.3 \times 0.3 + 0.6 \times 0.3 + 0.9 \times 0.2 = 0.47$ with $C = 470$, a 11.9% increase compared to the baseline. This finding indicates that the weights assigned to high-vulnerability structures exert the most significant influence on casualty estimation, as buildings within this category constitute a substantial proportion of the total structural inventory and demonstrate higher fragility. Consequently, adjustments to these weights typically result in a substantial elevation of casualty estimates.
- (2) Regarding variations in the weights of low-vulnerability structures (reinforced concrete), assuming that the reinforced concrete weight is reduced from 0.1 to 0.05 (a 50% decrease), while all other weights remain unchanged. The calculation yields $L = 0.05 \times 0.2 + 0.3 \times 0.3 + 0.5 \times 0.3 + 0.8 \times 0.2 = 0.41$, with $C = 410$, representing a 2.4% reduction compared to the baseline. This indicates that structural weights associated with low vulnerability exert a negligible influence on the overall results,

owing to their limited proportional representation and modest baseline magnitude, which consequently restricts their contribution to the composite index.

- (3) Concerning variations in structural weights for moderately vulnerable constructions (e.g., brick-concrete systems). Assuming the weight assigned to brick-concrete structures increases from 0.3 to 0.4 (a rise of 33%) while all other weights remain constant, the resulting computation yields: $L = 0.1 \times 0.2 + 0.4 \times 0.3 + 0.5 \times 0.3 + 0.8 \times 0.2 = 0.45$, with $C = 450$, representing a 7.1% increase compared to the baseline. These findings indicate that the structural weight impact for moderately vulnerable structures is intermediate between those of high and low vulnerability levels. Owing to its comparatively substantial proportion (30%), the adjustment of weight exerts a discernible amplifying effect on the resultant outcomes.
- (4) Concerning multi-weight collaborative variation: Assuming an increase in the weights assigned to high-vulnerability structures (brick-wood: $0.5 \rightarrow 0.6$, earth-wood: $0.8 \rightarrow 0.9$) concurrently with a decrease in the weights for low-vulnerability structures (steel-concrete: $0.1 \rightarrow 0.05$). The calculation yields $L = 0.05 \times 0.2 + 0.3 \times 0.3 + 0.6 \times 0.3 + 0.9 \times 0.2 = 0.46$, with $C = 460$, representing a 9.5% increase compared to the baseline. This finding indicates that during multi-weight collaborative adjustments, the dominant influence of high-vulnerability structures remains substantial, with the overall effect magnitude comparable to that observed under single high-vulnerability weight variations.

In summary, weight variations in high-vulnerability structures (e.g., brick-wood and earth-wood) exert the most pronounced influence on casualty estimation (exceeding $\pm 10\%$), followed by medium-vulnerability structures (e.g., brick-concrete), while low-vulnerability structures (e.g., steel-concrete) demonstrate the least impact. Should actual building structures exhibit positive spatial correlations—such as the clustering of vulnerable constructions—the conventional weighting framework may further underestimate casualty figures.

5.6. Simulation-Based Framework for Disaster Uncertainty Propagation

To mitigate the deficiency in uncertainty propagation within the integration of seismic intensity, population exposure, and lethality coefficients, a probabilistic framework employing Monte Carlo simulation was developed. Initially, a parameter uncertainty model was constructed, wherein probability distributions for key uncertain parameters were defined based on data characteristics and expert knowledge. Assume a lognormal distribution $PGA \sim LN(\mu_{PGA}, \sigma_{PGA}^2)$, where μ_{PGA} and σ_{PGA} are derived from regional seismic hazard analysis (e.g., using ground motion prediction equations and historical earthquake records). We define Model spatial-temporal variability population exposure P with a truncated normal distribution $P \sim N(\mu_P, \sigma_P^2)$, constrained by census data and diurnal population flow (e.g., higher variance in urban areas due to commuting). For each structure type (brick-wood, civil, reinforced concrete, etc.), define collapse probability (building structure vulnerability) as a function of PGA, e.g., logistic distribution $P_{collapse}$ is shown in Equation (12).

$$P_{collapse} = \frac{1}{1 + \exp(-(\alpha + \beta \cdot PGA))} \quad (12)$$

where α and β are calibrated via empirical vulnerability curves. Assume a Beta distribution Lethality Coefficient $L \sim \text{Beta}(\alpha_L, \beta_L)$ for collapse-induced mortality rates, parameterized by historical casualty data (e.g., α_L, β_L) estimated from post-earthquake field surveys). We then Incorporate dependencies between parameters PGA-collapse probability and population-exposure using copula functions: We build Positive correlation (higher PGA increases collapse risk) modeled via Gaussian copula with correlation coefficient ρ . Urban areas with high population density may have higher building density, so we use rank correlation to link population exposure and structural type proportions.

Monte Carlo simulation workflow includes sampling, casualty calculation (Equation (13)) and statistical analysis. By Sampling, we generate N (e.g., 10^4) random samples for each parameter, and draw $PGA_i \sim LN(\mu_{PGA}, \sigma_{PGA}^2)$, $P_i \sim N(\mu_P, \sigma_P^2)$, $L_i \sim \text{Beta}(\alpha_L, \beta_L)$, structure type proportions via multinomial distribution; compute $P_{collapse,i}$ for each structure type, using sampled PGA_i . For each sample i ,

$$Casualties_i = P_i \times \sum_{k=1}^m (Proportion_k \times P_{collapse,k,i} \times L_{k,i}) \quad (13)$$

where m means number of structure types, $Proportion_k$ means the proportion of structure type k .

Subsequently, we aggregate N casualty outcomes to derive the probability density function and cumulative distribution function of casualties, along with summary statistics including the mean, standard deviation, and key quantiles, to support risk-informed decision-making. This framework yields a probabilistic distribution of casualties, thereby superseding deterministic point estimates with probabilistic confidence intervals and facilitating a rigorous quantification of uncertainty inherent in disaster impact assessments.

5.7. Comparison of This Method with International Casualty Assessment Frameworks

Our method, akin to mainstream international frameworks (e.g., HAZUS-MH, PAGER, and ELER), adopts population exposure (density and distribution), building vulnerability (type and age), and seismic intensity (e.g., PGA) as core components, while also incorporating considerations of secondary disasters and indirect casualty risks, thereby reflecting the general logic underlying casualty assessment. Concurrently, the principal distinctions between the present approach and prevailing international methodologies reside in the following aspects:

- (1) In consideration of population dynamics and data timeliness, early iterations of HAZUS-MH relied primarily on static census data. Although recent updates have integrated dynamic adjustments, these often lack adaptability to migrant populations in developing regions undergoing rapid urbanization—such as mobile labor groups in southwestern China—frequently overlooking discrepancies between actual residential distribution and registered household data. In contrast, the proposed method explicitly acknowledges that “urbanization-driven migration leads to deviations between actual residence distribution and registration data,” and advocates for incorporating “real-time multi-source information” to address regional population dynamics, thereby mitigating the timeliness limitations inherent in existing international frameworks.
- (2) Regarding the region-specific building vulnerability customization, methods based on HAZUS typically employ vulnerability curves derived from building types common in Europe and North America (e.g., timber framing, steel structures), which exhibit limited applicability to traditional building typologies prevalent in many developing regions (e.g., bamboo, straw, or adobe dwellings; brick-wood hybrid structures). This discrepancy can introduce substantial bias in loss estimation. The present method focuses on the current building stock in southwestern China, where “brick, timber, adobe, and bamboo/straw structures collectively constitute over 50% of the inventory, while reinforced concrete buildings account for less than 20%.” By classifying buildings according to “construction era” and “structural type diversity,” and integrating spatial distributions of older buildings with demographic data, a regionally tailored vulnerability model is developed. This model demonstrates closer alignment

with actual building conditions in southwestern China compared to internationally generalized frameworks.

- (3) In consideration of the methodological orientation of assessment models, international frameworks predominantly employ a chain-model approach structured as “building damage → casualties” (e.g., HAZUS transforms ground motion → building damage states → casualty ratios), emphasizing refinement of intermediate processes (e.g., structural damage mechanisms) and suitability for decomposing casualties within complex building systems. The method presented here employs a “lethality model” that simplifies intermediate damage estimation by directly correlating mortality rates with population density, seismic intensity, and high-risk demographic groups (e.g., the elderly and children). While this enhances computational efficiency, it may overlook differential impacts of varying damage levels on casualties and lacks the mechanistic explanatory depth characteristic of international frameworks.
- (4) Regarding the integration of probabilistic risk with emergency response: Whereas PAGER emphasizes rapid post-earthquake impact assessment and HAZUS focuses on long-term risk zonation, few frameworks directly link probabilistic risk metrics (e.g., casualties under different exceedance probabilities) to emergency response activation thresholds. The proposed method innovatively analyzes “fatalities corresponding to 10% and 2% probability of exceedance in 50 years, as well as 0.1% probability of exceedance in one year,” while quantitatively evaluating “the likelihood of triggering a Level I national emergency response” (e.g., in counties such as Xichang where the 50-year 10% exceedance probability threshold is surpassed). By directly coupling risk assessment with emergency preparedness planning, this approach enhances practical applicability and extends the implementation scope of international frameworks within risk management practice.
- (5) Regarding the weighting of population structure vulnerability: While existing methodologies incorporate demographic factors, they predominantly treat these as ancillary attributes of exposed populations, rather than explicitly defining “high-risk groups such as the elderly and children” as independent adjustment coefficients. The present approach explicitly establishes “population structure—specifically, (the proportion of elderly and children as high-risk groups—as an impact coefficient in casualty rate estimation,” thereby directly integrating sociodemographic characteristics into the model parameters. This refinement enhances the vulnerability dimensions of exposed populations and more prominently highlights population-specific risks compared to conventional international frameworks.

Our method constitutes a regionally customized probabilistic casualty assessment model, representing an adaptive optimization and practical extension of international assessment frameworks within a specific regional context, namely, the Southwest Seismic Belt of China. Tailored to the characteristics of China’s southwestern mountainous regions, where traditional buildings are densely distributed and population mobility is high, the model optimizes the integration of population dynamics, building vulnerability, and emergency response coordination. It thereby addresses limitations inherent in internationally standardized frameworks when applied to localized contexts in developing countries. Nevertheless, the method remains comparatively limited in terms of refined building damage mechanisms—such as casualty conversion across varying damage states—and model generalizability, which hinders direct application to other regions, in contrast to modular designs employed in mature international frameworks such as HAZUS.

5.8. Problems with Our Method and Some Corresponding Improvement Objectives

To facilitate rapid assessment, this study concentrates on Southwest China as the research area. It calculates the residential mortality rate on a kilometer grid using county-level load-bearing housing data from the seventh national census. The study integrates population data to estimate casualties, taking into account earthquake site effect assessments and residential building conditions, thereby providing a reference for rapid disaster assessment and auxiliary decision-making post-earthquake. Given that the anti-seismic capacity varies across different regions, even with similar intensity and population density, the casualty rate will differ significantly. Theoretically, a lethal level matrix can be calculated and graded for the entire country, with each region having its own grade range. The model accurately evaluates the lethality matrices of different levels nationwide. This method is not only applicable to the study area but can also be used to estimate earthquake-related fatalities in other seismic risk areas across China. Furthermore, it can be extended to estimate economic losses and other post-earthquake indicators.

Below are problems with our method and some corresponding improvement objectives:

- (1) We have compiled a significant amount of pre-earthquake assessment data nationwide, covering indicators of casualties and economic losses for specific regions and magnitudes. However, discrepancies between the actual magnitudes of future earthquakes and those estimated beforehand can lead to challenges during post-earthquake assessments. Data for seismic intensities VI, VII, and VIII were primarily used to construct the casualty model. In contrast, mortality data for intensities above VIII were limited, and data for intensities above X were nearly nonexistent. Furthermore, due to the lack of statistical data on casualties from secondary geological disasters in historical earthquake cases in China, it is impractical to include secondary geological disasters in the model construction. Additionally, in some national disaster assessment reports, the statistical methods for calculating building damage rates vary, resulting in the omission of relevant data. Consequently, due to the limitations of the original data, steel and stone (wood) structures were not included in the theoretical calculations.
- (2) Although China's local statistics are traditionally reported from the grassroots level upwards, the village and community constitute the most fundamental tier of population statistics. Each village and community is required to maintain some basic data. However, it is known that the more granular the information, the less complete it tends to be. Provincial statistics bureaus typically publish official data only at the county levels, which are also the most valued and comprehensive. Consequently, the statistics on disaster-prone houses in most Chinese provinces, accessible through formal public channels, are limited to the county levels. Thus, the lethality at the village level in this paper is spatially distributed based on the lethality at the county levels, resulting in the actual spatial resolution of lethality being determined by the basic data at those levels. In the future, if China's official statistics were to standardize and preserve statistical data on house types at the village level for record-keeping, or if such data could be obtained through field surveys, it would be possible to enhance the actual resolution of lethality to a smaller scale and achieve greater accuracy in lethality assessments.
- (3) Regarding the time coefficient factor, our model primarily relies on the direct definition provided by expert experience and related outcomes. We calculate the proportion of house types based on the distribution of houses as indicated in the statistical yearbook. If an actual earthquake occurs at night, this method will yield more accurate results, as most people are likely to be inside houses. However, if an earthquake strikes during

the daytime, the population distribution in office buildings and other structures introduces some inaccuracies, suggesting a need for further research.

- (4) The mortality rate is influenced by secondary geological disasters, urban population density (which hinders timely evacuation), traffic congestion, the education level of the population, and their alertness, among other factors. Given the model's applicability, the factors affecting mortality vary across different regions. Consequently, the lethality/injury matrix utilized in this paper requires localized adjustments, and the localization of the matrix can be explored based on the national lethality/injury matrix.
- (5) Considering the extensibility of the model, there is still room for improvement. The spatialization of population and socio-economic data can effectively enhance data accuracy. For the rapid assessment model, it is essential to thoroughly study the regional death/injury matrix (including the southwest, northwest, east, etc.), the secondary disaster chain matrix, and the refined assessment model (which provides the spatial distribution of casualties). In terms of rapid assessment of secondary geological disasters, beyond the casualties from residential collapses, attention should also be given to the assessment and management of other secondary disasters such as landslides and rolling stones. It is important to consider and analyze the soft impact of public opinion and indirect losses. For the earthquake impact field, it is necessary to simulate the distribution of ground motion parameters based on rapid report information and real-time social perception data, such as PGA, PGV (peak ground velocity), Arias intensity, and China earthquake intensity [56,60]. The thematic dataset of disaster bearing institutions can be improved. For earthquake pre-assessment, risk assessment, and risk zoning, fine, quantitative, and periodic lethal level (mortality score) layers can be realized in the risk area or test area, based on remote sensing and machine learning techniques. Combined with the spatio-temporal dynamic distribution of the population, multi-time dynamic seismic disaster risk assessment can be conducted in key areas. Building on seismic risk research, and incorporating refined full-factor disaster bearing body data, full-factor full-chain seismic disaster simulation, including casualty assessment, can be achieved.
- (6) Population growth, urban expansion, and the evolution of architectural structures constitute fundamental factors influencing the estimation of earthquake casualties. To enhance estimation accuracy, it is imperative to establish a dynamic framework that integrates real-time demographic data, develops multidimensional coupling models, and applies region-specific adjustment coefficients. As regional populations increase, population density within epicentral and seismically affected zones rises accordingly, thereby expanding the potential baseline for earthquake-related casualties. Accurate population density data are critical for casualty estimation, necessitating differentiation between core seismic zones and areas of diminishing intensity, as well as consideration of diurnal population fluctuations. Demographic shifts accompanying population growth influence the calculation coefficients for casualty rates, given that elderly individuals and children represent high-risk groups. Urbanization-driven migratory populations further complicate estimation, as actual residential distribution often diverges from registered demographic data, underscoring the need for models to incorporate real-time multi-source information. During urban expansion, developed areas increasingly extend into zones of high seismic hazard. Precise delineation of risk zones within urban built-up areas is essential, integrating geological and hazard assessment data. High-density building clusters elevate the risk of secondary disasters, and models should accordingly integrate projections of secondary disaster casualties. The development of complex infrastructure systems increases the

potential for indirect casualties; thus, assessments should evaluate such risks based on seismic resistance capacity and population coverage. With the advancement of seismic design standards, newer construction types exhibit reduced collapse risks, thereby diminishing core earthquake casualties. Casualty estimation methodologies should classify buildings according to construction era and apply differentiated vulnerability coefficients. The diversity of building types necessitates tailored modeling approaches, while aging structures in urban environments continue to pose significant risks; their spatial distribution should therefore be analyzed in conjunction with demographic datasets.

- (7) The evaluation of statistical performance in earthquake casualty estimation models necessitates the application of quantitative metrics to assess prediction accuracy, reliability, and uncertainty. Core error measurement indicators include Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Bias. Classification performance metrics for the identification of high-risk areas comprise Accuracy, Precision, Recall, and the F1-score. Correlation and uncertainty are evaluated using the Pearson correlation coefficient, Spearman rank correlation coefficient, and Confidence Intervals. The current model lacks quantitative validation against key statistical performance metrics and primarily depends upon descriptive spatial distribution analyses. Input parameters such as “population density data” and “building vulnerability coefficients” contain inherent uncertainties, which may elevate MAE and RMSE values. For instance, a $\pm 5\%$ error in population density data can propagate to casualty estimation errors ranging from $\pm 5\%$ to $\pm 10\%$. Bias cannot be directly evaluated under the existing framework, and assumptions regarding new construction types may introduce negative bias in newly developed urban areas and positive bias in older districts. To enhance model reliability, historical earthquake casualty data should be incorporated to calibrate input errors through the calculation of MAE, RMSE, and Bias. Metrics such as Precision, Recall, and the F1-score should be introduced to validate the accuracy of high-risk county identification. Confidence intervals and the Area Under the Curve (AUC) should be included to quantify prediction uncertainty and discriminative capacity. Spearman rank correlation tests can be employed to examine associations between “building structure proportions” and “predicted casualty density.” Furthermore, Spearman correlation analyses should be conducted to evaluate relationships among variables such as building structures, population density, and casualty outcomes, thereby strengthening model interpretability.

5.9. Implications for Earthquake Emergency Management Strategies

In order to translate research findings into managerial practices encompassing “risk identification, vulnerability reduction, precise response, and long-term governance,” thereby enhancing earthquake emergency management, we formulated the following strategies:

- (1) A dynamic risk priority assessment system should be established to identify core zones. A three-dimensional model is developed, designating areas with high casualty density as Primary Emergency Core Zones; counties experiencing over 300 fatalities under a 10% probability in 50-year periods as Secondary Emergency Priority Zones; and regions with concentrated casualties under lower probabilities as Highly Vulnerable Building Zones for retrofitting. Counties with over 50% vulnerable buildings are prioritized for seismic upgrades. Annual drills are implemented in high-risk counties, and quarterly consultations are conducted for 100 counties.
- (2) Engineering defenses and policy strategies should be strengthened. In Highly Vulnerable Building Zones, the Seismic Reinforcement Initiative for old buildings is imple-

mented, with subsidies covering 30–50% of costs, aiming to achieve 70% retrofitting by 2025. The highest seismic standards are mandated for new constructions in areas with low reinforced concrete usage. In urban areas, high-density clusters are limited, and the construction of new schools and hospitals in high-risk zones is prohibited. In rural areas, seismic-resistant housing standards are promoted, and hazardous structures are to be eliminated by 2030.

- (3) Emergency resource allocation should be optimized to establish a dynamic and precise response system. Fifteen-minute supply depots are established in densely populated and fluctuating areas, equipped with supplies for the elderly and children. Mobile signaling and traffic data are utilized to generate real-time population heat maps, enabling the estimation of victims within one hour post-earthquake. Secondary disaster risks are addressed through integrated monitoring platforms in central-southern Sichuan and Guizhou, which generate high-risk lists within 30 min. Fatality and casualty zone probabilities are adopted as response criteria. An automated algorithm is developed to reduce decision-making time to 30 min.
- (4) Data integration and technology application should be promoted to establish a holistic risk governance system. Data from multiple authorities are integrated to create a Seismic Risk Map incorporating factors such as population density and building vulnerability. Based on PGA-casualty research, a 1 h post-earthquake casualty assessment model was developed to guide resource allocation. Seismic Emergency Smart Terminals are deployed in Primary Emergency Core Zones for early warning and grassroots reporting.
- (5) Seismic disaster prevention should be enhanced by incorporating the “Three-Dimensional Priority Evaluation Model” into the revised Seismic Disaster Prevention and Mitigation Law, thereby clarifying policies for high-risk zones. Seismic Emergency Resource Allocation Standards and a Special Fund for reinforcement projects are established. Building Vulnerability Reduction Rates and First-Level Emergency Response Timeliness are integrated into local government safety evaluations with a weight of at least 10%. A tri-provincial mechanism among Sichuan, Yunnan, and Guizhou is developed to share risk data, conduct joint drills, coordinate rescue efforts, and organize annual cross-provincial emergency drills.

6. Conclusions

- (1) The distribution of earthquake casualties in Yunnan, Guizhou, and Sichuan was determined using a lethality model. The casualty density is highest in the south and east of central Sichuan, with a more scattered pattern in the west. In Yunnan, the casualty density is low in the northwest and southeast, whereas in Guizhou, it is high in the eastern and western mountainous areas.
- (2) The casualty assessment results indicate that Xichang is the most severely affected, with Lancang also significantly impacted. When the exceeding probability exceeds 10% over 50 years, counties such as Xundian, Songming, and Dongchuan are projected to have a death toll of more than 300. As the exceeding probability increases to 2% over 50 years and 0.1% in 1 year, the fatalities are exclusively concentrated in Yunnan, western Guizhou, and central and western Sichuan.
- (3) With the increase in the intensity of ground motion, the maximum casualties per kilometer grid exhibit minimal variation, the casualties within the kilometer grid become more spatially concentrated, and the disaster-affected area expands. There is a nonlinear relationship between casualties and the PGA. The number of casualties increases several-fold when the 50-year probability of exceedance rises from 63% to 10%.

- (4) The likelihood of earthquakes hitting Xichang, Lancang, Xundian, Songming, and Dongchuan County, triggering a national level I emergency response within the next 50 years, exceeds 10%. During the same period, the probability of earthquakes in 100 counties causing a level I emergency response is greater than 2%. In the coming year, the chance of an earthquake in 146 counties resulting in a level I emergency response surpasses 0.5%.
- (5) In some counties of Western Sichuan and Yunnan, the total proportion of brick, wooden, civil, and bamboo/grass/adobe houses exceeds 50%, while the proportion of reinforced concrete houses is less than 20%. The proportion of civil and wood structure houses in some counties of northwest Yunnan exceeds 30%. In a few counties in western Guizhou, the proportion of brick–concrete structure houses exceeds 50%, reflecting a low seismic fortification capacity.

Author Contributions: Conceptualization, G.N. and C.X.; methodology, X.F.; software, X.F.; validation, N.Z., C.X. and G.N.; formal analysis, N.Z.; investigation, N.Z.; resources, X.F., N.X.; data curation, X.F.; writing—original draft preparation, N.Z.; writing—review and editing, N.Z., J.W.; visualization, N.Z.; supervision, X.F., J.W.; project administration, X.F.; funding acquisition, N.Z., X.F. All authors have read and agreed to the published version of the manuscript. Author Gaozhong Nie passed away prior to the publication of this manuscript. All other authors have read and agreed to the published version of this manuscript.

Funding: This research was jointly funded by the National Nonprofit Fundamental Research Grant of China, the Institute of Geology, China Earthquake Administration (IGCEA2633, IGCEA2106), the National Key Research and Development Program of China (2022YFC3003700), and the Key R&D Program Project of Ningxia Hui Autonomous Region (2024BEG02033).

Institutional Review Board Statement: Not applicable. This manuscript does not contain any studies with human or animal participants. There are no human participants in this manuscript, and informed consent is not required. All participants in this paper were informed and verbally agreed to the publication of the manuscript.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data analyzed in this manuscript involve Personally Identifiable Information (PII), and so are not provided to protect the identities of those involved. The population data used for calculating lethality is sourced from the Global Human Settlement Layer (GHSL), accessible via the European Space Agency (ESA) at: <https://human-settlement.emergency.copernicus.eu> (accessed on 1 January 2025). National bureau of statistics of China: *China Statistical Yearbook 2025*. Beijing: Chinese statistics press, 2025 [51]. China Earthquake Administration (2023) *1996~2022 Disaster Assessment Report*. Beijing: China Earthquake Press [55]. GB 18306-2015 (2015) Seismic Ground Motion Parameter Zoning Map of China [56]. National Bureau of Statistics of China, *China Population Census Yearbook-2020*. Beijing: China Statistics Press, 2022 [58]. National Bureau of Statistics of China (5 August 2009) *Compilation Rules for Statistical Zoning Codes and Urban Rural Zoning Codes* [N]. Beijing: Chinese statistics press [59]. GB/T 17742-2020 (2020) The Chinese seismic intensity scale [60]. GB/T 29425-2012 (2012) Basic Requirements for Classification of Emergency Response in Natural Disaster Relief [66]. General Office of the State Council of China (21 September 2012) [EB/OL] *National Earthquake Emergency Plan* [26 March 2024] [67].

Acknowledgments: The authors expressed their gratitude for the strong support from the CMEM, CEA, the Gansu and Qinghai Seismological Bureaus, Zhang Yingfeng from the EPCCEA for providing earthquake catalogs, and the village committees of affected areas for providing village formation information. They also extended their gratitude to Yongsheng Zhou from the Laboratory of Structural Physics at the Institute of Geology, China Earthquake Administration, for his valuable comments on the revision of this paper.

Conflicts of Interest: The authors declare no conflict of interest. The authors declare that they had no potential conflicts of interest with respect to the research, authorship, and/or publication of this manuscript.

References

1. Akimov, A.; Gemueva, K.; Semenova, N. The seventh population census in the PRC: Results and prospects of the country's demographic development. *Her. Russ. Acad. Sci.* **2021**, *91*, 724–735. [CrossRef]
2. Alel, M.; Ramly, N.; Ghafar, M. Estimation of Life Casualties for Seismic Vulnerability Assessment in Bukit Tinggi, Pahang, Malaysia. In Proceedings of the Geocon 2013 Proceedings, Bahru, Malaysia, 28–30 October 2013.
3. Alexander, D.E. Mortality and morbidity risk in the L'Aquila, Italy earthquake of 6 April 2009 and lessons to be learned. In *Human Casualties in Earthquakes: Progress in Modelling and Mitigation*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 185–197.
4. Allen, C.R.; Luo, Z.; Qian, H.; Wen, X.; Zhou, H.; Huang, W. Field study of a highly active fault zone: The Xianshuihe fault of southwestern China. *Geol. Soc. Am. Bull.* **1991**, *103*, 1178–1199.
5. Badal, J.; Samardzhieva, E. Estimation of the Expected Number of Casualties Caused by Strong Earthquakes. *Bull. Seismol. Soc. Am.* **2002**, *92*, 2310–2322. [CrossRef]
6. Badal, J.; Vazquez-Prada, M.; González, Á. Preliminary quantitative assessment of earthquake casualties and damages. *Nat. Hazards* **2005**, *34*, 353–374. [CrossRef]
7. Ceferino, L.; Kiremidjian, A.; Deierlein, G. Regional multiseverity casualty estimation due to building damage following a Mw 8.8 earthquake scenario in Lima, Peru. *Earthq. Spectra* **2018**, *34*, 1739–1761. [CrossRef]
8. Xia, C.; Nie, G.; Li, H.; Fan, X.; Yang, R.; Zeng, X. Comparative analysis of the earthquake disaster risk of cities in Eastern China based on lethal levels—a case study of Yancheng City, Suqian City and Guangzhou City. *Geomat. Nat. Hazards Risk* **2021**, *12*, 3224–3264.
9. Xia, C.; Nie, G.; Fan, X.; Li, H.; Zhou, J.; Zeng, X. A new model for the quantitative assessment of earthquake casualties based on the correction of anti-lethal level. *Nat. Hazards* **2022**, *110*, 1199–1226.
10. Cheng, J. Probability of the original time of earthquake affecting the casualty. *J. Catastrophology* **1993**, *8*, 78–85.
11. Cheng, Y.-Z.; Tang, J.; Chen, X.-B.; Dong, Z.-Y.; Xiao, Q.-B.; Wang, L.-B. Electrical structure and seismogenic environment along the border region of Yunnan, Sichuan and Guizhou in the south of the North-South seismic belt. *Chin. J. Geophys.* **2015**, *58*, 3965–3981.
12. Christoskov, L.; Samardjieva, E. An approach for estimation of the possible number of casualties during strong earthquakes. *Bulg. Geophys. J.* **1984**, *4*, 94–106.
13. Coburn, A.W.; Spence, R.J.S.; Pomonis, A. Factors determining human casualty levels in earthquakes: Mortality prediction in building collapse. In *Proceedings of the Tenth World Conference on Earthquake Engineering*; Balkema: Rotterdam, The Netherlands, 1992.
14. Cornell, C.A. Engineering seismic risk analysis. *Bull. Seismol. Soc. Am.* **1968**, *58*, 1583–1606. [CrossRef]
15. Du, Y.; Ding, Y.; Li, Z.; Cao, G. The role of hazard vulnerability assessments in disaster preparedness and prevention in China. *Mil. Med. Res.* **2015**, *2*, 27. [CrossRef] [PubMed]
16. Ferreira, M.; Oliveira, C.; de Sá, F.M. Estimating human losses in earthquake models: A discussion. In *Human Casualties in Earthquakes: Progress in Modelling and Mitigation*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 255–266.
17. Gao, X.; Ji, J. Analysis of the seismic vulnerability and the structural characteristics of houses in Chinese rural areas. *Nat. Hazards* **2014**, *70*, 1099–1114.
18. Nie, G.; Xia, C.; Fan, X.; Li, H. Research on the lethal level of buildings based on historical seismic data. *Chin. J. Geol.* **2021**, *56*, 1250–1266.
19. Goncharov, S.; Frolova, N. Casualty estimation due to earthquakes: Injury structure and dynamics. In *Human Casualties in Earthquakes: Progress in Modelling and Mitigation*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 141–152.
20. Guo, S.-L.; Liu, S.-Q.; Liu, B.-T.; Zhang, H.-Q.; Xie, F.-T. Evaluation of the degree of coordinated development of population, resource, environment and economy in contiguous areas of Sichuan, Yunnan and Guizhou, China. In *2012 International Conference on Management Science & Engineering 19th Annual Conference Proceedings*; IEEE: New York, NY, USA, 2012.
21. Hong, H.; You, J.; Tao, X. Comparative study of seismic damage induced by 2014 Ludian M S 6.5 and Jinggu M S 6.6 earthquakes in Yunnan Province. *China Earthq. Eng. J.* **2015**, *37*, 1013–1022.
22. Hongqing, G. Characteristics of time distribution of earthquakes in China and its application. *Earthq. Res. China* **1989**, *5*, 64–70.
23. Jaiswal, K.S.; Bausch, D.; Chen, R.; Bouabid, J.; Seligson, H. Estimating annualized earthquake losses for the conterminous United States. *Earthq. Spectra* **2015**, *31*, S221–S243. [CrossRef]
24. Jaiswal, K.S.; Petersen, M.D.; Rukstales, K.; Leith, W.S. Earthquake shaking hazard estimates and exposure changes in the conterminous United States. *Earthq. Spectra* **2015**, *31*, S201–S220. [CrossRef]

25. Hou, J.; Li, Y.; Song, L.; Lu, Y.; Yuan, Z. Cause analysis of M6. 6 Jinggu earthquake and M6. 5 Ludian earthquake in Yunnan in 2014. *J. Catastrophology* **2015**, *30*, 100–101.
26. Li, S.-Q.; Chen, Y.-S.; Liu, H.-B.; Del Gaudio, C. Empirical seismic vulnerability assessment model of typical urban buildings. *Bull. Earthq. Eng.* **2023**, *21*, 2217–2257. [CrossRef]
27. Li, X.; Li, Z.; Yang, J.; Li, H.; Liu, Y.; Fu, B.; Yang, F. Seismic vulnerability comparison between rural Weinan and other rural areas in Western China. *Int. J. Disaster Risk Reduct.* **2020**, *48*, 101576. [CrossRef]
28. Liu, Y.; Li, Z.; Wei, B.; Li, X.; Fu, B. Seismic vulnerability assessment at urban scale using data mining and GIScience technology: Application to Urumqi (China). *Geomat. Nat. Hazards Risk* **2019**, *10*, 958–985. [CrossRef]
29. Ceferino, L.; Kiremidjian, A.; Deierlein, G. *Probabilistic Model for Regional Multiseverity Casualty Estimation Due to Building Damage Following an Earthquake*; American Society of Civil Engineers: Reston, VA, USA, 2018.
30. Mansouri, B.; Hosseini, K.A.; Nourjou, R. Seismic human loss estimation in Tehran using, G.I.S. In Proceedings of the 14th World Conference on Earthquake Engineering, Beijing, China, 12–17 October 2008.
31. Miyakoshi, J.; Hayashi, Y.; Tamura, K.; Fukuwa, N. Damage ratio functions of buildings using damage data of the 1995 Hyogo-Ken Nanbu earthquake. In Proceedings of the 7th International Conference on Structural Safety and Reliability (ICOSSAR'97), Kyoto, Japan, 24–28 November 1997.
32. Okada, S.; Takai, N. Classifications of structural types and damage patterns of buildings for earthquake field investigation. In Proceedings of the 12th World Conference on Earthquake Engineering, Auckland, New Zealand, 30 January–4 February 2000.
33. Pomonis, A.; Kappos, A.; Karababa, F.; Panagopoulos, G. Seismic vulnerability and collapse probability assessment of buildings in Greece. In *Human Casualties in Earthquakes: Progress in Modelling and Mitigation*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 153–170.
34. So, E. Challenges in collating earthquake casualty field data. In *Human Casualties in Earthquakes: Progress in Modelling and Mitigation*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 231–254.
35. So, E.; Spence, R. Estimating shaking-induced casualties and building damage for global earthquake events: A proposed modelling approach. *Bull. Earthq. Eng.* **2013**, *11*, 347–363.
36. Spence, R.; So, E.; Cultrera, G.; Ansal, A.; Pitilakis, K.; Costa, A.C.; Tönük, G.; Argyroudis, S.; Kakderi, K.; Sousa, M.L. Earthquake loss estimation and mitigation in Europe: A review and comparison of alternative approaches. In Proceedings of the 14th World Conference on Earthquake Engineering, Beijing, China, 12–17 October 2008.
37. Spence, R.; So, E.; Scawthorn, C. *Human Casualties in Earthquakes: Progress in Modelling and Mitigation*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2011; Volume 29.
38. Sun, B.; Zhang, G. Study on seismic disaster risk distribution of buildings in mainland China. *China Civ. Eng. J.* **2017**, *50*, 1–7.
39. Tao, L.; Zhuo, Y. A county-level analysis of China's population change: Insights from the 7th population census. *Popul. Res.* **2022**, *46*, 72.
40. Wang, C.-Y.; Yang, W.-C.; Wu, J.-P.; Ding, Z.-F. Study on the lithospheric structure and earthquakes in North-South Tectonic Belt. *Chin. J. Geophys.* **2015**, *58*, 3867–3901.
41. Wang, X.-S.; Lü, J.; Xie, Z.-J.; Long, F.; Zhao, X.-Y.; Zheng, Y. Focal mechanisms and tectonic stress field in the North-South Seismic Belt of China. *Chin. J. Geophys.* **2015**, *58*, 4149–4162.
42. Wang, Z.; Zhao, D.; Wang, J. Deep structure and seismogenesis of the north-south seismic zone in southwest China. *J. Geophys. Res. Solid Earth* **2010**, *115*, jb007797.
43. Wei, B.; Nie, G.; Su, G.; Sun, L.; Bai, X.; Qi, W. Risk assessment of people trapped in earthquake based on km grid: A case study of the 2014 Ludian earthquake, China. *Geomat. Nat. Hazards Risk* **2017**, *8*, 1289–1305. [CrossRef]
44. Wen, X.-Z.; Ma, S.-L.; Xu, X.-W.; He, Y.-N. Historical pattern and behavior of earthquake ruptures along the eastern boundary of the Sichuan-Yunnan faulted-block, southwestern China. *Phys. Earth Planet. Inter.* **2008**, *168*, 16–36.
45. Wu, J.; Li, N.; Hallegatte, S.; Shi, P.; Hu, A.; Liu, X. Regional indirect economic impact evaluation of the 2008 Wenchuan Earthquake. *Environ. Earth Sci.* **2012**, *65*, 161–172.
46. Wu, S.; Jin, J.; Pan, T. Empirical seismic vulnerability curve for mortality: Case study of China. *Nat. Hazards* **2015**, *77*, 645–662. [CrossRef]
47. Xia, C.; Nie, G.; Fan, X.; Zhou, J.; Li, H.; Pang, X. Research on the rapid assessment of earthquake casualties based on the anti-lethal levels of buildings. *Geomat. Nat. Hazards Risk* **2020**, *11*, 377–398. [CrossRef]
48. Xia, C.; Nie, G.; Li, H.; Fan, X.; Yang, R. Study on the seismic lethal level of buildings and seismic disaster risk in Guangzhou, China. *Geomat. Nat. Hazards Risk* **2022**, *13*, 800–829. [CrossRef]
49. Wang, H.X.; Niu, J.X.; Wu, J.F. ANN model for the estimation of life casualties in earthquake engineering. *Syst. Eng. Procedia* **2011**, *1*, 55–60. [CrossRef]
50. Xiaowen, Z.; Maoyuan, Y. Prediction and reality of casualties caused by earthquake: Relationship between house damage and death toll. *Earthq. Sci. Technol. Inf.* **1998**, *32*–37.
51. National Bureau of Statistics of China. *China Statistical Yearbook 2025*; Chinese Statistics Press: Beijing, China, 2025.

52. Nie, G.; Xia, C.; Fan, X. Grading of anti-lethal level based on historical earthquake mortality data. *Chin. J. Geol.* **2020**, *55*, 1298–1314.
53. Xu, J.; Wei, F.; Zhang, L.; Fang, H.; Li, H. Preliminary study on evaluating the number of casualties and trapped victims by a earthquake—A case study of Zhangzhou City, Fujian Province. *J. Seismol. Res.* **2008**, *31*, 382–387.
54. Liao, X.; Gao, C.; Han, S. Forecast of direct casualty caused by earthquake in the different time of a day. *J. Nat. Disasters* **1999**, *8*, 92–96.
55. China Earthquake Administration. *1996~2022 Disaster Assessment Report*; China Earthquake Press: Beijing, China, 2023.
56. GB 18306-2015; Seismic Ground Motion Parameter Zoning Map of China. China Standards Press: Beijing, China, 2015.
57. Xu, X.; Tan, M.; Liu, X.; Wang, X.; Xin, L. Stability and changes in the spatial distribution of China's population in the past 30 years based on census data spatialization. *Remote Sens.* **2023**, *15*, 1674. [CrossRef]
58. National Bureau of Statistics of China. *China Population Census Yearbook-2020*; China Statistics Press: Beijing, China, 2022.
59. National Bureau of Statistics of China. *Compilation Rules for Statistical Zoning Codes and Urban Rural Zoning Codes (5 August 2009)*; Chinese Statistics Press: Beijing, China, 2009.
60. GB/T 17742-2020; The Chinese Seismic Intensity Scale. China Standards Press: Beijing, China, 2020.
61. Yang, S.-X.; Yao, R.; Cui, X.-F.; Chen, Q.-C.; Huang, L.-Y. Analysis of the characteristics of measured stress in Chinese mainland and its active blocks and North-South seismic belt. *Chin. J. Geophys.* **2012**, *55*, 4207–4217.
62. Yin, Z. A study for predicting earthquake disaster loss. *Earthq. Eng. Eng. Vib.* **1991**, *11*, 87–96.
63. Yuan, H.; Gao, X.; Qi, W. Assessing the seismic risk of cities at fine-scale: A case study of Haidian district in Beijing. *China Seismol. Geol.* **2016**, *38*, 197–210.
64. Zhang, Y.; Lin, Q.; Liu, Y.; Wang, Y. The quick assessment model of casualties for Asia based on the vulnerability of earthquake. *Nat. Hazards Earth Syst. Sci. Discuss.* **2018**, *2018*, 1–22. [CrossRef]
65. Zhenwu, Z.; Wenli, L. Data quality of the 7th population census and new developments of China's population. *Popul. Res.* **2021**, *45*, 46.
66. GB/T 29425-2012; Basic Requirements for Classification of Emergency Response in Natural Disaster Relief. China Standards Press: Beijing, China, 2012.
67. General Office of the State Council of China. *National Earthquake Emergency Plan [26 March 2024]*; General Office of the State Council of China: Beijing, China, 2024.
68. Du, F.; Wen, X.; Zhang, P. Post-earthquake slip and deformation of the Luhuo section of the Xianshuihe fault zone. *J. Geophys. Res.* **2010**, *53*, 2355–2366.
69. Han, Z.; Dong, S.; Mao, Z.; Hu, N.; Tan, X.; Yuan, R.; Guo, P. Geological and geomorphological evidence of Holocene activity and slip rate in the southern segment of the Xiaojiang fault zone. *Seismol. Geol.* **2017**, *39*, 1–19.
70. He, H.; Ikeda Yasuro Song, F.; Dong, X. Left-lateral slip rate and tectonic significance of the late Quaternary period in the Xiaojiang fault zone. *Seismol. Geol.* **2002**, *24*, 14–26.
71. He, H.; Yasuro, I. Discussion on the Late Quaternary Movement Characteristics and Patterns of the Anninghe Fault Zone. *Acta Seismol. Sin.* **2007**, *29*, 12.
72. Li, G. *Tectonic Activity of the Nujiang Fault Zone in Southwestern Yunnan Since the Quaternary Period*; Institute of Earthquake Prediction, China Earthquake Administration: Beijing, China, 2008.
73. Li, T.; Zhu, Y.; Yang, Y.; Xu, Y.; An, Y.; Zhang, Y.; Feng, S.; Huai, Y.; Yang, J. Comprehensive utilization of multiple crustal deformation observation data to calculate the current slip rate of the Xianshuihe fault zone. *Geophys. J. China* **2019**, *62*, 1323–1335.
74. Li, C.; Gan, W.; Qin, S.; Hao, M.; Liang, S.; Yang, F. Study on current deformation characteristics of the southern section of the southeastern margin of the Qinghai-Tibet Plateau. *J. Geophys. Res.* **2019**, *62*, 4540–4553.
75. Liu, X.; Yuan, D.; Zhang, B.; He, W.; Fang, L. Discussion on Holocene slip rate and strike-slip initiation time of Lancang Fault in southwestern Yunnan region. *J. Seismol. Eng.* **2016**, *38*, 413–422.
76. Shen, J.; Wang, Y. Estimating seismic hazard of the Xiaojiang fault zone using fault slip rate. *Seismol. Res.* **1999**, *22*, 251–259.
77. Shi, F.; You, W.; Fu, Y. Recent movement characteristics of the Xiaojiang Fault revealed by GPS data. *Seismol. Res.* **2012**, *35*, 207–212.
78. Xu, S.; Cong, Z.; Pei, Z.; Meng, Z. Study on deep slip rate of Anninghe fault zone and surrounding areas using repeated earthquakes. *J. Seismol.* **2023**, *45*, 658–670.
79. Tang, F.; Song, J.; Cao, Z.; Deng, Z.; Wang, M.; Xiao, G.; Chen, W. Motion characteristics of major fault zones around the eastern tectonic junction revealed by the latest GPS data. *J. Geophys. Res.* **2010**, *53*, 2119–2128.
80. Wang, H.; Ran, Y.; Chen, L.; Liang, M.; Gao, S.; Li, Y.; Xu, L. Estimation of sliding rate in the southern segment of the Anninghe fault zone. *Seismol. Geol.* **2018**, *40*, 967–979.
81. Wang, L.; Wang, Q.; Zhang, Y.; Wang, Y. Analysis of current movement characteristics of major fault zones in Yunnan region based on GPS. *J. Disaster Prev. Sci. Technol. Coll.* **2016**, *18*, 1–8.

82. Wang, Y.; Wang, E.; Shen, Z.; Wang, M.; Gan, W.; Qiao, X.; Meng, G.; Li, T.; Tao, W.; Yang, Y. Inversion of current activity rates of major faults in Sichuan-Yunnan region based on GPS data constraints. *China Sci. D Ser.* **2008**, *38*, 582–597.
83. Wang, Y.; Wang, M.; Shen, Z.; Shao, D.; Shi, F. Current displacement rate and seismic risk of the Nujiang Fault. *Seismol. Geol.* **2015**, *37*, 374–383.
84. Wei, W.; Jiang, Z.; Wu, Y.; Liu, X.; Zhao, J.; Li, Q.; Dong, M. Study on Motion and Strain Accumulation Characteristics of the Xiaojiang Fault Zone. *Geod. Earth Dyn.* **2012**, *32*, 11–15.
85. Wen, X.; Luo, Z.; Qian, H.; Zhou, H.; Huang, W. Segmental characteristics, geometric features and seismotectonic significance of the Holocene fault zone in the Xianshuihe River Basin. *J. Seismol.* **2011**, *11*, 362–372.
86. Zhong, K.; Liu, Z.; Shu, L.; Li, F.; Shi, Y. Kinematic characteristics of Cenozoic strike-slip movement in the Lancang River fault zone. *Geol. Rev.* **2004**, *50*, 1–8. [CrossRef]
87. Zhou, R.; He, Y.; Huang, Z.; Li, X.; Yang, T. Slip rate and strong earthquake recurrence interval of the Qian-Ning-Kangding segment of the Xianshuihe fault zone. *J. Seismol.* **2001**, *23*, 250–261.
88. Zhou, R.; Li, X.; Huang, Z.; He, Y.; Ge, T. Average slip rate of the Late Quaternary in the Daliangshan fault zone, Sichuan. *Seismol. Res.* **2003**, *26*, 191–196.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

A Physics-Informed Online Learning Framework for Landslide Displacement Prediction

Jie Zhou ^{1,2}, Nengpan Ju ^{1,2,*}, Chaoyang He ^{1,2} and Mingli Xie ^{1,2}

¹ State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu 610059, China

² College of Environmental and Civil Engineering, Chengdu University of Technology, Chengdu 610059, China

* Correspondence: jnp@cdut.edu.cn

Abstract

Current landslide displacement prediction models often suffer from insufficient integration between physical mechanisms and data-driven approaches, weak model generalizability, and limited operational applicability. To address these issues, this study develops a physics-informed online learning framework for landslide displacement prediction. The core of this framework is a Physics-informed Long Short-Term Memory network (Phys-LSTM). By embedding discretized forms of the stress balance, creep constitutive, and kinematic equations as hard constraints into the LSTM's gating mechanisms and loss function, the model ensures physically consistent predictions and enhanced interpretability throughout the learning process. Leveraging real-time data streams from the Sichuan Provincial Geological Hazard Monitoring and Warning Platform, we developed an online processing pipeline for real-time multi-source data ingestion, automated quality control, spatiotemporal alignment, and physics-informed feature engineering. A progressive three-stage learning algorithm was designed to support model cold-start, incremental training, and rolling prediction. Validation across 45 model-development landslide sites and one independent application case demonstrated the framework's significant superiority over traditional models in displacement prediction accuracy ($RMSE \leq 1.78$ mm, $R^2 \geq 0.96$), cross-site generalization stability, and its capability to capture accelerated deformation phases. This research indicates that deeply integrating geomechanical prior knowledge into an online learning framework can effectively improve the reliability, interpretability, and operational applicability of landslide displacement prediction models, thereby providing methodological support for subsequent landslide early warning applications.

Keywords: landslide prediction; multi-source monitoring; LSTM; physics-informed neural network

1. Introduction

Landslide hazards, due to their sudden onset and destructive nature, cause significant casualties and economic losses globally every year [1,2], remaining a core challenge in the field of geological disaster prevention and mitigation worldwide. Achieving accurate, timely, and operationally viable landslide displacement prediction is crucial for improving subsequent warning decision-making and reducing disaster risk. In recent years, with the proliferation of monitoring technologies, data-driven methods, particularly machine learning and deep learning models, have made significant progress in landslide displacement prediction and early warning research. However, existing studies still face two prominent

bottlenecks: First, most purely data-driven models act as “black boxes”; while they can fit the data, they lack constraints from physical mechanisms, leading to physically implausible predictions, weak extrapolation and generalization capabilities, and difficulties in explaining the internal deformation mechanisms [3]. Second, many advanced models have shown good performance only under offline and static validation settings based on fixed historical datasets and limited case studies. Here, “offline” means that the model is not continuously updated with newly acquired monitoring data, while “static” means that the trained parameters remain unchanged during deployment. As a result, such models often perform well in research validation but inadequately in real monitoring practice, where deformation patterns evolve over time and vary across landslide sites. This makes it difficult to meet the urgent operational needs of actual monitoring and warning services for model self-updating, self-adjustment, and continuous operation [4]. This stands in stark contrast to the urgent need of practical landslide monitoring and early warning systems for model self-updating and adaptive capabilities [5].

Specifically, current attempts to integrate physics and data mostly adopt “soft constraints” (e.g., using physical indicators as additional features) or post hoc verification, failing to deeply embed the core geomechanical principles governing landslide evolution into the model’s learning architecture itself [6]. Meanwhile, some studies have attempted to enhance predictive performance by integrating ensemble machine learning models [7], or by combining remote sensing data such as InSAR with deep learning for deformation monitoring and prediction [8]. However, these approaches still exhibit shortcomings in terms of the deep integration of physical mechanisms and online adaptive learning. Performance limitations in practical applications directly lead to severe challenges of insufficient warning accuracy, particularly manifested as persistently high false alarm rates [9].

To address these bottlenecks, this study aims not merely to propose an improved prediction algorithm, but to construct an integrated landslide displacement prediction framework that combines physical mechanisms with online adaptive learning. The core innovation of this research lies in: (1) Proposing the Physics-informed Long Short-Term Memory network (Phys-LSTM), which embeds discretized geomechanical equations as hard constraints directly into the network’s gating mechanisms and loss function, achieving directional guidance of the learning process by physical laws, and (2) Designing and implementing a complete adaptive online processing and learning framework, covering the entire workflow from real-time quality control of multi-source data and dynamic generation of physics-informed features, to progressive three-stage online training and rolling prediction of the model. Based on real operational data from the Sichuan Provincial Geological Hazard Monitoring and Warning Platform, this study validates the advantages of the framework in terms of prediction accuracy, generalization capability, and operational suitability, aiming to provide a theoretical foundation and engineering practice paradigm for intelligent and operational landslide displacement prediction, while offering technical support for subsequent research.

2. Data and Methods

To construct a deep learning model suitable for displacement prediction warning and constrained by physical laws, a high-quality, highly reliable, and physics-informed input dataset is the primary prerequisite. However, actual landslide monitoring data inherently presents challenges such as being multi-source and heterogeneous, noisy, and non-aligned [10]. These challenges are equally pronounced in monitoring systems based on Wireless Sensor Networks (WSNs), necessitating adaptive data processing strategies [11]. To address these issues and proactively encode prior physical knowledge into the input data level, this chapter designs a systematic online data processing and feature construction

pipeline, as detailed in Figure 1. The monitoring data of different sensors will be subject to abnormal interference in different ways, so differential pre-processing is required. GNSS data is easily affected by factors such as temperature and humidity changes, terrain obstacles, multipath effects, and receiver instability, which often lead to noise pollution and jump anomalies. Therefore, it is necessary to perform Savitzky–Golay filter jump detection filtering and EMD filter noise filtering on it. The continuous noise of data from crack gauges and micropiles is relatively low, but sudden jumps may occur due to human activities, construction interference, or local instantaneous interference. Therefore, Savitzky–Golay filters still require mutation detection filtering. Due to the fact that the disaster monitoring and early warning platform has already conducted quality control on rainfall data based on regional meteorological information, this study will not proceed with further anomaly handling. After exception handling and noise filtering, GNSS, Cracks, micro-pile inclinometers, and rainfall data are sent to the time alignment processing stage, and all monitoring data are resampled using PCHIP based on displacement prediction step size. Finally, first-order and second-order derivative features are created.

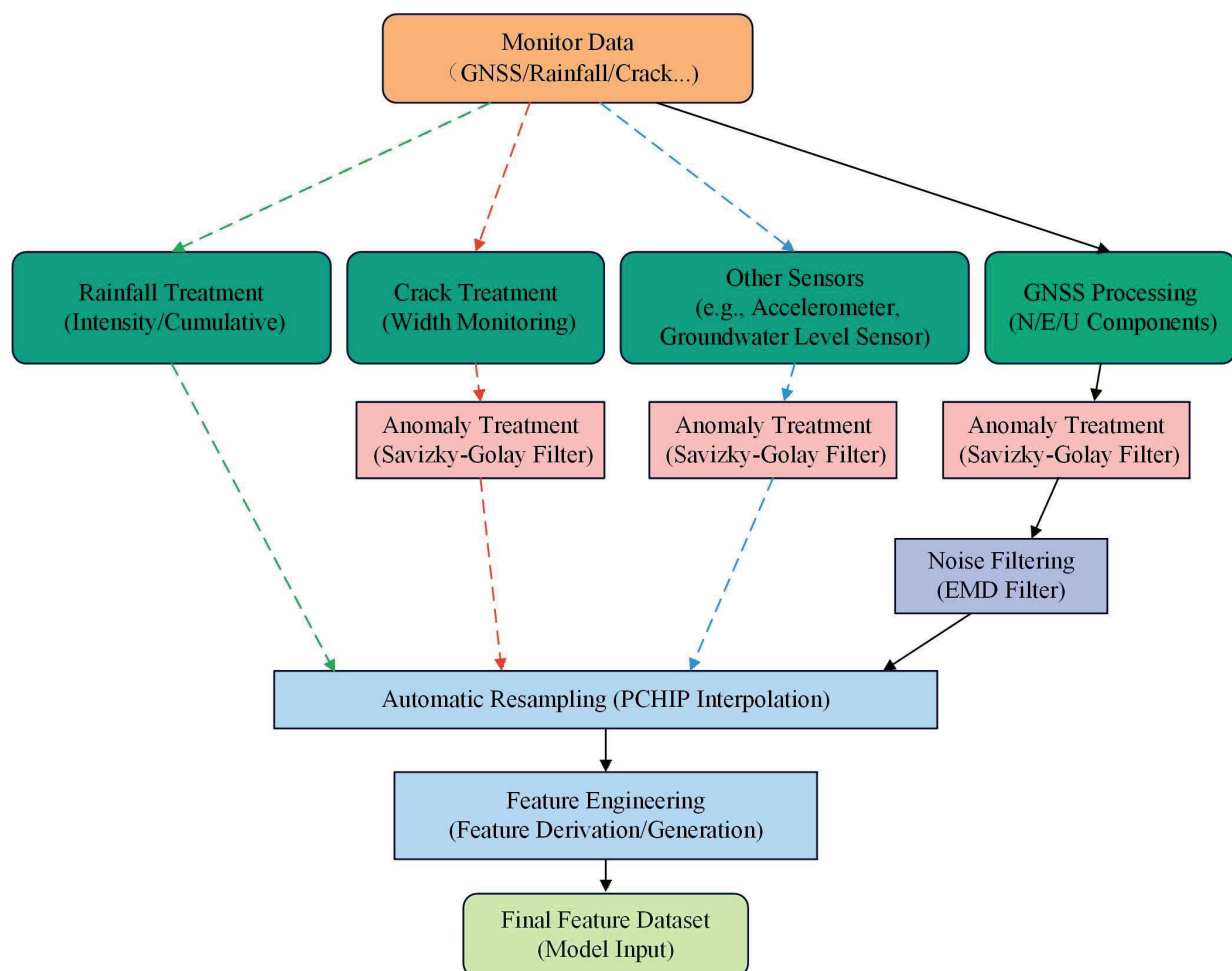


Figure 1. Feature Dataset Construction Pipeline.

2.1. Data Sources and Monitoring Equipment

The data for this study were sourced from the Sichuan Provincial Geological Hazard Monitoring and Warning Platform. For model development and cross-site validation, the dataset comprised multi-source heterogeneous monitoring data from 210 monitoring points across 45 typical landslide hazards between 2017 and 2025. In addition, the Huangcaoping deformation mass was used as an independent long-term application case, giving

46 landslide sites involved in the overall study. The monitoring system adopts an integrated space–air–ground architecture and primarily incorporates five sensor types: GNSS displacement monitoring stations, micro-pile inclinometers, crack gauges, rain gauges, and water-level gauges. The database contains 697 devices and 27,153,183 raw monitoring records, including surface displacement, deep deformation, crack dynamics, rainfall, and groundwater-level observations, and the detailed information is presented in Table 1. After quality control, resampling, and sliding-window construction, these records were transformed into 30-step historical input sequences with a 5-step forecasting horizon for model training, validation, cross-site testing, and application evaluation.

Table 1. Multi-source monitoring data details.

Data Category	Monitoring Equipment	Monitoring Parameter (Unit)	Physical Significance	Sampling Frequency	Data Volume (Records)
Surface Displacement	GNSS	3D Displacement (mm)	Direct representation of macroscopic deformation of the landslide body	once per day to once per hour	1,899,424
Deep Deformation	Micro-pile inclinometer/inclinometer	Inclination Angle (°), Deep Displacement (mm)	Indication of slip zone location and deep shear deformation	once per day to once per hour	639,727
Crack Dynamics	Crack Gauge	Crack Width (mm)	Quantitative indicator of surface tensile/shear deformation	once per day to once per second	15,808,074
Rainfall	Rain Gauge	Cumulative Rainfall (mm), Intensity (mm/h)	Primary triggering factor for landslides	once per day to once per hour	8,801,127
Groundwater Level	Water Level Gauge	Water Level Elevation (m), Pressure (kPa)	Key parameter for seepage-stress field coupling	once per hour	4831

2.2. Feature Selection

This study constructs a dataset based on all monitoring data from typical landslide monitoring sites within the Sichuan Provincial Geological Hazard Monitoring and Warning System. Each monitoring site must possess at least one GNSS displacement monitoring device. The time-series data it collects (including north, east, and vertical displacements, in mm/day) serves as the core output target for model training and prediction. GNSS displacement was selected as the core prediction target because it has much broader site coverage and better monitoring continuity in the provincial monitoring platform used in this study. Although crack gauges produced more records, this mainly resulted from their higher sampling frequency rather than longer or more complete monitoring periods. Compared with crack gauge data, GNSS observations provide more continuous and representative three-dimensional deformation information. Therefore, GNSS was used as the core output target, while crack gauge and other sensor data were treated as supplementary inputs. Auxiliary data originates from other devices at the monitoring site, such as additional GNSS devices, micro-piles (recording parameters like pile vibration acceleration and inclination angle, in m/s^2 and $^\circ$), crack gauges (recording crack width changes, in mm), and rain gauges (recording precipitation, in mm/h). These multi-source heterogeneous data collectively form the foundation of the model’s input features.

In the adaptive multi-parameter landslide displacement prediction model based on machine learning, feature selection is the key determinant of model performance and warning effectiveness. The feature selection for this model strictly adheres to three core principles: “physics-driven mechanism, data availability constraints, and prediction objective orientation.” It makes full use of the multi-source heterogeneous time-series data

collected by sensors deployed at landslide monitoring points. Feature selection is based on the actual equipment deployed at the monitoring site. If multiple devices of the same type exist (e.g., three crack gauges), they are all included in the feature set to preserve spatial variability information. The common feature is constructed in Table 2.

Table 2. Common Core Characteristics.

Feature Category	Data Source Equipment	Physical Significance	Relevance to Landslide Evolution
Core Displacement Features	GNSS Displacement Monitor	Three-dimensional surface displacement of landslide (N, E, U)	Directly characterizes the surface deformation state of the landslide body, serving as the core target and key input for model prediction.
Surface Deformation Features	Crack Gauge	Crack opening/closure width, dislocation amount	Indicates the expansion of surface tension/shear zones, serving as a sensitive indicator of impending landslide.
Triggering Factor Features	Rain Gauge	Real-time rainfall, cumulative rainfall	Main external factors triggering/accelerating landslide deformation, affecting pore water pressure and soil strength of landslide mass.
Deep Displacement Features	Micro-pile inclinometer	Inclination angle or displacement at different depths in the borehole	Reflects the location of the slip zone and deep deformation mechanism, providing direct evidence of slip surface formation.
Environmental Auxiliary Features	(Optional) Soil Moisture Content/Pore Water Pressure Sensor	Soil volumetric water content, matric suction, or pore water pressure	Reflects hydrological response state, closely related to rainfall infiltration and slip zone softening processes.

2.3. Data Preprocessing

The monitoring data from devices such as GNSS, micro-piles, and crack gauges exhibit diverse anomalous features. Among these, an outlier anomaly is defined as a spike-shaped abnormal value in the process curve, where the amplitudes on both sides are approximately comparable, while the spike value deviates from the local trend by more than three standard deviations of that feature. To avoid removing possible real deformation signals, jump-anomaly pre-filtering is applied only to historical data, and the latest observation is retained for online learning and displacement prediction. Outliers in GNSS data are often caused by multipath interference, unresolved cycle slips, or receiver faults [12,13], while outliers in micro-pile or crack gauge data are commonly induced by electromagnetic interference, instantaneous overload, or nearby disturbances (as shown in Figure 2a) [14]. Data gaps frequently result from signal obstruction, adverse weather, or network issues, which severely reduce temporal resolution and compromise the quality of periodic signal analysis in the monitoring data (as shown in Figure 2b) [15]. Direct modeling with unprocessed data would introduce significant bias. Frequent jump anomalies can seriously distort the statistical characteristics of time series, mask the true signal, and cause the denoising algorithm to retain false peaks as legitimate changes. Therefore, direct denoising of the raw data was avoided. Through the comparative experiments of different processing sequences, a systematic preprocessing pipeline is finally determined and constructed, which successively performs three core steps: Anomaly detection and repair, noise filtering, and spatiotemporal alignment. Finally, physics-guided feature enhancement is applied to improve the mechanistic interpretability of the data, providing high-quality input for the Phys-LSTM model.

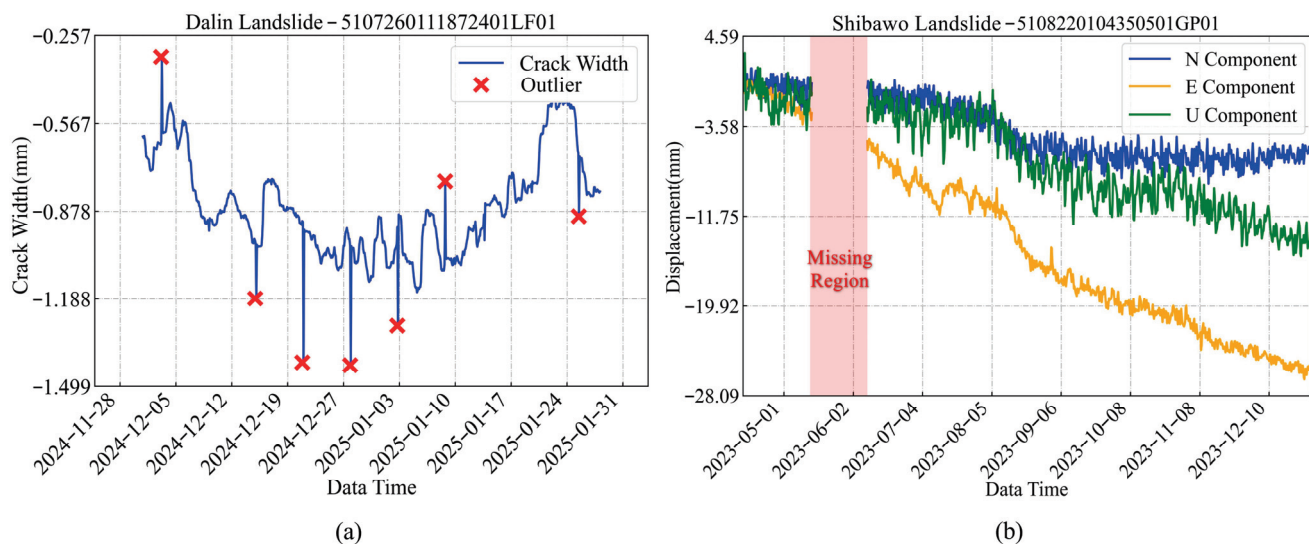


Figure 2. Typical anomalies in landslide monitoring data: (a) Jump anomaly caused by nearby disturbance, manifested as an abrupt deviation of crack-width observations from the local variation trend. (b) Continuous missing data caused by equipment malfunction; the shaded region indicates the period of observation interruption.

2.3.1. Anomaly Detection and Repair

Anomaly detection was organized into three method families: statistical tests, signal-processing residual methods, and machine-learning approaches. Statistical methods such as Z-Score [16], Modified Z-Score [17], and IQR [18] are simple and efficient, but they depend strongly on distributional assumptions and are less robust for non-stationary landslide monitoring series.

Signal-processing methods identify anomalies from local trend or residual changes. Moving-average residual detection [19] is easy to implement but may introduce phase delay, whereas the Savitzky–Golay residual method [20] better preserves local peak morphology while detecting abrupt jump anomalies.

Machine-learning methods, including Gaussian Process Regression [21], Isolation Forest [22], DBSCAN, LOF, autoencoders, and LSTM-based detectors, can represent more complex data distributions, but they are more sensitive to parameter selection and training-data representativeness. Therefore, the final preprocessing pipeline favored a residual-based method that balanced detection accuracy, computational simplicity, and operational robustness.

The comparative results in Figures 3 and 4 show that the Savitzky–Golay residual method achieved the best overall balance between missed detections and false positives, with an F1-score of 0.81. Figure 4 further illustrates four representative detection results selected from the compared algorithms. This method was therefore adopted in the preprocessing pipeline for transient jump-anomaly detection in landslide monitoring data.

2.3.2. Noise Filtering

Noise filtering is required because GNSS, micro-pile inclinometer, and crack-gauge observations are affected by multipath effects, environmental disturbance, mechanical vibration, and sensor noise. The filtering strategy therefore prioritizes signal fidelity and deformation-trend preservation rather than excessive smoothing.

For GNSS monitoring, major noise sources include multipath effects, temperature-related atmospheric delays, and unmodeled atmospheric loads. Micro-pile inclinometer records are more often affected by mechanical vibration and local disturbance [23], while crack-gauge and other sensor records can contain high-frequency sensor noise [24].

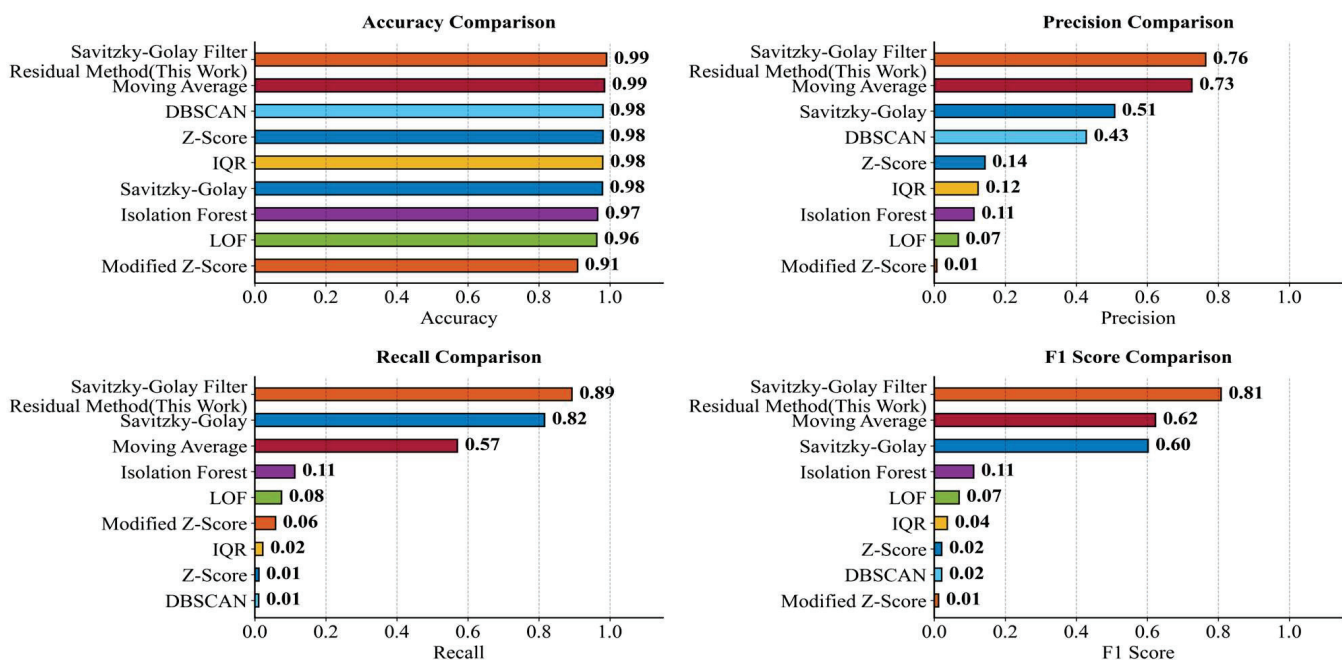


Figure 3. Performance comparison of outlier detection algorithms on representative landslide monitoring time series. Precision, recall, and F1-score were calculated against manually checked anomaly labels. The Savitzky–Golay residual method achieved the highest F1-score (0.81), indicating the best balance between missed detections and false positives for transient jump anomalies.

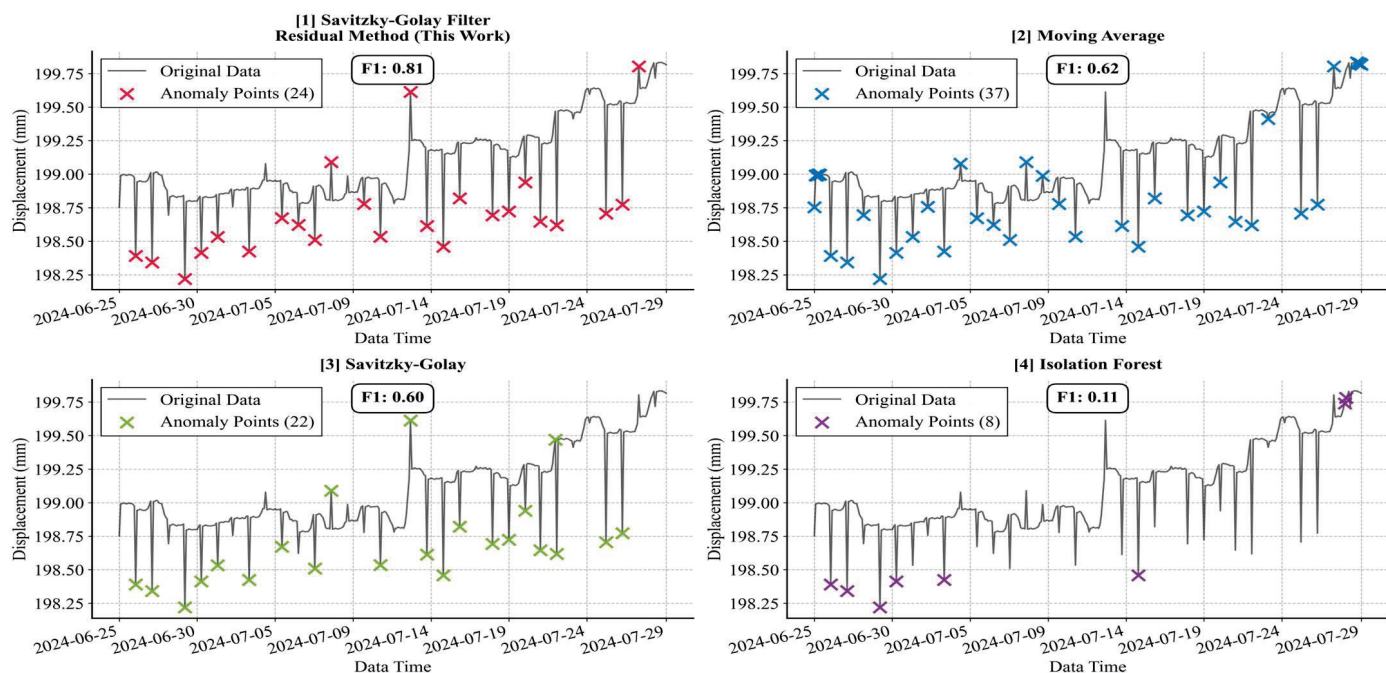


Figure 4. Comparison of anomaly detection results obtained by the four representative methods: Cross markers denote the anomalies identified by each method. Among the four methods, the Savitzky–Golay Filter Residual Method achieved the best overall performance, identifying 24 anomaly points with an F1-score of 0.81. The Moving Average method detected 37 anomalies but produced 16 false positives, yielding an F1-score of 0.62. The Savitzky–Golay method showed relatively few false detections but missed more true anomalies, with an F1-score of 0.60. Isolation Forest exhibited both frequent missed detections and false detections, resulting in the lowest performance among the four methods (F1 = 0.11).

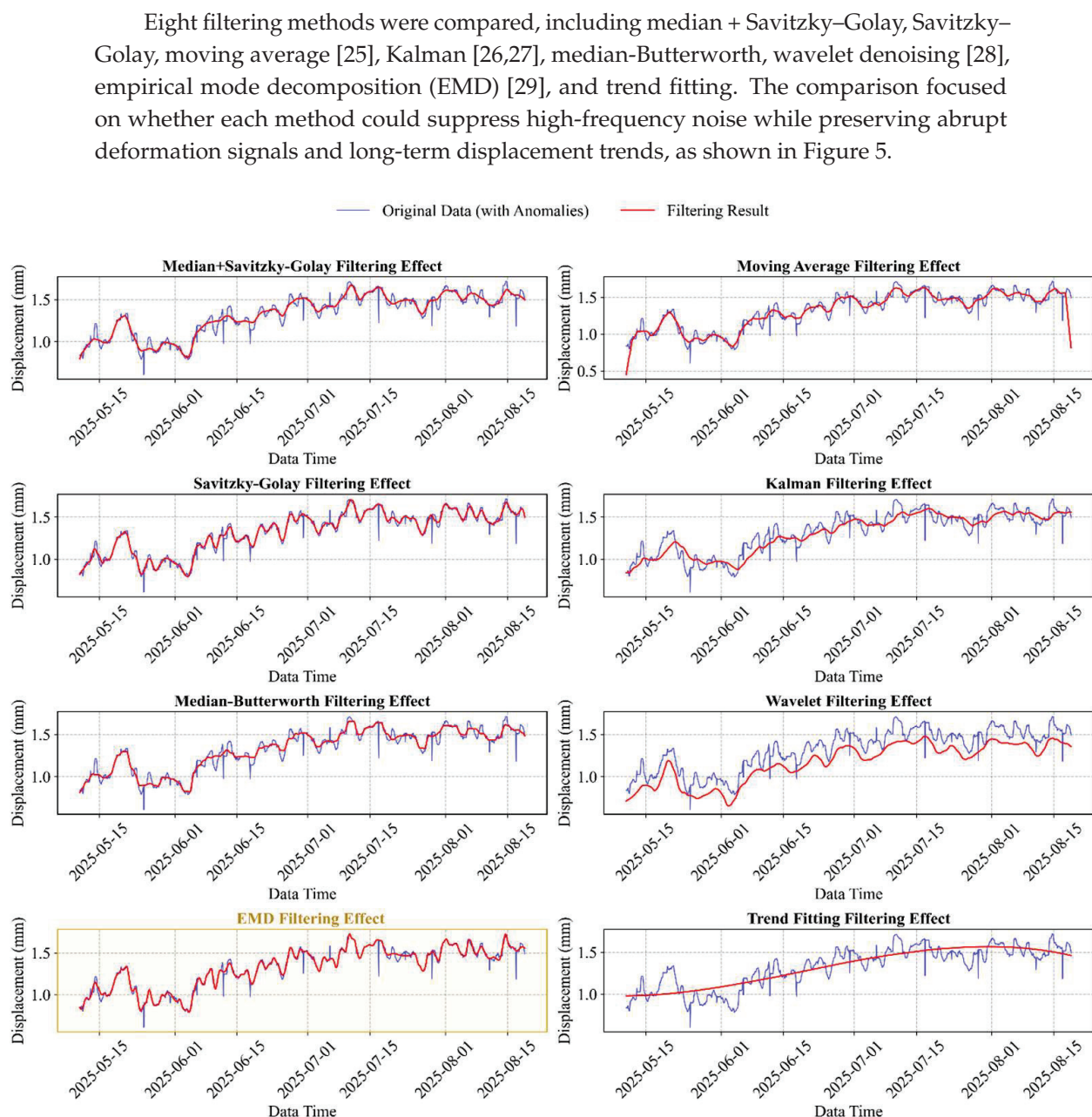


Figure 5. Comparison of denoising results obtained by eight filtering methods on the same displacement time series: All methods reduce high-frequency fluctuations and local abnormal spikes to different extents, but they differ markedly in trend preservation, local fluctuation retention, and smoothing strength. Median + Savitzky–Golay, Savitzky–Golay, Median-Butterworth, and EMD maintain high consistency with the main trend of the original sequence while preserving local variations. Among them, EMD shows the highest agreement with the original curve, indicating a better balance between noise suppression and detail preservation. By contrast, Moving Average, Kalman, and Wavelet produce stronger smoothing effects, whereas Trend Fitting mainly captures the long-term evolution trend but responds less effectively to short-term fluctuations and local anomalies.

The performance evaluation of noise filtering algorithms for landslide monitoring data prioritizes fidelity as the core criterion, supplemented by appropriate penalties for noise residual, based on three guiding principles for landslide monitoring: (a) Signal fidelity takes precedence, displacement trends and periodic characteristics must be preserved; (b) Tolerance for moderate noise, GNSS data inherently contain millimeter-level noise, and excessive smoothing is unnecessary; (c) Protection of abrupt signals, key warning

signals must not be obscured by smoothing-based metrics. The weighting scheme has been validated with actual data to ensure that high-fidelity algorithms achieve leading scores, while over-smoothing methods receive significantly lower scores, thereby aligning with engineering requirements. The comprehensive scoring formula is as follows:

$$Q = F \times (1 - \alpha \cdot \min(p_n, 1.0)) \quad (1)$$

where the fidelity F (Pearson correlation coefficient) holds the dominant weight (100%); the noise residual penalty term P_n (after normalization) sets an upper deduction limit of 10% ($\alpha = 0.1$). Here, $P_n = \frac{N_r}{\sigma(y_0 + \varepsilon)}$ (N_r : number of residual noise samples, σ : standard deviation, ε : a minimal positive value, 10^{-8}). The performance metrics of each algorithm derived from the above formula are shown in Figure 6.

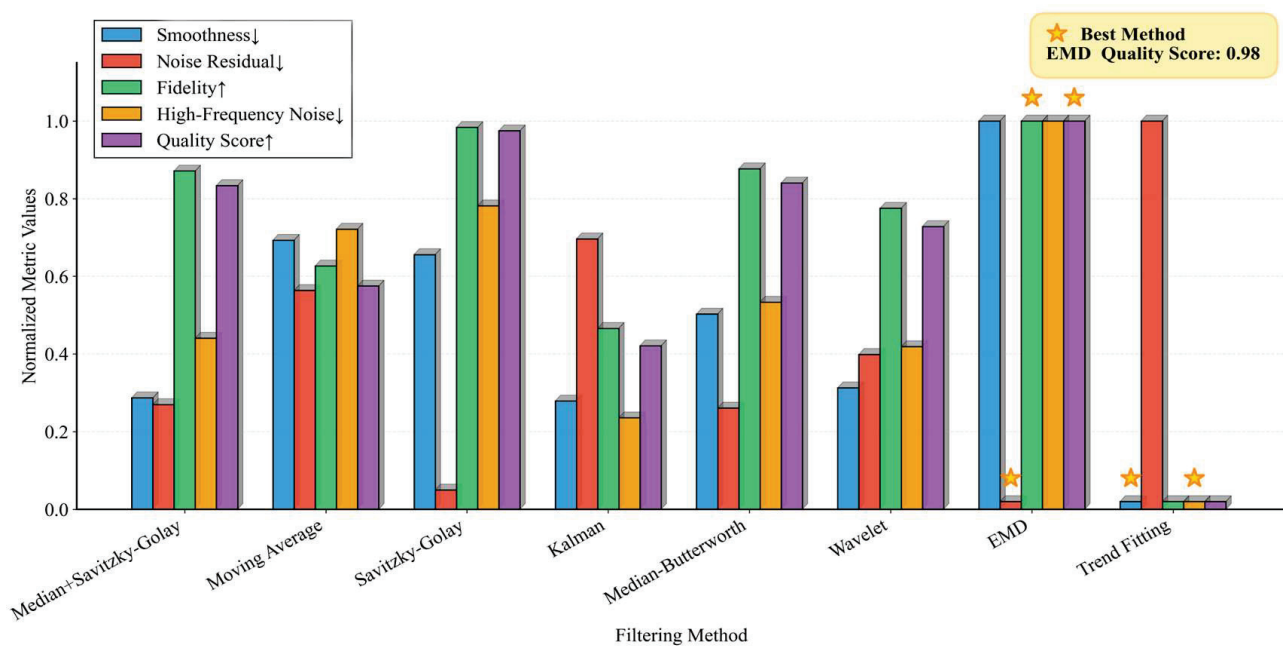


Figure 6. Performance Comparison of Different Noise Filtering Methods.

2.4. Data Alignment and Physics-Informed Feature Construction

To address sampling-frequency discrepancies among GNSS, rainfall, crack-gauge, and other monitoring data, Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) interpolation is used to resample the raw heterogeneous series to the target GNSS time series. PCHIP avoids the Runge phenomenon associated with high-order polynomial interpolation and preserves monotonic variation, making it suitable for landslide-monitoring variables. Spatially, each monitoring point is transformed into a local landslide coordinate system referenced to the main sliding direction, so displacement vectors, timestamps, and positions are aligned for cross-source coupled features such as cumulative rainfall-crack propagation rate.

To achieve deep integration of physical mechanisms with data-driven models, this section constructs a multi-dimensional feature system based on the preprocessed multi-source data. Through physics-informed feature enhancement techniques, the physical interpretability and predictive potential of the input features are significantly improved. The feature engineering framework is shown in Table 3, covering three categories of features: temporal decomposition, differential dynamics, and cross-coupling.

Table 3. Physics-Informed Feature System.

Feature Category	Feature Name	Calculation Method/Formula	Physical Meaning	Data Source
Temporal Decomposition Features	Displacement Trend Strength	Trend component $T(t)$ from STL decomposition	Long-term creep trend of the landslide, reflecting macroscopic deformation driven by gravity	GNSS
	Displacement Periodic Oscillation	Periodic component $S(t)$ from STL decomposition	Response to periodic external loads (e.g., seasonal rainfall)	GNSS, Rain gauge, Water level gauge
Differential Dynamics Features	Displacement Rate	$v(t) = \Delta D / \Delta t$ (First-order differentiation)	Instantaneous velocity of landslide movement, identifying the accelerated creep stage	GNSS
	Creep Acceleration	$a(t) = \Delta^2 D / \Delta t^2$ (Second-order differentiation)	Key indicator for accelerated creep, used for early warning and critical state judgment	GNSS
Cross-Coupling Features	Rainfall-Displacement Response Ratio	$RDRR = \Delta P / \Delta v$ (ΔP : Rainfall increment)	Quantifies the triggering efficiency of rainfall infiltration on displacement rate	Crack gauge, GNSS
	Crack-Displacement Synergy Coefficient	$C_{c-d} = \text{Corr}(\text{Crack}, \text{Velocity})$	Spatial correlation between surface rupture and deep-seated movement	Rain gauge, GNSS

The STL (Seasonal-Trend decomposition using Loess) method is employed to decompose the displacement time series into three constituent components: trend, seasonality, and residual. This decomposition facilitates the separation of deformation components attributable to distinct physical causes, which in turn enhances the accuracy of landslide displacement prediction models [30–33]:

$$D(t) = T(t) + S(t) + R(t) \quad (2)$$

where the trend component $T(t)$ reflects the long-term creep accumulation under gravitational forces and serves as a core indicator for stability prediction, and the seasonal component $S(t)$ captures repetitive deformation induced by periodic reservoir water level fluctuations or seasonal rainfall. This decomposition reduces the interference of high-frequency noise on trend prediction.

Differential Dynamics Features: Numerical differentiation is performed on the displacement trend component $T(t)$ to quantify the dynamic evolution process: displacement rate $v(t) = dT/dt$, and creep acceleration $a(t) = d^2T/dt^2$ serves as a key indicator for identifying the accelerated creep stage. A condition where $a(t) > 0$ and continues to rise indicates that the landslide is entering a critical unstable state, providing a physical basis for early warning thresholds [34–36].

Physics-Informed Cross-Coupling Features: Cross-features reflecting the coupling between external triggers and internal deformation are constructed based on multi-field coupling theory:

The Rainfall-Displacement Response Ratio (RDRR) is defined as $RDRR = \Delta P_{3h} / \Delta v_{24h}$, where ΔP_{3h} is the 3-h cumulative rainfall increment (mm), and Δv_{24h} is the 24-h displacement rate change (mm/day). This index quantifies the efficiency of rainfall infiltration in softening the slip zone, with a high RDRR value indicating strong regional sensitivity to rainfall [33,37–40].

The Crack-Displacement Synergy Coefficient (C_{c-d}) is calculated as:

$$C_{c-d} = \frac{\text{Cov}(\text{Crack}, \text{Velocity})}{\sigma_{\text{Crack}} \cdot \sigma_{\text{Velocity}}} \quad (3)$$

where $\text{Cov}(\text{Crack}, \text{Velocity})$ is the covariance between the crack width change series and the displacement rate series, σ_{Crack} and σ_{Velocity} are the standard deviations of the respective series. This coefficient links surface crack monitoring with deep-seated displacement, enhancing the understanding of the overall deformation mechanism. The value of C_{c-d} ranges from -1 to 1 , with its absolute value closer to 1 indicating stronger synergy [41].

This study first employs Pearson correlation analysis to rapidly eliminate features that show no significant correlation with the displacement target variable. Subsequently, the SelectKBest (F-regression) method is used to select features with the most significant linear relationship to the target variable from the preliminarily screened set. By comparing the average correlation and normalized F-scores of each feature, the contribution of each feature to displacement prediction is evaluated. The results, as shown in Figure 7, indicate that displacement rate, periodic component amplitude, and the GNSS elevation component rank as the top three contributors. Among these, the displacement rate achieves the highest F-score with an average correlation of 43.4%, which fully confirms the effectiveness of the physics-informed features. Furthermore, although rainfall lag features and crack-related features show relatively lower coverage, they still hold certain reference value. Based on the above analysis results, ten core features are ultimately selected as input for the Phys-LSTM model, thereby avoiding the curse of dimensionality while ensuring information density. Through this feature engineering process, the principles of geomechanics and multi-field coupling effects are transformed into quantifiable feature indicators, laying a solid foundation for the model to achieve “physics-constrained data-driven” learning.

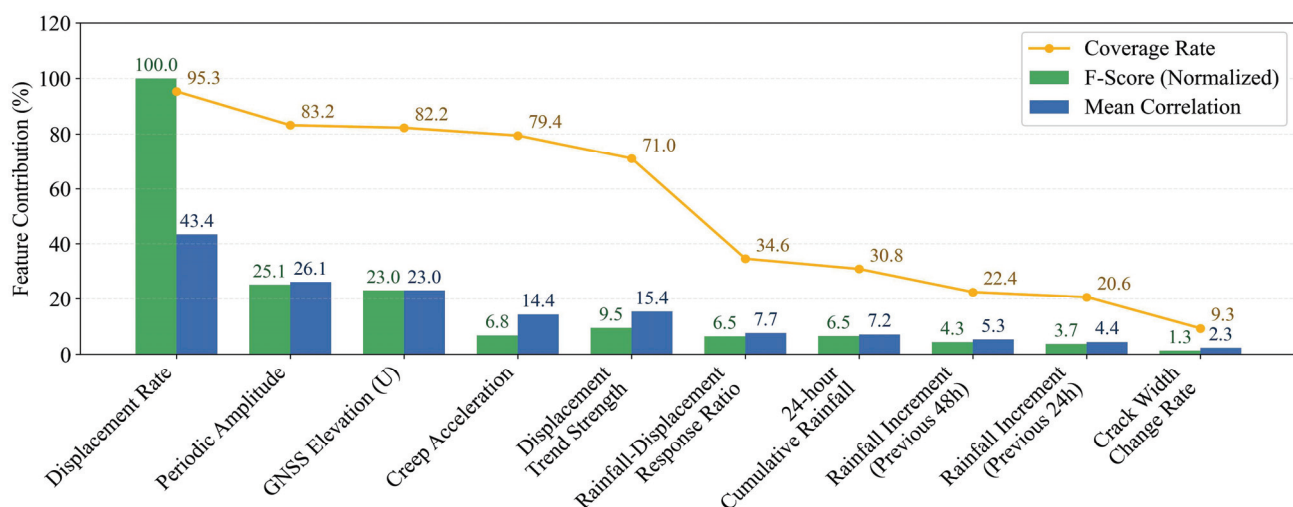


Figure 7. Feature contribution in displacement prediction. The bars summarize normalized F-scores and average correlations between candidate physics-informed features and the GNSS displacement target; higher values indicate stronger explanatory contribution to the displacement-prediction task.

3. Design of the Phys-LSTM Model Architecture and Integration of Physical Constraints

This chapter elaborates on the overall architecture of the Phys-LSTM model. Its core innovation lies in the discretization and computationally reconstruction of key physical laws governing landslide evolution (stress equilibrium, creep laws, kinematic equations) and deeply embedding them into the gating mechanisms and loss function of the LSTM network in the form of hard constraints. This achieves a deep integration of physical mechanisms and data-driven capabilities. Specifically, this chapter first improves three fundamental physical equations to make them suitable for the deep learning framework. Subsequently, a dual-branch collaborative model architecture is designed, enabling dynamic guidance of the data-driven process by physical laws through the parallel processing

and interaction of a physics-informed feature extraction branch and an LSTM temporal learning branch. Finally, a multi-task loss function incorporating physical constraints is defined to ensure the model's predictions comply with fundamental physical laws. The overall design aims to construct an adaptive prediction framework that integrates mechanisms and data, enhancing the model's prediction accuracy, generalization ability, and physical interpretability in complex environments.

3.1. Physical Reformulation of Landslide Dynamics

A reliable physical model provides the foundation for mechanism-guided prediction. To achieve a deep integration of physical laws and data-driven modeling, this study first systematically refines and computationally reformulates three core physical equations governing landslide evolution. Building on this basis, we propose a landslide displacement prediction architecture that integrates dynamic feature engineering with online adaptive learning. The central design philosophy of this framework is to transform stability evolution, creep accumulation, and displacement kinematic consistency into physically guided state variables and regularization principles that can be continuously updated from multi-source monitoring data through their coordinated interpretation. This reformulation is not intended to replace rigorous geotechnical analysis; rather, under conditions where only real-time monitoring data are available, it is designed to preserve, to the greatest extent possible, physically meaningful evolutionary trends and constraint relationships, thereby enabling accurate short-term and imminent prediction of landslide deformation trends.

3.1.1. Reformulation of the Factor of Safety for Online Prediction

In classical limit-equilibrium analysis, the factor of safety (*FOS*) is generally defined as the ratio between the total resisting force and the total driving force along a potential sliding surface, and it serves as a core indicator for evaluating slope or landslide stability. Its general form can be written as:

$$FOS = \frac{\text{total resisting force}}{\text{total driving force}} \quad (4)$$

where the numerator represents the overall resistance mobilized along the potential sliding surface, whereas the denominator represents the total downslope driving action. For geotechnical slopes, the resisting term is primarily controlled by the shear strength of the sliding zone soil or weak structural surfaces, which in turn is governed by cohesion c , internal friction angle ϕ , normal stress conditions, and pore-water pressure. The driving term is mainly associated with slope geometry, self-weight, and the gravitational components acting in the sliding direction. In essence, the classical factor of safety reflects the mechanical competition between resisting and driving actions within the slope mass. For clarity, the stress-dependent resistance discussed below is interpreted primarily in the effective-stress sense, in which the shear-strength relation can be expressed as $\tau_f = c' + \sigma'_n \tan \phi'$ and the available resistance depends on the effective normal stress $\sigma'_n = \sigma_n - u$ rather than on total normal stress alone.

Although this formulation has clear engineering significance in static stability analysis, its direct application becomes problematic in a monitoring-driven online prediction framework [42]. We therefore clarify that the discussion in this study refers primarily to an effective-stress-based interpretation of slope resistance. Under a conventional total-stress or effective-stress limit-equilibrium calculation, cohesion and friction angle can be treated as constant input parameters for a specified material state, drainage condition, and stress path. In the present online framework, however, rainfall infiltration changes pore-water pressure and hence effective normal stress, while crack propagation and cumulative deformation

indicate progressive damage. The time-varying quantity represented by the model is therefore not a direct variation in the intrinsic c' and ϕ' values themselves, but the evolution of mobilized shear resistance under changing effective-stress and damage conditions.

From this perspective, the temporal degradation term is governed by two observable processes. Rainfall intensity is used as a proxy for hydrological forcing because infiltration raises pore-water pressure and reduces effective normal stress, thereby lowering the mobilized frictional resistance even when c' and ϕ' are kept constant in the classical strength equation. Movement velocity is used as a proxy for progressive structural damage, crack development, and shear disturbance, which may appear as equivalent strength degradation at the monitoring scale. Their combined effect controls the temporal evolution of the stability state.

Because complete geotechnical parameters, pore-pressure fields, and sliding-surface conditions cannot be updated continuously from the monitoring data, the present study does not attempt to invert $c'(t)$ and $\phi'(t)$ independently. Instead, their combined influence on the effective-stress-controlled mobilized resistance is absorbed into a reduced-order stability degradation term, denoted as $\Gamma(t)$. This term should be interpreted as a monitoring-derived proxy for time-varying effective shear resistance and damage accumulation, rather than as a direct inversion of the fundamental strength parameters. Under this assumption, the temporal evolution of slope stability can be written as:

$$SI(t) \propto \frac{1}{1 + \Gamma(t)} \quad (5)$$

where a larger value of $\Gamma(t)$ indicates stronger degradation and thus lower stability. To retain physical interpretability while maintaining parameter identifiability and computational efficiency for online prediction, the composite degradation term is further approximated as a weighted combination of a rainfall-related term and a velocity-related term, namely

$$\Gamma(t) \approx k_r \tilde{R}(t) + k_v \tilde{v}(t)^2 \quad (6)$$

where $\tilde{R}(t)$ is the normalized rainfall intensity, or more generally the rainfall-related hydrological forcing, $\tilde{v}(t)$ is the normalized movement velocity magnitude, and k_r and k_v are scaling coefficients representing the relative contributions of rainfall forcing and deformation activity to stability degradation. A quadratic form is adopted for the velocity term because the destabilizing effect of deformation activity is typically amplified in a nonlinear manner: during the low-velocity stage, the slope may remain in a stable or quasi-stable creep state, whereas a marked increase in velocity usually indicates that the system is approaching an accelerated instability stage. The use of v^2 is therefore more appropriate for characterizing this enhancement in sensitivity.

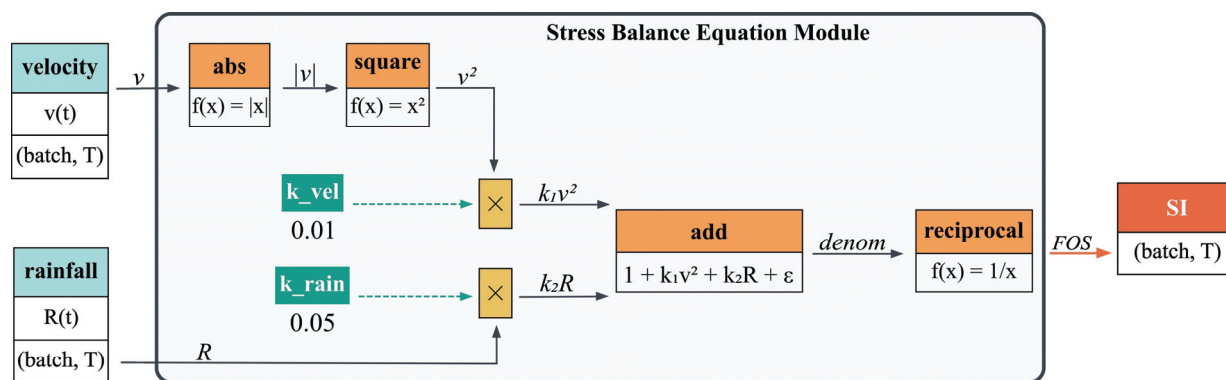
Substituting the above degradation term into the stability representation yields the dynamic stability index used in this study for online learning,

$$SI(t) = \frac{1}{1 + k_r \tilde{R}(t) + k_v \tilde{v}(t)^2 + \varepsilon} \quad (7)$$

where $SI(t)$ denotes the dynamic stability index at time t , k_r and k_v are scaling coefficients, and ε is a small positive constant introduced to avoid a zero denominator. This reciprocal formulation has several desirable properties. A larger degradation term results in a smaller index value, thereby capturing the physically meaningful trend that stronger rainfall forcing or more intense deformation activity corresponds to lower stability. The expression remains positive for all time steps, which makes it suitable for embedding into the network as

a stability-related state variable. At the same time, its compact form is convenient for continuous updating under an online monitoring framework.

In this way, the stability concept governed by the competition between resisting and driving actions in classical *FOS* theory is reformulated into a dynamic stability index $SI(t)$ that is suitable for monitoring-data-driven conditions. Rather than relying on the real-time updating of complete engineering geological parameters, this index provides a reduced-order description of stability degradation through rainfall forcing and deformation activity, thereby offering a computable physical characterization of the temporal evolution of slope stability. In the subsequent model, $SI(t)$ is introduced as a monitoring-derived physical state variable into the feature representation and temporal learning process, and it further provides mechanism guidance for network optimization through gate modulation and the design of constraint terms. Through this treatment, the physical insight from classical stability analysis is embedded into the online learning framework, allowing the model to retain the adaptability of data-driven learning while achieving stronger physical consistency and interpretability. We developed the dynamic stability index evolution model shown in Figure 8. This module maps dynamic features such as displacement velocity and rainfall into a time-series stability index through learnable parameters.



Formula: $SI = 1 / (1 + k_{vel} \times v^2 + k_{rain} \times R + \epsilon)$

Where: • v : velocity magnitude • R : rainfall • ϵ : $1e-8$ (numerical stability)

Figure 8. Computational flowchart of the dynamic stability index evolution model.

3.1.2. Discretization of the Creep Constitutive Relation

The classical Burgers creep model, which describes the time-dependent deformation of geomaterials, is expressed as a continuous function:

$$\epsilon(t) = \frac{\sigma}{E_1} + \frac{\sigma}{\eta_1} \cdot t + \frac{\sigma}{E_2} \cdot \left[1 - \exp\left(-\frac{E_2}{\eta_2} \cdot t\right) \right] \quad (8)$$

where E_1 and E_2 are elastic parameters, η_1 and η_2 are viscous parameters, and σ is the constant stress. This model simultaneously captures three deformation stages: instantaneous elastic deformation, decaying (primary) creep, and steady-state (secondary) creep. Such a constitutive description is particularly relevant for landslide deformation because monitored displacement commonly exhibits progressive accumulation, delayed response, and stage-dependent acceleration under sustained external disturbance. These characteristics make creep theory a physically meaningful basis for representing long-term deformation evolution. However, a fundamental discrepancy exists between its continuous time variable t and the input of discrete-time sequence models like LSTM, which require data at equally spaced time intervals. This paper employs the Euler method to discretize the continuous time t into a sequence with time step Δt . By applying a finite difference approximation to

the original equation and reformulating it into a recurrence relation based on the strain at the previous time step ε_t , a discrete-time framework is established. Furthermore, to account for complexities not fully captured by the discretization process or the model itself—such as non-uniform stress distribution—the hidden state H_t of the LSTM network is incorporated as a corrective term to capture the unmodeled complex nonlinear creep behavior [5,43–46]. The final discrete-time creep equation is obtained as:

$$\varepsilon_{t+1} = \varepsilon_t + \frac{\sigma}{E_1} \Delta t + \frac{\sigma}{\eta_1} (\Delta t)^2 + \frac{\sigma}{E_2} \left[\exp\left(-\frac{E_2}{\eta_2} \Delta t\right) - 1 \right] \varepsilon_t + \text{LSTM}_{\Delta\varepsilon}(H_t) \quad (9)$$

where $\text{LSTM}_{\Delta\varepsilon}(H_t)$ represents the output of the LSTM network's correction module, which learns to model the discrepancy between the ideal physical model and observed complex system behavior. The implementation of this formula within the Phys-LSTM landslide prediction model is illustrated in Figure 9, where the physical model is naturally embedded into the stepwise computation of the temporal network. This enables the LSTM to learn residuals within the framework of physical laws, thereby enhancing the model's interpretability and extrapolation capability.

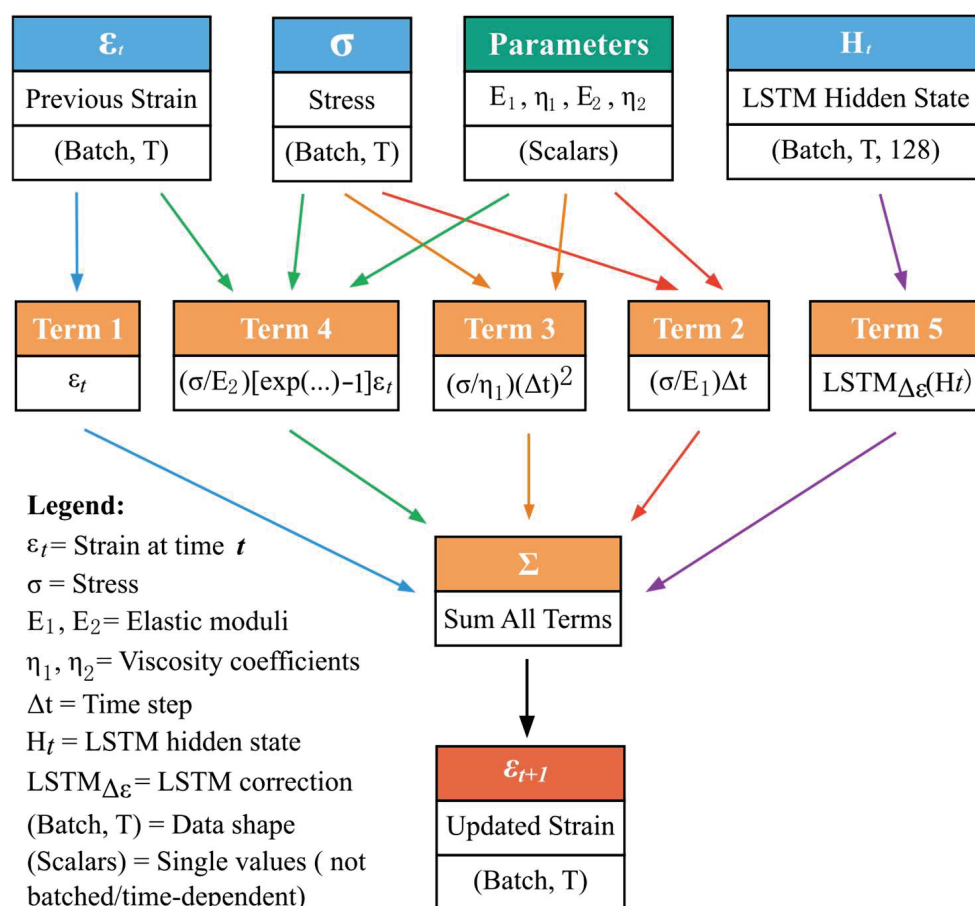


Figure 9. Computational Flowchart of the Time-Dependent Parameter Creep Constitutive Model.

3.1.3. Kinematic Constraint and Physics-Informed Loss

In classical uniformly accelerated linear motion, displacement results from the superposition of a uniform motion component due to initial velocity and a uniformly variable motion component due to acceleration. The kinematic relationship is expressed as:

$$x_{t+1} = x_t + \dot{x}_t \Delta t + \frac{1}{2} \alpha_t (\Delta t)^2 \quad (10)$$

where x_{t+1} is the position at time $t + 1$, x_t is the position at time t , $\dot{x}_t$ is the instantaneous velocity at time t , Δt is the time interval between the two moments, in seconds, and a_t is the instantaneous acceleration at time t . In the context of landslide displacement prediction, this kinematic relationship provides a direct physical link among displacement, velocity, and acceleration, and is therefore well suited for constraining the temporal consistency of predicted motion. Compared with a fully coupled mechanical simulation of the sliding body, such a representation retains a clear physical basis while remaining sufficiently compact for integration into a data-driven learning framework.

In this study, quantifying the uncertainty of machine learning-based predictions is crucial for reliable early warning and risk assessment. Uncertainty is introduced by treating the acceleration as a random variable, whose distribution parameters are learned by the LSTM from the data. This enables the model to predict uncertainty ranges [33,47]. The final improved probabilistic motion equation is as follows:

$$x_{t+1} = x_t + \dot{x}_t \Delta t + \frac{1}{2} \tilde{a}_t (\Delta t)^2, \quad \tilde{a}_t \sim \mathcal{N}(\mu_a(t), \sigma_a^2(t)) \quad (11)$$

where $\mu_a(t)$ and $\sigma_a^2(t)$ are output by the LSTM based on the historical sequence H_t . The model ultimately outputs the displacement for future time steps. To ensure that the predicted sequence remains not only numerically accurate but also physically consistent with the underlying motion state, this paper designs a multi-task loss function that integrates data fitting error and physical equation residual based on the improved discretized kinematic equation [43]. In this way, the kinematic equation is not used as a full deterministic simulation model, but as a physically grounded regularization principle that guides the network toward temporally coherent displacement prediction. The total loss function is defined as:

$$L_{\text{total}} = \lambda_{\text{data}} L_{\text{data}} + \lambda_{\text{physics}} L_{\text{physics}} + \lambda_{\text{bound}} L_{\text{bound}} \quad (12)$$

where L_{total} is the mean squared error between predicted and true values; L_{physics} is the physical equation residual loss, based on the discretized kinematic equation, ensuring the predicted trajectory conforms to motion laws. Its expression is:

$$L_{\text{physics}} = \frac{1}{T-1} \sum_{t=1}^{T-1} \left\| (x_{t+1} - x_t) - \left[\dot{x}_t \Delta t + \frac{1}{2} \tilde{a}_t (\Delta t)^2 \right] \right\|^2 \quad (13)$$

L_{bound} is the boundary constraint loss. Based on the physical meaning of the improved stress balance equation, a boundary constraint loss term is designed to ensure the predicted velocity does not exceed a maximum critical velocity and that the safety factor is non-negative and aligns with physical common sense. Its expression is as follows:

$$L_{\text{bound}} = \sum_{t=1}^T [\max(0, v_{\text{max}}) + \max(0, -FOS(t))] \quad (14)$$

The loss weights λ employ an adaptive strategy [48]. Higher weights are assigned to the physics constraints in the early stages of training to guide the network in quickly learning the fundamental laws. As training progresses, the weight for data fitting is gradually increased, allowing the model to better fit the real observed data. Based on the total loss function in Equation (10), the corresponding loss function constraint terms are constructed. The computational flow and the derivation relationships among the various physical quantities (displacement d , velocity v , acceleration a) are detailed in Figure 10.

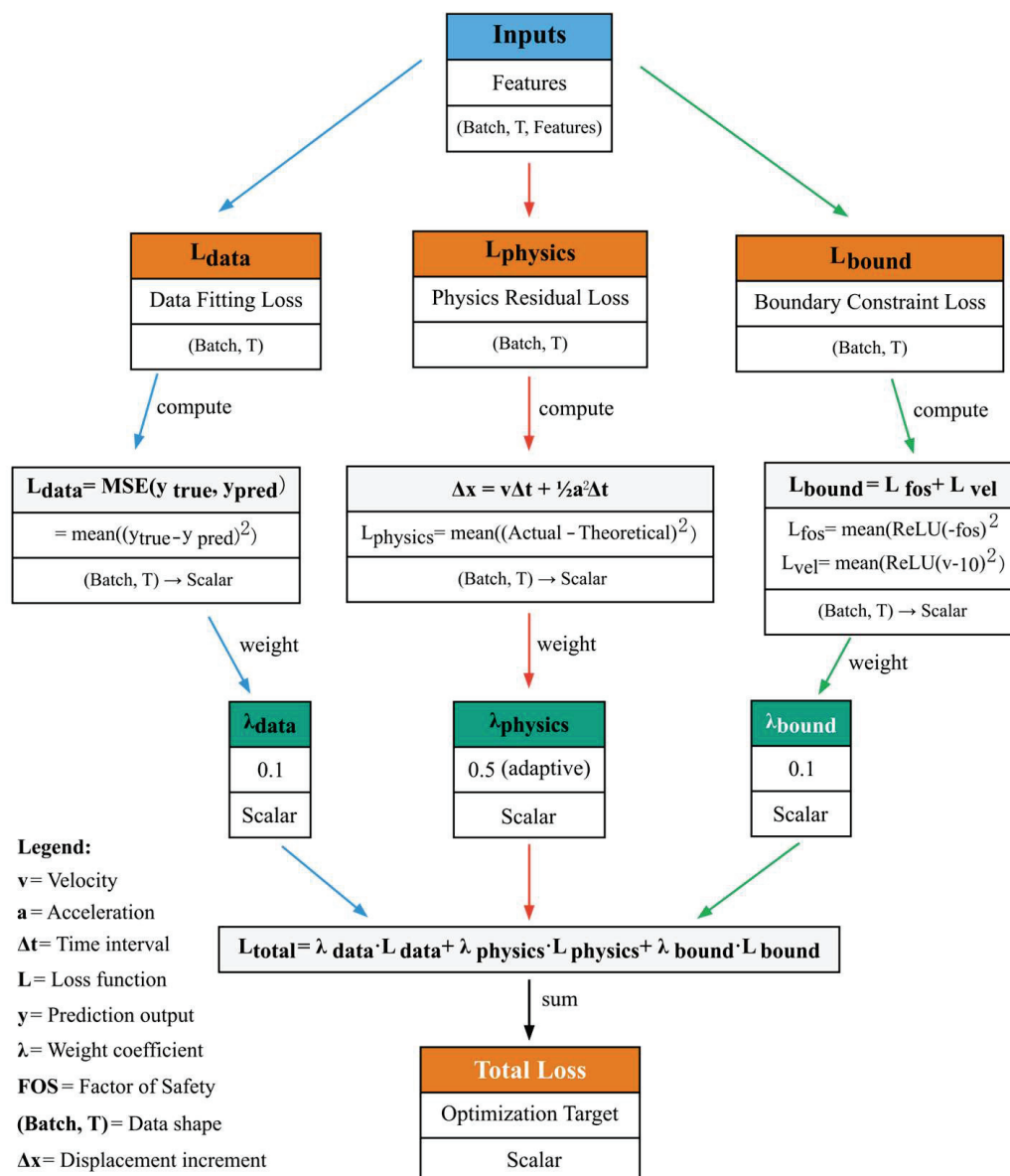


Figure 10. Neural Network Architecture Diagram with Loss Function.

3.2. Phys-LSTM Architecture

Based on the constraints of the aforementioned physical equations, this paper constructs the overall architecture of the Phys-LSTM model as shown in Figure 11. This architecture adopts a dual-branch hybrid design, achieving synergistic learning between mechanism and data through a physics-guided interaction mechanism.

3.2.1. Input Representation and Masked Feature Encoding

To accommodate the reality of potentially missing data in online monitoring scenarios, a Masked Input Layer is designed as the first layer of the model. This layer accepts the raw feature vector containing missing values, denoted as X_{input} , along with a binary mask vector M . It identifies valid features through element-wise multiplication, $X_{\text{masked}} = X_{\text{input}} \odot M$, thereby endowing the model with the ability to operate robustly under conditions of incomplete data. The resulting masked historical sequence is then treated as the input representation for the LSTM encoder–decoder forecasting path. The encoder aggregates the 30-step monitoring window into a latent deformation state that contains deformation trend, rainfall response, and stability information, and the decoder maps this state to

the 5-step future displacement trajectory. Thus, the encoder–decoder structure supports multi-step rolling prediction rather than single-step regression, while allowing the stability index, creep increment, and kinematic residual loss to guide physically consistent future trajectories.

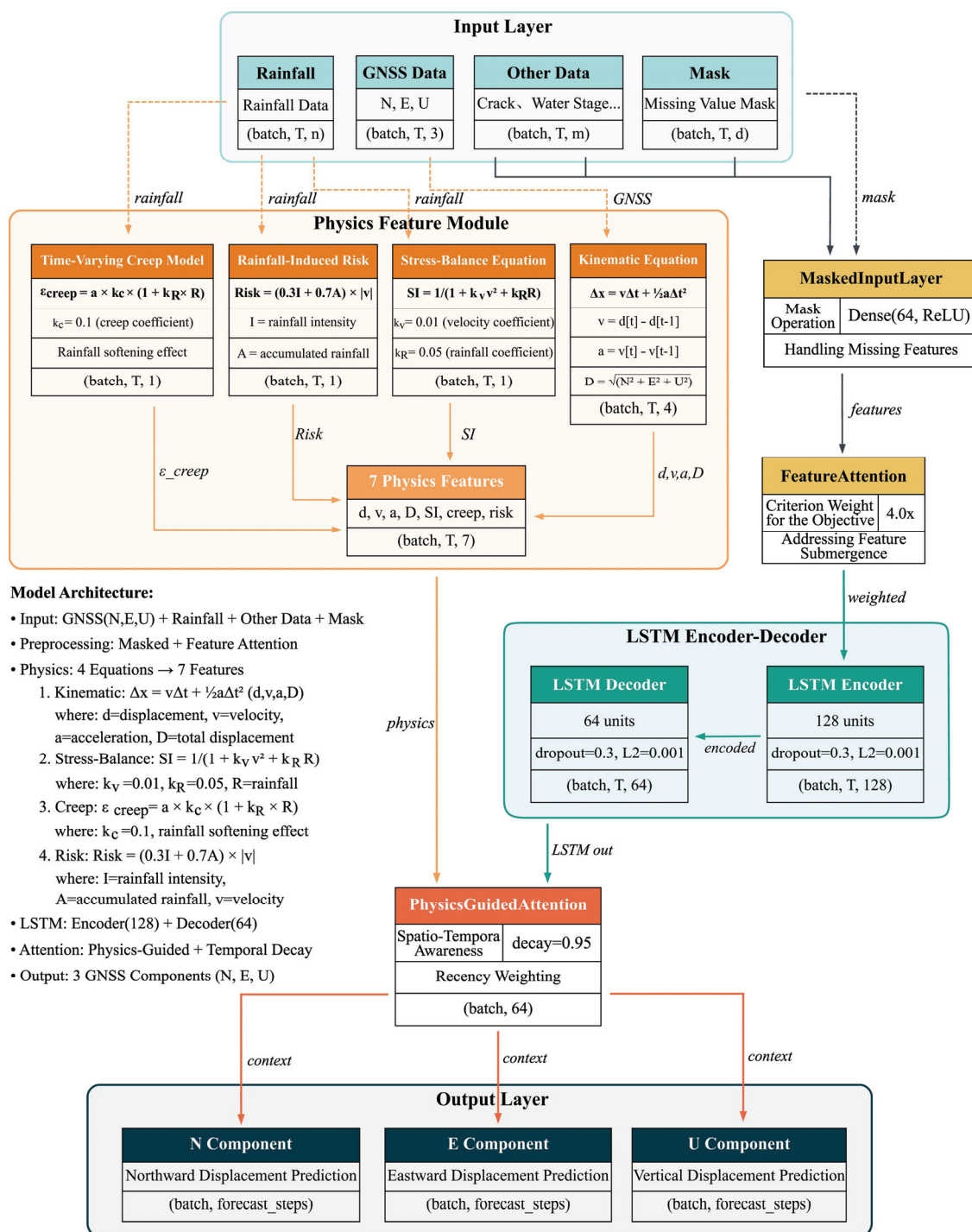


Figure 11. Architecture of the Phys-LSTM model, including masked input encoding, physics-informed feature extraction, LSTM encoder–decoder forecasting, physics-guided attention, and physics-consistency loss.

3.2.2. Physics Feature Module and Dynamic State Variables

The core of the model is a parallel processing structure consisting of two components: a physics-informed feature extraction branch and an LSTM temporal learning branch. The

physics-informed feature extraction branch executes the discretized and reconstructed physical equations from Section 3.1 (stress balance, creep constitutive equation). Based on real-time input data, it calculates a series of physics-informed intermediate features, such as the dynamic stability index ($SI(t)$) and creep increment (δ_{creep}). The LSTM temporal learning branch is a standard multi-layer LSTM network responsible for learning complex nonlinear temporal patterns from historical monitoring sequences.

3.2.3. Physics-Guided Temporal Learning

To achieve deep integration of information from the two branches, the Phys-LSTM model implements physics-guided interaction through a gating mechanism. Building upon the standard LSTM forget gate, a physics-informed prior guidance term based on the computed stability index (SI) is incorporated. This enables the model to decide on memory retention based on the real-time stability state of the landslide. The expression is as follows:

$$f_t = \sigma(W_f \cdot [H_{t-1}, X_t] + b_f + P_f \cdot SI(t)) \quad (15)$$

where σ is the Sigmoid activation function (output range 0–1), with values approaching 1 indicating “complete retention” and values approaching 0 indicating “complete forgetting”; W_f is the weight matrix of the forget gate, learning to extract feature patterns most relevant to memory retention from the historical state and current input; $[H_{t-1}, X_t]$ denotes the concatenation of the hidden state and the current input, where H_{t-1} encapsulates the summary of the historical deformation sequence, and X_t contains measured data such as displacement and rainfall at the current timestep; b_f is the bias term of the forget gate, adjusting the baseline activation level; and p_f is the physics-guided weight matrix, learning how the physical state influences the memory retention strategy.

The physics-guided term $P_f \cdot FOS(t)$ integrates real-time stability assessment into memory control. When $FOS(t)$ decreases (indicating reduced stability), the value of this term decreases, guiding the forget gate f_t to tend towards retaining more historical state information. This enhances the model’s memory of precursor sequences associated with danger, thereby aiding in more accurate early-warning judgments.

The temporally sequential data, enhanced by physics-informed features, is then fed into a multi-layer LSTM network. Following the LSTM layers, an attention mechanism is introduced to construct an attention-enhanced LSTM temporal encoder. The calculation of attention weights depends not only on the hidden state H_t but is also influenced by real-time physics-informed features. By directly linking the attention weight calculation to the importance of physical features characterizing landslide deformation and failure, the model is actively guided to focus on critical physical phases of the landslide, enabling earlier and more accurate warnings. Its expression is as follows:

$$\alpha_t = \frac{\exp(\text{Score}(H_t, \text{Physics}_t))}{\sum_{i=1}^T \exp(\text{Score}(H_i, \text{Physics}_i))} \quad (16)$$

where α_t is the attention weight (range 0–1), with higher weights indicating that the information at that timestep is considered more critical by the model when making decisions; H_t is the temporal hidden state, encoding dynamic information from the historical sequence, representing the data-driven learning component of the model; Physics_t is the physics-informed feature vector, enabling the model to actively focus on dangerous periods when the physical state undergoes abrupt changes; Score is the scoring function, comprehensively evaluating the significance of timestep t in terms of both data patterns and physical laws, defined as $\text{Score} = W[H_t || \text{Physics}_t] + b$ (where $||$ denotes vector concatenation).

3.3. Online Training Strategy

A core objective of the proposed framework is not only to improve prediction accuracy, but also to maintain adaptability under continuously incoming monitoring data. The training process is therefore designed as an online updating strategy rather than a one-time offline optimization procedure. This design allows the model to initialize rapidly, expand its effective learning window as more data become available, and subsequently maintain stable rolling prediction during long-term operation.

3.3.1. Three-Stage Online Learning Mechanism

To accommodate the practical constraints of landslide monitoring and the evolving availability of time-series data, a progressive three-stage online learning mechanism is adopted in the proposed framework, as illustrated in Figure 12. During the phase 1 stage, the model is initialized with the minimum amount of data, enabling prediction to start as early as possible after monitoring begins. As additional observations accumulate, the framework enters the phase 2 stage, in which the effective learning window expands progressively and the model develops a more complete representation of site-specific temporal characteristics. Once sufficient historical information has been accumulated, the training procedure transitions to the phase 3 stage, where a fixed rolling window is maintained for continuous updating and prediction. This staged design reflects the operational reality that emergency deployment and long-term monitoring impose different demands on the prediction system: The bootstrap stage prioritizes rapid initialization under limited data support, whereas the growth and sliding stages progressively strengthen robustness, temporal completeness, and adaptability during sustained operation, thereby providing a unified strategy for balancing model timeliness at deployment and prediction stability in long-term online learning.

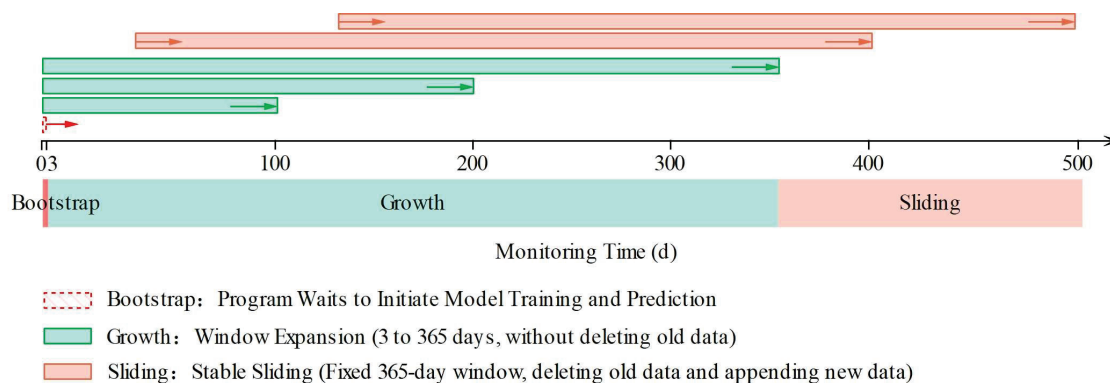


Figure 12. Three-Stage Model Training and Prediction Strategy.

3.3.2. Adaptive Physics-Constraint Weighting

The role of physical guidance is not uniform throughout the training process. When the available data are limited or the model is still in the early stage of convergence, stronger physical regularization is beneficial because it stabilizes optimization and suppresses implausible fitting to local noise. As the model gradually acquires richer site-specific information, excessive reliance on the physical prior may limit its ability to adapt to the actual deformation pattern. This observation motivates the use of an adaptive weighting strategy for the physics-constraint term.

In the present framework, the physics-constraint weight is updated according to an exponential decay rule:

$$\lambda_{\text{physics}}^{(t+1)} = \max\left(\lambda_{\min}, \text{decay_rate} \cdot \lambda_{\text{physics}}^{(t)}\right) \quad (17)$$

where $\lambda_{\text{physics}}^{(t)}$ denotes the physics-constraint weight at training step or epoch t , decay_rate controls the decay speed, and λ_{min} is the lower bound that prevents the physics regularization from vanishing completely. In the present implementation, the initial physics weight is set to 1.0, the decay rate is set to 0.99, and the lower bound is set to 0.1. This setting ensures stronger physical guidance in the early stage of training and a more balanced optimization between mechanism constraint and data fitting in the later stage.

The adaptive weighting strategy is consistent with the operational logic of the proposed three-stage online learning framework. During the phase 1 stage, a relatively high physics weight stabilizes model initialization under limited data support. During the phase 2 stage, the gradual reduction of λ_{physics} allows the model to incorporate newly accumulated site-specific observations more effectively. During the phase 3 stage, the lower-bounded physics constraint continues to provide physical regularization while preserving the flexibility required for rolling prediction. To improve interpretability, the evolution curve of λ_{physics} across the training process is presented in Figure 13, where the bootstrap, growth, and sliding stages are also indicated.

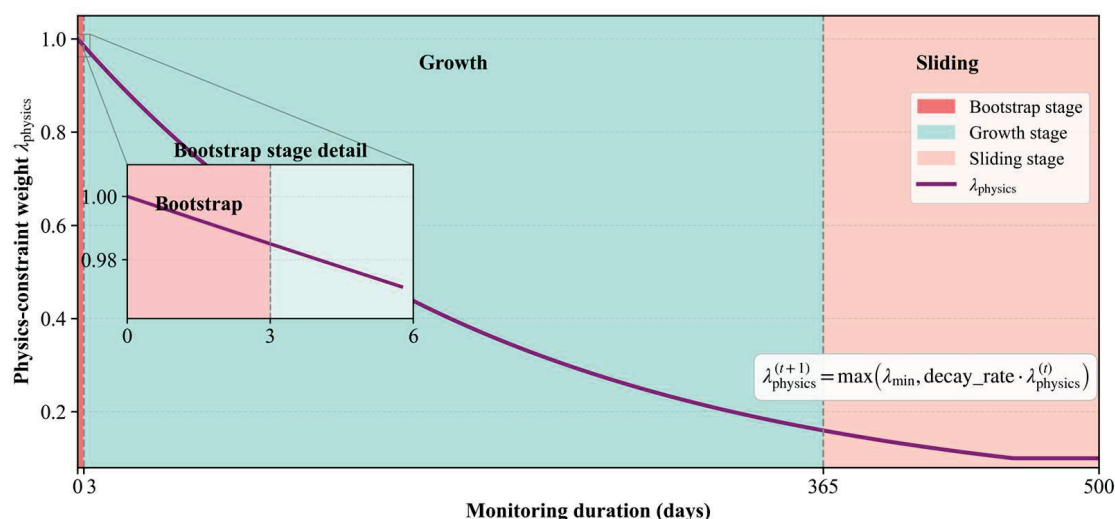


Figure 13. Evolution of the adaptive physics-constraint weight across the three-stage online learning process.

3.3.3. Rolling Updating and Prediction Workflow

Based on the above design, the complete prediction workflow of Phys-LSTM consists of three tightly connected processes, as illustrated in Figure 14: monitoring-data ingestion and preprocessing, physics-informed representation and temporal learning, and rolling updating with continuous prediction. Incoming observations are first transformed into structured feature sequences through quality control, spatiotemporal alignment, and physics-informed feature construction. These sequences are then processed by the Phys-LSTM architecture to generate multi-component displacement predictions while preserving physically meaningful temporal consistency. Once new monitoring data arrive, the rolling window is updated and the model parameters are incrementally adjusted according to the current training stage and the adaptive weighting strategy. In this way, the proposed framework operates as an online displacement-prediction model under realistic monitoring conditions, with its significance lying not only in predictive accuracy, but also in the integration of mechanism-oriented guidance, temporal learning, and online updating into a unified operational workflow.

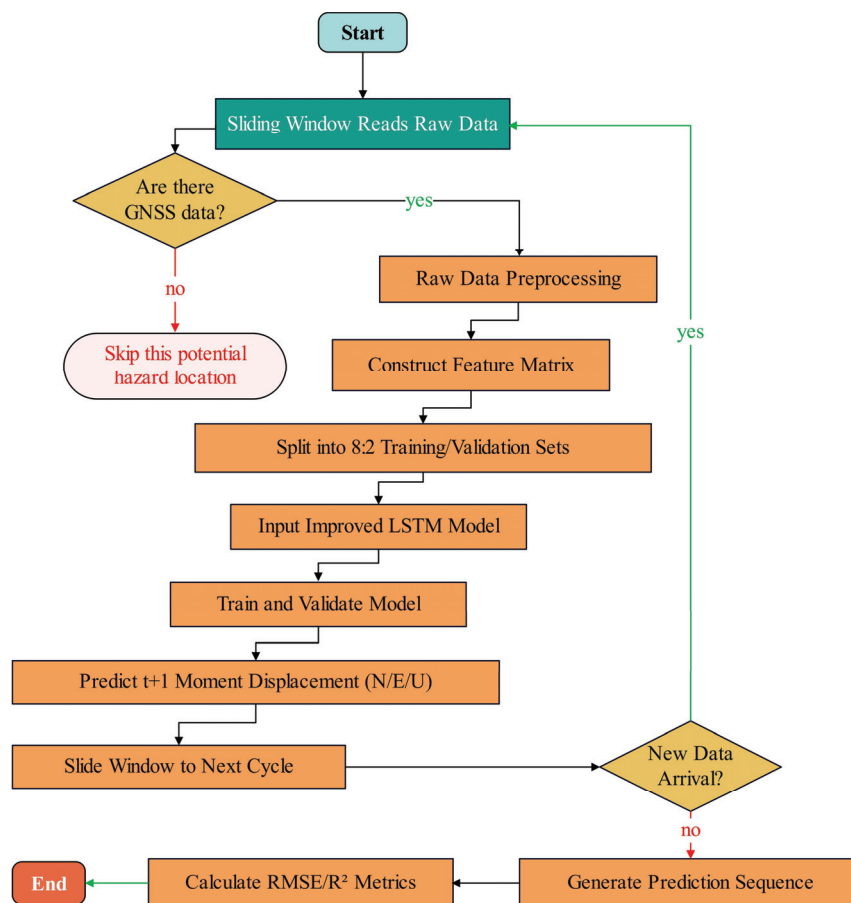


Figure 14. Online Learning Workflow of Phys-LSTM.

From an implementation perspective, the online framework consists of data ingestion, quality control, feature updating, incremental training, rolling prediction, and warning-interface output. Initial model training can be completed offline or in the background after a monitoring site has accumulated enough observations, whereas routine online operation only updates the rolling data window and performs lightweight incremental parameter adjustment. Because the forecasting horizon is five time steps and the input window is 30 time steps, single-site inference can be executed on a standard monitoring-platform server or edge node; the practical update frequency follows the incoming GNSS, rainfall, crack-gauge, and inclinometer sampling intervals.

4. Experiments and Results

To systematically evaluate the performance of the Phys-LSTM model, comprehensive experiments were conducted using real-world datasets from the Sichuan Province Geological Hazard Monitoring and Early Warning Platform. The experimental data includes core features such as complete GNSS displacement monitoring, rainfall monitoring, crack width monitoring, and water level gauge readings. The Phys-LSTM model architecture incorporates a masking layer, a physics-informed feature calculation module, a feature-level attention mechanism, an LSTM encoder–decoder (with 128 and 64 hidden units, respectively), and a physics-guided temporal attention mechanism. To enhance generalization capability, a Dropout rate of 0.3 and an L2 regularization coefficient of 0.001 were introduced. The model was trained using the Adam optimizer with an initial learning rate of 0.0005, employing a learning rate scheduling strategy that combines a 5-epoch warm-up phase followed by cosine decay. The batch size was set to 512, training epochs ranged from 100 to 300, and an early stopping mechanism was implemented to monitor

validation loss. The physics-informed feature module extracts a 7-dimensional feature set, including displacement velocity, acceleration, total displacement, dynamic factor of safety, creep increment, and rainfall risk index. The feature-level attention mechanism increases the weight of GNSS target features to approximately 22%, while the temporal attention mechanism assigns higher weights to recent data through exponential decay (decay rate of 0.95).

4.1. Experimental Models and Evaluation Metrics

To ensure both the breadth and rigor of the comparative analysis, five representative baseline models were selected, as listed in Table 4, spanning classical statistical approaches to state-of-the-art deep learning methods, and their predictive performance was evaluated using four widely adopted metrics, namely Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the Coefficient of Determination (R^2), so as to characterize model performance from multiple perspectives, including absolute error, relative deviation, and goodness of fit. To guarantee the comparability of the reported results, all experiments in the comparative study were conducted on the same test dataset for each landslide under a unified progressive three-stage framework for training, validation, and prediction. The reported results were derived from a single complete run for each model–landslide pair rather than from the average of repeated runs. For multi-component displacement prediction, each evaluation metric was first computed separately for the N, E, and U displacement components, after which the arithmetic mean of the three directional results was used as the integrated performance indicator reported in the manuscript. For instance, the comprehensive RMSE was calculated as the mean of RMSE_N, RMSE_E, and RMSE_U, such that the reported values represent the overall predictive capability of each model across the three-dimensional displacement field rather than the performance for any single component alone.

Table 4. Overview of Experimental Models.

Model Category	Model Name	Core Characteristics
Classical Statistical Model	SARIMA	Excels in capturing linear trends and seasonal patterns
Traditional Machine Learning	XGBoost	Gradient boosting decision-tree regression using flattened 30-step historical windows, squared-error objective, early stopping, and L2 regularization.
Deep Learning Sequence Model	LSTM	Standard Long Short-Term Memory network without physical constraints
	GRU	A lightweight variant of LSTM with higher computational efficiency
	Transformer	Modern sequence model based on the self-attention mechanism
Proposed Model in This Study	Phys-LSTM	Hybrid LSTM model enhanced by physical equations

4.2. Hyperparameter Setting and Tuning Strategy

To ensure a fair comparison, all models were evaluated under a unified validation protocol. The historical input window length and forecasting horizon were fixed at 30 and 5 time steps, respectively, and the samples were divided into training and validation sets with a ratio of 80%/20%. The architecture was not selected through an exhaustive global optimization algorithm; instead, it was determined using validation-guided empirical calibration and sensitivity comparison. The 128-64 hidden-unit configuration was retained because the 128-unit encoder provided sufficient capacity to represent nonlinear deformation memory and rainfall-lag response, while the 64-unit decoder reduced forecasting-head complexity and online computational burden. For XGBoost, the 30-step

historical window was flattened into supervised learning features and used to predict the N, E, and U displacement components under a squared-error regression objective; L2 regularization was explicitly represented by `reg_lambda`. For LSTM, GRU, Transformer, and Phys-LSTM, validation loss and early stopping were used to control overfitting, and adaptive learning-rate reduction was applied where appropriate. The general parameters and key model-specific settings are shown in Table 5.

Table 5. Key hyperparameter settings of the compared models. Common parameters define the unified input-output protocol. For XGBoost, `reg_lambda` denotes the L2 regularization coefficient. For Phys-LSTM, `temporal_decay_rate` denotes the temporal attention decay coefficient, `target_weight_ratio` denotes the feature-attention emphasis assigned to GNSS target features, and L2 denotes kernel regularization.

Model	Key Hyperparameters	Value
Common	Historical input window length/Forecasting horizon	30/5
Common	Training/validation split	80%/20%
SARIMA	Order selection	auto_arima
XGBoost	<code>n_estimators</code> , <code>max_depth</code> , <code>learning_rate</code> , <code>objective</code> , <code>reg_lambda</code>	800, 6, 0.03, squared error, 1.0
LSTM	Hidden units, dropout, learning rate, training control	128-64, 0.2, adaptive LR, early stopping
GRU	Hidden units, dropout, learning rate, training control	128-64, 0.2, adaptive LR, early stopping
Transformer	Heads, key dimension, dropout	4, 32, 0.2
Phys-LSTM	Encoder hidden units, decoder hidden units, dropout, L2, optimizer, learning rate, batch size, epochs, <code>temporal_decay_rate</code> , <code>target_weight_ratio</code>	128, 64, 0.3, 0.001, Adam, 5×10^{-4} with warm-up/cosine decay, 512, 100–300 with early stopping, 0.95, 4.0

4.3. Overall Performance Comparison

The comparative experiments were completed using the TensorFlow framework, employing the same dataset and a unified training/validation set split ratio (8:2) to ensure a fair comparison. The overall prediction performance of each model on the test set is shown in Figure 15. In this figure, both the z-axis and the bar-top labels denote relative performance scores obtained by min–max normalization of each evaluation metric across the compared models. For the error-based metrics RMSE, MAE, and MAPE, reverse min–max normalization was applied so that lower raw values correspond to higher normalized scores, whereas for R^2 , forward min–max normalization was adopted, such that higher raw values correspond to higher normalized scores. The Phys-LSTM model demonstrated optimal performance across all evaluation metrics.

The comparative results reveal the performance boundaries of different modeling paradigms in landslide displacement prediction. As shown in Figure 15, SARIMA and XGBoost are limited in representing nonlinear deformation memory when the monitoring sequence contains asynchronous sampling, missing observations, and abrupt local disturbances. The unconstrained Transformer baseline also showed unstable performance in the present online monitoring setting. Under the present site-specific sequence length, 30-step input window, 5-step forecasting horizon, and parameter budget, self-attention without physical priors was more easily affected by noisy or non-stationary observations. In addition, MAPE is highly sensitive during low-displacement periods, which can amplify local deviations. These findings support the need for physically constrained recurrent modeling in strongly mechanism-driven landslide forecasting tasks.

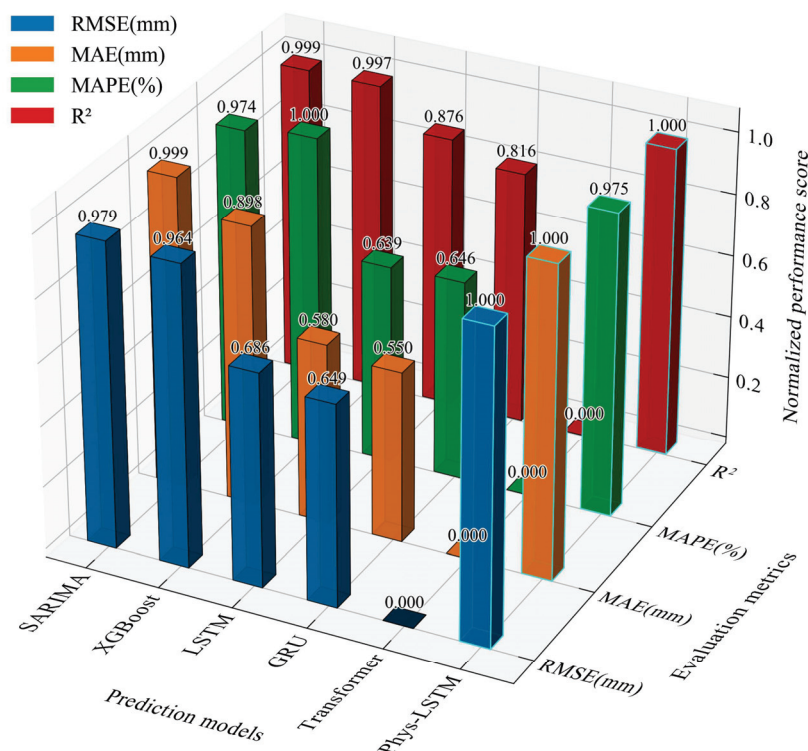


Figure 15. Comparative 3D Visualization of Multi-Metric Performance for Different Models.

In contrast, the Phys-LSTM model, derived from a deeply integrated “physics-informed and data-driven” architectural design, demonstrated comprehensive and stable superiority ($\text{RMSE} \leq 1.78 \text{ mm}$, $R^2 \geq 0.96$). Specifically, the hard-constraint loss function based on kinematic equations provides directional guidance for model optimization that conforms to fundamental mechanical laws, ensuring the physical plausibility of the predicted trajectory. Meanwhile, the physics-guided gating mechanism, which incorporates the stability index (SI), enables the model to adaptively filter and memorize critical temporal states, thereby achieving a unification of high accuracy and high robustness within complex systems.

To further evaluate the prediction stability of the models under different geological conditions, we calculated the standard deviation of the RMSE for the aforementioned models across six different landslide mass test sets. The six landslides used in this cross-site comparison cover diverse geomorphological and geological settings, including bedding-controlled rock landslides, accumulation-layer landslides, and cover-material landslides sliding along the bedrock interface. Their slope gradients, lithological compositions, structural conditions, and deformation styles differ markedly, thereby providing a representative basis for evaluating cross-site generalization stability. The monitoring configurations of these sites include combinations of GNSS displacement monitoring, crack gauges, rain gauges, and inclinometer-type sensors, and the available monitoring data comprise three-dimensional displacement, crack-width variation, rainfall records, and related auxiliary observations. Such heterogeneity makes the six-site comparison particularly suitable for testing whether the proposed model can maintain stable performance under varying geological conditions, monitoring configurations, and data characteristics. The geological conditions, deformation characteristics, and monitoring configurations of the six landslides used in the cross-site comparison are summarized in Table 6.

Table 6. Geological and Monitoring Background.

Landslide	Geological and Geomorphological Setting	Deformation/Structural Type	Monitoring Background
Shibanwo	Upper-steep and lower-gentle slope; Quaternary residual/colluvial deposits over mudstone and sandstone	Cover-material landslide developed above weak bedrock	3 GNSS sensors, 2 acceleration sensors, 2 tiltmeters and 1 rain gauge
Pengjiawan	Broken-line slope; loose cover deposits over weathered phyllite	Shallow accumulation-layer landslide sliding along a weak cover-bedrock contact	4 GNSS sensors, 1 deep-displacement sensor, 1 rain gauge and 1 soil-moisture sensor
Caiyuanzi	Steep bedding slope; shale, sandstone, and Quaternary colluvial gravelly soil	Bedding-controlled landslide with a clear slip surface along weak structural interfaces	3 GNSS sensors and 1 rain gauge
Guoluojiao	Gently descending slope; Quaternary loose deposits over bedrock	Cover-material landslide sliding along the cover-bedrock interface	1 crack gauge, 6 GNSS sensors, 1 acceleration sensor, 1 tiltmeter and 2 rain gauges
Houshan	Relatively straight but locally undulating slope; gravelly soil cover over phyllite bedrock	Large accumulation-layer landslide deforming along the cover-bedrock interface	4 crack gauges, 2 GNSS sensors, 2 acceleration sensors, 2 tiltmeters, 1 rain gauge
Dashuzi	Middle-mountain slope; Holocene gravelly soil over phyllite; stepped topography	Soil landslide with upper-steep and lower-gentle morphology, influenced by rainfall and human activities	1 crack gauge, 5 GNSS sensors and 1 rain gauge

The calculation results are presented in Table 7. It should be noted that the RMSE values summarized in Table 7 represent the arithmetic mean of the directional RMSE values for the N, E, and U displacement components at each site.

Table 7. Standard Deviation of RMSE for Various Models Across Different Landslide Masses.

Model	Shibanwo Landslide	Pengjiawan Landslide	Caiyuanzi Landslide	Guoluojiao Landslide	Houshan Landslide	Dashuzi Landslide	Standard Deviation
SARIMA	2.16	0.48	1.37	1.02	1.38	6.63	2.25
XGBoost	4.16	3.37	4.23	4.31	4.41	16.12	4.92
LSTM	0.76	29.02	0.72	0.90	1.23	12.27	11.49
GRU	0.78	32.23	0.57	0.64	1.22	13.53	12.84
Transformer	1.57	77.79	1.18	3.30	1.47	34.32	31.18
Phys-LSTM	2.13	0.49	1.36	1.01	1.38	4.29	1.34

Based on the experimental data in Table 7, it can be observed that Phys-LSTM maintained the optimal and most stable prediction performance across six landslide masses with diverse geological conditions and deformation characteristics (with the lowest RMSE standard deviation of only 1.34). This exceptional cross-site generalization stability is inherently attributed to the advantages embedded within its physically constrained architecture. The core physical equations (such as stress equilibrium and creep constitutive laws) describe universal mechanical principles governing landslide evolution. They provide the model with a unified optimization constraint and inductive bias that does not depend on the data distribution of any specific site. This fundamentally suppresses the model's tendency to overfit to incidental noise or local features within a single training dataset, enabling it to capture the common dynamic mechanisms underlying different landslides. In contrast, the performance of purely data-driven models (e.g., LSTM, Transformer) heavily relies on the similarity between the training and testing data distributions. The extreme error points and large performance standard deviations (up to 31.18) observed for these models on landslides with significant differences, such as Pengjiawan and Dashuzi, directly reflect their lack of physical prior knowledge and the fragility of their generalization foundation.

4.4. Generalizability and Operational Applicability

The proposed framework is most applicable to monitored landslides with progressive deformation behavior, where displacement, rainfall, deformation-rate, and auxiliary monitoring features can be continuously updated. The six-site comparison covers bedding-controlled rock landslides, accumulation-layer landslides, cover-material landslides, and soil landslides, indicating that the physics-informed constraints provide a transferable inductive bias across different geological and monitoring conditions.

Nevertheless, the framework is not intended to replace site-specific geotechnical assessment. Its transferability may be reduced for rapid failures with limited precursory displacement, earthquake-triggered failures, debris-flow processes, or sites where key hydrological or structural controls are not observed. In such cases, additional monitoring variables, recalibration, or more explicit physical modules should be introduced before operational deployment.

5. Application Case

The Huangcaoping deformation mass landslide is located on the left bank of the Dadu River in Luding County, Sichuan Province, China, near Xinhua Village, Detuo Town. It lies within the reservoir area of the Dagangshan Hydropower Station, approximately 16.8 km from the dam site. The deformation mass has a toe elevation of about 1115 m and a crown elevation of about 1600 m. It extends approximately 160–250 m along the river direction and is about 660 m wide, with a total volume of roughly 1.5 million m³. Geomorphologically, the slope has an average gradient of 30–45° and is widely covered by 5–20 m thick colluvial deposits of blocky gravelly soil. The deformation mass has developed within thin-layered sandstone and shale of the Triassic Baiguowan Formation, which is sandwiched within the Dadu River fault zone. The underlying bedrock consists of toppling deformation rock mass. This landslide deformation mass has been instrumented with monitoring equipment including rain gauges, crack gauges, surface displacement monitors, and fixed inclinometers. Since October 2015, surface cracking and deformation have been observed on the slope, as shown in Figure 16. Deformation intensifies annually during the rainy season. From July to November 2017, slope deformation accelerated, with cracking at the rear crown, widespread cracking across the slope body, and localized collapses. A relatively large-scale collapse, with a volume of approximately 50,000 m³, occurred on 2 July 2018. Subsequently, brief periods of intensified deformation were observed at the rear crown cracks in August 2019 and September 2020, followed by a return to a relatively stable state. Its failure mode is a toppling-tensile cracking landslide, formed on the basis of a toppling deformation rock mass under the combined influence of reservoir impoundment, rainfall, and agricultural water use.

To comprehensively validate the long-term performance and practical value of Phys-LSTM in real-world operational online learning scenarios, we conducted an in-depth case study using five consecutive years of monitoring data from the Huangcaoping deformation mass landslide. The monitoring data fully covers multiple critical stages of landslide evolution, including slow creep, uniform deformation, and accelerated deformation triggered by heavy rainfall. This provides an ideal testing environment for evaluating the model's adaptive evolution capability, continuous learning ability regarding physical mechanisms, and responsiveness to sudden external disturbances. Through continuous online learning, Phys-LSTM can not only accurately fit long-term trends but also make timely and reliable predictions at key physical inflection points. The specific results are shown in Figure 17.

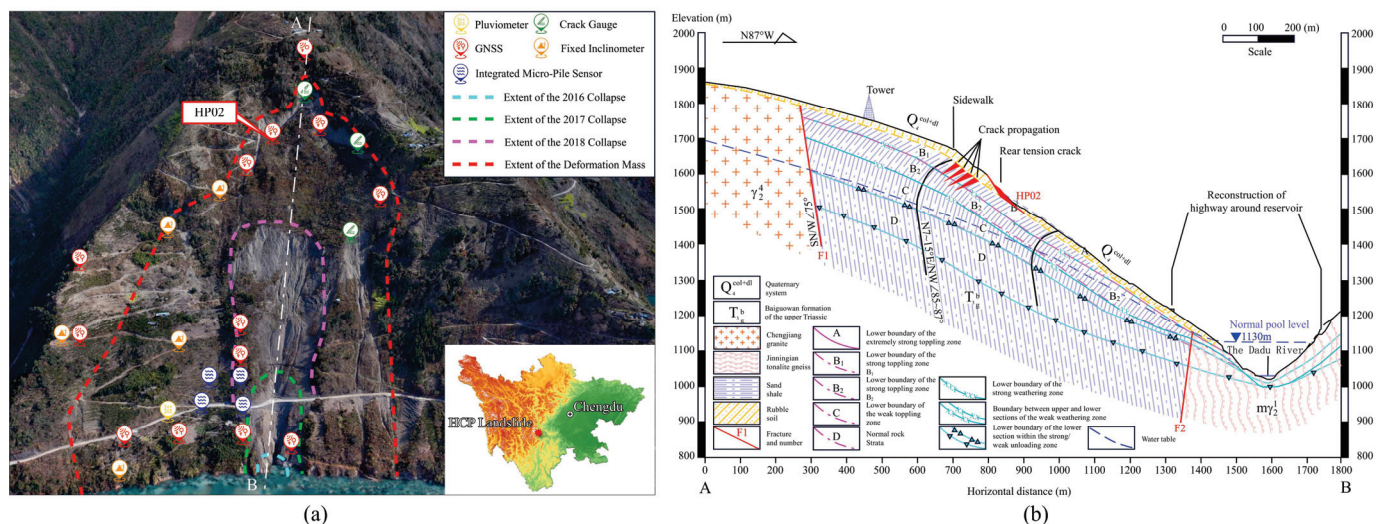


Figure 16. Huangcaoping deformation mass landslide. (a) Landslide zoning and monitoring point arrangement, with a lower-right DEM inset showing the regional position of the HCP landslide and Chengdu in Sichuan Province; (b) A–B geological profile.

During the initial phase of intensified deformation from July to November 2017, the model was in the extended learning stage. A comparison between the actual and predicted curves (Figure 17a,b) shows that the model could preliminarily identify the overall upward trend of displacement caused by rear crown crack settlement, widespread cracking, and localized collapses. However, its fitting of transient fluctuations and detailed deformation remained insufficient. During this stage, the RMSE, MAE, R^2 , and MAPE were 26.44 mm, 11.23 mm, 0.87, and 48.9% respectively, reflecting the model's limited accuracy in depicting complex deformation patterns during early learning.

By July 2018, a large-scale collapse occurred at the front, and the cumulative deformation at the rear reached 10.3 m within four months. At this point, the model had entered a stable sliding phase, and its prediction results showed good agreement with the measured data in terms of trend Figure 17c. For this period, the RMSE was 11.52 mm, MAE was 10.21 mm, R^2 was 0.92, and MAPE was 38.39%, indicating that the model had developed an effective capability to capture the overall displacement characteristics of intense deformation events.

In August 2019, deformation at the rear intensified again, reaching 98 mm within the month, after which the deformation rate gradually stabilized. Having been trained on two years of historical data, the model was in a stable prediction period. Although the actual sequence showed significant local fluctuations, the predicted curve consistently reflected the overall deformation trend Figure 17d. During this stage, the RMSE, MAE, R^2 , and MAPE were 9.64 mm, 7.50 mm, 0.94, and 4.58% respectively, confirming the model's high predictive reliability during periods of stable deformation rates.

As shown in Figure 17e, following a heavy rainfall event in September 2020, the landslide displacement rate increased sharply (peaking at 106.03 mm/d). At this physical inflection point, the prediction curve from Phys-LSTM and the actual observations exhibited highly synchronized surge trends and evolutionary trajectories (R^2 reached 0.97 for this stage). This demonstrates that the “rainfall-infiltration-softening of geotechnical material strength-accelerated deformation” coupling mechanism learned through the physics-constrained framework is effective and generalizable. It enables the model to go beyond simple fitting of historical data and genuinely achieve physical perception and reliable capture of key warning signals.

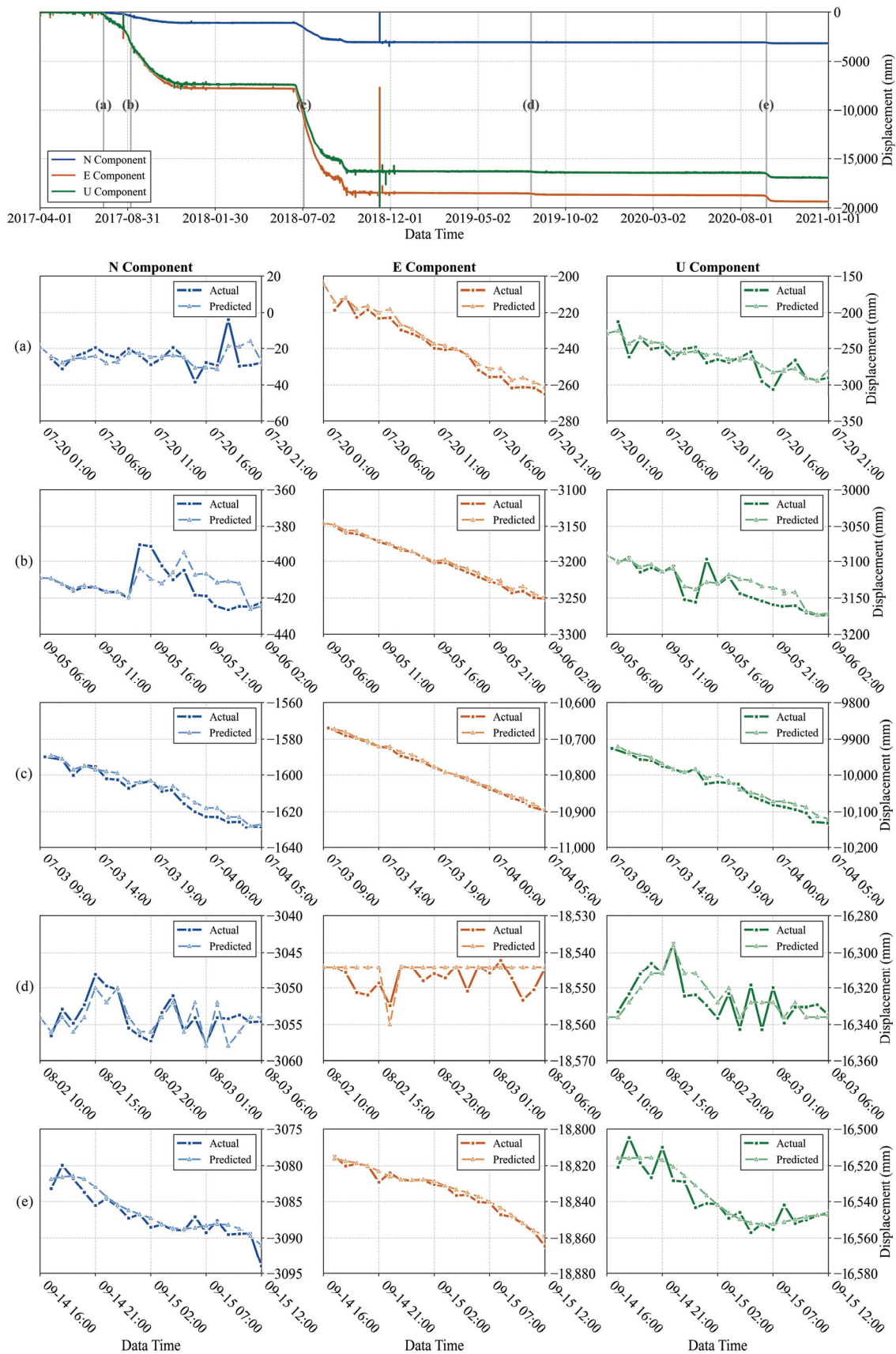


Figure 17. Comparison Chart of HP02 Landslide Displacement Prediction. (a) The July 2017 deformation event, (b) The September 2017 Deformation Event, (c) The July 2018 deformation event, (d) The deformation event in August 2019, and (e) The September 2020 Deformation Event.

6. Conclusions

This study developed a physics-informed online learning framework for landslide displacement prediction by integrating multi-source monitoring data, geomechanical prior knowledge, and progressive online updating. The framework addresses two limitations of conventional static data-driven models: weak physical consistency and insufficient adaptability to newly arriving monitoring observations.

The proposed Phys-LSTM achieved strong prediction accuracy and stability in comparative experiments, with overall RMSE ≤ 1.78 mm and $R^2 \geq 0.96$. In the six-site cross-condition comparison, Phys-LSTM produced the lowest RMSE standard deviation (1.34), indicating better robustness under different geological backgrounds, deformation modes, and monitoring configurations.

The main scientific contribution lies in embedding a dynamic stability index, creep-informed deformation representation, and kinematic residual loss into an LSTM encoder-decoder architecture. Together with the three-stage online learning strategy, these components allow the model to update continuously while maintaining physically plausible displacement trajectories.

From an engineering perspective, the framework provides a practical basis for operational landslide monitoring and short-term displacement forecasting. Future work should further connect the prediction output with explicit warning thresholds, incorporate hydrological and InSAR observations where available, and validate the method under a wider range of rapidly evolving or structurally controlled landslide scenarios.

Author Contributions: Conceptualization, J.Z., N.J. and C.H.; methodology, J.Z., N.J. and C.H.; software, J.Z., N.J. and C.H.; validation, J.Z., C.H. and M.X.; investigation, M.X.; resources, J.Z., C.H. and M.X.; data curation, J.Z.; writing—original draft preparation, J.Z.; writing—review and editing, J.Z., N.J., C.H. and M.X.; visualization, J.Z. and M.X.; supervision, N.J. and C.H.; project administration, N.J.; funding acquisition, N.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Scientific Research Project of Sichuan Provincial Department of Natural Resources, grant number KJ-2025-069, and the Contract for the Procurement of Automated Monitoring Instruments and Equipment and Related Labor Services for the Xiaojiawan Landslide during the Construction of the Hubei Section of the Newly Built Xi'an-Shiyan High-Speed Railway, contract number 80303-AH20240021.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Petley, D. Global Patterns of Loss of Life from Landslides. *Geology* **2012**, *40*, 927–930. [CrossRef]
2. Dai, F.C.; Lee, C.F.; Ngai, Y.Y. Landslide Risk Assessment and Management: An Overview. *Eng. Geol.* **2002**, *64*, 65–87. [CrossRef]
3. Zhou, C.; Yin, K.; Cao, Y.; Ahmed, B. Application of Time Series Analysis and PSO-SVM Model in Predicting the Bazimen Landslide in the Three Gorges Reservoir, China. *Eng. Geol.* **2016**, *204*, 108–120. [CrossRef]
4. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
5. Zhong, Z.; Hu, B.; Li, J.; Sheng, J.; Wan, C. Impact of Rainfall Dry-Wet Cycles on Slope Deformation and Landslide Prediction in Open-Pit Mines: A Case Study of Mohuandang Landslide, Emeishan, China. *Results Eng.* **2025**, *26*, 105011. [CrossRef]
6. Zhang, J.; Tang, H.; Li, C.; Gong, W.; Zhou, B.; Zhang, Y. Deformation Stage Division and Early Warning of Landslides Based on the Statistical Characteristics of Landslide Kinematic Features. *Landslides* **2024**, *21*, 717–735. [CrossRef]

7. Tumon, M.S.H.; Imran, H.M.; Chow, M.F.; Zahidi, I.; Ahmed, F.; Islam, M.S.; Islam, M.d.R.; Hafez, M. Integrating a Hybrid Machine Learning Model with Climate Projections for Landslide Susceptibility Assessment. *Results Eng.* **2025**, *28*, 108067. [CrossRef]
8. Deng, Z.; Wang, W.; Luo, Y.; Zhang, S.; Xu, L.; Su, Q. Slope deformation monitoring and prediction based on InSAR and deep learning model. *Results Eng.* **2025**, *28*, 107113. [CrossRef]
9. Froude, M.J.; Petley, D.N. Global Fatal Landslide Occurrence from 2004 to 2016. *Nat. Hazards Earth Syst. Sci.* **2018**, *18*, 2161–2181. [CrossRef]
10. Han, N.; Miao, W.; Li, M.; Mohamad Ismail, M.A.; Hu, Q.; Duan, L.; Tang, J. Integrating Multi-Source Monitoring Data and Deep Convolutional Autoencoder Technology for Slope Failure Pattern Recognition. *Front. Earth Sci.* **2025**, *13*, 1531857. [CrossRef]
11. Cinar, Y.G.; Mirisaee, H.; Goswami, P.; Gaussier, E.; Ait-Bachir, A. Period-Aware Content Attention RNNs for Time Series Forecasting with Missing Values. *Neurocomputing* **2018**, *312*, 177–186. [CrossRef]
12. Dong, D.; Wang, M.; Chen, W.; Zeng, Z.; Song, L.; Zhang, Q.; Cai, M.; Cheng, Y.; Lv, J. Mitigation of Multipath Effect in GNSS Short Baseline Positioning by the Multipath Hemispherical Map. *J. Geod.* **2016**, *90*, 255–262. [CrossRef]
13. Han, J.Q.; Huang, G.W.; Li, Z. Multipath Effect Analysis and Processing Method of GNSS Landslide Monitoring Under Complicated Environment. *J. Earth Sci. Environ.* **2018**, *40*, 355–362.
14. Le, T.; Vu, H.C.; Ponchet-Durupt, A.; Boudaoud, N.; Cherfi-Boulanger, Z.; Nguyen-Trang, T. Unsupervised Detecting Anomalies in Multivariate Time Series by Robust Convolutional LSTM Encoder–Decoder (RCLED). *Neurocomputing* **2024**, *592*, 127791. [CrossRef]
15. Šegina, E.; Peternel, T.; Urbančič, T.; Realini, E.; Zupan, M.; Jež, J.; Caldera, S.; Gatti, A.; Tagliaferro, G.; Consoli, A. Monitoring Surface Displacement of a Deep-Seated Landslide by a Low-Cost and near Real-Time GNSS System. *Remote Sens.* **2020**, *12*, 3375. [CrossRef]
16. Amani, A.-M.; Sandra, K.; Frank, K. Standardization with Zlog Values Improves Exploratory Data Analysis and Machine Learning for Laboratory Data. *J. Lab. Med.* **2024**, *48*, 215–222.
17. Yaro, A.S.; Maly, F.; Prazak, P.; Malý, K. Outlier Detection Performance of a Modified Z-Score Method in Time-Series RSS Observation with Hybrid Scale Estimators. *IEEE Access* **2024**, *12*, 12785–12796. [CrossRef]
18. Ilyushin, B.B. On Applicability of IQR Method for Filtering of Experimental Data. *J. Eng. Thermophys.* **2024**, *33*, 1–8. [CrossRef]
19. Vande Kamp, G.N.; Jordan, S. The Necessity of Moving Averages in Dynamic Linear Regression Models. *Am. J. Polit. Sci.* **2025**, *69*, 176–193. [CrossRef]
20. Cai, L.; Xu, G.; Niyato, D. ASG: An Adaptive Savitzky-Golay Method for Channel Estimation in Deep-Space Communications. *IEEE Trans. Aerosp. Electron. Syst.* **2025**, *61*, 19525–19536. [CrossRef]
21. Sanz-Alonso, D.; Yang, R. Gaussian Process Regression under Computational and Epistemic Misspecification. *SIAM J. Numer. Anal.* **2025**, *63*, 495–519. [CrossRef]
22. Gałka, Ł. Optimized Deep Isolation Forest. *Pattern Recognit. Lett.* **2025**, *197*, 88–94. [CrossRef]
23. Yang, H.; Wang, L.; Hao, S. Landslide Monitoring and Stabilization Process Based on Inclination Evolution. *Eng. Mech.* **2020**, *37*, 193–199.
24. He, C. Research on Key Technology and Application of Real-Time Monitoring and Early Warning System for Landslides. Ph.D. Thesis, Chengdu University of Technology, Chengdu, China, 2020.
25. Azami, H.; Mohammadi, K.; Bozorgtabar, B. An Improved Signal Segmentation Using Moving Average and Savitzky-Golay Filter. *J. Signal Inf. Process.* **2012**, *3*, 39–44. [CrossRef]
26. Liu, H.; Sun, X.; Bai, S. Modified Unscented Kalman Filter Considering Maximum Point of Probability Density Function. *IEEE Trans. Aerosp. Electron. Syst.* **2025**, *61*, 4926–4944. [CrossRef]
27. Chai, B.; Chan, S.C.; Hou, Y.H. Robust Adaptive Fading Unscented Kalman Filter for Decentralized Dynamic State Estimation in Power Systems under Unknown Inputs and Covariance Mismatches. *IEEE Trans. Power Syst.* **2025**, *40*, 4732–4745. [CrossRef]
28. Roopa, C.K.; Harish, B.S. An Empirical Evaluation of Savitzky-Golay (SG) Filter for Denoising ST Segment. In *International Conference on Cognitive Computing and Information Processing*; Springer: Singapore, 2018; Volume 801, pp. 18–28.
29. Hsu, C.-H.; Wu, Y.-N.; Hsu, C.-H.; Wu, Y.-N. Application of Empirical Mode Decomposition for Decoding Perception of Faces Using Magnetoencephalography. *Sensors* **2021**, *21*, 6235. [CrossRef] [PubMed]
30. Xing, Y.; Yue, J.; Chen, C.; Cong, K.; Zhu, S.; Bian, Y. Dynamic Displacement Forecasting of Dashuitian Landslide in China Using Variational Mode Decomposition and Stack Long Short-Term Memory Network. *Appl. Sci.* **2019**, *9*, 2951. [CrossRef]
31. Li, X.; Li, S. Large-Scale Landslide Displacement Rate Prediction Based on Multi-Factor Support Vector Regression Machine. *Appl. Sci.* **2021**, *11*, 1381. [CrossRef]
32. Xu, J.; Jiang, Y.; Yang, C. Landslide Displacement Prediction during the Sliding Process Using XGBoost, SVR and RNNs. *Appl. Sci.* **2022**, *12*, 6056. [CrossRef]
33. Sun, Y.; Liu, T.; Zhang, C.; Xi, N.; Wang, H. Spatiotemporal Prediction of Landslide Displacement Using Graph Convolutional Network-Based Models: A Case Study of the Tangjiao 1# Landslide in Chongqing, China. *Appl. Sci.* **2024**, *14*, 9288. [CrossRef]

34. Mufundirwa, A.; Fujii, Y.; Kodama, J. A New Practical Method for Prediction of Geomechanical Failure-Time. *Int. J. Rock Mech. Min. Sci.* **2010**, *47*, 1079–1090. [CrossRef]
35. Jiang, P.; Chen, J. Displacement Prediction of Landslide Based on Generalized Regression Neural Networks with K-Fold Cross-Validation. *Neurocomputing* **2016**, *198*, 40–47. [CrossRef]
36. Intrieri, E.; Carlà, T.; Gigli, G. Forecasting the Time of Failure of Landslides at Slope-Scale: A Literature Review. *Earth-Sci. Rev.* **2019**, *193*, 333–349. [CrossRef]
37. Yang, Z.; Shao, W.; Qiao, J.; Huang, D.; Tian, H.; Lei, X.; Uchimura, T. A Multi-Source Early Warning System of MEMS Based Wireless Monitoring for Rainfall-Induced Landslides. *Appl. Sci.* **2017**, *7*, 1234. [CrossRef]
38. Xiao, T.; Zhang, L.-M. Data-Driven Landslide Forecasting: Methods, Data Completeness, and Real-Time Warning. *Eng. Geol.* **2023**, *317*, 107068. [CrossRef]
39. Bai, D.; Lu, G.; Zhu, Z.; Zhu, X.; Tao, C.; Fang, J. A Hybrid Early Warning Method for the Landslide Acceleration Process Based on Automated Monitoring Data. *Appl. Sci.* **2022**, *12*, 6478. [CrossRef]
40. Park, J.-Y.; Lee, S.-R.; Kim, Y.-T.; Kang, S.; Lee, D.-H. A Regional-Scale Landslide Early Warning System Based on the Sequential Evaluation Method: Development and Performance Analysis. *Appl. Sci.* **2020**, *10*, 5788. [CrossRef]
41. Shihui, P.; Xiaoming, S.; Shanyong, W.; Liang, G.; Manchao, H.; Zhigang, T. A Multi-Parameter Monitoring System and Early Warning Model for Rainfall-Induced Landslides. *Bull. Eng. Geol. Environ.* **2024**, *83*, 312.
42. Duncan, J.M. Fellow State of the Art: Limit Equilibrium and Finite-Element Analysis of Slopes. *J. Geotech. Eng.* **1996**, *122*, 577–596. [CrossRef]
43. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations(Article). *J. Comput. Phys.* **2019**, *378*, 686–707. [CrossRef]
44. Mao, Z.; Jagtap, A.D.; Karniadakis, G.E. Physics-Informed Neural Networks for High-Speed Flows. *Comput. Methods Appl. Mech. Eng.* **2020**, *360*, 112789. [CrossRef]
45. Zhou, C.; Yin, K.; Cao, Y.; Ahmed, B.; Fu, X. A Novel Method for Landslide Displacement Prediction by Integrating Advanced Computational Intelligence Algorithms. *Sci. Rep.* **2018**, *8*, 7287. [CrossRef] [PubMed]
46. Antonelo, E.A.; Camponogara, E.; Seman, L.O.; Jordanou, J.P.; de Souza, E.R.; Hübner, J.F. Physics-Informed Neural Nets for Control of Dynamical Systems. *Neurocomputing* **2024**, *579*, 127419. [CrossRef]
47. Wang, B.; Li, T.; Yan, Z.; Zhang, G.; Lu, J. DeepPIPE: A Distribution-Free Uncertainty Quantification Approach for Time Series Forecasting. *Neurocomputing* **2020**, *397*, 11–19. [CrossRef]
48. Mochiutti, E.; Antonelo, E.A.; Camponogara, E. Physics-Informed Echo State Networks for Modeling Controllable Dynamical Systems. *Neurocomputing* **2025**, *639*, 130251. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Seismic Hazard Area Prediction Based on Slip and Locking Rates of Active Faults in Northern Ningxia, China

Nan Zhang ^{1,2,3,*}, Xiwei Fan ^{1,2,4}, Jing Wang ^{1,2,4} and Gaozhong Nie ^{1,2,4}

¹ Key Laboratory of Seismic and Volcanic Hazards, China Earthquake Administration, Beijing 100029, China; fanxiwei@ies.ac.cn (X.F.); jingwang@ies.ac.cn (J.W.); niegz@ies.ac.cn (G.N.)

² Institute of Geology, China Earthquake Administration, Beijing 100029, China

³ Ningxia Seismological Bureau, Yinchuan 750001, China

⁴ Xinjiang Pamir Intracontinental Subduction National Observation and Research Station, Beijing 100029, China

* Correspondence: zhangnan@ies.ac.cn

Abstract

This research employs the ITRF2014 framework to conduct an analysis of Eurasian GNSS data, thereby acquiring the recent velocity and strain rate fields in the Ningxia–Inner Mongolia border region, as well as in the Yinchuan and Lingwu regions of western China. It also provides in-depth insights into the relative movements, slip rates, and locking positions of active faults in these areas. Based on the elastic dislocation theory, a constraint model is constructed to perform numerical simulations of active faults. This enables the inversion of strike–slip, dip–slip, and slip deficit rates, along with the locking degree. Fault lockings are identified at the southern extremities of the northern and southern segments of the Niushoushan–Luoshan fault, the western ends of the Xiangshan–Tianjingshan fault zone, and the eastern segment of the Gancanling Fault. In Yinchuan, local locking is detected below 15 km along the southern part of the Helan Mountain eastern fault. The Bayannula Mountain Fault exhibits an uneven locking distribution, while the locking degree of the western margin fault of Zhuozi Mountain increases from the southwest to the northeast, with the fault being locked in the Etuoque Banner section.

Keywords: Chinese North–South seismic belt; seismic hazard; global navigation satellite system (GNSS); slip velocity; fault locking rate

1. Introduction

The velocity and strain rate variations of two plates in active faults are crucial for seismic hazard assessment [1–4]. By investigating their relationship with earthquake generation, constructing a multi-source data fusion monitoring network, and enhancing long-term continuous monitoring and dynamic analysis, the accuracy of seismic hazard assessment can be improved to mitigate earthquake losses.

The slip rate of active faults is a key indicator for assessing long-term fault zone activity [1,2,5–7]. Fault zones with stable and high slip rates have rapid strain energy release and higher seismic risks. Fault plate velocities are not static; their variations may signal changes in fault activity. Sudden slip rate changes could indicate abnormal stress processes. For example, long-term monitoring shows significant fault slip rate fluctuations before earthquakes, providing early warning for seismic risk assessment. Continuous monitoring and analysis of velocity changes can adjust seismic risk evaluations in time and improve prediction accuracy.

According to the elastic rebound theory, when crustal elastic strain energy exceeds a threshold, it may trigger fractures and earthquakes. Strain rate changes reflect fault zone locking. Abnormal strain rate accumulation in an area may indicate severe locking, high-stress concentration, and increased seismic risk. Earthquakes with magnitudes greater than 6.0 typically occur in the locked sections of block boundary faults, particularly at the junctions of strongly and weakly locked segments, compressive and tensile locked segments, and fault intersections with locking–positioning discrepancies. The locking ratio is defined as the proportion of the length of the locked segment of a fault to the total length of the fault, which reflects the frictional strength and energy accumulation capacity of the fault medium [1,2,8,9]. It correlates with seismic voids. When it exceeds a critical threshold, the fault enters a “high-stress accumulation–low-creep release” state, causing seismic voids [10]. In 2022, the Ninglang M6.0 earthquake occurred with the epicenter in the predicted high-locking area (error 3.2 km) [9]. An 80km rupture gap formed between the 1976 Yanyuan Ninglang M6.7 earthquake swarm and the 1996 Lijiang M7.0 earthquake, with a sustained locking rate in the area [8]. From 2009 to 2015, the locking rate and sliding loss rate of the middle section of the Lijiang Xiaojinhe fault zone increased. In 2015, the probability of an $M \geq 6.5$ earthquake within 5 years was estimated at 68%. The strain accumulated in the elastic deformation phase is a critical parameter in earthquake prediction [1–3,11,12].

It is crucial to acquire real-time information regarding the movement, energy accumulation, and locking states of major active faults in seismic-risk areas from the shallow to the deep crust. Deformation data from fixed-point, continuous, and mobile surfaces under inherited conditions are of great significance for evaluating stress accumulation and active locking near active faults. The temporal and spatial variations of the crustal strain field calculated by the Global Navigation Satellite System (GNSS) exhibit a strong correlation with strong earthquakes. GNSS observations can accurately record changes in the surface displacement field and obtain crustal deformation information over large areas. Therefore, it can monitor surface deformation responses to stress-field changes within the Earth, as well as crustal deformation, local deformation, and block movement during tectonic stress processes. By using large earthquakes as reference points, the relationship between ground motion and the seismogenic process can be investigated through the changes in the regional GNSS baseline length, velocity field, and strain field before and after the earthquake. The DEFNODE inversion program can quantify the relationship between the locking rate and the sliding loss rate.

The velocity field of the Chinese mainland within the ITRF framework, relative to the Eurasian plate, was established, which reflects the relative movement within the Chinese mainland under a non-net rotation benchmark [13–16]. The horizontal velocity field in the vicinity of the fault zone was projected onto cross-sections from various directions to investigate the differences in block movement on both sides of the fault. The regional strain-rate field was established based on the displacement and velocity of GNSS stations [17,18]. Research can concentrate on aspects such as the GNSS vertical deformation-rate gradient, strain-rate field [3,4], inversion of fault–plane slip amount using coseismic displacement [19], multi–factor integrated seismic risk assessment, multi-constrained inversion of the [20] seismic source spatio-temporal rupture process [21], wavelet transformation of the GNSS strain-rate field at different spatial scales [22], or inversion of (GNSS velocities at the intra-plate fault spatial scale to differentiate regional deformation from long-term fault creep [23]. Over the past few decades, scholars have employed stochastic models beyond the Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO) to formulate stochastic optimization algorithms [24–27]. These algorithms, characterized by rapid convergence and high stability, are extensively utilized in deep-crustal inversion

and serve to address the inversion problem of fault three-dimensional slip rates [6,28–30]. However, the main steps of these bionic algorithms involve random elements, and there is significant uncertainty in the iterative process. When dealing with high-dimensional or complex problems, they tend to converge prematurely and sub-optimally, unable to guarantee the optimal solution.

Regarding the combination of deformation methods, some scholars integrate Global Navigation Satellite System (GNSS) and leveling data [28], whereas others employ one or two stochastic algorithms to infer seismic fault slip parameters from the gravity field and GNSS velocity field [6]. Some researchers, based on GNSS and leveling data, combine stochastic algorithms and dislocation theory models [29] to invert for parameters such as fault slip rates. Certain academics extend the original network inversion filter to incorporate Interferometric Synthetic Aperture Radar (InSAR) observations, utilizing the Network Inversion Filter framework integrating GNSS and InSAR to model time-dependent slip with high spatial resolution [27]. There are also studies on the methodological improvement of combination methods, such as using GNSS deformation data, two-dimensional deformation data, and evaluating fault slip inversion accuracy with one-dimensional deformation [7]. For instance, the improved Smith 3D model is utilized to invert the parameters of the Haiyuan fault zone [31], and an improved screw dislocation model considering the dip angle is employed to invert the sliding rate, locking depth, and dip angle of the Haiyuan fault system [32]. Another example is the improved construction of a horizontal layered medium seismogenic structure model, constrained by displacement/deformation observation data from GNSS, leveling, and strong earthquakes, as well as leveraging the high time resolution of InSAR data to conduct coseismic slip distribution simulations after earthquakes such as the 2013 Lushan M7.0 earthquake [33] and the 2021 Yangbi M6.4 earthquake [34].

Within the realm of geodetic inversion, the classical dislocation model theory put forward by [35,36] [the relevant person] has been extensively applied and optimized on a global [37–40] scale. A significant amount of early research, particularly from Japan, concentrated on constructing coseismic rectangular fault models through the inversion of crustal deformation resulting from tectonic stress release, utilizing multiple geodetic data sources such as GNSS, leveling or InSAR. This process entails estimating the uniform slip models of rectangular dislocations within an elastic half-space and the slip distributions on fault planes [37,39,40]. Furthermore, coseismic and post-seismic slip distributions have been inverted by leveraging the high time-resolution characteristics of InSAR data in conjunction with GNSS and InSAR data [38]. The classical dislocation model has been repeatedly employed for risk discrimination in seismic fracture zones within China's North–South seismic belt [7,32–34,41–44]. These applications encompass the analysis of the West Kunlun mountain Pass earthquake [44] and the Haiyuan Liupanshan fault zone [32,41,43], as well as the inversion of fault parameters for the Longmenshan fault zone [42], to examine the three-dimensional movement and deformation characteristics of the fault. Through the utilization of hazard discrimination of the primary active faults in northern Ningxia, in conjunction with the velocity field and strain rate, and by integrating fault morphology data obtained from geological surveys, this paper adopts an improved fault dislocation model [35,36] that takes into account the movement of secondary blocks in the region. This method facilitates the determination of the slip rate and slip deficit, as well as the locking rate (the coupling fractions, the ratio of locked slip to total slip) and locking depth of the main active faults.

Section 2 introduces the seismological background of northern Ningxia and its adjacent area, and it presents findings from using GNSS and other surface deformation data to infer inter-seismic activity of active faults. Section 3 elaborates on GNSS data sources, main

research steps, and processing methods for evaluating active fault risk in northern Ningxia. Section 4 discusses the acquired strain rate field, local strain variations, slip rate, slip deficit rate of primary active faults, as well as locking depth and rate. Section 5 compares this study's results with those from previous geological methods and analyzes the reasons for large discrepancies in some faults. Conclusions are presented in Section 6.

2. Geological Structure Background

Figure 1a shows the tectonic framework of the northeastern margin of the Qinghai–Tibet Plateau, the western margin of the Ordos Block, and the northern segment of the China North–South Seismic Zone within the study area. The figure illustrates several major primary geological blocks surrounding the study area, including the Ordos Block in the east, the Bayan Har Block in the west, the South China Block in the south, and the Longzhong Basin tectonic region surrounded by several primary geological blocks. Ningxia is generally located on the western margin of the Ordos Block.

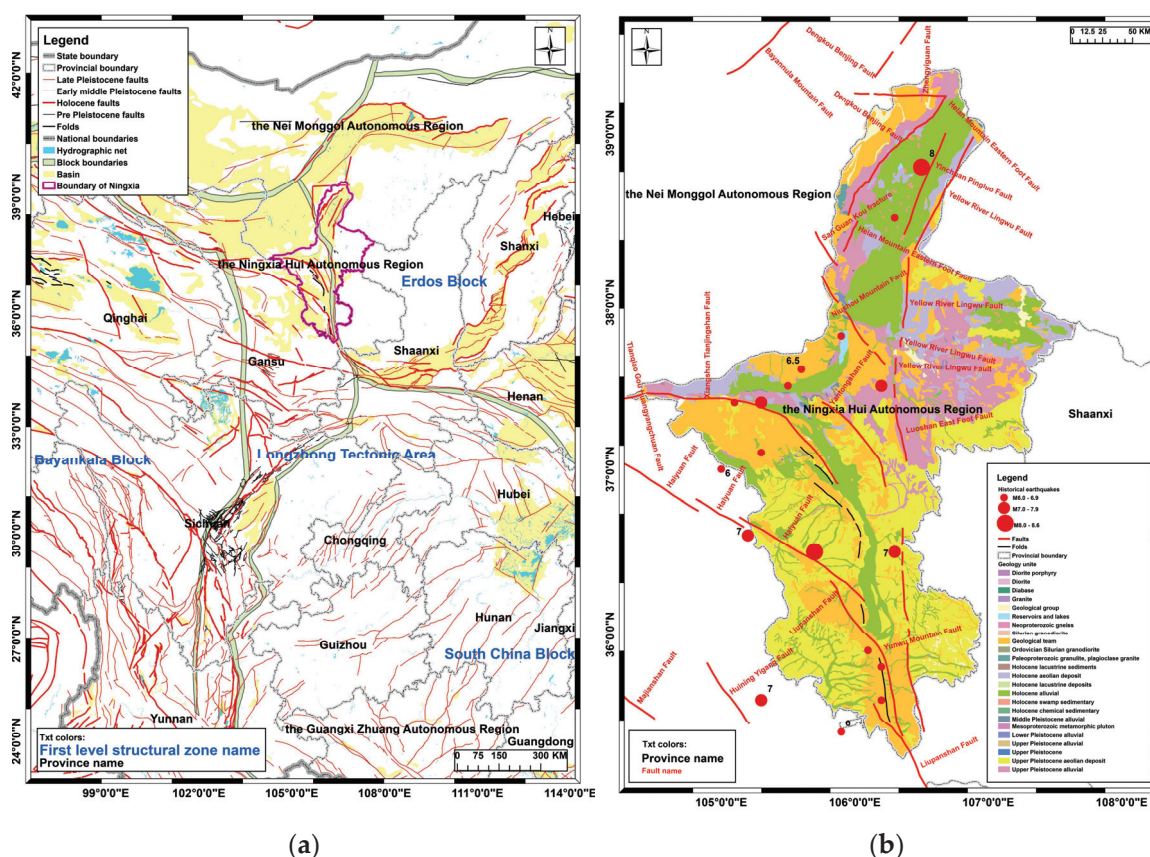


Figure 1. (a) The active tectonic structures surrounding the China North–South Seismic Belt, specifically those in the Ningxia Hui Autonomous Region situated at the western margin of the Ordos Block in the northern segment of the seismic belt, are delineated by dark purple lines as boundaries. (b) Geological Configuration and Historical Seismicity in Ningxia.

The structural belt along the western margin of the Ordos Block represents a region characterized by abrupt alterations in crustal thickness and geophysical fields within the Chinese mainland [45]. This western margin of the Ordos Block extends approximately 600 km in length, with a maximum width of about 200 km [45]. Over an extended period, scholars have conducted research on the structural characteristics of the western margin of the Ordos Block from multiple perspectives, including the Longxi System [46], Helan–Liupanshan Terrace Fold [47], the northern section of the North–South Seismic Belt [48,49], the Early Paleozoic Plate Suture [50], the Yinchuan Graben [51], the Mesoproterozoic

Rift [52], the Ordos Platform Fold [53], and the adjustment resulting from the collision between the North China and South China Plates [54], among others. The thrusting and subduction activities within the western margin thrust belt of the Ordos Basin are closely associated with the closure of the Paleotethys Ocean, the northward collision of geological entities such as the Qiangtang, Lhasa, and Gangdise, the southward subduction of the Okhotsk Ocean, the collision between the South China Block and the Sino–Korean Block, as well as the subduction and reduction of the Izanagi Block at the eastern margin of the Sino–Korean Block [55].

During the Late Triassic, the Ordos Block was subjected to compressional forces from the north–south (NS) and northeast (NE) directions, along with right-lateral stress, which led to the formation of the Liupanshan arc-type thrust structure system. Under NS-directed stress, the Alxa Block was pushed eastward, initiating the development of the thrust-fold belt along the western margin of the Ordos Block. In the mid-to-late Jurassic, the continental, Atlantic, Indian, and Pacific Oceans underwent expansion, connecting northward to the Bangong–Nujiang belt region [56]. Simultaneously, the northern section of the western margin of the Ordos Block experienced counterclockwise rotation, ultimately establishing the fundamental configuration of the thrust-fold structures [57]. During the Early Cretaceous, the Gangdise landmass underwent northward thrusting and collision [58,59], resulting in the extrusion of the southern section of the western margin of the Ordos Block and the formation of the Liupanshan anticline depression [57]. Geological data [60] indicate that the area west of the Helan Mountains belongs to the Alxa Block, while the region west of the Qingtongxia–Guyuan Fault is part of the Haiyuan–Liupanshan arc tectonic belt. During the Cenozoic, the Helan Mountains underwent disintegration, leading to the formation of a secondary terrestrial structure, the Yinchuan Fault Basin [61]. The final formation of the tectonic pattern along the western margin of the Ordos Block occurred approximately during the Late Himalayan period (8 Ma), when the Zhuozhi Mountain–Helan Mountain–Liupan Mountain line likely stabilized [56]. Around 7.2 Ma, sedimentation commenced in the eastern part of the Ordos Basin, while the western Liupanshan Mountains began to uplift. Ultimately, the present-day western margin exhibits a nearly north–south oriented structure, with the interior still displaying a clearly segmented foreland basin configuration.

The western boundary of the Ordos Block is defined by the Dengkou–Alxa Left Banner Fault–Sanguankou Fault–Qingtongxia–Guyuan Fault system located on the western foothills of the Helan Mountains [62]. The boundary between Ningxia and Inner Mongolia is located at the neotectonic boundary zone between the Alxa and Ordos blocks, including the eastern segment of the Bayin ‘ergong Fault Zone, the eastern foothills of the Bayin’ula Fault, the western foothills of the Helan Fault, the Zhengyiguan Fault, the western foothills of the Zhuozhi Fault, and the Dengkou–Benjing Fault. The Yinchuan–Helan Mountain basin–tectonic coupling structure in the northern part of the western margin of the Ordos belongs to the North China tectonic domain and is primarily controlled by normal faults [63]. As the north–south demarcation along the western margin of the Ordos region, the Helan Mountain West Slope Fault, the Sangguankou–Niushoushan Fault, and the Luoshan Mountain East Slope Fault have manifested right-lateral strike–slip and vertical displacements since the Late Quaternary. Geomorphological investigations and sediment dating indicate that since the Late Quaternary, the horizontal and vertical slip rates of the Sangguankou–Niushoushan Fault have been 0.35 mm/a and 0.09 mm/a, respectively, whereas those of the Helan Mountain West Slope Fault are 0.28 mm/a and 0.11 mm/a. The northern section of the Luoshan Mountain East Slope Fault shows a horizontal slip rate ranging from 1.31 to 1.54 mm/a, and its southern section approaches 1.4 mm/a [63]. The GNSS vertical velocity field reveals that the vertical slip rates of the Helan Mountain East Foothill Fault and the Liupanshan Fault range from 1 to 1.6 mm/a [64].

Figure 1b depicts the geological structures and historical seismic activities in Ningxia. Adjacent to the Niushoushan Mountain Range, the southern part of Ningxia is part of the Qilian Mountain fault zone, while the northern area is within the western margin fault zone of the Ordos Platform in the North China Block. The Helan Mountains, located in northern Ningxia within the study area, attain a maximum elevation of approximately 3500 m, with an average altitude of around 2000 m [65]. The Helan Mountain thrust system consists of a thrust fault structural framework composed of the Alxa crystalline thrust sheet and the North China foreland fold-thrust belt, with the Hengshanbao–Zuozishan fold-thrust belt serving as its leading thrust front [66]. The Zhuozi Mountain area is characterized by a series of west-trending thrust faults, while the Tiesumiao area presents a series of fault ridges and forelimbs, and the Taole–Hengshanbao area exhibits a series of east-trending thrust faults [67]. Figure 1b clearly demonstrates a distinct correlation between geological zoning boundaries and major active faults in terms of orientation and location within the study area. For instance, the Helan Mountain East Foothill Fault demarcates the Shanxi Formation from the Holocene alluvial deposits on the western foothills of the Helan Mountains. The Yinchuan–Pingluo concealed fault is situated within the Holocene alluvial layer of the Yinchuan Plain, and it aligns with the Holocene aeolian deposits in the eastern part of the plain. The Huanghe–Lingwu concealed fault lies beneath the transition zone of the Yinchuan Plain and is connected to the Linhe and Qingshuiying Formations. The Luoshan East Foothill Fault follows the Huanghe–Lingwu fault, bordered by different deposits and structural units, and forms intricate geological structures at the junction of Pliocene deposits. The Sangguankou and Niushoushan Faults are parallel strike-slip faults, representing the surface extensions of the boundary between the Ordos Block and the Longzhong Basin. The Niushoushan Fault serves as the southern boundary of the Holocene alluvial layer of the Yinchuan Plain, with complex stratigraphic units in its southwestern section. The middle section exposes the Qingtongxia Reservoir and the adjacent alluvial layers. The Yantongshan Fault is the southwestern boundary of the Upper Pleistocene aeolian deposits, with complex formations to its southwest. The Xiangshan–Tianjingshan Fault, which extends from Ningxia to Gansu, exhibits distinct geological divisions in its northern and southern sections. The northern segment contains various deposits, while the southern portion features complex geological clusters. These formations extend to Zhongwei City, with scattered Upper Pleistocene aeolian structures.

Integrating historical seismic data, it is found that primary tectonic plate boundaries and local geological structures are strongly correlated with active fault planes and intersections, and they have significant associations with historical earthquakes, particularly those of magnitude 8 or higher. Earthquake analysis and forecasting necessitate the monitoring of geophysical precursor signals and seismic void zones, as well as a comprehensive understanding of geological structures and active tectonic systems in high-risk areas. Attention should be directed towards local geological boundaries related to active fault heights, plate boundary movement patterns, and the dynamic states of major active faults. By employing geophysical observation techniques, direct observations are integrated with geophysical constraints to infer the deep crustal motion states, thereby providing crucial data for seismic hazard assessment.

There are four future high-risk areas for strong earthquakes in Ningxia and its adjacent regions: the northwest edge of the Ordos Block, the southeast section of the Xiangshan–Tianjingshan Fault and the southern section of the Yellow River–Lingwu Fault, the Tianqiaogou–Huangyangchuan Fault and the Laohushan Fault, and the Liupanshan–West Qinling northern edge Fault [68,69]. Throughout history, Ningxia’s northern region and adjacent areas have had significant seismic events, and modern strong seismic activities continue. Historically, M7.0 or higher magnitude earthquakes have occurred at the

eastern foot of the Helan Mountain Fault Zone and the Xiangshan–Tianjingshan Fault Zone in northern Ningxia. On 15 April 2015, an M5.8 earthquake happened in Azuoqi, Inner Mongolia, at the northwest edge of the Ordos Block within the North China Block, north of Ningxia. Then, on 3 June 2017, an M5.0 earthquake struck Azuoqi, Inner Mongolia, and its causative fault belongs to the Qilian Mountain Seismic Belt. After the 1976 M6.2 earthquake in Bayinmuren, Inner Mongolia, the Ningxia–Inner Mongolia border area has seen relative seismic quiescence for M6 earthquakes. The 2015 M5.8 earthquake in Azuoqi, Inner Mongolia, had low intensity and insufficient energy release, and there has been a lack of significant M6 earthquakes. On 2 January 2025, an M4.8 earthquake took place in Yongning County, northern Ningxia, nearly disrupting the period of seismic calm. In the next 200 years (after 2006), the southern section of the Yellow River Fault in northern Ningxia has a relatively high risk of characteristic earthquakes and may experience an M7.0 earthquake [70–73]. Moreover, the scarcity of fixed-point precursor observation points in this border region makes it difficult to assess regional seismic risk, suggesting a continuous risk of moderate-to-strong earthquakes [72,73].

3. Data and Methods

3.1. Data

The research scope of this project encompasses northern Ningxia and its adjacent regions. The primary data sources are the China Crustal Movement Observation Network (CMONOC, which has been in formal operation since 1999) and the China Continental Tectonic Environment Monitoring Network (CMTEMN, upgraded in 2011). Under the guidance of the China Earthquake Administration, Beijing, China and jointly implemented by multiple institutions, CMTEMN represents China's most significant scientific project in GNSS geoscience research in terms of investment, scale, and precision. It consists of 260 GNSS reference stations and 2000 GNSS regional stations [74]. Presently, the project has 8 continuous GNSS stations in northern Ningxia. Although the continuous observation of base stations is constrained by the insufficient density of stations, the spatial density of CMTEMN stations is compensated for through approaches such as the irregular re-measurement of regional stations, the utilization of mandatory antenna alignment devices, regional synchronous observation, and intensive spatial observation in significant tectonic areas via research projects. The current horizontal positioning accuracy ranges from 1 to 2 mm [75].

This study employs data from continuous and discontinuous GNSS stations, as well as global IGS stations since 1998, as the fundamental dataset. It selects 21 uniformly distributed IGS stations around the research area in Eurasia as reference points for one-day relaxation solutions and velocity adjustments under the ITRF2014 framework. In the one-day calculation, 54 continuous stations, 475 mobile stations, and 21 IGS reference stations are included, and global IGS stations are used for adjustment purposes. The boundary between Niushoushan and Luoshan in Wuzhong, Ningxia, is defined as the north–south boundary of Ningxia, and station data from N37.43° to N40°, E104° to E108.4° are processed.

All GNSS data utilized in this study are sourced from CMTEMN, with the observation time windows for all stations spanning from March 1998 to December 2024. Approaches for addressing inconsistent processing times (e.g., discontinuous observations) primarily encompass pre-processing, interpolation, and outlier handling. The data in this paper resolves inconsistent processing times in the following manner: for short interruptions (<less than 30 s), linear interpolation or satellite motion models are employed to complete the pseudorange/carrier phase; for long-term interruptions (>greater than 1 min), segmented processing is carried out with parameter reset. For time-overlapping or out-of-order data,

optimal observations are retained based on factors such as signal strength and satellite elevation angle. Python v3.8 scripts are used to automatically identify time jumps and call gLAB or Precise Point Positioning (PPP) tools for batch processing.

3.2. Workflow

The workflow consists of the following steps:

1. Apply the GAMIT v10.6 software to calculate the regional one-day relaxation solution for the reference station and the surrounding International GNSS Service (IGS) stations.
2. Employ the GLOBK v10.6 software to integrate regional and global one-day relaxation solutions sourced from *sopac*. Determine seven similarity transformation parameters in relation to ITRF2014 through IGS stations and obtain the station's initial coordinates and velocities within the ITRF2014 framework. It should be noted that velocities require estimation in the *glorg* command file.
3. Transform the station velocities to the Eurasian reference frame, and remove overall rotation, unreliable station data, trends, and coseismic deformations.
4. Ascertain the relative velocity of the two plates perpendicular to the active fault through velocity projection.
5. Acquire the meshed strain rate field and depict it; after refining the secondary block, the internal strain rate variations within the secondary block are determined.
6. Compute the length of geodetic baselines by means of the Bessel iteration method.
7. Construct an improved dislocation model based on the surface geological structure of local active faults, and determine the slip rate, slip rate loss, and locking degree of active faults through inversion of the GNSS velocity field.

3.3. Velocity Field and Strain Rate Field

The relatively stable Eurasian Plate is selected as the reference, and the least-squares method is utilized to solve for the Euler motion vector. Through this approach, the velocity of the Eurasian Plate within the ITRF2014 geocentric coordinate frame can be determined, and consequently, the velocity field of northern Ningxia and its adjacent areas relative to the Eurasian Plate is obtained. To reduce the influence of earthquakes with magnitudes greater than M6, the coseismic displacement is calculated using continuous GNSS observation data, and the coseismic effect is removed during the solution of the velocity field. Equation (1) is employed to calculate the movement rate, where r is the average radius of the Earth, and λ and φ are the longitude and latitude of the station, respectively. Several IGS core stations are selected, and the least-squares method is used to solve the Euler motion vector Ω (Equation (1)).

The GNSS velocity field in northern Ningxia and its adjacent areas within the Eurasian framework is presented in Figure 1. The Liupanshan Fault is located at the intersection of the northeastern margin of the Qinghai–Tibet Block and the Ordos Block, while the Yinchuan Basin is situated between the Alxa Block and the Ordos Block. These junctions of land blocks are the key monitoring areas of the CMONOC. The black circles with vector arrows denote the positions of GNSS stations, where the length of the vector represents the relative horizontal crustal movement rate during this period, and the diameter of the black circle represents the relative error.

To explore the activity trend of the fault zone in this region, cross-sectional analyses are conducted on the main faults in northern Ningxia and its adjacent areas since the Quaternary, including the Zhengyiguan Fault, Eastern Helan Fault, Yinchuan–Pingluo Fault, Yellow River–Lingwu Fault, and Niushoushan Fault. Velocity projections are calculated, and the deformation width, depth, and characteristics of compression and strike-slip defor-

mation on both sides of the fault are investigated. The coverage area extends to 200 km on either side of the fault.

$$\begin{bmatrix} V_n \\ V_e \end{bmatrix} = \begin{bmatrix} r \sin \lambda & -r \cos \lambda & 0 \\ -r \sin \varphi \cos \lambda & -r \sin \varphi \sin \lambda & r \cos \varphi \end{bmatrix} \begin{bmatrix} \Omega_x \\ \Omega_y \\ \Omega_z \end{bmatrix} \quad (1)$$

Strain rate denotes the strain accumulation related to regional stress concentrations and seismic activity. Leveraging the spatial distribution characteristics of strain rate can facilitate the analysis of strain accumulation and seismic risk within each region. We calculate the strain rate at any arbitrary point based on the velocity values of several adjacent stations, and the model algorithm can be expressed by Formulas (2) and (3).

v_e^i and v_n^i are the velocity components of station i ; v_e^j and v_n^j are the velocity components of the interpolation point j ; Δx_{ij} and Δy_{ij} are the distance components from station i to the interpolation point j ; ε_{ee}^j , ε_{en}^j , and ε_{nn}^j are the strain rate tensors of the interpolation point j ; ω^j is the rotation rate of the interpolation point j ; r_e^i and r_n^i are the velocity component residuals of station i .

Since the distances between the stations and the interpolation points are unequal, the weight value of each station is determined by the weight W_s obtained from the standard deviation of the station velocity and the distance weight between the stations and the interpolation points (Equation (3)). σ is the observation error, and the larger the error, the smaller the contribution; d is the distance between the station and the interpolation point. The shorter the distance, the greater the contribution; a represents the characteristic length, which is defined as the distance from the fourth-nearest point to the interpolation point.

Initially, we construct a grid within the study area with uniform spacing. Subsequently, each grid point is regarded as the center, and more than four observation stations are selected within the search radius for calculation. Calculations are not carried out if there are fewer than four observation stations.

The selection of grid spacing and search radius is contingent upon the actual distribution of monitoring stations. A series of parameters were tested to evaluate the sensitivity of strain rate estimation. An excessively small grid spacing (below the average distribution distance of mobile stations) results in redundant computations. Ideally, each grid should contain one mobile station; however, due to the uneven distribution, an approximation is required. Grid spacings ranging from 10 to 50 km and several search radii were examined. Search radii below 100 km may result in an inadequate number of stations, whereas those above 100 km introduce unnecessary redundancy. When a grid spacing of 25 km and a search radius of 100 km are employed, the strain rate varies relatively smoothly. Based on the distribution of survey points in the study area, a grid spacing of 25 km and a search radius of 100 km are utilized for calculations in this project.

$$\begin{bmatrix} v_e^j \\ v_n^j \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta x_{ij} & \Delta y_{ij} & 0 & -\Delta y_{ij} \\ 0 & 1 & 0 & \Delta x_{ij} & \Delta y_{ij} & -\Delta x_{ij} \end{bmatrix} \begin{bmatrix} v_e^j \\ v_n^j \\ \varepsilon_{ee}^j \\ \varepsilon_{en}^j \\ \varepsilon_{nn}^j \\ \omega^j \end{bmatrix} + \begin{bmatrix} r_e^i \\ r_n^i \end{bmatrix} \quad (2)$$

$$\begin{cases} W_s = 1/\sigma \\ W_d = \exp(-d^2/2a^2) \end{cases} \quad (3)$$

Within a unified reference framework, the alteration of GNSS baseline length can objectively and directly mirror the dynamic changes in the relative motion between two stations. Utilizing the continuous observation data points from GNSS, we directly computed the alterations of several baselines in northern Ningxia and its adjacent regions from 2010 to 2024. The longitude and latitude were calculated based on the new coordinates of the reference station, and subsequently, the baseline length was computed via the Bessel iteration method. The rate of change in the geodetic baseline length, which reflects crustal elongation and shortening, was obtained through linear fitting.

3.4. Deep Inversion of GNSS Crustal Deformation

The theory of geophysical inversion (Equation (4)) and the elastic dislocation theory [35,36] offer methodologies for estimating fault slip magnitude, slip angle, locking degree, and slip rate loss from observational data, such as the displacement of observation points. Equation (4) depicts the observed data, specifically the surface displacement induced by the sliding of rectangular faults within an elastic half-space. It represents the theoretical numerical relationship between observational data U and fault parameters. This relationship G is commonly referred to as Green's function. Additionally, D is the sliding parameter vector, and the term e represents the error.

$$U = G \cdot D + e \quad (4)$$

For point sources and finite rectangular sources, ref. [36] presented a comprehensive set of closed-form analytical expressions for the internal displacement and strain resulting from shear and tensile fractures within a half-space. These expressions systematically incorporate terms representing deformations in infinite media. Given that geophysical observations are restricted to the near-surface, it holds significant practical value to theoretically examine the expressions of various physical quantities on the surface of a half-space. By summarizing the surface displacement and strain caused by shear faults within a half-space, the static changes in the surface gravity field and piezomagnetic field induced by dislocation sources within a half-space, and the explicit expressions of surface displacement caused by mid-point shear sources within an isotropic layered half-space, previous research has been extended to the internal deformation field caused by shear and tensile faults within a half-space.

Based on Okada's half-space elastic dislocation theory [36], an active fault and secondary active block motion model for northern Ningxia was established. DEFNODE adopts an adaptive regularization method to solve the constrained inversion equations. It automatically adjusts the regularization factor according to the geometric complexity of fault segments to strike a balance between data fitting and model smoothness. Simultaneously, it calculates the long-term slip rates across the depth range of 0–30 km for each fault segment, thereby obtaining accurate slip loss rates. Subsequently, the locking rate of each fault segment is derived from the ratio of the slip loss rate to the long-term slip rate. Based on the joint estimation of the slip loss rate and geophysical/geodetic data, the long-term slip rate is calculated, and the locking rate for each segment is determined using the specified formula. The regularization inversion equation is presented as Equation (5):

$$D = (G^T G + \mathcal{E} L^T I)^{-1} G^T U \quad (5)$$

The regularization parameter $\mathcal{E}$, determined through the l (smooth matrix) curve method, is set to 0.01. The locking rate is calculated according to Equation (6):

$$l_r = \frac{d}{V_l} \times 100\% \quad (6)$$

The locking rate l_r , sliding loss rate d (mm/a), and long-term sliding rate V_l (mm/a) are presented, respectively. Using the GNSS velocity field from the Eurasian benchmark since 2018, the inter-seismic motion parameters of the main active faults (the eastern foot of Helan Mountain Fault, Liupanshan Fault, Haiyuan Fault, the northern edge of West Qinling Mountain Fault, and Zhuanglanghe Fault) were inverted, excluding coseismic effects. The velocity field accuracy is better than 0.5 mm/a.

Drawing on the findings of [76–81] and the research results of the Ningxia Seismological Bureau on active faults, the fault dip angle was obtained, and the fault structure was configured as follows: three isobaths were set for all faults, and several nodes were set on each isobath. The fault dip angle is detailed in Table 1.

Table 1. Dip angle setting of main active faults in northern Ningxia.

Fault	Dip Angle (Degrees)
Dengkou–Benjing Fault	70
Zhuozi western margin fault	62
Zhengyiguan Fault	75
Bayannula Mountain Fault	50
Eastern Helan Piedmont Fault	70
Sanguankou Fault	50
Chonggang–Luhuatai Fault	75
Yinchuan–Pingluo Fault	75
Yellow River–Lingwu Fault	75
Niushoushan–Luoshan Fault	45
Guanguanling fault	80
Xiangshan–Tianjingshan Fault	70
Qingshuihe Fault	65
Yantongshan Fault	55

The fault structure is configured as follows: a series of nodes (not exceeding 60 per segment) are selected along the strike direction of each fault line, and three nodes are established in the depth direction. During the inversion process, pre-existing Green’s functions are utilized, and new ones are generated only when necessary, thereby avoiding the mandatory generation of Green’s functions for all faults. The tolerance threshold for new node positions triggered by the calculation of new Green’s functions is set at 1 km. Stations within a distance of 1 km from any node employ the same Green’s function, while stations located more than 1500 km away from any node do not generate a Green’s function.

Simultaneous implementation of simulated annealing and grid search is carried out. The initial temperature is set at 100 °C, with 40 iterations and a mandatory locking rate ranging from 0 to 1. Gradient search is performed on parameters in a random sequence, with a grid search step size of 100 steps and 15 iterations across all parameters. The search step size for slip loss and slip rate is set at 0.001 mm/year.

The uncertainty of parameters is generated concurrently. Upon completion, the forward model is constructed, strain and rotation residuals are calculated, and the optimal solution is obtained by adjusting the error weight factors of the GNSS velocity field through GPS NE covariance estimation. The locking rate does not decrease with depth.

The fault strike–slip and dip–slip rates are defined as positive in the N/E direction and negative in the S/W direction. The horizontal velocity field resolution for inversion is 0.01 mm/year, the slip loss resolution is 0.001 mm/year, the slip rate resolution is 0.001 mm/year, and the locking rate resolution is 0.001. The mapping resolution is set as 1 km along the strike and 0.5 km along the dip.

During the inversion process, the fault locking degree decreases monotonically with depth, and the optimal χ^2 is achieved by adjusting the GNSS velocity field error weight factor.

4. Results

4.1. The Velocity Field

The collision and compression between the Indian Plate and the Eurasian Plate have resulted in the shortening of the Qinghai–Tibet Plateau’s crust in the north–south (NS) direction. Spatially, the deformation response pattern propagates from the south to the north and deepens from the cap rock to the basement, characterized by numerous folds and faults with diverse natures and orientations. Blocked by the rigid blocks in the north and east (e.g., the Alxa Block, Ordos Block, and South China Block), the crustal material has started to migrate in the northeast (NE) direction.

Within the interior of the plateau, the NE-directed material migration is relatively rapid. When encountering relatively rigid blocks, topographic uplift occurs, as demonstrated by the Helan Mountains and Liupan Mountains in Figure 2. In the study area, especially in the southern part, the faults are predominantly left-lateral, which is part of the long-term background inherited movement in the northeastern margin of the Qinghai–Tibet Plateau.

Subsequent to 2009, the density of GNSS stations of the China Seismological Bureau witnessed an increase. This research utilizes the velocity field spanning from 1999 to 2024. As depicted in Figure 2 within the Eurasian stability framework, the velocity vectors at each point in northern Ningxia and its adjacent regions generally exhibit a clockwise rotation in the NNE–NE–NEE direction. Along the NE–SW principal compressive stress, the magnitude of the velocity vector diminishes from the SW to the NE. Along the principal tensile stress direction, the magnitude increases from the NW to the SE, which reflects the attenuation of the compression exerted by the Indian Plate from the south–north and west–east directions.

Locally, the Liupanshan Fault is characterized by strong locking. The Haiyuan Fault Zone persists in left-lateral strike–slip movement. The Yinchuan Basin is in a state of east–west extension, with a steadily increasing extension velocity and simultaneous right-lateral strike–slip. Furthermore, the stress accumulation in the NE–NEE direction of tectonic faults in northern Ningxia and adjacent areas has been intensified.

During the period from 2014 to 2022, the velocity directions on both sides of the Yinchuan Basin, Liupanshan Fault, and Haiyuan Fault remained relatively stable. A slow-down was observed in the southwestern margin of the Alxa Block, the interior of the Ordos Block, and the northeastern wall of the Haiyuan Fault, accompanied by enhanced stress accumulation. The northern wall of the Altyn Tagh Fault and the southwestern edge of the Alxa Block continued to decelerate, whereas the extensional velocity of the Yinchuan Basin increased, facilitating the acceleration of material spillover between the Alxa and Ordos Blocks [4]. In the vicinity of the Eastern Helan Mountain Fault, Bayan–Wula Mountain Fault, Yellow River–Lingwu Fault, Yinchuan–Pingluo Fault, Xiangshan–Tianjingshan Fault, Luoshan Fault, Yantongshan Fault, and Sanguankou Fault, the east–west velocity of GNSS exhibited an increase, signifying heightened crustal movement activity.

To effectively distinguish the activities of each segment of the major active faults in the region, the active faults in the Chinese mainland were segmented based on the detection results of [14,76], with the aim of ensuring that the strike of the segment aligns as closely as possible with a straight line. The station velocities were projected onto a segmented or individual fault line, and the horizontal crustal movement on both sides of the fault was analyzed. When the displacement velocity of a station is opposite to the preset positive orientation, it is recorded as negative. The reference system determines the sign of slip values. For instance, when the fault strike is taken as the positive x-axis, left-lateral strike–slip movement is regarded as negative. By using the regional tectonic stress orientation as a reference, the rotation angle of the principal stress axis is determined through Mohr’s circle analysis. The selection of the reference system does not influence the identification of the fault mechanism; however, standardized data processing is crucial to avoid sign



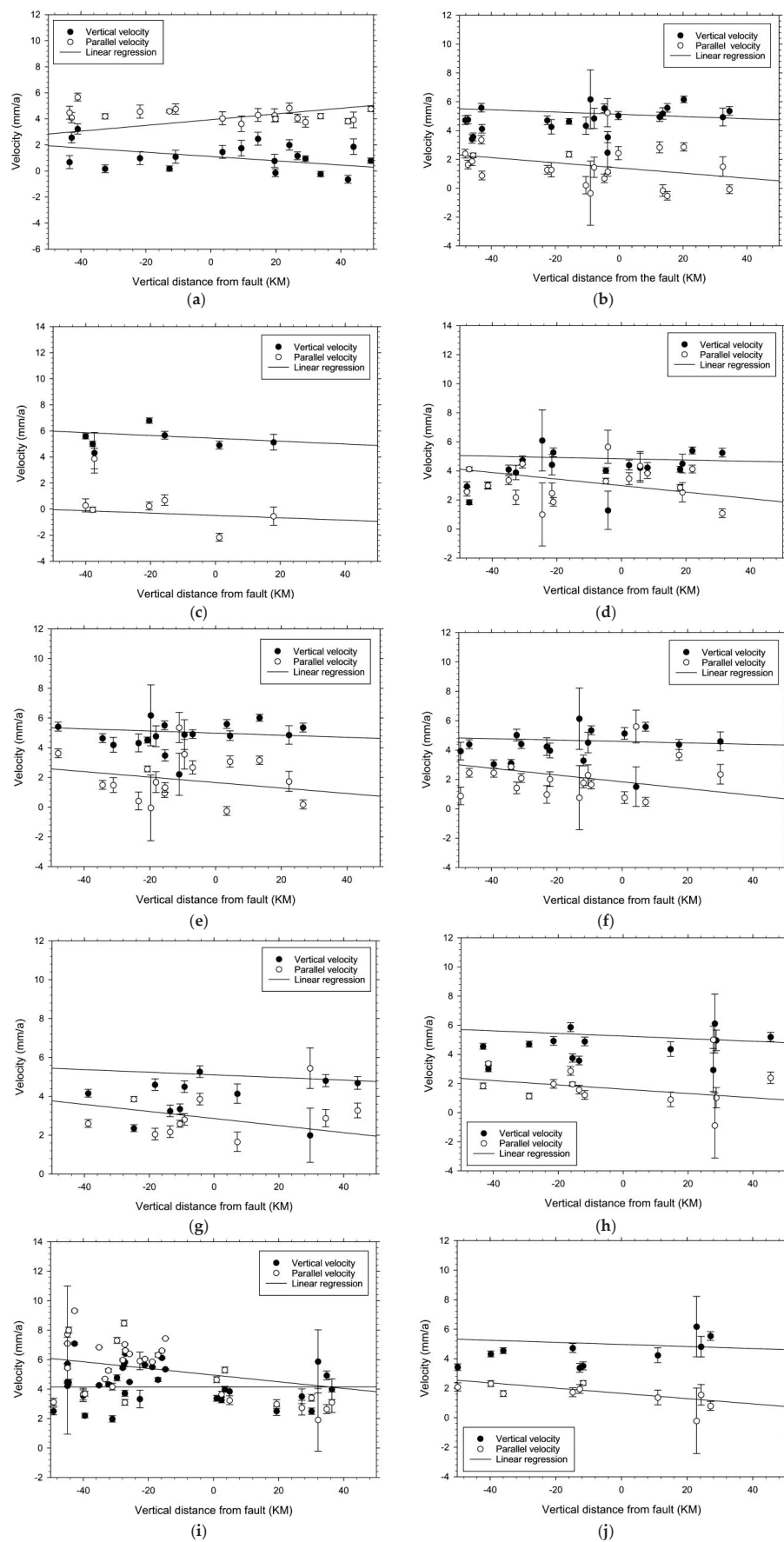


Figure 3. Cont.

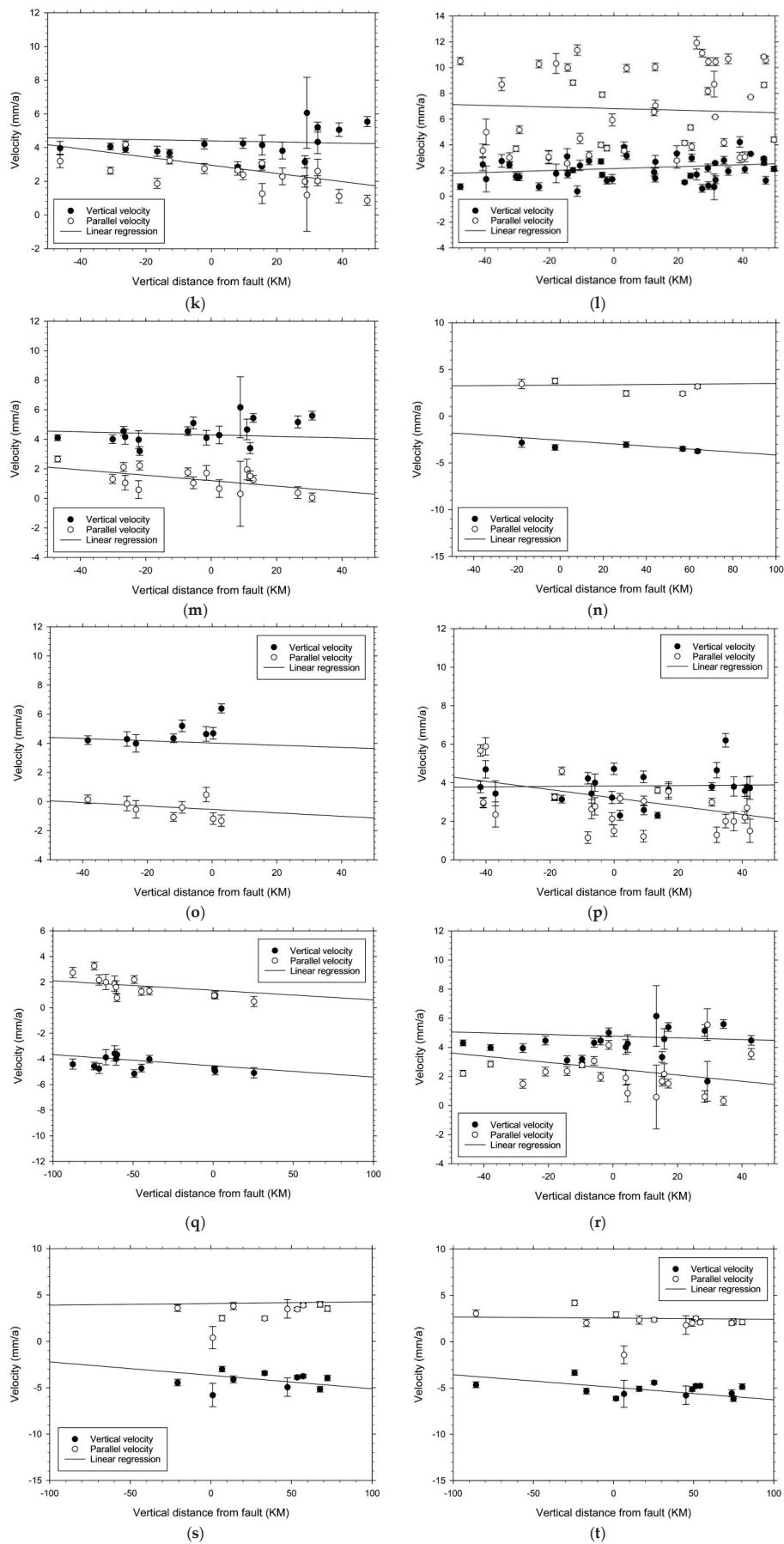


Figure 3. Projection of Global Navigation Satellite System (GNSS) velocity along (a) Zhengyiguan Fault, (b) Yinchuan–Pingluo Fault, (c) Yellow River–Lingwu Fault Section 1, (d) Yellow River–Lingwu

Fault Section 2, (e) Yellow River–Lingwu Fault Section 3, (f) Yellow River–Lingwu Fault Section 4, (g) Eastern Helan Piedmont Fault Section 1, (h) Eastern Helan Piedmont Fault Section 2, (i) Eastern Helan Piedmont Fault Section 3, (j) Eastern Helan Piedmont Fault Section 4, (k) Eastern Helan Piedmont Fault Section 5, (l) Eastern Helan Piedmont Fault Section 6, (m) Eastern Helan Piedmont Fault Section 7, (n) Sanguankou Fault, (o) Zhuozi West Fault, (p) Dengkou–Benjing Fault, (q) Bayannula Mountain Fault, (r) Chonggang–Luhuatai Fault, (s) Niushoushan Fault northern section, (t) Niushoushan Fault southern section.

Table 2. Observed velocity differential between the two plates of active faults in Northern Ningxia and adjacent regions from 1998 to 2024.

Fault	Observed Vertical Velocity Range Within a Distance of 50 km on Either Side of the Fault (mm/yr)	The Observed Vertical Velocity Disparity Between Two Plates at the Point Closest to the Fracture (mm/yr)	Observed Range of Parallel Velocity Within a Distance of 50 km on Either Side of the Fault (mm/yr)	The Observed Difference in Parallel Velocity Between Two Plates at the Point Closest to the Fracture (mm/yr)
Zhengyiguan Fault	−0.65–3.38	0.3 ^a	3.62–5.67	0.73 ^a
Yinchuan–Pingluo Fault	2.48–6.22	0.03	−0.55–5.28	0.34 ^a
Yellow River–Lingwu Fault Section 4 (South)	4.35–6.88	0.82	−2.15–0.73	2.85
Yellow River–Lingwu Fault Section 3 (Middle)	1.24–6.11	3.08 ^a	1.01–5.69	2.23
Yellow River–Lingwu Fault Section 2 (Middle)	2.3–6.22	0.74 ^a	−0.23–5.33	2.96
Yellow River–Lingwu Fault Section 1 (North)	1.54–6.19	0.16	0.47–5.59	0.86
Eastern Helan Piedmont Fault Section 1 (South)	2.04–5.29	1.19	1.69–5.47	2.19
Eastern Helan Piedmont Fault Section 2	3.03–6.12	0.54	0.92–4.96	0.34
Eastern Helan Piedmont Fault Section 3	1.92–7.05	1.91	1.92–9.26	2.76
Eastern Helan Piedmont Fault Section 4	3.4–6.19	0.67 ^a	−0.21–2.35	1.02
Eastern Helan Piedmont Fault Section 5	2.92–6.03	1.3	0.88–4.20	0.1
Eastern Helan Piedmont Fault Section 6	0.35–4.29	2.47	2.56–11.9	2.38
Eastern Helan Piedmont Fault Section 7 (North)	3.22–6.19	0.13 ^a	0.31–2.72	1.07
Sanguankou Fault	−3.24–2.79	0.24 ^a	2.43–3.78	1.35 ^a
Zhuozi West Fault	3.97–6.48	0.02 ^a	−1.28–0.50	1.68
Dengkou–Benjing Fault	2.33–6.25	2.47 ^a	1.13–5.90	1.68 ^a
Bayannula Mountain Fault	−5.18–4.0	0.66	0.50–2.17	0.35
Chonggang–Luhuatai Fault	1.67–6.17	1.00	0.31–5.56	2.28
Niushoushan Fault northern section	−5.81–2.83	1.39	0.44–3.83	3.14 ^a
Niushoushan Fault southern section	−6.10–3.36	0.82	−1.44–4.15	0.9 ^a

Superscript a denotes that the direction of the relative movement rate of the observation point closest to the fault is opposite to the overall movement trend of the fault's two plates, as indicated by the trend line.

Via a phased comparative analysis of the velocity field [1,4], it can be inferred that the Bayannula Mountain Fault (a normal fault) in Inner Mongolia exhibits no obvious movement. The vertical velocity shows an upward trend, whereas the velocity difference in the vicinity of the fault decelerates. The parallel velocity experiences a slight decrease, and the velocity difference near the fracture reduces.

The Dengkou–Benjing Fault (a normal fault) exhibits strike–slip movement. The vertical velocity experiences an upward trend, and the near-fault velocity differential shows a slight increase. The parallel velocity slightly increases with a notable disparity, and the near-fracture velocity differential increases.

The Western Zhuozi Fault (a normal fault) demonstrates both dip–slip and strike–slip motions. Both the vertical and parallel rates slightly decline with disparities, and the near-fracture velocity differential undergoes minimal change.

The movement of the ZHENGYIGUAN Fault (a concealed reverse fault) in Shizuishan City, Northern Ningxia, is not pronounced. The vertical velocity scarcely changes, and the near-fault velocity differential varies minimally. The parallel velocity slightly decreases, and the near-fracture velocity differential is not distinct.

In the vicinity of the first (South) segment of the Eastern Helan Eastern Piedmont Fault in the western part of the Yinchuan Basin, the velocity changes are minimal. The second segment exhibits strike–slip and dip–slip motions. The alterations in vertical and parallel velocities are not apparent, and the near-fault velocity differential increases. The third segment shows obvious strike–slip movement. The vertical velocity slightly increases, and the near-fracture velocity differential changes little. The parallel velocity changes minimally, and the near-fracture velocity differential increases. The fourth segment displays strike–slip and dip–slip motions. The vertical velocity slightly increases, and the near-fault velocity differential increases. The parallel velocity changes minimally, and the near-fracture velocity differential decreases. The fifth segment shows obvious strike–slip and some dip–slip motions. Both the vertical and parallel velocities slightly increase, and the near-fracture velocity differential increases. The 6th segment exhibits dip–slip motion. Both the vertical and parallel velocities slightly increase, and the near-fracture velocity differential changes little. The seventh (North) segment shows dip–slip motion, and the strike–slip motion decelerates. The vertical velocity changes minimally, and the near-fracture velocity differential exhibits few changes. The parallel velocity slightly decreases, and the near-fracture velocity differential decreases. There is a disparity in vertical velocity between the two fault blocks.

The Chonggang–Luhuatai Fault (concealed) in the Yinchuan Basin shows dip–slip motion. The vertical velocity and its differential change minimally. The parallel velocity slightly decreases, and the near-fault velocity differential slightly increases. The Yinchuan–Pingluo Fault (concealed) exhibits strike–slip and dip–slip motions. The vertical velocity increases, the parallel velocity changes minimally, and the near-fault velocity differential changes little.

The first segment of the Yellow River–Lingwu Fault (buried) in the eastern part of the Yinchuan Basin shows strike–slip motion. The vertical velocity slightly decreases, the parallel velocity slightly increases, and the near-fault velocity differential exhibits few changes. There is a disparity in horizontal velocity between the two fault blocks. The second segment shows weak dip–slip and strike–slip motions. Both the vertical and parallel rates slightly increase. The near-fracture velocity differential decelerates. The third segment shows slow dip–slip and obvious strike–slip motions. The vertical velocity changes minimally, and the near-fault velocity differential decelerates. The parallel velocity slightly increases, and the near-fracture velocity differential increases. The fourth segment shows slow dip–slip and no significant strike–slip motion. Both the vertical and parallel velocities increase, and the near-fault velocity differential decelerates.

The Sanguankou Fault in the southern part of the Helan Mountain shows no obvious strike–slip movement. Both the vertical and parallel velocities slightly increase. The near-fracture vertical velocity differential slightly increases, and the parallel velocity differential is small.

The North segment of the Niushoushan buried Fault passing through Wuzhong City shows accelerated strike–slip motion. The vertical velocity slightly increases, and the near-fault velocity differential decreases. The parallel velocity slightly decreases, and the near-fracture velocity differential increases. The South segment of the Niushoushan

shows accelerated strike–slip motion. The vertical velocity slightly decreases, and the near–fault velocity differential decreases. The parallel velocity changes minimally, and the near–fracture velocity differential increases.

4.2. Strain Rate

The strain accumulation, indicated by the strain rate, correlates with seismic activity and regional stress concentration. The strain rate field is computed by discretizing velocity field outcomes into a mesh, from which the tensile, compressive, first shear, and surface strain rates in northern Ningxia and adjacent regions from 1999–2024 are derived. A gradually color–changing isoline is plotted (Figure 4). All strain rate plots are generated using GMT’s vector plotting function, and the software’s built-in smoothing or spatial filtering commands are used for better visual clarity. Assuming a spatially negative correlation and uniform distribution of strain rates, gradient color coding shows their planar variations. The low-value region of the shear strain rate corresponds to the low-pressure strain area. The principal strain axis is calculated from the velocity field. Compared to before 2018, the principal compressive stress direction in Ningxia and surrounding areas has shifted northwest/north by northwest, with a relatively consistent direction. The compressive strain rate is from -2.5×10^{-8} to -0.3×10^{-8} per annum, and the tensile strain rate is from -0.3×10^{-8} to 0.8×10^{-8} per annum. In northern Ningxia, compressive and tensile strains are approximately equal, while in the south, compressive strain dominates. From 1999 to 2024, the compressive strain in southwestern Ningxia (south of the Yellow River–Lingwu Fault and Sanguankou Fault) was notably high.

The Bayanwula Mountain, Dengkou–Benjing, Eastern Helan Piedmont, Yinchuan–Pingluo, Yellow River–Lingwu, Xiangshan–Tianjingshan, Yantongshan, and Sanguankou Faults exhibit gradient variations in compressive strain, accompanied by distinct dip–slip characteristics. The Bayanwula Mountain, Dengkou–Benjing, Zhengyiguan, Eastern Helan, Yinchuan–Pingluo, Yellow River–Lingwu, Sanguankou, and Luoshan Piedmont Faults demonstrate gradient alterations in shear strain rate along their strike directions. Specifically, the overall shear strain gradient of the Dengkou–Benjing, Zhengyiguan, and Yellow River–Lingwu Faults ranges from 0.4×10^{-8} per annum to 0.6×10^{-8} per annum, suggesting notable strike–slip movement. The linear and shear strain rates of the Eastern Helan Piedmont, Sanguankou, Niushoushan, Zhengyiguan, and the southern section of the Yellow River, Luoshan Piedmont, and Xiangshan–Tianjingshan Faults are comparatively high.

4.3. Baseline Time Series

The time series, defined as the difference between a subsequent observation and the initial one, can be employed to monitor the dip–slip condition by utilizing data in the vicinity of the baseline length across the fault, facilitating inferences regarding the locking potential. Continuous observation is required for CORS; however, not all stations conduct full-time observations, resulting in a non-uniform baseline solution due to measurement deficiencies and data omissions. Figure 5 depicts the fracture location and the CORS baseline distribution. The GNSS baseline and base station codes are as follows: NMWT, NMWH, NMAZ, NXYS, YANC, NXZW, NXHY, GSMQ.

The baseline time-series data from the selected CORS in the study area over the past decade indicate that the NMAZ–NXYS baseline traverses the Eastern Helan Piedmont fault and is experiencing accelerated tensile deformation, which is consistent with the lateral-tension context of the Yinchuan Basin and the normal-fault properties of the fault. The NMAZ–YANC baseline intersects multiple faults and is under stable tensile stress, in line with the basin’s lateral-tension and fault characteristics.

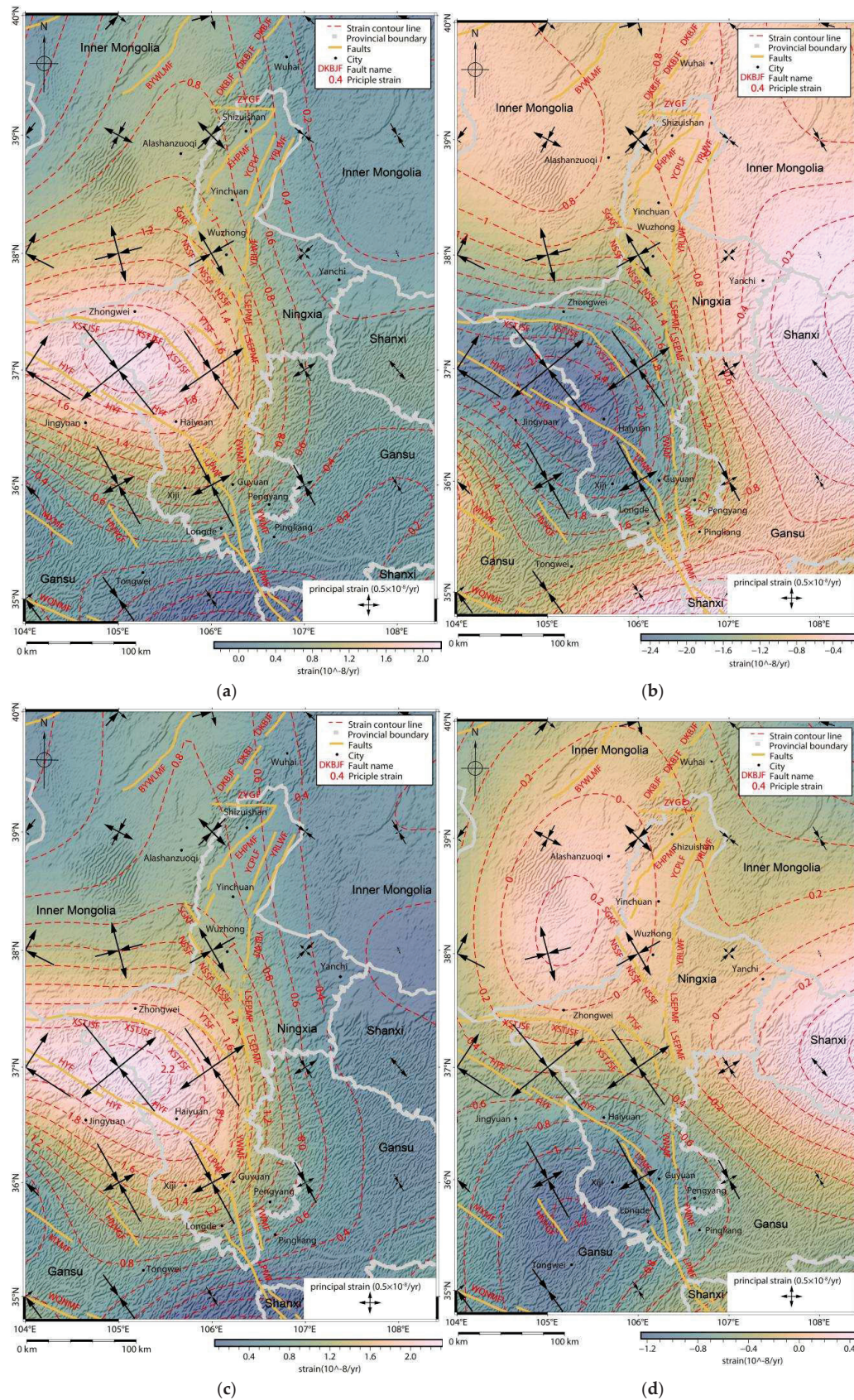


Figure 4. (a) Maximum principal strain (tensile strain); (b) minimum principal strain (compressive strain); (c) first shear strain; (d) plane strain in northern Ningxia and its adjacent areas from 1999 to 2024.

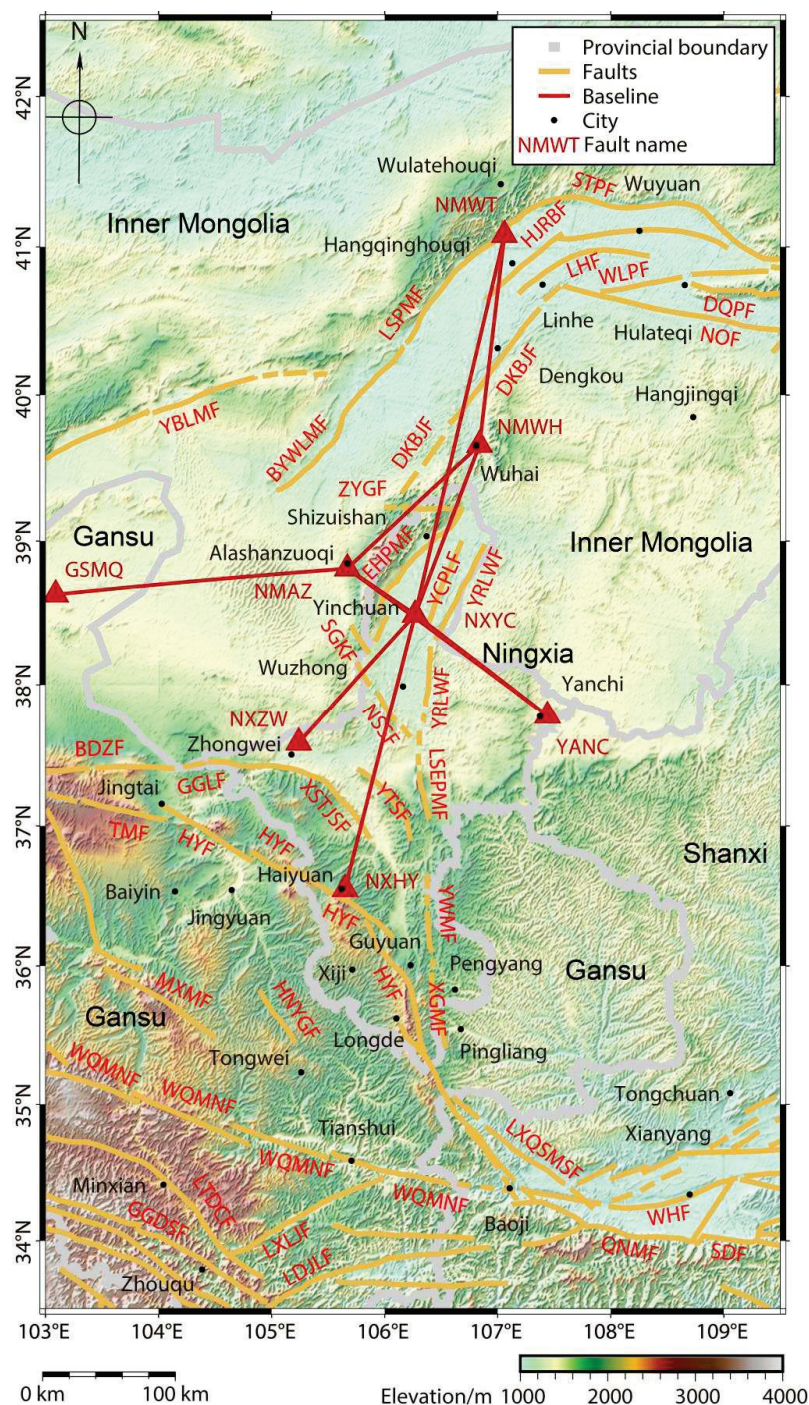


Figure 5. The distribution of major GNSS CORS and certain baselines in northern Ningxia and its adjacent regions.

The NMWH–YANC baseline crosses the Zhengyiguan Fault and has undergone stable compressive deformation for nearly five years, corresponding to the coexistence of lateral-tension and longitudinal compression in the basin. The NXYC–NXZW and NMWH–NMWT baselines have transitioned from continuous compression to a stable state, with no significant dip–slip displacement, suggesting the possibility of locking.

The NXYC–NMWT baseline crosses the Zhengyiguan Fault, changing from continuous tension to slow compression with dip–slip displacement. The NMAZ–NMWH baseline crosses the same fault and is under continuous compression with dip–slip displacement.

The NXYC–NXHY baseline intersects multiple faults, shifting from continuous compression to a stable state, where the faults may exhibit dip–slip displacement or be locked.

The NXYC–YANC baseline crosses two faults, transitioning from continuous tension to accelerated compression with dip–slip displacement. The continuous compression of the NMMQ–NMAZ baseline has adjusted to reverse extension, reflecting the extensional features of the Yinchuan Basin.

4.4. Fault Slip Deficit Rate, Locking Rate, and Slip Rate

We selected the major active faults in northern Ningxia and carried out geophysical inversion analyses on their slip rate deficit, locking rate, and slip velocity. The distribution of faults in the study area is depicted in Figure 6, where the active faults involved are marked in red. The results related to slip rate deficit, locking rate, and slip velocity are shown in Figures 7–10.

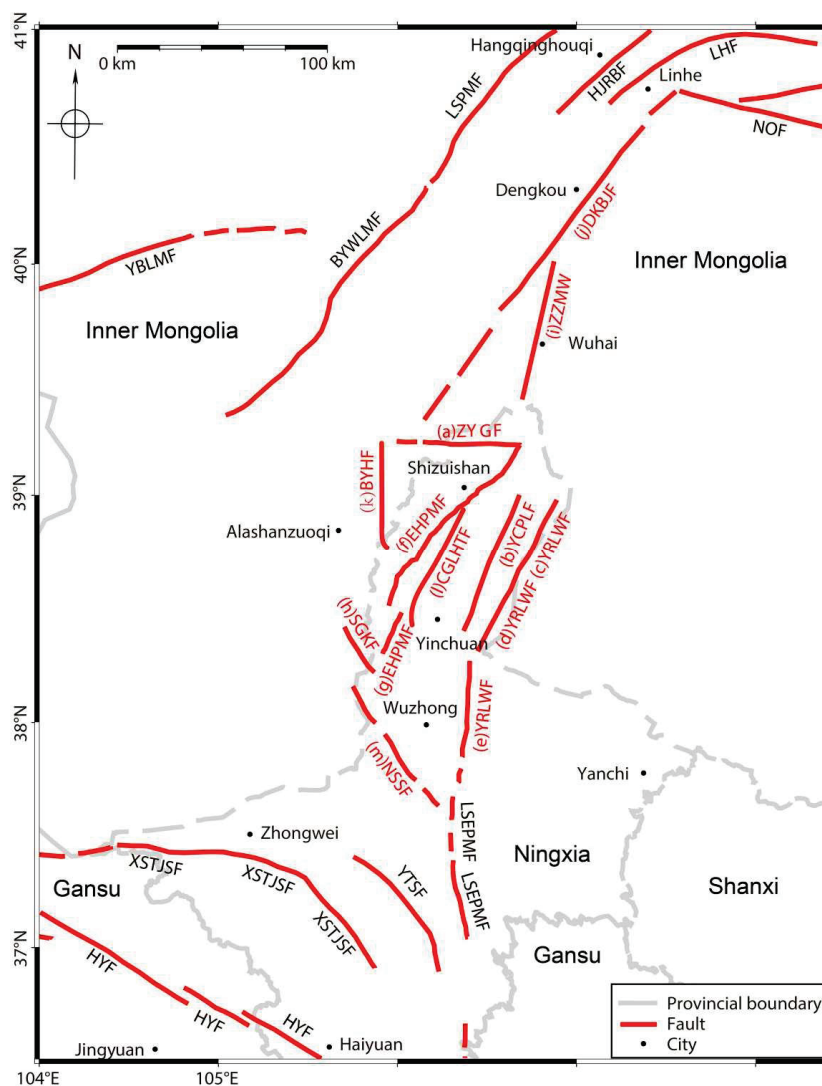


Figure 6. The distribution of active faults in northern Ningxia, which is the subject of geophysical inversion within the scope of this study (fault abbreviations are presented in red font), is as follows: (a) Zhengyiguan Fault (ZYGF), (b) Yinchuan–Pinguo Fault (YCPLF), (c) Section 1 of the Yellow River–Lingwu Fault (YRLWF), (d) Section 2 of the YRLWF, (e) Sections 3 and 4 of the YRLWF, (f) Sections 1 to 5 of the Eastern Helan Piedmont Fault (EHPMF), (g) Sections 6 to 7 of the EHPMF, (h) Sanguankou Fault (SGKF), (i) Zhuozhi West Fault (ZZWF), (j) Dengkou–Benjing Fault (DKBJF), (k) Bayanhaote Fault (BYHF), (l) Chonggang–Luhuatai Fault (CGLHTF), (m) the Niushoushan Fault (NSSF).

The Hetao seismic zone is characterized by a unified left-lateral strike-slip fault mechanism predominantly driven by horizontal movement. The Bayannula Mountain Fault exhibits non-prominent strain accumulation and a non-uniform distribution of fault slip deficit. Its dip-slip rate reaches up to 5 mm/yr at depths below 15 km in the crust, and the local strike-slip rate at the southern end is relatively high. The slip deficit and locking distribution are uneven, with a slip deficit (ranging from 0 to 1 mm/yr at a depth of approximately 18 km, and they reach up to 5 mm/yr in the northern section where most areas are locked).

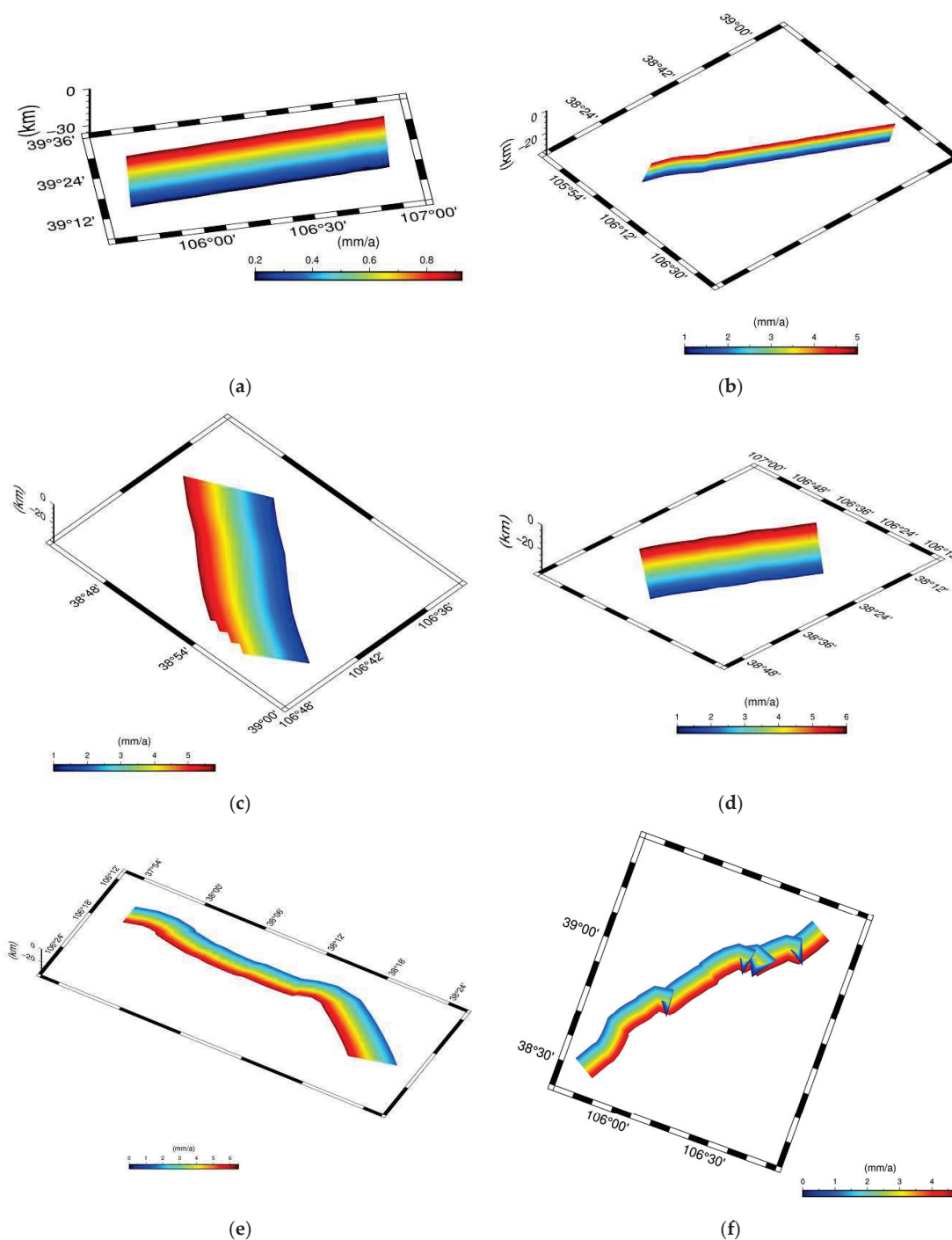


Figure 7. Cont.

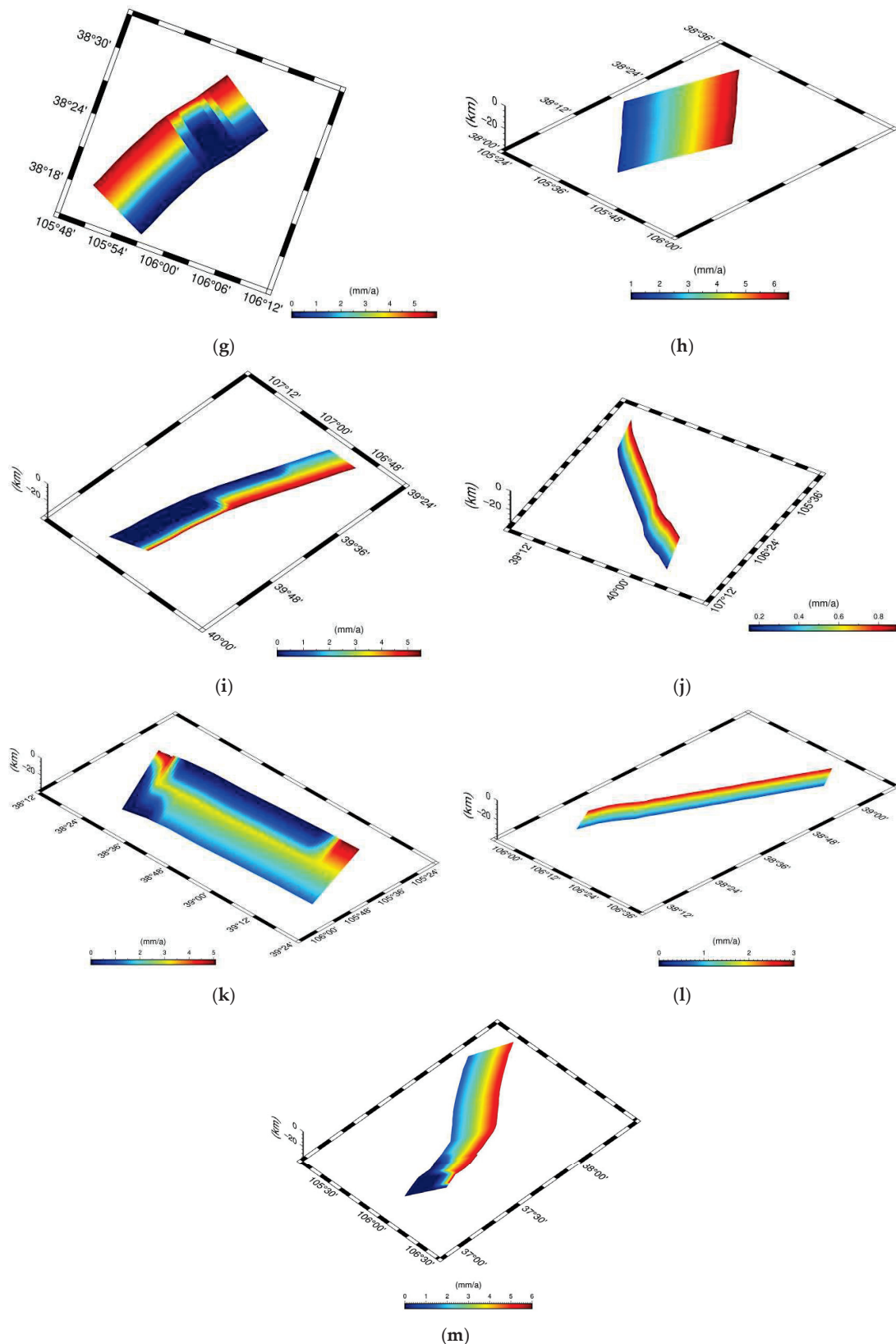


Figure 7. Deficit of fault slip rate during the period from 1999 to 2024 for (a) Zhengyiguan Fault (ZYGf), (b) Yinchuan–Pingluo Fault (YCPLF), (c) Section 1 of the Yellow River–Lingwu Fault (YRLWF), (d) Section 2 of the YRLWF, (e) Sections 3 and 4 of the YRLWF, (f) Sections 1 to 5 of the Eastern Helan Piedmont Fault (EHPMF), (g) Sections 6 to 7 of the EHPMF, (h) Sanguankou Fault (SGKF), (i) Zhuozi West Fault (ZZWF), (j) Dengkou–Benjing Fault (DKBJF), (k) Bayanhaote Fault (BYHF), (l) Chonggang–Luhuatai Fault (CGLHTF), (m) the Niushoushan Fault (NSSF).

The non-locking depth of the Dengkou–Benjing Fault is approximately 18 km. Its strike–slip rate is slightly higher in Etuoke Banner and Hangjin Banner, and it varies inversely with the fault strike at the junction of Zuoqi and Etuoke Banner, with the highest rate of 0.4 mm/yr at the westernmost end in Zuoqi. The slip deficit is not pronounced, and it demonstrates obvious left-lateral strike–slip strain accumulation with a low overall slip deficit, indicating a low risk.

The Western Zhuozzi Fault accumulates normal-fault dip–slip and left-lateral strike–slip strain. The distribution of slip deficit and locking depth in the southwest (Wuhai) is non-uniform. The locking degree increases from the southwest to the northeast, and the non-locking depth in Wuhai is 20 km. The slip deficit in Wuhai ranges from 3.8 to 5.3 mm/yr. The dip–slip rate is higher in Wuhai, and the strike–slip rate in Etuoke Banner reaches 1.5 mm/yr. The Otok Banner section may be locked, and there is locking and thrust strain accumulation in the Etuokeqi segment.

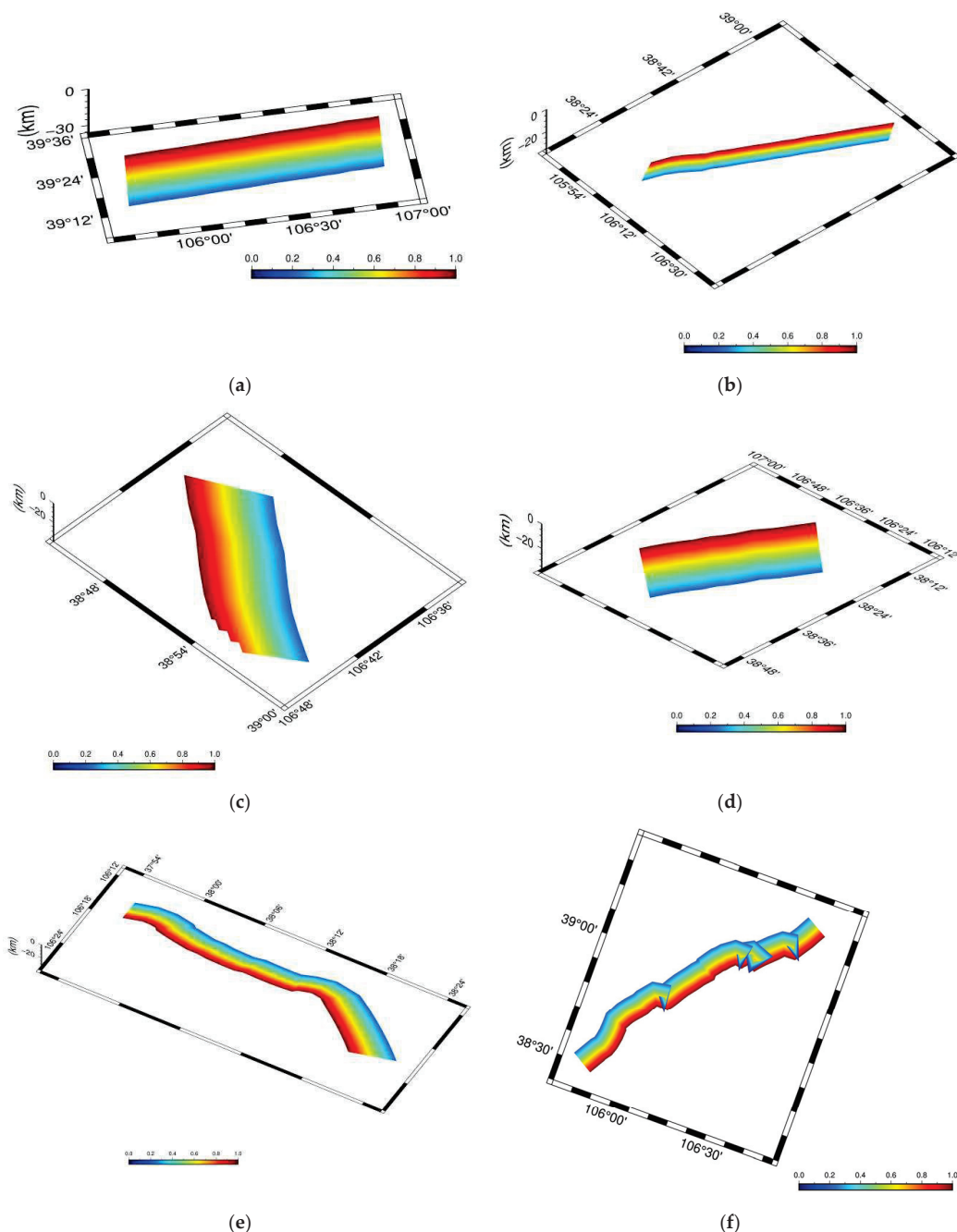


Figure 8. Cont.

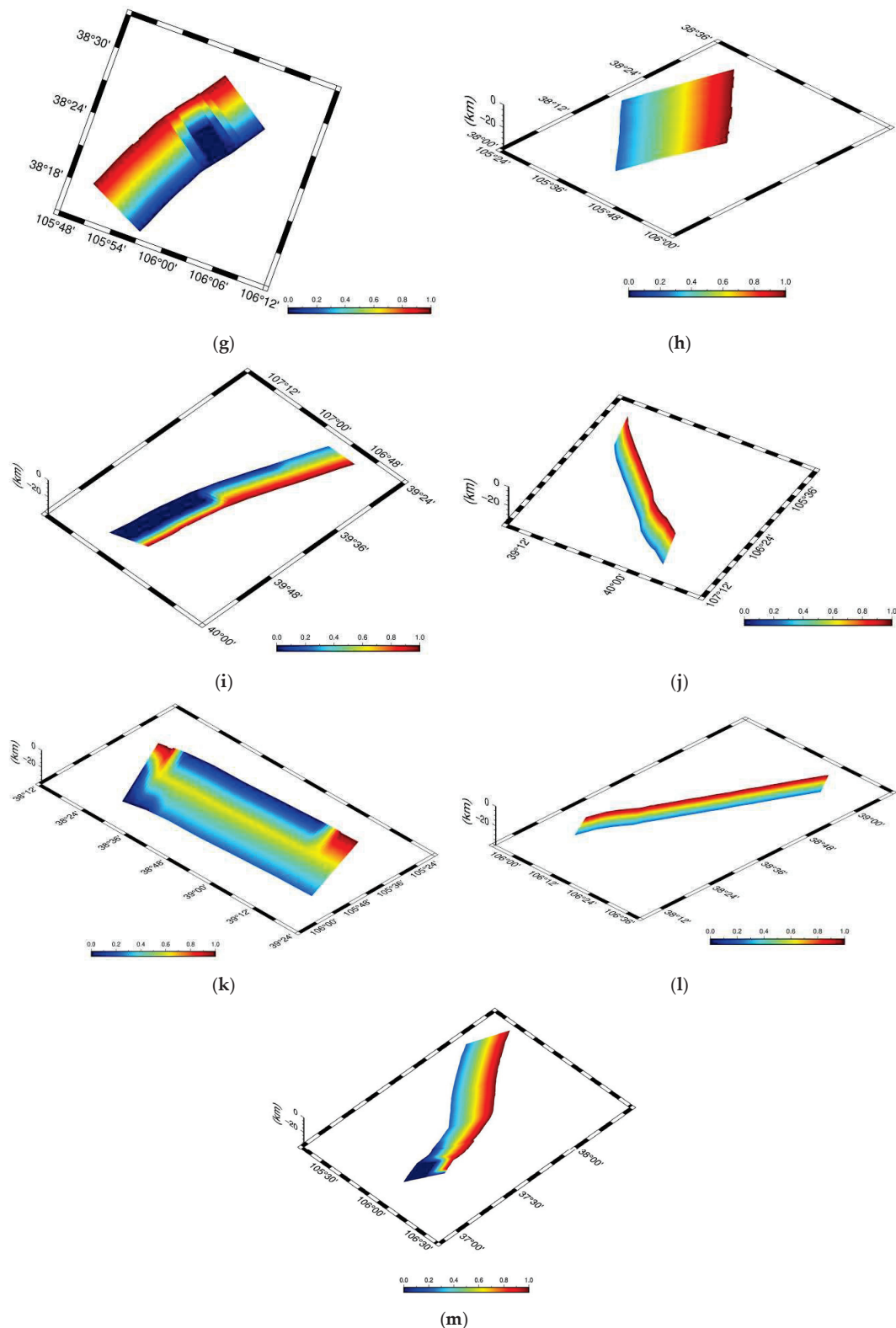


Figure 8. Fault locking rate during the period from 1999 to 2024 of (a) Zhengyiguan Fault (ZYGF), (b) Yinchuan–Pingluo Fault (YCPLF), (c) Section 1 of the Yellow River–Lingwu Fault (YRLWF), (d) Section 2 of the YRLWF, (e) Sections 3 and 4 of the YRLWF, (f) Sections 1 to 5 of the Eastern Helan Piedmont Fault (EHPMF), (g) Sections 6 to 7 of the EHPMF, (h) Sanguankou Fault (SGKF), (i) Zhuozhi West Fault (ZZWF), (j) Dengkou–Benjing Fault (DKBJF), (k) Bayanhaote Fault (BYHF), (l) Chonggang–Luhutai Fault (CGLHTF), (m) the Niushoushan Fault (NSSF).

In Shizuishan, the non-locking depth of the ZHENGYIGUAN fault is about 18 km. The slip deficit is slightly higher, and both the dip-slip and strike-slip rates increase slightly. The fault locking is weak, and the slip deficit is not large.

For segment 1 (North) of the Helan East Piedmont Fault, the non-locking depth is approximately 18 km. The local dip-slip rate at the fault strike adjustment is relatively high. The strike-slip rate is slightly higher in the northeastern segment and locally reversed (-1.0 mm/yr) in Yinchuan. The local strike-slip and dip-slip rates of the northern segment are 5 mm/yr, indicating relatively high activity. The slip deficit reaches 5 mm/yr, and the locking is not obvious.

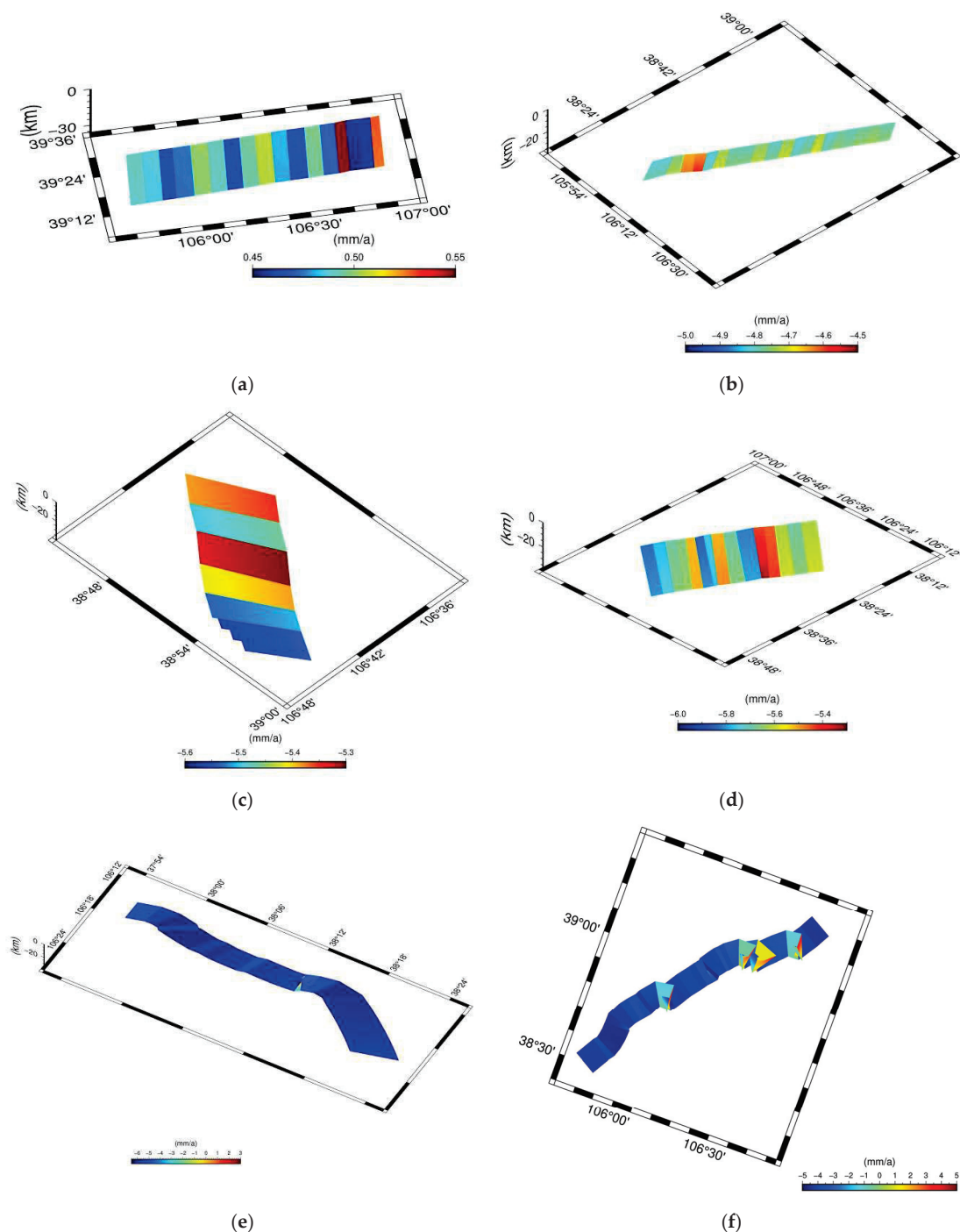


Figure 9. Cont.

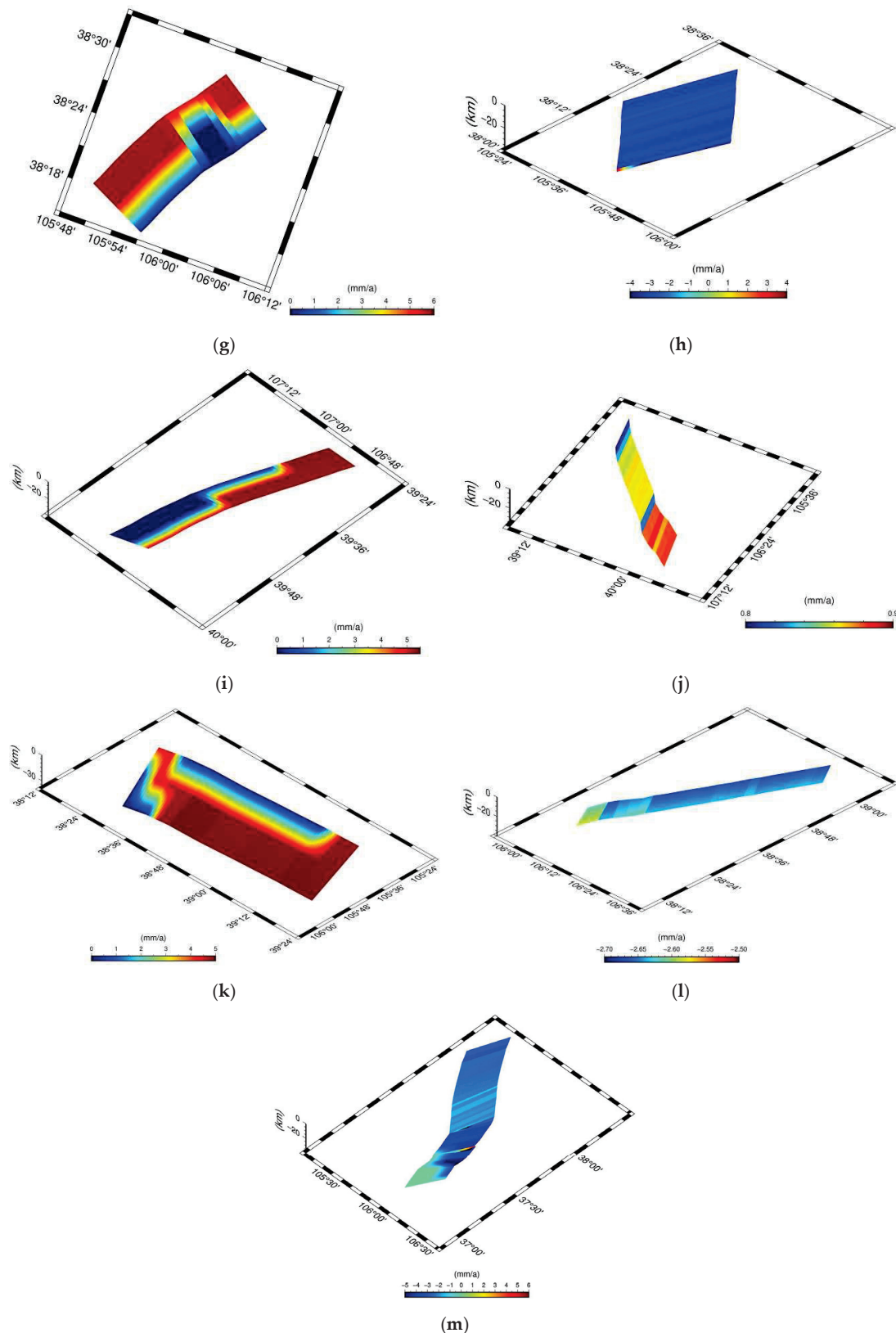


Figure 9. Fault dip–slip rates from 1999 to 2024 of (a) Zhengyiguan Fault (ZYGF), (b) Yinchuan–Pingluo Fault (YCPLF), (c) Section 1 of the Yellow River–Lingwu Fault (YRLWF), (d) Section 2 of the YRLWF, (e) Sections 3 and 4 of the YRLWF, (f) Sections 1 to 5 of the Eastern Helan Piedmont Fault (EHPMF), (g) Sections 6 to 7 of the EHPMF, (h) Sanguankou Fault (SGKF), (i) Zhuozi West Fault (ZZWF), (j) Dengkou–Benjing Fault (DKBJF), (k) Bayanhaote Fault (BYHF), (l) Chonggang–Luhuatai Fault (CGLHTF), (m) the Niushoushan Fault (NSSF).

Segment 2 (South) of the Helan Mountain Eastern Foot Fault traverses Yongning and Yinchuan northward and connects with the Sanguankou Fault. The southern-end strike-slip rate is close to 0 mm/yr, with local strong locking below 15 km in Yinchuan. Other areas have a non-locking depth of 20 km and a local strike-slip rate of 6 mm/yr, suggesting high activity.

The Sanguankou Fault connects to the southern Helan East Fault in the north. It has a non-locking depth of up to 18 km, uniform dip-slip and strike-slip rates, shows general activity, and has no locking.

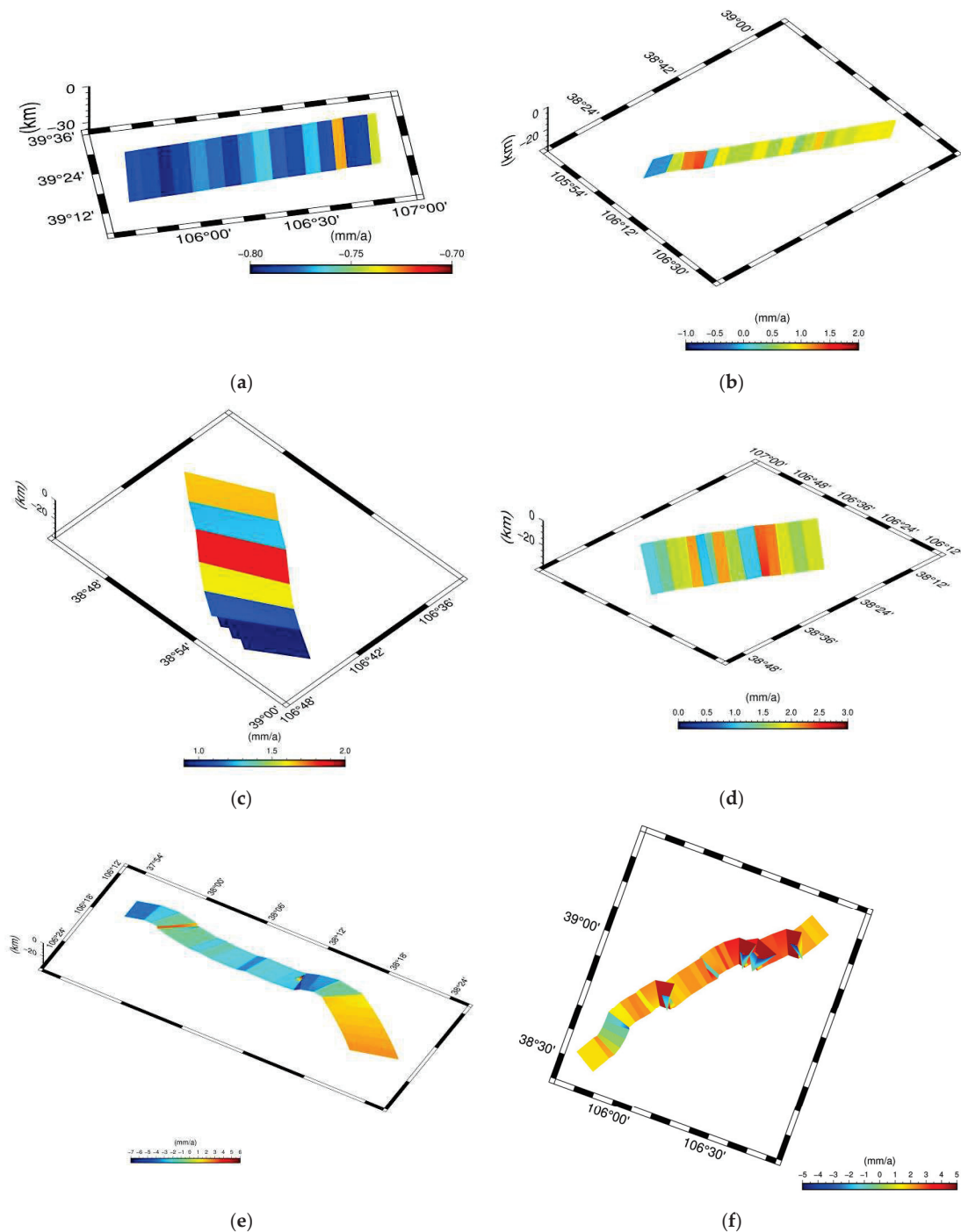


Figure 10. Cont.

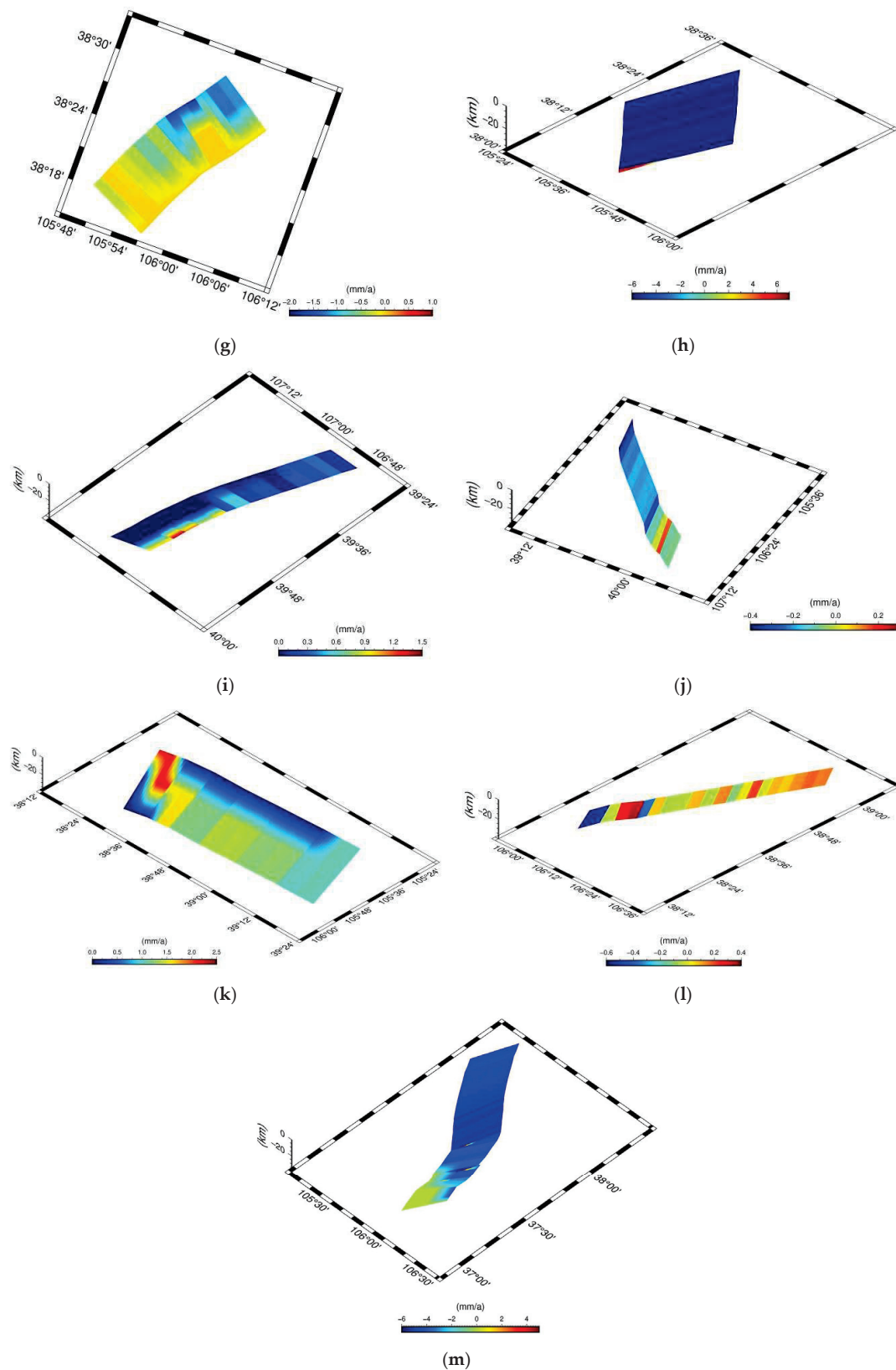


Figure 10. Strike-slip rates of faults from 1999 to 2024 for (a) Zhengyiguan Fault (ZYGF), (b) Yinchuan-Pingluo Fault (YCPLF), (c) Section 1 of the Yellow River-Lingwu Fault (YRLWF), (d) Section 2 of the YRLWF, (e) Sections 3 and 4 of the YRLWF, (f) Sections 1 to 5 of the Eastern Helan Piedmont Fault (EHPMF), (g) Sections 6 to 7 of the EHPMF, (h) Sanguankou Fault (SGKF), (i) Zhuozi West Fault (ZZWF), (j) Dengkou-Benjing Fault (DKBJF), (k) Bayanhaote Fault (BYHF), (l) Chonggang-Luhuatai Fault (CGLHTF), (m) the Niushoushan Fault (NSSF).

The Chonggang–Luhuatai Buried Fault extends southward across Huinong, Pingluo, Helan, and Yinchuan. It has a non-locking depth of about 20 km, no locking, a low dip–slip rate, a slightly higher dip–slip rate in Yinchuan, and an unremarkable strike–slip.

In Yinchuan, the Yinchuan–Pingluo buried fault traverses Pingluo, Helan, Yinchuan, and Yongning southward, with an unlocked depth of approximately 20 km. The dip–slip rate varies minimally, has a relatively low absolute value, and is slightly higher in Yinchuan, and the strike–slip movement is unremarkable.

Section 1 (North) of the Yellow River–Lingwu fault is located in northern Taole County, Shizuishan City, with an unlocked depth of about 18 km. The dip–slip rate varies slightly and is slightly higher at the southern end. The strike–slip rate has a small absolute value and is slightly higher in the middle (2.0 mm/yr). The northern segment is active in dip movement with thrust activity and no obvious locking.

The second (middle) segment of the Yellow River–Lingwu fault extends northward across Taole, Helan, and Yinchuan, with an unlocked depth of around 18 km. The variance of the dip–slip rate is small and is slightly higher at the northern end. The strike–slip rate has a small absolute value but is relatively higher in the middle. Similarly to the northern section, it has thrust activity and no obvious locking.

Section 3 (South) of the Yellow River–Lingwu fault is situated in north–central Lingwu City, trending northeast, with an unlocked depth of about 18 km. The variance of the dip–slip rate is small, and the strike–slip rate is higher in the south–central part. It is more active in strike–slip movement than the first and second segments, has thrust movement, and there is no obvious locking.

In Zhongwei, the second-largest city in Ningxia, the Niushoushan Fault Northern Section traverses Qingtongxia and Wuzhong in a northwestward direction. Its southern end (central Wuzhong) is locked, with a non-locked depth of approximately 18 km. The remaining segments display strike–slip and thrust kinematics. The Niushoushan Fault Southern Section intersects Azuoqi, Qingtongxia, Wuzhong, and Tongxin, also trending northwest. Its southern end (central Tongxin) is locked, with a non-locked depth of about 18 km, and strike–slip and thrust displacements are observed in other parts.

The Guanganling Fault crosses Jingtai, Azuoqi, and Zhongwei in an east–west orientation. Its western end (within Jingtai) is locked, with a non-locked depth of roughly 20 km. In the non-locked section, strike–slip and dip–slip motions are relatively inactive.

The Xiangshan–Tianjingshan Fault Section 1 (Western) is situated west of Zhongwei, striking east–west and connecting with the Guanganling Fault to the west. Its western end is locked. In the non-locked section, dip–slip movement is inactive, while strike–slip movement persists to a depth of 40 km.

The Xiangshan–Tianjingshan Fault Section 2 (Middle) is located in the central–eastern part of Zhongwei, with an east–west strike. Its strike–slip and dip–slip movements are active, and the non-locked depth is 18 km.

The Xiangshan–Tianjingshan Fault Section 3 (Eastern) crosses Zhongwei, Zhongning, Tongxin, and Haiyuan, trending northwest and connecting with the Miaoshan Fault to the east. The western end (east of Zhongwei) is partially locked. The non-locked depth is 18 km, with active dip–slip and strike–slip displacements and an uneven distribution of dip–slip rates. The dip–slip rate at the eastern end is relatively higher, whereas the strike–slip rate is relatively lower.

The Qingshuihe Fault crosses Zhongning and Tongxin County, striking northwest. The non-locked depth is 18 km, with distinct dip–slip and strike–slip displacements and no locking phenomenon.

The Yantongshan Fault strikes northwest, with a non-locked depth of 20 km. The dip-slip and strike-slip rates are unevenly distributed but exhibit regular variations. The dip-slip rate in the northwest section is low, while the strike-slip rate is high.

5. Discussion

This study summarizes the results of the third section. By analyzing the velocity field, strain rate field, and cross-fault baseline variations, along with the slip rates and locking rates of active faults computed through DEFNODE, it pinpoints potentially hazardous areas.

5.1. Well-Defined Outcomes for the Determination of Seismic Hazard

Our objective is to evaluate earthquake risk by defining outcomes that are independent of model and coordinate system selection. These include the deceleration of the velocity difference on both sides of the fault (excluding the increase in velocity difference) identified through periodic comparisons of velocity fields, the alterations in surface and linear strain across the entire study area, and the time-series variations near the cross-fault baseline length.

5.1.1. Locking Inferred from Velocity Field

A comparison was carried out between the contemporary sliding rates and those of major active faults during the late Quaternary period across the Chinese mainland [82]. The research findings demonstrated that, in most instances, the sliding rates of these faults across different time intervals were of the same order of magnitude or even highly comparable. Consequently, it was proposed that the current fault sliding rate could represent the long-term background fault sliding rate.

Generally, in recent years, an upward trend has been detected in the east–west (EW) velocity of the Global Navigation Satellite System (GNSS) in the vicinity of the eastern foothills of the Helan Mountain Fault, Bayannula Mountain Fault, Yellow River–Lingwu Fault, Yinchuan–Pingluo Fault, Xiangshan–Tianjingshan Fault, Luoshan Fault, Yantongshan Fault, and Sanguankou Fault.

Based on the station velocities within a range of 50–100 km on either side of the active faults in the study area from 1999 to 2024, particularly in the vicinity of the faults, a preliminary inference can be drawn regarding the recent activity trends of these fault segments. The majority of the active faults in the northern part of Ningxia in Inner Mongolia are normal faults. The dip-slip and strike-slip rates within a 50 km radius on both sides of the Bayannula Mountain Fault are relatively low. The dip-slip and strike-slip rates within the vicinity of the Dengkou–Benjing fault exhibit a sharp decline, indicating potential locking. No significant strike-slip and dip-slip movements were observed in the Western Zhuozi Fault.

This study primarily focuses on Shizuishan and Yinchuan, situated in the northern region of Ningxia. The analysis in the aforementioned table indicates that the Zhengyiguan Fault, a concealed fault at the northern tip of Ningxia, exhibits distinct dextral strike-slip characteristics, with a strike-slip rate of 2–5 mm/year and a relatively low dip-slip rate. The Eastern Helan Piedmont Fault is segmented into segments 1 through 7 from south to north. Within a 50 km radius of this fault, the extensional rate is 2–7 mm/year, while also demonstrating dextral strike-slip behavior. Significantly, the strike-slip rates of segments 1 and 3 across the fault experience a sharp decline followed by an increase, which might result in a locked state. The dip-slip rate along the Yellow River–Lingwu Fault ranges from 1 to 7 mm/year, and the parallel rate near the fault shows a sudden decrease and then a subsequent increase, suggesting a potential locked condition. The

Sanguankou Fault has a limited number of monitoring points on both sides, implying a left-lateral strike-slip movement. At the middle and southern ends of the Niushoushan Fault, a buried fault in Wuzhong, a city in central Ningxia, there is a phenomenon where both the strike-slip and dip-slip rates increase and then gradually decrease across the fault, indicating strain accumulation.

In summary, a phased comparison of the velocity fields in northern Ningxia reveals that the velocity difference across certain faults has diminished, which implies a potential locking risk. In Inner Mongolia, there exists a non-significant velocity disparity in the vicinity of the Xizhouzi Mountain Fault. In northern Ningxia, the movement of the Zhengyiguan Fault in Shizuishan City is not prominent. The strike-slip movement of the fourth (middle) and seventh (northernmost) segments of the Helan Mountain front fault in the western Yinchuan Basin has decelerated. Regarding the Lingwu Fault on the Yellow River in the eastern Yinchuan Basin, the tilting and sliding speeds of segment 1 exhibit a slight decline, the tilting and strike-slip speeds of segment 2 experience a slowdown, and segments 3 and 4 display a slow tilting motion. Additionally, there is no evident strike-slip movement on the Sanguankou Fault in the southern Helan Mountain. The strike-slip movement of the northern segment of the hidden fault of Niushoushan (passing through Wuzhong City) accelerates, while the dip-slip rate decreases, and the dip-slip velocity of the southern segment of the Niushoushan Fault has slightly decreased.

5.1.2. Locking Inferred from Strain Rate

Under inheritance conditions, the intensification and attenuation of local strain in the magnitude direction constitute a critical basis for evaluating stress accumulation and active locking near a fault. Large-scale earthquakes are most likely to occur in regions where local deformation is intensifying within the context of significant deformation. These regions also represent the deformation locking zones for in situ stress accumulation and correspond to the marginal gradient zones of local horizontal deformation anomalies.

The large-scale local strain rate field in the northern segment of the North–South Seismic Belt [1,4] reveals that the NE- and NS-oriented linear strains, the primary and secondary shear strains, tensile strain, and surface strain in the northern part of the structural area of the Longzhong Basin have significantly declined. In contrast, the tensile and maximum shear strains in the northern Longzhong Basin have exhibited an increase, and the surface strain has undergone a reversal. The primary frontal shear strain within the structural area of the central Longzhong Basin remains at a relatively high level. Additionally, the frontal EW-oriented linear, tensile, and maximum shear strains in the northern Alxa Block have increased, and the primary and maximum shear strains continue to be elevated.

Generally, the principal compressive stress in Ningxia and its adjacent regions is predominantly oriented from the northwest to the north, with a relatively uniform alignment. The strain field is primarily characterized by compression. The southwestern part of Ningxia is situated in a high-tension area, and the extensional strain rate has been relatively high since 1998. Areas such as the eastern foothills of the Helan Mountain and certain faults are located in high-pressure strain zones. The northern part of Ningxia and its neighboring areas are generally within high-shear strain zones, and the shear strain rate has increased but has not exceeded the high level recorded in 1998.

Compressive strains are present in some faults, while strike-slip strains are observed in others. The Yantongshan and Xiangshan–Tianjingshan Faults may exhibit inclined slip strains. Locally, the Zhengyiguan Fault does not accumulate strike-slip strain, demonstrates a stable compressive strain tendency, has weak locking, and a low overall slip deficit, indicating a relatively low seismic risk.

Within the study region, the tensile strain rate spans from 0.2 to 1.6×10^{-8} /yr, signifying a phase of relatively elevated values since 1999. The Sanguankou Fault, Xiangshan–Tianjingshan Fault, and Niushoushan Fault are situated within a high-value zone of 1 to 1.8×10^{-8} /yr. The compressive strain rate fluctuates between -1.6 and -0.6×10^{-8} /yr, failing to attain the high values witnessed from 2009 to 2014 [4]. The shear strain in northern Ningxia and its adjacent regions is 0.4 to 1.6×10^{-8} /yr, and the shear strain rates of the Sanguankou Fault, Niushoushan Fault, Luoshan Piedmont Fault, and Xiangshan–Tianjingshan Fault are 1 to 2×10^{-8} /yr. These regions have reached the high value of shear strain since 1999, comparable to the period from 2009 to 2014. Presently, the area strain rate in northern Ningxia and its adjacent regions is at a low value of -0.2 to 0.2×10^{-8} /yr, which falls within the normal range since 1999 [4]. Based on the strain rate, it can be deduced that the dip-slip of the Zhengyiguan Fault and Dengkou–Benjing Fault is not prominent, suggesting potential locking. Likewise, the strike-slip of the Yellow River–Lingwu Fault, Bayanwulashan Fault, Niushoushan Fault, Sanguankou Fault, Xiangshan–Tianjingshan Fault, and the southern segment of the Luoshan Piedmont Fault appears to be insignificant, indicating potential locking as well.

In conclusion, the strain rate within the study area is characterized as follows: The Bayannula Mountain Fault, Dengkou–Benjing Fault, Helan Mountain East Foot Fault, Yinchuan–Pingleuo Fault, Huanghe–Lingwu Fault, Xiangshan–Tianjingshan Fault, Yantongshan Fault, and Sanguankou Fault demonstrate gradient variations in compressive strain, with distinct dip-slip characteristics. The Bayannula Mountain Fault, Dengkou–Benjing Fault, Zhengyiguan Fault, Helan Mountain East Foot Fault, Yinchuan–Pingleuo Fault, Huanghe–Lingwu Fault, Sanguankou Fault, and Luoshan Mountain Front Fault display gradient changes in shear strain rate. Among these, notable strike-slip movements are observed in the Dengkou–Benjing Fault, Zhengyiguan Fault, and Huanghe–Lingwu Fault. The linear and shear strain rates of the Helan Mountain eastern foothills, Sanguankou, Niushoushan, Zhengyiguan, the southern section of the Yellow River, Luoshan Mountain front, and Xiangshan–Tianjingshan Faults are relatively high.

5.1.3. Locking Inferred from Baseline Variation

Over the past decade, the time series of the baseline length of the Zhengyiguan Fault exhibits a stable compressive trend. Simultaneously, the Yinchuan–Pingleuo and Yellow River–Lingwu Faults display dip-slip characteristics. The Yinchuan–Zhongwei and Yinchuan–Haiyuan baselines have transitioned from a compressive state to a stable one, and the Wuhai–Ulat rear banner baseline has shifted from a tensile state to a stable state. It is hypothesized that the Niushoushan and Dengkou–Benjing Faults may be partially locked, and the Niushoushan, Yantongshan, and Xiangshan–Tianjingshan faults may be undergoing dip-slip strain or partial locking.

That is to say, regarding the time-series interpretation near the cross-fault baseline length, the NXYC–NZZW baseline traverses the Niushoushan Fault, and the NMWH–NMWT baseline traverses the Dengkou–Benjing Fault. Both baselines have transitioned from a state of continuous compression to a stable state without significant tilting displacement, indicating that they may be in a locked state. The NXYC–NHV baseline intersects the Niushoushan, Yantongshan, and Xiangshan–Tianjingshan Faults. These faults have shifted from continuous compression to a stable state and may exhibit tilting displacement or be locked.

5.2. Model-Dependent Interpretations for Seismic Hazard Determination

Simultaneously, despite relying on the ITRF2014 fixed-framework velocity field, we utilize DEFNODE solutions, which encompass geophysical inversion regularization and

weighting schemes, optimal solution indicators, and other outcomes derived from the selection of fracture sliding rates, locking rates, etc., as explanations for model dependencies.

Based on the slip rate, slip deficit, and locking rate inverted by the DEFNODE method, it is hypothesized that the slip deficits of the Dengkou–Benjing and Tun–Benjing Faults in Inner Mongolia are relatively low and uniformly distributed. The Tuokeqi section of the Western Margin Fault of Zhuozi may be in a locked state. The slip deficit at the southwestern end (within Wuhai) is non-uniform, indicating potential locking, and the degree of locking increases from the southwest to the northeast. The slip deficit along the Bayannula Mountain Fault is non-uniform, and areas other than the northern section may be locked, with a deficiency of slip in the northern segment. Most of these faults exhibit deep-seated locking. The ZhengYiguan Fault around the Yinchuan Basin has a uniform distribution of slip deficit, weak locking, and a low overall slip deficit. The northern segment of the Helan Eastern Piedmont Fault has a uniform distribution with shallow overall locking. There is local strong locking below 15 km in the southern segment of Yinchuan. Local locking is observed in the northern and southern sections of the Niushoushan Fault in Zhongwei City (central Wuzhong City) and Tongxin County (central Tongxin County), the western end of the Cangcangling Fault (within Jingtai County), and the western and eastern ends of the Xiangshan–Tianjingshan Fault (eastern Zhongwei City).

5.3. Implications for Research

The locking rate of active faults is a critical parameter reflecting the energy accumulation of faults. By establishing a “monitoring–assessment–warning” technical system, it is feasible to quantitatively improve the seismic hazard assessment. This quantitative analysis offers guidance for the optimization of seismic hazard evaluation and monitoring strategies. Meanwhile, it is expected to accumulate accurate real-time, high-frequency, three-dimensional data and integrate it with GPS vertical and fixed-point deformation data for more in-depth research. The following are some further research directions:

- (1) Short-term forecasting requires high-frequency and long-term encrypted data. The changes in the velocity and strain rate of two active fractures are a long-term process, so long-term continuous monitoring should be strengthened. A long-term monitoring database should be established, and the monitoring data should be analyzed and processed in real time to detect abnormal changes in speed and strain rate promptly. Meanwhile, big data and artificial intelligence technologies should be used to conduct dynamic analysis of the monitoring data, establish earthquake hazard assessment models, adjust the assessment results in real time according to data changes, and improve the timeliness and accuracy of earthquake prediction.
- (2) A more comprehensive integration of GNSS and remote sensing technologies and a comparison between GNSS and large-scale fixed-point deformation observation techniques (especially tunnel strain gauges) are required. To precisely acquire information on active fault velocity and strain rate variations, a monitoring network integrating multiple data sources (e.g., GNSS continuous observation, borehole strain gauges, and geodetic leveling) should be established. GNSS continuous observation can acquire horizontal displacement field data, borehole strain gauges can record deep-seated strain dynamics, and geodetic leveling can monitor surface elevation changes. By applying multi-source data fusion technology, the accuracy of strain observation is enhanced to 10^{-9} , enabling precise capture of crustal strain and providing comprehensive and accurate data for earthquake hazard assessment. Moreover, the computation of vertical components at GNSS stations is prone to non-structural interference, leading to significant errors. It can be integrated with regional leveling data to obtain vertical anomaly information by analyzing vertically high-gradient areas.

- (3) Future research ought to concentrate on achieving breakthroughs in deep-seated locking state detection technologies and develop dynamic prediction models for locking rates by leveraging machine learning, thereby offering more scientific decision-making support for earthquake disaster prevention and mitigation.

6. Conclusions

This research centers on earthquake risk assessment. The results that are independent of model selection and coordinate system selection are considered as the definitive results. These are integrated with the outcomes of DEFNODE calculation, such as fault slip rate and locking rate, to provide explanations for model dependency. The joint inference reveals that the southern extremity of the northern segment of the Niushoushan–Luoshan Fault (located in the central region of Wuzhong City) and the southern extremity of the southern segment (situated in the central part of Tongxin County) are locked. The western extremity of the western segment and the western extremity of the eastern segment of the Xiangshan–Tianjingshan Fault Zone (east of Zhongwei City) display locking behavior. There is locking at the western extremity of the Cangcangling Fault (within Jingtai County). The southern part of the Helan Eastern Piedmont Fault encounters local obstruction at depths below 15 km in Yinchuan. The locking distribution of the Bayannula Mountain Fault is non-uniform. In the northern segment, the overall slip deficit only attains 5 mm/yr, and most parts of the fault are locked at depth. The locking degree of the Zhuozi Western Margin Fault gradually rises from the southwest (within Wuhai) to the northeast, and the fault is locked in the Etuohe Banner section.

Author Contributions: The original observational data of the “land network” continuous stations and mobile stations were copied by N.Z. at the Crustal Engineering Center of the China Seismological Bureau. Additionally, N.Z. downloaded the original observational data of the IGS stations serving as frame points and the global IGS station h-file. N.Z. completed tasks such as data pre-processing, gamit/globk solution, and adjustment. N.Z. also carried out subsequent data processing, including work on the velocity field, velocity projection, strain rate field, baseline calculation, geophysical inversion, etc., as well as mapping. X.F. and G.N. offered technical and project-related financial support for the writing of the articles. Meanwhile, J.W. provided certain support for the GNSS geophysical inversion algorithm. All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by N.Z. and Xiwei Fan. The first draft of the manuscript was written by N.Z., and all authors commented on previous versions of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was jointly funded by the Key R&D Program Project of Ningxia Hui Autonomous Region (2024BEG02033), the National Nonprofit Fundamental Research Grant of China, the Institute of Geology, China Earthquake Administration (IGCEA2106), and the National Key Research and Development Program of China (2022YFC3003700). The Sponsorship Agencies have absolutely no influence on the study design, data collection, analysis, interpretation, report writing, or the decision to submit the article for publication.

Institutional Review Board Statement: This manuscript does not contain any studies with human or animal participants. There are no human participants in this manuscript, and informed consent is not required.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets generated during or analyzed during the current study are available from the corresponding author on reasonable request. A part of the data is derived from the terrestrial network stations of the China Earthquake Administration and the GNSS data product service platform for high-precision data processing (ftp.cgps.ac.cn).

Acknowledgments: The authors sincerely express their gratitude to Zhao Bin from the Wuhan Institute of Seismology, China Earthquake Administration, for providing certain GNSS raw data. The velocity projection, strain rate field, baseline calculation DOS program were provided by Wu Yanqiang from the First Surveying and Mapping Center of the China Earthquake Administration. Gratitude is hereby expressed. Additionally, the authors are thankful to Gan Weijun from Institute of Geology of China Earthquake Administration for his substantial support of this article.

Conflicts of Interest: The authors declared that they had no potential conflicts of interest with respect to the research, authorship, and/or publication of this manuscript. They have no relevant financial or non-financial interests to disclose.

Abbreviations

The following abbreviations are used in this manuscript:

BYHF	Bayanwula Mountain Fault
LSPMF	Langshan Piedmont Fault
HJRBF	Hanggin Rear Banner Fault
LHF	Linhe Fault
NOF	Northern Ordos Fault
NNSF	Niushoushan Fault
YRLWF	Yellow River–Lingwu Fault
ZYGF	Zhengyiguan Fault
YCPLF	Yinchuan–Pingluo Fault
EHPMF	Eastern Helan Piedmont Fault
DKBJF	Dengkou–Benjing Fault
SGKF	Sanguankou Fault
LSEPMF	Luoshan East Piedmont Fault
HYF	Haiyuan Fault
YTSF	Yantongshan Fault
XSTJSF	Xiangshan–Tianjingshan Fault
LPMF	Liupan Mountain Fault
YWMF	Yunwu Mountain Fault
WQNMF	North Margin Fault of West Qinling
MXMF	Maxian Mountain Fault
HNHYGF	Huining–Yigang Fault
STF	Serteng Fault
ZZWF	Zhuozi West Fault
BYHF	Bayannula Mountain Fault
CGLHTF	Chonggang–Luhuatai Fault
XGMF	Xiaoguan Mountain Fault
BDZF	Bai Dunzi Fault
TMF	Tiger Mountain Fault
WQMNF	Fault on the Northern Edge of the Western Qinling Mountains
LXLJF	Lixian Luojiabao Fault
LTDCF	Lintan Dangchang Fault
GGDSF	Guanggaishan Dieshan Fault
LDJLF	Liangdang Jiangluo Fault
QNMF	Qinling Northern Margin Fault
WHF	Weihe Fault
SDF	Shangdan Fault
WLMFF	Wulashan Piedmont Fault
STPF	Serteng Piedmont Fault
DQSPF	Daqingshan Piedmont Fault
GNSS	Global Navigation Satellite System
ACO	Ant Colony Optimization

PSO	Particle Swarm Optimization
CMONOC	China Crustal Movement Observation Network
CMTEMN	China Continental Tectonic Environment Monitoring Network
IGS	International GNSS Service

References

- Zhang, N.; Nie, G.; Fan, X.; Wang, J.; Yu, S.; Sun, J.; Xie, H. Estimating the current movement rates and slip losses of major active faults in the northern segment of the North-South Seismic Belt based on GNSS velocity field data. *Earthquake* **2022**, *42*, 25–46.
- Zhang, N.; Wang, J.; Yu, S.; Sun, J.; Xie, X. Current block strain rates and slip deficit of active faults along the northeastern margin of the Qinghai-Tibet Plateau. *J. Geod. Geodyn.* **2023**, *43*, 661–668.
- Zhang, N.; Xu, W.; Wang, J.; Xie, H.; Li, M.; Ma, R. Study on dynamic variation characteristics of strain rate field in the northern segment of the North-South seismic belt. *Prog. Geophys.* **2020**, *35*, 1724–1735.
- Zhang, N.; Xu, W.; Wang, J.; Xie, H.; Li, M.; Ma, R. Study on crustal motion field and strong earthquake trends along the northeastern margin of the Qinghai-Tibet Plateau based on GNSS. *China Earthq. Eng. J.* **2022**, *44*, 649–660.
- Ji, L.; Zhu, L.; Li, N.; Wang, C. A review of fault movement studies based on geodetic observations. *J. Geod. Geodyn.* **2017**, *37*, 771–776.
- Li, T.; Zhu, Y.; Yang, Y.; Xu, Y.; An, Y.; Zhang, Y.; Feng, S.; Hui, Y.; Yang, J. Comprehensive utilization of multiple crustal deformation observation data to calculate the current slip rate of the Xianshuihe fault zone. *Chin. J. Geophys.* **2019**, *62*, 1323–1335.
- Yong, Q. Study on the Impact Characteristics of Seismic Fault Slip Inversion Using InSAR and GPS Geodetic Deformation Data. Master's Thesis, Southwest Jiaotong University, Chengdu, China, 2017.
- Li, N.; Kang, S.; Zhu, L. Study on current closure degree and seismic risk of the Honghe fault zone based on GPS data. *Geod. Earth Dyn.* **2019**, *39*, 700–705.
- Li, N.; Zhu, L.; Liu, L. Analysis of current closure degree and seismic risk of the Lijiang-Xiaojinhe fault zone. *J. Seismol. Res.* **2018**, *41*, 244–250.
- Thingbaijam, K.; Mai, P.; Goda, K. source scaling laws. *Bull. Seismol. Soc. Am.* **2017**, *107*, 2225–2246. [CrossRef]
- Wu, Y.; Jiang, Z.; Yang, G.; Fang, Y.; Wang, W. Method and Application of Global GPS Strain Field Calculation Using Least Squares Configuration on Spherical Surfaces. *Chin. J. Geophys.* **2009**, *52*, 1707–1714.
- Lv, Z.; Wu, J.; Meng, G.; Qiao, X.; Xu, K. Analysis of spatial distribution characteristics of ground strain along the eastern margin of the Qinghai-Tibet Plateau. *J. Geod. Geodyn.* **2014**, *34*, 19–23.
- Wang, Q.; Zhang, P.-Z.; Freymueller, J.T.; Bilham, R.; Larson, K.M.; Lai, X.A.; You, X.; Niu, Z.; Wu, J.; Li, Y. Present-day crustal deformation in China constrained by global positioning system measurements. *Science* **2001**, *294*, 574–577. [CrossRef]
- Wang, Q.; Zhang, P.; Niu, Z.; Lai, X.; Li, Y.; Zhu, W.; Liu, J. Chinese mainland current crustal movements and tectonic deformation. *Sci. China (Ser. D)* **2001**, *31*, 529–536.
- Wang, Q. Current research on crustal movements in Chinese mainland. *Acta Seismol. Sin.* **2003**, *25*, 541–547.
- Yang, G.; Jiang, Z.; Wu, Y.; Han, Y. Chinese mainland as a whole lacks a net rotation benchmark and its application. *Geod. Geodyn.* **2005**, *25*, 6–10.
- Shen, Z.; Wang, M.; Gan, W.; Zhang, Z. Study on the current tectonic strain rate field and its dynamic origin in the Chinese mainland. *Earth Sci. Front.* **2003**, *10*, 93–100.
- Jiang, Z.; Ma, Z.; Zhang, X.; Wang, Q.; Wang, S. Preliminary GPS results reveal the horizontal strain field and tectonic deformation in the Chinese mainland. *Chin. J. Geophys.* **2003**, *46*, 352–358.
- Wang, Y.; Gan, W.; Chen, W.; You, X.; Lian, W. Preliminary results of co-seismic displacement observed by GNSS during the 7.0-magnitude earthquake in Jiuzhaigou. *Chin. J. Geophys.* **2018**, *61*, 161–170.
- Du, F.; Wen, X.; Feng, J.; Liang, M.; Long, F.; Wu, J. Seismic tectonic characteristics and strong earthquake hazard background of the Liupanshan fault zone. *Chin. J. Geophys.* **2018**, *61*, 545–559.
- Meng, G.; Su, X.; Wang, Z.; Liao, H. Joint inversion of the spatiotemporal rupture process of the Wenchuan M_s8.0 earthquake in 2008 using near-field high-frequency GPS, strong ground motion, and far-field seismic waveform data. *Earthquake* **2018**, *38*, 11–27.
- Li, C.; Su, X.; Meng, G. Spatial distribution characteristics of GPS strain rate at the northeastern margin of the Bayan Har block and its relationship with the 2017 Jiuzhaigou M_s7.0 earthquake. *Earthquake* **2018**, *38*, 37–50.
- Cambiotti, G.; Palano, M.; Orecchio, B.; Marotta, A.M.; Barzaghi, R.; Neri, G.; Sabadini, R. New insights into long-term aseismic deformation and regional strain rates from GNSS data inversion: The case of the pollino and castrovillari faults. *Remote Sens.* **2020**, *12*, 2921. [CrossRef]
- Gambardella, L.M.; Dorigo, M. Solving symmetric and asymmetric TSPs by ant colonies. In *Proceedings of IEEE International Conference on Evolutionary Computation*; IEEE: New York, NY, USA, 1996; pp. 622–627.

25. Kennedy, J.; Eberhart, R. Particle swarm optimization. In *Proceedings of ICNN'95-International Conference on Neural Networks*; IEEE: New York, NY, USA, 1995; Volume 4, pp. 1942–1948.
26. Shi, Y.; Eberhart, R. A modified particle swarm optimizer. In *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No. 98TH8360)*; IEEE: New York, NY, USA, 1998; pp. 69–73.
27. Bekaert, D.; Segall, P.; Wright, T.J.; Hooper, A.J. A network inversion filter combining GNSS and InSAR for tectonic slip modeling. *J. Geophys. Res. Solid Earth* **2016**, *121*, 2069–2086. [CrossRef]
28. Jiang, G. Study on Crustal Deformation Characteristics and Seismic Fault Parameter Inversion of the Longmen Mountain Fault Zone. Ph.D. Thesis, Chang'an University, Xi'an, China, 2016.
29. Liu, J. Joint inversion of three-dimensional slip rates for faults along the northern margin of Qilian Mountains using particle swarm algorithm based on GPS and leveling data. *J. Geod. Geodyn.* **2019**, *39*, 906–909.
30. Smittarello, D.; Cayol, V.; Pinel, V.; Froger, J.-L.; Peltier, A.; Dumont, Q. Combining InSAR and GNSS to track magma transport at basaltic volcanoes. *Remote Sens.* **2019**, *11*, 2236. [CrossRef]
31. Cui, D.; Wang, Q.; Hu, Y.; Wang, W.; Zhu, G. Inversion of GPS data for slip rates and locking depths of the Haiyuan fault. *J. Acta Seismol. Sin.* **2009**, *31*, 516–525.
32. Peng, D.; Chen, J.; Zheng, S.; Nie, Y.; Lam, C.-C.; Hui, T.-W.; Feng, L. Crustal deformation in Hong Kong: Insights from GNSS observations and geophysical models. *Int. J. Appl. Earth Obs. Geoinf.* **2025**, *144*, 104946. [CrossRef]
33. Liu, Q.; Wen, X.; Shao, Z. Joint inversion of GPS, leveling, and strong-motion observation data for the distribution of co-seismic slip during the 2013 Lushan 7.0 magnitude earthquake. *Chin. J. Geophys.* **2016**, *59*, 2113–2125.
34. Zhang, K.; Gan, W.; Liang, S.M.; Xiao, G.R.; Dai, C.L.; Wang, Y.B.; Li, Z.J.; Zhang, L.; Ma, G.Q. Coseismic displacement and slip distribution of the 2021 May 21, M S 6.4, Yangbi Earthquake derived from GNSS observations. *Chin. J. Geophys.* **2021**, *64*, 2253–2266.
35. Okada, Y. Surface deformation to shear and tensile faults in a halfspace. *Bull. Seismol. Soc. Am* **1985**, *75*, 1135–1154. [CrossRef]
36. Okada, Y. Internal deformation due to shear and tensile faults in a half-space. *Bull. Seismol. Soc. Am.* **1992**, *82*, 1018–1040. [CrossRef]
37. Fujita, M.; Nishimura, T.; Miyazaki, S.i. Detection of small crustal deformation caused by slow slip events in southwest Japan using GNSS and tremor data. *Earth Planets Space* **2019**, *71*, 96. [CrossRef]
38. Meneses-Gutierrez, A.; Nishimura, T.; Hashimoto, M. Coseismic and postseismic deformation of the 2016 Central Tottori earthquake and its slip model. *J. Geophys. Res. Solid Earth* **2019**, *124*, 2202–2217. [CrossRef]
39. Nishimura, T.; Fujiwara, S.; Murakami, M.; Suito, H.; Tobita, M.; Yarai, H. Fault model of the 2005 Fukuoka-ken Seiho-oki earthquake estimated from coseismic deformation observed by GPS and InSAR. *Earth Planets Space* **2006**, *58*, 51–56. [CrossRef]
40. Nishimura, T.; Hirasawa, T.; Miyazaki, S.i.; Sagiya, T.; Tada, T.; Miura, S.; Tanaka, K. Temporal change of interplate coupling in northeastern Japan during 1995–2002 estimated from continuous GPS observations. *Geophys. J. Int.* **2004**, *157*, 901–916. [CrossRef]
41. Hao, M.; Li, Y.; Qin, S. Spatiotemporal distribution of slip rate deficit in the Haiyuan-Liupanshan fault zone based on GPS data. *Seism. Geol.* **2017**, *39*, 471–484.
42. Zhao, J.; Ren, J.; Jiang, Z.; Liu, X.; Niu, A.; Yan, W.; Yue, C.; Yuan, Z. Locking and deformation characteristics of the southwestern segment of the Longmenshan fault zone. *Seism. Res.* **2018**, *41*, 216–225.
43. Guo, N.; Wu, Y.; Su, G. Analysis of the fault slip, creep, and coupling characteristics of the Maomaoshan-Laohushan-Haiyuan Fault using InSAR and GNSS measurements. *Tectonophysics* **2023**, *863*, 229988. [CrossRef]
44. Wan, Y.; Shen, Z.; Wang, M.; Zhang, Z.; Gan, W.; Wang, Q.; Sheng, S. Inversion of co-seismic rupture distribution for the 2001 Kunlun Mountain Pass West earthquake based on GPS and InSAR data. *J. Geophys.* **2008**, *51*, 1074–1084.
45. Fu, J.; Zheng, C. Evolution of Ordovician Huabei Sea and Qilian Sea and Petrofacies Paleogeographic Characteristics in the Ordos Basin. *Acta Paleogeographica Sin.* **2001**, *3*, 25–34.
46. Li, S. *Tectonic Rotation and Other Related Issues on the Composite System of the Geotectonic System in Northwest China*; Science Press: Beijing, China, 1955.
47. Ma, X.; You, Z.; Tan, Y.; Yang, W.; Li, D.; Wu, Z. Several fundamental issues on the tectonics of China. *Acta Geol. Sin.* **1961**, *41*, 30–44.
48. Ma, Z. *Seismic Zoning and Its Dynamics in the Chinese Mainland*; Science Press: Beijing, China, 1990; pp. 86–95.
49. Ma, Z.; Zheng, D. Central axial structural belt and seismic activity in the China-Mongolia continent. *Seism. Res.* **1981**, *4*, 421–436.
50. Şengör, A.C. *The Cimmeride Orogenic System and the Tectonics of Eurasia*; Geological Society of America: Boulder, CO, USA, 1984.
51. National Earthquake Administration. *Peripheral Active Fault System of Ordos*; National Earthquake Administration: Beijing, China, 1988.
52. Zhang, K. *Structure and Resources of the Ordos Block*; Shaanxi Science and Technology Press: Xi'an, China, 1989.
53. Huang, J. *China's Main Geological Structural Units*; Geological Publishing House: Beijing, China, 1994.

54. Yin, A.; Nie, S. *20 A Phanerozoic Palinspastic Reconstruction of China and Its Neighboring Regions*; Cambridge University: Cambridge, UK, 1996.
55. Zhai, G.; Song, J.; Jin, J.; Gao, W. *Block Tectonic Evolution and Hydrocarbon-Bearing Basin Formation and Evaluation*; Petroleum Industry Press: Beijing, China, 2002.
56. Bai, Y.; Wang, X.; Liu, H.; Li, T. Determination of the western boundary of the Ordos Basin and its geodynamic background. *Acta Geol. Sin.* **2006**, *80*, 792–813.
57. Yang, S.; Geng, X.; Guo, Q.; Hou, G.; Liu, Z. Mesozoic tectonic evolution in the northern section of the western margin of the Ordos Basin. *Geol. Rev.* **2008**, *54*, 307–315.
58. Darby, B.J.; Ritts, B.D. Mesozoic contractional deformation in the middle of the Asian tectonic collage: The intraplate Western Ordos fold–thrust belt, China. *Earth Planet. Sci. Lett.* **2002**, *205*, 13–24. [CrossRef]
59. Yang, G.; Qian, X.; Guo, H. *Structural Evolution and Oil/Gas Prospect Evaluation of the Kuqa Intra-Continent Flexion Basin Along the Northern Margin of the Tarim Basin*; Geological Publishing House: Beijing, China, 2003.
60. Wang, S. Tectonic evolution and tectonic coal control in the Ordos Basin. *Geol. Bull.* **2011**, *30*, 544–552.
61. Zhong, S.; Wu, J.; Fang, L.; Wang, W.; Fan, L.; Wang, H. Surface wave tomographic imaging of the northeastern margin and surrounding regions of the Qinghai-Tibet Plateau based on the Cheng-Han equation. *Chin. J. Geophys.* **2017**, *60*, 2304–2314.
62. Zhang, J.; Ma, Z.; Ren, W. Structural characteristics of the thrust-fold belt along the western margin of Ordos and the formation mechanism of its north-south variations. *Acta Geol. Sin.* **2004**, *78*, 600–611.
63. Lei, Q. Extension of the Arc-Shaped Tectonic Belt Along the Northeastern Margin of the Qinghai-Tibet Plateau and Evolution of the Yinchuan Basin on the Western Margin of North China. Doctoral Dissertation, Institute of Geology, China Earthquake Administration, Beijing, China, 2016.
64. Hao, M. Study on Present Crustal Vertical Motion, Typical Seismic Isochronology, and Post-earthquake Vertical Deformation along the Eastern Margin of the Qinghai-Tibet Plateau Based on Precise Leveling Data. *Int. Seism. Dyn.* **2013**, *34*, 40–42.
65. Liu, C.-Y.; Zhao, H.-G.; Wang, F.; Chen, H. Mesozoic tectonic attributes of the western margin (part) of the Ordos Basin. *Acta Geol. Sin.* **2005**, *79*, 737–747.
66. Li, T. Tectonic Characteristics and Evolution of Thrusting and Folding Structures in the Western Margin of the Ordos Basin. Ph.D. Thesis, China University of Geosciences, Beijing, China, 2006.
67. Yang, H.; Tao, J.; Ouyang, Z.; Li, G.; Jing, C. Structural characteristics and their causative mechanisms of the western margin of the Ordos Basin. *J. Northwest Univ. Nat. Sci. Ed.* **2011**, *41*, 863–868.
68. Shao, Z.; Wang, W.; Liu, Q.; Dou, A.; Wang, X.; Wu, Y.; Ji, L.; Zhang, L.; Xu, W.; Xu, C. China continental earthquake risk prediction and seismic risk assessment for the next 10 years. *Earthquake* **2025**, *45*, 193–218.
69. Dong, P.; Shi, Y.; Cheng, H.; Qiao, X. Numerical analysis of future seismic activity trends in the Qinghai-Tibet Plateau and adjacent regions. *Chin. J. Geophys.* **2020**, *63*, 1155–1169.
70. Chai, C.; Meng, G.; Ma, G. *Active Fault Detection and Seismic Hazard Assessment in Yinchuan City*; Science Press: Beijing, China, 2011.
71. Xu, Y.; Zeng, X. Analysis of seismic activity and strong earthquake risk in the Wu Zhong–Lingwu region. *Prog. Earthq. Sci.* **2020**, *50*, 1–12.
72. Wang, J.; Ren, X.; Li, G. Study on seismic hazard in Ningxia based on GNSS. *Adv. Geosci.* **2025**, *15*, 1114–1122. [CrossRef]
73. Cai, X. Seismic risk analysis of northern Ningxia and adjacent areas in the coming years. *Adv. Geosci.* **2019**, *9*, 478. [CrossRef]
74. Li, Q.; You, X.; Yang, S.; Du, R.; Qiao, X.; Zou, R.; Wang, Q. A precise velocity field of tectonic deformation in China as inferred from intensive GPS observations. *Sci. China. Earth Sci.* **2012**, *55*, 695. [CrossRef]
75. Wang, M.; Shen, Z.K. Present-day crustal deformation of continental China derived from GPS and its tectonic implications. *J. Geophys. Res. Solid Earth* **2020**, *125*, e2019JB018774. [CrossRef]
76. Deng, Q.; Cheng, S.; Min, W.; Yang, G.; Ren, D. Discussion on Tectonic Activity and Dynamics of the Ordos Block during the Cenozoic Era. *J. Geomech.* **1999**, *5*, 13–21.
77. Yuan, D.; Liu, B.; Lü, T.; He, W.; Liu, X. Study on the slip rate of Maomao Mountain fault using paleosol age data from loess profiles. *Seism. Geol.* **1997**, *19*, 1–8.
78. He, W.; Yuan, D.; Ge, W.; Luo, H. Precise determination of the slip rate of the Lenglongling Fault in the middle-eastern section of the Qilian Mountain active fault zone. *Earthquake* **2010**, *30*, 131–137.
79. Li, C. Quantitative Study of Several Major Fault Zones in the Northeastern Part of the Qinghai-Tibet Plateau. Doctoral Dissertation, Institute of Geology, China Earthquake Administration, Beijing, China, 2005.
80. Chai, C.; Zhang, W.; Jiao, D. Time-and segmental study on the horizontal activity intensity of the late Quaternary in the Tianjingshan fault zone. *Earthq. Res. China* **1997**, *13*, 35–42.

81. Xiang, H.; Gao, S.; Zhang, B.; Zhang, W.; Chi, T.; An, L.; He, H. Active characteristics of the active reverse fault structural belt on the eastern foothills of Liupan Mountain since the late Quaternary. *Seismol. Geol.* **1998**, *20*, 321–327.
82. Zhang, P.; Xu, X.; Wen, X.; Ran, Y. Slip rate, recurrence period, and tectonic causes of the rupture that triggered the 2008 Wenchuan 8.0-magnitude earthquake. *Chin. J. Geophys.* **2008**, *51*, 1066–1073.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Article

Integrating SVR Optimization and Machine Learning-Based Feature Importance for TBM Penetration Rate Prediction

Halil Karahan * and Devrim Alkaya

Department of Civil Engineering, Pamukkale University, 20160 Denizli, Turkey; dalkaya@pau.edu.tr

* Correspondence: hkarahan@pau.edu.tr

Abstract

In this study, a Support Vector Regression (SVR) model was developed to predict the rate of penetration (ROP) during tunnel excavation, and its hyperparameters were optimized using Grid Search (GS), Random Search (RS), and Bayesian Optimization (BO). The results indicate that BO reached the optimal parameter set with only 30–50 evaluations, whereas GS and RS required approximately 1000 evaluations. In addition, BO achieved the highest predictive accuracy ($R^2 = 0.9625$) while reducing the computational time from 25.83 s (GS) to 17.31 s. Compared with the baseline SVM model, the optimized SVR demonstrated high accuracy ($R^2 = 0.9610$ – 0.9625), strong stability ($NSE = 0.9194$ – 0.9231), and low error levels ($MAE = 0.0927$ – 0.1099), clearly highlighting the critical role of hyperparameter optimization in improving model performance. To enhance interpretability, a feature importance analysis was conducted using four machine learning methods: Random Forest (RF), Bagged Trees (BT), Support Vector Machines (SVM), and the Generalized Additive Model (GAM). The relative contributions of BI, UCS, ALPHA, and DPW to ROP were evaluated, providing clearer insight into the model's decision-making process and enabling more reliable engineering interpretation. Overall, integrating hyperparameter optimization with feature importance analysis significantly improves both predictive performance and model explainability. The proposed approach offers a robust, generalizable, and scientifically sound framework for TBM operations and geotechnical modeling applications.

Keywords: Tunnel Boring Machine (TBM); TBM penetration rate; rate of penetration (ROP); Support Vector Regression (SVR); hyperparameter optimization; grid search; Random Search; Bayesian optimization; machine learning (ML)

1. Introduction

Modern transportation, energy, and infrastructure projects rely on tunnels to ensure the continuity of intercity and regional transportation, as well as to support water supply, wastewater management, and energy transmission. Tunnel Boring Machine (TBM) modeling is a critical engineering discipline that enables the practical implementation of governmental directives on infrastructure planning, urban development, environmental protection, and resource management. The realization of safe, economical, rapid, and environmentally efficient tunnel projects demanded by public authorities depends on the scientific modeling of TBM performance, early prediction of risks, and the optimization of energy and material use. In this context, TBM modeling aligns with sustainable development goals and governmental policies such as low-carbon infrastructure and urban resilience by making tunneling processes more efficient, environmentally safer, and technically more predictable.

Optimizing energy consumption, cutter wear, and project duration reduces resource use while lowering the carbon footprint; accurate prediction of ground behavior helps mitigate environmental risks such as surface settlements and groundwater impacts. All these improvements directly contribute to the development of sustainable, resilient, and environmentally compatible infrastructure by reducing costs, enhancing occupational safety, and minimizing adverse social impacts.

However, since tunneling operations are performed under complex geotechnical conditions and variable ground properties, significant costs, safety risks, and technical challenges emerge during the design, construction, and operation phases. Therefore, the efficient and safe excavation of tunnels depends directly on the accurate prediction of Tunnel Boring Machine (TBM) performance. Reliable estimation of the TBM rate of penetration (ROP) is of critical importance for planning excavation operations and controlling project costs.

Early research on TBM penetration rate prediction [1] began with empirical models based on rock strength, followed by studies focusing on TBM performance in sedimentary environments [2]. The first statistical regression model establishing the relationship between rock properties (UCS, PSI, DPW, α) and penetration rate was proposed by [3]. The brittleness index (BI), the mean distance between weak planes (DPW), and the angle between the tunnel axis and weak planes (α) are among the most significant variables influencing TBM performance. When evaluated together with other factors, these three parameters enable a holistic modeling of sustainable penetration capacity by accounting for the applied cutting load and cutter wear.

Thanks to advanced regression analysis, this model achieved a high correlation ($r = 0.82$) and provided an empirical estimation equation based on measured rock parameters [3]. However, these early approaches were mostly limited to simple correlations and were insufficient for capturing complex geotechnical interactions.

Heuristic optimization algorithms were first applied to TBM performance prediction by Yağiz and Karahan [4], who developed a particle swarm optimization (PSO)-based model aimed at minimizing the error between measured and predicted ROP values. Later studies [5] extended this approach using Harmony Search (HS), Differential Evolution (DE), and Gray Wolf Optimization (GWO) algorithms, demonstrating the applicability of heuristic methods to TBM performance prediction. Although these models offer interpretability through their analytical structure, they remain limited in capturing ground heterogeneity and complex interactions.

To overcome these limitations, artificial intelligence-based methods such as Artificial Neural Networks (ANN) [6–8], Support Vector Machines (SVM) [9–11], and fuzzy logic [12–14] have been increasingly used in TBM modeling. These methods can accurately predict penetration rates under multivariable and highly complex geotechnical conditions, effectively capturing nonlinear relationships and complex interactions among parameters. They can even produce reliable results with small or noisy datasets, although interpretability may sometimes be limited.

Support Vector Machine (SVM) and Support Vector Regression (SVR) are methods derived from Vapnik's Statistical Learning Theory and share the same mathematical foundation [15–17]. Both use the maximum-margin principle and can model nonlinear relationships through kernel functions. Their main difference lies in the type of problem they solve: while SVM aims to separate two classes with the widest margin in classification problems, SVR seeks to find the best-fitting function in regression problems using the ϵ -insensitive loss function [16–21]. Therefore, SVR can be considered a version of SVM adapted for continuous output prediction; both methods share the same mathematical structure and learning mechanisms, enabling effective modeling of nonlinear relationships, reduction in

uncertainties, and reliable predictions across various engineering and scientific applications. While SVM is used for classification, SVR is employed for regression problems. In this study, SVR was used because the model training utilized sequential samples collected along the tunnel alignment.

Accurate prediction of TBM penetration rate is critical for planning and optimizing tunneling projects. In recent years, AI-based models have been widely used for this purpose [22,23]. For example, study [22] developed an SVR model under challenging rock conditions and achieved high predictive accuracy. Similarly, Afradi and Ebrahimabadi [23] compared ANN, SVR, and GEP models for the Chamshir water transfer tunnel in Iran and reported that the GEP model produced relatively lower errors. These studies emphasize the influence of dataset size and geological variability on model performance.

A more recent study [24] compared MLR, ANN, and SVR techniques for predicting TBM penetration rate in the Kerman and Gavoshan tunnels. The models used variables such as RQD, UCS, BTS, density, Poisson's ratio, joint angle, and joint spacing as inputs. Although all methods achieved high accuracy, SVM predictions were found to be more precise and realistic than those of the other techniques.

The accuracy and generalization capability of SVM and SVR models critically depend on the proper selection of hyperparameters. In particular, the C parameter controls model fit to the training data, γ determines the influence of the kernel function, and ϵ defines the tolerance band in SVR [15–17]. If these parameters are not optimized properly, the model either becomes overly simplistic (underfitting) or overly complex (overfitting), resulting in reduced predictive accuracy and poor generalization. Dependence on hyperparameter optimization and data preprocessing steps can make modeling time-consuming and computationally expensive for ANN, SVM, and similar methods, especially with large datasets [25–28]. For SVM, partitioning the dataset into training and testing subsets or applying cross-validation plays a critical role in reducing overfitting risks [29]. Previous studies [22–24,30–32] have conducted extensive trial-and-error procedures to identify appropriate parameter values and achieve high model performance.

In this study, an SVR-based model was developed for predicting TBM penetration rate (ROP), and its performance was systematically enhanced using Grid Search, Random Search, and Bayesian Optimization strategies. Furthermore, a comprehensive comparison was conducted using several widely applied machine learning algorithms—including Random Forest (RF), Bagged Trees (BT), Support Vector Machines (SVM), and the Generalized Additive Model (GAM)—on the same dataset to benchmark the SVR results. In addition, the relative influence of the input variables on ROP was examined through a unified assessment of feature importance obtained from these four methods, providing the first holistic evaluation of variable contributions within this context. This dual focus on prediction accuracy and model interpretability demonstrates that combining an optimized SVR framework with feature-importance-based insights substantially enhances the reliability, consistency, and process-oriented understanding of ROP forecasting in TBM projects.

2. Materials and Methods

2.1. Study Area and Data Collection

This study is based on field and laboratory measurements obtained from a tunnel construction project located in the southwestern part of New York City, in the Queens area [3]. The dataset used in the analysis [4,5] has also been widely employed in subsequent TBM performance modeling studies [7,8]. It contains the key parameters required for predicting TBM rate of penetration (ROP). The measured parameters include uniaxial compressive strength (UCS), Brazilian tensile strength (BTS), brittleness index (BI), mean

distance between weak planes (DPW), and the angle (α) between the tunnel axis and the weak planes. The data used in the study are provided as Supplementary Material.

These data were used to model TBM performance and predict penetration rates, reliably representing field conditions and providing a practical reference for both academic research and real-world tunneling applications.

2.2. Input and Output Parameters

Five fundamental input parameters influencing TBM performance were selected for the prediction models: UCS, BTS, BI, DPW, and α° . The target output of the model was defined as the TBM rate of penetration (ROP, m/h). The correlation relationships between these input variables and ROP are presented in Figure 1.

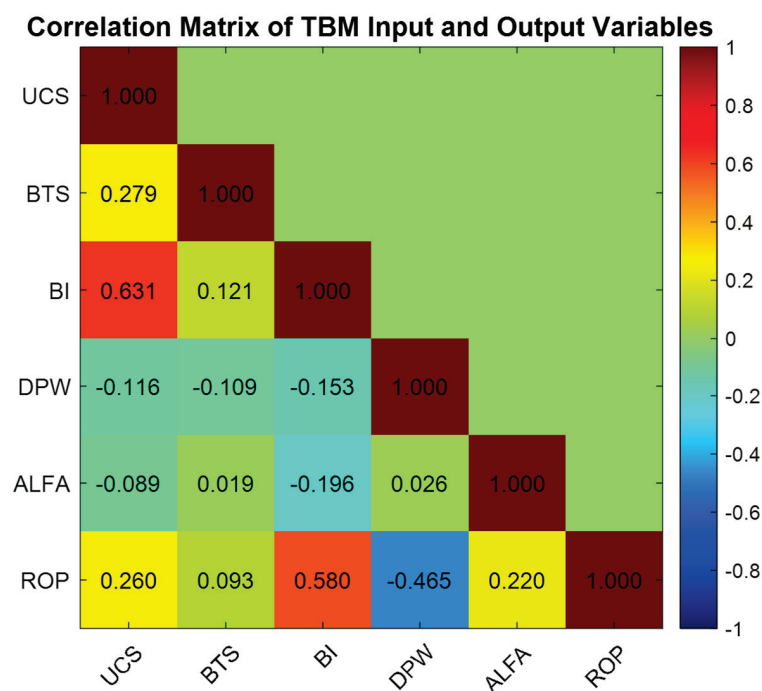


Figure 1. The correlation relationship between ROP and model parameters.

As shown in Figure 1, the strongest correlations with ROP are observed with BI, UCS, α , and BTS parameters, respectively, whereas DPW exhibits a negative correlation of -0.465 with ROP. The correlation matrix clearly illustrates the relationships between ROP and the other variables, providing a crucial foundation for constructing the network architectures of SVM and ANN models. This information enables the identification of which input variables exert the most significant influence on the output.

The variation in the model input variables and the measured ROP values along the tunnel length is presented in Figure 2. In this figure, the first five parameters represent the model inputs, while the last parameter corresponds to the target ROP to be predicted. It is evident that the relationship between the model inputs and ROP is highly variable and complex along the tunnel length. Consequently, conventional regression methods or empirically derived relationships optimized through heuristic approaches exhibit limited success in accurately capturing these intricate patterns.

To more accurately model the complex and nonlinear relationships illustrated in Figure 2, and to enable a performance comparison across different ML approaches, advanced machine learning methods such as RF, BT, SVM, and GAM were employed. During the parameter optimization and modeling stages, the dataset was divided into training

and test subsets to minimize the risks of overfitting and underfitting and to enhance the model's generalization capability.

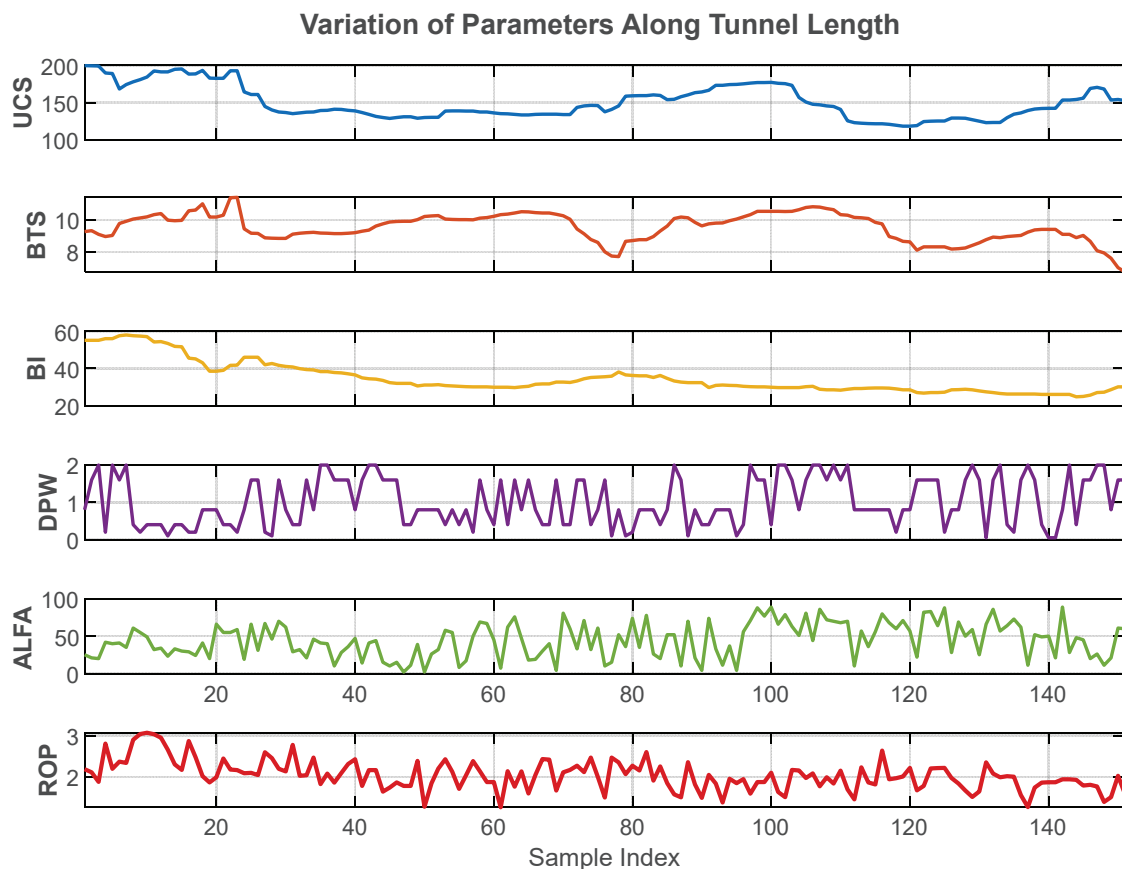


Figure 2. Variation in model variables and ROP values along the tunnel length.

In the SVM model, the hyperparameters C , γ , and ε were optimized using three different strategies: Grid Search, Random Search, and Bayesian Optimization. Subsequently, the predictive performance of the model was evaluated in a comparative manner to determine the most effective configuration among these optimization methods.

2.3. Evaluation of Model Performance

The performance of the models was assessed using several statistical indicators commonly adopted in the literature, including the coefficient of determination (R^2), mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and the Nash–Sutcliffe efficiency coefficient (NSE). The R^2 value represents the proportion of the measured ROP variance explained by the model and reflects its overall explanatory capability. MSE and RMSE quantify the prediction accuracy, with lower values indicating a higher level of precision. MAE defines the average magnitude of the prediction errors, providing insight into the model's overall error distribution. The NSE coefficient measures the agreement between predicted and observed values, where values approaching unity denote excellent predictive performance. Collectively, these metrics offer a comprehensive evaluation of model accuracy, robustness, and generalization capacity, enabling a meaningful comparison of the effects of different hyperparameter optimization strategies.

3. Methodology

3.1. General Approach

In this study, an SVR-based model was developed to predict the performance parameters of Tunnel Boring Machines (TBMs). SVR was selected as the modeling technique due to its strong generalization capability and its effectiveness in handling nonlinear and multivariate data structures, making it particularly suitable for TBM performance prediction.

To enhance model performance, the key SVR hyperparameters—the penalty parameter (C), kernel parameter (γ), and epsilon (ϵ)—were optimized using three different techniques: Grid Search, Random Search, and Bayesian Optimization. These methods were chosen based on their distinct strategies for exploring the hyperparameter space and their potential to identify the global optimum.

This comprehensive framework enabled a comparative evaluation of each optimization method in terms of prediction accuracy, computational cost, and model stability. Accordingly, the study aims to identify the most effective optimization strategy for TBM performance prediction, demonstrate the effectiveness of SVR-based modeling, and investigate the relative importance of input variables on model outputs.

In this study, a Support Vector Machine (SVM)-based model was developed to predict the performance parameters of a Tunnel Boring Machine (TBM). SVM was selected as the modeling technique due to its strong generalization capability in handling nonlinear and multivariate data structures, which makes it a suitable approach for TBM performance prediction.

To optimize the model performance, the main hyperparameters of SVM—namely the penalty parameter (C), kernel parameter (γ), and epsilon (ϵ)—were tuned using three different optimization techniques: Grid Search, Random Search, and Bayesian Optimization. These methods were chosen considering their distinct strategies for exploring the hyperparameter space and their potential to achieve a global optimum.

Through this approach, the effects of each optimization technique on model prediction accuracy, computational cost, and stability were comparatively evaluated. Consequently, the study aimed to identify the most suitable optimization strategy for TBM performance prediction and to demonstrate the effectiveness of SVM-based modeling in this context.

3.2. Dataset and Preprocessing

In the modeling process, BI, UCS, α , BTS, and DPW parameters were considered as independent variables, while the rate of penetration (ROP), representing the excavation performance, was designated as the dependent variable.

Although outlier removal is commonly applied to enhance model performance, in this study, extreme values were deliberately retained. Eliminating these data points could weaken the representation of natural excavation behavior in both hard and soft ground conditions. Therefore, the dataset was preserved in its entirety to ensure a more realistic reflection of actual field performance.

Additionally, the dataset was divided into two subsets, with 70% used for training and 30% for testing. In contrast to the commonly adopted 80–20 split in the literature, this study intentionally trained the model with a smaller portion of the data. This approach aimed to evaluate the model's generalization capability under limited training conditions and to obtain a performance assessment that better reflects real-world scenarios. Remarkably, despite being trained with less data, the proposed model achieved superior predictive accuracy compared to studies employing larger training ratios, demonstrating its strong generalization ability and the effectiveness of the adopted methodology. This confirms that

the SVM-based approach can effectively capture underlying patterns and maintain robust predictive power even when trained on a reduced dataset.

3.3. SVM Model

SVM is a powerful regression and classification technique capable of modeling both linear and nonlinear relationships. Since this study does not involve classification, SVR was employed to predict the TBM performance parameters. The flowchart of the SVR model is presented in Figure 3.

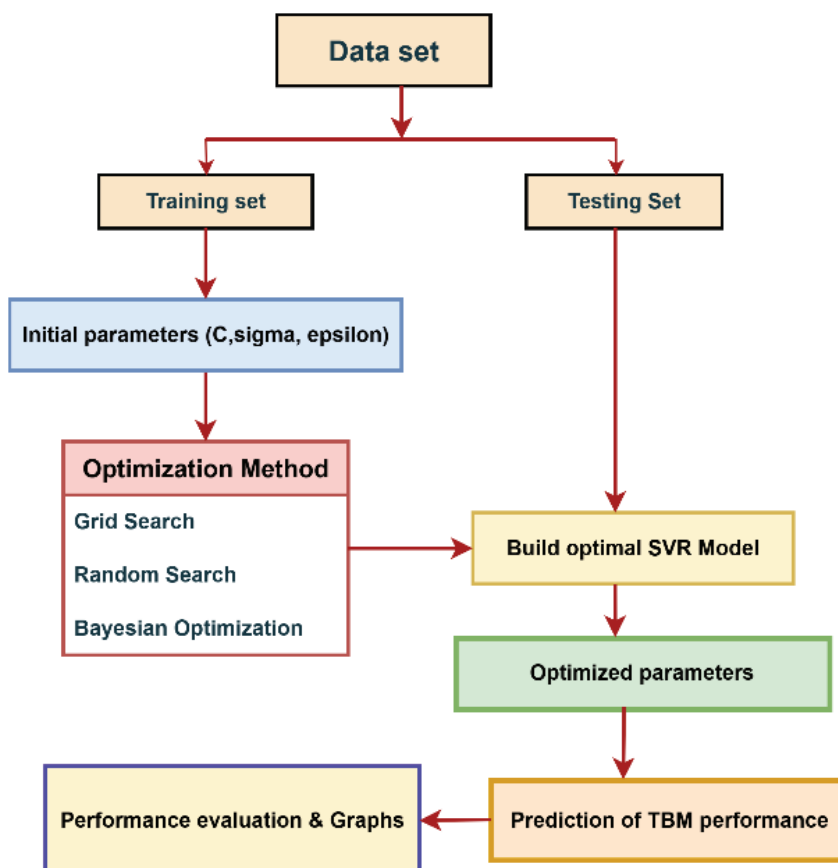


Figure 3. Flowchart of the proposed SVR model for predicting TBM penetration rate.

The SVR model illustrated in Figure 3 was configured using the Radial Basis Function (RBF) kernel, which enables effective learning by mapping nonlinear patterns into a high-dimensional feature space. The RBF kernel evaluates the similarity between data points based on Euclidean distance, allowing the model to capture complex structural relationships within the data. The performance of the SVR model is highly dependent on the proper selection of its key hyperparameters— C (penalty parameter), γ (gamma), and ϵ (epsilon).

Specifically, C governs the balance between model complexity and error tolerance, where higher values enforce stricter minimization of deviations and lower values offer greater model flexibility. The γ parameter determines the influence radius of the RBF kernel; smaller values yield smoother decision surfaces, whereas larger values create more localized and sharply defined regions. The ϵ parameter defines the width of the error-insensitive tube within which deviations are not penalized.

Because predictions in the SVR framework rely solely on support vectors, the model remains computationally efficient, requires relatively low memory, and maintains strong generalization capability. Additionally, the margin-maximization principle helps reduce the risk of overfitting.

In this study, the modeling procedure was conducted in three main stages. First, the SVR model was trained using the training dataset. Subsequently, the hyperparameters were optimized using Grid Search, Random Search, and Bayesian Optimization techniques. Finally, the optimal parameter combination was applied to the test dataset, and the predictive performance of the model was comprehensively evaluated.

3.4. Hyperparameter Optimization

The performance of an SVM model depends on three key parameters: the penalty parameter (C), the RBF kernel parameter (γ), and epsilon (ϵ) [15,19].

In this study, the hyperparameter ranges used in the SVR model were determined based on recommendations from previous research as well as preliminary experiments. The parameter C was defined within the range of 0.1–10,000 to balance model complexity and generalization capability, while γ , which controls the influence of individual training samples in capturing nonlinear relationships, was restricted to the interval between 1×10^{-6} and 0.1. The ϵ parameter, representing the tolerance for prediction errors, was explored within the range of 1×10^{-3} to 0.2. These ranges guided the hyperparameter search across the three optimization methods.

Hyperparameter optimization was performed using three different search strategies. Grid Search systematically evaluated all possible parameter combinations arranged on a logarithmically spaced grid within the selected ranges. Random Search tested 1000 randomly selected parameter combinations, enabling exploration of a broader parameter space with lower computational cost [33]. Bayesian Optimization was implemented as a probabilistic optimization technique that dynamically explores the parameter domain using a Gaussian Process-based surrogate model and selects new candidate combinations via the Expected Improvement (EI) acquisition function [34,35]. The performance of the SVM model depends on three key parameters: the penalty parameter (C), the RBF kernel parameter (γ), and epsilon (ϵ) [22,23].

4. Model Implementation

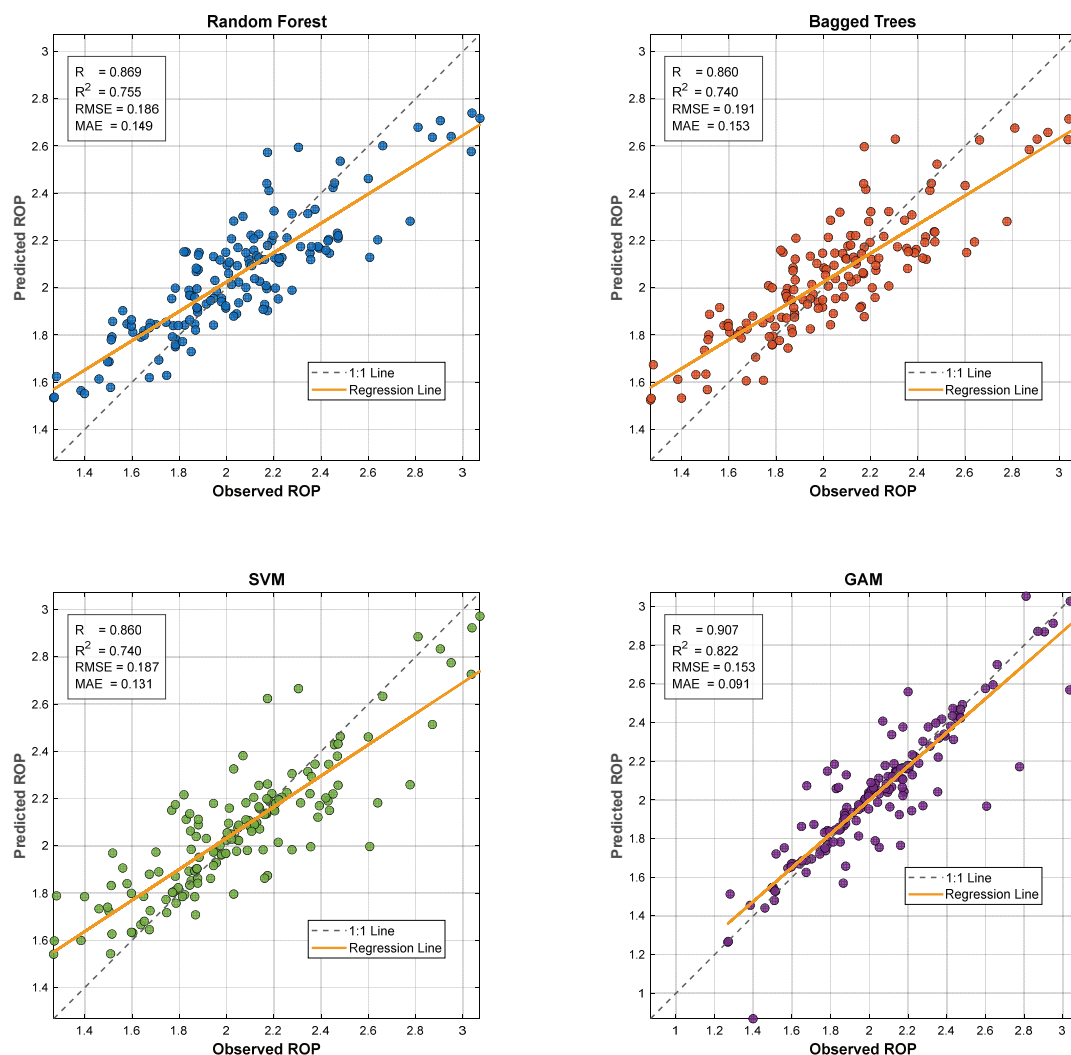
4.1. Machine Learning Algorithms

Machine learning is an artificial intelligence discipline that aims to mathematically represent complex processes and to generalize from past observations without requiring any predefined functional form, equation, or parametric structure. Instead, it derives patterns and relationships directly from data. This capability allows computer systems to learn how to perform a given task without being explicitly programmed to do so [36,37]. Machine learning algorithms are commonly categorized as supervised, unsupervised, or semi-supervised learning, and they can be applied to a wide range of problems, including regression, classification, clustering, and dimensionality reduction. Today, owing to their high predictive performance, flexibility, and minimal dependence on strict model assumptions, they have become essential analytical tools across numerous fields—from engineering and hydrology to medicine and financial modeling.

In this study, the predictive performance of the Support Vector Regression (SVR) method in estimating TBM penetration rate was evaluated through a comparative modeling framework using the same dataset. The SVR model was benchmarked against several classical machine learning algorithms, namely Random Forest (RF) [38,39], Bagged Trees (BT) [40], and the Generalized Additive Model (GAM) [41,42]. The quantitative performance indicators of these ML models are presented in Table 1, while Figure 4 illustrates scatter plots showing the relationships between observed and predicted values.

Table 1. Model Performance Comparison of Machine Learning Methods.

Model	R ²	RMSE	MAE	RSR	NSE
RF	0.731	0.186	0.149	0.517	0.731
BT	0.717	0.191	0.153	0.531	0.717
SVM	0.728	0.187	0.131	0.520	0.728
GAM	0.819	0.153	0.091	0.425	0.819

**Figure 4.** Observation–prediction scatter plots for the ML models.

According to Table 1, the GAM outperforms the other methods across all performance metrics and achieves the highest predictive accuracy. RF and SVM models show comparable and moderately high performance, whereas BT ranks below them with a relatively lower level of accuracy. These findings are also clearly reflected in Figure 4, where the positions of the 1:1 and regression lines, together with the distribution of the scatter points, visually confirm the relative performance differences.

Sensitivity Analysis of Machine Learning Algorithms

In this study, the developed code aims to systematically evaluate regression models based on Support Vector Machine (SVM), Random Forest (RF), Boosted Trees (BT), and Generalized Additive Model (GAM) using the same dataset. For the SVM model, the penalty parameter (C), kernel function parameter (γ), and error tolerance parameter (ϵ) were defined over different ranges, and each hyperparameter combination was assessed

using five-fold cross-validation, with the corresponding mean squared error (MSE) values calculated. Similarly, for the RF and BT models, the effects of the number of trees and the minimum leaf size on model performance were investigated, and prediction errors were obtained for each hyperparameter combination using the test dataset.

Through this approach, the optimal hyperparameters for each model were identified, and the relative influence of individual hyperparameters on predictive performance was quantified. In addition, a computational cost analysis was conducted for all models by separately measuring and comparing training and prediction times. Consequently, not only prediction accuracy but also computational efficiency and practical applicability of the models were evaluated, and the results are presented in Figure 5.

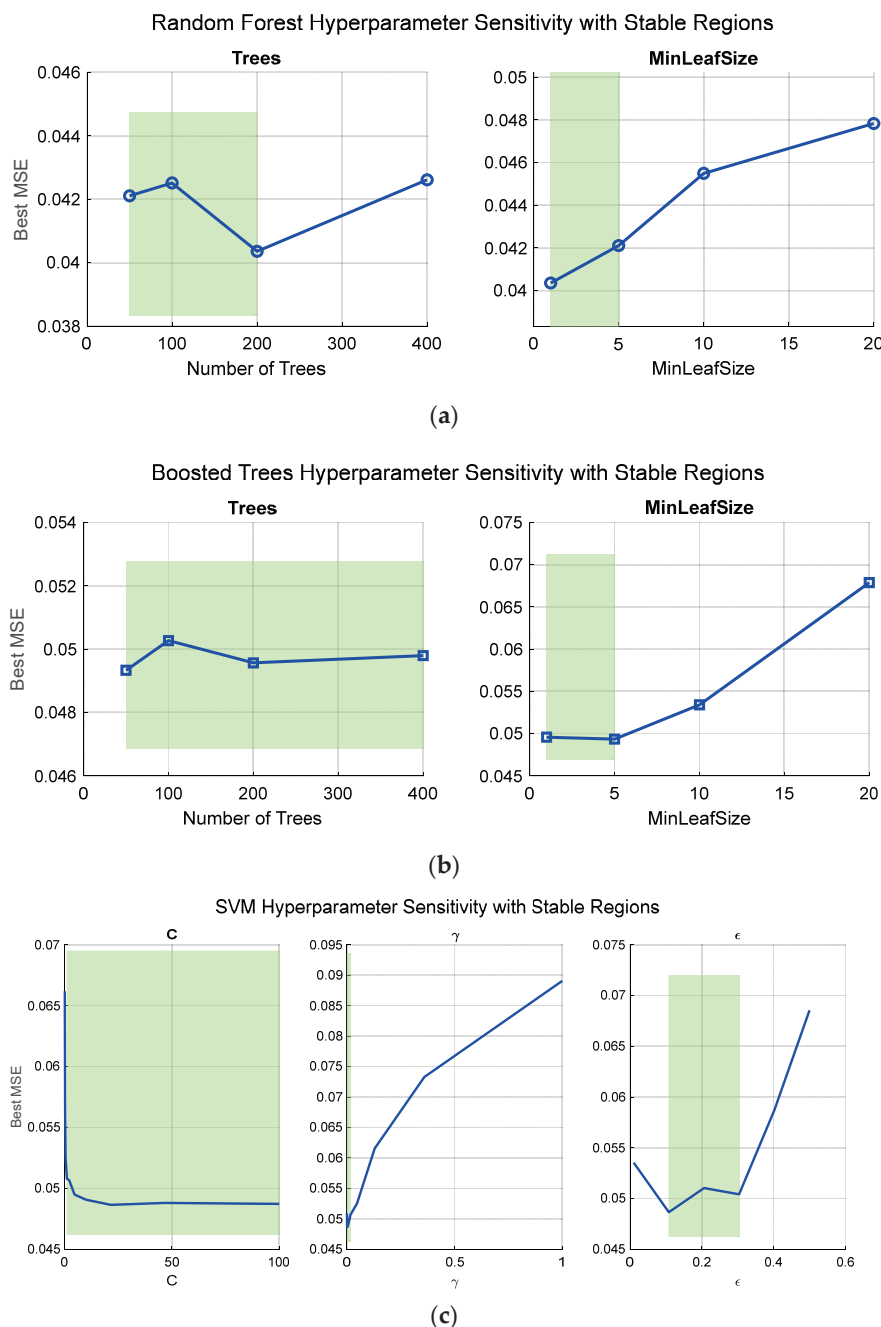


Figure 5. Sensitivity analysis and stability regions for ML algorithms.

The stability region analyses presented in Figure 5 reveal the hyperparameter sensitivities and performance tolerances of different machine learning models. In SVM, the γ

parameter of the RBF kernel exhibits a narrow stable range with high performance sensitivity; even small changes can significantly affect the mean squared error (MSE). In contrast, the C parameter has a broader stable range, and ϵ shows an intermediate range, making γ tuning critical while providing more tolerance for C and ϵ . In Random Forest and Boosted Trees models, the number of trees is relatively tolerant, whereas the MinLeafSize parameter requires careful adjustment.

GAM was primarily evaluated for interpretability, with partial dependence functions directly visualizing the effect of each input on ROP. Some variables display linear relationships, while others exhibit nonlinear and saturation effects. Computationally, SVM and GAM offer shorter training times, whereas Boosted Trees require the longest; prediction times are fast across all models. These findings enable a comparative assessment of the models in terms of both performance and computational efficiency, supporting their applicability in decision-making processes.

4.2. Optimized SVR Model

In this study, the developed SVM model was implemented using three different optimization strategies after splitting the dataset into 70% training and 30% testing, as illustrated in Figure 3. The predictive performance of the model was comprehensively evaluated using several performance indicators, including Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), the Coefficient of Determination (R^2), and the Nash–Sutcliffe Efficiency Coefficient (NSE).

During the hyperparameter optimization phase, Grid Search, Random Search, and Bayesian Optimization techniques were employed. Among these methods, Bayesian Optimization demonstrated a clear advantage in terms of computational efficiency by achieving high prediction accuracy with significantly fewer evaluations. The optimal values of C and γ obtained from each technique reflected the distinct search behaviors of the optimization methods, and the impact of these variations on model performance was visualized through the three-dimensional hyperparameter search surfaces shown in Figure 6.

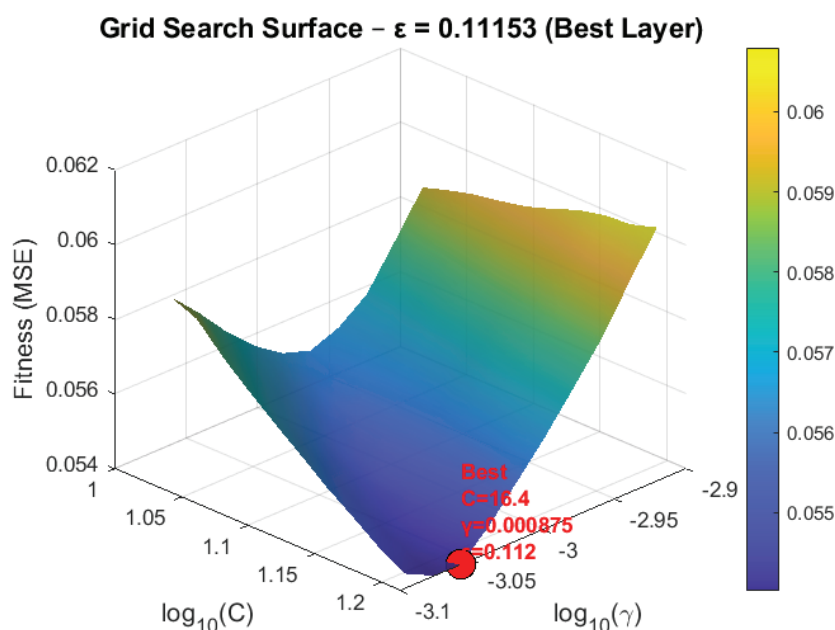


Figure 6. Cont.

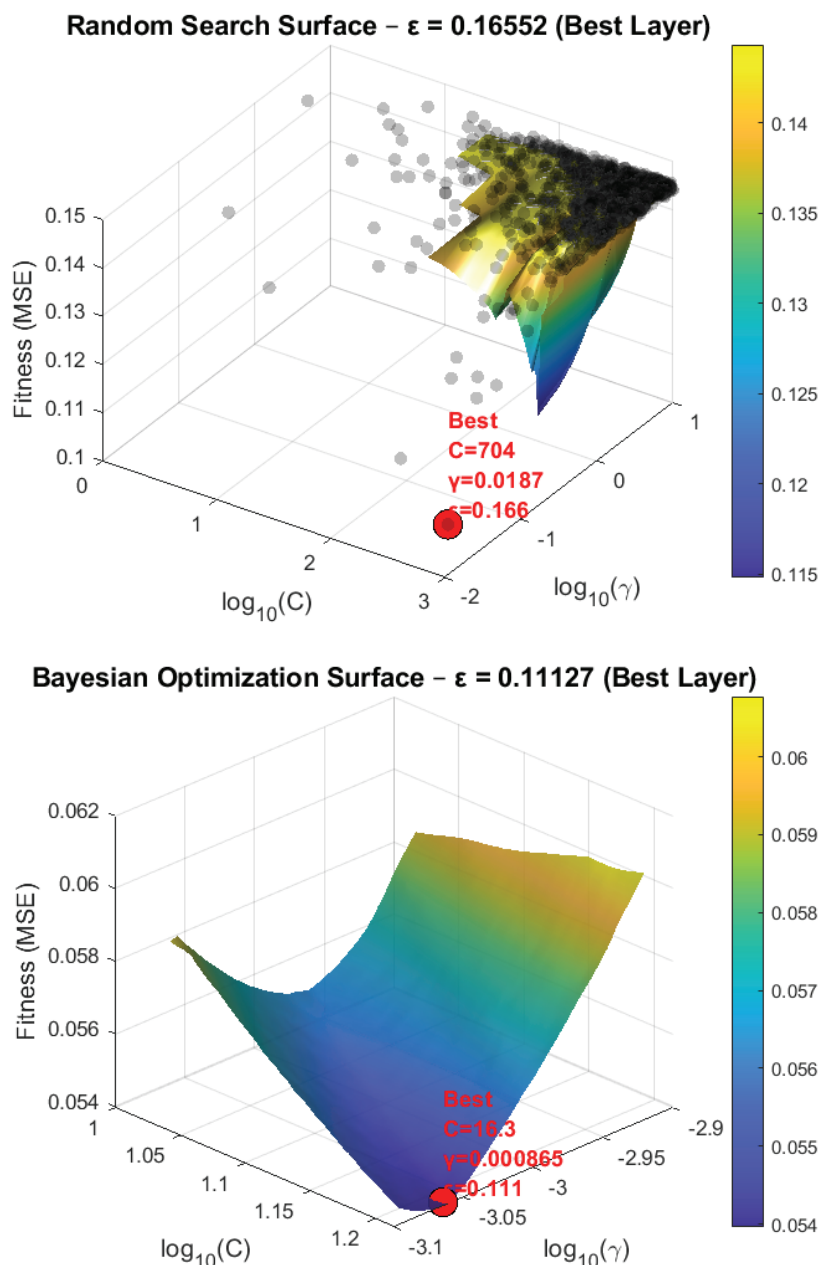


Figure 6. Optimized SVM parameters.

These visualizations clearly indicate that Bayesian Optimization reaches optimal performance with a limited number of evaluations and provides more efficient hyperparameter tuning compared with Grid Search and Random Search. While Grid Search yielded results close to those of Bayesian Optimization, Random Search exhibited greater deviations from the optimal region.

Although reducing grid density in Grid Search or increasing the number of iterations in Random Search could potentially enhance model performance, both approaches substantially increase computational time and memory requirements. In contrast, Bayesian Optimization offers a distinct advantage by delivering high predictive accuracy and computational efficiency with far fewer evaluations.

Overall, the predictions derived from the measured data demonstrate that the developed model exhibits high accuracy and stability. Furthermore, the comparative analysis of the optimization strategies highlights the critical role of hyperparameter selection in model

performance and provides a practical methodology for engineering applications that aim to achieve effective results under limited computational resources.

As shown in Figure 6, the optimal SVR parameters obtained through Bayesian Optimization were determined as $C = 16.322462$, $\gamma = 0.000865$, and $\varepsilon = 0.111265$. To examine the effect of these parameters on model performance, SVR models constructed using parameter sets obtained from the Grid Search and Random Search methods were evaluated and compared using the MSE, RMSE, MAE, NSE, and R^2 metrics. All computations were performed on a 64-bit computer running the Windows operating system, equipped with a 12th-generation Intel Core i7-12700H processor (2.30 GHz) and 32 GB of RAM. The results for the training, test, and entire datasets are presented in Table 2. In addition, Table 2 reports the number of function evaluations (NFE) required during model training, as well as the required CPU time in seconds.

Table 2. Comparison of SVR parameter optimization methods and model performance.

Opt. Method	Opt. Parameters	Dataset	CPU (s)	MSE	MAE	RMSE	R^2	NSE
Bayesian	$C = 16.32$ $\gamma = 0.000865$ $\varepsilon = 0.111$	Training	17.31	0.0096	0.0927	0.0981	0.9610	0.9194
		Testing		0.0121	0.1099	0.1100	0.9625	0.9231
		All		0.0101	0.0961	0.1006	0.9607	0.9213
Random Search	$C = 704$ $\gamma = 0.0187$ $\varepsilon = 0.166$	Training	19.19	0.0197	0.1301	0.1404	0.9277	0.8534
		Testing		0.0214	0.1409	0.1462	0.8709	0.7951
		All		0.0201	0.1323	0.1416	0.9183	0.8440
Grid Search	$C = 16.4$ $\gamma = 0.000875$ $\varepsilon = 0.112$	Training	25.83	0.0102	0.0967	0.1010	0.9558	0.9159
		Testing		0.0106	0.0974	0.1030	0.9726	0.9268
		All		0.0103	0.0968	0.1014	0.9600	0.9200

As presented in Table 2, the R^2 values for the training and test datasets were obtained as 0.9610 and 0.9625, respectively, while the NSE values were 0.9164 and 0.9231, and the MAE values were 0.0927 and 0.1099. These results indicate minimal prediction errors and demonstrate that the SVM model exhibits high accuracy and strong predictive capability. In addition, the comparatively lower-performing Random Search and Grid Search results still fall within the range of $0.75 \leq \text{NSE} (0.9231, 0.7951, 0.9275) \leq 1$, corresponding to the “Very good” classification according to the criteria reported in reference [43]. While the performance of Grid Search was found to be close to that of Bayesian Optimization, Random Search yielded comparatively lower predictive accuracy.

The convergence curves of the three optimization strategies used for SVR parameter tuning, with respect to the number of function evaluations, are presented in Figure 7. As shown in Figure 7, the Bayesian optimization algorithm reaches a stable solution after approximately 30 evaluations, whereas the Grid Search algorithm requires about 1000 function evaluations to achieve similar performance, and Random Search generally exhibits lower performance than the other two methods even after 1000 trials. These results indicate that the proposed Bayesian optimization approach is highly efficient for the SVR model, achieving optimal parameters with a limited number of iterations (30–50). In terms of computational time, as shown in Table 2, Grid Search requires 25.83 s, while Bayesian optimization completes in 17.31 s. In other words, Bayesian optimization is more advantageous in terms of model performance, number of iterations, and CPU time.

Figure 8 illustrates (a) the ROP predictions along the tunnel, (b) the distribution of the training data, and (c) the distribution of the test data.

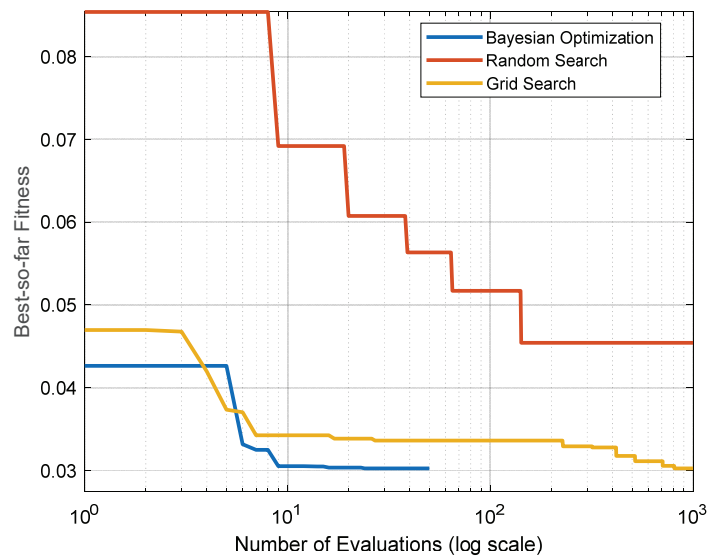


Figure 7. Convergence plots for the optimization strategies.

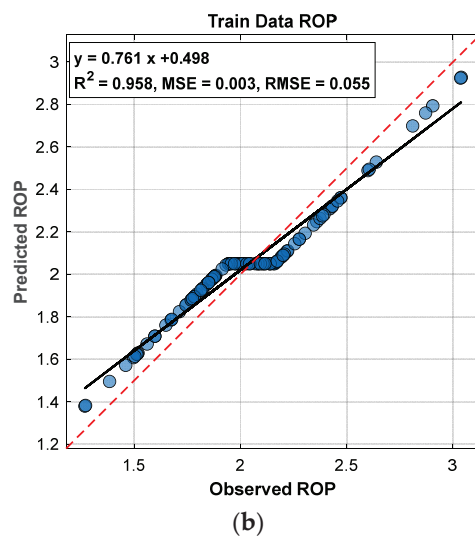
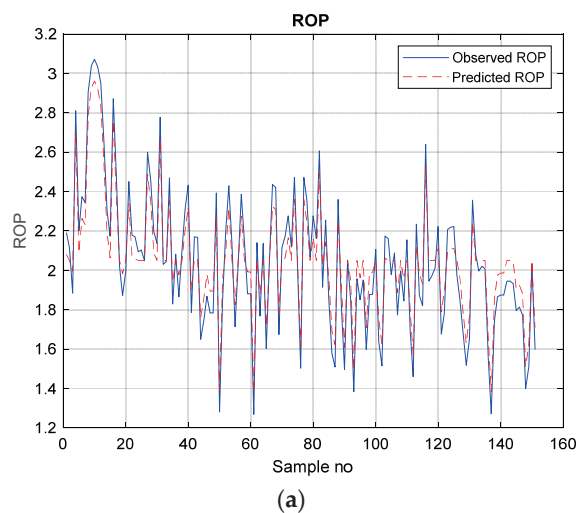


Figure 8. Cont.

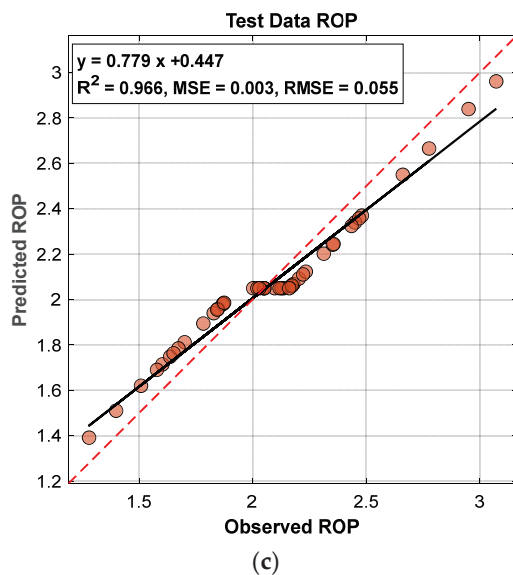


Figure 8. Predicted ROP values along the tunnel (a) and scatter plots of the training (b) and test (c) datasets.

When Table 2 and Figure 8 are evaluated together, it is evident that the model maintains high accuracy and stability across both datasets, and the predictions exhibit a strong agreement with the observed values.

4.3. Feature Importance Analysis

Feature importance analysis is a fundamental explainability approach that reveals the extent to which a machine learning model relies on each input variable when predicting the target output. In scientific studies, this analysis is crucial not only for demonstrating how well a model performs but also for identifying which variables contribute to this performance. The literature emphasizes that feature importance—particularly in tree-based methods such as Random Forest and Gradient Boosting—is a powerful tool for understanding a model’s decision-making mechanism and plays an essential role in interpretability, reliability, and decision-support processes [38,40].

The primary objective of determining feature importance is to clarify the relative impact of the variables in the dataset and enhance the transparency of the model’s internal operations. Through this analysis, it becomes evident which inputs the model relies on most and which variables contribute most significantly to prediction accuracy. As a result, the model moves away from being a “black box” and becomes more interpretable to users. Since not all inputs contribute equally in multivariate data structures, identifying truly informative variables and separating them from noise becomes possible only through such analyses.

Feature importance analysis also offers substantial advantages in terms of dimensionality reduction and model simplification. Removing variables with low contribution, redundancy, or high collinearity leads to a more concise model structure, shortens training time, reduces computational cost, and improves generalization performance. The literature indicates that effective feature selection reduces the risk of overfitting and makes models more robust against independent test data [44].

Additionally, evaluating feature importance facilitates comparisons across different models. Observing which variables receive higher emphasis in different algorithms applied to the same dataset enables an understanding of the learning dynamics and sensitivities of each model. When a specific physical or biophysical variable exhibits high importance in a model, it indicates that the model is more responsive to that process. This not only enables

statistical interpretation but also offers process-based insights, providing valuable scientific contributions, particularly in disciplines such as remote sensing, geosciences, hydrology, and environmental modeling.

For these reasons, this study presents—for the first time in TBM performance modeling—a detailed investigation of the influence of input variables on ROP using four different machine learning approaches. The analysis results are presented in Table 3 and Figure 9.

Table 3. Model-Based Variable Importances for ROP Prediction.

ML Method	Variable			
	BI	UCS	ALPHA	DPW
RF	31.03	10.89	33.77	24.31
BT	36.47	15.17	31.68	16.68
SVM	25.56	25.03	22.43	26.98
GAM	30.45	23.74	29.52	16.28

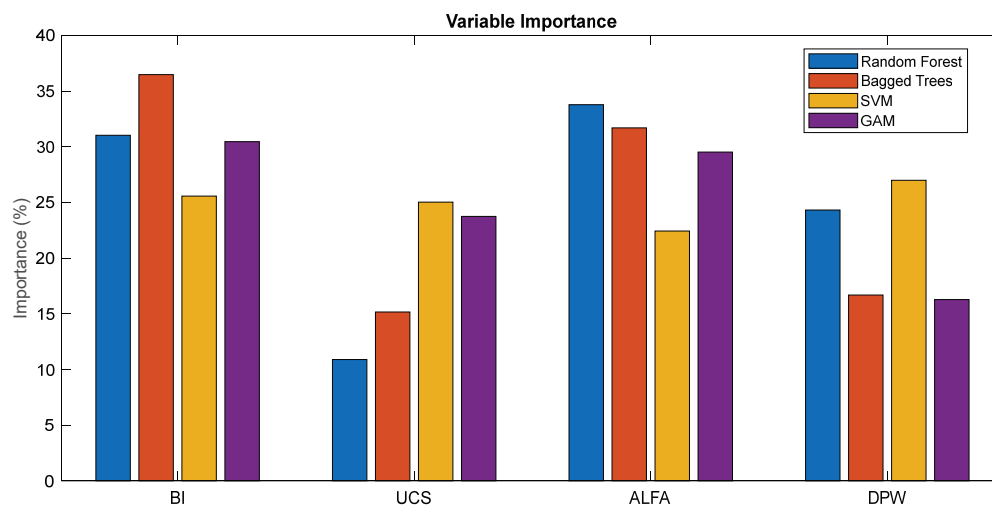


Figure 9. Relative importance of BI, UCS, ALPHA, and DPW variables according to ML models.

Table 3 presents a comparative overview of the relative importance of the input variables across four different machine learning methods used for ROP prediction (RF, BT, SVM, and GAM). The results indicate that variable importance varies distinctly between models, reflecting the differing sensitivities of each algorithm to the underlying data structure.

Tree-based methods (RF and BT) identified BI and ALPHA as the most influential variables. In the RF model, ALPHA (33.77%) and BI (31.03%) emerged as the dominant predictors, whereas BT highlighted BI (36.47%) as the primary variable. This suggests that tree-based methodologies are particularly sensitive to physical and geometric parameters related to penetration in the context of ROP prediction.

Unlike the tree-based algorithms, the SVM model distributed variable importance more evenly among the inputs, with BI, UCS, ALPHA, and DPW showing comparable contributions. Notably, the relatively high importance of UCS in SVM (25.03%) indicates that nonlinear, kernel-based approaches tend to assign greater relevance to strength-related properties of the rock mass.

The GAM, on the other hand, assigned similar importance levels to BI (30.45%), ALPHA (29.52%), and UCS (23.74%), while attributing lower importance to DPW. This pattern highlights GAM's semi-parametric flexibility in capturing smooth and gradual variations in predictor effects, consistently emphasizing the BI–ALPHA–UCS trio.

Overall, the findings show that BI and ALPHA are dominant predictors of ROP across all four models, whereas the importance of DPW varies considerably between methods. UCS, meanwhile, plays a more substantial explanatory role in GAM and SVM. These differences arise from the distinct ways in which each algorithm perceives and interprets relationships within the data, reinforcing that ROP prediction is a multidimensional problem sensitive to interactions among multiple parameters.

In conclusion, feature importance analysis is not merely a model performance assessment tool; it is a comprehensive approach that enhances model interpretability, supports performance optimization, provides deeper insight into variable–target relationships, and strengthens the scientific reliability of the results. The motivation behind conducting this analysis is not only to evaluate model accuracy but also to elucidate the foundations on which the model operates—thereby contributing meaningfully to both practical applications and the academic literature.

5. Discussion

The SVM model developed in this study demonstrated both high accuracy and remarkable stability in predicting the rate of penetration (ROP) during tunnel excavation. The performance metrics presented in Table 2 clearly indicate strong agreement between the predicted and observed values for both the training and test sets: the R^2 values were 0.9610 and 0.9625, NSE values were 0.9194 and 0.9231, and MAE values were 0.0927 and 0.1099, respectively. The training and test MSE values (0.0096 and 0.0121) are both very low and close to each other, providing strong evidence of the model's high generalization capability and absence of overfitting.

Fattahi and Babanouri [45] also employed a similar SVR model; however, they optimized the model parameters using a hybrid framework incorporating heuristic algorithms such as Differential Evolution (DE) [46], Artificial Bee Colony (ABC) [47], and the Gravitational Search Algorithm (GSA) [48]. In contrast, the present study performed hyperparameter tuning as a sequential search procedure, applying Grid Search, Random Search, and Bayesian Optimization consecutively, and selecting the optimal parameters obtained from the best-performing method. The SVR model trained with these optimized parameters exhibited significant improvements in test performance and enhanced consistency between training and test predictions. Therefore, compared to [45], the performance difference largely stems from the optimization strategy employed, indicating a stronger generalization capacity in the current model.

When comparing the two SVR configurations developed using the same dataset, the optimized parameter set obtained through the Grid–Random–Bayesian sequence in this study ($C = 16.32$, $\gamma = 0.000865$, $\epsilon = 0.111$) provides a lower complexity structure with a broader influence range, making it more robust to noise and yielding superior generalization capability. On the other hand, the parameter set reported in [45] ($C = 2605.86$, $\gamma \approx 0.067$, $\epsilon = 0.214$) contains an excessively large C value and a comparatively higher γ , which makes the model overly sensitive to the training data, increases the risk of overfitting, and ultimately limits its generalization performance.

Meanwhile, studies such as [22,23] and Afradi et al. [24] used different datasets and methodologies; therefore, performance comparisons can only be interpreted at a methodological level rather than directly. For instance, the SVR model developed in [22] using the Queens Water Tunnel dataset achieved high R^2 values of 0.99 for training and 0.95 for testing. However, the limited sample size and the use of eight input variables restricted the model's ability to generalize across varying geological conditions. Similarly, [23] compared ANN, SVM, and GEP models, reporting relatively lower performance for SVM. In these studies, SVR parameters were set as $C = 1000$, $\sigma = 0.5$ ($\gamma = 2$), and $\epsilon = 0.1$ in [23,24],

while in [22], ε was set to 0.1 and the other two parameters were determined through 10-fold cross-validation via trial-and-error. In other words, none of these studies employed a systematic optimization procedure for determining SVR hyperparameters.

To evaluate the impact of optimal SVR parameter selection on model performance, the parameter set reported in [45] ($C = 2605.86$, $\gamma = 0.067$, $\varepsilon = 0.111$) was applied to the same dataset using the SVR algorithm of the present study, and the resulting performance metrics were recalculated. A summary of the results is presented in Table 4.

Table 4. Comparative results demonstrating the impact of optimal SVR parameters on model performance.

Reference	Opt. Method	Dataset	MSE	MAE	RMSE	R ²	NSE
[45]	SVR-DE	Training	0.0302	0.1600	0.1737	0.8599	0.7549
		Testing	0.0299	0.1540	0.1730	0.9100	0.7836
		All	0.0301	0.1588	0.1736	0.8728	0.7655
This study	Bayesian	Training	0.0102	0.1008	0.0968	0.9164	0.9583
		Testing	0.1016	0.1008	0.0950	0.9292	0.9664
		All	0.0102	0.1008	0.0963	0.9209	0.9605

Examination of Table 4 shows that the optimal SVR parameters obtained through Bayesian optimization provide substantial improvements over those used in [45] across all performance metrics for the training, test, and complete datasets. In particular, the reductions in MSE, MAE, and RMSE, together with the increases in R^2 and NSE, clearly demonstrate that hyperparameter optimization significantly enhances the predictive capability of the SVR model.

Bayesian optimization reached the optimal parameter set within approximately 30–50 function evaluations, whereas Grid Search required nearly 1000 evaluations to achieve comparable results. Moreover, Bayesian optimization improved the test performance to $R^2 = 0.97$ while demonstrating superior efficiency in terms of the number of function evaluations and CPU time compared to the RS and GS algorithms. Consequently, the three optimization algorithms were applied sequentially, their results compared, and the optimal parameters yielding the highest model performance were selected for the final SVR model.

Although the influence of SVR hyperparameters may vary depending on the characteristics of the dataset, general tendencies are well established: γ strongly affects model complexity by controlling the effective radius of the RBF kernel; C adjusts the penalty for errors, thereby governing the model's flexibility; and ε defines the width of the insensitive zone, affecting the number of support vectors to a limited extent.

A key contribution of this study is the integrated treatment of the dataset, model inputs, and hyperparameter optimization strategy. The high R^2 and NSE values, along with very low RMSE and MAE values, demonstrate the model's robustness and strong generalization capability. Future studies may further enhance generalization performance by applying k-fold cross-validation.

As shown in Table 1, although baseline SVM performance is comparable to RF and BT, the SVR model exhibits significant improvements across all metrics once its hyperparameters are optimized using the proposed strategies. For example, the R^2 value increases from 0.728 to 0.966, while the MAE decreases from 0.131 to 0.095, clearly illustrating that carefully tuned hyperparameters are critical for improving model accuracy and stability.

A detailed examination of hyperparameter effects not only provides methodological insight for academic research but also facilitates the development of more reliable and computationally robust SVM models for TBM applications.

In conclusion, the comprehensive analysis of feature importance demonstrates that the contribution of this study extends beyond predictive performance, providing significant value for understanding underlying physical processes and for developing interpretable, engineering-relevant prediction frameworks. The findings offer both a methodological reference for future research and a practical basis for more reliable, computationally efficient, and physically meaningful TBM performance predictions.

6. Conclusions and Future Work

In this study, an SVM-based model was developed to predict the rate of penetration (ROP) during tunnel excavation, and hyperparameter optimization was systematically carried out using Grid Search (GS), Random Search (RS), and Bayesian Optimization (BO). The results indicate that BO reached the optimal parameter set with only 30–50 evaluations, whereas GS and RS required approximately 1000 evaluations. In addition, BO achieved the highest predictive accuracy ($R^2 = 0.9625$) while reducing the computational time from 25.83 s (GS) to 17.31 s. Compared with the baseline SVM model, the optimized SVR demonstrated high accuracy ($R^2 = 0.9610$ – 0.9625), strong stability (NSE = 0.9194–0.9231), and low error levels (MAE = 0.0927–0.1099), clearly highlighting the critical role of hyperparameter optimization in improving model performance.

One of the key contributions that distinguishes this study from previous work in the literature is not only the systematic execution of hyperparameter optimization but also the detailed examination of Model-Based Feature Importances. The variable importance values presented in Table 3 consistently indicate that BI, ALPHA, and DPW are the most influential parameters on ROP. This finding aligns with the physical processes governing the excavation mechanism and demonstrates that the model successfully captures the causal relationships within the data. Moreover, the relatively higher importance of UCS in the SVM and GAM highlights the deterministic influence of strength parameters. The similarity of feature importance trends across different models confirms that the prediction process is both mathematically and physically coherent.

Therefore, this study demonstrates that combining hyperparameter optimization with feature importance analysis enables the development of more reliable and interpretable models for TBM penetration rate prediction. This holistic approach not only provides a methodological reference for academic research in geotechnical engineering but also establishes a robust predictive framework for practical applications.

Future studies will aim to enhance the generalization capacity of the model by using larger and more up-to-date datasets representing different geological conditions and excavation machines. In addition, investigating how variable importances change across different geological environments will further strengthen the interpretability of SVM and other machine learning models.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/app16010355/s1>, Table S1: Supplementary Materials.

Author Contributions: Conceptualization, H.K. and D.A.; methodology, H.K.; software, H.K.; validation, H.K. and D.A.; formal analysis, H.K. and D.A.; investigation, H.K. and D.A.; resources, H.K. and D.A.; data curation, H.K. and D.A.; writing—original draft preparation, H.K. and D.A.; writing—review and editing, H.K. and D.A.; visualization, H.K.; supervision, H.K.; project administration, H.K.; funding acquisition, D.A. All authors have read and agreed to the published version of the manuscript.

Funding: No external funding was received for this study.

Data Availability Statement: The original contributions presented in this study are included in the Supplementary Material.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tarkoy, P.J. *Rock Hardness Index Properties and Geotechnical Parameters for Predicting Tunnel Boring Machine Performance*; University of Illinois at Urbana-Champaign: Champaign, IL, USA, 1975.
2. Nelson, P. *Tunnel Boring Machine Performance in Sedimentary Rocks*; Cornell University: Ithaca, NY, USA, 1983.
3. Yağız, S. Utilizing Rock Mass Properties to Predict TBM Performance in Hard Rock Conditions. *Tunn. Undergr. Space Technol.* **2008**, *23*, 326–339. [CrossRef]
4. Yağız, S.; Karahan, H. Prediction of Hard Rock TBM Penetration Rate Using Particle Swarm Optimization. *Int. J. Rock Mech. Min. Sci.* **2011**, *48*, 427–433. [CrossRef]
5. Yağız, S.; Karahan, H. Application of Various Optimization Techniques to Predict TBM Penetration Rate in Rock Mass and Comparison of Their Performance. *Int. J. Rock Mech. Min. Sci.* **2015**, *80*, 308–315. [CrossRef]
6. Benardos, A.G.; Kaliampakos, D.C. Modeling TBM Performance Using Artificial Neural Networks. *Tunn. Undergr. Space Technol.* **2004**, *19*, 597–605. [CrossRef]
7. Armaghani, D.J.; Koopialipoor, M.; Marto, A.; Yagiz, S. Application of Various Optimization Techniques to Predict TBM Advance Rate in Granitic Rocks. *J. Rock Mech. Geotech. Eng.* **2019**, *11*, 779–789. [CrossRef]
8. Koopialipoor, M.; Fahimifar, A.; Ghaleini, E.N.; Momenzadeh, M.; Armaghani, D.J. Development of a New Hybrid ANN to Solve a Geotechnical Problem Related to Tunnel Boring Machine Performance. *Comput. Geotech.* **2020**, *36*, 345–357.
9. La, Y.S.; Kim, M.I.; Kim, B. Prediction of Shield TBM Disc Cutter Replacement Time Using SVM. *J. Korean Tunn. Undergr. Space Assoc.* **2019**, *21*, 641–656.
10. Afradi, A.; Ebrahimabadi, A.; Hallajian, T. Prediction of TBM Penetration Rate Using Support Vector Machine. *Geosaberes Rev. Estud. Geoeducacionais* **2020**, *11*, 467–479. [CrossRef]
11. Liu, F.; Yang, M.; Ke, J. TBM Tunnel Prediction Method Surrounding Rock Classification Based on LSTM-SVM. *Adv. Mech. Eng.* **2024**, *16*, 16878132241255209. [CrossRef]
12. Minh, V.T.; Katushin, D.; Antonov, M.; Veinthal, R. Regression Models and Fuzzy Logic Prediction of TBM Penetration Rate. *Open Eng.* **2017**, *7*, 60–68. [CrossRef]
13. Arbabsiar, M.H.; Farsangi, M.A.E.; Mansouri, H. Fuzzy Logic Modeling to Predict the Level of Geotechnical Risks in Tunnel Boring Machine (TBM) Tunneling. *Rud.-Geološko-Naft. Zb.* **2020**, *35*, 1–14.
14. Afradi, A.; Ebrahimabadi, A.; Hallajian, T. Prediction of TBM Penetration Rate Using Fuzzy Logic, Particle Swarm Optimization, and Harmony Search Algorithm. *Geotech. Geol. Eng.* **2022**, *40*, 1513–1536. [CrossRef]
15. Cortes, C.; Vapnik, V. Support-vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
16. Drucker, H. Improving regressors using boosting techniques. In *Proceedings of the 14th International Conference on Machine Learning (ICML '97)*; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1997; Volume 97, p. e115.
17. Drucker, H.; Wu, D.; Vapnik, V.N. Support vector machines for spam categorization. *IEEE Trans. Neural Netw.* **1999**, *10*, 1048–1054. [CrossRef]
18. Drucker, H.; Shahrari, B.; Gibbon, D.C. Support vector machines: Relevance feedback and information retrieval. *Inf. Process. Manag.* **2002**, *38*, 305–323. [CrossRef]
19. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [CrossRef]
20. Tian, Y.; Shi, Y.; Liu, X. Recent advances on support vector machines research. *Technol. Econ. Dev. Econ.* **2012**, *18*, 5–33. [CrossRef]
21. Schölkopf, B.; Smola, A. Support vector machines and kernel algorithms. In *Encyclopedia of Biostatistics*; Wiley: Hoboken, NJ, USA, 2005; pp. 5328–5335.
22. Mahdevari, S.; Shahriar, K.; Yagiz, S.; Shirazi, M.A. A Support Vector Regression Model for Predicting Tunnel Boring Machine Penetration Rates. *Autom. Constr.* **2014**, *72*, 214–229. Available online: <https://www.sciencedirect.com/science/article/abs/pii/S1365160914002536> (accessed on 25 December 2025). [CrossRef]
23. Afradi, A.; Ebrahimabadi, A. Comparison of ANN, SVM, and GEP Approaches for Predicting TBM Penetration Rate. *SN Appl. Sci.* **2020**, *2*, 2004.
24. Afradi, A.; Ebrahimabadi, A.; Hedayatzadeh, M. Performance Prediction of a Hard Rock TBM Using Statistical and Artificial Intelligence Methods. *Min. Environ. J.* **2024**, *15*, 323–343.
25. Maier, H.R.; Dandy, G.C. Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications. *Environ. Model. Softw.* **2000**, *15*, 101–124. [CrossRef]
26. Karahan, H.; Ayvaz, M.T. Prediction of Aquifer Parameters Using Artificial Neural Networks. *J. Porous Media* **2006**, *9*, 429–444. [CrossRef]

27. Karahan, H.; Çetin, M.; Can, M.E.; Alsenjar, O. Development of a New ANN Model to Predict Daily Actual Evapotranspiration Using Limited Climate Data and Remote Sensing Techniques for Sustainable Water Management. *Sustainability* **2024**, *16*, 2481. [CrossRef]
28. Karahan, H.; Erkan Can, M. A New Method to Predict Nitrate Concentration Levels in Irrigation Fields for Sustainable Agriculture. *Agriculture* **2025**, *15*, 161. [CrossRef]
29. Karahan, H.; İplikci, S.; Yaşar, M.; Gurarslan, G. River Flow Prediction from Upstream Records Using Support Vector Machines. *J. Appl. Math.* **2014**, *2014*, 714213. [CrossRef]
30. Salimi, A.; Moormann, C.; Singh, T.; Jain, P. TBM performance prediction in rock tunneling using various artificial intelligence algorithms. In Proceedings of the 11th Iranian and 2nd Regional Tunnelling Conference Tunnels and the Future, Tehran, Iran, 2–5 November 2015.
31. Hu, M.; Li, B.; Zhang, B.; Wang, R.; Chen, L. Improved SVR method for predicting the cutting force of a TBM cutter using linear cutting machine test data. *KSCE J. Civ. Eng.* **2021**, *25*, 4425–4442. [CrossRef]
32. Liu, B.; Wang, R.; Guan, Z.; Li, J.; Xu, Z.; Guo, X.; Wang, Y. Improved support vector regression models for predicting rock mass parameters using tunnel boring machine driving data. *Tunn. Undergr. Space Technol.* **2019**, *91*, 102958. [CrossRef]
33. Bergstra, J.; Bengio, Y. Random Search for Hyper-Parameter Optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
34. Snoek, J.; Larochelle, H.; Adams, R.P. Practical Bayesian Optimization of Machine Learning Algorithms. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 2951–2959.
35. Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R.P.; De Freitas, N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *IEEE Proc.* **2015**, *104*, 148–175. [CrossRef]
36. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, UK, 2016.
37. Bahri, M.; Romero-Hernández, R.; Mascort-Albea, E.J.; Soriano-Cuesta, C.; Jaramillo-Morilla, A. Predicting Maximum Surface Displacement from Mechanized Twin Tunnel Excavation in Seville Using Machine Learning and FLAC3D Simulation. *Geotech. Geol. Eng.* **2025**, *43*, 70. [CrossRef]
38. Breiman, L.; Cutler, A.; Liaw, A.; Wiener, M.; Liaw, M.A. *Package ‘Randomforest’*; University of California, Berkeley: Berkeley, CA, USA, 2018; Volume 81, pp. 1–29.
39. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
40. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
41. Hastie, T.; Tibshirani, R. Generalized additive models: Some applications. *J. Am. Stat. Assoc.* **1987**, *82*, 371–386. [CrossRef]
42. Marra, G.; Wood, S.N. Practical variable selection for generalized additive models. *Comput. Stat. Data Anal.* **2011**, *55*, 2372–2387. [CrossRef]
43. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [CrossRef]
44. Guyon, I.; Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.
45. Fattahi, H.; Babanouri, N. Applying optimized support vector regression models for prediction of tunnel boring machine performance. *Geotech. Geol. Eng.* **2017**, *35*, 2205–2217. [CrossRef]
46. Storn, R.; Price, K. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* **1997**, *11*, 341–359. [CrossRef]
47. Karaboga, D. An İdea Based on Honey Bee Swarm for Numerical Optimization. 2005. Available online: https://abc.erciyes.edu.tr/pub/tr06_2005.pdf (accessed on 25 December 2025).
48. Rashedi, E.; Nezamabadi-Pour, H.; Saryazdi, S. GSA: A gravitational search algorithm. *Inf. Sci.* **2009**, *179*, 2232–2248. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Horizontal PGA Estimates for Varying Deep Geological Conditions—A Case Study of Banja Luka

Borko Bulajić ¹, Silva Lozančić ², Senka Bajić ^{1,*}, Anka Starčev-Ćurčin ¹, Miloš Šešlija ¹, Miljan Kovačević ³ and Marijana Hadzima-Nyarko ²

¹ Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia; borkobulajic@uns.ac.rs (B.B.); astarcev@uns.ac.rs (A.S.-Ć.); sele@uns.ac.rs (M.Š.)

² Faculty of Civil Engineering and Architecture Osijek, Josip Juraj Strossmayer University of Osijek, Vladimir Prelog St. 3, 31000 Osijek, Croatia; lozancic@gfos.hr (S.L.); mhadzima@gfos.hr (M.H.-N.)

³ Faculty of Technical Sciences, University of Pristina, Knjaza Milosa 7, 38220 Kosovska Mitrovica, Serbia; miljan.kovacevic@pr.ac.rs

* Correspondence: senka.bajic@uns.ac.rs

Abstract: In this study, the city of Banja Luka is used as a case study to evaluate horizontal PGA values in regions with a history of moderate to strong earthquakes and with different deep geological conditions. We present regional attenuation equations for PGA that can capture both the impacts of deep geology and local soil conditions. A PSHA study for a site in Banja Luka was carried out using the developed empirical scaling equations and compared to all previous seismic hazard estimations for the same region. The data indicate that variations in deep geological conditions may have a greater impact on PGA values than local soil effects. Given the scarcity of scaling equations that consider deep geology in addition to local soil conditions, we believe this case study is a step toward developing more accurate PGA estimates for comparable regions.

Keywords: PGA; low to medium seismicity regions; regional attenuation equations; deep geology; local soil conditions

1. Introduction

Peak ground acceleration (PGA) is still a very important parameter in earthquake engineering [1–3], and it is imperative to estimate it as reliably as possible. Most empirical formulae for scaling PGA values only take into account the effects of local soil, ignoring the effects of deep geology [4]. There are, however, strong motion studies [5–8] that show that evaluating the severity of surface ground motion requires taking into account both deep geology and local soil conditions to minimize bias. The 2004 version of Eurocode 8 [9] also acknowledges the significance of deeper geological conditions, stating in Clause 3.1.2(1) that the National Annex may specify the classification scheme that will account for deep geology. The term “deep geology” refers to geological site surroundings on a scale of a few hundreds of meters or even kilometers [10], whereas “local soil” refers more to a geotechnical description on a scale of a few tens of meters, usually down to the first layer, with $V_s \geq 800$ m/s.

Furthermore, most GMPEs (Ground Motion Prediction Equations) classify local soil conditions solely based on the top 30 m of the stratigraphic profile, regardless of research showing that the average shear-wave velocity of the top 30 m does not correlate with a site resonance [11,12], particularly for deep soil sites [13]. This method is still extensively employed, because data on deep geology is limited, and investigating a soil profile to

depths of at least 100 m would be prohibitively costly for the design of common civil engineering structures.

In Bosnia and Herzegovina, seismic hazard maps for use in the application area of Eurocode 8 [14,15] are defined for ground type A, which is defined as a rock or “other rock-like geological formation” with up to 5 m of weaker surface material and $V_S > 800$ m/s in the top 30 m of the soil profile. The PGA value from the official hazard map is then multiplied by the soil factor, S , to obtain the PGA at the top of local soil. Table 1 shows how S varies by ground type. S is additionally dependent on the type (Type 1 or Type 2) of spectrum, which is determined by the magnitude of the “most contributing earthquakes” [9]. It should be noted that Eurocode 8 defines the magnitudes of the most significant earthquakes using the M_S scale. However, as the majority of modern GMPEs and seismic source models utilize the M_W magnitude, the majority of seismic hazard maps in Europe were created using this magnitude. Furthermore, recent studies have demonstrated the M_W scale’s significant shortcomings and support the implementation of the M_{wg} scale in its place [16–18]. Since our hazard analyses utilize a European source zone model defined by other authors (see Section 4), converting magnitude scales will be outside the purview of this research. We do, however, intend to include magnitude conversion of the seismicity data into our upcoming studies.

Table 1. PGA for different ground types according to Eurocode 8 [9], where a_g represents the PGA value indicated on the accompanying seismic hazard map for the investigated site.

Eurocode 8 [9] Ground Types	Type 1 Spectrum: “Most Contributing” Earthquakes with $M_S > 5.5$	Type 2 Spectrum: “Most Contributing” Earthquakes with $M_S \leq 5.5$
Ground type A, $V_{S,30} > 800$ m/s “Rock, at the surface up to 5 m of weaker material.”	a_g	a_g
Ground type B, $V_{S,30} = 360$ – 800 m/s “At least several tens of meters thick deposits. Very dense sand, gravel, or very stiff clay.”	$a_g \times 1.2$	$a_g \times 1.35$
Ground type C, $V_{S,30} = 180$ – 360 m/s “Deep deposits, several tens of meters up to hundreds of meters thick. Dense or medium dense sand, gravel, or stiff clay.”	$a_g \times 1.15$	$a_g \times 1.5$
Ground type D, $V_{S,30} < 180$ m/s “Deposits. Loose-to-medium cohesionless soil, or predominantly soft-to-firm cohesive soil.”	$a_g \times 1.35$	$a_g \times 1.8$
Ground type E “Alluvium layer at the surface, between 5 and 20 m thick, above stiffer material. $V_{S,30} < 360$ m/s.”	$a_g \times 1.4$	$a_g \times 1.6$

The goal of this study was to estimate PGA values at different deep geological conditions for regions with moderate to strong earthquakes. The case-study region encompasses the city of Banja Luka, Bosnia and Herzegovina, for which our associated studies [19–21] examined vertical PGA values as well as uniform hazard spectral estimates.

We calculated the seismic hazard for a location in Banja Luka using regional GMPEs for PGA and compared the results to all previous and current estimates of PGA and macroseismic intensity. In this study, local soil will be classified in compliance with Seed et al. [22,23]. The same local soil classification was used in former Yugoslavia to define site conditions for recorded strong motion data and to develop regional GMPEs [24–26]. According to this classification, “rock” soil sites are those with $V_S > 800$ m/s, stiff soil has a 15 to 75 m thick soil layer atop that with $V_S > 800$ m/s, while deep soil has a soil layer with

a thickness of more than 100 m. For classifying the deep geology conditions, we will use the classification of Trifunac and Brady [10], who divided deep geology into three types: the geological rock, sediments, and complex or intermediate geological site surroundings.

2. Area Surrounding the Case-Study Site—Seismicity and Geology

2.1. Seismicity of the Area Surrounding the Analyzed Location in Banja Luka

The area surrounding Banja Luka is characterized by intraplate seismicity. The city of Banja Luka is located between the south-west part of the Pannonian basin, with a rare occurrence of stronger earthquakes [27,28] and the Dinaric Alps—see the top left plot in Figure 1. The Adriatic Platform and Dinarides collide, causing most regional earthquakes to occur along the Adriatic coast (see Figure 1). Moho depths in the western Balkans range from 25 km in the area of the Pannonian basin to 45 km below the Dinarides [29,30].

The metropolitan area of Banja Luka encompasses the Banja Luka valley and surrounding hills. The area has deep fault zones that are visible in the terrain [31–33]. The largest MCS intensity assessed in Banja Luka was VIII, caused by the $M_W = 6.1$ Banja Luka earthquake on 27 October 1969.

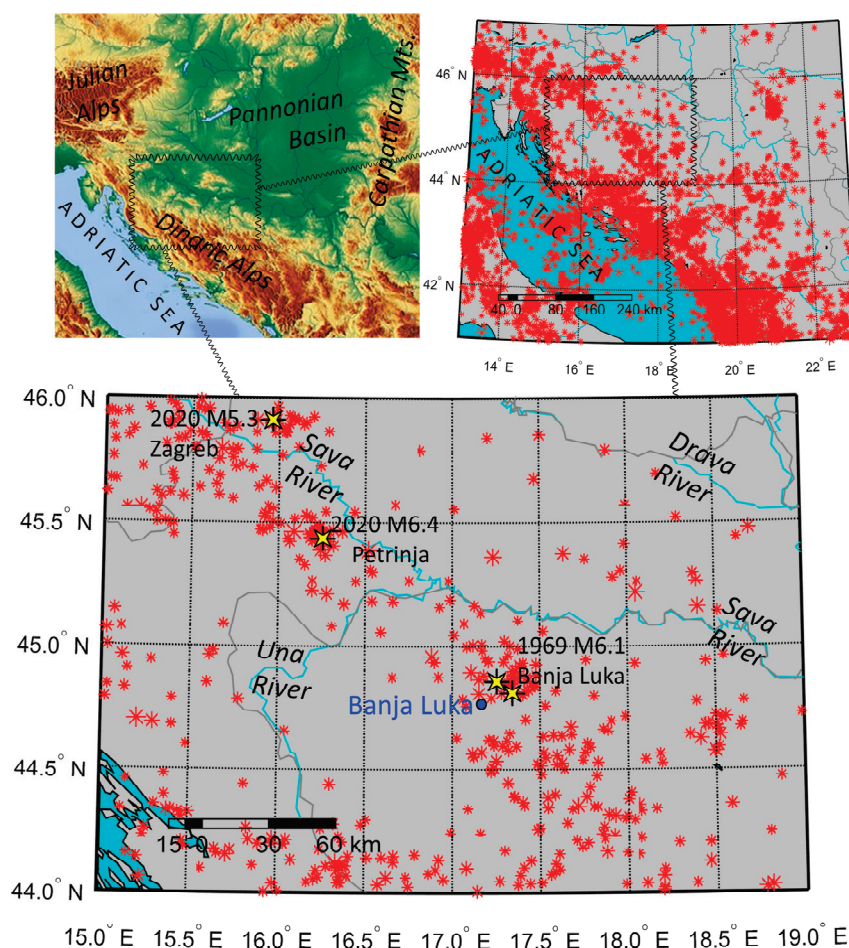


Figure 1. (Top left) Dinaric Alps and Pannonian basin, (top right) epicenters of the earthquakes with $M_W \geq 3$ recorded between 1900 and 2025 [34], (bottom) the analyzed site (solid blue circle) and the epicenters of the largest historical earthquake in the vicinity of Banja Luka and two recent destructive earthquakes in neighboring Croatia.

The 1969 earthquakes began with a series of small-magnitude earthquakes between late 1968 and early 1969. This activity resumed in September and culminated with the largest foreshock on 26 October 1969 ($M_W = 6.1$ and $M_L = 6.0$, epicentral intensity of VII–VIII

°MCS) and the main shock the next day (27 October 1969, $M_W = 6.1$ and $M_L = 6.4$, epicentral intensity of VIII–IX °MCS). The earthquakes' impact on Banja Luka (over 1117 people were injured and 15 were killed [32]) would have been even greater if each of the schools, firms, and private as well as public institutions had not been closed following the foreshock and if authorities had not prevented residents from returning to their homes [35]. The major earthquake was felt across a vast area. The VII °MCS isoseismal encompassed an area of 9000 km² [32], and 60% of all buildings were irreparably damaged, including the majority of schools. Over 76,000 people were rendered homeless, and financial losses were significant.

In neighboring Croatia, an $M_W = 5.3$ earthquake hit Croatia's capital, the city of Zagreb, on 22 March 2020, with the epicenter only 7 km north [36] and a hypocentral depth of 10 km. This was the strongest event recorded near Zagreb in the past 140 years—an $M_W = 6.3$ earthquake was recorded in 1880. Later that very year, on 29 December 2020, yet another catastrophic earthquake hit Croatia, this time near the city of Petrinja, around 100 km northwest of Banja Luka and 40 km south of Zagreb, with $M_W = 6.4$ [37]. Although the maximum intensity of the 2020 Zagreb earthquake was not larger than VII, it still caused significant damage to the historical city center, with over 1900 buildings rendered uninhabitable as a result of the catastrophe. There were 27 people injured, and one of them died as a result of their injuries. Interestingly, even though the 1880 Zagreb earthquake was thought to be stronger (maximum felt intensity, VIII–IX) than the 2020 one, and the population within today's city limits increased tenfold (from 80,000 to 800,000), the effects of the 1880 event were strikingly similar: approximately 1800 buildings were damaged, 29 people were seriously injured, and one person died as a result of injuries.

Today, the wider area of Banja Luka has a population of around 180,000 inhabitants. Although many buildings in the city center were built after the 1969 earthquake in compliance with earthquake-resistant design codes, in the wider area of Banja Luka, there are still many older and vulnerable buildings. Some of the recent regional seismic risk studies [38–42] reveal a significant proportion of buildings older than the 1960s in the entire region of former Yugoslavia (the first earthquake-resistant code was enacted in 1964 [43]). Similar structures are common in other countries of the Mediterranean and may sustain severe damage during moderate to large earthquakes [44–48].

2.2. Geological Surroundings of the Analyzed Location in Banja Luka

Banja Luka is situated on the south-west boundary of the Pannonian basin and the north boundary of the Dinaric Alps (see Figure 1). The basin encompasses parts of nine different countries, including northeast Croatia. The deep geological sediments of the Pannonian basin date back to the Pannonian Sea and have a thickness of up to several kilometers [49]. When the Pannonian Sea emptied (after reaching its maximum geographical extent during the Pliocene), the basin's lowlands formed the remaining plain. The basin represents a back-arc basin filled with sediments, and it spread during the Miocene [50].

In Banja Luka, besides geological sediments, there are also scattered outcrops of geological rock and areas of complex deep geological strata. In the vicinity of the analyzed location, the geological sediments have thicknesses of up to several hundred meters and consist of marls, clays, and sands, with a river terrace atop it, while a spilite formation represents geological rock [51]. There is also a diabase-cherty formation (or metamorphic ophiolitic melange), which can be classified as intermediate deep geology [10]. For a more detailed geological description of the Banja Luka region, please refer to [21,51–53].

2.3. Official PGA Estimates for Banja Luka

The ex-SFRY issued its first seismic-resistant design code in 1964 [43]. There was a seismic zoning map developed in 1950 based on the maximum (MCS—Mercalli–Cancani–Sieberg Scale) intensities observed at the time, which was utilized in conjunction with the 1964 code. In 1981, a new earthquake-resistant design code was approved [54], and a year later, a seismic zoning map was developed to be used temporarily with the new code [55]. The 1982 map was based on the maximum measured MCS intensities once more. In 1990, the 1981 code was modified with six new seismic zoning maps (for 50, 100, 200, 500, 1000, and 10,000 year return periods) [56]. Although the seismic hazard was again expressed in MCS intensity and for average soil conditions, these were the first seismic hazard maps in ex-Yugoslavia created utilizing the PSHA (Probabilistic Seismic Hazard Assessment) approach. Please take note that we do not have return periods for the 1950 and 1982 maps, because they were developed using the largest observed intensities rather than a probabilistic hazard analysis.

Table 2 shows MCS intensities for Banja Luka, as given in 1950, 1982, and 1990 official hazard maps. Table 2 also shows the corresponding empirical PGA estimates, calculated using the following equation [24]:

$$\log(\text{PGA}) = -0.079 + 0.290 \cdot I \pm P \cdot \sigma, \sigma = 0.049, \quad (1)$$

where PGA is the horizontal peak ground acceleration [cm/s^2], I is the macroseismic intensity [$^\circ\text{MCS}$], σ represents the standard deviation, and $P = 0$ for median estimates. For each intensity degree, Table 2 shows the PGA values calculated by Equation (1) for $P = -1$ and $P = 1$.

Table 2. Macroseismic (MCS) intensity degrees given in the official seismic zoning maps of former Yugoslavia [43,54–56] for the selected location in Banja Luka, together with the associated median – σ and median + σ empirical PGA estimates computed using Equation (1). The PGA values of the 2018 official maps, which were used in conjunction with Eurocode 8 [14], are also shown.

Map	I [$^\circ\text{MCS}$]	PGA
1950	VII	0.081–0.102 g
1982	VIII	0.159–0.199 g
1990—50 years	VII	0.081–0.102 g
1990—100 years	VIII	0.159–0.199 g
1990—200 years	IX	0.309–0.388 g
1990—500 years	IX	0.309–0.388 g
1990—1000 years	IX	0.309–0.388 g
1990—10,000 years	IX	0.309–0.388 g
2018—95 years	VI–VII	0.080 g
2018—475 years	VIII	0.170 g

In 2018, two new hazard maps were included in Bosnia and Herzegovina’s National Annex to Eurocode 8 [14]. The maps were created with the PSHA approach. The first map is for a 95-year return period, or a 10% chance of exceeding in 10 years, while the second is for a 475-year return period, or a 10% chance of exceeding within 50 years. The 95-year map correlates to Eurocode 8’s “damage limitation” requirement, whilst the 475-year map corresponds to the “no-collapse” requirement. The 2018 maps were generated for ground type A (see Table 1), and on these two maps, it is the horizontal PGA that represents seismic hazard numerically. Furthermore, both maps disregard how deep geology affects surface ground motion. On these maps, the horizontal PGA values for the analyzed location in Banja Luka are 0.08 g for the 95-year and 0.17 g for the 475-year return period. We added these values in Table 2, together with the corresponding intensities estimated using

Equation (1). Table 2 shows that official hazard estimates from 1950 and 1982 are similar to those from 2018. However, the empirical PGA estimates based on the intensities provided in the ex-Yugoslav 1990 hazard map for the 500-year return period are substantially larger than the 475-year PGA value from the 2018 hazard map, even after multiplying it by the greatest soil factors from Table 1. We presume that following the disastrous 1969 Banja Luka earthquake, the hazard level for Banja Luka was purposefully increased several-fold on the 100-, 200-, and 500-year official maps from 1990 [56] to assure better seismic safety in the future.

3. GMPEs for Horizontal PGA and Varying Deep Geology

The northwestern Balkans are one of the few areas in the entire world with available information on deep geological conditions from numerous recording stations. Several researchers used those regional strong motion data to develop GMPEs that take into consideration both local soil (up to 100 m in depth) and deep geology [7,25,26]. The GMPEs were employed in a series of recent microzonation studies and shown to be in great agreement with recorded macroseismic intensities [6,52,57–65].

We are now going to present regional empirical equations for estimating PGA values that account for both local soil conditions and deep geological site environment. The attenuation equations will be defined in this mathematical form:

$$\log[\text{PGA}] = c_1 + c_2 \cdot M + c_3 \cdot \log(\sqrt{R^2 + R_0(T)^2}) + c_4 \cdot S_{L1} + c_5 \cdot S_{L2} + c_6 \cdot S_{G1} + c_7 \cdot S_{G2} + \sigma_{\log} \cdot P. \quad (2)$$

In Equation (2), PGA denotes horizontal peak ground accelerations (in g), M is the earthquake magnitude (refer to [66] for more information on magnitude type), and R is the epicentral distance. Equations for the hypocentral distance were also developed in our previous studies [57–60]. S_{L1} and S_{L2} are categorical variables for local soil, while S_{G1} and S_{G2} are for deep geological conditions. Table 3 summarizes these categorical variables. In Table 3, we use the same values for the categorical (dummy) variables (S_{L1} , S_{L2} , S_{G1} , and S_{G2}) and site parameters (s_L and s) as defined in [8]. We assumed that the data had a log-normal distribution. Consequently, $\sigma_{\log}$ is the standard deviation of the common logarithm of PGA, and ε is equal to 0 for the median estimate and ± 1 for the median ± 1 $\sigma_{\log}$ estimates.

Table 3. Categorical variables used in this study for different types of local soil and deep geological site settings.

Local Soil Parameters	Local Soil Categorical Variables	Deep Geology Parameters	Deep Geology Categorical Variables
“Rock” soil sites: $s_L = 0$	$S_{L1} = S_{L2} = 0$	Geological rock: $s = 2$	$S_{G1} = S_{G2} = 0$
Stiff soil sites: $s_L = 1$	$S_{L1} = 1$ and $S_{L2} = 0$	Intermediate (or complex) sites: $s = 1$	$S_{G1} = 1$ and $S_{G2} = 0$
Deep soil sites: $s_L = 2$	$S_{L1} = 0$ and $S_{L2} = 1$	(Deep geological) sediments: $s = 0$	$S_{G1} = 0$ and $S_{G2} = 1$

The database for prediction equations includes 436 horizontal components of strong-motion accelerograms from 112 earthquakes with a magnitude of $3 < M \leq 6.8$. Our analysis is predominantly based on the EQINFOS database [67], which contains records from earthquakes in the northwest Balkan region. The authors of the database used the

M_L magnitude as defined by Lee et al. [66] for all events with a magnitude smaller than around 6.0, while only for the strongest earthquakes, they used the M_S magnitude. They purposefully chose the local magnitude, M_L , for the majority of events, since it was thought to be most suited for the so-called “strong motion”—the frequency content between 0.1 and 30 Hz that civil engineers are most interested in [68]. According to a subsequent study that converted the M_L magnitudes of the earthquakes in Croatia and the adjacent areas to moment magnitudes (M_W) [69], the difference between M_L and M_W was less than 2% for magnitudes between 3.5 and 6.5. Furthermore, Scordilis [70] found that the gap between M_S and M_W magnitudes is roughly 0.2% for magnitudes between 6.2 and 7.0. Therefore, we believe that while accounting for other limitations of the provided database, we can use our GMPEs in seismic hazard calculations with the seismic source zones identified by using the M_W -based seismicity data. Every one of the accelerograms in our database was recorded in the northwestern Balkans. The majority (418) were recorded between 1976 and 1987 [67,71,72], with the remaining (18) in 2010. Maps that show the locations of recording stations and the epicenters of contributing earthquakes can be found in our previous papers [7,58,59].

MATLAB® (version 8.5) scripts were used for calculating the GMPE scaling coefficients $c_1, c_2, c_3, c_4, c_5, c_6, c_7$, and $\sigma_{\log}$ by multiple linear regression analyses. For maximizing the R^2 statistics, R_0 was iteratively adjusted. Because around two thirds of the data came from distances shorter than around 30 km, we performed a supplementary analysis that only included data collected at epicentral distances of less than 30 km. This was conducted with the intention of determining whether more reliable GMPEs might be obtained for shorter source-to-site distances. Finally, because the regression was also performed for pseudo-spectral accelerations across different vibration periods, the scaling coefficients were smoothed using the MATLAB® function “smooth” and the weighted linear least squares and a second-degree polynomial model.

The finalized GMPE is as follows:

$$\begin{aligned} \log[\text{PGA}] = & -1.2957 + 0.3946 \cdot M - 1.3818 \cdot \log(\sqrt{R^2 + 19.5^2}) + 0.1772 \cdot S_{L1} \\ & - 0.0953 \cdot S_{L2} - 0.1469 \cdot S_{G1} - 0.1059 \cdot S_{G2} + 0.2691 \cdot P, \end{aligned} \quad (3)$$

The GMPE based on data from epicentral distances shorter than 30 km is as follows:

$$\begin{aligned} \log[\text{PGA}] = & 3.1868 + 0.3919 \cdot M - 3.8728 \cdot \log(\sqrt{R^2 + 40.0^2}) + 0.1276 \cdot S_{L1} - 0.0953 \\ & \cdot S_{L2} - 0.2229 \cdot S_{G1} - 0.0899 \cdot S_{G2} + 0.2675 \cdot P. \end{aligned} \quad (4)$$

By comparing Equations (3) and (4), we can observe that using only data recorded at distances less than 30 km results in a slightly smaller standard deviation.

Figure 2 displays the attenuation of PGA as a function of distance, as calculated with Equations (3) and (4). The coefficients pertaining to S_L and S_G can be used to compute the variations in values calculated by the same equation under different site conditions (refer to Table 3). For example, the ratio between PGA at the geological rock ($s = 2$) and intermediate deep geology sites ($s = 1$) is calculated as $1/10^{c_6}$, while the ratio between the geological rock ($s = 2$) and deep geological sediments ($s = 0$) is calculated as $1/10^{c_7}$. If we use Equation (3), the PGA estimates at the geological rock will be $1/10^{-0.1469} = 1.40$ times larger than at intermediate sites and $1/10^{-0.1059} = 1.28$ times bigger than at geological sediments. The fact that short-period waves travel faster through harder and more compact geological formations (like basalts and granites) than through sediments may explain the amplification at geological rock sites.

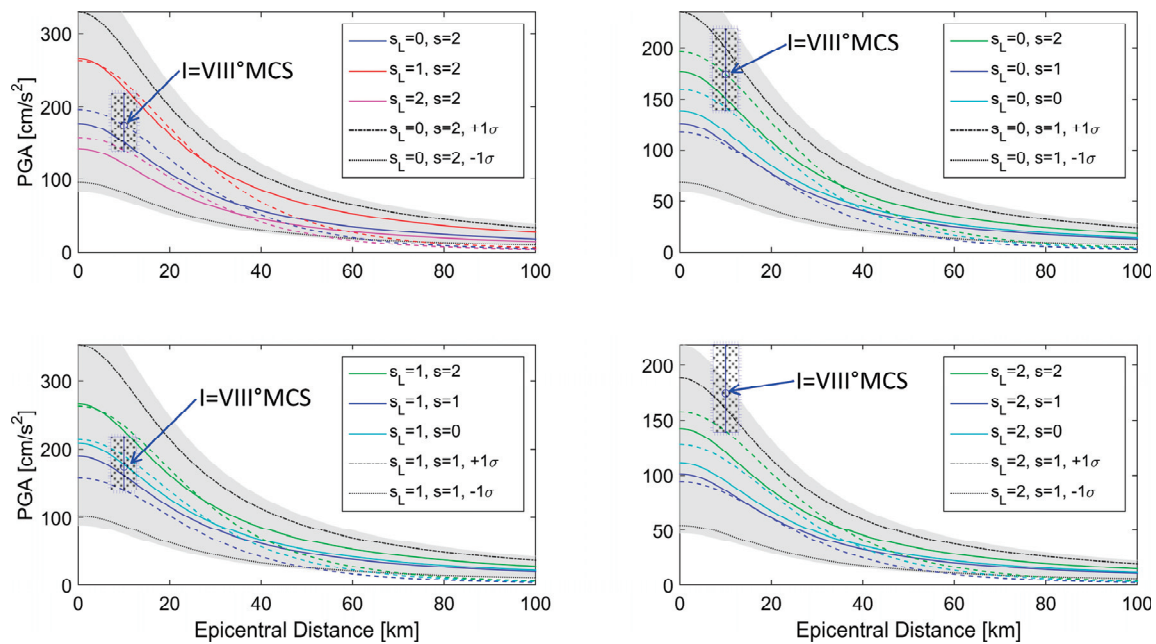


Figure 2. Empirical GMPEs for horizontal PGA, calculated using Equations (3) (solid lines) and (4) (dashed lines) for $M = 6.0$ and different combinations of local soil types and deep geology (see Table 3). The gray area shows the 95% confidence interval (median $\pm 2 \cdot \sigma_{\log}$ empirical estimates) for $s_L = 2$.

If the deep geology conditions remain constant, we can calculate ratios between the PGA values at stiff soil ($s_L = 1$) and “rock” soil ($s_L = 0$) by calculating 10^{c^4} , while the PGA ratios between deep soil ($s_L = 2$) and “rock” soil ($s_L = 0$) is calculated as 10^{c^5} . If Equation (3) is used, the PGA value for stiff soil sites is $10^{0.1772} = 1.50$ times larger than for “rock” soil sites. Horizontal PGA for deep soil sites will be $10^{-0.0953} = 0.80$ times the PGA for “rock” soil, according to the same equation. Put another way, horizontal short-period seismic waves at the surface will be less amplified (or even de-amplified) at deep soil sites than at “rock” soil sites. This implies that the energy dissipation of short-period waves in deep soils may outweigh a local soil amplification generated by the impedance difference between the deep soil and the harder rocks beneath.

What should be noted here, regardless of the type of deep geological formation, is that the effects of deep geology are comparable in scale to those of local soil and must not be overlooked.

Figure 2 illustrates PGA ranges (median $\pm$ one standard deviation) calculated with Equation (1) and VIII °MCS, which was the estimated predominant intensity in Banja Luka during the 1969 earthquake with the epicenter around 10 km from the city center. We can observe that only the combined effects of stiff soil sites and deep geological rock can result in median empirical PGA values comparable to the ones associated with VIII °MCS.

Figure 3 compares the horizontal PGA values recorded in Banja Luka at four separate accelerograph stations [67] with the median and median $\pm \sigma_{\log}$ empirical PGA estimates that were computed using Equation (3). As seen in Figure 3, the recorded horizontal PGA values and the empirical predictions coincide very well.

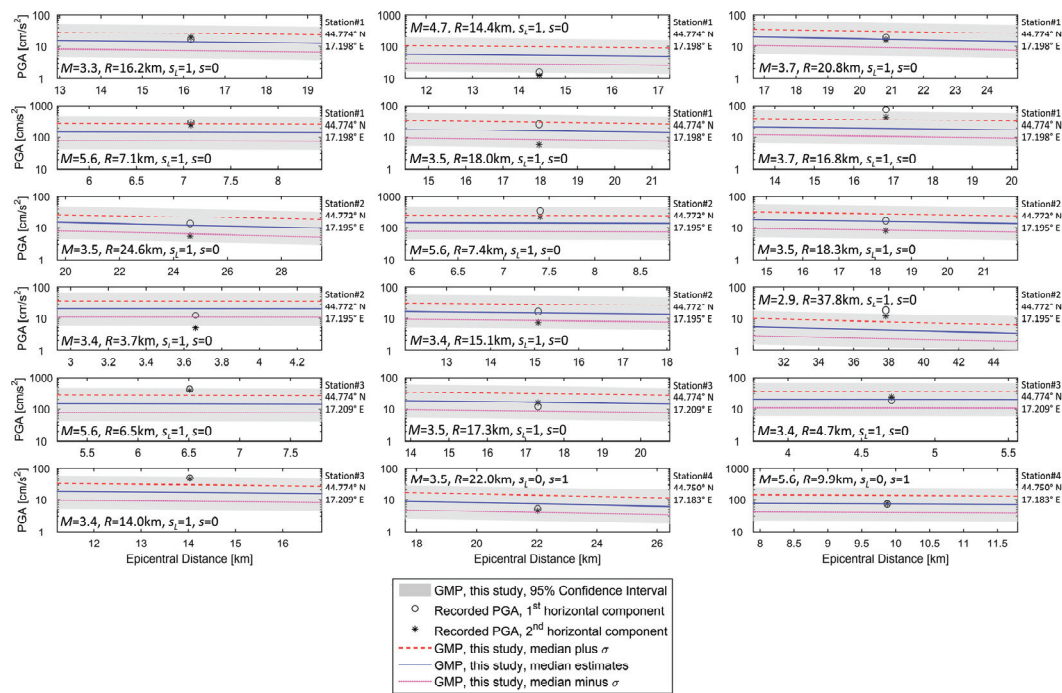


Figure 3. The horizontal PGA values recorded at four separate accelerograph stations in Banja Luka, compared to the median and median $\pm \sigma_{log}$ empirical horizontal PGA estimates derived using Equation (3). The 95% confidence interval (median $\pm 2 \cdot \sigma_{log}$ empirical PGA estimates) is displayed by the gray area.

The presented scaling equations do not account for potential PGA variations as a function of rupture direction relative to the investigated site. According to a recent study, the directivity phenomenon (the presence of high-velocity pulses) can have a significant impact on masonry structures, such as those in Banja Luka [73]. Additionally, near-field and far-field earthquake effects on low-rise buildings, such as those that are common in Banja Luka, are found to be dependent on local soil characteristics [74]. However, due to a lack of sufficient directivity and near-field data in regional strong motion recordings, this will fall outside the scope of our current research.

In this study, we used empirical GMPEs to directly predict site response. Though we did not compare our PGA estimates to the soil response modeled by pattern recognition methods [75] or linear and nonlinear soil response analyses [13,76–81], we intend to do so in the future using geotechnical data. Because Banja Luka lacks a network of soil profiles, we may need to interpolate a few available soil profile data [82] to cover the city's territory.

4. PSHA Analysis for the Case Study Location in Banja Luka

Using Equation (3) as the GMPE and the SHARE Project's pan-European seismic source zone model [83–85], we performed a PSHA analysis for the site with coordinates 44°46.5' N, 17°15' E, denoted by the solid blue circle in Figures 1 and 4. The PSHA analysis will be based on Cornell's [86] and McGuire's [87] approaches and will adhere to Chioccarelli et al.'s [88] procedure. For all hazard calculations, we used the REASSESS V2.1 software [88]. The total hazard is the sum of the contributions from each source zone, i (from the full set of zones, I). Figure 4 depicts the borders of the selected seismic zones used for hazard calculations in this research. Although not the subject of this investigation, seismic source zones can be established using publicly accessible regional seismological data [89–91]. In this particular study, we use the source zone model that was based on the new homogeneous earthquake catalogue ("SHARE European Earthquake Catalogue"—SHEEC [92]), created for the scope

of the SHARE Project. The SHEEC catalogue's completeness, homogeneity of magnitudes with respect to M_W , and other significant aspects are explained in depth in [84,93,94].

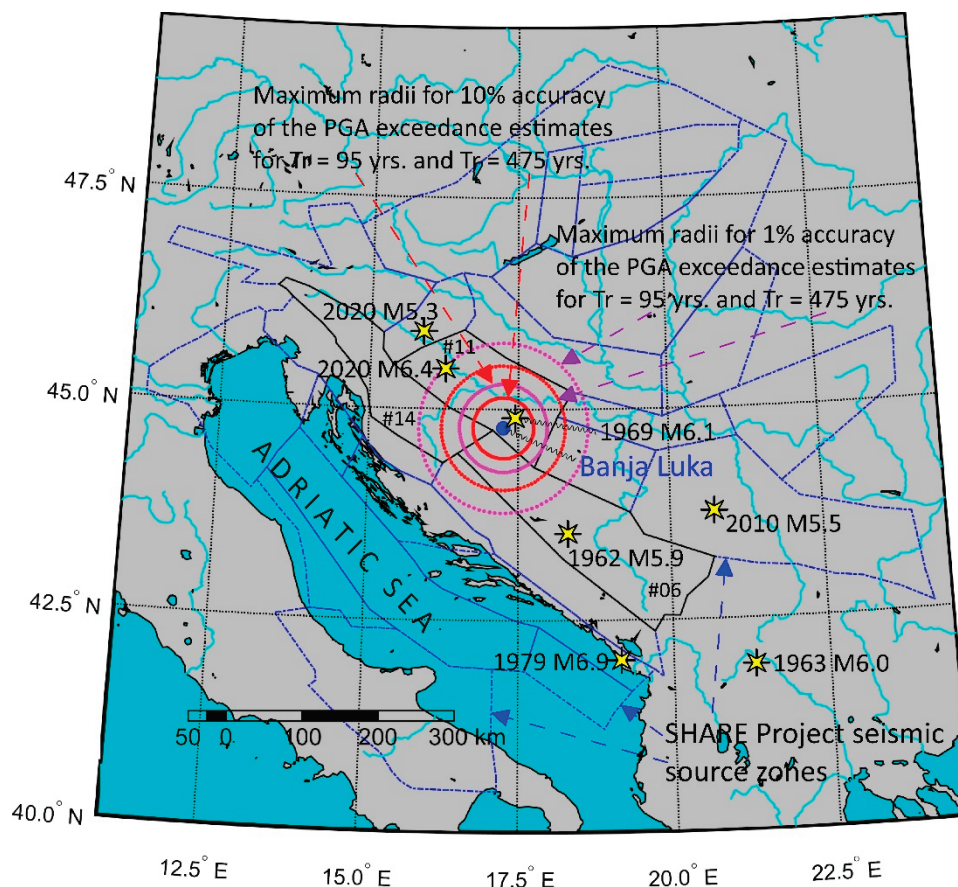


Figure 4. The SHARE Project earthquake source zones [83–85] used in this research for seismic hazard assessments.

Figure 4 displays the source zones included in the hazard calculations for Banja Luka. The two larger circles, with radii of 111.75 and 81.75 km, represent the source-to-site distances that must be included in the PSHA computations to ensure 1% accuracy of computed PGA estimates with a probability of 10% in 10 and 50 years (the 95- and 475-year return periods), respectively. The two smaller circles, with radii of 57.75 and 40.00 km, indicate the maximum distances that will ensure a PSHA accuracy of 10% for the same two probabilities. Figure 4 also depicts the epicenters of several of the major historical earthquakes as well as a few recent disastrous earthquakes in the northwestern Balkan region [34]. As illustrated in Figure 4, local seismicity dominates in the PGA probabilistic estimates, and only the $M_W = 6.1$ 1969 Banja Luka earthquake and $M_W = 6.4$ 2020 Petrinja earthquake are within the depicted radii, while all other shown events occur in areas that do not significantly contribute to the seismic hazard. It is also worth mentioning that some recent regional studies have found that PGA amplitudes are primarily influenced by local seismicity and are not particularly affected by very strong and distant events, such as those in Romania's Vrancea source zone [61,63–65,95,96]. This is because high-frequency seismic waves attenuate rapidly. As a result, while traveling from great distances, their final contribution is less than that of the waves created by local events. Seismic microzonation studies for the cities in Serbia and North Macedonia that are all closer to the Vrancea zone than Banja Luka demonstrate that the contribution of Vrancea earthquakes to PGA values is negligible when compared to local seismicity [61,63–65]. As a result, in this study, we will also overlook the Vrancea source zone.

For the calculation of the total hazard, we will employ the law (or formula) of total probability [97], which defines the probability of an event, A , in case that a series of n mutually exclusive events, H_i , occurs, and H_i comprises a full system of hypotheses about the event, A , as follows:

$$P(A) = \sum_{i=1}^n P(A|H_i) \cdot P(H_i). \quad (5)$$

Based on Equation (5), we can apply the following equation to calculate the average yearly rate of occurrence of seismic events that may cause PGA to exceed the expected pga :

$$N(pga) = \sum_{i \in I} \nu_i \int_{M_{min}}^{M_{max}} \int_{R_{min}}^{R_{max}} f_{gmpe}(PGA > pga|M, R) \cdot f_{m_i}(M) \cdot f_{r_i|m_i}(R|M) dM dR, \quad (6)$$

where i denotes the seismic source zone serial number (from the set I), and ν is the annual rate of events exceeding M_{min} , which is defined as follows:

$$\nu = e^{aln10 - (bln10)M_{min}}. \quad (7)$$

In Equation (7), a and b are the coefficients that characterize the total degree of seismicity in the analyzed area and the ratio between events of smaller and larger sizes, respectively, and are defined as follows:

$$\log N_{GR}(M) = a - b \cdot M, \quad (8)$$

where N_{GR} represents the rate of occurrence of earthquakes of different magnitudes in each zone, i ; M_{max} is the maximum magnitude for each zone, i ; M_{min} is the minimum magnitude (although the SHARE seismic-hazard map for all of Europe was calculated with $M_{min} = 4.5$ [84], we chose to use $M_{min} = 4.0$ for all zones in this study to obtain more conservative PGA estimates for shorter return periods); and R_{min} and R_{max} are the minimum and maximum epicentral distances equal to 0 and 300 km, respectively.

Furthermore, f_{gmpe} is the conditional cumulative distribution function, which is based on the GMPE and determines the likelihood that PGA will surpass pga if an event of magnitude, M , occurs at a source-to-site distance, R ; $f_{r|m}$ is the probability density function for distance; and f_m is the probability density function for magnitudes, which is defined as follows:

$$f_m(M) = \beta \frac{e^{-\beta(M-M_{min})}}{1 - e^{-\beta(M_{max}-M_{min})}}, M_{min} \leq M \leq M_{max}, \quad (9)$$

where $\beta = (\ln 10)b$.

We will now explain the alternative metrics for seismic hazard, shown in Table 4. The so-called “return period” is calculated as follows:

$$Tr = N(pga)^{-1}. \quad (10)$$

Although Tr is commonly used in the engineering literature as an instinctive measure of hazard levels (a larger Tr indicates stronger events that occur less often and vice versa), it lacks a distinct physical meaning in the sense that it does not (generally) correspond to a single earthquake of engineering concern. This is because $N(pga)$ for a single location is calculated by combining and summing data from all analyzed seismic source zones across all magnitudes and distances.

When homogeneous Poisson distributions are assumed, the annual probability, or the likelihood of at least one yearly exceedance of the expectation pga , may be expressed as follows [97]:

$$P(pga) = 1 - e^{-N(pga)}. \quad (11)$$

Finally, if a binomial distribution is assumed, the probability that the expectation pga will be exceeded at least once during t years can be computed as follows [97]:

$$p(pga) = 1 - [1 - P(pga)]^t. \quad (12)$$

Table 4. Four different probability metrics: $P(pga)$ is the yearly probability of at least one exceedance of the expectation pga ; $p(pga)$ is the likelihood of at least one exceedance of the expectation pga during $t = 10$ or 50 years, and $Tr(pga)$ is the “return period”.

$P(pga)$	$p(pga)$ [%] in $t = 10$ years	$p(pga)$ [%] in $t = 50$ years	$Tr(pga) = \frac{1}{N(pga)} \cong \frac{1}{P(pga)} = \frac{1}{1 - \sqrt[t]{1 - p(pga)}}$ [years]
0.020000	18.29	63.58	50.00
0.010481	10.00	40.95	95.41
0.010000	9.56	39.50	100
0.005000	4.89	22.17	200.00
0.002105	2.09	10.00	475.06
0.002000	1.98	9.52	500
0.001025	1.02	5.00	975.29
0.001000	1.00	4.88	1000.00
0.000404	0.40	2.00	2475.42
0.000100	0.10	0.50	10,000.00

A procedure that is called “seismic hazard disaggregation” can be used to identify the earthquakes that contribute the most to the computed hazard [98,99]. It is effectively the inverse of the process described in Equation (6), and it allows us to distinguish independent contributions from distinct magnitude and distance combinations.

The four plots of Figure 5 show the PSHA disaggregation for $44^\circ 46.5' \text{ N}$, $17^\circ 15' \text{ E}$ (refer to the solid blue circles in Figures 1 and 4) and $Tr = 95, 475, 975$, and 2475 years. These plots show that the most contributing magnitudes increase with the return period, Tr . The cumulative disaggregation for M and four different return periods is depicted in the top left plot of Figure 6. For the return periods of 95, 475, 975, and 2475, the magnitudes that contribute to 50% of the anticipated likelihood of exceedance are equal to 4.9, 5.2, 5.35, and 5.5, respectively. The three other plots in Figure 6 show the magnitude recurrence relations for the three most contributing source zones (in Figure 4, these three zones are designated with serial numbers). As can be seen from these plots, the real recurrence (return) periods of the earthquakes that contribute the most vary with the “return period”, Tr , which represents just a reciprocal value of $N(a)$ (see Equations (6) and (10)). What is even more important, the real return periods are significantly shorter than the corresponding return periods, Tr . This is very important to understand, because Eurocode 8 [9] defines Types 1 and 2 spectra not in terms of maximum credible magnitudes but rather in terms of those magnitudes that “contribute most to the seismic hazard defined for the site for the purpose of probabilistic hazard assessment” (please refer to Eurocode 8 [9]: 3.2.2.1, (4), Note 2 and 3.2.2.2, (2)P, Note 1). The selection of the spectral type based on the most contributing earthquakes can be extremely confusing for average civil engineers. The type of spectrum is not specified in Bosnia and Herzegovina’s National Annex to Eurocode 8 [14], and users are free to select it without being given access to the hazard disaggregation data. Given Banja Luka’s history of experiencing earthquakes with magnitudes in excess of 6.0, it stands to reason that the majority of the contributing magnitudes would be greater than 5.5. However, Figures 5 and 6 show that the magnitudes of the most contributing earthquakes are smaller than 5.5, with the exception of the most contributing magnitude

for $T_r = 2475$ years. Even for $T_r = 2475$ years, this magnitude is equal to 5.5 or little less than 5.5.

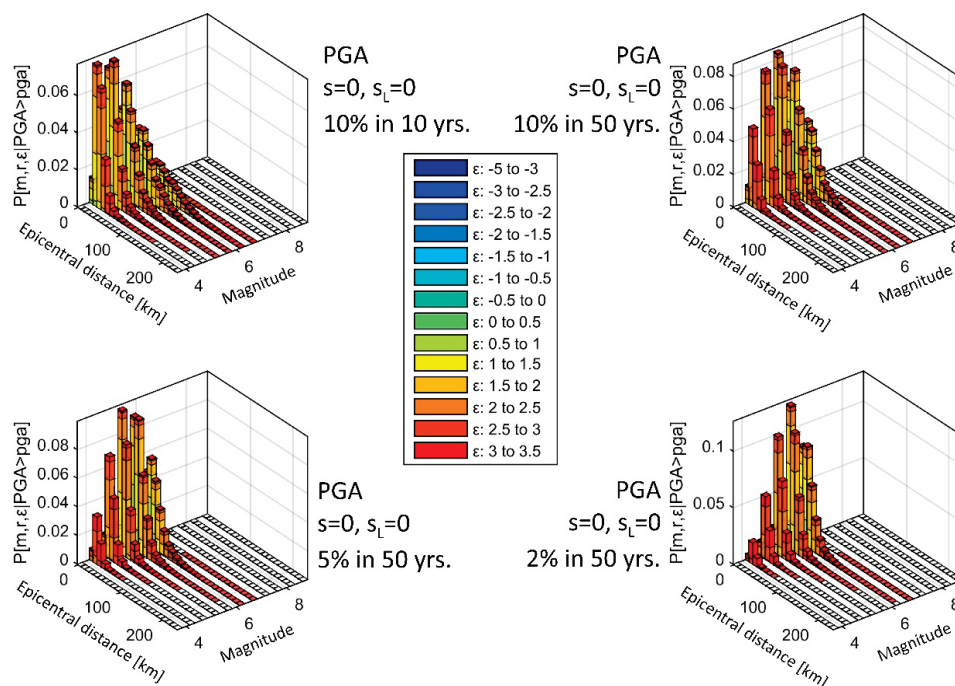


Figure 5. Seismic hazard disaggregation for four different return periods, $T_r = 95$, 475, 975, and 2475 years (see Table 4), for the location $44^{\circ}46.5' \text{ N}$, $17^{\circ}15' \text{ E}$ (see Figures 1 and 4).

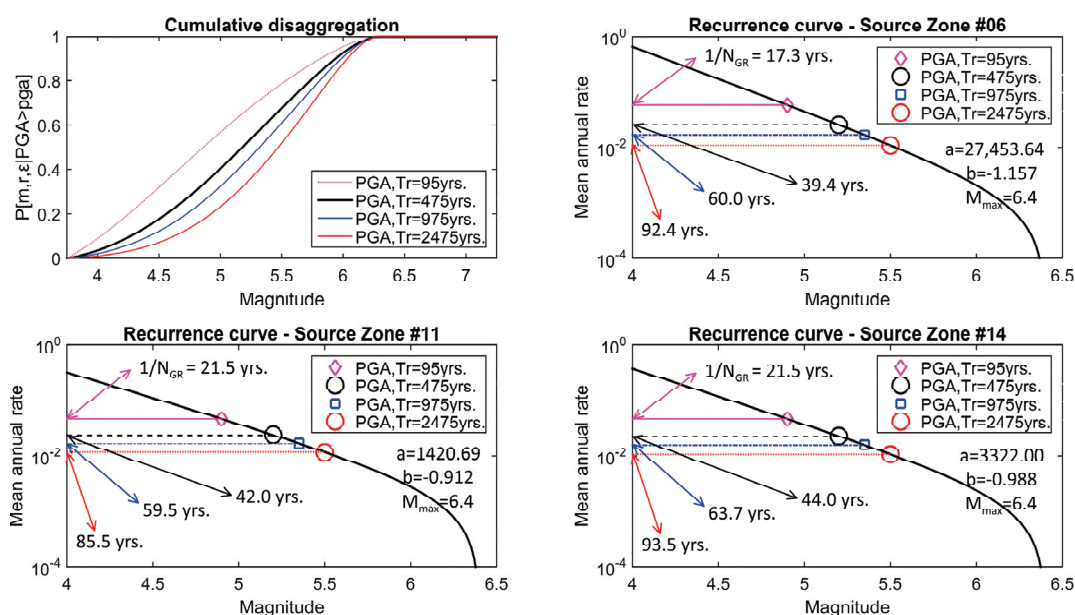


Figure 6. Top left plot—cumulative seismic hazard disaggregation for different values of magnitude and four different return periods, T_r , for the location $44^{\circ}46.5' \text{ N}$, $17^{\circ}15' \text{ E}$ (see Figures 1 and 4); top right and bottom plots—magnitude recurrence curves for the three most contributing source zones around Banja Luka (see Figure 2). For each source zone, the maximum magnitude and Gutenberg–Richter [100] a and b coefficients (see Equation (8)) are also shown.

The seismic hazard curves for $44^{\circ}46.5' \text{ N}$, $17^{\circ}15' \text{ E}$ (see Figures 1 and 4) were compared to the PGA values calculated based on regulations in former Yugoslavia [56] and the values from current official seismic hazard maps for Bosnia and Herzegovina [14] to validate the findings of our PSHA study, as shown in Figure 7.

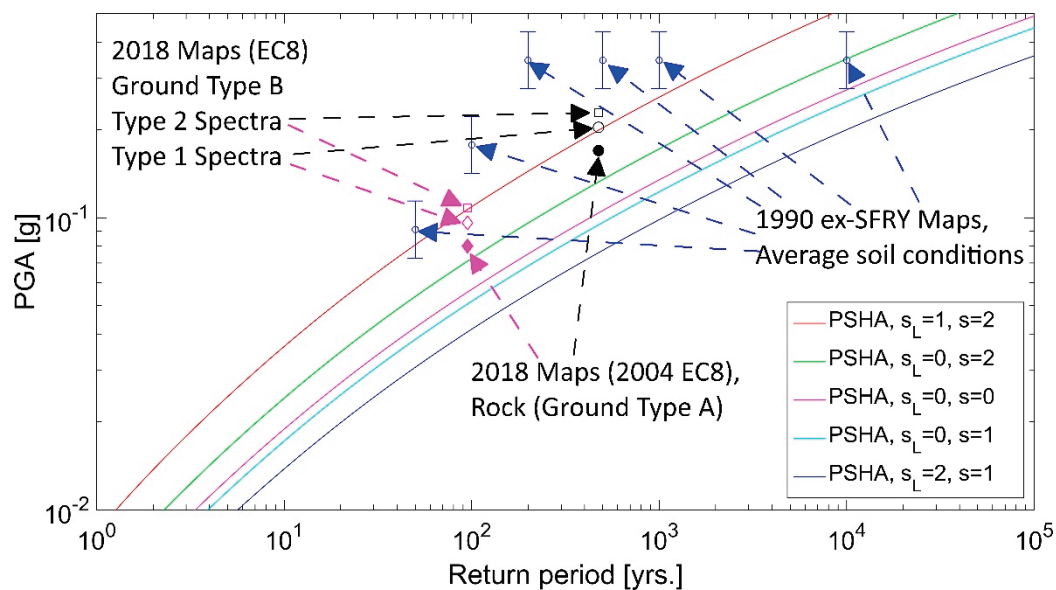


Figure 7. Various probabilistic PGA estimates for the location in Banja Luka with coordinates $44^{\circ}46.5' \text{ N}$, $17^{\circ}15' \text{ E}$, computed in this study for different return periods, to the PGAs from current official seismic hazard maps of Bosnia and Herzegovina [14] and PGAs estimated based on the macroseismic intensity degrees from ex-Yugoslavian maps [56].

Before we compare different hazard estimates, it is important to remember that the PGA values on the hazard maps provided in the National Annex to Eurocode 8 are for “rock or comparable geological formations” [9]. Under former Yugoslavian standards, official hazard maps were created for average local soil conditions, and the hazard was calculated using the MCS intensity scale degrees [43,54–56]. The impact of deep geologic conditions was not explored in the creation of any of these maps.

As shown in Figure 6, our PSHA estimates of horizontal PGAs are in fair agreement with those from 1990 macroseismic intensities [56] only for the return periods of 50 and 10,000 years. For all other return periods (100, 200, 500, and 1000 years), our estimates are much smaller. However, our estimates in the case of (deep) geological rock and stiff soil sites are in good agreement with the PGA values for Banja Luka from the 2018 official maps of Bosnia and Herzegovina [14].

5. Discussion and Conclusions

In our previous studies, we either exclusively examined the geological rock sites that exist in Kopaonik, Serbia [6], or the deep soil sites that are found in Osijek, Croatia [57–60]. This is the first time we have examined a region that has varying deep geological and local soil conditions. Our study’s primary weakness is the size of our database, which only includes 436 horizontal components of regional strong motion records. Unfortunately, no other sets of data in the world include deep geology data in parallel with local soil data, with the exception of California, which has different seismo-tectonic properties than the Balkans [8]. We are, therefore, constrained to rely on our current database.

In this study, we examined horizontal PGAs for Banja Luka, which served as a case study area for regions with varying deep geological conditions and with a history of catastrophic earthquakes. This city is situated between the Pannonian basin and the Dinaric Alps. In addition to deep geological sediments, Banja Luka has scattered outcrops of geological rock and also areas with complex deep geology. This city was also devastated by the 27 October 1969 Banja Luka earthquake with $M_W = 6.1$, and the epicenter around 10

km from the city center. The estimated predominant intensity in Banja Luka during the 1969 earthquake was VIII °MCS.

Regional empirical GMPEs for horizontal PGAs are first presented. These GMPEs consider the effects of local soil and deep geology simultaneously. Despite the modest number of the strong-motion accelerograms from former Yugoslavia, which were used to derive the GMPEs, the empirical PGA estimates are shown to be in excellent agreement with the PGA values that were recorded in the past in Banja Luka.

Our conclusions can be summarized as follows:

- Regardless of the type of deep geological formation, the effects of deep geology are comparable in scale to those of local soil and must not be overlooked.
- Only the combined effects of stiff soil sites and (deep) geological rock can result in median empirical PGA estimates comparable to the ones associated with the intensity of the 1969 Banja Luka earthquake (VIII °MCS).
- Real re-occurrence (return) periods of the earthquakes that are most contributing to seismic hazard vary with the so-called “return period”, T_r , which represents just a reciprocal value of $N(a)$ (see Equations (6) and (10)); moreover, real return periods are significantly shorter than the corresponding return periods, T_r .
- PGA probabilistic hazard estimates are dominated by local seismicity and are less influenced by earthquakes at distances larger than around 100 km.
- The latest official (2018) hazard maps [14] provide PGA estimates that are comparable to our estimates of horizontal PGAs for Banja Luka.

Table 5 shows that depending on the GMPE used (Equation (3) or Equation (4)), the effects of deep geology may even outweigh the effects of local soil. For example, if Equation (4) is used, the largest site amplification of 1.67 is obtained for deep geology when comparing the PGA at geological rock to that at intermediate sites. Compared to the local soil amplification of 1.34, which is obtained at stiff soil sites (compared to the “rock” soil), the deep geology amplification of 1.67 is 25% higher.

Table 5. Empirical ratios between the PGA for different site conditions, calculated using scaling coefficients of Equations (3) and (4) for local soil (c_4/c_5) and deep geology (c_6/c_7).

	Stiff Soil $s_L = 1/“Rock”Soils_L = 0$	“Rock” Soil $s_L = 0/Stiff Soils_L = 1$	Deep Soil $s_L = 2/“Rock” Soils_L = 0$	“Rock” Soil $s_L = 0/Deep Soils_L = 2$	Intermed. Sites: $s = 1/GeologicalRock: s = 2$	Geological Rock: $s = 2/Intermed.Sites, s = 1$	Deep Geol. Sediments: $s = 0/GeologicalRock: s = 2$	Geological Rock: $s = 2/Deep Geol.Sediments:s = 0$
GMPE	10^{c4}	$1/10^{c4}$	10^{c5}	$1/10^{c5}$	10^{c6}	$1/10^{c6}$	10^{c7}	$1/10^{c7}$
Equation (3)	1.50	0.66	0.80	1.25	0.71	1.40	0.78	1.28
Equation (4), $R < 30$ km	1.34	0.75	0.80	1.25	0.60	1.67	0.81	1.23

We believe that our PSHA estimates, which were computed using regional GMPEs, can be perceived as a first step toward defining more realistic PGA probabilistic estimates in Banja Luka. When the number of recorded acceleration time histories increases in this region, the obtained scaling equations will be further assessed and, if necessary, calibrated or modified, as well as the PSHA analyses.

Author Contributions: Conceptualization, B.B., S.L., and S.B.; methodology, B.B. and S.L.; formal analysis, B.B., S.L., S.B., and M.K.; investigation, B.B., S.B., A.S.-Ć., and M.Š.; data curation, B.B.; writing—original draft preparation, B.B., S.L., S.B., and M.H.-N.; writing—review and editing, B.B., S.L., S.B., A.S.-Ć., M.Š., and M.H.-N.; visualization, B.B. All authors have read and agreed to the published version of the manuscript.

Funding: For the first, third, fourth, and fifth authors, this research has been supported by the Ministry of Science, Technological Development and Innovation (Contract No. 451-03-137/2025-03/200156), and the Faculty of Technical Sciences, University of Novi Sad, through the project “Scientific and Artistic Research Work of Researchers in Teaching and Associate Positions at the Faculty of Technical Sciences, University of Novi Sad 2025” (No. 01-50/295).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, and further inquiries can be directed to the authors.

Acknowledgments: The authors would like to thank the four anonymous reviewers for their insightful comments and suggestions, which lead to a significantly improved manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

a_g	Horizontal PGA value from the official seismic hazard map
EC8	Eurocode 8
GMPEs	Ground Motion Prediction Equations
PGA	Peak ground acceleration
PSHA	Probabilistic Seismic Hazard Assessment
s	Deep geology parameter
S	Soil factor according to Eurocode 8
s_L	Local soil parameter
Tr	Return period
V_s	Average shear wave velocity

References

1. Bysiec, D.; Maleska, T. Influence of the mesh structure of geodesic domes on their seismic response in applied directions. *Arch. Civ. Eng.* **2023**, *69*, 65–78. [CrossRef]
2. Araz, O. Effect of PGV/PGA ratio on seismic-induced vibrations of structures equipped with parallel tuned mass dampers considering SSI. *Structures* **2024**, *68*, 107188. [CrossRef]
3. Wang, J.; Wu, Y.; Yang, J.; Ruan, B.; Hou, Y. Simulation for non-uniform seismic motion based on frequency-wavenumber spectrum and its application in seismic analysis of long tunnels. *Tunn. Undergr. Space Technol.* **2025**, *161*, 106537. [CrossRef]
4. Douglas, J. Earthquake ground motion estimation using strong-motion records: A review of equations for the estimation of peak ground acceleration and response spectral ordinates. *Earth-Sci. Rev.* **2003**, *61*, 43–104. [CrossRef]
5. Trifunac, M.D. How to model amplification of strong earthquake motions by local soil and geologic site conditions. *Earthq. Eng. Struct. Dyn.* **1990**, *19*, 833–846. [CrossRef]
6. Bulajić, B.Đ.; Bajić, S.; Stojnić, N. The effects of geological surroundings on earthquake-induced snow avalanche prone areas in the Kopaonik region. *Cold Reg. Sci. Technol.* **2018**, *149*, 29–45. [CrossRef]
7. Bulajić, B.Đ.; Manić, M.I.; Lađinović, Đ. Effects of shallow and deep geology on seismic hazard estimates: A case study of pseudo-acceleration response spectra for the northwestern Balkans. *Nat. Hazards* **2013**, *69*, 573–588. [CrossRef]
8. Lee, V.W. *Influence of Local Soil and Geologic Site Conditions on Pseudo Relative Velocity Spectrum Amplitudes of Recorded Strong Motion Accelerations*; Department of Civil Engineering, University of Southern California: Los Angeles, CA, USA, 1987.
9. EN 1998-1:2004; Eurocode 8: Design of Structures for Earthquake Resistance. Part 1: General Rules, Seismic Actions and Rules for Buildings. European Committee for Standardization: Brussels, Belgium, 2004.
10. Trifunac, M.D.; Brady, A.G. On the correlation of seismic intensity scales with the peaks of recorded strong ground motion. *Bull. Seismol. Soc. Am.* **1975**, *65*, 139–162. [CrossRef]
11. Peng, Y.; Wang, Z.; Woolery, E.W.; Lyu, Y.; Carpenter, N.S.; Fang, Y.; Huang, S. Ground-motion site effect in the Beijing metropolitan area. *Eng. Geol.* **2020**, *266*, 105395. [CrossRef]
12. Tavakoli, H.R.; Amiri, M.T.; Abdollahzade, G.; Janalizade, A. Site effect microzonation of Babol, Iran. *Geomech. Eng.* **2016**, *11*, 821–845. [CrossRef]

13. Jakka, R.S.; Hussain, M.; Sharma, M.L. Effects on amplification of strong ground motion due to deep soils. *Geomech. Eng.* **2015**, *8*, 663–674. [CrossRef]
14. EN 1998-1/NA:2018; Seismic Zone Maps and Reference Ground Accelerations Therein of B&H, BAS. Maps Accompanying National Annexes (NA). 2018. Available online: <http://eurokodovi.ba/seizmika/> (accessed on 1 May 2025).
15. Ademović, N.; Demir, V.; Cvijić-Amulić, S.; Málek, J.; Prachař, I.; Vackář, J. Compilation of the seismic hazard maps in Bosnia and Herzegovina. *Soil Dyn. Earthq. Eng.* **2021**, *141*, 106500, Erratum in *Soil Dyn. Earthq. Eng.* **2023**, *164*, 107633. <https://doi.org/10.1016/j.soildyn.2022.107633>. [CrossRef]
16. Das, R.; Das, A. Limitations of Mw and M Scales: Compelling Evidence Advocating for the Das Magnitude Scale (Mwg)—A Critical Review and Analysis. *Indian Geotech. J.* **2025**, *in press*. [CrossRef]
17. Das, R.; Menesis, C.; Urrutia, D. Regression relationships for conversion of body wave and surface wave magnitudes toward Das magnitude scale, Mwg. *Nat. Hazards* **2023**, *117*, 365–380. [CrossRef]
18. Das, R.; Sharma, M.L.; Wason, H.R.; Choudhury, D.; Gonzalez, G. A Seismic Moment Magnitude Scale. *Bull. Seismol. Soc. Am.* **2019**, *109*, 1542–1555. [CrossRef]
19. Bulajić, B.Đ.; Lozančić, S.; Bajić, S.; Radu, D.; Işık, E.; Negovanović, M.; Hadzima-Nyarko, M. UHS estimates for vertical strong ground motion and varying deep geological site surroundings—A case study of the city of Banja Luka. 2025; *submitted for publication*.
20. Bulajić, B.Đ.; Lozančić, S.; Bajić, S.; Radu, D.; Işık, E.; Negovanović, M.; Hadzima-Nyarko, M. Horizontal UHS predictions for varying deep geology conditions—A case study of the city of Banja Luka. *Sustainability*, **2025**; *submitted for publication*.
21. Bulajić, B.Đ.; Lozančić, S.; Bajić, S.; Starčev-Ćurčin, A.; Šešlija, M.; Kovačević, M.; Hadzima-Nyarko, M. PGA estimates for vertical ground motion and varying deep geology site surroundings—A case study of Banja Luka. *Appl. Sci.* **2025**, *15*, 6542. [CrossRef]
22. Seed, H.B.; Murarka, R.; Lysmer, J.; Idriss, I.M. Relationships of maximum acceleration, maximum velocity, distance from source, and local site conditions for moderately strong earthquakes. *Bull. Seismol. Soc. Am.* **1976**, *66*, 1323–1342. [CrossRef]
23. Seed, H.B.; Ugas, C.; Lysmer, J. Site-dependent spectra for earthquake-resistant design. *Bull. Seismol. Soc. Am.* **1976**, *66*, 221–243. [CrossRef]
24. Trifunac, M.; Lee, V.; Živčić, M.; Manić, M. On the correlation of Mercalli-Cancani-Sieberg intensity scale in Yugoslavia with the peaks of recorded strong earthquake ground motion. *Eur. Earthq. Eng.* **1991**, *5*, 27–33.
25. Lee, V.; Trifunac, M. Empirical scaling of Fourier amplitude spectra in former Yugoslavia. *Eur. Earthq. Eng.* **1993**, *7*, 47–61.
26. Lee, V.W.; Manić, M.I. Empirical scaling of response spectra in former Yugoslavia. In Proceedings of the 10th European Conference on Earthquake Engineering, Vienna, Austria, 28 August–2 September 1994; pp. 2567–2572.
27. Morales-Esteban, A.; Martínez-Álvarez, F.; Scitovski, S.; Scitovski, R. Mahalanobis clustering for the determination of incidence-magnitude seismic parameters for the Iberian Peninsula and the Republic of Croatia. *Comput. Geosci.* **2021**, *156*, 104873. [CrossRef]
28. Morales-Esteban, A.; Martínez-Álvarez, F.; Scitovski, S.; Scitovski, R. A fast partitioning algorithm using adaptive Mahalanobis clustering with application to seismic zoning. *Comput. Geosci.* **2014**, *73*, 132–141. [CrossRef]
29. Skoko, D.; Prelogović, E.; Alinović, B. Geological structure of the Earth's crust above the Moho discontinuity in Yugoslavia. *Geophys. J. Int.* **1987**, *89*, 379–382. [CrossRef]
30. Bielik, M.; Makarenko, I.; Csicsay, K.; Legostaeva, O.; Starostenko, V.; Savchenko, A.; Šimonová, B.; Dérerová, J.; Fojtková, L.; Pašteka, R. The refined Moho depth map in the Carpathian-Pannonian region. *Contrib. Geophys. Geod.* **2018**, *48*, 179–190.
31. Stojković, M.B. Seismic microzoning of the city of Banja Luka area. In Proceedings of the International conference on earthquake engineering, Banja Luka, Bosnia and Herzegovina, 26–28 October 2009; pp. 345–357.
32. Trkulja, D. Earthquakes in Banja Luka. In Proceedings of the International Conference on Earthquake Engineering, Banja Luka, Bosnia and Herzegovina, 26–28 October 2009; pp. 43–57.
33. Trkulja, D. *Earthquakes of Banja Luka region*; Institute of Construction Banja Luka—ZIBL: Banja Luka, Bosnia and Herzegovina, 2009.
34. *Earthquake Catalogue for All Earthquakes with Mw ≥ 3.0 in the Period Between 1900 and April 2025 for the Geographic Region Between 41° N and 47° N, and 13° E and 23° E*; USGS: Reston, VA, USA, 2025.
35. Arsovski, M.; Bouwkamp, J.; Cismigiu, A.; Izumi, M.; Napetvaridze, S.G.; Petrovski, J.; Sonobe, Y.; Stojković, M.; Velkov, M. *The Banja Luka Earthquakes of 26 and 27 October 1969*; UNESCO: Paris, France, 1970.
36. EMSC-CSEM. M 5.4—CROATIA—2020-03-22 05:24:02 UTC. Available online: <https://www.emsc-csem.org/Earthquake/earthquake.php?id=840695#scientific> (accessed on 17 April 2025).
37. Ganas, A.; Elias, P.; Valkaniotis, S.; Tsironi, V.; Karasante, I.; Briole, P. Petrinja Earthquake Moved Crust 10 Feet. Available online: <https://temblor.net/earthquake-insights/petrinja-earthquake-moved-crust-10-feet-12410/> (accessed on 17 April 2025).
38. Pavić, G.; Bulajić, B.; Hadzima-Nyarko, M. The Vulnerability of Buildings from the Osijek Database. *Front. Built Environ.* **2019**, *5*, 66. [CrossRef]

39. Pavić, G.; Hadzima-Nyarko, M.; Bulajić, B. A Contribution to a UHS-Based Seismic Risk Assessment in Croatia—A Case Study for the City of Osijek. *Sustainability* **2020**, *12*, 1796. [CrossRef]
40. Pavić, G.; Hadzima-Nyarko, M.; Bulajić, B.; Jurković, Ž. Development of Seismic Vulnerability and Exposure Models—A Case Study of Croatia. *Sustainability* **2020**, *12*, 973. [CrossRef]
41. Blagojević, N.; Brzev, S.; Petrović, M.; Borozan, J.; Bulajić, B.; Marinković, M.; Hadzima-Nyarko, M.; Koković, V.; Stojadinović, B. Residential building stock in Serbia: Classification and vulnerability for seismic risk studies. *Bull. Earthq. Eng.* **2023**, *21*, 4315–4383. [CrossRef]
42. Işık, E.; Hadzima-Nyarko, M.; Bilgin, H.; Ademović, N.; Büyüksaraç, A.; Harirchian, E.; Bulajić, B.; Özmen, H.B.; Aghakouchaki Hosseini, S.E. A Comparative Study of the Effects of Earthquakes in Different Countries on Target Displacement in Mid-Rise Regular RC Structures. *Appl. Sci.* **2022**, *12*, 12495. [CrossRef]
43. SFRY. *Temporary Technical Regulations for Construction in Seismic Areas*; 39/64; Official Gazette of SFRY: Belgrade, Serbia, 1964.
44. Inel, M.; Ozmen, H.B.; Bilgin, H. Re-evaluation of building damage during recent earthquakes in Turkey. *Eng. Struct.* **2008**, *30*, 412–427. [CrossRef]
45. Kaplan, H.; Bilgin, H.; Yilmaz, S.; Binici, H.; Öztas, A. Structural damages of L'Aquila (Italy) earthquake. *Nat. Hazards Earth Syst. Sci.* **2010**, *10*, 499–507. [CrossRef]
46. Bilgin, H.; Huta, E. Earthquake performance assessment of low and mid-rise buildings: Emphasis on URM buildings in Albania. *Earthq. Struct.* **2018**, *14*, 599–614. [CrossRef]
47. Bilgin, H.; Korini, O. Seismic capacity evaluation of unreinforced masonry residential buildings in Albania. *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 3753–3764. [CrossRef]
48. Işık, E.; Kutanis, M.; Bal, İ.E. Displacement of the Buildings According to Site-Specific Earthquake Spectra. *Period. Polytech. Civ. Eng.* **2016**, *60*, 37–43. [CrossRef]
49. Timkó, M.; Kovács, I.; Weber, Z. 3D P-wave velocity image beneath the Pannonian Basin using traveltimes tomography. *Acta Geod. Et Geophys.* **2019**, *54*, 373–386. [CrossRef]
50. Balázs, A.; Matenco, L.; Magyar, I.; Horváth, F.; Cloetingh, S. The link between tectonics and sedimentation in back-arc basins: New genetic constraints from the analysis of the Pannonian Basin. *Tectonics* **2016**, *35*, 1526–1559. [CrossRef]
51. Mojičević, M.; Vilovski, S.; Tomić, B. *Basic Geological Map—Banja Luka*, L33-119; Federal Geological Survey: Belgrade, Serbia, 1976.
52. Lee, V.W.; Manić, M.I.; Bulajić, B.Đ.; Herak, D.; Herak, M.; Trifunac, M.D. Microzonation of Banja Luka for performance-based earthquake-resistant design. *Soil Dyn. Earthq. Eng.* **2015**, *78*, 71–88. [CrossRef]
53. Manić, M.I.; Bulajić, B.Đ.; Trifunac, M.D. A note on peak accelerations computed from sliding of objects during the 1969 Banja Luka earthquakes in former Yugoslavia. *Soil Dyn. Earthq. Eng.* **2015**, *77*, 164–176. [CrossRef]
54. SFRY. *Book of Rules on Technical Norms for Construction of High-Rise Buildings in Seismic Regions*; 31/81; Official Gazette of SFRY: Belgrade, Serbia, 1981.
55. SFRY. *Book of Rules on Technical Norms for Construction of High-Rise Buildings in Seismic Regions*; 31/81, 49/82; Official Gazette of SFRY: Belgrade, Serbia, 1982.
56. SFRY. *Book of Rules on Technical Norms for Construction of High-Rise Buildings in Seismic Regions*; 31/81, 49/82, 29/83, 21/88, 52/90; Official Gazette of SFRY: Belgrade, Serbia, 1990.
57. Bulajić, B.Đ.; Hadzima-Nyarko, M.; Pavić, G. Vertical to Horizontal UHS Ratios for Low to Medium Seismicity Regions with Deep Soil atop Deep Geological Sediments—An Example of the City of Osijek, Croatia. *Appl. Sci.* **2021**, *11*, 6782. [CrossRef]
58. Bulajić, B.Đ.; Hadzima-Nyarko, M.; Pavić, G. Horizontal UHS Amplitudes for Regions with Deep Soil Atop Deep Geological Sediments—An Example of Osijek, Croatia. *Appl. Sci.* **2021**, *11*, 6296. [CrossRef]
59. Bulajic, B.Đ.; Hadzima-Nyarko, M.; Pavic, G. PGA estimates for deep soils atop deep geological sediments -An example of Osijek, Croatia. *Geomech. Eng.* **2022**, *30*, 233–246. [CrossRef]
60. Bulajić, B.Đ.; Pavić, G.; Hadzima-Nyarko, M. PGA vertical estimates for deep soils and deep geological sediments—A case study of Osijek (Croatia). *Comput. Geosci.* **2022**, *158*, 104985. [CrossRef]
61. Lee, V.W.; Trifunac, M.D.; Bulajić, B.Đ. Seismic microzoning in Novi Sad, Serbia—A case study in a low-seismicity region that is exposed to large and distant earthquakes. *J. Seismol.* **2023**, *27*, 979–997. [CrossRef]
62. Lee, V.W.; Trifunac, M.D.; Bulajić, B.Đ. Seismic hazard mapping for peak ground velocity: Microzonation of Novi Sad, Serbia—A case study in a low-seismicity region exposed to large and distant earthquakes. *J. Seismol.* **2024**, *29*, 85–105. [CrossRef]
63. Lee, V.W.; Trifunac, M.D.; Bulajić, B.Đ.; Manić, M.I.; Herak, D.; Herak, M. Seismic microzoning of Belgrade. *Soil Dyn. Earthq. Eng.* **2017**, *97*, 395–412. [CrossRef]
64. Lee, V.W.; Trifunac, M.D.; Bulajić, B.Đ.; Manić, M.I.; Herak, D.; Herak, M.; Dimov, G. Seismic microzoning in Skopje, Macedonia. *Soil Dyn. Earthq. Eng.* **2017**, *98*, 166–182. [CrossRef]
65. Lee, V.W.; Trifunac, M.D.; Bulajić, B.Đ.; Manić, M.I.; Herak, D.; Herak, M.; Dimov, G.; Gičev, V. Seismic microzoning of Štip in Macedonia. *Soil Dyn. Earthq. Eng.* **2017**, *98*, 54–66. [CrossRef]

66. Lee, V.; Trifunac, M.; Herak, M.; Živčić, M.; Herak, D. M_L^{SM} computed from strong motion accelerograms recorded in Yugoslavia. *Earthq. Eng. Struct. Dyn.* **1990**, *19*, 1167–1179. [CrossRef]
67. Jordanovski, L.R.; Lee, V.W.; Manić, M.I.; Olumčeva, T.; Sinadnovski, C.; Todorovska, M.I.; Trifunac, M.D. *Strong Earthquake Ground Motion Data in EQINFOS: Yugoslavia. Part I*; Department of Civil Engineering, University of Southern California: Los Angeles, CA, USA, 1987.
68. Datta, T.K. *Seismic Analysis of Structures*; John Wiley & Sons (Asia) Pte Ltd.: Singapore, 2010; p. 464.
69. Herak, M. Conversion between the local magnitude (M_L) and the moment magnitude (M_w) for earthquakes in the Croatian Earthquake Catalogue. *Geofizika* **2020**, *37*, 197–211. [CrossRef]
70. Scordilis, E.M. Empirical Global Relations Converting M_S and m_b to Moment Magnitude. *J. Seismol.* **2006**, *10*, 225–236. [CrossRef]
71. Ambraseys, N.; Douglas, J.; Margaris, B.; Sigbjörnsson, R.; Berge-Thierry, C.; Suhadolc, P.; Costa, G.; Smit, P. Dissemination of European strong-motion data. In Proceedings of the 13th World Conference on Earthquake Engineering, Vancouver, BC, Canada, 1–6 August 2004.
72. Ambraseys, N.; Douglas, J.; Margaris, B.; Sigbjörnsson, R.; Smit, P.; Suhadolc, P. Internet site for European strong motion data. In Proceedings of the 12th European Conference on Earthquake Engineering, London, UK, 9–13 September 2002.
73. Bilgin, H.; Hysenliu, M. Comparison of near and far-fault ground motion effects on low and mid-rise masonry buildings. *J. Build. Eng.* **2020**, *30*, 101248. [CrossRef]
74. Sönmezer, Y.; Celiker, M. Determination of seismic hazard and soil response of a critical region in Turkey considering far-field and near-field earthquake effect. *Geomech. Eng.* **2020**, *20*, 131–146.
75. Javadi, A.A.; Rezaia, M. Applications of artificial intelligence and data mining techniques in soil modeling. *Geomech. Eng.* **2009**, *1*, 53–74. [CrossRef]
76. Onturk, K.; Bol, E.; Ozocak, A.; Edil, T.B. Effect of grain size on the shear strength of unsaturated silty soils. *Geomech. Eng.* **2020**, *23*, 301–311.
77. Saffarian Mohammad, A.; Bagheripour Mohammad, H. Seismic response analysis of layered soils considering effect of surcharge mass using HFTD approach. Part II: Nonlinear HFTD and numerical examples. *Geomech. Eng.* **2014**, *6*, 531–544. [CrossRef]
78. Saffarian Mohammad, A.; Bagheripour Mohammad, H. Seismic response analysis of layered soils considering effect of surcharge mass using HFTD approach. Part I: Basic formulation and linear HFTD. *Geomech. Eng.* **2014**, *6*, 517–530. [CrossRef]
79. Sahin, A. Dynamic simulation models for seismic behavior of soil systems—Part II: Solution algorithm and numerical applications. *Geomech. Eng.* **2015**, *9*, 169–193. [CrossRef]
80. Sahin, A. Dynamic simulation models for seismic behavior of soil systems—Part I: Block diagrams. *Geomech. Eng.* **2015**, *9*, 145–167. [CrossRef]
81. Sonmezer Yetis, B.; Bas, S.; Isik Nihat, S.; Akbas Sami, O. Linear and nonlinear site response analyses to determine dynamic soil properties of Kirikkale. *Geomech. Eng.* **2018**, *16*, 435–448. [CrossRef]
82. Aziz, M.; Ahmed, T. Spatial interpolation of geotechnical data: A case study for Multan City, Pakistan. *Geomech. Eng.* **2017**, *13*, 475–488. [CrossRef]
83. Giardini, D.; Woessner, J.; Danciu, L.; Crowley, H.; Cotton, F.; Grünthal, G.; Pinho, R.; Valensise, L.; Consortium, S. European Seismic Hazard Map for Peak Ground Acceleration, 10% Exceedance Probabilities in 50 years. *Swiss Seismol. Serv.* **2013**.
84. Woessner, J.; Laurentiu, D.; Giardini, D.; Crowley, H.; Cotton, F.; Grünthal, G.; Valensise, G.; Arvidsson, R.; Basili, R.; Demircioglu, M.B.; et al. The 2013 European Seismic Hazard Model: Key components and results. *Bull. Earthq. Eng.* **2015**, *13*, 3553–3596. [CrossRef]
85. Pagani, M.; Garcia-Pelaez, J.; Gee, R.; Johnson, K.; Poggi, V.; Styron, R.; Weatherill, G.; Simionato, M.; Viganò, D.; Danciu, L.; et al. *Global Earthquake Model (GEM) Seismic Hazard Map (Version 2018.1–December 2018)*; GEM Foundation: Pavia, Italy, 2018.
86. Cornell, C.A. Engineering seismic risk analysis. *Bull. Seismol. Soc. Am.* **1968**, *58*, 1583–1606. [CrossRef]
87. McGuire, R.K. *FORTTRAN Computer Program for Seismic Risk Analysis*; Series number 76-67; US Geological Survey: Reston, VA, USA, 1976.
88. Chioccarelli, E.; Cito, P.; Iervolino, I.; Giorgio, M. REASSESS V2.0: Software for single- and multi-site probabilistic seismic hazard analysis. *Bull. Earthq. Eng.* **2019**, *17*, 1769–1793. [CrossRef]
89. Amaro-Mellado, J.L.; Melgar-García, L.; Rubio-Escudero, C.; Gutiérrez-Avilés, D. Generating a seismogenic source zone model for the Pyrenees: A GIS-assisted triclustering approach. *Comput. Geosci.* **2021**, *150*, 104736. [CrossRef]
90. Amaro-Mellado, J.L.; Morales-Esteban, A.; Asencio-Cortés, G.; Martínez-Álvarez, F. Comparing seismic parameters for different source zone models in the Iberian Peninsula. *Tectonophysics* **2017**, *717*, 449–472. [CrossRef]
91. Amaro-Mellado, J.L.; Tien Bui, D. GIS-Based Mapping of Seismic Parameters for the Pyrenees. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 452. [CrossRef]
92. Stucchi, M.; Rovida, A.; Gomez Capera, A.A.; Alexandre, P.; Camelbeeck, T.; Demircioglu, M.B.; Gasperini, P.; Kouskouna, V.; Musson, R.M.W.; Radulian, M.; et al. The SHARE European Earthquake Catalogue (SHEEC) 1000–1899. *J. Seismol.* **2013**, *17*, 523–544. [CrossRef]

93. Grünthal, G.; Wahlström, R.; Stromeyer, D. The SHARE European Earthquake Catalogue (SHEEC) for the time period 1900–2006 and its comparison to the European-Mediterranean Earthquake Catalogue (EMEC). *J. Seismol.* **2013**, *17*, 1339–1344. [CrossRef]
94. Grünthal, G.; Wahlström, R. The European-Mediterranean Earthquake Catalogue (EMEC) for the last millennium. *J. Seismol.* **2012**, *16*, 535–570. [CrossRef]
95. Lee, V.W.; Trifunac, M.D.; Bulajić, B.; Manić, M. A preliminary empirical model for frequency-dependent attenuation of Fourier amplitude spectra In Serbia from the Vrancea earthquakes. *Soil Dyn. Earthq. Eng.* **2016**, *83*, 167–179. [CrossRef]
96. Lee, V.W.; Trifunac, M.D.; Bulajić, B.D.; Manić, M.I. Preliminary empirical scaling of pseudo relative velocity spectra in Serbia from the Vrancea earthquakes. *Soil Dyn. Earthq. Eng.* **2016**, *86*, 41–54. [CrossRef]
97. Ang, A.H.; Tang, W.H. *Probability Concepts in Engineering Planning: Emphasis on Applications to Civil and Environmental Engineering*; John Wiley and Sons: Hoboken, NJ, USA, 2007.
98. Bazzurro, P.; Allin Cornell, C. Disaggregation of seismic hazard. *Bull. Seismol. Soc. Am.* **1999**, *89*, 501–520. [CrossRef]
99. McGuire, R.K. Probabilistic seismic hazard analysis and design earthquakes: Closing the loop. *Bull. Seismol. Soc. Am.* **1995**, *85*, 1275–1284. [CrossRef]
100. Gutenberg, B.; Richter, C.F. Frequency of earthquakes in California. *Bull. Seismol. Soc. Am.* **1944**, *34*, 185–188. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

PGA Estimates for Vertical Ground Motion and Varying Deep Geology Site Surroundings—A Case Study of Banja Luka

Borko Bulajić ¹, Silva Lozančić ², Senka Bajić ^{1,*}, Anka Starčev-Ćurčin ¹, Miloš Šešlija ¹, Miljan Kovačević ³ and Marijana Hadzima-Nyarko ²

¹ Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia; borkobulajic@uns.ac.rs (B.B.); astarcev@uns.ac.rs (A.S.-Ć.); sele@uns.ac.rs (M.Š.)

² Faculty of Civil Engineering and Architecture Osijek, Josip Juraj Strossmayer University of Osijek, Vladimir Prelog St. 3, 31000 Osijek, Croatia; lozancic@gfos.hr (S.L.); mhadzima@gfos.hr (M.H.-N.)

³ Faculty of Technical Sciences, University of Pristina, Knjaza Milosa 7, 38220 Kosovska Mitrovica, Serbia; miljan.kovacevic@pr.ac.rs

* Correspondence: senka.bajic@uns.ac.rs

Abstract: Vertical PGA is frequently included in civil engineering regulations simply by multiplying the horizontal PGA by a constant. Moreover, most design codes, including Eurocode 8, do not consider the impact of the local soil on vertical ground motion at all. In this study, we demonstrate that such practices increase earthquake risks. The article examines vertical PGA strong-motion estimations for the city of Banja Luka. Banja Luka serves as a case study for areas with records of moderate to strong earthquakes and diverse deep geological conditions. Regional equations for scaling vertical PGA are presented. The vertical PGA values and vertical to horizontal PGA ratios are calculated and analyzed. The findings indicate that the vertical to horizontal PGA ratios for the rock sites depend on the source-to-site distance and deep geology and fall between 0.30 and 0.66. Hence, these ratios cannot be approximated by a single value of 0.90 and 0.45, as specified by Eurocode 8 for Type 1 and Type 2 spectra, respectively. Moreover, the results show that the deep geology effects on vertical ground motion can exceed the local soil effects. When the amount of recorded data from comparable areas increases, we will be able to properly calibrate the existing scaling equations and obtain more reliable estimates of vertical PGA.

Keywords: vertical PGA; Eurocode 8; regional attenuation equations; deep geology; local soil; vertical to horizontal PGA ratios

1. Introduction

Civil engineering regulations often incorporate vertical PGA by simply multiplying the horizontal PGA by a constant. In Eurocode 8 [1], the vertical design PGA values are simply calculated by multiplying the horizontal PGA by a constant that depends only on the spectrum type [1]. Furthermore, the effect of local soil on vertical ground motion is not taken into account at all in many design regulations worldwide. In Eurocode 8 [1], vertical design spectra remain the same for all ground types. However, a recent study on the vertical ground motion at deep soils in Osijek, Croatia, showed that the vertical PGA values can be approximated as 0.61 of the horizontal PGA values, which is a 37% higher ratio than the one given by Eurocode 8 for Type 2 spectra [2]. Another study on the vertical ground motion recorded during the two 6 February 2023 Kahramanmaraş earthquakes showed that it was higher than expected and that the vertical earthquake effects were not assessed correctly in the Turkish code as well [3].

The reason why most codes do not take into account local soil effects on the vertical PGA can be found in the so-called H/V methodology, first presented by Nogoshi and Igarashi [4,5] and later popularized by Nakamura [6–8]. This methodology is widely used to determine the natural period of soil vibrations using microtremors. The H/V approach assumes that local soil amplification happens only horizontally, not vertically. Although recent research has shown that this method will not always provide realistic amplification factors [9], most codes continue to be based on the assumption that local soil effects are not existent in the vertical ground motion direction.

In this study, we will define site effects in terms of “local soil” and “deep geology”. The term “deep geology” will be used to describe the geological site conditions on a scale of at least a few hundred meters, while the deep geology categories are defined as rock, intermediate sites, and sediments, according to Trifunac and Brady’s classifications [10]. The term “local soil” will be used to describe soil with a thickness of up to 100 or so meters, and the local soil types are defined in this study according to classifications by Seed et al. [11,12]. We use the same classifications as in our previous studies [2,13–17] for the same key reasons. First, we acquire the associated deep geology data from the EQINFOS database [18], which contains the majority of our input strong motion data and is the only strong motion database besides the Californian database [19] for which such data exist. Following the classification suggested by Trifunac and Brady [10], an international team of 13 geologists and earthquake engineers interpreted the description of the corresponding site geology and defined the deep geology site condition parameters for the recording sites [20]. As for local soil classification, Trifunac et al. [20] initially presented the local soil classification of the accelerograph locations in the former Yugoslavia, following the classifications of Seed et al. [11,12]. Lee and Trifunac [21] and Lee and Manić (1994) [22] subsequently revised and improved these classifications. Rock soil sites are defined by an average shear wave velocity, V_S , of more than 800 m/s and a soil layer thickness of less than 10 m. Stiff soil sites feature a soil layer that is 15–75 m deep on top of a layer with a $V_S > 800$ m/s. Deep soil locations have an overlaying soil layer that is more than 100 m deep. The second key reason why we use these classifications is that many recent regional microzonation research studies [23–27], which used GMPEs that take into account deep geology and are based only on the data recorded in former Yugoslavia, demonstrated that the resulting empirical predictions were in excellent agreement with newer regional ground motion records and observed macro-seismic intensities. This is why, despite the constraints of the database, we will use the scaling equation described here for the probabilistic hazard analyses in Banja Luka. It will be simple to update the GMPEs and rerun the PSHA analyses as the number of acceleration records increases.

This study is focused on the deep geology effects on the vertical PGA values in regions with a history of moderate to strong earthquakes. For the case study, we chose a location in the city of Banja Luka in Bosnia and Herzegovina, where all three types of deep geology can be found. The paper is structured as follows. In the next Section, we first describe the regional seismicity, Banja Luka geological site surroundings, as well as all past official estimates of vertical PGA. In Section 3, we present regional empirical GMPEs that can predict the effects of the deep geology and compare them to real strong motion records. We also compare the empirical vertical to horizontal PGA ratios (hereinafter, VH_{pga}) to the VH_{pga} ratios suggested by ex-Yugoslav codes, Eurocode 8, and other researchers. In Section 4 we will compute PSHA estimates for vertical PGA, and finally we will compare them to the vertical PGA values given for Banja Luka in various previous ex-Yugoslav codes as well as in the Eurocode 8 National Annex for Bosnia and Herzegovina.

2. Regional Seismicity, Geological Site Surroundings, and Official Vertical PGA Estimates for Banja Luka

Banja Luka is located between the Pannonian basin and the Dinaric Alps fold-and-thrust belt. According to a recent study, the contractional reactivation of Late Palaeogene to Middle Miocene normal faults, which postdates the deposition of the youngest strata in the Pannonian Basin of the Pontian age, may be the cause of earthquakes in the Banja Luka epicentral area [28]. Figure 1 depicts the epicenters of earthquakes with $M_W \geq 3.0$ detected in the region between 1900 and April 2025 [29]. Also visible are the epicenters of two recent major earthquakes in Croatia and a few of the largest regional historical earthquakes. The strongest earthquake close to Banja Luka occurred on 27 October 1969, as part of the seismic activity from 1968 to 1969, with an epicenter around 10 km northeast of Banja Luka (see Figure 1) and a magnitude of $M_W = 6.1$, while the intensity in Banja Luka was predominantly VIII ($^{\circ}$ MCS). On 22 March 2020, a $M_W = 5.3$ earthquake struck the Croatian capital, Zagreb, and this was the strongest earthquake near Zagreb since the $M_W = 6.3$ earthquake near Zagreb in 1880 [30]. On 29 December 2020, a destructive $M_W = 6.4$ event devastated the city of Petrinja, with the epicenter located 40 km south of Zagreb and 100 km northwest of Banja Luka [31]. Both earthquakes had a focal depth of around ten kilometers. Ustaszewski et al. [28] proposed reverse faulting along ESE–WNW striking nodal planes after analyzing fault plane solutions for the two 1969 Banja Luka earthquakes. Furthermore, Ustaszewski et al. [28] discovered evidence of an ongoing shortening in the interior Dinarides, which accounts for a component of the current Adria–Europe convergence, based on comparable findings for a number of minor earthquakes in the area. To learn more about Banja Luka seismotectonics, readers are referred to this exceptionally comprehensive study.

In this study, we will base our insights into the geological settings around the analyzed site on the geological map compiled by Mojičević et al. for Banja Luka and neighboring areas [33]. In a previous microzonation study for Banja Luka [32], this map was divided into 30×30 s cells for the area shown in the bottom plot of Figure 1. The geology of each of these 1200 cells was then defined using the major lithostratigraphic formations and their depths. Site geology was evaluated for each cell per Trifunac and Brady’s classification [10]. Each cell’s geology categorization was assigned as “alluvial and sedimentary deposits” ($s = 0$), “intermediate sites” ($s = 1$), or “basement rock” ($s = 2$)—see Figure 1. We have deliberately chosen the location in Banja Luka with coordinates $44^{\circ}46.5' \text{ N}$, $17^{\circ}15' \text{ E}$ (see blue circle in Figure 1) as this location is surrounded by all three types of deep geology. To the west, there is a lower terrace of the river Vrbas atop of marl, clay, and sand, and we defined this formation with $s = 0$. To the north, there is a diabase-cherty formation (or metamorphic ophiolitic melange), which we defined with $s = 1$. To the east, there is a spilite formation, which we designated by $s = 2$.

In 2018, Bosnia and Herzegovina introduced new seismic hazard maps as part of the National Annex to Eurocode 8 [34,35]. The horizontal PGA values represent the seismic hazard, whereas the vertical PGA is obtained from the horizontal PGA and the ratios provided by Eurocode 8. The two 2018 maps were generated using the PSHA approach, for ground type A (rock), one for the 95-year return period and the other for the 475-year return period. These maps show that the horizontal PGA values for the site in Banja Luka are 0.08 g for the 95-year return period and 0.17 g for the 475-year return period. Multiplying the horizontal values by the Eurocode 8 values of 0.90 and 0.45, the vertical PGA values for the Type 1 Spectrum (with $M_S > 5.5$ for the “most contributing” earthquakes) become 0.072 and 0.153, and 0.036 and 0.077 for the Type 2 Spectrum ($M_S \leq 5.5$), for the 95- and 475-year return periods, respectively. Table 1 shows these values and the corresponding

macroseismic intensities, as well as the MCS intensities for Banja Luka, as provided by seismic zoning maps in the former Yugoslavia.

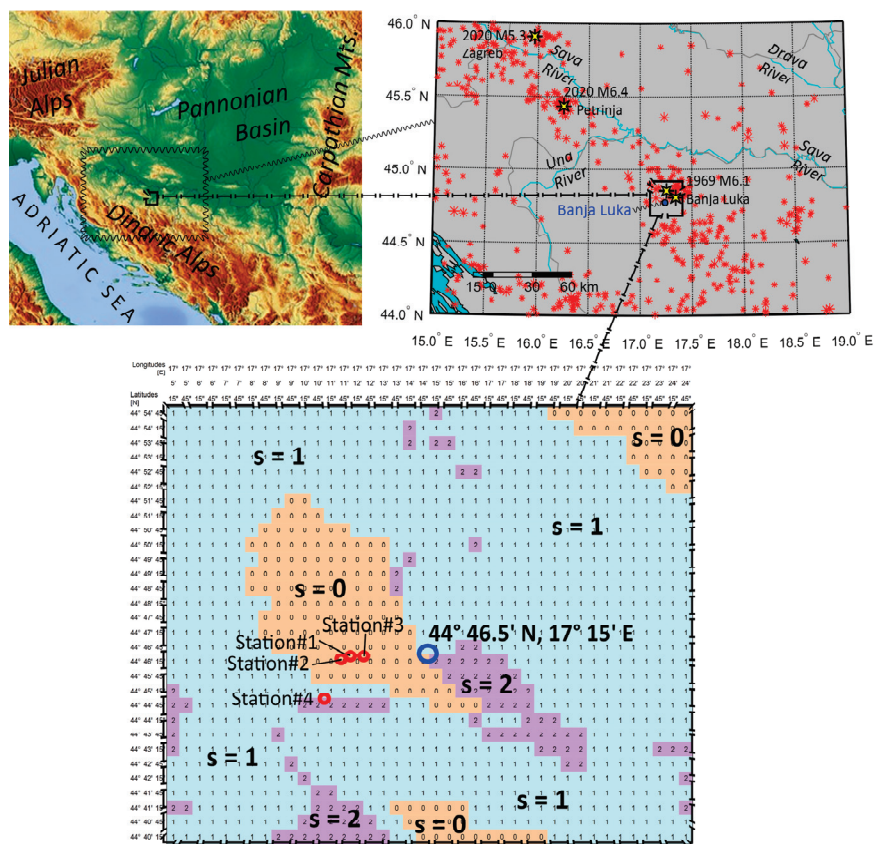


Figure 1. (Top left)—Dinaric Alps and Pannonian Basin, (top right)—epicenters of earthquakes with $M_W \geq 3.0$ recorded in the period of 1900–April 2025 [29] (red stars), including the largest historical event occurring close to Banja Luka and two recent destructive events in Croatia (yellow stars), (bottom)—the distribution of deep geology [10] parameters ($s = 0$ designates deep geological sediments, $s = 1$ complex or intermediate deep geology conditions, and $s = 2$ the rocks) for the rectangle between $44^{\circ}40' N$ and $44^{\circ}55' N$, and $17^{\circ}05' E$ and $17^{\circ}25' E$, modified from [32]. Blue circle shows the analyzed location in Banja Luka. Red circles show four accelerograph stations in Banja Luka that recorded strong ground motion time histories in the past.

Table 1. The vertical PGA estimates calculated using Equation (1) for Banja Luka, based on the intensities depicted in the official zoning maps of the former Yugoslavia [36–39]. The vertical PGA values estimated using horizontal values from the 2018 official maps, which are used in conjunction with Eurocode 8 [34], are also given.

Official Zoning Map	I [°MCS]	PGA [g]
1950	VII	0.052–0.056
1982	VIII	0.112–0.120
1990—50 yrs.	VII	0.052–0.056
1990—100 yrs.	VIII	0.112–0.120
1990—200 yrs.	IX	0.239–0.257
1990—500 yrs.	IX	0.239–0.257
1990—1000 yrs.	IX	0.239–0.257
1990—10,000 yrs.	IX	0.239–0.257
2018—95 yrs.	VI–VIII	0.036–0.072
2018—475 yrs.	VII–IX	0.077–0.153

All former Yugoslav maps expressed the hazard in terms of seismic intensity and for average soil conditions [36–39]. The map that was issued in 1950 was used in conjunction with the first design code for earthquake-resistant structures in the former Yugoslavia, enacted in 1964 [36]. This map was compiled based on the data on the greatest MCS intensities ever recorded. The 1982 interim seismic zoning map was based on the greatest observed intensities at the time and utilized in conjunction with the 1981 code [37,39]. In 1990, six new maps, for the 50-, 100-, 200-, 500-, 1000-, and 10,000-year return periods, were added to the 1981 code [38]. Although the seismic hazard was still displayed in terms of MCS intensity and for average ground conditions, the 1990 maps were the first in the former Yugoslavia to be created using a PSHA method.

For each MCS intensity, Table 1 also displays the associated empirical vertical PGA estimates, obtained using the equation [20]:

$$\log(PGA_{vert}) = -0.593 + 0.331 \cdot I \pm P \cdot \sigma, \sigma = 0.016, \quad (1)$$

where PGA_{vert} is the vertical peak ground acceleration [cm/s^2] and I is the MCS (Mercalli–Cancani–Sieberg Macroseismic Scale) intensity degree. Furthermore, $P = 0$ for the median empirical estimate and ± 1 for the median ± 1 standard deviation, σ . Table 1 displays the vertical PGA values computed by Equation (1) for P equal to -1 and 1 , for each intensity degree.

The vertical PGA predictions for Banja Luka based on seismic intensities from the 1950 and 1982 maps are fairly consistent with the vertical PGA values obtained from the current official hazard maps. However, the 2018 vertical PGAs for the 475-year return period are significantly lower than the 1990 vertical PGAs for the 500-year return period. Even the vertical PGA values for the Type 1 Spectrum, which imply higher seismic activity, are significantly smaller than the 1990 values. Unfortunately, we do not have any information regarding the seismic source models and attenuation functions that were used to compile the 1990 maps. Hence, we were unable to provide a systematic interpretation of a significant discrepancy between the vertical PGA values from the 1990 maps and the more recent 2018 Eurocode-based estimates.

Table 2 shows the vertical PGA values (divided by the gravitational constant, g) estimated by Equation (1) for various intensities, as well as the empirical VHpga [20]. Table 3 includes the VHpga values used by the earthquake-resistant design codes in the former Yugoslavia [36,39], as well as those defined by two empirical GMPEs that incorporate the simultaneous effects of deep geology and local soil, one for California [19] and the other for the former Yugoslavia [22]. While both GMPEs neglect the local soil effect on vertical ground motion, the equation constructed for California, USA, which was based on 1482 acceleration components, takes into account the effects of deep geology on the amplification of vertical ground motion [22].

Table 2. Vertical empirical estimates of PGA and empirical VHpga computed by Equation (1) for different macroseismic intensities.

I	PGA [g]			VHpga		
	-1σ	Median	$+1\sigma$	-1σ	Median	$+1\sigma$
V °MCS	0.0113	0.0118	0.0122	0.530	0.491	0.455
VI °MCS	0.0243	0.0252	0.0261	0.582	0.540	0.500
VII °MCS	0.0521	0.0540	0.0560	0.640	0.593	0.550
VIII °MCS	0.1115	0.1157	0.1201	0.703	0.652	0.604
IX °MCS	0.2390	0.2480	0.2573	0.773	0.716	0.664
X °MCS	0.5122	0.5315	0.5514	0.849	0.787	0.729

Table 3. The vertical PGA according to various design codes and researchers based on the horizontal PGA value, a_g , as provided for the analyzed site on the corresponding seismic hazard map.

Eurocode 8 [1] Ground Types	Vertical PGA
Eurocode 8, Type 1 Spectrum ($M_5 > 5.5$), all ground types [1]	$a_g \times 0.900$
Eurocode 8, Type 2 Spectrum ($M_5 \leq 5.5$), all ground types [1]	$a_g \times 0.450$
1964 code of former Yugoslavia [36]	$a_g \times 0.330$
1981 code of former Yugoslavia [39]	$a_g \times 0.700$
GMPEs for California, USA [19], geological rocks	$a_g \times 0.760$
GMPEs for California, USA [19], intermediate geological sites	$a_g \times 0.718$
GMPEs for California, USA [19], geological sediments	$a_g \times 0.681$
GMPEs for former Yugoslavia [22], all deep geology types	$a_g \times 0.720$

We can see from Table 2 that the VH_{pga} values differ significantly from the ones suggested by Eurocode 8. However, they are in good agreement with the ratios suggested by the 1981 former Yugoslav regulations [39], shown in Table 3, as well as the ratios suggested by the empirical GMPEs for the former Yugoslavia [22] and California [19].

3. GMPEs for Vertical PGA Values for Varying Deep Geology

We will next present empirical GMPEs for estimating vertical PGA values when local soil and deep geology effects are to be considered together. The equations will be expressed in the following mathematical form:

$$\log[PGA_{vert}] = c_1 + c_2 \cdot M + c_3 \cdot \log(\sqrt{R^2 + R_0(T)^2}) + c_4 \cdot S_{L1} + c_5 \cdot S_{L2} + c_6 \cdot S_{G1} + c_7 \cdot S_{G2} + \sigma_{log} \cdot P. \quad (2)$$

In Equation (1), PGA_{vert} is the vertical maximum acceleration (in g), M represents the earthquake magnitude, and R represents the epicentral distance. Equations for hypocentral distance were also generated in our prior studies [2,14–16]. Table 4 lists the categorical factors S_{L1} , S_{L2} , S_{G1} , and S_{G2} . The local soil conditions were classified as per Seed et al. [11,12], and the deeper geological conditions were in line with Trifunac and Brady's classification [10].

Table 4. Categorical variables from Equations (2)–(4) for shallow and deeper geological site conditions.

Local Soil Parameters	Local Soil Categorical Variables	Deep Geology Parameters	Deep Geology Categorical Variables
“Rock” soil sites: $s_L = 0$	$S_{L1} = S_{L2} = 0$	Basement (geological) rock: $s = 2$	$S_{G1} = S_{G2} = 0$
Stiff soil sites: $s_L = 1$	$S_{L1} = 1$ and $S_{L2} = 0$	Intermediate (or complex) sites: $s = 1$	$S_{G1} = 1$ and $S_{G2} = 0$
Deep soil sites: $s_L = 2$	$S_{L1} = 0$ and $S_{L2} = 1$	(Deep geological) Sediments: $s = 0$	$S_{G1} = 0$ and $S_{G2} = 1$

Before we proceed to discuss the magnitude types we used, we should reiterate that the EQINFOS database [18], which includes records from earthquakes in the northwest Balkan region, serves as the foundation for our analysis. For the largest events (magnitudes that exceeded 6.0), the database's authors employed the M_5 magnitude; for all other events, they employed the M_L magnitude as specified by Lee et al. [40]. The local magnitude, M_L , was deliberately selected for the majority of events, as it was deemed most suitable for

the so-called “strong motion”—the frequency content between 0.1 and 30 Hz that civil engineers are most interested in (see, for example, the related discussion in [41]). A later study that converted the M_L magnitudes of the earthquakes in Croatia and the surrounding regions to moment magnitudes (M_W) revealed that the difference in magnitudes between 3.5 and 6.5 was less than 2% [42]. Additionally, according to Scordilis [43], for magnitudes between 6.2 and 7.0, the difference between M_S and M_W magnitudes is approximately 0.2%. We therefore think that we can utilize our GMPEs in seismic hazard calculations with the seismic source zones determined by using the M_W -based seismicity data, taking into account the other constraints of the given database. Here, we must note that even the M_W magnitude has significant restrictions. As shown in recent studies by Das et al. [44] and Das and Das [45], the M_{wg} scale should be used to adjust the seismicity data in order to improve the accuracy of seismic hazard estimates. In our future research, we plan to modify our calculations by using other magnitude scales. However, this will wait until additional regional seismicity and strong motion data are collected, and further discussions on this topic will have to be left out of the purview of this study.

The database used to derive equations for scaling PGA_{vert} consists of 218 vertical components of ground motion records from the region of the former Yugoslavia from 112 earthquakes of magnitude $3 < M \leq 6.8$ [18,46,47]. Detailed descriptions of the data can be found in [2,15,17]. The database has no records from events with magnitudes above 7, and less than 20% of the data is from events with magnitudes greater than 6.0. As a result, the generated predictive equations will be most useful in areas with low to moderate seismic activity. However, a comparison to real records from a 6.8 magnitude earthquake [17] demonstrates that these equations can be generalized to zones of moderate-to-high seismicity, assuming that magnitudes do not exceed 7. It should also be noted that the proposed attenuation equations do not account for potential PGA_{vert} variations as a function of fault rupture directionality (relative to the investigated site). As shown in a recent study, the directivity phenomenon (the presence of high-velocity pulses) can have a significant impact on low- and mid-rise masonry structures like those found in Banja Luka [48]. However, due to the limited quantity of near-fault pulse data in regional strong motion records, this will stay outside the scope of our current research.

We conducted multiple linear regression analyses to compute the scaling coefficients, with R_0 systematically adjusted to lower the root mean squared error. Since the majority of the data was obtained over short distances, we performed a supplementary analysis that only included data acquired at epicentral distances of less than 30 km. Finally, because we also conducted regression analyses for the pseudo-spectral accelerations at 61 vibration periods (and in both horizontal and vertical directions), in all our analyses the scaling coefficients were smoothed in MATLAB® (version 8.5, release 2015a) by “weighted linear least squares and a second degree polynomial model” [49–51]. We utilized a log-normal distribution for all regression analyses, with σ_{log} being the standard deviation of the decadic logarithm of PGA and ε being equal to 0 for median estimations. The resulting scaling equation for the PGA_{vert} , based on the data recorded at all distances, is as follows:

$$\log[PGA_{vert}] = -1.6833 + 0.3947 \cdot M - 1.3186 \cdot \log(\sqrt{R^2 + 16.8^2}) + 0.1575 \cdot S_{L1} - 0.0165 \cdot S_{L2} - 0.1204 \cdot S_{G1} - 0.0678 \cdot S_{G2} + 0.2428 \cdot P, \quad (3)$$

while based solely on the data recorded at distances of not more than 30 km, we got the following equation:

$$\log[PGA_{vert}] = 3.9665 + 0.4045 \cdot M - 4.5446 \cdot \log(\sqrt{R^2 + 40.0^2}) + 0.1328 \cdot S_{L1} - 0.0165 \cdot S_{L2} - 0.1949 \cdot S_{G1} - 0.0630 \cdot S_{G2} + 0.2506 \cdot P. \quad (4)$$

As illustrated in Figure 2, both deep geology and local soil influence PGA_{vert} . The coefficients of the variables S_L and S_G can be used to compute the differences in the PGA calculated for different site conditions. For example, if the deep geology parameters remain constant and Equation (3) is utilized, the PGA value for stiff soil sites is $10^{0.1575} = 1.44$ times that for rock sites and $10^{-0.0165} = 0.96$ times that for deep soil. In other words, for deep soil, seismic waves are de-amplified as energy dissipation outweighs local soil amplification.

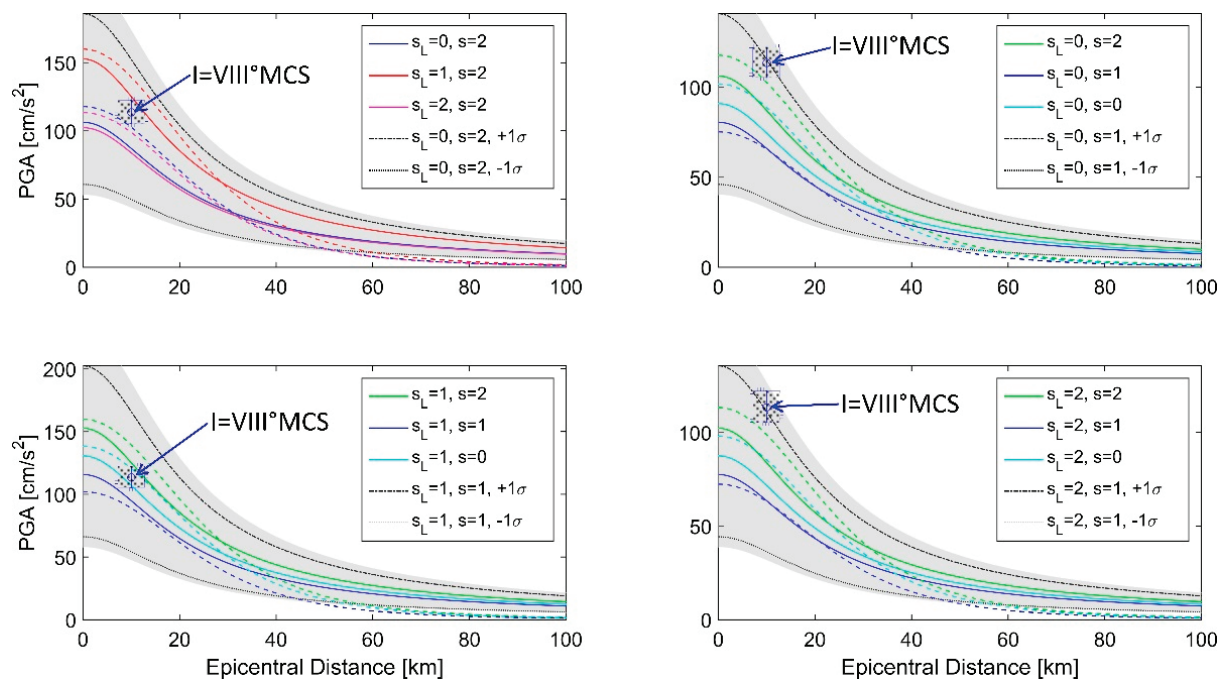


Figure 2. The empirical GMPEs for PGA_{vert} , defined by Equations (3) (solid lines) and (4) (dashed lines), for different deep geology and local soil combinations (see Table 4). The gray area shows the 95% confidence interval (median $\pm 2 \cdot \sigma_{log}$ empirical PGA_{vert} estimates) for $s_L = 2$.

Likewise, if the local soil type remains the same, the PGA values will be lower at geological sediments and intermediate sites than at geological rock sites, since short-period waves travel more easily through compact hard rocks like granites or basalts. Specifically, the PGA estimates at geological rock will be $1/10^{-0.1949} = 1.57$ times larger than at intermediate locations, and $1/10^{-0.0630} = 1.16$ times larger than at geological sediment locations. All of these statistics demonstrate not only that site conditions have a significant impact on PGA_{vert} but also that the deep geology effects are even greater than the local soil effects.

For the epicentral distance of 10 km, Figure 2 shows PGA_{vert} ranges estimated with Equation (1) (median $\pm$ one standard deviation) for VIII °MCS. The intensity of VIII °MCS was predominant during the 1969 earthquake in Banja Luka, and the epicenter was at a distance of around 10 km. We can see that only the combination of stiff soil and deep geology rocks may produce a median empirical PGA_{vert} equivalent to that associated with VIII °MCS.

Figure 3 compares 18 PGA_{vert} values recorded at the four different accelerograph stations in Banja Luka [18] (see the red circles in the bottom plot of Figure 1) with the median and median $\pm \sigma_{log}$ empirical estimates, which were obtained using Equation (3). The empirical estimates for PGA_{vert} are in good agreement with the actual records in Banja Luka, as Figure 4 illustrates.

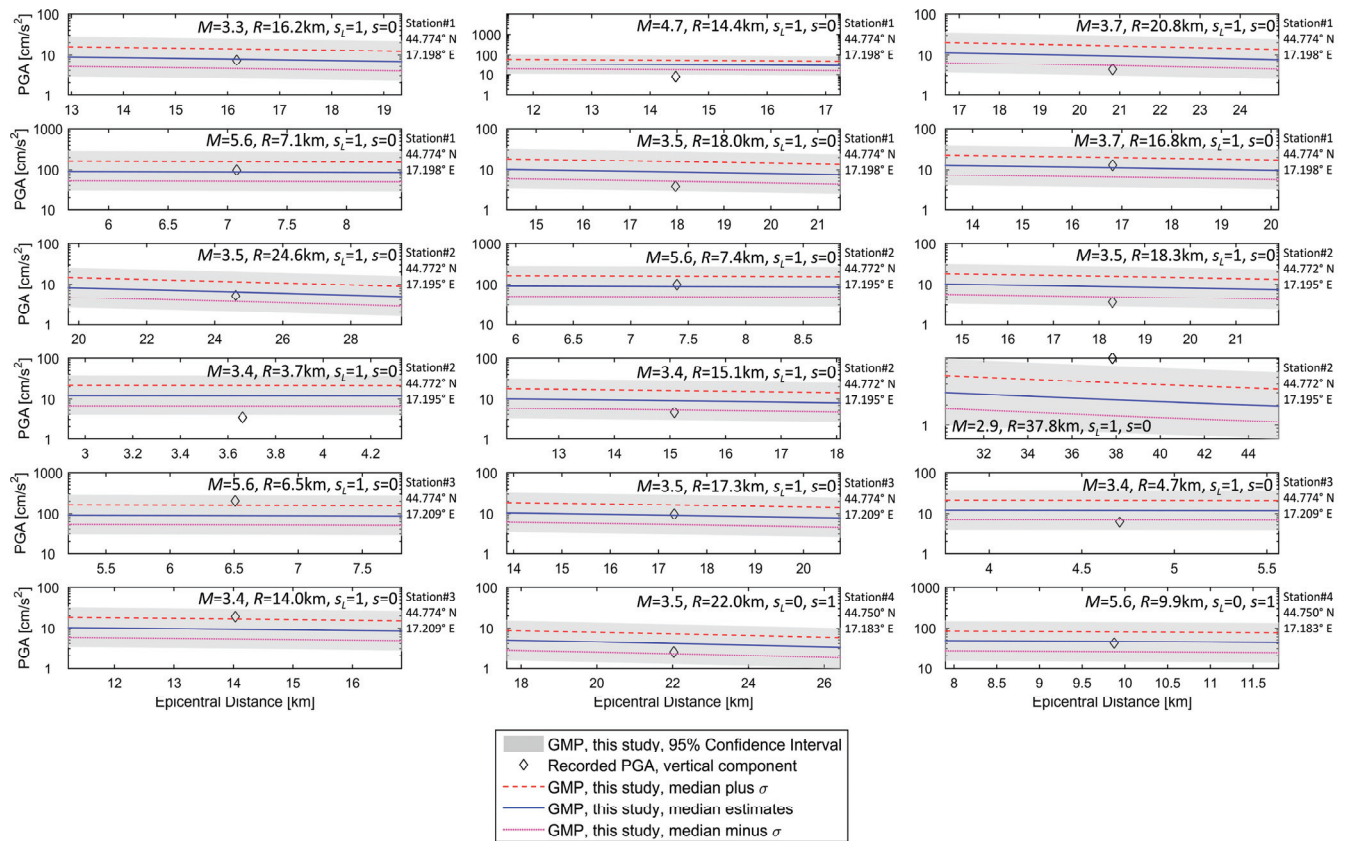


Figure 3. The PGA_{vert} values recorded by four different accelerograph stations in Banja Luka (see the red circles in Figure 1) vs. the median and median $\pm \sigma_{log}$ empirical PGA_{vert} estimates, which were calculated using Equation (3). The 95% confidence interval (median $\pm 2 \cdot \sigma_{log}$ empirical PGA_{vert} estimates) is displayed by the gray area.

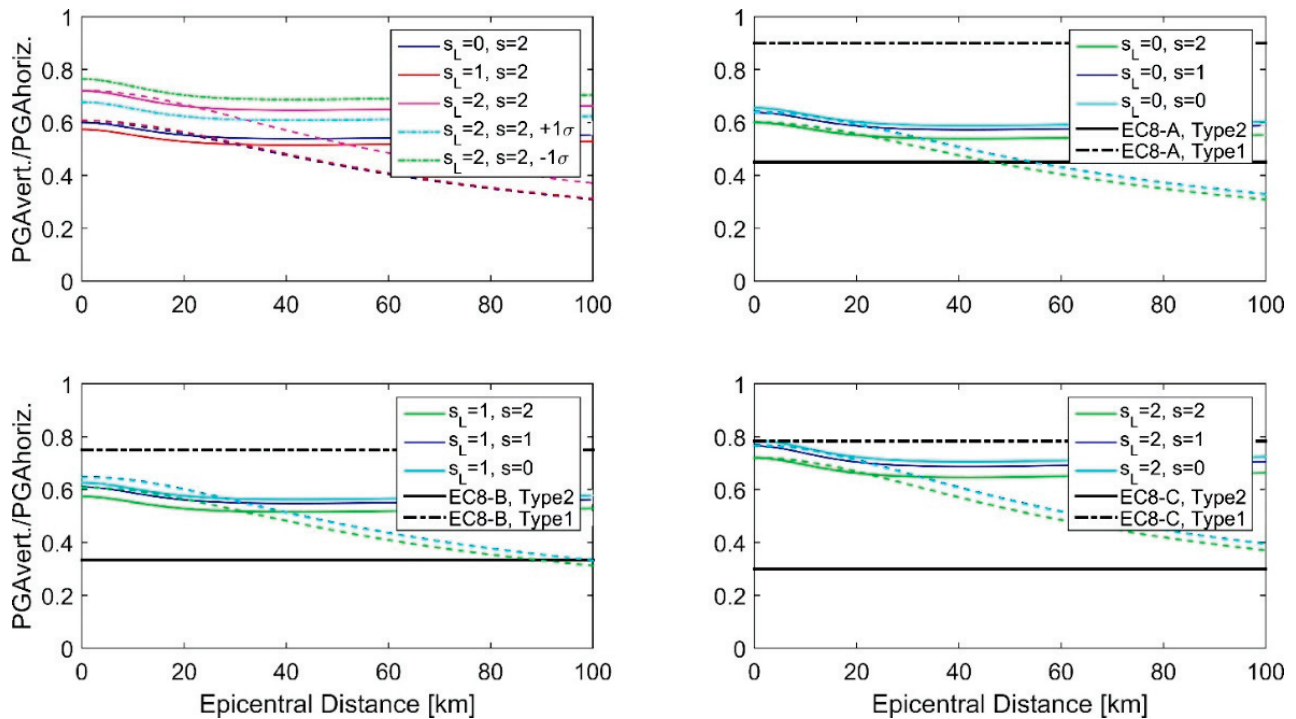


Figure 4. VH_{pga} values for source-to-site distances up to 100 km and different deep geology and local soil types, calculated by using Equations (3) (solid lines) and (4) (dashed lines), and corresponding equations for the horizontal direction [51].

Figure 4 displays the VH_{pga} for different site conditions and distances. The VH_{pga} values, as shown, are influenced by deep geology, local soil, source-to-site distance, and the probability level of empirical estimates. For distances up to 100 km, the empirical VH_{pga} ranges from 0.30 to 0.66 for the rock soil sites, from 0.30 to 0.65 for the stiff soil sites, and from 0.36 to 0.78 for the deep soils.

4. PSHA for the Case Study Site in Banja Luka

In this section, we will conduct a PSHA analysis for the location in Banja Luka with coordinates $44^{\circ}46.5' N$, $17^{\circ}15' E$, using Equation (3) as the GMPE. PSHA estimations will be generated with the REASSESS V2.1 software [52], which is based on Cornell–McGuire “deductive” methodology [53,54]. The PSHA computations are used to evaluate the likelihood of encountering a certain value of PGA_{vert} while accounting for all earthquakes that may occur in the analyzed area over the expected life of a building. The seismic sources in the surrounding area of the analyzed site are first identified. In the present study, we will employ the SHARE Project’s pan-European seismic source zone model [55–57]—see Figure 5, which also displays the epicenters of several large historical earthquakes, as well as a few recent catastrophic earthquakes in the northwestern Balkans [29]. This source zone model was based on the new homogeneous earthquake catalogue (SHEEC—the “SHARE European Earthquake Catalogue” [58]), which was compiled for the scope of the SHARE project. Thorough explanations and details regarding the completeness, homogeneity with respect to M_W , and other important features of this catalog can be found in [56,58–60] and will not be repeated here.

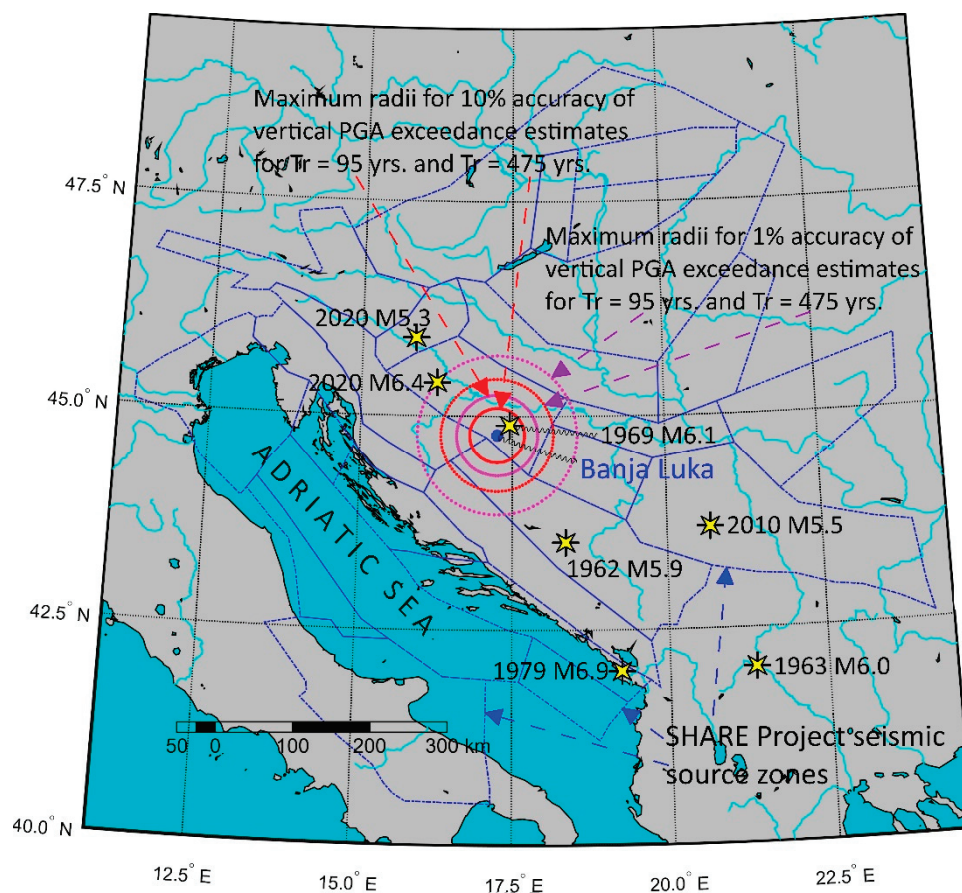


Figure 5. The seismic source zones from the SHARE Project [55–57] used to evaluate the PSHA for the site with coordinates $44^{\circ}46.5' N$, $17^{\circ}15' E$ in Banja Luka.

In Figure 5, the two larger circles, whose radii are 103.5 and 73.0 km, respectively, show the maximum distances that we must include in the PSHA computations to achieve 99 percent of the estimated hazard at the analyzed location for the return periods of 95 and 475 years; the two smaller circles with radii of 52.5 and 32.5 km, respectively, show the maximum distances needed to obtain 90 percent of the total hazard for the same return periods. Figure 5 shows that the *PGAvert* probabilistic estimates are dominated by local seismicity, with only the M6.4 1969 Banja Luka and M6.4 2020 Petrinja earthquakes occurring within the depicted radii, while all other events occur in areas that are not significantly contributing to the seismic hazard.

After the designation of seismic sources, the next step in the PSHA analysis will be to estimate the severity of ground motion at different distances. In the present study, we shall apply Equation (3) to this aim. Finally, the contributions from all analyzed source zones are aggregated and tallied across all magnitudes and distances. Recent regional microzonation investigations [23–27] found that horizontal PGA is dominated by local seismicity and is not particularly responsive to distant and very strong earthquakes, such as those that occur in the Vrancea area in Romania. It is reasonable to suppose that the *PGAvert* in Banja Luka will be dominated by local seismicity as well. Hence, we did not use a different scaling equation for the Vrancea region [61,62].

Earthquakes that contribute the most to the seismic hazard can be assessed using the technique commonly known as seismic hazard disaggregation [63]. Figure 6 shows the seismic hazard disaggregation at $44^{\circ}46.5' \text{ N}$, $17^{\circ}15' \text{ E}$ (refer to Figures 1 and 5) for the 95-, 475-, 975-, and 2475-year return periods.

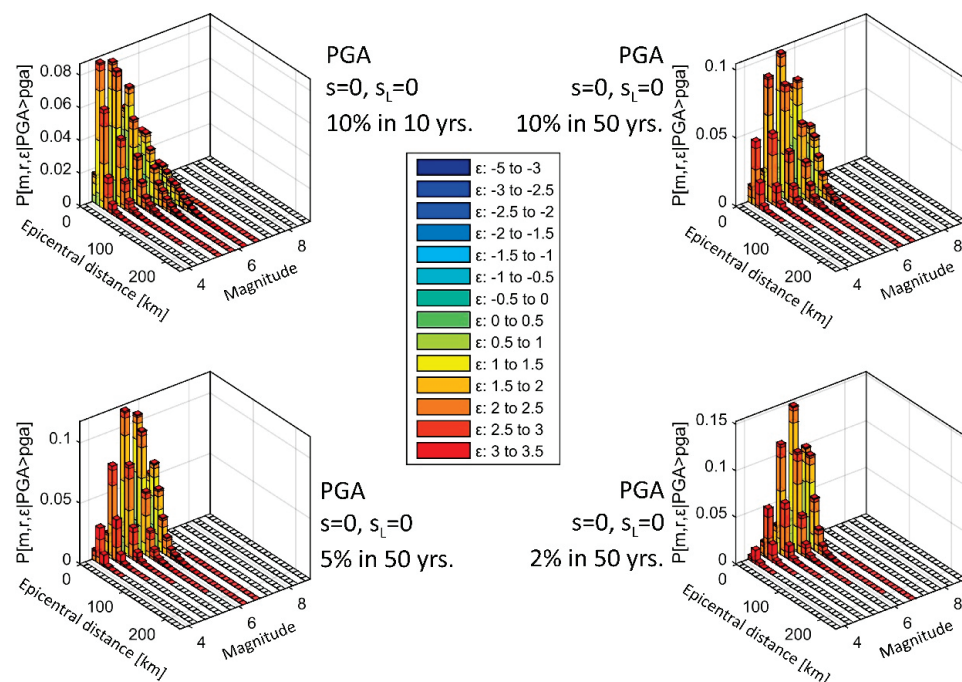


Figure 6. Seismic hazard disaggregation for location $44^{\circ}46.5' \text{ N}$, $17^{\circ}15' \text{ E}$ (see Figures 1 and 5) and four different probabilities of exceedance of PGA estimates.

The four plots displayed in Figure 6 show the PSHA disaggregation of *PGAvert* estimates for the location $44^{\circ}46.5' \text{ N}$, $17^{\circ}15' \text{ E}$ (solid blue circles in Figures 1 and 5) and the probabilities of exceedance of 10% in 10 years, and 10%, 5%, and 2% in 50 years. These probabilities correspond to the so-called return period, T_r , with values of 95, 475, 975, and 2475 years. These plots show that the most contributing magnitudes increase with the return period, T_r .

To verify our PSHA estimates of PGA_{vert} , in Figure 7 we compare the hazard curves for the analyzed location in Banja Luka (coordinates $44^{\circ}46.5' N$, $17^{\circ}15' E$ —see the solid blue circles in Figures 1 and 5) with the PGA_{vert} values given in Table 1. We will also compare our PSHA estimates to the PGA_{vert} calculated based on the horizontal PGA from the 2018 hazard maps [34]. It is worth noting here that former Yugoslav official maps were given for the average local soil conditions [36–39] and that none of the regional official seismic hazard maps so far considered the deeper geological formations beneath the local soil.

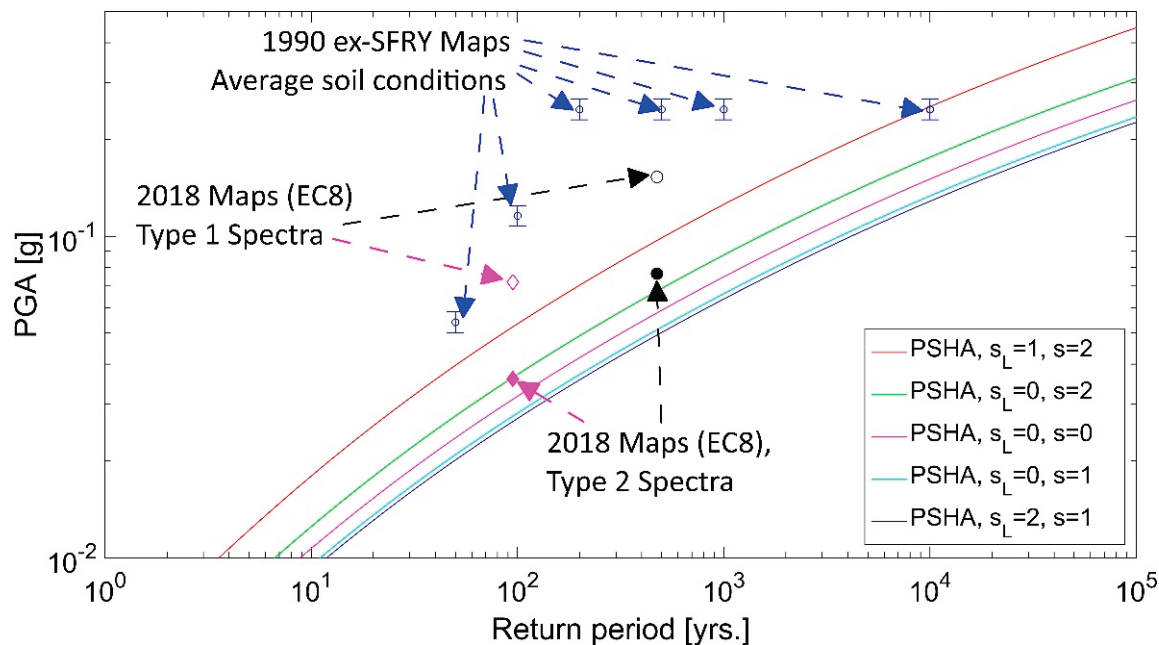


Figure 7. A comparison of the PGA_{vert} values for the location with coordinates $44^{\circ}46.5' N$, $17^{\circ}15' E$, in Banja Luka, calculated for different return periods, to the PGAs based on current official hazard maps for Bosnia and Herzegovina [34] and the VH_{pga} suggested by Eurocode 8 [1]. PGA_{vert} estimates based on the macroseismic intensity degrees from ex-Yugoslav maps [38] are also shown.

We can see in Figure 7 that the PGA_{vert} values pertaining to the former Yugoslav maps from 1990 [38] are significantly larger than our PGA_{vert} estimates for all return periods except for 10,000 years. However, our PGA_{vert} probabilistic estimates for deep geological rocks and “rock” soil sites agree very well with the 2018 maps [34].

5. Conclusions

With a population of around 180,000, Banja Luka is the cultural and financial center of the Republic of Srpska and the second most populated city in Bosnia and Herzegovina. The city is located between the Pannonian Basin and the Dinaric Alps and has scattered outcrops of geological rocks as well as areas of complex deep geology, in addition to areas with deep geological sediments. The city was struck by the catastrophic $M_W = 6.1$ Banja Luka earthquake on 27 October 1969, when 60% of all buildings were damaged beyond repair [64]. This demands a constant re-evaluation and update of Banja Luka’s seismic hazard estimations, including vertical PGA values.

Regional attenuation equations for vertical PGA are first described, and then PSHA analyses were done using these GMPEs, which take into account both local soil conditions and deep geology. Despite the small number of accelerograms from the northwestern Balkans utilized to produce the GMPEs, the empirical estimates are highly consistent with observed intensities in Banja Luka.

The most important findings of this study can be summarized as follows:

- Site conditions have a significant impact on the vertical PGA,
- The deep geology effects are even greater than the local soil effects,
- VHp_{ga} values are influenced by deep geology, local soil, source-to-site distance, and the probability level of empirical estimates,
- Only the combination of stiff soil and deep geology rocks may produce a median empirical vertical PGA equivalent to that associated with VIII °MCS, which was the predominant intensity of the 1969 Earthquake in Banja Luka,
- Vertical PGA probabilistic hazard estimates are dominated by local seismicity.

Furthermore, we should mention that the latest official (2018) hazard maps [34] provide PGA estimates that are comparable to our estimates of horizontal PGA for Banja Luka. It will be possible to obtain considerably better GMPEs only if the number of strong motion records in this region is increased. Until then, our results can be seen as a first step toward more realistic PHSA estimates for Banja Luka.

Author Contributions: Conceptualization, B.B., S.L. and S.B.; methodology, B.B. and S.L.; formal analysis, B.B., S.L., S.B. and M.K.; investigation, B.B., S.B., A.S.-Ć. and M.Š.; data curation, B.B.; writing—original draft preparation, B.B., S.L., S.B. and M.H.-N.; writing—review and editing, B.B., S.L., S.B., A.S.-Ć., M.Š. and M.H.-N.; visualization, B.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the authors.

Acknowledgments: The authors would like to thank the two anonymous reviewers for their insightful comments and suggestions, which lead to a significantly improved manuscript. For the first, third, fourth, and fifth authors, this research has been supported by the Ministry of Science, Technological Development and Innovation (Contract No. 451-03-137/2025-03/200156) and the Faculty of Technical Sciences, University of Novi Sad, through the project “Scientific and Artistic Research Work of Researchers in Teaching and Associate Positions at the Faculty of Technical Sciences, University of Novi Sad 2025” (No. 01-50/295).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

a_g	Horizontal PGA value from the official seismic hazard map
EC8	Eurocode 8
GMPE	Ground Motion Prediction Equations
PGA	Peak Ground Acceleration
PGA_{vert}	Peak Ground Acceleration for vertical ground motion direction
PSHA	Probabilistic Seismic Hazard Assessment
s	Deep geology parameter
S	Soil factor according to Eurocode 8
s_L	Local soil parameter
T_r	Return period
VHp _{ga}	Vertical to horizontal PGA ratio
V_s	Average shear wave velocity

References

1. EN 1998-1:2004; Eurocode 8: Design of Structures for Earthquake Resistance. Part 1: General Rules, Seismic Actions and Rules for Buildings. CEN: Brussels, Belgium, 2004.
2. Bulajić, B.Đ.; Pavić, G.; Hadzima-Nyarko, M. PGA vertical estimates for deep soils and deep geological sediments—A case study of Osijek (Croatia). *Comput. Geosci.* **2022**, *158*, 104985. [CrossRef]
3. Nemutlu, Ö.F.; Sari, A. Seismic Performance of the Kahramanmaraş Earthquakes: Ground Motion Prediction Models and the Role of Vertical Components. *Bitlis Eren Üniversitesi Fen Bilim. Derg.* **2025**, *14*, 610–632. [CrossRef]
4. Nogoshi, M.; Igarashi, T. On the Amplitude Characteristics of Microtremor (Part 1). *Zisin (J. Seismol. Soc. Japan. 2nd Ser.)* **1970**, *23*, 281–303. [CrossRef]
5. Nogoshi, M.; Igarashi, T. On the Amplitude Characteristics of Microtremor (Part 2). *Zisin (J. Seismol. Soc. Japan. 2nd Ser.)* **1971**, *24*, 26–40. [CrossRef]
6. Nakamura, Y. Clear identification of fundamental idea of Nakamura's technique and its applications. In Proceedings of the 12th World Conference on Earthquake Engineering, Auckland, New Zealand, 30 January–4 February 2000; pp. 1–8.
7. Nakamura, Y. *A method for Dynamic Characteristics Estimation of Subsurface Using Microtremor on The Ground Surface*; Railway Technical Research Institute: Tokyo, Japan, 1989; pp. 25–30.
8. Nakamura, Y. Real-time information systems for hazards mitigation. In Proceedings of the 11th World Conference on Earthquake Engineering, Acapulco, Mexico, 23–28 June 1996.
9. Bonnefoy-Claudet, S.; Köhler, A.; Cornou, C.; Wathélet, M.; Bard, P.-Y. Effects of Love Waves on Microtremor H/V Ratio. *Bull. Seismol. Soc. Am.* **2008**, *98*, 288–300. [CrossRef]
10. Trifunac, M.D.; Brady, A.G. On the correlation of seismic intensity scales with the peaks of recorded strong ground motion. *Bull. Seismol. Soc. Am.* **1975**, *65*, 139–162.
11. Seed, H.B.; Murarka, R.; Lysmer, J.; Idriss, I.M. Relationships of maximum acceleration, maximum velocity, distance from source, and local site conditions for moderately strong earthquakes. *Bull. Seismol. Soc. Am.* **1976**, *66*, 1323–1342.
12. Seed, H.B.; Ugas, C.; Lysmer, J. Site-dependent spectra for earthquake-resistant design. *Bull. Seismol. Soc. Am.* **1976**, *66*, 221–243. [CrossRef]
13. Bulajić, B.Đ.; Bajić, S.; Stojnić, N. The effects of geological surroundings on earthquake-induced snow avalanche prone areas in the Kopaonik region. *Cold Reg. Sci. Technol.* **2018**, *149*, 29–45. [CrossRef]
14. Bulajić, B.Đ.; Hadzima-Nyarko, M.; Pavić, G. Vertical to Horizontal UHS Ratios for Low to Medium Seismicity Regions with Deep Soil atop Deep Geological Sediments—An Example of the City of Osijek, Croatia. *Appl. Sci.* **2021**, *11*, 6782. [CrossRef]
15. Bulajić, B.Đ.; Hadzima-Nyarko, M.; Pavić, G. Horizontal UHS Amplitudes for Regions with Deep Soil Atop Deep Geological Sediments—An Example of Osijek, Croatia. *Appl. Sci.* **2021**, *11*, 6296. [CrossRef]
16. Bulajic, B.Đ.; Hadzima-Nyarko, M.; Pavic, G. PGA estimates for deep soils atop deep geological sediments—An example of Osijek, Croatia. *Geomech. Eng.* **2022**, *30*, 233–246. [CrossRef]
17. Bulajić, B.Đ.; Manić, M.I.; Ladinović, Đ. Effects of shallow and deep geology on seismic hazard estimates: A case study of pseudo-acceleration response spectra for the northwestern Balkans. *Nat. Hazards* **2013**, *69*, 573–588. [CrossRef]
18. Jordanovski, L.R.; Lee, V.W.; Manić, M.I.; Olumčeva, T.; Sinadnovski, C.; Todorovska, M.I.; Trifunac, M.D. *Strong Earthquake Ground Motion Data in EQINFOS: Yugoslavia. Part I*; Department of Civil Engineering, University of Southern California: Los Angeles, CA, USA, 1987.
19. Lee, V.W. *Influence of Local Soil and Geologic Site Conditions on Pseudo Relative Velocity Spectrum Amplitudes of Recorded Strong Motion Accelerations*; Department of Civil Engineering, University of Southern California: Los Angeles, CA, USA, 1987.
20. Trifunac, M.; Lee, V.; Živčić, M.; Manić, M. On the correlation of Mercalli-Cancani-Sieberg intensity scale in Yugoslavia with the peaks of recorded strong earthquake ground motion. *Eur. Earthq. Eng.* **1991**, *5*, 27–33.
21. Lee, V.; Trifunac, M. Empirical scaling of Fourier amplitude spectra in former Yugoslavia. *Eur. Earthq. Eng.* **1993**, *7*, 47–61.
22. Lee, V.W.; Manić, M.I. Empirical scaling of response spectra in former Yugoslavia. In Proceedings of the 10th European Conference on Earthquake Engineering, Vienna, Austria, 28 August–2 September 1994; pp. 2567–2572.
23. Lee, V.W.; Trifunac, M.D.; Bulajić, B.Đ. Seismic microzoning in Novi Sad, Serbia—A case study in a low-seismicity region that is exposed to large and distant earthquakes. *J. Seismol.* **2023**, *27*, 979–997. [CrossRef]
24. Lee, V.W.; Trifunac, M.D.; Bulajić, B.Đ. Seismic hazard mapping for peak ground velocity: Microzonation of Novi Sad, Serbia—A case study in a low-seismicity region exposed to large and distant earthquakes. *J. Seismol.* **2025**, *29*, 85–105. [CrossRef]
25. Lee, V.W.; Trifunac, M.D.; Bulajić, B.Đ.; Manić, M.I.; Herak, D.; Herak, M. Seismic microzoning of Belgrade. *Soil Dyn. Earthq. Eng.* **2017**, *97*, 395–412. [CrossRef]
26. Lee, V.W.; Trifunac, M.D.; Bulajić, B.Đ.; Manić, M.I.; Herak, D.; Herak, M.; Dimov, G. Seismic microzoning in Skopje, Macedonia. *Soil Dyn. Earthq. Eng.* **2017**, *98*, 166–182. [CrossRef]
27. Lee, V.W.; Trifunac, M.D.; Bulajić, B.Đ.; Manić, M.I.; Herak, D.; Herak, M.; Dimov, G.; Gičev, V. Seismic microzoning of Štip in Macedonia. *Soil Dyn. Earthq. Eng.* **2017**, *98*, 54–66. [CrossRef]

28. Ustaszewski, K.; Herak, M.; Tomljenović, B.; Herak, D.; Matej, S. Neotectonics of the Dinarides–Pannonian Basin transition and possible earthquake sources in the Banja Luka epicentral area. *J. Geodyn.* **2014**, *82*, 52–68. [CrossRef]
29. USGS. *Earthquake Catalogue for all Earthquakes with $M_w \geq 3.0$ in the Period Between 1900 and April 2025 for the Geographic Region Between 41° N and 47° N, and 13° E and 23° E*; USGS: Reston, VA, USA, 2025.
30. EMSC-CSEM. M 5.4–CROATIA–2020-03-22 05:24:02 UTC. Available online: <https://www.emsc-csem.org/Earthquake/earthquake.php?id=840695#scientific> (accessed on 17 April 2025).
31. Ganas, A.; Elias, P.; Valkaniotis, S.; Tsironi, V.; Karasante, I.; Briole, P. Petrinja Earthquake Moved Crust 10 Feet. Available online: <https://temblor.net/earthquake-insights/petrinja-earthquake-moved-crust-10-feet-12410/> (accessed on 17 April 2025).
32. Lee, V.W.; Manić, M.I.; Bulajić, B.Đ.; Herak, D.; Herak, M.; Trifunac, M.D. Microzonation of Banja Luka for performance-based earthquake-resistant design. *Soil Dyn. Earthq. Eng.* **2015**, *78*, 71–88. [CrossRef]
33. Mojičević, M.; Vilovski, S.; Tomić, B. *Basic Geological Map–Banja Luka, L33-119*; Federal Geological Survey: Belgrade, Serbia, 1976.
34. BAS EN 1998-1/NA:2018; Seismic Zone Maps and Reference Ground Accelerations Therein of B&H. Maps Accompanying National Annexes (NA): Bosnia and Herzegovina, 2018. Available online: <http://eurokodovi.ba/seizmika/> (accessed on 1 May 2025).
35. Ademović, N.; Demir, V.; Cvijić-Amulić, S.; Málek, J.; Prachař, I.; Vackář, J. Compilation of the seismic hazard maps in Bosnia and Herzegovina. *Soil Dyn. Earthq. Eng.* **2021**, *141*, 106500, Erratum in *Soil Dyn. Earthq. Eng.* **2023**, *164*, 107633. [CrossRef]
36. SFRY. *Temporary Technical Regulations for Construction in Seismic Areas*; 39/64; Official Gazette of SFRY: Belgrade, Serbia, 1964.
37. SFRY. *Book of Rules on Technical Norms for Construction of High-Rise Buildings in Seismic Regions*; 31/81, 49/82; Official Gazette of SFRY: Belgrade, Serbia, 1982.
38. SFRY. *Book of Rules on Technical Norms for Construction of High-Rise Buildings in Seismic Regions*; 31/81, 49/82, 29/83, 21/88, 52/90; Official Gazette of SFRY: Belgrade, Serbia, 1990.
39. SFRY. *Book of Rules on Technical Norms for Construction of High-Rise Buildings in Seismic Regions*; 31/81; Official Gazette of SFRY: Belgrade, Serbia, 1981.
40. Lee, V.; Trifunac, M.; Herak, M.; Živčić, M.; Herak, D. M_L^{SM} computed from strong motion accelerograms recorded in Yugoslavia. *Earthq. Eng. Struct. Dyn.* **1990**, *19*, 1167–1179. [CrossRef]
41. Datta, T.K. *Seismic Analysis of Structures*; John Wiley & Sons (Asia) Pte Ltd.: Singapore, 2010; p. 464.
42. Herak, M. Conversion between the local magnitude (M_L) and the moment magnitude (M_w) for earthquakes in the Croatian Earthquake Catalogue. *Geofizika* **2020**, *37*, 197–211. [CrossRef]
43. Scordilis, E.M. Empirical Global Relations Converting M_S and m_b to Moment Magnitude. *J. Seismol.* **2006**, *10*, 225–236. [CrossRef]
44. Das, R.; Sharma, M.L.; Wason, H.R.; Choudhury, D.; Gonzalez, G. A Seismic Moment Magnitude Scale. *Bull. Seismol. Soc. Am.* **2019**, *109*, 1542–1555. [CrossRef]
45. Das, R.; Das, A. Limitations of M_w and M Scales: Compelling Evidence Advocating for the Das Magnitude Scale (M_{wg})—A Critical Review and Analysis. *Indian Geotech. J.* **2025**, *in press*. [CrossRef]
46. Ambraseys, N.; Douglas, J.; Margaris, B.; Sigbjörnsson, R.; Berge-Thierry, C.; Suhadolc, P.; Costa, G.; Smit, P. Dissemination of European strong-motion data. In Proceedings of the 13th World Conference on Earthquake Engineering, Vancouver, BC, Canada, 1–6 August 2004.
47. Ambraseys, N.; Douglas, J.; Margaris, B.; Sigbjörnsson, R.; Smit, P.; Suhadolc, P. Internet site for European strong motion data. In Proceedings of the 12th European Conference on Earthquake Engineering, London, UK, 9–13 September 2002.
48. Bilgin, H.; Hysenlliu, M. Comparison of near and far-fault ground motion effects on low and mid-rise masonry buildings. *J. Build. Eng.* **2020**, *30*, 101248. [CrossRef]
49. Bulajić, B.Đ.; Lozančić, S.; Bajić, S.; Radu, D.; Işık, E.; Negovanović, M.; Hadzima-Nyarko, M. UHS estimates for vertical strong ground motion and varying deep geological site surroundings—A case study of the city of Banja Luka. *Appl. Sci.* **2025**, submitted for publication.
50. Bulajić, B.Đ.; Lozančić, S.; Bajić, S.; Radu, D.; Işık, E.; Negovanović, M.; Hadzima-Nyarko, M. Horizontal UHS predictions for varying deep geology conditions—A case study of the city of Banja Luka. *Sustainability* **2025**, submitted for publication.
51. Bulajić, B.Đ.; Lozančić, S.; Bajić, S.; Starčev-Ćurčin, A.; Šešlija, M.; Kovačević, M.; Hadzima-Nyarko, M. Horizontal PGA estimates for varying deep geological conditions—A case study of Banja Luka. *Appl. Sci.* **2025**, submitted for publication.
52. Chioccarelli, E.; Cito, P.; Iervolino, I.; Giorgio, M. REASSESS V2.0: Software for single- and multi-site probabilistic seismic hazard analysis. *Bull. Earthq. Eng.* **2019**, *17*, 1769–1793. [CrossRef]
53. McGuire, R.K. *FORTTRAN Computer Program for Seismic Risk Analysis*; US Geological Survey: Reston, VA, USA, 1976; pp. 67–76.
54. Cornell, C.A. Engineering seismic risk analysis. *Bull. Seismol. Soc. Am.* **1968**, *58*, 1583–1606. [CrossRef]
55. Giardini, D.; Woessner, J.; Danciu, L.; Crowley, H.; Cotton, F.; Grünthal, G.; Pinho, R.; Valensise, L.; Consortium, S. *European Seismic Hazard Map for Peak Ground Acceleration, 10% Exceedance Probabilities in 50 Years*; Swiss Seismological Service: Zurich, Switzerland, 2013.

56. Woessner, J.; Laurentiu, D.; Giardini, D.; Crowley, H.; Cotton, F.; Grünthal, G.; Valensise, G.; Arvidsson, R.; Basili, R.; Demircioglu, M.B.; et al. The 2013 European Seismic Hazard Model: Key components and results. *Bull. Earthq. Eng.* **2015**, *13*, 3553–3596. [CrossRef]
57. Pagani, M.; Garcia-Pelaez, J.; Gee, R.; Johnson, K.; Poggi, V.; Styron, R.; Weatherill, G.; Simionato, M.; Viganò, D.; Danciu, L.; et al. *Global Earthquake Model (GEM) Seismic Hazard Map (Version 2018.1–December 2018)*; GEM Foundation: Pavia, Italy, 2018.
58. Stucchi, M.; Rovida, A.; Gomez Capera, A.A.; Alexandre, P.; Camelbeeck, T.; Demircioglu, M.B.; Gasperini, P.; Kouskouna, V.; Musson, R.M.W.; Radulian, M.; et al. The SHARE European Earthquake Catalogue (SHEEC) 1000–1899. *J. Seismol.* **2013**, *17*, 523–544. [CrossRef]
59. Grünthal, G.; Wahlström, R.; Stromeyer, D. The SHARE European Earthquake Catalogue (SHEEC) for the time period 1900–2006 and its comparison to the European-Mediterranean Earthquake Catalogue (EMEC). *J. Seismol.* **2013**, *17*, 1339–1344. [CrossRef]
60. Grünthal, G.; Wahlström, R. The European-Mediterranean Earthquake Catalogue (EMEC) for the last millennium. *J. Seismol.* **2012**, *16*, 535–570. [CrossRef]
61. Lee, V.W.; Trifunac, M.D.; Bulajić, B.; Manić, M. A preliminary empirical model for frequency-dependent attenuation of Fourier amplitude spectra In Serbia from the Vrancea earthquakes. *Soil Dyn. Earthq. Eng.* **2016**, *83*, 167–179. [CrossRef]
62. Lee, V.W.; Trifunac, M.D.; Bulajić, B.D.; Manić, M.I. Preliminary empirical scaling of pseudo relative velocity spectra in Serbia from the Vrancea earthquakes. *Soil Dyn. Earthq. Eng.* **2016**, *86*, 41–54. [CrossRef]
63. Bazzurro, P.; Allin Cornell, C. Disaggregation of seismic hazard. *Bull. Seismol. Soc. Am.* **1999**, *89*, 501–520. [CrossRef]
64. Trkulja, D. Earthquakes in Banja Luka. In Proceedings of the International Conference on Earthquake Engineering, Banja Luka, Bosnia and Herzegovina, 26–28 October 2009; pp. 43–57.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Integrating Radon/Thoron and Gamma Radiation Exposure for a Realistic Estimation of Dose Arising from Building Materials

Mirsina M. Aghdam ^{1,2} and Quentin Crowley ^{1,2,3,*}

¹ Geology, School of Natural Sciences, Trinity College Dublin, D08 XW7X Dublin, Ireland; senan.aghdam@aol.com

² Trinity Center for the Environment, Trinity College Dublin, D08 XW7X Dublin, Ireland

³ R&D Department, Geochron Ltd., D18 C6FW Dublin, Ireland

* Correspondence: crowleyq@tcd.ie

Featured Application: The dose prediction and air ventilation models for conventional homes can be utilized by construction professionals and regulators to evaluate and manage radiation risks from radon, thoron, and gamma rays emitted from building materials. These models assist in establishing safety standards and guidelines for indoor environments, promoting safer construction practices and enhancing public health.

Abstract: Long-term exposure to radon, thoron, and gamma radiation from building materials poses a significant health risk to occupants. Current methods for estimating radiation doses often fail to consider the combined impact of these sources. Based on commonly used building materials available on the Irish market, this paper advocates for the development of a comprehensive dose estimation model that accounts for radon, thoron, and gamma radiation. To achieve this, several models and various scenarios (e.g., ventilation conditions and building characteristics) are integrated to convert radon and thoron gas doses into a common unit recognized in the existing literature. This approach enables the comparison of combined dose values with accepted radiation thresholds for building materials, typically set at 1 mSv, alongside data on material compositions. Previous studies suggested gamma radiation doses in Irish materials are unlikely to exceed 1 mSv annually. Our findings confirm this, showing gamma doses <0.4 mSv for all materials. However, combined radon–thoron doses exceeded thresholds in altered granites (e.g., Galway granite: 3.90 mSv), with thoron contributing $\leq 93\%$ of total exposure due to uranium/thorium-rich minerals (e.g., monazite, zircon). Ventilation proved critical—high airflow (10 m³/h) reduced thoron doses by 90–95%, while current gamma-focused safety indices (I-index ≤ 1) inadequately addressed combined risks. These results highlight the previously underestimated importance of thoron and the necessity of multi-parameter models for regulatory compliance. The study establishes a novel framework to evaluate holistic radiation risks, urging revised standards that prioritize ventilation strategies and material mineralogy to protect public health in residential and commercial built environments.

Keywords: radon; thoron; building material; radiation dose; ventilation; simulation models

1. Introduction

Radon (Rn-222), a naturally occurring radioactive gas from the decay of uranium in rocks, soils, and water, can accumulate to high concentrations indoors. Prolonged exposure can pose significant health risks, specifically lung cancer [1]. Thoron (Rn-220),

like radon, is a geogenically produced radioactive gas but is a decay product of thorium [1]. While both gases pose considerable health risks due to their radioactive nature, thoron is often less discussed than radon. Recent assessments indicate that the contribution of thoron to radiation dose was previously misunderstood and, therefore, underestimated [2]. This underestimation stems from a historical focus on radon and gamma radiation in regulatory frameworks, neglecting the synergistic effects of thoron (UNSCEAR, 2022 [3]). One reason for this omission is the shorter half-life of thoron (55.6 s), compared to that of radon (3.8 days), which limits its ability to diffuse far from its source. Consequently, thoron concentrations typically decrease rapidly with distance from emitting surfaces, making it more challenging to detect and measure accurately compared to radon [4].

Despite these challenges, recent assessments indicate that the decay products of thoron, which are the actual contributors to the dose, have long enough half-lives to reach indoor environments even from the soil beneath building foundations [5]. However, it is understood that building materials and natural stones used in interiors are the primary sources of thoron [6]. Radon, including exhalation from soil, building materials, and water, remains the main source of radiation exposure for the general public, contributing almost 60% to the annual average dose in Ireland, as highlighted by the Environmental Protection Agency (EPA) and the Health Information and Quality Authority (HIQA) in their recent ‘Ionizing Radiation—National Dose Report—June 2024’ [7]. The report, which assesses the radiation exposure received from various sources, estimates over 8% contribution from thoron. Notably, Ireland’s thoron contribution exceeds the global average of 5% reported by the World Health Organization (WHO, 2021 [8]), while approximately 6% of the radiation dose comes from gamma radiation originating from the ground.

More homes are at risk from radon than previously thought, and about half a million (10%) of the total Irish population may be affected by high indoor radon levels. Furthermore, the average radiation dose received is 4.2 mSv [7], with more than two-thirds arising from radon and thoron. This remains among the highest in Europe [9], highlighting the importance of addressing both radon and thoron exposure, the major contributors to the dose, in indoor environments. This study introduces a novel methodology to bridge critical gaps in radiation risk assessment by integrating gamma radiation with radon and thoron exhalation dynamics under varied ventilation scenarios—an approach not yet standardized in existing guidelines [10]. Both the methodology and findings are directly applicable to radiation protection in other countries.

Rock and aggregate samples were collected from materials of both Irish origin (quarries and stone markets) and non-Irish origin (imported), covering a wide range of rock types. As there is no legal requirement on radon and thoron levels in building materials, models were developed for different housing types to convert radon and thoron activity to effective dose (mSv), which is the standard metric in scientific and regulatory contexts to assess and verify building and construction limits. A 1 mSv annual dose level is suggested as the threshold for building materials used indoors [11]. Existing studies, like [12], focus primarily on radon or gamma radiation in isolation, whereas our model quantifies combined thresholds and ventilation efficacy. Finally, mathematical models were examined to determine the optimal amount of rock surface coverage or appropriate ventilation rates needed to maintain the total effective dose (from both radon/thoron and gamma radiation) below the recommended levels.

The model developed in this study presents a novel methodology to quantify the effective dose from natural radioactive sources in building materials, which is beneficial for further strengthening guidelines and regulations to protect public health. This work addresses the lack of harmonized frameworks for multi-source radiation risk assessment, as identified by the International Commission on Radiological Protection (ICRP, 2007 [13]).

Future research could expand these models to include radon and thoron emanating from the soil and rock [14] beneath building foundations, as well as radon from water sources. Such expansions would contribute to a comprehensive view of the doses from all natural radioactive sources, further enhancing the applicability of the radiation dose assessments conducted in this research.

2. Materials and Methods

2.1. Sample Collection

There are limited studies relating to gamma radiation in building materials in Ireland [15], and the previous research does not discuss the radon/thoron concentrations, exhalation rates of these gases, or the associated doses. Therefore, in the present study, we aimed to include as many samples as possible to provide a reasonable coverage of the most popular, available, and abundant materials in Ireland. These materials, which are also widely available in the EU and some of which are available worldwide, include the following: (1) natural stone available in the Irish market, both locally sourced and imported rock (38 samples); (2) products from active quarries with potential for elevated radon/thoron activity (20 samples); (3) aggregate samples, which are the main components of concrete production (these samples were tested but are not included in this paper); (4) construction and demolition waste intended for recycling (8 samples); and (5) rock core samples provided by Geological Survey Ireland (12 samples). It is noteworthy that estimating radon and thoron exhalation rates could be of interest in other applications, such as geothermal energy potential characterization, as rocks with high radon/thoron potential may indicate a suitable environment for geothermal drilling. Supplementary Table S1 contains a full list of materials tested, their images and selected properties.

2.2. Gamma Radiation Testing

To quantify the levels of uranium (U-238 mg/kg), potassium (K-40%), and thorium (Th-232 mg/kg), and the dose rates in the selected materials, we utilized the Georadis (Brno, Czech Republic) GT40s equipped with a NaI (TI 76 × 76 mm, 345 cm³, with bi-alkali PMT gamma detector in assay mode. In our experimental setup (Figure 1), a lead collimator was placed over each sample to ensure that the gamma radiation detected was primarily from a defined area on the sample's surface. Specifically, the lead collimator restricts the gamma radiation detection to a one-dimensional surface, to isolate a circular surface area with a diameter of 2.5 cm. This configuration allowed for an approximate measurement of gamma radiation emanating from the targeted section of the sample.

2.3. Radon and Thoron Exhalation Rate Testing

Seventy-eight representative rock samples underwent radon and thoron testing using the RTM1688-2 (SARAD, Dresden, Germany) Monitor Setup, which is calibrated according to DIN ISO/IEC EN 17025:2018 standards [16]. The samples were placed in an airtight vacuum desiccator (Figure 1) to prevent gas leakage, verified by a gas leak detector spray. A small fan was used to ensure uniform air circulation within the chamber. We allowed between 36 and 48 h for radon and thoron equilibrium to be established between the samples and the surrounding air, with measurements taken at 1-h intervals to accurately capture gas emissions during this period. Considering Equations (1) and (2), the exhalation rates of ²²²Rn (E222Rn, Bq m⁻² h⁻¹) and ²²⁰Rn (E220Rn, Bq m⁻² h⁻¹) were calculated by extrapolating the slope of the growth curve (m) (Bq m⁻³ h⁻¹) and the equilibrium Rn-220

concentration (C_m) (Bq m^{-3}), respectively [17]. Figure 2 shows an example of $^{222}\text{Rn}/^{220}\text{Rn}$ concentrations as a function of time and the derived exhalation rates.

$$E_{222\text{Rn}} = \frac{(m + \lambda_{222} \times C_0) \times V}{S} \quad (1)$$

$$E_{220\text{Rn}} = \frac{\lambda_{220} \times V_0}{S} \frac{C_m}{e^{-\lambda_{220} \times (V_1/Q)}} \quad (2)$$

where λ_{222} and λ_{220} are ^{222}Rn and ^{220}Rn decay constants (h^{-1}), C_0 is the initial radon concentration (Bq m^{-3}), V is the free total volume of the analytical system (m^3), S is the surface of the sample, V_0 and V_1 (m^3) are the free volume of the accumulation chamber and the volume between the outflow of the accumulation chamber and the inflow of the radon monitor, respectively. Q (15 L h^{-1}) is the flow rate in the system.

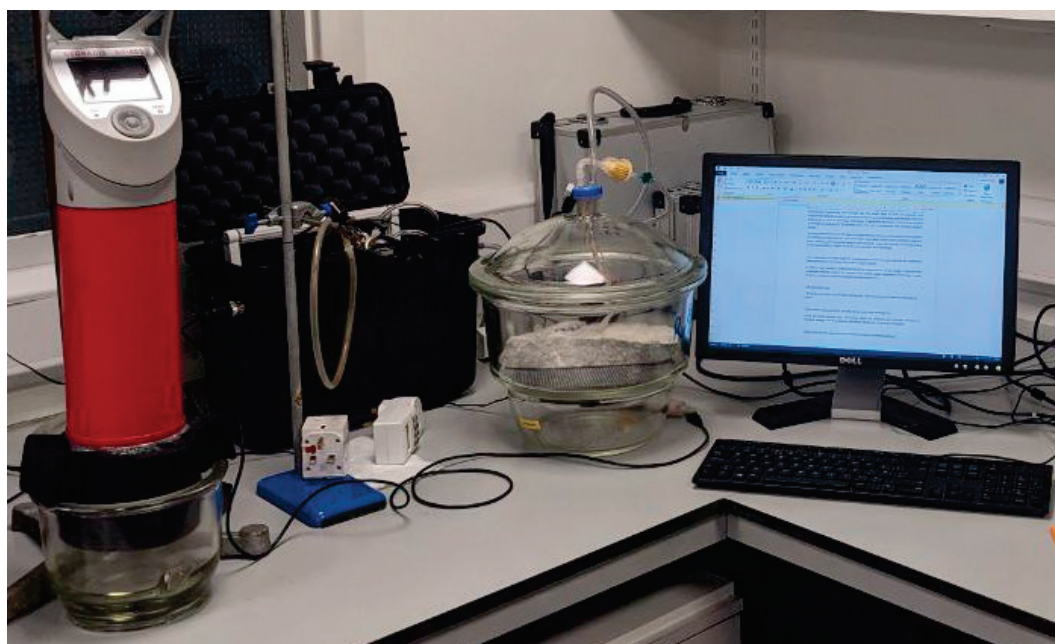


Figure 1. Radon/thoron and gamma measurement setup at the Environmental Radioactivity Laboratory, Trinity Center for the Environment.

2.4. Dose Conversion and Ventilation Models

2.4.1. Dose from Gamma Radiation

The gamma dose index, or I-index, is a simplified radiological assessment tool used to estimate the potential external gamma radiation exposure from building materials [18]. It is calculated using the concentrations of the three radionuclides as follows (Equation (3)):

$$(\text{mSv}) = C_U/300 (\text{Bqkg}^{-1}) + C_{\text{Th}}/200 (\text{Bqkg}^{-1}) + C_K/3000 (\text{Bqkg}^{-1}) \quad (3)$$

where C_U is the concentration of uranium (U-238) in Bqkg^{-1} , and 12.35 Bqkg^{-1} corresponds to 1 mg/kg of uranium. Similarly, C_{Th} is the concentration of thorium (Th-232) in Bqkg^{-1} , with 4.06 Bqkg^{-1} equating to 1 mg/kg of thorium. For potassium (K-40), C_K represents the concentration in Bqkg^{-1} , and 313 Bqkg^{-1} corresponds to 0.1% of potassium [19].

2.4.2. Dose from Radon and Thoron

Calculating the dose from radon and thoron in indoor environments is a complex process due to the various factors influencing the concentration and distribution of these gases [5]. Radon and thoron levels depend on several variables, including the type of

building materials used, the ventilation rate, the design of the building, and the time occupants spend indoors. To address these complexities, three common housing scenarios are considered: a one-bedroom house, a two-bedroom house, and a three-bedroom house. A model of each housing scenario was created to estimate the effective (net) volumes for different building type scenarios. This technique helped to visualize and calculate the surface areas of potential radon sources in homes, determine the total volume of furniture inside the house, estimate the net air volumes for different housing types, and analyze data from numerous real-world cases. By studying dozens of actual instances, we approximated the range of surface areas of natural stones—recognized sources of radon and thoron in indoor environments—that could be present in each housing type (Table 1). Additionally, to calculate the effective volume of indoor spaces, we subtracted the volume occupied by furniture and appliances from the total volume of the house. This allowed an estimation of the effective space in which radon and thoron can accumulate, providing a more accurate calculation of the potential exposure doses for the occupants.



Figure 2. Schematic model for the visualization of radon and thoron emission and the sources of radioactive gases and gamma radiation (including kitchen worktops, kitchen island, floor, tiles, and hand basin top) within a one-bedroom house. In this instance, the effective volume is estimated at 121.6 m^3 , with stone (e.g., granite, limestone and sandstone) coverage ranging from 42 to 80 m^2 .

Table 1. Annual gamma, radon, and thoron doses (mSv/year) for Irish building materials. Altered granites show the highest combined doses.

Model	1-Bedroom House (See Figure 2)	2-Bedroom House	3-Bedroom House
V effective (V_{eff}) m^3	121.16	230	380
Stone Surface Area (s) range m^2 (10 to 30% of the total Surface Area of the house)	42–80	72–123	118–180

$F_{\text{Rn-222}}$ and $F_{\text{Rn-220}}$ are radon and thoron conversion factors equal to 2.7×10^{-6} and 7.0×10^{-6} mSv per Bqm^{-3} , respectively. $C_{\text{Rn-222}}$ and $C_{\text{Rn-220}}$ represent the average radon and thoron exhalation per surface Area (Bqm^{-2}).

Equation (4) for calculation of Dose [20,21]:

$$\text{Dose (mSv)} = (F_{\text{Rn-222}} C_{\text{Rn-222}} (\text{Bqm}^{-2}) + F_{\text{Rn-220}} C_{\text{Rn-220}} (\text{Bqm}^{-2})) S(\text{m}^2) / V_{\text{eff}} (\text{m}^3) \quad (4)$$

2.4.3. Total Effective Dose and Suggested Dose Limit

The I-index is used to assess whether building material is safe for use based only on its gamma radiation levels. According to the European Union guidelines and other international standards,

$I \leq 1$: The material is considered safe for use in building construction without restrictions.

$I > 1$: The material may pose a higher radiological risk, and further evaluation may be necessary to determine whether it is suitable for use, particularly in bulk quantities or in buildings where occupants might spend prolonged periods [22].

At the time of writing, there is no recognized index for the dose of radon and thoron emitted from building materials. It is possible that for some materials, the I index is below 1 mSv but adding the dose from radon and thoron might increase the total dose over the 1 mSv limit.

2.4.4. Modeling the Effect of Air Ventilation

Models were developed to simulate the effectiveness of ventilation rates in reducing exposure to radon and thoron gas for samples where the estimated annual dose from radon and thoron (in the worst-case scenario from Section 2.4.2) exceeded 0.3 mSv. Three ventilation scenarios were considered: low ($0.5 \text{ m}^3/\text{h}$), medium ($2 \text{ m}^3/\text{h}$), and high ($10 \text{ m}^3/\text{h}$) ventilation rates. Mathematical expressions were then derived to calculate the dose after ventilation, considering variables, such as ventilation time, house condition, ventilation rate, and the amount of stone surface area used. The formulae developed here are useful for adjusting ventilation rates or the amount of radon/thoron-emitting surface area in various housing conditions to keep the total annual effective dose, including the effect of gamma radiation, below the recommended limit of 1 mSv [23].

$$\text{Dose after Ventilation (mSv)} = T \times \left(\frac{S \times C_{\text{Rn220}} \times \text{Eq}_{\text{Rn220}} \times \text{DCF}_{\text{Rn220}}}{Q \times V_{\text{eff}}} + \frac{S \times C_{\text{Rn222}} \times \text{Eq}_{\text{Rn222}} \times \text{DCF}_{\text{Rn222}}}{Q \times V_{\text{eff}}} \right) \quad (5)$$

In Equation (5), Q is the ventilation rate per hour, T is the system running time (assuming 8 h per day), Eq_{Rn220} and Eq_{Rn222} are Equilibrium Factors = 0.4 for radon and 0.01 thoron.

Also, dose conversion factors (DCFs) for radon are equal to 9×10^{-6} and 40×10^{-6} mSv per Bqm^{-3} per h, respectively.

The theoretical framework for the model integrates radon/thoron exhalation dynamics with gamma dose calculations, assuming steady-state conditions for gas diffusion (Fick's law) and equilibrium between parent isotopes (U-238, Th-232) and their decay

products. Key limitations include the exclusion of seasonal humidity effects on material exhalation rates and the assumption of homogeneous material composition, which may oversimplify real-world heterogeneity. Ventilation rates were modeled as constant, though actual airflow varies diurnally. Additionally, gamma dose estimates rely on mass activity concentrations without granular spatial resolution. Recent work by Kumar et al. (2014) [24] highlights similar challenges in reconciling lab-measured exhalation rates with field observations. This emphasizes the need for probabilistic modeling in future studies to account for environmental variability.

2.4.5. Elemental Composition and Mineral Mapping

Elemental composition and mineral mapping were conducted using a Tescan TIGER MIRA3 Field Emission Gun Scanning Electron Microscope (FEG-SEM) (Tescan, Brno, Czech Republic) equipped with backscattered electron (BSE) imaging and energy-dispersive X-ray spectroscopy (EDS). Samples were prepared by mounting polished thin sections in epoxy resin and coating them with a conductive layer of carbon to optimize imaging and analysis.

BSE imaging was used to investigate the microstructure and mineralogical matrix, with emphasis on identifying accessory minerals associated with uranium and thorium, such as uraninite (UO_2), thorianite (ThO_2), monazite $[(\text{Ce}, \text{La}, \text{Nd}, \text{Th})\text{PO}_4]$, and zircon (ZrSiO_4). Energy-dispersive X-ray spectroscopy was utilized to generate high-resolution elemental maps, providing spatial distributions of uranium and thorium across the sample surfaces. Regions of interest were identified based on preliminary gamma spectrometry and radon/thoron exhalation results. Analytical parameters, including accelerating voltage (15–20 kV), beam current, and working distance, were optimized to ensure the precise quantification of elemental concentrations and the accurate identification of mineral phases. This methodology provided detailed spatial and compositional data on uranium- and thorium-bearing minerals, facilitating the understanding of their contribution to radon and thoron emissions from the tested building materials.

3. Results and Discussion

3.1. Radon/Thoron Exhalation Rates vs. Radionuclide Activity

For each sample, radon and thoron exhalation rates were calculated based on the equilibrium thoron activity and radon growth model as shown in Figure 3. Supplementary Table S2 shows the detailed measurements of gamma radiation and radon/thoron exhalation rates for each sample.

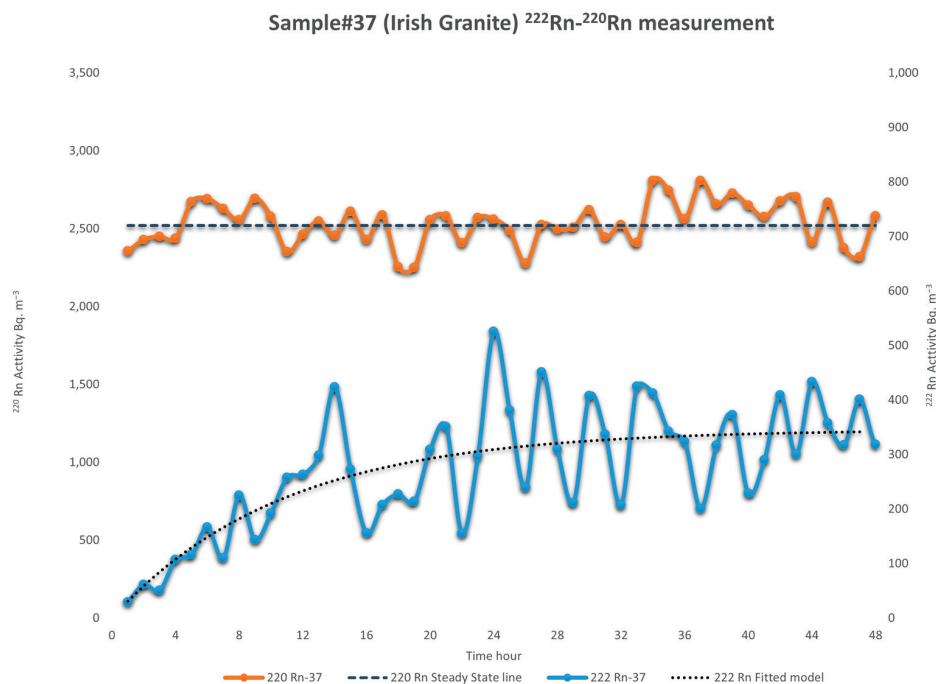


Figure 3. Radon and thoron growth models for an Irish granite (sample #37).

Table 2 shows the statistics of the radionuclide activity of the samples grouped based on their geology. Along with Figure 4, the table provides a comprehensive comparison of radionuclide activity and radon/thoron exhalation rates across various building materials, with an emphasis on the range and mean values to highlight differences in radiological properties. Granite displays the widest range in thorium (Th) concentrations (2.09 to 7.94 mg/kg) and the highest variability in thoron (Rn-220) exhalation rates, ranging from 10 to 45,889 Bq m⁻² h⁻¹, with a mean of 6396 Bq m⁻² h⁻¹. This significant variation is largely due to the differing mineralogical composition of granite, which can include minerals, like thorite and monazite, that are rich in thorium and contribute to higher thoron emissions.

Some of the granites with high thoron exhalation rates were found to be highly altered (e.g., Donegal and Galway areas). Higher modal % of Th-bearing accessory minerals and the secondary porosity and fractured nature of altered granite could facilitate the escape of thoron gas, enhancing its potential to contribute to indoor radiation exposure [25]. Quartzite also exhibits notable radiological properties, with thoron exhalation rates ranging from 103 to 6091 Bq m⁻² h⁻¹ and a mean of 1637 Bq m⁻² h⁻¹. The variability in thoron exhalation from quartzite can be attributed to the presence of thorium-bearing minerals, such as monazite and zircon, which contribute to higher emissions. This suggests that quartzite, like granite, can potentially contribute to indoor radiation levels, particularly when used extensively in building construction [17,26].

In contrast, materials, like gypsum and rock-types, such as limestone, show much lower and more consistent thoron exhalation rates. For instance, gypsum has a uniform thoron exhalation rate of 21 Bq m⁻² h⁻¹, with a mean thorium concentration of 2.31 mg/kg, indicating a lower radiological risk. The consistency in the thoron exhalation rates of gypsum can be linked to its homogeneous mineral composition. Limestone exhibits a broader range of thoron exhalation (15 to 15,213 Bq m⁻² h⁻¹), but with a much lower mean of 1352 Bq m⁻² h⁻¹, which is still significantly lower than that of granite. This is likely due to the predominance of CaCO₃ with minimal thorium content, leading to reduced thoron emissions [2].

The Pearson correlation coefficients in Table 3 provide important insights into the relationships between various radon and thoron exhalation rates, surface emissions, and radionuclide activities. The strong positive correlations between thoron exhalation rates (E220Rn) and its surface emission ($r = 0.96$), as well as between radon exhalation rates (E222Rn) and its surface emission ($r = 0.89$), underscore a direct and robust link between the amount of these gases emitted and their measurable surface presence. This indicates that increases in the exhalation rates of thoron and radon correspond to proportional increases in their respective surface emissions.

In contrast, the moderate correlation between E220Rn and E222Rn ($r = 0.60$) suggests a more nuanced relationship. While there is some association between thoron and radon exhalation rates, the correlation is less pronounced compared to the correlation between each gas's exhalation rate and its surface emission. This implies that factors influencing the emission of radon may not directly impact thoron, and vice versa. Similarly, the moderate to strong correlation between Average Rn-222 Surface Emission and Average Rn-220 Surface Emission ($r = 0.71$) reflects a partial association between the emissions of these gases, though this relationship is weaker compared to their exhalation rates.

Moreover, the exceptionally strong correlation between the I Index and the Dose Rate ($r = 0.99$) underscores a significant relationship between this index—likely representing a composite measure of radon and thoron activities—and the radiation dose rate. This indicates that fluctuations in the I Index are closely tied to variations in the detected radiation dose. In contrast, the weaker correlations between radon and thoron exhalation rates with uranium (U mg/kg) and thorium (Th mg/kg) activities (generally below 0.5) suggest a less direct relationship when considering each element individually. This weaker association implies that other factors might play a more substantial role in influencing surface emissions and exhalation rates.

Table 2. Thoron exhalation rates (Bq/m²/h) and mineral properties of Irish rock samples.

Radionuclide Activity (Min, Max, and Average)									
Rock Type	#	U mg/kg	Th mg/kg	K %	I Index	Exhalation Rates of ²²⁰ Rn (E _{220Rn} , Bq m ⁻² h ⁻¹)		Exhalation Rates of ²²² Rn (E _{222Rn} , Bq m ⁻² h ⁻¹)	
Breccia	1	1.64	2.78	0.63	0.19	71	0.170	0.008	0.031
Brick	2	1.76	3.11	0.75	0.21	514	0.031	0.001	0.019
Concrete	3	1.76	3.13	0.75	0.21	106	0.007	0.001	0.005
Granite	25	1.34	7.94	0.56	0.40	10	3.534	0.070	0.514
Gypsum	6	1.69	2.37	0.66	0.17	21	0.034	0.034	0.034
Limestone	14	1.48	2.95	0.68	0.19	15	0.983	0.018	0.210
Phosphate	1	1.78	2.20	0.67	0.18	9	0.048		
Porcelain	3	1.45	2.92	0.65	0.23	73	0.098	0.010	0.064
Tile						1394			
Quartzite	8	1.60	2.78	0.60	0.22	103	0.232	0.030	0.088
Rhyolite	1	1.83	2.34	0.62	0.18	2207	0.065		
Sandstone	11	1.41	2.92	0.55	0.21	6	0.137	0.030	0.071
Volcanic	2	1.58	2.51	0.67	0.18	15	0.123	0.105	0.114
Greywacke	1	1.38	2.57	0.66	0.17	115	0.021		

Table 3. Similarity test results between radon/thoron exhalation rates of materials with mineralogical drivers (U, Th, and K).

Pearson Correlation Coefficient):	E _{220Rn}	E _{222Rn}	Average ²²⁰ Rn Surface Emission		I index	U mg/kg	Th mg/kg	K %	Dose Rate nSv/hr
			Average ²²² Rn Surface Emission	Average ²²⁰ Rn Surface Emission					
E220Rn	1.00	0.60	0.73	0.96	0.46	0.35	0.35	0.49	0.49
E222Rn	0.60	1.00	0.89	0.61	0.67	0.74	0.50	0.50	0.69
Average Rn-222 Surface Emission	0.73	0.89	1.00	0.71	0.77	0.71	0.64	0.64	0.79
Average Rn-220 Surface Emission	0.96	0.61	0.71	1.00	0.48	0.36	0.36	0.51	0.50
I index	0.46	0.67	0.77	0.48	1.00	0.78	0.89	0.88	0.99
U mg/kg	0.35	0.74	0.71	0.36	0.78	1.00	0.50	0.57	0.78
Th mg/kg	0.35	0.50	0.64	0.36	0.89	0.50	1.00	0.68	0.86
K %	0.49	0.50	0.64	0.51	0.88	0.57	0.68	1.00	0.88
Dose rate nSv/h	0.49	0.69	0.79	0.50	0.99	0.78	0.86	0.88	1.00

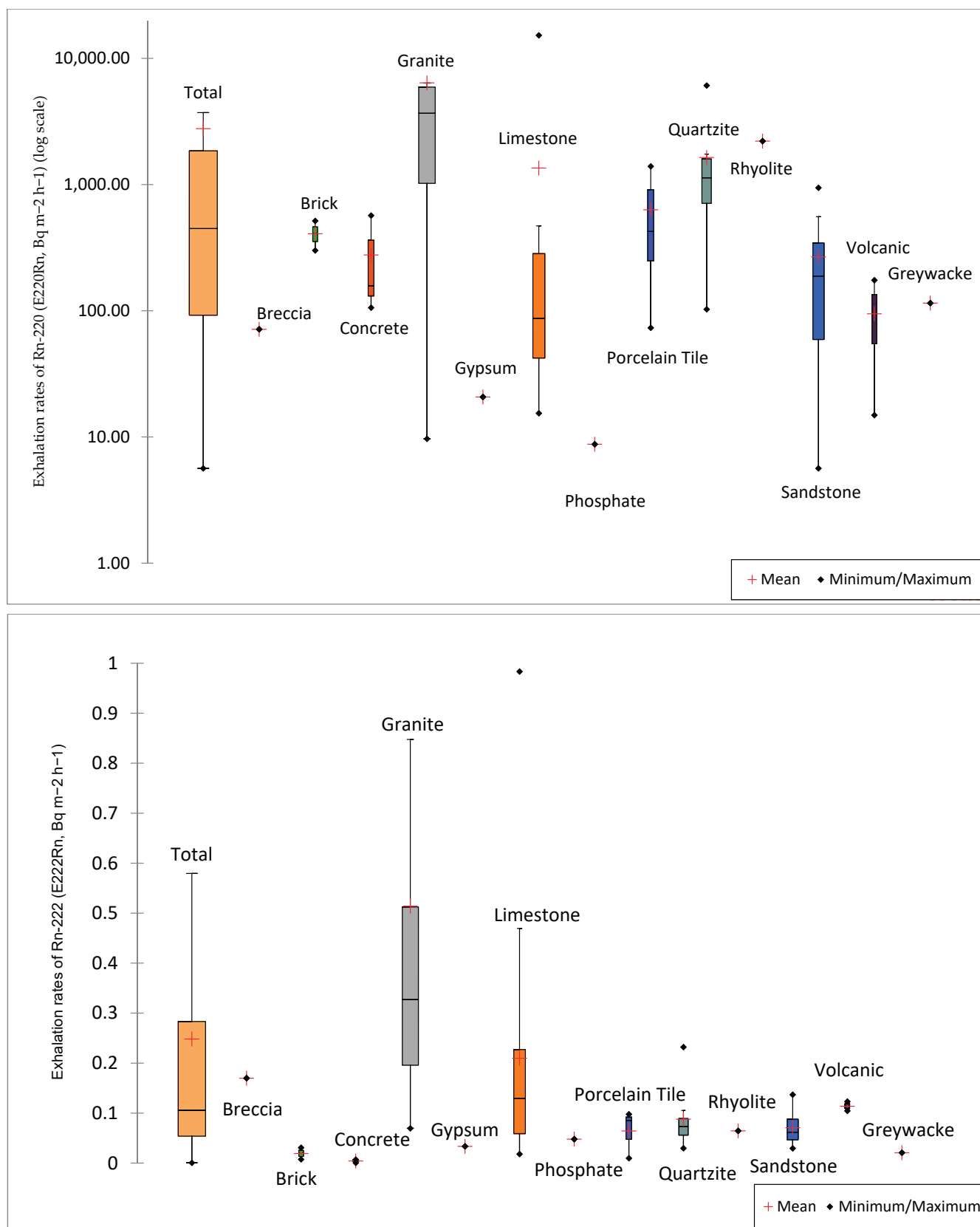


Figure 4. Box plot showing the statistics of radon (**bottom**) and thoron (**top**) exhalation rates for different rock types and building materials.

Furthermore, the correlations between radon and thoron exhalation rates with uranium and thorium suggest some degree of interference between these elements (i.e., Th is a

significant input in radon prediction, and similarly, U is significant when predicting thoron exhalation rates. This interference indicates that interactions among these radionuclides could be affecting the observed exhalation rates. To better understand and model these interactions, advanced modeling techniques could be applied. By incorporating the effects of thorium and uranium interference, it may be possible to enhance the accuracy of predictions for radon and thoron exhalation rates. Such advanced models [27] could account for the combined influences of multiple radionuclides, leading to more refined assessments of radiation levels and improved monitoring methods.

3.2. Implications for Radiometric Surveys and Radionuclide Scanning

The observed correlations have significant implications for radiation monitoring and assessment technologies. Airborne or drone-borne vehicles equipped with gamma detectors [28–30] are effective for covering large areas, such as quarries or mines. Given the correlations between radon and thoron exhalation rates or surface emissions with radioelement activities, drones can be utilized to map radiation levels and identify high-activity zones efficiently. This capability is crucial for environmental monitoring and safety in areas with potential radon and thoron exposure.

Handheld gamma detectors are well-suited for localized monitoring at distribution centers, markets, and residential or workplace settings. The strong correlation between the I Index and the dose rate indicates that handheld devices can provide accurate measurements of radiation levels corresponding to radon and thoron activities. Regular use of these devices can help pinpoint high-radiation areas and ensure that safety measures are in place to mitigate potential health risks associated with elevated radiation exposure. Potential drawbacks of using handheld gamma detectors include potential instrument calibration issues, and the high sensitivity of some instruments, which necessitates adequate lead shielding to prevent overestimating gamma dose rates and uranium, thorium, and potassium concentrations.

Cross-Validation of Data from Handheld Gamma Detector (GT40) and Drone-Based Surveys with D230A Detector

As part of a supplementary research project [30], drone-based surveys were conducted to map gamma radiation variations across four quarries in County Donegal, Ireland: one sandstone, two quartzite, and one granite quarry. A D230A detector mounted on a drone was utilized for these surveys. Before the drone surveys, a GT40 handheld gamma detector (in assay measurement mode) was employed to collect dose rate, uranium, thorium, and potassium activity data at specific locations.

Table 4 presents the results of a cross-validation analysis of radioelement data (see Supplementary Table S3) collected using both handheld and drone-borne surveys, employing the Pearson correlation test ($p < 0.05$). Overall, there is a good agreement between the datasets collected using both methods. Dose rates exhibit a strong correlation ($R^2 = 0.73$), while the highest agreement ($R^2 = 0.79$) was observed for thorium values. This enhanced agreement for thorium is likely due to its higher activity at the source, minimizing the impact of the background radiation effect. It is noteworthy that regression methods were applied to the drone-based gamma data to interpolate values between flight lines and generate smooth surface maps. Therefore, it is expected that the interpolation of data may have slightly affected (decreased) the correlation coefficients between the data.

3.3. Converted Doses

Supplementary Figure S1 provides insights into various stone materials sourced from different regions, focusing on their exposure to gamma radiation and radon/thoron exhalation. These factors contribute significantly to the total effective radiological dose.

The dataset includes a wide range of stone types from Ireland, other EU countries, and non-EU sources.

Table 4. Cross-validation of dose rate (DR) uranium (U), thorium (Th), and potassium (K) collected using a GT40 hand-held gamma detector with those measured during the drone-borne surveys using a D230A gamma detector.

D230A	GT-40			
	Dose Rate (DR) nSv/h	U [mg/kg]	Th [mg/kg]	K [%]
DR	0.73			
U		0.53		
Th			0.79	
K				0.50

The total annual effective dose and the I index, which indicates radiation hazard levels, show considerable variation among these materials. Notably, in the worst-case scenario, the highest total annual effective dose recorded for Irish-origin materials comes from granite sourced in the Galway region, with a dose of 3.90 mSv and an I index of 0.29 mSv, highlighting its significant radiological importance.

The accompanying figure illustrates the contributions of gamma radiation and radon/thoron exhalation rates for samples that had the highest dose contributions.

In the tested samples, which originate from Ireland, gamma radiation contributes between 0.17 mSv and 0.40 mSv to the total effective dose, accounting for about 7% to 33% of the overall exposure. In contrast, radon and thoron exhalation play a much more substantial role, contributing between 0.01 mSv and 3.61 mSv—approximately 7% to 93% of the total effective dose (Figure 5). These data underscore that for the measured samples, radon and thoron are the primary sources of radiation exposure in Irish samples, while gamma radiation acts as a secondary contributor.

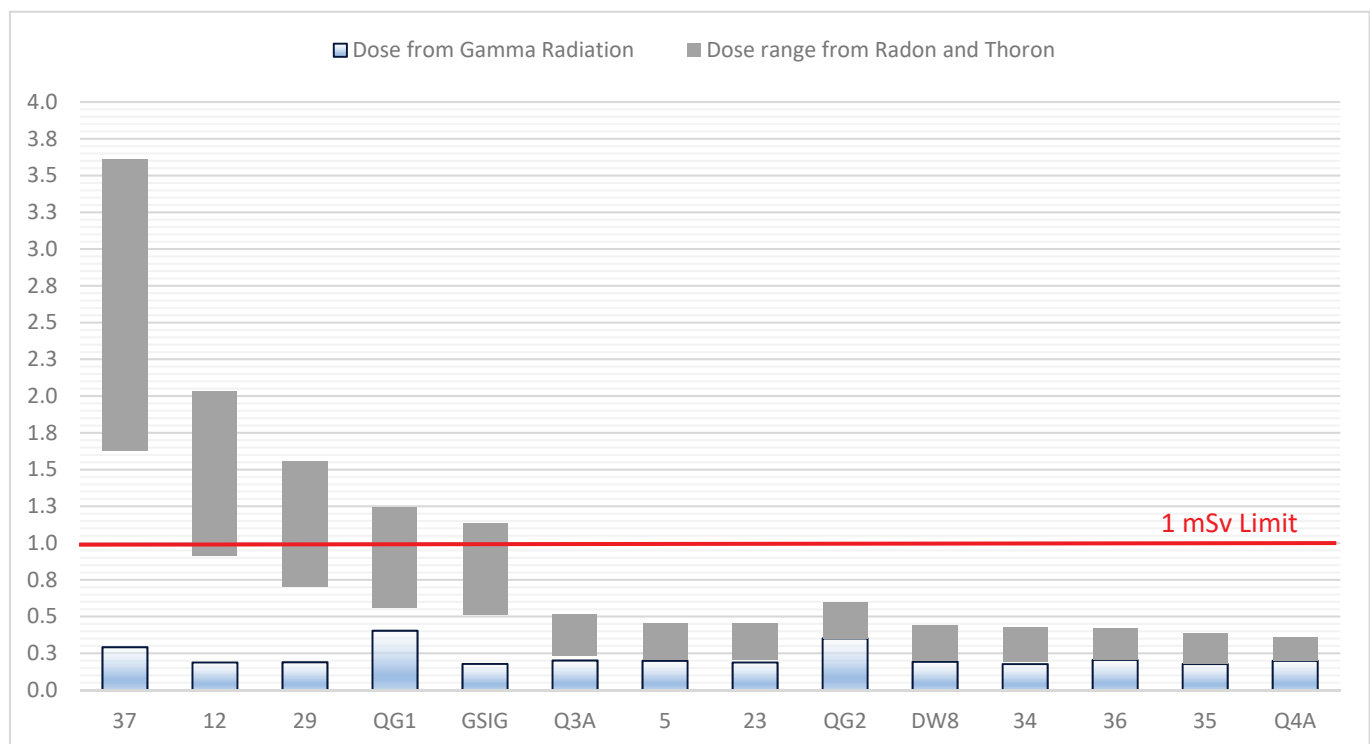


Figure 5. Contributions to the total effective dose (dose from gamma radiation plus dose from radon and thoron exhalations).

For the samples tested, which originated within the EU (but excluding Ireland), gamma radiation contribution ranges from 0.18 mSv to 0.20 mSv, making up about 15% to 32% of the total dose. The total annual effective dose for these samples varies from 0.18 mSv to 0.68 mSv, with radon and thoron doses ranging from 0.00 mSv to 0.46 mSv, contributing roughly 0% to 85% of the total dose. This re-emphasizes that while gamma radiation constitutes a modest fraction of the total dose in EU samples, radon and thoron remain the more significant contributors.

For the tested samples, which originated from outside the EU, gamma radiation contributes between 0.16 mSv and 0.19 mSv, accounting for about 9% to 44% of the total dose. The total annual effective dose ranges from 0.17 mSv to 2.22 mSv, with radon and thoron contributions varying from 0.00 mSv to 2.03 mSv—representing approximately 56% to 91% of the total dose. This variation highlights the diverse impact of gamma radiation across non-EU samples, where radon and thoron generally emerge as the predominant sources of overall radiation exposure.

3.4. Effect of Ventilation

Table 5 presents a worst-case scenario model for radon and thoron dose reduction in a one-bedroom house, focusing on samples with radon and thoron dose contributions higher than 0.3 mSv. The analysis considers two configurations: 10% and 30% of the surface area covered by stone, and the impact of varying air ventilation rates. At a ventilation rate of 0.5 m³/h, with 10% surface coverage, doses range from approximately 0.041 to 1.327 mSv. Increasing the ventilation to 2 m³/h reduces these doses to a range of 0.079 to 0.662 mSv, representing a reduction of about 50–80%. When the surface coverage is increased to 30%, the doses are generally higher due to the greater exposure, but the relative reduction remains similar. At the highest ventilation rate of 10 m³/h, the doses decrease to a range of 0.005 to 0.126 mSv, corresponding to a reduction of approximately 90–95%, regardless of surface coverage.

Table 5. Details of dose ranges, surface activity, and air ventilation rates for the investigated samples.

Dose mSv. (Worst Case Scenario with No Ventilation)		Surface Activity Bq m ^{−2}				Air Ventilation Rate					
						Low		Medium		High	
Sample	I index	Radon	Thoron	Thoron	Radon	0.5 m ³ /h		2 m ³ /h		10 m ³ /h	
37	0.291	1.627	3.612	782	95.28	1.327	2.528	0.332	0.632	0.066	0.126
12	0.188	0.916	2.034	440	45.09	0.685	1.305	0.171	0.326	0.034	0.065
29	0.189	0.701	1.556	337	23.81	0.446	0.850	0.112	0.212	0.022	0.042
QG1	0.404	0.562	1.248	270	129.42	1.162	2.213	0.290	0.553	0.058	0.111
GSIG	0.178	0.511	1.134	97	46.59	0.418	0.797	0.105	0.199	0.021	0.040
Q3A	0.201	0.234	0.519	112	6.42	0.138	0.262	0.034	0.066	0.007	0.013
5	0.199	0.205	0.455	99	11.67	0.165	0.314	0.041	0.078	0.008	0.016
23	0.188	0.205	0.205	98	12.79	0.173	0.329	0.043	0.082	0.009	0.016
QG2	0.351	0.203	0.450	96	3.50	0.104	0.197	0.026	0.049	0.005	0.010
DW8	0.192	0.200	0.445	92	10.74	0.153	0.291	0.038	0.073	0.008	0.015
34	0.177	0.192	0.426	84	13.39	0.165	0.315	0.041	0.079	0.008	0.016
36	0.205	0.174	0.387	81	10.89	0.145	0.276	0.036	0.069	0.007	0.014
35	0.177	0.168	0.374	61	18.45	0.184	0.350	0.046	0.088	0.009	0.018
Q4A	0.199	0.132	0.294	64	13.43	0.149	0.284	0.037	0.071	0.007	0.014

Notably, the reduction in thoron dose is more significant than the reduction in radon dose due to ventilation. This can be attributed to the shorter half-life of thoron (approximately 55.6 s) compared to radon (3.8 days). Thoron decays more rapidly, meaning that effective ventilation can remove thoron gas from indoor environments before it has a chance to decay and contribute significantly to radiation dose. As a result, higher ventilation rates are particularly effective in reducing thoron levels, leading to the observed substantial

reductions in thoron dose across different ventilation scenarios. Based on the charts of Figure 6, a formula was developed to calculate the reduced dose from radon and thoron because of active air ventilation.

$$\text{Reduced Dose after Ventilation} = 0.1192Q^{-1.001} \times e^{(0.7805 \cdot \text{Dose Initial})} \quad (6)$$

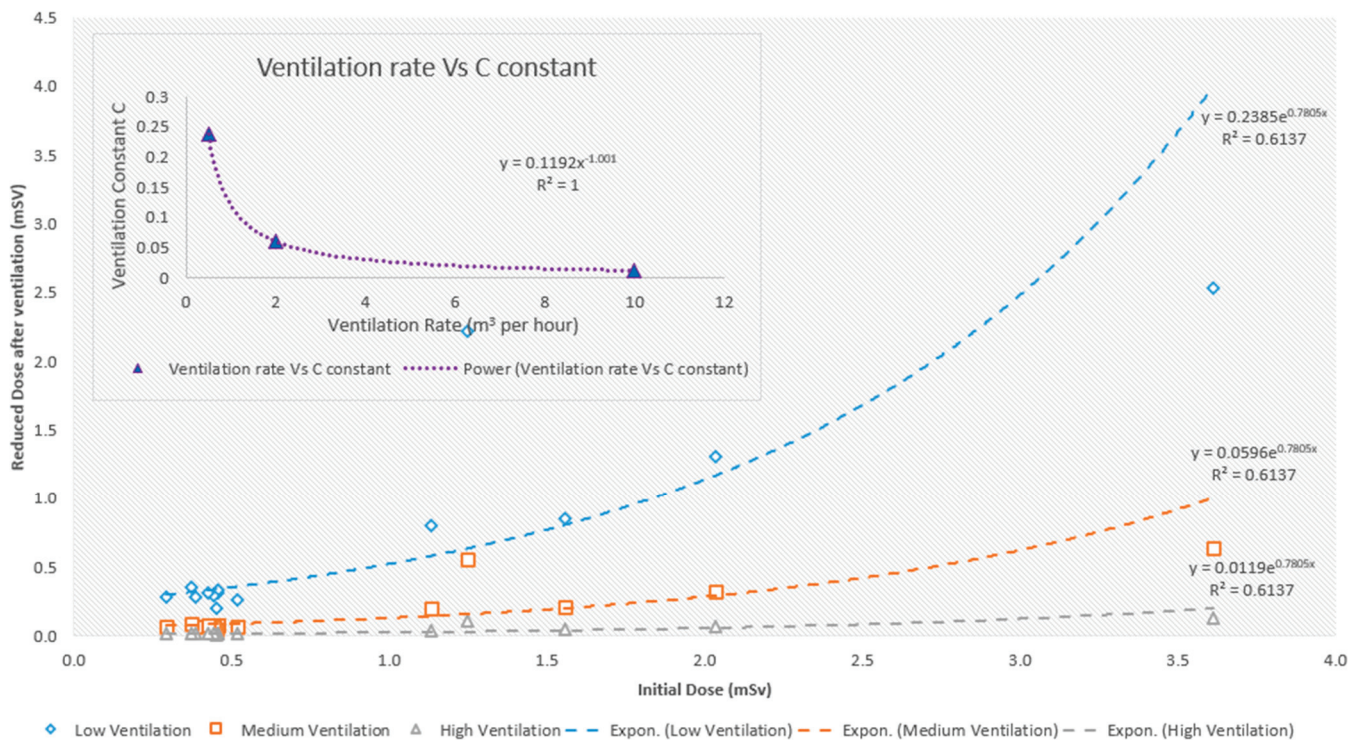


Figure 6. Combined doses calculation models of radon and thoron exhalation from Irish building materials under standard (0.5 h^{-1}) and high ventilation ($10 \text{ m}^3/\text{h}$) scenarios. High ventilation reduces combined doses by 85–95%, demonstrating its critical role in mitigating risks.

To mitigate indoor radiation exposure, practitioners first measure the annual radiation dose from radon, thoron, and gamma sources using detectors, such as the RAD7 (DurrIDGE, Billerica, MA, USA). In Ireland, where the average annual radiation dose is 4.2 mSv , approximately 60% (2.5 mSv/year) comes from radon and thoron. If 0.4 mSv/year is also attributed to gamma radiation (total = 2.9 mSv/year), to meet the recommended safety threshold of 1 mSv/year , the radon/thoron contribution must be reduced to 0.6 mSv/year (because $1 \text{ mSv/year total} - 0.4 \text{ mSv/year gamma} = 0.6 \text{ mSv/year}$). Using an empirical model derived from experimental data, the required ventilation rate (Q) is calculated to be $1.4 \text{ m}^3/\text{h}$. This airflow can be achieved with a small exhaust fan or mechanical ventilation system. Because of the extremely short half-life of thoron (55.6 s), even modest ventilation drastically reduces its dose contribution. However, the longer half-life of radon (3.8 days) means additional measures—such as sealing foundation cracks, installing radon barriers, or substituting high-exhalation building materials (e.g., granite) with low-emission alternatives (e.g., limestone)—are often necessary for long-term mitigation. Post-installation validation using RAD7 detectors confirms compliance with the 1 mSv/year threshold and accounts for real-world variables, like uneven airflow or material heterogeneity.

3.5. Effect of Mineralogy

While a detailed investigation into the effect of mineralogy on high radon or thoron exhalation rates is of interest, it falls outside the main scope of this paper. However, to illustrate how mineral composition can influence radon/thoron potential, a high-resolution

mineral scan of a Galway granite sample (shown here to have high radon and thoron potential) was evaluated. Prior to analysis, minerals of interest were manually introduced to the microscope. Some minerals commonly associated with radon and thoron emission (Figure 7) include those containing uranium and thorium, such as the following:

- Uraninite (UO_2): A primary uranium ore with high radioactivity (gamma radiation source) and radon potential (alpha radiation source).
- Thorianite (ThO_2): A thorium-rich mineral with high radioactivity (gamma radiation source) that contributes to thoron emissions (alpha radiation source).
- Monazite ($\text{Ce,La,Nd,Th}\text{PO}_4$): Contains significant thorium (gamma radiation source) and can be a source of thoron and, to a lesser extent, radon (alpha radiation sources).
- Zircon (ZrSiO_4): Often contains trace uranium and, to a lesser extent thorium (gamma radiation sources), contributing to radon and thoron (alpha radiation sources).

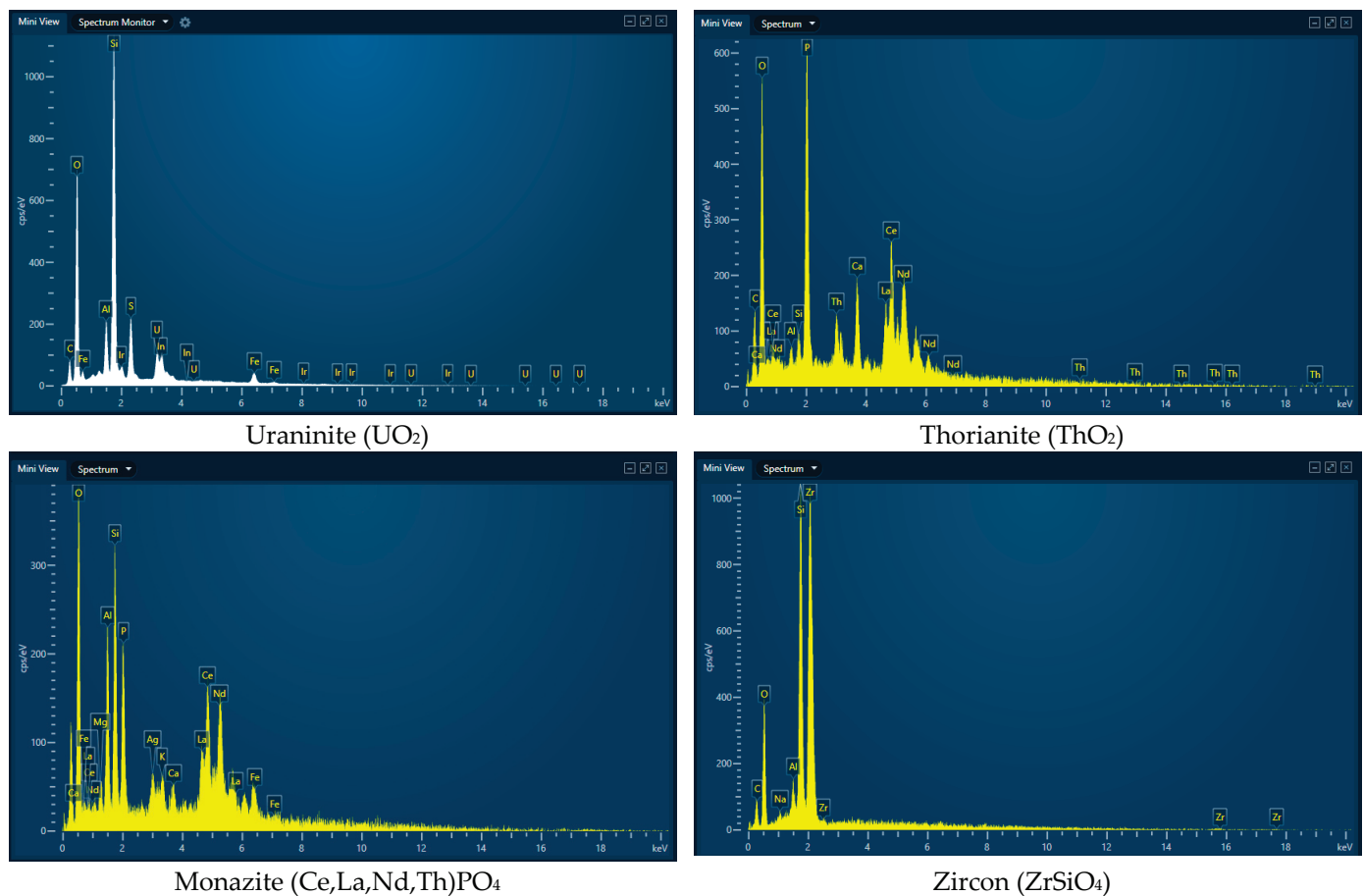


Figure 7. Spectra of the most common uranium–thorium bearing minerals used for capturing the minerals which can produce high radon or thoron emanation rates.

Figure 8 presents a composite of backscattered electron (BSE) images of the Galway granite sample and the corresponding mineral maps. BSE imaging reveals the mineralogical matrix and texture, while the elemental maps highlight the spatial distribution of uranium- and thorium-bearing minerals within the sample. The granite sample scan revealed a considerable number of rutile, zircon, monazite, and hematite minerals. These kinds of detailed imaging and mapping are beneficial for future research aimed at identifying the specific minerals contributing to elevated radon and thoron levels and the degree of their contribution. By analyzing these mineralogical features and the chemical composition of the samples, it is anticipated that a better understanding of how the distribution and concentration of uranium and thorium-bearing minerals influence overall radon and thoron

emissions can be achieved. This approach provides mineralogical insights, which enable a refined understanding of radiation risks associated with various building materials. The approach demonstrated here highlights a novel method for material testing and certification for radon and thoron potential studies, ultimately facilitating the development of more effective radiation protection resource management strategies for building materials.

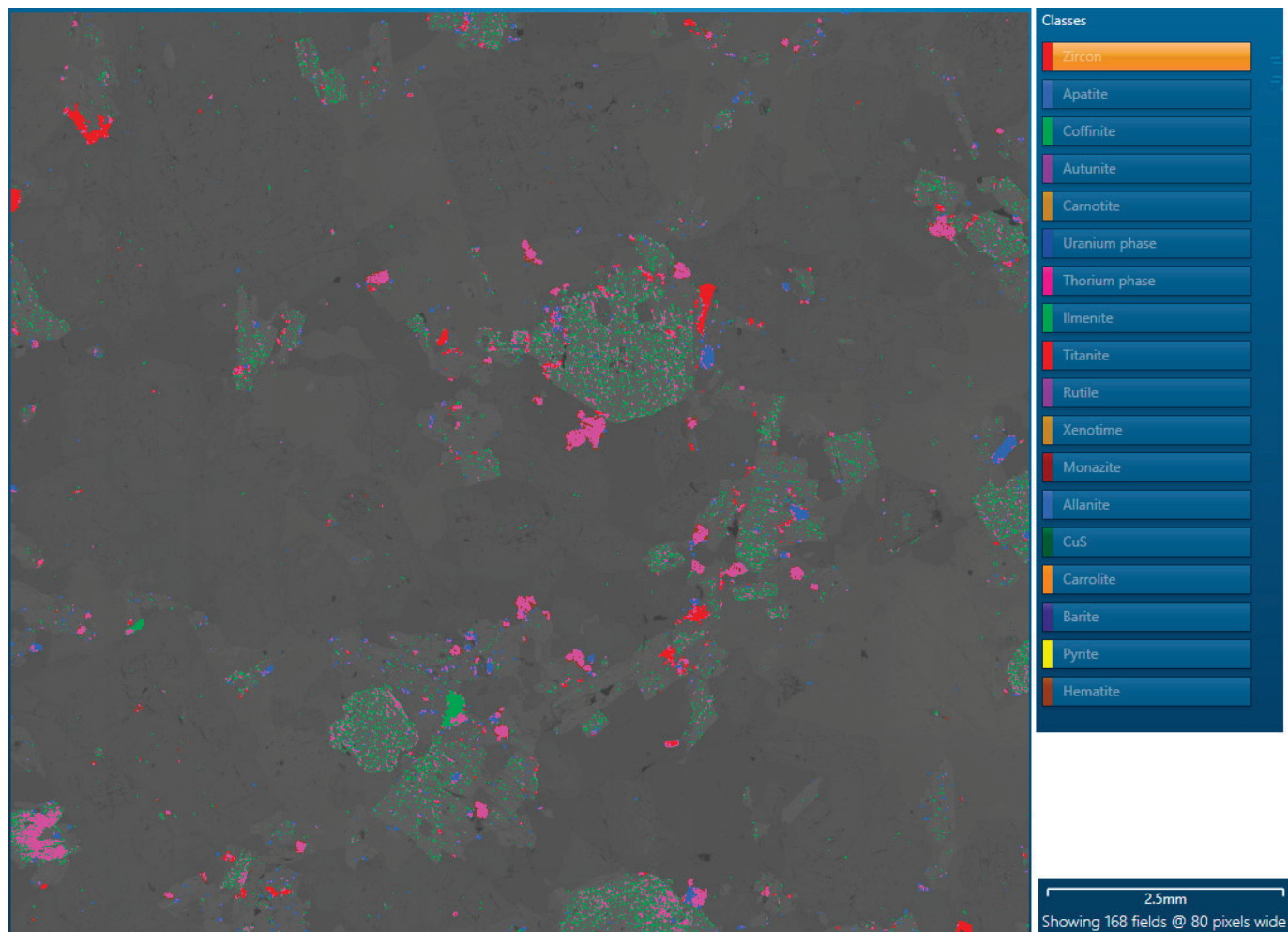


Figure 8. Elemental composition and mineral mapping with the Tescan TIGER MIRA3 FEG-SEM.

4. Conclusions

This study presents a comprehensive assessment of radon and thoron exhalation rates, alongside associated gamma radiation exposure, from various building materials commonly utilized in Ireland, and widely available in countries around the world. Results indicate substantial variability in the radiological properties of these materials, with granite and quartzite exhibiting notably high thoron exhalation rates due to their mineralogical composition. These findings highlight the potential for certain natural rock materials to contribute significantly to indoor radiation exposure, particularly in poorly ventilated environments. Building on our measurements and modeling, the following comparative contextualization places our key radiological metrics—radon ($E_{(222)}$), thoron ($E_{(220)}$) exhalation rates and gamma dose contributions—alongside published values for the same natural building materials:

Granite: In this study, the mean thoron exhalation rate of tested granites was $6396 \pm 3534 \text{ Bq m}^{-2} \text{ h}^{-1}$, considerably higher than the $3500 \pm 800 \text{ Bq m}^{-2} \text{ h}^{-1}$ reported by Frutos-Puerto et al. (2020) [2]. Its radon exhalation ($0.51 \pm 0.46 \text{ Bq m}^{-2} \text{ h}^{-1}$) also exceeds

the literature value of $0.20 \pm 0.05 \text{ Bq m}^{-2} \text{ h}^{-1}$. The annual gamma dose from granite in our work ($0.17\text{--}0.40 \text{ mSv y}^{-1}$) aligns with, though slightly under the upper bound of, the $0.20\text{--}0.50 \text{ mSv y}^{-1}$ range found by Sas et al. (2017) [31].

Quartzite: Our measured mean thoron flux for quartzite ($1637 \pm 277 \text{ Bq m}^{-2} \text{ h}^{-1}$) is modestly above the $1200 \pm 500 \text{ Bq m}^{-2} \text{ h}^{-1}$ documented by Lucchetti et al. (2019) [32], while its radon exhalation ($0.75 \pm 0.36 \text{ Bq m}^{-2} \text{ h}^{-1}$) similarly surpasses their $0.50 \pm 0.20 \text{ Bq m}^{-2} \text{ h}^{-1}$ value. The gamma dose ($0.18\text{--}0.20 \text{ mSv y}^{-1}$) matches closely with the 0.19 mSv y^{-1} they reported.

Limestone: We observed thoron exhalation in tested limestones at $1352 \pm 1213 \text{ Bq m}^{-2} \text{ h}^{-1}$, relative to the $1000 \pm 400 \text{ Bq m}^{-2} \text{ h}^{-1}$ from Nunes et al. (2023) [33]. Its radon flux ($0.35 \pm 0.15 \text{ Bq m}^{-2} \text{ h}^{-1}$) is likewise above the $0.25 \pm 0.10 \text{ Bq m}^{-2} \text{ h}^{-1}$ figure in that study. The resulting gamma dose ($0.16\text{--}0.19 \text{ mSv y}^{-1}$) is in good agreement with the 0.15 mSv y^{-1} noted by Lee et al. (2004) [34].

Correlation analyses suggest a potential link between the factors controlling uranium and thorium distribution. Some accessory minerals (e.g., zircon) contain both uranium and thorium, potentially influencing this correlation. Conversely, other accessory minerals (e.g., monazite) contain thorium but not uranium, leading to variability in the radon–thoron correlation across samples with contrasting accessory mineralogy. This variability is further influenced by the specific chemistry of accessory minerals, such as the range of thorium concentrations observed in zircon (from negligible to relatively high). These observations indicate a degree of interaction between these factors, emphasizing the need for advanced models that consider both uranium and thorium concentrations, as well as mineralogical composition when evaluating radiological risks associated with building materials. Furthermore, the gamma dose index (I-index) suggests that materials deemed safe based solely on gamma radiation may still pose health risks due to the combined effects of radon, thoron, and gamma radiation if not adequately managed.

The ventilation models proposed in this study offer insights to develop practical strategies for mitigating radon and thoron exposure in indoor settings. Adjusting ventilation rates or limiting the use of materials with high emission rates can help to maintain radiation doses below recommended safety thresholds. However, further research is needed to refine these models, particularly by incorporating additional sources of radon and thoron, such as soil and water, for a more comprehensive assessment of indoor radiation exposure. In conclusion, this research provides valuable data and methodological advancements that can inform future guidelines and regulations concerning building materials. Given the importance of the construction sector in Ireland, and across the EU, particularly with an emphasis on the extensive use of locally sourced building materials, it is important to consider the radiological implications of these materials to ensure the safety and well-being of building occupants. From a policy perspective, these results argue strongly for revising European construction standards. We recommend that the EU Construction Products Regulation be updated to require manufacturers to report radon and thoron exhalation rates alongside radionuclide concentrations, and to adopt a combined radon + thoron + gamma dose index capped at 1 mSv y^{-1} under typical occupancy and ventilation. Building codes should also mandate minimum ventilation rates—such as a 0.5 h^{-1} air-change rate in rooms built with high-emission materials—and encourage the use of lower-emission alternatives wherever possible.

Future research should expand upon these findings by investigating the long-term health effects of low-level radiation exposure and developing more robust models to predict indoor radon and thoron concentrations under diverse environmental conditions. A particular focus should be placed on elucidating the mineralogical factors influencing exhalation rates.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/app15126470/s1>.

Author Contributions: M.M.A.: conceptualization, methodology, investigation, resources, data curation, writing—original draft preparation, and visualization. Q.C.: validation, formal analysis, writing—review and editing, supervision, project administration, and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This study is part of a postdoctoral research project funded by the Irish Research Council and the Geological Survey Ireland. Funding reference code: EPSPD/2022/141. Geochron Ltd. acknowledges financial support from EOSC Futures, administered by GÉANT under contract SER-23-131.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Acknowledgments: The authors acknowledge the assistance of Paul Guyett at Trinity College Dublin, for his assistance in acquiring SEM data.

Conflicts of Interest: Author Quentin Crowley is the Director of Geochron Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Makumbi, T.; Breustedt, B.; Raskob, W. Parameter uncertainty analysis of the equivalent lung dose coefficient for the intake of radon in mines: A review. *J. Environ. Radioact.* **2024**, *276*, 107446. [CrossRef] [PubMed]
2. Frutos-Puerto, S.; Pinilla-Gil, E.A.-O.; Andrade, E.A.-O.; Reis, M.; Madruga, M.J.; Rodríguez, C.M. Radon and thoron exhalation rate, emanation factor and radioactivity risks of building materials of the Iberian Peninsula. *PeerJ* **2020**, *8*, e10331. [CrossRef] [PubMed]
3. United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR). *Sources, Effects and Risks of Ionizing Radiation*; United Nations: New York, NY, USA, 2022. Available online: https://www.unscear.org/unscear/en/publications/2020_2021_1.html (accessed on 24 May 2024).
4. Tokonami, S. Characteristics of Thoron (^{220}Rn) and its progeny in the indoor environment. *Int. J. Environ. Res. Public Health* **2020**, *17*, 8769. [CrossRef] [PubMed]
5. Semwal, P.; Agarwal, T.K.; Joshi, M.; Kumar, A.; Singh, K.; Ramola, R.C. Effective dose estimation of radon, thoron and their progeny concentrations in the environs of Himalayan belt, India. *Int. J. Environ. Sci. Technol.* **2023**, *20*, 4127–4138. [CrossRef]
6. Dentoni, V.; Da Pelo, S.; Aghdam, M.M.; Randaccio, P.; Loi, A.; Careddu, N.; Bernardini, A. Natural radioactivity and radon exhalation rate of Sardinian dimension stones. *Constr. Build. Mater.* **2020**, *247*, 118377. [CrossRef]
7. Health Information and Quality Authority (HIQA). Ionising Radiation National Dose Report. 2024. Available online: <https://www.epa.ie/publications/compliance--enforcement/radiation/ionising-radiation-national-dose-report.php> (accessed on 24 May 2024).
8. World Health Organization (WHO). *Radon and Health*; WHO: Geneva, Switzerland, 2021. Available online: <https://www.who.int/news-room/fact-sheets/detail/radon-and-health> (accessed on 24 May 2024).
9. European Cancer Observatory. *Ireland Is the Country with the Highest Cancer Incidence in the EU*; European Cancer Observatory: Lyon, France, 2020.
10. Katase, A.; Shimo, M. (Eds.) Radon and thoron in the human environment. In Proceedings of the 7th Tohwa University International Symposium, Fukuoka, Japan, 23–25 October 1997. [CrossRef]
11. Mohamed, T.G.; Hassan, M.; El-Midany, A.A.; Ismail, M.A.; El-Samrah, M.G. Inherent radiological hazard and γ -rays shielding properties of black sand minerals. *Egypt. J. Pet.* **2024**, *33*, 1. [CrossRef]
12. Chen, B.; Wang, Q.; Zhuo, W. Assessment of gamma dose rate in dwellings due to decorative stones. *Radiat. Prot. Dosim.* **2015**, *167*, 251–254. [CrossRef]
13. International Commission on Radiological Protection. The 2007 recommendations of the International Commission on Radiological Protection (ICRP Publication 103). *Ann. ICRP* **2007**, *37*, 1–332.

14. Aghdam, M.M.; Dentoni, V.; Da Pelo, S.; Crowley, Q. Detailed geogenic radon potential mapping using geospatial analysis of multiple geo-variables—A case study from a high-risk area in SE Ireland. *Int. J. Environ. Res. Public Health* **2022**, *19*, 15910. [CrossRef]
15. Lee, E.M.; Menezes, G.; Finch, E.C. Natural radioactivity in building materials in the Republic of Ireland. *Health Phys.* **2004**, *86*, 378–383. [CrossRef]
16. DIN ISO/IEC EN 17025:2018; General Requirements for the Competence of Testing and Calibration Laboratories. Beuth Verlag GmbH: Berlin, Germany, 2018.
17. Lucchetti, C.; Tzortzis, M.; Tzortzis, A. Integrating radon and thoron flux data with gamma radiation mapping in radon-prone areas: The case of volcanic outcrops in a highly-urbanized city (Roma, Italy). *J. Environ. Radioact.* **2019**, *202*, 41–50. [CrossRef] [PubMed]
18. Sas, Z.; Doherty, R.; Kovács, T.; Soutsos, M.; Sha, W.; Schroeyers, W. Radiological evaluation of by-products used in construction and alternative applications; Part I. Preparation of a natural radioactivity database. *Constr. Build. Mater.* **2017**, *150*, 227–237. [CrossRef]
19. Reinhardt, N. Gamma-Ray Spectrometry as Auxiliary Information for Soil Mapping and Its Application in Research for Development. Ph.D. Dissertation, University of Bonn, Bonn, Germany, 2019.
20. Marsh, J.W.; Harrison, J.D.; Laurier, D.; Blanchardon, E.; Paquet, F.; Tirmarche, M. Dose conversion factors for radon: Recent developments. *Health Phys.* **2010**, *99*, 511–516. [CrossRef] [PubMed]
21. Allen, D. Dose Conversion Factors for Radon WLM. National Institute for Occupational Safety and Health, Technical Information Bulletin DC-T11-R5; 2021. Available online: <https://www.cdc.gov/niosh/ocas/pdfs/tibs/dc-t11-r5-508.pdf> (accessed on 24 May 2024).
22. Mas, J.L.; Ramírez, J.R.C.; Bermúdez, S.H.; Fernández, C.L. Assessment of natural radioactivity levels and radiation exposure in new building materials in Spain. *Radiat. Prot. Dosim.* **2021**, *194*, 178–185. [CrossRef]
23. Harrison, J.D.; Marsh, J.W. ICRP recommendations on radon. *Ann. ICRP* **2020**, *49*, 68–76. [CrossRef]
24. Kumar, A.; Chauhan, R.P.; Joshi, M.; Sahoo, B.K. Modeling of Indoor Radon Concentration from Radon Exhalation Rates of Building Materials and Validation through Measurements. *J. Environ. Radioact.* **2014**, *127*, 50–55. [CrossRef]
25. Skierszkan, E.K.; Dockrey, J.W.; Mayer, K.U.; Bondici, V.F.; McBeth, J.M.; Beckie, R.D. Geochemical controls on uranium release from neutral-pH rock drainage produced by weathering of granite, gneiss, and schist. *Minerals* **2020**, *10*, 1104. [CrossRef]
26. Nunes, L.J.R.; Curado, A.; Lopes, S.I. The relationship between radon and geology: Sources, transport and indoor accumulation. *Appl. Sci.* **2023**, *13*, 7460. [CrossRef]
27. Joel, E.S.; Omeje, M.; Olawole, O.C.; Adeyemi, G.A.; Akinpelu, A.; Embong, Z.; Saeed, M.A. In-situ assessment of natural terrestrial-radioactivity from uranium-238 (^{238}U), thorium-232 (^{232}Th) and potassium-40 (^{40}K) in coastal urban-environment and its possible health implications. *Sci. Rep.* **2021**, *11*, 17555. [CrossRef]
28. Manfreda, S.; Dor, E.B. Remote sensing of the environment using unmanned aerial systems. In *Unmanned Aerial Systems for Monitoring Soil, Vegetation, and Riverine Environments*; Sabins, P.A., Sabins, M.F., Eds.; Springer: Cham, Switzerland, 2023; pp. 3–36. [CrossRef]
29. Dardac, M.; Elío, J.; Aghdam, M.M.; Banrion, M.; Crowley, Q. Application of airborne geophysical survey data in a logistic regression model to improve the predictive power of geogenic radon maps: A case study in Castleisland, County Kerry, Ireland. *Sci. Total Environ.* **2023**, *894*, 164965. [CrossRef]
30. Aghdam, M.M.; O'Brien, M.K.; Crowley, Q. Utilization of radiometric data for mapping primary and secondary sources of gamma radiation and radon/thoron release potentials in Ireland. *Front. Public Health Radiat. Health* **2024**, *12*, 1443332. [CrossRef] [PubMed]
31. Sas, Z.; Kovács, T.; Somlai, J. External gamma-ray dose rate and radon concentration in indoor environments covered with Brazilian granites. *Radiat. Prot. Dosim.* **2017**, *175*, 75–81. [CrossRef]
32. Lucchetti, C.; Tzortzis, M.; Tzortzis, A. Measurement of natural radioactivity and radon exhalation rates in building materials used in the construction industry. *Radiat. Prot. Dosim.* **2019**, *184*, 497–503. [CrossRef]
33. Nunes, L.; Silva, J.; Ferreira, A. Natural radioactivity and radiological risks of common building materials: A case study in Portugal. *J. Environ. Radioact.* **2023**, *255*, 106961. [CrossRef]
34. Lee, S.Y.; Kim, H.S.; Lee, J.K. Assessment of gamma radiation dose rates and radon concentrations in indoor environments covered with granite materials. *J. Environ. Radioact.* **2004**, *77*, 293–304. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Thermal Damage Characterization and Modeling in Granite Samples Subjected to Heat Treatment by Leveraging Machine Learning and Experimental Data

Gabit Sansyzbekov, Amoussou Coffi Adoko * and Paul Mathews George

School of Mining and Geosciences, Nazarbayev University, Astana 010000, Kazakhstan; gabit.sansyzbekov@nu.edu.kz (G.S.); george.mathews@nu.edu.kz (P.M.G.)

* Correspondence: amoussou.adoko@nu.edu.kz; Tel.: +77-172706456

Featured Application: The predictive model developed in this study can be used to assess the thermal damage of granite for general applications such as geothermal energy extraction, deep mining, and nuclear waste disposal. By leveraging machine learning and experimental data, the model enables engineers and researchers to evaluate the effect of high temperature on rock properties in underground excavations.

Abstract: High temperatures significantly affect the physical and mechanical properties of rocks in deep geoengineering projects, such as geothermal development, deep mining, and the geological disposal of nuclear waste. Therefore, it is essential to explore the relationship between the thermal damage (TD) of granite and its influencing factors. This paper characterizes the TD of granite specimens subjected to high temperatures of up to 800 °C and proposes a predictive model for this thermal damage. A database, which includes publicly available experimental data of advanced microscopic observations of granite specimens exposed to high-temperature treatments and their changes in physical and mechanical properties, was compiled and analyzed. The collected data revealed a consistent trend: crack development among quartz, feldspar, and biotite minerals was observed to intensify notably between 400 °C and 600 °C, as indicated by changes in the mechanical properties. Based on these characteristics, the relationships between TD and its influential parameters were determined using regression models and several machine learning algorithms. The derived models indicated good predictability performance with a coefficient of determination (R^2) varying between 0.60 and 0.97, with the boosted ensemble tree model being the best. Nevertheless, mineral contents were not found to be good predictors of TD, even if they control the evolution of the crack during the heat treatment. It was concluded that the findings of this study could serve as a valuable tool for assessing the thermal damage of rocks.

Keywords: granite; crack development; thermal damage; predictive model; machine learning

1. Introduction

Granite is one of the most common rocks in deep geoengineering applications involving high-temperature treatments such as geothermal energy extraction, deep mining, and nuclear waste geological disposal [1–3]. For example, enhanced geothermal systems reach 200 °C, while the temperature in underground chambers for nuclear disposal storage now varies from 100 °C to 300 °C and will increase over the storage period [4]. The high temperature damages the reservoir rocks and weakens their physical and mechanical

properties, which is a pressing concern in these engineering projects [5–7]. In essence, the weakening in macroscopic physical and mechanical properties results from changes in mineral composition and characteristics at high temperatures [8]. Different minerals expand at different rates at high temperatures, and the same mineral can expand at different rates and in many directions [6,9,10]. Such expansion can cause stress accumulation at the edges of minerals, damaging the structure when heated. Therefore, the heated rock may produce intergranular and intragranular microcracks among and within mineral crystals [3]. In addition to these intergranular microcracks in igneous rock, intragranular cracks in some of the weaker mineral elements, like feldspar and biotite grains, also play a role in crack propagation. High temperatures can also induce thermal reactions in minerals that form rocks, altering their composition and structure, resulting in different types of microcracks. As a result, microcrack formation brought on by heat treatment alters the physico-mechanical behaviors of rocks [11–13], which eventually impacts any rock engineering design subjected to high temperatures. Therefore, extensive research is required to fully comprehend the impact of high temperatures on rocks spanning from microscopic mineral composition to macroscopic physical and mechanical properties, including mineral composition, dehydration, crystalline state, and the formation of new phases.

Motivated by this importance, extensive studies have been conducted over the past few decades, taking advantage of the availability of technologies such as scanning electron microscopy (SEM), optical microscopy, X-ray microcomputed tomography (CT), polarizing microscope (PM), nuclear magnetic resonance imaging (NMR), and acoustic emission (AE) to reveal the mechanisms of physico-mechanical property deterioration [9]. Additionally, in most existing studies, researchers often combine SEM, thermogravimetric analysis (TGA), and differential scanning calorimetry (DSC) with X-ray diffraction (XRD) pattern analysis to examine the microstructural properties of rock minerals [5]. For example, Zhang, Sun [14] investigated the heat influence on the microstructure of limestone using XRD, SEM, TGA, and DSC tests. Their results indicated that thermal treatment alters the crystallization in addition to causing mineral disintegration. Shen, Zhang [15] conducted comparable research on sandstone samples and reached similar conclusions. Many other studies corroborated how high temperatures affect the pore structure, mechanical characteristics, and mineralogy of rocks, including the study of acoustic emission and fracture morphology characteristics of thermally damaged granite [7], analysis of physical and mechanical behaviors and microscopic mineral characteristics of thermally damaged granite [16], experimental and numerical simulation study of the evolution of mechanical properties of granite after thermal treatment [6], and experimental investigation on the physical/thermal properties of Nahan granite [2].

Nevertheless, despite the numerous studies on the thermal damage (TD) of rock, the relationship between the TD of granite and its influencing factors has not been comprehensively quantified yet [5,12]. Hence, this paper aims to characterize the TD of granite specimens exposed to high temperatures and then propose a predictive model for TD. The wide temperature range (from 25 °C to 800 °C) is selected to describe the physical and mechanical changes in granite, along with the crack development study to better explain the effect of high temperatures on rock.

2. Granite Thermal Damage Dataset Description

The data used for this study consist of the physical and mechanical properties and the microscopic observations of granite specimens exposed to high-temperature treatments that were compiled from the experimental results of existing studies [2,4,10,17–29]. The rock samples were cylindrical without pre-existing holes and primarily sourced from various international projects, including those in Australia, India, Germany, the UK, Canada, and

China, with the experiments conducted under nearly identical conditions. The dataset includes parameters such as temperature (P1), mineral contents—specifically feldspar (P2), quartz (P3), biotite (P4), and other minerals (P5)—elastic modulus (P6), porosity (P7), density (P8), uniaxial compressive strength (P9), tensile strength (P10), and P-wave velocity (P11), all obtained during these experiments. The compiled dataset is available in Supplementary Materials, while sample data are presented in Table 1. To study the damage evolution during the testing, SEM observations of the samples were also collected to study the damage evolution during the testing.

As can be seen in Table 1, the original dataset has many missing points; hence, they need to be handled. In this study, the multiple imputation (MI) technique was implemented by using the mice package in the R program. This procedure consists of seven steps: (1) loading required packages to enable data processing and multiple imputation, (2) exploration of missing data to inform the imputation strategy, (3) setting the imputation method, (4) running the imputation using chained equations over 50 iterations, (5) checking convergence to confirm that stable imputations are generated, (6) analysis of imputed data, and (7) exporting completed dataset. Instead of simply discarding the missing values, MI creates several different datasets by imputing the missing values multiple times and proposes the average value [30]. There are three key stages in MIs: data imputation, data analysis, and results pooling [31]. The first stage involves the generation of several complete datasets. Then, each of the imputed datasets is analyzed separately using the same statistical model. Finally, these multiple analyses' results are averaged to produce final estimates. MI has many advantages and is often preferred where missing data are common. The resulting dataset is used for the TD.

Table 1. A sample of the dataset.

Temp, °C	Feldspar Content, %	Quartz Content, %	Biotite Content, %	Other Minerals, %	Elastic Modulus, GPa	Porosity, %	Density g/cm ³	UCS, MPa	Tensile Strength, MPa	P-Wave Velocity, km/s
25	50	40	5	5	NA	0.2	2.63	NA	9.03	4.71
100	75	20	5	0	12.96	NA	2.64	134.1	7.62	4.04
200	72.21	11.89	15.9	0	22.25	NA	NA	149	NA	3.57
300	59.85	11.12	21.56	7.47	11.12	0.68	2.63	102.04	59.85	3.52
400	83.51	9.04	0	7.45	7.36	NA	NA	72.51	4.85	3.03
500	60.59	34.09	5.32	0	20.5	1.43	NA	187	NA	3.09
600	63.59	27.72	4.94	3.75	5.69	NA	2.61	54.88	1.51	NA
700	29	50	15	6	1.82	NA	NA	22.43	8.69	NA
800	48.34	41.78	4.29	5.39	6.96	NA	2.49	38.07	NA	0.49

3. Crack Evolution

The influence of temperature on crack development in granite specimens was revealed through the collected data. Microcracks in the rock are generally classified as intragranular (contained within a single grain), intergranular (along the grain boundaries), and transgranular (affecting more than one grain) [19]. This study examined the intragranular cracks of feldspar, quartz, and biotite minerals at various temperatures and the intergranular cracks at the boundaries between minerals. Samples of the analyzed microscopic images are provided in Figure 1, where red lines indicate the intergranular cracks.

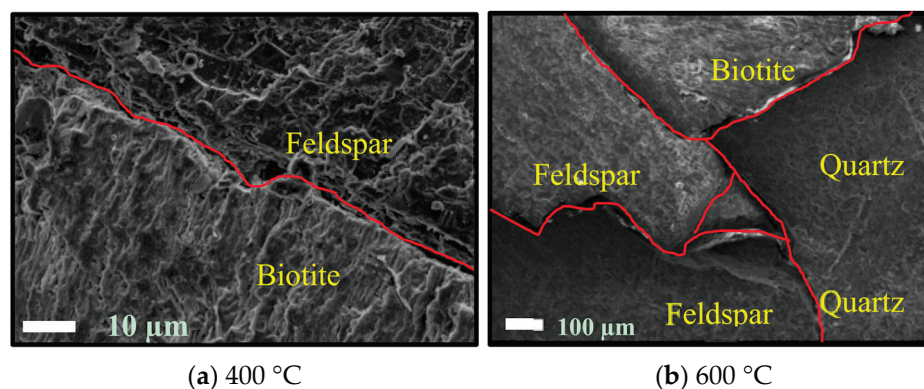


Figure 1. SEM images of heat-treated granites with intergranular cracks highlighted in red [32].

The gathered data and observations indicate that three clear patterns emerge as the temperature increases [18,23,33–40]. From room temperature to about 200 °C, there are generally no visible intergranular cracks in any of the minerals (quartz, feldspar, and biotite). The minerals exhibit a close arrangement, devoid of initial pores or fissures [41]. However, some instances show intergranular cracks between quartz and feldspar minerals when the temperature reaches 200 °C. The initiation of grain boundaries is first observed at the triple junctions of quartz and feldspar minerals, which are typically considered the weakest grain boundaries due to their longer and weaker intergranular bonds. Intergranular cracks between two quartz minerals at 200 °C were also detected. Therefore, this temperature can be regarded as the threshold for the initiation of intergranular cracks in granite.

Upon increasing the temperature to 400 °C, more intergranular cracks appear between two quartz minerals and between quartz and feldspar minerals. More specifically, the intergranular boundary between biotite and feldspar minerals only weakens at 400 °C. These intergranular cracks are caused by strong bound water loss, dihydroxylation loss of constitution water, and solid mineral expansion between 100 and 500 °C, as suggested by previous studies [34]. In contrast, biotite has shown a supporting effect, since the boundaries of biotite with quartz or feldspar did not result in intergranular cracks at 400 °C.

The increase in the temperature to 600 °C leads to an increase in the widths of intergranular cracks between two quartz minerals, quartz and feldspar minerals, and biotite and feldspar minerals. The width of the crack could reach up to 30 μm [34]. However, the intergranular cracks between quartz and biotite minerals initiate at 600 °C. Finally, when the temperature was increased to 800 °C, the width of intergranular cracks between all minerals also increased. At this temperature, more thermal microcracks, which gradually propagate and coalesce within the granite specimens, are observed. Based on these observations, the intergranular crack evolution of granite is summarized in Table 2.

On the other hand, in most experiments, the intragranular cracks initiate starting from 400 °C in feldspar grains with an average width of less than 30 μm [34]. The same goes for large quartz minerals. Crack initiation occurs due to feldspar's lower strength compared to quartz and biotite. Upon increasing the temperature to 600 °C, intragranular cracks occur not only in the feldspar and quartz minerals but also in the biotite minerals, leading to a significant increase in the number of microcracks [37]. This is mostly because of the α/β transition of quartz minerals, which results in a large increase in the volume of quartz minerals, leading to the generation of a substantial amount of cracks. Heating quartz minerals causes a reversible structural phase transformation referred to as the α/β transition of quartz. Furthermore, the remarkable increase in intragranular cracks is due to the dehydration of minerals, lattice reorganization, shrinkage, and decomposition of minerals. Finally, when the temperature was increased to 800 °C, the width of intergranular cracks between all minerals significantly increased. Biotite affects the formation of cracks inside quartz and

feldspar crystals and plays a major role because it expands more when heated compared to quartz and feldspar. The temperature increase led to the observation of more thermal microcracks, which gradually propagated and coalesced within the granite specimens. Based on these microscopic observations, Table 3 summarizes the characteristics of intragranular crack development in granite samples exposed to high-temperature development.

Table 2. The development of intergranular cracks due to temperature.

Temperature (°C)	The Neighboring Minerals			
	Quartz and Quartz	Quartz and Biotite	Quartz and Feldspar	Biotite and Feldspar
25	No intergranular cracks in the mineral boundaries			
200	Crack initiation	No crack initiation	Crack initiation	No crack initiation
400	The intergranular cracks occur between two quartz minerals	No intergranular cracks in the mineral boundaries	The intergranular cracks occur between quartz and feldspar minerals	The intergranular cracks occur between biotite and feldspar minerals
600	Widening of the intergranular cracks in quartz minerals	The intergranular cracks occur between quartz and biotite minerals	Widening of the intergranular cracks between quartz and feldspar minerals	Widening of the intergranular cracks between biotite and feldspar minerals
800	Intergranular cracks fully developed with large apertures.			

Table 3. The development of intragranular cracks due to temperature.

Temperature (°C)	Mineral		
	Quartz	Feldspar	Biotite
200	No intragranular cracks		
400	Intragranular cracks are initiated in large quartz minerals (crystals > 0.5 mm)	Intragranular cracks are initiated in feldspar minerals	Almost intact
600	Widening of the intragranular cracks in quartz minerals	Widening of intragranular cracks in feldspar grains	Initiation of intragranular cracks in large weak biotite minerals (parallel to the grain boundaries and grain boundary cracks)
800	Macroscopic intragranular cracks with large apertures		

4. Physical and Mechanical Parameters

4.1. Porosity

Figure 2a shows the changes in porosity with temperature. At room temperature, granite exhibits low porosity, indicating minimal visibility of the original pores and fissures. Overall, as the temperature rises, granite porosity increases. This is attributed to the formation of thermal microcracks. However, a slight decrease in porosity was observed, mostly at the beginning of the heat treatment, suggesting mineral grain expansion. Below 100 °C, the loose bonding of rock minerals allows water trapped in tiny pores and between layers to escape. Simultaneously, heat treatment prompts mineral particles to expand, leading to the closure of primary cracks. Consequently, rock porosity slightly decreases. Within the 100–300 °C range, physically combined water within the rock evaporates, exerting pressure that facilitates the expansion of cracks and pores. Notably, differences in thermal expansion between feldspar and quartz become apparent, leading to uncoordinated thermal expansion and crack initiation. In the 300–600 °C range, nearly all trapped and physically combined water evaporates, while chemically bound water dissociates, disrupting the mineral lattice. As a result, erratic trends in porosity are observed within the 300–600 °C

temperature range, as shown in Figure 2a. This temperature range may correspond to a transition zone. Despite the continued thermal expansion of mineral particles and the initiation of new cracks, some cracks and pores may close due to particle expansion, resulting in little to no increase in porosity. Overall, porosity exhibits only slight growth up to 600 °C, implying minor structural changes due to existing and new microcrack development. Beyond 600 °C, porosity escalation becomes more pronounced, with significant increases in thermally induced cracks, reaching 5.6% at 800 °C. Thus, the 600–800 °C range marks a critical threshold for temperature-induced changes.

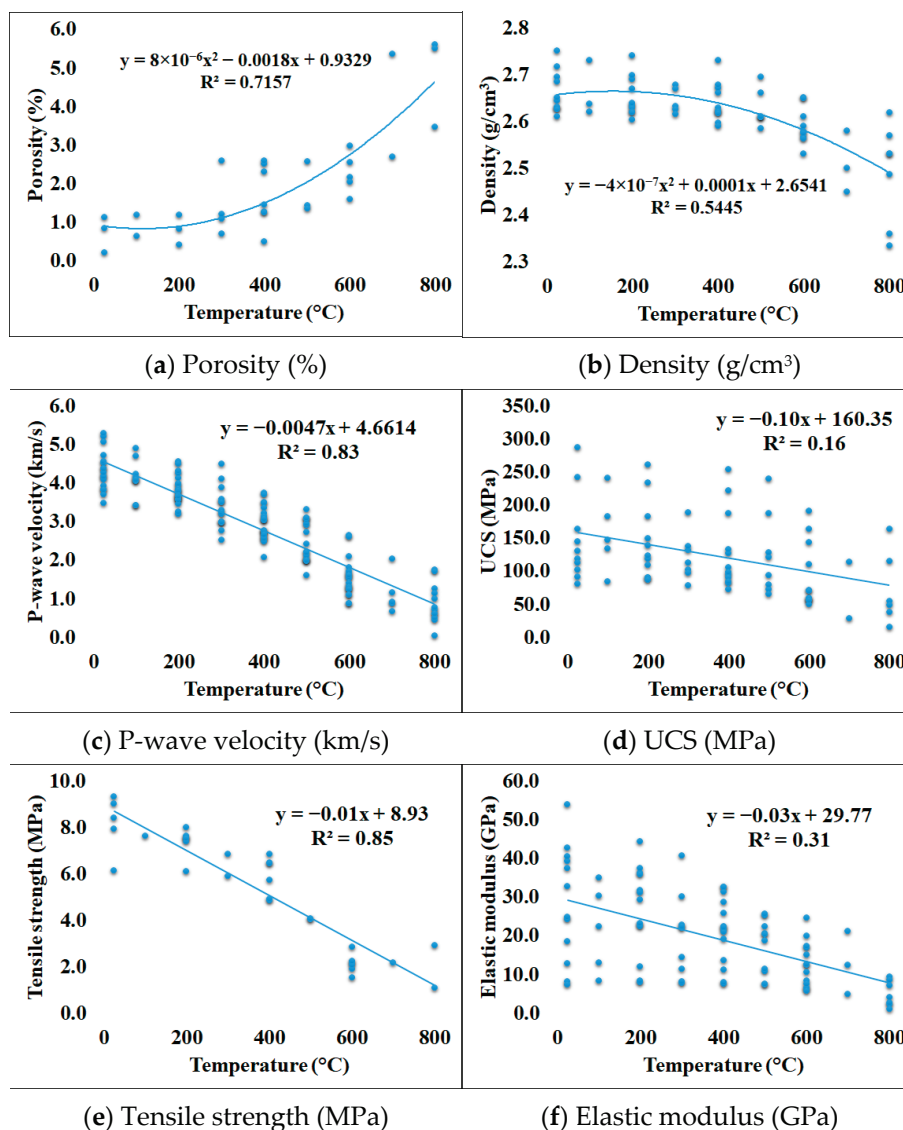


Figure 2. The effect of temperature on the physical and mechanical properties: (a) porosity; (b) density; (c) P-wave velocity; (d) UCS; (e) tensile strength; (f) elastic modulus.

4.2. Density

As shown in Figure 2b, untreated granite has a starting density of about 2.6 g/cm³ and decreases smoothly up to 600 °C. Density decreases by around 0.1 g/cm³ at 700 °C and 0.2 g/cm³ at 800 °C, with the drop being most noticeable. Hence, there are two clear regions: a very low decrease up to 600 °C and a visible decrease between 600 °C and 800 °C, respectively. However, the water content of granite is minimal. Hence, the mass of granite samples does not vary much during heating. Since the rate of mass decrease is far less than

the rate of volume increase, the volume increase due to mineral expansion is the leading factor for the decrease in granite density.

4.3. P-Wave Velocity

Figure 2c shows the effect of high temperature on the P-wave velocity. As the heating temperature rises, the P-wave velocity decreases linearly. Below 600 °C, there are minor changes in P-wave velocity. However, beyond 600 °C, there is a substantial decrease, up to 92% compared to untreated samples, indicating the generation of numerous new thermally induced cracks and significant damage to the granite samples. According to many studies, the relationship between P-wave velocity and temperature (T) can be divided into two segments: a mild decline up to 600 °C and a pronounced decline following treatment temperatures exceeding 600 °C. However, it is difficult to observe these regions in Figure 2c.

4.4. UCS and Tensile Strength

Figure 2d,e shows the effect of high temperature on the UCS and tensile strength, respectively. Both parameters decrease with increasing temperature, consistent with thermally induced degradation in rock strength. UCS indicates low correlation with temperature, while the tensile strength shows a high correlation with temperature. This can be explained by the fact that high temperature tends to transform granite samples from a more brittle material to a more ductile material.

4.5. Elastic Modulus

Finally, Figure 2f shows the effect of high temperature on the elastic modulus of granite samples. The EM decreases slowly from 25 °C to 600 °C, followed by a relatively significant reduction after 600 °C. This implies that the emerging and rapid expansion of microcracks, because of internal thermal stress appearing at higher temperatures, significantly reduces the elastic modulus to some degree. Thus, high temperatures damage rocks.

4.6. Physical and Mechanical Parameters Versus Crack Development

According to Sections 4.1–4.5, two distinct regions were observed: very low changes up to 600 °C and notable changes between 600 °C and 800 °C. Consequently, physical and mechanical parameters exhibited a steady change followed by a rapid alteration. After 600 °C, rock properties have been significantly degraded. These findings do not contradict the observations on crack evolution (Section 3) but rather reinforce the results. They highlight the link between microstructural changes and the degradation of mechanical properties. However, the only difference is that two stages of crack evolution are involved at temperatures up to 600 °C. Since this crack evolution occurs at a microscopic level, defining two stages through physical experiments instead of three stages is acceptable. To summarize, the effect of high temperature on the physical and mechanical parameters of rocks can be understood through studying the crack development.

5. Thermal Damage

In this section, the predictive modeling of granite thermal damage (TD) is discussed.

5.1. Definition of TD and Its Influencing Parameters

The TD is often used to characterize the degree of damage to rock after heating, based on variations in the P-wave velocity with temperature, defined in Equation (1) as [42,43]

$$TD_{in} = 1 - (V_{pT}/V_{p0})^2 \quad (1)$$

where V_{PT} stands for the P-wave velocity (V_p) of samples after heating treatment, and V_{p0} stands for the V_p of the same specimens at room temperature. There are other definitions of thermal damage, but Equation (1) is chosen since P-wave velocity correlates quite well with the crack evolution, which reflects the amount of damage experienced by the sample.

Among the six parameters studied in Section 4, P-wave velocity and tensile strength show a good correlation to the TD_{in} . Figure 3 shows the negative correlation between the predicted TD_{in} and P-wave velocity (Figure 3a) and tensile strength (Figure 3b). Generally, the intact rock sample has a higher P-wave velocity than the fractured one because cracks, air, or fluid gaps slow down the waves. Hence, the trend is that intuitive thermal damage often degrades material properties, reducing the ability to transmit seismic waves. The coefficient of determination R^2 is 0.85 and 0.90 for P-wave velocity and the tensile strength, respectively, which shows a strong inverse relationship between these two parameters. Moreover, tensile strength has a higher negative correlation to the TD_{in} . Hence, tensile strength is a good parameter to define TD_{in} , indicating that strong tensile strength prevents TD in rocks. High-temperature treatment can transform brittle material into a more ductile material. Therefore, it is important to employ all key influencing parameters when evaluating the thermal damage.

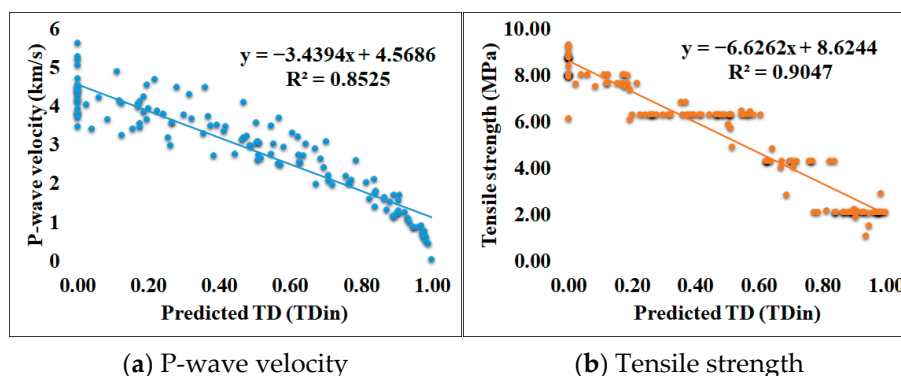


Figure 3. The correlation between the predicted thermal damage (TD_{in}) and (a) P-wave velocity and (b) tensile strength.

5.2. Multivariate Regression Model

The imputed dataset is used to derive an empirical equation (TD_{emp}) that can correlate thermal damage with its influencing factors in a more comprehensive way. The correlation between the parameters P1–P10 and TD_{emp} was determined using SPSS (version 24), and after several trials, the following equation was obtained. The coefficient of determination R^2 was 0.66. The equation can be described as follows:

$$TD_{emp} = 0.09 + 0.0003 \cdot P1 + 0.002 \cdot (P2 + P3 + P4) + 0.024 \cdot P7 - 0.0003 \cdot P9 - 0.036 \cdot P10 \quad (2)$$

The equation indicates that TD_{emp} is directly proportional to the temperature, mineral content, and porosity of granite specimens, while higher UCS and tensile strength have an opposing effect. By observing the parameters' coefficients, the mineral contents (quartz, feldspar, and biotite) may have almost the same impact on the overall TD, while porosity has the most influence on TD among all parameters. Finally, UCS and BTS have negative effects on TD. In other words, strong granite specimens tend to have lower TD, which makes more sense.

Figure 4 shows a good correlation between the actual (defined in Equation (1)) and predicted TD. Hence, the proposed equation can estimate the thermal damage accurately,

where R^2 is 0.59. Since a fair agreement of the coefficient of correlation $R = 0.77$ was reached, in the next section, machine learning algorithms are implemented to enhance the results.

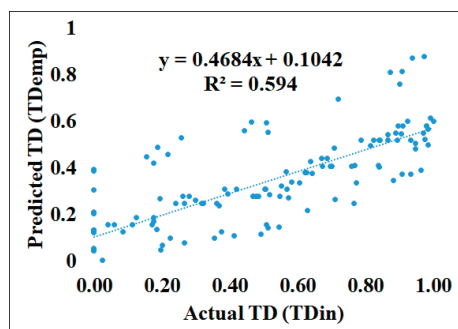


Figure 4. The correlation between the actual and predicted thermal damage.

5.3. Machine Learning Techniques

The compiled dataset was utilized to develop predictive models for thermal damage estimation using several supervised machine learning algorithms within the MATLAB (R2024b) environment, specifically via the Regression Learner Toolbox. The machine learning workflow consisted of the following key steps: data preprocessing, model selection, training, cross-validation, and performance evaluation. Initially, the dataset was prepared by ensuring consistency and removing any remaining anomalies or extreme outliers. No additional feature scaling or transformation was applied as the toolbox handles these internally when required. Several regression models were then trained, including Regression Tree (Fine Tree), Regression Tree (Coarse Tree), Artificial Neural Network (ANN), Gaussian Process Regression, Support Vector Machine (SVM), SVM (Medium Gaussian), Ensemble Boosted Trees, and Ensemble Bagged Trees. Since the original dataset had many missing points, only 136 full data points were used for model training. All models were trained using the default parameter settings provided by the toolbox. To ensure robust model evaluation and prevent overfitting, 10-fold cross-validation was employed during training. The performance of each model was assessed based on the coefficient of determination (R^2) and the root mean square error (RMSE), which measure the proportion of variance explained and the average prediction error, respectively. The results, summarized in Table 4, indicated that model performance varied from fair to excellent, with R^2 values ranging between 0.69 and 0.97, and RMSE values ranging from 0.0633 to 0.1902. Figure 5a–c provides representative plots comparing predicted and actual thermal damage values for selected models, visually illustrating the accuracy and reliability of the predictions.

Table 4. The performance of the predictive models.

Model	R^2	RMSE
Regression Tree (Fine Tree)	0.94	0.0837
Regression Tree (Coarse Tree)	0.69	0.1902
ANN	0.89	0.1110
Gaussian Process Regression	0.97	0.0633
SVM	0.95	0.0790
SVM (Medium Gaussian)	0.91	0.1445
Ensemble Boosted Trees	0.95	0.0741
Ensemble Bagged Trees	0.95	0.0767

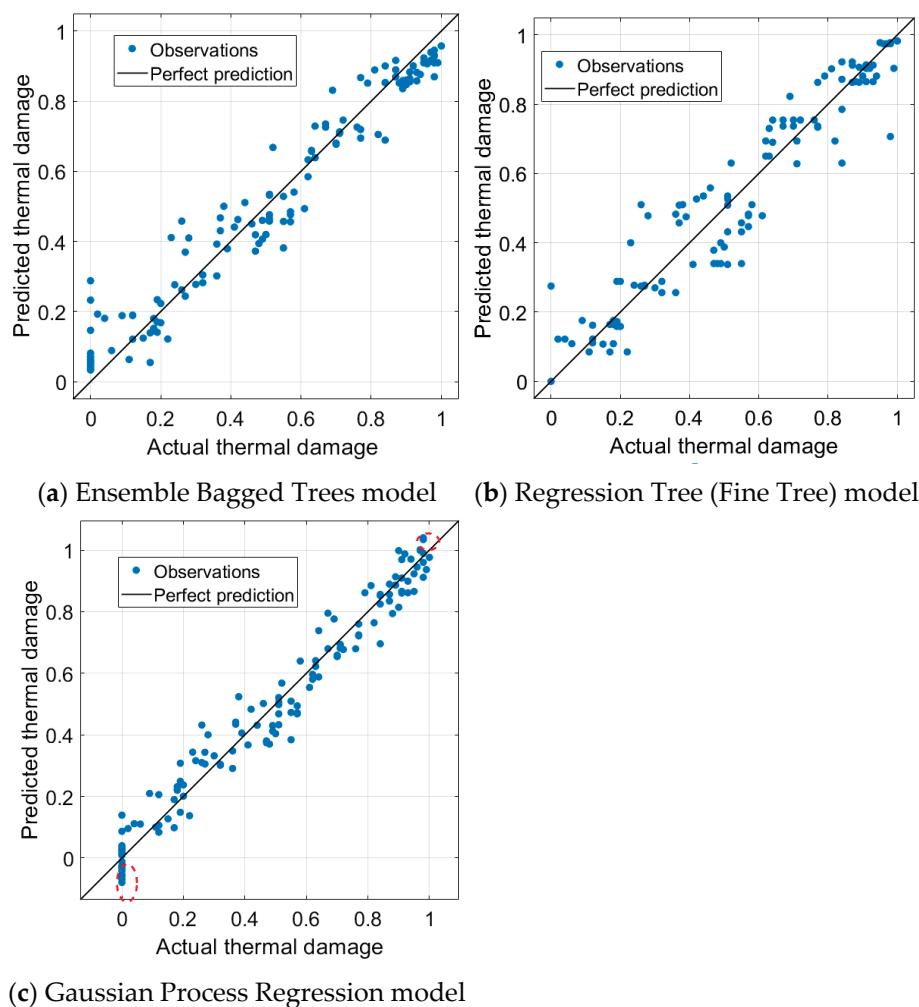


Figure 5. The correlation between the actual and predicted TD of the best predictive models: (a) Ensemble Bagged Trees model; (b) Regression Tree (Fine tree) model; (c) Gaussian Process Regression model.

From Table 4, it can be seen that the Gaussian Process Regression is the best-performing model overall, with the highest R^2 and lowest RMSE values. Then, SVM, Ensemble Boosted Trees, and Ensemble Bagged Trees models equally showed the second-best prediction performance. Next, the Regression Tree (Fine Tree) model performed quite well, indicating strong predictive capability with relatively low error. However, among these top-five models, Gaussian Process Regression, SVM, and Ensemble Boosted Trees models predicted values slightly outside of the TD range: negative values or larger values than 1 (note that the TD, as per the definition, is between 0 and 1). Hence, despite their good prediction capabilities, these models cannot be suggested for TD prediction. As a result, the Ensemble Bagged Trees model (Figure 5a) and the Regression Tree (Fine Tree) model (Figure 5b) are considered the best. As an illustration of prediction outside the range, the Gaussian Process Regression model (Figure 5c) is given.

5.4. Sensitivity Analysis (Importance of the Input Parameters)

To identify and rank the most influential input parameters affecting thermal damage (TD), a sensitivity analysis was conducted using the ReliefF algorithm implemented in MATLAB (R2024b). This step forms part of the broader machine learning workflow, specifically focusing on feature importance analysis—a critical phase used to interpret model behavior and prioritize variables for further investigation or simplification. The ReliefF method is a distance-based algorithm suitable for regression problems. It estimates

the importance (or weight) of each predictor by repeatedly sampling instances and evaluating how well feature values distinguish between instances with similar or different target values. This approach captures not only individual feature effects but also potential interactions with other variables. In this study, the ReliefF function was applied to the input dataset used in the machine learning models. The resulting feature weights indicated the relative importance of each predictor to TD. Figure 6 illustrates the ranked weights, showing the predictors in the following order of importance: P1, P10, P7, P9, P8, P6, P2, P5, P3, and P4 (see Figure 6). Among these, the three most important parameters were identified as (1) temperature, (2) tensile strength, and (3) porosity. These results suggest that mineralogical properties (such as P2 to P5), although expected to be among the top influencing parameters of thermal damage due to differential expansion and microcrack formation during heating, exhibited negative weights and are influenced by TD. Instead, TD mechanics cause changes in mineralogical properties. Interestingly, UCS showed a near-zero weight, implying a limited direct role in the observed TD variations. This may seem counterintuitive, as UCS often initiates cracking. However, it is possible that heating may initiate and close existing microcracks, thereby masking the effects of UCS alone. This outcome highlights the complex interactions among material properties under thermal stress and suggests that a broader set of mechanical and mineralogical features should be considered to fully understand TD behavior.

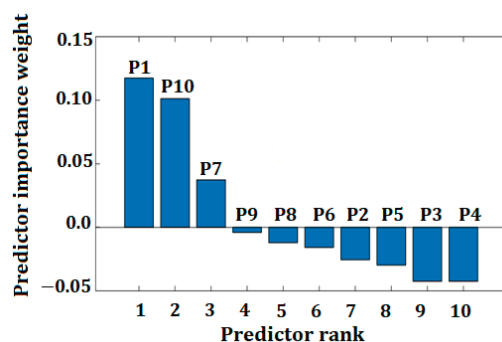


Figure 6. The order of importance of the TD input parameters.

Since mineralogical properties contribute in a reverse way in the TD process, the mineral content parameters were removed from Equation (2), and the equation was redefined as

$$TD_{emp2} = -0.1115 + 0.0007 \cdot P1 + 0.018 \cdot P7 + 0.0001 \cdot (P9/P10) \quad (3)$$

Similarly, Figure 7 shows the updated correlation between the actual (defined in Equation (1)) and predicted TD (defined in Equation (3)). The updated equation can still estimate the thermal damage very accurately, with an R^2 close to 0.94. Compared to the previous equation, a better coefficient of correlation, $R = 0.97$, was reached.

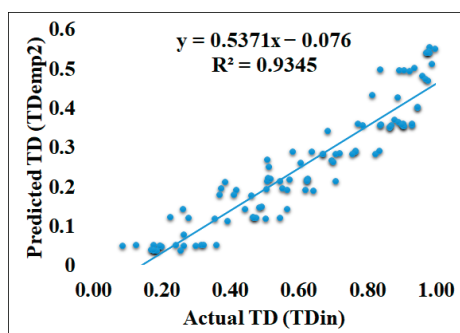


Figure 7. The updated correlation between the actual and predicted thermal damage.

Further removing the porosity term from Equation (3) leads to

$$TD_{emp3} = -0.0991 + 0.0008 \cdot P1 - 0.0001 \cdot (P9/P10) \quad (4)$$

Now, Figure 8 shows the updated correlation between the actual (defined in Equation (1)) and predicted TD (defined in Equation (4)). The updated equation can still estimate the thermal damage very accurately, where R^2 is close to 0.93.

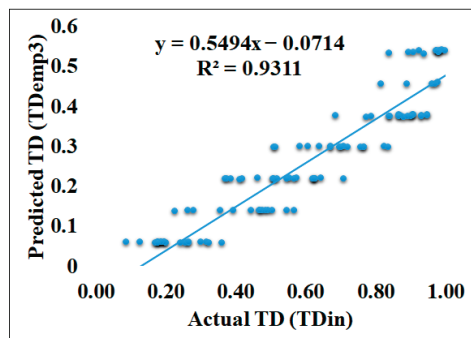


Figure 8. The updated correlation between the actual and predicted thermal damage considering temperature, UCS, and tensile strength only.

5.5. Discussions

Since P-wave velocity and tensile strength showed a good correlation to the TD_{in} , these two parameters could be solely used to evaluate the thermal damage. However, granite has different physical, mechanical, and mineralogical properties. It is known that these factors impact the thermal damage process. Therefore, additional parameters need to be included in the TD model for a more comprehensive assessment. As a result, ten parameters were considered to determine TD empirically. Equation (2) indicates that TD_{emp} is directly proportional to the temperature, mineral content, and porosity of granite specimens, while higher UCS and tensile strength have an opposing effect. Hence, the proposed empirical equation contains not only physico-mechanical parameters but mineral content as well. The correlation between the TD_{in} and TD_{emp} is provided in Figure 4.

Since a fair agreement of $R = 0.77$ was reached, machine learning algorithms were implemented to enhance the results. As a result, among the eight proposed models, the Ensemble Bagged Trees model (Figure 5a) and the Regression Tree (Fine Tree) model (Figure 5b) were considered the best prediction models and outperformed the Gaussian Process Regression model (Figure 5c). The reasons behind this better performance are (1) they are more robust to noise and (2) they make fewer assumptions about the data distribution. They could do even better if the dataset implemented were larger [44,45].

The order of importance of the ten input parameters to evaluate TD in predictive models was assessed, and the results are presented in Figure 6. The three most important parameters were temperature, tensile strength, and porosity. Hence, predictive models with these data can perform well. However, this process indicated that mineralogical properties and UCS exhibited negative weights, being influenced by TD rather than causing it. Therefore, Equation (2) was updated to Equation (3), and Figure 7 was provided, where the updated correlation between the actual and predicted thermal damage is shown. Moreover, Equation (4) was also proposed, where the porosity term is eliminated. Finally, Figure 8 shows the new correlation between the actual and predicted thermal damage considering temperature, UCS, and tensile strength only.

To sum up, three new equations were proposed to determine TD in granites in this study. The main difference is that the new equations do not contain the P-wave velocity parameter as in Equation (1). Figure 3 shows the good correlation between the predicted

thermal damage using Equation (1) and tensile strength. However, tensile strength data in the graph are clustered into 4–5 groups (not continuous); therefore, it alone cannot be used to determine the TD of rocks. Figures 4, 7 and 8 show the correlation between the actual and predicted thermal damage based on Equations (2)–(4), respectively. Based on the order of importance of the TD input parameters shown in Figure 6, Equation (3) is better suited to determine the TD. The novelty of this work is the ability to determine the TD at high temperatures by using temperature, porosity, UCS, and tensile strength data. These findings can be very helpful in geothermal energy extraction, deep mining, and nuclear waste disposal applications where rock is exposed to very high temperatures.

As limitations of this work, the samples in the published literature were heated at different heating rates and then cooled either in an oven or in the open air. The duration of the heating process ranged from half an hour to 24 h. Therefore, variations in the heating rate, heating time, and cooling method could affect the test outcomes. Also, the diversity of the collected data could be an issue. Since the experiments were conducted in different environments, the experimental results could be inconsistent. These shortcomings of the present study can be eliminated either by adding more data, including data related to other rock types, to the dataset or by selecting the data from studies with a similar methodology. As a recommendation, other tests, including heat conductivity, specific heat capacity, thermal diffusivity, and thermal expansion, might yield insightful results and should be the focus of further investigations. Despite these limitations, the current study has elaborated on existing data, leveraging machine learning algorithms to establish a comprehensive thermal damage model.

6. Conclusions

This paper studied crack development (intergranular and intragranular) and the physical and mechanical characteristics of granite specimens at high temperatures. Predictive models based on machine learning algorithms and linear regression were implemented to model the relationships between thermal damage and its influencing factors. Existing data, including mineralogical contents and basic physical and mechanical properties of granite specimens, were collected from the literature and re-interpreted.

In conclusion, the findings of this study can be summarized as follows:

- (1) Crack development analysis revealed that no intra- or intergranular cracks occur below 400 °C; however, intergranular cracks appear between quartz–quartz, quartz–feldspar, and biotite–feldspar at 400 °C, and additional intergranular and intragranular cracks involving biotite, quartz, and feldspar emerge at 600 °C.
- (2) Based on the observed cracking behavior, three temperature zones have been identified: up to 200 °C, between 200 °C and 600 °C, and above 600 °C.
- (3) Physical and mechanical tests indicated a steady decrease in density, P-wave velocity, UCS, tensile strength, and elastic modulus, alongside an increase in porosity as the temperature rose from 25 °C to 800 °C.
- (4) Two distinct zones of property change were noted: minimal changes occurred up to 600 °C, while significant changes were observed between 600 °C and 800 °C, supporting the crack development pattern.
- (5) Predictive models developed using machine learning and linear regression demonstrated excellent performance in estimating TD based on comprehensive input parameters, including mineralogical, physical, and mechanical variables.
- (6) These models can serve as practical tools for planning underground operations in granite-rich environments subjected to thermal effects.
- (7) New equations were proposed to determine the TD of rocks at high temperatures.

- (8) Finally, the novelty of this research lies in the use of soft computing techniques to model and predict thermal damage in granite specimens heated up to 800 °C, using a globally sourced dataset.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/app15116328/s1>, Table S1: Thermal damage raw dataset. References [1,4,10,17–24,26,28,29,32,34–40,46–58] are cited in the supplementary materials.

Author Contributions: Conceptualization, G.S. and A.C.A.; methodology, A.C.A.; software, G.S.; validation, G.S. and A.C.A.; formal analysis, A.C.A. and P.M.G.; investigation, G.S.; data curation, G.S.; writing—original draft preparation, G.S.; writing—review and editing, A.C.A.; supervision, A.C.A. and P.M.G.; project administration, A.C.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors wish to acknowledge the contributions of the anonymous reviewers.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SEM	Scanning electron 57 microscopy
CT	X-ray microcomputed tomography
PM	Polarizing microscope
NMR	Nuclear magnetic resonance imaging
AE	Acoustic emission
TGA	Thermogravimetric 61 analyses
DSC	Differential scanning calorimetry
XRD	X-ray diffraction
MI	Multiple imputation
TD	Thermal damage
UCS	Unconfined compressive strength

References

- Alcock, T.; Bullen, D.; Benson, P.M.; Vinciguerra, S. Temperature-driven micro-fracturing in granite: The interplay between microstructure, mineralogy and tensile strength. *Heliyon* **2023**, *9*, e13871. [CrossRef] [PubMed]
- Zhu, Z.; Yang, S.; Wang, R.; Xie, J.; Tian, N.; Tian, H.; Zheng, J.; Jiang, G.; Dou, B. Experimental investigation on the physical-thermal properties of Nanang granite after air and water cooling for deep geothermal heat extraction. *Renew. Energy* **2024**, *223*, 119963. [CrossRef]
- Wang, L.; Wu, Y.; Huang, Z.; Lin, J.; Wang, Y.; Zhang, W. Effects of temperature and confining pressure on the permeability of Beishan granite from high-level radioactive waste disposal repository. *Case Stud. Therm. Eng.* **2024**, *55*, 104111. [CrossRef]
- Shao, S.; Ranjith, P.; Wasantha, P.; Chen, B. Experimental and numerical studies on the mechanical behaviour of Australian Strathbogie granite at high temperatures: An application to geothermal energy. *Geothermics* **2015**, *54*, 96–108. [CrossRef]
- Shang, X.; Zhang, Z.; Xu, X.; Liu, T.; Xing, Y. Mineral Composition, Pore Structure, and Mechanical Characteristics of Pyroxene Granite Exposed to Heat Treatments. *Minerals* **2019**, *9*, 553. [CrossRef]
- Jin, Y.; He, C.; Yao, C.; Sun, Z.; Wang, J.; Zhang, X.; Yang, J.; Jiang, Q.; Zhou, C. Experimental and numerical simulation study on the evolution of mechanical properties of granite after thermal treatment. *Comput. Geotech.* **2024**, *172*, 106464. [CrossRef]
- Shi, Z.; Li, J.; Ranjith, P.G.; Wang, M.; Lin, H.; Han, D.; Li, K. Acoustic emission and fracture morphology characteristics of thermal-damage granite under mixed mode I/III loading. *Theor. Appl. Fract. Mech.* **2024**, *133*, 104524. [CrossRef]

8. Dang, Y.; Yang, Z.; Liu, X. Three-dimensional numerical study on failure mechanism of granite after thermal treatment based on moment tensor inversion. *Comput. Geotech.* **2023**, *155*, 105185. [CrossRef]
9. Zhu, Z.; Tian, H.; Mei, G.; Jiang, G.; Dou, B. Experimental investigation on physical and mechanical properties of thermal cycling granite by water cooling. *Acta Geotech.* **2020**, *15*, 1881–1893. [CrossRef]
10. Kang, F.; Jia, T.; Li, Y.; Deng, J.; Huang, X. Experimental study on the physical and mechanical variations of hot granite under different cooling treatments. *Renew. Energy* **2021**, *179*, 1316–1328. [CrossRef]
11. Wu, X.; Guo, Q.; Zhu, Y.; Ren, F.; Zhang, J.; Wu, X.; Cai, M. Pore structure and crack characteristics in high-temperature granite under water-cooling. *Case Stud. Therm. Eng.* **2021**, *28*, 101646. [CrossRef]
12. Gao, M.; Li, T.; Zhu, J.; Yin, H.; Yang, Y. An Analysis of Relationship between the Microfracture Features and Mineral Morphology of Granite. *Adv. Civ. Eng.* **2021**, *2021*, 4765731. [CrossRef]
13. Wong, L.N.Y.; Zhang, Y.; Wu, Z. Rock strengthening or weakening upon heating in the mild temperature range? *Eng. Geol.* **2020**, *272*, 105619. [CrossRef]
14. Zhang, Y.; Sun, Q.; Geng, J. Microstructural characterization of limestone exposed to heat with XRD, SEM and TG-DSC. *Mater. Charact.* **2017**, *134*, 285–295. [CrossRef]
15. Shen, Y.-J.; Zhang, Y.-L.; Gao, F.; Yang, G.-S.; Lai, X.-P. Influence of Temperature on the Microstructure Deterioration of Sandstone. *Energies* **2018**, *11*, 1753. [CrossRef]
16. Peng, L.; Li, X.; Peng, X.; Gan, Y.; Wang, J. Analysis of physical and mechanical behaviors and microscopic mineral characteristics of thermally damaged granite. *Sci. Rep.* **2024**, *14*, 14776. [CrossRef]
17. Kumari, W.G.P.; Ranjith, P.G.; Perera, M.S.A.; Shao, S.; Chen, B.K.; Lashin, A.; Arifi, N.A.; Rathnaweera, T.D. Mechanical behaviour of Australian Strathbogie granite under in-situ stress and temperature conditions: An application to geothermal energy extraction. *Geothermics* **2017**, *65*, 44–59. [CrossRef]
18. Xiao, P.; Zheng, J.; Dou, B.; Tian, H.; Cui, G.; Kashif, M. Mechanical Behaviors of Granite after Thermal Shock with Different Cooling Rates. *Energies* **2021**, *14*, 3721. [CrossRef]
19. Zhang, B.; Tian, H.; Dou, B.; Zheng, J.; Chen, J.; Zhu, Z.; Liu, H. Macroscopic and microscopic experimental research on granite properties after high-temperature and water-cooling cycles. *Geothermics* **2021**, *93*, 102079. [CrossRef]
20. Inserra, C.; Biwa, S.; Chen, Y. Influence of thermal damage on linear and nonlinear acoustic properties of granite. *Int. J. Rock Mech. Min. Sci.* **2013**, *62*, 96–104. [CrossRef]
21. Su, G.; Chen, Z.; Ju, J.W.; Jiang, J. Influence of temperature on the strainburst characteristics of granite under true triaxial loading conditions. *Eng. Geol.* **2017**, *222*, 38–52. [CrossRef]
22. Wang, F.; Frühwirt, T.; Konietzky, H. Influence of repeated heating on physical-mechanical properties and damage evolution of granite. *Int. J. Rock Mech. Min. Sci.* **2020**, *136*, 104514. [CrossRef]
23. Jin, P.; Hu, Y.; Shao, J.; Zhao, G.; Zhu, X.; Li, C. Influence of different thermal cycling treatments on the physical, mechanical and transport properties of granite. *Geothermics* **2019**, *78*, 118–128. [CrossRef]
24. Kang, F.; Li, Y.; Tang, C.a. Grain size heterogeneity controls strengthening to weakening of granite over high-temperature treatment. *Int. J. Rock Mech. Min. Sci.* **2021**, *145*, 104848. [CrossRef]
25. Miao, S.; Pan, P.-Z.; Yu, P.; Zhao, S.; Shao, C. Fracture analysis of Beishan granite after high-temperature treatment using digital image correlation. *Eng. Fract. Mech.* **2020**, *225*, 106847. [CrossRef]
26. Chen, Y.-L.; Ni, J.; Shao, W.; Azzam, R. Experimental study on the influence of temperature on the mechanical properties of granite under uni-axial compression and fatigue loading. *Int. J. Rock Mech. Min. Sci.* **2012**, *56*, 62–66. [CrossRef]
27. Xu, Z.; Zhong, X.; Zhang, S.; Wang, Y.; Liu, K.; Liu, X.; Meng, Y.; Hou, X.; Chen, C. Experimental study on mechanical damage and creep characteristics of Gonghe granite under real-time high temperature. *Geothermics* **2024**, *123*, 103100. [CrossRef]
28. Li, C.; Hu, Y.; Meng, T.; Jin, P.; Zhao, Z.; Zhang, C. Experimental study of the influence of temperature and cooling method on mechanical properties of granite: Implication for geothermal mining. *Energy Sci. Eng.* **2020**, *8*, 1716–1728. [CrossRef]
29. Pan, P.-Z.; Shao, C.-Y. Experimental studies on the physical and mechanical properties of heated rock by air, water and high-viscosity fluid cooling. *Geomech. Energy Environ.* **2022**, *31*, 100315. [CrossRef]
30. Austin, P.C.; White, I.R.; Lee, D.S.; van Buuren, S. Missing Data in Clinical Research: A Tutorial on Multiple Imputation. *Can. J. Cardiol.* **2021**, *37*, 1322–1331. [CrossRef]
31. Woods, A.D.; Gerasimova, D.; Van Dusen, B.; Nissen, J.; Bainter, S.; Uzdavines, A.; Davis-Kean, P.E.; Halvorson, M.; King, K.M.; Logan, J.A.R.; et al. Best practices for addressing missing data through multiple imputation. *Infant Child Dev.* **2024**, *33*, e2407. [CrossRef]
32. Isaka, B.L.A.; Gamage, R.P.; Rathnaweera, T.D.; Perera, M.S.A.; Chandrasekharam, D.; Kumari, W.G.P. An Influence of Thermally-Induced Micro-Cracking under Cooling Treatments: Mechanical Characteristics of Australian Granite. *Energies* **2018**, *11*, 1338. [CrossRef]
33. Li, M.; Wang, D.; Shao, Z. Experimental study on changes of pore structure and mechanical properties of sandstone after high-temperature treatment using nuclear magnetic resonance. *Eng. Geol.* **2020**, *275*, 105739. [CrossRef]

34. Yang, S.-Q.; Ranjith, P.G.; Jing, H.-W.; Tian, W.-L.; Ju, Y. An experimental investigation on thermal damage and failure mechanical behavior of granite after exposure to different high temperature treatments. *Geothermics* **2017**, *65*, 180–197. [CrossRef]
35. Chen, S.; Yang, C.; Wang, G. Evolution of thermal damage and permeability of Beishan granite. *Appl. Therm. Eng.* **2017**, *110*, 1533–1542. [CrossRef]
36. Gautam, P.K.; Verma, A.K.; Jha, M.K.; Sharma, P.; Singh, T.N. Effect of high temperature on physical and mechanical properties of Jalore granite. *J. Appl. Geophys.* **2018**, *159*, 460–474. [CrossRef]
37. Kumari, W.G.P.; Ranjith, P.G.; Perera, M.S.A.; Chen, B.K.; Abdulagatov, I.M. Temperature-dependent mechanical behaviour of Australian Strathbogie granite with different cooling treatments. *Eng. Geol.* **2017**, *229*, 31–44. [CrossRef]
38. Fan, L.F.; Gao, J.W.; Wu, Z.J.; Yang, S.Q.; Ma, G.W. An investigation of thermal effects on micro-properties of granite by X-ray CT technique. *Appl. Therm. Eng.* **2018**, *140*, 505–519. [CrossRef]
39. Shao, Z.; Wang, Y.; Tang, X. The influences of heating and uniaxial loading on granite subjected to liquid nitrogen cooling. *Eng. Geol.* **2020**, *271*, 105614. [CrossRef]
40. Griffiths, L.; Heap, M.J.; Baud, P.; Schmittbuhl, J. Quantification of microcrack characteristics and implications for stiffness and strength of granite. *Int. J. Rock Mech. Min. Sci.* **2017**, *100*, 138–150. [CrossRef]
41. Mnzool, M.; Al-Mukhtar, A.; Majeed, A.J.; Arafat, A.; Gomaa, E. Simulation and performance characteristics of rock with borehole using Visual Finite Element Analysis. *Min Min. Depos* **2024**, *18*, 33–41. [CrossRef]
42. Wang, Z.; Feng, C.; Wang, J.; Song, W.; Wang, H. An Improved Statistical Damage Constitutive Model for Rock Considering the Temperature Effect. *Int. J. Geomech.* **2022**, *22*, 04022203. [CrossRef]
43. Liu, W.; Yin, S.; Thanh, H.V.; Soltanian, M.R.; Yu, Q.; Yang, S.; Li, Y.; Dai, Z. Advancements and development trend in statistical damage constitutive models for rock: A comprehensive review. *Nat. Hazards* **2025**, *121*, 3703–3744. [CrossRef]
44. Liu, H.; Cai, J.; Ong, Y.-S.; Wang, Y. Understanding and comparing scalable Gaussian process regression for big data. *Knowl. Based Syst.* **2019**, *164*, 324–335. [CrossRef]
45. Hultquist, C.; Gang, C.; Zhao, K. A comparison of Gaussian process regression, random forests and support vector regression for burn severity assessment in diseased forests. *Remote Sens. Lett.* **2014**, *5*, 723–732. [CrossRef]
46. Zhu, Z.; Kempka, T.; Ranjith, P.G.; Tian, H.; Jiang, G.; Dou, B.; Mei, G. Changes in thermomechanical properties due to air and water cooling of hot dry granite rocks under unconfined compression. *Renew. Energy* **2021**, *170*, 562–573. [CrossRef]
47. Fan, L.F.; Wu, Z.J.; Wan, Z.; Gao, J.W. Experimental investigation of thermal effects on dynamic behavior of granite. *Appl. Therm. Eng.* **2017**, *125*, 94–103. [CrossRef]
48. Wu, Q.; Weng, L.; Zhao, Y.; Guo, B.; Luo, T. On the tensile mechanical characteristics of fine-grained granite after heating/cooling treatments with different cooling rates. *Eng. Geol.* **2019**, *253*, 94–110. [CrossRef]
49. Zhu, Z.; Tian, H.; Mei, G.; Jiang, G.; Dou, B.; Xiao, P. Experimental investigation on mechanical behaviors of Nanai granite after thermal treatment under conventional triaxial compression. *Environ. Earth Sci.* **2021**, *80*, 46. [CrossRef]
50. Yin, T.-B.; Shu, R.-H.; Li, X.-B.; Wang, P.; Dong, L.-J. Combined effects of temperature and axial pressure on dynamic mechanical properties of granite. *Trans. Nonferrous Met. Soc. China* **2016**, *26*, 2209–2219. [CrossRef]
51. Zhu, S.; Zhang, W.; Sun, Q.; Deng, S.; Geng, J.; Li, C. Thermally induced variation of primary wave velocity in granite from Yantai: Experimental and modeling results. *Int. J. Therm. Sci.* **2017**, *114*, 320–326. [CrossRef]
52. Sun, Q.; Zhang, W.; Xue, L.; Zhang, Z.; Su, T. Thermal damage pattern and thresholds of granite. *Environ. Earth Sci.* **2015**, *74*, 2341–2349. [CrossRef]
53. Liu, S.; Xu, J. An experimental study on the physico-mechanical properties of two post-high-temperature rocks. *Eng. Geol.* **2015**, *185*, 63–70. [CrossRef]
54. Shao, S.; Wasantha, P.; Ranjith, P.; Chen, B. Effect of cooling rate on the mechanical behavior of heated Strathbogie granite with different grain sizes. *Int. J. Rock Mech. Min. Sci. Géoméch. Abstr.* **2014**, *70*, 381–387. [CrossRef]
55. Xu, X.L.; Karakus, M. A coupled thermo-mechanical damage model for granite. *Int. J. Rock Mech. Min. Sci.* **2018**, *103*, 195–204. [CrossRef]
56. Zhang, F.; Zhang, Y.; Yu, Y.; Hu, D.; Shao, J. Influence of cooling rate on thermal degradation of physical and mechanical properties of granite. *Int. J. Rock Mech. Min. Sci. Géoméch. Abstr.* **2020**, *129*, 104285. [CrossRef]
57. Németh, A.; Antal, Á.; Török, Á. Physical Alteration and Color Change of Granite Subjected to High Temperature. *Appl. Sci.* **2021**, *11*, 8792. [CrossRef]
58. Guo, P.; Zhang, N.; He, M.; Bai, B. Effect of water saturation and temperature in the range of 193 to 373K on the thermal conductivity of sandstone. *Tectonophysics* **2017**, *699*, 121–128. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Pseudotachylyte Formation in Brittle–Ductile Transition of the Anning River Fault Zone: Implications for Seismic Processes

Wenhao Dai ^{1,2}, Yongsheng Zhou ^{2,*}, Huiru Lei ^{2,*}, Xi Ma ², Jiaxiang Dang ², Sheqiang Miao ², Shimin Liu ^{2,3,*} and Changrong He ²

¹ Jiangxi Earthquake Agency, Nanchang 330026, China; daiwh19@163.com

² State Key Laboratory of Earthquake Dynamics and Forecasting, Institute of Geology, China Earthquake Administration, Beijing 100029, China; maxi@ies.ac.cn (X.M.); dangjiaxiang@ies.ac.cn (J.D.); miaosq@ies.ac.cn (S.M.); crhe@ies.ac.cn (C.H.)

³ School of Fire Protection Engineering, China People's Police University, Langfang 065000, China

* Correspondence: zhouyush@ies.ac.cn (Y.Z.); leihuiru@ies.ac.cn (H.L.); liushimin@cpgu.edu.cn (S.L.)

Abstract: Pseudotachylytes and cataclasites record transient seismic slips within the brittle–ductile transition zone and ductile flow layers. Investigating the mechanisms of pseudotachylytes can provide the most direct geological evidence for revealing seismic fault slip and coseismic processes. We investigate the deformation and chemical composition of pseudotachylytes, cataclasites, and mylonites collected from the Anning River fault zone in this study. Three kinds of pseudotachylyte veins were found in granite gneiss and cataclasite. Microstructural analyses show that pseudotachylytes and cataclasites developed within granitic gneiss and mylonites, and EBSD analysis indicates granitic gneiss deformed at temperatures of 250–350 °C. All of the pseudotachylytes are enriched in Fe and Ca, with SiO₂ content closely resembling that of the wall rock of granitic gneiss. The geochemical results indicate that pseudotachylytes originated from the in situ melting of granitic gneiss, which was produced during coseismic frictional heating. Based on the deformation and geochemical data of mylonites, cataclasites, and pseudotachylytes, a simple model of the seismogenic layer is established for rock deformation during coseismic, post-seismic relaxation, and interseismic periods. Mylonite represents the rheological flow of the brittle–ductile transition zone during interseismic periods, cataclasites display brittle fracturing during coseismic rupture, and pseudotachylytes stand for localized melting induced by coseismic frictional heating. During the post-seismic relaxation, crack healing and static recrystallization of quartz occur.

Keywords: pseudotachylyte; cataclasite; mylonite; Anning River fault zone; coseismic deformation

1. Introduction

Seismic fault mechanics have traditionally focused on the brittle upper crust of tectonically active regions. In classical models, deformation below the brittle–ductile transition (BDT) is thought to be governed by crystal plasticity, while earthquakes are attributed to rapid brittle failure. Consequently, seismic activity was not expected to occur in the deeper, ductile portions of the crust. However, recent observations challenge this assumption. Seismic nucleation has been documented in strong lower crustal rocks within the continental lithosphere, particularly beneath the BDT [1]. In the upper crust, deformation primarily occurs via brittle processes such as frictional sliding and faulting, whereas dislocation creep dominates in the middle and lower crust under thermally activated conditions [2].

During the seismic cycle, the deformation mechanisms of fault rocks are significantly influenced by temperature, strain rate, and pore fluid pressure [3–14]. The BDT is thus recognized as a critical zone for controlling the depth extent of seismicity in the continental crust [15,16], with its behavior highly sensitive to temperature, strain rate, and pore fluid pressure [11–14,17]. Importantly, deformation mechanisms within this transition zone vary across different seismic phases—including coseismic loading, post-seismic relaxation, and interseismic creep—due to variations in strain rate and fluid conditions [11–14,18]. During interseismic periods, low strain rates promote ductile flow, while higher strain rates during coseismic and post-seismic phases favor brittle failure.

To fully understand the transient creep behavior and strain rate sensitivity of rocks in the BDT during the seismic cycle, it is essential to investigate the microstructures and deformation mechanisms preserved in deep fault rocks [19]. Key evidence for such processes comes from pseudotachylyte and mylonite, which together record the complex interplay of brittle and ductile deformation in the lower crust [20]. Mylonites form under ductile conditions in shear zones, while pseudotachylytes typically originate from high-speed coseismic slip, producing solidified frictional melt along fault surfaces [21–25].

Traditionally, pseudotachylyte has been considered a product of dry, high-velocity frictional melting [23,26]. However, new studies suggest that pseudotachylyte may also form during slower processes such as ductile creep, microseismic rupture, and slow slip events [1,19,27,28]. Furthermore, pseudotachylyte formation has been observed in fluid-rich fault zones [29] and may be preceded by intense fragmentation and grain size reduction [30]. The preservation of pseudotachylyte is also problematic due to its susceptibility to post-seismic alteration and weathering during exhumation [26,31,32]. Nevertheless, pseudotachylyte has now been documented at depths of 22–50 km [33–35], indicating that earthquake processes can extend well below the traditionally defined BDT. Microscale amorphization structures observed in both natural fault gouges and experimental granite deformation further suggest that intense shear, rather than melting alone, can produce pseudotachylyte-like textures [36–39].

The Anning River fault zone provides an ideal natural laboratory for such research. We aim to investigate the chemical composition, microstructures, and deformation mechanisms of pseudotachylytes, cataclasites, and mylonites from the brittle–ductile transition zone in this study. Based on the deformation and petrogenetic origin of pseudotachylyte, a deformation model was established that captures the behavior of fault rocks during coseismic, post-seismic, and interseismic periods in the brittle ductile transition zone, thereby contributing to a more comprehensive understanding of seismic fault dynamics.

2. Materials and Methods

2.1. Samples

In the Xichang segment of the Anning River fault zone, pseudotachylyte has been found in the pre-Avalonian granite gneiss, representing coseismic deformation at the fault source depth during a geological historical period. This study collected a large number of pseudotachylyte and related tectonic rock samples, which are the main research objects of this study (Figure 1).

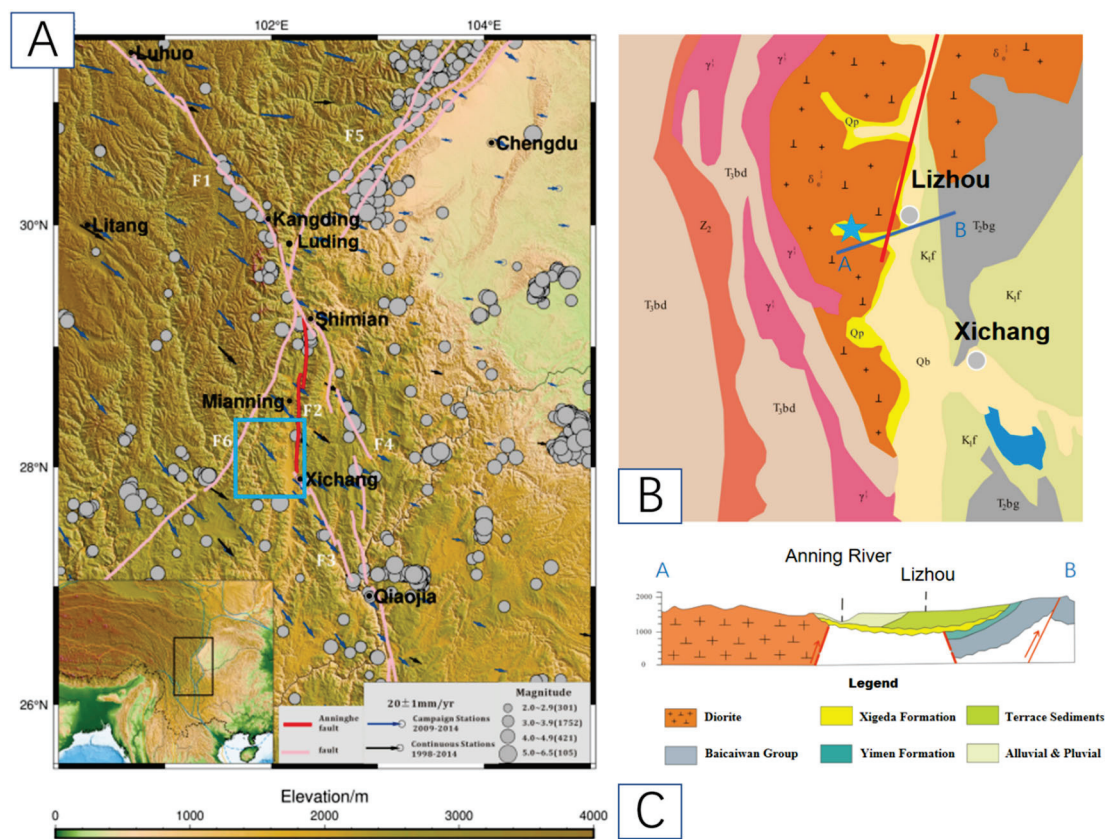


Figure 1. Simplified geological map of the sample collection area along the Anning River fault zone. (A) Schematic diagram of the Anning River fault zone [40]; (B) simplified geological map near Lizhou, where the red line represents the location of the Anning River fault zone and the blue line indicates the position of geological cross-section AB; and (C) geological cross-section AB.

Samples AF19-43A, AF19-43D, AF19-43G, AF19-44A, AF19-44C, and AF21-4 (of which 43D, 43G, and 44C are rolled stone samples, and the rest are in situ samples, as shown in Figures 2 and 3) are mylonite and shattered rock containing pseudotachylyte. These were all collected from the west side of Lizhoushan, Minsheng Township, Changshitou Village, Xichang City, Sichuan Province, with the coordinates N28°05′27.94″ E102°02′55.71″ (Figures 2 and 3). The elevation is 1725 m.



Figure 2. Sections of samples from the sample collection area. (A) Schematic diagram of the sampling profile on the south bank of the river at Changshitou Village (AF19-43A to AF21-1); (B) schematic diagram of the sampling profile on the south bank of the river at Changshitou Village (AF21-1 to AF21-4).

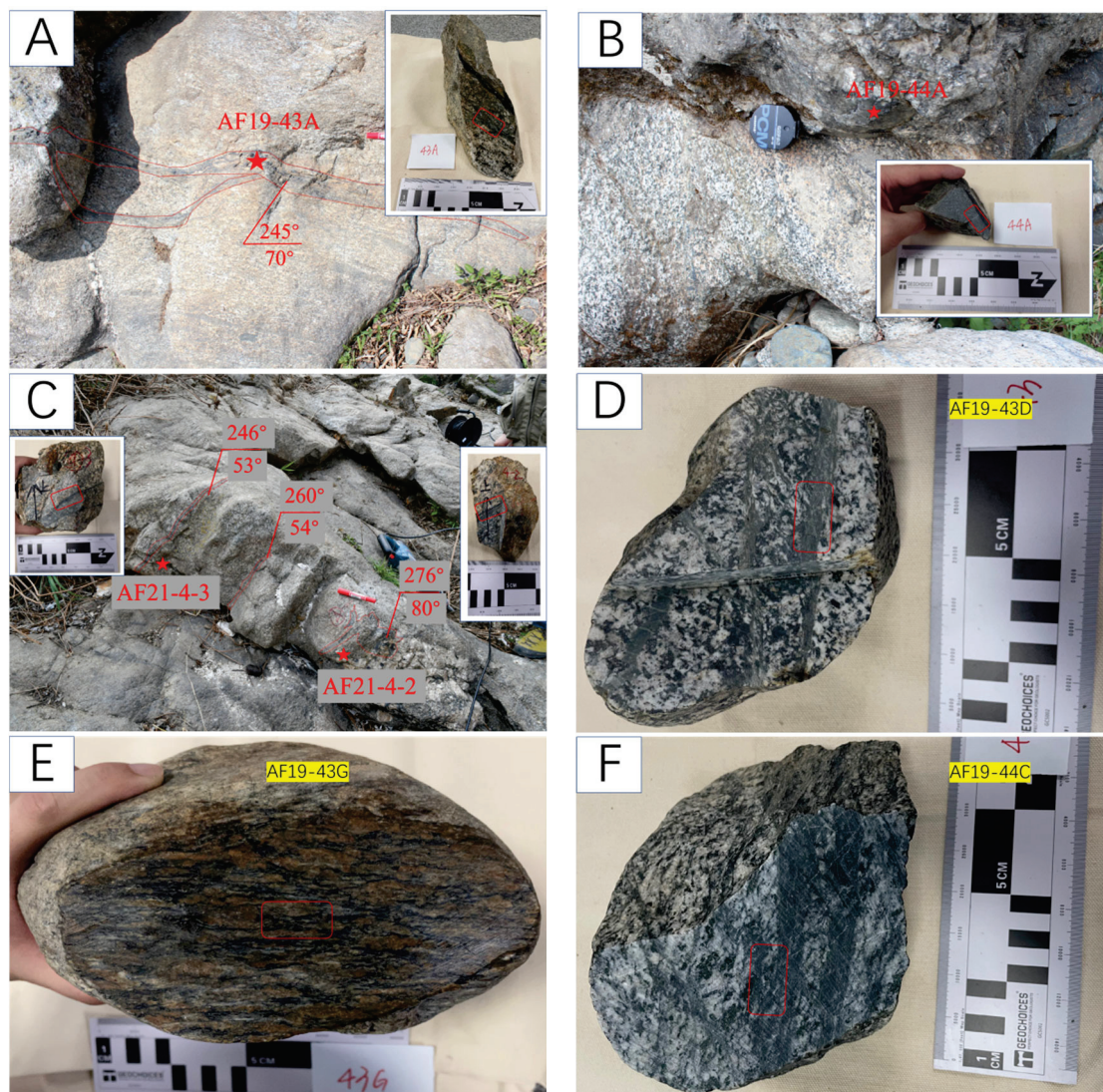


Figure 3. Images of specimens. (A) AF19-43A granitic gneiss containing pseudotachylyte; (B) AF19-44A gneiss containing pseudotachylyte; (C) AF21-4-2 and AF21-4-3 breccia containing pseudotachylyte; (D) AF19-43D breccia containing pseudotachylyte; (E) AF19-43G mylonite containing pseudotachylyte; and (F) AF19-44C breccia containing pseudotachylyte.

The sampling site is located on the southern riverbed of a river channel trending nearly east–west, where the bedrock is composed of early Proterozoic granite gneiss. Locally, mylonite or shattered rocks develop, with several branching pseudotachylyte veins along the mylonitic foliation (Figure 3). The vein widths are approximately 1–2 cm, and they extend in east–west and southwest directions, with traceable widths of about 20–30 m. These veins cut each other, showing features of synchronous activity. The attitudes of the pseudotachylyte veins are $246^\circ < 53^\circ$, $260^\circ < 54^\circ$, and $276^\circ < 80^\circ$. The surrounding mylonitic foliation in the host rock has an attitude of $72^\circ < 80^\circ$. A 20 cm wide pegmatite vein is also developed in the host rock (Figure 3), with an attitude of $310^\circ < 54^\circ$, and pseudotachylyte is observed along the vein, with an attitude of $302^\circ < 47^\circ$.

In the riverbed, mylonite cobbles containing pseudotachylyte veins can be seen. These cobbles provide the three-dimensional spatial distribution characteristics of the pseudotachylyte veins. Some veins are filled in clear, straight fractures (Figure 3), while others are injected along the fractured network formed by the shattered rocks (Figure 3).

2.2. Methods

This study used a variety of analytical techniques to examine the microstructure and chemical composition of the samples, including Optical Microscopy, scanning electron microscopy (SEM)–Energy Dispersive Spectroscopy (EDS), Electron Backscatter Diffraction (EBSD), X-ray Diffraction (XRD), X-ray Fluorescence Spectroscopy (XRF), and Electron Probe Micro-analysis (EPMA).

Polarizing Microscope & SEM-EDS: Microstructures of the samples were observed using a polarizing optical microscope (LEICA) and a scanning electron microscope (Zeiss SIGMA) (SEM) at the State Key Laboratory of Earthquake Dynamics, Institute of Geology, China Earthquake Administration, Beijing, China. The SEM equipped with an Oxford X-Max 50N energy-dispersive X-ray spectroscopy (EDS) were used for measuring element composition, respectively.

EBSD: The quartz fabrics were measured by SEM-EBSD analysis using a scanning electron microscope (SEM) coupled with an Oxford Nordlys nano electron backscattered diffraction (EBSD) detector at the State Key Laboratory of Earthquake Dynamics, Institute of Geology, China Earthquake Administration, Beijing, China. Polished thin sections were carbon-coated prior to EBSD analysis. EBSD allows for the determination of the full crystallographic orientation of crystals in a thin section using a backscattered electron (BSE) signal. Automatic EBSD measurements were used to obtain orientation maps and crystallographic preferred orientation (CPO) patterns. Working conditions during acquisition of the EBSD patterns were 20 kV accelerating voltage, 12 nA probe current, 70° sample tilt, ~20 mm working distance, and high vacuum. EBSD patterns of quartz were acquired and indexed with HKL AZtec software (Oxford Instruments), which selected the direction of the stretching lineation as the X-axis direction; that is, the quartz–mica strip direction is the X-axis direction, and the plane-oriented direction is the XY-plane direction. CHANNEL 5 software (Oxford Instruments) was used for noise reduction and for filling the missing data with at least eight identical neighbors with similar orientation. All stereographic diagrams of EBSD are lower hemisphere projections.

XRD: XRD analysis was performed to qualitatively identify mineral compositions in pseudotachylyte samples and host rocks, XRD data were collected using Bruker D8 ADVANCE X-ray diffractometers with Cu K α radiation over the 2 θ range of 3°–70°, with a step size of 0.02° at the Beida Zhihui Microstructure Analytical Center in Beijing.

XRF: Whole-rock chemical analysis of major elements in pseudotachylyte and host rocks using a Magix-pro2440 XRF machine at the Institute of Geology, Chinese Academy of Geological Sciences.

EPMA: Major element compositions of mineral grains were determined using a JEOL JXA-8230 electron probe micro-analyzer (EPMA) at the Institute of Tibetan Plateau Research, Chinese Academy of Sciences. The analytical conditions were 15-kV accelerating voltage, 20-nA beam current, 5- μ m probe diameter, and 5- to 10-s counting time (background and peak) per element. EPMA analyses were performed on 10 grains for each mineral.

The above methods are described in detail in the following Table 1.

Table 1. Analytical methods and co-application necessity.

Analytical Method	Research Objective	Technical Features	Co-Application Necessity
Polarizing microscope	Pseudotachylyte morphology observation Micro-deformation analysis	Macro-scale structural resolution (50–500 μ m)	Morphology composition co-analysis
SEM-EDS	Mineral assemblage quantification Micro-fracture filling characterization	Micron-scale morphology imaging (1–50 μ m) Elemental mapping	

Table 1. Cont.

Analytical Method	Research Objective	Technical Features	Co-Application Necessity
EBSD	Quartz crystallographic orientation determination Deformation mechanism inversion	Nanoscale crystallographic orientation imaging CPO pattern reconstruction	The EBSD crystallographic orientation data and XRD phase analysis jointly constrain the deformation history and the formation depth of the melt
XRD	Mineral phase identification	Crystal structure fingerprint recognition	
XRF	Whole-rock major element analysis	Bulk chemical composition determination	XRF bulk chemistry coordinates with EPMA micro-chemistry to reveal element migration patterns
EPMA	Mineral micro-domain chemical variation analysis	Micron-scale element quantification (0.1–5 μm)	

3. Results

This study focuses on the pseudotachylyte found in the granite gneiss within the Anning River fault zone. Based on the contact relationship between the pseudotachylyte and the host rock, three kinds of pseudotachylyte veins were identified: (A) veins with a bifurcated pattern, splitting from the main vein and gradually tapering as it fills the fractures in the host rock (AF19-43A, AF19-43D, and AF19-44C); (B) some veins have undergone later alteration, exhibiting significant plastic deformation features (AF19-43G); and (C) multi-stage veins with a combination of various characteristics (AF21-4-2).

3.1. Microstructures of Pseudotachylyte

3.1.1. Straight Vein-Type Pseudotachylyte in Granite Gneiss and Cataclasite

The straight vein-type pseudotachylyte appears in the cataclasite, with a straight, black pseudotachylyte vein clearly delineated from the host rock.

(1) Sample AF19-44A is a pseudotachylyte vein developed in granite gneiss, with a distinct boundary between the vein and the host rock. The vein has a banded appearance, with a width of approximately 0.5 cm (Figure 4A). This is the dominant type of pseudotachylyte distribution in the study area. The pseudotachylyte vein is dark to black with a clastic texture both in field samples and under Optical Microscopy and SEM. The angular clasts consist of plagioclase and quartz. The plagioclase clasts in the host rock are angular or sub-angular, while the quartz clasts are sub-angular or sub-rounded (Figure 4B–D). In the host rock, brittle fracturing to form cataclasite (Figure 4A, red) and localized S-C structures from initial mylonitization (Figure 4A, yellow) are observed. Under scanning electron microscopy, plagioclase, quartz, and dark minerals are seen to be flattened and elongated in a directional manner, with folding (Figure 4E–G).

The microstructure of this sample suggests that the granite gneiss, with fractured rocks forming cataclasite, produced the pseudotachylyte vein during the coseismic rupture process. It is hypothesized that the mylonitized and cataclastic granite gneiss surrounding the pseudotachylyte vein serves as its protolith. Due to the high-temperature melting and coseismic deformation, the contact area between the pseudotachylyte and the host rock underwent plastic deformation.

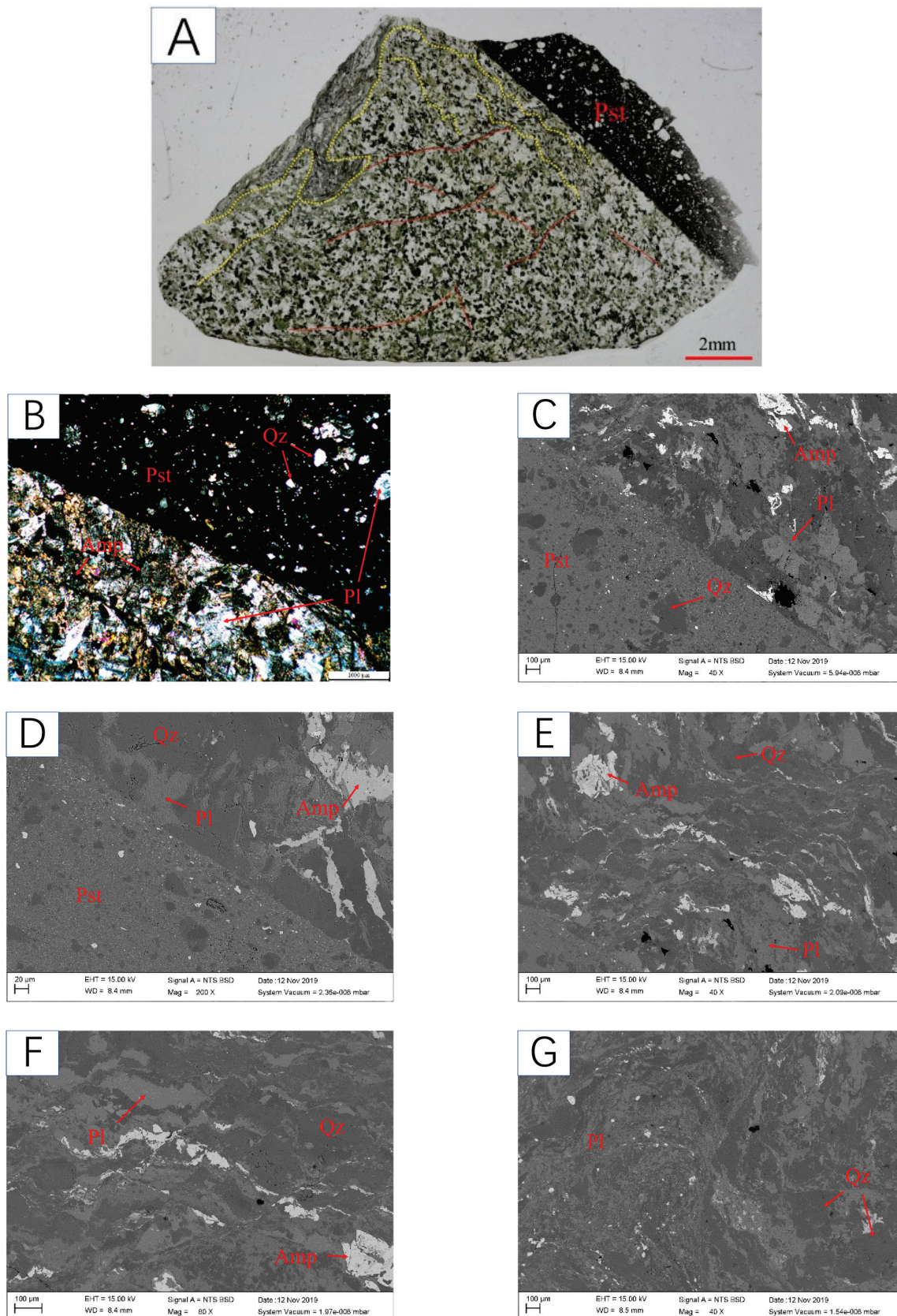


Figure 4. Microstructures of a straight vein of pseudotachylyte in AF19-44A. (A) AF19-44A under plane-polarized light (PPL). Red solid lines demarcate brittle fractures; yellow dashed lines indicate domains of plastic deformation. (B) AF19-44A under cross-polarized light (XPL). (C–G) Scanning electron micrographs (SEM) of AF19-44A.

(2) In sample AF21-4-3, the boundary of the black pseudotachylyte vein is straight and distinct (Figure 5A–C), representing the deformation of the coseismic rupture stage. The host rock is an equigranular granite gneiss, where numerous brittle fractures are observed (Figure 5A, red) and localized plastic deformation is present around the vein (Figure 5A, yellow). The host rock exhibits localized fracturing to form cataclasite, with the fractured clastic angular clasts being angular or sub-angular plagioclase and sub-angular or sub-rounded quartz grains (Figure 5D,E). This indicates that the host rock of the pseudotachylyte is cataclasite fractured from granite gneiss.

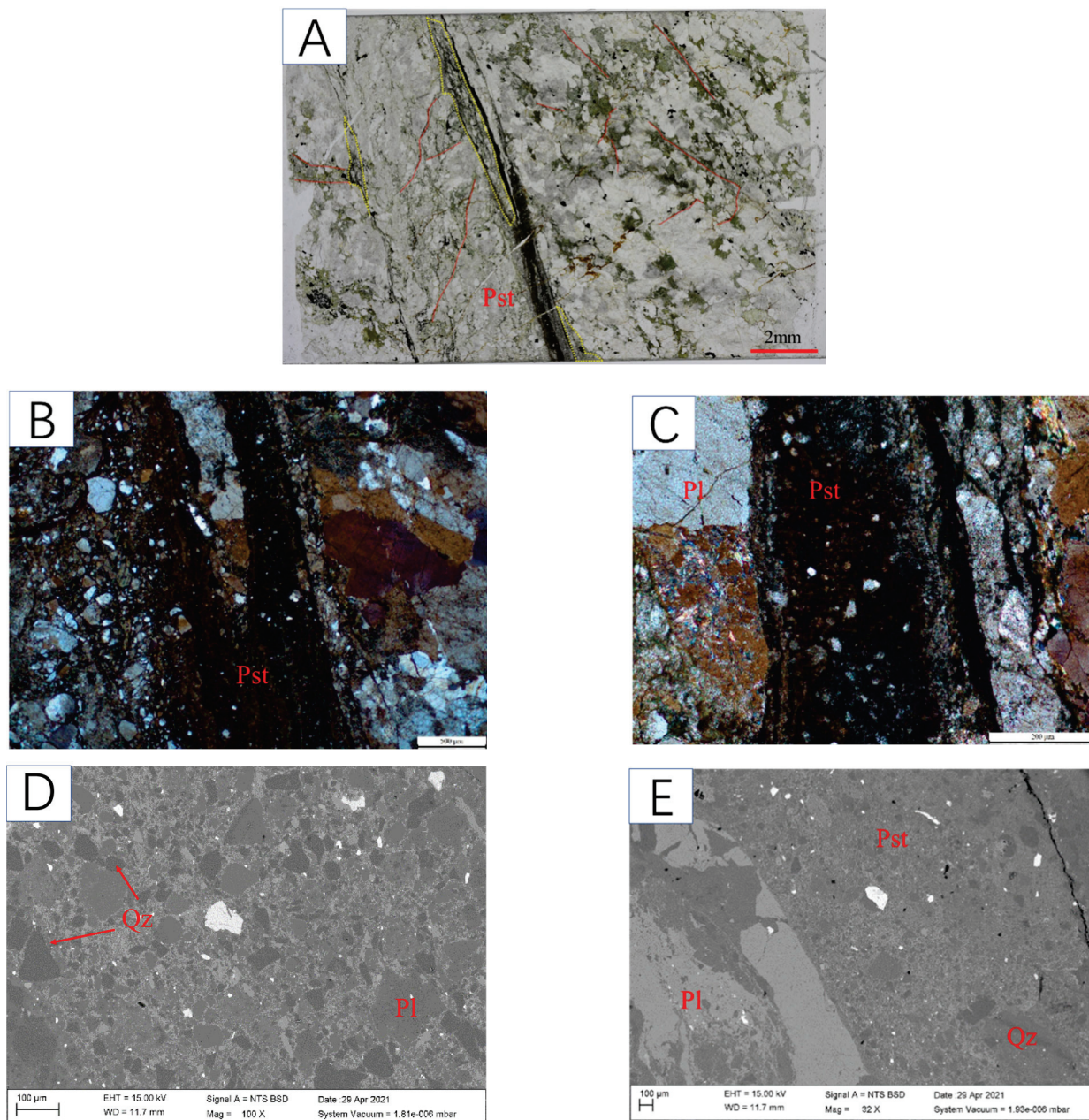


Figure 5. Microstructures of pseudotachylyte with secondary veins in AF21-4-3. (A) AF21-4-3 under PPL. (B,C) AF21-4-3 under XPL. (D,E) SEM of AF21-4-3.

(3) In sample AF19-43A, the boundary between the pseudotachylyte vein and the surrounding fractured granite gneiss is clear. The glass exhibits a dendritic shape, where two narrow veins converge into a main vein, with a width of approximately 0.3–1 cm (Figure 6A, black). The plagioclase clasts in the host rock are angular or sub-angular, and

some clasts show internal fractures. The quartz clasts are sub-angular or sub-rounded, with a grain size slightly smaller than that of the plagioclase clasts (Figure 6B,C). Under scanning electron microscopy, the de-glassified mineral aggregates with cryptocrystalline growth can be observed (Figure 6D,E).

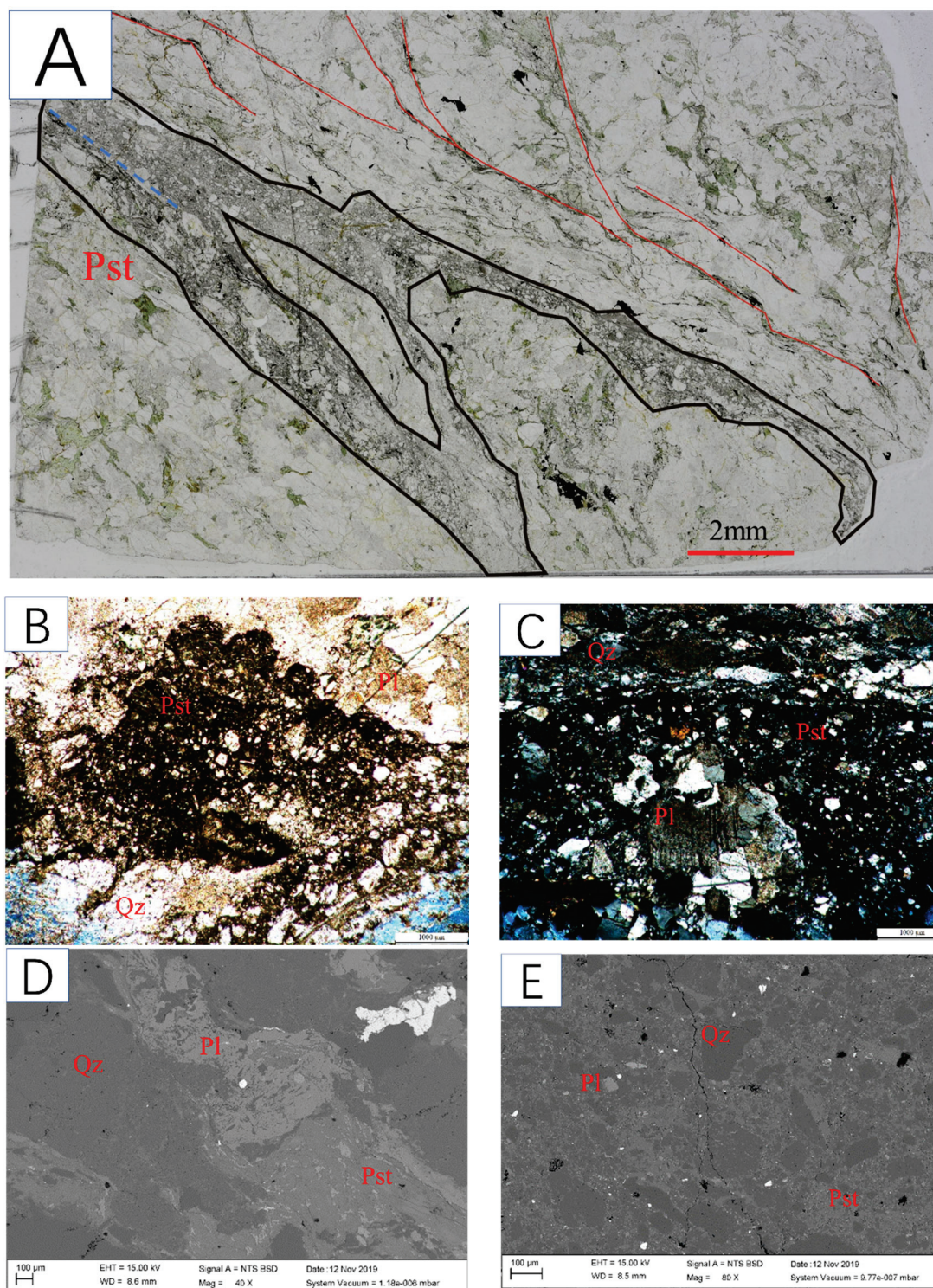


Figure 6. Microstructures of multi-branching veins of pseudotachylyte in AF19-43A. (A,B) AF19-43A under PPL. (C) AF19-43A under XPL. (D,E) SEM of AF19-43A.

In the host rock, multiple sets of secondary brittle fractures are observed at high angles to the main fracture surface (Figure 6A, red). These secondary fractures are mostly filled with fragmented clasts, with a small amount of melt glass. The intersection of

the secondary fractures forms a clear and straight boundary (Figure 6A, blue), which is presumed to be a rupture surface containing melt glass, formed by simultaneous sliding along two intersecting slip surfaces. This is likely the product of the initial sliding stage during coseismic motion.

(4) In sample AF19-43D, the pseudotachylyte primarily exhibits a straight vein morphology, with localized injection into fractured fissures. The boundary between the main pseudotachylyte vein and the surrounding fractured granite gneiss is distinct, and the secondary veins are irregularly distributed in a net-like pattern (Figure 7A), gradually tapering off. The clasts in the host rock are euhedral or subhedral plagioclase and quartz (Figure 7A,B). The plagioclase clasts are angular, with visible internal fractures (Figure 7C,D). Quartz grains exhibit undulatory extinction (Figure 7C,D). Under scanning electron microscopy, mineral aggregates with de-glassified growth can be observed within the pseudotachylyte vein (Figure 7E–G).

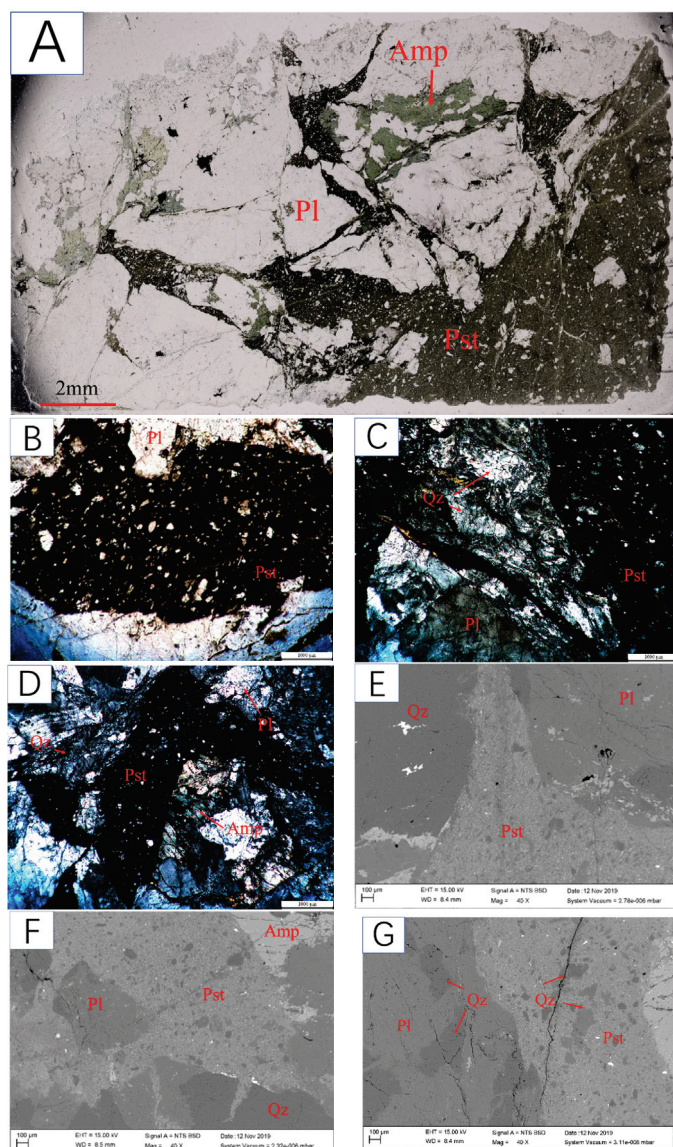


Figure 7. Microstructures of pseudotachylyte with secondary injected veins in AF19-43D. (A,B) AF19-43D under PPL. (C,D) AF19-43D under XPL. (E–G) SEM of AF19-43D.

This sample represents the main developmental stage of coseismic rupture, where sliding friction along the fault surface reached its peak, and the frictional temperature reached its maximum. The melt produced along the fault surface increased, forming

the main body of the pseudotachylite vein, which also injected into secondary fissures near the main fault. The cryptocrystalline or vitreous mineral aggregates within the vein underwent de-glassification and growth during later deformation, indicating a process of stress unloading.

3.1.2. Distributed Fine Vein-Type Pseudotachylite Along the Foliation of Mylonite

This type of pseudotachylyte appears in mylonite, with discontinuous fine veins of pseudotachylyte distributed along the foliation.

(1) In sample AF19-44C, the banded pseudotachylyte fine veins are distributed along the foliation of the granite mylonite, with clear but discontinuous boundaries. The veins are intermittent or taper off, and their shape is not straight but rather curved (Figure 8A). Two types of quartz grains are observed in the surrounding mylonite: one type is larger, with deformed edges bending to form bulging structures, and under crossed polarized light, sheared deformation bands are visible (Figure 8B); the other type consists of quartz aggregates with smaller grain size, showing allotriomorphic or subhedral shapes, and are arranged in bands between larger mineral grains. These quartz grains exhibit undulatory extinction (Figure 8B–D). The fine-grained quartz aggregates form banded structures along the foliation, exhibiting distinct plastic flow features (Figure 8D–G). The pseudotachylyte veins and the quartz bands alternate in a directional arrangement and exhibit plastic deformation characteristics (Figure 8A,E–G), indicating that the pseudotachylyte and fine quartz aggregates underwent plastic deformation together.

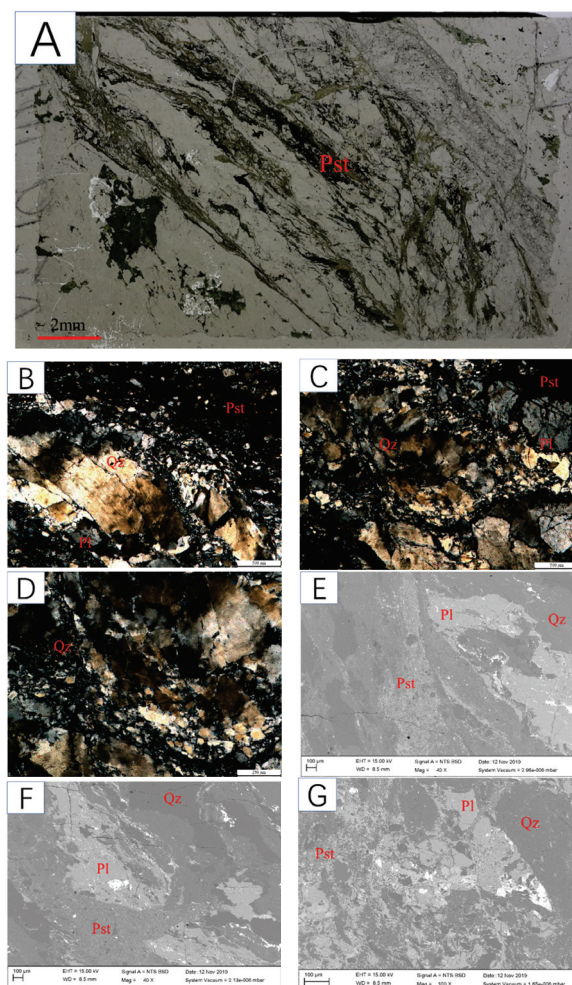


Figure 8. Microstructures of pseudotachylite distributed along the foliation in AF19-44C. (A) AF19-44C under PPL. (B–D) AF19-44C under XPL. (E–G) SEM of AF19-44C.

(2) Sample AF19-43G also contains fine veins of pseudotachylyte with plastic deformation. The host rock in this sample shows a porphyritic structure under polarized light, with plagioclase phenocrysts exhibiting σ -shaped or eye-shaped textures (Figure 9A–C). Under scanning electron microscopy, crushed structures are visible inside the phenocrysts, with fractures filled by quartz veins. The tailing ends of the phenocrysts show elongated fine-grained quartz and biotite aggregates, which surround the plagioclase grains (Figure 9D,E), indicating that the plagioclase has undergone hydrolysis to form quartz. The fine-grained aggregates of plagioclase, quartz, and biotite form a banded structure, with quartz subgrains developing, and the edges of the grains show bulging and subgrain rotation recrystallization (Figure 9B,C), indicating medium-temperature dislocation creep in quartz.

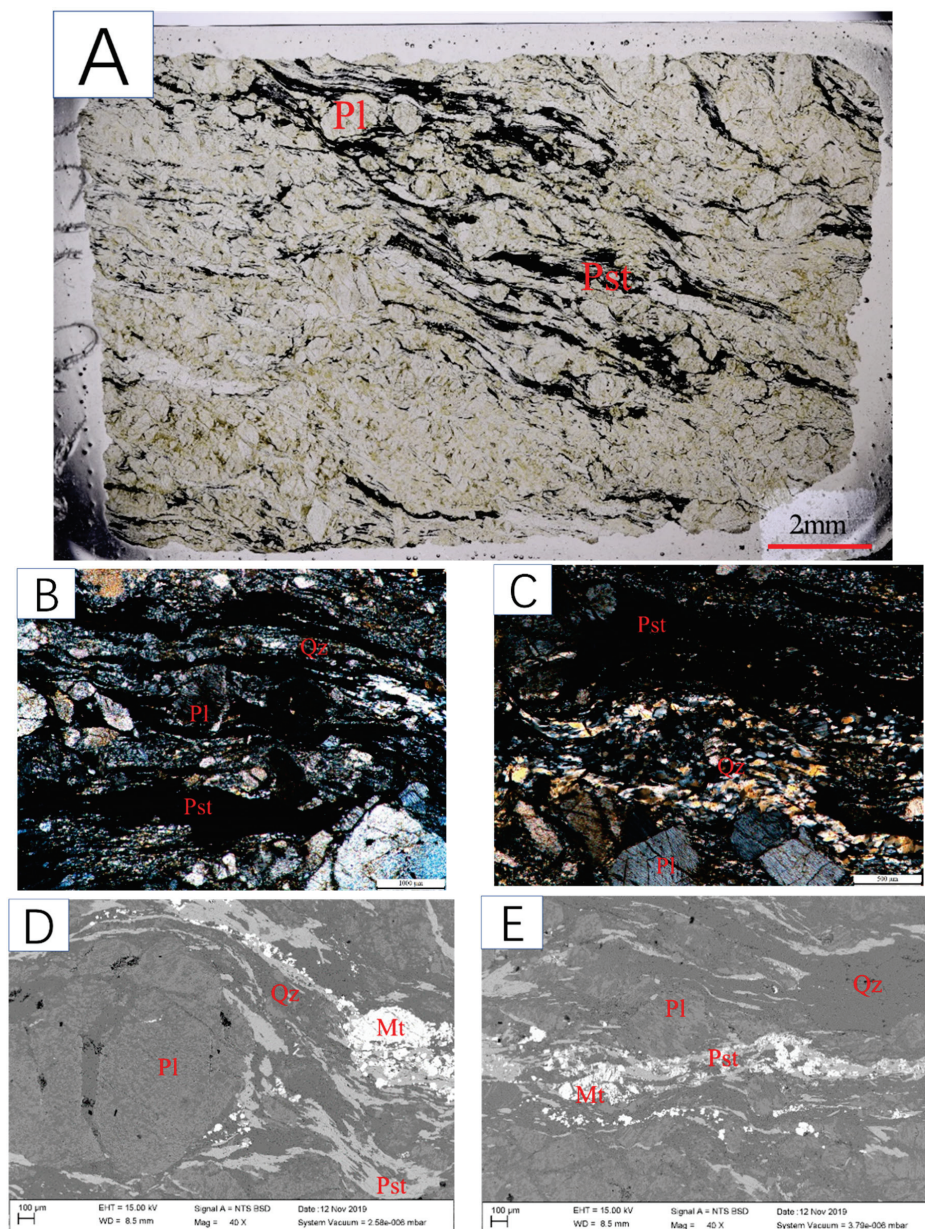


Figure 9. Microstructures of pseudotachylyte that underwent plastic deformation in AF19-43G. (A) AF19-43G under PPL. (B,C) AF19-43G under XPL. (D,E) SEM of AF19-43G.

Along the foliation, fine pseudotachylyte veins exhibit a directional distribution, with discontinuous boundaries, interruptions, or tapering. The deformation features of the pseudotachylyte veins are consistent with those of the quartz bands, and the veins deform

plastically around the eye-shaped plagioclase phenocrysts (Figure 9A–C). These features suggest that after the formation of pseudotachylyte, it and the fine quartz aggregates underwent plastic deformation together. Under scanning electron microscopy, granular magnetite inclusions were observed within the pseudotachylyte (Figure 9D,E).

3.1.3. Multi-Stage Pseudotachylyte Veins in Granite Gneiss

The host rock of this type of pseudotachylyte is fractured granite gneiss, which is similar to the host rock of the straight vein-type pseudotachylyte.

(1) In sample AF21-4-2, three stages of pseudotachylyte veins can be observed (Figure 10A). The boundary between each vein and the host rock is clearly visible, and the veins have a banded morphology, with distinct cutting relationships between them. The first-stage pseudotachylyte vein is about 0.2 cm wide, developed within a shear zone, with an irregular shape and injection characteristics. The vein is black, with a high matrix content and fewer clastic clasts, and it is cut by a wider fracture zone greater than 1 cm. The second-stage pseudotachylyte vein is lighter in color, black–gray, and clearly cuts through the first-stage vein at a large angle. This vein contains a higher proportion of clastic clasts, with the clasts being plagioclase and quartz. The plagioclase clasts are angular or sub-angular, while the quartz clasts are sub-angular or sub-rounded (Figure 10C). The matrix shows de-glassified growth of cryptocrystalline mineral aggregates under scanning electron microscopy (Figure 10D,E), parallel to the fracture zone.

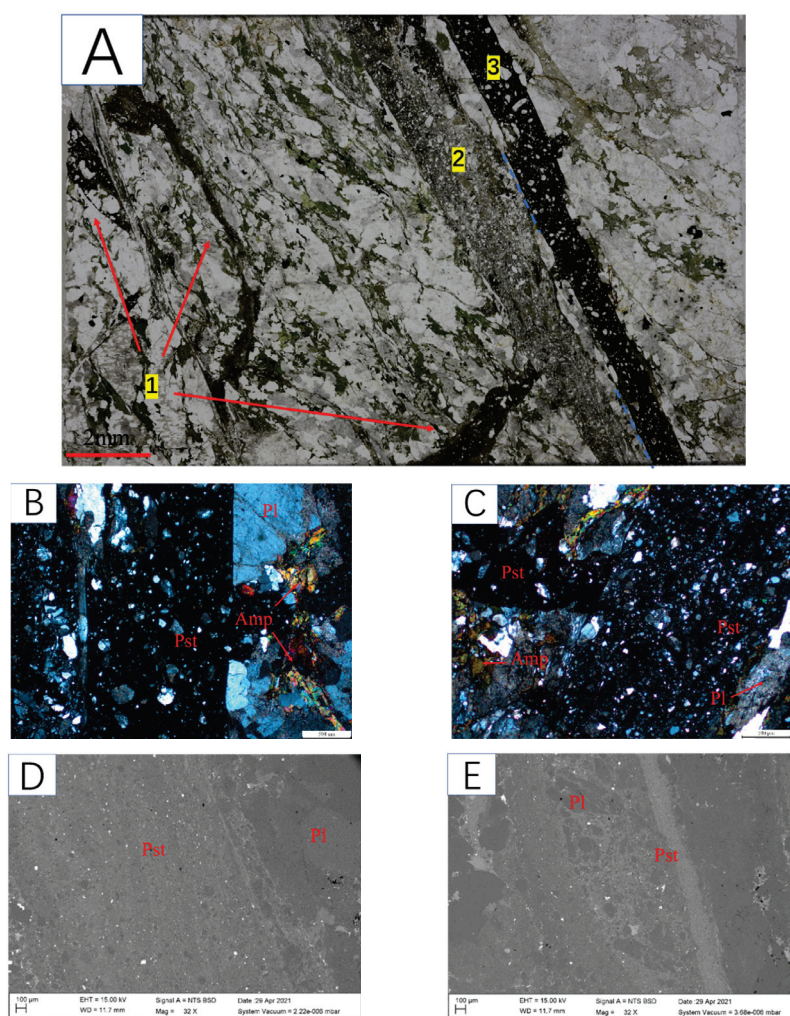


Figure 10. Microstructures of multi-period pseudotachylytes in AF21-4-2. (A) AF21-4-2 under PPL, (1)–(3) are three stages of pseudotachylyte veins. (B,C) AF21-4-2 under XPL. (D,E) SEM of AF21-4-2.

The third-stage pseudotachylyte vein is nearly parallel to the second-stage vein but locally cuts through it (Figure 10A, blue). The boundary between the two stages is straight and clear. The third-stage vein has a higher matrix content (Figure 10B), is black in color macroscopically, and contains clastic clasts of plagioclase and quartz. However, no de-glassified growth of cryptocrystalline minerals is observed in the matrix. The microstructure and composition of the third-stage vein are similar to those of the straight vein-type pseudotachylyte.

It is evident that the exposed fault rocks have undergone multiple coseismic slip events throughout geological history. Due to the limited distribution and retention of information in the first two stages of pseudotachylyte veins, this study focuses on the last stage of fresh pseudotachylyte, which is more widely exposed in outcrops.

Table 2 summarizes the key characteristics of the various types of pseudotachylyte discussed above.

Table 2. Summary of pseudotachylyte vein types.

Vein Type	Host Rock	Morphology	Key Characteristics	Representative Samples
Straight vein-type	Granite gneiss and cataclasite	Straight, banded appearance with clear boundaries	Dark to black color with clastic texture Angular/sub-angular plagioclase and quartz clasts Evidence of brittle fracturing forming cataclasite Localized S-C structures from initial mylonitization Flattened and elongated minerals with folding under SEM Formed during coseismic rupture with high-temperature melting	AF19-44A, AF21-4-3, AF19-43A, AF19-43D
Fine vein-type	Mylonite	Discontinuous fine veins along foliation	Clear but discontinuous boundaries Intermittent or tapering, curved shapes Alternates with quartz bands in the directional arrangement Exhibits plastic deformation features Two types of quartz grains: larger deformed grains and fine aggregates Shows medium-temperature dislocation creep features	AF19-44C, AF19-43G
Multi-stage veins	Fractured granite gneiss	Multiple generations with distinct cutting relationships	Three stages are visible with clear boundaries First stage: irregular shape, high matrix content Second stage: lighter color, higher clast proportion Third stage: similar to straight vein-type, higher matrix content De-glassified cryptocrystalline mineral aggregates in matrix Evidence of multiple coseismic slip events	AF21-4-2

3.2. Fabric of Quartz in Granite Gneiss and Mylonite Around Pseudotachylyte Under Electron Backscatter Diffraction (EBSD)

Under different temperature conditions, the intracrystalline slip system of quartz varies, leading to different crystal preferred orientations (CPOs) [41–43]. When the temperature increases or the strain rate decreases, the main slip systems of quartz are basal slip, rhombic slip, and prismatic slip [44,45]. Previous studies have shown that during the plastic deformation of quartz under non-coaxial shear strain, at low temperatures (250–350 °C), deformation primarily occurs through basal $\langle a \rangle$ dislocation slip creep; at medium–low temperatures (350–450 °C), it primarily occurs through rhombic $\langle a \rangle$ dislocation slip creep; at medium temperatures (450–600 °C), it mainly occurs through prismatic $\langle a \rangle$ dislocation slip creep; and at high temperatures (600–700 °C), it mainly occurs through prismatic $\langle c \rangle$ dislocation slip creep [44–48]. Among these, low-temperature fabrics can overlap with early high-temperature fabrics, and the evolution from high to low temperature in ductile shear zones can form composite fabric models [49]. Due to the limited number of slip systems in quartz, the slip systems can be analyzed simply from the c-axis pole figure. By examining the distribution of the extreme densities in the pole figure, the slip

system of quartz deformation can be identified, thus allowing for the determination of the deformation temperature.

In this study, the host rock of pseudotachylyte consists of granite gneiss and granitic mylonite. To determine the depth (temperature) conditions of pseudotachylyte formation, the most typical straight vein-type pseudotachylyte samples, AF21-4-2 and AF21-4-3, were selected for EBSD analysis of quartz in the host rock (parent rock of the melt), fractured granite gneiss (Figure 11).

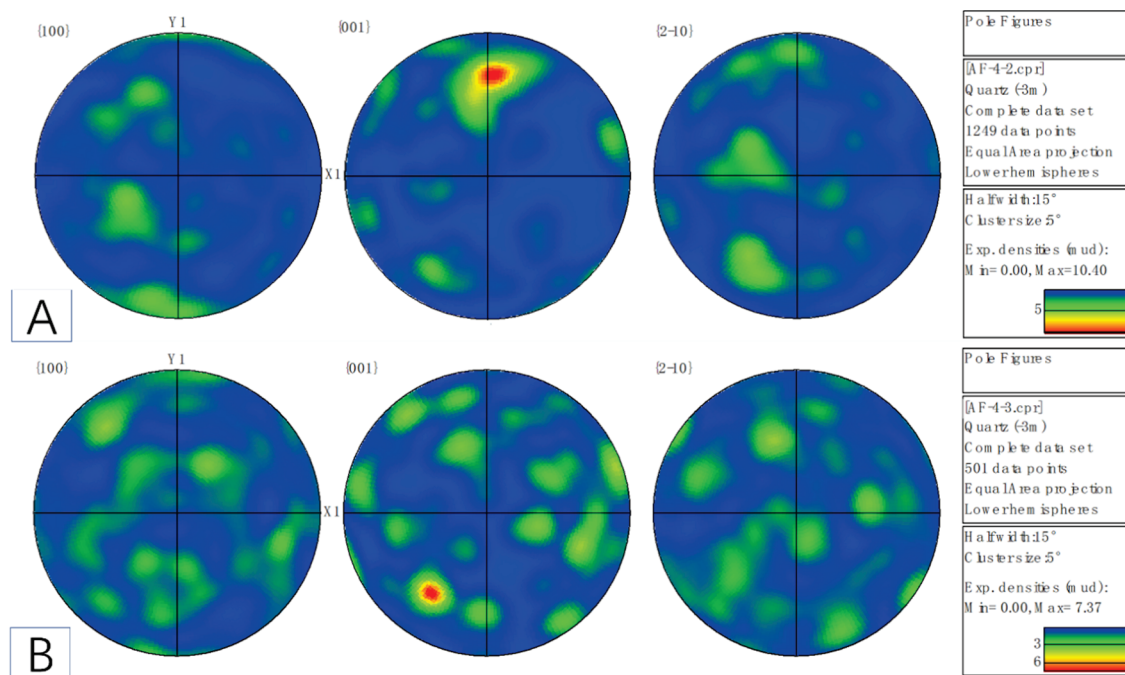


Figure 11. EBSD for quartz in the host rock of pseudotachylyte. (A) AF21-4-2; (B) AF21-4-3.

For sample AF21-4-2, EBSD analysis of fine-grained quartz in the host rock identified 1249 distinct grains. The pole figure of the c-axis {001} shows a clear orientation, consistent with low-temperature basal dislocation slip creep.

For sample AF21-4-3, a surface scan was performed on quartz in the host rock, and noise reduction was carried out using Channel 5 software. A total of 501 distinct quartz grains were identified, and a pole figure was generated (Figure 11B). Similarly, the pole figure of the c-axis {001} of quartz in this sample shows clear orientation, and the extreme density pattern indicates low-temperature basal $\langle a \rangle$ dislocation slip creep characteristics.

EBSD analyses of quartz in the granitic host rock reveal a pronounced c-axis alignment characteristic of low-temperature basal $\langle a \rangle$ dislocation creep, indicating deformation temperatures of roughly 250–350 °C in the fractured granite gneiss surrounding straight vein-type pseudotachylyte. By correlating these temperatures with typical geothermal gradients, the host-rock deformation—and hence melt injection—depth is constrained to the brittle–ductile transition zone at approximately 15–20 km. Importantly, this melt is not derived from high-temperature magmatic intrusion but is instead generated in situ by coseismic, high-velocity frictional heating along fault planes. Thus, EBSD-derived quartz slip system characterization provides a robust thermobarometric framework for estimating both host rock deformation conditions and the depth of pseudotachylyte formation.

3.3. Chemical Composition

To analyze the genesis of pseudotachylyte and its relationship with the host rock, typical pseudotachylyte and host rock samples were selected for electron probe composition

analysis. To obtain pure chemical data of pseudotachylyte, this study focuses on the straight vein-type pseudotachylyte samples, which are widely distributed at outcrops, for analysis.

In samples AF19-43D and AF19-44A, an SEM-EDS surface scan and electron probe major element analysis were conducted on the pseudotachylyte vein. A cross-sectional measurement was performed on AF19-43D (Figure 7), with approximately 20 measurement points for each profile, and two cross-section profiles on AF19-44A, with approximately 20 measurement points per profile (Figure 4). Based on the spatial distribution of pseudotachylyte and its relationship with the surrounding fractured granite, it is preliminarily concluded that the host rock on both sides of the pseudotachylyte vein is the parent rock of the melt.

Major minerals, such as plagioclase and a small amount of amphibole in the parent rock, as well as plagioclase clasts in the pseudotachylyte, were analyzed using electron probe major element analysis (Figures 12 and 13).

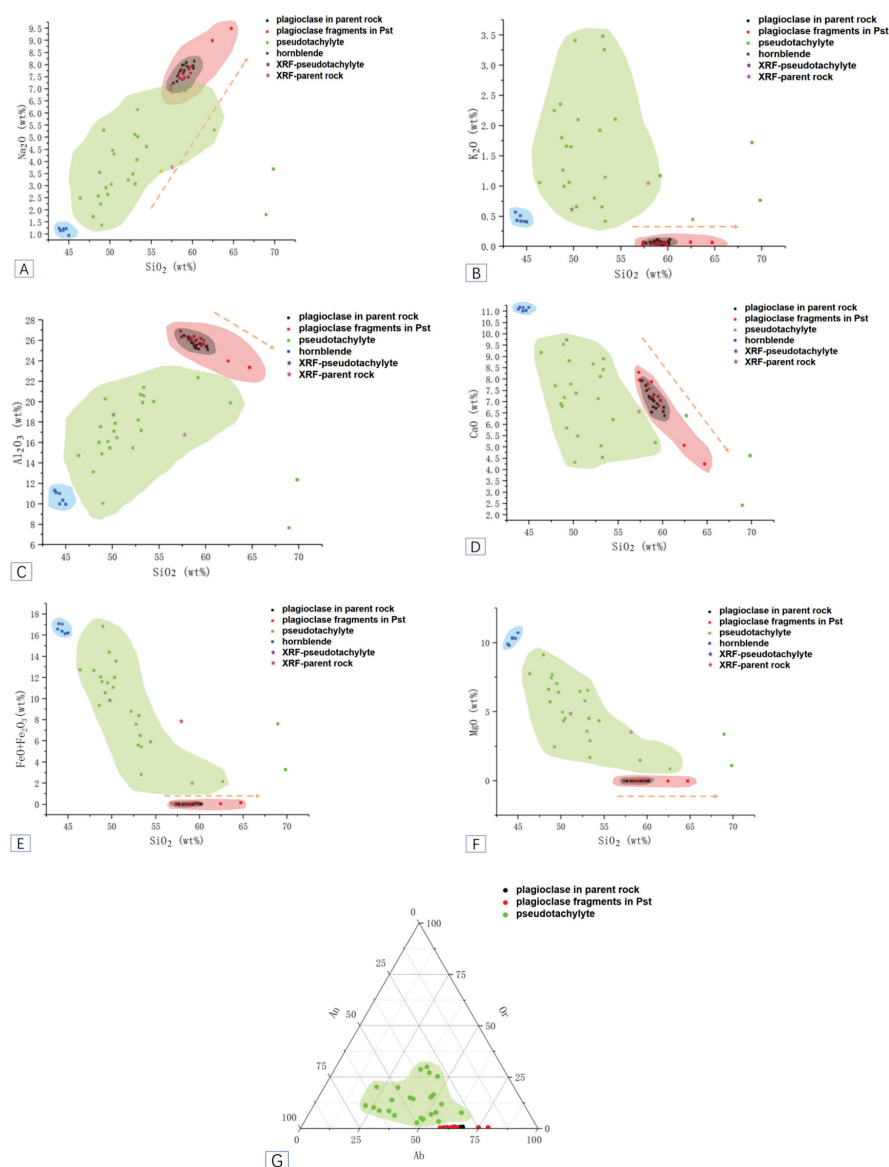


Figure 12. The results of electron microprobe and XRF for pseudotachylyte in AF19-43D. (A–F) Harker diagrams correlating Na₂O, K₂O, Al₂O₃, CaO, FeO, and MgO versus SiO₂ content, respectively. The red fields designate plagioclase fragment compositions within pseudotachylyte, black fields delineate parent rock plagioclase compositions, green fields represent pseudotachylyte compositions, and blue fields correspond to hornblende phase chemistries. (G) Ternary phase diagram of the feldspar system (Ab–An–Or).

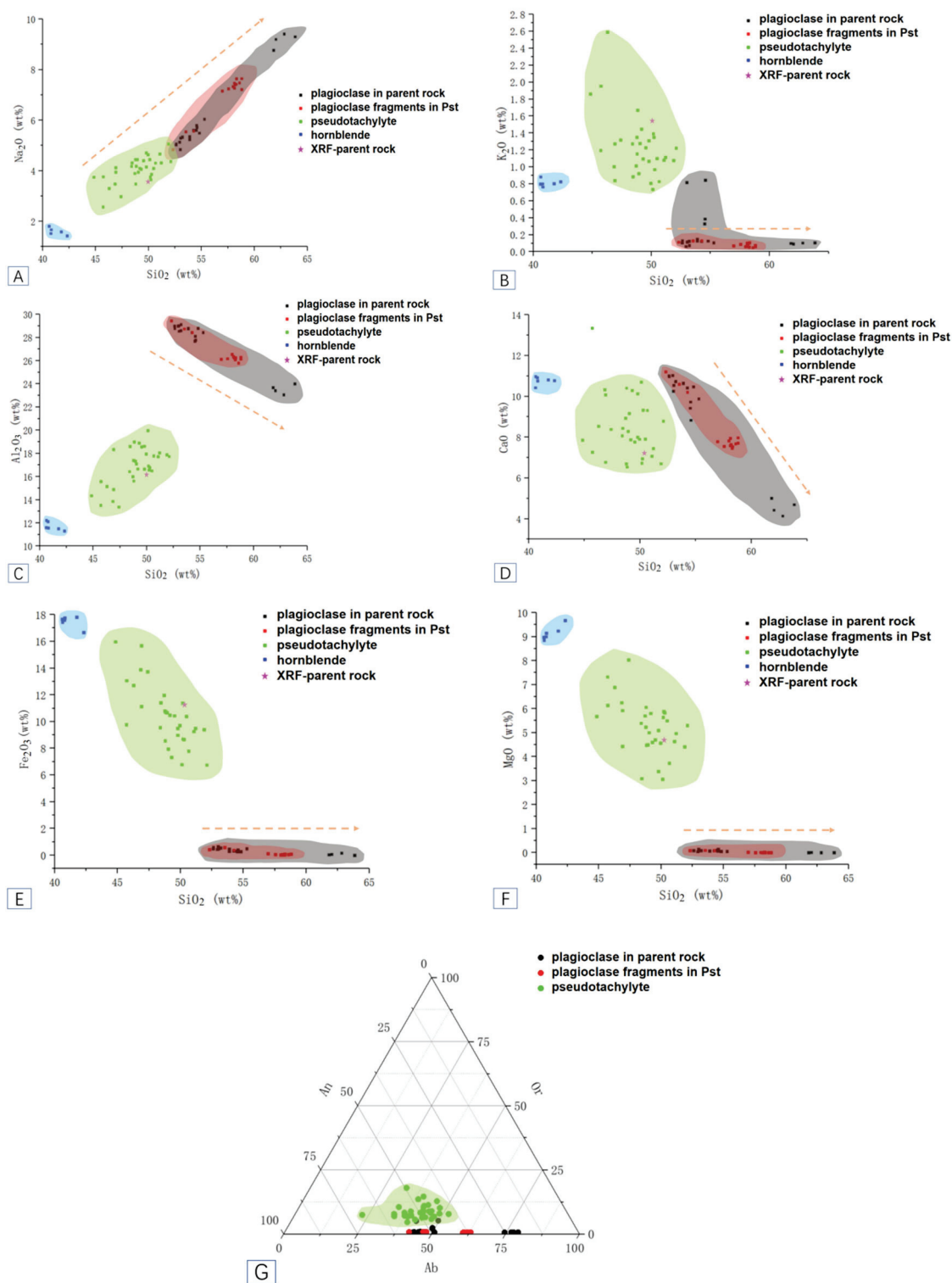


Figure 13. The results of electron microprobe and XRF for pseudotachylyte in AF19-44A. (A–F) Harker diagrams correlating Na₂O, K₂O, Al₂O₃, CaO, FeO, and MgO versus SiO₂ content, respectively. The red fields designate plagioclase fragment compositions within pseudotachylyte, black fields delineate parent rock plagioclase compositions, green fields represent pseudotachylyte compositions, and blue fields correspond to hornblende phase chemistries. (G) Ternary phase diagram of the feldspar system (Ab–An–Or).

Evidence from petrochemical analysis, combined with XRF results (Table 3), indicates the genetic relationship between pseudotachylyte and host rocks.

Table 3. The results of XRF analysis (wt%).

Sample	SiO ₂	TiO ₂	Al ₂ O ₃	Fe ₂ O ₃	FeO	MnO	MgO	CaO	Na ₂ O	K ₂ O	P ₂ O ₅	LOI	Total
AF19-43D1	56.67	0.87	17.16	3.63	5.22	0.13	3.48	6.30	3.52	1.23	0.21	0.70	99.12
AF19-43D2	49.92	0.94	18.87	3.63	6.12	0.20	4.70	9.17	4.70	0.63	0.25	0.94	99.02
AF19-43D	57.63	0.78	16.49	7.97	-	0.12	3.31	6.52	3.71	1.06	0.20	1.55	99.34
AF19-44A	50.62	2.17	15.72	11.69	-	0.15	4.55	7.14	3.50	1.51	0.32	2.26	99.64

AF19-43D1: pseudotachylyte; AF19-43D2: residual restite of host rock; AF19-43D: parent rock, and AF19-44A: parent rock.

Plagioclase fragments within pseudotachylyte exhibit highly consistent evolutionary trends in major elements (Na₂O, CaO, and Al₂O₃) with wall rock plagioclase. For instance, a positive correlation was observed between Na₂O and SiO₂ (Figures 12A and 13A). Negative correlations were observed between CaO and Al₂O₃ and SiO₂ (Figures 12C,D and 13C,D). Overlapping An values (An33–41 for pseudotachylyte versus An20–54 for wall rocks) indicate direct inheritance from wall rock minerals.

The matrix of pseudotachylyte shows significantly higher K₂O content than the parent rocks, plagioclase and hornblende (Figures 12D and 13B), which is attributed to potassium feldspar decomposition in the parent rocks. FeO and MgO concentrations in the matrix exceed those in parent rock plagioclase but remain lower than in hornblende (Figures 12 and 13), indicating preferential incorporation of Fe-Mg components from mafic minerals (e.g., hornblende, biotite) during melting.

Intermediate Na₂O and Al₂O₃ levels (between parent rock plagioclase and hornblende) and CaO consistency with plagioclase suggest multi-mineral phase mixing during melting.

Pseudotachylyte whole-rock SiO₂ content aligns with intermediate rocks (e.g., AF19-43D), while melt residues exhibit Si depletion and Na-Al-Ca enrichment. This reflects selective element migration during partial melting.

Fe-Mg enrichment and Na-Al-Ca depletion: Pseudotachylyte matrix is enriched in Fe-Mg, whereas residual components show elevated Na, Al, and Ca (Figures 12 and 13), consistent with melt-residue compositional complementarity.

Matrix mineral assemblages display diffusion toward central An values (e.g., An12–47 in AF19-43D; An33–66 in AF19-44A), indicating homogenization and re-equilibration of plagioclase during melting (Figures 12G and 13G).

Whole-rock compositional consistency and Fe-Mg-SiO₂ complementary trends indicate pseudotachylyte as an autochthonous product of host rock (granitic cataclasite) melting.

The pseudotachylyte's chemical inheritance from host rocks, matrix element partitioning, and whole-rock complementary evolution collectively demonstrate its formation via in situ partial melting of granitic cataclasite under transient high-temperature/pressure conditions in seismogenic fault zones. This process involves selective melting of parent rock minerals (e.g., K-feldspar, hornblende) and rapid quenching, preserving critical records of seismic slip dynamics.

4. Discussion

The depth distribution of shallow-source earthquakes in the crust is controlled by the strength of fault zones and the stability of sliding friction [18]. Destructive earthquakes generally occur in the regions where the lithospheric strength of the crust is the highest, i.e., the brittle–ductile transition zone [15,16,50,51]. The brittle–ductile transition zone typically lies between depths of 10 and 15 km in the crust, where the temperature ranges from approximately 250 °C to 450 °C. This zone is a semi-brittle deformation domain, where quartz can undergo plastic deformation, while feldspar primarily deforms through brittle

fracture. The granite in the brittle–ductile transition zone exhibits unstable sliding [52], and its strength is controlled by temperature, pressure, fluids, and other conditions [11–14].

The coseismic processes triggered by unstable sliding and high-speed frictional heat generation may be the main mechanisms for the formation of pseudotachylyte. During the coseismic rupture phase, due to the sharp increase in strain rate, the depth of the brittle–ductile transition zone in the fault zone will extend deeper [1]. In the interseismic period, the ductile shear zone, which is in a plastic deformation state, can undergo brittle failure during the coseismic process, forming cataclasites and pseudotachylytes. During the post-seismic relaxation phase, the strain rate gradually decreases, returning to the background values of the normal region. The depth of the brittle–ductile transition zone will gradually become shallower, and at this time, semi-brittle non-steady-state flow becomes the main deformation feature of the fault. This is accompanied by hydrolysis of feldspar and static recrystallization, dynamic recrystallization, and other plastic deformation processes of quartz and other minerals, while the fractures generated by the earthquake begin to heal [17,18,53].

This paper analyzes the microstructures of granite gneiss, mylonite, fractured rock, and pseudotachylyte, providing geological evidence for the deformation processes during each stage of the earthquake cycle.

4.1. Constraints of Microstructure and Rock Chemical Composition on the Genesis of Pseudotachylyte

The microstructures indicate that the rocks collected from the Anning River fault zone underwent deformation during different stages, including interseismic plastic flow, coseismic brittle fracturing, pseudotachylyte formation, and even post-seismic relaxation and fracture healing.

The deformation of mylonite represents interseismic plastic deformation. Mylonite develops an S-C fabric, with feldspar showing a porphyroclastic texture, while quartz and mica are significantly flattened, elongated, and oriented, forming foliations. The pervasive occurrence of feldspar twins and the undulose extinction of quartz indicate that the rock started to be stress-controlled, and high dislocation densities began to develop within the quartz grains. The fractured structures in feldspar grains, as well as the fine-grained quartz particles developing subgrains, grain rotation, and grain boundary migration, suggest that while quartz started to undergo dislocation creep and other plastic deformations, feldspar deformed mainly by brittle fracturing.

The deformation of fractured rocks and pseudotachylytes represents coseismic rupture. The primary deformation characteristic is the presence of fractured feldspar and other minerals, with angular or sub-angular clasts, and fractures filled by fine-grained quartz aggregates. Based on the EBSD analysis, the deformation temperature of the fractured rocks is about 250–350 °C, which corresponds to the brittle–ductile transition temperature (depth) range of the Anning River fault zone.

In sample AF21-4-2, multiple generations of pseudotachylytes are visible. The first generation of veins, formed earlier, is not related to the current coseismic rupture. The second and third generations of veins were both formed during the coseismic rupture phase. After the second-generation vein formation, the viscosity of the melt temporarily restored the fault strength (which may explain why the clasts in the second-generation vein are larger than those in other generations). After the fault was cemented, sliding friction did not stop but continued along the adjacent pre-existing fault surface, forming the third-generation veins.

Samples AF19-43G and AF19-44C all show fine-veined pseudotachylytes along foliations. The localization of plastic strain leads to crystal growth, recrystallization, and crystal preferred orientation in the pseudotachylyte matrix. The clasts remaining in the pseudotachylyte may also undergo plastic deformation. These features correspond to the

first-generation vein in AF21-4-2, indicating that multiple seismic events occurred in the Anning River fault zone, and each coseismic pseudotachylyte formation underwent later geological deformation and transformation, eventually becoming the “mother rock” in the next earthquake event.

In samples AF19-43A and AF19-43D, localized devitrification of the pseudotachylyte is evident. Due to the glassy or cryptocrystalline matrix in the pseudotachylyte, the pseudotachylyte is especially prone to recrystallization under fluid (water) influence. Newly formed minerals may partially or completely cover the original microstructure. These observations suggest that pseudotachylytes are difficult to preserve after formation and can be easily modified in later geological processes, which may be one of the reasons why pseudotachylytes are rarely found or reported.

In the fractured rock, the feldspar phenocrysts' fractures are healed by quartz veins, indicating post-seismic relaxation. Fine-grained quartz particles exhibit elongation and orientation, showing plastic flow deformation, subgrain development, grain boundary migration, and recrystallization.

The chemical composition analysis of pseudotachylyte and its host rocks reveals critical insights into the melting behavior of minerals during coseismic frictional processes. The comparative study of major elements (e.g., Na_2O , K_2O , Al_2O_3 , CaO , FeO , and MgO) and SiO_2 content demonstrates that amphibole, biotite, and feldspar were preferentially melted under transient high-temperature/pressure conditions, while quartz remained largely unmelted and was preserved as clasts within the pseudotachylyte matrix. This selective melting pattern indicates non-equilibrium melting, where thermodynamically unstable phases (e.g., hydrous minerals like hornblende and biotite) and alkali feldspars decomposed rapidly, whereas quartz resisted melting.

The low SiO_2 content in the pseudotachylyte melt (e.g., 49.92–57.63 wt.% in AF19-43D samples) further supports limited quartz participation in melting. Instead, quartz fragments were mechanically incorporated into the melt as relics, consistent with the observed clast-rich microstructure. Notably, the pseudotachylyte matrix exhibits chemical complementarity with the host rock residues: elevated FeO and MgO (derived from mafic minerals like hornblende) and enriched K_2O (from K-feldspar breakdown) in the melt contrast with residual Na-Al-Ca enrichment in the restite (e.g., AF19-43D2: $\text{CaO} = 9.17$ wt.%, $\text{Na}_2\text{O} = 4.70$ wt.%). This dichotomy aligns with partial melting processes, where melt and residue compositions diverge but retain genetic links to the parent rock.

The inheritance of plagioclase chemistry (An_{33-41} in pseudotachylyte vs. An_{20-54} in host rocks) and the overlapping trends in Harker diagrams (e.g., Na_2O - SiO_2 : positive correlation; CaO - SiO_2 : negative correlation) underscore a local origin for the melt. Moreover, the spatial distribution of pseudotachylyte veins—intimately associated with fractured granite and showing sharp, non-reacted contacts—rules out an external injection mechanism. The homogenization of plagioclase compositions (e.g., An_{12-47} in AF19-43D) and the ternary feldspar system (Ab-An-Or) equilibration further attest to rapid melting–quenching cycles, preserving disequilibrium signatures.

In summary, the geochemical coherence between pseudotachylyte and host rocks, coupled with microstructural evidence (e.g., clast–melt relationships), conclusively supports an in situ melting genesis within seismogenic faults. The pseudotachylyte formed via frictional heating-induced selective melting of hydrous and alkali-rich phases, while refractory minerals (e.g., quartz) survived as clasts, leaving a distinct chemical imprint of coseismic slip dynamics.

4.2. Model of Genesis Pseudotachylyte and Deformation During Seismic Period Cycle in the Anning River Fault Zone

Through the study of the microstructures, deformation fabrics, and chemical compositions of pseudotachylyte and its associated fractured rocks and mylonites in the Anning River fault zone, the process of pseudotachylyte formation can be summarized in the following stages (Figure 14):

1. **Pre-peak Friction Strength of the Fault:** Before the fault reaches its peak friction strength, the friction within the fault zone is relatively high, resulting in the occurrence of numerous coseismic micro-fractures. Multiple sets of high-angle brittle fractures extend along the fault plane, forming fractured rocks associated with coseismic rupture. During this phase, quartz in the parent rock begins to fragment. The temperature rise induced by coseismic sliding causes the quartz particles in the parent rock to develop high dislocation densities, which leads to dislocation creep in the quartz (manifested by undulose extinction in quartz grains).
2. **Fault Reaching Peak Friction Strength:** After the fault reaches its peak friction strength, rapid sliding along the fault plane occurs. Due to the frictional shear heating during coseismic movement, the temperature continuously increases, causing the rock to melt. The molten material is distributed along the fault plane and injected into the fractures of the host rocks (e.g., sample AF19-43D). During this phase, the sustained increase in temperature promotes the development of dislocations and the reorganization of crystal lattices in the quartz particles in the parent rock. This leads to dynamic recrystallization of quartz, forming subgrains and new crystal grains, with grain size controlled by both high stress and high strain rates.
3. **Fault Sliding Weakening Phase:** In the weakening phase of fault sliding, coseismic shear continues for several tens of seconds, and the molten material keeps forming, leading to the main body of the pseudotachylyte (e.g., sample AF19-44A, AF21-4-3). At this stage, in the microshear zone formed concurrently with the unstable fault sliding, quartz particles and subgrain aggregates that have already undergone recrystallization continue to undergo dynamic recrystallization through grain boundary sliding mechanisms.
4. **Cessation of Coseismic Shear Sliding:** After the cessation of coseismic shear sliding, the molten material begins to cool. However, the high temperatures generated by the molten material continue to heat the host rock, causing quartz in the host rock to continue undergoing plastic deformation. As stress and strain rates begin to decrease, quartz begins to undergo static recrystallization, forming a bubble-like structure and creating triple junctions (where the boundary angles between quartz grains are 120°), which restricts grain growth. Most of the fractures in the host rock are healed during this process.

From this model, it can be seen that pseudotachylyte formed in the brittle–ductile transition zone of the fault is the result of a non-steady-state rheological process. During this process, the ability of quartz to capture the transient non-steady-state rheology records a series of events: fracturing, dynamic recrystallization, and static recrystallization. This also explains why most pseudotachylytes exhibit the systematic feature of containing both quenched melt phases and recrystallized fine-grained aggregates simultaneously.

This model is cyclical and repetitive. For each seismic cycle, the pseudotachylyte generated during coseismic rupture undergoes metamorphic and deformational alteration during the post-seismic relaxation phase and interseismic phase. This results in the formation of fine vein pseudotachylyte with plastic deformation overprinted along the foliation, as mentioned earlier. In the next seismic cycle, these deformed products serve as the parent rock, undergoing

further seismic modification. The multiple generations of pseudotachylyte veins present in sample AF21-4-2 provide a good correspondence to this model.

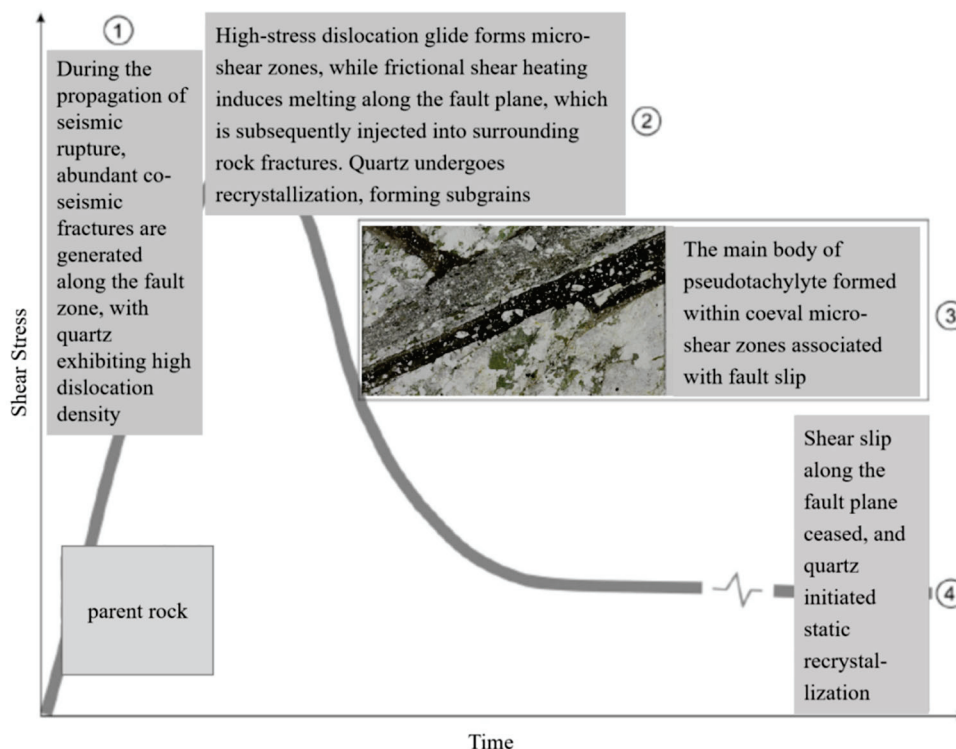


Figure 14. A model of the pseudotachylyte and quartz microstructure during the seismic fault sliding process (modified from [25]).

4.3. Limitations and Future Directions

While this study acknowledges the alteration of pseudotachylytes post-formation, additional limitations—such as sampling constraints and uncertainties in temperature estimation—warrant explicit discussion. Sampling is restricted to specific fault zones (e.g., Anning River fault zone), potentially limiting the generalizability of microstructural and geochemical conclusions to other tectonic settings with varying lithologies or stress regimes. Temperature estimates, primarily derived from quartz EBSD slip systems (250–350 °C for fractured rocks), rely on empirical relationships between crystallographic preferred orientations and thermobarometric frameworks. However, these correlations may not fully account for strain rate effects or fluid-induced recrystallization, which could obscure original deformation temperatures.

Future research could mitigate these gaps through in situ thermometry techniques (e.g., Raman spectroscopy of carbonaceous phases or laser microprobe analysis of fluid inclusions), which provide direct constraints on peak temperatures during pseudotachylyte formation. Additionally, integrating multi-scale structural transects across diverse fault segments (e.g., from shallow brittle to deeper ductile domains) would clarify how temperature–depth relationships vary spatially and temporally within seismogenic zones. Such approaches would enhance the robustness of thermomechanical models linking microstructural observations to macroscale seismic processes, ensuring more reliable interpretations of fault zone evolution and earthquake dynamics.

5. Conclusions

This paper focuses on pseudotachylytes, cataclasites, and mylonites exposed in the brittle–ductile transition zone along the Anning River fault zone, conducting microstruc-

tural analysis, rock chemical composition measurements, mineral deformation, and fabric orientation analysis. The main results are as follows:

The mylonites exposed along the Anning River fault zone represent deformation in the brittle–ductile transition zone during the interseismic period. Cataclasites represent coseismic fractures in the brittle–ductile transition zone. The quartz fabric from EBSD showed deformation temperature in the cataclasites is between 250 and 350 °C, corresponding to a depth range of 15–20 km. Three kinds of black pseudotachylyte veins presented in the cataclasites and mylonites represent frictional melting at coseismic slip.

Rock chemical analysis was performed on typical straight vein-type pseudotachylytes and their host rocks, showing that pseudotachylytes were produced directly from the frictional heating process of granitic host rocks during coseismic slip.

A simple model was proposed for the deformation characteristics and mechanisms of fault zone rocks during the coseismic, post-seismic relaxation, and interseismic periods. This model provides a reference for understanding the deformation processes of rocks at the depth of the seismic source in earthquake fault zones.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/app15115870/s1>. Section S1: Geological Setting.

Author Contributions: Y.Z. contributed to the conception and design of this study. W.D. conducted the experiments. W.D. and Y.Z. wrote the manuscript. H.L. and X.M. guided the use of instruments and discussed the results of SEM and EBSD. S.L. participated in sample preparations and illustration. All authors contributed to the field work, collection of samples, manuscript revision, read and approved the submitted version. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the National Natural Science Foundation of China (Grant No. U2139202) and the National Key Research and Development Program of China (Grant No. 2023YFC3012001).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article/Supplementary Materials, further inquiries can be directed to the corresponding authors.

Acknowledgments: We would like to express our sincere gratitude to the editors of this journal for their kind assistance and professional support throughout the review process. We also deeply appreciate the constructive and insightful comments provided by the two anonymous reviewers, which greatly improved the quality of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Campbell, L.R.; Menegon, L.; Fagereng, A.; Pennacchioni, G. Earthquake nucleation in the lower crust by local stress amplification. *Nat. Commun.* **2020**, *11*, 1322. [CrossRef] [PubMed]
2. Brace, W.F.; Kohlstedt, D.L. Limits on lithospheric stress imposed by laboratory experiments. *J. Geophys. Res.* **1980**, *85*, 6248–6252. [CrossRef]
3. Chernak, L.J.; Hirth, G.; Selverstone, J.; Tullis, J. Effect of aqueous and carbonic fluids on the dislocation creep strength of quartz. *J. Geophys. Res.* **2009**, *114*, B04201. [CrossRef]
4. Gueydan, F.; Leroy, Y.M.; Jolivet, L. Mechanics of low-angle extensional shear zones at the brittle-ductile transition. *J. Geophys. Res.* **2004**, *109*, B12407. [CrossRef]
5. Holyoke, C.W.; Tullis, J. The interaction between reaction and deformation: An experimental study using a biotite+plagioclase+quartz gneiss. *J. Metamorph. Geol.* **2006**, *24*, 743–762. [CrossRef]
6. Ikari, M.J.; Niemeijer, A.R.; Marone, C. The role of fault zone fabric and lithification state on frictional strength, constitutive behavior, and deformation microstructure. *J. Geophys. Res.* **2011**, *116*, B08404. [CrossRef]
7. Kohlstedt, D.L.; Evans, B.; Mackwell, S.J. Strength of the lithosphere: Constraints imposed by laboratory experiments. *J. Geophys. Res.* **1995**, *100*, 17587–17602. [CrossRef]

8. Montési, L.G.J.; Zuber, M.T. A unified description of localization for application to large-scale tectonics. *J. Geophys. Res.* **2002**, *107*, 1–21. [CrossRef]
9. Sibson, R.H. Continental fault structure and the shallow earthquake source. *J. Geol. Soc.* **1983**, *140*, 741–767. [CrossRef]
10. Tullis, J.; Yund, R.A. Chapter 4 The Brittle-Ductile Transition in Feldspar Aggregates: An Experimental Study. *Int. Geophys.* **1992**, *51*, 89–117. [CrossRef]
11. Trepmann, C.A.; Stockhert, B. Quartz microstructures developed during non-steady state plastic flow at rapidly decaying stress and strain rate. *J. Struct. Geol.* **2003**, *25*, 2035–2051. [CrossRef]
12. Trepmann, C.A.; Stockhert, B.; Dorner, D.; Moghadam, R.H.; Küster, M.; Roller, K. Simulating coseismic deformation of quartz in the middle crust and fabric postseismic stress relaxation: An experimental study. *Tectonophysics* **2007**, *442*, 83–104. [CrossRef]
13. Trepmann, C.A.; Stockhert, B. Short-wavelength undulatory extinction in quartz recording coseismic deformation in the middle crust: An experimental study. *Solid Earth* **2013**, *4*, 263–276. [CrossRef]
14. Trepmann, C.A.; Chen, H.; Felix, H.; Katharina, D.; Christian, S. Recrystallization of quartz after low-temperature plasticity: The record of stress relaxation below the seismogenic zone. *J. Struct. Geol.* **2017**, *95*, 77–92. [CrossRef]
15. Tse, S.T.; Rice, J.R. Crustal earthquake instability in relation to the depth variation of frictional slip properties. *J. Geophys. Res.* **1986**, *91*, 9452–9472. [CrossRef]
16. Scholz, C.H. The brittle-pl and depth of seismic faulting. *Geol. Rundsch.* **1988**, *77*, 319–328. [CrossRef]
17. Dai, W.H.; Zhou, Y.S. Deformation of the Brittle-Plastic Transition Zone at the Post-Seismic Relaxation Period: A Case Study of the Red River Fault. *Seismol. Geol.* **2019**, *41*, 996–1011, (In Chinese with English Abstract). [CrossRef]
18. Zhou, Y.S.; Han, L.; Jing, C.; He, C.-R.; Dang, J. The Rheological Structures of Brittle-Plastic Transition in Longmenshan Fault Zone and Seismogenic Mechanism of Wenchuan Earthquake. *Seismol. Geol.* **2014**, *36*, 882–895, (In Chinese with English Abstract). [CrossRef]
19. Campbell, L.R.; Menegon, L. Transient high strain rate during localized viscous creep in the dry lower continental crust (Lofoten, Norway). *J. Geophys. Res. Solid Earth* **2019**, *124*, 10240–10260. [CrossRef]
20. Sibson, R.H. Transient discontinuities in ductile shear zones. *J. Struct. Geol.* **1980**, *2*, 165–174. [CrossRef]
21. McKenzie, D.; Brune, J.N. Melting on fault planes during large earthquakes. *Geophys. J. Int.* **1972**, *29*, 65–78. [CrossRef]
22. Sibson, R.H. Generation of Pseudotachylite by Ancient Seismic Faulting. *Geophys. J. Int.* **1975**, *43*, 775–794. [CrossRef]
23. Cowan, D.S. Do faults preserve a record of seismic slip? A field geologist's opinion. *J. Struct. Geol.* **1999**, *21*, 995–1001. [CrossRef]
24. Bestmann, M.; Pennacchioni, G.; Frank, G.; Goken, M.; de Wall, H. Pseudotachy in muscovite-bearing quartzite: Coseismic friction-induced melting and plastic deformation of quartz. *J. Struct. Geol.* **2011**, *33*, 169–186. [CrossRef]
25. Bestmann, M.; Pennacchioni, G.; Nielsen, S.; Goken, M.; de Wall, H. Deformation and ultrafine dynamic recrystallization of quartz in pseudotachylite-bearing brittle faults: A matter of a few seconds. *J. Struct. Geol.* **2012**, *38*, 21–38. [CrossRef]
26. Sibson, R.H.; Toy, V.G. The Habitat of Fault-Generated Pseudotachylite: Presence vs. Absence of Friction-Melt. In *Radiated Energy and the Physics of Faulting*; Abercrombie, R., McGarr, A., Di Toro, G., Kanamori, H., Eds.; American Geophysical Union: Washington, DC, USA, 2006; pp. 153–166. [CrossRef]
27. Wang, L.; Barbot, S. Excitation of San Andreas tremors by thermal instabilities below the seismogenic zone. *Sci. Adv.* **2020**, *6*, eabb2059. [CrossRef]
28. Papa, S.; Pennacchioni, G.; Menegon, L.; Thielmann, M. High-stress creep preceding coseismic rupturing in amphibolite-facies ultramylonites. *Earth Planet. Sci. Lett.* **2020**, *541*, 116271. [CrossRef]
29. Song, B.R.; Johnson, S.E.; Song, W.J.; Gerbi, C.C.; Yates, M.G. Coseismic damage runs deep in continental strike-slip faults. *Earth Planet. Sci. Lett.* **2020**, *539*, 116241. [CrossRef]
30. Magloughlin, J.F. Microstructural and chemical changes associated with cataclasis and frictional melting at shallow crustal levels: The cataclasite-pseudotachylite connection. *Tectonophysics* **1992**, *204*, 243–260. [CrossRef]
31. Kirkpatrick, J.D.; Shipton, Z.K.; Persano, C. Pseudotachylites: Rarely generated, rarely preserved, or rarely reported? *Bull. Seismol. Soc. Am.* **2009**, *99*, 382–388. [CrossRef]
32. Kirkpatrick, J.D.; Rowe, C.D. Disappearing ink: How pseudotachylites are lost from the rock record. *J. Struct. Geol.* **2013**, *52*, 183–198. [CrossRef]
33. Moecher, D.P.; Steltenpohl, M.G. Petrological evidence for co-seismic slip in extending middle-lower continental crust: Heier's zone of pseudotachylite, north Norway. *Geol. Soc. Lond. Spec. Publ.* **2011**, *359*, 169–186. [CrossRef]
34. Clerc, A.; Renard, F.; Austrheim, H.; Jamtveit, B. Spatial and size distributions of gamets grown in a pseudotachylite generated during a lower crust earthquake. *Tectonophysics* **2018**, *733*, 159–170. [CrossRef]
35. Orlandini, O.F.; Mahan, K.H.; Williams, M.J.; Regan, S.P.; Mueller, K.J. Evidence for deep crustal seismic rupture in a granulite-facies, intraplate, strike-slip shear zone, northern Saskatchewan, Canada. *Geol. Soc. Am. Bull.* **2019**, *131*, 403–425. [CrossRef]
36. Janssen, C.; Wirth, R.; Rybacki, E.; Naumann, R.; Kemnitz, H.; Wenk, H.R.; Dresen, G. Amorphous material in SAFOD core samples (San Andreas Fault): Evidence for crush-origin pseudotachylites? *Geophys. Res. Lett.* **2010**, *37*, L01302. [CrossRef]

37. Pec, M.; Stünitz, H.; Heilbronner, R. Semi-brittle deformation of granitoid gouges in shear experiments at elevated pressures and temperatures. *J. Struct. Geol.* **2012**, *38*, 200–221. [CrossRef]
38. Pec, M.; Stünitz, H.; Heilbronner, R.; Drury, M.; de Capitani, C. Origin of pseudotachylites in slow creep experiments. *Earth Planet. Sci. Lett.* **2012**, *355–356*, 299–310. [CrossRef]
39. Marti, S.; Stünitz, H.; Heilbronner, R.; Plümper, O. Amorphous material in experimentally deformed mafic rock and its temperature dependence: Implications for fault rheology during aseismic creep and seismic rupture. *J. Struct. Geol.* **2020**, *138*, 104129. [CrossRef]
40. Miao, S.Q.; Zhou, Y.S.; Ma, X.; Dang, J.X. Lithologic Structure of the Anninghe Fault Zone: Constraints From High-Pressure Wave Velocity Experiments. *Front. Earth Sci.* **2022**, *10*, 853474. [CrossRef]
41. Brunel, M.; Maliakov, Y.K. Crystal slip systems in quartz under varying temperature conditions. *J. Struct. Geol.* **1972**, *5*, 123–135.
42. Hara, I.; Takeda, K.; Kimura, T. Quartz fabric transitions in shear zones. *Tectonophysics* **1973**, *18*, 287–301.
43. Mainprice, D.; Bouchez, J.L.; Casey, M. Crystallographic preferred orientations of quartz in mylonites. *J. Geophys. Res.* **1986**, *91*, 3857–3869.
44. Schmid, S.M.; Paterson, M.S.; Boland, J.N. High temperature flow and dynamic recrystallization in Carrara Marble. *Tectonophysics* **1980**, *65*, 245–280. [CrossRef]
45. Toy, V.G.; Prior, D.J.; Norris, R.J. Quartz fabrics in the Alpine Fault mylonites: Influence of strain rate and temperature. *J. Struct. Geol.* **2008**, *30*, 1172–1187. [CrossRef]
46. Ji, S.C. Experimental studies on quartz plasticity under non-coaxial deformation. *Acta Geophys. Sin.* **1988**, *31*, 213–225. (In Chinese with English Abstract)
47. Heidelbach, F.; Stretton, I.; Langenhorst, F. Slip system activity in quartz during plastic deformation. *J. Geophys. Res.* **2000**, *105*, 13345–13356. [CrossRef]
48. Ábalos, B.; Puelles, P.; Fernández-Armas, S.; Sarrionandia, F. EBSD microfabric study of pre-Cambrian deformations recorded in quartz pebbles from the Sierra de la Demanda (N Spain). *J. Struct. Geol.* **2011**, *33*, 500–518. [CrossRef]
49. Ni, J.L.; Liu, J.L. Composite fabric models in ductile shear zones: Insights from quartz CPO evolution. *Chin. J. Geol.* **2014**, *49*, 756–769. (In Chinese with English Abstract)
50. Zhou, Y.S.; He, C.R. The rheological structures of crust and mechanics of high-angle reverse fault slip for Wenchuan Ms8.0 earthquake. *Chin. J. Geophys.* **2009**, *52*, 474–484. (In Chinese with English Abstract)
51. Zhang, Y.Y.; Zhou, Y.S. Strength and deformation mechanisms of fault brittle-plastic transition zones with the influence of fluids and strain rates. *Seismol. Geol.* **2012**, *34*, 172–194. (In Chinese with English Abstract) [CrossRef]
52. Zhang, L.; He, C.R. Frictional properties of phyllosilicate-rich mylonite and conditions for the brittle-ductile transition. *J. Geophys. Res. Solid Earth* **2016**, *121*, 3017–3047. [CrossRef]
53. Niu, L.; Zhou, Y.S.; Yao, W.M.; Ma, X.; He, C.R. An experimental study on the transient creep of granite. *Seismol. Geol.* **2021**, *43*, 20–35. (In Chinese with English Abstract) [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Field Test and Numerical Simulation Study of Bearing Characteristics of Combined Post-Grouted Piles for Railway Bridges

Runze Zhang ¹, Zilong Guo ², Weiming Gong ¹ and Zhihui Wan ^{1,2,*}

¹ School of Civil Engineering, Southeast University, Nanjing 211189, China; 230199011@seu.edu.cn (R.Z.); wmgong@seu.edu.cn (W.G.)

² College of Transportation Engineering, Nanjing Tech University, Nanjing 211816, China; guozilong0217@njtech.edu.cn

* Correspondence: wanzhihui@njtech.edu.cn

Abstract: Utilizing the results of static load tests using the self-balancing method on two large-diameter bored piles from the Huaiyang Left Line Special Bridge Project of the Lianyungang–Zhenjiang Railway, this study aims to investigate the effect of combined tip-and-side post-grouting on the bearing characteristics of post-grouted piles in railway bridges. The difference in bearing performance between individual piles before and after grouting was evaluated using a comparative analysis. The results show that the bearing capacity of the pile foundations is greatly increased by combined tip-and-side post-grouting. In particular, following grouting, a single pile’s maximum bearing capacity rises from 32.99% to 38.42%. The combined post-grouting produces a compressed grout that enhances the mechanical characteristics of the pile–soil contact, resulting in a significant increase in side resistance all the way along the pile. The combined post-grouting also optimizes the performance of the tip resistance, resulting in a more rapid response as the pile tip displacement increases. Additionally, the combined post-grouting modifies the pile shaft’s load transfer mechanism by increasing the tip resistance’s contribution to the pile foundation’s ultimate bearing capacity and moving the bearing’s center of gravity closer to the pile end.

Keywords: railway bridges; bored piles; field static load test; self-balancing method; combined post-grouting; numerical simulation

1. Introduction

Super-long and large-diameter bored piles are frequently utilized in large-scale bridges due to their exceptional single-pile capacity. Nonetheless, conventional bored piles frequently confront challenges, including the formation of mud cakes along their perimeters, the accumulation of drilling residues at their termini, and the disruption of the underlying bearing stratum, all of which negatively influence their load-bearing capabilities. The quality of pile formation is crucial in determining the bearing capacity of bored piles, leading to a relatively variable vertical bearing capacity. A substantial body of research has consistently shown that the application of post-grouting methodologies to pile foundations can proficiently mitigate these aforementioned issues [1–5]. Specifically, pile end post-grouting creates an enlarged pile head at the pile bottom, which pre-compresses and reinforces the surrounding soil, significantly enhancing the single-pile bearing performance. Conversely, pile side post-grouting strengthens the soil adjacent to the pile through upward and downward seepage of the pressurized grout. The formation of a grout–soil blend at both the

pile's terminal and lateral regions effectively removes the residual slag at the base and the mud cake adhering to the pile sides. The resultant benefits of employing post-grouting techniques notably enhance the load-transfer efficiency of the pile foundation, leading to a considerable augmentation in its bearing capacity while concurrently mitigating settlement. Consequently, post-grouting techniques have found widespread application in high-rise buildings, bridges, harbors, railways, and water conservancy projects [6–11].

To investigate the bearing characteristics of post-grouting bored piles, numerous scholars have conducted extensive research. Ma et al. [10] investigated the effects of post-grouting at the pile bases by performing static load tests on six drilled piles, and they combined finite element analysis software with an exploration of how the cement grout's diffusion radius at the pile tip, along with its penetration depth, influences the pile's load-bearing capacity. Thiyyakkandi et al. [12] detailed the construction process and results of field static load tests conducted on an innovative jet-grouted precast pile in cohesionless soil conditions, and their findings revealed that, when utilizing this grouting technique, the bearing capacity of these pile foundations surpassed that of similar-sized bored piles by a factor of over 2.5 times. However, there is a paucity of reports on the application of combined tip-and-side post-grouting in railway bridge projects.

Traditional static load testing typically encompasses the accumulation load method and anchor pile method. However, in railroad bridge projects, site constraints often render both methods impractical. Furthermore, large-diameter post-grouted bored piles exhibit substantial bearing capacities, rendering traditional static load testing inadequate in providing sufficient upper loads to fulfill testing requirements. Consequently, numerous scholars have adopted the self-balancing method for static load testing in pile foundation projects with significant loading demands. Nguyen and Fellenius [13] observed superior outcomes from a self-balancing static load test conducted on two large-diameter piles that underwent post-grouting at their terminals for the Binh Le Bridge in Vietnam. Similarly, Zhou et al. [14] conducted an on-site self-balancing static load test on large-diameter post-grouted piles at the Guanyin Temple Yangtze River Bridge, showing that the post-grouting technique enhances the ultimate bearing capacity and reduces pile top settlement. In summary, the self-balancing static load test is a proven method for assessing the bearing capacity of large-diameter post-grouted piles with substantial loading requirements in field conditions. This method is especially advantageous for assessing the load-bearing capacity of bored piles that have undergone post-grouting at both their tips and sides in practical railway and bridge construction projects.

This paper focuses on the Huaiyang Left Line Bridge project of the Lianyungang–Zhenjiang Railway. Given the load requirements and site constraints, this study employed a self-balancing method for static load testing on two large-diameter bored piles installed within this project. The objective is to analyze the difference in the bearing performance of the two piles before and after the application of combined post-grouting. By comparing these results and integrating findings from numerical simulations, this study investigated the impact of combined post-grouting on the pile foundation bearing capacity and load transfer mechanisms. The results obtained from these tests offer significant insights that inform the design of pile foundations in comparable railway bridge endeavors, thereby fostering the utilization of the combined tip-and-side post-grouting methodology within the realm of foundation engineering.

2. Project Overview

2.1. Overview of Site Stratigraphy

The Lianyungang–Zhengjiang Railway's Huaiyang Left Line Bridge is situated in Fairy Town, Jiangdu District, Yangzhou City. A comprehensive analysis of the drilling data

and regional geological information revealed that the geological stratigraphy at the bridge site was relatively complex. Based on genesis and age, the strata primarily consisted of Quaternary Holocene (Q_4^{al}) and Quaternary Upper Pleistocene (Q_3^{al}) alluvial formations, which encompassed clay and sandy soils. At the bridge site in question, the geological layers were predominantly composed of silty clay and clayey materials. To meet the loading requirements of the superstructure, the project employed large-diameter piles that have undergone post-grouting as the foundational design approach. Table 1 presents the parameters of each soil layer on the test site.

Table 1. Parameters of each soil layer on the test site.

Layer No.	Soil Name	Basic Permissible Value of Load Capacity (kPa)	Standard Value of Side Resistance (kPa)	Grading of Geotechnical Construction Work
1-1	Miscellaneous fill			I
2-1	Silty clay	120	40	II
2-2	Silty clay	160	55	II
2-3	Silty clay	150	50	II
2-4	Silty clay	160	50	II
2-5	Silty clay	200	65	II
3-1	Clay	180	60	II
3-2	Clay	120	35	II
3-3	Clay	180	60	II
3-4	Clay	180	55	II

2.2. Design of Test Piles

In accordance with the design specifications, two engineering piles, designated as TP1 and TP2, were selected as test piles at Pier 85# of the site. The pertinent parameters of these two test piles are detailed in Table 2. Both piles were constructed using the rotary excavation technique, and the elastic modulus of the concrete for the test piles was 31.5 GPa. Figure 1 depicts the soil layer distribution within the pile lengths at the test location, along with the arrangement of the testing apparatus. During pile construction, measures were taken to enhance the mud casing on the pile sides and eliminate debris from the pile tips. Furthermore, to bolster the load-bearing capacity of individual piles while minimizing settlement, both test piles underwent reinforcement through the application of combined tip-and-side post-grouting methodologies. The test piles were plugged 12 h after concrete placement, and the combined post-grouting process was initiated subsequent to the completion of the pile bearing capacity test before grouting. To guarantee the quality and uniformity of the grouting, three grouting pipes were arranged for pile end grouting, and two pipes were placed for the pile side grouting. Additionally, grouting reinforcement was performed at the load box following the completion of the testing. The pressurized grout was composed of P.O.42.5 ordinary silicate cement mixed with a water–cement ratio of 0.55. The specific grouting parameters are detailed in Table 2.

Table 2. Parameters of test piles and grouting parameters.

Pile No.	Diameter (m)	Length (m)	Bearing Layer	Grouting Method	Grouting Pressure (MPa)		Water–Cement Ratio	Quantity of Cement (10^3 kg)
					Pile End	PILE SIDE		
TP1	1.5	56.5	Clay	Combined	2.7	3.8	0.55	6.69
TP2	1.5	56.5	Clay	Combined	2.5	3.3	0.55	6.58

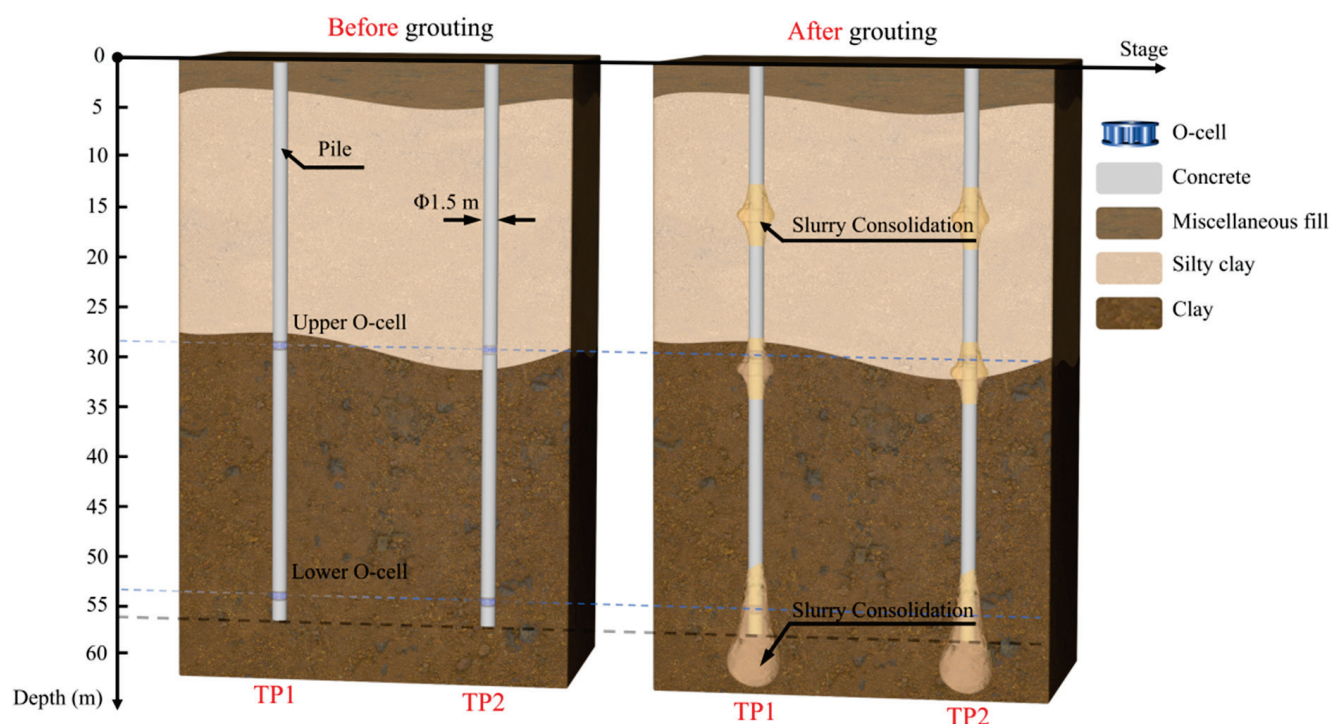


Figure 1. Distribution profile of soil layers of test site and shaft instrumentation.

3. Field Test Method

3.1. Self-Balancing Static Load Test Inspection Principle

The self-balancing method represents a novel approach for conducting static load tests on foundation piles. Different from traditional pile loading and anchor pile methods, this technique involves embedding a load box (O-cell) and reinforcing cage, which are welded together, inside the pile. The high-pressure oil pipe of the load box is then routed to the surface before pouring concrete into the pile. The load box is subsequently filled with oil and loaded using a high-pressure oil pump located on the ground. This force is transmitted to the pile, with the side resistance of the upper pile segment balancing the side and tip resistance of the lower segment to maintain the applied load. Given the considerable length of the test piles, the rated load capacity of the jack poses limitations on the use of a single load box, making its arrangement challenging. To address this issue, the double load box self-balancing test method was employed in this study. Furthermore, this method offers the following advantages: (1) It enables effective assessments of pile foundation bearing capacity before and after grouting. (2) The loading sequence of the double load box can be flexibly adjusted, allowing for the measurement of bearing capacity across different pile sections. In accordance with the design specifications, the upper load box (Upper O-cell) of test piles TP1 and TP2 was positioned 28.5 m from the pile end, while the lower load box (Lower O-cell) was located 3 m from the pile end. The equipment arrangement for the double load box self-balancing test method is depicted in Figure 2.

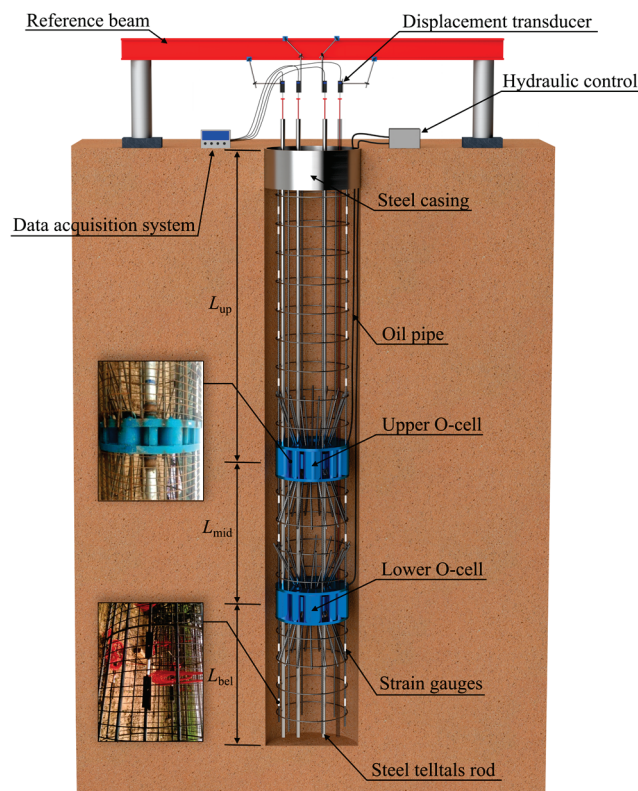


Figure 2. Schematic diagram of self-balancing test with double-level O-cells.

3.2. Loading and Unloading Method

Each test pile incorporated two annular load boxes, which were pressurized using a high-pressure oil pump and a precision pressure gauge of grade 0.4. The maximum pressure value indicated by the gauge was 60 MPa. An I32 I-beam served as the reference beam and was driven into the soil to a depth of 2 m. One end of the reference beam was hinged to the reference pile, while the other end was welded to it, with the beam spanning a length of 9 m. During the testing process, six electronic displacement meters were utilized per pile to measure pile displacement. Specifically, two meters were dedicated to measuring the upward displacement of the load box's top plate, another two to measure the downward displacement of the load box's bottom plate, and the remaining two meters measured the vertical displacement at the pile base. The test loading was conducted using the slow sustained load method, where each loading increment was set to 1/15 of the estimated ultimate load capacity, with the initial loading increment being double the subsequent levels. Similarly, unloading was performed in increments matching the loading levels, with each unloading step being three times the corresponding loading increment.

4. Analysis of Test Results

4.1. Pile Top Load–Top Displacement Curve

The curve depicting the pile top load (Q) against its corresponding top displacement (s) provides a clear and intuitive illustration of the pile foundation's bearing capacity and deformation behavior [15]. It serves as a crucial reference for engineering inspectors to assess the safety and stability of the test piles. The results obtained from the self-balancing test were converted into equivalent top load–top displacement curves, which correspond to those derived from traditional static load tests. These conversions yield the equivalent top load (Q) versus top displacement (s) curves for test piles TP1 and TP2, as illustrated in Figure 3.

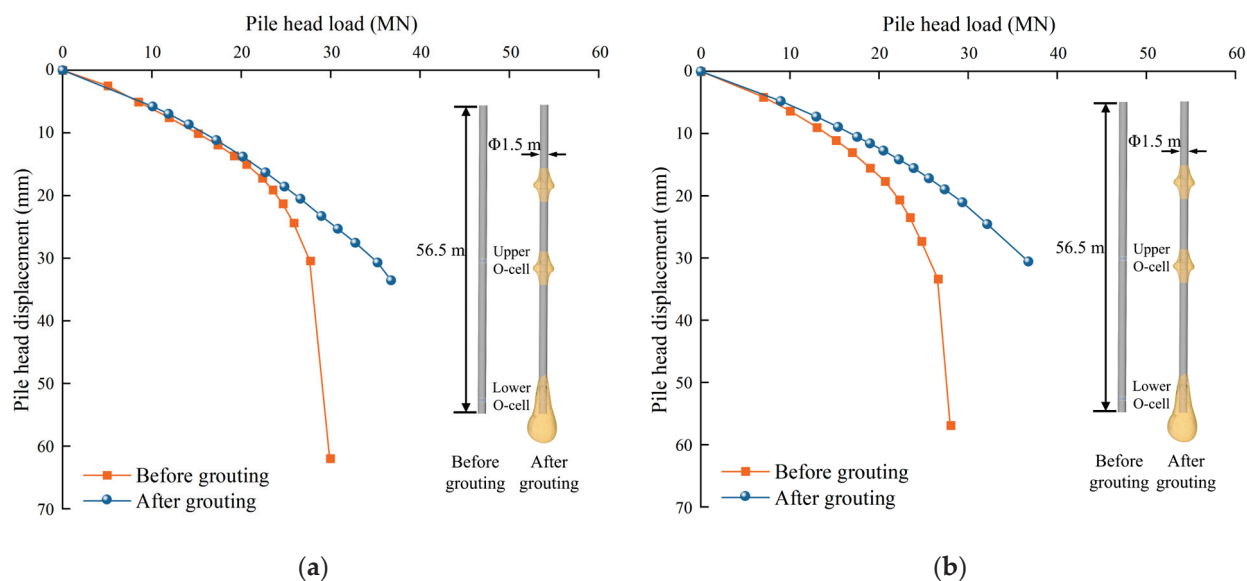


Figure 3. Equivalent conversion Q - s curves of test piles: (a) TP1; (b) TP2.

As depicted in Figure 3, the curves of both test piles exhibit notable steep drop points prior to grouting, whereas the curves after grouting display a gradual change, indicating a significant enhancement in the single-pile bearing characteristics. Under low loading conditions, the variation in pile top displacement prior to and subsequent to grouting is comparatively negligible, as the primary support for the pile top load originates from the lateral resistance of the pile, and the strengthening influence of the grout has yet to manifest itself. However, as the load gradually increases, a significant disparity emerges between the Q - s curves before and after grouting. During this phase, the pile top load is distributed along the pile shaft down to the pile tip, with the pile's lateral resistance gradually nearing its ultimate capacity, while the contribution of the pile tip resistance starts to increase progressively. Prior to grouting, soil damage occurs upon reaching the ultimate bearing capacity, resulting in a steep decline in the load–displacement curve. Conversely, following the combined grouting procedure, the grout injected into both the pile sides and pile tip significantly improves the mud layer on the pile sides and the slag accumulation at the pile tip, respectively. This reinforcement leads to an improvement in the performance of both the side and tip resistance as the load at the pile top increases. Throughout this process, the grout applied to the pile sides notably boosts the bearing capacity of the individual pile under moderate loads while also reducing the pile top displacement. As the load further increases, the reinforcing effect of the grout at the pile tip becomes increasingly prominent, significantly augmenting the bearing capacity of the individual pile under heavier loads. Notably, upon reaching the ultimate bearing capacity post-grouting, there is no abrupt, steep decline observed.

At a particular loading stage, the Q - s curve manifests a distinct point of inflection, signaling that the pile under test has attained its vertical ultimate bearing capacity. The load value recorded at the stage preceding this inflection point can be considered the pile's ultimate bearing capacity [16,17]. According to this criterion, the ultimate bearing capacities of test piles TP1 and TP2 before grouting are 27,644 kN and 26,560 kN, respectively. However, the curves of the test piles after grouting do not display any obvious inflection points; thus, the ultimate bearing capacity is taken as the last applied load, which is 36,765 kN for both TP1 and TP2. The ultimate bearing capacities of the test piles in different phases along with their corresponding pile top displacements are presented in Table 3. The ultimate bearing capacities of test piles TP1 and TP2 increase by 32.99% and 38.42%, respectively, after undergoing combined grouting. This underscores the profound

influence of combined post-grouting on the load-carrying capacity of individual piles. It is worth noting that the settlement of the pile foundation does not increase in tandem with the augmentation of its load-bearing capacity; indeed, the pile top displacement under the ultimate load condition of test pile TP2, post-combined grouting, undergoes a slight decrement. The test results demonstrate that combined post-grouting improves the bearing performance of bored piles. The cement grout injected at the pile side effectively reduces the impact of the mud skin formed during construction, thereby enhancing the exertion of side resistance. Meanwhile, the grout infused at the pile tip consolidates the slag present at the borehole's base, resulting in a substantial enhancement in the pile tip resistance's expression. Additionally, the reinforcing influence of combined post-grouting on the soil adjacent to the pile serves to further elevate the individual pile's bearing capacity and diminish settlement.

Table 3. Ultimate bearing capacity and corresponding pile top settlement of test piles.

Pile No.	Before Grouting		After Grouting	
	Ultimate Bearing Capacity (kN)	Displacement of Pile Top (mm)	Ultimate Bearing Capacity (kN)	Displacement of Pile Top (mm)
TP1	27,644	30.33	36,765	33.53
TP2	26,560	33.23	36,765	30.56

4.2. Distribution of Pile Side Resistance

The combined post-grouting technique represents a significant advancement over traditional pile end grouting as it integrates both pile end and pile side post-grouting technologies. This combination results in substantial improvements in both pile tip-and-side resistance. The bearing capacity of piles is primarily derived from these two components. During the combined post-grouting process, the pile side grout strengthens the surrounding soil and alleviates issues related to the mud skin formed during pile installation. Figure 4 illustrates the variation in pile side resistance (q_{si}) along depth (H) for each soil layer of test piles at different stages under ultimate load. The figures reveal that, although the distribution patterns of pile side resistance before and after grouting remain largely unchanged, the pile side resistance at the same depth is significantly enhanced due to the injected grout. It is important to note that when a single pile is loaded, the pile side resistance gradually increases from the top to the bottom until it reaches its ultimate value. Under smaller loads, the deep pile side resistance may not be fully exerted. However, as shown in Figure 4, under ultimate load, the pile side resistance at all depths has reached its ultimate value. Table 4 provides statistics on the ultimate pile side resistance at each depth. The data indicate that the pile side resistance within the entire pile length of test piles TP1 and TP2 is increased after grouting compared to before. Due to the similar nature of the soil layer, there is no significant difference in the shallow pile side resistance, and the increase in side resistance is consistent. Notably, the increase in pile side resistance is particularly pronounced at approximately 1 m from the pile end. Specifically, the side resistance of test piles TP1 and TP2 at this depth increased by 78.72% and 80.21%, respectively. This enhancement is attributed to the pre-pressurization effect of the cement grout injected into the pile end, which extends to a certain range of the surrounding soil, thereby increasing the stress in the pile side soil within this range. The results demonstrate that combined post-grouting not only reinforces the soil on the pile side and improves the mud skin through pile side grouting but also further densifies the soil within a specific range of the pile end through the upward return of the grout at the pile end. This comprehensive approach significantly enhances the bearing performance of the pile foundation.

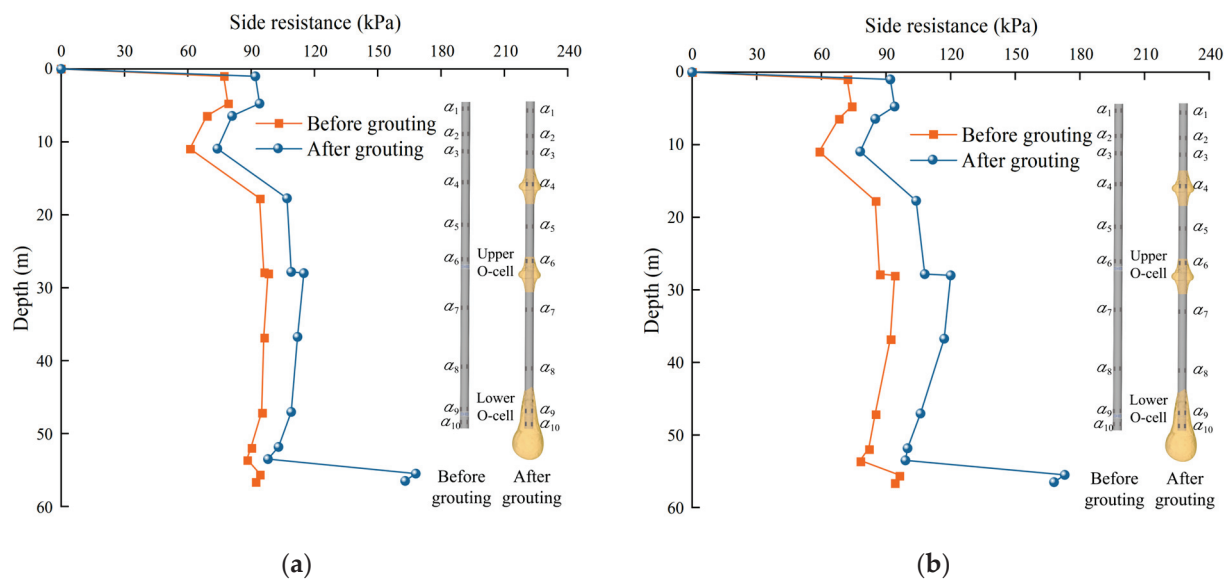


Figure 4. Distribution curves of side resistance along the depth of test piles under ultimate load: (a) TP1; (b) TP2.

Table 4. Pile side resistance statistics.

Soil Name	Elevation of Soil Layer	Average Side Resistance of TP1 (kPa)		Average Side Resistance of TP2 (kPa)	
		Before Grouting	After Grouting	Before Grouting	After Grouting
2-1 Silty clay	8.572–7.572 m	77	92	72	92
2-1 Silty clay	7.572–3.84 m	79	94	74	94
2-2 Silty clay	3.84–2.14 m	69	81	68	85
2-3 Silty clay	2.14–2.36 m	61	74	59	78
2-4 Silty clay	−2.36–9.16 m	94	107	85	104
2-5 Silty clay	−9.16–19.26 m	96	109	87	108
3-1 Clay	−19.26–19.428 m	98	115	94	120
3-1 Clay	−19.428–28.16 m	96	112	92	117
3-2 Clay	−28.16–38.46 m	95	109	85	106
3-3 Clay	−38.46–43.26 m	90	103	82	100
3-3 Clay	−43.26–44.928 m	88	98	78	99
3-4 Clay	−44.928–46.928 m	94	168	96	173
3-4 Clay	−46.928–47.928 m	92	163	94	168

4.3. Tip Resistance–Tip Displacement Curve

Pile tip resistance is a crucial factor in determining the bearing capacity of pile, particularly under heavy loads. Its role becomes even more prominent as it directly influences the pile's overall load-bearing capacity. Notably, the displacement of the pile tip, which is necessary for pile tip resistance to develop, is significantly impacted by grout compression. Figure 5 presents a comparison of the pile tip resistance (q_b) versus pile tip displacement (s_b) curves for test piles, both before and after grouting. Initially, when the pile top load is minimal, whether before or after grouting, the pile tip resistance for both TP1 and TP2 is relatively low. At this phase, the load is predominantly borne by the lateral resistance of the pile, leading to minimal displacement at the pile's termination point. As the load applied to the pile's apex progressively intensifies, the resistance offered by the pile tip starts to become evident, which is accompanied by a corresponding augmentation in the displacement at the pile's termination. However, once the top load exceeds a certain threshold, the growth rate of the tip resistance for the test pile before grouting slows down and stabilizes, while the pile tip displacement continues to rise. This indicates that the pile top load has surpassed the ultimate bearing capacity of the piles, leading to soil damage.

Conversely, after combined grouting, the pile tip resistance of TP1 and TP2 does not exhibit any noticeable slowdown in the growth rate. The grout injected at the pile end consolidates the pile end slag, reinforces the soil, and forms a consolidated grout body at the pile end. This results in a significant improvement in tip resistance. When providing the same pile tip resistance, the pile tip displacement of the test piles after combined grouting is reduced, leading to decreased settlement of the pile foundation under the same pile top loading condition. Furthermore, Figure 5 reveals that after combined grouting, the pile tip resistance responds more swiftly, with a faster growth rate. When the test pile reaches its ultimate bearing capacity, the proportion of pile tip resistance increases. This is attributed to the pre-pressurization effect of the grout injected into the soil at the pile end after grouting, which allows for the tip resistance to develop earlier. This not only enhances the bearing capacity of the single pile but also significantly reduces the settlement of the pile foundation. In summary, combined grouting effectively improves pile tip resistance, reduces pile tip displacement, and enhances the overall bearing capacity and stability of the pile foundation.

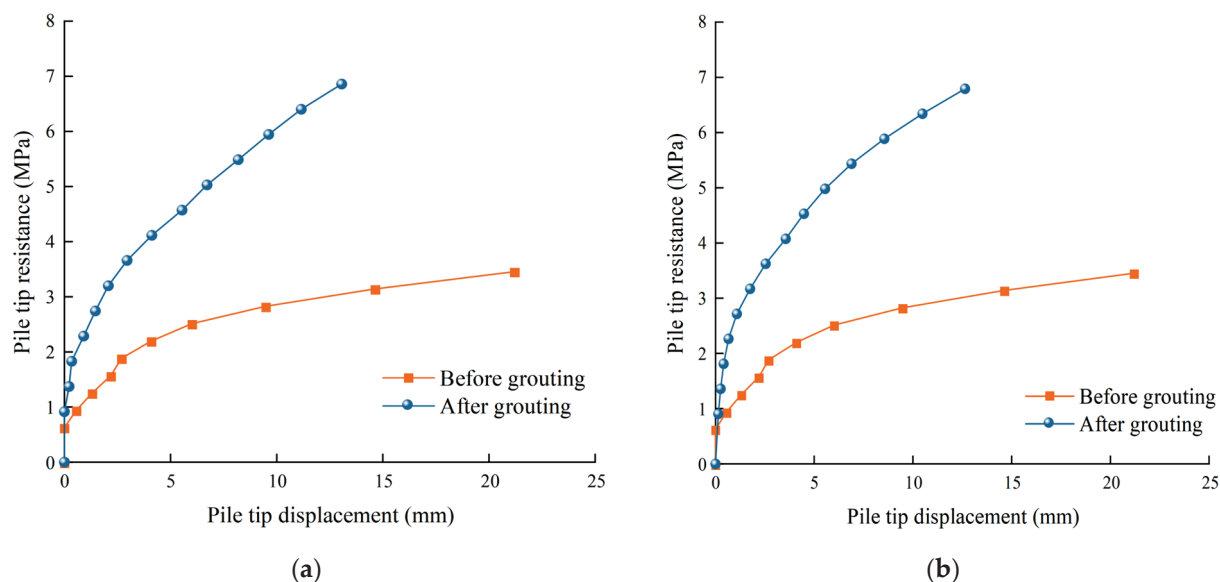


Figure 5. q_b - s_b curves of test piles: (a) TP1; (b) TP2.

4.4. Pile Load Transfer Mechanism

The ratio of pile side and tip resistance to the load capacity of a single pile serves as a quantitative indicator of the pile's load transfer mechanism [18]. Table 5 presents the proportions of pile tip-and-side resistance relative to the top load of the pile when the test piles, TP1 and TP2, reach their ultimate bearing capacity state. Upon examination of the table, it is evident that following combined post-grouting, both the pile side resistance and tip resistance of the test piles increased to a certain degree in their ultimate bearing capacity state. Notably, the increase in tip resistance is significantly greater than that of pile side resistance, amounting to approximately twice the pile tip resistance before grouting. Consequently, the ratio of pile tip resistance to the ultimate load for TP1 and TP2 increased from 12.49% and 12.99% to 18.65% and 18.48%, respectively. The improvement observed can be attributed to the mitigation of problems such as mud cake formation along the pile sides and slag accumulation at the pile ends through the application of combined post-grouting. Furthermore, the reinforcement of the soil surrounding the pile contributed to this enhancement. These combined factors resulted in a modification of the load transfer mechanism within the pile. Prior to grouting, the pile top loads of TP1 and TP2 were primarily borne by pile side resistance. However, following combined

post-grouting, although pile side resistance was effectively improved, its proportion in the ultimate bearing capacity decreased notably. Conversely, pile tip resistance exhibited a more substantial improvement, with its proportion in the ultimate bearing capacity also increasing significantly. This shift indicates that the center of gravity of the pile load bearing moved towards the pile end after combined post-grouting, resulting in an earlier and increased contribution of pile tip resistance to the pile top load. In conclusion, combined post-grouting not only enhances the bearing capacity of a single pile but also alters the load transfer mechanism of the pile. This change enables pile tip resistance to participate in bearing the pile top load earlier and respond more rapidly. Consequently, pile tip resistance is greatly improved, and its proportion of the ultimate bearing capacity is significantly increased. These findings further improve the performance of single-pile bearing capacity in cohesive soils.

Table 5. The ratio of pile resistance to pile top load under ultimate bearing capacity.

Pile No.	Stage	Side Resistance (kN)	Side Resistance as a Proportion of the Ultimate Load	End Resistance (kN)	End Resistance as a Proportion of the Ultimate Load	Ultimate Bearing Capacity (kN)
TP1	Before grouting	24,191	87.51%	3453	12.49%	27,644
	After grouting	29,907	81.35%	6858	18.65%	36,765
TP2	Before grouting	23,110	87.01%	3450	12.99%	26,560
	After grouting	29,972	81.52%	6793	18.48%	36,765

5. Finite Element Analysis of Post-Grouted Piles

5.1. Model Building

Due to the limited computational capabilities of computers in the past, numerical simulations in traditional post-grouted pile engineering primarily relied on simplified 2D models [19,20]. However, this approach resulted in a lack of refinement in the models and neglected the three-dimensional geometric shapes of actual objects. With advancements in computer technology, constructing more detailed and accurate 3D models has become the preferred method to conduct numerical simulation tests on post-grouted piles [21]. Compared to 2D models, 3D models can more comprehensively represent the geometric shapes of actual individual piles and lay the foundation for future, even more refined numerical simulation tests of post-grouted piles. Research has indicated that numerical simulation analysis using PLAXIS 3D V22 to evaluate the bearing capacity of pile foundations can yield effective and regular conclusions [21,22]. Therefore, a three-dimensional full-scale finite element model was developed in PLAXIS 3D V22 software using the field-tested pile TP1 as the foundational blueprint. In this model, the soil and additional solids are simulated utilizing the Moore–Cullen model, while the pile is defined as a linear elastic material. Since the bored piles were symmetric cylinders, and the setup of the combined post-grouting device was relatively uniform in the horizontal direction, the monopile models before and after grouting were considered symmetric bodies. The test pile possesses a diameter of 1.5 m and a length of 56.5 m. The dimensions for pile end grout consolidation are established with reference to the coring test results [23], with the diameter and height set to 1.9 times the pile radius. Additionally, the height of the grout's upward return is designated as 12 m, in accordance with the JTG 3363-2019 “Code for Design of Highway Bridge and Culvert Foundations and Fundamentals” [24]. The grout on the pile side is represented as an interface unit, with a virtual thickness factor of 0.2. To simulate the force exerted by the self-balancing load box in the pile at a distance of 28.5 m from the pile end, a pair of vertically opposed forces is implemented in the model. The calculation domain for the soil extends 20 times the pile diameter (D) horizontally and 1.5 times the pile length (L)

vertically from the pile's tip. PLAXIS 3D automatically meshes the pile by configuring the appropriate parameters. The granularity of mesh delineation has a direct impact on both calculation accuracy and speed; thus, 15-node cells are employed for mesh delineation. After comprehensive consideration, the piles and soil within a specific range are refined and encrypted. Figure 6 illustrates the mesh division of the calculation model. The blue area in the figure is the encrypted area of the mesh division, which improves calculation efficiency while ensuring calculation accuracy. Table 6 shows the specific parameters of each material in the model.

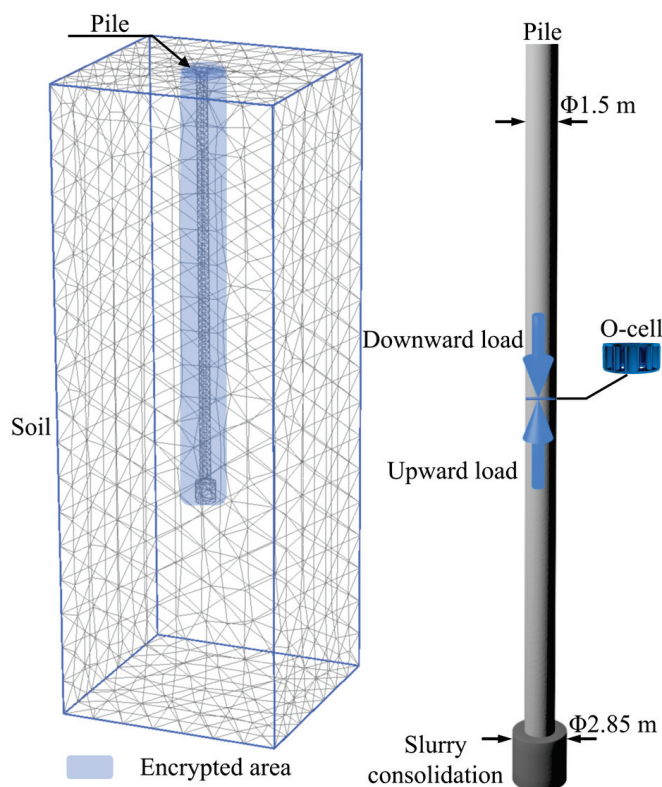


Figure 6. Computational modeling and mesh segmentation.

Table 6. Specific parameters of material.

Material Name	Stage	E (MPa)	ν	c (kPa)	φ ($^{\circ}$)
Pile		3.15×10^4	0.2		
soil		30	0.3	4	30
Grout	Before grouting	30	0.3	3.2	24
consolidation	After grouting	450	0.3	40	35

Note: E , modulus of elasticity of the material; ν , Poisson's ratio of the material; c , the cohesion of each soil; φ , the internal friction angle of each soil.

5.2. Analysis of Calculation Results

To validate the rationality of the model, a comparative analysis was conducted between the load–displacement curves of the test piles, as calculated through numerical simulation prior to grouting, and the corresponding curves measured during field testing. The comparison is depicted in Figure 7. Observing Figure 7, it is evident that the load–displacement curves derived from the numerical simulation align well with those recorded in the field, despite the presence of minor discrepancies. Overall, the results from both the numerical simulation and field testing exhibit a consistent trend, indicating that the established calculation model and parameter selection are reasonably accurate.

Furthermore, this analysis underscores the feasibility of utilizing PLAXIS 3D software to analyze the force characteristics of pile foundations prior to grout injection.

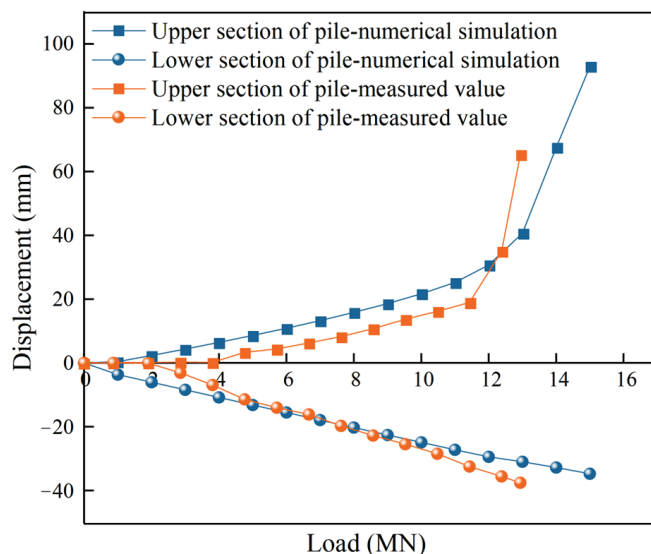


Figure 7. Curves of calculated and measured results.

Numerical simulation serves as an efficient tool for investigating issues related to the post-grouting of pile foundations. As illustrated in Figure 8, load–displacement curves for the self-balancing tests of pile foundations, both before and after grouting, have been computed. Upon examining Figure 8, it is apparent that, after grouting, under identical loading conditions, the displacements of both the upper and lower pile sections are significantly reduced compared to before grouting. Prior to grouting, the upper pile section bears the entire load primarily through pile side resistance. However, post-grouting, this resistance is substantially enhanced due to soil reinforcement surrounding the pile and improved pile–soil interface mechanical properties. Both before and after grouting, the pile side resistance exhibits a notable steep decline as the load increases, eventually reaching a limit value where the upper pile section can no longer sustain additional load, leading to pile–soil displacement and significant relative movement. Differently, the lower pile section supports the load from the load box through a combination of pile side resistance and pile tip resistance. Comparing the load–displacement curves of the lower pile before and after grouting reveals a significant reduction in displacement under the same load following combined grouting, indicating a substantial improvement in the single-pile bearing capacity. Furthermore, Figure 8 demonstrates that the curves for the lower pile section, both pre- and post-grouting, do not exhibit a steep decline. Upon reaching the maximum resistance of the upper pile segment’s lateral surface, the lower pile segment’s bearing capacity remains unmaximized, indicating that the pile tip resistance has yet to be fully utilized and is still in a phase of gradual growth. This observation indicates that pile tip resistance and pile lateral resistance develop asynchronously, and the capacity of a single pile to support heavier loads is predominantly influenced by the extent of pile tip resistance mobilization.

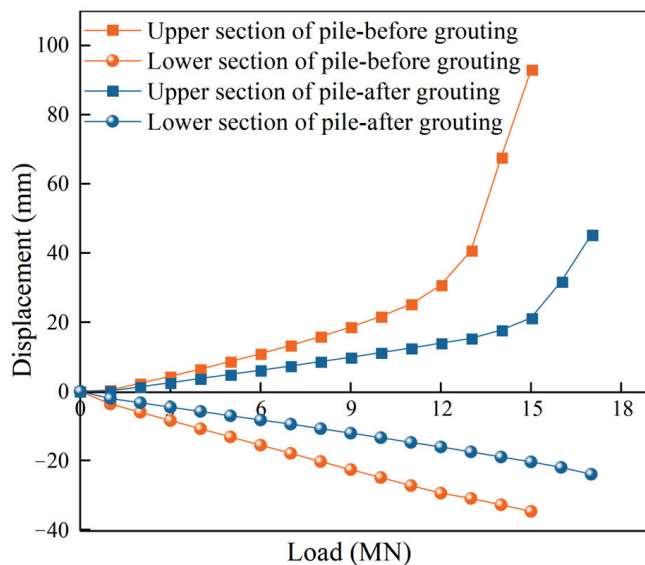


Figure 8. Self-balance test of numerical simulation load–displacement curve.

6. Conclusions

This paper presented a comparative analysis of the bearing performance of two large-diameter bored piles in the Huaiyang Left-Line Bridge Project of the Lianyungang–Zhenjiang Railway, focusing on the results obtained from static load tests conducted using the self-balancing method. The analysis compared the pile foundation’s performance before and after the application of combined grouting. It also examined the changes in the pile load transfer mechanism before and after grouting. Additionally, the paper integrated numerical simulation results to investigate the bearing characteristics of piles subjected to combined post-grouting and compared these findings with those from field tests. The key conclusions drawn are as follows:

- (1) The application of the combined post-grouting technique significantly increased the ultimate bearing capacity of the two test piles of the railway bridge project by 32.99% and 38.42%, respectively. Moreover, the pile top displacement is significantly decreased under the same applied load. It is worth noting that the load–displacement curve at the pile top exhibited a gradual and smoothly varying trend, rather than a sharp decline, following the combined grouting treatment. This observation underscores the substantial impact of the post-grouting intervention.
- (2) The injection of grout into the pile sides during the combined grouting process significantly enhances the mechanical characteristics of the pile–soil interface, leading to an increase in pile lateral resistance along the entire length of the pile. The grout injected at the pile end creates a pre-compression effect on the soil within a specific proximity of the pile tip. Consequently, the most pronounced increase in pile side resistance is observed within approximately 1 m from the pile end, with an enhancement ranging from 78.72% to 80.21%.
- (3) Under identical pile tip displacements, test piles subjected to combined post-grouting demonstrate an augmented pile tip resistance. The pile tip resistance responds more promptly to increasing pile tip displacements after grouting. Furthermore, the application of combined tip-and-side post-grouting alters the load transfer mechanism of the pile. Specifically, there is a substantial increase in both pile lateral resistance and pile tip resistance in the vicinity of the pile end. Consequently, the proportion of pile tip resistance to the ultimate bearing capacity of the pile foundation increases, from a range of 12.49% to 12.99% before grouting to a range of 18.14% to 18.65%

after grouting. This shift signifies a relocation of the center of gravity of the pile foundation's bearing capacity towards the pile end.

- (4) The numerical simulation results indicate that the implementation of combined post-grouting enhances the mechanical properties of the pile–soil interface through the introduction of grout into the pile sides, thereby increasing the pile side resistance. Moreover, the consolidation of grout injected into the pile end significantly improves the pile tip resistance performance. It is worth highlighting that the development of pile lateral resistance and pile tip resistance occurs asynchronously; the pile lateral resistance reaches its ultimate capacity before the pile tip resistance is fully mobilized. This observation aligns with the findings obtained from the field testing.

Author Contributions: Methodology, R.Z.; Software, Z.G.; Resources, W.G. and Z.W.; Writing—original draft, R.Z., Z.G. and Z.W.; Writing—review & editing, R.Z., Z.G., W.G. and Z.W.; Project administration, Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (52008100, 52178317), Natural Science Foundation of Jiangsu Province (BK20200400), China Postdoctoral Science Foundation (2022M723534), and the Natural Science Foundation of Jiangsu Higher Education Institutions of China (23KJA560005). And The APC was funded by Zhihui Wan.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Acknowledgments: Lei Wang of Nanjing Dongtu Construction Technology Co., Ltd. was also acknowledged for his contribution to this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gouvenot, D.; Gabaix, J.C. A new foundation technique using piles sealed by cement grout under high pressure. In *Offshore Technology Conference*; OTC: Houston, TX, USA, 1975.
2. Safaqah, O.; Bittner, R.; Zhang, X.G. Post-grouting of drilled shaft tips on the Sutong Bridge: A case study. In *Proceedings of the Geo-Denver 2007 Congress: Contemporary Issues in Deep Foundation*, Denver, CO, USA, 18–21 February 2007.
3. Sze, J.W.C.; Chan, K.M. Application of shaft grouting technique in deep foundations-Hong Kong experience. In *Grouting and Deep Mixing*; ASCE: Denver, CO, USA, 2012.
4. Zhai, D.Z.; Zhou, F.; Zhou, J.F.; Zhu, R. Experimental study and numerical analysis on bearing characteristics of large diameter ultra long bored piles with post grouting on the tip and side. In *IOP Conference Series: Earth and Environmental Science*; IOP Publishing: Bristol, UK, 2018; Volume 189, p. 022024.
5. Nguyen, M.H.; Fellenius, B.H. Bidirectional cell tests on not-grouted and grouted large-diameter bored piles. *J. Geo-Eng. Sci.* **2015**, *2*, 105–117. [CrossRef]
6. Lin, S.S.; Lin, T.; Chang, L.T. A case study for drilled shafts base mud treatment. In *New Technological and Design Developments in Deep Foundations*; ASCE: Denver, CO, USA, 2000; pp. 46–58.
7. Ho, C.E. Base grouted bored pile on weak granite. In *Grouting and Ground Treatment*; ASCE: New Orleans, LA, USA, 2003; pp. 716–727.
8. Liu, K.-F.; Xie, X.-Y.; Luo, Z.; Hu, Q.-H.; Huang, C.-X. Full-scale field load testing of long drilled shafts with enlarged base constructed in marine sediment. *Mar. Georesour. Geotechnol.* **2017**, *35*, 346–356. [CrossRef]
9. Fiscina, L.F.G.; Barbosa, Y.; de Albuquerque, P.J.R.; de Carvalho, D. Field study on axial behavior of instrumented post-grouted steel pipe micropiles in tropical lateritic soil. *Innov. Infrastruct. Solut.* **2021**, *56*, 127–135. [CrossRef]
10. Ma, H.; Ma, Y.; Yao, W.; Chen, C. Research on bearing capacity of post-grouting super-long bored pile. In *IOP Conference Series: Earth and Environmental Science*; IOP Publishing: Bristol, UK, 2021; Volume 634, p. 012111.
11. Lai, V.Q.; Keawsawasvong, S.; Shiau, J. Analysis of shaft-grouted piles using load-transfer method. *Int. J. Geosynth. Ground Eng.* **2022**, *8*, 4. [CrossRef]
12. Thiyyakkandi, S.; McVay, M.; Neeraj, C.R. Full-scale axial load response of jetted and grouted precast piles in cohesionless soils. *J. Geotech. Geoenviron. Eng.* **2022**, *148*, 04022030. [CrossRef]

13. Nguyen, M.H.; Fellenius, B.H. Bidirectional-cell tests on two 70 m long bored piles in Vietnam. In *GeoInstitute Geo Congr*; Geotechnical Special Publication: Reston, VA, USA, 2014; pp. 482–496.
14. Zhou, X.; Wang, C.C.; Liu, X.H.; Xu, X. Research on load bearing performance of post-grouting bored piles in deep and thick pebble layer. *Bridge Constr.* **2023**, *53* (Suppl. S1), 31–37. (In Chinese)
15. Cui, Y.; Qi, C.; Zheng, J.; Wang, X.; Zhang, S. Field test research on post-grouting effect for super-long cast-in-place bored pile in thick soft foundation. *Geotech. Geol. Eng.* **2021**, *39*, 4833–4842. [CrossRef]
16. Prakash, S.; Sharma, H.D. *Pile Foundations in Engineering Practice*; John Wiley & Sons: New York, NY, USA, 1990.
17. Zhang, Q.Q.; Li, S.C.; Li, L.P. Field Study on the Behavior of Destructive and Non-Destructive Piles Under Compression. *Mar. Georesour. Geotechnol.* **2014**, *32*, 18–37. [CrossRef]
18. Ruiz, M.E.; Pando, M.A. Load transfer mechanisms of tip post-grouted drilled shafts in sand. In *Contemporary Topics in Deep Foundation*; ASCE: Denver, CO, USA, 2009; pp. 23–30.
19. Youn, H.; Tonon, F. Numerical analysis on post-grouted drilled shafts: A case study at the Brazo River Bridge, TX. *Comput. Geotech.* **2010**, *37*, 456–465. [CrossRef]
20. Thiyyakkandi, S.; McVay, M.; Bloomquist, D.; Lai, P. Measured and predicted response of a new jetted and grouted precast pile with membranes in cohesionless soils. *J. Geotech. Geoenviron. Eng.* **2013**, *139*, 1334–1345. [CrossRef]
21. Xu, Y.F.; Wan, Z.H.; Dai, G.L.; Gong, W.; Gao, L. Experimental study on long-term bearing behaviors of base post-grouting bored piles. *J. Build. Struct.* **2021**, *42*, 139–146. (In Chinese)
22. Liang, Y.W.; Bian, X.C.; Dong, L. Analysis on viaduct pile-group foundation deformation caused by lateral stack load. *Railw. Eng.* **2017**, *57*, 5–12. (In Chinese)
23. Dai, G.L.; Gong, W.M.; Tong, X.D. Detection technology for grouting effects of pile-end post-grouting piles. *Constr. Technol.* **2005**, *34*, 74–77. (In Chinese)
24. *JTG 3363-2019*; Code for Design of Ground Base and Foundation of Highway Bridges and Culverts. China Communications Press: Beijing, China, 2019. (In Chinese)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Element Migration of Mineralization-Alteration Zones and Its Geological Implication in the Beiya Porphyry–Skarn Deposit, Northwestern Yunnan, China

Fei Liu ¹, Runsheng Han ^{2,3,*}, Yuxinyue Guo ¹, Mingzhi Wang ^{2,3} and Wei Tan ⁴

¹ College of Mining Engineering, Guizhou University of Engineering Science, Bijie 551700, China; liufei727@gues.edu.cn (F.L.)

² Faculty of Land Resources Engineering, Kunming University of Science and Technology, Kunming 650093, China

³ Southwest Institute of Geological Survey, Geological Survey Center for Non-Ferrous Mineral Resources, Kunming 650093, China

⁴ Sino-Zijin Resources Ltd., Beijing 100012, China

* Correspondence: 11301124@kust.edu.cn

Abstract: Porphyry and the associated skarn-type deposit is one of the most important types of ore deposits worldwide, which usually exhibit significant zoning of mineralization-alteration, but the research on element migration in these mineralization-alteration zones is relatively weak. The Beiya porphyry–skarn gold-polymetallic deposit is a super-large Cenozoic deposit located in the Sanjiang metallogenic belt, northwestern Yunnan, China. In this paper, through a detailed analysis of mineralization and alteration zoning and its element migration regularity, the findings are as follows: (1) Three types of hydrothermal alteration—porphyry alteration, contact alteration, and wall-rock alteration—are developed, and porphyry alteration includes potassic, phyllic, propylitic, and argillic alteration; (2) five types of mineralization—porphyry-type Cu–Au–(Mo), skarn-type Au–Fe–(Cu), hydrothermal vein-type Au–Fe, distal hydrothermal-type Pb-polymetallic, and oxidizing-leaching enriched-type Au—occur in a diversity of forms, which are dominantly controlled by structures and lithologies; (3) concentric-banded mineralization-alteration zones are exhibited centrally from the alkaline porphyry outward or upward, namely [porphyry alteration] potassic → phyllic → propylitic → argillic → [contact alteration] skarnitization–marbleization → [wall-rock alteration] marbleization–silicification–calcitization; (4) porphyry-type mineralization predominantly forms within potassic and phyllic zones, while skarn-type mineralization occurs in contact alteration zones, and proximal and distal hydrothermal (vein)-type mineralization are commonly distributed in marbleization–silicification–calcitization alteration zones; and (5) element migration analysis demonstrates a significantly lateral and vertical zoning in the metallogenic element association of Cu–Mo → Cu–Au → Au–Fe–Cu → Au–Fe → Pb–Zn–Au–Ag–Fe from alkaline porphyry outward to the wall-rock. The mineralization-alteration zoning model indicates the Beiya deposit has similar mineralization and alteration zone characteristics to the typical porphyry copper system; and element migration within mineralization-alteration zones provides new scientific information for understanding the metallogenic regularity and prospecting at Beiya, as well as the similar types of deposits in the Sanjiang metallogenic belt and elsewhere in the world.

Keywords: element migration; mineralization-alteration zone; porphyry copper system; Beiya porphyry–skarn deposit; Sanjiang metallogenic belt

1. Introduction

Porphyry deposits are important sources of the metals Cu, Mo, Au, Ag, and Re [1–3]; they are formed as a result of dynamic interactions between magmatic, hydrothermal, and tectonic processes [4,5]. These deposits exhibit a particular mineralization and alteration

zoning that commonly comprises sodic–calcic, potassic, phyllic (sericitic), propylitic (chloritic–epidotic–carbonation), and (advanced) argillic, as well as greisen and skarn [2,6–9]. Hydrothermal alteration minerals and assemblages are zoned spatially, temporally, and thermally [2,9]: (1) sodic–calcic alteration usually occurs at the deepest level but is poorly preserved or uncommon [2]; potassic alteration is centrally located and deeper and is overlain by phyllic alteration, while (advanced) argillic alteration forms near the surface and generally overlaps the potassic and phyllic zones; propylitic alteration can develop at deep to shallow parts surrounding the central zones of potassic, phyllic, and argillic alteration—it may also grade downward into deeper zones of sodic–calcic alteration or upward into the argillic zone; skarn is commonly generated in contact with wall-rocks; (2) in general, these alteration types form progressively younger upward or outward [2]; and (3) these alterations are the result of the declining temperature of hydrothermal fluids [2]. These hydrothermal alterations record important information of ore-forming processes and provide the basis for the exploration of porphyry deposits [10–12]. For instance, potassic and phyllic alteration are predominantly associated with sulfide mineralization, whereas propylitic and argillic alteration are sulfide-free or contain minor amounts of sulfides [2,6,9]. Hydrothermal alteration is characterized by ionic metasomatism that consists of alkali metasomatism, hydrolytic (or acidic) reactions, oxidation/reduction reactions, and hydration–carbonation reactions, among others [9]. Moreover, element gains and losses always occur during metasomatic processes. Investigating the element migration associated with hydrothermal alteration has crucial significance for the establishment of mineralization and alteration zoning and for a deep understanding of the ore-forming process [13–15]. However, the element migration characteristics in distinctive mineralization and alteration zones of porphyry deposits are relatively unclear and lack a detailed case study.

The Beiya gold-polymetallic deposit is typical and representative of Cenozoic porphyry–skarn-type deposits in the Sanjiang metallogenic belt, northwestern Yunnan, China, with gold reserves of more than 200 tonnes and a gold grade of about 2.45 g/t [16]. Three stages of magmatic activities are mainly developed at Beiya: (1) quartz albite porphyry (QBP) and lamprophyre are formed in 65–60 Ma [17,18]; (2) quartz syenite porphyry (QSP) or monzogranitic porphyry (MGP) are formed at about 37–34 Ma [19–22]; and (3) there is 3.78–3.66 Ma of biotite syenite porphyry (BSP) [17]. The Re–Os isotope ages of molybdenite are 36–34 Ma [21,22], indicating that mineralization is closely related to QSP or MGP in the late Paleogene, and the sources of ore-forming fluids and materials are mainly derived from magma [23–26]. Five types of mineralization—porphyry, skarn, hydrothermal vein, stratabound, and supergene-enriched—are well developed in the Beiya deposit [16,25–30] and are intensely controlled by various structures and lithologies [16,27,30]. There are three main types of hydrothermal alteration—porphyry alteration, contact zone alteration, and wall-rock alteration—that occur in the Beiya deposit [16,19], and He et al. [16] proposed the characteristics of a metallogenic element zone. Nevertheless, the element migration of these mineralization–alteration zones has not been discussed.

In this paper, we identify the various types of mineralization and alteration and their occurrence features, carry out element migration regularity of different mineralization–alteration zones, and then establish the model of mineralization and alteration zoning, which provides new understandings of metallogenic regularity and a scientific basis for prospecting prediction of the Beiya deposit (Figure 1), as well as the similar porphyry deposits in the region and elsewhere in the world.

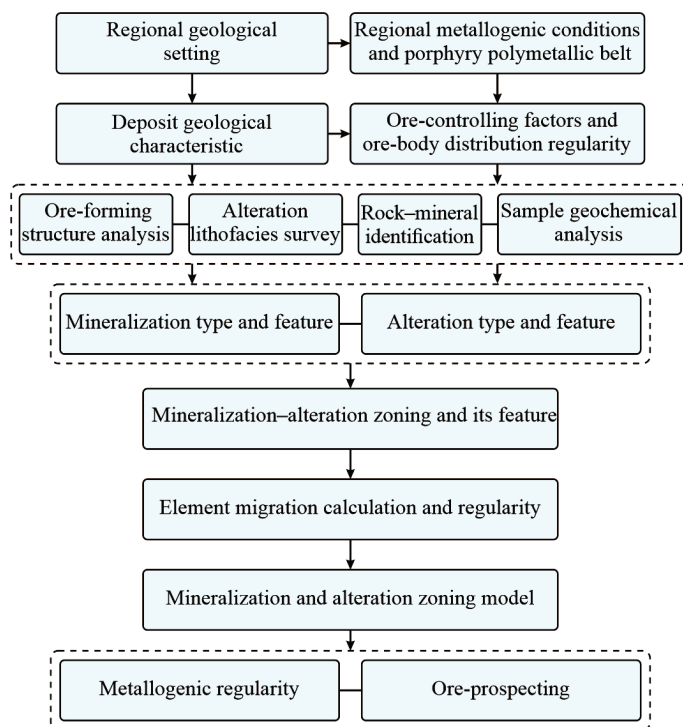


Figure 1. The framework plot of the study in the Beiya deposit.

2. Regional and Deposit Geology

The Beiya porphyry–skarn gold-polymetallic deposit is located in the middle part of Jinshajiang–Ailaoshan alkaline porphyry belt, on the western margin of the Yangtze Plate, where there is a combination of several geotectonic units (Figure 2a). The Jinshajiang–Ailaoshan Fault, the Honghe Fault, the Chenghai Fault, and the Lijiang–Muli Fault are a few examples of regional deep faults dispersed in neighboring areas (Figure 2a). These faults restrict the evolution of geological structures, magmatic activities, and mineralization in the area, particularly the formation and distribution of the Cenozoic alkaline porphyry and its genetically related porphyry-type polymetallic deposits.

The exposed strata at Beiya are composed of the Permian Emeishan Formation ($P_3\beta$) of dark-green basalts, the Early Triassic Qingtianbao Formation (T_1q) of amaranth sandstones, the Middle Triassic Beiya Formation (T_2b), and the Quaternary sediments (Q_1s , Q_p , and Q_h) (Figure 2b). The primary wall-rock of magmatic rocks and ore-bodies is the Beiya Formation, which is composed mainly of light-grey to dark-grey limestones and dolomites.

Structural activities are intense in the Beiya deposit. The SN-trending Beiya syncline, along with the dominant SN- and EW-trending faults, form the structural framework (Figure 2b). In addition, the secondary faults and fractures in various directions are developed. These structures together constitute a fold-fault-joint rock- and ore-controlling structural system [27,31,32], which controls the emplacement of alkaline porphyries and the formation and distribution of ore-bodies [16,30,31].

The Beiya deposit has three stages of alkaline porphyries, as previously mentioned. Most of these alkaline porphyries have a small scale, and their extended direction is primarily in the SN and EW orientations (Figure 2b). The geochemical characteristics of these porphyries indicate that they are silicon-supersaturation and high-K alkali-rich porphyries, which originated from the deep crust-mantle interaction caused by the intense tectonic deformation in the collision and orogeny of the Indo-Eurasian plates [17–19,21].

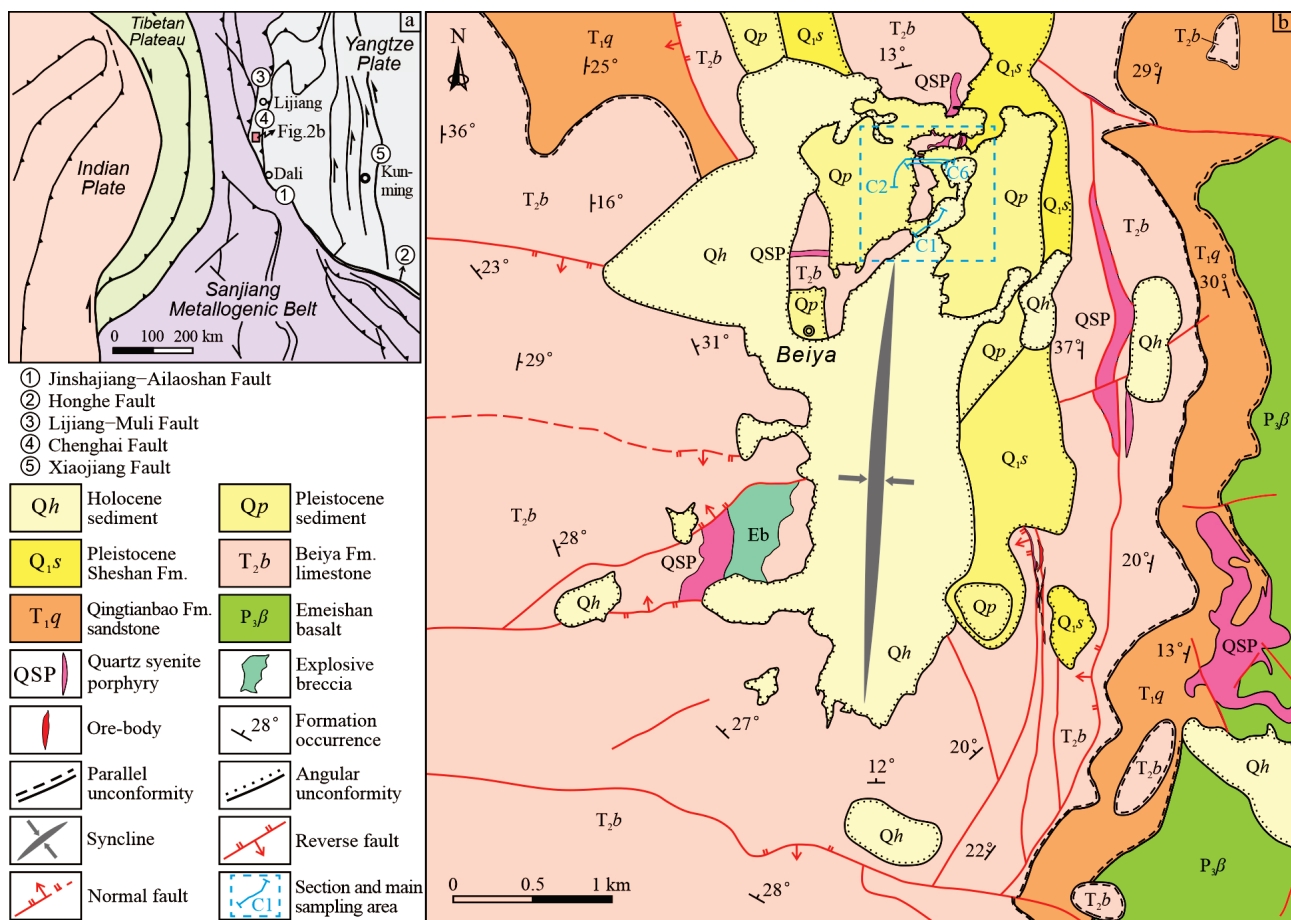


Figure 2. The regional geological setting (a) and geology of the Beiya deposit (b) (modified after Liu et al. [27]).

3. Samples Selected and Analytical Methods

The analysis of whole-rock major and trace elements was performed on 77 samples mostly collected from the measured geological sections (Figures 3 and 4) to study the element migration in different mineralization-alteration zones. These samples include the fresh and altered/mineralized limestones from the Beiya Formation, the exoskarns (protoliths are limestones), the unaltered quartz syenite porphyries, the altered quartz syenite porphyries (from potassic, phyllic, propylitic, and argillic alteration, respectively), and the endoskarns (protoliths are quartz syenite porphyries) (Table 1). To ensure the initial chemical compositions of the fresh limestone and the unaltered quartz syenite porphyry, these samples were collected from the periphery of the mining area to avoid the effects of hydrothermal alteration or mineralization. The altered samples were subjected to extensive macroscopic and microscopic rock-mineral identification, as well as electron probe microanalysis of typical minerals [33] to determine the types and characteristics of alteration and/or mineralization.

Table 1. The average geochemical characteristics of selected samples from the Beiya deposit.

Elements (10 ^{−6})/ Oxides (wt.%)	Al ₂ O ₃	CaO	Fe ₂ O ₃ [*]	K ₂ O	MgO	MnO	Na ₂ O	P ₂ O ₅	SiO ₂	TiO ₂	W	Sn	Mo	Bi	Cu	Au	Ag	
UL (<i>n</i> = 6)	0.60	34.64	0.43	0.12	17.67	0.02	0.08	0.050	0.72	0.04	0.90	1.30	1.00	0.10	3.60	0.01	0.02	
AML (<i>n</i> = 14)	2.29	19.50	16.21	0.44	5.02	1.28	0.04	0.053	6.79	0.20	21.69	6.80	6.17	99.32	1540.00	3.45	15.70	
EXS (<i>n</i> = 2)	1.53	14.10	30.24	0.75	9.92	0.26	0.01	0.040	8.48	0.12	158.50	11.50	8.04	30.40	615.00	0.57	5.73	
UQSP (<i>n</i> = 3)	16.98	0.18	2.70	7.11	0.33	0.04	0.33	0.060	67.78	0.33	2.50	1.83	0.73	0.27	14.27	0.01	0.66	
PZ (<i>n</i> = 13)	8.60	0.10	10.53	7.68	0.34	0.03	0.33	0.061	61.01	0.15	15.00	12.70	81.50	477.31	20,770.00	5.88	84.84	
PLZ (<i>n</i> = 16)	8.68	0.09	9.67	4.63	0.24	0.02	0.63	0.064	69.80	0.16	31.94	8.00	93.60	207.70	16,647.00	3.68	57.50	
PTZ (<i>n</i> = 6)	12.12	0.23	8.88	5.17	1.38	0.03	0.92	0.033	57.40	0.41	23.83	7.50	56.80	207.37	3394.00	3.82	26.82	
AZ (<i>n</i> = 13)	15.04	0.26	5.13	6.91	0.34	0.04	0.42	0.067	63.40	0.17	32.31	3.40	53.30	37.51	1497.00	1.10	21.22	
ENS (<i>n</i> = 4)	4.31	2.98	26.66	2.18	10.98	0.49	0.18	0.073	45.80	0.19	39.88	12.70	16.37	2.43	64.90	0.02	5.32	
Elements (10 ^{−6})/ Oxides (wt.%)	Pb	Zn	As	Sb	La	Ce	Pr	Nd	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
UL (<i>n</i> = 6)	10.40	17.00	2.70	0.80	1.90	3.40	0.64	1.97	0.36	0.13	0.38	0.11	0.32	0.12	0.21	0.10	0.18	0.10
AML (<i>n</i> = 14)	17,190.00	7234.00	400.00	30.59	7.00	12.60	1.53	6.10	1.32	0.60	1.29	0.20	1.05	0.22	0.60	0.12	0.51	0.11
EXS (<i>n</i> = 2)	172.15	335.00	361.00	15.20	2.00	3.60	0.48	1.95	0.45	0.27	0.47	0.07	0.42	0.09	0.26	0.05	0.29	0.06
UQSP (<i>n</i> = 3)	103.40	82.90	6.47	2.68	21.47	42.83	3.94	14.67	2.64	0.66	2.11	0.30	1.60	0.30	0.83	0.14	0.75	0.14
PZ (<i>n</i> = 13)	219.70	111.00	283.10	61.45	14.30	29.50	3.28	12.90	2.76	0.52	2.42	0.33	1.67	0.31	0.80	0.12	0.66	0.11
PLZ (<i>n</i> = 16)	514.00	209.00	745.00	73.80	16.70	35.50	3.84	14.40	2.69	0.62	2.07	0.27	1.40	0.26	0.67	0.10	0.58	0.09
PTZ (<i>n</i> = 6)	320.70	118.00	285.40	42.80	27.90	51.00	5.21	19.50	3.92	0.83	3.10	0.41	2.02	0.37	0.97	0.15	0.87	0.14
AZ (<i>n</i> = 13)	422.50	148.00	286.80	35.37	23.00	37.00	3.97	14.30	2.74	0.68	2.24	0.31	1.56	0.29	0.79	0.13	0.71	0.13
ENS (<i>n</i> = 4)	1110.00	814.00	27.20	9.22	17.80	29.90	3.58	13.00	2.22	0.51	1.88	0.25	1.21	0.22	0.56	0.08	0.48	0.08

UL—Unaltered limestone (Beiya Fm.); AML—Altered/mineralized limestone (Beiya Fm.); EXS—Exoskarn (protolith is limestone); UQSP—Unaltered quartz syenite porphyry; PZ—Potassic zone; PLZ—Phyllic zone; PTZ—Propylitic zone; ENS—Endoskarn (protolith is quartz syenite porphyry). Fe₂O₃^{*} = Fe₂O₃ + FeO. *n* represents the number of samples.

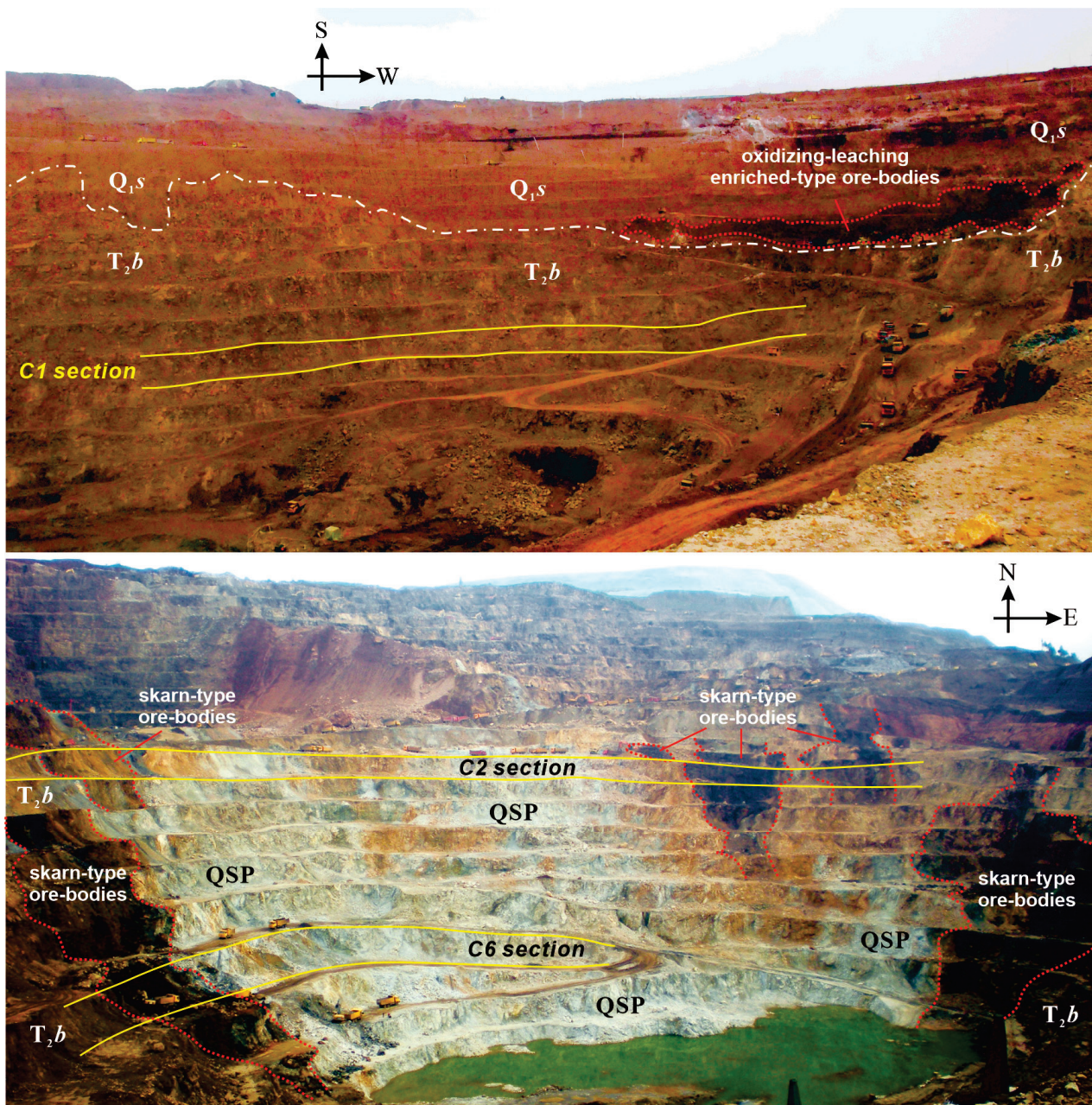


Figure 3. The location of geological sections in the Beiya deposit (modified after Liu et al. [31]).

The analysis of whole-rock major and trace elements was conducted at ALS Chemex (Guangzhou, China) Co., Ltd. The samples were cleaned and crushed to <200 mesh using an agate mill. The whole-rock major elements were analyzed by using X-ray fluorescence spectrometry (XRF), and the analysis precision was better than $1\% \pm 2\%$. The trace elements were analyzed by using inductively coupled plasma-mass spectrometry (ICP-MS) with analytical uncertainties of <5%.

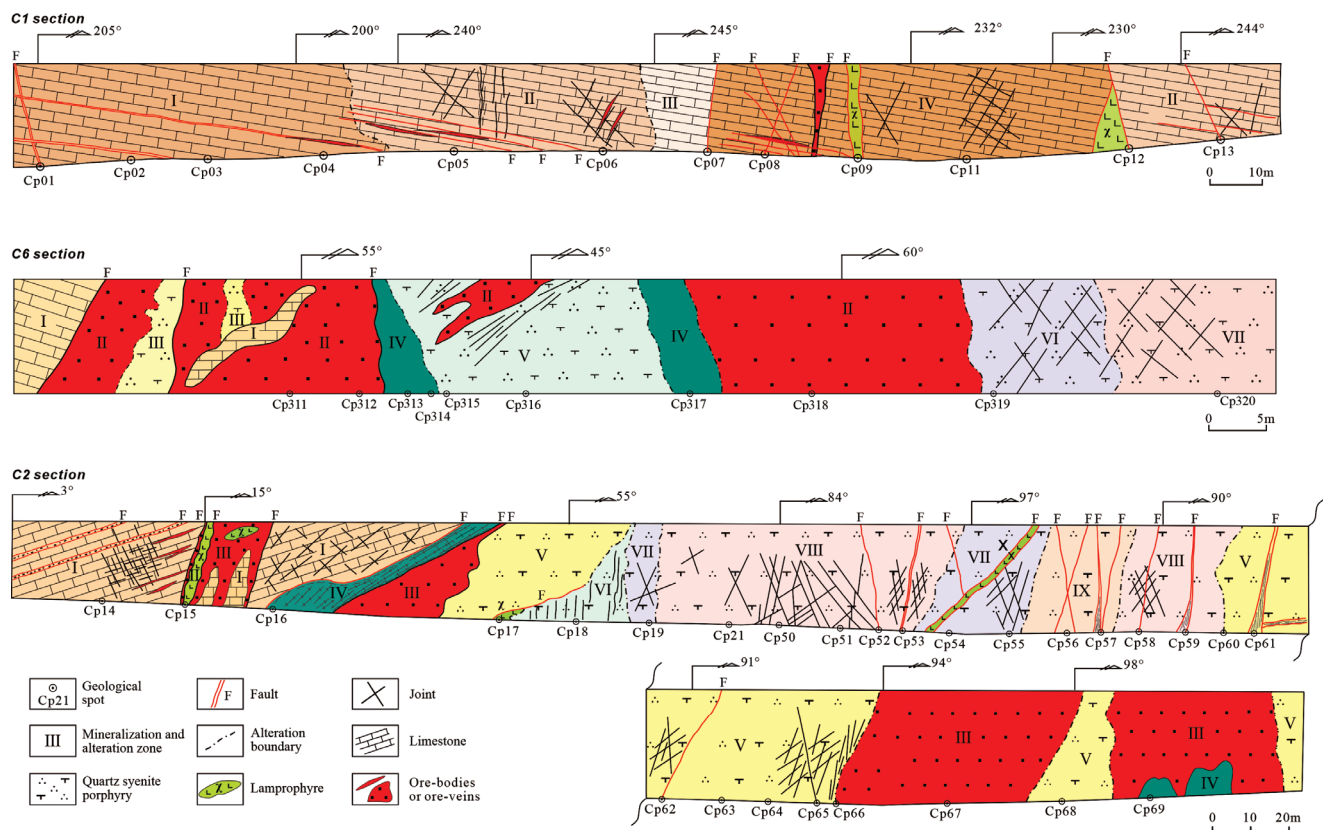


Figure 4. The mineralization-alteration zones of selected geological sections in the Beiya deposit. The description of the mineralization and alteration zones identified numerically is given in the text (C1 and C2 sections are modified after Liu et al. [31]).

4. Result

4.1. Alteration Types

Multi-stage magmatic and hydrothermal activities are developed in the Beiya deposit, resulting in significant hydrothermal alteration in the porphyries, wall-rocks, and their contact zones.

4.1.1. Porphyry Alteration

Potassic Alteration

Potassic alteration usually occurs in the central part of alkaline porphyry (Figure 4), which typically displays a pale red color (Figure 5A) and mainly contains hydrothermal orthoclase (K-feldspar) and hydrothermal biotite (Figure 5B,C), as well as a small amount of quartz, sericite (Figure 5B,C) and some metal minerals such as pyrite, chalcopyrite, molybdenite, native gold, etc. The potassic alteration has certain characteristics: (1) plagioclase or orthoclase phenocrysts are replaced by hydrothermal orthoclase or hydrothermal biotite, forming a ring edge of hydrothermal orthoclase (Figure 5B) and even replacing them completely; (2) hydrothermal orthoclase or hydrothermal biotite replace groundmass; (3) hydrothermal orthoclase or hydrothermal biotite fill in fractures (Figure 5C); (4) a small degree of sericitic and pyritic alteration telescopes on potassic alteration.

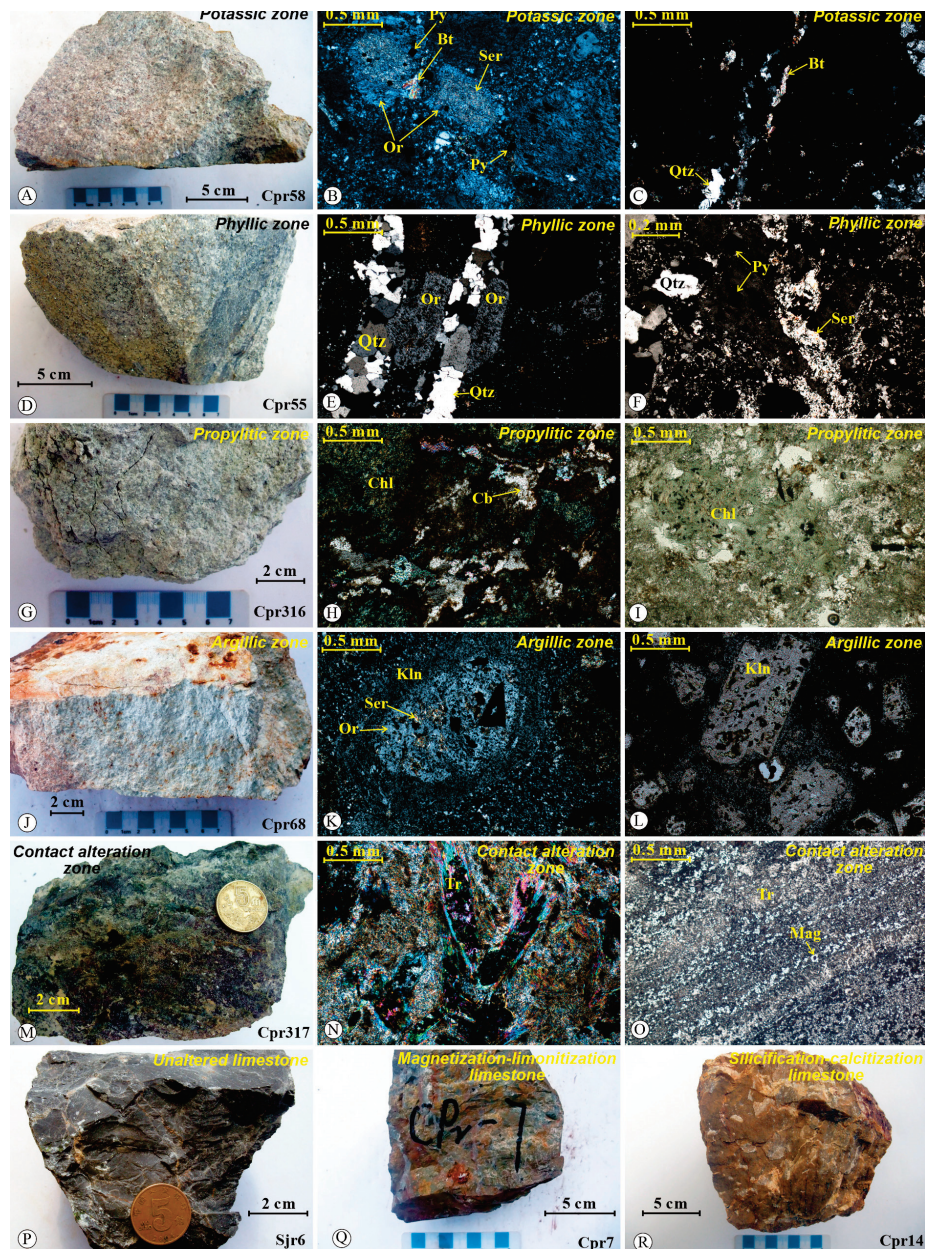


Figure 5. Alteration characteristics in the Beiya deposit. (A)—potassic alteration QSP, with light red to pink in color; (B)—hydrothermal orthoclase (K-feldspar) and hydrothermal biotite occur, and the hydrothermal orthoclase are replaced by the later sericites; (C)—hydrothermal biotite fill in fracture; (D)—phyllic alteration QSP; (E)—quartz occur in fracture and cut the previous orthoclase; (F)—quartz, sericite, and pyrite are developed in phyllic zone; (G)—propylitic alteration QSP, with light greenish in color; (H)—chlorite and some carbonate minerals are developed in propylitic zone; (I)—chlorite replace groundmass; (J)—argillic alteration QSP; (K)—kaolinite metasomatize orthoclase along the edge; (L)—kaolinite completely metasomatize plagioclase; (M)—endoskarn, with greenish-gray in color; (N)—tremolite are developed in skarn; (O)—tremolite and magnetite are developed in skarn; (P)—unaltered limestone of the Beiya Fm.; (Q)—magnetization and limonitization limestone of the Beiya Fm.; (R)—silicification and calcitization limestone of the Beiya Fm. Bt = Biotite, Or = Orthoclase, Ser = Sericite, Qtz = Quartz, Chl = Chlorite, Cb = Carbonate mineral, Kln = kaolinite, Tr = Tremolite, Py = Pyrite, Mag = Magnetite.

Phyllic Alteration

Phyllic alteration commonly occurs outside of potassic alteration (Figure 4), mostly including quartz, sericite, and pyrite (Figure 5D), and a small amount of K-feldspar, chal-

copyrite, bornite, native gold, etc. The phyllic alteration is characterized by (1) lots of sericite and quartz occurring, replacing mineral phenocryst and groundmass; (2) sericite or quartz in veins or veinlet fillings (Figure 5E,F); (3) large amounts of pyrite disseminated or in quartz sulfide veins or veinlets. Phyllic alteration is later than potassic alteration, in which the new hydrothermal alteration minerals are cut by those of the former. For instance, the quartz vein that formed in the fracture cut the previous hydrothermal orthoclase (Figure 5E).

Propylitic Alteration

Propylitic alteration usually occurs on the outside of phyllic alteration (Figure 4). These propylitic rocks are characterized by a light green color (Figure 5G), mainly containing epidote and chlorite (Figure 5H,I), with a small amount of carbonate minerals (Figure 5H), sericite, quartz, pyrite, etc. The propylitic alteration has the following characteristics: (1) many chlorites and epidotes are generated, replacing mineral phenocryst and groundmass (Figure 5H,I); (2) some carbonate minerals such as siderite and calcite occurred; (3) the content and variety of metal minerals are reduced, while pyrite veins or veinlets are dominant.

Argillic Alteration

Argillic alteration occurs on the outside and upside of porphyry alteration (Figure 4). The argillic rocks are predominantly light white (Figure 5J) and contain mainly kaolinite (Figure 5K,L), as well as a little sericite and quartz. The argillic alteration zone in the Beiya deposit has a relatively narrow range and is mainly distributed near the surface, with strong ferritization accompanying it.

4.1.2. Contact Alteration

The contact zone between alkaline porphyry and the wall-rock is primarily where calcareous skarn develops. The skarn rocks are dark green to greyish green (Figure 5M). Altered residues of limestone and alkaline porphyry can be observed in the skarn alteration zone. Two sub-types of tremolite–diopside skarn and apatite–grossularite–diopside skarn can be divided according to mineral association. The former mainly contains garnet, tremolite, diopside, epidote, chlorite, quartz, pyrite, magnetite, and arsenopyrite (Figure 5N,O), etc. The latter commonly contains apatite, grossularite, diopside, diopside, epidote, quartz, and a small amount of pyrite and magnetite.

4.1.3. Wall-Rock Alteration

Compared with the fresh limestones (Figure 5P) of the Beiya Fm., the altered wall-rocks mainly include marbleization, silicification, and calcitization, with some pyritization and magnetization (Figure 5Q,R). The wall-rock alteration is characterized by the recrystallization of calcite and the formation of calcite veins or clumps, as well as quartz–carbonate–(pyrite or magnetite) veinlets.

4.2. Mineralization Types

In consideration of the close relationship with the alkaline porphyry and structure, five types of mineralization are classified according to ore-controlling structures by Liu et al. [27]: (1) porphyry-type Cu–Au–(Mo), controlled by faults, fractures, and breccia pipes within the alkaline porphyries, while the metal minerals are mainly molybdenite, chalcopyrite, bornite, pyrite, and native gold (Figure 6A,E–G); (2) skarn-type Au–Fe–(Cu), controlled by contact zone structures between alkaline porphyries and wall-rocks (Figure 3), while the metal minerals are mainly magnetite, chalcopyrite, bornite, pyrite, and native gold (Figure 6B,H); (3) hydrothermal vein-type Au–Fe, controlled by interlayer fracture zones, faults, and fractures in the wall-rocks, while the metal minerals are mainly magnetite, pyrite, and native gold (Figure 6C,I); (4) distal low- to medium-temperature hydrothermal-type Pb-polymetallic, controlled by interlayer fracture zones between the Beiya Fm. and the Qingtianbao Fm. and interlayer faults within the Beiya Fm. limestone, while the metal

minerals are mainly galena, sphalerite, and pyrite; (5) oxidizing-leaching enriched-type Au, controlled by paleo-karst caves and angular unconformity between the Beiya Fm. and the Q₁s Fm. (Figure 3), while the metal minerals mainly contain magnetite, hematite, pyrite, and native gold (Figure 6D,J).

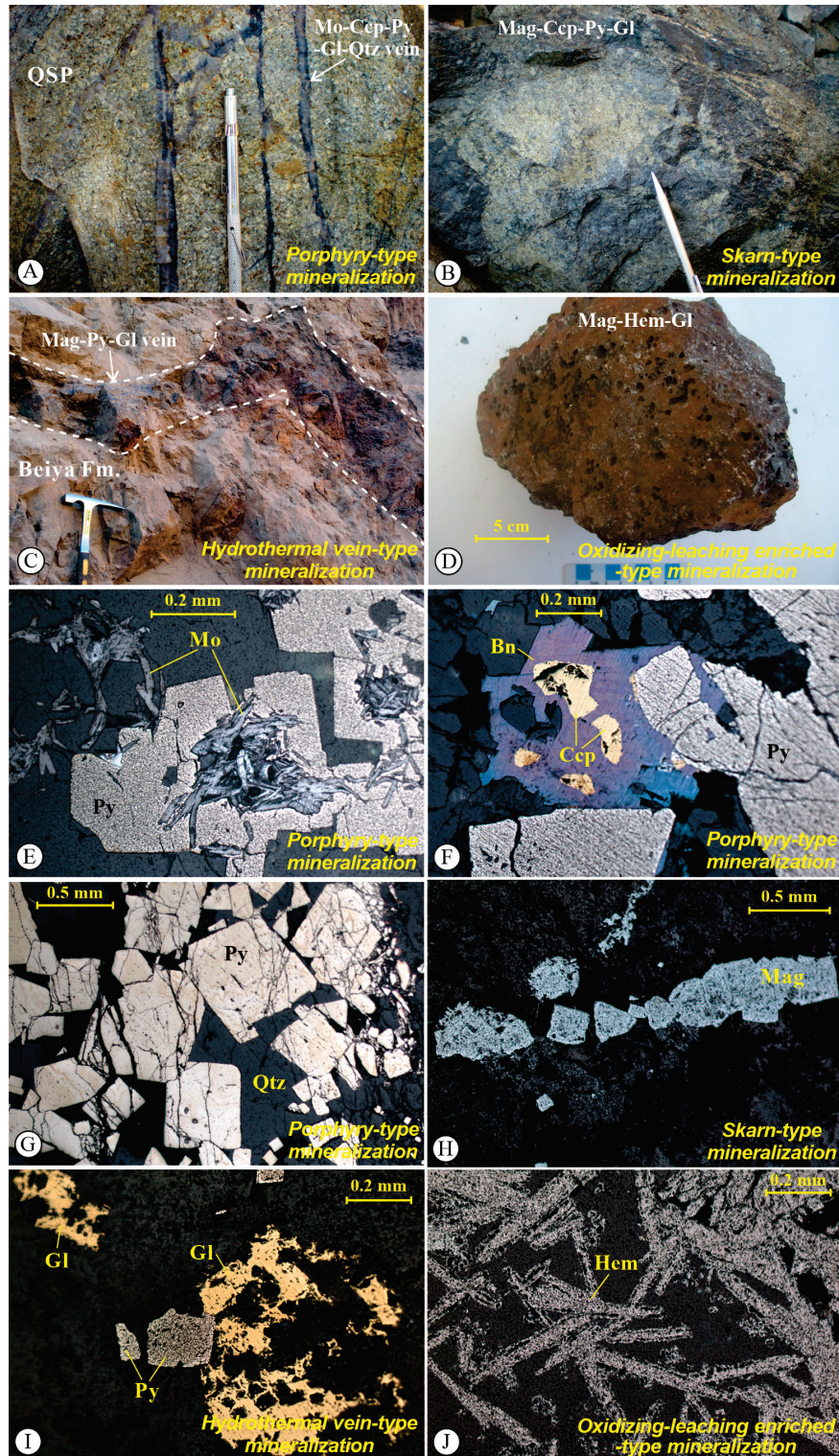


Figure 6. Mineralization characteristics in the Beiya deposit. (A)—porphyry-type mineralization, Au-containing molybdenite, chalcopyrite, pyrite and quartz occur in veins or veinlets; (B)—skarn-type

mineralization, Au-containing magnetite, chalcopyrite and pyrite occur in massive structure; (C)—hydrothermal vein-type mineralization, Au-containing magnetite and pyrite occur within fracture in the Beiya Fm. limestone; (D)—oxidizing-leaching enriched-type mineralization, Au-containing magnetite and hematite generate in honeycomb structure; (E)—porphyry-type mineralization, molybdenite coexist with pyrite; (F)—porphyry-type mineralization, bornite coexist with chalcopyrite and pyrite; (G)—porphyry-type mineralization, pyrite coexist with quartz; (H)—skarn-type mineralization, magnetite occur in vein; (I)—hydrothermal vein-type mineralization, gold coexist with pyrite; (J)—oxidizing-leaching enriched-type mineralization, hematite are developed. Mo = Molybdenite, Ccp = Chalcopyrite, Py = Pyrite, Gl = Gold, Mag = Magnetite, Hem = Hematite, Bn = Bornite, Qtz = Quartz, QSP = Quartz syenite porphyry, Beiya Fm. = Beiya Formation limestone.

4.3. Mineralization and Alteration Zone

Three representative measured sections show the characteristics of mineralization and alteration zoning at Beiya in Figure 3.

4.3.1. C1 Section

The elevation of the C1 section is about 1707 m and it is located within the limestone of the Beiya Fm. (Figure 3). From NE to SW, the mineralization and alteration zones successively are (Figure 4) (I) s cataclastic intense magnetization limestone zone, with Au-bearing magnetite–limonite veins occurring in the interlayer faults; (II) s weak magnetization limestone zone, with a few Au-bearing magnetite–limonite veins filling within interlayer faults or fractures; (III) no mineralized or altered limestone zones; (IV) a cataclastic intense magnetization–limonitization limestone zone, with abundant Au-bearing magnetite–limonite veins or ore-bodies occurring in interlayer faults or fractures. Marbleization, silicification, and calcitization usually develop around ore-bodies or ore-veins, which mainly contain metal minerals such as magnetite, hematite, pyrite, chalcopyrite, bornite, and native gold.

4.3.2. C2 Section

The elevation of the C2 section is about 1758 m and it is located in the limestone of the Beiya Fm. and QSP (Figure 3). From S to E, the mineralization and alteration zones successively are (I) a weak magnetization–limonitization–carbonation limestone zone, with Au-bearing magnetite–limonite veins occurring in the interlayer faults; (II) lamprophyre vein, with weak chloritization and epidotization; (III) skarn-type Au-bearing magnetite–limonite ore-bodies; (IV) skarn zone, formed in the contact zone structure; (V) argillic zone, mainly kaolinization QSP and commonly presenting a degree of limonitization; (VI) propylitic zone, mainly chloritization and epidotization QSP, with weak calcitization, sericitization, and silicification; (VII) phyllic zone, mainly sericitization and silicification QSP; (VIII) potassic zone, mainly potassium feldspathization and biotitization QSP, with weak sericitization and silicification; (IX) telescoped zone of phyllic and potassic alteration.

4.3.3. C6 Section

The elevation of the C6 section is about 1690 m and it is located in the limestone of the Beiya Fm. and QSP (Figure 3). From SW to NE, the mineralization and alteration zones successively are (I) a cataclastic magnetization–limonitization limestone zone; (II) Au-bearing magnetite–limonite ore-bodies; (III) argillic zone, mainly kaolinization QSP, with a few limonitizations; (IV) skarn zone, mainly tremolite–diopside skarn, with an amount of pyrite, magnetite, and arsenopyrite; (V) propylitic zone, mainly chloritization and epidotization QSP, with weak calcitization, sericitization, and silicification; (VI) phyllic zone, mainly sericitization and silicification QSP; (VII) potassic zone, mainly potassium feldspathization and biotitization QSP, with weak sericitization and silicification.

Metallogenic element associations also exhibit a significant zone. From alkaline porphyry outward to the wall-rock in the C2 section, the content of high-temperature met-

allogenic element associations of Cu–Au–Fe tends to gradually decrease, whereas the content of low-temperature metallogenic element associations of Pb–Zn tends to increase. From the C2 section to the C1 section, the ratios of Cu + Au/Cu + Mo, Au + Fe/Cu + Au, Pb + Zn + Ag/Au + Fe increase from 0.90, 0.01, and 103.53 to 0.99, 0.04, and 835.58, respectively (Table 2). These reflect that the metallogenic element associations have a lateral zoning of Cu–Mo → Cu–Au → Au–Fe–Cu → Au–Fe → Pb–Zn–Au–Ag–Fe from alkaline porphyry outward to the wall-rock. From the C6 section to the C2 section, the ratios of Cu + Au/Cu + Mo and Pb + Zn + Ag/Au + Fe increase from 0.88 and 34.04 to 0.90 and 103.53, respectively (Table 2), indicating that the metallogenic element associations have a similar zoning to the lateral zoning from the deep to the shallow surface.

Table 2. Characteristics of major mineralization elements at C1, C2, and C6 section in the Beiya deposit (10^{-6}).

Sections	Sample No.	Mo	Cu	Au	Fe	Pb	Zn	Ag	Ratios of Element Association		
									Cu + Au/ Cu + Mo	Au + Fe/ Cu + Au	Pb + Zn + Ag/Au + Fe
C1	Cp01	3.49	381.00	0.06	3.74	2740.00	838.00	5.72	0.991	0.010	942.588
	Cp03	2.92	251.00	0.57	47.30	2710.00	101.00	26.50	0.991	0.190	59.273
	Cp04	17.15	945.00	2.16	41.20	48,600.00	2390.00	40.70	0.984	0.046	1176.907
	Cp05	5.70	524.00	0.68	12.35	3960.00	829.00	13.20	0.991	0.025	368.578
	Cp07	6.66	1300.00	0.78	36.40	41,900.00	35,000.00	41.30	0.995	0.029	2069.427
	Cp08-1	6.62	3740.00	2.67	40.10	34,300.00	15,000.00	23.10	0.999	0.011	1153.217
	Cp08-2	4.86	1985.00	17.85	19.40	14,050.00	8440.00	17.30	1.006	0.019	604.223
	Cp11	4.89	7630.00	16.00	41.80	27,600.00	3930.00	17.40	1.001	0.008	545.803
	Cp13-1	10.45	914.00	2.35	17.95	7180.00	8260.00	12.30	0.991	0.022	761.197
	Cp13-3	1.49	126.50	0.05	4.79	425.00	2980.00	6.67	0.989	0.038	704.599
mean value		6.42	1779.70	4.32	26.50	18,346.50	7776.80	20.42	0.99	0.04	835.58
C2	Cp14	2.17	29.50	0.01	0.55	214.00	204.00	0.52	0.932	0.019	752.734
	Cp15-1	28.50	180.50	0.02	2.61	673.00	670.00	2.80	0.864	0.015	510.934
	Cp16-3	9.60	1040.00	0.09	5.32	256.00	592.00	7.99	0.991	0.005	158.311
	Cp17-2	165.50	919.00	0.08	7.56	102.50	132.00	7.46	0.847	0.008	31.674
	Cp18-2	94.70	9400.00	16.95	19.8	771.00	52.00	92.20	0.992	0.004	24.903
	Cp21-2	64.60	27,400.00	7.88	23.4	1020.00	84.00	123.00	0.998	0.001	39.226
	Cp50-3	236.00	9100.00	40.8.0	22.3	1310.00	96.00	550.00	0.979	0.007	30.998
	Cp52	86.50	160,500.00	4.24	17.65	119.50	61.00	606.00	0.999	0.0001	35.930
	Cp53-1	51.80	11,350.00	10.10	38.7	1350.00	169.00	129.00	0.996	0.004	33.770
	Cp54-2	54.10	5320.00	4.53	14.4	233.00	256.00	42.20	0.991	0.004	28.061
	Cp56-1	6.87	17,550.00	10.25	32.5	1860.00	26.00	58.60	1.000	0.002	45.488
	Cp57-2	16.35	168,500.00	4.92	18.7	904.00	135.00	442.00	0.100	0.0001	62.701
	Cp58-1	14.55	5650.00	19.25	23.7	1380.00	115.00	63.60	1.001	0.008	36.289
	Cp59-2	70.20	26,400.00	46.60	45	787.00	436.00	285.00	0.999	0.003	16.463
	Cp61-1	196.50	620.00	0.09	1.35	70.30	20.00	2.00	0.759	0.002	64.231
	Cp62	49.10	1540.00	8.42	23.7	302.00	71.00	69.80	0.974	0.021	13.786
	Cp63	115.5	907.00	6.67	22.5	2750.00	141.00	22.70	0.894	0.032	99.887
	Cp65	143.00	2680.00	0.71	34.1	435.00	279.00	5.94	0.950	0.013	20.681
	Cp66-1	482.00	2540.00	17.50	20.8	1570.00	176.00	98.40	0.846	0.015	48.157
	Cp67	62.60	4330.00	0.17	46.20	5160.00	369.00	31.30	0.986	0.011	119.925
mean value		93.00	21,712.00	9.00	20.00	1013.00	194.00	126.00	0.90	0.01	103.53
C6	Cp311	10.45	937.00	0.21	50.00	1135.00	207.00	3.74	0.989	0.054	26.801
	Cp314	2.96	111.50	0.02	16.60	141.50	279.00	1.16	0.974	0.149	25.375
	Cp315	7.92	89.70	0.01	12.20	213.00	377.00	0.91	0.919	0.136	48.384
	Cp316	6.50	180.50	0.01	2.19	46.20	108.00	0.63	0.965	0.012	70.281
	Cp317-1	55.7	46.30	0.02	23.30	31.80	637.00	0.56	0.454	0.503	28.708
	Cp318-1	6.48	190.00	1.05	35.10	88.30	78.00	3.46	0.972	0.189	4.696
mean value		15.00	259.20	0.20	23.20	276.00	281.00	1.70	0.88	0.17	34.04

4.4. Geochemical Characteristics of Mineralization and Alteration Zone

The mean compositions of major and trace elements of the analyzed samples are listed in Table 1. For six unaltered limestones from the Beiya Fm., they have a high content of CaO (34.64 wt.%) and MgO (17.67 wt.%), and a low content of Fe₂O₃* (0.43 wt.%), Al₂O₃ (0.60 wt.%), SiO₂ (0.72 wt.%), MnO, K₂O, Na₂O, and TiO₂ (Table 1). The contents of metallogenic and rare earth elements in unaltered limestones are relatively low, with Cu,

Pb, and Zn contents of 3.60, 10.40, and 17.00×10^{-6} , respectively, while the content of Au and Ag is between 0.01 and 0.02×10^{-6} (Table 1).

While for fourteen altered or mineralized limestones in the wall-rock alteration zone, the abundances of CaO and MgO are 19.50 and 5.02 wt.%, respectively, there is also a high content of Fe_2O_3^* (16.21 wt.%), SiO_2 (6.79 wt.%), Al_2O_3 (2.29 wt.%), and MnO (1.28 wt.%) (Table 1). The contents of major metallogenic and rare earth elements from altered or mineralized limestones are high, with Cu, Pb, and Zn contents of 1540.00, 17,190.00, and 7234.00×10^{-6} , respectively, and Au (3.45×10^{-6}) and Ag (15.70×10^{-6}) also showing a high content (Table 1). Significantly, the content of P_2O_5 between unaltered limestones (0.050 wt.%, $n = 6$) and altered or mineralized limestones (0.053 wt.%, $n = 14$) is almost consistent (Table 1).

Unaltered quartz syenite porphyry has a high content of Al_2O_3 and SiO_2 (16.98 and 67.78 wt.%, respectively, $n = 3$), and a low content of Fe_2O_3^* (2.70 wt.%, $n = 3$), CaO, MgO, MnO, Na_2O , and TiO_2 , with the content of K_2O and P_2O_5 being about 7.11 and 0.060 wt.%, respectively (Table 1). The Cu, Pb, Zn, Au, and Ag contents of unaltered quartz syenite porphyry are about 14.27, 103.40, 82.90, 0.01, and 0.66×10^{-6} , respectively, showing the characteristics of low content.

With regard to the altered quartz syenite porphyry from four different porphyry alteration zones, they have various content and compositions of major and trace elements (Table 1). For instance, the altered quartz syenite porphyries in the potassic zone contain the highest amounts of Fe_2O_3^* (10.53 wt.%), K_2O (7.68 wt.%), Cu ($20,770.00 \times 10^{-6}$), Au (5.88×10^{-6}), and Ag (84.84×10^{-6}) (Table 1), while they have the highest amounts of SiO_2 (69.80 wt.%) and the lowest amounts of K_2O (4.63 wt.%) in the phyllic zone (Table 1). The content of P_2O_5 of altered quartz syenite porphyries, in addition to that in the propylitic zone ($\text{P}_2\text{O}_5 = 0.033$ wt.%), is about 0.061, 0.064, and 0.067 wt.% in the potassic, phyllic, and argillic zones (Table 1), respectively, which is consistent with that of unaltered quartz syenite porphyry.

In the contact alteration zone, exoskarn and endoskarn show different element compositions due to the difference of protoliths. Altered rocks in the exoskarn zone have a relatively high content of CaO (14.10 wt.%) and MgO (9.92 wt.%), while they have a high content of SiO_2 (45.80 wt.%) in the endoskarn zone, which is preferred to the compositional features of the respective protoliths. Both exoskarn and endoskarn have a high content of Fe_2O_3^* (Table 1). The P_2O_5 content varies greatly in the exoskarn zone (0.040 wt.%) and endoskarn zone (0.073 wt.%) (Table 1).

4.5. Element Migration of Mineralization and Alteration Zone

We adopting the equation of rock composition changing with volume change in the process of hydrothermal alteration proposed by Gresens [34] and Grant [35,36], to estimate the introduction and removal of elemental components from different alteration zone. Gresens [34] considered that some components in the rock were stable during the process of hydrothermal alteration, and Grant [35,36] established the corresponding equation of rock volume change, i.e., Equation (1):

$$C_i^A = M^O/M^A (C_i^O + \Delta C_i) \quad (1)$$

For the immobile component, M^O/M^A ratio is constant and $\Delta C_i = 0$, so it can obtain Equation (2):

$$C_i^A = (M^O/M^A) C_i^O \quad (2)$$

C_i^A and C_i^O represent the concentration of a component in the rock after and before alteration, respectively; M^A and M^O represent the equivalent mass of rock after and before alteration, respectively; ΔC_i represents the change in concentration of a component in the rock after alteration. Assuming that the volume and mass of rock remain constant before and after hydrothermal alteration and immobile components in the rock do not vary in concentration, by plotting the values of the bivariable C_i^A and C_i^O , we can create a straight

line of the immobile component through the origin with a slope of M^O/M^A and $\Delta C_i = 0$, i.e., the Isocon line [35,36].

As mentioned above, with the exception of propylitic, exoskarn, and endoskarn zones, the content of P_2O_5 is essentially constant between fresh rocks and altered rocks. For instance, the contents of P_2O_5 from unaltered and altered or mineralized limestones are 0.050 wt.% ($n = 6$) and 0.053 wt.% ($n = 14$), respectively; the same features are also shown between unaltered and altered quartz syenite porphyry, showing 0.060 wt.% ($n = 3$) of P_2O_5 in the former and 0.061–0.067 wt.% in the potassic, phyllic, and argillic zones. Therefore, P_2O_5 is selected as the immobile component by comparing unaltered samples with altered samples. The element migrations of six mineralization and alteration zones, including the wall-rock alteration zone, contact alteration zone, potassic zone, phyllic zone, propylitic zone, and argillic zone, are plotted by the Isocon method in Figure 7.

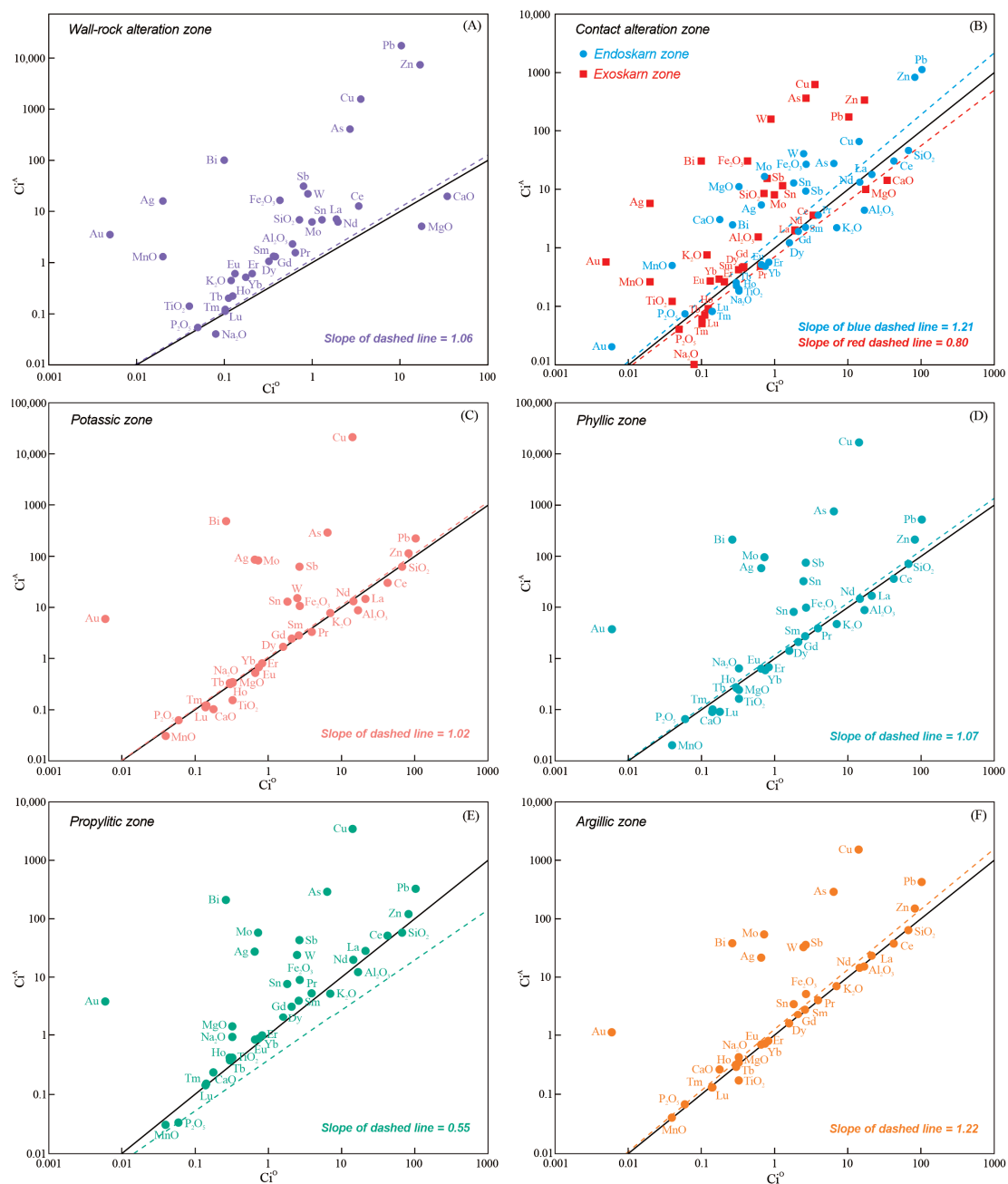


Figure 7. Element migration characteristics of different alteration zones in the Beiya deposit. The slope of the solid line is 1, and the slope of the dashed line is seen in each diagram (A–F).

In the wall-rock alteration zone, the Na₂O, MgO, and CaO contents significantly decrease, while Fe₂O₃^{*}, SiO₂, K₂O, and Al₂O₃ increase. Compared to unaltered limestone, the content of metallogenic elements from altered or mineralized limestones increases, particularly for Pb and Zn, as well as rare earth elements.

In the potassic zone, K₂O, MgO, and Fe₂O₃^{*} are slightly enriched, while SiO₂, Al₂O₃, and CaO slightly decrease; the Cu, Mo, and Bi contents significantly increase, whereas Pb and Zn slightly increase; there is little change in the content of Na₂O and rare earth elements.

In the phyllic zone, Na₂O, SiO₂, and Fe₂O₃^{*} are enriched; however, K₂O, Al₂O₃, CaO and MgO decrease; it is similar to the potassic zone in that the Cu, Mo, and Bi contents significantly increase, while Pb and Zn slightly increase, as well as rare earth elements.

In the argillic zone, Na₂O, CaO, and Fe₂O₃^{*} are enriched, but K₂O, Al₂O₃, and SiO₂ decrease; Au predominates even though the amount of metallogenic elements increases; the contents of MgO and rare earth elements have a minimal change.

In the contact zone and propylitic zone, probably due to the small number of samples, the slope of the Isocon line deviates greatly from 1. However, it can still provide some information about element migration; for instance, CaO and MgO decrease in the exoskarn zone, but they increase in the endoskarn zone; SiO₂, Al₂O₃, and K₂O increase in the exoskarn zone, while decreasing in the endoskarn zone. In a word, the element migration in these two zones is for reference only.

5. Discussion

5.1. Regularity of Mineralization and Alteration Zoning

The typically lateral and vertical alteration zone of porphyry copper deposits was first proposed by Lowell and Guibert [6], including concentric-banded alteration zones of potassic, phyllic, argillic, and propylitic zones, with typical hydrothermal mineral associations within each zone, as well as metal sulfide minerals and related mineralization zones displayed in various occurrences. Sillitoe [2] and John et al. [9] summarized the consistent mineralization and alteration zoning pattern in porphyry copper systems, which comprises sodic–calcic, potassic, chlorite–sericite, sericitic, and advanced argillic zones centrally from the bottom upward or outward. Similar to that of typical porphyry deposits, the Beiya deposit exhibits mineralization–alteration zoning of potassic, phyllic (pyritic–sericitic), propylitic, and argillic alteration from the center outward or upward (Figure 4). However, the propylitic zone is not fully developed and is absent or overprinted by the argillic zone in some areas (Figure 4). Wall-rocks, especially the Beiya Fm. limestones, have an influential impact on mineralization and alteration in the Beiya deposit, just as they do in other porphyry copper deposits worldwide [2], which results in contact and wall-rock alteration and the corresponding hydrothermal vein-type mineralization.

Additionally, it displays mineralization occurrence zoning at Beiya. Porphyry-type mineralization commonly occurs in veins and veinlets; skarn-type mineralization appears as long and thick veins or irregular-shaped veins; hydrothermal vein-type and distal hydrothermal-type mineralization occur in veins, bedded, near bedded, and lenticular shapes; oxidizing-leaching enriched mineralizations are mainly sack-like or irregular in shape. These various occurrences are dominantly controlled by structures and lithologies, which are generally observed in many porphyry and associated deposits [2,4,37,38].

The element migrations of different zones further reflect their alteration and mineralization characteristics. In the potassic and phyllic zone, the high temperature metallogenic element associations of Cu, Mo, Bi, and Au are significantly enriched, while the low temperature metallogenic element association of Pb, Zn, and Au significantly increases outward in wall-rock alteration zone (Table 1; Figure 7), demonstrating an obvious metallogenic element association zoning that consistent with the spatial interrelationships of mineralization in the porphyry copper system [2] and confirming a recent study that the alteration zones in the porphyry deposit result from gradual cooling of a single magma–hydrothermal continuous interaction with porphyry intrusion [12]. The gain and/or loss of oxides in

different zones reflects the primary generation or replacement of hydrothermal minerals. In the wall-rock mineralization-alteration zone, there is a decrease in Na_2O , MgO , and CaO with an increase in Fe_2O_3^* , SiO_2 , K_2O , and Al_2O_3 , reflecting the development of an amount of pyrite, magnetite, quartz, and some clay minerals and the replacement of calcite and dolomite. In the potassic zone, K_2O , MgO , and Fe_2O_3^* are slightly enriched, while SiO_2 , Al_2O_3 , and CaO slightly decrease, indicating the formation of K-feldspar, hydrothermal biotite, and pyrite and the replacement of plagioclase. In the phyllic zone, Na_2O , SiO_2 , and Fe_2O_3 are enriched, with K_2O , Al_2O_3 , CaO , and MgO decrease, reflecting the formation of some quartz, sericite, and pyrite and the replacement of plagioclase or mafic minerals.

5.2. Mechanism of the Mineralization-Alteration Zone

The temperature, sulfur fugacity, and solution pH are the principal controls of sulfide assemblages and the alteration type [2]. And the evolution of ore-forming fluids is not only a crucial component in the formation of various alterations but also a reflection of the ore-forming process [39–42].

Fluid inclusions and isotopes studies indicate that the ore-forming fluids and materials of the Beiya gold-polymetallic deposit originate from the magmatic–hydrothermal system related to alkaline porphyry [23–26,43]. In the early stage of the magmatic–hydrothermal system, the ore-forming fluid is at a high temperature (400–560 °C) and high salinity (38–67 wt.% NaCl equiv.) fluid [24,26]. This early-stage high temperature and salinity ore-forming fluid, on the one hand, causes osmotic potassium metasomatism in the alkaline porphyry [44], resulting in the formation of potassium silicate minerals of K-feldspar and hydrothermal biotite, along with enrichments of K_2O , MgO , and Fe_2O_3^* and losses of SiO_2 , Al_2O_3 , and CaO ; the early garnet-containing skarn alteration-mineralization occurs synchronously [45], which is predominantly induced by fluids with high temperature and salinity of 360–575 °C and 36–69 wt.% NaCl equiv., respectively, as well as a low to moderate salinity of 2–25 wt.% NaCl equiv. conducted in epidotes or garnets within the skarn [26,46]. On the other hand, the early-stage ore-forming fluid extracts metallogenic elements such as Au, Cu, and Mo from the alkaline porphyry, generating Cu–Au–Mo mineralization in the potassic zone and initial Au mineralization in the wall-rocks [44,45].

With the processes of cooling, boiling, and/or mixing with meteoric waters, the ore-forming fluid at Beiya transforms to a predominant low to moderate temperature (310–440 °C) and low salinity (1–15 wt.% NaCl equiv.) fluid [26,46], creating a general phyllic and propylitic alteration, which is similar to sericitic alteration occurring at temperatures between 370 and 450 °C in the Butte porphyry copper–molybdenum deposit [39]. The early garnet-containing skarn undergoes retrograde alteration as well, producing epidote, chlorite, and quartz through fluid metasomatism [45]. In addition, the diffusion and migration ability of Au and Cu decreases at this stage in the thermodynamics state; when the hydrothermal fluid migrates to the lithologic and structural discontinuity interface, Au, Cu, and other metallogenic elements precipitate in these special locations, mostly in phyllic and contact alteration zones (Figures 4 and 6).

In the late-stage evolution of ore-forming fluid, cooling and meteoric water mixing are the main processes to form a low temperature (160–340 °C) and low salinity (1–12 wt.% NaCl equiv.) of the hydrothermal fluid [26,46], which promotes the formation of argillic alteration and migrates Au, Ag, Pb, and Zn outward to form low- to medium-temperature hydrothermal vein- and distal hydrothermal-type mineralization in the wall-rocks (Figures 4 and 6).

5.3. Geological Implication of Element Migration for the Beiya Deposit

Element migration of different mineralization-alteration zones roughly describes the ore-forming process at Beiya; meanwhile, a reasonably inferred mineralization and alteration zoning model of the Beiya gold-polymetallic deposit is established as well (Figure 8). In Figure 8, along with the continuous evolution of ore-forming fluids, it generates concentric-banded alteration zone from the center of the alkaline porphyry

outward or upward, i.e., [porphyry alteration] potassic zone → phyllic zone → propylitic zone → argillic zone → [contact alteration] skarnitization-marbleization zone → [wall-rock alteration] marbleization-silicification-calcitization zone. In potassic and phyllic zones, it develops porphyry-type native gold–chalcopyrite–bornite–molybdenite–pyrite mineralization. Skarn-type native gold–pyrite–magnetite–chalcopyrite–bornite–siderite mineralization occurs outward in the contact alteration zone. Proximal hydrothermal vein-type native gold–pyrite–magnetite mineralization mainly distributes in the marbleization Beiya Fm. limestone, while distal hydrothermal-type galena–sphalerite–native gold–native silver mineralization predominantly forms in silicification and calcitization alteration zones. Metallogenic element association also exhibits a significant zone, laterally and vertically, of Cu–Mo → Cu–Au → Au–Fe–Cu → Au–Fe → Pb–Zn–Au–Ag–Fe from alkaline porphyry outward to the wall-rock.

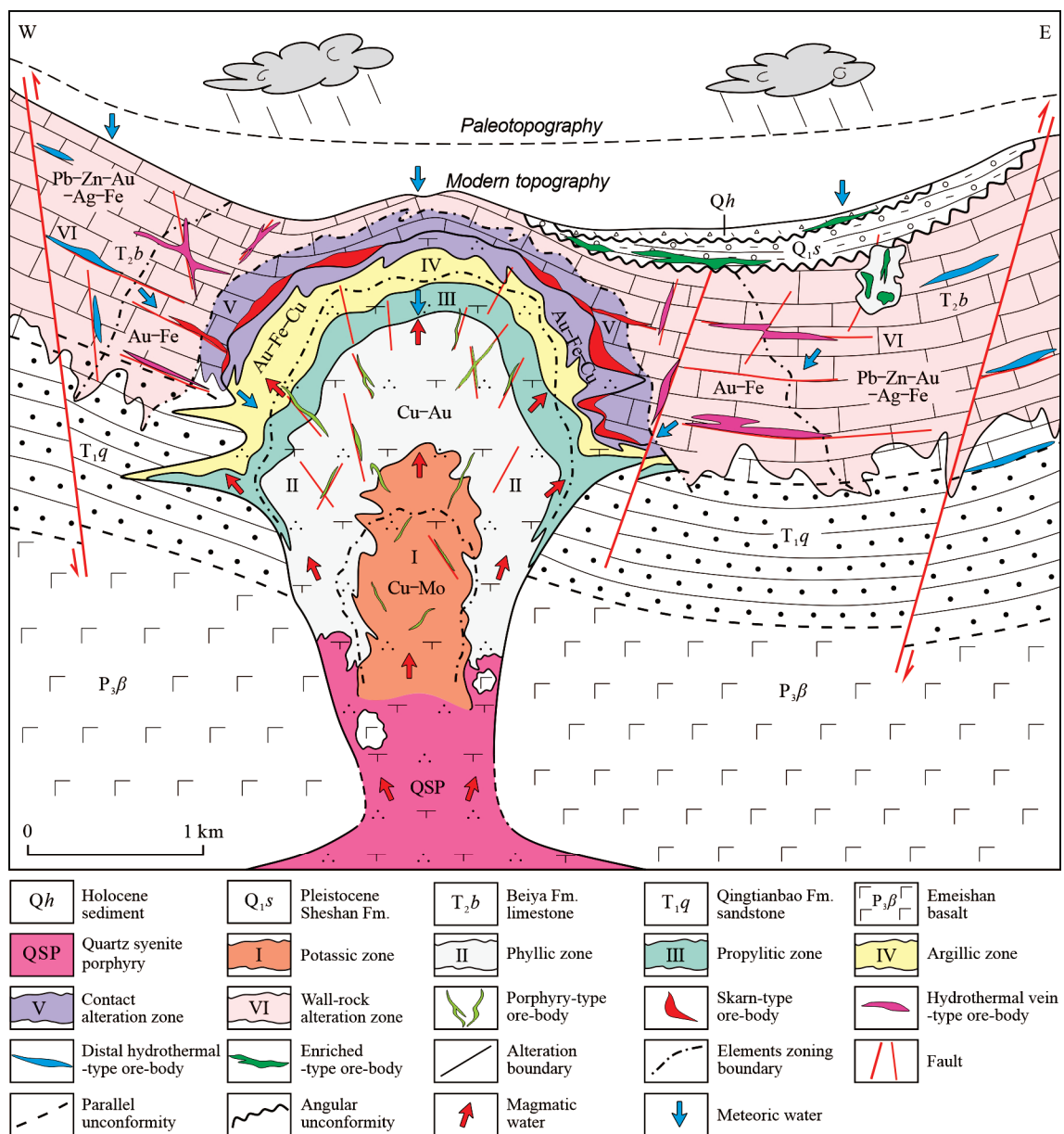


Figure 8. Mineralization and alteration zoning model of the Beiya deposit.

This provides a new basis for exploration in the Beiya deposit from the element migration and its related mineralization and alteration zoning model (Table 3). In the deep area and center of the porphyry that associated with mineralization, it is beneficial to find

the porphyry-type Cu–Mo–Au ore-bodies or deposits. In the space between wall-rock and ore-forming porphyry, skarn-type Au–Fe–Cu ore-bodies or deposits are the most favorable prospecting object. Distal to the ore-forming porphyry, the hydrothermal vein-type Au–Fe and Pb–Zn–Au–Ag–Fe ore-bodies or deposits are formed and distributed; moreover, the Pb-polymetallic deposit found at Luping [47,48] located 2 km north of the Beiya deposit is a good example.

Table 3. The brief information for prospecting prediction of the Beiya deposit.

Metallogenic System/Space	Ore-Controlling Structure	Mineralization and Alteration Type	Metallogenic Element Association	Mineralization Occurrence
oxidizing-leaching metallogenic system	angular unconformity interface, paleo-karst cave	calcitization silicification limonitization	Au–(Fe)	capsule-like irregular-shaped
porphyry metallogenic system	wall-rock (distal)	interlayer fracture zone, interlayer fault	Pb–Ag–Zn–(Au)–(Fe)	layered banded
	wall-rock (proximal)	interlayer fault, fracture and joint	Au–Fe	vein lenticular
	contact zone	contact zone structure	Au–Fe–(Cu)	irregular-shaped
	porphyry	fault, fracture and joint, breccia pipe	Cu–Au–(Mo)	stockwork veinlet

6. Conclusions

Based on the identification of types and features of mineralization and alteration, with a detailed study on the element migration of mineralization-alteration zones, allowing us to establish the mineralization and alteration zoning model at Beiya, we draw the following conclusions.

- (1) Three types of porphyry alteration, contact alteration, and wall-rock alteration are developed; and the porphyry alteration comprises four alteration zones of potassic, phyllic, propylitic, and argillic.
- (2) Five types of mineralization with various occurrence characteristics and distinctive metallogenic element associations are dominantly controlled by structures and lithologies.
- (3) Concentric-banded mineralization-alteration zones occur centrally from alkaline porphyry outward or upward, presenting common alteration zones of potassic → phyllic → propylitic → argillic → skarn alteration → wall-rock alteration, accompanying with individual elements gains and losses in different zones.
- (4) An inferred mineralization and alteration zoning model indicates the Beiya deposit is similar to the typical porphyry copper systems, and it also deepens the understanding of metallogenic regularity and provides scientific information for prospecting at Beiya, as well as at the same type of deposits in the Sanjiang metallogenic belt and elsewhere worldwide.

In addition, we suggest that in conjunction with the study of the properties and evolution of ore-forming fluids, isotopes and alteration minerals' microanalysis, among

others, studies will obtain a more reliable research results with respect to the element migration of mineralization-alteration zones.

Author Contributions: Conceptualization, R.H.; methodology, F.L.; software, Y.G.; validation, R.H., F.L., and Y.G.; formal analysis, F.L.; investigation, R.H., F.L., Y.G., M.W. and W.T.; resources, R.H. and F.L.; data curation, F.L. and Y.G.; writing—original draft preparation, F.L.; writing—review and editing, F.L.; visualization, F.L., Y.G., M.W. and W.T.; supervision, R.H.; project administration, R.H.; funding acquisition, R.H. and F.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Project funded by China Postdoctoral Science Foundation (2022M711437), the Program of China Geological Survey (12120113095900), the Program of “Yunling Scholar” of Yunnan Province (2014), the Projects of Yunnan Engineering Laboratory of Mineral Resources Prediction and Evaluation (YM Lab) (2010), and the Innovation Team of Yunnan province and KMUST (2008, 2012). And the APC was funded by the Projects of Yunnan Engineering Laboratory of Mineral Resources Prediction.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: Author Wei Tan was employed by the company Sino-Zijin Resources Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Richards, J.P. Tectono-magmatic precursors for porphyry Cu-(Mo-Au) deposit formation. *Econ. Geol.* **2003**, *98*, 1515–1533. [CrossRef]
- Sillitoe, R.H. Porphyry copper systems. *Econ. Geol.* **2010**, *105*, 3–41. [CrossRef]
- John, D.A.; Taylor, R.D. By-products of porphyry copper and molybdenum deposits. In *Rare Earth and Critical Elements in Ore Deposits*; Verplanck, P.L., Hitzman, M.W., Eds.; Society of Economic Geologists, Inc.: Littleton, CO, USA, 2016; Volume 18, pp. 137–164.
- Tosdal, R.M.; Richards, J.P. Magmatic and structural controls on the development of porphyry Cu $\pm$ Mo $\pm$ Au deposits. *Soc. Econ. Geol. Rev.* **2001**, *14*, 157–181.
- Piquer, J.; Sanchez-Alfaro, P.; Pérez-Flores, P. A new model for the optimal structural context for giant porphyry copper deposit formation. *Geology* **2021**, *49*, 597–601. [CrossRef]
- Lowell, J.D.; Guilbert, J.M. Lateral and vertical alteration-mineralization zone in porphyry ore deposits. *Econ. Geol.* **1970**, *65*, 373–408. [CrossRef]
- Hollister, V.F.; Potter, R.R.; Barker, A.L. Pophyry-type deposit of Appalachian Orngen. *Econ. Geol.* **1974**, *69*, 618–630. [CrossRef]
- Hollister, V.F. Models for prospect eraluation of porphyry copper deposits. In *Porphyry Copper, Molybdenum, and Gold Deposit, Volcanogenic Deposits (Massive Sulfides), and Deposits in Layered Rock*; Hollister, V.F., Ed.; Society for Mining, Metallurgy and Exploration: New York, NY, USA, 1991; Volume 3, p. 5210.
- John, D.A.; Ayuso, R.A.; Barton, M.D.; Blakely, R.J.; Bodnar, R.J.; Dilles, J.H.; Gray, F.; Graybeal, F.T.; Mars, J.C.; McPhee, D.K.; et al. *Porphyry Copper Deposit Model. Chap. B of Mineral Deposit Models for Resource Assessment*; U.S. Geological Survey Scientific Investigations Report 2010–5070–B; U.S. Geological Survey: Reston, VA, USA, 2010; 169p.
- Ulrich, T.; Heinrich, C.A. Geology and Alteration Geochemistry of the Porphyry Cu-Au Deposit at Bajo de la Alumbrera, Argentina. *Econ. Geol.* **2002**, *97*, 1865–1888. [CrossRef]
- Cannell, J.; Cooke, D.R.; Walshe, J.L.; Stein, H. Geology, mineralization, alteration, and structural evolution of the El Teniente porphyry Cu-Mo deposit. *Econ. Geol.* **2005**, *100*, 979–1003. [CrossRef]
- Zhang, S.Y.; He, W.Y.; Gao, X.; Tian, C.H.; Xiao, Y.W. Alteration zonation and metallogenic mechanism of porphyry copper deposits: A case study of thermodynamic equilibrium simulation of fluid-rock interactions in Yulong deposit. *Acta Petrol. Sin.* **2024**, *40*, 1837–1852. (In Chinese with English abstract) [CrossRef]
- Kubrakova, I.V.; Fortygin, A.V.; Lobov, S.G.; Koshcheeva IYa Tyutyunnik, O.A.; Mironenko, M.V. Migration of platinum, palladium, and gold in the water systems of platinum deposits. *Geochem. Int.* **2011**, *49*, 1072–1084. [CrossRef]
- Ye, R.; Zhang, B.; Wang, Y. Mechanism of the migration of gold in desert regolith cover over a concealed gold deposit. *Geochem. Explor. Environ. Anal.* **2015**, *15*, 62–71. [CrossRef]
- Jiu, B.; Huang, W.H.; Li, Y. The origin, migration, and accumulation mechanism of germanium and the metallogenic model of coal-hosted Ge ore deposits in Wulantuga, Erlan Basin, China. *J. Geochem. Explor.* **2021**, *226*, 106779. [CrossRef]

16. He, Z.H.; Zhou, Y.M.; He, W.Y.; Su, G.S.; Li, W.H.; Yang, S.W. Genetic types and metallogenic regularity of Beiya superlarge gold-polymetallic deposit, northwestern Yunnan. *Miner. Depos.* **2013**, *32*, 244–258. (In Chinese with English abstract)
17. Xu, X.W.; Cai, X.P.; Song, B.C.; Zhang, B.L.; Ying, H.L.; Xiao, Q.B.; Wang, J. Petrologic, chronological and geochemistry characteristics and formation mechanism of alkaline porphyries in the Beiya gold district, western Yunnan. *Acta Petrol. Sin.* **2006**, *22*, 631–642. (In Chinese with English abstract)
18. Xu, X.W.; Zhang, B.L.; Qin, K.Z.; Mao, Q.; Cai, X.P. Origin of lamprophyres by the mixing of basic and alkaline melts in magma chamber in Beiya area, western Yunnan, China. *Lithos* **2007**, *99*, 339–362. [CrossRef]
19. He, W.Y.; Yu, X.H.; Mo, X.X.; He, Z.H.; Li, Y.; Huang, X.K.; Su, G.S. Genetic types and the relationship between alkaline intrusion and mineralization of Beiya gold-polymetallic ore field, western Yunnan Province, China. *Acta Petrol. Sin.* **2012**, *28*, 1401–1412. (In Chinese with English abstract)
20. Deng, J.; Wang, Q.F.; Li, G.J.; Hou, Z.Q.; Jiang, C.Z.; Danyushevsky, L. Geology and genesis of the giant Beiya porphyry-skarn gold deposit, northwestern Yangtze Block, China. *Ore Geol. Rev.* **2015**, *70*, 457–485. [CrossRef]
21. Liu, B.; Liu, H.; Zhang, C.Q.; Mao, Z.H.; Zhou, Y.M.; Huang, H.; He, Z.H. Gangsheng Su Geochemistry and geochronology of porphyries from the Beiya gold-polymetallic orefield, western Yunnan, China. *Ore Geol. Rev.* **2015**, *69*, 360–379. [CrossRef]
22. Fu, Y.; Sun, X.M.; Lin, H.; Zhou, H.Y.; Li, X.; Ouyang, X.Q.; Jiang, L.Y.; Shi, G.Y.; Liang, Y.H. Geochronology of the giant Beiya gold-polymetallic deposit in Yunnan Province, Southwest China and its relationship with the petrogenesis of alkaline porphyry. *Ore Geol. Rev.* **2015**, *71*, 138–149. [CrossRef]
23. Wu, K.X.; Hu, R.Z.; Bi, X.W.; Peng, J.T.; Su, W.C.; Chen, L. Study of the fluid inclusions in altered porphyries in the Beiya gold deposit, Western Yunnan. *J. Mineral. Petrol.* **2005**, *25*, 20–26. (In Chinese with English abstract)
24. Xiao, X.N.; Yu, X.H.; Mo, X.X.; Yang, G.L.; Li, Y.; Huang, X.L. A study of fluid inclusions from Beiya gold-polymetallic deposit in Western Yunnan. *Earth Sci. Front.* **2009**, *16*, 250–261. (In Chinese with English abstract)
25. Li, W.C.; Wang, J.H.; He, Z.H.; Dou, S. Formation of Au-polymetallic ore deposits in alkaline porphyries at Beiya, Yunnan, Southwest China. *Ore Geol. Rev.* **2016**, *73*, 241–252. [CrossRef]
26. He, W.Y.; Yang, L.Q.; Brugger, J.; McCuaig, T.C.; Lu, Y.J.; Bao, X.S.; Gao, X.Q.; Lu, Y.G.; Xing, Y.L. Hydrothermal evolution and ore genesis of the Beiya giant Au polymetallic deposit, western Yunnan, China: Evidence from fluid inclusions and H–O–S–Pb isotopes. *Ore Geol. Rev.* **2017**, *90*, 847–862. [CrossRef]
27. Liu, F.; Wang, L.; Han, R.S.; Gou, Y.X.Y.; Wang, M.Z.; Tan, W. Ore-field structural system and its ore-controlling processes of the Beiya porphyry gold polymetallic deposit in northwestern Yunnan, China. *Earth Sci. Front.* **2017**, *24*, 208–224.
28. Mao, J.W.; Zhou, Y.M.; Liu, H.; Zhang, C.Q.; Fu, D.G.; Liu, B. Metallogenic setting and ore genetic model for the Beiya porphyry-skarn polymetallic Au orefield, western Yunnan, China. *Ore Geol. Rev.* **2017**, *86*, 21–34. [CrossRef]
29. Zhou, Y.M.; Zhang, C.Q.; He, Z.H.; Liu, B.; Wang, L.D. The characteristics marks of metallogenesis of Beiya gold-polymetallic deposit in the Northwest Yunnan province. *Contrib. Geol. Miner. Resour. Res.* **2018**, *33*, 1–14. (In Chinese with English abstract)
30. Zhou, Y.M.; Zhou, G.W.; Zhang, C.Q.; Wang, L.D.; Yu, H.P.; Li, W.H.; Liu, Z.R. Characteristics of Metallogenetic Structures in the Beiya Gold-polymetallic Deposit in Northwestern Yunnan Province and Geological and Exploration Applications. *Geotecton. Et Metallog.* **2021**, *45*, 308–326. (In Chinese with English abstract)
31. Liu, F.; Han, R.S.; Wang, L.; Gou, Y.X.Y.; Wang, M.Z.; Tan, W. Mechanism of rock- and ore-controlling structures in the Wandongshan ore block of the Beiya super-large porphyry-type poly-metallic gold deposit, northwestern Yunnan. *Geotecton. Et Metallog.* **2016**, *40*, 266–280. (In Chinese with English abstract)
32. Sun, J.C.; Han, R.S. *Theory and Method of ore-Field Geomechanics*; Science Press: Beijing, China, 2016; pp. 332–349. (In Chinese)
33. Guo, Y.X.Y.; Han, R.S.; Wang, L.; Liu, F.; Wang, M.Z.; Tan, W. Mineralogical characteristics of feldspar and its metallogenetic significance in the Beiya porphyry-type gold polymetallic deposit, NW Yunnan. *Miner. Explor.* **2016**, *7*, 951–957. (In Chinese with English abstract)
34. Gresens, R.L. Composition–volume relationships of metasomatism. *Chem. Geol.* **1967**, *2*, 47–55. [CrossRef]
35. Grant, J.A. The isocon diagram—a simple solution to Gresens’ equation for metasomatic alteration. *Econ. Geol.* **1986**, *81*, 1976–1982. [CrossRef]
36. Grant, J.A. Isocon analysis: A brief review of the method and applications. *Phys. Chem. Earth* **2005**, *30*, 997–1004. [CrossRef]
37. Cooke, D.R.; Hollings, P.; Walshe, J.L. Giant porphyry deposits: Characteristics, distribution, tectonic controls. *Econ. Geol.* **2005**, *100*, 801–818. [CrossRef]
38. Simmons, S.F.; White, N.C.; John, D.A. *Geological Characteristics of Epithermal Precious and Base Metal Deposits*; Economic Geology 100th Anniversary Volume; Society of Economic Geologists, Inc.: Littleton, CO, USA, 2005; pp. 485–522.
39. Rusk, B.G.; Reed, M.H.; Dilles, J.H. Fluid Inclusion Evidence for Magmatic-Hydrothermal Fluid Evolution in the Porphyry Copper-Molybdenum Deposit at Butte, Montana. *Econ. Geol.* **2008**, *103*, 307–334. [CrossRef]
40. Li, N.; Ulrich, T.; Chen, Y.J.; Thomsen, T.B.; Pease, V.; Pirajno, F. Fluid evolution of the Yuchiling porphyry Mo deposit, East Qinling, China. *Ore Geol. Rev.* **2012**, *48*, 442–459. [CrossRef]
41. Catchpole, H.; Kouzmanov, K.; Putlitz, B.; Seo, J.H.; Fontbote, L. Zoned base metal mineralization in a porphyry system: Origin and evolution of mineralizing fluids in the Morococha district, Peru. *Econ. Geol.* **2015**, *110*, 39–71. [CrossRef]
42. Audétat, A. The metal content of magmatic-hydrothermal fluids and its relationship to mineralization potential. *Econ. Geol.* **2019**, *114*, 1033–1056. [CrossRef]

43. Liu, H.Q.; Bi, X.W.; Lu, H.Z.; Hu, R.Z.; Lan, T.G.; Wang, X.S.; Huang, M.L. Nature and evolution of fluid inclusions in the Cenozoic Beiya gold deposit, SW China. *J. Asian Earth Sci.* **2018**, *161*, 35–56. [CrossRef]
44. Xu, S.M. Metallogenic Modeling of the Beiya Gold Deposit in Western Yunnan and Its Relation to the Cenozoic Alkali-Rich Porphyries. Ph.D. Thesis, China University of Geoscience (Beijing), Beijing, China, 2007; pp. 1–115. (In Chinese with English abstract)
45. He, W.Y. The Beiya Giant Gold-Polymetallic Deposit: Magmatism and Metallogenic Model. Ph.D. Thesis, China University of Geoscience (Beijing), Beijing, China, 2014; pp. 1–154. (In Chinese with English abstract)
46. Wang, J.H.; Li, W.C.; Wang, K.Y.; Yin, G.H.; Wu, S.; Jiang, W.T. The characteristics and evolution of the ore-forming fluids in the Beiya porphyry Au-polymetallic deposit, western Yunnan. *Acta Petrol. Sin.* **2015**, *31*, 3269–3280. (In Chinese with English abstract)
47. Dou, S.; Liu, J.S.; Guo, Y.S.; Yan, J. Ore mineralogical characteristics of the Luping lead polymetallic deposit in Heqing of Yunnan province and their geological implications. *Acta Geosci. Sin.* **2013**, *34*, 87–94. (In Chinese with English abstract)
48. Zhou, J.X.; Dou, S.; Huang, Z.L.; Cui, Y.L.; Ye, L.; Li, B.; Gan, T.; Sun, H.R. Origin of the Luping Pb deposit in the Beiya area, Yunnan Province, SW China: Constraints from geology, isotope geochemistry and geochronology. *Ore Geol. Rev.* **2016**, *72*, 179–190. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Review

Frictional Experiments on Granitic Faults: New Insights into Continental Earthquakes and Micromechanical Mechanisms

Huiru Lei ¹, Shimin Liu ^{1,2,*} and Wenhao Dai ^{1,3,*}

¹ State Key Laboratory of Earthquake Dynamics and Forecasting, Institute of Geology, China Earthquake Administration, Beijing 100029, China; leihuiru@ies.ac.cn

² School of Fire Protection Engineering, China People's Police University, Langfang 065000, China

³ Jiangxi Earthquake Agency, Nanchang 330026, China

* Correspondence: liushimin@cpgu.edu.cn (S.L.); daiwh19@163.com (W.D.)

Abstract: Granitic faults within the crystalline upper-to-middle continental crust play a critical role in accommodating tectonic deformation and controlling earthquake nucleation. To better understand their frictional behavior, we review experimental studies conducted under both dry and hydrothermal conditions using velocity-stepping (VS), constant-velocity (CV), and slide-hold-slide (SHS) tests. These approaches allow the quantification of frictional strength, velocity dependence, and healing behavior across a range of conditions. Our synthesis highlights that the friction coefficient of granite gouges decreases with increasing temperature and pore fluid pressure, decreasing slip velocity, and increasing slip displacement. The velocity-weakening regime shifts to higher temperatures with increasing slip velocity or decreasing pore fluid pressure. Temperature, normal stress, pore fluid pressure, and slip velocity interact to modulate frictional stability. In particular, microstructural observations reveal that grain size reduction, pressure solution creep, and fluid-assisted chemical processes are key mechanisms governing transitions between velocity-weakening and velocity-strengthening regimes. These insights support the growing application of microphysical-based models, which integrate micromechanical processes and offer improved extrapolation from the laboratory to natural fault systems compared to classical rate-and-state friction laws. The collective evidence underscores the importance of considering fault rheology in a temperature- and fluid-sensitive context, with implications for interpreting seismic cycle behavior in continental regions.

Keywords: continental earthquake; granitic fault; frictional strength; fault stability; microphysical model

1. Introduction

Earthquakes on crustal faults are dominated by granitoid gouges, which mostly range between depths of 2 and 18 km and provide information about the location and style of deformation of the fault in this upper-crustal “seismogenic” zone (e.g., the San Andreas Fault). With increasing temperature (depth), the frictional (brittle)–viscous (ductile) transition in these fault gouges plays a crucial role, as it represents a significant shift in the rheological behavior of the Earth’s lithosphere. This transition also aids in defining the lower limit of the seismogenic zone, which is the region where large-magnitude earthquakes typically initiate [1,2], such as the depth distribution of earthquakes in the central creeping section of the San Andreas Fault [3–5], the Kellyland Fault in eastern Maine [6], the Alpine Fault in New Zealand [7], and the Alps Fault [8], as well as the Anninghe–Xianshuihe Fault zone on the southeastern boundary of the Tibetan Plateau [9,10]. Understanding

the depth distribution and sliding characteristics of seismogenic zones is crucial, as it not only delineates where large-magnitude earthquakes are likely to initiate but also provides insights into the diverse slip mechanisms that characterize fault zones. Nevertheless, the direct observation of fault mechanics at seismogenic depths remains technically challenging due to the extreme in situ conditions. In this context, laboratory friction experiments provide a valuable approach to simulate and systematically investigate the physical and chemical processes that control fault strength and stability under controlled temperature, pressure, pore fluid, and slip rate conditions that are representative of various depths within the continental crust.

Fault slip events can span a range of scales, from a few millimeters to several hundred kilometers, and occur over time scales ranging from fractions of a second to millions of years, which causes the distribution of slip thickness on field faults to range from ~1 mm to 1 cm [11,12]. A complex multi-mode slip is associated with various types of earthquake phenomena [13], including regular earthquakes [14], slow-slip events [15,16], low-frequency earthquakes [17,18], episodic tremor and slip [19], and aseismic creep [20,21]. In terms of crustal deformation, the loading cycle is often divided into four phases, preseismic, coseismic, postseismic, and interseismic, called the earthquake periods. The concept of the seismic cycle was first formalized by H.F. Reid [22], who proposed the elastic rebound theory based on his investigation of the 1906 San Francisco earthquake along the San Andreas Fault. According to this model, tectonic motion causes elastic strain to accumulate on faults locked by friction until failure occurs, and stored energy is suddenly released during an earthquake [23]. While this conceptual framework provides a first-order understanding of earthquake occurrence, the detailed physical processes governing strain accumulation, fault strength, and instability remain incompletely understood. Laboratory friction experiments offer a controlled means to investigate these processes by replicating the stress conditions, slip behaviors, and microstructural evolution of fault zones, thereby providing empirical constraints on the mechanics of the seismic cycle at different depths. Bowden and Tabor [24] developed the adhesion friction theory, which revolutionized the understanding of the frictional behavior of stick–slip instability at the microscopic level. Brace and Byerlee [25] inferred that the recurrence of stick–slip could be a possible mechanism for shallow-focus earthquakes. Moreover, studies on the preseismic period focus on the analysis of frictional weakening or strengthening before failure, the development of nucleation zones, and the study of slow-slip events [26–28]. Coseismic deformation is replicated by rapidly increasing the slip rate to simulate fault rupture, including the role of dynamic weakening, temperature changes, and the potential formation of a fluid film or molten material at the fault interface [29,30]. The processes of the postseismic period examine afterslip and time-dependent healing processes, such as fault restrengthening and viscoelastic effects [2], aftershock distributions [31], and the interaction between small repeated earthquakes [32,33]. Research in the interseismic period focuses on the rate-and-state friction (RSF) laws, which describe how friction evolves with time and the slip rate [26,34,35].

There are differences in shear strength between different shape types in rock and gouges [4]. Experiments have demonstrated the effects of temperature, sliding rate, pore fluid pressure, normal stress, and shear displacement, as well as microstructure, on friction properties and associated seismic behaviors [9,36–41]. Friction experiments have been conducted using different types of instruments, including the rotary shear apparatus [42,43]; saw-cut apparatus [44]; direct-shear apparatus [45]; and double-direct shear [28] (Figure 1). Within the RSF framework, two experimental approaches are particularly important in investigating frictional properties: (a) velocity stepping, which is ideal for investigating how friction varies with sliding velocity and understanding velocity dependence [46,47],

and (b) slide-hold-slide, which focuses on the healing or strengthening of the fault interface during periods of inactivity [48,49].

The frictional behavior of granitoid material under dry and hydrothermal conditions has been extensively studied through laboratory experiments, limited in situ studies, and numerical modeling. However, there appears to be a lack of review articles focusing on frictional strength and slip stability in relation to continental earthquake nucleation and the latest friction models. Marone [27] reviewed the standard RSF model, highlighting frictional healing at room temperature and discussing earthquake afterslip in relation to laboratory data within this model framework. Hu and Sun [50] investigated the impact of temperature and pore fluid pressure on the nominal friction coefficient of rocks but did not consider the rate-and-state effects and frictional stability. Also, Mei et al. [51] emphasized the influence of temperature on hydrothermal conditions but did not consider the effects of pore fluids, slip rate, and healing mechanisms on the seismic cycle. In contrast to these reviews, the present study specifically addresses the temperature, rate, and pore fluid pressure dependence of frictional stability, focusing on the velocity dependence of rock friction and physically based friction models.

We first review the standard RSF laws, followed by an extensive review of the results from hydrothermal friction experiments on granitic gouges, emphasizing the steady-state friction (including healing) and frictional stability parameters. Next, we analyze the deformation mechanisms that influence frictional strength and slip stability and summarize physically based constitutive friction laws. Finally, we propose several open questions for future research and highlight the key conclusions.

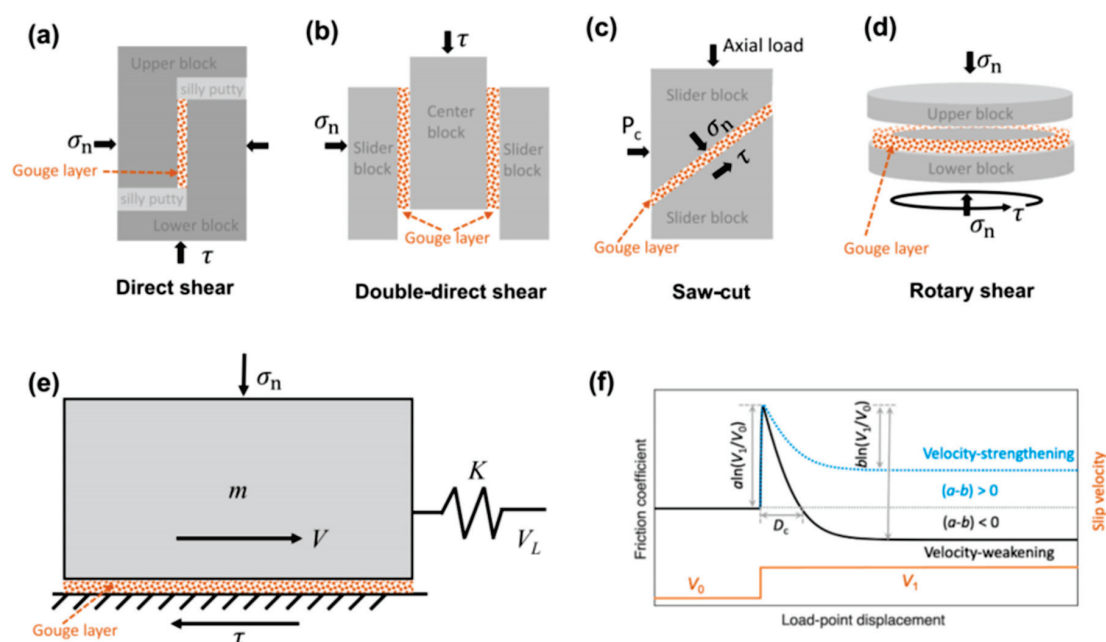


Figure 1. The laboratory fault systems. The sketches of (a) the direct shear setup [48], (b) double-shear setup [28], (c) saw-cut setup [37], and (d) rotary shear setup used for the laboratory friction experiments. The frictional interface is assumed to be filled with granular gouges or simply a bare rock surface. (e) The simulated fault systems can be simplified into a spring–slider system. The slider is driven by a load point velocity of V_L and slides along a frictional interface at a slip velocity of V via a spring with stiffness K . The frictional resistance is τ . (f) A schematic drawing of the rate- and state-dependent friction parameter estimation during a change of load–point velocity from V_0 to V_1 , Figure reproduced from [51] with permission.

2. Friction Motions and Friction Law

Through extensive experiments, Dieterich [26] and Ruina [35] demonstrated that rock friction is influenced by both sliding velocity (V), a state variable (θ), and normal stress (σ). Dieterich interpreted D_c as representing the slip necessary to renew surface contacts; in that case, the ratio $\frac{D_c}{V}$ defines the average contact lifetime θ . They proposed that the frictional shear stress τ could be described as

$$\tau = \bar{\sigma} \left[\mu_0 + a \ln \left(\frac{V}{V_0} \right) + b \ln \left(\frac{V_0 \theta}{D_c} \right) \right] \quad (1)$$

where $\bar{\sigma}$ is the effective normal stress equal to the difference between the total normal stress (σ_n) and pore fluid pressure (P_f), and μ_0 is the steady-state friction coefficient at a reference sliding velocity of V_0 .

The meanings of the parameters a , b and D_c can be illustrated by an idealized velocity stepping test (Figure 1f) of a spring block system (Figure 1e). If the block is sliding at a steady velocity of V_0 and the velocity is suddenly increased to V_1 (a sudden change requires an infinitely stiff spring), the friction suddenly increases by $a \ln(\frac{V_1}{V_0})$. The friction stress then decays by an amount $b \ln(\frac{V_1}{V_0})$ roughly exponentially over a characteristic distance D_c . If $(a - b) > 0$, the new steady-state shear stress is greater than the old one, and the behavior is said to be “velocity strengthening”. If $(a - b) < 0$, the new steady-state shear stress is below the old one, and the behavior is “velocity weakening”.

The decay of the shear stress is governed by the evolution of the state variable θ . This is described by either the aging law [26]

$$\frac{d\theta}{dt} = 1 - \frac{V\theta}{D_c} \quad (2)$$

or the slip law (Ruina, 1983) [35]

$$\frac{d\theta}{dt} = -\frac{V\theta}{D_c} \ln \left(\frac{V\theta}{D_c} \right) \quad (3)$$

The main difference between the two descriptions is that for zero velocity, θ increases linearly with time for the aging law, but for the slip law, it changes only with non-zero velocity. Nevertheless, both laws give the same value for the steady state $\theta = 0$, $\theta_{ss} = \frac{D_c}{V}$ and linearize to the same expression. Notwithstanding this difference, the two evolution laws (2a) and (2b) also have several common features. First, under a steady state, the state variable is a constant given by $\theta_{ss} = \frac{D_c}{V}$ and in particular, $\theta_* = \frac{D_c}{V_*}$. Second, the steady-state frictional strength, Equation (1) becomes

$$\mu_{ss} = \mu_0 + (a - b) \ln \left(\frac{V_{ss}}{V_0} \right), \quad (4)$$

and the behavior is “velocity weakening” or “velocity strengthening” depending on the sign of $(a - b)$.

The spring block system is an adequate representation of experiments in which slip occurs roughly simultaneously over the frictional surface within the precision of measurements. However, for a slip in a continuum, as in the earth, the stiffness depends on the length of the slipping zone [52,53] and, in this case, the results of the spring block system are only an analog to earthquake behavior. While the classical RSF framework effectively describes fault slip behavior at low-to-intermediate velocities and over a broad range of geological conditions, it becomes increasingly limited at seismic slip rates ($V > 1$ m/s), especially under hydrothermal conditions. At such high velocities, additional dynamic

weakening mechanisms must be considered to account for the dramatic reductions in friction observed in laboratory experiments and inferred from field observations [54,55]. However, the present review primarily focuses on experimental studies conducted under the slower slip rate ($V < 1$ mm/s) relevant to earthquake nucleation conditions.

3. Friction Experiments of Granitoid Fault

The standard RSF law, derived from laboratory friction experiments, has been extensively utilized to quantify the frictional properties of rock faults and simulate the various stages of both natural and induced seismic events [27,56–58], as well as laboratory earthquakes [37,59–61]. This includes the preseismic, nucleation, coseismic, and interseismic healing processes. The RSF framework effectively captures a broad spectrum of fault slip behaviors, ranging from stable sliding to slow-slip events and dynamic ruptures. In this section, we summarize laboratory friction experiments under both dry and hydrothermal conditions for diverse granites and their gouges (Table 1). Friction experiments were conducted using various apparatus, such as saw-cut, direct-shear, double-direct-shear, and rotary-shear configurations. These various experimental setups facilitated the investigation of frictional strength variations under constant-velocity slip, the assessment of velocity dependence parameters ($a - b$) during velocity-stepping tests, and the evaluation of fault healing behavior through slide-hold-slide experiments [26].

As a fundamental basement rock in the continental upper crust, granite and its fault gouges have been extensively investigated in laboratory experiments since the 1950s. We focused on the frictional properties of rocks/gouges under varying temperatures, normal stress, pore pressure, and sliding rate, which are key parameters controlling granite fault mechanical behavior. Given that in situ crustal conditions exhibit linear-like gradients in temperature and pressure with depth (geothermal gradient: 25–30 °C/km), understanding earthquake nucleation and induced seismic under coupled thermos-hydro-mechanical conditions remains a central scientific challenge. Early friction experiments were primarily conducted on Westerly granite, which is composed predominantly of ~28% quartz, ~35% plagioclase, ~34% K-feldspar, and <5% biotite [27,62] but with biotite replaced by chlorite and muscovite. Subsequent studies have incorporated natural fault gouges retrieved from continental drilling projects and fault cores, which are characterized by higher proportions of illite- and chlorite-rich components [7,9,40,41,63].

Table 1. Friction experiments on granitoid gouge/rock.

Gouge/Rock Sample	Temperature T (°C)	Effective Normal Stress $\bar{\sigma}$ (MPa)	Pore Fluid Pressure P_f (MPa)	Slip Rate V ($\mu\text{m/s}$)	Shear Displacement S (mm)	Experiment Type	References
Westerly granite	ambient	2–10	dry	0.1–10	~1		[64]
Westerly granite	ambient	27–84	dry	0.01–10	~70	Bare surface, CV/VS	[65]
Westerly granite	ambient–845	CP: 250	dry	0.0476–0.476–4.76	~4	VS	[36]
Westerly granite	ambient	50	dry	0.001–3000	~376	VS	[66]
Westerly granite gouge	25–600	400	dry, 100	0.01–0.1–1	~4	VS	[37]
Westerly granite gouge	ambient	25	dry	1–10	~400	VS	[38]
Granitoid gouge	100–600	200	30	0.04–0.2–1	~4	VS	[67]
Granitoid gouge	100–600	100, 200, 300	30	0.04–0.2–1	~4	VS	[40]

Table 1. Cont.

Gouge/Rock Sample	Temperature T (°C)	Effective Normal Stress σ (MPa)	Pore Fluid Pressure P_f (MPa)	Slip Rate V ($\mu\text{m/s}$)	Shear Displacement S (mm)	Experiment Type	References
Westerly granite, gouge	25–600	5–40	minor water	0.1–30	~2.5 or ~112	VS	[45]
Natural granite gouge	25–600	200	30	0.04–0.2–1	~4	VS	[68]
Westerly granite gouge	25–250	CP: 20	10	0.1	~5	HSH	[69]
Natural granitoid gouge	20–650	100	100	0.1–100	~15	CV	[41]
Natural granite gouge	25–600	100, 200	30,100	0.04–0.2–1 and 1–3–10–30–100	~25	VS and CV	[9]

Note. VS = velocity stepping, CV = constant velocity, SHS = slide-hold-slide, CP = confining pressure.

3.1. Velocity-Stepping Experiment

Velocity-stepping (VS) experiments are widely used in the study of fault friction properties. In VS friction experiments, variations in shear velocity are systematically applied to investigate the rate-dependence of frictional behavior in fault materials, which is critical for assessing slip stability in earthquake nucleation zones (e.g., the values of $(a - b)$). These experiments are typically conducted using high-stiffness shear apparatuses, most commonly direct-shear, saw-cut, and rotary shear configurations [64,70]. The direct-shear apparatus is mechanically straightforward and allows for easy sample assembly and force alignment (Figure 1a,d). It is well-suited for low-to-moderate slip rates ($\mu\text{m/s}$ – mm/s) and small displacement (<10 mm) experiments, making it ideal for studying quasi-static frictional properties. Its primary advantages include the direct application of shear load and ease of data interpretation. However, its limitations include relatively small cumulative displacements and potential boundary effects, such as stress concentrations at the sample edges, which can affect the uniformity of deformation. In the saw-cut test, the servo-controlled triaxial machine is capable of independently controlling axial stress, confining pressure, and pore pressure. This setup allows the simulation of realistic in situ stress conditions relevant to crustal fault zones, particularly under elevated pressure and temperatures. However, the total shear displacement achievable is typically constrained (<4 – 5 mm). In contrast, the rotary shear apparatus employs torsional motion to impose shear displacement, allowing for continuous sliding at a wide range of velocities and large total displacements. This makes it particularly suitable for simulating seismic slip conditions and capturing thermomechanical processes during rapid fault motion. Despite these advantages, rotary systems are generally more complex, require careful alignment, and involve non-uniform stress distributions due to their radial geometry. Data interpretation also requires the conversion of torque and angular displacement into shear stress and slip, which adds to the analytical complexity. The choice of apparatus depends on the experimental objectives and the specific slip conditions of interest.

We summarized the VS test on granitic examples and discussed the mechanisms that influence their sliding stability. At dry conditions ($P_f = 0$ MPa), the velocity dependence of friction, that is, on $(a - b)$ values only, shows the positive (V-S behavior) under all temperatures (T) and velocities (V) (Figure 2) [36]. However, V-W behavior can also be observed in the nominally dry tests, although such results should be interpreted with greater caution [45]. Beeler et al. [38] conducted friction experiments under room T at an effective normal stress (σ_n^{eff}) of 25 MPa, and granite fault gouges show the V-W behavior occurring up to ~40–100 mm displacement, then giving way to V-S behavior.

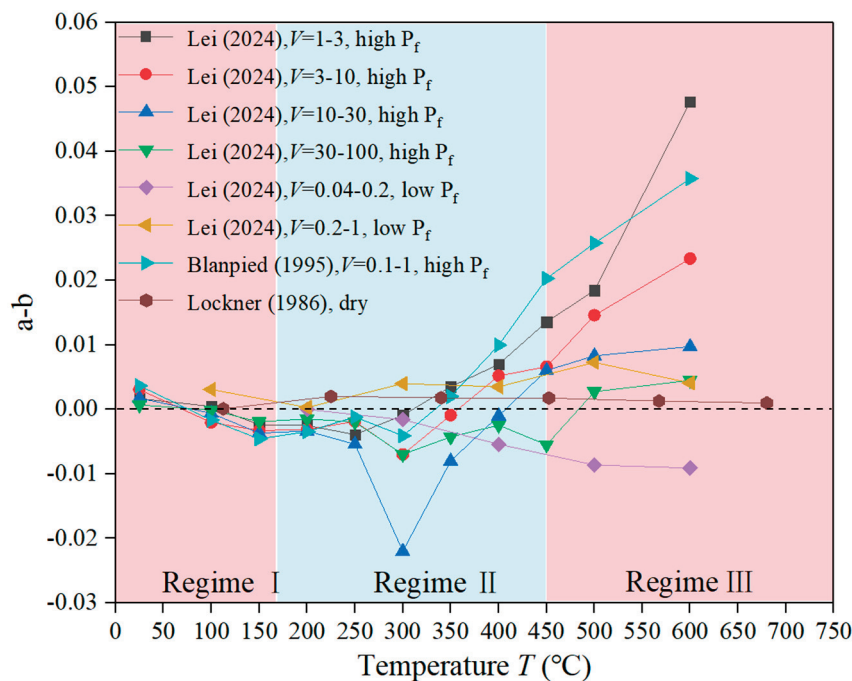


Figure 2. Compilation of published experimental data for velocity dependence ($a - b$) as function of temperature under different velocities and pore fluid pressures. Velocity (V) is expressed in $\mu\text{m/s}$ [9,36,37].

On the contrary, three T regimes were observed under hydrothermal conditions for granitic gouge and rock: (a) (Regime I) velocity-strengthening (V-S) at a low T of 25–100 °C; (b) (Regime II) velocity-weakening (V-W) at an intermediate T of 150–300 °C; (c) (Regime III) velocity-strengthening (V-S) at a high T of 350–600 °C. Notably, the boundaries of these T ranges may shift toward a higher T with different V and P_f . Blanpied et al. [37] found that at 350 °C, lower V (0.01–0.1 $\mu\text{m/s}$) transition from the V-W regime to a high- T , V-S regime earlier than higher V (0.1–1 $\mu\text{m/s}$). A similar trend is observed in the natural granite gouge [9]. They conducted friction experiments over a wider V range of 1–100 $\mu\text{m/s}$ and found that the V-W regime occurs at 150–300 °C for low $V = 1\text{--}3 \mu\text{m/s}$, while it extends to 100–450 °C as the V increases to 30–100 $\mu\text{m/s}$ (Figure 2). In addition to T and V , pore pressure (P_f) and shear displacement influence the occurrence of the three regimes. For example, increasing P_f shifts all three regimes toward lower T , particularly narrowing the V-W regime to within 100–350 °C. This behavior has also been observed in phyllosilicate and quartz-rich fault gouges [71]. Regarding shear displacement, large cumulative slips can promote an earlier transition from the V-W regime to a high- T , V-S regime [9]. Note that in tests conducted over a narrow range of V , the ($a - b$) values may capture only two of the three frictional regimes [40,68,72].

Recent microstructural studies suggest that the transitions among the three frictional regimes are primarily governed by underlying microphysical mechanisms, particularly the intergranular flow and deformation of fine-grained gouge facilitated by intergranular pressure solutions, often accompanied by healing processes, such as cavitation–creep behavior, and CNS models [42,72–75]. Under hydrothermal conditions, the strain rate of the fine grains is strongly dependent on temperature (T), differential stress (σ), and grain size (d). These relationships are commonly described by the flow laws of the diffusion creep

$$\dot{\epsilon} = A\sigma^n d^{-m} \exp\left(\frac{-Q}{RT}\right) \quad (5)$$

where A is a constant value, n is the stress exponent, m is the grain size exponent, Q is the activation energy, R is the gas constant, and T is the absolute temperature. Accordingly, elevated temperature and differential stress, along with finer grain size, contribute to an increased strain rate [76,77]. The typical microstructural observations of dissolution reveal widespread intergranular indentation features and grain-to-grain contacts in granite gouges at T above 600 °C [9,41]. Pressure solutions in the field are commonly recognized by the development of stylolites, solution seams, and sutured grain boundaries. These structures are most frequently observed in sedimentary rocks, especially in carbonates and quartz-rich gouges, where mineral solubility and the presence of pore fluids promote the pressure solution process [78,79]. However, to quantitatively resolve the understanding of microphysical controls on the transition between stable and unstable sliding, more extensive frictional experiments on granite are required under a broader range of conditions, particularly at lower velocities and higher pore pressures.

3.2. Constant-Velocity Experiment

Constant-velocity friction experiments are fundamental for characterizing the steady-state friction strength of faults or the continental crust. These tests enable systematic observations of the shear stress–strain response, including peak strength, strain weakening, and steady sliding behavior. In recent decades, advances in experimental techniques have allowed for the simultaneous monitoring of porosity, volumetric strain, and acoustic emissions during shear. These additional datasets provide critical insights into the micromechanical processes, such as dilation, compaction, and grain-scale damage, that govern fault zone evolution during slip. Here, we mainly summarize the friction strength or coefficient changed by temperature (T), normal stress (σ_n), velocity (V), and pore pressure (P_f).

Under dry conditions, granitic and quartz–feldspar materials exhibit high friction coefficients (μ) typically ranging from 0.6 to 0.8 across both low- and high-temperature (T) regions, consistent with Byerlee’s law. For example, Lockner et al. (1986) [36] showed a high μ , which increased slightly over a T range of ~20–800 °C (normal stress ranged from 380 to 460 MPa). Also, in a study by Stesky [80], dry westerly granite exhibited relatively high μ at 300–700 °C (normal stress ranged from 370 to 450 MPa). Tullis and Weeks [65] conducted a shear test on dry granite rock with a bare rock surface at $V = 0.01$ –10 $\mu\text{m/s}$, further confirming high μ . The experimental results on natural granite gouges indicate that at $T = 600$ °C and an effective normal stress (σ_n^{eff}) of 100 MPa, the μ in dry conditions remains at a value of 0.68 [9]. Compared to dry fault gouges, wet granite gouges exhibit a broader range of friction coefficients. At T s below 400 °C, there is no difference between dry and wet conditions; however, above 400 °C, the μ decreases significantly with increasing pore pressure (P_f). In particular, at 600 °C and a low velocity (V) of <1 $\mu\text{m/s}$, the μ values can decrease to ~0.4 [9,37]. In addition to T , P_f , and V , the shear under large displacement also changed the μ at 600–650 °C. The μ decreases from 0.75 to 0.39 with increasing shear displacement from 4 to ~20 mm [9,41] (Figure 3).

The microstructural evidence indicates that high μ (>0.6) corresponds to a temperature-insensitive regime, which is dominated by frictional deformation, while low μ (<0.6) corresponds to a temperature-sensitive regime, which is mainly governed by frictional deformation, potentially accompanied by plastic deformation [81]. In high μ , the deformed granite gouge shows a cataclastic frictional granular flow, and the angular grain boundaries suggest a process of grain crushing and fracturing in the slip zone. In that case, high strength is mainly determined by the actively deformed slip zones, where frictional resistance is caused by grain rotations and localized sliding along angular grains and on the fracture array. However, at lower μ , microstructural observations reveal evidence of localized shear bands, including Riedel shear, Boundary shear, and Y-shear [82], accompanied by

intergranular pressure solution and granular flow. In particular, shear localization in granular materials plays a critical role in controlling both their frictional strength and deformation characteristics. In this regime, the extremely fine-grained gouge exhibited few fracture arrays, and many nanopores were located along the grain contact surfaces, and these features are commonly observed in naturally deformed rocks from mid-crustal shear zones [83–85]. These nanopores likely involve a combination of processes of intergranular sliding and dissolution–precipitation creep (DPC), and they infer that such low friction strength caused by the finer-grain size within the principal slip zones allows DPC to operate at rates that are sufficiently rapid to accommodate most of the sliding and ultimately resulting in a reduction in the gouge strength [9,41,42].

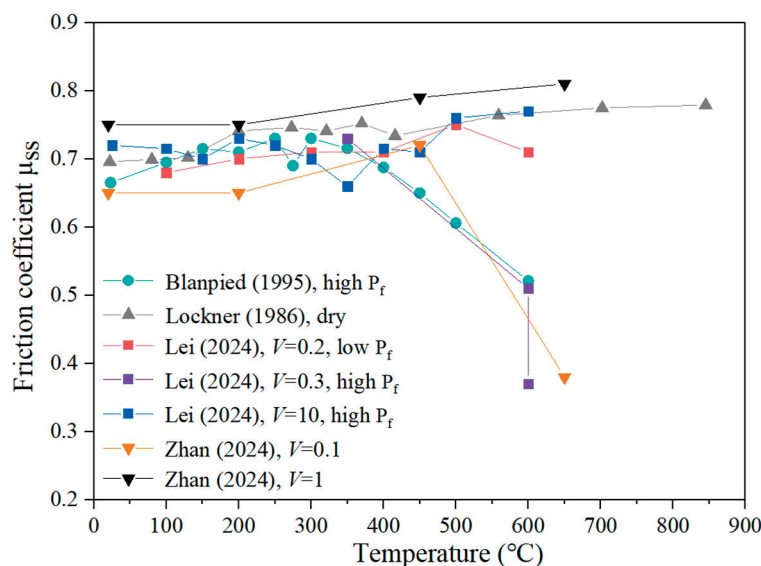


Figure 3. Compilation of published experimental data for steady-state friction coefficient as function of temperature. Velocity (V) is expressed in $\mu\text{m/s}$ [9,36,37,41].

3.3. Slide-Hold-Slide Experiment

Laboratory slide-hold-slide (SHS) experiments serve as a fundamental analog for studying the seismic cycle and are widely employed to measure strength recovery in simulated fault systems. SHS tests consist of alternating periods of induced stick–slip and quasi-static holds during which time-dependent strength recovery occurs. The strength changes according to secant friction $\mu = \tau / (\sigma_n - P_f)$, τ and σ_n are shear and normal stress measured on the fault surface, and P_f is pore pressure. The magnitude of restrengthening ($\Delta\mu$) is calculated by the difference between the peak failure strength upon the resumption of the slip and the steady-state sliding strength preceding the hold (Figure 4). With subsequent shearing, the shear stress gradually declines to its original steady-state value (τ_{ss}). Seismic observations reveal that the drop in earthquake stress rises logarithmically over time, while studies on rock friction demonstrate that restrengthening increases with the log of hold time, an equation that can be characterized by the logarithmic relation defined in previous studies [49,64,86].

Previous studies under room temperature (T) indicate that the frictional healing rate is sensitive to time and the slip rate. Marone [27] conducted an SHS test under constant normal stress of 25 MPa using quartz sand sheared between rough westerly granite surfaces. The results indicate that the friction decreased during hold periods, reaching a peak value (the steady-state friction coefficient) when reloaded. Additionally, friction increased almost linearly and logarithmically with the hold time. This trend is consistent with many other earth materials [87–90]. However, when the T increases to levels corresponding to the

shallow crust, the friction coefficient of westerly granite linearly increases with T , but T has little effect on the rate of change in static friction with hold time. It is cautioned that extrapolating laboratory results to natural faults should be undertaken carefully. Neglecting the effect of T could lead to an underestimation of the fault strength under dry conditions by up to 10% [45]. Recent studies show that frictional weakening was observed in hydrothermal SHS tests at $T_s \geq 200$ °C, indicating that SHS tests are more effective in promoting weakening than VS tests [37]. The simulation of long-duration holds at 250 °C further suggests that hydrothermal alterations occur at this T because the negative state variable is not required to simulate shorter hold periods or lower T_s . Although the precise mechanisms controlling frictional behavior remain unclear, the data imply that the processes involved are water-assisted, temperature-dependent, and chemically complex with a characteristic time delay in the order of 1000 s [69]. Jeppson et al. [91] measured the restrengthening of wet quartzite samples under a broader range of effective normal stress of 20–200 MPa. The restrengthening ($\Delta\mu$) increases with the log of hold duration, the 200 °C healing rate, and the 0.014 per e-fold increase in time, which is comparable to that determined from seismological observations along the Calaveras Fault. Overall, they indicate that cohesive healing is negligible at $T_s < 200$ °C, and strength recovery is predominantly due to frictional healing processes.

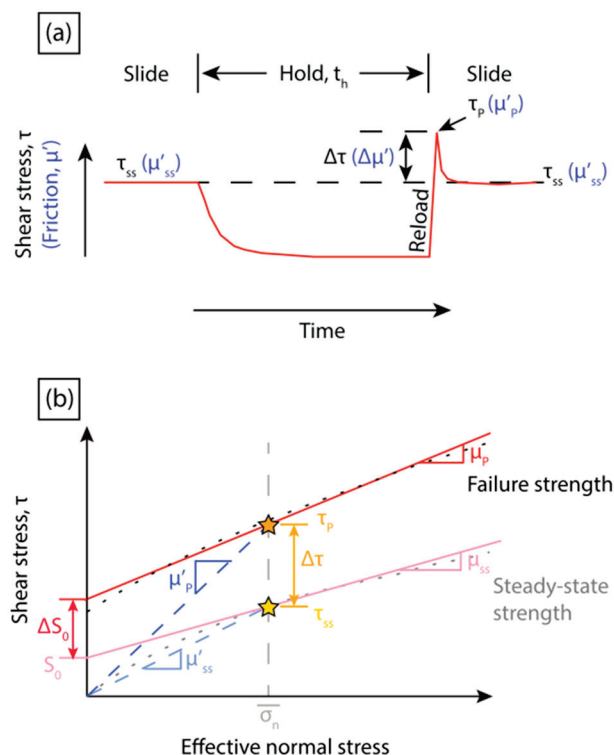


Figure 4. (a) Schematic diagram of restrengthening or secant friction evolution during slide-hold-slide (SHS) tests. (b) Diagram showing how secant friction (blue lines), tangent friction (red lines), and cohesion (y-intercept of red lines) are determined. Lower gray dashed line represents steady-state sliding stress conditions; upper black dashed line represents stress conditions at failure. Figure reproduced from [91] with permission.

4. Velocity Dependence Behavior Explained Using Microphysical Model

To describe the frictional behaviour of granular fault materials, Chen, Niemeijer, and Spiers developed a microphysical model (CNS) [42,73–75]. This model attributes friction to two key factors: porosity and grain boundary friction. It captures how velocity-dependent transitions arise from the competition between porosity generation via granular flow and

porosity reduction via time-dependent, thermally activated creep processes. In the CNS framework, the rates of change of shear stress (τ), porosities of shear band (φ^{sb}), and bulk gouge (φ^{bulk}) are described by:

$$V_{imp} - \dot{\tau}/K = L_t \left[\lambda \dot{\gamma}_{ps}^{sb} + (1 - \lambda) \dot{\gamma}_{ps}^{bulk} \right] + L_t \lambda \dot{\gamma}_{gr}^{sb} \quad (6)$$

$$\tau = \bar{\sigma} \frac{\tilde{\mu} + \tan \varphi}{1 - \tilde{\mu} \tan \varphi} \quad (7)$$

$$\dot{\varphi}^{sb} / (1 - \varphi^{sb}) = \tan \varphi \dot{\gamma}_{gr} - \dot{\varepsilon}_{pl}^{sb} \quad (8)$$

$$\dot{\varphi}^{bulk} / (1 - \varphi^{bulk}) = -\dot{\varepsilon}_{pl}^{bulk} \quad (9)$$

Here, V_{imp} represents the velocity imposed at the loading point, and the gouge layer has a thickness of L_t , with a shear strain localizing within a fraction λ of that thickness. The dot notation indicates a time derivative. Granular flow (gr) is assumed to be confined to the localized shear zone with a width of $L_t \lambda$, whereas intergranular plastic creep (pl) is considered to occur both inside this localized region and throughout the remaining bulk gouge material, covering a thickness of $L_t(1 - \lambda)$. The dilatancy angle is denoted by φ . Superscripts “sb” and “bulk” refer to the shear band and surrounding bulk material, respectively. Shear strain and normal strain are indicated by γ and ε . The frictional resistance at grain boundaries, represented by $\tilde{\mu}$, depends on the strain rate associated with the granular flow within the shear band.

Figure 5 shows a schematic diagram outlining the three distinct temperature-dependent regimes in terms of the velocity-dependence parameter. This schematic links the apparent friction coefficient and its velocity dependence to the underlying micro-mechanical deformation mechanisms proposed by the model. In Regime I, deformation is characterized by significant grain size reduction and a cataclastic flow involving quartz, albite, and K-feldspar, without any microstructural evidence for the creep or mass transfer processes, consistent with model A in Figure 5. For Regime II, the microstructures indicate a more distributed deformation pattern with multiple sharp Y-shears, in contrast to the localization observed in Regimes I and III. While cataclastic processes remain active, features such as grain-to-grain indentation and grain truncation, particularly within ultrafine-grained zones, suggest that pressure solution mechanisms began to operate, likely triggered by grain size reduction [79]. This regime reflects the interplay between dilatant granular flow and solution-assisted deformation, in line with model B of Figure 5. In Regime III, the presence of active B-shear bands containing ultrafine, equant grains of quartz, albite, and K-feldspar—accompanied by intergranular pores or healed cavities with low overall porosity and lacking shape-preferred orientation (SPO)—indicates a shift towards grain boundary sliding assisted by solution-precipitation creep and pore healing [73]. These characteristics support a deformation mechanism involving granular flow coupled with enhanced diffusive mass transfer, as illustrated by model C in Figure 5.

Under hydrothermal conditions, the extended CNS model provides a more physically consistent explanation of frictional behavior than classical RSF laws. It captures three velocity-dependent regimes—strengthening at low temperatures, weakening at intermediate temperatures, and restrengthening at high temperatures—by accounting for competing deformation and healing mechanisms (Figure 6). Unlike RSF or empirical flow laws, CNS incorporates the effects of slip velocity, effective normal stress, and evolving microstructures, such as porosity, grain size, and shear band thickness. These features make it well-suited for interpreting experimental observations of granite gouges across a range of temperatures and velocities [92].

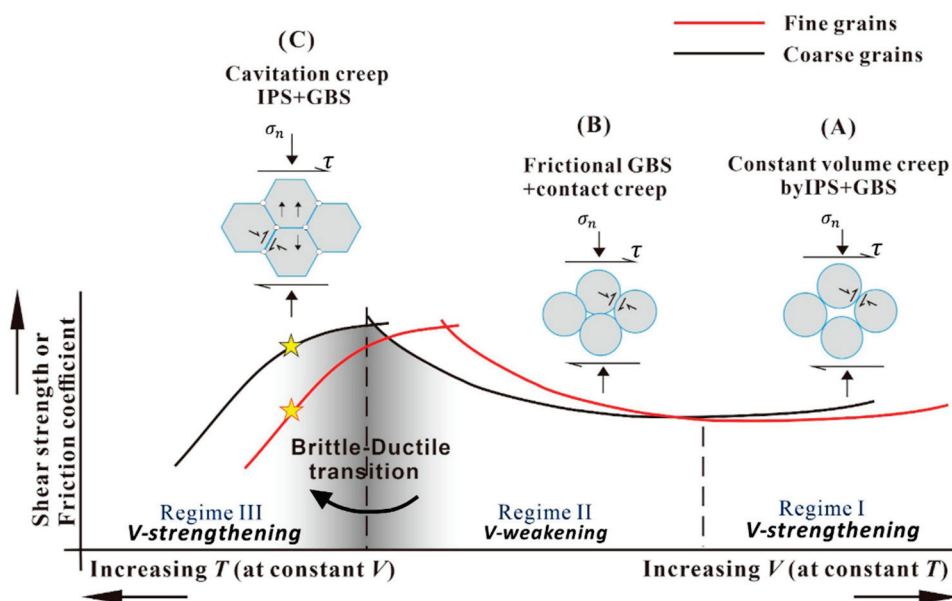


Figure 5. Schematic diagram of our sample summarizing the relationship between steady-state shear strength (or friction coefficient), its rate dependence, and the dominant deformation mechanisms and associated microstructures as a function of sliding velocity (V) and temperature (T), as predicted by the CNS model [74]. Figure reproduced from [9] with permission.

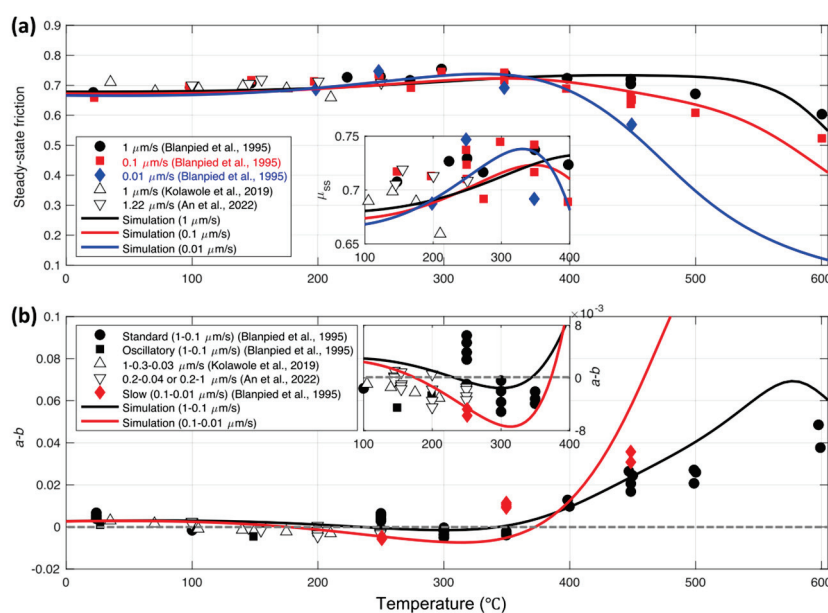


Figure 6. Comparisons between numerical and experimental results of (a) steady-state friction coefficient and (b) frictional stability parameter ($a - b$) for simulated granite gouges at sliding velocities of 0.01–10 $\mu\text{m/s}$ under wet conditions [37,93,94]. Figure reproduced from [92] with permission.

5. Implications for Nucleation of Continental Earthquakes

It has long been known that earthquake activity is a symptom, if not the agent, of active tectonics. The cyclic nature of earthquakes, as described by Reid [22], depends on the schizosphere, plastosphere, and asthenosphere. This coupling of mechanical processes across these distinct layers influences the timing and magnitude of seismic events. Three primary models of strain accumulation exist, each differing significantly in how the schizosphere and plastosphere are coupled and their relative strengths. In the vis-

coelastic coupling model, the fault traverses the schizosphere but does not extend into the plastosphere, which is modeled as a viscous layer [95,96]. However, the strong plastosphere model suggests the fault is weak, and the plastosphere is strong; in that case, the plastosphere drives all systems during interseismic processes [97]. In the deep slip model, the fault may be active through both the schizosphere and plastosphere, being seismogenic in the schizosphere and aseismically slipping in a ductile shear zone, weaker than the bulk plastosphere, in the lower crust. In this model, the deep shear zone loads the schizosphere during the interseismic period, and the deeper zone is subsequently reloaded by the coseismic slip [98]. The deep slip model explored by Tse and Rice [99] not only provides more comprehensive details than earlier formulations but also accounts for several key characteristics documented in prior studies. This model investigated how the slip is distributed throughout repeated loading cycles on a fault governed by the RSF law (Equations (1)–(4)), using parameters derived from laboratory experiments. The values they used for the velocity dependence friction term ($a - b$), with the temperature gradient they assumed, predict a transition to a stable slip at 11 km (as $(a - b)$ values transition from negative to positive) (Figure 7).

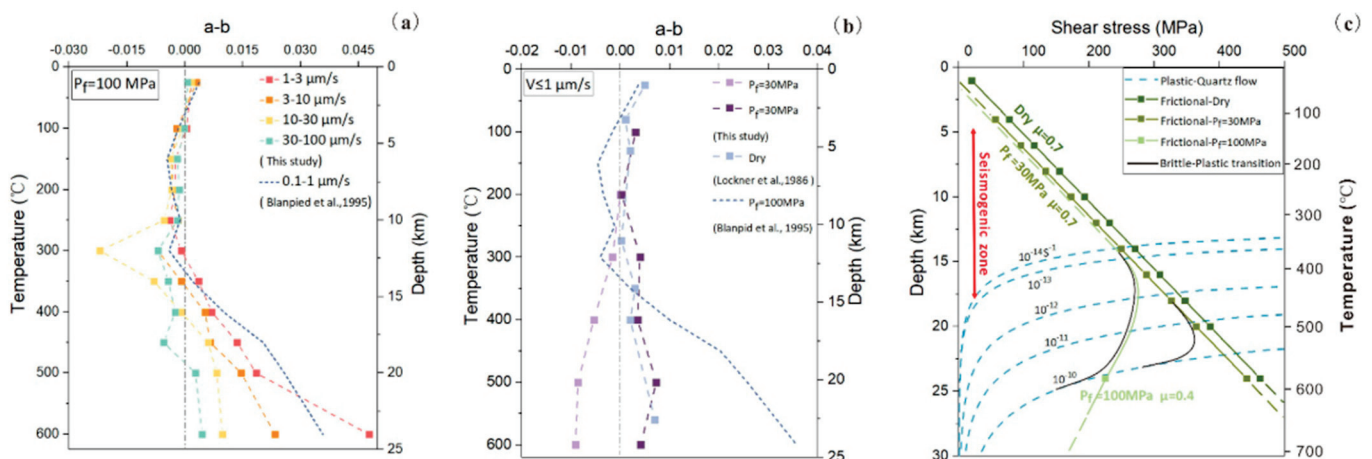


Figure 7. (a,b) Depth profile of $(a - b)$ values for granite gouges at different V and P_f , respectively. (c) Shear strength versus depth profile. Geothermal gradient is $25\text{ }^{\circ}\text{C/km}$ as representative average value for continental crust [36,37,100]. Figure reproduced from [9] with permission.

To our knowledge, the formation of seismological zones is linked to the thermomechanical and hydrological conditions within the lithosphere. Variations in temperature (T), normal stress (σ_n), slip rate (V), shear displacement (S), and pore fluid pressure (P_f) can significantly alter the mechanical properties and rheological behavior of both the brittle upper crust (schizosphere) and the underlying ductile lower crust or upper mantle (plastosphere). The experimental results on quartz-rich and granitic gouges indicate that with increasing T_s , fault slips exhibit VW behavior ($(a - b) < 0$) at intermediate T_s ($\sim 100\text{--}400\text{ }^{\circ}\text{C}$), corresponding to mid-crustal depths ($\sim 4\text{--}16$ km), suggesting stable slips for earthquake nucleation. In contrast, at both lower and higher T_s , the fault shows VS behavior ($(a - b) > 0$), indicative of stable sliding (Figure 5, A). However, the T range of the intermediate VW regime is not fixed; increasing the V tends to shift the transition from the VW to VS regime toward higher T_s (Figure 5, A). This is likely due to the reduced time available for thermally activated deformation mechanisms at higher strain rates. In contrast, elevated P_f facilitates all three frictional regimes toward lower T_s (Figure 5, B), a result consistent with the reduction in effective normal stress and hence frictional resistance. Furthermore, increasing cumulative slip allows more time for grain-scale deformation processes such as pressure solutions and ductile creep to operate, promoting a rheological transition from brittle to plastic behavior.

These findings underscore the complex interplay among T , P_f , V , and S in controlling the micromechanical processes governing fault strength and stability (Figure 5, C).

Elevated pore fluid pressure is widely recognized as a critical factor in fault weakening and the initiation of unstable slip, thereby serving as a potential trigger for earthquake nucleation. In continental settings, numerous studies have demonstrated a strong correlation between high fluid pressure zones and seismogenic fault segments, particularly in regions such as the Tibetan Plateau and the North China Craton, where deep crustal fluids are actively involved in fault zone dynamics. Advances in microstructural analysis, including transmission electron microscopy (TEM) and synchrotron-based X-ray imaging, have revealed fluid-assisted grain boundary weakening, pressure solution features, and enhanced mineral reaction kinetics at asperity contacts under high fluid pressure conditions. These microscale observations provide compelling evidence that elevated pore fluid pressure facilitates the activation of ductile deformation mechanisms and reduces effective normal stress, thus promoting the transition from stable to unstable sliding. As such, incorporating fluid–rock interaction processes into fault mechanics models is essential for a more realistic understanding of earthquake nucleation in continental lithosphere. Future work should aim to bridge the gap between laboratory-derived frictional behavior and field-scale observations, with a particular focus on the spatiotemporal evolution of pore pressure and its role in modulating fault slip behavior across the seismic cycle.

6. Conclusions

This study provides a comprehensive synthesis of the frictional behavior of granitic faults, integrating key insights from theoretical frameworks and experimental observations. We first reviewed the classical rate-and-state friction (RSF) law, establishing its foundational role in modeling fault dynamics. Through a systematic comparison of dry and hydrothermal experiments on granitic fault gouges, we highlight distinct frictional responses (velocity-stepping, constant-velocity, and slide-hold-slide tests) under varying conditions. Our analysis emphasizes the critical roles of temperature, normal stress, pore fluid pressure, and slip velocity in modulating friction strength and sliding stability, revealing how these parameters interact to govern fault behavior across different regimes. Notably, we demonstrate that microstructural evolution—such as grain size reduction, pressure solution, and fluid-assisted healing—acts as a key mechanism driving transitions between velocity-weakening and velocity-strengthening regimes, with direct implications for earthquake nucleation and arrest. The key conclusions drawn are as follows:

- (1) Granitic faults in the crystalline upper-to-middle continental crust represent key structures controlling the seismogenic behavior and long-term deformation of the continental lithosphere. Their frictional properties are thus central to understanding earthquake initiation and rupture dynamics.
- (2) Laboratory friction experiments, including velocity-stepping (VS), constant-velocity (CV), and slide-hold-slide (SHS) protocols, consistently demonstrate that granitic gouges exhibit measurable and reproducible frictional behavior. Notably, the VS tests reveal frictional stability regimes (velocity-strengthening or -weakening), while SHS tests provide quantitative estimates of fault healing and restrengthening during interseismic periods.
- (3) Microstructural observations indicate that shear deformation induces localized shear, significant grain size reduction, pressure solution creep, and mineralogical alteration within the fault gouge. These microphysical processes are sensitive to temperature, normal stress, pore fluid pressure, and strain rate and play a pivotal role in modulating the rheology and frictional response of the fault zone.

- (4) In the continental crust, the interplay between microstructural evolution and frictional properties suggests that nucleation is favored in zones where brittle–frictional weakening mechanisms dominate but are bounded by ductile, velocity-strengthening regions that arrest rupture.

Author Contributions: Methodology, H.L.; Resources, S.L. and W.D.; Data curation, H.L., S.L., and W.D.; Formal analysis, H.L. and W.D.; Writing original draft, H.L., S.L., and W.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (Grant 2023YFC3012001) and the National Nonprofit Fundamental Research Grant of China, Institute of Geology, China Earthquake Administration (Grant JB-24-31).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Acknowledgments: We gratefully acknowledge Yongsheng Zhou for his constructive reviews, which significantly improved the clarity of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Scholz, C.H. Earthquakes and friction laws. *Nature* **1998**, *391*, 37–42. [CrossRef]
2. Sibson, R.H. Fault zone models, heat flow, and the depth distribution of earthquakes in the continental crust of the United States. *Bull. Seismol. Soc. Am.* **1982**, *72*, 151–163.
3. Scholz, C.H.; Molnar, P.; Johnson, T. Detailed studies of frictional sliding of granite and implications for the earthquake mechanism. *J. Geophys. Res.* **1972**, *77*, 6392–6406. [CrossRef]
4. Blanpied, M.L.; Lockner, D.A.; Byerlee, J.D. Fault stability inferred from granite sliding experiments at hydrothermal conditions. *Geophys. Res. Lett.* **1991**, *18*, 609–612. [CrossRef]
5. Blanpied, M.L.; Lockner, D.A.; Byerlee, J.D. An earthquake mechanism based on rapid sealing of faults. *Nature* **1992**, *358*, 574–576. [CrossRef]
6. Sullivan, W.A.; Peterman, E.M. Pulverized granite at the brittle-ductile transition: An example from the Kellyland fault zone, eastern Maine, USA. *J. Struct. Geol.* **2017**, *101*, 109–123. [CrossRef]
7. Niemeijer, A.R.; Boulton, C.; Toy, V.G.; Townend, J.; Sutherland, R. Large-displacement, hydrothermal frictional properties of DFDP-1 fault rocks, Alpine Fault, New Zealand: Implications for deep rupture propagation. *J. Geophys. Res. Solid Earth* **2016**, *121*, 624–647. [CrossRef]
8. Guermani, A.; Pennacchioni, G. Brittle precursors of plastic deformation in a granite: An example from the Mont Blanc massif (Helvetic, western Alps). *J. Struct. Geol.* **1998**, *20*, 135–148. [CrossRef]
9. Lei, H.; Niemeijer, A.R.; Zhou, Y.; Spiers, C.J. Frictional properties of natural granite fault gouge under hydrothermal conditions: A case study of strike-slip fault from Anninghe Fault zone, southeastern Tibetan Plateau. *J. Geophys. Res. Solid Earth* **2024**, *129*, e2024JB028760. [CrossRef]
10. Roger, F.; Calassou, S.; Lancelot, J.; Malavieille, J.; Mattauer, M.; Zhiqin, X.; Liwei, H. Miocene emplacement and deformation of the Konga Shan granite (Xianshui He fault zone, west Sichuan, China): Geodynamic implications. *Earth Planet. Sci. Lett.* **1995**, *130*, 201–216. [CrossRef]
11. Marone, C.; Blanpied, M.L. Faulting, Friction and Earthquake Mechanics. *Pure Appl. Geophys.* **1994**, *142*.
12. Savage, H.M.; Rowe, C.D. Localization and delocalization during seismic slip. *Geophys. Res. Lett.* **2024**, *51*, e2024GL110058. [CrossRef]
13. Collettini, C.; Niemeijer, A.; Viti, C.; Smith, S.A.F.; Marone, C. Fault structure, frictional properties and mixed-mode fault slip behavior. *Earth Planet. Sci. Lett.* **2011**, *311*, 316–327. [CrossRef]
14. Froment, B.; Campillo, M.; Chen, J.; Liu, Q. Deformation at depth associated with the 12 May 2008 mw 7.9 wenchuan earthquake from seismic ambient noise monitoring. *Geophys. Res. Lett.* **2013**, *40*, 78–82. [CrossRef]
15. Schwartz, S.Y.; Rokosky, J.M. Slow slip events and seismic tremor at circum-Pacific subduction zones. *Rev. Geophys.* **2007**, *45*, RG3004. [CrossRef]
16. Wallace, L.M. Slow slip events in New Zealand. *Annu. Rev. Earth Planet. Sci.* **2020**, *48*, 175–203. [CrossRef]

17. Shelly, D.R.; Beroza, G.C.; Ide, S.; Nakamura, S. Low-frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip. *Nature* **2006**, *442*, 188–191. [CrossRef] [PubMed]
18. Shelly, D.R.; Beroza, G.C.; Ide, S. Non-volcanic tremor and low-frequency earthquake swarms. *Nature* **2007**, *446*, 305–307. [CrossRef]
19. Ito, Y.; Obara, K.; Shiomi, K.; Sekine, S.; Hirose, H. Slow earthquakes coincident with episodic tremors and slow slip events. *Science* **2007**, *315*, 503–506. [CrossRef]
20. Reinen, L.A.; Weeks, J.D.; Tullis, T.E. The frictional behavior of serpentinite: Implications for aseismic creep on shallow crustal faults. *Geophys. Res. Lett.* **1991**, *18*, 1921–1924. [CrossRef]
21. Kaduri, M.; Gratier, J.P.; Renard, F.; Cakir, Z.; Lasserre, C. The implications of fault zone transformation on aseismic creep: Example of the North Anatolian Fault, Turkey. *J. Geophys. Res. Solid Earth* **2017**, *122*, 4208–4236. [CrossRef]
22. Reid, H.F. The Mechanics of the Earthquake, The California Earthquake of April 18, 1906: Report of the State Investigation Commission, Volume 2. Carnegie Institution of Washington: Washington, DC, USA, 1910; pp. 16–28. [CrossRef]
23. Lawson, A.C. The California Earthquake of April 18, 1906: Report of the State Earthquake Investigation Commission. Available online: https://oac.cdlib.org/view?docId=hb1h4n989f&brand=oac4&doc.view=entire_text (accessed on 5 May 2025).
24. Bowden, F.P.; Tabor, D. The friction and lubrication of solids. *Am. J. Phys.* **1951**, *19*, 428–429. [CrossRef]
25. Brace, W.F.; Byerlee, J.D. Stick-slip as a mechanism for earthquakes. *Science* **1966**, *153*, 990–992. [CrossRef]
26. Dieterich, J.H. Modeling of rock friction: 1. Experimental results and constitutive equations. *J. Geophys. Res. Solid Earth* **1979**, *84*, 2161–2168. [CrossRef]
27. Marone, C. Laboratory-derived friction laws and their application to seismic faulting. *Annu. Rev. Earth Planet. Sci.* **1998**, *26*, 643–696. [CrossRef]
28. Scuderi, M.M.; Niemeijer, A.R.; Collettini, C.; Marone, C. Frictional properties and slip stability of active faults within carbonate-evaporite sequences: The role of dolomite and anhydrite. *Earth Planet. Sci. Lett.* **2013**, *369–370*, 230–232. [CrossRef]
29. Tsutsumi, A.; Shimamoto, T. High-velocity frictional properties of gabbro. *Geophys. Res. Lett.* **1997**, *24*, 699–702. [CrossRef]
30. Di Toro, G.; Han, R.; Hirose, T.; De Paola, N.; Nielsen, S.; Mizoguchi, K.; Ferri, F.; Cocco, M.; Shimamoto, T. Fault lubrication during earthquakes. *Nature* **2011**, *471*, 494–498. [CrossRef]
31. Sibson, R.H. Brecciation processes in fault zones: Inferences from earthquake rupturing. *Pure Appl. Geophys.* **1986**, *124*, 159–175. [CrossRef]
32. Chen, K.H.; Burgmann, R.; Nadeau, R.M.; Chen, T.; Lapusta, N. Postseismic variations in seismic moment and recurrence interval of repeating earthquakes. *Earth Planet. Sci. Lett.* **2010**, *299*, 118–125. [CrossRef]
33. Lui, S.K.; Lapusta, N. Repeating microearthquake sequences interact predominantly through postseismic slip. *Nat. Commun.* **2016**, *7*, 13020. [CrossRef] [PubMed]
34. Dieterich, J.H. Constitutive properties of faults with simulated gouge. *Mech. Behav. Crustal Rocks Handin Vol.* **1981**, *24*, 103–120.
35. Ruina, A. Slip instability and state variable friction laws. *J. Geophys. Res. Solid Earth* **1983**, *88*, 10359–10370. [CrossRef]
36. Lockner, D.A.; Summers, R.; Byerlee, J.D. Effects of Temperature and Sliding Rate on Frictional Strength of Granite. In *Friction and Faulting*; Tullis, T.E., Ed.; Birkhäuser: Basel, Switzerland, 1986. [CrossRef]
37. Blanpied, M.L.; Lockner, D.A.; Byerlee, J.D. Frictional slip of granite at hydrothermal conditions. *J. Geophys. Res.* **1995**, *100*, 13045–13064. [CrossRef]
38. Beeler, N.M.; Tullis, T.E.; Weeks, J.D. Frictional behavior of large displacement experimental faults. *J. Geophys. Res.* **1996**, *101*, 8697–8715. [CrossRef]
39. Bedford, J.D.; Faulkner, D.R. The role of grain size and effective normal stress on localization and the frictional stability of simulated quartz gouge. *Geophys. Res. Lett.* **2021**, *48*, e2020GL092023. [CrossRef]
40. Zhang, L.; He, C.R. Frictional properties of phyllosilicate-rich mylonite and conditions for the brittle-ductile transition. *J. Geophys. Res. Solid Earth* **2016**, *121*, 3017–3047. [CrossRef]
41. Zhan, W.; Niemeijer, A.R.; Berger, A.; Nevskaya, N.; Spiers, C.J.; Herwegh, M. The frictional-viscous transition in experimentally deformed granitoid fault gouge. *J. Geophys. Res. Solid Earth* **2024**, *129*, e2024JB028797. [CrossRef]
42. Niemeijer, A.R.; Spiers, C.J.; Peach, C.J. Frictional behaviour of simulated quartz fault gouges under hydrothermal conditions: Results from ultra-high strain rotary shear experiments. *Tectonophysics* **2008**, *460*, 288–303. [CrossRef]
43. Ma, S.; Shimamoto, T.; Yao, L.; Togo, T.; Kitajima, H. A rotary-shear low to high-velocity friction apparatus in Beijing to study rock friction at plate to seismic slip rates. *Earthq. Sci.* **2014**, *27*, 469–497. [CrossRef]
44. He, C.; Ma, X.; Yao, S. Unstable sliding of plagioclase gouge and deformation mechanisms under hydrothermal conditions with effective normal stresses of 100–300 MPa. *J. Geophys. Res. Solid Earth* **2024**, *129*, e2024JB028883. [CrossRef]
45. Mitchell, E.K.; Fialko, Y.; Brown, K.M. Velocity-weakening behavior of Westerly granite at temperature up to 600 °C. *J. Geophys. Res. Solid Earth* **2016**, *121*, 6932–6946. [CrossRef]
46. Verberne, B.A.; He, C.; Spiers, C.J. Frictional properties of sedimentary rocks and natural fault gouge from the Longmen Shan fault zone, Sichuan, China. *Bull. Seismol. Soc. Am.* **2010**, *100*, 2767–2790. [CrossRef]

47. Ma, X.; Zhou, Y.; Dang, J.; Zhang, L.; He, C. Mineralogical and frictional properties of fault rocks from the Laohushan creep segment of the Haiyuan fault, Northeastern Tibet. *J. Struct. Geol.* **2024**, *178*, 105017. [CrossRef]
48. Chen, J.; Verberne, B.A.; Spiers, C.J. Interseismic re-strengthening and stabilization of carbonate faults by “non-Dieterich” healing under hydrothermal conditions. *Earth Planet. Sci. Lett.* **2015**, *423*, 1–12. [CrossRef]
49. Nakatani, M.; Scholz, C.H. Frictional healing of quartz gouge under hydrothermal conditions: 1. Experimental evidence for solution transfer healing mechanisms. *J. Geophys. Res.* **2004**, *109*, B07201. [CrossRef]
50. Hu, J.; Sun, Q. The effect of high temperature and pressure on rock friction coefficient: A review. *Int. J. Earth Sci.* **2020**, *109*, 409–419. [CrossRef]
51. Mei, C.; Mercuri, M.; Rudnicki, J.W. Rock friction experiments and modeling under hydrothermal conditions. *Earth-Sci. Rev.* **2024**, *254*, 104824. [CrossRef]
52. Rice, J.R. Spatio-temporal complexity of slip on a fault. *J. Geophys. Res. Solid Earth* **1993**, *98*, 9885–9907. [CrossRef]
53. Segall, P. Earthquake and volcano deformation. In *Earthquake and Volcano Deformation*; Princeton University Press: Princeton, NJ, USA, 2010.
54. Rice, J.R. Heating and weakening of faults during earthquake slip. *J. Geophys. Res. Solid Earth* **2006**, *111*. [CrossRef]
55. Noda, H.; Dunham, E.M.; Rice, J.R. Earthquake ruptures with thermal weakening and the operation of major faults at low overall stress levels. *J. Geophys. Res. Solid Earth* **2009**, *114*. [CrossRef]
56. Veedu, D.M.; Barbot, S. The Parkfield tremors reveal slow and fast ruptures on the same asperity. *Nature* **2016**, *532*, 361–365. [CrossRef] [PubMed]
57. Wang, L.; Barbot, S. Excitation of San Andreas tremors by thermal instabilities below the seismogenic zone. *Sci. Adv.* **2020**, *6*, eabb2057. [CrossRef]
58. Im, K.; Saffer, D.; Marone, C.; Avouac, J.P. Slip-rate-dependent friction as a universal mechanism for slow slip events. *Nat. Geosci.* **2020**, *13*, 705–710. [CrossRef]
59. He, C.; Tan, W.; Zhang, L. Comparing dry and wet friction of plagioclase: Implication to the mechanism of frictional evolution effect at hydrothermal conditions. *J. Geophys. Res. Solid Earth* **2016**, *121*, 6365–6383. [CrossRef]
60. Leeman, J.R.; Saffer, D.M.; Scuderi, M.M.; Marone, C. Laboratory observations of slow earthquakes and the spectrum of tectonic fault slip modes. *Nat. Commun.* **2016**, *7*, 11104. [CrossRef] [PubMed]
61. Leeman, J.R.; Marone, C.; Saffer, D.M. Frictional mechanics of slow earthquakes. *J. Geophys. Res. Solid Earth* **2018**, *123*, 7931–7949. [CrossRef]
62. Tullis, J.; Yund, R.A. Experimental deformation of dry Westerly granite. *J. Geophys. Res.* **1977**, *82*, 5705–5718. [CrossRef]
63. Zoback, M.; Hickman, S.; Ellsworth, W. Scientific drilling into the San Andreas fault zone. *Eos Trans. Am. Geophys. Union* **2010**, *91*, 197–199. [CrossRef]
64. Dieterich, J.H. Time-dependent friction and the mechanics of stick-slip. *Rock Frict. Earthq. Predict.* **1978**, *6*, 445–469.
65. Tullis, T.E.; Weeks, J.D. Constitutive behavior and stability of frictional sliding of granite. *Pure Appl. Geophys.* **1986**, *124*, 383–414.
66. Blanpied, M.L.; Tullis, T.E.; Weeks, J.D. Frictional behavior of granite at low and high sliding velocities. *Geophys. Res. Lett.* **1987**, *14*, 554–557. [CrossRef]
67. Ren, F.; He, C. An experimental study of frictional sliding of granitic mylonite under hydrothermal conditions. *Chin. J. Geophys.* **2014**, *57*, 877–883.
68. Lei, H.; Zhou, Y.; Yao, W.; Ma, X.; He, C.; Dang, J.; Miao, S.; Dai, W. Earthquake nucleation conditions of Anninghe Fault: Constrains from frictional experiments on natural granite gouge. *Chin. J. Geophys.* **2022**, *65*, 978–991. (In Chinese) [CrossRef]
69. Jeppson, T.N.; Lockner, D.A.; Beeler, N.M.; Moore, D.E. Time-dependent weakening of granite at hydrothermal conditions. *Geophys. Res. Lett.* **2023**, *50*, e2023GL105517. [CrossRef]
70. Dong, L.; Luo, Q. Investigations and new insights on earthquake mechanics from fault slip experiments. *Earth-Sci. Rev.* **2022**, *228*, 104019. [CrossRef]
71. den Hartog, S.A.; Spiers, C.J. Influence of subduction zone conditions and gouge composition on frictional slip stability of megathrust faults. *Tectonophysics* **2013**, *600*, 75–90. [CrossRef]
72. Bos, B.; Spiers, C.J. Frictional-viscous flow of phyllosilicate-bearing fault rock: Microphysical model and implications for crustal strength profiles. *J. Geophys. Res. Solid Earth* **2002**, *107*, ECV-1. [CrossRef]
73. den Hartog, S.A.; Spiers, C.J. A microphysical model for fault gouge friction applied to subduction megathrusts. *J. Geophys. Res. Solid Earth* **2014**, *119*, 1510–1529. [CrossRef]
74. Chen, J.; Spiers, C.J. Rate and state frictional and healing behavior of carbonate fault gouge explained using microphysical model. *J. Geophys. Res. Solid Earth* **2016**, *121*, 8642–8665. [CrossRef]
75. Chen, J.; Niemeijer, A.R.; Spiers, C.J. Seismic fault slip behavior predicted from internal microphysical processes. *J. Geophys. Res. Solid Earth* **2022**, *127*, e2022JB024530. [CrossRef]
76. Rutter, E.H. The kinetics of rock deformation by pressure solution. *Philos. Trans. R. Soc. Lond. A* **1976**, *283*, 203–219.

77. Gratier, J.P.; Dysthe, D.K.; Renard, F. The role of pressure solution creep in the ductility of the Earth's upper crust. *Adv. Geophys.* **2013**, *54*, 47–179.
78. De Boer, R.B.; Nagtegaal, P.J.C.; Duyvis, E.M. Pressure solution experiments on quartz sand. *Geochim. Cosmochim. Acta* **1977**, *41*, 257–264. [CrossRef]
79. Spiers, C.J.; Schutjens, P.M.T.M.; Brzesowsky, R.H.; Peach, C.J.; Liezenberg, J.L.; Zwart, H.J. Experimental determination of constitutive parameters governing creep of rocksalt by pressure solution. *Geol. Soc. Lond.* **1990**, *54*, 215–227. [CrossRef]
80. Stesky, R.M. The Mechanical Behavior of Faulted Rock at High Temperature and Pressure. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 1975.
81. Pec, M.; Stünitz, H.; Heilbronner, R.; Drury, M. Semi-brittle flow of granitoid fault rocks in experiments. *J. Geophys. Res. Solid Earth* **2016**, *121*, 1677–1705. [CrossRef]
82. Logan, J.M.; Dengo, C.A.; Higgs, N.G.; Wang, Z.Z. Fabrics of experimental fault zones: Their development and relationship to mechanical behavior. In *International Geophysics*; Academic Press: Cambridge, MA, USA, 1992; Volume 51, pp. 33–67.
83. Fousseis, F.; Regenauer-Lieb, K.; Liu, J.; Hough, R.M.; De Carlo, F. Creep cavitation can establish a dynamic granular fluid pump in ductile shear zones. *Nature* **2009**, *459*, 974–977. [CrossRef]
84. Précigout, J.; Ledoux, E.; Arbaret, L.; Spriet, C. Porosity induced by dislocation dynamics in quartz-rich shear bands of granitic rocks. *Sci. Rep.* **2022**, *12*, 6141. [CrossRef]
85. Duan, Q.; Fagereng, A.; Chen, J.; Blenkinsop, T. Fluid environment controls along-strike variation in slip style: Midcrustal geological signatures from the Red River fault, China. *Geology* **2024**, *52*, 405–410. [CrossRef]
86. Berthoud, P.; Baumberger, T.; G'sell, C.; Hiver, J.M. Physical analysis of the state- and rate-dependent friction law: Static friction. *Phys. Rev. B* **1999**, *59*, 14313–14327. [CrossRef]
87. Beeler, N.B.; Tullis, T.E.; Weeks, J.D. The roles of time and displacement in the evolution effect in rock friction. *Geophys. Res. Lett.* **1994**, *21*, 1987–1990. [CrossRef]
88. Carpenter, B.; Ikari, M.; Marone, C. Laboratory observations of time-dependent frictional strengthening and stress relaxation in natural and synthetic fault gouges. *J. Geophys. Res. Solid Earth* **2016**, *121*, 1183–1201. [CrossRef]
89. Dieterich, J.H. Time-dependent friction in rocks. *J. Geophys. Res.* **1972**, *77*, 3690–3697. [CrossRef]
90. Dieterich, J.H.; Conrad, G. Effect of humidity on time- and velocity-dependent friction in rocks. *J. Geophys. Res.* **1984**, *89*, 4196–4202. [CrossRef]
91. Jeppson, T.; Lockner, D.; Beeler, N.; Hickman, S. Strength recovery in quartzite is controlled by changes in friction in experiments at hydrothermal conditions up to 200 °C. *J. Geophys. Res. Solid Earth* **2023**, *128*, e2022JB025663. [CrossRef]
92. Mei, C.; Rudnicki, J.W. Microphysical modeling of fault slip and stability transition in hydrothermal conditions. *Geophys. Res. Lett.* **2023**, *50*, e2023GL103730. [CrossRef]
93. Kolawole, F.; Johnston, C.; Morgan, C.; Chang, J.; Marfurt, K.; Lockner, D.; Reches, Z.; Carpenter, B.M. The susceptibility of Oklahoma's basement to seismic reactivation. *Nat. Geosci.* **2019**, *12*, 839–844. [CrossRef]
94. An, M.; Zhang, F.; Min, K.-B.; Elsworth, D.; He, C.; Zhao, L. Frictional stability of metamorphic epidote in granitoid faults under hydro- thermal conditions and implications for injection-induced seismicity. *J. Geophys. Res. Solid Earth* **2022**, *127*, e2021JB023136. [CrossRef]
95. Nur, A.; Mavko, G. Postseismic viscoelastic rebound. *Science* **1974**, *183*, 204–206. [CrossRef]
96. Savage, J.C.; Prescott, W.H. Asthenosphere readjustment and the earthquake cycle. *J. Geophys. Res. Solid Earth* **1978**, *83*, 3369–3376. [CrossRef]
97. Lachenbruch, A.H.; Sass, J.H. Heat flow from Cajon Pass, fault strength, and tectonic implications. *J. Geophys. Res. Solid Earth* **1992**, *97*, 4995–5015. [CrossRef]
98. Hearn, E.H.; Thatcher, W.R. Reconciling viscoelastic models of postseismic and interseismic deformation: Effects of viscous shear zones and finite length ruptures. *J. Geophys. Res. Solid Earth* **2015**, *120*, 2794–2819. [CrossRef]
99. Tse, S.T.; Rice, J.R. Crustal earthquake instability in relation to the depth variation of frictional slip properties. *J. Geophys. Res. Solid Earth* **1986**, *91*, 9452–9472. [CrossRef]
100. Richter, B.; Stünitz, H.; Heilbronner, R. The brittle-to-viscous transition in polycrystalline quartz: An experimental study. *J. Struct. Geol.* **2018**, *114*, 1–21. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Applied Sciences Editorial Office
E-mail: appls@mdpi.com
www.mdpi.com/journal/appls



Disclaimer/Publisher's Note: The title and front matter of this reprint are at the discretion of the Guest Editor. The publisher is not responsible for their content or any associated concerns. The statements, opinions and data contained in all individual articles are solely those of the individual Editor and contributors and not of MDPI. MDPI disclaims responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-8952-5