A Review of Research on Algorithmic Modeling in Landslide Hazards - Visualization and Analysis Based on VOSviewer

Xiaomin Dai[1,a], Liuyang Xing[2,b], Zijie Ye[2,c], Rong He[3,d], Shengqiang Ma[4,e*]

[1] School of Traffic and Transportation Engineering, Xinjiang University, Saybag District, Ürümqi, Xinjiang, China
[2] School of Business, Xinjiang University, Saybag District, Ürümqi, Xinjiang, China
[3] Xinjiang Transportation Construction Group Co, Saybag District, Ürümqi, Xinjiang, China
[4]College of Civil Engineering and Architecture, Xinjiang University, Saybag District, Ürümqi, Xinjiang, China

[a]email: 44848282@qq.com, [b]email: 1208647315@qq.com, [c]email: 1354216736@qq.com, [d]email: 64921014@qq.com
*Corresponding author: msq-2000@163.com

Abstract. Accompanied by global climate variability and the increase in engineering construction, the problem of landslide disaster is becoming more and more prominent because landslide disaster has a wide range of hazards, high risk, and other characteristics, which makes the application of intelligent algorithms in the field of landslide disaster is particularly important. In this paper, we use VOSviewer knowledge mapping analysis software to count all the databases of Web of Science (WOS) from 2000 to 2023, about the literature related to the research of intelligent algorithms in landslide disaster, and jointly decipher the research status and predict the research trend through high-frequency keywords, high correlation coefficient keywords, keyword clustering map, keyword superposition map, and keyword density map. The research results show that the typical model of landslide disaster mainly includes the application of landslide disaster monitoring and landslide disaster assessment; the scenario primarily provides for the analysis of the cause mechanism of geohazard, geohazard monitoring methods, and geohazard prevention exploration in three aspects. Based on this, it is believed that the field of landslide disasters will have more universality and special geological conditions.

Keywords: Landslide hazard; Intelligent algorithm; VOSviewer

1 Introduction

Landslides are the most hazardous category of geologic hazards after earthquakes and floods [1] due to the characteristics of landslide disasters, such as high danger, suddenness, and a wide range of injuries, which leads to about 300 million people and 3.7 million square kilometers of land suffering from their hazards worldwide[2]. With the

D. Li et al. (eds.), *Proceedings of the 2023 9th International Conference on Architectural, Civil and Hydraulic Engineering (ICACHE 2023)*, Advances in Engineering Research 228,
https://doi.org/10.2991/978-94-6463-336-8_88

development of artificial intelligence, machine learning models and intelligent algorithms are widely used in geologic disaster stability analysis, early warning prediction, and other fields and have achieved good results [3].

As a typical tool for drawing a knowledge map, VOSviewer can reveal a particular field's development status, research hotspots, and future trends. Therefore, this paper uses VOSviewer to analyze the application of intelligent algorithms in the field of landslide disaster, combine the development status of algorithmic models in the field of landslide disaster, explore the application scenarios of intelligent algorithms in the field of landslide disaster, and looking forward to the use and development of intelligent algorithms in the field, which is of great practical significance for the healthy development of various regions of the country.

2 Data and methodology

2.1 Data sources and processing

In this paper, we use all databases of Web of Science (WOS) as the data source of foreign literature and use geologic disaster and model as the search terms to conduct advanced search in all databases of WOS with the search formula TS = (geologic disaster) AND TS=(model), the time range of 2000-2023, and the retrieval date of August 19, 2023, and a total of 474 pieces of literature were obtained by searching, which were taken as the primary data for the research of this paper.

2.2 Analytical tools and methods

Using VOSviewer 1.6.18, export the information such as “author”, “title”, “source”, and “abstract” in the WOS database in a “tab-separated file” format, and name it as “saved-recs. txt” and import it into VOSViewer. By analyzing the co-occurrence of keywords in the input information, we can better understand the current research hotspots and development trends.

3 Analysis of keyword results

3.1 co-occurrence

Keywords are the reflection of the article's theme and the extraction of core content. When a keyword appears many times in a particular research field, it can reflect that the theme and content represented by this keyword is the research hotspot. In this paper, we imported a total of 474 English documents from all the databases of WOS, and there are 14,728 keywords, 127 keywords with a frequency of 15 times or more, the top 30 high-frequency keywords are shown in Table 1, and the top 30 keywords with high correlation coefficient are shown in Table 2.

Table 1. Top 30 high-frequency keywords

Term	Occurrences	Term	Occurrences	Term	Occurrences
model	290	event	97	assessment	79
study	209	effect	93	disaster	76
area	206	system	91	impact	75
analysis	175	hazard	89	approach	74
factor	147	region	89	structure	72
data	145	characteristic	82	value	70
landslide	128	paper	82	level	68
condition	107	slope	82	source	68
process	101	time	82	year	64
sediment	99	water	80	research	63

Table 2. Top 30 high correlation coefficient keywords

Term	Relevance score	Term	Relevance score
landslide susceptibility map	5.598	dem	2.6737
landslide susceptibility mapping	5.3968	water	2.4869
concentration	4.9585	lithology	2.3268
landslide susceptibility	4.4201	distance	2.2363
drought	4.3893	digital elevation model	2.0785
field survey	4.1612	study area	2.055
landslide occurrence	3.9947	climate change	1.9967
geographic information system	3.9826	geology	1.8737
aspect	3.2943	presence	1.7698
gis	3.2872	flood	1.7181
day	3.1387	climate	1.5601
landslide hazard	2.9794	accuracy	1.366
flooding	2.9594	increase	1.2887
road	2.9524	catchment	1.2845
sediment	2.8299	change	1.2603

3.2 Cluster analysis

The English keyword clustering clusters are divided into three categories, and the results of the visualization analysis are shown in Figure 1.

The red area mainly focuses on the research on the conditions for forming different types of geological disasters and their impacts on different fields, including the impacts and differences caused by geological disasters such as drought and floods. The green area mainly evaluates and predicts different types of surface deformation generated in different areas, including landslides, slopes, and other assessments and predictions. The blue area mainly discusses which models are used for evaluation and research under different geological disaster backgrounds.

3.3 Superposition analysis

The keyword overlay view is shown in Figure 2, where the year represents the average year in which keywords appear. The year distribution of this view can be used to display research hotspots and trends in the field. Navy and yellow represent before 2013 and after 2016, respectively. In recent years, accuracy, scale of geological hazards, prediction, evaluation, and evaluation have been the main keywords in the research field of digital models in geological disasters. This means that the accuracy of models has gradually been valued, and models' evaluation and prediction ability have been affirmed.

3.4 Density analysis

The density analysis reveals that the research in geohazards mainly focuses on modeling, region, and prediction, indicating that this is the invariable topic and research focus of the research in this field, and the keyword density view is shown in Figure 3.

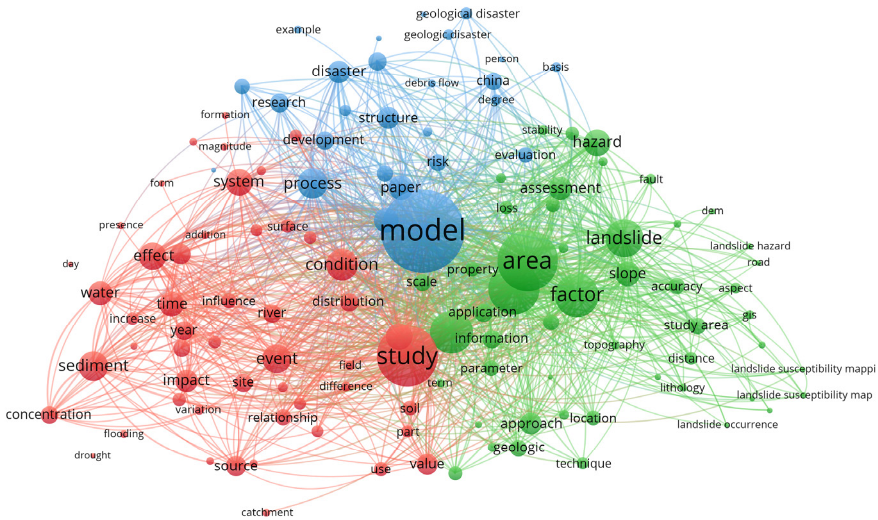


Fig. 1. Keyword Co-occurrence and Clustering View

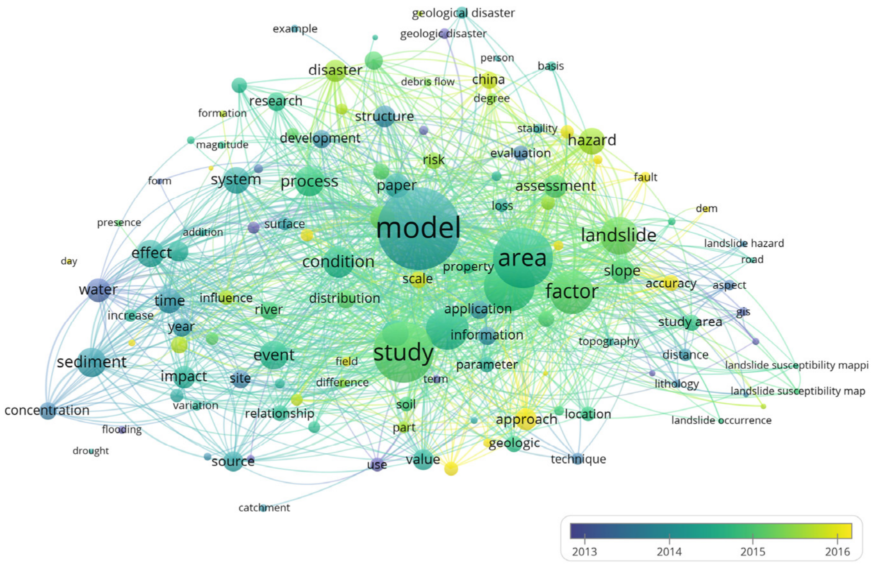


Fig. 2. Keyword Overlay View

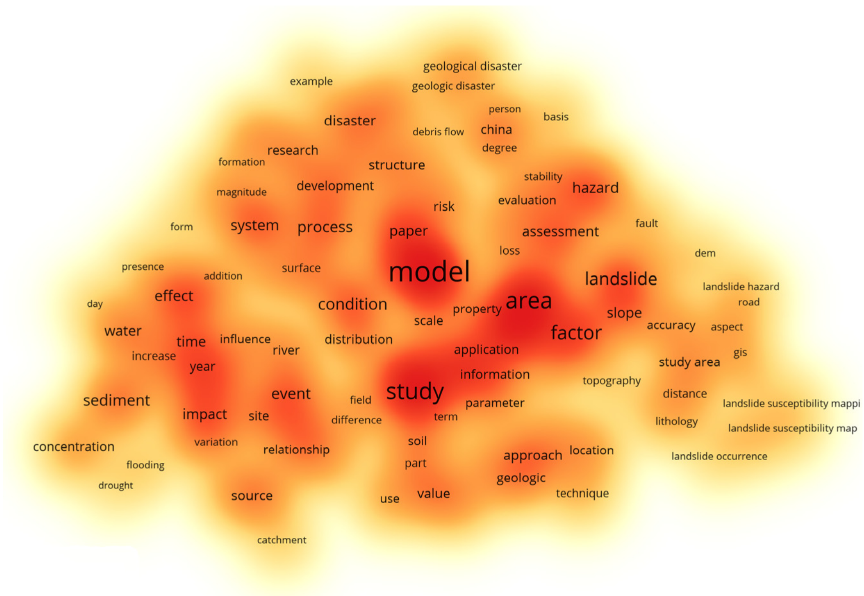


Fig. 3. Keyword Density View

4 Standard Models for Landslide Disaster Prevention

4.1 Landslide disaster monitoring

In landslide disaster monitoring mainly through the landslide early warning model and landslide displacement prediction model, two methods for landslide monitoring and early warning, which can be subdivided into physical models, mathematical models, empirical models, artificial intelligence models, and other categories of models, which are as follows Table 3.

Table 3. Landslide Hazard Detection Model

Model Category			Working Principle
Landslide Early Warning Model	Data threshold model [4]		Thresholds are set by comparing with historical data, and when the monitoring data exceeds the entries, the system issues a landslide warning.
	Geographic Information System (GIS) model [5]		Remote sensing and geographic information data were used to compare the landslide features; they were recognized as a landslide when the features were similar.
	Physical and numerical simulation models		Based on the physical characteristics of landslides, they were analyzed by numerical simulations and biological experiments and were identified as landslides when the traits were similar.
	Empirical model		A curve-fitting method is applied to make predictions based on the empirical rule of cyclic changes in landslide displacement.
Landslide Displacement Prediction Model	chronological model [6]-[7]	Autoregressive moving average model (ARMA model) Autoregressive integral moving average model (ARIMA model)	Landslide analysis based on temporal variation characteristics by monitoring data.
	physical model	Finite element model [8] Slope stability analysis model	Prediction of landslide displacements by numerical simulations and physical experiments based on the physical characteristics of landslides.
	artificial intelligence model	Support vector machine model [9] Random forest model [10] BP neural network model [11] Convolutional neural network model [12]	Based on artificial neural networks and machine learning algorithms, landslide displacements are predicted by training and learning from monitoring data.

4.2 Landslide hazard assessment

Landslide disaster prevention, in addition to monitoring, also needs susceptibility assessment and risk assessment through the analysis of landslide geology and related disaster-causing factors. The data obtained by a series of simulation modeling analyses can determine the probability of occurrence of landslides and the degree of risk, and its commonly used models are as follows: Table 4.

Table 4. Landslide Hazard Evaluation Model

Model Category		Working Principle	Common Models
Susceptibility assessment	Statistical regression model	A regression model of landslide susceptibility is developed by statistically analyzing data on landslides that have occurred. The model calculates landslide risk probability or grade based on geological and topographical factors.	Binary logistic regression model Decision tree modeling [13]
	Physical model	Numerical simulations or physical experiments predict landslide susceptibility based on geomechanics principles and mechanical modeling, combined with the region's geological, topographical, and hydrological information.	Finite Element Models Particle flow model
	Multi-indicator evaluation model (MIME)	The conditions and factors for the occurrence of landslides are quantitatively assessed, and a landslide hazard assessment index is obtained through statistical analysis.	Gray correlation model Fuzzy integrated evaluation model [14]
Hazard assessment	Evaluation of index models	Based on the physico-mechanical properties of landslides, combined with geological, topographical, and hydrological factors, the stability analysis is carried out by establishing a mechanical model to predict the probability of occurrence and the danger of landslides.	Landslide hazard index method Information entropy model [15]
	Physical-mechanical model	Based on the physico-mechanical properties of landslides, combined with geological, topographical, and hydrological factors, the stability analysis is carried out by establishing a mechanical model to predict the probability of occurrence and the danger of landslides.	Slope Stability Analysis Model Finite element analysis model
	Statistical model	Based on the statistical data of historical landslide events and related factors, a mathematical model was developed to assess the risk of landslides through regression analysis, time series, and other methods.	Logistic regression modeling [16] Hierarchical analysis (AHP) [17]

5 Several Application Scenarios of Intelligent Algorithms in Geological Hazards of Infrastructure Slopes

Research on slope geological hazards based on intelligent algorithms mainly focuses on macro-level driving forces, mid-level modeling and evaluation, and micro-level scenario analysis of disaster warning, prevention, and emergency response.

5.1 Scenario 1: Analysis of the internal driving force of slope geological disasters

In recent years, slope geohazards have been dominated by landslides, avalanches, and mudslides, and it is crucial to explore the breeding process of disaster-causing through intelligent algorithms to sort out the disaster characteristics of the occurrence of geohazards, which is essential for the prevention and control of the construction and maintenance safety of infrastructure projects. For example, Wu Shaoyuan et al. used informative simulation to give the data of collapse and landslide geohazards in Xiamen City as an evaluation index. They concluded that topography and geomorphology are more likely to trigger collapse and landslides than precipitation [18]. Yang Longwei et al. derived the characteristics of high landslides, the structural features of landslide-prone geology, and the three stages of high remote landslide chain disaster through reading literature, field geological investigation, and generalized classification, which provided a reference for the study of high landslide disaster risk [19]. Di Jingyue et al. found that the precipitation extremity factor is more likely to cause geologic disasters than the precipitation duration factor by investigating precipitation-caused disasters in Enshi Prefecture [20]. The above scholars effectively explored the topography, geomorphology, and each influencing factor through intelligent algorithms. They put forward the internal driving force of slope geohazards under different geographical conditions, which provides the reference degree index for the subsequent research.

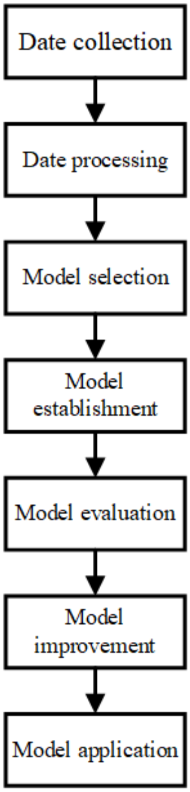
Research on slope geological hazards based on intelligent algorithms mainly focuses on macro-level driving forces, mid-level modeling and evaluation, and micro-level scenario analysis of disaster warning, prevention, and emergency response.

5.2 Scenario 2: Slope Geohazard Modeling and Evaluation

With increasingly severe disaster categories, application-oriented disaster risk assessment has to guide solid significance. The traditional risk assessment and analysis methods cost a lot of workforce and resources, and the data obtained is also limited. With the development of science and technology, scholars combine different modeling methods to analyze geologic hazards in detail to improve the prediction accuracy of geologic hazards to reduce the risk of disaster to the population. For example, Dai Jun et al. used a two-layer weight model to refine the classification of disaster-hazardous areas. They used the optimization AHP method to construct a multi-indicator analysis in the spatial optimization of clusters to improve the forecasting accuracy of disasters [21]. Feng Fan et al. used the information quantity and hierarchical analysis methods to statistically analyze the occurrence of geologic disasters in Wubao County, Shaanxi Province, using the slopes divided by the hydrological method as the evaluation unit. They formulated refined management measures to prevent and control geologic disasters, which provided a basis for managing the relevant governmental service departments [22]. He Shuhong, Huang Zhenxiong, and Zheng Shangping established a Copula model with the annual statistical data of geologic disasters in Yunnan Province from 1991 to 2020 as a support and proposed a model of risk apportionment of geologic disasters in Yunnan Province [23].

Table 5. Model classification diagram

	advantage	disadvantage	Model Category	assessment process
Deterministic coefficient model	Easy to understand Strong stability Strong interpretability Wide application range	Sensitive to Outlier Large limitations Multiple constraints	Linear regression Polynomial regression Exponential smoothing	
Frequency ratio model	Accurate data processing Strong predictive ability High scalability	Model complexity Weak interpretability High data requirements Possible overfitting	Time Series Model (ARMA) Vector autoregression (VAR)	
Copula model	Wide application range Can effectively quantify risks Robust targeting of different sub-models	Model complexity High data requirements The process is complex and challenging to understand	Gaussian Copula [24] T-Copula	



A comprehensive analysis of the above geohazard modeling and their respective advantages and disadvantages are shown in Table 5 below.

5.3 Scenario 3: Early Warning and Prevention of Slope Geological Hazards

The early warning and prevention of slope geological disasters, as the core element, are directly related to the safety of people's lives and property. In the current context of rapid technological development, the early warning of geological disasters is guided by speed and accuracy, effectively constructing advanced devices combined with intelligent algorithms.

Unlike traditional surveying and mapping techniques, drone remote sensing technology reduces workload. It provides work efficiency, especially for digital surveying and mapping in high mountain areas with large mountain slopes. Wu Mingyuan, Luo

Ming, and others identified the geology of counties in western Yunnan, China, established optical interpretation and InSAR identification of regional-level landslides and potential landslides, and constructed a combination model of optical remote sensing and InSAR technology [25]. For areas where numerous disaster-causing factors cannot directly evaluate and predict geological hazards, neural networks can enable machines to conduct deep learning, which can then be used for subsequent judgment and analysis of similar terrain and geology. For example, Zhang Linfan and others used data mining technology to analyze the controlling factors of field loess landslides in Yining County, Xinjiang, built a prediction model of regional landslide susceptibility through the BP neural network model, and combined with the DEM data and remote sensing interpretation data of the study area, obtained the landslide susceptibility zoning map of the area [26]. Chen Yinfeng used an elevation change model to calculate the amount of soil change, fitted the trend of geological disasters through a polynomial mathematical model, and figured out their differences, providing a theoretical basis for predicting geological disasters in the study area [27].

6 Conclusion

Intelligent algorithms in landslide disasters for prevention, monitoring, governance, and other work provide a good foundation and are a solid guarantee for carrying out landslide disaster research. Comprehensive analysis of its research status, future trends, and results in the current level of study at home and abroad, mainly reflected in the following aspects:

Application model: The model types are divided into monitoring and evaluation models. Among them, the detection models include physical, mathematical, empirical, and artificial intelligence models. The evaluation models are mainly hazard evaluation and susceptibility evaluation. At present, the model types are wide and varied with high accuracy. However, the generalization ability could be better, and it is necessary to research a more universal model to help people prevent geohazards in the future.

Application Scenario: By analyzing the use of intelligent algorithms in landslide mechanism, landslide early warning and risk assessment from three perspectives: macro, meso, and micro. Geological conditions are different in different areas, so it is one of the future development directions to explore more suitable models according to particular geological conditions.

Intelligent algorithms have been increasingly emphasized in the study of landslide disasters, with scholars focusing on monitoring, early warning, and assessment to have the ability to cope with landslide disasters gracefully.

Acknowledgments

This study was supported by Xinjiang Key R&D Program Projects (grant numbers 2022B03033-1), Natural Science Foundation of Xinjiang (grant number 2020D01C027), Xinjiang Uygur Autonomous Region "Dr. Tianchi" Project.

References

1. Yiping Wu, Kunlong Yin, Lixia Chen. Comparison and application of mathematical models for landslide spatial prediction [J]. Geological Science and Technology Intelligence, 2007 (06): 95-100.
2. Shu Jiang, Yifeng Wang, Ke Liu, et al. Summary of research on spatial prediction methods for landslide disasters [J]. People's Yangtze River, 2017,48 (21): 67-73.
3. Changfu Chen, Wei Li, Jiarui Zhang, et al. Progress in Intelligent Analysis and Design of Mountain Highway Slope Engineering [J]. Journal of Hunan University (Natural Science Edition), 2022,49 (07): 15-31.
4. Xiepan Liu, Kunlong Yin, Changgui Xiao et al. Landslide Meteorological Warning Based on I-D-R Threshold Model [J/OL]. Earth Science: 1-15 [2023-08-22].
5. Fuxi Jin, Quanwei Yuan, Shuaishuai Qin. A Landslide Meteorological Warning Model Based on GIS and K-means Clustering Method [J]. Metal Mines, 2018 (06): 162-166.
6. Yongqi Zhao, Jiancheng Li, Xinyu Xu, et al. Using GOCE and GRACE satellite observation data to determine a static gravity field model [J]. Journal of Geophysics, 2023,66 (06): 2322-2336.
7. Min Li, Mingfeng Zhang, Liming Zhu, et al. Hydrological drought prediction based on Bayesian theory[J/OL]. Hydrology:1-8[2023-05-06].
8. Yanhui Liu, Junbao Huang, Ruihua Xiao, et al. Research on a Regional Landslide Disaster Warning Model in Fujian Province Based on Random Forests [J]. Journal of Engineering Geology, 2022,30 (03): 944-955.
9. Chang Zhang, Zhiwen Chen, Meng Wei. et al. 3D landslide geohazard monitoring method and application based on BIM[J]. Journal of Chengdu University of Technology (Natural Science Edition),2017,44(03):377-384.
10. Haishen Geng, Wenbin Lv, Sun Li. et al. Analysis of the destabilization mechanism of landslides and research on early warning prediction[J]. Journal of Northwest Normal University (Natural Science Edition),2021,57(06):103-109.
11. Gorsevski, P.V., Brown, M.K., Panter, K. et al. Landslide detection and susceptibility mapping using LiDAR and an artificial neural network approach: a case study in the Cuyahoga Valley National Park[J]. Ohio. Landslides 13, 467–484 (2016).
12. Youssef, A.M., Pourghasemi, H.R., Pourtaghi, Z.S. et al. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparing their performance at Wadi Tayyah Basin[J]. Asir Region, Saudi Arabia. Landslides 13, 839–856 (2016).
13. Yanhui Liu, Ranco Fang, Yongchao Su. et al. Research on regional landslide disaster warning model based on machine learning[J]. Journal of Engineering Geology,2021,29(01):116-124.
14. Cong Li, Qinghui Jiang, Chuangbing Zhou. et al. Research on landslide early warning criterion based on instance inference system[J]. Geotechnics,2011,32(04):1069-1076.
15. Zhifu Zhu, Shu Gan, Jianming Zhang. et al. Identifying geological landslide potentials in Cangshan Mountain by combining SBAS-InSAR technology and information entropy[J]. Surveying and Mapping Bulletin,2022(11):13-19.
16. Felicísimo, Á.M., Cuartero, A., Remondo, J. et al. Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study[J]. Landslides 10, 175–189 (2013).
17. Zhao, H., Kowalski, J. Bayesian active learning for parameter calibration of landslide runout models[J] Novel Slope Monitoring Technology Methods and Their Combined Models Landslides 19, 2033–2045 (2022).

18. Jiehui Huang, Mowen Xie. Research on the selection of data source parameters of satellite data source for geohazard monitoring onboard SAR satellites[J]. Advances in Geophysics,2019,34(06):2169-2172.
19. Shaoyuan Wu. Evaluation of geologic hazard susceptibility of landslide and landslide in Xiamen City based on informativeness modeling[J]. Safety and Environmental Engineering,2019,26(03):22-27.
20. Longwei Yang, Yangqing Xu, Jingxuan Gao, Qiangqiang Jiang, Guoqiang Yan, Nan Zhang. Progress of research on the geological environment of high-level landslide disaster [J/OL]. Water Conservancy and Hydropower Technology (In English and Chinese):1-8[2023-05-05].
21. Jingyue Di, Fengwen Xu, Huanlian Li, Yumei Li, Guoping Zhang, Chunli Chen, Zhiwei Fang, Hongjun Bao. Analysis of the causes of heavy rainfall-induced geological disasters in Enshi Prefecture[J]. Meteorology,2022,48(10):1321-1332.
22. Jun Dai, Wenjun Chen, Shujuan Shen. Spatial optimization of rural settlements in highland mountainous areas based on comprehensive disaster risk assessment - A case study of Ledu District, Haidong City, Qinghai Province[J]. Disaster Science,2021,36(04):119-125+132.
23. Fan Feng, Yaming Tang, Xueshu Pan, Xiaohao Wang, Yuxuan Zhao, Xuan Bai. Discuss the risk evaluation methods of geologic hazards at different scales--Take Wubao County, Shaanxi, as an example[J]. Chinese Journal of Geological Hazards and Prevention,2022,33(02):115-124.
24. Yu Huang, Bingzhe Zhang, Huizhen Pang, et al. Research on wind speed prediction method based on hybrid Copula optimization algorithm[J]. Journal of Solar Energy,2022,43(10):192-201.
25. Junxia Zhang, Wubin Huang, Antai Li, Xiumei Yang, Qian Li, Hongwei Bian. Refined meteorological risk early warning forecast of major geologic hazards in Gansu Province[J/OL]. Arid Zone Geography:1-11[2023-05-05].
26. Mingyuan Wu, Ming Lu, Touhai Liu. Comprehensive identification of potential and old landslides based on optical remote sensing and InSAR technology - A case study in Northwest Yunnan[J]. Chinese Journal of Geological Hazards and Prevention,2022,33(03):84-93.
27. Linfan Zhang, Jiayun Wang, Maozhou Zhang, Shebin Chen, Tao Wang. Evaluation of regional landslide susceptibility based on BP neural network[J]. Northwest Geology,2022,55(02):260-270.