https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

Review article: Research progress on influencing factors, data, and methods for early identification of landslide hazards

Heng Lu [a,b], Zhengli Yang [a,b,*], Kai Song [c], Zhijie Zhang [d], Chao Liu [a,b], Ruihua Nie [a,b], Lei Ma [e], Wanchang Zhang [f], Gang Fan [a,b], Chen Chen [a,b], Min Zhang [g]

a State Key Laboratory of Hydraulics and Mountain River Engineering, Sichuan University, Chengdu, Sichuan, China.

b College of Hydraulic and Hydroelectric Engineering, Sichuan University, Chengdu, Sichuan, China.

c Changjiang Survey, Planning, Design and Research Co., Ltd., Wuhan, China.

d School of Geography, Development & Environment, University of Arizona, Tucson, AZ, USA.

e School of Geography and Ocean Science, Nanjing University, Nanjing 210093, PR China

f Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, PR China

g School of Environmental and Geographical Sciences, Shanghai Normal University, Shanghai 200234, China.

* Corresponding author. Tel.: +86 13882016220. E-mail address: yangzhengli@scu.edu.cn (Zhengli Yang).

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

Abstract

The early identification of potential landslide hazards has always been a hot and difficult issue in the field of international landslide research. In recent years, many scholars have conducted extensive and beneficial explorations in this field, making significant contributions to the effective prevention of landslide disasters. However, until now, there are very few review documents on summarizing such valuable experience in the system, which makes it difficult to meet the ever-increasing demand of researchers in scientific documents. To address the gap, this paper systematically reviews 843 documents collected by the two data platforms of Web of Science (WOS) and Scopus from 1971 to 2023 by using the bibliometric analysis software. This paper first figures out the internal relationship between documents by analysing their spatial and temporal distribution characteristics, and then emphatically analyses the application, advantages and disadvantages of different early identification methods based on the influencing factors of landslide disaster formation and multi-source data acquisition links. And finally, this paper discusses the challenges and development trends in this field from four aspects of cooperative analysis, multi-source data, topic analysis and research trends, and puts forward some suggestions. This research can help researchers to use various early identification methods reasonably and provide summary and integration services of scientific document achievements for efficient research in this field.

Keywords: Landslide hazards; early identification; multi-source data fusion; machine learning; bibliometrics

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

1 Introduction

The early identification of potential landslide hazards is to identify the possible location and range of landslides before the occurrence of such disasters. The possibility of occurrence of potential landslides is determined, and the corresponding prevention and control measures are taken to effectively avoid the risk of landslide disasters (Li et al., 2021; Chang et al., 2023). Just like the research conducted by Yannick Thiery et al. (2020) in France, the use of practical mapping tools, remote sensing, network services and other new technologies can provide higher quality landslide data information, thereby improving prediction accuracy. Most of the landslides follow their evolution laws. Therefore, the effectiveness of early identification of potential landslide hazards largely determines the success rate of landslide disaster prediction and early warning (Yin et al., 2023). They often undergo creep with different deformation characteristics before failure (Scoppettuolo et al., 2020), and have a uniform deformation process from deceleration to acceleration, providing time for data collection and analysis in early identification (Saito M. 1969). As shown in Figure 1, the uniform deformation stage is the main stage of early identification.

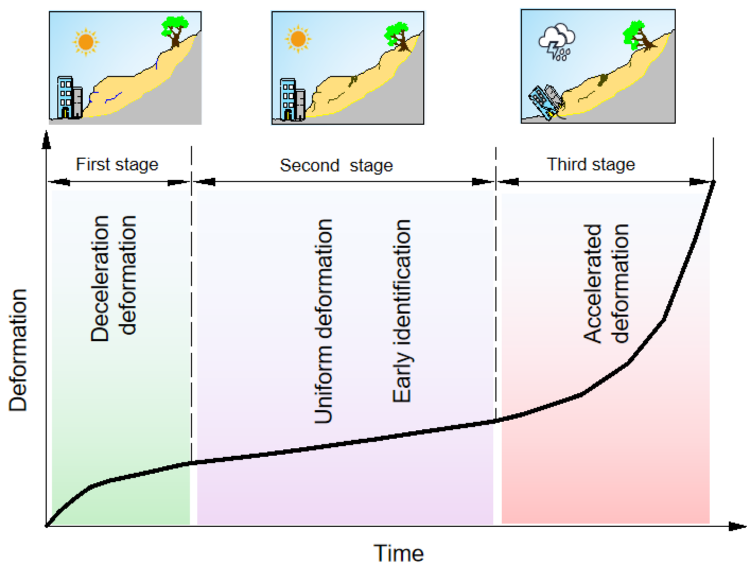


Fig. 1. Schematic diagram of landslide deformation stage.

Facts have proved that scientific and reasonable early potential disaster identification can effectively avoid the occurrence of landslide disasters or reduce losses to a large extent (Sreelakshmi et al., 2022). For example, He et al. (2005) applied the quantitative theory to

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences Discussions

determine the role and correlation of the main dynamic factors, and compared with the actual dynamic pattern of Xintan landslide, showing that it coincides with the actual slip mode and formation mechanism of Xintan landslide. Wang et al. (2023) discovered that there is an obvious spatial and temporal relationship between vegetation anomaly and landslide deformation in the creep stage of landslides, which indirectly reflects the gradually unstable evolution process of landslides in southwest China and provides a theoretical basis for the identification of high vegetation coverage areas. The newly-developed machine learning model by Bui et al. (2019) has become a powerful tool to mitigate and manage landslide disasters in Lang Son, Vietnam. For the convenience of the author's reading and inquiry, the main abbreviations used in this article are shown in Table 1.

Table 1 List of the main abbreviations.

Abbreviation	Description	Abbreviation	Description
AI	Artificial Intelligence	IoT	Internet of ThingS
ANN	Artificial Neural Network	InSAR	Interferometric Synthetic Aperture Radar
AUC	Area Under Curve	LLR	Log-likelihood rate
CF	Certainty Factor	IOT	Internet of Things
CNN	Convolutional Neural Network	LSI	Landslide Susceptibility Index
DEM	Digital Elevation Model	LSM	Landslide susceptibility mapping
DL	Deep Learning	ML	Machine learning
GIS	Geographic Information Systems	RS	Remote Sensing
GLiM	The new global lithological map	SRTM	Shuttle Radar Topography Mission
GPS	Global Positioning System	USGS	The United States Geological Survey

Landslide disasters often occur in hidden and high-elevation mountain areas. There are many influencing factors, which have complex interrelationships, and the landslide data is very scarce. The uncertainties have increased the difficulty of early potential disaster identification. In the past few years, scholars all over the world have done a lot of research, and they will carry out more research based on these documents in the future. Therefore, it is necessary to systematically summarize the previous experience. However, so far, there are few review papers on early identification of potential landslide hazards, especially those based on bibliometric analysis methods. Only 20 related review documents were retrieved from WOS and Scopus, and several of them are representative, as shown in Figure 2 (Chou et

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

al., 2015; Lagomarsino et al., 2015; Reichenbach et al., 2018; Whiteley et al., 2019; Imani et al., 2021; Li et al., 2021; Jiang et al., 2022; Bagwari et al., 2021; Xu et al., 2023).

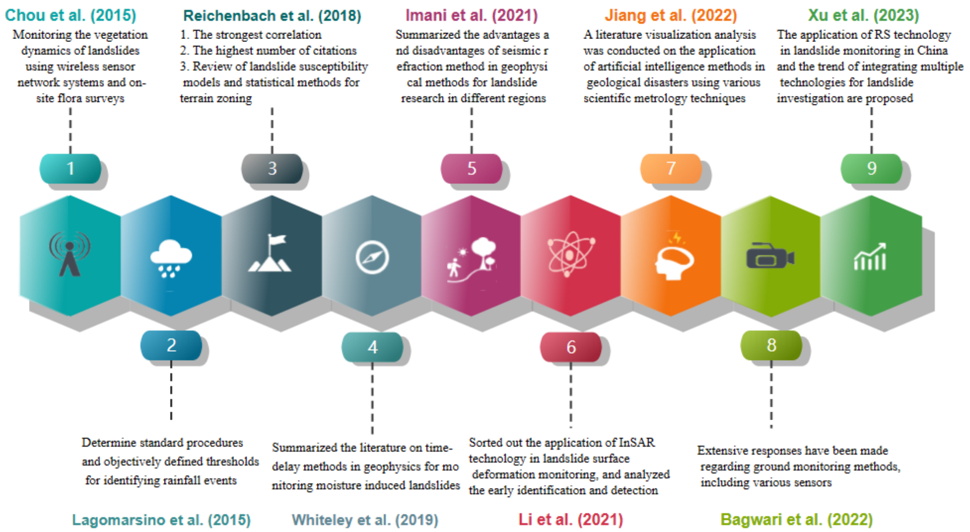


Fig. 2. Several representative literature reviews.

Figure 2 shows that the review documents only appeared in this filed in recent years, and it is not until 2022 that scholars began to make document review on the application of artificial intelligence methods in geological disasters by bibliometric methods. Through further analysis, it is not difficult to find that there are many problems in the existing review documents, such as narrow research area, long retrieval time of the document database, and divergent research direction. In order to help scholars quickly grasp the development process and research direction of this field, based on the existing review documents and combined with the background knowledge and early identification research experience of landslides, this paper establishes a document sample database through the document data platforms of WOS and Scopus, analyzes the spatial and temporal distribution characteristics of documents with VOSviewer, SCImago graphica and other tools, and emphatically analyzes the application, advantages and disadvantages of 5 main early identification methods on the basis

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

of clarifying the influencing factors and data acquisition of landslides. Meanwhile, this paper discusses the existing problems in this field with CiteSpace from the aspects of cooperative analysis, multi-source data, topic analysis and research trends, and puts forward some suggestions. It also outlines future research to provide help for optimizing the early recognition method and improving the recognition accuracy in this research.

2 Document Statistical Analysis and Results

2.1 Establishment of document sample database

In this research, two recognized, reliable and highly reputable academic publishing databases, namely WOS and Scopus are adopted, which have multi-disciplinary fields and can cover all the topics involved in the research of early potential landslide hazard identification by scholars around the world. On May 8, 2023, all documents published in this field from 1971 to 2023 were comprehensively retrieved. The process of data retrieval, processing and analysis is shown in Fig. 2.

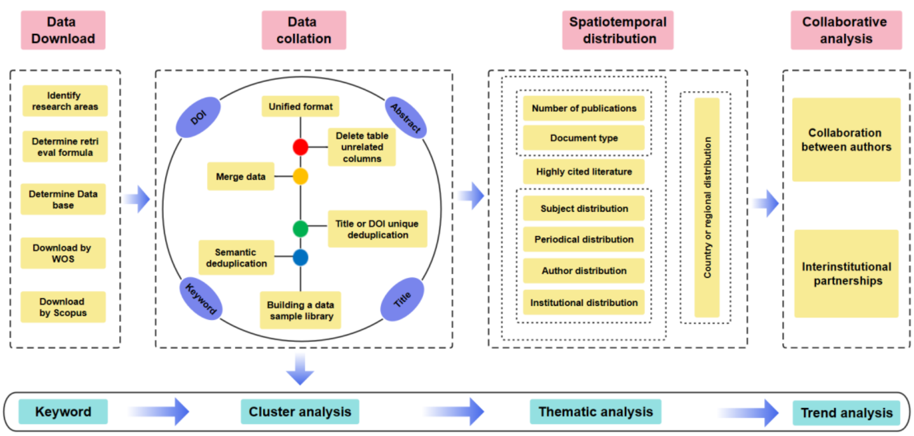


Fig. 3. Flowchart for bibliometric analysis of literature.

According to the research objective, the search strategy TS =“ landslide* ” AND “ early ” AND “ identif* ”. In WOS, a total of 599 records were retrieved. In Scopus, a total of 791 records were retrieved. A total of 1,390 records were retrieved. Due to row errors, omissions

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

and duplications and other phenomena of the raw data, Microsoft 365 was first adopted for the format adjustment of the data downloaded from the two databases. Then, the EndNote X9 reference manager was used to remove duplicate documents based on the uniqueness of the title and DOI. The Manual deduplication was applied to the situations in which the title or the author has different spelling versions. The publication years, journals, citation frequency, countries, institutions, co-citation references and co-citation authors, etc. were selected as the data sample database in this paper. After that, documents that are not related to the topic of the paper such as campaigns, cultural heritage, isotopes and epidemic diseases were deleted to form a final document sample database, and a total of 843 records were generated.

2.2 Analysis method

This research is carried out by the bibliometric method, and the mainly used software are VOSviewer, SCImago graphica and CiteSpace, etc., as shown in Table 2.

Table 2 Quantitative analysis tools involved

NO.	Tool Name	Link	Functional application	Version adopted
1	WOS	https://www.webofscience.com/	Download literature data	Online version
2	Scopus	https://www.scopus.com/	Download literature data	Online version
3	Microsoft 365	https://www.microsoft.com/	Literature data organization, unified format, text writing, etc	Latest version
4	WPS Office	https://www.wps.cn/	Calculation and drawing of statistical charts Fig. 4; Fig. 5; Fig. 6	12.1.0.15.120
5	EndNote	https://endnote.com/	Document duplication, document reading, and citation in reference format	X9
6	CiteSpace	https://citespace.podia.com/	Create a collaborative network visualization graphFig. 9; Knowledge Graph of Multiple Source Data Fig. 10; Keyword clustering analysisFig. 11; Keywords with citation bursts Fig. 12	R6.2.R7
7	VOSviewer	https://www.vosviewer.com/	Data Text Format in Country or Region Distribution Fig. 6; Create keyword cloud maps of landslide influencing factors Fig. 7	1.6.16
8	SCImago graphica	https://www.scimagojr.com/	Create partial maps of country or regional distribution Fig. 6	1.0.23
9	Snipaste	https://www.snipaste.com/	Screenshot software: Try to maintain the original image quality as much as possible	Latest version

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Discussions

10	Micro word Cloud	https://design.weiciyun.com/	Intelligent automatic recognition model word cloud Fig. 9	Online version

As the above-mentioned software have unique functions, different advantages and limitations, this paper selects some of these functions and applies them comprehensively to give full play to advantages of the software and do a better job in bibliometric analysis. For example, VOSviewer can provide a brightly colored knowledge mapping, and help researchers to quickly view the frequency of occurrence of different countries, authors and keywords (van Eck 2010). The number of connection lines between nodes can be used to visually show their relevance degrees. However, users are not allowed to edit, so they cannot handle the problem of overlapped nodes and labels. As a popular free graph maker software, SCImage graphica can generate a text file into a symbiosis network diagram to achieve geographic visualization, so as to understand the partnership between countries (Wang et al., 2023). SCImage graphica is a lightweight software with limited functions, so it is a bit complicated for S to convert the text file required for graphic drawing. VOSviewer can be first used to cluster the national document publication information in the document database, and then to convert the *.txt format to the *.gml format in the data sample database for S recognition. CiteSpace, as a commonly used bibliometric software, can visualize the information of keywords, authors, countries, and institutions, etc. to visually identify the internal connections between scientific documents and display the development trends and dynamics of the research field (Chen et al., 2017). Although the software allows users to edit, it needs to update its version frequently and pay to use its advanced version. In addition, the operation is relatively complicated, and the computer configuration requirements are high, so the overlap of nodes and labels is still a thorny problem.

2.3 Spatial and temporal distribution characteristics of documents

The data of the document sample database established in this paper has been provided

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

with a series of cleaning operations such as deduplication, merging and manual screening. Therefore, the statistical distribution data of the number of publications is determined based on the document sample database, which may be slightly different from the document visualization analysis results displayed on the WOS and Scopus platforms.

2.3.1 Time distribution

As shown in Fig. 4, the annual number of publications shows an overall upward trend, indicating that the research on early identification of potential landslide hazards has received more and more attention from scholars around the world. The research in this field can be roughly divided into three stages: The first stage is the embryonic period (1971-1995), in which the research was few. There was no research in some years, and the total number of publications was only 11. Quigley et al. (1971) was the first to conduct some research. They focused on the landslides in the gorge areas, Toronto, and then analyzed from the three dimensions of geology, mineralogy and engineering characteristics, discovering that the landslide is mainly caused by the expansive clay minerals contained in the clay layer. The research enlightened subsequent scholars that the analysis of soil mineralogy shall be included as a valuable part of this field.

The second stage is the slow development period (1996-2008): Since 1996, there has been continuous research in this field, but the annual number of publications was less than 10, and the total number of publications was only 63. Some areas began to establish landslide databases (Devoli et al., 2007), and field survey and ground monitoring were the main identification means at this stage (Fritsche et al., 2006). With the rise of satellite remote sensing technology, some new technologies have been initially tried. For example, the technology based on synthetic aperture radar has been applied in volcanic landslide map (Weissel et al., 2004), and the experimental landslide prediction model are being established (Chen et al., 2008).

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

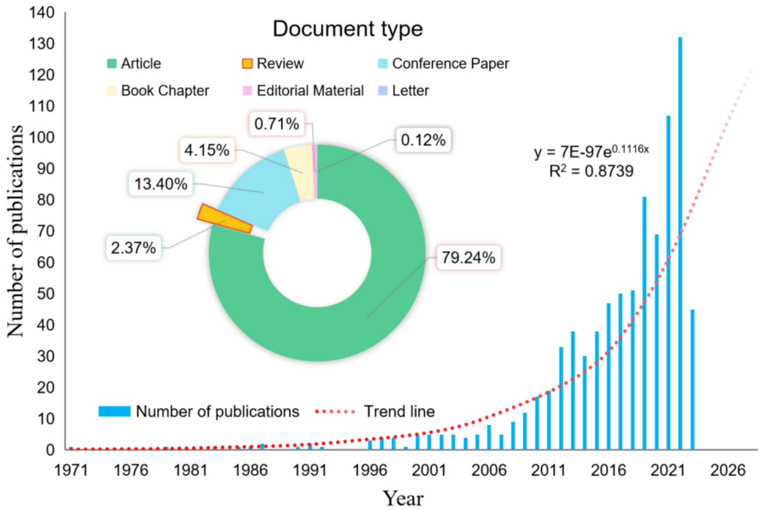


Fig. 4. Number of publications.

The third stage is the rapid development period (2009-present), in which the number of publications increased, with an average of more than 10 per year. In 2009, the number of publications was 12, which was the least but exceeded the sum of the number of publications in the first stage. After 2018, the number of publications increased significantly, especially in the past three years, it showed a blowout trend, and exceeded 100 for the first time in 2021. In 2022, a peak year, the number of publications was as high as 132, which was not only 12 times the total number of publications in the first stage, but also more than 2 times the total number of publications in the second stage. In this paper, only the data before May 8, 2023 are retrieved, and the number of publications this year has reached 45. According to the trend line of the total number of publications, the trend line of the first five cycles was predicted, indicating that the number of publications in this filed will continue to show an increasing trend for at least the next five years. This trend corresponds roughly with the exponential growth function (the calculating formula can be fitted by the trend line of Microsoft Excel2022).

In addition, Fig. 4 indicates the distribution of document types. There are 6 types in the document sample database, of which Review only accounts for 2.37%. Although there are not many Review papers, they are often highly cited. Table 1 shows that in the TOP 10 highly cited papers, two are Review papers, accounting for 20%. Both of the two Review papers were published in 2018, and ranked first and third in the ranking of the number of papers cited in this field in just five years. From here, it can be seen that researchers have paid great attention to the Review papers in the research field of early identification of potential landslide hazards. For Review papers, there is a big gap between the number of publications and the demand. Therefore, it is necessary to strengthen the paper review in this field. Meanwhile, the highly cited documents listed in Table 3 are also the core references in this research, which inspires the research in later chapters of this paper.

Table 3 Highly cited literature information table（TOP 10）

NO.	Cite Frequency	References	document type	Main Contribution
1	707（WOS） 805（Scopus）	Reichenbach et al. (2018)	Review	A systematic review of 565 literature from 1983 to 2016 was conducted on statistical methods for landslide susceptibility modeling and terrain zoning
2	424（WOS） 465（Scopus）	Gorum et al., (2011)	Article	The distribution map of landslides caused by earthquakes was introduced, and a large number of satellite images and aerial photos before and after the earthquake were used to make up for the shortcomings of the early landslide inventory
3	272（WOS） 279（Scopus）	Segoni et al., (2018)	Review	Provide literature references for commonly used or advanced methods to identify standard procedures and threshold definitions for rainfall events, filling the gaps in validation processes and program explanations
4	221（WOS） 229（Scopus）	Intrieri et al., (2018)	Article	Through the analysis of Sentinel 1 satellite data, precursor signals of landslide deformation in Maoxian County were identified, indicating that satellite radar data is expected to monitor large-scale and short-term landslides
5	203（WOS） 214（Scopus）	Pourghasemi and Kerle (2016)	Article	Tested the success and prediction rates of data-driven models, providing a new approach for mapping landslide susceptibility at the regional scale due to the lack of data and methods

6	184（WOS） 205（Scopus）	Pradhan (2011)	Article	Propose a method based on fuzzy logic relationships to draw landslide susceptibility maps, and combine GIS to screen out the main factors affecting landslides and analyze their relationships
7	181（WOS） 199（Scopus）	Zhao et al. (2012)	Article	Using InSAR data to investigate regional landslide activity and verifying it with adjacent satellite orbit data, in order to solve the problem of landslide activity in areas that are difficult to identify using traditional methods
8	169（WOS） 194（Scopus）	Harp E.L. et al. 2011	Article	This article summarizes the development of landslide inventory maps caused by earthquakes from aspects such as landslide inventory standards, mapping standards, early inventory maps, digital inventory maps, and disaster maps
9	170（WOS） 183（Scopus）	Barnard et al. (2001)	Article	Directly use magnetic tape measurement and simple triangulation technology to measure active slopes after earthquakes and analyze the impact of human activities on landslides
10	142（WOS） 169（Scopus）	Hong et al. (2006)	Article	Global rainfall intensity duration thresholds were specified based on landslide events and TMPA rainfall characteristics to identify rainfall induced landslides

2.3.2 Spatial distribution

Fig. 5 indicates the TOP 10 publication status of discipline, journal, author and institution. It can be seen from Fig. 5(a) that “Geology” is the discipline with the highest number of publications, which is 374. The long research history and scholars focus have contributed to high output in this direction. As early as 1997, people began to study the differences between collapse structures and geometric shapes in fracture systems, as well as the mechanisms by which they lead to catastrophic landslides. After that, scholars discovered from different geological perspectives that areas with landslide hazards such as natural vegetation caves (Kuriakose et al., 2009), shallow seismic structures (Camerlenghi et al., 2009), lithology (Del Gaudio et al., 2011), and structures are the main disaster-pregnant environmental factor of landslides. Moreover, the number of publications in Engineering, Environmental Sciences Ecology and Water Resources is greater than 100, showing that these four disciplinary directions are also important research directions in this field. Statistics show that in the document data sample database, 113 papers belong to the same discipline direction,

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

accounting for 13.40%. The remaining 730 papers cover more than 2 discipline directions, which indicates that the research on early identification of potential landslide hazards covers multiple disciplines, and the continuous development of these interdisciplines has driven progress in this field.

In terms of the journals in which landslide papers are published, most of the documents were published in professional journals, but less in comprehensive journals. As shown in Fig. 5(b), the TOP 10 journals, as the core source of documents, have a total number of publications of 230, accounting for 27.28%. They are all professional journals in landslide prevention and control, but not comprehensive journals. For example, among the "LANDSLIDES", "REMOTE SENSING" and "ENGINEERING GEOLOGY", etc., "LANDSLIDES" ranked first, and published 61 papers on early landslide identification, accounting for 7.24 %, showing that subsequent researchers may choose to publish their research results in comprehensive journals.

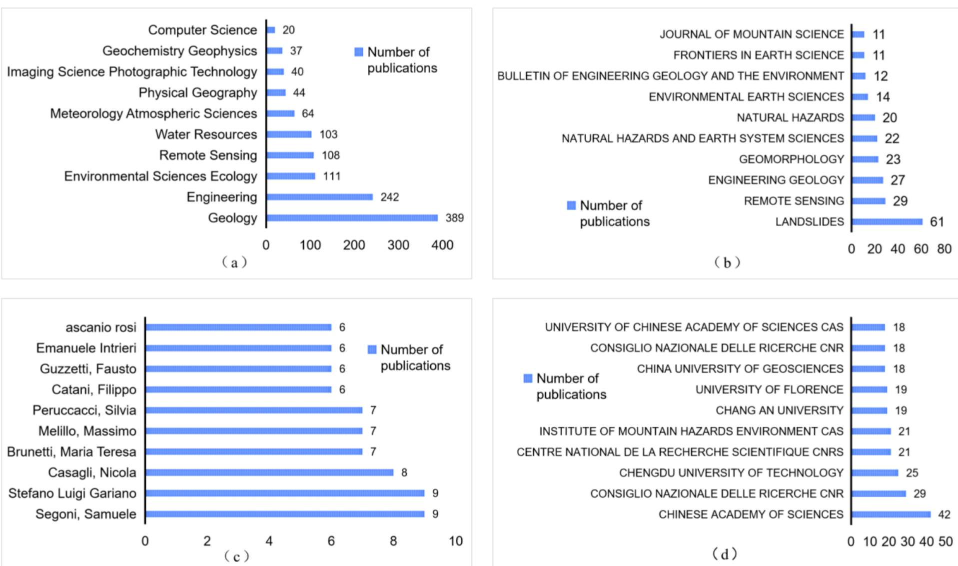


Fig. 5. Top 10 spatial distribution：a）subject distribution, b）periodical distribution, c）author distribution, d）institutional distribution.

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

As the author with the highest productivity in Fig. 5(c), Samuele Segoni has published a total of 9 papers in this field, one of which is a review paper entitled “A”. It is also a highly cited paper in the field of early identification of potential landslide hazards, which has been cited for 272 times. The more details are shown in Table 1. Samuele Segoni, one of the world-renowned leading peer experts in this field, is from the University of Florence. The two high productivity authors, namely Casagli, Nicola (8 papers) and Brunetti, Maria Teresa (6 papers) are both from the University of Florence, and they cooperated to published many papers on landslides. The three high productivity authors have accounted for 32.39 % of the total productivity of the TOP 10 authors. S, M and G are three high productivity authors from CNR and published a total of 22 papers in this field, accounting for 30.98 % of the total number of papers of the high productivity authors, which means the early identification of potential landslide hazards have been valued by the top experts and mainstream institutions.

By combining with the TOP 10 of the number of papers published by high productivity institutions in Fig. 5(d), it can be further found that the high productivity authors are often from high productivity institutions, such as U and C, showing that these institutions have achieved great success in attracting and cultivating high-level scholars’, and will help further enhance the academic strength and influence of these institutions in this field. On the contrary, the situation will be different. Although the high productivity institutions such as C1, C2, C3, I1, C4 and C5, etc. have published a large number of papers, but they lack of high productivity authors. It is not difficult to find that there are 6 high productivity institutions from China, accounting for 60% of the total number of high productivity institutions, but none of the high productivity authors are from China. And it is necessary for China with serious landslide disasters to strengthen the team building and improve the quality of academic research in this field.

Through the above analysis of authors and institutions with high productivity, the

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

spatial distribution of the number of publications in this field is further analyzed, the bibliometric tools V and S are adopted and the threshold of the number of publications is set to 10, so as to obtain the national clustering map shown in the upper right of Fig. 6. The circle node represents the country. The larger the node is, the greater the number of publications is. The connection between the nodes represents the cooperation between the countries. It can be found that five major national scientific research groups have been initially formed at present. For example, the cluster in orange is the largest national scientific research group, which consists of five countries, that is, China, USA, Canada, Switzerland and New Zealand. The total number of publications is 264, accounting for 31.32 % of the document sample database. Moreover, the cooperation relationship in this field is mainly reflected in the connection between national scientific research groups. The larger the national scientific research group is, the more frequent the cooperation with other groups is, and the lack of mutual cooperation between individual countries is not an individual phenomenon. For instance, Belgium, South Korea, Malaysia, etc. only cooperate with the countries in the cluster group, and the cooperation relationship is weak.

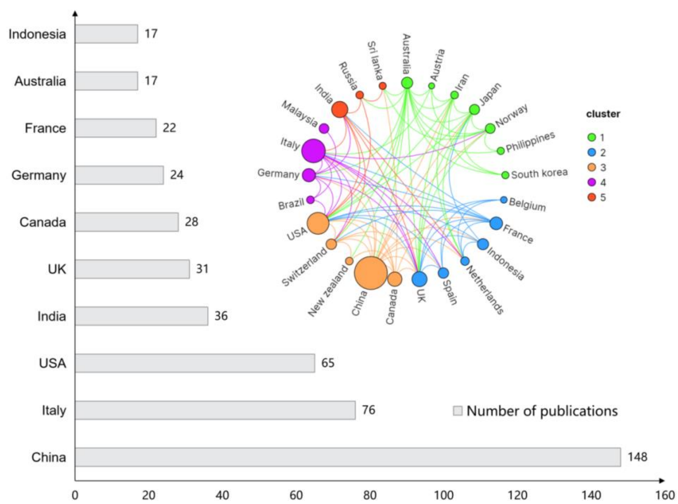


Fig. 6. Country or regional distribution.

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

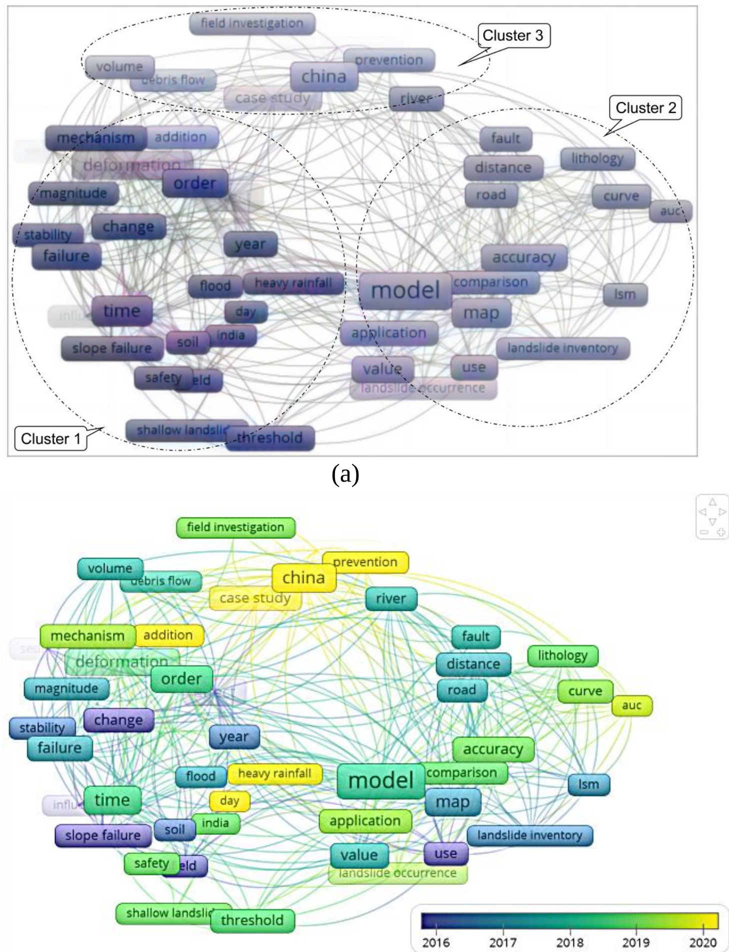
In order to facilitate comparative analysis, the clustering map of the number of papers published by countries is added to the map of the number of papers published by countries generated by the WPS Office through knowledge mapping superposition. Fig. 6 shows that high productivity countries often play a leading role in national research groups, such as China and USA in the orange group, Italy in the purple group, India in the red group, UK and France in the blue group and Australia in the green group. The total number of papers published by China ranked first, which is consistent with the number of papers published by high productivity institutions calculated in Fig. 5(d). The fundamental reason is that China is a country with frequent landslide disasters. As these landslides often occur in high-elevation and hidden places, most of the newly discovered landslides are not included in the known landslide database. The research of Petley (2012) shows that the less developed countries often have enormous numbers of deaths, and these countries have little investment in this field due to the lack of corresponding resources. Therefore, it is necessary to carry out cooperative research on early identification of potential landslide hazards and increase research efforts to help most countries in the world improve their ability and level of landslide prevention and control.

3 Landslide Formation Factors and Data Acquisition

In order to explore the research on landslide formation factors, further screening was carried out in the document sample database to refine the documents with the keywords of "factor*", and V was used to extract the keywords related to landslide formation factors. Keywords with an occurrence frequency of more than 5 times were selected to form a cloud map, as shown in Fig. 7. Each rectangular node represents a keyword. The larger the area of nodes is, the higher the occurrence frequency of keywords is. The connection line between nodes represents the correlation between the keywords. The shorter the connection line is, the closer the correlation between the keywords is.

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions



(a)

(b)

Fig. 7. Keyword cloud map of landslide influencing factors：a）cluster of keywords, b）timeline of keywords.

The collinear function of keywords can be used for cluster analysis to figure out the research hotspots and their relationships in this field, as shown in Fig. 7(a). The cluster in red contains the most keywords, mainly involving landslide disaster factors, such as disaster-causing factors (“heavy rainfall”), disaster-pregnant environmental factors (“soil”) and disaster-bearing object factors (“shallow landslide”). The cluster in green contains fewer

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

keywords than the cluster in red, which are mainly related to the early identification methods of potential landslide hazards, such as “model”, “auc”, “accuracy” and other high-frequency computational terms. The cluster in blue contains the fewest keywords, which are related to the early identification of landslides. For example, “China” represents the research area, and “field investigation” represents the monitoring methods. Generally, the keywords in Fig. 7 are important terms in landslide research, representing a hot topic in early landslide identification research.

Fig. 7(b) is another visual presentation mode of keyword cloud maps, which includes the information on the order of occurrence of keywords and also reflects the research progress in this field. Keywords such as “slope failure”, “change”, “use”, etc. occurred in 2016 or before, which was relatively early, while keywords including “model”, “order” and “threshold”, etc. roughly appeared in the same period, that is, around 2018. Seen from the total number of papers published in this field and the year in which the highly cited documents were generated, 2018 is a milestone year in this field, and the landslide research in the past few years after 2018 is worthy of special attention and in-depth analysis. A review of the past major landslide events suggests that the major landslides such as Xinmo (Intrieri et al. 2018), Baige (Wei et al. 2021), and Sedongpu (Chen et al. 2020), occurred from 2018 to 2019, and the huge landslide disasters have sparked a rapid increase in scholars’ attention to this field. Fig. 7(b) also indicates that keywords such as “China”, “prevention”, “heavy rainfall”, etc. appeared after 2020. Rainfall induced landslide is one of the most important and dangerous landslide types (Varnes et al. 1978). In recent years, with the frequent occurrence of rainstorm events, Chinese scholars have attached great importance to landslide prevention, and achieved continuous research achievements in this field.

Moreover, landslide data is vitally important for the analysis of landslide formation

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

reasons and even the early identification of potential landslide hazards. The website link of data acquisition has brought great convenience to researchers, but at present, only some links of data acquisition can be found sporadically in the previous documents. Therefore, this paper sorts out the link information of the main landslide data acquisition in this field, and the references of these data links are shown in Table 4. This paper first collates the integrated data (e.g., landslide inventory, InSAR data, Landsat Data, SRTM data, etc.), and then sorts them by the three elements of landslide formation, namely, disaster-causing factors, disaster-pregnant environment and disaster-bearing objects.

Table4 Data acquisition of potential factors for landslide hazards

Description	Website Name	Data acquisition link	References
Landslide inventory	The NASA Global Landslide Catalog	https://data.nasa.gov/Earth-Science/Global-Landslide-Catalog-Not-updated-/h9d8-neg4	Culler et al. (2021)
	The United States Geological Survey (USGS)	http://landslides.usgs.gov/regional/inventory	Kirschbaum et al. (2010)
InSAR data （Sentinel-1）	ESA's Copernicus Earth Observation datacenter	https://scihub.copernicus.eu/	Carlà et al. (2022)
		https://scihub.copernicus.eu/gnss/#/hom	Li et al. (2023)
Landsat Data	the US Geological Survey	https://landsat.usgs.gov/index.php	Dietz et al. (2013)
SRTM data	CGIARCSI SRTM 90 m Database	http://www.cgiar-csi.org/	
Precipitation Data Directory	The NASA Global Precipitation Measurement Mission	https://gpm.nasa.gov/data/directory	Culler et al. (2021)
	Rainfall Recording and Analysis, Department of Agriculture Maharashtra State	https://maharain.maharashtra.gov.in/	Patil and Panhalkar (2023)
	Earth System Research Laboratories	https://www.esrl.noaa.gov/	Ruiz-Villanueva et al., (2017)
	NASA's EarthData site	https://disc.gsfc.nasa.gov/datasets/NLDAS_FORA0125_H_002/summary?keywords=NLDAS/	Culler et al. (2021)
	Department of Hydrology and Meteorology （DHM）	http://www.dhm.gov.np/	Morin et al., (2018)
Seismic data	the SEISCOPE Consortium	https://seiscope2.osug.fr/	Lavoue et al. (2021)
	the AlpArray Seismic Network	http://www.alparray.ethz.ch/en/home/	
Human activity data	Resource and Environmental Science and Data Center of Chinese Academy of Sciences	https://www.resdc.cn/	Lin and Wang. (2018)
Topography data	SRTM 90 m DEM resolution	http://srtm.csi.cgiar.org	

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

	the OpenTopography website	https://portal.opentopography.org/datasets	Li et al. (2023)
	NASA official website	https://search.asf.alaska.edu/#/	Ab Rahman et al. (2018)
Geotechnical data	Global scale studies is provided in the PANGEA database	https://doi.pangaea.de/10.1594/PANGAEA.788537	Hartmann and Moosdorf (2012)
	the solid Earth and Ocean	https://www.pmel.noaa.gov/eoi/	Embley, RW et al. (2014)
	the System for Earth Sample Registration	https://www.geosamples.org/	
	EarthChem	www.earthchem.org	Hartmann and Moosdorf (2012)
	ORR & ASSOCIATES	www.orrbodies.com	
Environment Data	Resource and Environmental Science and Data Center of Chinese Academy of Sciences	https://www.resdc.cn/	Lin and Wang. (2018)
	the United Nations Environment Programme	https://www.unep.org/	Overland et al. (2008)
Map Catalog	Food and Agriculture Organizationof the United Nations	https://data.apps.fao.org/map/catalog/srv/eng/catalog.search#/home	FAO (2010).
Affected data	The Asian Disaster Reduction Center (ADRC)	https://www.adrc.asia/	Dietz et al. (2023)
	International Programme on Landslides	http://iplhq.org/	Kirschbaum et al. (2010)

4 Early Identification Methods

There are many kinds of methods for early identification of potential landslide hazards, which can be divided into qualitative and quantitative methods according to the type of landslide inventory. The data by qualitative analysis is generally text data, focusing on whether the landslide deformation has qualitative changes, and the experience-based identification method is qualitative analysis. The information by quantitative analysis is often numerical data, with an emphasis on the deformation of landslide body, as well as the methods based on physical-mechanical mechanism. Meanwhile, they can be divided into traditional identification model and intelligent automatic identification model according to the identification means. For example, the above-mentioned methods belong to traditional identification models, and intelligent automation mainly achieves the early identification of potential landslide hazards by virtue of artificial intelligence. The development process of these typical methods is described separately as follows.

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

4.1 Experience-based qualitative assessment method

Due to complicated mechanisms of landslide formation, experience-based qualitative assessment methods are mainly used to determine the value of evaluation index by applying expert experience and professional knowledge and combining with mathematical analysis and theoretical analysis (Federico et al., 2012; Newcomen et al., 2015). Researchers often need to carry out a large number of field investigations to analyze the disaster-pregnant environmental characteristics such as topography, geology and distance to river in the research area, and evaluate whether there are characteristic conditions for potential landslide hazards based on the existing landslide data and their own knowledge and abilities. This method requires researchers to have a high degree of experience accumulation and professional background knowledge requirements, but the identification results may vary from individual to individual. In addition, the factors affecting landslide deformation are diversified, so it is difficult to find accurate index values to judge their impacts on landslide deformation, causing an imbalance in the accuracy of landslide identification results. However, the experience-based qualitative assessment method has less requirements for basic data, and can be operated simply and efficiently, which is convenient for people to make some targeted achievements based on the actual situation of the research area, as shown in the typical cases in Table 5.

From Table 5, it is not difficult to find that it is problematic and challenging to implement the early identification method of potential landslide hazards by relying solely on researchers' experience and judgment. In order to improve the effectiveness and applicability of such identification, other methods (e.g., determination method based on physical-mechanical mechanism, statistical probability model method and artificial intelligence method, etc.) and models of multiple methods are often combined to reduce the uncertainty of landslide prediction to a certain extent.

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

Table 5 Typical cases of empirical qualitative identification method.

Study area	References	Question	Method	Limitations and Suggestions
the Lishanyuan landslide in southern China	Dongxin et al. (2022)	Relying on experience to select landslide velocity and acceleration threshold is a traditional approach of differential algorithm for landslide acceleration, which is prone to false positives	Combining conventional warning methods with critical sliding warning based on normalized tangent angle, and proposing a dynamic semi quantitative and semi empirical threshold determination method	The determination of the parameters of this method still depends on experience, and it is recommended to conduct future research on the internal physical mechanisms of landslides in conjunction
many rock slope in Australia	Tommaso et al. (2017)	The warning standards established based on empirical deduction of displacement or velocity thresholds are often too conservative	Adopting a new "review based" method to conduct statistical analysis on existing databases to determine the impending failure of rock slopes	The background information of the database is not comprehensive, lacks complete time series records, and is limited to slope velocity analysis for partial time. Suggest combining with similar databases for verification
the Hollin Hill Landslide Observatory in the UK	Whiteley et al. (2021)	The analysis results of the underground characteristic model of a single landslide obtained by Geoelectrical and seismic geophysical surveys are isolated, and the accuracy highly depends on experience and skills	Using an unsupervised machine learning form to classify geophysical data into cluster groups, aiming to compare borehole data and obtain slope scale ground models.	Single use of geophysical methods may have high uncertainty, and it is recommended to rely on prior or follow-up information to determine the root cause of anomalies
landslides in Papua New Guinea	Robbins (2018)	In areas with complex weather conditions, empirical identification models encounter difficulties because rainfall and landslides do not always have a causal relationship	Using TRMM Multi Satellite Analysis Product (TMPA) and Landslide Inventory to Solve the Critical Rainfall Problem of Landslides in Areas with Sparse Data	At present, the dataset can only provide a small landslide event feature for evaluation. It is recommended to identify rainfall event standards for certain triggering events

4.2 Physical methods

This method is to digitize the physical properties and external factors of the landslide itself, take the classical slope stability theory as the evaluation basis and adopt the methods of numerical simulation, indoor experiment and field investigation to accurately analyze the information of the research area (Mebrahtu et al., 2022). Although this method has the

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

advantages of rapidity and accurate in early identification of single potential landslide hazard, it requires a large number of detailed and reliable data concerning terrain, geology, soil characteristics, which limits its use in a relatively large area. In recent decades, with the rapid development of remote sensing technology and large-scale finite element software, the stability analysis of landslide slopes with complex boundary and interface conditions has been realized (Hammah et al., 2007). Scholars have put forward appropriate research plans according to different types of research objects, and made a lot of contributions in the practice process, as shown in Table 6.

Table 6 Practical Cases

References	Research object	Method	Main Contribution
Tanmoy et al. (2022)	Rainfall induced landslides	Numerical analysis of landslide mechanism using fluid solid coupling and comparison with on-site observation results	The necessity of conducting coupled flow deformation analysis based on conventional saturation stability analysis was determined, and the maximum cumulative rainfall in the study area was determined
Sunbul et al. (2018)	Earthquake-induced landslides	Using microstructure and finite element modeling of geotechnical engineering to determine the extent to which soil structure and elements contribute to the mechanism of landslides	Proposed a systematic evaluation method for earthquake induced landslide disasters and conducted long-term monitoring work to provide data for future research
Youssef et al. (2015)	Landslides triggered by human activities	Conduct on-site investigations to determine geological units and different structures along the road, use rock mass rating programs to determine rock mass characteristics, and identify key areas that affect instability	The impact of road cutting on landslides was studied, and the orientation of structural planes and rock mass strength parameters have a significant impact on the overall stability of protected objects
Du et al. (2021)	Landslide mass after engineering treatment	The response surface method and finite element method were used to identify the main controlling factors and deformation characteristics of topic deformation after engineering treatment, and the transformation mechanism of landslide failure mode was explored	Previous studies have mostly focused on the failure mechanism of landslides, and this article reveals the control factors and failure modes of landslides based on failure examples
Xiao et al. (2018)	reservoir landslide	Using the geometric shape and wave distribution of landslides as constraints, apply the Tsunami Squares method to simulate landslides and shock waves	The numerical results obtained by Tsunami Squares method can quickly view and simulate potential landslides

4.3 Method based on statistical probability model

This method analyzes the existing landslides as the statistical samples, obtains landslide information via GIS spatial analysis technology and mathematical model, and establishes the relationship between landslide influencing factors and inducing factors (rainfall, earthquake, human activities, etc.). Meanwhile, this method calculates the quantitative value of the disaster-prone area as the threshold value, compares with the evaluation units of the research area and speculates on the possible area and scale of landslide disasters based on the historical data of the disaster. The method based on the statistical probability model does not require accurate parameters, and the statistical method does not have obvious physical process of landslide deformation. The probability method can consider the physical process, and the probability distribution of parameter values can be input to quantify the reliability of the identification results (Salvatici et al., 2018). The movement pattern of landslides is significantly controlled by internal inhomogeneity and discontinuity (Zhang et al., 2018). Therefore, compared with the deterministic method based on physical-mechanical mechanism, this method is more suitable for early identification of regional potential landslide hazards. In recent decades, the development of probability models and statistical methods has greatly improved the performance of evaluation models used to measure the occurrence probability and deformation characteristics of potential landslide hazards (Corominas et al. 2014). Table 7 shows several common methods for calculating probability, which have their own advantages and disadvantages. The appropriate methods shall be selected according to the specific conditions, and attention shall be paid to the definition, nature and application of probability to avoid wrong calculation results.

Table 7 Highly cited literature information table（TOP 5）

Typical methods	**Advantage**	**disadvantage**	**References**	**summarize experience**
Fuzzy Mathematics	Accurate mathematical language can be used to	The design of the plan lacks	Xie et al. (2013)	Adopting the principle of two-level fuzzy

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Method	describe the characteristics of landslide deformation, which is suitable for dealing with the uncertainty of landslide recognition	systematicity, and the identification is subjective, and the accuracy depends on the amount of landslide data		comprehensive evaluation, the landslide influencing factors are graded from high to low, resulting in more accurate evaluation results
Monte Carlo method	Suitable for problems that are difficult or even impossible to solve using analytical methods, recognition errors are independent of dimensionality, and there is no need to discretize continuity problems	It is necessary to convert certainty into randomness, and the calculation steps are quite cumbersome	Liu et al. (2022)	The diffusion angle is used to describe the propagation process of landslides. The larger the diffusion angle, the more iterations MC simulation takes
Bayesian method	Statistical methods for determining whether two population means are equal	Require a large amount of repeated observation data	Robbins (2016)	When data is available, the Bayesian method can be easily updated by changing the posterior probability
Weight of Evidence	It can improve the predictive performance of the model, improve class comprehension, and directly compare the values within the independent variable	Only considering the recognition ability of landslide hazards, the monotonic variation of eigenvalues	Kumar and anbalagan 2019	Applied to the relationship between influencing factors and landslide formation, based on positive and negative correlation values, important factors leading to landslides can be identified
logistic regression	Suitable for classification scenarios, easy to calculate, without assuming data distribution in advance	Low classification accuracy, underfitting, and missing data features	Budimir et al. (2015)	Identified common covariates and their frequency of inclusion, provided a list of covariates

4.4 Intelligent automatic identification method of the model

Taking the document sample database as the basis and the “artificial intelligence”, “machine learning” and “deep learning” as search terms, the documents concerning artificial intelligence, machine learning and deep learning in this field are further extracted. The occurrence frequency of these keywords in the documents is displayed in the form of a word cloud to realize the visualized analysis of intelligent automatic identification model in the research field of this paper, as shown in Fig. 8. The outer contour shape of the word cloud represents the first letters “AI”, “ML” and “DL” of artificial intelligence, machine learning and deep learning. The higher the frequency of the keyword is, the greater the occupied area is, and the more obvious it shows on the map.

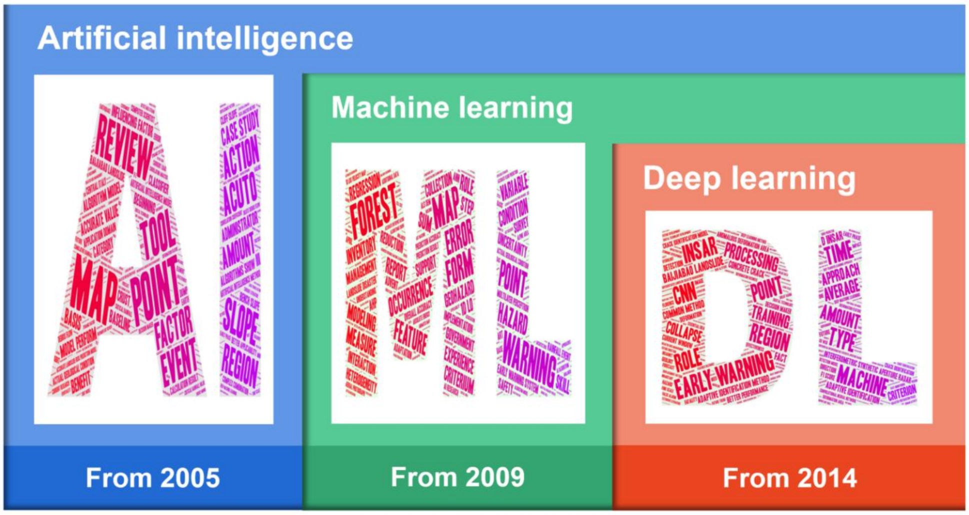


Fig. 8. Intelligent automatic recognition model word cloud map.

From Fig. 8, it can be seen that the AI related keywords are dispersedly distributed, indicating that the research on documents directly related to AI in this field is not sufficient. Although AI made a breakthrough in the initial development stage, which turned out to be higher than expected, it slowly entered in the trough of the development due to the lack of computing power and theory. AI did not start a new round of application and development until its ML began to appear in this field. In Fig. 8, ML displays the densest and non-repetitive keywords, and the reference parameters are significantly enriched, showing that ML has promoted the vigorous development of this field. The ML method is often a good supplement to traditional empirical and statistical methods, and it is also a main quantitative analysis method with high accuracy. However, at least 80% of ML is data preprocessing work (Ma et al., 2021), and the performance of the ML method depends on data quality. The landslide data is often hard to be obtained in real time, which greatly affects the identification accuracy of the ML model (Cao et al., 2022). For example, noise and outliers have a significant impact on the accuracy of pixel-based methods, so a significant number of parameters need to be adjusted (Sameen and pradhan et al. 2019). As a new research direction

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

of ML, DL can effectively process big data and simplify data preprocessing steps, and its emergence has inspired the construction of intelligent automatic identification models. In recent years, people have become increasingly interested in methods of intelligent automatic identification models, and frequently published the research documents on early identification of landslides by ML and its derivative method (such as DL). For early identification of potential landslide hazards, it is first necessary to build a training sample database for landslide distribution, which can be divided into positive samples (landslide point) and negative samples (non-landslide point). Then, the feature extraction automation is realized in the DL-based training process, the input data is directly sent to the ANN, and the specific features of these data are learned hierarchically in each network. Finally, the specific features are associated with labels, categories, and decisions（Tehrani et al., 2022）. Lu et al. (2023) designed a semantic segmentation model with dual encoder architecture and feature fusion functions, which can represent the in-depth features of optical bands and DEM data by levels to predict landslides. There are many factors affecting the formation of landslides, such as slope, elevation, formation lithology, etc., which means that there are many parameters available for model selection, and there is a certain correlation between the parameters. At present, a unified model parameter calibration system has not yet been formed, and targeted choices need to be made based on the actual situation and application scenarios of the research area. The principles of parameter selection are closely related to the formation mechanism of landslides or the environmental characteristics of danger-hidden areas, all of which are practicable (*Wang* et al., 2020). The performance of the ML model method is usually evaluated according to the existing landslide inventory, and various indicators are determined based on the area under the ROC curve and confusion matrix. In recent studies, Sevgen et al. (2019) has evaluated the model performance by introducing photogrammetry database before and after landslide events. In general, different documents have significant

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

differences in data samples, parameter calibration, model robustness and generalization, so it is very important to select appropriate ML model methods for early identification research of landslides.

4.5 Fusion method of the model

The fusion model that combines two or more models is used for early identification of potential landslide hazards through selecting landslide samples, screening landslide deformation characteristics, extracting disaster-causing factors and other processes. It synthesizes the advantages of all models, and can effectively improve the accuracy of accelerated deformation identification of landslides. Bibliometric analysis shows that 44 documents on fusion models have been applied in this field, most of which are concentrated in the past five years. The main forms of the fusion model include coupling model of multiple physical methods, coupling model of statistical method-deep learning and coupling model of deterministic coefficient-DNN, etc. The specific promotion and application examples are shown in Table 8.

Table 8 Integration model promotion and application cases

References	Integration method	Function	Research objective	Main results
Yue et al. (2018)	the Newmark model	Permanent displacement used to identify potential landslide areas	Determine the impact of landslides caused by earthquakes on roads	Used to analyze the probability of earthquake landslides, simulate the movement trajectory of dangerous rock masses, and identify the danger of landslides along roads and other routes. The simulation results are highly consistent with known results
	the RockFall Analyst model	Simulate the possible impact of mass movement on roads		
Do Minh et al. (2022)	statistical models	the five Landslide Susceptibility Index (LSI) mapswere established	Establishing a large-scale regional landslide sensitivity zoning map	The output of the fusion model has good predictive ability and is not significantly different from published research
	machine learning models	Creating Landslide Sensitivity Maps		

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Ma et al. (2023)	the coefficient of determination method (CF)	Determine the weights of various landslide condition factors to analyze the correlation between each factor and landslide occurrence	Established a regional landslide spatial database and analyzed the main influencing factors of landslides	Using the confusion matrix to evaluate the accuracy of three methods, it was found that CF-DNN is more suitable for evaluating the landslide susceptibility in the region
	deep neural network (DNN)	Mining deep features of samples to provide accuracy in magnetic susceptibility models		
Lin et al. (2023)	convolutional neural network (CNN) model	Extract local features of data	Solving the defect of neglecting random displacement in traditional landslide displacement models	Constructing a CNN BiLSTM fusion model, extracting the spatiotemporal features of landslide displacement prediction data, and testing results show that it can be promoted in landslide prone areas
	a bidirectional long short-term memory network (BiLSTM) model	Ability to process time series data		

5 Discussion and research trends

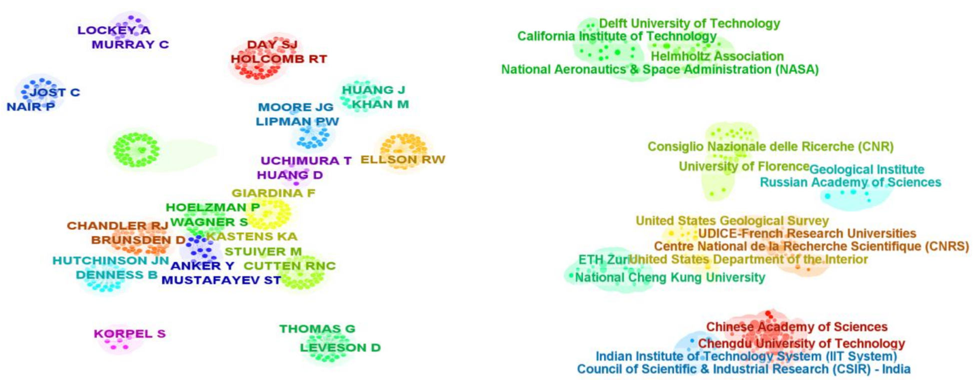
Through the analysis of previous chapters, it is found that the field has the characteristics of not close cooperation, large amount of data, and diversified early identification methods. In this section, the four aspects including cooperation analysis, multi-source data, topic analysis, and research trends are further analyzed and discussed by combining with the document sample database and relying on CiteSpace software.

5.1. Cooperation analysis: discrete nodes and unformed core groups

The data of the sample database can clearly indicate that there are few cases in which a paper is completed by an independent author, and most of the papers are co-signed by multiple authors and institutions. The number of authors and institutions in a paper is gradually increasing over time, indicating that the awareness of cooperation among scholars in this field is gradually increasing. The collaborative relationships between interrelated authors and institutions can be explored based on knowledge graph technology (*Yu* et al., 2021a). The further analysis requires to import the data in the document sample database into CiteSpace software in a “download_*.txt” format, set the time span to 1971-2023 and select

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

the nodes as "Author" and "Institution", so as to obtain visualized maps of the cooperation network of authors and institutions, as shown in Fig. 9. Each node represents an author or institution, and a color block represents a group. The nearer the distance between color blocks is, the closer the cooperation between the research groups is. On the whole, both authors and institutions cooperate in a form of small-scale groups. The location distribution of the map spot is relatively discrete, and no obvious core group has been formed. This shows that the research in this field is relatively independent, but the foreign academic cooperation is weak. It is suggested that researchers should break through barriers of the research group, seek more extensive cooperation and get more ideas on early identification of potential landslide hazards from different perspectives according to their research interests.



(a) (b)

Fig. 9. Network of cooperation：a）author cooperation, b）institutional cooperation.

5.2. Multi-source data: inherent defects and insufficient fusion

The multi-source data plays an indispensable role in the research on early identification of potential landslide hazards (Juang et al., 2019; Nappo et al., 2019). In Section 3, by sorting out the data to obtain links, it is found that there is a wide range of data sources suitable for early identification of potential landslide hazards. However, these multi-source data often have inherent defects in their quality, such as strong periodic fluctuation, multiple outliers, and different sampling frequencies (Qian et al., 2023), which seriously restricts the

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

application of landslide identification methods and the improvement of identification accuracy (Chen et al., 2021). This section extracts documents related to data from the document sample database for visualized analysis, and classifies the similar landslide data documents automatically by LLR (logarithmic likelihood test) clustering analysis algorithm to get a total of 43 clusters. The cluster name is determined by the keywords related to landslide data. To identify the internal structure of clusters and the connections between multi-source data more clearly, the first 17 cluster information was extracted as the main analysis content, as shown in Fig. 10, from which it can be seen that the distribution of each cluster of the multi-source data is discrete, and the connection between the data is not obvious. Different identification methods often have great differences in the selection of data samples and influencing factors (Achour et al., 2019; Kalantar et al., 2018; Lu et al., 2020). Optimizing data quality and integrating multi-source data are the key to improving model robustness and application. For example, spatial analysis and average regional statistics are made to form the attribute fusion data sets of landslide disaster deformation concentration areas to improve the applicability of data (Zheng et al., 2021). Alternatively, the discretized multi-source data can be quantified and grouped, and external factors influencing data and landslide displacement response data can be integrated to identify potential landslide hazards through trend sequence models and sensitive state models (Liu et al., 2020).

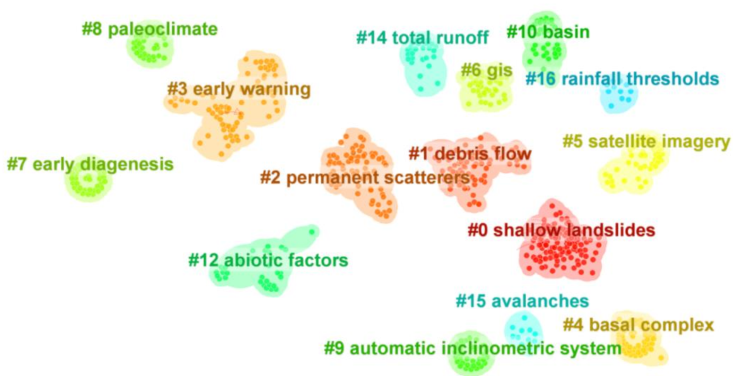


Fig. 10. Knowledge graph of multiple source data.

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

5.3. Topic analysis: complex contents and diversified clusters

The collinear network knowledge map of all keywords in the data sample database can be obtained by running CiteSpace software and selecting collinear analysis functions of keywords. A total of 1071 nodes has been formed in the map, and there are 5,777 connection lines, with a density of 0.0101. The clustering information module value Modlularity Q=0.6064 (＞0.3), indicating that the clustering community structure is significant. The map contour coefficient Weighted Mean Silhouette S=0.8365 (＞0.5), showing that the overall homogeneity of this cluster is high and credible. The "keywords" is selected as the cluster analysis source, and the LLR algorithm is used to obtain a total of 46 clusters. The top 11 cluster maps are selected for visualized analysis according to relevance, as shown in Fig. 11. Each cluster is composed of several closely-related keywords. The larger the cluster number is, the more keywords the cluster contains. As the largest cluster, "synthetic aperture radar" is located at the center of the cluster, suggesting that this topic deserves attention.

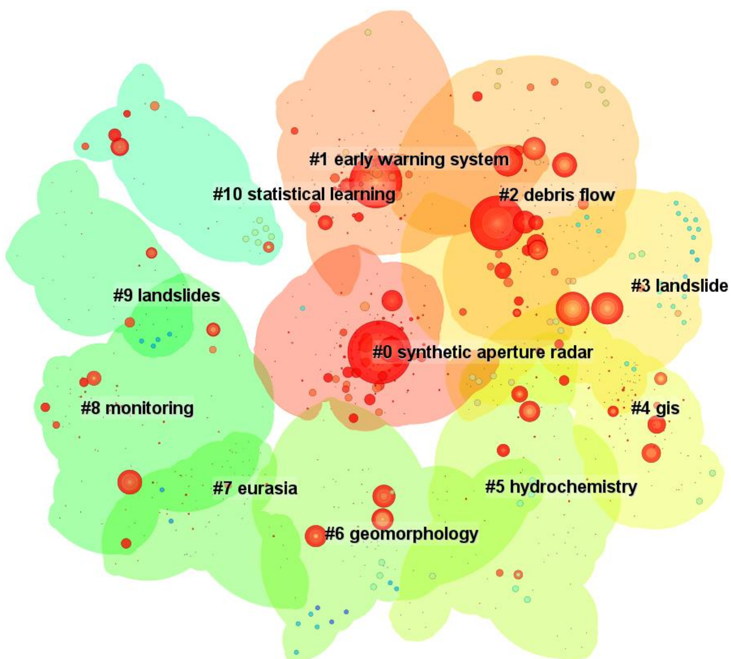


Fig. 11. Keyword clustering analysis graph.

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

5.4. Research trend: simplified-to-diversified trend

The mutation detection was conducted for keywords occurred from 1971 to 2023 to reflect the research process and development trend in this field. The "Burstterms" function in CiteSpace software was run to get a total of 25 keywords with the maximum burst strength in this field, as shown in Fig. 12.

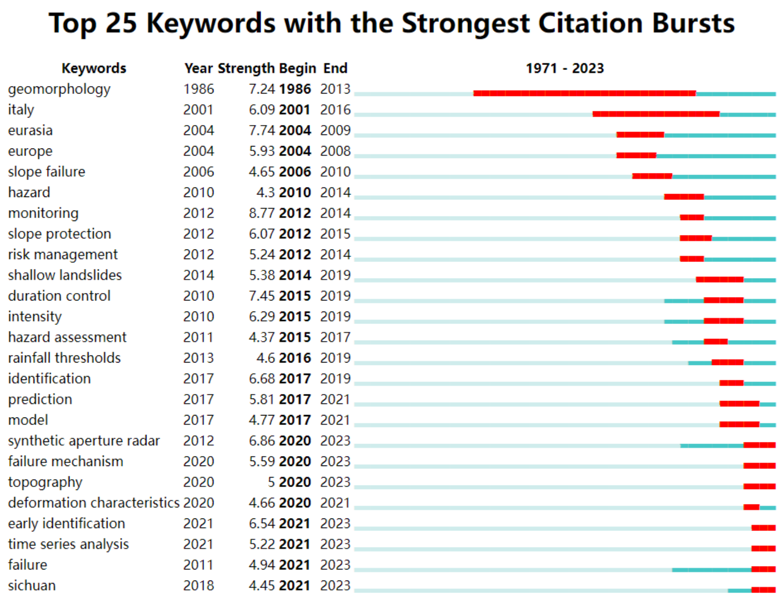


Fig. 12. Keywords with citation bursts.

The "Strength" column represents the burst strength of the keyword, and can also be understood as the quantized value of the influence of the keyword in this field. The blue segment represents the timeline, while the red segment represents the duration of occurrence of the keyword. As one of the earliest keywords, "geomorphology" has the longest duration of 27 years, lasting from 1983 to 2013, which indicates that this field focuses on the factors of landslide formation from the perspective of disaster-pregnant environment. The keyword "monitoring" has the maximum burst strength of 8.77, meaning that the reference of documents related to landslide monitoring may help find out innovative academic viewpoints,

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

and the landslide monitoring is also a hot topic for scholars. The latest 6 keywords such as "synthetic aperture radar", "failure mechanism", "topography", "time series analysis failure" and "Sichuan" show the research trends in this field, which will be the focus of future research in this field. It is also suggested that the future research in this field will include more disciplinary contents, and a single method is difficult to solve complex problems such as nonlinear correlation analysis of influencing factors, multi-source heterogeneous data fusion, massive landslide information processing, etc. The transition from simplification to diversification is an inevitable trend in the development of this field, and scholars may integrate multiple identification methods to carry out related research.

Fig. 13 shows a timeline view of co cited references, describing the development and evolution of the cited references over time in each cluster. The nodes represent different collinear keywords, the link lines between nodes represent collaborative relationships, and their thickness represents the strength of the links *(Yu et al., 2017)*. The duration of 0#, "synthetic aperture radar (SAR)" is relatively long, and the distribution is the most dense node, indicating that SAR is the focus of current research and will continue to be active in the future. Furthermore, it is worth noting that 8#, "*monitoring*" is one of the latest clusters, indicating that landslide monitoring is an emerging research topic in recent years.

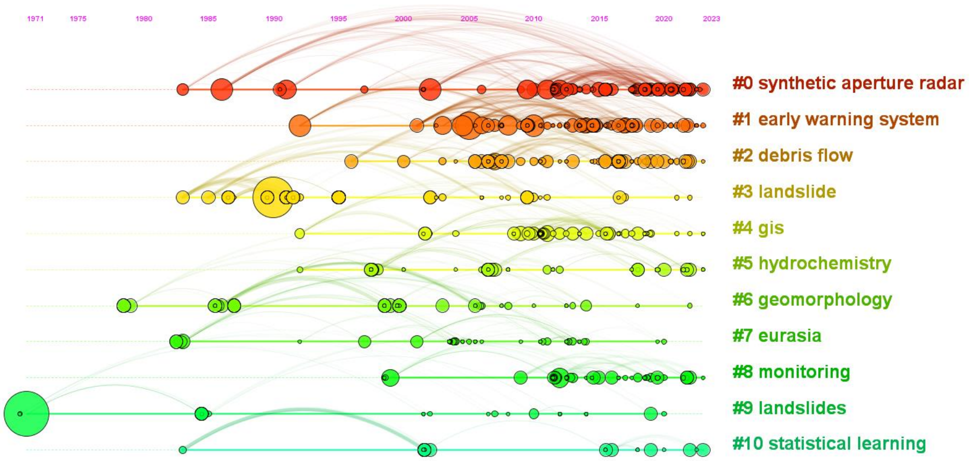


Fig. 13. Timeline view for document co-citation clusters.

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Although this article identifies future development trends *(Yu et al., 2021b)*, research in this field still faces many challenges, such as uncertainty in data sources, diversity of influencing factors, and immature identification methods. Therefore, it is necessary to continuously update the literature database and keep track of the field *(Yu et al., 2022)*.

6 Conclusions

The research on early identification of potential landslide hazards is a very important and complex problem, facing with great challenges now and in the future. In the past 30 years, many researchers have made continuous contributions and formed a rich document basis in this field. The conclusion drawn from this article is as follows:

(1) The number of documents published: Three stages (1971-1995 is the embryonic period, 1996-2008 is the slow development period and 2009-2023 is the rapid development period) can be roughly divided, which shows an exponential growth trend over time on the whole. The most influential journal is “LANDSLIDES”. “Samuele Segoni” is not only an author with the highest productivity, but also a highly cited author. As the institution with the highest number of publications, “CHINESE ACADEMY OF SCIENCES” is ranked as a high productivity institution together with other 5 institutions from China.

(2) Factors and data: There are many factors affecting the formation of landslides. The “heavy rainfall”, “soil”, “shallow landslide” and other keywords are important factors. 2018 is a turning point in this research field, and more attention will be paid to the research of heavy rainfall in the future. The data acquisition links related to the early identification of landslides are sorted out. At present, these websites can be accessed. If you want to know more information during use, you can view it by referring to the documents mentioned in this paper.

(3) Early identification methods: Through exemplifying 5 main methods and making document analysis, it is found that the empirical qualitative method mainly depends on the experience and professional knowledge of researchers, so it is hard to guarantee the accuracy.

Although the physical determination method is very accurate, it can only be used on a small scale. The statistical probability method has obvious advantages in solving the uncertainty problem, but it has a large amount of computation and low identification efficiency. The intelligent model can address complex problems that are difficult to be solved by traditional methods, but it is difficult to be popularized due to its cumbersome modeling process. The fusion model synthesizes the advantages of the above four methods and will become the most promising method in the field of early identification of potential landslide hazards.

(4) Others: Most of the authors and institutions only cooperate in the form of research groups, and they lack foreign exchanges and cooperation. The multi-source landslide data has inherent defects and is difficult to be fused, therefore scholars use new methods like machine learning to improve data quality and meet scientific research needs through quantification, classification, and database establishment. The cluster analysis shows that the number of clusters in this field is as high as 46, and the largest cluster is “synthetic aperture radar”, indicating that landslide monitoring has always been the key topic for scholars.

(5) Trends: From the keyword mutation analysis, it is understood that the research focus in the future involves multiple disciplines, and the transition of corresponding research methods and modes from simplification to diversification is the future trend. This article attempts for the first time to use a timeline view to track the evolutionary patterns in this field, identify future research trends, and provide a dynamic analysis perspective as a reference for early landslide identification research.

In general, this paper systematically reviews the research process and hotspots in this field from the perspective of bibliometrics. These panoramic results will help researchers quickly understand the field, save time and reduce research costs. In the future, the real-time update of the document database can lay a foundation for further tracking the research frontiers in this field.

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

Data availability. No data sets were used in this article.

Author contributions. We use the *CRediT* Contributor Roles Taxonomy to categorise author contributions. Methodology: HL; Writing – original draft preparation: ZY; Data curation: ZZ; Funding acquisition: CL; Investigation: CC, GF; Supervision: RN, WZ; Validation: KS; Writing – review & editing: HL, MZ. All authors have read and agreed to the published version of the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Acknowledgments. This research was supported by the National Key R&D Program of China (2019YFC1510700), and the Sichuan Science and Technology Program (2023YFS0380, 2023YFS0377, 2023NSFSC1989, 2022YFS0539).

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

References

Ab Rahman, A.A., Abd Majid, N., Ahli, N.A., Ab Latip, A.S., Taib, A.M., 2023. The capability of SNAP software application to identify landslide using InSAR technique. Physics and Chemistry of the Earth, Parts A/B/C, 103427. https://doi.org/10.1016/j.pce.2023.103427.

Achour, Y., Pourghasemi, H.R., 2020. How do machine learning techniques help in increasing accuracy of landslide susceptibility maps? Geoscience Frontiers, 11, 3, 871-883. https://doi.org/10.1016/j.gsf.2019.10.001.

Bagwari, S., Gehlot, A., Singh, R., Priyadarshi, N., Khan, B., 2021. Low-cost sensor-based and LoRaWAN opportunities for landslide monitoring systems on IoT platform: a review. IEEE Access, 10, 7107-7127. https://doi.org/10.1109/ACCESS.2021.3137841.

Bai, D., Lu, G., Zhu, Z., Zhu, X., Tao, C., Fang, J., 2022. A hybrid early warning method for the landslide acceleration process based on automated monitoring data. Applied Sciences, 12(13), 6478. https://doi.org/10.3390/app12136478.

Barnard, P.L., Owen, L.A., Sharma, M.C., Finkel, R.C., 2001. Natural and human-induced landsliding in the Garhwal Himalaya of northern India. Geomorphology, 40, 1-2, 21-35. https://doi.org/10.1016/S0169-555X(01)00035-6.

Budimir, M.E.A., Atkinson, P.M., Lewis, H.G., 2015. A systematic review of landslide probability mapping using logistic regression. Landslides, 12, 419-436. https://doi.org/10.1007/s10346-014-0550-5.

Bui, D.T., Hoang, N.D., Nguyen, H., Tran, X.L. (2019). Spatial prediction of shallow landslide using Bat algorithm optimized machine learning approach: A case study in Lang Son Province, Vietnam. Advanced Engineering Informatics, 42, 100978. https://doi.org/10.1016/j.aei.2019.100978.

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Camerlenghi, A., Accettella, D., Costa, S., Lastras, G., Acosta, J., Canals, M., Wardell, N., 2009. Morphogenesis of the SW Balearic continental slope and adjacent abyssal plain, Western Mediterranean Sea. International Journal of Earth Sciences, 98, 735-750. https://doi.org/10.1007/s00531-008-0354-8.

Cao, C., Zhu, K., Xu, P., Shan, B., Yang, G., Song, S., 2022. Refined landslide susceptibility analysis based on InSAR technology and UAV multi-source data. Journal of Cleaner Production, 368, 133146. https://doi.org/10.1016/j.jclepro.2022.133146.

Carlà, T., Intrieri, E., Farina, P., Casagli, N., 2017. A new approach to assess the stability of rock slopes and identify impending failure conditions. In Advancing Culture of Living with Landslides: Volume 2 Advances in Landslide Science, pp. 733-739. Springer International Publishing. https://doi.org/10.1007/978-3-319-53498-5_84

Carlà, T., Intrieri, E., Raspini, F., Bardi, F., Farina, P., Ferretti, A., Colombo, D., Novali, F., Casagli, T., 2019. Perspectives on the prediction of catastrophic slope failures from satellite InSAR. Sci Rep 9, 14137. https://doi.org/10.1038/s41598-019-50792-y.

Chang, M., Sun, W., Xu, H., Tang, L., 2023. Identification and deformation analysis of potential landslides after the Jiuzhaigou earthquake by SBAS-InSAR. Environmental Science and Pollution Research, 30, 13, 39093-39106. https://doi.org/10.1007/s11356-022-25055-5.

Chen C., 2017. Science mapping: a systematic review of the literature. Journal of Data and Information Science, 2,. 1-40, https://doi.org/10.1515/jdis-2017-0006.

Chen, C., Zhang, L., Xiao, T., He, J., 2020. Barrier lake bursting and flood routing in the Yarlung Tsangpo Grand Canyon in October 2018. Journal of Hydrology, 583, 124603. https://doi.org/10.1016/j.jhydrol.2020.124603.

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Chen, Y., Ming, D., Ling, X., Lv, X., Zhou, C. 2021. Landslide susceptibility mapping using feature fusion-based CPCNN-ML in Lantau Island, Hong Kong. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 3625-3639. https://doi.org/10.1109/JSTARS.2021.3066378.

Chen, Z., Zhang, J.M., Ho, K., Wu, F.Q., 2008. Prediction of the spatiotemporal distribution of landslide: Integrated landslide susceptibility zoning techniques and real-time satellite rainfall. In Landslides and Engineered Slopes. From the Past to the Future, Two Volumes+ CD-ROM (pp. 2035-2038). CRC Press.

Chou, F.S., Lin, W.C., Chen, Y.H., Liao, C.K., 2015. Monitoring the vegetation dynamics of early succession following a landslide on Shanping forest road. Taiwan J For Sci, 30, 4, 217-28.

Corominas, J., van Westen, C., Frattini, P., Cascini, L., Malet, J.P., Fotopoulou, S., Catani, F., Van Den Eeckhaut, M., Mavrouli, O., Agliardi, F., Pitilakis, K., Winter, M.G., Pastor, M., Ferlisi, S., Tofani, V., Hervás, J., Smith, J.T., 2014. Recommendations for the quantitative analysis of landslide risk, B. Eng. Geol. Environ., 73, 209 – 263, https://doi.org/10.1007/s10064-013-0538-8.

Culler, E.S., Badger, A.M., Minear, J.T., Tiampo, K.F., Zeigler, S.D., Livneh, B., 2021. A multi‐sensor evaluation of precipitation uncertainty for landslide‐triggering storm events. Hydrological processes, 35, 7, e14260. https://doi.org/10.1002/hyp.14260.

Das, T., Rao, V.D., Choudhury, D., 2022. Numerical investigation of the stability of landslide-affected slopes in Kerala, India, under extreme rainfall event. Natural Hazards, 114, 1, 751-785. https://doi.org/10.1007/s11069-022-05411-x.

Del Gaudio, V., Wasowski, J., 2011. Advances and problems in understanding the seismic response of potentially unstable slopes. Engineering geology, 122, 1-2, 73-83.

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

https://doi.org/10.1016/j.enggeo.2010.09.007.

Devoli, G., Strauch, W., Chávez, G., Høeg, K. 2007. A landslide database for Nicaragua: a tool for landslide-hazard management. Landslides, 4, 163-176. https://doi.org/10.1007/s10346-006-0074-8.

Dietz, A.J., Kuenzer, C., Conrad, C., 2013. Snow-cover variability in central Asia between 2000 and 2011 derived from improved MODIS daily snow-cover products. International journal of remote sensing, 34, 11, 3879-3902. https://doi.org/10.1080/01431161.2013.767480.

Do Minh, H., Nguyen, V.H., Mai, L.D., Luong, H.D., Ngo, T.T., Van Thi, H., 2022. Large-scale Mapping of Landslide and Debris Flow using Flowr Model with Statistical and Machine Learning Methods. VNU Journal of Science: Earth and Environmental Sciences, 38, 4. https://doi.org/10.25073/2588-1094/vnuees.4872.

Du, Y., Yan, E., Gao, X., Mwizerwa, S., Yuan, L., Zhao, S., 2021. Identification of the main control factors and failure modes for the failure of Baiyuzui landslide control project. Geotechnical and Geological Engineering, 39, 3499-3516. https://doi.org/10.1007/s10706-021-01707-0.

Embley, R.W., Merle, S.G., Baker, E.T., Rubin, K.H., Lupton, J.E., Resing, J.A., Dziak, R.P., Lilley, Ma.D., Chadwick Jr. W.W., Shank, T., Ron Greene, Walker, S.L., Haxel, J., Olson, E., Baumberger, T., 2014. Eruptive modes and hiatus of volcanism at West Mata seamount, NE Lau basin: 1996–2012. Geochemistry, Geophysics, Geosystems, 15, 10, 4093-4115. https://doi.org/10.1002/2014GC005387.

FAO. 2010. Hydrological basins in Near East (Derived from HydroSHEDS) Version 1, available online from the FAO GeoNetwork Database. Accessed February 4, 2013. http://www.fao.org/geonetwork/srv/en/metadata.show?id=37299.

Federico, A., Popescu, M., Elia, G., Fidelibus, C., Internò, G., Murianni, A., 2012. Pr ediction of time to slope failure: a general framework. Environmental Earth Sciences, 66, 245-256. https://doi.org/10.1007/s12665-011-1231-5.

Fritsche, S., Fäh, D., Gisler, M., Giardini, D., 2006. Reconstructing the damage field of the 1855 earthquake in Switzerland: historical investigations on a well-documented event. Geophysical Journal International, 166, 2, 719-731. https://doi.org/10.1111/j.1365-246X.2006.02994.x.

Gorum, T., Fan, X., van Westen, C.J., Huang, R.Q., Xu, Q., Tang, C., Wang, G., 2011. Distribution pattern of earthquake-induced landslides triggered by the 12 May 2008 Wenchuan earthquake. Geomorphology, 133, 3-4, 152-167. https://doi.org/10.1016/j.geomorph.2010.12.030.

Hammah, R.E., Yacoub, T., Corkum, B., Wibowo, F., Curran, J.H., 2007. Analysis of blocky rock slopes with finite element shear strength reduction analysis. Proc 1st Canada-US Rock Mech Symp - Rock Mech Meet Soc Challenges Demands 1:329 – 334. https:// doi. org/ 0. 1201/ noe04 15444 019-c40.

Harp, E.L., Keefer, D.K., Sato, H.P., Yagi, H., 2011. Landslide inventories: the essential part of seismic landslide hazard analyses. Engineering Geology, 122, 1-2, 9-21. https://doi.org/10.1016/j.enggeo.2010.06.013.

Hartmann, J., Moosdorf, N., 2012. The new global lithological map database GLiM: a representation of rock properties at the earth surface. Geochem Geophys Geosyst 13:1 – 37. https://doi.org/10.1029/2012GC004370.

Hong, Y., Adler, R., Huffman, G., 2006. Evaluation of the potential of NASA multi – satellite precipitation analysis in global landslide hazard assessment. Geophysical Research Letters, 33, 22. https://doi.org/10.1029/2006GL028010.

Imani P, El-Raouf AA, Tian G., 2021. Landslide investigation using Seismic Refraction Tomography method: a review. Annals of Geophysics. 64, 6, SE657-SE657. http s://doi.org/10.4401/ag-8633.

Intrieri, E., Raspini, F., Fumagalli, A., Lu, P., Del Conte, S., Farina, P., Allievi J. Ferretti A., Casagli, N., 2018. The Maoxian landslide as seen from space: detecting precursors of failure with Sentinel-1 data. Landslides, 15, 123-133. https://doi.org/10.1007/s10346-017-0915-7.

Jiang, S., Ma, J., Liu, Z., Guo, H., 2022. Scientometric Analysis of Artificial Intelligence (AI) for Geohazard Research Sensors 22, no. 20, 7814. https://doi.org/10.3390/s22207814

Juang, C.S., Stanley, T.A., Kirschbaum, D.B., 2019. Using citizen science to expand the global map of landslides: Introducing the Cooperative Open Online Landslide Repository (COOLR). PloS one, 14, 7, e0218657. https://doi.org/10.1371/journal.pone.0218657.

Kalantar, B., Pradhan, B., Naghibi, S.A., Motevalli, A., Mansor, S., 2018. Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). Geomatics, Natural Hazards and Risk, 9, 1, 49-69. https://doi.org/10.1080/19475705.2017.1407368.

Keqiang, H., Jibao, Y., Sijing, W., 2005. Analysis of dynamic factors of debris landslide by means of the model of quantitative theory—using the Xintan landslide, China, as an example. Environ Geol 48, 676–681. https://doi.org/10.1007/s00254-005-0002-6.

Kirschbaum, D.B., Adler, R., Hong, Y., Hill, S., Lerner-Lam, A., 2010. A global landslide catalog for hazard applications: method, results, and limitations. Nat. Hazards 52, 561–575. http://dx.doi.org/10.1007/s11069-009-9401-4.

Kumar, R., Anbalagan, R., 2019. Landslide susceptibility mapping of the Tehri reservo

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

ir rim area using the weights of evidence method. Journal of Earth System Science, 128, 1-18. https://doi.org/10.1007/s12040-019-1159-9.

Kuriakose, S.L., Sankar, G., Muraleedharan, C., 2009. History of landslide susceptibility and a chorology of landslide-prone areas in the Western Ghats of Kerala, India. Environmental geology, 57, 1553-1568. https://doi.org/10.1007/s00254-008-1431-9.

Lagomarsino, D., Segoni, S., Rosi, A., Rossi, G., Battistini, A., Catani, F., Casagli, N., 2015. Quantitative comparison between two different methodologies to define rainfall thresholds for landslide forecasting, Nat. Hazards Earth Syst. Sci., 15, 2413–2423, https://doi.org/10.5194/nhess-15-2413-2015, .

Lavoué, F., Coutant, O., Boué, P., Pinzon‐Rincon, L., Brenguier, F., Brossier, R., Dales, P.; Rezaeifar, M., Bean, C.J., 2021. Understanding seismic waves generated by train traffic via modeling: Implications for seismic imaging and monitoring. Seismological Society of America, 92, 1, 287-300. https://doi.org/10.1785/0220200133.

Li, C., Li, L., Zhang, C., 2023. Deformation Analysis of Guobu Slope based on SBAS-InSAR. Academic Journal of Science and Technology, 5, 3, 126-131. https://doi.org/10.54097/ajst.v5i3.7803.

Li, X.E., Zhou, L., Su, F.Z., Wu, W.Z., 2021. Application of InSAR technology in landslide hazard: Progress and prospects. Natl. Remote Sens. Bull, 25, 02, 614-629. https://doi.org/10.11834/jrs.20209297.

Li, Y., Zhang, Y., Su, X., Zhao, F., Liang, Y., Meng, X., Jia, J., 2021. Early identification and characteristics of potential landslides in the Bailong River Basin using InSAR technique. Yaogan Xuebao/Journal of Remote Sensing, 25, 2, 677-690. https://doi.org/10.11834/jrs.20210094.

Lin, Q., Wang, Y., 2018. Spatial and temporal analysis of a fatal landslide inventory in China from 1950 to 2016. Landslides, 15, 2357-2372. https://doi.org/10.1007/s10

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

346-018-1037-6.

Lin, Z., Ji, Y., Sun, X. 2023. Landslide Displacement Prediction Based on CEEMDAN Method and CNN–BiLSTM Model. Sustainability, 15, 13, 10071. https://doi.org/10.3390/su151310071.

Liu, J., Wu, Y., Gao, X., Zhang, X., 2022. A Simple Method of Mapping Landslides Runout Zones Considering Kinematic Uncertainties. Remote Sensing, 14, 3, 668. https://doi.org/10.3390/rs14030668.

Liu, Y., Xu, C., Huang, B., Ren, X., Liu, C., Hu, B., Chen, Z. 2020. Landslide displacement prediction based on multi-source data fusion and sensitivity states. Engineering Geology, 271, 105608. https://doi.org/10.1016/j.enggeo.2020.105608.

Lu, H., Ma, L., Fu, X., Liu, C., Wang, Z., Tang, M., Li, N. 2020. Landslides information extraction using object-oriented image analysis paradigm based on deep learning and transfer learning. Remote Sensing, 12, 5, 752. https://doi.org/10.3390/rs12050752.

Lu, W., Hu, Y., Zhang, Z.Cao, W., 2023. A dual-encoder U-Net for landslide detection using Sentinel-2 and DEM data. Landslides 20, 1975–1987. https://doi.org/10.1007/s10346-023-02089-5.

Ma, W., Dong, J., Wei, Z., Peng, L., Wu, Q., Wang, X., Dong, Y. Wu, Y., 2023. Landslide susceptibility assessment using the certainty factor and deep neural network. Frontiers in Earth Science, 10, 1091560. https://doi.org/10.3389/feart.2022.1091560.

Ma, Z., Mei, G., Piccialli, F., 2021. Machine learning for landslides prevention: a survey. Neural Comput & Applic 33, 10881–10907 . https://doi.org/10.1007/s00521-020-05529-8.

Mebrahtu, T.K., Heinze, T., Wohnlich, S., Alber, M., 2022. Slope stability analysis of

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

deep-seated landslides using limit equilibrium and finite element methods in Debre Sina area, Ethiopia. Bulletin of Engineering Geology and the Environment, 81, 10, 403. https://doi.org/10.1007/s10064-022-02906-6.

Morin, G.P., Lavé, J., France‐Lanord, C., Rigaudier, T., Gajurel, A.P., Sinha, R., 2018. Annual sediment transport dynamics in the Narayani basin, Central Nepal: assessing the impacts of erosion processes in the annual sediment budget. Journal of Geophysical Research: Earth Surface, 123, 10, 2341-2376. https://doi.org/10.1029/2017JF004460.

N.J. van Eck, L., 2010. Waltman Software survey: VOSviewer, a computer program for bibliometric mapping Scientometrics, 84, 2, pp. 523-538, https://doi.org/10.1007/s11192-009-0146-3.

Nappo, N., Peduto, D., Mavrouli, O., van Westen, C.J., Gullà, G., 2019. Slow-moving landslides interacting with the road network: Analysis of damage using ancillary data, in situ surveys and multi-source monitoring data. Engineering geology, 260, 105244. https://doi.org/10.1016/j.enggeo.2019.105244.

Newcomen W., Dick G., 2015. An update to strain-based pit wall failureprediction method and a justification for slope monitoring. In: Proceedings of slope stability 2015. Cape Town, South Africa, 116, 5, pp. 139–150. https://doi.org/10.17159/2411-9717/2016/v116n5a3.

Overland, J.E., Walsh, J.E., Wang, M., 2008. Global Outlook for Ice and Snow: Why are Ice and Snow Changing? United Nation's Environment Program. Accessed March 21, 2012. http://www.unep.org/geo/geo_ice/PDF/GEO_C3_LowRes.pdf

Patil, A.S., Panhalkar, S.S., 2023. Remote sensing and GIS-based landslide susceptibility mapping using LNRF method in part of Western Ghats of India. Quaternary Science Advances, 100095. https://doi.org/10.1016/j.qsa.2023.100095.

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Petley, D.N., 2012. Landslides and engineered slopes: protecting society through improved understanding. Landslides and engineered slopes, 1, 3-13.

Pradhan, B., 2011. Use of GIS-based fuzzy logic relations and its cross application to produce landslide susceptibility maps in three test areas in Malaysia. Environmental Earth Sciences, 63, 2, 329-349. https://doi.org/10.1007/s12665-010-0705-1.

Qian, J., Li, L., Wu, S., Liu, J., Zhang, Y. 2023. Improving Railway Alignment Selection in Mountainous Areas with Complex Vegetation: A Multisource Data Landslide Identification Approach for Assisted Decision-Making Research. Sustainability, 15, 14, 11388. https://doi.org/10.3390/su151411388.

Quigley, R.M., Matich, M.A.J., Horvath, R.G., Hawson, H.H., 1971. Swelling clay in two slope failures at Toronto, Canada. Canadian Geotechnical Journal, 8, 3, 417-424. https://doi.org/10.1139/t71-043.

Reichenbach, P., Rossi, M., Malamud, B.D., Mihir, M., Guzzetti, F., 2018. A review of statistically-based landslide susceptibility models. Earth-science reviews, 180, 60-91. https://doi.org/10.1016/j.earscirev.2018.03.001.

Robbins, J.C., 2016. A probabilistic approach for assessing landslide-triggering event rainfall in Papua New Guinea, using TRMM satellite precipitation estimates. Journal of Hydrology, 541, 296-309. https://doi.org/10.1016/j.jhydrol.2016.06.052.

Robbins, J.C., 2018. Rainfall characteristics and critical rainfall for landslides in Papua New Guinea. Landslides and Engineered Slopes. Experience, Theory and Practice, 1731-1738. https://doi.org/10.1201/9781315375007-204.

Ruiz-Villanueva, V., Allen, S., Arora, M., Goel, N.K., Stoffel, M., 2017. Recent catastrophic landslide lake outburst floods in the Himalayan mountain range. Progress in Physical Geography, 41, 1, 3-28. https://doi.org/10.1177/0309133316658614.

Saito M. 1969. Forecasting time of slope failure by tertiary creep. Proceedings of the

7th International Conference on Soil Mechanics and Foundation Engineering Mexico city , 2, 677-683.

Salvatici, T., Tofani, V., Rossi, G., D'Ambrosio, M., Tacconi Ste-fanelli, C., Masi, E. B., Rosi, A., Pazzi, V., Vannocci, P., Petrolo,M., Catani, F., Ratto, S., Stevenin, H., Casagli, N., 2018. Applica-tion of a physically based model to forecast shallow landslides at a regional scale. Nat. Hazards Earth Syst. Sci., 18, 1919–1935, https://doi.org/10.5194/nhess-18-1919-2018.

Sameen M. I., and Pradhan B., 2019. Landslide Detection Using Residual Networks and the Fusion of Spectral and Topographic Information. IEEE, 7, 114363-114373. doi: 10.1109/ACCESS.2019.2935761.

Scoppettuolo, M.R., Cascini, L., Babilio, E., 2020. Typical displacement behaviours of slope movements. Landslides, 17, 1105–1116. https://doi.org/10.1007/s10346-019-01327-z.

Segoni, S., Piciullo, L., Gariano, S. L., 2018. A review of the recent literature on rainfall thresholds for landslide occurrence. Landslides, 15, 8, 1483-1501. https://doi.org/10.1007/s10346-018-0966-4.

Sevgen, E., Kocaman S., Nefeslioglu H.A., Gokceoglu C., 2019. A Novel Performance Assessment Approach Using Photogrammetric Techniques for Landslide Susceptibility Mapping with Logistic Regression, ANN and Random Forest. Sensors. 19, 18: 3940. https://doi.org/10.3390/s19183940.

Sreelakshmi S., Vinod Chandra S.S., Shaji, E, 2022. Landslide identification using machine learning techniques: Review, motivation, and future prospects. Earth Sci Inform 15, 2063–2090. https://doi.org/10.1007/s12145-022-00889-2.

Sunbul, F., Haner, B., Mungan, H., Akarsu, V., Sunbul Guner, A.B., Temiz, C., 2021. Stability analysis of a landslide: a view with implications of microstructural soil

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

characters. Indian Geotechnical Journal, 51, 647-660. https://doi.org/10.1007/s40098-020-00467-7.

Tehrani, F. S., Calvello, M., Liu, Z., Zhang, L., Lacasse, S., 2022. Machine learning and landslide studies: recent advances and applications. Nat Hazards 114, 1197–1245. https://doi.org/10.1007/s11069-022-05423-7.

Thiery, Y., Terrier, M., Colas, B., Fressard, M., Maquaire, O., Grandjean, G., Gourdier, S., 2020. Improvement of landslide hazard assessments for regulatory zoning in France: STATE–OF–THE-ART perspectives and considerations. Int. J. Disaster Risk Reduct. 47, 101562. https://doi.org/10.1016/j.ijdrr.2020.101562.

Varnes, D.J., 1978. Slope Movement Types and Processes, SpecialReport, 176, 11–33, https://doi.org/10.1016/j.mser.2018.11.001.

Wang, B., He, L., He, Z., Qu, R., Kang, G., 2023. Study of early identification method for large landslides in high vegetation coverage areas of Southwest China. Frontiers in Ecology and Evolution, 11, 1169028. https://doi.org/10.3389/fevo.2023.1169028.

Wang, C., Kong, D., Song, H., Liu, J., Qi, M., Li, L., 2023. Characterization of Global Research Trends and Prospects on Moyamoya disease: Bibliometric Analysis. World Neurosurgery. 173, e329-e340, https://doi.org/10.1016/j.wneu.2023.02.047.

Wang H.,Zhang L., Yin K., Luo H., Li J., 2021. Landslide identification using machine learning,Geoscience Frontiers,12, 1, 351-364. https://doi.org/10.1016/j.gsf.2020.02.012.

Wei, X., Wenkai, F., 2021. Application of Slope Radar (S-SAR) in Emergency Monitoring of the “11.03” Baige Landslide. Mathematical Problems in Engineering, 2021, 1-12. https://doi.org/10.1155/2021/2060311.

Weissel, J.K., Czuchlewski, K.R., Kim, Y. 2004. Synthetic aperture radar (SAR)-based

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

mapping of volcanic flows: Manam Island, Papua New Guinea. Natural Hazards and Earth System Sciences, 4, 2, 339-346. https://doi.org/10.5194/nhess-4-339-2004, 2004.

Whiteley, J.S., Chambers, J.E., Uhlemann, S., Wilkinson, P. B., Kendall, J. M., 2019. Geophysical monitoring of moisture‐induced landslides: a review. Reviews of Geophysics, 57, 1, 106-145. https://doi.org/10.1029/2018RG000603.

Whiteley, J.S., Watlet, A., Uhlemann, S., Wilkinson, P., Boyd, J.P., Jordan, C., Kendall, J.M., Chambers, J.E., 2021. Rapid characterisation of landslide heterogeneity using unsupervised classification of electrical resistivity and seismic refraction surveys. Engineering Geology, 290, 106189. https://doi.org/10.1016/j.enggeo.2021.106189.

Xiao, L., Wang, J., Ward, S.N., Chen, L., 2018. Numerical modeling of the June 24, 2015, Hongyanzi landslide generated impulse waves in Three Gorges Reservoir, China. Landslides, 15, 2385-2398. https://doi.org/10.1007/s10346-018-1057-2.

Xie, Z.H., Luan, T.T., & He, N., 2013. Safety evaluation technology for waste dump landslide of open-pit mine based on fuzzy mathematics. Applied Mechanics and Materials, 353, 2245-2250. https://doi.org/10.4028/www.scientific.net/AMM.353-356.2245.

Xu, Q., Zhao, B., Dai, K., Dong, X., Li, W., Zhu, X., Yang Y., Xiao X., Wang X., Huang J., Lu H., Deng B., Ge, D., 2023. Remote sensing for landslide investigations: A progress report from China. Engineering Geology, 107156. https://doi.org/10.1016/j.enggeo.2023.107156.

Yin, W., Niu, C., Bai, Y., Zhang, L., Ma, D., Zhang, S., Zhou, X., Xue, Y, 2023. An Adaptive Identification Method for Potential Landslide Hazards Based on Multisource Data. Remote Sensing 15, 7, 1865. https://doi.org/10.3390/rs15071865.

Youssef, A.M., Pradhan, B., Al-Harthi, S.G., 2015. Assessment of rock slope stability

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

and structurally controlled failures along Samma escarpment road, Asir Region (Saudi Arabia). Arabian Journal of Geosciences, 8, 6835-6852. https://doi.org/10.1007/s12517-014-1719-x.

Yu, D., Kou, G., Xu, Z., Shi, S., 2021a. Analysis of collaboration evolution in AHP research: 1982-2018. International Journal of Information Technology & DecisionMaking, 20, 1, 7-36. https://doi.org/10.1142/S0219622020500406.

Yu D, and Pan T., 2021b. Tracing knowledge diffusion of TOPSIS: A historical perspective from citation network. Expert Syst Appl, 168, 114238. https://doi.org/10.1016/j.eswa.2020.114238.

Yu, D., Sheng, L., Xu, Z., 2022. Analysis of evolutionary process in intuitionistic fuzzy set theory: A dynamic perspective. Information Sciences, 601, 175-188. https://doi.org/10.1016/j.ins.2022.04.019.

Yu, D., Xu, Z., Pedrycz, W., Wang, W., 2017. Information sciences 1968-2016: A retrospective analysis with text mining and bibliometric. Information Sciences, 418, 619-634. https://doi.org/10.1016/j.ins.2017.08.031.

Yue, X., Wu, S., Yin, Y., Gao, J., Zheng, J. 2018. Risk identification of seismic landslides by joint Newmark and RockFall analyst models: a case study of roads affected by the Jiuzhaigou earthquake. International Journal of Disaster Risk Science, 9, 392-406. https://doi.org/10.1007/s13753-018-0182-9.

Zhang, S., Zhao, L., Delgado-Tellez, R., Bao, H., 2018. A physics-based probabilistic forecasting model for rainfall-induced shal-low landslides at regional scale, Nat. Hazards Earth Syst. Sci., 18, 969–982, https://doi.org/10.5194/nhess-18-969-2018.

Zhao, C., Lu, Z., Zhang, Q., de La Fuente, J., 2012. Large-area landslide detection and monitoring with ALOS/PALSAR imagery data over Northern California and Southern Oregon, USA. Remote sensing of environment, 124, 348-359. https://doi.org

https://doi.org/10.5194/nhess-2024-68
Preprint. Discussion started: 17 June 2024

/10.1016/j.rse.2012.05.025.

Zheng, X., He, G., Wang, S., Wang, Y., Wang, G., Yang, Z., Yu, J., Wang, N., 2021. Comparison of machine learning methods for potential active landslide hazards identification with multi-source data. ISPRS International Journal of Geo-Information, 10, 4, 253. https://doi.org/10.3390/ijgi10040253.