REVIEW ARTICLE

A bibliometric and content analysis of research trends on GIS-based landslide susceptibility from 2001 to 2020

Junpeng Huang[1] · **Xiyong Wu**[1,2] · **Sixiang Ling**[1,2] · **Xiaoning Li**[3] · **Yuxin Wu**[1] · **Lei Peng**[1] · **Zhiyi He**[1]

Abstract
To assess the status of hotspots and research trends on geographic information system (GIS)–based landslide susceptibility (LS), we analysed 1142 articles from the Thomas Reuters Web of Science Core Collection database published during 2001–2020 by combining bibliometric and content analysis. The paper number, authors, institutions, corporations, publication sources, citations, and keywords are noted as sub/categories for the bibliometric analysis. Thematic LS data, including the study site, landslide inventory, conditioning factors, mapping unit, susceptibility models, and mode fit/prediction performance evaluation, are presented in the content analysis. Then, we reveal the advantages and limitations of the common approaches used in thematic LS data and summarise the development trends. The results indicate that the distribution of articles shows clear clusters of authors, institutions, and countries with high academic activity. The application of remote sensing technology for interpreting landslides provides a more convenient and efficient landslide inventory. In the landslide inventory, most of the sample strategies representing the landslides are point and polygon, and the most frequently used sample subdividing strategy is random sampling. The scale effects, lack of geographic consistency, and no standard are key problems in landslide conditioning factors. Feature selection is used to choose the factors that can improve the model's accuracy. With advances in computing technology and artificial intelligence, LS models are changing from simple qualitative and statistical models to complex machine learning and hybrid models. Finally, five future research opportunities are revealed. This study will help investigators clarify the status of LS research and provide guidance for future research.

Keywords Bibliometric analysis · Content analysis · Landslide susceptibility · GIS · Research trends · Machine learning

Introduction

Landslides are the most common natural hazards and cause damage to infrastructure and natural ecosystems; they result in serious casualties and tremendous property losses and affect the sustainable development of the human living environment in mountainous regions (Lee et al. 2007; Hungr et al. 2014; Sassa, 2019; Ling et al. 2021). Therefore, the effective prediction and susceptibility mapping of landslides are regarded as urgent tasks to reduce the related detrimental impacts, which is of great significance for promoting the sustainable development of society and the environment (Zhao et al. 2019). Landslide susceptibility (LS) assessments have been performed since the 1970s to solve practical problems at the small-catchment, regional, national, and global scales (Neuland 1976; Lin et al. 2017; Tang et al. 2020). Essentially, LS is defined as the likelihood of landslide occurrence in a given study area on the basis of the local environmental conditions by predicting where landslides are likely to occur (Brabb 1984). All LS approaches and methods are generally based on the following assumptions: (1) The deformation and failure signs of landslides can be identified through field investigation or remote sensing images; (2) conditions that affect landslide occurrence are directly or indirectly

---

✉ Sixiang Ling
  lingsx@swjtu.edu.cn

  Junpeng Huang
  huangjp90@163.com

[1] Present Address: Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, 611756 Chengdu, China

[2] Ministry of Education, Key Laboratory of High-Speed Railway Engineering, Southwest Jiaotong University, Chengdu 610031, China

[3] School of Emergency Management, Xihua University, Chengdu 610039, China

linked, and thus, they can be used to build predictive models (Reichenbach et al. 2018); and (3) future slope failures will be more likely to occur under the conditions that led to previous instabilities (Kanungo 2015).

In recent years, geographic information system (GIS) technology has been extensively used to produce LS mapping because of the strong capability of GIS for spatial data collection, storage, processing, and visualisation (Bui et al. 2012; Ling et al. 2022). Generally, LS calculations based on GIS environment are performed in four complex steps, including (i) landslide inventory preparation, (ii) conditioning factor selection, (iii) model construction, and (iv) model validation and evaluation (Pradhan 2011; Huang et al. 2022a). LS approaches can be broadly categorised into empirically based, process-based, statistically based, and machine learning methods. The empirically based method (e.g. heuristic approach (HA), fuzzy logic (FL), and multicriteria decision analysis (MCDA)) is an indirect and qualitative approach that relies on expert opinions and judgements to rank and weight the instability factors. The process-based method (e.g. stability index mapping (SINMAP)) is based on a physically based modelling scheme to construct the physical process for landslide occurrence, which can compute the safety factor using a range of geotechnical, topographic, and hydrological parameters (Kim et al. 2014). The statistical models (e.g. information value (IV), statistical index (SI), certainty factor (CF), logistic regression (LR), and frequency ratio (FR)) rely on statistical analysis theory to reveal the spatial relationship between variables and landslides. Finally, machine learning attempts to build LS models by learning from complex data without banking on rules-based functions (Merghadi et al. 2020). Generally, machine learning algorithms can be divided into supervised learning (e.g. artificial neural network (ANN), decision tree (DT), random forest (RF), support vector machine (SVM), and naive Bayesian (NB)) and unsupervised learning models. Supervised learning can handle classification problems that rely on learning from labelled training data, whereas unsupervised learning attempts to predict landslides in unlabelled data (Chang et al. 2020).

In the past two decades, numerous investigators have developed models and techniques to construct maps of LS worldwide and have formed an abundant literature base. The literature, which is composed of several hundred papers, has provided a valuable overview of LS for scholars. In these studies, only a few investigators have attempted to conduct scientific reviews. In general, these reviews can be classified into the following four categories: (1) a summary of the methods of generating landslide inventory maps (Guzzetti et al. 2012; Jaboyedoff et al. 2012; Scaioni et al. 2014) and discussion of parameters derived from and/or used with digital elevation models for LS (Saleem et al. 2019; Kakavas & Nikolakopoulos 2021); (2) systematic and critical reviews of LS assessment systems that discuss landslide inventory mapping, mapping units, conditioning factors, and the different models used while also exploring the intrinsic and/or specific advantages and disadvantages of these approaches (Guzzetti et al. 1999; Wang et al. 2005; Lee & Pradhan 2007; van Westen et al. 2008; Pardeshi et al. 2013; Kanungo 2015; Saleem et al. 2019; Shano et al. 2020); (3) an overview of the availability of machine and deep learning techniques for landslide detection and/or LS assessment (Huang & Zhao 2018; Ma et al. 2021; Mohan et al. 2021); and (4) a discussion of the methods for LS model fitting and for the evaluation of the models' prediction performances (Brenning 2005; Begueria 2006; Chacon et al. 2006).

Although the above reviews provide valuable insights for the LS field, most of them may tend to be qualitative and subjective analyses. With a growing body of literature, the knowledge structure of the domain is not completely provided in these reviews. For example, it is difficult to answer the following questions using traditional review techniques: (i) What is the trend in terms of the number of publications in this domain? (ii) How are the most-productive and influential stakeholders (authors, institutions, countries, and journals) interconnected in this field? (iii) What are the evolutionary patterns of research hotspots? and (iv) How can the dynamic development of research frontiers be tracked? However, this information is very useful for investigators to understand the structure of this domain. Bibliometric analysis, as a modern technique in computer engineering, database management, and statistics, has become a prominent method for analysing published literature (Qin et al. 2022). It utilises a scientific and structured method to quantitatively analyse the distribution structure and internal relationship among numerous publications to determine the research hotspots and assess the development trends in a certain field (Chen 2017; Zhou & Song 2021). It has been applied to identify current statuses and development trends and assess advanced topics in the landslide (Wu et al. 2015; Yang et al. 2019), marine geohazard (Camargo et al. 2019), natural hazard (Fan et al. 2020), and deep learning (Li et al. 2020) domains. The relevant conclusions are significant for researchers to determine the key areas, explore future research directions, and pursue cooperation with other institutions or countries in a particular research domain. Previous literature reviews in the LS field, such as Pourghasemi et al. (2018), Reichenbach et al. (2018), and Lee (2019), reviewed the status of LS according to the authors, journals, study areas, landslide inventories, conditioning factors, models, model evaluations, and number of publications based on articles published from 2005–2016, 1983–2016, and 1999–2018,

respectively. These reviews spend considerable time manually collecting information about LS, which may cause the loss of certain key information. Furthermore, these three reviews do not analyse the advantages and disadvantages of certain methods used in LS (e.g. modelling methods and model fit/prediction performance evaluation methods). In addition, Budimir et al. (2015) and Merghadi et al. (2020) reviewed the application of LR and machine learning in LS based on articles published from 2001–2013 to 2000–2019, respectively. Both studies mostly focused on a particular model or method. Significantly, these five reviews covered an inadequate number of articles and lacked recent research findings. Moreover, cooperation network, co-citation network, and keyword co-occurrence analyses have not previously been adopted to investigate GIS-based LS research, which results in an incomplete understanding of this domain. Therefore, further systematic literature reviews that explore research expert contributions, evolution, themes, and future scholarly opportunities in the LS field are needed.

To enrich the study of GIS-based LS, a systematic and objective bibliometric analysis is conducted on a sample of 1142 publications obtained from the Thomas Reuters Web of Science Core Collection (WoSCC) database from 2001 to 2020. In addition, we also perform a content analysis on the key information (e.g. study area, landslide inventory, conditioning factors, mapping units, susceptibility models, and validation methods) from these 1142 publications to further enhance the objectiveness and comprehensiveness of the bibliometric analysis, thus identifying the research frontiers and trends in the GIS-based LS field. The contributions of this paper include combining bibliometric and content analyses to (i) present publication trends in GIS-based LS from 2001 to 2020; (ii) identify the influential authors, institutions, countries and journals; (iii) analyse the collaboration relationship of authors, institutions, and countries; (iv) reveal the main research hot themes and dynamic developments; (v) discuss the advantages and limitations for key information of LS research; and (vi) address the challenges ahead and research directions. Accordingly, it will help professional and nonprofessional investigators elucidate the status and development trends, the core author groups, and the study hotspots of LS research in the last two decades, providing guidance for future research.

Materials and methodology

Data sources and filtering strategies

Peer-reviewed publications between 2001 and 2020 were retrieved online through the WoSCC database until January 29, 2021. The search and analysis statistical flow charts
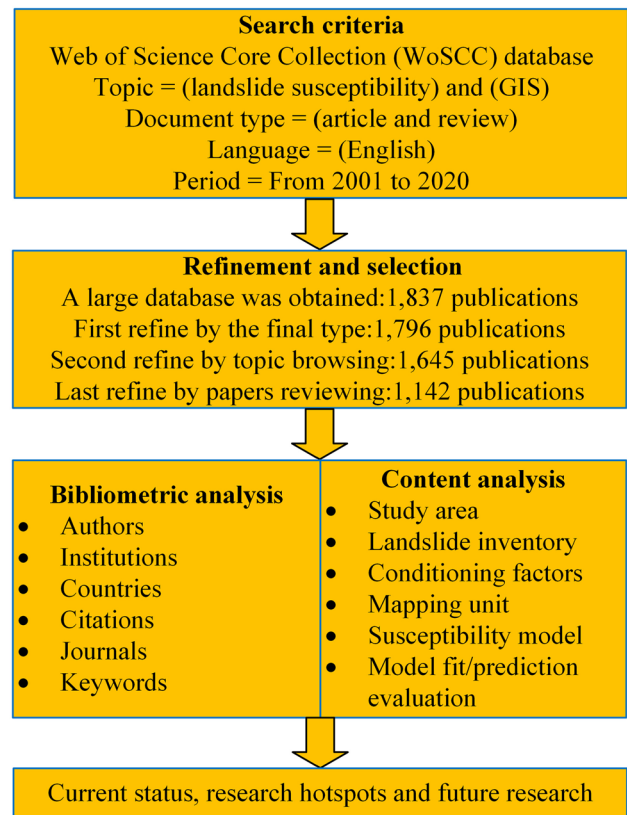
Fig. 1 Flowchart of systematic bibliometric and content analyses in the GIS-based LS field

are presented in Fig. 1. The search keywords were entered into the database as follows: Topic: (landslide susceptibility) AND Topic: (GIS) AND Language: (English) AND Document types: (article, review). Through this step, a "large database" of 1837 publications was acquired. Then, we refined the search results through the following three filtering steps. First, according to the final type of publication, we excluded publications with early online access (41 publications). Second, we checked the titles, abstracts, and keywords and excluded articles from the flood (47 publications), land subsidence (17 publications), snow avalanche (5 publications), gully or soil erosion (47 publications), and groundwater (35 publications) susceptibility domains, which identified 1645 valid records. Finally, the data were downloaded directly from the literature database, and we carefully read every publication and excluded those about landslide vulnerability and risk assessment. A total of 1142 publications in which GIS-based LS generally included the abovementioned four main steps were selected (Appendix Table S1).

Analysis methods

A bibliometric analysis was conducted using HistCite (Garfield et al. 2003) and VOSviewer (version 1.6.15, van Eck

& Waltman 2010) software. HistCite software can analyse the trends in the field of GIS-based LS and calculate the total local citation score (TLCS). The TLCS is the number of times that the papers by an author in the GIS-based LS database were cited by other papers in the database (Zhang & Chen 2020). The collaboration of authors, institutions, and countries and keyword co-occurrences were analysed using a social network in VOSviewer (Tao et al. 2020). A unified approach to mapping and clustering techniques is used to study the structure of a network. VOS (visualisation of similarities) is a distance-based map, and the idea of the mapping technique is to minimise a weighted sum of the squared Euclidean distances between all pairs of items (van Eck & Waltman, 2010). VOSviewer uses the association strength algorithm to measure the similarity between nodes and determine the thickness of the connection between nodes. In social network maps, the node size indicates the number of items (e.g. counties, institutions, authors, and keywords), and the links that connect the nodes denote cooperative relationships (Xiao et al. 2022). The thicker the line is, the greater the relationship between the nodes is (Yang et al. 2020). The clustering technique proposed in VOSviewer is a kind of weighted variant of modularity-based clustering, which contains a resolution parameter γ (Waltman et al. 2010). This clustering operates based on the same principles as node positioning (Leydesdorff & Rafols 2012). The parameter γ can be changed interactively to overcome the problem that a small cluster cannot be identified. The cluster results are automatically coloured into groups to facilitate the interpretation of relationships. The clustering analysis reveals research themes.

To identify the research content and explore the research trends, systematic content analysis was adopted for the literature data. First, six critical datasets related to GIS-based LS mapping were extracted and included in the studied database, including the (1) study area, (2) landslide inventory, (3) conditioning factors, (4) mapping units, (5) susceptibility model, and (6) model fit/prediction performance evaluation data. We then counted the number of these parameters by year and divided them into three time periods (2001–2010, 2011–2015, and 2016–2020) for a comprehensive analysis based on the number of articles.

Results

Publication trends

Figure 2 describes the annual number, total citations, and average citations of the research publications related to GIS-based LS mapping over the past 20 years. It is evident that an increasing interest in LS research and the relationship between the publication year and the cumulative
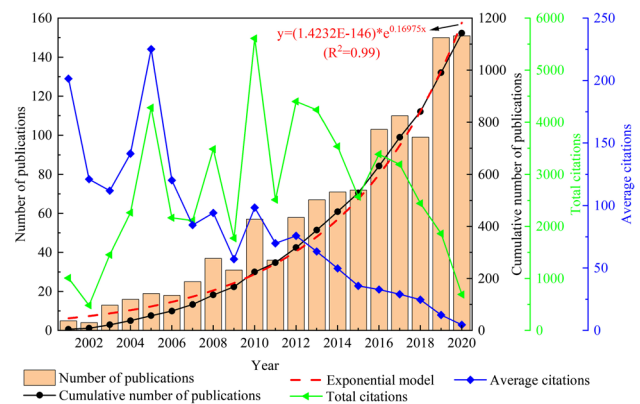

Fig. 2 Annual number and citations of research publications on GIS-based LS from 2001 to 2020

number of publications are exponential. The number of publications slowly increased from 2001 to 2009, with fewer than 40 publications each year. A steady increase in the number of publications between 2010 and 2015 can be observed, but there were fewer than 80 publications per year. Subsequently, the number of publications dramatically increased, reaching 151 in 2020. In the 20-year period, a total number of citations of publications in 2010 were the highest (5610), but the average number of citations of publications reached a maximum (225) in 2005. One factor that contributed to the development of the field may be that governments and decision-makers need LS maps as valuable decision-support tools in land use infrastructural planning and management (Ciampalini et al. 2016). Second, the "Guidelines for Landslide Susceptibility, Hazard and Risk Zoning for Land Use Planning", published by the Joint ISSMGE, ISRM and IAEG Technical Committee on Landslides and Engineered Slopes in 2008 (Fell et al. 2008), unified and standardised this field. Third, the rapid development of GIS, digital photogrammetry, global positioning systems, digital image processing, and artificial intelligence reduces the technical threshold and increases the availability of data (e.g. digital elevation data and landslide inventories), which drives more researchers to conduct more in-depth research (van Westen et al. 2008). GIS technology with a strong capability to visualise spatial data and 3D spatial analysis has made great contributions to the development of this field. The application of GIS involves any stage of LS assessment. For example, GIS can obtain terrain parameters, conduct overlay analysis between landslide and terrain parameters, and implement bivariate statistical modelling. Furthermore, with a general increase in the number of academic publications and journals related to geoscience and the environment over time, such as *Land*, the number of articles in this field has also increased. We believe that if the number of

journals and/or researchers decreases, the number of publications may also decrease.

Bibliometric analysis

Influential authors

Significantly, the analyses of cooperation networks of authors, institutions, and countries in VOSviewer are not confined to the first authors and their institutions and countries, but all signed authors, institutions, and countries are included. According to the 1142 publications, a total of 2570 authors have contributed to GIS-based LS research. However, 72.41% of all authors published only one paper, 29 authors (1.13%) published more than 10 papers, and 12 authors (1.05%) published more than 20 papers. The top 10 most-productive authors together published 457 articles (Table 1), which accounted for 40.02% of all articles. This finding suggests that although many investigators are involved in LS work, there are very few productive authors. Perhaps these investigators have switched research fields or focused on other topics, and only a few scholars have focused on one field. The most-productive and cited author is Biswajeet Pradhan from the University of Technology Sydney, with 91 publications and 10,898 citations, followed by Saro Lee from the Korea Institute of Geoscience and Mineral Resources (KIGAM), with 73 publications and 7037 citations, and Dieu Tien Bui from Duy Tan University, with 64 publications. Of them, Saro Lee, Wei Chen, Binh Thai Pham, Biswajeet Pradhan, and Dieu Tien Bui are the most-productive first authors with 31, 26, 20, 19, and 17 publications, respectively. This may be because with the development of their labs, researchers manage more graduate students, and more articles may change first authors to corresponding authors or other signed authors.

Candan Gokceoglu from Hacettepe University is the most-cited author on average, with 134.67 citations per paper. Notably, six of the top ten authors are from China, Iran, and Vietnam. The research interests of Biswajeet Pradhan focus on the field of GIS, remote sensing and image processing, machine learning and soft-computing applications, and natural hazards and environmental modelling. Based on average citations, Biswajeet Pradhan's articles rank second, indicating that the overall quality of the articles published by Pradhan is high. In his earliest research as the first author (Pradhan et al. 2006), he used remote sensing data to obtain the stress orientation and terrain variables, and then the SI model was used to predict landslides. Since then, as first author, he has published more relevant articles from 2008 to 2014, and the study area was Malaysia. Of the 10 most-cited publications in the field of GIS-based LS (Appendix Table S2), Pradhan participated in four, two as first author, indicating his significant influence in this domain. In his most-cited publication (Pradhan 2013), he compared the predication ability of different factors in the DT, SVM and adaptive neuro-fuzzy inference system (ANFIS) models. Saro Lee contributed the three most-cited publications, two of which were published by the first author. For example, he discovered that the FR model has better prediction accuracy than the LR model for predicting landslides in Malaysia (Lee & Pradhan 2007). Dieu Tien Bui also contributed the most-cited article (Bui et al. 2016), in which he introduced a framework for the training and validation of shallow LS models using the latest statistical methods.

The threshold for the number of publications was at least 5, with a total of 125 authors meeting this criterion; this is illustrated in Fig. 3. Figure 3 indicates that authors can be grouped into 6 categories based on cluster analysis with different colours for international

Table 1 Top 10 most-productive authors

Authors	Country/Institute in 2020	Total publications	Total citations	Average citations	TNLs	TLS
Pradhan, Biswajeet	Australia/University of Technology Sydney	91	10,898	119.76	45	195
Lee, Saro	South Korea/KIGAM	73	7037	96.40	30	126
Bui, Dieu Tien	Vietnam/Duy Tan University	64	4225	66.02	42	222
Chen, Wei	China/Xi'an University of Science and Technology	48	2252	46.92	38	154
Pham, Binh Thai	Vietnam/Ton Duc Thang University	45	2285	50.78	27	160
Pourghasemi, Hamid Reza	Iran/Shiraz University	39	3994	102.41	26	61
Hong, Haoyuan	China/Nanjing Normal University	28	1438	51.36	32	118
Shahabi, Himan	Iran/University of Kurdistan	26	1725	66.35	27	129
Prakash, Indra	India/Bhaskaracharya Institute for Space Applications and Geo-Informatics	22	993	45.14	8	49
Gokceoglu, Candan	Turkey/Hacettepe University	21	2828	134.67	8	43

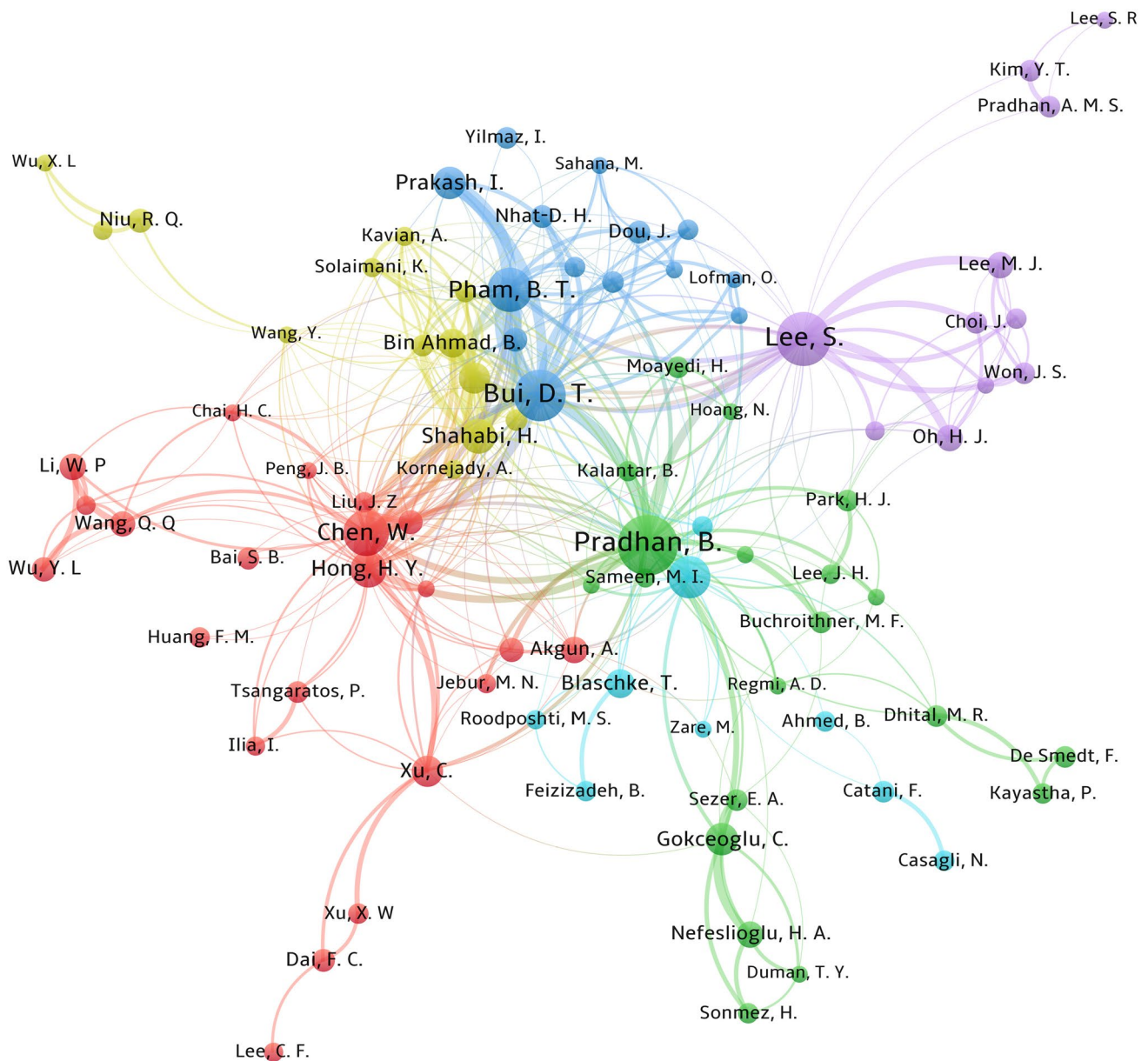TNLs, total number of links; TLS, total link strength

Fig. 3 Author collaboration network on GIS-based LS

cooperation, total number of links (TNLs) and total link strength (TLS). Generally, clusters of the same and different colours show domestic and international cooperation, respectively. TLS indicates the total strength of the links of an item to other items. When the TLS is larger, there is more co-authorship between a given author and other authors (Yang et al. 2020). The first key groups (green) can be identified around Biswajeet Pradhan in Fig. 3, where the largest TNLs of 45 and the second TLS value of 195 are displayed. This group mainly comprises Malaysian scholars. Their articles were mainly published from 2010 to 2015, and the modelling methods are primarily traditional machine learning and

statistical methods (Pradhan & Lee 2010a). A second key group (blue) mainly includes scholars from Vietnam and India, such as Dieu Tien Bui (TNLs = 30, TLS = 222, Table 1), Binh Thai Pham (TNLs = 27, TLS = 160) and Indra Prakash (TNLs = 8, TLS = 49). Their research is mainly concentrated in the last 5 years, and the models adopted are mostly machine learning, deep learning, and ensemble models. The third key group (red) mainly includes Wei Chen (TNLs = 38, TLS = 154) and Haoyuan Hong (TNLs = 8, TLS = 49) from China. They proposed an LS modelling and optimisation system based on the optimisation of conditioning factors, mapping units, and model parameters under multisource and heterogeneous

conditions (Hong et al. 2018). In addition, the cluster centred on Saro Lee (TNLs = 30, TLS = 126) is another key group in purple, comprising Korean scholars. They generally used statistical models and machine learning to predict shallow landslides caused by rainfall (Lee et al. 2006). The top 10 authors have exhibited direct or indirect cooperation, and co-authorship is common. For example, the groups led by Biswajeet Pradhan and Saro Lee have continuously focused on GIS-based LS, and they have collaborated with one another on 13 papers.

Influential research institutions

The 1142 retrieved publications involve 948 institutions. Table 2 shows the top 10 most-productive institutions, which together account for 451 publications and 39.49% of all publications. China and Iran have more research institutions (6 institutions) in this group than other countries, which suggests that a small number of institutions dominate this field. KIGAM is the most-productive and most-cited institution, with 77 publications and 7511 citations, which has promoted the development of this field. The institution has focused on the verification of landslide susceptibility mapping since 2001 (Lee & Min 2001), and the most commonly used verification methods are the receiver operating characteristic (ROC) curve and field surveys at an early stage. University Putra Malaysia ranks second in total publications and citations but first in average citations (124.11); it usually uses heuristic models, statistical models and machine learning models to predict landslides (Pradhan et al. 2008). The Chinese Academy of Science has published the third most publications (54), with 2292 citations, but the average number of citations ranks almost the lowest. This indicates that the overall quality of articles published by China needs to be improved. Heuristic, statistical, and machine learning models are frequently used by this institution

to predict landslides (Yi et al. 2019). The co-authorship network for institutions that have published GIS-based LS studies with a > 10 publication threshold is shown in Fig. 4. KIGAM is the core of the yellow cluster with 29 TNLs and 133 TLSs. The purple clusters are the centre of University Putra Malaysia for the main geographical co-institutions in Malaysia. The red clusters represent co-institutions from China, such as the Chinese Academy of Science, Xi'an University of Science & Technology, and China University of Geosciences. The institutions that form the core of the green cluster are from Middle East and Southeast Asian countries, such as Duy Tan University, which has 32 TNLs and 150 TLSs. The blue clusters are mainly collaborations among Sejong University, Shiraz University, and Islamic Azad University. The dominant co-institutions occur not only at the national level but also at the international level (Fig. 4).

Geographical distribution and international cooperation analysis

The geographical distribution of the number of articles published in the 79 countries is shown in Fig. 5. The map clearly shows that there are clear geographical clusters of high academic activity related to GIS-based LS. The most-productive countries are mainly in Asia, Europe, and North America, and most published papers are concentrated in a few countries with frequent geological disasters. China is the most-productive country, with 282 publications, accounting for 24.69% of all articles. The next most-productive countries are India, Iran, and South Korea, with 154, 148, and 139 publications, respectively. The remaining countries among the top 10 most-productive countries are Malaysia (8.32%), Turkey (8.23%), Italy (7.88%), Vietnam (7.18%), the USA (7.09%), and Norway (4.38%). The top 10 countries accounted for 66.72% of the global total document volume with the

Table 2 Top 10 most-productive research institutions for GIS-based LS studies

| Institution	Country	Total publications	Total citations	Average citations	TNLs	TLS
KIGAM	South Korea	77	7511	97.55	29	133
University Putra Malaysia	Malaysia	56	6950	124.11	26	89
Chinese Academy of Science	China	54	2292	42.44	21	66
Xi'an University of Science & Technology	China	47	2248	47.83	32	148
China University of Geosciences	China	46	738	16.04	19	34
Duy Tan University	Vietnam	44	1276	29.00	32	150
Sejong University	South Korea	35	1965	56.14	27	103
Shiraz University	Iran	33	1815	55.00	21	39
Tarbiat Modares University	Iran	30	2544	84.80	19	54
University Kurdistan	Iran	29	1693	58.38	28	121
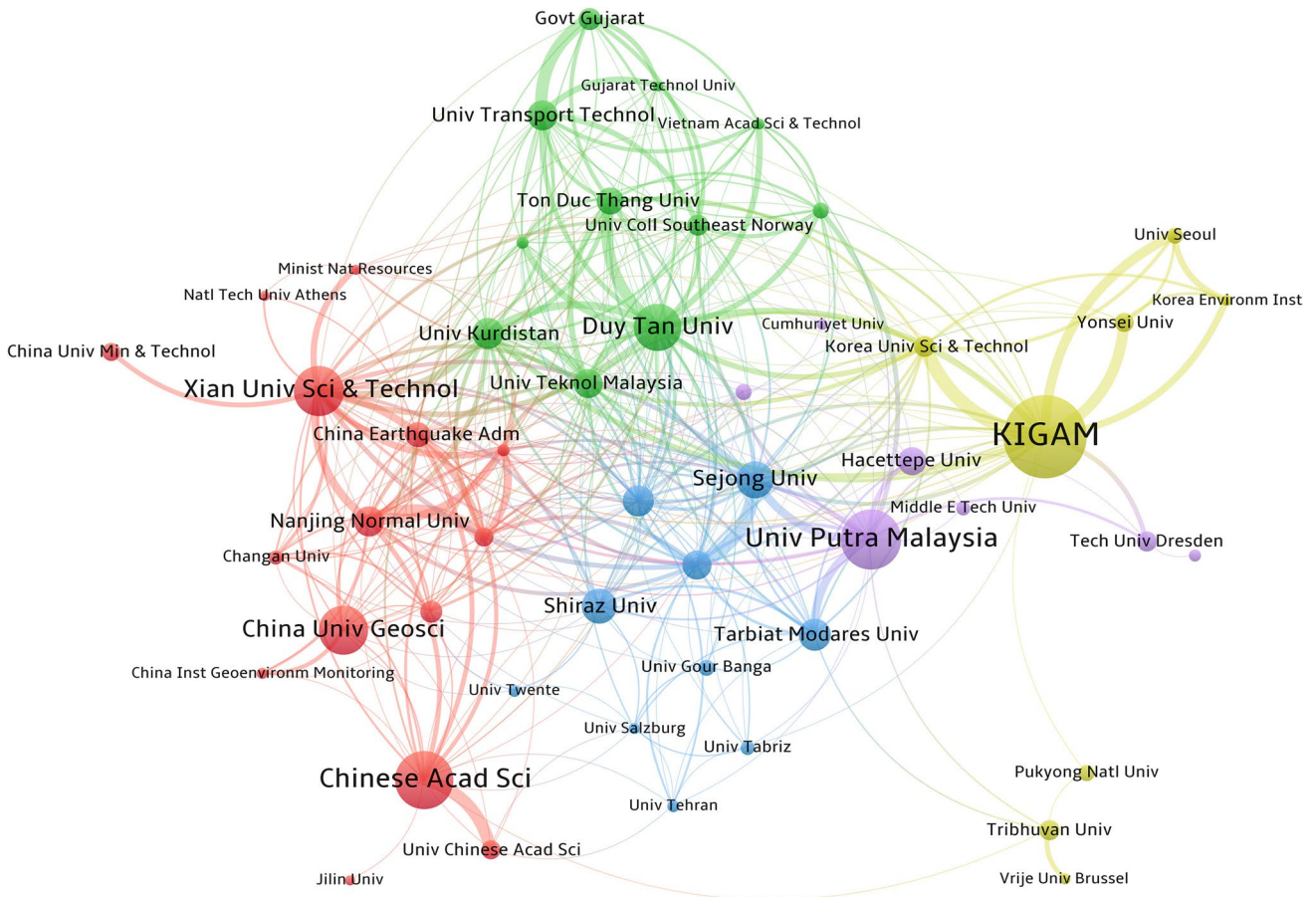
**Fig. 4** Institutional collaboration network on GIS-based LS

Fig. 5 Geographical distribution of publications on GIS-based LS
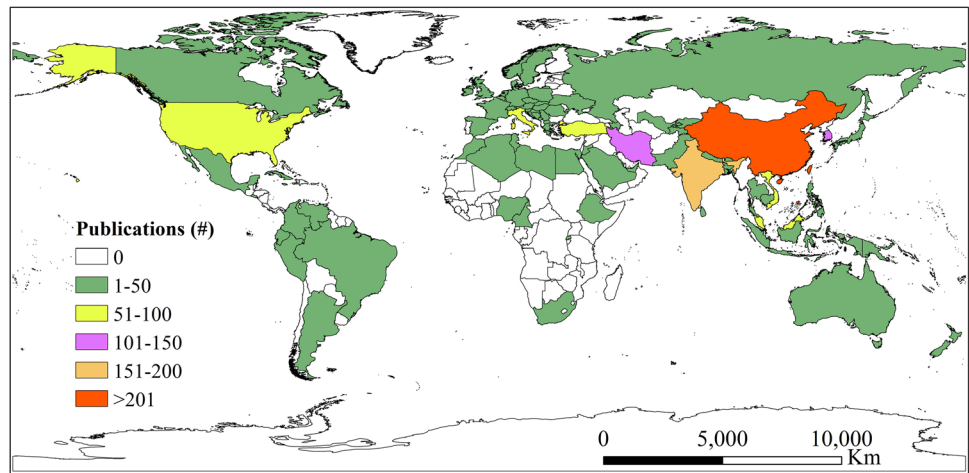
development of the economy, which means there was an increasing focus on geohazards in these countries. The number of articles of these countries or territories is not equal to the number of LS study areas in a country or territory. This is because the country or territory frequency statistics are based on all the signed authors' affiliations in the articles. Some articles have similar contents, such as the same prediction models being used in different study areas or different models being used to evaluate the LS of the same study areas. Therefore, the information of the study area, landslide inventory, conditioning factors, LS model, and evaluation method for each publication

Fig. 6 Collaboration network between countries on GIS-based LS
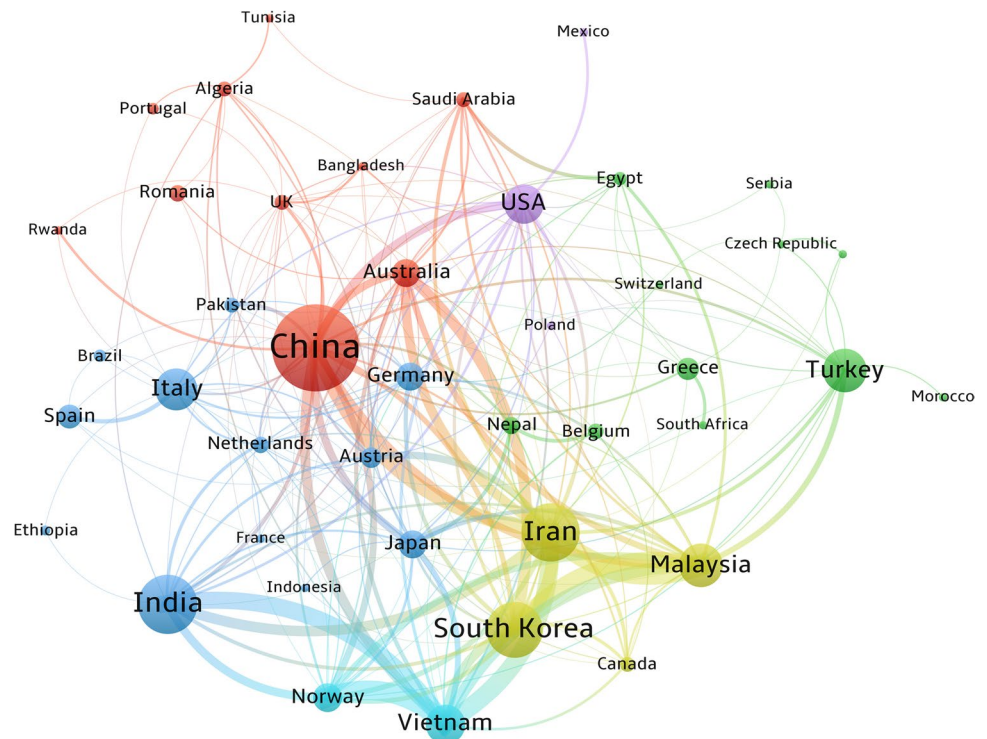
are extracted to reveal the differences in the field, and the results are shown in Section Content analysis.

The network of country collaboration for GIS-based LS studies for countries that meet the publication threshold of > 5 articles is illustrated in Fig. 6, which can be divided into six clusters. The largest cluster (red) is led by China, which occupies the central position on the map and forms an international collaborative network with 27 countries (TNLs = 27), including Australia, the UK, and Saudi Arabia. The China cluster maintains co-authorship ties with the other five international collaborative networks (Fig. 6). The blue cluster reflects an international collaboration network with three main nodes with the centres of India, Italy, and Germany. In the bottom-right part of the map, two clusters (green and yellow) include a variety of countries that are mostly located in Asia and Europe, with abundant connections between these clusters. Iran and Malaysia have the highest number of collaborations, with 41 joint papers in the yellow clusters. The cluster at the top of the diagram (purple) includes countries such as the USA, Mexico, and Poland. However, the remaining cluster (cyan) only includes Norway and Vietnam, which cooperated on 38 publications. It is apparent that the main collaborations between countries in this field include visiting scholars, visiting Ph.D. students, postdoctoral exchanges, and affiliation adjustments, such as Biswajeet Pradhan, who changed his affiliation at least 4 times at the Dresden University of Technology, University Putra Malaysia, Sejong University (South Korea), and the University of Technology Sydney (Australia).

Analysis of published sources and highly cited publications

To identify the journals that are most influential and highly published in this domain and to help investigators find suitable journals for their articles, the number of documents from different publication sources and the corresponding citation parameters were analysed. From 2001 to 2020, 183 journals published 1142 articles related to GIS-based LS, which included 92 journals that published only one article in this field.

The top 15 source publications in terms of the number of publications are listed in Table 3, and these journals published 63.92% of all articles in the last two decades. *Environmental Earth Sciences* (impact factor = 2.748) published 128 publications, which accounted for 11.21% of the total publications and made it the most predominant journal, as reflected by the maximum TLCS (3516) and average citations (183.65). The next most popular journals were *Natural Hazards* (92 publications) and *Geomorphology* (77 publications). *Geomorphology* (impact factor = 4.139) had the highest number of citations (8291) and the third-highest number of average citations (107.68) among the top 15 journals. *Landslides* (64 publications) had the third-highest number of citations (4715), and *Computers & Geosciences* (22 publications) had the second-highest number of average citations (142.59). Obviously, studies of GIS-based LS meet the requirements of these journals that involve earth science or algorithms.

Table 3 Publications in the top 15 most-productive journals from 2001 to 2020

Journal	Total publications	TLCS	Total citations	Average citations	Impact factor (2020)
Environmental Earth Sciences	128	3516	7821	183.65	2.748
Natural Hazards	92	1854	4553	49.49	3.102
Geomorphology	77	3065	8291	107.68	4.139
Landslides	64	2033	4715	73.67	6.578
Arabian Journal of Geosciences	48	530	1556	32.42	1.827
Geomatics Natural Hazards & Risk	47	400	1093	23.26	3.528
Catena	44	1322	3518	79.95	5.198
Bulletin of Engineering Geology and the Environment	40	417	984	24.60	4.298
Engineering Geology	37	1675	3745	101.22	6.755
Journal of Mountain Science	31	171	431	13.90	2.071
Remote Sensing	29	0	411	14.17	4.848
Natural Hazards and Earth System Sciences	27	433	1275	47.22	4.345
ISPRS International Journal of Geo-Information	26	14	261	10.04	2.899
Computers & Geosciences	22	1291	3137	142.59	3.372
Geocarto International	18	179	436	24.22	4.889

TLCS, total local citation score

The top 10 most-cited publications related to GIS-based LS are displayed in Appendix Table S2, indicating that these publications received 4993 citations and 2201 TLCSs. Of the top 10 most-cited articles, three were published in *Geomorphology*. The most-cited article was published in *Geomorphology* by Ayalew and Yamagishi (2005), with 890 citations and 435 TLCSs. It attempts to extend the application of LR combined with bivariate statistical analyses for LS mapping, which has a notable effect on the LS field and provides an integrated method for LS. We found that highly cited journals seem to prefer to publish new algorithms or integrated models in this research field. Therefore, investigators can easily publish articles in high-quality journals to promote this field for rapid development and obtain high citations if their results provide ground-breaking findings.

Co-occurrence network analysis of the main keywords

In this analysis, a total of 2051 terms were identified by VOSviewer from the keywords, titles, and abstracts of the 1142 articles. To eliminate the interference caused by some research topics, we removed the terms GIS and LS mapping. To assess the temporal evolution in the field, Fig. 7 shows a temporal overlay of the keyword co-occurrence map, and Table 4 lists the top 20 terms used in this research domain during three periods (2001–2010, 2011–2015, and 2016–2020) from 2001 to 2020.

The primary terms before 2010 included certain study areas, LS models, zonation, validations, and predictions. A validation and zonation of the results were considered in this period (Lee 2005). The LS models included LR, FR, ANN, and IV. The main study areas were Lantau Island (China) (Dai et al. 2001), Turkey (Yalcin & Bulut, 2007), Apennines (Italy) (Clerici et al. 2010), Malaysia (Zulhaidi et al. 2010), and Boun (South Korea) (Lee et al. 2003). We note that these sites are located in urban areas adjacent to mountains and by the sea. Because of the rapid expansion of cities, slope instability is often triggered by rainfall or extreme climate. For example, over 800 landslides occurred on Lantau Island, China, after the 1993 Ira typhoon (Dai & Lee 2001). Therefore, there have been many related studies to manage urban construction. Some study-scale terms, such as area, basin, island, region, and mountains, were also common keywords before 2010.

From 2011 to 2015, conventional models, such as LR, FR, ANN, and IV, were still popular methods to predict shallow landslides, rainfall-induced landslides, and debris flows. Only a few new or more intensely studied research topics emerged in this period, which include conditional probability, AHP, FL, and SVM. The study areas were the Wenchuan earthquake-impacted areas in China (Xu et al. 2012), the lesser Himalayas and Himalayas in India and Nepal (Das et al. 2011), and the south coast of the Black Sea in Turkey (Ercanoglu & Temiz 2011). These sites are in mostly tectonically active zones and alpine gorge areas that are prone to landslides. In addition, different sampling strategies were considered for LS mapping (Sujatha et al. 2012).

In the third period (2016 to 2020), the research topics were similar to those in the previous two periods. Thus, this period signifies a continuation of the research topics. Notably, novel LS models that use machine learning, DT,

Fig. 7 Temporal overlay of the keyword co-occurrence map. AHP, analytical hierarchy process; EBF evidential belief function; LiR, likelihood ratio; RoF, rotation forest; WLC, weighted linear combination; WoE, weight of evidence

RF, RoF, ensemble models, hybrid or integrated models, neural networks, and ANFIS received increasing attention in GIS-based LS studies during this period. The study sites are mainly at the province scale and are located in the Three Gorges Reservoir in China (Zhou et al. 2018), Hoa Binh/Yen Bai Province in Vietnam (Pham et al. 2018), and some neighbouring provinces of the Alborz Mountains in Iran (Aghdam et al. 2016).

Content analysis

Study area

Information on the study areas, such as the country, location (latitude and longitude), number of study areas, and spatial extent, was extracted from 1142 publications. However, some articles fail to provide the location and extent of the research areas directly. The latitude and longitude are determined in Google Earth based on the location names of the research areas described in the articles or the location names marked in the location map of the research areas.

These spatial extents are calculated by the polygon drawn by the unknown extent map in GIS and the scale provided in the article. A total of 1198 study areas were extracted from 1142 publications, and the numbers of articles with 1, 2, 3, and 4 study areas were 1105, 20, 15, and 2, respectively. The 1142 articles in the database include 4 country group-level, 1 continent-scale (Africa) (Broeckx et al. 2018) and 5 global-level study areas (Hong et al. 2007a, b; Hong & Adler 2008; Lin et al. 2017; Stanley & Kirschbaum, 2017). The remaining 1188 study areas in the remaining 1132 publications were distributed in 72 countries. Based on the longitudinal and latitudinal coordinates given in the publications, the spatial distribution of the study sites is presented in Fig. 8. We find that the distribution of study sites has a significant geographical bias. The study areas are mainly distributed in China (257 sites), India (126 sites), Iran (97 sites), South Korea (89 sites), and Italy (81 sites), which are all located in Asia and Europe. In the first decade of the twenty-first century, the hotspots in the study areas were in Italy and France in Europe and South Korea and Japan in Asia. The study area hotspots shifted to China, India, and

Table 4 Temporal distributions of the top 20 terms

Global		2001–2010		2011–2015		2016–2020	
Terms	N (#)	Terms	N (#)	Terms	N (#)	Terms	N (#)
LR	697	Landslide	100	LR	182	LR	420
FR	415	Hazard	95	Landslide	116	FR	281
Landslide	408	LR	95	Hazard	116	AHP	250
Hazard	404	Area	66	Area	102	Area	194
Area	362	Validation	45	FR	94	Hazard	193
AHP	321	Model	43	ANN	84	Landslide	192
Model	293	Zonation	43	Model	81	SVM	170
ANN	275	FR	40	Turkey	67	Model	169
SVM	213	ANN	39	AHP	54	ANN	152
Remote sensing	170	Remote sensing	39	Remote sensing	53	DT	123
Hazard assessment	166	Hazard assessment	32	Validation	52	Spatial prediction	116
Zonation	156	Prediction	30	Hazard assessment	50	Hazard assessment	84
Validation	149	Lantau-island	29	Zonation	46	RF	78
Turkey	140	Turkey	29	Conditional probability	45	Remote sensing	78
DT	140	Apennines	24	FL	41	Machine learning	78
Prediction	134	IV	23	SVM	41	Neural network	76
Spatial prediction	132	Region	22	IV	33	Prediction	72
Neural network	116	Earthquake	20	Prediction	32	Zonation	67
FL	107	Maps	20	Neural network	30	Bivariate statistical models	62

Fig. 8 The geographical distribution of the study areas
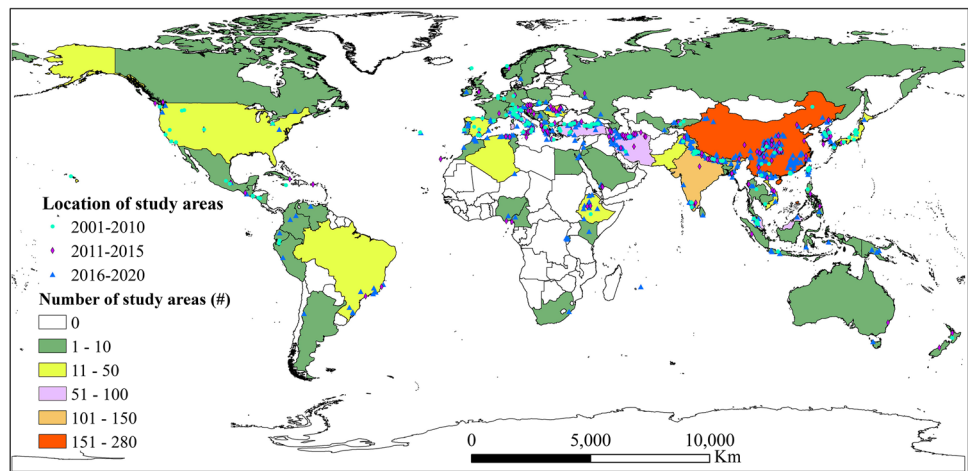
Iran from 2011 to 2015. In the last 5 years, i.e. from 2016 to 2020, the hotspots in the study areas were basically the same as those in the previous stage, but some African and South American countries (e.g. Ethiopia and Brazil) were given increasing attention in this field.

After excluding the continental- and global-scale assessments, the distribution size of the remaining 1192 study areas is illustrated in Fig. 9, with a total coverage area of 30.77 million km². The extent of the study areas varies from 0.05 km² to 9.6 million km². The size of the study areas was largest in Asia, particularly from 2011 to 2015. The median size of the study

areas showed an increasing trend. Based on the size of the study area, we divided the study area scale for LS assessment into the detailed scale (0 ~ 10 km²), medium scale (10 ~ 100 km²), large scale (100 ~ 1000 km²), regional scale (1000 ~ 100,000 km²), and national scale (> 100,000 km²). The 57 detailed-scale assessments in the database mainly focused on small watersheds and specific areas in South Korea (11), Italy (9) and China (5). The medium-scale LS assessments encompassed 275 study areas, including small watersheds, basins, valleys, towns, and cities in South Korea (62), India (40), Italy (34), Turkey (23), and China (21). Large-scale
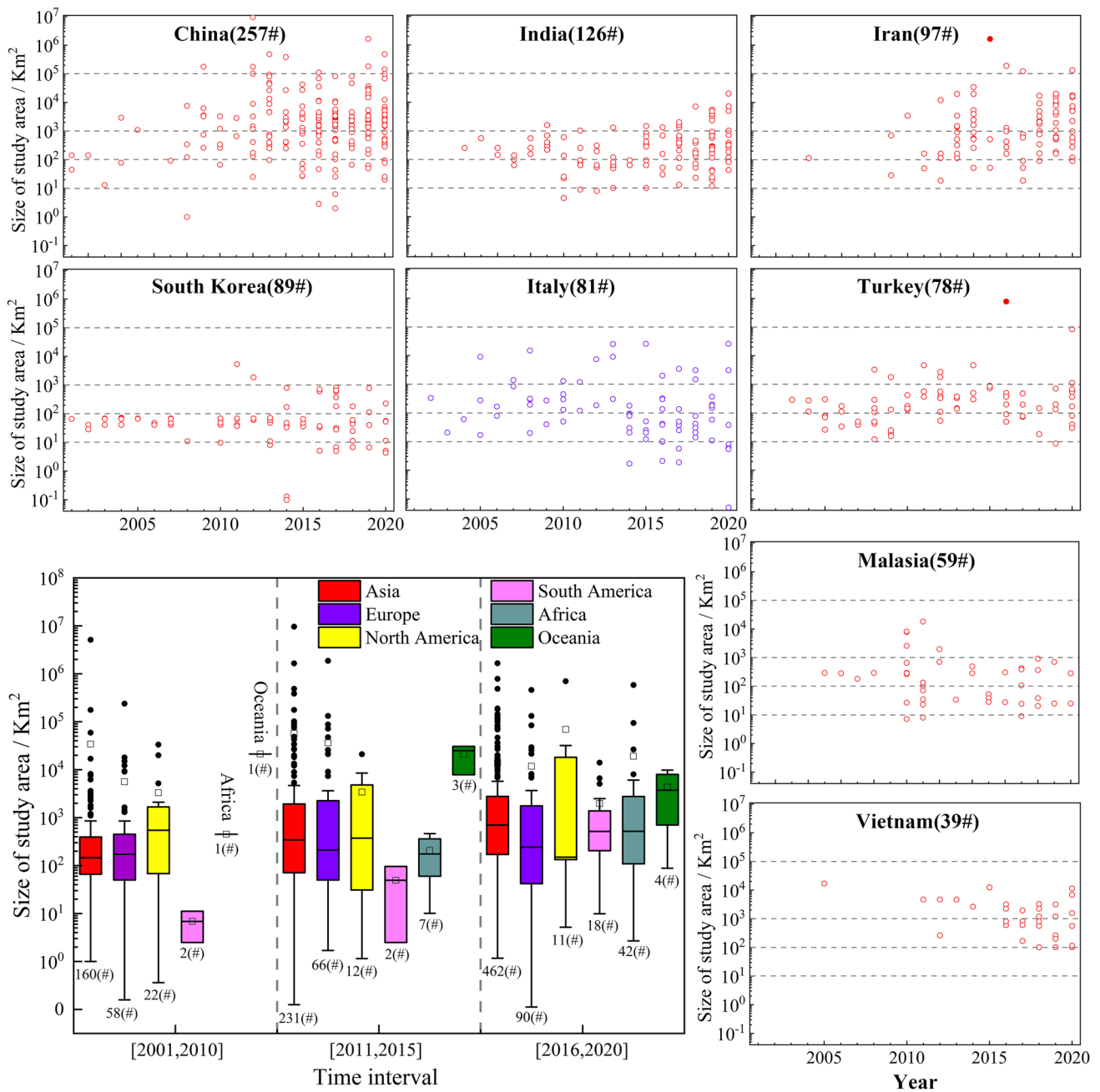
Fig. 9 Size of the study areas. The solid dots in the scatter chart represent the extent of the entire country. The symbol # indicates the number of study sites

assessments were the most common, spanning 461 study areas, including large watersheds and cities in China (67), India (63), Turkey (44), Iran (43), Malaysia (33), and Italy (23). Regional-scale assessments were performed in 371 areas and mainly included large watersheds, multiple watersheds, large cities, and provinces in China (150), Iran (40), Vietnam (21) and India (20). Finally, national-scale studies focused on 28 areas, including country groups, countries, and provinces.

Landslide inventory

A study area in an article might have one or more landslide inventories. Most studies (92.74%, 1111 out of 1198 study areas) used a single inventory for a given area, whereas the other 7.26% either used > 2 landslide inventories or did not describe the detailed inventory. Figure 10 shows the geographic distribution of the number of landslides in the 1101 single-inventory cases, whereas the remaining 10 landslide

Fig. 10 Number of landslides per landslide inventory in the 1101 study areas
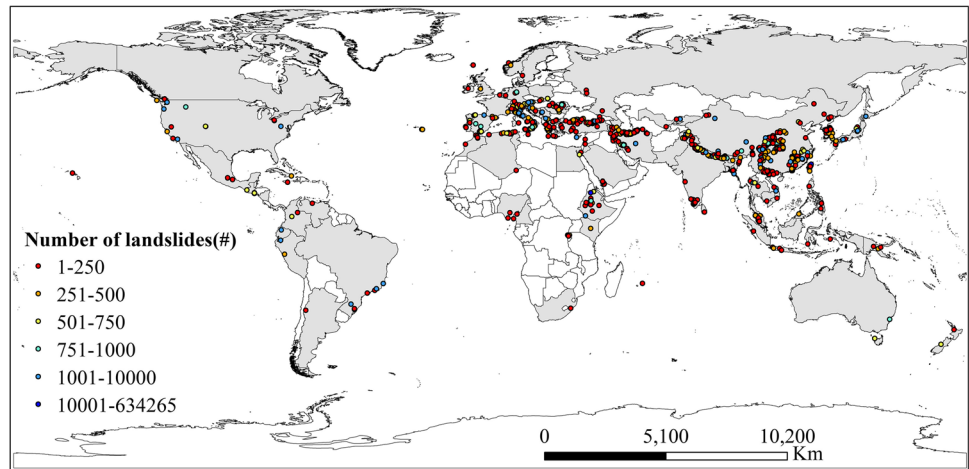
inventories were excluded due to the lack of specific longitudinal and latitudinal coordinates for the national groups, Africa, and global scales. A total of 2,048,882 landslides occurred at 1111 study sites, and an average of 1843 landslides were used in each study area to evaluate susceptibility. Huang et al. (2013) and Dumperth et al. (2016) included the maximum (634,265) and minimum (1) numbers of landslides, respectively, in the evaluations of LS in their study areas. Of the 1111 single inventories, 614 (55.27%) included 1 to 250 landslides, 227 (20.43%) included 251 to 500 landslides, and only 16 contained more than 10,000 landslides.

In addition, the different sampling strategies were applied to prepare landslide inventory maps in the database. Of the 1142 articles, 1116 articles (97.72%) had one sampling strategy, 13 articles (1.14%) used two or more sampling strategies, and the remaining articles did not describe the sampling strategy. Through the understanding of different terms and principles of various sampling strategies, we identified four sampling strategies to determine the landslide boundary, namely, point, polygon, seed cell, and circle. A point-based sampling strategy was used in 55.43% of the articles, which uses the initiation point (5 times), the highest point (5 times), and the mass (587 times) or scarp (34 times) centroid as the landslide location (a single pixel). The polygon-based sampling strategy is the second most used, accounting for 41.77% of the articles. This strategy draws the main scarp zone (38 times), the depletion zone (5 times), or the whole area composed of accumulation and depletion zones (428 times) as the landslide boundary. The other sampling strategies were less frequent, with seed cells used in 1.93% of articles and circles used in 0.61% of articles. A seed cell-based sampling strategy was used to represent the possible prefailure conditions by adding a buffer around the crown and flanks of the present landslides (Suzen & Doyuran 2004). A similar approach is that the landslide representative element is extracted on and around the landslide crown, which basically is the upper edge of the landslide scarp area,

the so-called the main scarp upper edge (MSUE) (Clerici et al. 2006). Finally, the circle-based sampling strategy uses a buffer zone around the highest point or the centroid of the scarp/polygon of the landslide as the landslide boundary.

Conditioning factors

A total of 10,980 conditioning factors were involved in 1142 publications. Overall, 1 to 34 factors were used per article, with an average of 9.61 factors. Of the 10,980 conditioning factors, we identified 430 different factors based on the following two main criteria (Reichenbach et al. 2018). First, synonymous terms are grouped together, such as "precipitation" and "rainfall". Second, variables that have different descriptions but similar meanings are grouped together. For example, "slope position" can be quantified by the "topographic position index", so "slope position" and "topographic position index" were grouped into the class "topographic position index". Ultimately, we distinguished 430 different conditioning factors involved in 1142 publications. These factors can be categorised into six main groups, namely, topographical factors, geological factors, land cover (e.g. soil and forest), hydrological factors, anthropogenic factors, and environmental factors (Fig. 11). Topographical factors were used most frequently and reached 4322 usage counters. Geological, land cover, and hydrological factors were used 1861, 1849, and 1752 times, respectively. The numbers of anthropogenic factors sharply increased and were used from 59 to 153 times and then to 387 times in the three periods. Likewise, environmental factors sharply increased from 66 to 136 to 395 times in the three periods. The doughnut analysis reveals that the proportions of the topographical, hydrological, anthropogenic, and environmental factors present increasing tendencies in the three periods, whereas the proportions of geological factors and land cover have degressive tendencies.
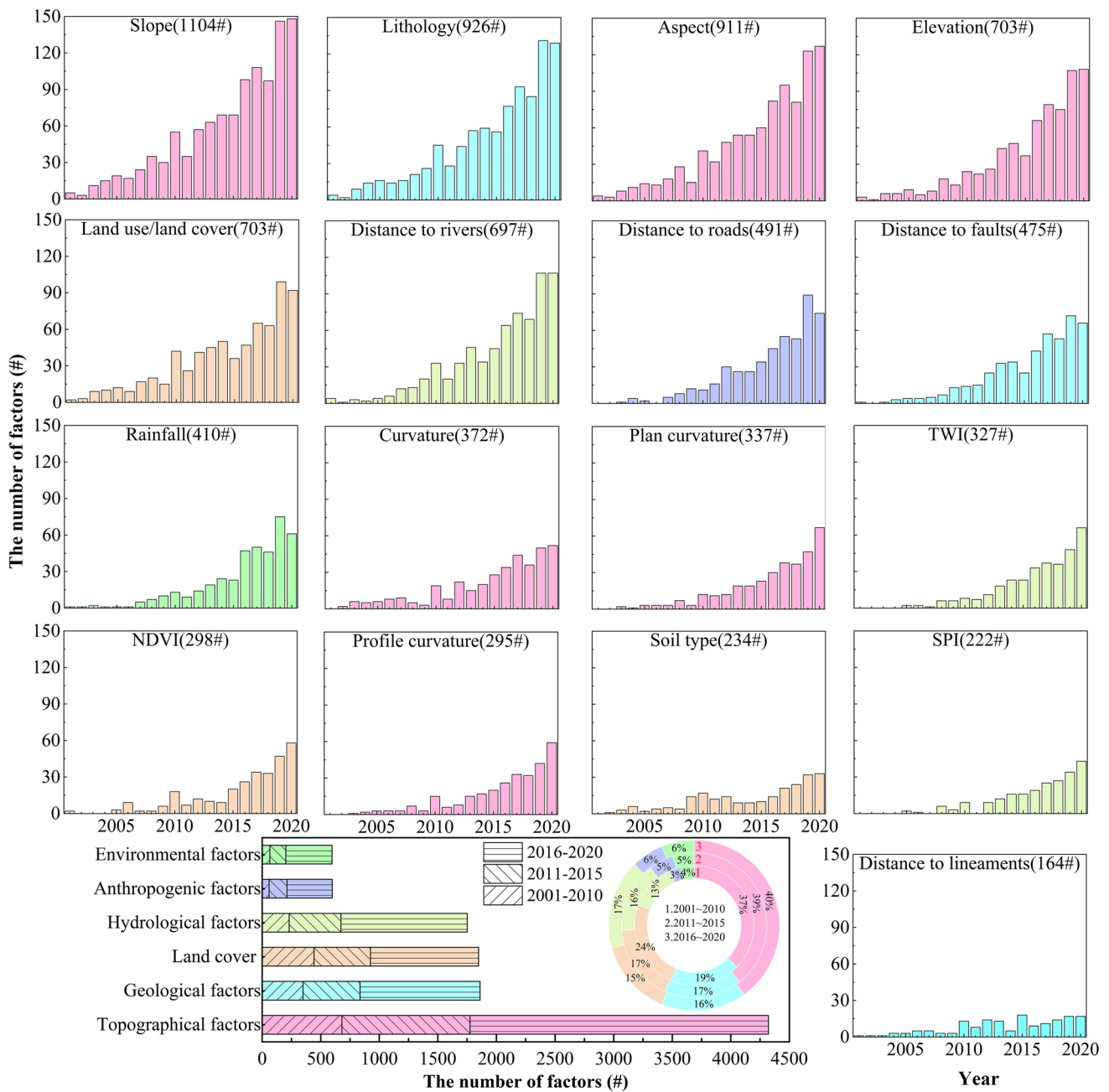
Fig. 11 Conditioning factors. The horizontal bar chart shows the count of six factor classes used in the 2001–2010, 2011–2015, and 2016–2020 periods. The doughnut shows the percentage of the clus-ters of variables used in these three periods. The vertical bar chart shows the count of the top 17 factors used per year. TWI, topographic wetness index; NDVI, normalised difference vegetation index

Slope was used most frequently, with 1104 uses in 96.67% of all articles, followed by lithology (926, 81.09%), aspect (911, 79.77%), elevation (702, 61.47%), land use/land cover (699, 61.21%), distance to roads (491, 42.99%), distance to faults (472, 41.33%), distance to rivers (429, 37.57%), rainfall (384, 33.63%), and curvature (368, 32.22%) (Fig. 11). These factors were steadily applied over time in LS assessments. In addition, the soil type, distance to drainages and stream power index (SPI) were used 40, 80, and 148 times

in the 2001–2010, 2010–2015, and 2016–2020 periods, respectively.

In addition, the conditioning factors used in different study scales were extracted from 1192 study areas, and the results are shown in Fig. 12. We found that the commonly used factors are differences for the five study scales. Topographical factors are the most applied at all study scales. For the detailed scale (0 ~ 10 km²) and medium scale (10 ~ 100 km²), the second most frequently used factor is
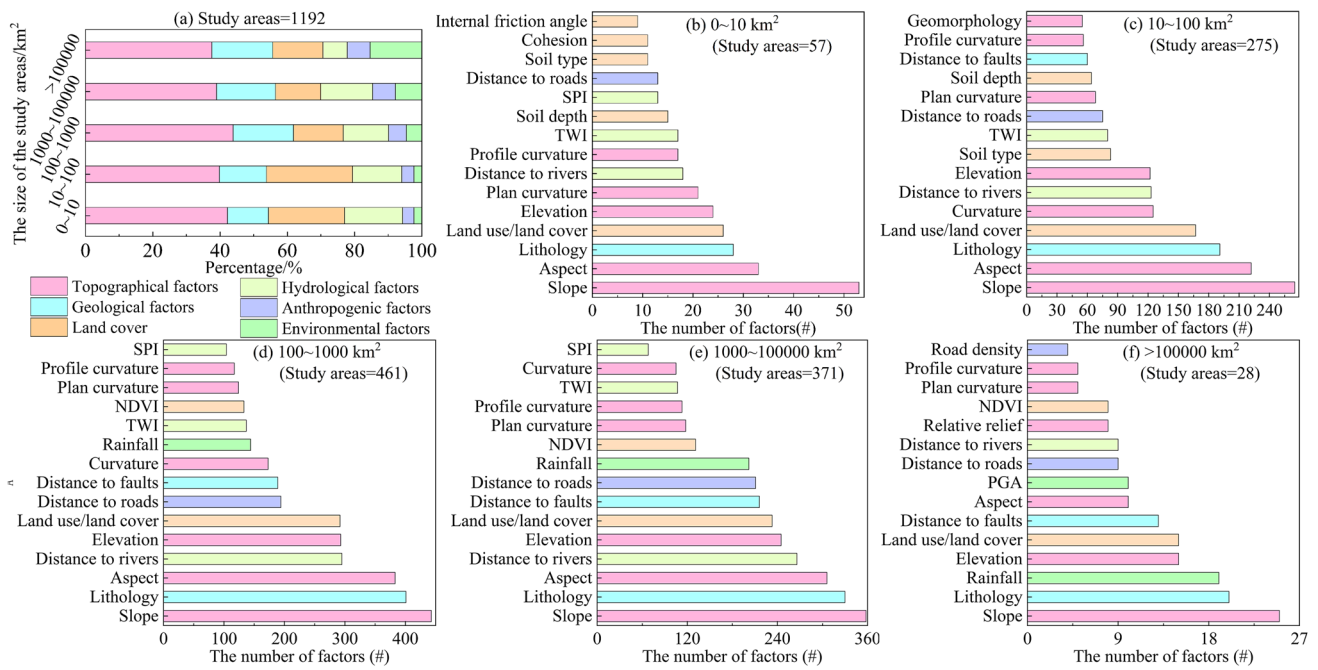
Fig. 12 Landslide conditioning factors in different size study areas. PGA, peak ground acceleration

land cover (e.g. land use/land cover, soil type, soil depth, and soil cohesion). However, geological factors are the second most applied at the large scale ($100 \sim 1000$ km^2), regional scale ($1000 \sim 100,000$ km^2), and national scale ($> 100,000$ km^2).

Mapping unit

The selection of the mapping unit, as the smallest indivisible unit, is crucial to the accuracy and practicality of LS mapping (Qin et al. 2019). Mapping units were extracted from 1142 articles. Most of the articles (98.69%) used a single mapping unit, the remaining used two or more mapping units, and a total of 1157 mapping units in the database (Fig. 13). These mapping units can be divided into five units: grid, slope, unique condition, watershed, and other types. Grid units were widely adopted in the database, accounting for 95.53% of all articles (1091 out of 1142). The range of the grid size is from 0.25 m $\sim$ 2 km. The grid sizes of 30 m, 10 m, 20 m, 25 m, and 5 m were the most commonly used, accounting for 26.12%, 21.36%, 13.29%, 11.09%, and 9.90% of the 1091 articles, respectively. However, the other mapping units were used less frequently, with slope units, unique condition units, and watershed units used in 2.80%, 1.66%, and 0.61% of the articles, respectively.

Mapping units were produced from digital elevation data with different spatial resolutions. We extracted the spatial resolution of the digital elevation data used in 1142 articles for further analysis. Of the 1142 articles, 1096
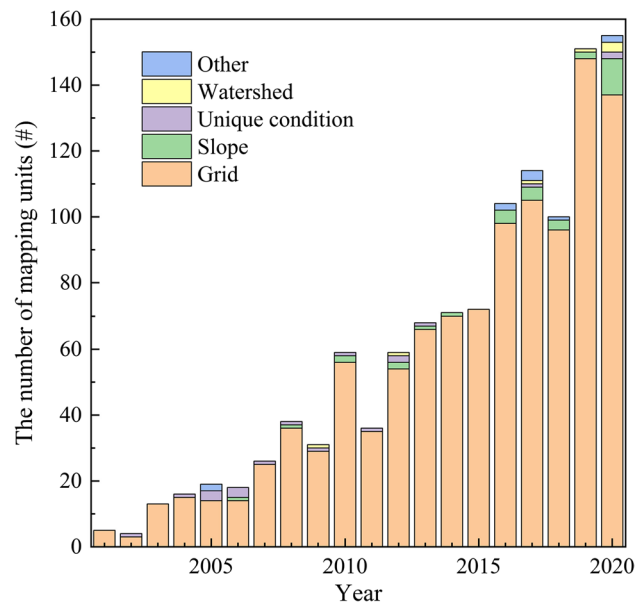
Fig. 13 Mapping units used in the 1142 articles

articles used a single spatial resolution, 12 articles used two or three spatial resolutions, and the remaining did not describe the spatial resolution. The range of the resolution is from 0.25 m $\times$ 0.25 m to 1 km $\times$ 1 km, and the resolution data of < 30 m $\times$ 30 m are used 1007 times. Thirty-metre digital elevation data are most frequently applied, with 26.62% of the 1142 articles, followed by resolutions of

10 m (19.70%), 20 m (12.35%), 25 m (10.51%), and 5 m (10.51%). These resolutions are basically consistent with the grid units.

Susceptibility models

Based on the literature statistics, a total of 2220 models were used in 1142 articles. A total of 27.88%, 20.36%, 20.81%, and 16.04% of the publications used one, two, three, and four model types, respectively. Based on the 2220 model names as given by the authors and their algorithm principles, we identified the repeated or combined use of 429 different model types. We reclassified these model types into six groups based on an algorithm or model principles as follows: qualitative models (28 types), bivariate statistical models (39 types), multivariate statistical models (29 types), machine learning models (96 types), deterministic models (19 types), and hybrid models (218 types). A hybrid model

is defined as the combination of two or more methods (e.g. the combination of machine learning and ensemble learning) (Huang et al. 2022b). Figure 14 shows the six groups of model types and the top 13 model types used from 2001 to 2020. The results show that bivariate statistical models were used most often (562 times), accounting for 25.32% of the model use from 2001 to 2020. Machine learning models (553 times, 24.91%), multivariate statistical models (389 times, 17.52%), hybrid models (331 times, 14.91%), and qualitative models (312 times, 14.05%) were the next most common, and deterministic models were only used 73 times (3.29%). The most frequently used models in the three periods were multivariate statistical models, bivariate statistical models, and machine learning models, respectively. Increasing attention has been given to hybrid models, which increased from 6 times in the first period to 32 times in the second period and 293 times in the third period. Similarly, machine learning models increased from 42 to 75 times and
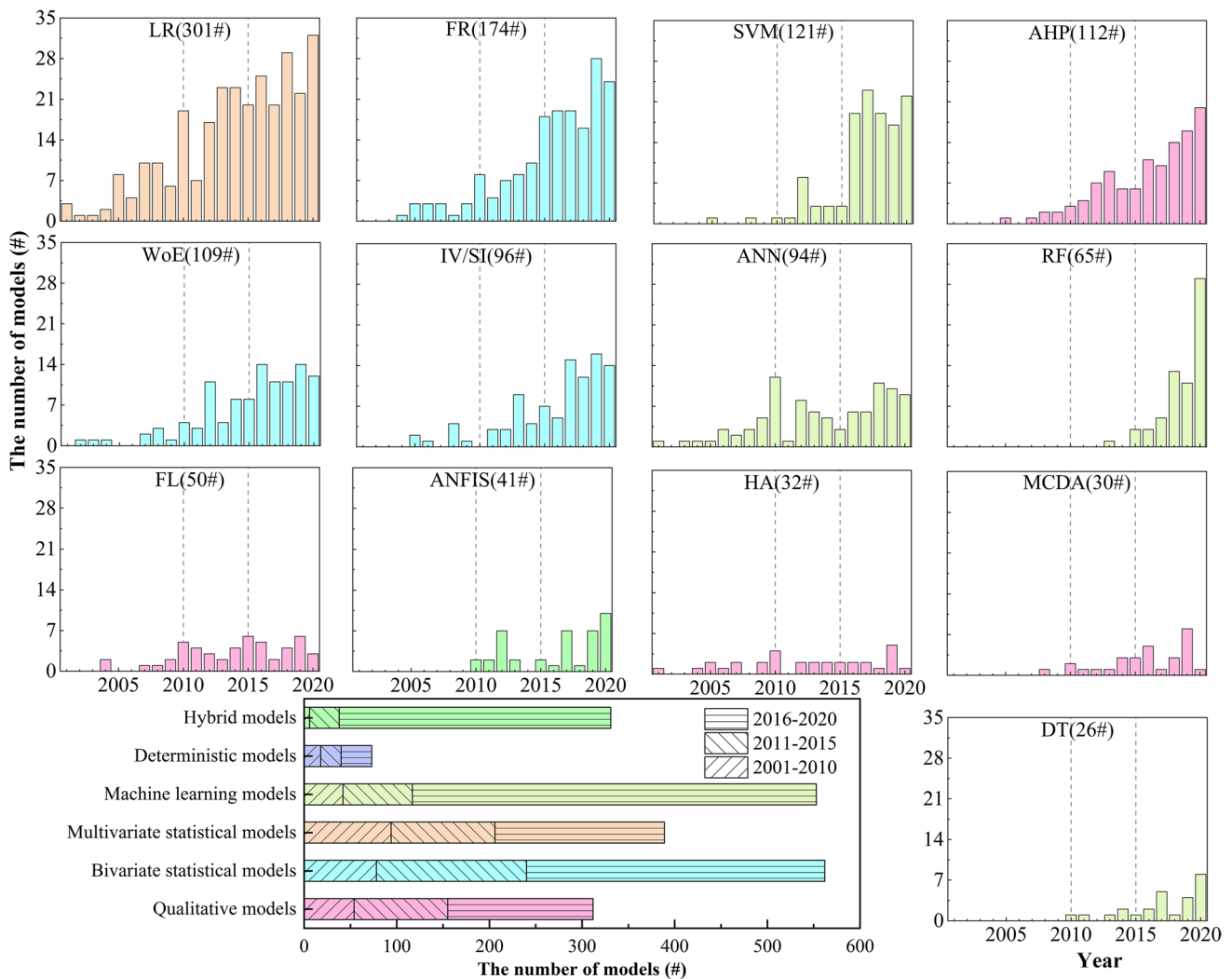

Fig. 14 LS model used in publications from 2001 to 2020

then to 436 times during the 2001–2010, 2011–2015, and 2016–2020 periods, respectively.

Among the top 13 model types in the 1142 publications, LR, one type of multivariate statistical model, was used most frequently, specifically 301 times (26.36%). Bivariate statistical models such as FR (174 times, 15.24%), WoE (109 times, 9.54%), and IV/SI (96 times, 8.41%) were frequently used from 2011 to 2020. Machine learning models such as SVM (121 times, 10.60%) and ANN (94 times, 8.23%) were more common from 2001 to 2020, and RF and DT became increasingly popular in the third period of 2016–2020. ANFIS, as a hybrid model, was used frequently in the 2010–2020 period.

Furthermore, the proportion of different factors and the top 10 most-used factors in the five study scales are shown in Fig. 15. The results indicated that deterministic models (e.g. TRIGRS (Transient Rainfall Infiltration and Grid-based Regional Slope-stability) and SINMAP) are widely applied, and hybrid models are the least used in the detail scale (0 ~ 10 km^2). Bivariate statistical models (e.g. FR, WoE, and IV) and multivariate statistical (e.g. LR) models are commonly used at the medium scale (10 ~ 100 km^2). Machine learning and bivariate statistical models are used more often in dealing with large (100 ~ 1000 km^2) and regional (1000 ~ 10,000 km^2) scales. At the national scale, machine learning and hybrid models are frequently applied.

Model performance evaluation

To verify the robustness of the constructed LS models, the geospatial data (landslide inventories and conditioning factors) are usually divided into training and testing datasets or training, validation and testing datasets using holdout split, k-fold cross-validation, and bootstrap sampling. These three datasets have different functions; for example, the training dataset is used to model development and model fit performance, the validation dataset is used to optimise the parameters of the learning algorithms, and the testing dataset is used to obtain the model prediction performance. An analysis of the literature database revealed that 57.97% of all articles divided the input data using a holdout strategy, 1.93% a k-fold cross-validation, 0.44% a bootstrap sampling, 0.18% a combination of the three sample subdividing strategies, and the remaining no divided the data. For the popular holdout split, temporal, spatial, and random sampling techniques have been commonly applied. Temporal validation denotes that the landslide events are divided into training and testing datasets based on time information. When adopting a spatial validation, the landslide dataset was geographically divided into two region groups. In the random sampling, the landslide dataset is segmented based on the proportion between the training and testing samplings. Of the 662 articles that described the model performance validation using the holdout strategy, 565, 52, 31, and 14 articles adopted the
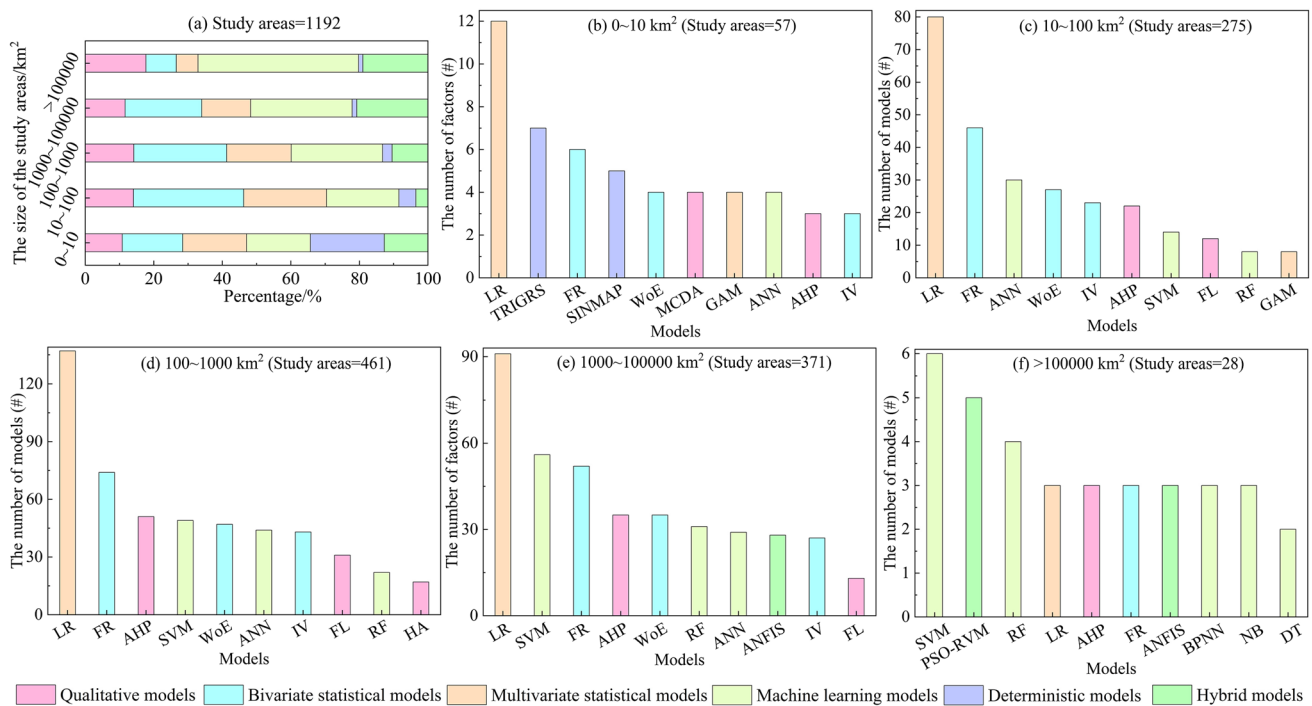

Fig. 15 LS models in different size study areas. GAM, generalised additive model; PSO-RVM, particle swarm optimisation-relevance vector machine; BPNN, back propagation neural network

random selection, temporal selection, spatial selection, and other techniques (e.g. the combination of split techniques, or Mahalanobis distance), respectively. We found that 563 articles separated the landslide inventory into two groups as training and testing datasets based on a certain ratio by means of a random selection technique and only 9 articles randomly divided the landslide inventory into training, validation, and testing datasets. Regarding the ratio between the training and testing samples, a total of 550 publications used one ratio, and 11 publications used two or more ratios, with the maximum number being 11 ratios. We rearranged the training and testing sample ratios and found that the most common training–testing ratio was 70/30 (287 times), followed by 80/20 (75 times), 50/50 (67 times), 75/25 (58 times), 60/40 (20 times), and 90/10 (19 times).

In addition, different metrics were employed in the articles to evaluate the model fit and the model prediction performance. We found that 33.71% of all articles used one metric to measure the model fit and prediction performance, 56.48% used two or more metrics, and 9.81% did not use a metric. To assess the model fit performance, approximately 40.37% (461 publications) of the 1142 publications used only one metric, 10.16% (116 publications) used two metrics, and 15.50% (177 publications) used more than two metrics, with a maximum of 14 metrics (Guzzetti et al. 2006). The remaining 388 publications (33.98%) did not perform any assessment of the model fit performance. A total of 1603 metrics were used for model fit in

754 publications, which can distinguish 93 unique metrics based on the author's description. According to the purpose and previous studies (Reichenbach et al. 2018), these unique metrics were arranged into six groups: probability, class label, regression, landslide density, signification test, and others (Fig. 16a). We found that the metrics of probability performance were commonly used in the articles, with 37.87% of 1603 metrics. The output result of the model is a format of probabilities ranging between 0 and 1, and the probabilistic outputs can be used as class predictions. The success rate curve (35.73% of 1142 articles) is the most widely applied metric in classification prediction, followed by ROC (17.16%). The model output category value is discrete, which is represented as a class label (landslide or nonlandslide). The metrics of class label performance are the second most used, which include accuracy (9.81%), sensitivity (8.67%), specificity (6.04%), precision (5.60%), and kappa index (5.17%). Root mean square error (4.47% of 1142 articles), −2Log likelihood (2.19%), Nagelkerke R^2 (1.93%), Cox and Snell R^2 (1.58%), mean square error (1.14%), and mean absolute error (0.79%) were commonly used regression problem metrics. For the groups of landslide density, landslide density/percentage is the most used in 10.77% of articles, followed by frequency ratio plots (1.40%). In addition, significance test metrics such as the chi-squared test (2.98% of 1142 articles) can compare the significance difference of the proposed models using the training data, which also reveals the model performance.
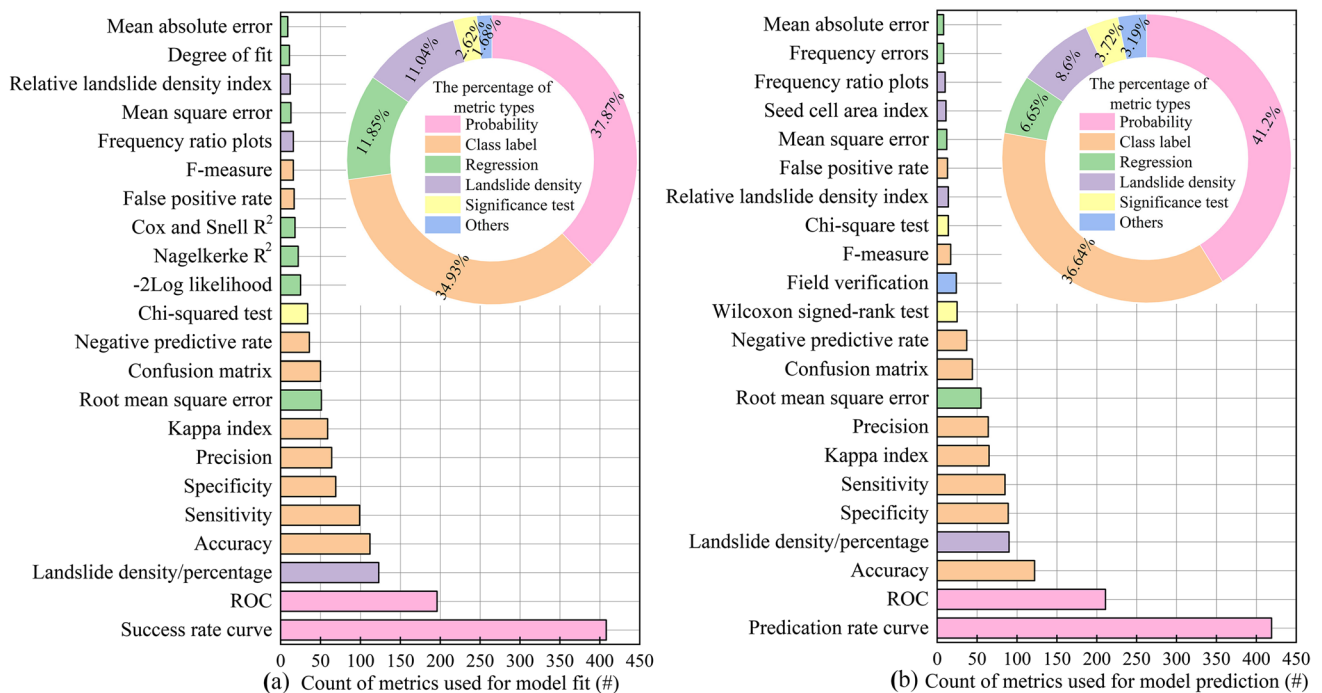

Fig. 16 The top 22 metrics used for model fit and model prediction. The two donuts illustrate the type of metrics used for model fit (left) and model prediction (right), respectively

Similarly, the results of the metrics adopted to assess the model prediction performance revealed that 738 articles (64.62%) used 1534 metrics to validate model prediction performance, and the remaining (404, 35.38% of all articles) did not perform any model verification. We identified 74 unique metrics, which were also reclassified into six classes (Fig. 16b). The results indicated that the most common were the predication rate curve (36.69% of 1142 articles), ROC (18.48%), accuracy (10.68%), landslide density/percentage (7.88%), specificity (7.79%), sensitivity (7.44%), kappa index (5.69%), precision (5.60), and root mean square error (4.82%).

Discussion

Popular research topics and development trends

Trend of the study sites

Both the keywords and study site statistical analysis indicated that the variation in the study sites and study scale were basically similar from 2001 to 2020; however, the focus of the research site has shifted from developed countries (South Korea, Italy, and Japan) in the period 2001–2010 to a few developing and mountainous countries (China, Iran, India, and Malaysia) in the period 2011–2020. One potential reason for the shift in the study sites may be that landslides occur frequently in mountainous countries. Infrastructure construction in developed countries is relatively mature, which indicates that engineering construction is basically completed. However, infrastructure construction is ongoing in many developing countries. Therefore, LS is needed to guide engineering construction and management in developing countries. Moreover, these developing countries are increasingly focusing on environmental protection and have sufficient funding to research and manage the LS field. For example, the National Natural Science Foundation of China supported the highest number of articles (149), which accounted for 13.05% of all 1142 articles. China is at the centre of this research field with the greatest number of publications and research sites and has launched international cooperation with most countries (Figs. 5, 6, and 8). Many institutions (Figs. 3 and 4), such as the Chinese Academy of Sciences, China University of Geosciences and Chengdu University of Technology, have specialised departments to carry out research on landslide prevention and mitigation and have trained many engineering geological experts, which has promoted the development of this field. Certainly, India, and Iran have also carried out much research in this field. In addition, the number of LS

assessments remains low in Africa, South America, and Oceania, which reveals a significant geographical bias. Therefore, we recommend that researchers devote more attention to Africa, South America, and Oceania.

Landslide inventory

Landslide inventory is the basic data used to determine landslide susceptibility, hazard, vulnerability, and risk (Jimenez-Peralvarez et al. 2017). Generally, a complete landslide inventory map contains the location, type, volume, activity, main causes, and occurrence data of landslides. Landslide inventory can fall into archives or geomorphological (e.g. historical, event, seasonal, or multitemporal) inventories according to the mapping type (Guzzetti et al. 2012). A detailed definition of these inventories can be found in Guzzetti et al. (2012). Analysis of the literature database reveals that 1111 out of 1198 study areas used one inventory, which are mainly historical inventories with landslide age not differentiated (e.g. Jaafari et al. (2015); Chen et al. (2018b)) or event inventories with landslides induced by a single event such as an earthquake, rainfall, and snowmelt (e.g. Reichenbach et al. (2014); Dou et al. (2020); Ling & Chigira (2020)). However, seasonal and multitemporal landslide inventories in which landslides are triggered by multiple events over longer periods have rarely been analysed in the literature. This is because analysing seasonal and multitemporal inventories is difficult and time-consuming, requiring abundant resources (Guzzetti et al. 2006; Reichenbach et al. 2018).

(1) **Source of landslide inventory** In the literature database, landslide inventories are produced by the following techniques: historical reports, field surveys, and remote sensing analysis. This finding is basically consistent with previous studies (Reichenbach et al. 2018; Zou & Zheng 2022). Landslide information from historical records and literature can be used to prepare an archive inventory, but it lacks sufficient data to support regional landslide mapping. Therefore, it is more suitable for small-scale inventories (< 1:200,000). Through the field investigation of landslide characteristics with the assistance of a global positioning system and drilling technology, detailed landslide information (e.g. type, volume, scarp, accumulation of landslides) was delineated on the topographic map. Field surveys are the best method for preparing large-scale inventories (> 1:25,000). However, they have limitations in mapping a large number of landslides or old landslides, especially in mountainous areas where access is difficult or even impossible (Choi et al. 2012). For instance, landslides are partially or totally covered by forests, which makes it difficult to determine the landslide boundary in the field (Guzzetti et al. 2012). Emerging technology based on remote sensing can quickly and contactlessly obtain surface displacement

information in the area covered by remote sensing images, which improves the efficiency and quality of landslide identification. In the literature database, the data sources of remote sensing technology mainly include aerial photographs, satellite images, unmanned aerial vehicles (UAVs), interferometric synthetic aperture radar (InSAR), light detection and ranging (LiDAR), and high-resolution digital elevation models (DEMs). The manual interpretation of aerial photographs and high-resolution satellite images is used to identify landslides according to the shape, size, colour, tone, texture, and site topography of geomorphic surfaces, which is the most used landslide detection technology (Suzen & Doyuran 2004; Lee et al. 2018). The visual interpretation of aerial photography is intuitive and low cost but is usually based on expert experience with uncertainty. A new generation of high-resolution remote sensing images, such as IKONOS (Youssef et al. 2015), Quickbird (Sharma et al. 2011), SPOT 5 (Pradhan & Buchroithner 2010), and World View, offers resolutions ranging from 0.5 to 2 m, which provides a very powerful tool for the quick mapping of regional landslide distributions (Rosi et al. 2018). Recently, Google Earth images have been used to draw landslides due to their easy availability and high resolution (Pourghasemi & Rahmati 2018; Nhu et al. 2020; Pandey et al. 2020). UAV technology is not controlled by the terrain and can flexibly and conveniently obtain high-precision surface information, which can compensate for the limitation of shadows generated by satellite images in alpine gorges. The application of InSAR provides a high precision and allows the coverage of large areas, which contributes to the creation and updating of landslide inventory maps (Rosi et al. 2018). We found that differential InSAR (Pareek et al. 2013), small baseline subset InSAR (Xie et al. 2017), and persistent scatter interferometry (Huang et al. 2020b) have produced a few successful case studies dealing with landslide detection, but these technologies were less common in the literature databases. The limitations of InSAR are that the quality of the results depends on optimal conditions concerning slope orientation, only slow deformation can be detected, and the temporal resolution of historical SAR images is limited (Jimenez-Peralvarez et al. 2017). LiDAR can penetrate through vegetative cover to capture the surface terrain changes and generate a high-resolution DEM, although it is expensive; furthermore, it is difficult to manipulate the point cloud. More LiDAR-derived products, such as shaded relief, slope, and contour line maps, can be calculated by a high-resolution DEM in a GIS environment, which improves the interpretation of landslides (Hess et al. 2017). Therefore, many studies have used LiDAR to generate landslide inventories (e.g. Chen et al. (2013); Jebur et al. (2014); Dou et al. (2019)).

With the rapid development of remote sensing and GIS technologies, remote sensing data are becoming increasingly abundant. The application of remote sensing technology to interpret landslides has the following two trends:

(i) The information source is from a single remote sensing image to multitemporal and multisource images. A single image can prepare a historical or an event inventory, whereas seasonal and multitemporal inventories can be obtained using multiple sets of images on different dates (Guzzetti et al. 2012). In the literature database, we found that some studies have used multisource satellite images to map landslides (e.g. Bui et al. (2018); Wu et al. (2020)), but new technologies, such as spatial analysis of DEM and 3D visualisation of images, are less applied in this process.

(ii) The method to extract landslide information is changed from visual interpretation to semiautomatic/automatic interpretation (Zou & Zheng 2022). Visual interpretation of landslides is still popular in the literature database (e.g. Tanoli et al. (2017); Yi et al. (2020)), but it is time-consuming and subjective. The automatic method is based on the change in morphometric data (e.g. slope angle, surface roughness, semi-variance, and fractal dimension) and spectral parameters (e.g. normalised difference vegetation index (NDVI), spectral angle, principal or independent components) using index thresholding (Shou & Yang 2015), change detection methods (Xu et al. 2013), and machine or deep learning algorithms (Gorsevski et al. 2016) to provide rapid landslide inventory. The automatic method has the advantages of objectivity and high efficiency. Therefore, we recommend the application of an automatic method to regularly update the landslide inventory, which can provide unique information on the spatial and temporal evolution of landslides.

(2) Sampling strategies An inspection of the literature database reveals that most of the sample strategies representing landslides are points (55.43%) and polygons (41.77%), followed by seed cells (1.93%) and circles (0.61%). This indicates that the sample strategies are not consistent in the landslide inventory. The first reason is that multiple sources provide very heterogenic information about landslides (Canavesi et al. 2020). Second, a standard and recognised operating procedure is not introduced to prepare landslide maps (Guzzetti et al. 2012). Last, identifying the location on the hillslope where the factor triggering the landslides exceeds the stability threshold and makes the hillslope unstable is a difficult problem (Regmi et al. 2014). Landslides may occur anywhere on hillslopes (e.g. hillslope toe, landslide centroid) and then can extend downslope, sidewise,

and upslope from the initial location. If a large region has abundant landslides, it would be more difficult to determine the locations of initiations through remote sensing images or field surveys (Wang et al. 2014). However, while the locations of landslide initiations are not accurately determined, a significant uncertainty can be introduced (Regmi et al. 2014). Therefore, adopting an appropriate sampling technique is necessary to minimise the uncertainty and generate a reliable result (Bordoni et al. 2020).

In the point-based sampling strategy, there seems to be a consensus on representing the area of landslides by the pixel unit at the centroid locations of landslides. This location is easy to obtain in the GIS environment, which only needs to convert landslides mapped as polygons into centroids. In addition, some studies used data from centroids of the landslide scarps or detachment zones and the highest locations of landslides (Park & Chi 2008). For polygon sampling, inventories of complete landslide areas (from scarps to accumulation zones) are far more extensive. This may be due to the difficulty in dividing the accumulation zones and the depletion zones. In some studies, only the scarps, the depletion zones, the detachment zones (Das et al. 2012), or the deposit areas (Mahalingam et al. 2016) are analysed to determine LS zones. What is the best sampling strategy? Although there is no agreement on the sampling strategy, we can develop insights from previous studies. For instance, some studies (Simon et al. 2017; Zezere et al. 2017; Pourghasemi et al. 2020) show that the centroid of the landslide rupture zone performs best, followed by the landslide centroid and landslide polygon in sampling strategies. Under the condition that landslides are small in size, a single point per landslide located in the centroid of the landslide rupture zone or in the landslide centroid can minimise possible heterogeneity of inducing factors within the landslide boundary, which is sufficient to produce a better prediction result (Zezere et al. 2017). According to Regmi et al. (2014), when the sizes of landslides are large, landslide scarps perform better than point sampling in generating reliable results. Similarly, the comparative result of the effect of different sampling strategies from Yilmaz (2010) showed that the scarp is better than the seed cell and point sampling strategies. In addition, Conoscenti et al. (2008) found that LS assessment using scarps instead of landslide areas for rotational slides generates better prediction results, whereas using areas uphill from crown instead of scarps for flow slide landslides can obtain a better result. In conclusion, the sampling strategies of landslide scarps or rupture zones may obtain a better prediction result because they are the most diagnostically unstable landforms.

Therefore, we recommend that landslide scarps or rupture zones as far away as possible be used to sample landslide data. The selection of sampling strategies can be guided by the following conditions: (i) the landslide types, (ii) the

positional accuracy of landslides; (iii) the scale of the study area; and (iv) the limitations of software and hardware.

(3) Sample subdividing strategy The number of landslides in a landslide inventory usually ranges from 1 to 250, with the areas of the study sites being the nonlinear correlation in the 1142 publications. A total of 60.51% of all articles used a sample subdividing strategy to divide the landslide dataset into training and testing sets. The different samples input into the model have different results, and the quantity of landslide samples available and sample subdividing strategy determine the accuracy of the LS mapping. For example, Maxwell et al. (2020) summarised that increasing the training sample size tended to improve the model accuracy, with the largest improvement at lower sample sizes (1 ~ 250); however, this benefit diminishes as the sample size increases. Random selection is the most frequently used sample subdividing strategy. It allows different ratios to be easily calculated using Hawth's analysis tool for ArcGIS. A random sample cannot guarantee selecting the best subset to train a model. Compared to the whole samples, the distribution of the training samples should not change much. The inconsistent results are mainly due to the differences in the training samples created by a random process, without investigating the attributes of the created samples (Sameen et al. 2020a). To improve this method, a stratified random sample is applied to each data layer. Landslide/nonlandslide training and testing datasets were constructed by this approach to allow all classes of variables to be represented with the same proportion in these constructed datasets (Eker et al. 2015). This can alleviate the problem that excessive amounts of samples would lead to overfitting of the model. When the sizes of the available original samples are limited, stratified random sampling can augment an existing sample to reduce the standard errors of the regional estimates, thus improving the accuracy of the results (Stehman et al. 2011). However, stratified random sampling is rarely studied in the literature database. Some studies attempt to search the best representative dataset by adopting other methods, such as cross-validation, bootstrap, and Mahalanobis distance (Wan 2009; Goetz et al. 2011; Kornejady et al. 2017; Chen et al. 2018a). For instance, when samples come from the same slide event, multiple points within the same polygon could be grouped into the same partition so as not to bias the assessment. Cross-validation was used to randomly partition the landslide dataset into k subsets. One of the subsets is used as testing data, whereas the remaining subsets are used to train the model and optimise the parameters. It is considered the gold standard for machine learning with many advantages, such as effectively reducing the randomness of the train-test split, better evaluating the modelling methods with limited data, and generating a less biased model.

In the literature database, the commonly used training–testing ratios are ordered from more to less: 70/30, 80/20, 50/50, 75/25, 60/40, and 90/10. What is the best ratio? Many studies directly choose a ratio without any proper explanation to predict landslides and show that the results are satisfactory (Luo et al. 2019; Akinci et al. 2020). However, there is no evidence in these studies that the results can be attributed to the subdividing strategy, or the results would not be optimised under alternative strategies (Pourghasemi et al. 2020). Some studies considered different training/testing ratios for landslide modelling to obtain a better division (Mancini et al. 2010; Paulin et al. 2010). For instance, Jaafari et al. (2019) and Sahin et al. (2020) compared the effect of nine sampling ratios (from 10/90 to 90/10, with intervals of 10%) on the accuracy of prediction models and found that the best results were achieved with a ratio of 70/30, and the lowest performance was achieved with a 90/10 ratio. However, Shirzadi et al. (2018) investigated four different sample sizes and different raster resolutions for preparing LS mapping, and the results revealed that the random subspace algorithm obtained the highest prediction accuracy with sample ratios of 60/40 and 70/30 and a raster resolution of 10 m, whereas the multiboost ensemble algorithm achieved the highest prediction accuracy with ratios of 80/20 and 90/10 and a raster resolution of 20 m. Moreover, Vakhshoori & Zare (2018) considered three ratios (50/50, 60/40, 70/30) for landslide prediction and found that none of the ratios meaningfully reduces/increases the validity of the models. Therefore, there is no agreement on which ratio produces better results. However, we noticed that a small training dataset may not cover the spatial variability of the conditioning factors. In contrast, a large training dataset will more likely violate the independent observation assumptions because of spatial autocorrelation (Heckmann et al. 2014).

Therefore, the selection of the sample subdividing strategy must consider (i) the availability of landslide data, (ii) the model type, (iii) the raster resolution, and (iv) the characteristics of the study area. We recommend that when the training dataset is limited, cross-validation and bootstrap subdividing strategies should be used; and when the training dataset is large, stratified random sampling should be considered.

Conditioning factors and their trends

Of the 476 conditioning factors, the use trends of slope, lithology, aspect, elevation, land use/land cover, distance to rivers, distance to roads, distance to faults, and rainfall in the three periods steadily increased with the number of articles (Fig. 11). These findings are basically consistent with Pourghasemi et al. (2018), Reichenbach et al. (2018), and Lee (2019). Topographical factors are the most frequently

used in the literature database because these parameters describing the geometric characteristics of hillslope landforms have an important impact on the occurrence of landslides and are easily extracted from DEMs in GIS. Maxwell et al. (2020) revealed that incorporating measures of lithology, soils, and distance to roads and streams did not improve the model performance in comparison to just using 14 topographical factors, which highlighted the value of topographical factors. Therefore, in remote mountainous areas, if geological factors are limited, DEM-derived terrain variables should be used to predict landslides in the preliminary study. In the literature, the range of the resolution of the DEM is from 0.25 m × 0.25 m to 1 km × 1 km, and the commonly used resolutions are 10–30 m. Digital elevation data with a spatial resolution of 30 m are most frequently applied in the literature and are generally available over large spatial extents, such as Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model version 3, NASA Shuttle Radar Topography Mission version 3, and Advanced Land Observing Satellite World 3D-30 m. The resolution of DEMs used for LS modelling has increased significantly in recent years, with most of the articles that use DEM resolutions better than 30 m × 30 m. Very high resolution DEMs (< 5 m) derived from the LiDAR, InSAR, and UAV techniques are becoming popular because they are easy to acquire and provide a high level of detail. The source and spatial resolution of DEM files directly affect the quality of the final predictive maps. For instance, Kaminski (2020) found that a 20-m LiDAR DEM obtained a better model accuracy than a 20-m Land Parcel Identification System DEM. Some studies (Jaafari 2018; Arabameri et al. 2019) demonstrated improved precision of results when high-resolution DEMs were utilised, while other studies (Mahalingam & Olsen 2016; Chang et al. 2019) reported that 30 m DEMs provided better LS mapping than their finer counterparts. Due to the scale effects of DEM-derived topographic variables, it remains a challenging task to select an optimal DEM resolution. A coarser DEM has a lower accuracy in describing the terrain, and the secondary derivatives of DEM (such as slope, aspect, and curvature) are highly dependent on resolution. However, very high-resolution DEMs describe the terrain variations at the microscale, which are unlikely to be related to mesoscale processes that induce landslides (Merghadi et al. 2020). Therefore, the appropriate choice of DEM is dependent on the availability of data, the study region area, and the topographic characteristics of the study region.

Geological factors, such as lithology and distance to faults and lineaments, are frequently used in the literature. Generally, these geological data were produced from geological maps with different scales. If the geological boundaries determined by the mesoscale and small-scale geological maps are inaccurate, it would lead to error

in the result. Therefore, field verification of geological boundaries is an essential step. The lithology of bedrock is usually described in the geological map, while the surface sediments are rarely mapped into the lithology. However, some rainfall- or earthquake-inducing landslides often occur in these surface sediments. Therefore, LS assessment does not consider the surface sediments, resulting in the uncertainty for the result.

Land cover mainly includes vegetation-, soil-, and land use-related factors. Among them, land use/land cover is the most used, which can be obtained from the visual or automatic interpretation of satellite images. Land use/land cover is a static variable in the short term and a dynamic variable in the long term. Reichenbach et al. (2014) evaluated and quantified the effect of land use change from 1954 to 2009 on LS zoning and found that an increase in unstable zones over a period of 56 years due to the expansion of bare soil and forested areas was destroyed. NDVI quantitatively describes the physical characteristics of the vegetation. We found that the application of NDVI rapidly increased from 2016 to 2020. This specific trend is due to the increased availability of remote sensing-derived products (e.g. NASA's Landsat satellites and ESA's Sentinel-2 satellites) (Pourghasemi et al. 2018). Many soil-related factors (soil type, soil depth, soil cohesion, and internal friction angle) are used in the detail scale ($0 \sim 10$ km^2) and medium scale ($10 \sim 100$ km^2) because these parameters need to be input into the deterministic models. The consistency of soil data also affects the use of soil at different study scales. For example, in the USA, detailed soil data were generated at the county level and, therefore, were not consistent between countries.

For hydrological factors, the application of distance to rivers, topographic wetness index (TWI), and SPI increased rapidly from 2016–2020. These three factors were also calculated by adopting the DEMs and certain algorithms in GIS. The river erodes and softens the hillslope toe for a long time, thus disturbing the stability of the slopes. When using distance to rivers, the distance that the river affects the instability of hillslope should be determined.

Recently, the use of environmental factors (e.g. rainfall and PGA) has shown an increasing trend. A rainfall map was constructed using the kriging spatial interpolation method in the GIS environment based on rainfall data (Huang et al. 2022a). The accuracy of rainfall depends on the number of available weather stations and the size of the mapping units. The use of rainfall increases with increasing study scale. There may be no weather station at the detailed and medium scales, and the rainfall value obtained by interpolation cannot describe the change at the microscale. However, regional and national scales contain many weather stations and usually use mapping units with grids larger than 100 m; thus, the change in

rainfall values on a small scale does not seem important. For the same reason, PGA, one of the main indicators of an earthquake, is often used in national-scale research. In addition, the keyword analysis also indicated that rainfall and earthquakes are research hotspots. Rainfall and earthquakes are the main triggering factors of landslides. The susceptibility mapping of rainfall and earthquake landslides has always been a research hotspot (Kamp et al. 2008; Zhao et al. 2019).

Significantly, input conditioning factors have some key problems in susceptibility assessment as follows: (i) The scale effects of landslide conditioning factors are due to diverse sources of factor data being used at a broad range of scales (from 1:5000 to 1:1,000,000). Ideally, the input data must have the same resolution and quality, (ii) a lack of geographic consistency of the different conditioning factors, and (iii) more importantly, there is no standard rule or global agreement that can be correctly adopted to select conditioning factors and number of factors. The selection of landslide conditioning factors has a significant impact on the performance of final predictive maps. The analysis of the literature database found that 430 different factors were identified in all articles and 1 to 34 factors were used per article. However, it is impossible to collect all these factors and to apply them in a single LS assessment. In addition, the use of a number of conditioning factors in a model increases the cost of data collection and the processing time, introduces noise, and possibly reduces prediction accuracy (Pradhan & Lee 2010b; Zhou et al. 2018). Generally, the selection of landslide conditioning factors must take into account (i) the availability of data, (ii) the scale of the study area, (iii) the characteristics of the study area, (iv) the landslide type, and (iv) the genetic mechanisms of landslides (van Westen et al. 2008). However, this method involves subjective experience analysis of experts and is a qualitative method. Recently, a quantitative feature selection method (e.g. correlation attribute evaluation, chi-square attribute evaluation, information gain, gain ratio, One-R classifier, Relief-F algorithm, subset evaluators, symmetrical uncertainty, and random forest) was adopted to rank the weight of the input factors and eliminate less important factors (Chen et al. 2018a; Bui et al. 2020; Huang et al. 2022b). These techniques can find an effective combination of geoenvironmental factors and determine the number of effective factors. For instance, Tang et al. (2020) used principal component analysis to select the relatively important factors, and the results revealed that the model accuracy that uses 9 factors with relatively high weights was better than that obtained from 11 factors. Therefore, we recommend identifying and selecting all conditioning factors available for the study area in the preliminary study. Then, a quantitative feature selection method can be used to eliminate irrelevant or less important factors to obtain highly accurate results.

The advantages and limitations of mapping units

Analysis of the 1142 articles revealed that grid units (95.53% of 1142), slope units (2.80%), unique condition units (1.66%), and watershed units (0.61%) were the common types of mapping units. Each unit has certain advantages and limitations. The grid unit subdivides the entire area into regular squares of predetermined dimensions and then assigns a value to each grid to represent different factors (Liu et al. 2020). Its data structure is in matrix form, which is convenient for storage and calculation. The grid unit is by far the most used mapping unit because it is easy to subdivide the landscape at all resolutions and geographical scales (Reichenbach et al. 2018). The resolutions of the grid units are from 0.25 to 2 km in the literature, of which 5 ~ 30 m was the most common. These resolutions often coincide with DEM data. Some studies have explored the effects of different grid resolutions on the performance of landslide modelling (Paulin et al. 2010; Cama et al. 2016; Shirzadi et al. 2019). The choice of grid size is an issue. Small grids can capture the micromorphological characteristics of small shallow landslides in great detail but reduce the computational efficiency. However, grids that are too large smooth the local terrain details, which may be more suitable to describe large deep landslides (Reichenbach et al. 2018). In addition, grid units divide the landscape with a fine to very fine resolution, but there is no physical relationship between grid units and landslides.

The slope unit divides the land surface into thousands of independent slopes by ridge and valley lines, which is the basic terrain unit of landslide occurrence. Slope units reflect the physical relationships between landslides and morphological features and accurately represent the geological and geomorphological characteristics of the study area. Currently, many slope unit division methods are proposed, such as manual division, GIS-based hydrological, and curvature watershed methods (Huang et al. 2021). However, there are a few limitations for slope units: (i) The division process is complicated, time consuming, and error prone, and (ii) the results of different division methods are different and subject to great subjective influence. Some studies have demonstrated that the assessment result based on the slope unit is better than that based on the grid unit (Ba et al. 2018; Yu & Chen 2020).

Unique condition units are obtained by overlaying all landslide conditioning factors; thus, every unit depends on the combination of different properties. The size and total number of units are determined by the number of and the classification criteria of influencing factors, respectively (Guzzetti et al. 1999). This unit reflects geomorphological and geological differences. The main advantages of unique condition units are their conceptual simplicity and ease of numeric computation (Federici et al. 2007). However, it has some limitations in LS mapping; for example, the continuous factors must be classified before overlaying, which is a subjective procedure. In addition, when intersecting all the geo-environmental layers, many small topographic units will be generated (Reichenbach et al. 2018), which can result in errors in the results. Therefore, a large spatial extent (regional and national scales) is not applicable for these units.

Watershed units, derived from DEM files in GIS, partition the whole region into independent small watershed units based on ridge lines. This unit is more suitable for the evaluation of debris flows but less suitable for landslides. It represents the actual geomorphic features. Compared with grid units, the watershed unit can achieve better prediction results in debris flow susceptibility mapping (Qin et al. 2019).

Therefore, we suggest selecting appropriate topographic units according to the following conditions: landslide type, spatial extent of the study area, number and quality of factors, and computing power of computer hardware and software.

Trend of susceptibility models

The results of the keyword analysis were consistent with those of the statistical analysis of the models, which have always been the key themes of landslide prediction (Fig. 7). Similar to the results of statistical studies by Pourghasemi et al. (2018), Reichenbach et al. (2018), and Lee (2019), we found that conventional LS models such as the LR, FR, SVM, AHP, IV/SI, and ANN were frequently used. Among the 429 different models in the 1142 articles, 11 models were used more than 30 times, and 286 models were used only once. This finding indicates that researchers have a largely unjustified and excessive interest in certain models but could not reach a consensus on the most suitable and efficient techniques or methods for all regions until now (Reichenbach et al. 2018; Chen et al. 2020). LS models can be qualitative and quantitative or direct and indirect. In this study, we divided the models for LS into qualitative models, deterministic models, bivariate statistical models, multivariate statistical models, machine learning models, and hybrid models. Every susceptibility model has its own advantages and disadvantages. The selection of an appropriate model in a given region often remains a challenging task.

The qualitative models are characterised by expert opinions and judgements about terrain conditions to determine the susceptible levels. The designating weights of factors are highly subjective, and the results of qualitative models vary depending on the knowledge of experts. AHP is the most used quantitative method in the literature database. The synthesis ideology of AHP is to compare pairs of decision factor hierarchies to assign the weights and examine

the consistency ratio (Althuwaynee et al. 2014). The merits of using AHP in LS assessment are as follows: (i) all types of information related to landslides can be included in the discussion (Kayastha et al. 2013); (ii) it is well suited to make collective judgements in complex research; (iii) it does not require a sufficient number of landslides; and (iv) it is often useful for studies at large, regional, and national scales (Fig. 15).

Deterministic models are used the least in the literature database. The proportion of deterministic models is the highest at the detailed scale ($0 \sim 10$ km^2), and its proportion decreases with the increase in the study scales. These models are quite complex and not applicable for large-scale areas. The deterministic method is a methodology that combines the hydrological model and infinite slope-stability models to estimate the potential or relative instability of hillslopes (Pradhan & Kim 2016). Hydrological models can analyse the pore-water pressure, and infinite slope-stability models can compute the safety factors. Therefore, deterministic models require a large quantity of geotechnical and hydrological input data, such as cohesion and internal friction angle. In the process of obtaining geotechnical and hydrological parameters, uncertainty arises due to the spatial variability of the surface and complex geological conditions, which increases the possibility of obtaining erroneous analysis results (Lee et al. 2020). This also limits its scope of application, and it is only applicable when ground conditions are fairly homogenous, landslide types are simple, and potential landslide mechanisms are also known (Akgun & Erkan 2016). In addition, deterministic methods consider the mechanism and process of landslide occurrence and do not need landslide inventories. Therefore, deterministic models are often used at detailed and medium scales (Fig. 15). In the literature, the most frequently used methodologies for regional deterministic modelling are SINMAP, TRIGRS, and the shallow slope stability model (Yilmaz & Keskin 2009; Zhuang et al. 2017) (Fig. 15).

Bivariate statistical models are the most used from 2001 to 2020 because they can effectively analyse the statistical relationship between historical landslides and influencing factors. The weights of the landslide conditioning factors are assigned based on landslide density using different algorithms (Shahabi et al. 2013), which are simple to operate and easy to understand. However, these models have shortcomings in that (i) they cannot represent the landslide mechanism; (ii) they cannot reflect the relationship between factors; (iii) the classification of each landslide conditioning factor affects the quality of the LS map; and (iv) landslide conditioning factors might be overemphasised regarding their proneness to landslide activity. We found that these models are most frequently applied to medium-scale and large-scale areas, whereas national-scale areas are less frequently used (Fig. 15). This is because a large number

of landslides in medium-scale and large-scale regions can objectively analyse the relationship between landslides and inducing factors, while there may not be so many landslides in national-scale regions. Among the bivariate methods, FR, WoE, and IV/SI have been applied widely and efficiently in solving the task of interest. FR is the ratio of the area where landslides occurred in the total study area and is also the ratio of the probabilities of a landslide occurrence to a nonoccurrence for a given attribute (Lee 2005; Lee & Pradhan 2006). Its calculation is simple and efficient. The results of the FR model show almost a high accuracy when sufficient and well-distributed landslide data are provided (Saadatkhah et al. 2015). WoE is the Bayesian approach in a log-linear form using posterior (conditional) probability and prior (unconditional) probability. The WoE can directly estimate the overall importance of a factor map class and can combine the subjective choice of the classified factors by the expert with the objective data-driven statistical analysis of the GIS (Hussin et al. 2016).

Multivariate statistical methods determine the weights of landslide conditioning factors according to the relative contribution of each in the presence or absence of historical landslides within a defined mapping unit. In the literature, LR is a commonly used multivariate approach. The algorithm of LR is used for maximum likelihood estimation after the conversion of the dependent variable into logit variable (Suzen & Doyuran 2004). The main advantage of this method is that it has no strict requirements on data types, which can be either discrete data or continuous data, either normal distribution data or binary classification data (Lee et al. 2007). The LR model is very effective and reliable for describing problems with variables (e.g. the presence or absence of landslides). Different comparative studies have shown that LR performs better than bivariate methods (Mahdadi et al. 2018) and qualitative models (Shahabi et al. 2014).

In the past 5 years (2016–2020), the use of machine learning has increased rapidly. As the study scale area increases, the use of machine learning also increases. This is mainly due to their high ability to handle complex, nonlinear data, and predictive accuracy problems. Generally, machine learning algorithms can be divided into supervised learning and unsupervised learning (Chang et al. 2020). Unsupervised learning can automatically classify landslide susceptibility without given labels (landslide inventory), which helps to find hidden structures and relationships among data. However, due to the lack of a landslide inventory, the modelling accuracy of unsupervised learning is difficult to verify. Therefore, common unsupervised learning methods, such as k-means clustering and self-organisation mapping, are less frequently used in the LS literature. Supervised learning can handle classification problems that rely on learning from labelled training data, which usually achieves a more

accurate LS map (Liang et al. 2021). The most frequently used supervised learning models include SVM, ANN, RF, DT, maximum entropy, NB, and logistic model tree in the literature database. These techniques have been reported to outperform conventional methods since they seem to be sufficient for handling nonlinear data with different scales and from different types of sources (Aditian et al. 2018). In recent years, deep learning, as a new subdiscipline of machine learning, has emerged to achieve artificial intelligence. Deep learning is an extension of ANNs, which utilise multiple layers of deep neural networks to model complex relations among data. Similar to machine learning, deep learning can also be supervised and unsupervised. Compared with traditional models, deep learning can extract more complex and advanced hidden features although hierarchical analysis of features (Xiong et al. 2021), but the structure is more complex. In the literature database, convolutional neural networks (CNNs) and deep learning neural networks have been successfully used for LS mapping (Bui et al. 2020; Sameen et al. 2020b). CNN is a typical supervised deep learning model that automatically learns complex nonlinear mapping from raw input to given labels or ground truth through a series of convolutional layers (Ji et al. 2020). The descriptions, advantages, and drawbacks of the main machine learning are listed in Table 5. We found that these machine learning-based predictive models are sensitive to some issues, such as overfitting, generalisation error, training dataset quality, and the model configuration parameters. In addition, machine learning models are black boxes in nature, and, thus, it is difficult to interpret or identify the underlying logic.

In the literature database, a total of 218 different hybrid models were used 331 times with the fastest growth. This is because the combination of two or more models can build a new hybrid model, resulting in the diversity of hybrid models. Hybrid models can take full advantage of each single model, overcome the drawback of each single model, and improve the analysis precision and the prediction capacity of the model. For different purposes, the main hybrid models can be divided into three groups. (i) To improve the quality of the data, hybrid models were developed by combining machine learning with statistical models, i.e. RF combined with SI, certainty factor, and index of entropy (Chen et al. 2018c). (ii) To optimise the parameters of machine learning, various metaheuristic optimisation algorithms were hybridised with machine learning algorithms, such as particle swarm optimisation-SVM (Yu et al. 2016), genetic algorithm-SVM (Niu et al. 2014), and artificial bee colony-ANN (Hoang et al. 2019). (iii) To solve the overfitting and generalisation error, hybrid models that combine ensemble learning techniques (e.g. AdaBoost, bagging, boosting, dagging, Decorate, MultiBoost, stacking, and random subspace) with machine learning algorithms (e.g. SVM, DT,

and logistic model tree) as base classifiers were proposed (Dou et al. 2020; Pham et al. 2020). The latter is the focus of current research. Ensemble learning is supervised machine learning that generates many diverse base classifiers on a training dataset and then combines base classifiers' decisions together by weighted or unweighted voting to obtain the final classification result (Huang et al. 2022b). Ensemble learning is mainly proposed from bagging, boosting, and stacking classification techniques. Bagging is a homogeneous ensemble model using parallel training, which obtains multiple training subsets from the original training set by resampling. Because each training subset is independent, it reduces the variance of the base classifier and boosts generalisation capabilities. Moreover, each base learner can be generated in parallel to improve the operating efficiency. Boosting is a homogeneous ensemble model using sequential training. The training set used by boosting in each training is unchanged, but each sample in the training set will be adjusted according to the last learning result so that the new learner can learn from the samples that have been judged wrong by the existing learner. This method can significantly improve the learning effect of weak learners, and each base learner can only be generated sequentially, so the training efficiency is relatively poor. Stacking, a heterogeneous ensemble method using parallel training, trains a model for combining all individual learners. Ensemble learning can overcome the limitations of the training dataset and improve the diversity of base classifiers, resulting in reduced generalisation error. The main disadvantages of ensemble learning are that (i) the complexity of the model increases, (ii) the parameters require careful tuning, and (iii) the output result cannot be interpreted.

Significantly, previous studies have pursued model accuracy too much, while its interpretability has not received enough attention. Model accuracy is related to model complexity, which in turn is the opposite of model interpretability. Generally, a model with a simple structure has good interpretability but poor fitting ability and often low accuracy. The model with a complex structure (e.g. CNN, ensemble learning) has a strong fitting ability and high prediction performance. However, due to the complex working mechanism and low transparency of the model, its interpretation is poor. Based on interpretation principles, interpretability methods can be divided into ante hoc interpretability and post hoc interpretability (global or local) methods (Carvalho et al. 2019). The ante hoc interpretability is the built-in interpretability of the model, which can explain the decision-making process of the model without additional information. In the literature database, the frequently used ante hoc interpretability models include NB, decision tree, and rule-based models. Post hoc interpretability explains the working mechanism and decision-making basis of a given trained learning model using an explanation method or building a

Table 5 Summary of machine learning algorithms

Models	Descriptions	Advantages	Drawbacks
Support vector machine (SVM)	A supervised learning algorithm is commonly used in classification problems by constructing an optimal separating hyperplane to distinguish between the two classes (i.e. landslides and no landslides) and the set of training data	(i) It is efficient in high dimensional spaces (ii) Effective when the dimension is larger than the number of samples (iii) It can improve the generalisation performance (iv) The nonlinear problem can be solved (v) The local minimum problem can be avoided	(i) The high-dimensional mapping interpretation of kernel functions is not strong, especially for radial basis functions (ii) Low modelling efficiency and not user-friendly
Artificial neural network (ANN)	A general nonlinear function approximation algorithm that performs complex calculations by connecting multiple neurons	(i) High classification accuracy and strong learning ability (ii) It has strong robustness and fault tolerance to noise data (iii) It has a self-learning function and associative storage function (iv) It can fully approach complex nonlinear relations	(i) It needs many parameters, such as network topology, weight and initial and value of threshold (ii) Unable to observe the learning process, with black box characteristics (iii) Local minima may occur
Random forest (RF)	A supervised learning algorithm can construct many decision trees to explain the spatial relationship between landslide occurrences	(i) The training can be highly parallelised, which has an advantage over the training speed of large samples (ii) It can process high dimensional data (iii) The contributions of each feature are provided (iv) Because of random sampling, the trained model has small variance and strong generalisation ability (v) It is not sensitive to the loss of some features	(i) When there are many decision trees, the space and time required for training will be large (ii) When dealing with some noisy sample sets, the model is prone to overfitting
Decision tree (DT)	A tree structure-based method for developing a model to classify or predict the values of dependent variables by learning decision rules established from all available data	(i) Simple to build and to interpret (ii) Trees can be visualised and rules can be easily extracted (iii) Able to process nominal and numerical data (iv) Data preprocessing is simple (v) It is a white box model, which can easily infer the corresponding logical expression according to the generated decision tree	(i) It can create overcomplex trees that do not generalise the data well, resulting in overfitting (ii) DTs may be unstable due to small changes in the data may result in a completely different tree being produced. This problem can be alleviated using DTs in an ensemble (iii) Ignoring the correlation between variables in the dataset
Maximum entropy	It is a rule of probability model learning. The principle holds that among all possible probability models that satisfy the existing facts, the model with the largest entropy is the best one	(i) The model with maximum information entropy has higher accuracy (ii) Flexible setting of the restraint conditions can adjust the adaptability of the model to unknown data and the degree of fitting to known data	(i) The number of constraint functions is related to the number of samples, resulting in calculation amount is too large in the iterative process (ii) When the samples are large, the iterative process of dual function optimisation is very slow
Naive Bayesian	A supervised learning algorithms based on Bayes theorem and independent assumption of feature conditions	(i) It is not sensitive to missing data, and the algorithm is simple (ii) A small amount of training data is required to estimate the necessary parameters (iii) It can be very fast compared with more complex methods	(i) Strong assumptions are made about the independence of variables, which is often not true in practical applications (ii) Hypothetical prior models may make predictions less effective (iii) There is a certain error rate in classification decision

Table 5 (continued)

Models	Descriptions	Advantages	Drawbacks
Logistic model tree	A relatively new classification algorithm that combines LR and DT in which the information gain is employed to split and the LogitBoost algorithm is employed to fit the logistic regression at a tree node	(i) It has good generalisation and low classification deviation (ii) Logic regression function is added to the leaves of the decision tree, and logic regression has a strong tolerance for small noise, and it is generally difficult to appear over fitting	(i) It is difficult to deal with the missing data (ii) Ignore the correlation of variables in the dataset
Convolutional neural network (CNN)	It is a typical supervised deep learning method which hierarchically combines simple local features into complex models	(i) It can automatically extract and classify features from high-dimensional data (ii) The performance of model depends on the different architectures involving various operations (iii) High classification accuracy	(i) It needs to normalise the dataset (ii) It has no memory function (iii) There is no corresponding theoretical guidance for the selection of the number of network layers and neurons, so parameter adjustment is needed (iv) It is a black box model with complex structure
K-means clustering	An unsupervised learning algorithm automatically partitions all input data into k clusters. The core of the algorithm is to optimise the distortion function so that it converges to the local minimum but not the global minimum	(i) Simple algorithm idea and fast convergence speed (ii) The clustering effect is better (iii) The main parameter to be adjusted is only the number of clusters k (iv) The interpretability of the algorithm is strong	(i) By using iterative method, the clustering results tend to converge to the local optimal solution and cannot obtain the global optimal solution (ii) The recognition effect of nonconvex shape clusters is poor (iii) Easily affected by noise, edge points and isolated points (iv) The data types that can be processed are limited, and the clustering effect for high-dimensional data objects is poor
Self-organisation map	An unsupervised learning algorithm is designed as two-dimensional arrangement of neurons, the input variables are mapped to two-dimensional with two main stages: training stage and classification stage	(i) The training set and the test set are not distinguished (ii) It can map high-dimensional input to low-dimensional input	(i) The network structure is fixed and cannot be changed dynamically (ii) The networks cannot be added to new categories without complete relearning

model. Some post hoc interpretability approaches have been proposed to explain the contribution of features to predicted values, such as Shapley additive explanation and partial dependency plots (Zhou et al. 2022). The Shapley additive explanation method shows the contribution of each feature to the predicted value by calculating the Shapley value of each feature in an individual sample and then explains the black box model from the global and local levels. Partial dependency plots can show the contribution of one or two features to the predicted value of the model and help to determine what will happen to the model predictions when many features are adjusted. However, we found that these post hoc interpretability methods are rarely used in the interpretation of landslide prediction models.

The selection of landslide susceptibility models is changing from simple qualitative and statistical models to complex machine learning and hybrid models. The performance of different model algorithms varied at different study sites (Ghasemain et al. 2020). The choice of the appropriate model depends on the fundamental quality of the model algorithm, the desired scale of analysis, the landslide inventory used, and the conditioning factors considered (Yilmaz 2009). For detailed scale research, the results obtained using deterministic models to predict shallow landslides may be more reliable. More machine learning and hybrid models are applied to regional and national study scales. The factors are more complex and heterogeneous at regional and national study scales. Machine learning is a rapidly developing field with rapid technology iteration. Deep learning can effectively extract useful information from big data to reveal complex nonlinear relationships between conditioning factors and landslides without requiring additional prior knowledge or assumptions. Therefore, deep learning, especially methods based on CNNs, may require additional research on LS mapping. When landslide samples are insufficient in a region-scale area, unsupervised and semisupervised deep learning is indispensable. We recommend using deep learning and hybrid models to solve the problems of machine learning algorithms, which improve model accuracy. Machine learning with black box properties cannot interpret the statistical regularity between landslides and environmental conditions. Therefore, the interpretability of machine learning models should also be considered.

Model performance validation analysis

The analysis of keywords in the GIS-based LS research showed that the terms "validation", "prediction" and "spatial prediction" are popular. Validation or zonation of the results began to be considered in early research from 2001 to 2010 (Fig. 7). The validation of LS maps is an important task that reflects the accuracy of the study and confirms the usability

of the model (Hong et al. 2017). We found that 112 articles in the 1142 publications did not perform the model fit and prediction and 404 articles did not validate model prediction performance. LS maps have no practical significance without model validation. Six groups of metrics were used in the literature database. A single metric only provides limited insight into the model performance; therefore, appropriate metrics should be selected to comprehensively evaluate the model performance based on the advantages and limitations of each metric (Reichenbach et al. 2018).

Technically, LS assessment is a classification problem with the binary outcome of the presence or absence of a landslide (Merghadi et al. 2020). Therefore, the types of probability and class label metrics were frequently used to measure the model fit and prediction performance. These metrics are applicable to classification problems, which were calculated from the confusion matrix (Table 6). The differences between these metrics in model fit and prediction can determine whether the model is underfitting or overfitting. The success/prediction rate and ROC curves are the most used summary metrics to measure model performance using a form of class probabilities. A ROC curve is a threshold-independent curve that exhibits sensitivity against 1-specificity for many different thresholds of classification scores. The area under the ROC curve (AUC) corresponds to the integral of an ROC curve. The AUC value can show the model success rate and prediction rate by engaging the training and test datasets, respectively. The success rate curve reveals the model fitting rate, which indicates how well the LS mapping separates the landslides among its susceptibility zones (Vakhshoori & Zare 2018). The prediction rate curve properly indicates the model performance in the prediction of future landslides. The advantage of ROC is that it yields a threshold-independent measure of predictive ability. A threshold-independent measure requires that the classifier generate some sort of scores from which the dataset can be divided into positively and negatively predicted classes, rather than simply providing a static partition (Saito & Rehmsmeier 2015). However, class label metrics are designed to assess hard classifications at a single probability decision threshold (e.g. accuracy, sensitivity, specificity, precision, and F-measure). These single threshold metrics cannot outline performance ranges with varying thresholds.

In addition, some classification metrics are sensitive to class imbalance (i.e. a difference in the relative proportion of positive and negative data). Due to the intuitive but incorrect interpretation of specificity, the visual interpretability of ROC curves can deceive conclusions about the reliability of classification performance when applied to imbalanced data (nonlandslides outnumbering landslides). In this case, precision-recall curves can provide an accurate prediction for future classification performance because they evaluate the fraction of true positives among positive predictions (Saito

Table 6 The main metrics used in model evaluation

Metrics	Formulas	Descriptions
AUC	$AUC = \frac{\sum TP + \sum TN}{P + N}$	The area under the ROC curve is defined as the AUC value. A higher value of AUC denotes better model performance
Accuracy	$Accuracy = \frac{TP + TN}{TP + FP + TN + PN}$	Accuracy is the proportion of landslide and nonlandslide pixels that are classified correctly
Sensitivity	$Sensitivity = \frac{TP}{TP + FN}$	Sensitivity (or recall or true positive rate) means the proportion of landslide pixels that are correctly classified
Specificity	$Specificity = \frac{TN}{FP + TN}$	Specificity (or true negative rate) indicates the ability of the test to correctly distinguish nonlandslide pixels
Precision	$Precision = \frac{TP}{TP + FP}$	Precision (or positive predictive rate) represents the proportion of actual landslide pixels to those classified as landslides by the model
Kappa index	$kappa = \frac{Accuracy - Accuracy_{exp}}{1 - Accuracy_{exp}} \quad Accuracy_{exp} = \frac{(TP+FN)(TP+FP)+(FP+TN)(TN+FN)}{(TP+TN+FP+FN)^2}$	Measure the agreement between true classes and the classifications
Negative predictive rate	$Negative\ predictive\ rate = \frac{TN}{TN + FN}$	The probability that predictive nonlandslide pixels have actual non-landslide occurrences
False positive rate	$False\ positive\ rate = \frac{FP}{FP + TN}$	The rate of nonlandslide instances falsely classified as landslide pixels
F-measure	$F = \frac{2 * Sensitivity * Precision}{Sensitivity + Precision}$	The harmonic mean of precision and recall
Root mean square error	$Root\ mean\ square\ error = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}_i\right)^2}$	Show the error measure between the predicted and actual values
Mean square error	$Mean\ square\ error = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}_i\right)^2$	Show the error measure between the predicted and actual values
Mean absolute error	$Mean\ absolute\ error = \frac{1}{n}\sum_{i=1}^{n}\left|X_i - \overline{X}_i\right|$	An error-based measure to check the performance of the models

TP true positive, *TN* true negative, *FN* false negative, *FP* false positive. *P* and *N* represent the total number of landslides and nonlandslides, respectively. *n* is the total number of samples. X_i and $\overline{X}_i$ are the predicted and actual values of outputs, respectively

& Rehmsmeier 2015). Similarly, class label metrics such as sensitivity, specificity, precision, and F-measure are suitable for imbalanced datasets. However, accuracy is suitable for balanced data. This is because in the context of imbalanced data, accuracy is misleading when the model is inaccurate in predicting negative class labels.

Regression metrics, such as root mean square error, mean square error, and mean absolute error, are used to measure the training and testing error of the implemented models. These metrics are applicable to only regression problems. In addition, landslide density calculated by the spatial overlay of landslides and different susceptible zones was used to compare whether the constructed susceptibility maps were reliable. The comparison between the landslide location obtained from the field survey and the landslide susceptibility map can reflect the efficiency of the model. At present, field verification is only used as an auxiliary metric and is rarely used.

Therefore, we recommend that the following be considered to evaluate model performance: (i) the case study scenario and the type (i.e. classification, regression); (ii) the landslide inventory with balanced data or imbalanced data; and (iii) use of more metrics to verify and compare the reliability of different models and to provide the confidence interval and standard error of the metrics (e.g. AUC) to capture the uncertainty inherent in the measures. In addition, investigators should concentrate on the design of novel and more reliable methods and metrics for the evaluation of model/algorithm quality, which will improve their credibility and practicality.

Challenges ahead and research directions

Although research in GIS-based LS is in the rapid development stage, opportunities for future research are still significant. Based on key phrase co-occurrence and content analysis, this study identifies several future research opportunities.

Exploring comprehensive factors that induce landslides

Determining the appropriate number and combination of conditioning factors according to their corresponding contributions to landslides is still a challenging task (Abdulwahid & Pradhan 2017). Currently, there is no universal guideline for the selection of conditioning factors because the factors involve the specific study site and model. In characteristic scenarios, there are differences in landslide failure mechanisms, and the inducing factors of landslides are also different. However, certain factors (e.g. underground water, rock weathering, and bedding-slope relations) that have a significant impact on landslide occurrence are rarely involved in model analysis (Huang & Zhao 2018; Huang et al. 2020a). Therefore, a more comprehensive factor database

for different landslide types should be established by deeply analysing the mechanism of landslides and exploring the controlling factors of landslides in specific areas. Generally, feature selection methods are adopted to evaluate the estimation ability of input factors and eliminate irrelevant or partially relevant factors before constructing the LS model.

Distinguishing landslide types

Most of the previous studies did not distinguish the types of landslides (e.g. rock falls/avalanches, debris avalanches, rockslides, flow slides, and debris flows), but the mechanisms of different landslide types are not the same. Then, the conditioning factors considered will inevitably be different. Therefore, future research should distinguish the types of landslides and adopt effective models for different landslide types. For example, rockslides/avalanches, rock falls, and debris flow susceptibility are best confirmed using physically based models; slow-moving slides, flows, and complex or compound LS are best ascertained using statistical models (Reichenbach et al. 2018).

Research on the uncertainty of multisource data

Landslide factors have diverse sources, and the scale and quality of data layers are different, which brings uncertainty to landslide modelling. Therefore, to avoid uncertainties associated with the input parameters in modelling, multisource data fusion analysis is very important.

Optimising and improving machine learning and hybrid models

Machine learning and hybrid models are clearly popular topics and are also possible directions for future research (Fig. 7). However, some limits exist to these methods, which are mainly associated with the complexity of the models and problems that arise from the effects of the curse of dimensionality and overfitting (Chen et al. 2020). A model that better reveals the failure mechanism associated with landslides should be deeply explored. Post hoc interpretability methods can be used for the interpretation of landslide prediction models. Furthermore, an accurate LS map generated using machine learning and hybrid models requires a balance between satisfactory performance and actual computing time (Merghadi et al. 2020). However, there are only a limited number of studies on using calculation time to evaluate performance level. Therefore, these methods must be further improved and optimised so that they operate within a reasonable timeframe rather than simply applying models to obtain results.

Establishing a global artificial intelligence LS evaluation system

Recent LS studies are stationary, and display assessment results that are invariant with time (Reichenbach et al. 2018). However, the conditions may change dynamically with extreme rainfall, earthquakes, volcanic activities, and human actions. Therefore, the landslide inventory will change, which results in previous LS maps being outdated. In the past, many researchers have conducted repeated studies at the same sites, which cause a considerable waste of human and material resources. In addition, it remains difficult to conduct LS research in some countries in Africa and South America, particularly due to the lack of sufficient and adequate landslide information. Accordingly, it is urgent to develop a global artificial intelligence LS assessment system to predict the risk in adjacent and distant areas. First, it is necessary to strengthen global cooperation to build a basic database. Then, the landslide inventory can be updated in real time through multiphase remote sensing images, multispectral satellite images, InSAR, LiDAR, or hyperspectral data methods or technologies. Later, machine learning and deep learning can be used to learn landslide occurrences in real time. Finally, according to the characteristics of the study area, the appropriate factors and models should be selected to intelligently evaluate LS at the local, regional, national, and global scales.

Conclusions

In this study, a systematic bibliometric and content analysis of the evolution and development of GIS-based LS research was performed with VOSviewer based on 1142 articles from WoSCC published over the past 20 years. Thematic data on the LS, including the landslide inventory, conditioning factors, mapping unit, evaluation model, and validation metrics, were extracted. The paper determines the topics and/ or hotspots of GIS-based LS, reveals the advantages and limitations of common approaches used in thematic data of LS, and summarises the development trends, which will guide investigators to quickly and fully master this research field. The main conclusions are as follows:

(1) The number of articles has increased greatly from 5 to 151 from 2001 to 2020, and they were published in 183 journals with 2570 authors, 948 institutions, and 79 countries. The distribution of articles shows clear authors, institutions, and geographical clusters with high levels of academic activity. Both the top 10 most-productive authors and institutions contributed more than 39% of all articles. The collaboration network maps reflect that these authors, institutions, and coun-

tries regularly collaborated with others. The analysis of the keyword co-occurrence reveals that the main study topics are validation and zonation, conventional models, and machine learning and hybrid models in the three periods of 2001–2010, 2011–2015, and 2016–2020, respectively.

(2) The distribution of study sites has a significant geographical bias. Popular study sites are mainly concentrated in China, India, Iran, South Korea, and Turkey, while the study areas in Africa, South America, and Oceania remain few. With the rapid development of remote sensing and GIS technologies, the application of remote sensing technology to prepare landslide inventories has changed from visual interpretation to semiautomatic/automatic interpretation. Most of the sample strategies representing the landslides are point and polygon. The sampling strategies of landslide scarps or rupture zones may obtain a better prediction result because they are the most diagnostically unstable landforms. Random selection is the most frequently used sample subdividing strategy, and cross-validation and bootstrap sampling can solve the limitations of small training samples that are not fully used.

(3) The scale effects, a lack of geographic consistency, and the lack of standards are key problems in landslide conditioning factors. We recommend identifying and selecting all geoenvironmental factors available for the study area and then using feature selection to eliminate irrelevant or less important factors. With advances in computing technology and artificial intelligence, LS models are changing from simple qualitative and statistical models to complex machine learning and hybrid models. A complex model improves model accuracy, but its interpretability has not received enough attention.

(4) The keyword co-occurrence and content analysis found the following five future research opportunities: exploring the comprehensive factors that induce landslides; distinguishing landslide types; researching the uncertainty of multisource data; optimising and improving machine learning and integration models; and establishing a global artificial intelligence LS evaluation system.

Finally, the demonstrated conclusions have significant potential implications for GIS-based LS, even for landslide effects on sustainable environmental development. Therefore, we recommend that a comprehensive/effective landslide database, intelligence models, and effective verification are necessary for LS maps at different geographical scales in real time, especially for extreme events (e.g. rainstorms, earthquakes, and volcanic eruptions). Additionally, we also expect that LS maps will be used for some landslide early

warning systems at different geographical scales, which will realise geohazard prevention and environmental protection management for the goal of sustainable social and environmental development.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Junpeng Huang, Yuxin Wu, Lei Peng, and Zhiyi He. The first draft of the manuscript was written by Junpeng Huang and reviewed by Sixiang Ling and Xiaoning Li. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data availability The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

Abdulwahid WM, Pradhan B (2017) Landslide vulnerability and risk assessment for multi-hazard scenarios using airborne laser scanning data (LiDAR). Landslides 14:1057–1076. https://doi.org/10.1007/s10346-016-0744-0

Aditian A, Kubota T, Shinohara Y (2018) Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia. Geomorphology 318:101–111. https://doi.org/10.1016/j.geomorph.2018.06.006

Aghdam IN, Varzandeh MHM, Pradhan B (2016) Landslide susceptibility mapping using an ensemble statistical index (Wi) and adaptive neuro-fuzzy inference system (ANFIS) model at Alborz Mountains (Iran). Environ Earth Sci 75:553. https://doi.org/10.1007/s12665-015-5233-6

Akgun A, Erkan O (2016) Landslide susceptibility mapping by geographical information system-based multivariate statistical and deterministic models: in an artificial reservoir area at Northern Turkey. Arabian J Geosci 9:165. https://doi.org/10.1007/s12517-015-2142-7

Akinci H, Kilicoglu C, Dogan S (2020) Random forest-based landslide susceptibility mapping in coastal regions of Artvin. Turkey ISPRS Int J Geo-Inf 9:553. https://doi.org/10.3390/ijgi9090553

Althuwaynee OF, Pradhan B, Park H-J, Lee JH (2014) A novel ensemble bivariate statistical evidential belief function with knowledge-based analytical hierarchy process and multivariate statistical logistic regression for landslide susceptibility mapping. Catena 114:21–36. https://doi.org/10.1016/j.catena.2013.10.011

Arabameri A, Pradhan B, Rezaei K, Lee C-W (2019) Assessment of landslide susceptibility using statistical- and artificial intelligence-based FR-RF integrated model and multiresolution DEMs. Remote Sens 11:999. https://doi.org/10.3390/rs11090999

Ayalew L, Yamagishi H (2005) The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology 65:15–31. https://doi.org/10.1016/j.geomorph.2004.06.010

Ba Q, Chen Y, Deng S, Yang J, Li H (2018) A comparison of slope units and grid cells as mapping units for landslide susceptibility assessment. Earth Sci Inf 11:373–388. https://doi.org/10.1007/s12145-018-0335-9

Begueria S (2006) Validation and evaluation of predictive models in hazard assessment and risk management. Nat Hazards 37:315–329. https://doi.org/10.1007/s11069-005-5182-6

Bordoni M, Galanti Y, Bartelletti C, Persichillo MG, Barsanti M, Giannecchini R, Avanzi GD, Cevasco A, Brandolini P, Galve JP, Meisina C (2020) The influence of the inventory on the determination of the rainfall-induced shallow landslides susceptibility using generalized additive models. Catena 193:104630. https://doi.org/10.1016/j.catena.2020.104630

Brabb EE (1984) Innovative approaches to landslide hazard mapping. Proceedings of the 4th International Symposium of Landslides, Toronto, Canada. pp 307–324

Brenning A (2005) Spatial prediction models for landslide hazards: review, comparison and evaluation. Nat Hazards Earth Syst Sci 5:853–862. https://doi.org/10.5194/nhess-5-853-2005

Broeckx J, Vanmaercke M, Duchateau R, Poesen J (2018) A data-based landslide susceptibility map of Africa. Earth Sci Rev 185:102–121. https://doi.org/10.1016/j.earscirev.2018.05.002

Budimir MEA, Atkinson PM, Lewis HG (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 DT, Pradhan B, Lofman O, Revhaug I, Dick OB (2012) Spatial prediction of landslide hazards in Hoa Binh province (Vietnam): a comparative assessment of the efficacy of evidential belief functions and fuzzy logic models. Catena 96:28–40. https://doi.org/10.1016/j.catena.2012.04.001

Bui DT, Tuan TA, Klempe H, Pradhan B, Revhaug I (2016) Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides 13:361–378. https://doi.org/10.1007/s10346-015-0557-6

Bui DT, Shahabi H, Shirzadi A, Chapi K, Alizadeh M, Chen W, Mohammadi A, Bin Ahmad B, Panahi M, Hong H, Tian Y (2018) Landslide detection and susceptibility mapping by AIRSAR data using support vector machine and index of entropy models in Cameron Highlands. Malaysia Remote Sens 10:1527. https://doi.org/10.3390/rs10101527

Bui DT, Tsangaratos P, Nguyen VT, Liem NV, Trinh PT (2020) Comparing the prediction performance of a deep learning neural network model with conventional machine learning models in landslide susceptibility assessment. Catena 188:104426. https://doi.org/10.1016/j.catena.2019.104426

Cama M, Conoscenti C, Lombardo L, Rotigliano E (2016) Exploring relationships between grid cell size and accuracy for debris-flow susceptibility models: a test in the Giampilieri catchment

(Sicily, Italy). Environ Earth Sci 75:238. https://doi.org/10.1007/s12665-015-5047-6

Camargo JMR, Silva MVB, Ferreira AV, Araujo TCM (2019) Marine geohazards: a bibliometric-based review. Geosciences 9:100. https://doi.org/10.3390/geosciences9020100

Canavesi V, Segoni S, Rosi A, Ting X, Nery T, Catani F, Casagli N (2020) Different approaches to use morphometric attributes in landslide susceptibility mapping based on meso-scale spatial units: a case study in Rio de Janeiro (Brazil). Remote Sens 12:1826. https://doi.org/10.3390/rs12111826

Carvalho DV, Pereira EM, Cardoso JS (2019) Machine learning interpretability: a survey on methods and metrics. Electronics 8:832. https://doi.org/10.3390/electronics8080832

Chacon J, Irigaray C, Fernandez T, El Hamdouni R (2006) Engineering geology maps: landslides and geographical information systems. Bull Eng Geol Environ 65:341–411. https://doi.org/10.1007/s10064-006-0064-z

Chang KT, Merghadi A, Yunus AP, Pham BT, Dou J (2019) Evaluating scale effects of topographic variables in landslide susceptibility models using GIS-based machine learning techniques. Sci Rep 9:12296. https://doi.org/10.1038/s41598-019-48773-2

Chang ZL, Du Z, Zhang F, Huang FM, Chen JW, Li WB, Guo ZZ (2020) Landslide susceptibility prediction based on remote sensing images and GIS: comparisons of supervised and unsupervised machine learning models. Remote Sens 12:502. https://doi.org/10.3390/rs12030502

Chen CM (2017) Science mapping: a systematic review of the literature. J Data Info Sci 2:1–40. https://doi.org/10.1515/jdis-2017-0006

Chen W, Li X, Wang Y, Liu S (2013) Landslide susceptibility mapping using LiDAR and DMC data: a case study in the Three Gorges area, China. Environ Earth Sci 70:673–685. https://doi.org/10.1007/s12665-012-2151-8

Chen W, Peng JB, Hong HY, Shahabi H, Pradhan B, Liu JZ, Zhu AX, Pei XJ, Duan Z (2018a) Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China. Sci Total Environ 626:1121–1135. https://doi.org/10.1016/j.scitotenv.2018.01.124

Chen W, Shahabi H, Zhang S, Khosravi K, Shirzadi A, Chapi K, Binh Thai P, Zhang T, Zhang L, Chai H, Ma J, Chen Y, Wang X, Li R, Bin Ahmad B (2018b) Landslide susceptibility modeling based on GIS and novel bagging-based kernel logistic regression. Applied Sciences-Basel 8:2540. https://doi.org/10.3390/app8122540

Chen W, Xie XS, Peng JB, Shahabi H, Hong HY, Dieu Tien B, Duan Z, Li SJ, Zhu AX (2018c) GIS-based landslide susceptibility evaluation using a novel hybrid integration approach of bivariate statistical based random forest method. Catena 164:135–149. https://doi.org/10.1016/j.catena.2018.01.012

Chen W, Chen YZ, Tsangaratos P, Ilia I, Wang XJ (2020) Combining evolutionary algorithms and machine learning models in landslide susceptibility assessments. Remote Sens 12:3854. https://doi.org/10.3390/rs12233854

Choi J, Oh H-J, Lee H-J, Lee C, Lee S (2012) Combining landslide susceptibility maps obtained from frequency ratio, logistic regression, and artificial neural network models using ASTER images and GIS. Eng Geol 124:12–23. https://doi.org/10.1016/j.enggeo.2011.09.011

Ciampalini A, Raspini F, Lagomarsino D, Catani F, Casagli N (2016) Landslide susceptibility map refinement using PSInSAR data. Remote Sens Environ 184:302–315. https://doi.org/10.1016/j.rse.2016.07.018

Clerici A, Perego S, Tellini C, Vescovi P (2006) A GIS-based automated procedure for landslide susceptibility mapping by the conditional analysis method: the Baganza valley case study (Italian

Northern Apennines). Environ Geol 50:941–961. https://doi.org/10.1007/s00254-006-0264-7

Clerici A, Perego S, Tellini C, Vescovi P (2010) Landslide failure and runout susceptibility in the upper T. Ceno valley (Northern Apennines, Italy). Nat Hazards 52:1–29. https://doi.org/10.1007/s11069-009-9349-4

Conoscenti C, Di Maggio C, Rotigliano E (2008) GIS analysis to assess landslide susceptibility in a fluvial basin of NW Sicily (Italy). Geomorphology 94:325–339. https://doi.org/10.1016/j.geomorph.2006.10.039

Dai FC, Lee CF (2001) Terrain-based mapping of landslide susceptibility using a geographical information system: a case study. Can Geotech J 38:911–923. https://doi.org/10.1139/t01-021

Dai FC, Lee CF, Li J, Xu ZW (2001) Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. Environ Geol 40:381–391. https://doi.org/10.1007/s002540000163

Das I, Stein A, Kerle N, Dadhwal VK (2011) Probabilistic landslide hazard assessment using homogeneous susceptible units (HSU) along a national highway corridor in the northern Himalayas, India. Landslides 8:293–308. https://doi.org/10.1007/s10346-011-0257-9

Das I, Stein A, Kerle N, Dadhwal VK (2012) Landslide susceptibility mapping along road corridors in the Indian Himalayas using Bayesian logistic regression models. Geomorphology 179:116–125. https://doi.org/10.1016/j.geomorph.2012.08.004

Dou J, Yunus AP, Bui DT, Sahana M, Chen C-W, Zhu Z, Wang W, Pham BT (2019) Evaluating GIS-based multiple statistical models and data mining for earthquake and rainfall-induced landslide susceptibility using the LiDAR DEM. Remote Sens 11:638. https://doi.org/10.3390/rs11060638

Dou J, Yunus AP, Bui DT, Merghadi A, Sahana M, Zhu ZF, Chen CW, Han Z, Pham BT (2020) Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. Landslides 17:641–658. https://doi.org/10.1007/s10346-019-01286-5

Dumperth C, Rohn J, Fleer A, Wang JG, Xiang W, Zimmermann K (2016) An easy approach to assess the susceptibility of a landslide by utilizing simple raster overlay methods: a case study on Huangtupo landslide (PR China). J Mt Sci 13:1701–1710. https://doi.org/10.1007/s11629-016-4011-6

Eker AM, Dikmen M, Cambazoglu S, Duzgun SHB, Akgun H (2015) Evaluation and comparison of landslide susceptibility mapping methods: a case study for the Ulus district, Bartin, northern Turkey. Int J Geogr Inf Sci 29:132–158. https://doi.org/10.1080/13658816.2014.953164

Ercanoglu M, Temiz FA (2011) Application of logistic regression and fuzzy operators to landslide susceptibility assessment in Azdavay (Kastamonu, Turkey). Environ Earth Sci 64:949–964. https://doi.org/10.1007/s12665-011-0912-4

Fan JL, Shen S, Wang JD, Wei SJ, Zhang X, Zhong P, Wang H (2020) Scientific and technological power and international cooperation in the field of natural hazards: a bibliometric analysis. Nat Hazards 102:807–827. https://doi.org/10.1007/s11069-020-03919-8

Federici PR, Puccinelli A, Cantarelli E, Casarosa N, Avanzi GDA, Falaschi F, Giannecchini R, Pochini A, Ribolini A, Bottai M, Salvati N, Testi C (2007) Multidisciplinary investigations in evaluating landslide susceptibility - an example in the Serchio River valley (Italy). Quat Int 171–72:52–63. https://doi.org/10.1016/j.quaint.2006.10.018

Fell R, Corominas J, Bonnard C, Cascini L, Leroi E, Savage WZ, Eng J-JTCL (2008) Guidelines for landslide susceptibility, hazard and risk zoning for land-use planning Commentary. Eng Geol 102:99–111. https://doi.org/10.1016/j.enggeo.2008.03.014

Garfield E, Pudovkin AI, Istomin VS (2003) Why do we need algorithmic historiography? J Am Soc Inf Sci Tech 54:400–412. https://doi.org/10.1002/asi.10226

Ghasemain B, Asle DT, Pham BT, Avand M, Nguyen HD, Janizadeh S (2020) Shallow landslide susceptibility mapping: a comparison between classification and regression tree and reduced error pruning tree algorithms Vietnam J. Earth Sci 42:208–227. https://doi.org/10.15625/0866-7187/42/3/14952

Goetz JN, Guthrie RH, Brenning A (2011) Integrating physical and empirical landslide susceptibility models using generalized additive models. Geomorphology 129:376–386. https://doi.org/10.1016/j.geomorph.2011.03.001

Gorsevski PV, Brown MK, Panter K, Onasch CM, Simic A, Snyder J (2016) Landslide detection and susceptibility mapping using LiDAR and an artificial neural network approach: a case study in the Cuyahoga Valley National Park, Ohio. Landslides 13:467–484. https://doi.org/10.1007/s10346-015-0587-0

Guzzetti F, Carrara A, Cardinali M, Reichenbach P (1999) Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31:181–216. https://doi.org/10.1016/s0169-555x(99)00078-1

Guzzetti F, Reichenbach P, Ardizzone F, Cardinali M, Galli M (2006) Estimating the quality of landslide susceptibility models. Geomorphology 81:166–184. https://doi.org/10.1016/j.geomorph.2006.04.007

Guzzetti F, Mondini AC, Cardinali M, Fiorucci F, Santangelo M, Chang K-T (2012) Landslide inventory maps: new tools for an old problem. Earth Sci Rev 112:42–66. https://doi.org/10.1016/j.earscirev.2012.02.001

Heckmann T, Gegg K, Gegg A, Becht M (2014) Sample size matters: investigating the effect of sample size on a logistic regression susceptibility model for debris flows. Nat Hazards Earth Syst Sci 14:259–278. https://doi.org/10.5194/nhess-14-259-2014

Hess DM, Leshchinsky BA, Bunn M, Mason HB, Olsen MJ (2017) A simplified three-dimensional shallow landslide susceptibility framework considering topography and seismicity. Landslides 14:1677–1697. https://doi.org/10.1007/s10346-017-0810-2

Hoang N, Mehrabi M, Kalantar B, Moayedi H, MaM A (2019) Potential of hybrid evolutionary approaches for assessment of geohazard landslide susceptibility mapping. Geomat Nat Haz Risk 10:1667–1693. https://doi.org/10.1080/19475705.2019.1607782

Hong Y, Adler RF (2008) Predicting global landslide spatiotemporal distribution: Integrating landslide susceptibility zoning techniques and real-time satellite rainfall estimates. Int J Sediment Res 23:249–257. https://doi.org/10.1016/s1001-6279(08)60022-0

Hong Y, Adler R, Huffman G (2007a) Use of satellite remote sensing data in the mapping of global landslide susceptibility. Nat Hazards 43:245–256. https://doi.org/10.1007/s11069-006-9104-z

Hong Y, Adler RF, Huffman G (2007b) An experimental global prediction system for rainfall-triggered landslides using satellite remote sensing and geospatial datasets. IEEE Trans Geosci Remote Sens 45:1671–1680. https://doi.org/10.1109/tgrs.2006.888436

Hong HY, Pradhan B, Sameen MI, Chen W, Xu C (2017) Spatial prediction of rotational landslide using geographically weighted regression, logistic regression, and support vector machine models in Xing Guo area (China). Geomat Nat Haz Risk 8:1997–2022. https://doi.org/10.1080/19475705.2017.1403974

Hong HY, Liu JY, Bui DT, Pradhan B, Acharya TD, Binh Thai P, Zhu AX, Chen W, Bin Ahmad B (2018) Landslide susceptibility mapping using J48 Decision Tree with AdaBoost, Bagging and Rotation Forest ensembles in the Guangchang area (China). Catena 163:399–413. https://doi.org/10.1016/j.catena.2018.01.005

Huang Y, Zhao L (2018) Review on landslide susceptibility mapping using support vector machines. Catena 165:520–529. https://doi.org/10.1016/j.catena.2018.03.003

Huang HP, Yang KC, Lin BW (2013) Statistical evaluation of the effect of earthquake with other related factors on landslide susceptibility: using the watershed area of Shihmen reservoir in Taiwan as a case study. Environ Earth Sci 69:2151–2166. https://doi.org/10.1007/s12665-012-2044-x

Huang JP, Sun CW, Wu XY, Ling SX, Wang S, Deng R (2020a) Stability assessment of tunnel slopes along the Dujiangyan City to Siguniang Mountain Railway, China. Bull Eng Geol Environ 79:5309–5327. https://doi.org/10.1007/s10064-020-01913-9

Huang JX, Xie MW, Atkinson PM (2020b) Dynamic susceptibility mapping of slow-moving landslides using PSInSAR. Int J Remote Sens 41:7509–7529. https://doi.org/10.1080/01431161.2020.1760398

Huang FM, Tao SY, Chang ZL, Huang JS, Fan XM, Jiang SH, Li WB (2021) Efficient and automatic extraction of slope units based on multi-scale segmentation method for landslide assessments. Landslides 18:3715–3731. https://doi.org/10.1007/s10346-021-01756-9

Huang JP, Ling SX, Wu XY, Deng R (2022a) GIS-based comparative study of the Bayesian network, decision table, radial basis function network and stochastic gradient descent for the spatial prediction of landslide susceptibility. Land 11:436. https://doi.org/10.3390/land11030436

Huang JP, Ma N, Ling SX, Wu XY (2022b) Comparing the prediction performance of logistic model tree with different ensemble techniques in susceptibility assessments of different landslide types. Geocarto Int 1-31. https://doi.org/10.1080/10106049.2022.2087751

Hungr O, Leroueil S, Picarelli L (2014) The Varnes classification of landslide types, an update. Landslides 11:167–194. https://doi.org/10.1007/s10346-013-0436-y

Hussin HY, Zumpano V, Reichenbach P, Sterlacchini S, Micu M, van Westen C, Balteanu D (2016) Different landslide sampling strategies in a grid-based bi-variate statistical susceptibility model. Geomorphology 253:508–523. https://doi.org/10.1016/j.geomorph.2015.10.030

Jaafari A (2018) LiDAR-supported prediction of slope failures using an integrated ensemble weights-of-evidence and analytical hierarchy process. Environ Earth Sci 77:42. https://doi.org/10.1007/s12665-017-7207-3

Jaafari A, Najafi A, Rezaeian J, Sattarian A, Ghajar I (2015) Planning road networks in landslide-prone areas: a case study from the northern forests of Iran. Land Use Policy 47:198–208. https://doi.org/10.1016/j.landusepol.2015.04.010

Jaafari A, Panahi M, Pham BT, Shahabi H, Bui DT, Rezaie F, Lee S (2019) Meta optimization of an adaptive neuro-fuzzy inference system with grey wolf optimizer and biogeography-based optimization algorithms for spatial prediction of landslide susceptibility. Catena 175:430–445. https://doi.org/10.1016/j.catena.2018.12.033

Jaboyedoff M, Oppikofer T, Abellan A, Derron MH, Loye A, Metzger R, Pedrazzini A (2012) Use of LIDAR in landslide investigations: a review. Nat Hazards 61:5–28. https://doi.org/10.1007/s11069-010-9634-2

Jebur MN, Pradhan B, Tehrany MS (2014) Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. Remote Sens Environ 152:150–165. https://doi.org/10.1016/j.rse.2014.05.013

Ji SP, Yu DW, Shen CY, Li WL, Xu Q (2020) Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks. Landslides 17:1337–1352. https://doi.org/10.1007/s10346-020-01353-2

Jimenez-Peralvarez JD, El Hamdouni R, Palenzuela JA, Irigaray C, Chacon J (2017) Landslide-hazard mapping through multi-technique activity assessment: an example from the Betic

Cordillera (southern Spain). Landslides 14:1975–1991. https://doi.org/10.1007/s10346-017-0851-6

Kakavas MP, Nikolakopoulos KG (2021) Digital elevation models of rockfalls and landslides: a review and meta-analysis. Geosciences 11:256. https://doi.org/10.3390/geosciences11060256

Kaminski M (2020) The impact of quality of digital elevation models on the result of landslide susceptibility modeling using the method of weights of evidence. Geosciences 10:488. https://doi.org/10.3390/geosciences10120488

Kamp U, Growley BJ, Khattak GA, Owen LA (2008) GIS-based landslide susceptibility mapping for the 2005 Kashmir earthquake region. Geomorphology 101:631–642. https://doi.org/10.1016/j.geomorph.2008.03.003

Kanungo DP (2015) Landslide susceptibility zonation (LSZ) mapping — a review. Landslides 12:631–640. https://doi.org/10.1007/s10346-015-0586-1

Kayastha P, Dhital MR, De Smedt F (2013) Application of the analytical hierarchy process (AHP) for landslide susceptibility mapping: a case study from the Tinau watershed, west Nepal. Comput Geosci 52:398–408. https://doi.org/10.1016/j.cageo.2012.11.003

Kim J, Lee K, Jeong S, Kim G (2014) GIS-based prediction method of landslide susceptibility using a rainfall infiltration-groundwater flow model. Eng Geol 182:63–78. https://doi.org/10.1016/j.enggeo.2014.09.001

Kornejady A, Ownegh M, Bahremand A (2017) Landslide susceptibility assessment using maximum entropy model with two different data sampling methods. Catena 152:144–162. https://doi.org/10.1016/j.catena.2017.01.010

Lee S (2005) Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data journals. Int J Remote Sens 26:1477–1491. https://doi.org/10.1080/01431160412331331012

Lee S (2019) Current and future status of GIS-based landslide susceptibility mapping: a literature review. Korean J Remote Sens 35:179–193. https://doi.org/10.7780/kjrs.2019.35.1.12

Lee S, Min K (2001) Statistical analysis of landslide susceptibility at Yongin, Korea. Environ Geol 40:1095–1113. https://doi.org/10.1007/s002540100310

Lee S, Pradhan B (2006) Probabilistic landslide hazards and risk mapping on Penang Island, Malaysia. J Earth Syst Sci 115:661–672. https://doi.org/10.1007/s12040-006-0004-0

Lee S, Pradhan B (2007) Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. Landslides 4:33–41. https://doi.org/10.1007/s10346-006-0047-y

Lee S, Ryu JH, Lee MJ, Won JS (2003) Use of an artificial neural network for analysis of the susceptibility to landslides at Boun, Korea. Environ Geol 44:820–833. https://doi.org/10.1007/s00254-003-0825-y

Lee S, Ryu JH, Lee MJ, Won JS (2006) The application of artificial neural networks to landslide susceptibility mapping at Janghung, Korea. Math Geol 38:199–220. https://doi.org/10.1007/s11004-005-9012-x

Lee S, Ryu J-H, Kim I-S (2007) Landslide susceptibility analysis and its verification using likelihood ratio, logistic regression, and artificial neural network models: case study of Youngin, Korea. Landslides 4:327–338. https://doi.org/10.1007/s10346-007-0088-x

Lee J-H, Sameen MI, Pradhan B, Park H-J (2018) Modeling landslide susceptibility in data-scarce environments using optimized data mining and statistical methods. Geomorphology 303:284–298. https://doi.org/10.1016/j.geomorph.2017.12.007

Lee S, Jang J, Kim Y, Cho N, Lee M-J (2020) Susceptibility analysis of the Mt. Umyeon landslide area using a physical slope model and probabilistic method. Remote Sens 12:2663. https://doi.org/10.3390/rs12162663

Leydesdorff L, Rafols I (2012) Interactive overlays: a new method for generating global journal maps from Web-of-Science data. J Informetr 6:318–332. https://doi.org/10.1016/j.joi.2011.11.003

Li Y, Xu ZS, Wang XX, Wang XZ (2020) A bibliometric analysis on deep learning during 2007–2019. Int J Mach Learn Cybern 11:2807–2826. https://doi.org/10.1007/s13042-020-01152-0

Liang Z, Wang CM, Duan ZJ, Liu HL, Liu XY, Khan KUJ (2021) A hybrid model consisting of supervised and unsupervised learning for landslide susceptibility mapping. Remote Sens 13:1464. https://doi.org/10.3390/rs13081464

Lin L, Lin QG, Wang Y (2017) Landslide susceptibility mapping on a global scale using the method of logistic regression. Nat Hazards Earth Syst Sci 17:1411–1424. https://doi.org/10.5194/nhess-17-1411-2017

Ling SX, Chigira M (2020) Characteristics and triggers of earthquake induced landslides of pyroclastic fall deposits: an example from Hachinohe during the 1968 M7.9 tokachi-Oki earthquake, Japan. Eng Geol 264:105301. https://doi.org/10.1016/j.enggeo.2019.105301

Ling SX, Sun CW, Li XN, Ren Y, Xu JX, Huang T (2021) Characterizing the distribution pattern and geologic and geomorphic controls on earthquake-triggered landslide occurrence during the 2017 Ms 7.0 Jiuzhaigou earthquake, Sichuan. China Landslides 18:1275–1291. https://doi.org/10.1007/s10346-020-01549-6

Ling SX, Zhao SY, Huang JP, Zhang XT (2022) Landslide susceptibility assessment using statistical and machine learning techniques: a case study in the upper reaches of the Minjiang River, southwestern China. Front Earth Sci 10:986172. https://doi.org/10.3389/feart.2022.986172

Liu HH, Li XG, Meng T, Liu YY (2020) Susceptibility mapping of damming landslide based on slope unit using frequency ratio model. Arabian J Geosci 13:790. https://doi.org/10.1007/s12517-020-05689-w

Luo X, Lin F, Chen Y, Zhu S, Xu Z, Huo Z, Yu M, Peng J (2019) Coupling logistic model tree and random subspace to predict the landslide susceptibility areas with considering the uncertainty of environmental features. Sci Rep 9:15369. https://doi.org/10.1038/s41598-019-51941-z

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

Mahalingam R, Olsen MJ (2016) Evaluation of the influence of source and spatial resolution of DEMs on derivative products used in landslide mapping. Geomat Nat Haz Risk 7:1835–1855. https://doi.org/10.1080/19475705.2015.1115431

Mahalingam R, Olsen MJ, O'Banion MS (2016) Evaluation of landslide susceptibility mapping techniques using lidar-derived conditioning factors (Oregon case study). Geomat Nat Haz Risk 7:1884–1907. https://doi.org/10.1080/19475705.2016.1172520

Mahdadi F, Boumezbeur A, Hadji R, Kanungo DP, Zahri F (2018) GIS-based landslide susceptibility assessment using statistical models: a case study from Souk Ahras province, N-E Algeria. Arabian J Geosci 11:476. https://doi.org/10.1007/s12517-018-3770-5

Mancini F, Ceppi C, Ritrovato G (2010) GIS and statistical analysis for landslide susceptibility mapping in the Daunia area, Italy. Nat Hazards Earth Syst Sci 10:1851–1864. https://doi.org/10.5194/nhess-10-1851-2010

Maxwell AE, Sharma M, Kite JS, Donaldson KA, Thompson JA, Bell ML, Maynard SM (2020) Slope failure prediction using random forest machine learning and LiDAR in an eroded folded mountain belt. Remote Sens 12:486. https://doi.org/10.3390/rs12030486

Merghadi A, Yunus AP, Dou J, Whiteley J, ThaiPham B, Bui DT, Avtar R, Abderrahmane B (2020) Machine learning methods for landslide susceptibility studies: a comparative overview of

algorithm performance. Earth Sci Rev 207:103225. https://doi.org/10.1016/j.earscirev.2020.103225

Mohan A, Singh AK, Kumar B, Dwivedi R (2021) Review on remote sensing methods for landslide detection using machine and deep learning. Trans Emerging Telecommun Technol 32:e3998. https://doi.org/10.1002/ett.3998

Neuland H (1976) A prediction model of landslips. Catena 3:215–230. https://doi.org/10.1016/0341-8162(76)90011-4

Nhu VH, Shirzadi A, Shahabi H, Chen W, Clague JJ, Geertsema M, Jaafari A, Avand M, Miraki S, Asl DT, Pham BT, Bin Ahmad B, Lee S (2020) Shallow landslide susceptibility mapping by random forest base classifier and its ensembles in a semi-arid region of Iran. Forests 11:421. https://doi.org/10.3390/f11040421

Niu RQ, Wu XL, Yao DK, Peng L, Ai L, Peng JH (2014) Susceptibility assessment of landslides triggered by the Lushan earthquake, April 20, 2013, China. IEEE J Sel Top Appl Earth Obs Remote Sens 7:3979–3992. https://doi.org/10.1109/jstars.2014.2308553

Pandey VK, Pourghasemi HR, Sharma MC (2020) Landslide susceptibility mapping using maximum entropy and support vector machine models along the highway corridor, Garhwal Himalaya. Geocarto Int 35:168–187. https://doi.org/10.1080/10106049.2018.1510038

Pardeshi SD, Autade SE, Pardeshi SS (2013) Landslide hazard assessment: recent trends and techniques. Springerplus 2:523. https://doi.org/10.1186/2193-1801-2-523

Pareek N, Pal S, Sharma ML, Arora MK (2013) Study of effect of seismic displacements on landslide susceptibility zonation (LSZ) in Garhwal Himalayan region of India using GIS and remote sensing techniques. Comput Geosci 61:50–63. https://doi.org/10.1016/j.cageo.2013.07.018

Park NW, Chi KH (2008) Quantitative assessment of landslide susceptibility using high-resolution remote sensing data and a generalized additive model. Int J Remote Sens 29:247–264. https://doi.org/10.1080/01431160701227661

Paulin GL, Bursik M, Lugo-Hubp J, Orozco JJZ (2010) Effect of pixel size on cartographic representation of shallow and deep-seated. landslide, and its collateral effects on the forecasting of landslides by SINMAP and Multiple Logistic Regression landslide models. Phys Chem Earth 35:137–148. https://doi.org/10.1016/j.pce.2010.04.008

Pham BT, Bui DT, Prakash I (2018) Bagging based support vector machines for spatial prediction of landslides. Environ Earth Sci 77:146. https://doi.org/10.1007/s12665-018-7268-y

Pham BT, Phong TV, Nguyen-Thoi T, Trinh PT, Tran QC, Ho LS, Singh SK, Duyen TTT, Nguyen LT, Le HQ, Le HV, Hanh NTB, Quoc NK, Prakash I (2020) GIS-based ensemble soft computing models for landslide susceptibility mapping. Adv Space Res 66:1303–1320. https://doi.org/10.1016/j.asr.2020.05.016

Pourghasemi HR, Rahmati O (2018) Prediction of the landslide susceptibility: which algorithm, which precision? Catena 162:177–192. https://doi.org/10.1016/j.catena.2017.11.022

Pourghasemi HR, Yansari ZT, Panagos P, Pradhan B (2018) Analysis and evaluation of landslide susceptibility: a review on articles published during 2005–2016 (periods of 2005–2012 and 2013–2016). Arabian J Geosci 11:193. https://doi.org/10.1007/s12517-018-3531-5

Pourghasemi HR, Kornejady A, Kerle N, Shabani F (2020) Investigating the effects of different landslide positioning techniques, landslide partitioning approaches, and presence-absence balances on landslide susceptibility mapping. Catena 187:104364. https://doi.org/10.1016/j.catena.2019.104364

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. Environ Earth Sci 63:329–349. https://doi.org/10.1007/s12665-010-0705-1

Pradhan B (2013) A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. Comput Geosci 51:350–365. https://doi.org/10.1016/j.cageo.2012.08.023

Pradhan B, Buchroithner MF (2010) Comparison and validation of landslide susceptibility maps using an artificial neural network model for three test areas in Malaysia. Environ Eng Geosci 16:107–126. https://doi.org/10.2113/gseegeosci.16.2.107

Pradhan AMS, Kim YT (2016) Evaluation of a combined spatial multi-criteria evaluation model and deterministic model for landslide susceptibility mapping. Catena 140:125–139. https://doi.org/10.1016/j.catena.2016.01.022

Pradhan B, Lee S (2010a) Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. Environ Earth Sci 60:1037–1054. https://doi.org/10.1007/s12665-009-0245-8

Pradhan B, Lee S (2010b) Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. Environ ModeLl Softw 25:747–759. https://doi.org/10.1016/j.envsoft.2009.10.016

Pradhan B, Singh RP, Buchroithner MF (2006) Estimation of stress and its use in evaluation of landslide prone regions using remote sensing data. Adv Space Res 37:698–709. https://doi.org/10.1016/j.asr.2005.03.137

Pradhan B, Lee S, Mansor S, Buchroithner M, Jamaluddin N, Khujaimah Z (2008) Utilization of optical remote sensing data and geographic information system tools for regional landslide hazard analysis by using binomial logistic regression model. J Appl Remote Sens 2:023542. https://doi.org/10.1117/1.3026536

Qin SW, Lv JF, Cao C, Ma ZJ, Hu XY, Liu F, Qiao SS, Dou Q (2019) Mapping debris flow susceptibility based on watershed unit and grid cell unit: a comparison study. Geomat Nat Haz Risk 10:1648–1666. https://doi.org/10.1080/19475705.2019.1604572

Qin Z, Zhao Z, Xia L, Ohore OE (2022) Research trends and hotspots of aquatic biofilms in freshwater environment during the last three decades: a critical review and bibliometric analysis. Environ Sci Pollut Res 29:47915–47930. https://doi.org/10.1007/s11356-022-20238-6

Regmi NR, Giardino JR, McDonald EV, Vitek JD (2014) A comparison of logistic regression-based models of susceptibility to landslides in western Colorado, USA. Landslides 11:247–262. https://doi.org/10.1007/s10346-012-0380-2

Reichenbach P, Busca C, Mondini AC, Rossi M (2014) The influence of land use change on landslide susceptibility zonation: the Briga catchment test site (Messina, Italy). Environ Manage 54:1372–1384. https://doi.org/10.1007/s00267-014-0357-0

Reichenbach P, Rossi M, Malamud BD, Mihir M, Guzzetti F (2018) A review of statistically-based landslide susceptibility models. Earth Sci Rev 180:60–91. https://doi.org/10.1016/j.earscirev.2018.03.001

Rosi A, Tofani V, Tanteri L, Stefanelli CT, Agostini A, Catani F, Casagli N (2018) The new landslide inventory of Tuscany (Italy) updated with PS-InSAR: geomorphological features and landslide distribution. Landslides 15:5–19. https://doi.org/10.1007/s10346-017-0861-4

Saadatkhah N, Kassim A, Lee LM (2015) Susceptibility assessment of shallow landslides in Hulu Kelang Area, Kuala Lumpur, Malaysia using analytical hierarchy process and frequency ratio. Geotech Geol Eng 33:43–57. https://doi.org/10.1007/s10706-014-9818-8

Sahin EK, Colkesen I, Kavzoglu T (2020) A comparative assessment of canonical correlation forest, random forest, rotation forest and logistic regression methods for landslide susceptibility mapping. Geocarto Int 35:341–363. https://doi.org/10.1080/10106049.2018.1516248

Saito T, Rehmsmeier M (2015) The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS ONE 10:e0118432. https://doi.org/10.1371/journal.pone.0118432

Saleem N, Huq ME, Twumasi NYD, Javed A, Sajjad A (2019) Parameters derived from and/or used with digital elevation models (DEMs) for landslide susceptibility mapping and landslide risk assessment: a review. ISPRS Int J Geo-Inf 8:545. https://doi.org/10.3390/ijgi8120545

Sameen MI, Pradhan B, Bui DT, Alamri AM (2020a) Systematic sample subdividing strategy for training landslide susceptibility models. CATENA 187:104358. https://doi.org/10.1016/j.catena.2019.104358

Sameen MI, Pradhan B, Lee S (2020b) Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment. CATENA 186:104249. https://doi.org/10.1016/j.catena.2019.104249

Sassa K (2019) Registered speakers of the Fifth World Landslide Forum-Implementing and Monitoring the Sendai Landslide Partnerships 2015–2025-Voluntary contribution to the Sendai Framework 2015–2030 and the Agenda 2030-Sustainable Development Goals. Landslides 16:1423–1424. https://doi.org/10.1007/s10346-019-01213-8

Scaioni M, Longoni L, Melillo V, Papini M (2014) Remote sensing for landslide investigations: an overview of recent achievements and perspectives. Remote Sens 6:9600–9652. https://doi.org/10.3390/rs6109600

Shahabi H, Ahmad BB, Khezri S (2013) Evaluation and comparison of bivariate and multivariate statistical methods for landslide susceptibility mapping (case study: Zab basin). Arabian J Geosci 6:3885–3907. https://doi.org/10.1007/s12517-012-0650-2

Shahabi H, Khezri S, Bin Ahmad B, Hashim M (2014) Landslide susceptibility mapping at central Zab basin, Iran: a comparison between analytical hierarchy process, frequency ratio and logistic regression models. CATENA 115:55–70. https://doi.org/10.1016/j.catena.2013.11.014

Shano L, Raghuvanshi TK, Meten M (2020) Landslide susceptibility evaluation and hazard zonation techniques - a review. Geoenvironmental Disasters 7:18. https://doi.org/10.1186/s40677-020-00152-0

Sharma LP, Patel N, Ghose MK, Debnath P (2011) Landslide vulnerability assessment and zonation through ranking of causative parameters based on landslide density-derived statistical indicators. Geocarto Int 26:491–504. https://doi.org/10.1080/10106049.2011.598951

Shirzadi A, Soliamani K, Habibnejhad M, Kavian A, Chapi K, Shahabi H, Chen W, Khosravi K, Binh Thai P, Pradhan B, Ahmad A, Bin Ahmad B, Dieu Tien B (2018) Novel GIS based machine learning algorithms for shallow landslide susceptibility mapping. Sensors 18:3777. https://doi.org/10.3390/s18113777

Shirzadi A, Solaimani K, Roshan MH, Kavian A, Chapi K, Shahabi H, Keesstra S, Bin Ahmad B, Dieu Tien B (2019) Uncertainties of prediction accuracy in shallow landslide modeling: Sample size and raster resolution. Catena 178:172–188. https://doi.org/10.1016/j.catena.2019.03.017

Shou K-J, Yang C-M (2015) Predictive analysis of landslide susceptibility under climate change conditions - a study on the Chingshui River Watershed of Taiwan. Eng Geol 192:46–62. https://doi.org/10.1016/j.enggeo.2015.03.012

Simon N, De Roiste M, Crozier M, Rafek AG (2017) Representing landslides as polygon (areal) or points? How different data types influence the accuracy of landslide susceptibility maps. Sains Malays 46:27–34. https://doi.org/10.17576/jsm-2017-4601-04

Stanley T, Kirschbaum DB (2017) A heuristic approach to global landslide susceptibility mapping. Nat Hazards 87:145–164. https://doi.org/10.1007/s11069-017-2757-y

Stehman SV, Hansen MC, Broich M, Potapov PV (2011) Adapting a global stratified random sample for regional estimation of forest cover change derived from satellite imagery. Remote Sens Environ 115:650–658. https://doi.org/10.1016/j.rse.2010.10.009

Sujatha ER, Kumaravel P, Rajamanickam VG (2012) Landslide susceptibility mapping using remotely sensed data through conditional probability analysis using seed cell and point sampling techniques. J Indian Soc Remote Sens 40:669–678. https://doi.org/10.1007/s12524-011-0192-1

Suzen ML, Doyuran V (2004) A comparison of the GIS based landslide susceptibility assessment methods: multivariate versus bivariate. Environ Geol 45:665–679. https://doi.org/10.1007/s00254-003-0917-8

Tang YM, Feng F, Guo ZZ, Feng W, Li ZG, Wang JY, Sun QY, Ma HN, Li YN (2020) Integrating principal component analysis with statistically-based models for analysis of causal factors and landslide susceptibility mapping: a comparative study from the loess plateau area in Shanxi (China). J Cleaner Prod 277:124159. https://doi.org/10.1016/j.jclepro.2020.124159

Tanoli JI, Chen N, Regmi AD, Jun L (2017) Spatial distribution analysis and susceptibility mapping of landslides triggered before and after Mw7.8 Gorkha earthquake along Upper Bhote Koshi, Nepal. Arabian J Geosci 10:277. https://doi.org/10.1007/s12517-017-3026-9

Tao J, Qiu DY, Yang FQ, Duan ZP (2020) A bibliometric analysis of human reliability research. J Cleaner Prod 260:121041. https://doi.org/10.1016/j.jclepro.2020.121041

Vakhshoori V, Zare M (2018) Is the ROC curve a reliable tool to compare the validity of landslide susceptibility maps? Geomat Nat Haz Risk 9:249–266. https://doi.org/10.1080/19475705.2018.1424043

van Eck NJ, Waltman L (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 84:523–538. https://doi.org/10.1007/s11192-009-0146-3

van Westen CJ, Castellanos E, Kuriakose SL (2008) Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview. Eng Geol 102:112–131. https://doi.org/10.1016/j.enggeo.2008.03.010

Waltman L, van Eck NJ, Noyons ECM (2010) A unified approach to mapping and clustering of bibliometric networks. J Informetr 4:629–635. https://doi.org/10.1016/j.joi.2010.07.002

Wan S (2009) A spatial decision support system for extracting the core factors and thresholds for landslide susceptibililot map. Eng Geol 108:237–251. https://doi.org/10.1016/j.enggeo.2009.06.014

Wang HB, Liu GJ, Xu WY, Wang GH (2005) GIS-based landslide hazard assessment: an overview. Prog Phy Geog 29:548–567. https://doi.org/10.1191/0309133305pp462ra

Wang X, Zhang L, Wang S, Lari S (2014) Regional landslide susceptibility zoning with considering the aggregation of landslide points and the weights of factors. Landslides 11:399–409. https://doi.org/10.1007/s10346-013-0392-6

Wu XL, Chen XY, Zhan FB, Hong S (2015) Global research trends in landslides during 1991–2014: a bibliometric analysis. Landslides 12:1215–1226. https://doi.org/10.1007/s10346-015-0624-z

Wu YL, Ke YT, Chen Z, Liang SY, Zhao HL, Hong HY (2020) Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. Catena 187:104396. https://doi.org/10.1016/j.catena.2019.104396

Xiao PF, Wu DD, Wang JQ (2022) Bibliometric analysis of global research on white rot fungi biotechnology for environmental application. Environ Sci Pollut Res 29:1491–1507. https://doi.org/10.1007/s11356-021-15787-1

Xie Z, Chen G, Meng X, Zhang Y, Qiao L, Tan L (2017) A comparative study of landslide susceptibility mapping using weight of evidence, logistic regression and support vector machine and

evaluated by SBAS-InSAR monitoring: Zhouqu to Wudu segment in Bailong River Basin. China Environ Earth Sci 76:313. https://doi.org/10.1007/s12665-017-6640-7

Xiong YB, Zhou Y, Wang FT, Wang SX, Wang JM, Ji JW, Wang ZQ (2021) Landslide susceptibility mapping using ant colony optimization strategy and deep belief network in Jiuzhaigou Region. IEEE J Sel Top Appl Earth Obs Remote Sens 14:11042–11057. https://doi.org/10.1109/Jstars.2021.3122825

Xu C, Dai FC, Xu XW, Lee YH (2012) GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China. Geomorphology 145:70–80. https://doi.org/10.1016/j.geomorph.2011.12.040

Xu C, Xu XW, Yu GH (2013) Landslides triggered by slipping-fault-generated earthquake on a plateau: an example of the 14 April 2010, Ms 7.1, Yushu. China Earthquake Landslides 10:421–431. https://doi.org/10.1007/s10346-012-0340-x

Yalcin A, Bulut F (2007) Landslide susceptibility mapping using GIS and digital photogrammetric techniques: a case study from Ardesen (NE-Turkey). Nat Hazards 41:201–226. https://doi.org/10.1007/s11069-006-9030-0

Yang JG, Cheng CX, Song CQ, Shen S, Ning LX (2019) Visual analysis of the evolution and focus in landslide research field. J Mt Sci 16:991–1004. https://doi.org/10.1007/s11629-018-5280-z

Yang YF, Chen GH, Reniers G, Goerlandt F (2020) A bibliometric analysis of process safety research in China: understanding safety research progress as a basis for making China's chemical industry more sustainable. J Cleaner Prod 263:121433. https://doi.org/10.1016/j.jclepro.2020.121433

Yi YN, Zhang ZJ, Zhang WC, Xu Q, Deng C, Li QL (2019) GIS-based earthquake-triggered-landslide susceptibility mapping with an integrated weighted index model in Jiuzhaigou region of Sichuan Province, China. Nat Hazards Earth Syst Sci 19:1973–1988. https://doi.org/10.5194/nhess-19-1973-2019

Yi YN, Zhang ZJ, Zhang WC, Jia HH, Zhang JQ (2020) Landslide susceptibility mapping using multiscale sampling strategy and convolutional neural network: a case study in Jiuzhaigou region. Catena 195:104851. https://doi.org/10.1016/j.catena.2020.104851

Yilmaz I (2009) Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat-Turkey). Comput Geosci 35:1125–1138. https://doi.org/10.1016/j.cageo.2008.08.007

Yilmaz I (2010) The effect of the sampling strategies on the landslide susceptibility mapping by conditional probability and artificial neural networks. Environ Earth Sci 60:505–519. https://doi.org/10.1007/s12665-009-0191-5

Yilmaz I, Keskin I (2009) GIS based statistical and physical approaches to landslide susceptibility mapping (Sebinkarahisar, Turkey). Bull Eng Geol Environ 68:459–471. https://doi.org/10.1007/s10064-009-0188-z

Youssef AM, Pradhan B, Jebur MN, El-Harbi HM (2015) Landslide susceptibility mapping using ensemble bivariate and multivariate statistical models in Fayfa area, Saudi Arabia. Environ Earth Sci 73:3745–3761. https://doi.org/10.1007/s12665-014-3661-3

Yu CL, Chen JP (2020) Application of a GIS-based slope unit method for landslide susceptibility mapping in Helong city: comparative assessment of ICM, AHP, and RF model. Symmetry 12:1848. https://doi.org/10.3390/sym12111848

Yu X, Wang Y, Niu R, Hu Y (2016) A combination of geographically weighted regression, particle swarm optimization and support vector machine for landslide susceptibility mapping: a case study at Wanzhou in the Three Gorges Area, China. Int J Environ Res Public Health 13:487. https://doi.org/10.3390/ijerph13050487

Zezere JL, Pereira S, Melo R, Oliveira SC, Garcia RAC (2017) Mapping landslide susceptibility using data-driven methods. Sci Total Environ 589:250–267. https://doi.org/10.1016/j.scitotenv.2017.02.188

Zhang Y, Chen YP (2020) Research trends and areas of focus on the Chinese Loess Plateau: a bibliometric analysis during 1991–2018. CATENA 194:104798. https://doi.org/10.1016/j.catena.2020.104798

Zhao Y, Wang R, Jiang YJ, Liu HJ, Wei ZL (2019) GIS-based logistic regression for rainfall-induced landslide susceptibility mapping under different grid sizes in Yueqing, Southeastern China. Eng Geol 259:105147. https://doi.org/10.1016/j.enggeo.2019.105147

Zhou CB, Song WY (2021) Digitalization as a way forward: a bibliometric analysis of 20 Years of servitization research. J Cleaner Prod 300:126943. https://doi.org/10.1016/j.jclepro.2021.126943

Zhou C, Yin KL, Cao Y, Ahmed B, Li YY, Catani F, Pourghasemi HR (2018) Landslide susceptibility modeling applying machine learning methods: a case study from Longju in the Three Gorges Reservoir area, China. Comput Geosci 112:23–37. https://doi.org/10.1016/j.cageo.2017.11.019

Zhou XZ, Wen HJ, Li ZW, Zhang H, Zhang WG (2022) An interpretable model for the susceptibility of rainfall-induced shallow landslides based on SHAP and XGBoost. Geocarto Int 1–32. https://doi.org/10.1080/10106049.2022.2076928

Zhuang JQ, Peng JB, Wang GH, Iqbal J, Wang Y, Li W, Xu Q, Zhu XH (2017) Prediction of rainfall-induced shallow landslides in the Loess Plateau, Yan'an, China, using the TRIGRS model. Earth Surf Processes Landforms 42:915–927. https://doi.org/10.1002/esp.4050

Zou Y, Zheng C (2022) A scientometric analysis of predicting methods for identifying the environmental risks caused by landslides. Appl Sci 12:4333. https://doi.org/10.3390/app12094333

Zulhaidi H, Shafri M, Zahidi IM, Abu Bakar S (2010) Development of landslide susceptibility map utilizing remote sensing and Geographic Information Systems (GIS). Disaster Prev Manag 19:59–69. https://doi.org/10.1108/09653561011022144