Review on the progress and future prospects of geological disasters prediction in the era of artificial intelligence

Article *in* Natural Hazards · May 2024

DOI: 10.1007/s11069-024-06673-3

13 authors, including:

Zhang Xiang
China University of Geosciences

Berhanu Keno
Addis Ababa University

Gu Xihui
China University of Geosciences

Chao Wang
Wuhan University

Natural Hazards
https://doi.org/10.1007/s11069-024-06673-3

REVIEW ARTICLE

Review on the progress and future prospects of geological disasters prediction in the era of artificial intelligence

Xiang Zhang[1,2,3] · Minghui Zhang[1,2,3] · Xin Liu[1,2,3] · Berhanu Keno Terfa[4] · Won-Ho Nam[5] · Xihui Gu[6,7,8] · Xu Zhang[1,2,3] · Chao Wang[9] · Jian Yang[10] · Peng Wang[11] · Chenghong Hu[12] · Wenkui Wu[1] · Nengcheng Chen[1,2]

Received: 20 August 2023 / Accepted: 7 May 2024

Abstract

Geological disasters such as landslide, debris flow and collapse are major natural disasters faced by both China and the world, which seriously threaten people's lives, property security and the socio-economic development. Although the method of using the paradigm of traditional mathematical statistics and physical model to predict the low-probability events of geological disasters have been developed for decades, the difficulty of accurate prediction still remains significant, which is recognized as a major and urgent scientific challenge in the field of Earth science. Artificial intelligence is an important driving force for a new round of scientific and technological revolution and industrial transformation. However, how to systematically establish the AI prediction paradigm for low-probability events of geological disasters and deeply coupled with the physical mechanisms of geological disaster evolution and AI learning models still remains as a scientific bottleneck at the intersection of Earth science and information science. In order to clarify the latest research progress of AI prediction of geological disasters such as landslide, collapse and debris flow, this paper first quantifies the current status of global geological disasters and the urgency of prediction, and then summarizes the overall methodology of AI prediction of geological disasters. In particular, prediction feature selection, data set collection and AI prediction models have been detailly reviewed. Moreover, this review discussed the approaches in establishing the physical-informed AI model for higher accurate, robust, and explainable prediction performance. Subsequently, this paper summarizes the recent research achievements of AI prediction for landslide, collapse, and debris flow. Based on these progresses, we also analyzed the existing problems in the field of AI prediction of geological disasters, and indicated the key directions of AI prediction of geological disasters in the future. This review work is believed to be a critical guidance for future intelligent prediction on the severe geological disasters.

Keywords Landslide · Prediction · Artificial intelligent · Geology

Extended author information available on the last page of the article

1 The background and demand for geological disaster prediction

Geological disasters refer to natural hazards caused by geodynamic activity or abnormal changes in the geological environment, usually including Earth's internal processes (such as earthquakes, volcanic activity, and emissions) and related geophysical processes (such as landslides, mudflows, rockfalls, and avalanches). Geological disasters are the instability of nonlinear complex geological systems, which are characterized by strong suddenness, difficulty in observation and rapid evolution. Geological disasters are difficult to prevent in advance, usually resulting in serious consequences, causing social and economic chaos, damaging the environment, and causing huge losses in life and property (Froude and Petley 2018; Grahn and Jaldell 2016; Guzzetti et al. 2012). In the context of climate change, some characteristics and impacts of geological hazards may undergo alterations. Climate change contributes to environmental instability and an upsurge in extreme weather events, potentially exacerbating the frequency and severity of geological disasters (Wang et al. 2023a, b, c). For instance, increased precipitation can elevate the risk of mountainous area mudslides and landslides, rising sea levels may heighten geological hazards along coastal regions, while droughts can lead to groundwater depletion, triggering ground subsidence and related issues (Casagli et al. 2023; Le Breton et al. 2021; Xue et al. 2021). According to the Global Disaster Data Platform's statistics of global geological disasters notification results from 2003 to 2022, a total of 350 significant geological disasters occurred in the world, although the number only accounted for 5.28% of the global disasters, the population affected by the disasters was as high as 4.43 billion, and the death toll was as high as 16,050 (Fig. 1). Among them, Philippines, India, Indonesia, Mexico, China, and even United States belong to the top countries with the highest frequency of geological disasters in the world. For example, 45 geological disasters have occurred in China in the past 20 years, affecting at least 2.424 million people, killing 3,508 people, and causing direct economic losses of nearly 2.1 billion US dollars. In order to mitigate or prevent geological disasters, it is necessary to make a prediction of them by using data collected from various sources.

Over the past few decades, the prediction models for geological disasters have mainly consisted of four types: physics models, heuristic models, mathematical statistics models, and machine learning (ML) models. Physics models require detailed mapping and monitoring data of local geological disasters, so that the prediction accuracy is theoretically high. Representative physical models include: Saito model (Saito 1965), physical and mechanical model of locked type landslide initiation (Chen et al. 2018a, b), landslide cusp catastrophe model (Li et al. 2009; Qin et al. 2006), displacement dynamics prediction model, Scoops3D model (He et al. 2021a, b; Rashid et al. 2020), etc. However, such methods require a large amount of detailed data, so they are generally suitable for the early prediction of imminent geological disasters where monitoring facilities are comprehensive and are not suitable for large-scale analysis (Bergstra et al. 2013).

Heuristic model is to build a model based on limited information, and then rank and weight landslide influencing factors according to expert opinions and expertise, then parameterize the model (Hansen et al. 1995). Representative heuristic models include exponential method (Ruff and Czurda 2008), fuzzy method (Gorsevski et al. 2006; Stanley and Kirschbaum 2017), etc. This method is generally difficult to quantify or evaluate results objectively.

Mathematical statistics model is also widely used in geological disaster prediction. The main idea is to make mathematical inferences about the future development and change

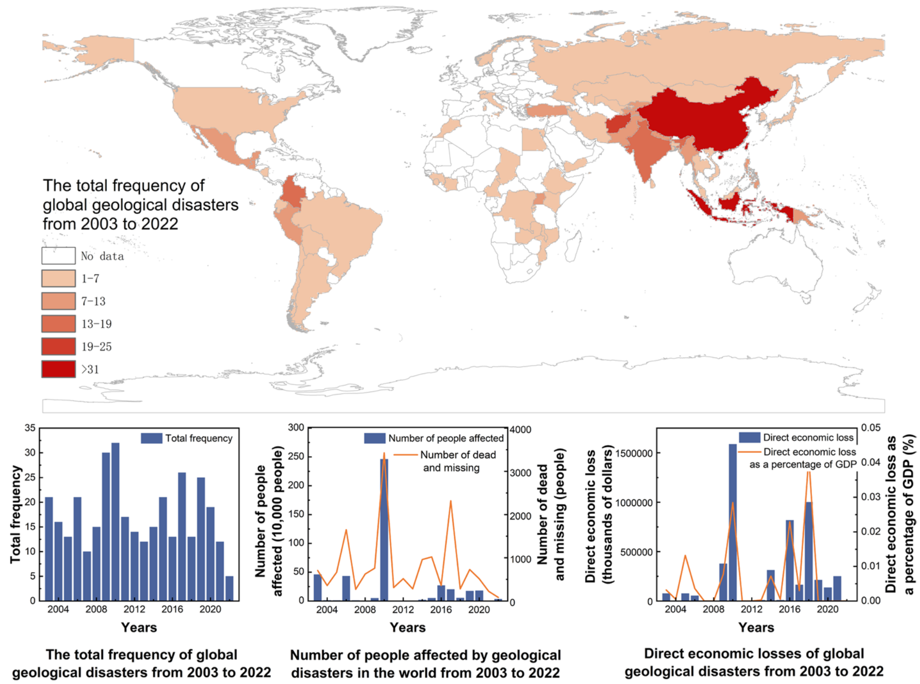

Fig. 1 Global geological disaster distribution map 2003–2022. The above statistics meet at least one of the three conditions of population loss, economic loss, and the government has declared a national emergency or requested international assistance in response to the disaster event. Population loss requires that five or more people have been reported dead due to the disaster, regardless of the affected population, and economic losses must be 0.1% (in relative terms) of local GDP or more

according to the digital data reflecting the development law of one geological disaster in the past. Statistical prediction models usually include weight of evidence method, analytic hierarchy process, information content model, and data superposition method (Murillo-García et al. 2019; Reichenbach et al. 2018; Tang et al. 2020; Yalcin 2008).

In the past two decades, with the rapid development of technologies such as the Internet of Things, geospatial information technology, big data mining and artificial intelligence, machine learning methods have begun to be introduced into geological disaster prediction. The latest machine learning algorithms represented by AlphaGo (Wang et al. 2016), Alphafold2 (Jones and Thornton 2022), Dalle2 (Nichol 2022), Graph Cast (Zhang et al. 2020) and ChatGPT (Thorp 2023), fully reflect the robustness, efficiency and independent learning ability when processing massive and complex data. In theory, this kind of intelligent algorithm is also suitable for deep information mining and efficient processing of geological disaster database to eliminate the uncertainty of information so as to be better applied in geological disaster prevention and control (Chang et al. 2019). However, how to deeply integrate artificial intelligence model algorithm and domain data knowledge and systematically establish an AI paradigm for predicting small probability events of geological disasters has become the bottleneck of intersecting frontier science between geoscience and computer science.

Therefore, this paper starts with the analysis of the complexity of geological disaster prediction, and then focuses on the latest progress of AI prediction of three kinds of geological disasters, such as landslide, collapse and debris flow so as to form a systematic understanding of this frontier direction and provide a theoretical basis for the further development of AI prediction paradigm of geological disasters. A landslide refers to a geological phenomenon where a portion of the Earth's surface or rocks moves downhill along a slope, typically occurring in areas with complex terrain such as mountains, hills, or riverbanks (Dai et al. 2002). Landslide is the most common geological disaster, accounting for more than 50% of the total number of geological disasters every year. The reasons for the frequent occurrence of landslide disasters are very complicated, which are influenced by various external environmental factors, such as heavy rainfall, and they are also influenced by human activities, such as manual cutting of the margins of the tableland. Of course, they are also closely related to regional geological conditions. Collapse refers to the sudden instability of a portion of the Earth's surface or rock, causing it to tilt, slide, or collapse downward, typically occurring in steep terrain such as steep slopes, cliffs, or riverbanks (Hewitt et al. 2008). This phenomenon may also be caused by geological structures, soil types, rainfall, among other factors. While similar to landslides, collapses often occur at a faster pace, with the trajectory of the collapsed material being steeper. A debris flow is a highly hazardous geological disaster that typically occurs in steep mountainous areas or valleys at the foot of mountains (Takahashi 2009). It consists of a large amount of debris such as mud, gravel, rocks, and a mixture of fluids, rapidly flowing along riverbeds or valleys. Debris flows are often triggered by natural factors such as heavy rainfall, snowmelt, or earthquakes, but can also result from human activities such as deforestation or improper mining practices. In addition to these three types of geological disasters, there are also GLOF and avalanches, among others, but predicting them is extremely difficult due to their suddenness and localized nature. For instance, the outburst of glacial lakes is typically a sudden event, and the dynamic uncertainty of lake evolution adds to the prediction challenge. Avalanche snowpack structures are complex, and avalanches are typically localized events that often occur without clear precursors, making prediction a major challenge. Therefore, this review paper focuses on landslides, collapses, and debris flows, while not addressing GLOF and avalanches.

2 The challenge of geological disaster prediction

One of the reasons for the high incidence and low prediction success rate of geological disasters is the high complexity of its evolution process. Taking landslide geological disaster as an example, the complexities mainly include the following four aspects (Fig. 2).

2.1 High variability of rock and soil medium and geological structure

Rock and soil medium and geological structure are one of the most important factors affecting slope stability. The complexity of landslide evolution and occurrence first comes from the high variability of rock and soil medium and geological structure, and the complexity is composed of three factors:

2.1.1 Macroscopic multiphase nature of rock and soil

Due to the effect of geological factors such as geological structure and stress history, the natural rock and soil mass is cut into a semi-continuous state by various geological

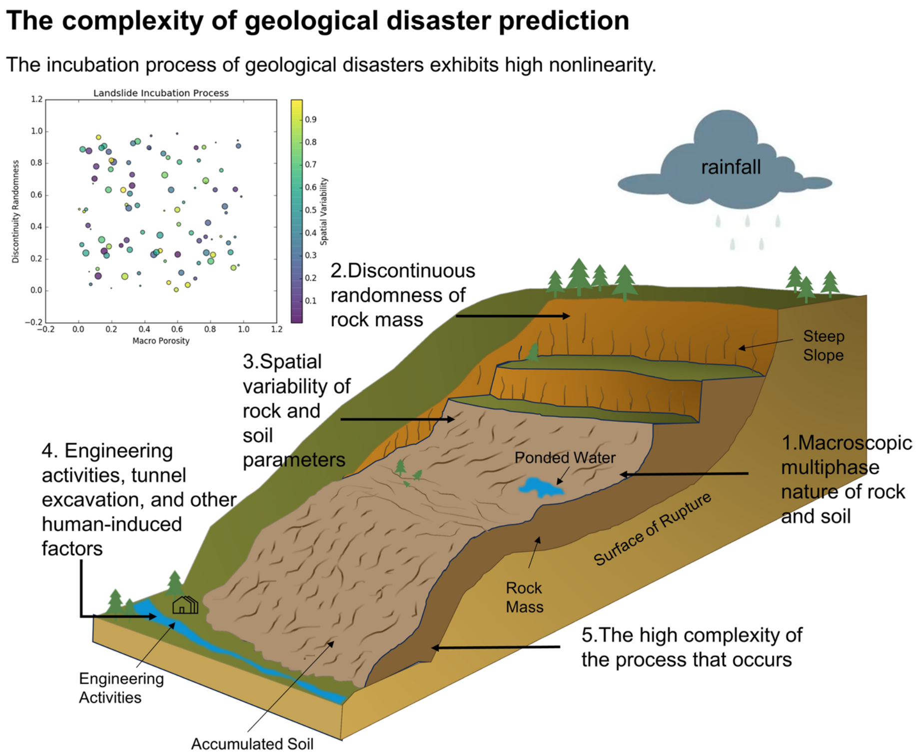

Fig. 2 Complexity of geological disaster prediction problem

movements, so that it is manifested as a two-phase or three-phase random distribution state composed of gas phase (cavity, karst cave), liquid phase (water capsule, formation water) and solid phase (different kinds of geotechnical media) at the macro scale. This state is called the macroscopic multiphase nature of rock and soil.

2.1.2 Discontinuous randomness of rock mass

Under the influence of geological processes, the rock mass shows obvious structural features, which are characterized by obvious discontinuity and randomness. Taking the joint characteristics of rock mass as an example, its distribution characteristics such as occurrence, length, spacing and undulation angle are obvious discontinuity and randomness. Due to the small scale of monomeric, and the large number and density overall, the current engineering investigation scale can't determine the exact spatial distribution state.

2.1.3 Spatial variability of rock and soil parameters

Due to the spatial inconsistency of mineral composition, sedimentary conditions, stress history, ground temperature, moisture content and other geological processes, rock and soil parameters in the three-dimensional space are usually different. Although the properties of different points are different, there is a certain correlation. This correlation is mainly controlled by some factors such as sedimentation and depositional environment, and it mainly manifests that the correlation between two points decreases with the increase of their spacing, and it generally shows the characteristics of gradual and stationary. Therefore, the natural randomness and spatial correlation of dynamic properties of rock and soil mass are called the spatial variability of rock and soil parameters.

2.2 High uncertainty of external disturbance

From the perspective of systematic analysis, the pregnant sliding system of slope is a high-level uncertainty and open system composed of slope surface vegetation, buildings or structures and rock and soil medium by certain ways, and it shows the characteristics of openness, nonlinear, chaotic, infinite dimension, distributed and multi-level. In terms of its openness, the slope system and the external environment are exchanging material and energy all the time. The disturbance of the external environment plays a crucial role in the stability of the pregnant sliding system and even the occurrence of landslide. The uncertainty of the external environment disturbance can be divided into two categories:

2.2.1 Rainfall and other natural factors

The natural factor that has the greatest influence on landslide is rainfall. The rainwater infiltration makes the moisture content of slope soil or the groundwater level increases, so the shearing strength of soil reduces. A large amount of rainwater infiltration leads to saturation of soil and rock layer on the slope, and even causes water accumulation on the waterproof layer at the lower part of the slope, thus it increases the weight of the sliding body and even results in landslide. Other natural factors such as earthquake and long-term high temperature will also have an impact on the slope system and even the occurrence of landslide. The external disturbance of these natural factors shows a high uncertainty, and it cannot be perfectly analyzed now.

2.2.2 Engineering activities and other human factors

Human factors are also one of the important components of external disturbance. Human factors such as the construction of railways, highways, houses, factories and other excavation projects; Tunnel blasting excavation and other mountain excavation engineering and blasting engineering; Overflow and leakage of canals and pools, discharge of waste water of industrial production, agricultural irrigation such as storage and drainage projects; Agricultural projects such as deforestation on slopes and crop cultivation These anthropogenic external disturbances also show a high uncertainty. If the above-mentioned human factors are coupled with adverse natural factors, it is easier to promote the occurrence of geological disasters.

2.3 High nonlinearity of the disaster evolution process

In fact, the occurrence of geological disasters such as landslides is determined by the joint effect of the control variables such as rock and soil medium and external disturbance (Chen et al. 2020a, b). Among these effected factors, which are the dominant ones? Is it geological structure, groundwater, ground stress, rock and soil medium, topography or external disturbances? Different factors play an important role in different regions. The dominance of some factors continuously changes over time through cooperation and competition, and it makes the complex evolutionary behavior of pregnant sliding system (Nowicki Jessee et al. 2018). The evolution of landslide is quite long with a history of decades or even hundreds of years. There must be nonlinear interaction within the pregnancy sliding system which is away from equilibrium (Cui et al. 2022; Zhang et al. 2023a, b). According to the theory of nonlinear dynamics, the nonlinear dynamic equation of landslide evolution can lead to bifurcation and non-uniqueness of the solution. It means that when the control variable of the system changes to a certain critical value, the system will suddenly and randomly choose a path in a variety of development paths and when the relationship between environmental excitation and system response is not strictly proportional as a linear system, the small change of some parameters may lead to the huge response of the system. This is also the main reason why landslide disasters are difficult to predict in the long term.

2.4 High complexity of the occurrence process

Geological disasters are instantaneous or long-term destabilization processes of complex systems, characterized by significant suddenness and uncertainty. Firstly, in terms of scale, the occurrence of geological disasters is often influenced by various factors such as geological structures, changes in groundwater levels, and rainfall, making accurate prediction of their scale challenging. Secondly, the uncertainty of occurrence time is also a major concern. Although we can issue warnings for areas prone to geological disasters based on geological and seismological knowledge, the specific occurrence time is often difficult to predict. Additionally, the movement path of geological disasters is extremely complex, influenced by various factors such as topography, geological structures, and soil characteristics, making their paths difficult to predict. Moreover, in the natural environment, various geological disasters often do not occur in isolation but are interconnected and mutually influential. One type of geological disaster is usually coupled with other types, leading to more severe damage. The occurrence of this disaster chain complicates prediction and assessment (Cao et al. 2020; Lan et al. 2022; Liu et al. 2015). For example, earthquakes may trigger landslides, collapses, or ground liquefaction, as well as other types of disasters such as floods, fires, or building collapses (Fan et al. 2019; Kargel et al. 2016; Zhang et al. 2024). There may be mutually reinforcing or superimposing relationships between these different types of disasters, further exacerbating the extent of damage. All of these factors pose challenges to disaster prevention and control.

3 Methodology framework of AI prediction paradigm for geological disasters

To address the above complexities in predicting geological disasters, AI paradigm is introduced in recent years for compensating the limitations of current methods. As shown in Fig. 3, there are eight steps for AI prediction of geological disasters (Baghbani et al. 2022; Ma et al. 2021; SS and Shaji 2022).

1. Determine the physical mechanism: Determine the physical mechanism of the occurrence of geological disasters and the physical characteristics of geological bodies;
2. Data acquisition: Acquire geological disaster data and data related to physical processes, including input parameters and corresponding outputs, the data can come from experi-

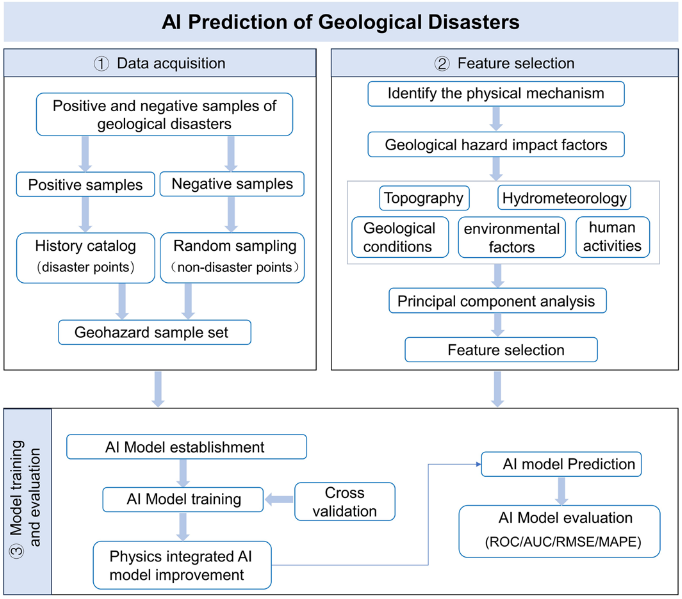

Fig. 3 Overall approach to AI prediction of geological disasters

mental observation, numerical simulation or other sources. Geological hazard sample dataset includes positive samples (disaster points) and negative samples (non-disaster points), and sampling should consider the ratio between the quantity of positive and negative samples (Chen et al. 2023; Jiang et al. 2017; Wang et al. 2023a, b, c);

3. Feature selection: Select the prediction features according to the characteristics of geological disasters, transform the collected data into feature vectors that can be input into the machine learning model, conduct principal component analysis on the feature factors to reduce the dimension of the feature factors, obtain the impact factors with high correlation, and resample the database to create a training dataset;
4. AI Model establishment: Select appropriate machine learning models, such as neural networks, support vector machines and decision trees;
5. AI Model training: Train the machine learning model by using the collected data and feature vectors. In the training process, cross-validation and other technologies can be used to optimize the parameters of the model and improve the accuracy of prediction;
6. Physics integrated AI model improvement: Combine the physical model with machine learning model to form physical informed neural network (PINN) with physical priori, it can improve the accuracy and efficiency of prediction. For example, the output of the machine learning model can be used as input to the physical model, or the output of the physical model can be used to optimize the parameters of the machine learning model;
7. AI model prediction: Predict with the trained model, input specific input parameters, and predict the corresponding spatial, temporal and intensity results;
8. AI Model evaluation: Evaluate and compare the predictive ability of ML algorithms by using receiver operating characteristic (ROC) curve, area under ROC curve (AUC), root mean square error (RMSE) and mean absolute percentage error (MAPE).

Through the analysis of this paradigm, we can find that feature selection, data set construction and AI model are the core of it. Therefore, the following chapters will carry out the analysis from these three aspects.

4 Features and data sets of AI prediction of geological disasters

The occurrence of geological disasters is closely related to the changes of geological structure, surface factors and hydrological conditions in the local regions. In different regions, the geological environment is different, the regulatory factors affecting geological disasters are also different. Therefore, the selection of prediction features should be considered firstly in AI prediction of geological disasters (Kavzoglu et al. 2019). The features of AI prediction of geological disasters mainly include six aspects: topography and landforms, hydrology and meteorology, geological conditions, soil features, vegetable features and human activities. They are critical for AI-based prediction, as shown in Table 1 and the following sections.

4.1 Topography and landforms

Topography refers to the change of terrain, that is, the form of the surface, divided into five basic terrain including plateau, mountain, plain, hill, and basin. The features of topography and geomorphology in AI prediction of geological disasters include slope, slope direction, curvature, relief degree and topographic index. For example, Sarkhoon Basin in Iran, which is located in mountains, selected topographic and geomorphic feature factors such as elevation, slope Angle, slope direction, total curvature, profile curvature, plane curvature, longitudinal curvature, tangential curvature, topographic position index (TPI), topographic wetness index (TWI), and topographic roughness index (TRI). As landslide regulatory factors to predict the spatial distribution of landslides in the study area (Tien Bui et al. 2019). In the prediction of landslide susceptibility in the Vietnamese region, three topographic factors—slope aspect, slope gradient, and curvature—have been selected as partial input data. The Relief-F feature selection method has been employed to quantify the efficacy of various regulating factors within the landslide prediction model (Van Dao et al. 2020).

Topographic data package contains slope, slope direction, curvature, relief degree, surface cutting depth, slope height, slope length, geomorphic type, etc. These data can be extracted from DEM data (Pourghasemi et al. 2013; Vaze et al. 2010). The commonly used DEM data include the GEBCO Colorized World Map, ETOPO1 Global Relief Model, SRTM (Shuttle Radar Topography Mission) data, and ALOS (Advanced Land Observing Satellite) data from Japan (Li and Zhao 2018). InSAR ground deformation data are also commonly used in topographic data. Interferometric synthetic aperture radar (InSAR) uses the phase information of radar images to detect the small deformation in the visual range direction of radar satellites, and can obtain the large-scale slow surface deformation of a long time series. InSAR technology can monitor the surface deformation of some geological disasters, such as landslides, and identify landslides. Providing data basis for geological disaster data analysis (Hu et al. 2014; Osmanoğlu et al. 2016).

Table 1 Features and data sets of AI prediction of geological disasters

Feature category	Predictive features	Description	Feature application
Topography and landforms	Slope	The degree of steepness of the surface element.	Bui et al. (2020a, b), Luo et al. (2019)
	Aspect	The direction of the projection of the slope normal on the horizontal plane.	Chang et al. (2023), Guo et al. (2015)
	Curvature	A measure of the curvature of the ground, plane including curvature and profile curvature.	Hong et al. (2019a, b), Sun et al. (2020)
	Degree of relief	The difference in elevation of adjacent areas.	Chen et al. (2019), Wang et al. (2015)
Hydrology and meteorology	Rainfall	The depth of the water layer that falls to the ground in a certain period of time and it is an important basis for the calculation of regional water resources.	Di et al. (2019), Harsa et al. (2023)
	Reservoir water level	One of the induced factors of reservoir landslide.	Nava et al. (2023), Zhang et al. (2022a, b)
	Groundwater table	Temporal and spatial variation of groundwater table.	Li et al. (2014), Liu et al. (2021a, b)
	River network	A buffer zone around Grade I river or eroding river.	Conforti et al. (2014), de Oliveira et al. (2019)
Geological conditions	Fault	The distance from an active fault or the width of the fault zone is an important factor in predictive mapping.	Demir (2019), Qi et al. (2021)
	Lithology	Lithology maps based on engineering features rather than stratigraphic classification.	Lei et al. (2020), Segoni et al. (2020)
	Geological structure	The relationship among geological structure, slope angle and orientation are relevant to predicting rock slide.	Gomez and Kavzoglu (2005), Zhao et al. (2022a, b)
Soil features	Soil moisture	Temporal and spatial variations in soil water content are one of the main components of stability analysis.	Lin et al. (2017), Ray et al. (2010)
	Soil type	Engineered soil type based on genetic or geotechnical classification	Bui et al. (2020a, b), Lee et al. (2015)
	Soil depth	Soil depth based on Borehole, geophysical, and outcrop is a key data layer in stability analysis.	Qiu et al. (2019), Saha and Saha (2022)
	Soil parameters	Hydraulic conductivity, soil bulk density, soil friction angle and effective cohesion, etc.	Hung et al. (2012), Ray et al. (2020)

Table 1 (continued)

Feature category	Predictive features	Description	Feature application
Vegetation features	NDVI	Detect vegetation growth status, vegetation coverage and eliminate some radiation errors, etc.	Nicu (2017), Sezer et al. (2011)
	NDWI	Strong water body detection capability enables the identification of potential water-related disaster risks.	Bhagya et al. (2023), Xu et al. (2015)
	SAVI	It can reduce soil influences and improve vegetation monitoring.	Javed (2019), Saad et al. (2021)
Human activities	Land use	The way and condition of human's use of the natural property of land.	Hu et al. (2020), Pham et al. (2022)
	Distance from road	The closer the distance to the river, the larger the area of landslide development.	Chen et al. (2020a, b), Di Napoli et al. (2020)
	Distance from river	The closer the distance to the river, the larger the area of landslide development.	Chen et al. (2018a, b), Selamat et al. (2022)
	Distance from settlements	Human intrusion has impact on the environment.	Abu El-Magd et al. (2021), Yilmaz (2010)

4.2 Hydrology and meteorology

Hydrological and meteorological factors that have a great impact on geological disasters include rainfall, reservoir water level, groundwater level and river network. Rainfall is an important factor that inducing landslide deformation and it has a certain lag. Considering the lag effect of rainfall in landslide displacement prediction, the reliability of prediction can be improved. The reservoir water level has a certain influence on landslide deformation. For example, taking Langshuwan landslide as an example, based on the mechanism of landslide deformation and destruction, grey relational analysis method was used to evaluate the influence of reservoir water level on the periodic displacement of landslide (Zhang et al. 2021a, b). As the main groundwater supply of Jinle landslide area is atmospheric precipitation, and the slope in this area is steep, it is conducive to the centralized drainage of surface water, so rainfall intensity and water level difference are selected as hydrological prediction factors (Li et al. 2014).

Hydrometeorological data comprises rainfall, reservoir levels, groundwater levels, river networks, among other elements. Precipitation data such as GPM (Global Precipitation Measurement) and TRMM (Tropical Rainfall Measuring Mission) can be obtained from the official website of NASA, offering global precipitation measurement data at various temporal and spatial scales (Groisman and Legates 1994; Huffman et al. 1997). The water level of lakes is recognized as Essential Climate Variable (ECV) by the Global Climate Observing System (GCOS). The number of water levels can be set by Copernicus Global Land Service, and the number of reservoir water levels can be set by Global Reservoirs and Lakes Monitor (G-REALM) and Open Government Data (OGD) Platform India. The groundwater level data can be obtained through the DWR continuous groundwater level measurement dataset. The dataset contains the continuous temporal series data of recorder from the Ministry of Water Resources operation site (https://data.cnra.ca.gov/dataset/continuous-groundwater-level-measurements). The river network is available from WWF Global River Data HydroSHEDS, which is a mapping product that provides hydrological information in a consistent format for regional and global-scale applications, and it is developed by the WWF Conservation Science Program, United States Geological Survey, International Center for Tropical Agriculture, Nature Conservancy and the Center for Environmental Systems Research at the University of Kassel, Germany (Wickel et al. 2007).

4.3 Geological condition

Geological factors are crucial to the AI prediction of geological disasters. In order to improve the accuracy of AI prediction, it is necessary to consider the physical and mechanical processes and the destruction mechanisms of geological disasters. At present, the geological factors involved in research include fault, lithology, geological structure and so on (Segoni et al. 2020; Van Westen et al. 2008). The geological structures can be extracted by a combination of UAV photogrammetry and aerial LiDAR. And the slope geometry can be evaluated with slope, slope direction, profile and plane curvature and cliff height (Pham et al. 2019; Van Dao et al. 2020). However, for the landslide prediction of steep coast, as the main influencing factors of spatial variation of coastal landslide, marine erosion conditions and slope geometry conditions should be considered (He et al. 2021a, b). Since landslides are related to faults and fractures, fault information is often used as one of the environmental factors in landslide disaster statistical assessment (Luo et al. 2019).

The occurrence of geological disasters is closely related to the geological structure of the local region. Geological bodies are the basic material components of geological disasters, and the physical and mechanical behaviors of the whole evolution process of geological disasters have an intrinsic mapping relationship with the macro and micro mechanical behaviors of geological bodies. Geological data include lithologic structure, fault, geological structure, etc., which can be extracted from geological maps (Chen et al. 2017; Kawabata and Bandibas 2009). The global geological maps can be acquired from the U.S. National Center for Environmental Information, and the fault data can be obtained from the U.S. Geological Survey.

4.4 Soil features

Among the environmental factors that affect the AI prediction of geological disasters, NDVI has the greatest influence, it is also one of the most frequently selected environmental features in the study. For example, Catani developed a semiempirical method based on geomorphology to obtain the spatial distribution of soil thickness (Catani et al. 2010). Lee determined the spatial distribution of soil thickness in the analysis of slope instability in Taiwan by using the humidity index (Lee and Ho 2009). The spatial distribution of soil thickness related to the moisture index significantly influences the prediction of shallow landslides. However, there is limited information available on soil thickness in landslide-prone areas. Soil parameters such as hydraulic conductivity, soil bulk density, soil friction angle, and effective cohesion can be utilized when soil thickness information is scarce. Alternatively, the spatial distribution of soil thickness can be used to forecast shallow landslides (Ho et al. 2012). The increase of soil water will lead to the decrease of landslide stability, so soil water is one of the important factors of landslide prediction (Lin et al. 2017; Ray et al. 2010; Zhang et al. 2015).

Soil related information such as soil type, soil depth, and soil parameters can be obtained from the World Soil Database (HWSD), which is a grid data with spatial resolution of kilometers and provides information about soil type, soil phase, soil physical and chemical properties of each grid node (Nachtergaele et al. 2010). Soil moisture data is available from Harmonized World Soil Database (HWSD) (Nachtergaele et al. 2010).

4.5 Vegetable features

Vegetation is sensitive to changes in the geological environment. Vegetation indices can utilize remote sensing data to monitor vegetation coverage, assess soil retention capacity, and the stability of the geological environment. They can also be used to evaluate changes in the hydrological environment and the potential risks of hydrological disasters. In the field of geological disaster prediction, commonly used vegetation indices include normalized difference vegetation index (NDVI), normalized difference water index (NDWI) and the soil adjusted vegetation index (SAVI). Among the vegetable features that affect the AI prediction of geological disasters, NDVI has the greatest influence, it is also one of the most frequently selected vegetable features in the study. NDVI can reflect the health and growth status of vegetation, helping to monitor changes in vegetation coverage, thereby assessing the stability of the land surface and the impact of vegetation on the geological environment (Liu et al. 2022; Wang et al. 2021). NDWI has a high sensitivity to water bodies, enabling effective detection and differentiation of water bodies from other land features in remote sensing images. This capability allows NDWI to be utilized in identifying potential risk areas for water-related geological hazards such as floods and debris flows (Ullah et al. 2022).

SAVI can adjust for soil brightness, making it particularly suitable for regions with high soil brightness or where soil exposure may affect traditional vegetation indices. It enhances the contrast between vegetation and bare soil, aiding in more accurate identification of geological hazard risk areas and improving the accuracy of early warning and prediction (Fathi et al. 2015).

4.6 Human activities

On the one hand, human activities such as tunnel construction may have a direct impact on the occurrence of geological disasters. On the other hand, human activities will also change the classification of land use and cause indirect effects. Therefore, human activities are also one of the factors of AI prediction of geological disasters. The factors of human activities in AI prediction of geological disasters include distance from roads, distance from rivers, distance from settlements, etc. And these can be used to evaluate the intensity of human engineering activities (Hu et al. 2020).

The data of human activities mainly include land use data, distance from road data, distance from river data, distance from fault data, distance from settlement data, etc. Road route and river line can be extracted from the city map and fault line from the topographic map. After obtaining the three-line data, the airborne orthogonal distance between the pixel and the line data can be calculated and to create the distance map from the road, the distance from the river and the distance from the fault line (Akgun 2012). Land use data is a kind of data that can reflect the status, characteristics, dynamic changes and distribution characteristics of land use system and land use elements and the human's development utilization, governance, transformation, management, protection, planning of land. At present, free open data with different resolutions can be easily downloaded from the official websites of the European Aviation Administration (Rahman and Szabó, 2021; Vali et al. 2020; Yin et al. 2021). Road data, settlement data are available from OpenStreetMap.

4.7 Geological disaster catalog

The collection of the geological disaster catalog library is the premise of geological disaster prediction. Geological disaster data is mainly collected through historical geological disaster data records, news media reports and geological disaster detection. The current geological disaster detection methods include visual interpretation, field investigation and automatic interpretation. At present, there are few catalogues of geological disasters on the global scale, and the statistical information is not complete. A representative example is an open global landslide catalog proposed by Kirschbaum. The is obtained based on media reports, online databases and other sources, it has a total of 5,741 landslide disaster sites, the time scale is 2007–2013, and the most of the events occurred in Asia, North America and Southeast Asia. It can be used to describe the global pattern of landslide occurrence and assess the relationship of extreme precipitation at regional and global scales. The catalog is currently available at http://ojo-streamer.herokuapp.com/ (Kirschbaum et al. 2015).

5 Machine learning algorithms in AI prediction of geological disasters

5.1 Common machine learning models

At present, AI prediction of geological disasters is mainly implemented by various machine learning algorithms, and its robustness is one of the ideal technologies to solve nonlinear geological environment problems (Baghbani et al. 2022; Ma and Mei 2021a; Tehrani et al. 2022). By using regression or classification, machine learning is able to learn associations between predictors of occurrence and causes of geological disasters without the need of hypothetical structural models. At the same time, machine learning algorithms can handle a wide variety of input data of different types and sizes without the need of any predefined data structures (Goetz et al. 2015; Merghadi et al. 2020). As a result, machine learning currently has a rate of 75–95% in successfully predicting geological disasters such as landslides (Korup and Stolle 2014). However, the performance of different machine learning algorithms will vary significantly with various factors such as study area, geological disaster type and training data (Tehrani et al. 2022). Therefore, it is necessary to first analyze various types of AI algorithms for geological disaster prediction from a theoretical perspective. This paper mainly introduce thirteen machine learning models used for geological disaster AI prediction: logistic regression, decision tree, artificial neural network, support vector machine, random forest, k-nearest neighbors, least squares support vector machine, kernel extreme learning machine, recurrent neural network, long short-term memory network, adaptive neuro-fuzzy inference system, Bayesian network and maximum entropy model (Table 2).

5.2 Improved machine learning models integrating physical mechanisms of geological disasters

In recent years, studies have shown that the accuracy of geological disasters prediction by using only data-driven methods (directly using the aforementioned machine learning models) has reached a bottleneck, and new AI prediction paradigms need to be innovated (Orland et al. 2020). For this reason, the current research focuses on building a machine learning model integrating the physical mechanism of geological disasters by using the PINN strategy and utilizing the advantages of data-driven and model-driven to achieve prediction capability with higher accuracy, lower computational consumption and less data dependent, as shown in Table 3 (Depina et al. 2022; Lu and Mei 2022). For example, Wu et al. found landslide surge propagation process simulation method based on the physics-informed neural network well simulate the following stages of landslide surge propagations (Wu et al. 2022). Chen and Zhang proposed a mechanism simulation network structure designed based on the geological mechanics equation, as illustrated in Fig. 4. Although different geomechanical parameters need to be calculated by different physical models or empirical models, the input information of these physical models basically includes P-wave velocity Vp, S-wave velocity Vs, true vertical depth TVD and density RHOZ. However, in practical, the two wave velocities (Vp and Vs) are often unknown information. Therefore, it is reasonable to assume that the mathematical model used to calculate the geomechanical parameters actually contains two intermediate variables. Based on this assumption, the researchers added a physical constraint layer to the conventional long short term memory model LSTM, which contains two empty neurons and two input variables

Table 2 Machine learning models used in geological disasters

Models	Main principles	Advantages	Disadvantages	Cases	Case references
Logistic regression	Establish regression formulas for classification boundaries based on existing data	It does not require much computing resource; It is easy to understand and implement; It does not require any invocation	It is easy to underfit; The classification accuracy may not be high	Rainfall-type landslide in Ningdu County, Jiangxi Province from 1980 to 2001	Pourghasemi et al. (2013)
Decision tree	The decision model is established in a tree structure model and based on data attributes	Low complexity; Insensitive to the absence of medians; It can deal with irrelevant feature data	Overmatching problems may happen	A typical stepped landslide in the Three Gorges Reservoir area Zhujiadian landslide	Huang et al. (2020), Ma et al. (2017)
Artificial neural network	Simulating biological neural networks is a type of pattern matching algorithm and a huge branch of machine learning, with hundreds of different algorithms	High class accuracy; strong learning ability	A large number of parameters are required; the learning process cannot be observed, and the results are difficult to interpret; and the learning time is long	Caiyuan River Basin, Nanping City, Fujian Province	Moayedi et al. (2019)
Support vector machine	Support vector machine is a supervised learning algorithm used for classification and regression analysis. It identifies an optimal hyperplane within a dataset, effectively separating different categories of data	Low generalization error rate; Low calculation overhead; The results are easy to interpret	It is sensitive to parameter adjustment and kernel function selection; The original classifier is suitable for dealing with second class problems without modification	National landslide disasters from 2016 to 2019	Han et al. (2021)
Random forest	Build a forest in a random way, the forest is composed of many independent decision trees, and finally synthesize the optimal classification result	The limited sample can be fully used; Advantages of variety and accuracy	Can overfit for some noisy problems	Landslide in Qingshan County, Sichuan Province	Hu et al. (2021)

Table 2 (continued)

Models	Main principles	Advantages	Disadvantages	Cases	Case references
K-nearest neighbors	Suppose there is a training data set and the class of the instances is determined. Predict the new instance according to the training instance class of its k nearest neighbors by majority voting	Solve the complex problems by using simple Euclidean measures; Simple design; non-parametric method; It is usually successful in classification cases with very irregular decision boundaries	Non-generalized learning algorithms; Simply stores instances of training data makes it computationally expensive; Classification is calculated from a simple majority vote of each point's nearest neighbors so the value of K needs to be determined	Cox's Bazar district in Bangladesh	Adnan et al. (2020), Tien Bui et al. (2017)
Least squares support vector machine	It is an improvement of the support vector machine. It changes the inequality constraints in the traditional support vector machine to equality constraints, and uses the error sum of squares loss function as the empirical loss of the training set	Convert the solution of quadratic programming problems into the problem of solving a system of linear equations to improve the speed and convergence accuracy of solving the problem	Sensitive to noise; the larger the number of samples, the more difficult the calculation process; the kernel parameters and regularization parameters are unknown; the solution is not sparse	Landslide displacement prediction in the Three Gorges Reservoir area	Wen et al. (2017), Zhu et al. (2017)
Kernel extreme learning machine	It is an improved algorithm based on extreme learning machine and combined with kernel function. KELM can improve the prediction performance of the model while retaining the advantages of ELM	There is no need to set the number of hidden layer nodes in the network; the learning rate is fast and the generalization performance is good	Failure to consider structural risks may lead to overfitting; users cannot fine-tune based on the characteristics of the data set, resulting in poor controllable performance	A long-term deformed landslide along the Jinsha River	Zhou et al. (2018a, b)

Table 2 (continued)

Models	Main principles	Advantages	Disadvantages	Cases	Case references
Recurrent neural network	Remember the previous information and use the previous information to affect the output of subsequent nodes. That is, the input of the hidden layer not only includes the output of the input layer, but also includes the output of the hidden layer at the previous moment	Parameters are shared at different times; sequence data can be processed efficiently	It has long-term dependency issues	Typical loess landslide area in China Gansu Heifangtai Dangchuan 6# landslide	Ngo et al. (2021)
Long short-term memory	It is a special recurrent neural network that introduces a gate mechanism to control the flow and loss of features, allowing information to selectively affect the state of the neural network at each moment	It inherits most of the characteristics of the RNN model, can solve the long-term dependency problem of RNN, and at the same time solves the gradient disappearance problem caused by the gradual reduction of the gradient backpropagation process	It cannot represent short-term, rapidly changing and non-periodic data	Baishuihe landslide and Bazimen landslide in the Three Gorges Reservoir area	Yang et al. (2019)

Table 2 (continued)

Models	Main principles	Advantages	Disadvantages	Cases	Case references
Adaptive neuro fuzzy inference system	The model structure of ANFIS is composed of an adaptive network and a fuzzy inference system. It functionally inherits the interpretability characteristics of the fuzzy inference system and the learning ability of the adaptive network	The system parameters can be changed based on prior knowledge to make the system output closer to the real output	The trained model is not interpretable	West of Seoraksan Mountain, Gangwon-do, South Korea	Lee et al. (2015), Panahi et al. (2020)
Bayesian network	Is a probabilistic graph model without conditional probability independence. It is structured as a directed acyclic graph, in which each node represents a random variable, each node has a corresponding probability distribution table, and directed edges represent dependencies between nodes	It is a simple and fast process to predict the prediction sample, and it is more effective for the multiclassification problem	The training of network structure is complicated	Landslide in Chongren County, Jiangxi Province, China	Chen et al. (2018a, b), Chen and Zhang (2021)
Maximum entropy model	When predicting the probability distribution of a random event, the prediction should satisfy all known constraints and do not make any subjective assumptions about unknown situations	As a classic classification model, it has high accuracy and can flexibly set constraints	Since the number of constraint functions is related to the number of samples, the iterative process requires a huge amount of calculation and is difficult to apply in practice	Landslide in Wanyuan City, Sichuan Province	Chen et al. (2017), Felicísimo et al. (2013), Kornejady et al. (2017)

Table 3 Centralized types that integrate physical prior knowledge

Centralized types that integrate physical prior knowledge	Advantages	Disadvantages
Customizing loss function	It is convenient to e design and implement and multiple regular terms can be added to express various priors.	As a soft constraint, the strict realization of spatiotemporal priors can't be guaranteed.
Customizing network structure	It can strictly meet the specific priors and it designed to be flexible and scalable.	It is technically difficult and requires a deep understanding of deep learning and research problems.
Describing geometric feature	Introducing reasonable geometric features as inductive preferences can significantly improve generalization ability and training efficiency.	The existing definitions of invariance and isotropic are too rough, and the specific dynamic systems lack refined geometric characteristics definitions.
Describing dynamic system	It can refined descript local and remote spatiotemporal coherence and correlation and flexible characterize complex dynamic system.	It lacks of clear governing equation and it is difficult to define and integrate the spatiotemporal prior.
Designing auxiliary task	It can design auxiliary tasks under the framework of meta-learning and integrate prior knowledge by making full use of correlation between tasks.	The reasonable design of auxiliary task is difficult and the auxiliary task may affect the main task is more difficult if the correlation is not strong.
Describing spatiotemporal heterogeneity	It can better describe the geographical dynamic system through the spatiotemporal heterogeneity reflected by the characterization of location and direction.	The model of spatiotemporal heterogeneity is complex, and there are few studies in non-geoscience fields, and there is a lack of methods for reference.

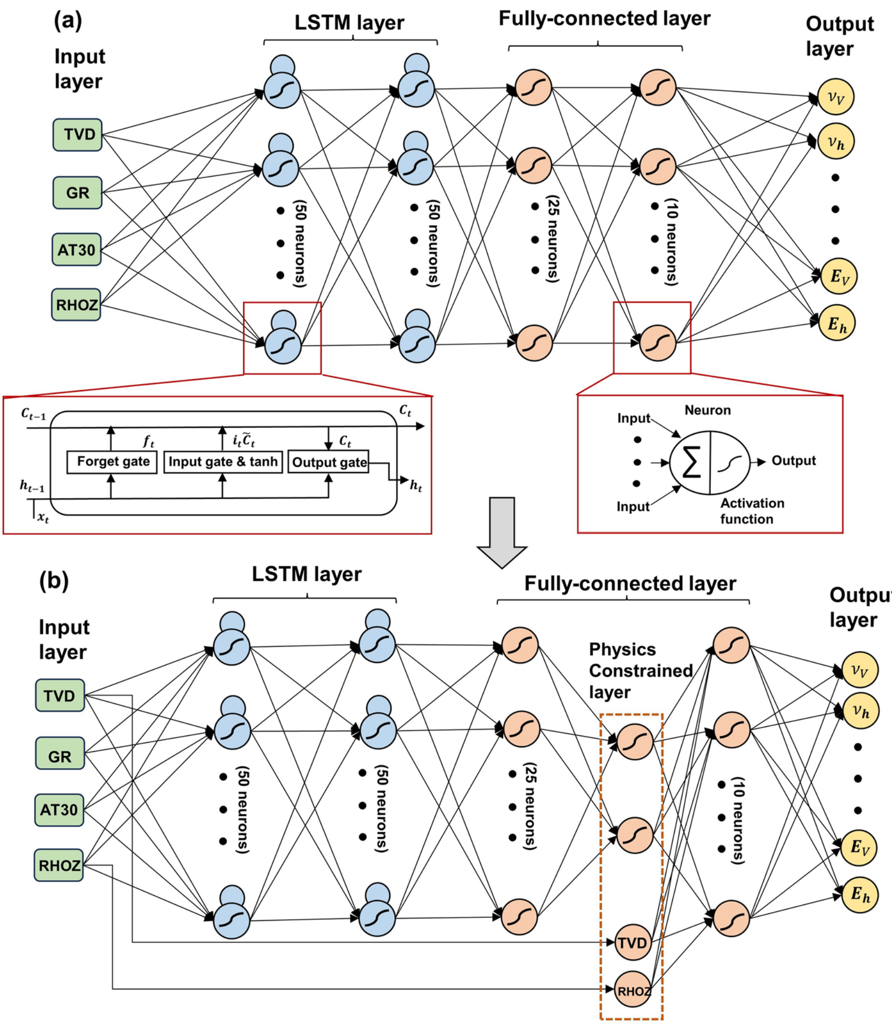

Fig. 4 Mechanism simulation network structure designed based on geomechanical equations

connected from the input middle skip layer. In the training session, supervised learning of the two empty neurons was not possible due to the lack of wave velocity, but the model could be compressed into a 2-dimensional space. In other words, the two empty neurons in the model would be equivalent basis of Vp and Vs in space. Finally, with the adjustment of the network structure, the researchers embedded the knowledge about dimensions in the geomechanical equations into the model. Experiments show that under the same network complexity (similar number of parameters), the physically neural network embedded with knowledge has higher prediction accuracy and it successfully generates costly 3D physical fields with 10 kinds of geomechanical parameters like Young's modulus and Poisson's ratio based on low-cost physical parameters that are easily obtained.

Four typical strategies have been summarized for this kind of machine learning model that integrates the physical mechanism of geological disasters:

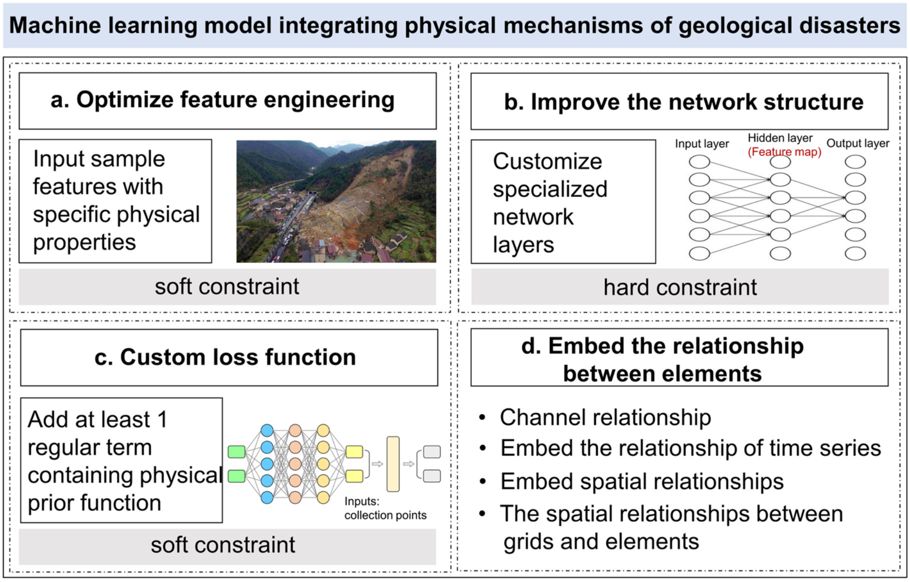

Fig. 5 Four machine learning model strategies integrating the physical mechanisms of geological hazards

1. Optimize feature engineering: It belongs to soft constraint, as shown in part a of Fig. 5. Focus on the selection and input of sample features with specific physical properties like structural control and inducible factors. For example, rainfall landslide focuses on precipitation intensity, and dynamic water landslide focuses on reservoir level. Therefore, it is necessary to select factors related to landslide occurrence and analyze the landslide susceptibility using pattern classification by looking at the relationship between the various aspects and landslide location.
2. Improve the network structure (model structure design): It belongs to hard constraints, as shown in part b of Fig. 5. It includes introducing defined special physical variables in the network structure, solidifying the network parameters into specific physical variable values, customizing special network layers to maintain some characteristics of physical and geometric, introducing special operators to describe dynamic systems, and using special numerical integration methods to improve the stability and robustness of network learning of dynamic systems. Among them, the most popular method is to customize a special network layer and use a variety of specially defined convolution or operators in it. For example, continuous convolution in Lagrange fluid simulation can describe the movement characteristics of dynamic fluid particles more accurately than discrete convolution in space–time. Hard constraints can not only provide the degree of deviation from the predicted value, but also provide a strategy for adjusting the predicted value to a result that conform to the physical constraint. One of the advantages of hard constraint map is that it is only directly related to the governing equation of the problem to be solved, and it will contain no parameters to be trained after the specific problem

has been determined. Therefore, the hard constraint map can be regarded as a special activation function in the neural network.

3. Custom loss function: It belongs to soft constraint, as shown in part c of Fig. 5. Add at least 1 regular term containing physical prior function on the basis of the original loss, which can punish the violation of physical priors or introduce the data prior mode. At present, there are many loss functions that have introduced the principle of fluid mechanics. For example, Zhang et al. provide a new paradigm for incorporating hybrid deep learning frameworks better than the classical Verruijt–Booker solution and a conventional purely data-driven model. In particular, the physical laws are expressed by incorporating the residual of physical governing equations into the loss function, which can be further minimized to search for the optimal neural network structure. It is found that the proposed PINN model can reasonably reproduce ground deformation fields obtained numerically with only a small amount of training data (Zhang et al. 2023a, b).
4. Embedding interrelationships: As shown in part d of Fig. 5, the relationship between elements in a single grid is a channel relationship, so we can use the channel attention mechanism; Embed the relationship of time series: LSTM model or Transformer model is the baseline model; Embed spatial relationships: The spatial relationships between grids and elements can be achieved by using the Attention mechanism.

6 Recent progress in AI prediction of geological disasters

Artificial intelligence (AI) is a key driver of the fourth industrial revolution and has the potential to bring about significant interdisciplinary transformations. Its application in the field of geological disasters holds immense potential, driving disaster prevention and mitigation towards intelligence (Ayawah et al. 2022; Mukherjee et al. 2021). Prediction models based on artificial intelligence incorporate domain knowledge or rules as inputs, include causal relationships, and are supported by powerful computing capabilities. Therefore, they are expected to break through the generalization ability beyond the distribution and surpass the accuracy of simple statistical learning (Yonggang et al. 2016; Zhang et al. 2022a, b).

It should be noted that machine learning is a branch of artificial intelligence, and its application in geological disaster modeling has been rapidly evolving. From current trends, the use of machine learning models has become mainstream in geological disaster prediction (Dikshit et al. 2021). Therefore, this paper primarily focuses on analyzing the specific applications of machine learning methods in geological disaster prediction.

6.1 AI prediction for landslide

Landslide is sudden and destructive, which is a great threat to people's life and property safety. Therefore, it is particularly important to predict and prevent landslide. Landslide prediction encompasses single prediction and regional prediction. Single prediction focuses on forecasting specific landslide events in a particular geographical location or region. Its goal is to proactively identify and predict the likelihood of single landslides, enabling timely preventive and management measures. On the other hand, regional prediction aims to assess the potential risk of landslides occurring across an entire area. The objective is to generate landslide probability or susceptibility maps for the regional scope, guiding comprehensive land-use planning and risk management decisions. In terms of AI methods, both single and regional predictions can employ various artificial intelligence techniques. However, regional landslide prediction typically involves considering a broader range of data and may utilize multi-model integration or hybrid models to comprehensively evaluate the potential occurrence of landslides across the entire region. There are two main aspects of the application of machine learning technology in landslide prediction: landslide susceptibility assessment and landslide displacement prediction.

6.1.1 Landslide susceptibility assessment

Landslide susceptibility assessment refers to inputting landslide adjustment factors as input datasets into prediction models, and dividing the susceptibility maps generated by each prediction model into landslide risk levels, such as extremely low, low, medium, high, and extremely high, in order to predict which areas are more prone to landslides in the future (Hong et al. 2019a, b; Ma et al. 2021; Zezere et al. 2017; Zhou et al. 2021). Susceptibility assessment based on machine learning algorithms is to analyze the spatial relationship between past events and inducing factors by analyzing data characteristics and predicting the spatial probability of landslide occurrence, and the common machine learning algorithms include random forest and support vector machine (GUO et al. 2021). The random forest model can flexibly capture the nonlinear relationships between features, thus better adapting to complex geological disaster prediction tasks (Yang et al. 2023). With the characteristics of small sample size, nonlinearity and high dimension, support vector machine can achieve better results in landslide susceptibility assessment and mapping when combined with other technologies (Ballabio and Sterlacchini 2012; Huang and Zhao 2018; Kavzoglu et al. 2013; Peng et al. 2014). For example, for the landslide disaster in Jinsha River basin, the landslide susceptibility map generated by the mixed model composed of FT model and SVM model is the final output product of landslide prediction (Hu et al. 2020). This falls under the category of regional landslide prediction. Deep learning models such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), long short-term memory (LSTM), etc., can handle complex geological data such as remote sensing images, terrain data, etc., thereby achieving more accurate geological disaster prediction. CNN is widely used for processing image data. In geological disaster prediction, it can be employed to analyze and process remote sensing images, satellite images, and other terrain data (Jiang et al. 2023). RNN is suitable for processing time series data. Through RNN, dynamic prediction and monitoring of geological disasters can be achieved, capturing the patterns and trends of geological disaster occurrences (Ji et al. 2023). LSTM is a special type of RNN that can effectively handle long-term dependencies. Through LSTM, long-term trend prediction and short-term variation monitoring of geological disasters can be achieved (Jiang et al. 2024). In addition to these, there are other ensemble learning methods that can combine multiple base models or use genetic algorithms to optimize predictive models, improving the accuracy and robustness of geological disaster prediction (Daviran et al. 2023; Huang et al. 2023).

6.1.2 Landslide displacement prediction

Landslide displacement prediction mainly focuses on short-term behavior of landslide, which is an important basis of landslide warning system. In general, early warning systems are used to provide actionable warning information before landslide accidents occur, thereby mitigating potential heavy losses (Piciullo et al. 2018). Landslide displacement is a finite and noisy time series, and extreme learning machine is one of the common prediction methods, and the prediction effect is better if combining with kernel function improvement or optimizing particle swarm algorithm (Du et al. 2020; Lian et al. 2014; Zhou et al. 2018a, b). In addition, long short-term memory network has also been proved to be useful in landslide displacement prediction, which is a dynamic model that can remember historical information and apply it to the current output (Jiang et al. 2020; Wang et al. 2022; Xu and Niu 2018). However, the traditional machine learning models mainly determine the input and output variables empirically, and do not solve the non-stationary characteristics of displacement time series. It is difficult for these prediction models to obtain the appropriate input and output variables, determine the appropriate model parameters, and obtain satisfactory prediction performance. The wavelet analysis (WA) method can decompose the displacement time series into low frequency and high frequency components to solve the non-stationary characteristics of the displacement time series, and then use chaos theory to obtain the appropriate input and output variables of the prediction model to improve the prediction accuracy (Li et al. 2019; Lin et al. 2022). Due to the different influencing factors of various geological disasters, no single or specific model can be considered as the most suitable for all geological disaster prediction, and the prediction method should be selected according to the characteristics of the disaster itself. Therefore, the mixed model is often used in landslide displacement prediction (Ma et al. 2017; Zhang et al. 2021a, b). The landslide predictions described above focus on forecasting the likelihood of a specific location or a particular landslide occurrence. These predictions require extensive geological surveys, monitoring, and data collection to perform in-depth analysis of the specific geological conditions and historical records at that site. This analysis helps predict the probability and potential risk of landslides at that specific location.

The team of Professor Tang Huiming from China University of Geosciences (Wuhan) conducted prediction and forecast research on major landslides. Focusing on the core scientific issues of landslide evolution process and physical and mechanical mechanism, the team summarizes two key scientific issues of "spatial–temporal correlation law of landslide multi-field characteristic information" and "Major landslide prediction and forecast theory based on physical and mechanical process" and develops the numerical prediction model and real-time prediction platform of landslide based on the physical and mechanical model of the whole process of evolution. The research utilized decision tree models and a least squares support vector machine optimized by genetic algorithms to predict landslide displacement. Evaluation confirmed that the selected models performed well for the current study area (Ma et al. 2017; Wen et al. 2017; Zhang et al. 2021a, b). Based on the field of disaster monitoring and warning, the research team of Chang'an University has developed a low-cost and high-precision BeiDou/GNSS monitoring system based on cloud platform, which has successfully warned three loess landslides (Liao et al. 2022; Ling et al. 2020, 2021). Machine learning models such as neural networks, adaptive neuro-fuzzy inference systems, support vector machines, were employed to generate landslide susceptibility maps (Chen et al. 2021; Fan et al. 2017; Li and Lan 2020; Liu et al. 2021a, b). The model performances were evaluated by receiver operating characteristic (ROC) curves and areas under the curves (AUC). The team of Professor Xu Qiang's in Chengdu University of Technology constructed an early identification system of landslide hidden danger organically combined geology and technology, established a comprehensive landslide early warning system based on time–space deformation law, broke through the problem of landslide warning, developed a real-time geological disaster monitoring and warning system, brought the identification, monitoring and warning of landslide hidden danger into the practical and operational stage and led the world (Fan et al. 2020; Li et al. 2018, 2020; Sun et al. 2021).

The study employed Artificial Bee Colony optimization of support vector regression for forecasting periodic displacement components, resulting in a 54% enhancement in accuracy (Zhang et al. 2021a, b). Chengdu University of Technology effectively identified over 3000 major landslide disasters, successfully forecasted over 40 landslide incidents, safeguarding the lives and property of tens of thousands of people. This achievement has resulted in significant social impact and economic benefits. The National Aeronautics and Space Administration (NASA) utilizes satellite precipitation data for mountain landslide hazard assessment and risk prediction. They have developed a global landslide disaster assessment software capable of identifying the potential for rainfall-induced mountain landslides in real time (Mirus et al. 2020; Psomiadis et al. 2020a, b; Psomiadis et al. 2020a, b).

6.2 AI prediction for collapse

Due to the rapid occurrence of collapse, lack of easily recognized premonition, sudden collapse and complex mechanism, there are few researches on the prediction of collapse by using machine learning methods.

The team of Wang Xueliang from the Institute of Geology and Geophysics, Chinese Academy of Sciences conducted numerical simulation and experiments on the surface process of landslide flow, studied the dynamic risk assessment and prevention methods of major engineering geological disasters, constructed the spatial–temporal distribution and evolution and the disaster-causing mechanism of rock mass structure in tectonically active areas, carried out relevant research on the identification and prediction methods of collapse (rolling stone) disasters in power transmission and transformation projects in mountainous areas with the help of aerial remote sensing, UAV aerial survey, rock mass structure analysis tools, numerical simulation software of movement of collapse (rolling stone) and other new technology methods. For example, the UAV aerial photography is used to identify the collapse body, and then the three-dimensional collapse simulation method is used to quickly identify the arrival of the collapse in the range of the tower, determine the impact characteristics of the collapse that can reach the range of the tower by comparing various methods, calculate the disaster risk of the collapse of the power transmission and transformation project in the mountain area to achieve the purpose of prediction.

6.3 AI prediction for debris flow

The flow velocity of debris flows is extremely fast, reaching speeds of tens to hundreds of kilometers per hour, with tremendous destructive power capable of destroying houses, washing away roads, and damaging farmland, among other things. Debris flows can also result in casualties and severe ecological damage. Due to their suddenness and danger, it is crucial to have early warning, monitoring, and prevention measures in place for potential debris flow disasters. Rainfall is one of the primary triggering factors for mudslides (Intrieri et al. 2017). For rainfall-induced mudslides, a threshold condition based on rainfall parameters can be constructed through statistical analysis of mudslide events and corresponding rainfall occurrences (Iadanza et al. 2016). When the precipitation exceeds this threshold, a mudslide may occur. Currently, rainfall parameter-based thresholds mainly encompass four categories: threshold based on rainfall intensity-duration (De Luca and Versace 2017); threshold relies on antecedent rainfall conditions (Yang et al. 2020); threshold using total rainfall accumulated during the event and its duration (Kanjanakul et al. 2016); and the rest presents various different parameters used to define rainfall thresholds (Vallet et al. 2016; Zhao et al. 2023).

Using rainfall thresholds is a method for predicting mudslides. Chen et al. proposed a technical process to analyze the interaction of rainfall parameters by using three-dimensional regression analysis. The regression results were converted into multiple Im-D rainfall warning regression charts containing different Re amounts to issue debris flow warning (Chen 2020). Zhao et al. used a rain gauge to record rainfall and established a rainfall prediction model for debris flow. The advantage is that it can automatically evaluate the probability that whether any given rainfall event will trigger debris flow (Zhao et al. 2022a, b). According to rainfall predictions, the hydrological quantitative rainfall forecast generated using the random forest method is applicable for analyzing mudslide models in the region of South Korea (Oh et al. 2021). Aside from using rainfall predictions to forecast mudslides, predictions can be made using other factors influencing mudslides. For instance, artificial neural network models can predict mudslide volume (Lee et al. 2021), Bayesian networks and logistic regression models can forecast rainfall probability (Kern et al. 2017). As sediment concentration is a crucial factor in evaluating peak mudslide flow, a novel hybrid model based on Bayesian networks, support vector regression-particle swarm optimization, and blended with hydrological simulation systems can predict mudslide peak flow (Banihabib et al. 2020). Additionally, employing machine learning models for susceptibility assessment using influencing factors of mudslides is another approach for mudslide prediction (Di et al. 2019; Qiu et al. 2022).

The team of Professor Meng Xingmin of Lanzhou University explained the superiority of "dredging" and "ecological management" in the prevention measures of debris flow disasters. Based on a large number of recorded debris flow events, Dr. Zhao Yan's team establish an intelligent rainfall prediction model for debris flow disasters by testing 17 machine learning models. The model can automatically analyze the rainfall records monitored by rain gauges in real time and predict whether the debris flow events will be triggered in advance and it greatly improved the prediction accuracy and warning time of debris flow (Qing et al. 2020; Zhao et al. 2021; Zhao et al. 2022a, b). The research result lays a foundation for establishing a reliable early warning system for debris flow and it has important scientific value and application prospects of disaster prevention and reduction. The Swiss Federal Institute of Technology in Zurich utilizes machine learning algorithms to distinguish ground vibrations caused by debris flows from other seismic signals, providing early warning for debris flows in Illgraben, Valais Canton (Belli et al. 2022; Chmiel et al. 2021; Hirschberg et al. 2021).

7 Opportunities and challenges

7.1 Scientifically construct AI paradigm of geological disaster prediction driven by physical knowledge in the geological field

The occurrence of geological disasters is closely related to the evolution of geological bodies. For example, the evolution and development of landslide is not a simple mathematical or mechanical process, but a result of the evolution of slope geological bodies with the coupling of internal and external dynamic factors. Evolution and physical and mechanical processes are the basic properties of landslides. At present, the domestic and foreign prediction models of landslides are mainly established based on the time-displacement characteristics of landslides (Guo et al. 2020; Huang et al. 2017; Liu et al. 2020; Wang et al. 2019). The evolution process of geological disasters is characterized by stages, nonlinearity and model diversity. The premise of geological disasters prediction and forecast is to effectively elucidate the physical and mechanical mechanism of geological disasters with the action of complex factors. However, the physical and mechanical mechanism of geological disasters is complex, and the traditional AI method only learns statistical rules from massive data and does not directly involve physical modeling. At present, the relevant physical and mechanical knowledge system in the field of geological disasters has been gradually established, especially for the typical geological disaster events such as landslide, collapse and debris flow, which basically clarified the occurrence mechanism of geological disasters. However, at present, the knowledge in these fields is relatively scattered and isolated. Not only effective semantic correlation and self-learning mechanism has not been established, but also effective coupling of physical prior knowledge in AI prediction paradigm has not been realized. It results in limited prediction accuracy, so it is urgent to couple multi-source monitoring data of geological disasters with disaster evolution and physical and mechanical processes and provide a theoretical framework for constructing an AI paradigm for predicting small probability geological disasters.

Recommendations for future research priorities include: (1) Reveal the evolution mechanism of geological disasters under the influence of rainfall, irrigation, reservoir water level fluctuation and other main factors. Build accurate and systematic geological disaster prediction criteria; (2) Study the expression mechanism of geological disaster physical models in artificial intelligence learning models based on logical constraints, prediction criteria and differential equations in typical geological disaster physical models; (3) Study the coupling model of the cause of geological disasters and machine learning and establish three models coupled with machine learning from the three aspects of geological disaster causality, space–time process and mechanism model; (4) Study the geological disaster prediction AI paradigm, which is based on the coupling framework of geological disaster evolution mechanism and artificial intelligence algorithm, and which take geological disaster domain knowledge map, common prediction rules, prediction sample set and artificial intelligence model as the main body. Clarify the interaction relationship and cooperation mode.

7.2 Establish the sample data set of the entire life cycle of geological disasters

The key of geological disaster prediction and forecast is to obtain multi-source big data based on three-dimensional synthesis, identify the main control factors of landslide evolution process combined with geological process, reveal internal and external dynamic factors, geological disaster evolution state and key characteristic attributes, and obtain the evolution rule of landslide behavior. The monitoring methods for geological disasters include InSAR, high-precision GPS monitoring, multispectral imaging technology, groundwater level monitoring networks, acoustic wave monitoring technology, and intelligent monitoring sensors. Based on these monitoring technologies, artificial intelligence methods are utilized to identify geological disasters and enhance the accuracy of the sample dataset (Auflič et al. 2023; Ge et al. 2023; Ma and Mei 2021b; Zhu et al. 2023). Furthermore, remote sensing technology, especially satellite technology, can provide assistance due to its wide spatial and temporal coverage (Antoine et al. 2020; Sousa et al. 2021). Particularly, the penetrative nature of remote sensing technology allows us to delve into the deeper layers of the Earth, understanding the distribution, structure, characteristics, and dynamic behavior of subsurface materials (Wang et al. 2023a, b, c). Different types of sensors enable us to monitor and simulate geological disasters in areas of different scales, providing high spatial, temporal, and spectral resolution, as well as stereoscopic mapping and all-weather imaging capabilities (Tomás and Li 2017). Remote sensing technology can also provide high-resolution data for environmental variables required for geological disaster prediction, such as soil moisture, precipitation, etc., thereby enhancing prediction accuracy (Brocca et al. 2023).

Geological disaster monitoring data has the characteristics of multi-source heterogeneous, cross-scale and multi-modal. The existing multi-source data fusion methods are unable to effectively characterize the process of geological disaster evolution. For example, at present, there are more than 330,000 hidden geological disaster points in the mainland of China, and abundant big data of "Sky-Earth-Human Network" three-dimensional monitoring in some key areas. However, these multi-modal data do not contain the complete characteristics of the entire life process of disaster, which results in relatively sparse high-quality learning samples. To effectively solve the problem that AI prediction models require a large number of sample data, it is necessary to break through the representation model and augment technology of geological disaster sample data, establish the sample data set of the entire life cycle of geological disasters and provide the support of high-quality input for the prediction system of geological disasters with low probability.

Recommendations for future research priorities include: (1) To study the feature description model of sample data set of geological disasters with small probability orienting the entire physical and mechanical life cycle process of geological disasters, the process includes latent, occurrence, development and completion; (2) Conduct automatic intensive labeling and sample generation of multi-source and multi-field observation data of geological disasters based on cataloging data of geological disasters, weakly supervised learning and data enhancement technology; (3) Establish a unified expression and fusion method of multisource monitoring data by combining with the geological process of disasters and comprehensively applying big data, artificial intelligence, data mining and modern information technology to improve the accuracy of geological disaster prediction and forecast.

7.3 Explore common rules of AI prediction of geological disaster events with small probability

Due to the complexity of geological bodies and their evolution process, the applicability of existing prediction models is poor and the success rate of the prediction is low. The different machine learning prediction models are not unsuitable for all geological disasters and also unsuitable for the same kind of geological disasters in different regions. Geological disaster has the characteristic of spatial random distribution, so it is difficult to be accurately determined (Gong et al. 2014; Li et al. 2016). Temporal and spatial uncertainties also exist in the occurrence frequency and intensity of external dynamic factors of geological disasters, such as rainfall, irrigation and changes of reservoir water level. The traditional AI prediction models mainly study and predict for specific spatial–temporal regions. However, geological disasters are not only widely distributed but also diverse in types. Therefore, it is urgent to build an AI paradigm of geological disaster prediction with the ability of spatial–temporal generalization, and it can be further extended into other disaster fields in earth science. The common prediction rules of medium-long term, short term and imminent term of the geological disasters with small probability are explored through domain knowledge map and large sample data set, which can provide the support of rule input for realizing high-precision prediction of geological disasters with small probability.

Recommendations for future research priorities include: (1) Analyze the response model of long time series in the entire process of "stability-start-destruction-recovery" of geological disasters with small probability and under the environmental disturbance. Study the common rules of medium-long term prediction based on evolution trend; (2) Study the continuous change curve, sudden change mode and cumulative damage limit of key parameters of various disaster-causing factors. Study the common rule of short-term prediction based on the extreme value of the dynamic evolution process of geological disasters; (3) Study and restore the short-term disaster imminent process from start-up to abrupt change of different geological disasters with high temporal resolution. Form the common rules of imminent disaster prediction based on the spatiotemporal symptoms of geological disasters with small probability.

8 Conclusions

At present, geological disasters occur frequently all over the world, which seriously threaten the people's safety and social development. However, due to the complexity of geological disaster prediction, it is difficult for the traditional geological disaster prediction model to achieve accurate prediction. With the development of technologies such as the Internet of Things, big data mining and artificial intelligence, machine learning methods have been introduced into geological disaster prediction. Therefore, this paper provides a review and future prospects of AI prediction of geological disasters such as landslide, debris flow and collapse. The conclusions are as follows:

1. Due to the frequent occurrence of geological disasters, there is a great demand of geological disaster prediction. However, the geological disaster prediction methods based on traditional mathematical statistics or physical models have the problems of low efficiency and low accuracy. Therefore, it is necessary to carry out AI prediction of geological disasters.
2. The overall ideas of AI prediction of geological disasters include the selection of prediction features, data collection, and the construction of AI prediction models. The prediction features of geological disaster AI prediction mainly include topography, hydrometeorology, geological conditions, environmental factors, and human activities. The geological disaster dataset is also included in these five categories of prediction features, including slope, aspect, curvature, terrain undulation, precipitation, reservoir water level, river network, fault, lithology, soil parameters, land use and so on. Currently used machine learning models for AI prediction of geological disasters include logistic regression, decision trees, support vector machines, random forests, K-nearest neighbors, Bayesian networks, etc. However, the accuracy of these models has reached a bottleneck, AI machine learning models that integrate the physical mechanism of geological disasters can achieve higher accuracy, lower computational consumption, and less data dependence in predictions.
3. Domestic and foreign teams have conducted numerous research in the field of AI prediction of geological disasters such as landslides, debris flows, and collapses. They have

achieved a series of research results and successfully applied them to multiple geological disasters, resulting in significant social and economic benefits.

4. Currently, the main challenges of AI prediction of geological disasters are how to scientifically construct an AI paradigm of geological disaster prediction driven by physical knowledge in the geological field, how to establish a sample dataset for the entire life cycle of geological disasters and how to explore the common rules of AI prediction for geological disaster events with low probability. These are also the key focuses for future research in AI prediction of geological disasters.

Funding This work was supported by grants from the National Natural Science Foundation of China (Grant U2340230), Open Fund of Hubei Luojia Laboratory (2201000059), Special Fund of Hubei Luojia Laboratory (220100034), Pre-research Project of SongShan Laboratory (YYYY062022001), Open Fund of State Key Laboratory of Remote Sensing Science (OFSLRSS202114), and KartoBit Research Network (KRN2301GK).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This work did not involve any live subjects (human or animal).

References

Abu El-Magd SA, Ali SA, Pham QB (2021) Spatial modeling and susceptibility zonation of landslides using random forest, naïve bayes and K-nearest neighbor in a complicated terrain. Earth Sci Inform 14(3):1227–1243

Adnan MSG, Rahman MS, Ahmed N, Ahmed B, Rabbi MF, Rahman RM (2020) Improving spatial agreement in machine learning-based landslide susceptibility mapping. Remote Sens 12(20):3347

Akgun A (2012) A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: a case study at İzmir, Turkey. Landslides 9(1):93–106

Antoine R, Lopez T, Tanguy M, Lissak C, Gailler L, Labazuy P, Fauchard C (2020) Geoscientists in the sky: unmanned aerial vehicles responding to geohazards. Surv Geophys 41(6):1285–1321

Auflič MJ, Herrera G, Mateos RM, Poyiadji E, Quental L, Severine B, Peternel T, Podolszki L, Calcaterra S, Kociu A, Warmuz B (2023) Landslide monitoring techniques in the Geological Surveys of Europe. Landslides 20(5):951–965

Ayawah PE, Sebbeh-Newton S, Azure JW, Kaba AG, Anani A, Bansah S, Zabidi H (2022) A review and case study of Artificial intelligence and Machine learning methods used for ground condition prediction ahead of tunnel boring Machines. Tunn Undergr Space Technol 125:104497

Baghbani A, Choudhury T, Costa S, Reiner J (2022) Application of artificial intelligence in geotechnical engineering: a state-of-the-art review. Earth Sci Rev 228:103991

Ballabio C, Sterlacchini S (2012) Support Vector machines for landslide susceptibility mapping: the Staffora River Basin case study, Italy. Math Geosci 44(1):47–70. https://doi.org/10.1007/s11004-011-9379-9

Banihabib ME, Jurik L, Kazemi MS, Soltani J, Tanhapour M (2020) A hybrid intelligence model for the prediction of the peak flow of debris floods. Water 12(8):2246

Belli G, Walter F, McArdell B, Gheri D, Marchetti E (2022) Infrasonic and seismic analysis of debris-flow events at Illgraben (Switzerland): relating signal features to flow parameters and to the seismo-acoustic source mechanism. J Geophys Res Earth Surf 127(6):e2021JF006576

Bergstra J, Yamins D, Cox D (2013) Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures. Paper presented at the International conference on machine learning

Bhagya SB, Sumi AS, Balaji S, Danumah JH, Costache R, Rajaneesh A, Gokul A, Chandrasenan CP, Quevedo RP, Johny A, Sajinkumar KS (2023) Landslide susceptibility assessment of a part of the

Western Ghats (India) employing the AHP and F-AHP models and comparison with existing susceptibility maps. Land 12(2):468
Brocca L, Zhao W, Lu H (2023) High-resolution observations from space to address new applications in hydrology. Innovation (camb) 4(3):100437. https://doi.org/10.1016/j.xinn.2023.100437
Bui DT, Tsangaratos P, Nguyen V-T, Van Liem N, Trinh PT (2020a) Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment. CATENA 188:104426
Bui X-N, Nguyen H, Choi Y, Nguyen-Thoi T, Zhou J, Dou J (2020b) Prediction of slope failure in open-pit mines using a novel hybrid artificial intelligence model based on decision tree and evolution algorithm. Sci Rep 10(1):9939
Cao J, Zhang Z, Du J, Zhang L, Song Y, Sun G (2020) Multi-geohazards susceptibility mapping based on machine learning—a case study in Jiuzhaigou, China. Nat Hazards 102:851–871
Casagli N, Intrieri E, Tofani V, Gigli G, Raspini F (2023) Landslide detection, monitoring and prediction with remote-sensing techniques. Nat Rev Earth Environ 4(1):51–64
Catani F, Segoni S, Falorni G (2010) An empirical geomorphology-based approach to the spatial prediction of soil thickness at catchment scale. Water Res Res 46(5)
Chang K-T, 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(1):12296
Chang Z, Catani F, Huang F, Liu G, Meena SR, Huang J, Zhou C (2023) Landslide susceptibility prediction using slope unit-based machine learning models considering the heterogeneity of conditioning factors. J Rock Mech Geotech Eng 15(5):1127–1143
Chen C-Y (2020) Event-based rainfall warning regression model for landslide and debris flow issuing. Environ Earth Sci 79(6):127. https://doi.org/10.1007/s12665-020-8877-9
Chen W, Zhang S (2021) GIS-based comparative study of Bayes network, Hoeffding tree and logistic model tree for landslide susceptibility modeling. CATENA 203:105344
Chen W, Pourghasemi HR, Kornejady A, Zhang N (2017) Landslide spatial modeling: introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. Geoderma 305:314–327. https://doi.org/10.1016/j.geoderma.2017.06.020
Chen H, Qin S, Xue L, Yang B, Zhang K (2018a) A physical model predicting instability of rock slopes with locked segments along a potential slip surface. Eng Geol 242:34–43
Chen W, Peng J, Hong H, Shahabi H, Pradhan B, Liu J, Zhu AX, Pei X, Duan Z (2018b) Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China. Sci Total Environ 626:1121–1135
Chen W, Fan L, Li C, Pham BT (2019) Spatial prediction of landslides using hybrid integration of artificial intelligence algorithms with frequency ratio and index of entropy in Nanzheng County, China. Appl Sci 10(1):29
Chen W, Sun Z, Zhao X, Lei X, Shirzadi A, Shahabi H (2020a) Performance evaluation and comparison of bivariate statistical-based artificial intelligence algorithms for spatial prediction of landslides. ISPRS Int J Geo Inf 9(12):696
Chen Z, He C, Yang W, Guo W, Li Z, Xu G (2020b) Impacts of geological conditions on instability causes and mechanical behavior of large-scale tunnels: a case study from the Sichuan–Tibet highway, China. Bull Eng Geol Env 79:3667–3688
Chen W, Chen X, Peng J, Panahi M, Lee S (2021) Landslide susceptibility modeling based on ANFIS with teaching-learning-based optimization and Satin bowerbird optimizer. Geosci Front 12(1):93–107. https://doi.org/10.1016/j.gsf.2020.07.012
Chen J, Wang Z, Chen W, Wan C, Liu Y, Huang J (2023) The influence of the selection of non-geological disasters sample spatial range on the evaluation of environmental geological disasters susceptibility: a case study of Liulin County. Environ Sci Pollut Res 30(15):44756–44772
Chmiel M, Walter F, Wenner M, Zhang Z, McArdell BW, Hibert C (2021) Machine learning improves debris flow warning. Geophys Res Lett 48(3):e2020GL090874
Conforti M, Pascale S, Robustelli G, Sdao F (2014) Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo River catchment (northern Calabria, Italy). CATENA 113:236–250
Cui P, Ge Y, Li S, Li Z, Xu X, Zhou GG, Chen H, Wang H, Lei Y, Zhou L, Yi S (2022) Scientific challenges in disaster risk reduction for the Sichuan–Tibet Railway. Eng Geol 309:106837
Dai F, Lee CF, Ngai YY (2002) Landslide risk assessment and management: an overview. Eng Geol 64(1):65–87

Daviran M, Shamekhi M, Ghezelbash R, Maghsoudi A (2023) Landslide susceptibility prediction using artificial neural networks, SVMs and random forest: hyperparameters tuning by genetic optimization algorithm. Int J Environ Sci Technol 20(1):259–276

De Luca DL, Versace P (2017) A comprehensive framework for empirical modeling of landslides induced by rainfall: the Generalized FLaIR Model (GFM). Landslides 14:1009–1030

de Oliveira GG, Ruiz LFC, Guasselli LA, Haetinger C (2019) Random forest and artificial neural networks in landslide susceptibility modeling: a case study of the Fão River Basin, Southern Brazil. Nat Hazards 99(2):1049–1073

Demir G (2019) GIS-based landslide susceptibility mapping for a part of the North Anatolian Fault Zone between Reşadiye and Koyulhisar (Turkey). CATENA 183:104211

Depina I, Jain S, Mar Valsson S, Gotovac H (2022) Application of physics-informed neural networks to inverse problems in unsaturated groundwater flow. Georisk: Assess Manag Risk Eng Syst Geohazards 16(1):21–36

Di Napoli M, Carotenuto F, Cevasco A, Confuorto P, Di Martire D, Firpo M, Pepe G, Raso E, Calcaterra D (2020) Machine learning ensemble modelling as a tool to improve landslide susceptibility mapping reliability. Landslides 17(8):1897–1914

Di B, Zhang H, Liu Y, Li J, Chen N, Stamatopoulos CA, Luo Y, Zhan Y (2019) Assessing susceptibility of debris flow in southwest China using gradient boosting machine. Sci Rep 9(1):1–12

Dikshit A, Pradhan B, Alamri AM (2021) Pathways and challenges of the application of artificial intelligence to geohazards modelling. Gondwana Res 100:290–301

Du H, Song D, Chen Z, Shu H, Guo Z (2020) Prediction model oriented for landslide displacement with step-like curve by applying ensemble empirical mode decomposition and the PSO-ELM method. J Clean Prod 270:122248

Fan W, Wei X-S, Cao Y-B, Zheng B (2017) Landslide susceptibility assessment using the certainty factor and analytic hierarchy process. J Mt Sci 14(5):906–925. https://doi.org/10.1007/s11629-016-4068-2

Fan X, Scaringi G, Korup O, West AJ, van Westen CJ, Tanyas H, Hovius N, Hales TC, Jibson RW, Allstadt KE, Zhang L (2019) Earthquake-induced chains of geologic hazards: patterns, mechanisms, and impacts. Rev Geophys 57(2):421–503

Fan X, Yang F, Siva Subramanian S, Xu Q, Feng Z, Mavrouli O, Peng M, Ouyang C, Jansen JD, Huang R (2020) Prediction of a multi-hazard chain by an integrated numerical simulation approach: the Baige landslide, Jinsha River, China. Landslides 17:147–164

Fathi MH, Khohdel K, Kandi AS, Ashrafifeini Z, Khaliji MA (2015) The combination of spectral and spatial data in zoning of landslide susceptibility (Case study: Sangorchay reservoir). J Biodivers Environ Sci (JBES) 6(2):515–527

Felicísimo ÁM, Cuartero A, Remondo J, Quirós E (2013) Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. Landslides 10:175–189

Froude MJ, Petley DN (2018) Global fatal landslide occurrence from 2004 to 2016. Nat Hazard 18(8):2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Ge Q, Sun Q, Zhang N, Hu J (2023) Evaluation of geological hazard susceptibility of Baiyin City based on MT-InSAR deformation measurements. Geomatics and Information Science of Wuhan University

Goetz J, Brenning A, Petschko H, Leopold P (2015) Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Comput Geosci 81:1–11

Gomez H, Kavzoglu T (2005) Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela. Eng Geol 78(1–2):11–27

Gong W, Luo Z, Juang CH, Huang H, Zhang J, Wang L (2014) Optimization of site exploration program for improved prediction of tunneling-induced ground settlement in clays. Comput Geotech 56:69–79

Gorsevski PV, Jankowski P, Gessler PE (2006) An heuristic approach for mapping landslide hazard by integrating fuzzy logic with analytic hierarchy process. Control Cybern 35(1):121–146

Grahn T, Jaldell H (2016) Assessment of data availability for the development of landslide fatality curves. Landslides 14(3):1113–1126. https://doi.org/10.1007/s10346-016-0775-6

Groisman PY, Legates DR (1994) The accuracy of United States precipitation data. Bull Am Meteor Soc 75(2):215–228

Guo C, Montgomery DR, Zhang Y, Wang K, Yang Z (2015) Quantitative assessment of landslide susceptibility along the Xianshuihe fault zone, Tibetan Plateau, China. Geomorphology 248:93–110

Guo Z, Chen L, Gui L, Du J, Yin K, Do HM (2020) Landslide displacement prediction based on variational mode decomposition and WA-GWO-BP model. Landslides 17:567–583

Guo C, Xu Q, Dong X, Liu X, She J (2021) Geohazard recognition by airborne LiDAR technology in complex mountain areas. 武汉大学学报● 信息科学版, 46(10):1538–1547

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(1–2):42–66. https://doi.org/10.1016/j.earscirev.2012.02.001

Han H, Shi B, Zhang L (2021) Prediction of landslide sharp increase displacement by SVM with considering hysteresis of groundwater change. Eng Geol 280:105876

Hansen A, Franks C, Kirk P, Brimicombe A, Tung F (1995) Application of GIS to hazard assessment, with particular reference to landslides in Hong Kong. Geogr Inform Syst Assess Nat Hazards 273–298

Harsa H, Anistia MH, Mulsandi A, Suprihadi B, Kurniawan R, Habibie MN, Hutapea TD, Swarinoto YS, Makmur EES, Fitria W, Sudewi RSS (2023) Machine learning and artificial intelligence models development in rainfall-induced landslide prediction. IAES Int J Artif Intell 12(1):262

He J, Qiu H, Qu F, Hu S, Yang D, Shen Y, Cao M (2021a) Prediction of spatiotemporal stability and rainfall threshold of shallow landslides using the TRIGRS and Scoops3D models. CATENA 197:104999

He L, Coggan J, Francioni M, Eyre M (2021b) Maximizing impacts of remote sensing surveys in slope stability—a novel method to incorporate discontinuities into machine learning landslide prediction. ISPRS Int J Geo Inf 10(4):232

Hewitt K, Clague JJ, Orwin JF (2008) Legacies of catastrophic rock slope failures in mountain landscapes. Earth Sci Rev 87(1–2):1–38

Hirschberg J, Badoux A, McArdell BW, Leonarduzzi E, Molnar P (2021) Evaluating methods for debris-flow prediction based on rainfall in an Alpine catchment. Nat Hazard 21(9):2773–2789

Ho J-Y, Lee KT, Chang T-C, Wang Z-Y, Liao Y-H (2012) Influences of spatial distribution of soil thickness on shallow landslide prediction. Eng Geol 124:38–46

Hong H, Liu J, Zhu A-X (2019a) Landslide susceptibility evaluating using artificial intelligence method in the Youfang district (China). Environ Earth Sci 78:1–20

Hong H, Miao Y, Liu J, Zhu A-X (2019b) Exploring the effects of the design and quantity of absence data on the performance of random forest-based landslide susceptibility mapping. CATENA 176:45–64

Hu J, Li Z, Ding X, Zhu J, Zhang L, Sun Q (2014) Resolving three-dimensional surface displacements from InSAR measurements: a review. Earth Sci Rev 133:1–17

Hu Q, Zhou Y, Wang S, Wang F (2020) Machine learning and fractal theory models for landslide susceptibility mapping: case study from the Jinsha River Basin. Geomorphology 351:106975. https://doi.org/10.1016/j.geomorph.2019.106975

Hu X, Wu S, Zhang G, Zheng W, Liu C, He C, Zhang H (2021) Landslide displacement prediction using kinematics-based random forests method: a case study in Jinping Reservoir Area, China. Eng Geol 283:105975

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 F, Huang J, Jiang S, Zhou C (2017) Landslide displacement prediction based on multivariate chaotic model and extreme learning machine. Eng Geol 218:173–186

Huang F, Cao Z, Guo J, Jiang S-H, Li S, Guo Z (2020) Comparisons of heuristic, general statistical and machine learning models for landslide susceptibility prediction and mapping. CATENA 191:104580

Huang W, Ding M, Li Z, Yu J, Ge D, Liu Q, Yang J (2023) Landslide susceptibility mapping and dynamic response along the Sichuan–Tibet transportation corridor using deep learning algorithms. CATENA 222:106866

Huffman GJ, Adler RF, Arkin P, Chang A, Ferraro R, Gruber A, Schneider U (1997) The global precipitation climatology project (GPCP) combined precipitation dataset. Bull Am Meteorol Soc 78(1):5–20

Hung W-C, Hwang C, Liou J-C, Lin Y-S, Yang H-L (2012) Modeling aquifer-system compaction and predicting land subsidence in central Taiwan. Eng Geol 147:78–90

Iadanza C, Trigila A, Napolitano F (2016) Identification and characterization of rainfall events responsible for triggering of debris flows and shallow landslides. J Hydrol 541:230–245

Intrieri E, Bardi F, Fanti R, Gigli G, Fidolini F, Casagli N, Capparelli G (2017) Big data managing in a landslide early warning system: experience from a ground-based interferometric radar application. Nat Hazards Earth Syst Sci 17(10):1713–1723

Javed S (2019) Landslide hazard mapping of Bagh district in Azad Kashmir. Int J Econ Environ Geol 8(3):47–50

Ji J, Zhou Y, Cheng Q, Jiang S, Liu S (2023) Landslide susceptibility mapping based on deep learning algorithms using information value analysis optimization. Land 12(6):1125

Jiang W, Rao P, Cao R, Tang Z, Chen K (2017) Comparative evaluation of geological disaster susceptibility using multi-regression methods and spatial accuracy validation. J Geog Sci 27:439–462

Jiang H, Li Y, Zhou C, Hong H, Glade T, Yin K (2020) Landslide displacement prediction combining LSTM and SVR algorithms: a case study of Shengjibao Landslide from the Three Gorges Reservoir Area. Appl Sci 10(21):7830

Jiang Z, Wang M, Liu K (2023) Comparisons of convolutional neural network and other machine learning methods in landslide susceptibility assessment: a case study in Pingwu. Remote Sens 15(3):798

Jiang Y, Wang W, Zou L, Cao Y (2024) Regional landslide susceptibility assessment based on improved semi-supervised clustering and deep learning. Acta Geotech 19(1):509–529

Jones DT, Thornton JM (2022) The impact of AlphaFold2 one year on. Nat Methods 19(1):15–20

Kanjanakul C, Chub-uppakarn T, Chalermyanont T (2016) Rainfall thresholds for landslide early warning system in Nakhon Si Thammarat. Arab J Geosci 9:1–11

Kargel JS, Leonard GJ, Shugar DH, Haritashya UK, Bevington A, Fielding EJ, Fujita K, Geertsema M, Miles ES, Steiner J, Anderson E, Bajracharya S, Bawden GW, Breashears DF, Byers A, Collins B, Dhital MR, Donnellan A, Evans TL, Geai ML, Glasscoe MT, Green D, Gurung DR, Heijenk R, Hilborn A, Hudnut K, Huyck C, Immerzeel WW, Jiang Liming R, Jibson A, Kääb NR, Khanal D, Kirschbaum PDA, Kraaijenbrink DL, Shiyin L, Lv Mingyang D, McKinney NK, Nahirnick NZ, Ojha S, Olsenholler J, Painter TH, Pleasants M, Pratima KC, Yuan QI, Raup BH, Regmi D, Rounce DR, Sakai A, Shangguan Donghui JM, Shea AB, Shrestha A, Shukla D, van der Stumm M, Kooij KV, Wang Xin B, Weihs DW, Lizong Wu, Yao Xiaojun MR, Yoder NY (2016) Geomorphic and geologic controls of geohazards induced by Nepal's 2015 Gorkha earthquake. Science 351(6269):aac8353

Kavzoglu T, Sahin EK, Colkesen I (2013) Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. Landslides 11(3):425–439. https://doi.org/10.1007/s10346-013-0391-7

Kavzoglu T, Colkesen I, Sahin EK (2019) Machine learning techniques in landslide susceptibility mapping: a survey and a case study. Landslides Theory Pr Modell 50:283–301

Kawabata D, Bandibas J (2009) Landslide susceptibility mapping using geological data, a DEM from ASTER images and an Artificial Neural Network (ANN). Geomorphology 113(1–2):97–109

Kern AN, Addison P, Oommen T, Salazar SE, Coffman RA (2017) Machine learning based predictive modeling of debris flow probability following wildfire in the intermountain Western United States. Math Geosci 49:717–735

Kirschbaum D, Stanley T, Zhou Y (2015) Spatial and temporal analysis of a global landslide catalog. Geomorphology 249:4–15. https://doi.org/10.1016/j.geomorph.2015.03.016

Kornejady A, Ownegh M, Bahremand A (2017) Landslide susceptibility assessment using maximum entropy model with two different data sampling methods. CATENA 152:144–162

Korup O, Stolle A (2014) Landslide prediction from machine learning. Geol Today 30(1):26–33

Lan H, Peng J, Zhu Y, Li L, Pan B, Huang Q, Li J, Zhang Q (2022) Research on geological and surfacial processes and major disaster effects in the Yellow River Basin. Sci China Earth Sci 65(2):234–256

Le Breton M, Bontemps N, Guillemot A, Baillet L, Larose É (2021) Landslide monitoring using seismic ambient noise correlation: challenges and applications. Earth Sci Rev 216:103518. https://doi.org/10.1016/j.earscirev.2021.103518

Lee KT, Ho J-Y (2009) Prediction of landslide occurrence based on slope-instability analysis and hydrological model simulation. J Hydrol 375(3–4):489–497

Lee M-J, Park I, Lee S (2015) Forecasting and validation of landslide susceptibility using an integration of frequency ratio and neuro-fuzzy models: a case study of Seorak mountain area in Korea. Environ Earth Sci 74(1):413–429. https://doi.org/10.1007/s12665-015-4048-9

Lee D-H, Cheon E, Lim H-H, Choi S-K, Kim Y-T, Lee S-R (2021) An artificial neural network model to predict debris-flow volumes caused by extreme rainfall in the central region of South Korea. Eng Geol 281:105979

Lei X, Chen W, Pham BT (2020) Performance evaluation of GIS-based artificial intelligence approaches for landslide susceptibility modeling and spatial patterns analysis. ISPRS Int J Geo Inf 9(7):443

Li L, Lan H (2020) Integration of spatial probability and size in slope-unit-based landslide susceptibility assessment: a case study. Int J Environ Res Public Health 17(21):8055. https://doi.org/10.3390/ijerph17218055

Li H, Zhao J (2018) Evaluation of the newly released worldwide AW3D30 DEM over typical landforms of China using two global DEMs and ICESat/GLAS data. IEEE J Sel Top Appl Earth Obs Remote Sens 11(11):4430–4440

Li C, Tang H, Hu X, Li D, Hu B (2009) Landslide prediction based on wavelet analysis and cusp catastrophe. J Earth Sci 20(6):971

Li C, Tang H, Ge Y, Hu X, Wang L (2014) Application of back-propagation neural network on bank destruction forecasting for accumulative landslides in the three Gorges Reservoir Region, China. Stoch Env Res Risk Assess 28:1465–1477

Li X, Zhang LM, Li J (2016) Using conditioned random field to characterize the variability of geologic profiles. J Geotech Geoenviron Eng 142(4):04015096

Li H, Xu Q, He Y, Deng J (2018) Prediction of landslide displacement with an ensemble-based extreme learning machine and copula models. Landslides 15:2047–2059

Li Y, Sun R, Yin K, Xu Y, Chai B, Xiao L (2019) Forecasting of landslide displacements using a chaos theory based wavelet analysis-Volterra filter model. Sci Rep 9(1):19853. https://doi.org/10.1038/s41598-019-56405-y

Li H, Xu Q, He Y, Fan X, Li S (2020) Modeling and predicting reservoir landslide displacement with deep belief network and EWMA control charts: a case study in Three Gorges Reservoir. Landslides 17(3):693–707

Lian C, Zeng Z, Yao W, Tang H (2014) Extreme learning machine for the displacement prediction of landslide under rainfall and reservoir level. Stoch Env Res Risk Assess 28:1957–1972

Liao K, Zhang W, Zhu H-H, Zhang Q, Shi B, Wang J-T, Xu W-T (2022) Forecasting reservoir-induced landslide deformation using genetic algorithm enhanced multivariate Taylor series Kalman filter. Bull Eng Geol Env 81(3):104

Lin L, Lin Q, Wang Y (2017) Landslide susceptibility mapping on a global scale using the method of logistic regression. Nat Hazard 17(8):1411–1424

Lin Z, Sun X, Ji Y (2022) Landslide displacement prediction based on time series analysis and Double-BiLSTM Model. Int J Environ Res Public Health 19(4):2077

Ling Q, Qu W, Zhang Q, Kong L, Zhang J, Zhu L (2020) Improved Kalman filter method considering multiple factors and its application in landslide prediction. Front Earth Sci 14(3):625–636. https://doi.org/10.1007/s11707-019-0796-2

Ling Q, Zhang Q, Zhang J, Kong L, Zhang W, Zhu L (2021) Prediction of landslide displacement using multi-kernel extreme learning machine and maximum information coefficient based on variational mode decomposition: a case study in Shaanxi, China. Nat Hazards 108(1):925–946. https://doi.org/10.1007/s11069-021-04713-w

Liu M, He Y, Wang J, Lee HP, Liang Y (2015) Hybrid intelligent algorithm and its application in geological hazard risk assessment. Neurocomputing 149:847–853

Liu Y, Xu C, Huang B, Ren X, Liu C, Hu B, Chen Z (2020) Landslide displacement prediction based on multi-source data fusion and sensitivity states. Eng Geol 271:105608

Liu Q, Jian W, Nie W (2021a) Rainstorm-induced landslides early warning system in mountainous cities based on groundwater level change fast prediction. Sustain Cities Soc 69:102817

Liu R, Peng J, Leng Y, Lee S, Panahi M, Chen W, Zhao X (2021b) Hybrids of support vector regression with grey wolf optimizer and firefly algorithm for spatial prediction of landslide susceptibility. Remote Sens 13(24):4966. https://doi.org/10.3390/rs13244966

Liu W, Zhang Y, Liang Y, Sun P, Li Y, Su X, Wang A, Meng X (2022) Landslide risk assessment using a combined approach based on InSAR and random forest. Remote Sens 14(9):2131

Lu Y, Mei G (2022) A deep learning approach for predicting two-dimensional soil consolidation using physics-informed neural networks (PINN). Mathematics 10(16):2949

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(1):15369

Ma Z, Mei G (2021a) Deep learning for geological hazards analysis: data, models, applications, and opportunities. Earth Sci Rev 223:103858. https://doi.org/10.1016/j.earscirev.2021.103858

Ma Z, Mei G (2021b) Deep learning for geological hazards analysis: data, models, applications, and opportunities. Earth Sci Rev 223:103858

Ma J, Tang H, Liu X, Hu X, Sun M, Song Y (2017) Establishment of a deformation forecasting model for a step-like landslide based on decision tree C5.0 and two-step cluster algorithms: a case study in the Three Gorges Reservoir area, China. Landslides 14(3):1275–1281. https://doi.org/10.1007/s10346-017-0804-0

Ma Z, Mei G, Piccialli F (2021) Machine learning for landslides prevention: a survey. Neural Comput Appl 33:10881–10907

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

Mirus BB, Jones ES, Baum RL, Godt JW, Slaughter S, Crawford MM, Schmitt RG, Lindsey KO, McCoy KM, Lancaster J, Stanley T, Kirschbaum DB, Burns WJ (2020) Landslides across the USA: occurrence, susceptibility, and data limitations. Landslides 17:2271–2285

Moayedi H, Mehrabi M, Mosallanezhad M, Rashid ASA, Pradhan B (2019) Modification of landslide susceptibility mapping using optimized PSO-ANN technique. Eng Comput 35:967–984

Mukherjee A, Sarkar S, Chakraborty M, Duttagupta S, Bhattacharya A, Saha D, Bhattacharya P, Mitra A, Gupta S (2021) Occurrence, predictors and hazards of elevated groundwater arsenic across India through field observations and regional-scale AI-based modeling. Sci Total Environ 759:143511

Murillo-García FG, Steger S, Alcántara-Ayala I (2019) Landslide susceptibility: a statistically-based assessment on a depositional pyroclastic ramp. J Mt Sci 16(3):561–580

NachtergaeleF, Velthuizen HV, Verelst L, Batjes N, Dijkshoorn K, Engelen VV, Fischer G, Jones A, Petri M, Prieler S, Montanarela L (2010) The harmonized world soil database. Paper presented at the Proceedings of the 19th World Congress of soil science, soil solutions for a changing world, Brisbane, Australia, 1–6 Aug 2010

Nava L, Carraro E, Reyes-Carmona C, Puliero S, Bhuyan K, Rosi A, Monserrat O, Floris M, Meena SR, Galve JP, Catani F (2023) Landslide displacement forecasting using deep learning and monitoring data across selected sites. Landslides 20(10):2111–2129

Ngo PTT, Panahi M, Khosravi K, Ghorbanzadeh O, Kariminejad N, Cerda A, Lee S (2021) Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. Geosci Front 12(2):505–519

Nichol A (2022) Dalle 2 pre-training mitigations.

Nicu IC (2017) Frequency ratio and GIS-based evaluation of landslide susceptibility applied to cultural heritage assessment. J Cult Herit 28:172–176

Nowicki Jessee M, Hamburger MW, Allstadt K, Wald DJ, Robeson SM, Tanyas H, Hearne M, Thompson EM (2018) A global empirical model for near-real-time assessment of seismically induced landslides. J Geophys Res Earth Surf 123(8):1835–1859

Oh C-H, Choo K-S, Go C-M, Choi J-R, Kim B-S (2021) Forecasting of debris flow using machine learning-based adjusted rainfall information and RAMMS model. Water 13(17):2360

Orland E, Roering JJ, Thomas MA, Mirus BB (2020) Deep learning as a tool to forecast hydrologic response for landslide-prone hillslopes. Geophys Res Lett 47(16):e2020GL088731

Osmanoğlu B, Sunar F, Wdowinski S, Cabral-Cano E (2016) Time series analysis of InSAR data: methods and trends. ISPRS J Photogramm Remote Sens 115:90–102

Panahi M, Gayen A, Pourghasemi HR, Rezaie F, Lee S (2020) Spatial prediction of landslide susceptibility using hybrid support vector regression (SVR) and the adaptive neuro-fuzzy inference system (ANFIS) with various metaheuristic algorithms. Sci Total Environ 741:139937

Peng L, Niu R, Huang B, Wu X, Zhao Y, Ye R (2014) Landslide susceptibility mapping based on rough set theory and support vector machines: a case of the Three Gorges area, China. Geomorphology 204:287–301

Pham BT, Prakash I, Singh SK, Shirzadi A, Shahabi H, Bui DT (2019) Landslide susceptibility modeling using Reduced Error Pruning Trees and different ensemble techniques: hybrid machine learning approaches. CATENA 175:203–218

Pham QB, Chandra Pal S, Chakrabortty R, Saha A, Janizadeh S, Ahmadi K, Khedher KM, Anh DT, Tiefenbacher JP, Bannari A (2022) Predicting landslide susceptibility based on decision tree machine learning models under climate and land use changes. Geocarto Int 37(25):7881–7907

Piciullo L, Calvello M, Cepeda JM (2018) Territorial early warning systems for rainfall-induced landslides. Earth Sci Rev 179:228–247. https://doi.org/10.1016/j.earscirev.2018.02.013

Pourghasemi HR, Moradi HR, Fatemi Aghda SM (2013) Landslide susceptibility mapping by binary logistic regression, analytical hierarchy process, and statistical index models and assessment of their performances. Nat Hazards 69(1):749–779. https://doi.org/10.1007/s11069-013-0728-5

Psomiadis E, Charizopoulos N, Efthimiou N, Soulis KX, Charalampopoulos I (2020a) Earth observation and GIS-based analysis for landslide susceptibility and risk assessment. ISPRS Int J Geo Inf 9(9):552

Psomiadis E, Papazachariou A, Soulis KX, Alexiou D-S, Charalampopoulos I (2020b) Landslide mapping and susceptibility assessment using geospatial analysis and earth observation data. Land 9(5):133

Qi T, Zhao Y, Meng X, Chen G, Dijkstra T (2021) Ai-based susceptibility analysis of shallow landslides induced by heavy rainfall in Tianshui, China. Remote Sens 13(9):1819

Qin S, Jiao JJ, Li Z (2006) Nonlinear evolutionary mechanisms of instability of plane-shear slope: catastrophe, bifurcation, chaos and physical prediction. Rock Mech Rock Eng 39:59–76

Qing F, Zhao Y, Meng X, Su X, Qi T, Yue D (2020) Application of machine learning to debris flow susceptibility mapping along the China–Pakistan Karakoram Highway. Remote Sens 12(18):2933

Qiu H, Cui P, Regmi AD, Hu S, Hao J (2019) Loess slide susceptibility assessment using frequency ratio model and artificial neural network. Q J Eng Geol Hydrogeol 52(1):38–45

Qiu C, Su L, Zou Q, Geng X (2022) A hybrid machine-learning model to map glacier-related debris flow susceptibility along Gyirong Zangbo watershed under the changing climate. Sci Total Environ 818:151752

Rahman MM, Szabó G (2021) Multi-objective urban land use optimization using spatial data: a systematic review. Sustain Cities Soc 74:103214

Rashid B, Iqbal J, Su L-J (2020) Landslide susceptibility analysis of Karakoram highway using analytical hierarchy process and scoops 3D. J Mt Sci 17(7):1596–1612

Ray RL, Jacobs JM, Cosh MH (2010) Landslide susceptibility mapping using downscaled AMSR-E soil moisture: a case study from Cleveland Corral, California. US Remote Sens Environ 114(11):2624–2636

Ray A, Kumar V, Kumar A, Rai R, Khandelwal M, Singh T (2020) Stability prediction of Himalayan residual soil slope using artificial neural network. Nat Hazards 103(3):3523–3540

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

Ruff M, Czurda K (2008) Landslide susceptibility analysis with a heuristic approach in the Eastern Alps (Vorarlberg, Austria). Geomorphology 94(3–4):314–324

Saad N, Bahari N, Talib N, Zaki NM, Rasam AA (2021) Landslide hazard analysis using Landsat-8 OLI and AHP technique in Tanjung Bungah, Penang. Paper presented at the IOP Conference Series: earth and environmental science

Saha A, Saha S (2022) Integrating the artificial intelligence and hybrid machine learning algorithms for improving the accuracy of spatial prediction of landslide hazards in Kurseong Himalayan Region. Artif Intell Geosci 3:14–27

Saito M (1965) Forecasting the time of occurrence of a slope failure. Paper presented at the Proceedings of 6th international conference soil mechanics and foundation eng

Segoni S, Pappafico G, Luti T, Catani F (2020) Landslide susceptibility assessment in complex geological settings: Sensitivity to geological information and insights on its parameterization. Landslides 17:2443–2453

Selamat SN, Majid NA, Taha MR, Osman A (2022) Landslide susceptibility model using artificial neural network (ANN) approach in Langat river basin, Selangor, Malaysia. Land 11(6):833

Sezer EA, Pradhan B, Gokceoglu C (2011) Manifestation of an adaptive neuro-fuzzy model on landslide susceptibility mapping: Klang valley, Malaysia. Expert Syst Appl 38(7):8208–8219

Sousa JJ, Liu G, Fan J, Perski Z, Steger S, Bai S, Wei L, Salvi S, Wang Q, Tu J, Tong L (2021) Geohazards monitoring and assessment using multi-source earth observation techniques. Remote Sens 13(21):4269

SS VC, Shaji E (2022) Landslide identification using machine learning techniques: review, motivation, and future prospects. Earth Sci Inform 1–28

Stanley T, Kirschbaum DB (2017) A heuristic approach to global landslide susceptibility mapping. Nat Hazards 87:145–164

Sun X, Chen J, Han X, Bao Y, Zhou X, Peng W (2020) Landslide susceptibility mapping along the upper Jinsha River, south-western China: a comparison of hydrological and curvature watershed methods for slope unit classification. Bull Eng Geol Env 79:4657–4670

Sun X, Zeng P, Li T, Wang S, Jimenez R, Feng X, Xu Q (2021) From probabilistic back analyses to probabilistic run-out predictions of landslides: a case study of Heifangtai terrace, Gansu Province, China. Eng Geol 280:105950. https://doi.org/10.1016/j.enggeo.2020.105950

Takahashi T (2009) A review of Japanese debris flow research. Int J Eros Control Eng 2(1):1–14

Tang Y, Feng F, Guo Z, Feng W, Li Z, Wang J, Sun Q, Ma H, Li Y (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 Clean Prod 277:124159

Tehrani FS, Calvello M, Liu Z, Zhang L, Lacasse S (2022) Machine learning and landslide studies: recent advances and applications. Nat Hazards 114(2):1197–1245

Thorp HH (2023) ChatGPT is fun, but not an author. Science 379(6630):313. https://doi.org/10.1126/science.adg7879

Tien Bui D, Nguyen QP, Hoang N-D, Klempe H (2017) A novel fuzzy K-nearest neighbor inference model with differential evolution for spatial prediction of rainfall-induced shallow landslides in a tropical hilly area using GIS. Landslides 14:1–17

Tien Bui D, Shahabi H, Omidvar E, Shirzadi A, Geertsema M, Clague J, Khosravi K, Pradhan B, Pham BT, Chapi K, Barati Z, Lee S (2019) Shallow landslide prediction using a novel hybrid functional machine learning algorithm. Remote Sens 11(8):931. https://doi.org/10.3390/rs11080931

Tomás R, Li Z (2017) Earth observations for geohazards: present and future challenges. Remote Sens 9:194

Ullah I, Aslam B, Shah SHIA, Tariq A, Qin S, Majeed M, Havenith H-B (2022) An integrated approach of machine learning, remote sensing, and GIS data for the landslide susceptibility mapping. Land 11(8):1265

Vali A, Comai S, Matteucci M (2020) Deep learning for land use and land cover classification based on hyperspectral and multispectral earth observation data: a review. Remote Sens 12(15):2495
Vallet A, Varron D, Bertrand C, Fabbri O, Mudry J (2016) A multi-dimensional statistical rainfall threshold for deep landslides based on groundwater recharge and support vector machines. Nat Hazards 84:821–849
Van Dao D, Jaafari A, Bayat M, Mafi-Gholami D, Qi C, Moayedi H, Van Phong T, Ly HB, Le TT, Trinh PT, Luu C (2020) A spatially explicit deep learning neural network model for the prediction of landslide susceptibility. CATENA 188:104451
Van Westen CJ, Castellanos E, Kuriakose SL (2008) Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview. Eng Geol 102(3–4):112–131
Vaze J, Teng J, Spencer G (2010) Impact of DEM accuracy and resolution on topographic indices. Environ Model Softw 25(10):1086–1098
Wang Q, Wang D, Huang Y, Wang Z, Zhang L, Guo Q, Chen W, Chen W, Sang M (2015) Landslide susceptibility mapping based on selected optimal combination of landslide predisposing factors in a large catchment. Sustainability 7(12):16653–16669
Wang F-Y, Zhang JJ, Zheng X, Wang X, Yuan Y, Dai X, Zhang J, Yang L (2016) Where does AlphaGo go: from church-turing thesis to AlphaGo thesis and beyond. IEEE/CAA J Autom Sin 3(2):113–120
Wang Y, Tang H, Wen T, Ma J (2019) A hybrid intelligent approach for constructing landslide displacement prediction intervals. Appl Soft Comput 81:105506
Wang X, Zhang C, Wang C, Liu G, Wang H (2021) GIS-based for prediction and prevention of environmental geological disaster susceptibility: From a perspective of sustainable development. Ecotoxicol Environ Saf 226:112881. https://doi.org/10.1016/j.ecoenv.2021.112881
Wang Y, Tang H, Huang J, Wen T, Ma J, Zhang J (2022) A comparative study of different machine learning methods for reservoir landslide displacement prediction. Eng Geol 298:106544. https://doi.org/10.1016/j.enggeo.2022.106544
Wang F, Harindintwali JD, Wei K, Shan Y, Mi Z, Costello MJ, Grunwald S, Feng Z, Wang F, Guo Y, Wu X, Tiedje JM (2023a) Climate change: Strategies for mitigation and adaptation. Innov Geosci 1(1):100015–100061. https://doi.org/10.59717/j.xinn-geo.2023.100015
Wang L, Zuo B, Le Y, Chen Y, Li J (2023b) Penetrating remote sensing: next-generation remote sensing for transparent earth. The Innovation 4(6):100519
Wang Z, Chen J, Chen W, Wan C, Liu Y, Huang J (2023) Susceptibility assessment of environmental geological disasters in Liulin County based on RF: from the perspective of positive and negative sample proportion. Environ Sci Pollut Res 30:122245–122261
Wen T, Tang H, Wang Y, Lin C, Xiong C (2017) Landslide displacement prediction using the GA-LSSVM model and time series analysis: a case study of Three Gorges Reservoir, China. Nat Hazard 17(12):2181–2198. https://doi.org/10.5194/nhess-17-2181-2017
Wickel B, Lehner B, Sindorf N (2007) HydroSHEDS: a global comprehensive hydrographic dataset. Paper presented at the AGU Fall Meeting Abstracts
Wu Y, Shao K, Piccialli F, Mei G (2022) Numerical modeling of the propagation process of landslide surge using physics-informed deep learning. Adv Model Simul Eng Sci 9(1):1–15
Xu S, Niu R (2018) Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorges area, China. Comput Geosci 111:87–96
Xu K, Guo Q, Li Z, Xiao J, Qin Y, Chen D, Kong C (2015) Landslide susceptibility evaluation based on BPNN and GIS: a case of Guojiaba in the Three Gorges Reservoir area. Int J Geogr Inf Sci 29(7):1111–1124
Xue Y, Kong F, Li S, Zhang Q, Qiu D, Su M, Li Z (2021) China starts the world's hardest "Sky-High Road" project: challenges and countermeasures for Sichuan–Tibet railway. Innovation (camb) 2(2):100105. https://doi.org/10.1016/j.xinn.2021.100105
Yalcin A (2008) GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): comparisons of results and confirmations. CATENA 72(1):1–12
Yang B, Yin K, Lacasse S, Liu Z (2019) Time series analysis and long short-term memory neural network to predict landslide displacement. Landslides 16:677–694
Yang H, Wei F, Ma Z, Guo H, Su P, Zhang S (2020) Rainfall threshold for landslide activity in Dazhou, southwest China. Landslides 17:61–77
Yang C, Liu L-L, Huang F, Huang L, Wang X-M (2023) Machine learning-based landslide susceptibility assessment with optimized ratio of landslide to non-landslide samples. Gondwana Res 123:198–216
Yilmaz I (2010) Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine. Environ Earth Sci 61:821–836

Yin J, Dong J, Hamm NA, Li Z, Wang J, Xing H, Fu P (2021) Integrating remote sensing and geospatial big data for urban land use mapping: a review. Int J Appl Earth Obs Geoinf 103:102514

Yonggang J, Chaoqi Z, Liping L, Dong W (2016) Marine geohazards: review and future perspective. Acta Geologica Sin Engl Ed 90(4):1455–1470

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 J, Gurung DR, Liu R, Murthy MSR, Su F (2015) Abe Barek landslide and landslide susceptibility assessment in Badakhshan Province, Afghanistan. Landslides 12:597–609

Zhang P, Wu H-N, Chen R-P, Chan TH (2020) Hybrid meta-heuristic and machine learning algorithms for tunneling-induced settlement prediction: a comparative study. Tunn Undergr Space Technol 99:103383

Zhang J, Tang H, Tannant DD, Lin C, Xia D, Liu X, Zhang Y, Ma J (2021a) Combined forecasting model with CEEMD-LCSS reconstruction and the ABC-SVR method for landslide displacement prediction. J Clean Prod 293:126205

Zhang Y-G, Tang J, Liao R-P, Zhang M-F, Zhang Y, Wang X-M, Su Z-Y (2021b) Application of an enhanced BP neural network model with water cycle algorithm on landslide prediction. Stoch Env Res Risk Assess 35:1273–1291

Zhang W, Gu X, Tang L, Yin Y, Liu D, Zhang Y (2022a) Application of machine learning, deep learning and optimization algorithms in geoengineering and geoscience: comprehensive review and future challenge. Gondwana Res 109:1–17

Zhang Y, Tang J, Cheng Y, Huang L, Guo F, Yin X, Li N (2022b) Prediction of landslide displacement with dynamic features using intelligent approaches. Int J Min Sci Technol 32(3):539–549

Zhang S, Tan S, Zhou J, Sun Y, Ding D, Li J (2023a) Geological disaster susceptibility evaluation of a random-forest-weighted deterministic coefficient model. Sustainability 15(17):12691

Zhang Z, Liu M, Tan YJ, Walter F, He S, Chmiel M, Su J (2024) Landslide hazard cascades can trigger earthquakes. Nat Commun 15(1):2878

Zhang Z, Pan Q, Yang Z, Yang X (2023) Physics-informed deep learning method for predicting tunnelling-induced ground deformations. Acta Geotech 18:4957–4972

Zhao Y, Meng X, Qi T, Chen G, Li Y, Yue D, Qing F (2021) Modeling the spatial distribution of debris flows and analysis of the controlling factors: a machine learning approach. Remote Sens 13(23):4813

Zhao P, Masoumi Z, Kalantari M, Aflaki M, Mansourian A (2022a) A GIS-based landslide susceptibility mapping and variable importance analysis using artificial intelligent training-based methods. Remote Sens 14(1):211

Zhao Y, Meng X, Qi T, Li Y, Chen G, Yue D, Qing F (2022b) AI-based rainfall prediction model for debris flows. Eng Geol 296:106456. https://doi.org/10.1016/j.enggeo.2021.106456

Zhao Y, Meng X, Qi T, Chen G, Li Y, Yue D, Qing F (2023) Estimating the daily rainfall thresholds of regional debris flows in the Bailong River Basin, China. Bull Eng Geol Env 82(2):46

Zhou C, Yin K, Cao Y, Ahmed B, Fu X (2018a) A novel method for landslide displacement prediction by integrating advanced computational intelligence algorithms. Sci Rep 8(1):7287. https://doi.org/10.1038/s41598-018-25567-6

Zhou C, Yin K, Cao Y, Intrieri E, Ahmed B, Catani F (2018b) Displacement prediction of step-like landslide by applying a novel kernel extreme learning machine method. Landslides 15:2211–2225

Zhou X, Wen H, Zhang Y, Xu J, Zhang W (2021) Landslide susceptibility mapping using hybrid random forest with GeoDetector and RFE for factor optimization. Geosci Front 12(5):101211

Zhu X, Xu Q, Tang M, Nie W, Ma S, Xu Z (2017) Comparison of two optimized machine learning models for predicting displacement of rainfall-induced landslide: a case study in Sichuan Province, China. Eng Geol 218:213–222

Zhu C, Xu Y, Tao Z, Zhu H-H, Cao C, He M (2023) Overview of comprehensive risk assessment methods and hazards early warning system for geological hazards in the mountain area. Remote Sens 15:2239

Authors and Affiliations

Xiang Zhang[1,2,3] · Minghui Zhang[1,2,3] · Xin Liu[1,2,3] · Berhanu Keno Terfa[4] · Won-Ho Nam[5] · Xihui Gu[6,7,8] · Xu Zhang[1,2,3] · Chao Wang[9] · Jian Yang[10] · Peng Wang[11] · Chenghong Hu[12] · Wenkui Wu[1] · Nengcheng Chen[1,2]

✉ Xiang Zhang
zhangxiang76@cug.edu.cn

Berhanu Keno Terfa
berhanu.keno@aau.edu.et

Won-Ho Nam
wonho.nam@hknu.ac.kr

Xihui Gu
guxihui421@163.com

Xu Zhang
zhangxufm@126.com

Chao Wang
c.wang@whu.edu.cn

Jian Yang
jian.yang@tum.de

Peng Wang
i3dgis@qq.com

Chenghong Hu
huchenghong.hch@alibaba-inc.com

Wenkui Wu
349974434@qq.com

Nengcheng Chen
chennengcheng@cug.edu.cn

1 National Engineering Research Center of Geographic Information System, School of Geography and Information Engineering, China University of Geosciences, Wuhan 430074, China

2 Hubei Luojia Laboratory, Wuhan 430079, China

3 SongShan Laboratory, Zhengzhou 450046, China

4 Institute of Geophysics, Space Science and Astronomy, Addis Ababa University, Addis Ababa, Ethiopia

5 School of Social Safety and Systems Engineering, Institute of Agricultural Environmental Science, National Agricultural Water Research Center, Hankyong National University, Anseong, Republic of Korea

6 Department of Atmospheric Science, School of Environmental Studies, China University of Geosciences, Wuhan 430074, China

7 Centre for Severe Weather and Climate and Hydro-geological Hazards, Wuhan 430074, China

8 School of Geography and the Environment, University of Oxford, Oxford OX1 3QY, UK

9 State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing (LIESMARS), Wuhan University, Wuhan 430079, China

10 School of Geospatial Information, Information Engineering University, Zhengzhou 450052, China

11 GAEA Space Time Co., Ltd, Wuhan 430079, China

12 AutoNavi Software Co., Ltd, Beijing 100012, China