Advancements, Challenges, and Future Directions in Rainfall-Induced Landslide Prediction: A Comprehensive Review

Do Van Vung¹, The Viet Tran¹, Nguyen Duc Ha² & Nguyen Huy Duong²

¹Department of Civil Engineering, Thuyloi University, 175 Tay Son, Dong Da, Hanoi, Vietnam
²Vietnam Institute of Geosciences and Mineral Resources, 67 Chien Thang, Van Quan, Ha Dong, Hanoi, Vietnam

Corresponding author: Trantheviet@tlu.edu.vn

Abstract

Rainfall-induced landslides threaten lives and properties globally. To address this, researchers have developed various methods and models that forecast the likelihood and behavior of rainfall-induced landslides. These methodologies and models can be broadly classified into three categories: empirical, physical-based, and machine-learning approaches. However, these methods have limitations in terms of data availability, accuracy, and applicability. This paper reviews the current state-of-the-art of rainfall-induced landslide prediction methods, focusing on the methods, models, and challenges involved. The novelty of this study lies in its comprehensive analysis of existing prediction techniques and the identification of their limitations. By synthesizing a vast body of research, it highlights emerging trends and advancements, providing a holistic perspective on the subject matter. The analysis points out that future research opportunities lie in interdisciplinary collaborations, advanced data integration, remote sensing, climate change impact analysis, numerical modeling, real-time monitoring, and machine learning improvements. In conclusion, the prediction of rainfall-induced landslides is a complex and multifaceted challenge, and no single approach is universally superior. Integrating different methods and leveraging emerging technologies offer the best way forward for improving accuracy and reliability in landslide prediction, ultimately enhancing our ability to manage and mitigate this geohazard.

Keywords: climate change; empirical method; landslide monitoring; machine-learning method; physical-based method; rainfall-induced landslides.

Introduction

Rainfall-induced landslides are a type of mass movement that occurs when excess water from rainfall infiltrates into soil or a rock slope, leading to a loss of stability and mass movement [1-4]. This category of disaster bears potentially significant ramifications, encompassing extensive infrastructure impairment, human casualties, and substantial economic consequences. The susceptibility to rainfall-induced landslides is influenced by numerous factors, including topography, geology, soil properties, land use, and climate [4-7]. Areas with steep slopes, weak or weathered rock, and intense or prolonged rainfall are particularly vulnerable to rainfall-induced landslides [3,6,8]. Recently, the frequency and severity of rainfall-induced landslides have increased due to climate change, urbanization, and other human activities [7,9,10]. Therefore, accurate prediction and early warning of rainfall-induced landslides are crucial for mitigating the risks associated with these events and for informing disaster management strategies [1,11].

The field of predicting rainfall-induced landslides encompasses a range of methodologies and models, which can be broadly classified into three overarching categories: empirical models, physically-based models, and machine-learning approaches [1,12,13]. Empirical models are based on statistical relationships between landslides and the environmental variables that trigger them [13-15]. These models use historical landslide data and rainfall measurements to estimate the probability of landslide occurrence under certain rainfall conditions [16]. Physically-based models are based on mathematical equations that describe the physical processes that control landslide behavior, such as soil mechanics, hydrology, and slope stability [3,11,17,18]. These models use rainfall data, soil properties, and other environmental variables to simulate the behavior of slopes and predict landslide occurrence [1,11,18]. Machine learning approaches are based on algorithms that learn from data and make predictions based on patterns and relationships in the data [19-21]. These models use large datasets of
landslide and environmental data to develop predictive models. Each of these methods has its own advantages and limitations, and the choice of method depends on the specific application and available data [22,23]. In recent years, there have been efforts to integrate these different prediction methods to improve the accuracy and reliability of landslide predictions [11,17,18,22,24].

This paper presents a comprehensive review of the current state of rainfall-induced landslide prediction methods, including their methods, models, strengths, limitations, and challenges. Our objective was to provide an overview of the different approaches that have been developed and highlight their effectiveness or limitations in practical application. Furthermore, we also discuss the key challenges associated with current landslide prediction methods, such as the need for detailed data on soil properties and geology, validation with real-world data, and the lack of real-time prediction and warning systems. Additionally, prospective areas were identified for further investigation and potential solutions to address these challenges.

**Overview of Relevant Literature on Various Methods and Models Used for Rainfall-Induced Landslide Prediction**

The prediction of rainfall-induced landslides constitutes a key scientific question with significant social implications. In recent years, various methods and models have been proposed and applied for rainfall-induced landslide prediction, each with its advantages and disadvantages [1,13,25,26]. Empirical methods use historical data and are simple to apply but depend on data availability and the assumption of past patterns [1,5,25]. Physically-based methods use physical laws and provide a mechanistic understanding of landslide processes but need detailed input data and face difficulties in calibration/validation [1,2,18,24]. Machine learning methods use data-driven algorithms and can capture complex relationships from data but also face issues related to data quality and model interpretability [27]. Further research and innovation are needed to enhance the accuracy and reliability of landslide prediction methods and to overcome the challenges associated with these approaches [13,17,19,28].

**Empirical Methods**

Empirical methods in rainfall-induced landslide prediction have developed over the decades as a simple and practical way to assess landslide hazards based on the statistical relationship between rainfall characteristics and landslide occurrence [13]. Empirical methods are widely used in rainfall-induced landslide prediction because they are simple and easy to apply [1,25,29,30]. The basic principle of empirical methods is to use historical data to identify patterns and relationships between rainfall and landslides [25,31]. The empirical threshold method is one such method and uses regression lines/curves to identify rainfall intensity-duration thresholds that can trigger landslides [25]. Figure 1 illustrates the definition of rainfall intensity-duration or intensity-cumulative rainfall thresholds using an empirical approach and historical data.

In accordance with the prioritization of safety assurance or accuracy in the context of landslide warning endeavors, the empirical threshold can be slightly modified to more appropriately correspond with the ratio of non-landslide occurrences or landslide incidents. Establishing the threshold at an excessively low level may lead to an excessive number of false alarms, whereas setting it too high may incur supplementary damages [32,33]. Empirical methods have several advantages, such as being easy to use, requiring minimal data inputs, and providing quick results [25,34,35]. The input data utilized in empirical methods encompass a wide range of variables, including rainfall measurements, time, and the distance from the rainfall measurement station to the location of the landslide (Figure 2). However, the most prevalent approach involves employing combinations of parameters such as rainfall intensity, duration, antecedent rainfall conditions, and cumulative rainfall [26].

However, empirical methods have limitations as well [34]. They are based on historical data and may not be applicable to future events that have different characteristics [25,34]. Empirical methods also do not account for the complex interactions between rainfall, soil properties, and topography that can influence landslide occurrence [25,37]. Some studies have also pointed out the differences and the necessity of establishing different rainfall thresholds for different geological types [38] or for different meteorological, hydrological, and geomorphological characteristics [32]. Therefore, empirical methods should be used with caution and in conjunction with other methods such as physically-based models that can account for these complex interactions [37,39].
Recent advances in empirical methods for predicting rainfall-induced landslides have shown promising results by focusing on improving the accuracy and reliability of landslide forecasts. These methods have been refined and improved through advancements in technology and data availability [25,35]. One notable development is the use of machine learning algorithms to analyze large datasets, including rainfall patterns, soil properties, topography, and historical landslide events. These algorithms have enabled researchers to develop predictive models that can assess the probability of landslides occurring in specific areas during rainfall events [22,40]. In addition, remote sensing techniques such as satellite-based rainfall measurements and LiDAR (light detection and ranging) data have been utilized to gather high-resolution information on rainfall and terrain characteristics, providing more comprehensive and up-to-date information for landslide prediction accuracy [41-43]. Overall, these recent empirical methods have the potential to enhance landslide prediction capabilities, leading to improved landslide risk management and mitigation strategies in vulnerable areas.

Empirical methods for predicting rainfall-induced landslides hold great promise for the future. With advancements in technology such as improved remote sensing techniques, more sophisticated machine learning algorithms, and increased availability of high-resolution rainfall data, empirical methods are expected to become even more accurate and reliable in predicting landslides [5]. For example, real-time monitoring data from sensors installed in landslide-prone areas combined with advanced data analytics may enable more precise and
timely landslide predictions [44]. Additionally, numerical modeling and simulation techniques may allow for a better understanding of landslide mechanisms and the incorporation of more complex factors such as climate change impacts into prediction models. As a result, empirical methods are expected to bring about significant advancements in predicting rainfall-induced landslides, ultimately leading to more effective landslide risk reduction strategies and increased resilience in vulnerable regions.

**Physically-Based Methods**

Methods based on the fundamental principles of physics and mechanics, known as physical-based methods, are utilized to assess the potential occurrence of landslides by considering the physical properties of soil, slope characteristics, and hydrological conditions [1,17,24,25,45]. These approaches employ mathematical models to simulate the behavior of soil and rock layers under varying rainfall conditions [1,2,46]. Physical-based methods offer several advantages in predicting rainfall-induced landslides. They tend to provide more accurate predictions of landslide occurrence and magnitude compared to empirical models [47]. Additionally, physical-based models can estimate the rate of landslide displacement and evaluate the effectiveness of different mitigation measures [35]. The rapid development of unmanned aerial vehicle (UAV) technology has significantly contributed to the effective application of physical models in simulating landslide events. In addition to their remote sensing capabilities for identifying areas susceptible to landslides, UAVs facilitate the expeditious generation of high-resolution digital surface models and the assessment of pre- and post-landslide terrain changes (Figure 3).

![Elevation change (m) (Before - After)](image)

-34.4 - -19.8
-19.8 - -13.2
-13.2 - -7.3
-7.3 - 0
0 - 1.7
1.7 - 5.1
5.1 - 9.6
9.6 - 14.7
14.7 - 20.5
20.5 - 33.2

**Figure 3** Changes in terrain determined using UAV technology for a landslide-affected area in Son La province, Vietnam.
These models play a crucial role as input data for deploying physical models. Therefore, the expedited and cost-efficient implementation of UAV technology holds great potential for enhancing the comprehension and analysis of landslide phenomena.

One of the most advanced models for simulating rainfall-induced landslides is the LS-RAPID model developed by Sassa & Nagai [48]. The model has two advantages: it can simulate the entire process of the landslide event, from the initiation process (stability analysis) to the sliding process (dynamic analysis), including the volume increase during movement; and it can be applied with model parameters determined by laboratory experiments without requiring calibration efforts. However, physical-based models for rainfall-induced landslides have limitations such as the requirement for detailed soil properties and geology data, validation using real-world data, and the absence of real-time prediction and warning systems [3,49]. These limitations are also evident in the usage requirements of the LS-RAPID model. Important parameters of the model such as friction angle during motion, steady-state shear resistance, shear displacement at the start of strength reduction, and shear displacement at the start of steady state (Figure 4) can only be determined using advanced devices such as the undrained dynamic-loading ring-shear apparatus [50].

In recent years, physical-based methods for predicting rainfall-induced landslides have undergone significant advancements. Researchers have been refining and integrating various physical models, such as the infinite slope model [51], the transient infiltration-excess model [11,18], and the coupled hydro-mechanical model [11,18,24,52], to better simulate the complex interactions between rainfall, soil properties, and slope behavior. An example of combining the use of two hydrological and landslide models was tested in a catchment located in Halong City, Vietnam by Ha, Sayama [53]. The LS-RAPID model was applied to identify locations at risk of landslides and the water level that triggers landslides and simulate the entire process of landslide material development. The RRI hydrological model [54] was used to simulate the water level, above and below the ground, of the area during rainfall events. The process of coupling these two models is illustrated in Figure 5. The results of the combined application of the two models were verified with high accuracy, but to apply them directly in an early warning system, further testing work needs to be expanded both in space and time.

The advancements mentioned above have resulted in more accurate predictions of landslide occurrence, magnitude, and potential impact area [42]. Furthermore, the integration of remote sensing data, such as rainfall radar and satellite measurements, into physical-based models has improved the temporal and spatial resolution of rainfall inputs, leading to more reliable and timely landslide predictions [42,55-57]. Additionally, the development of real-time monitoring and warning systems, utilizing technologies such as IoT (Internet of Things) and sensor networks, has enabled early detection and prediction of rainfall-induced landslides, providing valuable time for implementing timely mitigation measures [58]. These recent advancements in physical-based
methods have greatly improved our ability to predict and mitigate rainfall-induced landslides, contributing to better landslide risk management strategies and increasing the resilience of vulnerable areas.

Physical-based methods for predicting rainfall-induced landslides are expected to advance significantly in the coming years with new technologies such as LiDAR and radar, which will enable accurate and timely monitoring of key landslide parameters such as soil moisture, slope stability, and pore pressure. Advancements in numerical modeling techniques such as finite element analysis and computational fluid dynamics will allow for more detailed simulations of landslide processes, including complex interactions between rainfall, soil properties, and slope geometry. Coupling these models with real-time monitoring data will enhance the accuracy and reliability of landslide predictions, enabling early warning systems to be developed for vulnerable regions. The integration of machine learning algorithms and artificial intelligence into landslide prediction models will enable more precise and automated predictions by analyzing large datasets and identifying patterns and trends that may not be apparent to human observers.

![Diagram](image.png)

**Figure 5** Major steps for coupling an LS-RAPID model and an RRI model to produce a rainfall-induced risk index (RI) map ($ru$ = pore water pressure ratio, RI = risk index) (retraced following Ha, Sayama [53]).

**Machine Learning Methods**

Machine learning methods are increasingly being utilized in rainfall-induced landslide prediction due to their ability to analyze large datasets and identify complex patterns that may not be apparent through traditional methods [22,23,49,59-63]. The basic principles of machine learning methods in this context involve training algorithms on historical landslide and rainfall data to learn patterns and relationships between various parameters, such as rainfall intensity, duration, antecedent soil moisture, slope angle, and vegetation cover. These algorithms then use the learned patterns to make predictions on new, unseen data [22,64]. Figure 6 illustrates the main steps of seven advanced machine-learning techniques applied in landslide susceptibility mapping.

Machine learning methods and models can be classified into supervised, unsupervised, and semi-supervised, depending on the availability and use of labeled data. Supervised learning techniques, such as decision tree, support vector machines, and random forest, are commonly used in landslide prediction [22]. Unsupervised learning methods, such as clustering and anomaly detection, can be employed for identifying regions with similar characteristics that may be prone to landslides [65]. It is important to note that the accuracy and reliability of machine learning methods in landslide prediction depend on the quality and representativeness of the input data as well as the appropriate selection and tuning of the algorithms [23,65]. To compare the differences in the abilities and accuracies of various machine learning models, many studies have applied multiple models to build landslide susceptibility maps using the same dataset [59,66,67]. The results of these experiments are often noted to have higher accuracy than conventional techniques such as the analytical hierarchy process [28,68], but the
The selection of the most suitable model is typically recommended on a case-by-case basis, depending on the specific dataset and study area [21].

Figure 6  Process of applying machine learning models to construct landslide susceptibility maps for the Abha Basin, Asir Region, Saudi Arabia was carried out by Ahmed & Hamid [59].

The major advantage of machine learning methods is their ability to handle complex and nonlinear problems, to deal with high-dimensional and heterogeneous data, and to adapt to new data and situations [20,59]. However, this method also has several limitations, such as requiring a large amount of representative and reliable data for training and testing, being prone to overfitting or underfitting problems, being difficult to interpret or explain their results, and being sensitive to the choice of algorithms and parameters [69]. Furthermore, machine learning models are often considered black boxes because they do not provide explanations for their predictions [35].

The development of machine learning approaches for rainfall-induced landslide modeling has been facilitated by the availability of large datasets, such as satellite images, aerial photographs, and remote sensing data [49]. Advanced machine learning algorithms such as deep learning, decision tree, and support vector machines have been employed to analyze and extract patterns from these data, allowing for accurate and timely prediction of rainfall-induced landslides [12]. The future of machine learning methods in rainfall-induced landslide prediction is expected to witness continued advancements. Ongoing research in data acquisition, algorithm development, and model interpretability is likely to make machine-learning models more accurate, reliable, and transparent. The incorporation of diverse data sources such as real-time meteorological data, high-resolution remote sensing data, and historical landslide data is expected to improve the predictive capabilities of these models. Advancements in probabilistic modeling and uncertainty quantification are expected to enhance the reliability and confidence of landslide predictions. The future of machine learning methods in rainfall-induced landslide prediction holds great potential in providing valuable tools for landslide risk assessment, early warning systems, and decision support, contributing to better landslide management and mitigation efforts.
Challenges, Opportunities, and Directions for Future Research on Rainfall-Induced Landslide

Challenges

Despite the progress made, predicting rainfall-induced landslides remains a challenge due to difficulties in obtaining detailed and reliable data on soil properties, geology, and other relevant factors. The integration of multiple data sources can be a complex endeavor, necessitating advanced data fusion techniques. Accurate prediction models that employ advanced statistical and machine-learning techniques require a comprehensive understanding of the underlying physical and geological processes involved in landslides [70]. Furthermore, the incorporation of potential impacts of climate change on landslide risk is complicated by uncertainties in future climate scenarios. Additionally, the development of real-time prediction and warning systems may encounter technical, logistical, and institutional challenges. Overcoming these challenges demands interdisciplinary collaborations among researchers, practitioners, and stakeholders from various fields. Innovative approaches can provide new insights and solutions to improve the accuracy and reliability of rainfall-induced landslide prediction. The integration of expert knowledge with data-driven approaches can enhance the applicability and effectiveness of landslide prediction models in different regions and conditions. Addressing these challenges is crucial to advance the field of rainfall-induced landslide prediction.

Opportunities for Future Research

To effectively comprehend and mitigate rainfall-induced landslide risks, interdisciplinary approaches that synthesize knowledge from various fields are imperative. Opportunities for interdisciplinary research include developing improved methods for predicting landslide susceptibility and hazard, as well as integrating data and models from geology, hydrology, soil science, and remote sensing. Engineers, geologists, ecologists, and social scientists can collaborate to develop more effective and sustainable landslide mitigation strategies, such as slope stabilization measures and suitable vegetation selection. Interdisciplinary research is crucial for understanding and managing the risks associated with rainfall-induced landslides, and for developing comprehensive strategies to reduce impacts on human lives and infrastructure. Furthermore, the advent of sensor technology and the Internet of Things (IoT) presents novel opportunities for the real-time monitoring of slope conditions and the development of early warning systems.

Directions for Future Research

The field of rainfall-induced landslide prediction offers numerous potential avenues for future research. One such direction involves advanced data integration, entailing the exploration of innovative methods to integrate and analyze diverse datasets, including rainfall data, terrain characteristics, geotechnical properties, vegetation cover, and historical landslide occurrences. Another direction is remote sensing and GIS, which involves investigating the use of remote sensing techniques such as LiDAR and synthetic aperture radar (SAR), coupled with geographic information systems (GIS) to enhance landslide prediction. The third direction is climate change impacts, which involves investigating the influence of climate change on rainfall patterns and its impact on landslide occurrence. Numerical modeling represents another direction, seeking to enhance numerical modeling techniques for landslide prediction by incorporating more precise representations of rainfall infiltration, slope stability, and landslide initiation and propagation. Real-time monitoring is a further avenue, aiming to develop robust and cost-effective monitoring systems that can provide real-time data on rainfall, pore pressure, groundwater levels, and slope movements. Machine learning and artificial intelligence also hold promise as potential directions for future research in this field. Improving the effectiveness of early warning systems for rainfall-induced landslides is another direction that emphasizes the integration of meteorological data, ground monitoring data, and predictive models to issue timely warnings to vulnerable communities. Finally, conducting comprehensive risk assessments to identify areas susceptible to rainfall-induced landslides is another direction that can be incorporated into land use planning and urban development policies to minimize exposure and vulnerability.
Conclusion

Rainfall-induced landslides require ongoing research for predicting, monitoring, and mitigating their impacts. This research paper provides a comprehensive review of existing methods and models for predicting rainfall-induced landslides, highlighting their strengths, limitations, and challenges. It identifies opportunities for future research to improve the accuracy and reliability of landslide prediction methods. This paper emphasizes the need to incorporate detailed data on soil properties and geology, use advanced statistical and machine learning techniques, integrate multiple data sources, and consider the potential impacts of climate change on landslide risk. Additionally, developing real-time prediction and warning systems is crucial for effective landslide risk management. By addressing these research gaps and leveraging emerging technologies and approaches, we can advance the field of rainfall-induced landslide prediction and contribute to a better understanding and improved mitigation of this geohazard.

It is important to note that there is no single best method or model for predicting rainfall-induced landslides. Each approach has its strengths and weaknesses, and its performance may vary depending on factors such as data availability, quality, scale, and problem complexity. Therefore, selecting the most suitable method or model for a given case study requires a comprehensive evaluation and comparison of different approaches. Moreover, combining or integrating different methods or models may offer more robust and reliable predictions than relying on a single approach.

References


