Santos *et al.*
Smart Construction and Sustainable Cities (2025) 3:21
https://doi.org/10.1007/s44268-025-00071-9

Smart Construction and Sustainable Cities

REVIEW

Open Access

Machine learning-based hydrological models for flash floods: a systematic literature review

Leonardo B. L. Santos[1,2,3*], Elton V. Escobar-Silva[1], Luiz F. Satolo[1,2], Ricardo S. Oyarzabal[1], Michael M. Diniz[4], Rogério G. Negri[5], Glauston R. T. Lima[1], Stephan Stephany[2], Jaqueline A. J. P. Soares[1,2], Johan S. Duque[2,6], Fernando L. Saraiva Filho[3], Angélica N. Caseri[7] and Luiz Bacelar[8]

Abstract

Flash floods are critical events for emergency management, yet their modeling remains highly challenging, even in smart cities approaches. Physically based hydrological models are often unsuitable at small spatiotemporal scales due to their computational complexity and dependence on detailed local parameters, which are rarely available during flash floods. With the growing availability of hydrological data, machine learning (ML) has emerged as a promising alternative. This work performs a Systematic Literature Review (SLR) to improve our understanding of the research landscape on ML applications for flash flood forecasting, a significant subset of flash flood modeling. From more than 1,200 papers published until January 2024 in Web of Science, SCOPUS/Elsevier, Springer/Nature, and Wiley, 50 were selected following PRISMA guidelines. The inclusion and exclusion criteria removed reviews, retractions, papers focused on post-flood damage assessment (not forecasting), and those with time resolutions of 6 hours or more, retaining only studies with fine-scale temporal data (<6 hours). For each paper, we extracted information on forecasting horizon, study area size, input data, ML techniques, and outcomes (regression or classification). Results show a sharp rise in ML-based flash flood research, with China leading (38%). Nearly all studies rely on rainfall, discharge, and water level data - often in combination. Long short-term memory (LSTM) networks dominate (60%). Unfortunately, only 10% of the selected studies provide access to their datasets. This lack of transparency poses a major barrier to reproducibility, inhibits fair comparative evaluation of models, and ultimately slows methodological progress in flash flood forecasting. Furthermore, our review highlights that no method consistently outperforms others. This variability in performance is likely influenced by factors such as regional hydrological characteristics (e.g., differences between arid and tropical basins), variations in input data quality, and the length of the forecast horizon (e.g., 1- vs. 6-hour prediction). Lastly, we recommend advancing this field through integration with early warning systems, creation of benchmarks, open data practices, and stronger multidisciplinary collaboration.

Keywords Artificial intelligence, Machine learning, Flash floods, Hydrological modeling, Disasters

*Correspondence:
Leonardo B. L. Santos
leonardo.santos@cemaden.gov.br
Full list of author information is available at the end of the article

1 Introduction

Nearly 44% of disasters worldwide have been associated with floods, and different types of floods account for 31% of economic losses [1]. It is estimated that from 2000 to 2024, floods affected over 1.8 billion people and caused a global annual average economic loss of US$ 38.88 billion [2]. Compounding this issue, ongoing global climate change is anticipated to raise the frequency and severity of such events [3, 4].

Flash floods are among the most common types of natural disasters worldwide [5, 6]. They are often defined based on observations of river streamflow or water levels, with several quantitative metrics emphasizing the significance of a high peak discharge or a rapid rise in water level (typically within<6 h) [7], usually triggered by heavy rainfall [8], quick snowmelt [9], or induced by dam and levee breaks [10]. One example of a flash flood metric, for instance, is the Flashiness-Intensity-Duration-Frequency (F-IDF) curve, which is based on the frequency and duration of various rainfall events [11]. Although there is no general consensus among the scientific community regarding a metric that defines flash floods, their triggering rainfall events typically occur on a small spatiotemporal scale. Regardless of the climatic study area, they are predominantly observed in urban locations with steep terrain or inadequate drainage systems, particularly in regions prone to severe weather events [4, 6]. Smaller and steeper watersheds respond more rapidly to intense precipitation, resulting in a shorter time lag between the onset of heavy rainfall and the rise of water levels or river discharge. This can provide less warning time to residents and emergency responders [8].

Hydrological models are employed to study the hydrologic cycle, representing a component (or stage) of it [12]. There are many forms of hydrological models since they are designed to deal with different problems. These models take into account multiple factors, such as catchment characteristics and the spatial and temporal variations in rainfall [13], which can effectively characterize flash flood behaviors. Consequently, they serve as crucial tools for flash flood prediction and for issuing timely warnings.

Despite advances in physically-based hydrological models [14], such models are typically applied to flood forecasting in larger watersheds with slower responses. They are not designed to detect rainfall and runoff variations that occur on a small spatiotemporal scale, which can lead to flash floods. To monitor trigger mechanisms, operational flash flood forecasting relies on high-resolution remote sensing data, such as weather radar, to estimate accumulated rainfall volumes or utilize weather numerical models to forecast precipitation at short lead times [6]. The increased availability of observed hydrological data (e.g., water levels and discharge) has led to an increase in the usage of data-driven hydrological models, in which time series of river levels or discharges are predicted without needing to know the physical parameters related to the watershed [15]. Given the availability of good-quality observed data, data-driven models can more accurately predict river dynamics responses, requiring less computational time and calibration than physically-based hydrological models [16].

Artificial intelligence (AI) is a broad field of research dedicated to creating systems that use computer programs to mimic human intelligence and cognitive processes. These systems aim to perform tasks such as reasoning, learning, adaptively interacting with their environment, and making decisions without explicit instructions [17]. Among the various approaches to designing AI systems, the most notable is the application of machine learning (ML) techniques, which generally rely on the principle of learning exclusively from data [18]. It is important to highlight that the significant progress AI has made in recent years can be largely attributed to improvements in the predictive capabilities of ML techniques, particularly through deep learning (DL) (see Fig. 1).

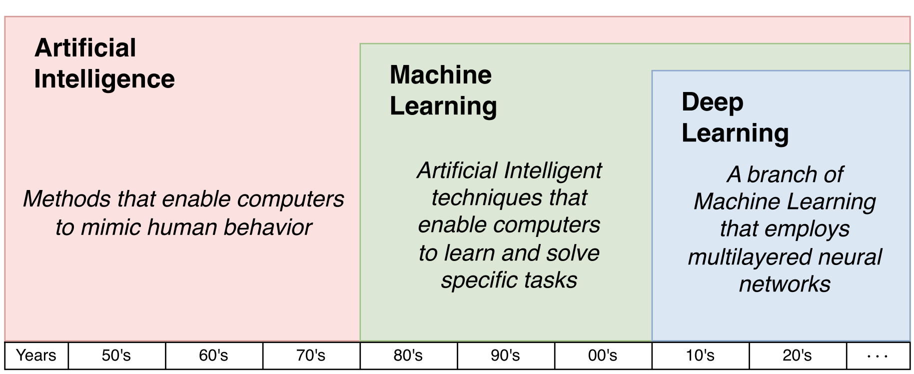

Fig. 1 Artificial intelligence (AI) trend towards deep learning (DL) for hydrological forecast/prediction over the years

Although ML has gained significant traction in hydrological modeling, it is not the only approach applied to flood risk assessment. Multi-criteria decision-making (MCDM) methods, such as the Analytic Hierarchy Process (AHP), its fuzzy extensions (FAHP), and the Analytic Network Process (ANP), have also been widely employed, often in combination with GIS, to evaluate flood risk under multiple factors [19–22]. Compared with ML, MCDM methods are more transparent, computationally simple, and suitable when expert judgment and stakeholder participation are critical. However, they are limited by subjectivity and potential inconsistencies in expert assessments. In contrast, ML techniques can process large-scale heterogeneous datasets and achieve higher predictive accuracy in flood forecasting, which explains their growing adoption in recent years.

As a result of the remarkable growth of ML methodologies in hydrological modeling, there is a need for periodic literature reviews aimed at identifying significant advancements and challenges within this area. In 2014, a considerable contribution to this field was presented by [23], where the authors conducted a comprehensive examination of contemporary advancements and the potential utility of support vector machine (SVM) techniques within hydrology. In the following year, [24] investigated the use of ML for streamflow forecasting from 2000 to 2015. The research revealed that over the examined years, ML methods showed substantial advancements in hydrological forecasting and simulation, effectively capturing complex information in the data that previous methods could not.

Since 2021, there has been a significant increase in the publication of review articles focused on applying ML in the field of hydrology. Notably, we highlight the work by [25], in which the authors explored the progress of employing ensemble methods across various hydrological application domains. Their findings indicate a general trend of superior performance compared to traditional ML models. In the context of runoff, a thorough examination is presented in [26], where the authors evaluated the specific use of adaptive neuro-fuzzy inference systems (ANFIS), artificial neural networks (ANN), and SVM for runoff simulations. The primary goal of this review was to clarify the main advantages and limitations of each of these methodologies. Additional reviews on the use of ML in hydrological contexts can be found in [27–29].

Furthermore, with the advancement of scientific repository search tools, the potential for methodically organizing and reproducing literature review protocols has emerged, leading to the establishment of a paradigm known as *Systematic Reviews* [30]. In [31], a systematic review is conducted on the state-of-the-art ML and DL methods in predicting hydrological processes, climate changes, and earth systems. Other more general systematic reviews involving hydrology can be found in [32]. Given this context, the primary objective of this paper is to conduct a Systematic Literature Review (SLR) to examine the current landscape of ML applications in flash flood forecasting (a key subset of flash flood modeling), assessing the extent and frequency with which such approaches have been employed and how significant they have become. More specifically, this review seeks to identify the most commonly applied ML techniques, highlighting those with superior predictive skill and assessing the frequency of their use as regressors or classifiers. To the best of our knowledge, this represents the first comprehensive literature review on ML models for flash floods. We outlined the scope of the review to address key questions regarding flash flood forecasting while maintaining conciseness.

2 Methodology

This review covers articles on ML and hydrological models through an extensive search in large scientific databases; for this purpose, it follows the process suggested by [33, 34], and the resources of *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* - also known as PRISMA 2020 - to make the review transparent and replicable. A table with all processed steps and the PRISMA 2020 checklist is available at https://github.com/rogerionegri/iFAST.

Our search strategy employed keywords relevant to the research questions, utilizing Boolean operators (AI, ML, DL, hydrology, hydrological model, hydrological forecast, flood, rainfall-runoff, fast response, fast dynamic, rapid response, rapid dynamic, short lead time, and/or short-term forecast). These terms were organized into overarching concepts or tiers. We used "OR" to encompass synonyms and alternative spellings while using "AND" to connect primary terms with secondary ones. Articles published in peer-reviewed journals in the English language were considered up until December 2023. The following databases were consulted: Web of Science, SCOPUS/Elsevier, Springer/Nature, and Wiley. The searches included paper titles, keywords, and abstracts. No limit was imposed on the number of articles returned in the query. Additionally, we included 20 other papers based on our prior knowledge of the literature.

During the initial screening process, we eliminated duplicate papers. During the screening stage, we applied strict inclusion and exclusion criteria to ensure the relevance and quality of the studies selected for this review (ML applications for flash flood forecasting, a key subset of flash flood modeling). Reviews, retractions, and papers clearly outside the scope were removed. In particular, we excluded studies mainly focused on rainfall forecasting, groundwater prediction, flood mapping, coastal flooding, or tsunami forecasting, since these topics, although related, do not directly target flash flood modeling. We also discarded papers relying on data with a temporal resolution coarser than six hours, as such datasets are insufficient for capturing the rapid dynamics of flash flood events. After this systematic filtering, the final dataset for detailed analysis comprised 50 papers, which collectively offer a representative overview of recent advances in applying machine learning techniques to flash flood forecasting. From this final set of ML models for flash floods, we analyzed various characteristics. A datasheet containing all 50 selected papers and their attributes can be found at https://github.com/rogerionegri/iFAST. A summary of the attributes considered in this study is presented in Table 1. Lastly, the PRISMA diagram for this systematic review is presented in Fig. 2.

3 Results and discussion

Figure 3 shows the number of published works, both annually and cumulatively, related to the topic of this review.

There has been a recent significant increase in publications on ML techniques for flash flood forecasting, from 9 articles over 21 years (2000-2020) to 41 articles in just 3 years (2021-2023). This growth may be attributed to either the increasing frequency of flash floods or the newly available ML methods.

Table 1 Summary of attributes observed in the reviewed papers

Attribute	Description
Area of study	Country in which the research is carried out
Data availability	If data is public
Input data	Input data used in the model(rainfall, water level, or discharge)
Lead time (min)[h]	Minimum forecast horizon
Lead time (max)[h]	Maximum forecast horizon
ML main method	Type of ML method
Model output data	Level, discharge, or both
Regression, classification, or both	The model predicts categories or classes for each element, respectively
Remote sensing	If the paper uses remote sensing data (radar or satellite)
Temporal resolution (min)	Temporal resolution of input data

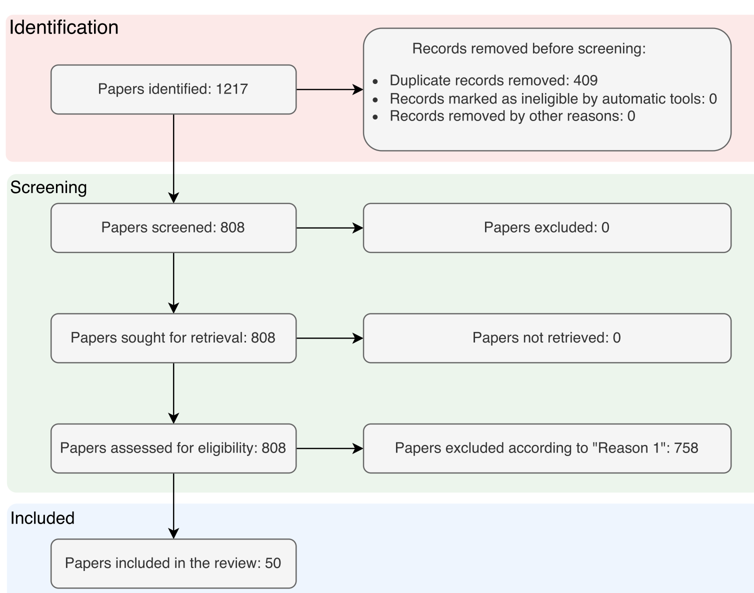

Fig. 2 The resulting PRISMA 2020 workflow diagram

The top seven journals encompass a diverse range of fields, from hydrology to applications of computer science. Regarding the frequency of articles reviewed per journal, as shown in Fig. 4, *Water* (MDPI) and *Journal of Hydrology* (Elsevier) are the two main sources of research on ML for modeling flash floods.

3.1 In which countries is research on machine learning (ML) and flash floods most commonly found?

Figure 5 illustrates a spatial representation of the number of studies conducted across various regions of the globe. This representation highlights that the revised studies encompass 20 countries spread throughout Asia, Europe, and North and South America. The majority of the study areas are situated in Asia,

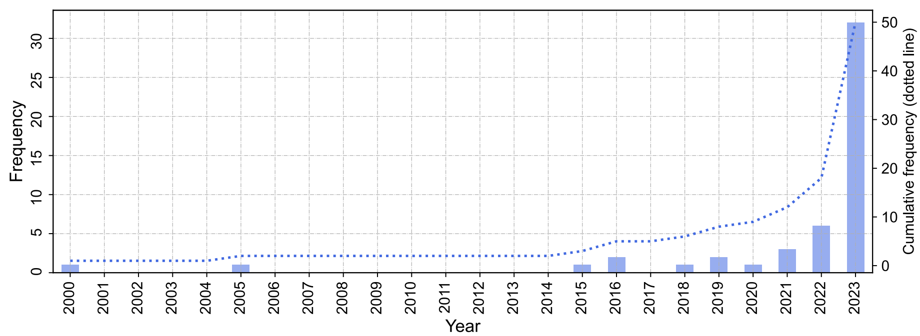

Fig. 3 Absolute and cumulative number of publications about machine learning (ML) applied to flash floods per year

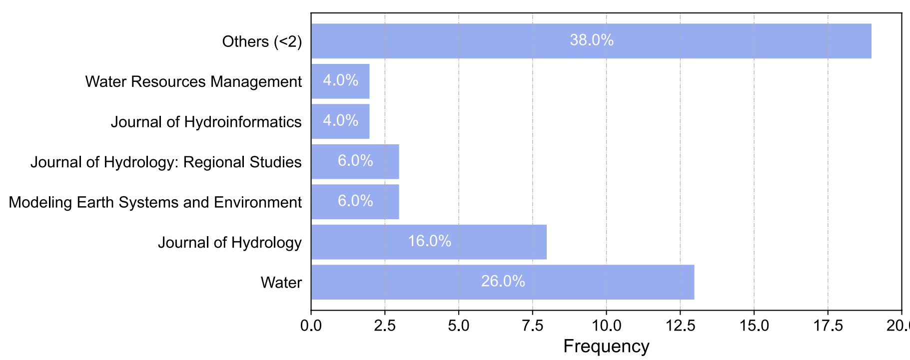

Fig. 4 Frequency of articles using machine learning (ML) for flash flood hydrological modeling by journal

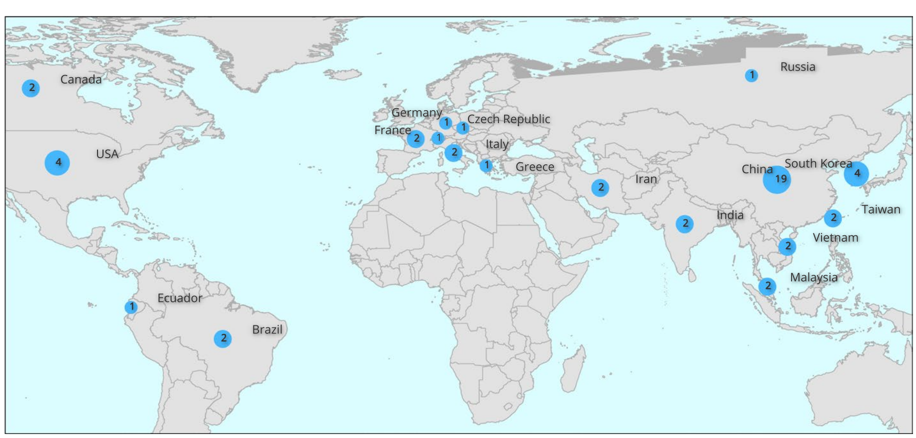

Fig. 5 Number of articles applying ML to flash flood prediction that had an area of study in a particular country

particularly in China (38%) and the Republic of Korea (8%), as well as in the United States (8%).

This scenario may reflect the fact that China and the United States are among the countries with the most experience in dealing with floods worldwide, alongside India, Indonesia, the Philippines, and Brazil [35]. China is the country most severely threatened by flood disasters worldwide, with damages from these events between 1990 and 2017 accounting for approximately 10% of the total global damage [36]. Flash floods, in particular, are widely recognized as a significant cause of human casualties and economic losses in China [37, 38]. From the American perspective, flash floods result in the highest number of deaths among various flood events in the U.S. [39, 40]. American national assessments have shown that the eastern U.S. frequently experiences flood events, accounting for a substantial proportion of the country's flood-induced fatalities. This is partly due to tropical cyclone-related precipitation, which contributes nearly 30% of the annual rainfall in the region, given its geographic position [41]. Finally, looking at the Republic of Korea, most small urban river basins in the country have a very short concentration time, which leads to frequent and deadly accidents caused by flash floods during located heavy rainfall [42].

3.2 What are the most commonly used input and output data in machine learning (ML) models for flash floods?

Among the 50 selected articles, rainfall is the most commonly applied input variable, appearing in 44 studies (88%). Discharge data is included in 23 studies (46%), and water level data is employed in 19 studies (38%). 49 articles employed only one or a specific combination of the following measurements: discharge, rainfall, and water level. Only one study used runoff as the sole input data [43]. Notably, four papers (8%) combined rainfall, water level, and discharge data simultaneously. Additionally, four studies relied solely on rainfall data, three studies (6%) on water level data, and two studies (4%) on discharge data. Figure 6 illustrates the distribution of the input data used in the 50 studies analyzed in this work.

It is important to note that discharge and rainfall are the most common combinations in these studies and are also suitable for physically based models. Studies that utilize water level data in ML applications for flash flood prediction hold significant potential, as obtaining water level data is often easier than acquiring discharge data [44].

Among the selected articles, 30 papers (60%) reported discharge as output, while 19 works(38%) indicated water level as output. Remarkably, only one article (2%) did not present either the discharge or water level as output. In [45], dynamic clustering and random forest techniques were employed to identify flood types and select suitable

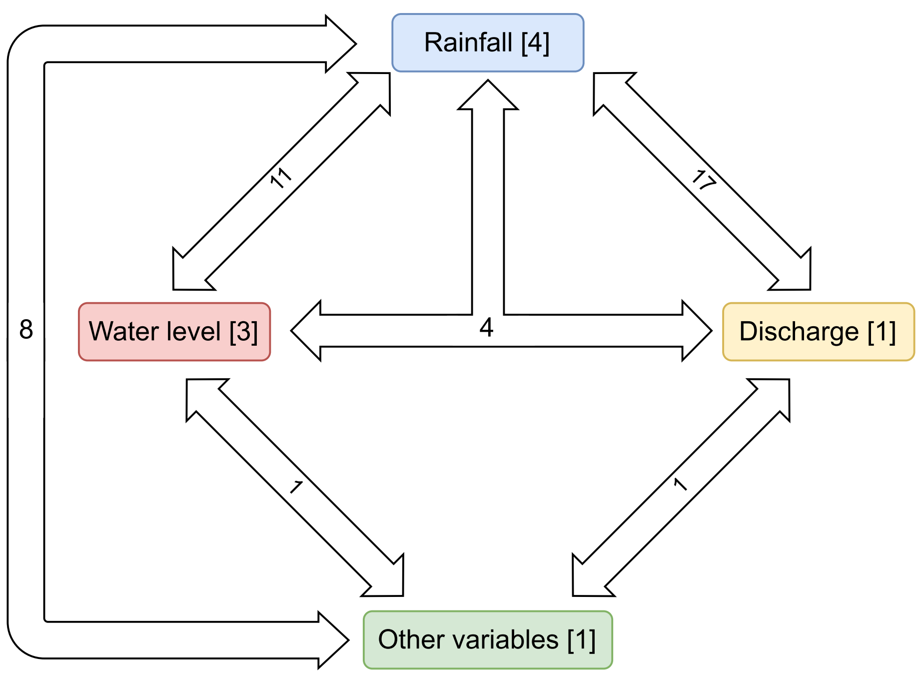

Fig. 6 Distribution of the input data applied by the 50 analyzed articles. Values between brackets mean the number of papers where the input data was employed solely; values within the arrows mean the number of papers that share the input data

model parameters. Following this, the Xinanjiang model for real-time flood forecasting was implemented. In this method, three flood indicators were recognized as the most crucial factors characterizing flooding: flood duration, peak discharge, and runoff depth.

Lastly, while more than half of the articles provided discharge and water level as the outputs of their studies instead of a flood extent map, most utilized historical flooding event records in their simulations, for instance, to train a neural network or to endorse and validate their findings, as referenced in [46–51].

3.3 What are the most common machine learning (ML) techniques for modeling flash floods?

Table 2 presents a list of the most commonly used ML methods for modeling flash floods. It also provides a brief description of each method and the respective papers that feature them as one of the main methods (those demonstrating the best performance). Furthermore, the frequency of each ML method applied to hydrological modeling is illustrated in Fig. 7. LSTM is the most utilized method, appearing in 60% of the works, followed by MLP, which is used in 28%. The revised studies also employed CNN, tree-based methods (decision trees or random forests), and SVM, used in 16%, 16%, and 14% of the papers, respectively. Other methods, such as k-Means, KNN, Extreme ML, Particle Swarm Optimization, and Fuzzy-based methods, were employed less frequently.

Figure 8 shows a comparison of all ML methods presented in the articles, including both result comparison methods and the main methods, which represent the highest performance in each paper. Notably, LSTM was utilized and ranked as one of the best methods in this set of papers. However, no single method consistently outperforms all others. Therefore, it is essential to explore various methods for each research problem to determine the most suitable approach for each case study.

3.4 What are the minimum and maximum lead times and what temporal resolution have scientists used to investigate flash flood forecasting?

The lead time values ranged from 5 minutes [49] to 720 hours [53]. Most sub-hourly predictions utilized multiple variables for training, typically a combination of water level and rainfall [49–51, 88, 93, 95, 104, 107]. A mix of hourly rainfall and discharge was primarily used to forecast lead times beginning at 1 hour and extending to a maximum of 720 hours (e.g., [53]). The majority of studies applying LSTM methods projected discharge for lead times ranging from 1 to at least 6 hours [9, 46, 47, 51, 53, 65, 89].

3.5 Is remote sensing commonly used in machine learning (ML) hydrological models?

Despite being a common data source in many environmental studies and applications [78], remotely sensed data were applied in only seven papers (14%) (Fig. 9). This limited usage may be due to the coarse spatial resolution often associated with meteorological products (e.g., precipitation and other environmental descriptors) derived from remote sensing data, as well as uncertainties related to their estimates. Additionally, the frequent unavailability of meteorological RADAR sensors may further contribute to this limited use. Consequently, studies might favor or depend on other data sources, such as ground-based measurements, hydrological models, or historical flood records. Lastly, the temporal resolution of remotely sensed data may not align well with the temporal dynamics of flash floods, which necessitate high-frequency data for accurate modeling.

However, while not widely utilized in the literature, it is important to emphasize that remotely sensed data, particularly those acquired by RADAR sensors, can offer valuable insights and support for ML-based approaches aimed at predicting flash floods [117].

3.6 Data availability

Among the reviewed articles, only 10% of them made the data used in the research publicly available ([59, 64, 76, 88, 89], while 90% of them did not share the data (Fig. 9). Although it is essential to respect the data confidentiality policies of companies and institutions, this result is concerning as it hinders the ability to replicate and validate findings. Furthermore, it limits collaborations within the scientific community that could advance research in this field. Lastly, data sharing accelerates the pace of discovery and its benefits to society.

3.7 What is the most common problem: regression or classification?

Regression is the most common method for predicting flash floods, as indicated by the selected papers. As shown in Fig. 9, 41 out of the 50 articles utilized at least one regression algorithm to forecast flash floods. Among these, four articles also incorporated a classification algorithm to address this issue. Furthermore, five articles employed both regression and classification algorithms to predict flash floods.

The dominance of regression algorithms can be explained by the fact that the variable of interest, i.e., the output data, is continuous in most of the articles included in this review. Basically, regression analysis is an ML approach that aims to predict the values of continuous output variables using input variables.

Table 2 List of the most used ML methods for modeling flash floods

Method	Short description	Papers in this review
AE [52]	Neural network architecture designed for unsupervised learning that learns to encode input data into a latent representation and reconstruct it with minimal loss.	[53]
ANFIS [54]	Hybrid system that combines fuzzy logic and NN techniques for adaptive modeling and inference.	[55]
ARMA [56]	Combines autoregressive and moving average components to predict a time series based on its own past values and error terms, balancing short and long-term dependencies	[57]
BMA [58]	Statistical technique that combines Bayesian models in a temporal framework, considering changes in relationships between variables over time.	[59]
CGBR [60]	Advanced ensemble model that incorporates ordered boosting for categorical features. It employs minimal variance sampling to balance tree growth, enhancing prediction accuracy and computational efficiency.	[61, 62]
CNN [63]	Deep learning architectures adept at processing structured grid data, utilizing convolutional layers to learn hierarchical features automatically.	[51, 64–67]
Conv-LSTM [68]	Integrates convolutional operations within LSTM units. It processes input sequences by convolving spatial features and capturing temporal dependencies simultaneously, enhancing the model's ability to learn spatiotemporal patterns efficiently.	[51, 61, 65]
DANN [69]	The Dynamic that adjusts the structure of the neural network during training	[70]
DNN [71]	Deep Neural Networks learn complex features by passing data through multiple layers of interconnected nodes, or neurons, mimicking human brain function for tasks like image recognition and natural language processing	[50]
DSTGNN [72]	Method for modeling dynamic spatiotemporal data, leveraging GNN to capture spatial dependencies and temporal dynamics efficiently	[73]
DT [74]	A machine learning algorithm that recursively partitions data based on feature values to create a predictive model represented by a tree-like structure	[75–78]
ELGBDT [79]	An ensemble learning technique that combines the strengths of Extreme Learning Machines and Gradient Boosted Decision Trees for efficient and accurate predictive modeling	[80]
Enconder-Decoder (ED) [52]	NN architecture consisting of an encoder and decoder, trained to learn a compressed representation of input data by minimizing the reconstruction error between input and output	[67]
GAN [81]	Deep learning framework consisting of two neural networks, the generator and the discriminator, engaged in a minimax game. The generator synthesizes data while the discriminator distinguishes between real and generated samples, aiming to achieve equilibrium in generating realistic data distributions	[43]
GRU [82]	Type of RNN, designed to capture long-range dependencies in sequential data, featuring simplified memory cells and gating mechanisms for efficiency in training	[61, 62, 67, 83, 84]
k-Means [85]	Clustering algorithm that partitions data into K clusters based on similarity, iteratively adjusting cluster centroids until convergence	[45, 76, 78]
k-NN [86]	Lazy supervised learning method where a data point is classified by a majority vote of its *k* nearest neighbors.	[57]
LSTM [87]	RNN designed to capture long-term dependencies in sequential data by utilizing specialized memory cells and gating mechanisms	[9, 43, 46, 47, 51, 53, 61, 62, 64–67, 67, 73, 83, 84, 88–101]
MARS [102]	Statistical method for non-linear regression analysis, employing piecewise linear segments to model complex relationships between multiple predictor variables and a response variable.	[76]
MLP [103]	NN with multiple layers of interconnected neurons, including an input layer, one or more hidden layers, and an output layer. It utilizes backpropagation for supervised learning.	[50, 51, 57, 67, 77, 89, 99, 101, 104–107]
OPENML [108]	Technique in machine learning that efficiently prunes irrelevant neurons from extreme learning machines to enhance model performance and reduce computational complexity	[76]
Random Forest (RF) [109]	An ensemble learning method in machine learning, consisting of multiple decision trees during training, resulting in improved accuracy and reduced overfitting through the aggregation of predictions.	[45, 51, 75, 110]

Table 2 (continued)

Method	Short description	Papers in this review
RNN [111]	Process sequential data by retaining information from previous inputs, making them suitable for tasks involving sequences such as time series prediction and natural language processing.	[50, 112]
SVM [113]	Supervised ML algorithm that constructs a hyperplane in high-dimensional space to classify data points by maximizing the margin between different classes while minimizing classification error.	[62, 66, 67, 89, 105, 114, 115]
Transformer [116]	NN architecture based on self-attention mechanisms, enabling parallel processing of sequential data by capturing long-range dependencies without recurrent connections, yielding significant advancements in various natural language processing tasks	[101]
XGBoost [91]	Gradient boosting algorithm that efficiently handles various regression and classification tasks by sequentially adding weak learners, employing regularization techniques to prevent overfitting	[49, 77]

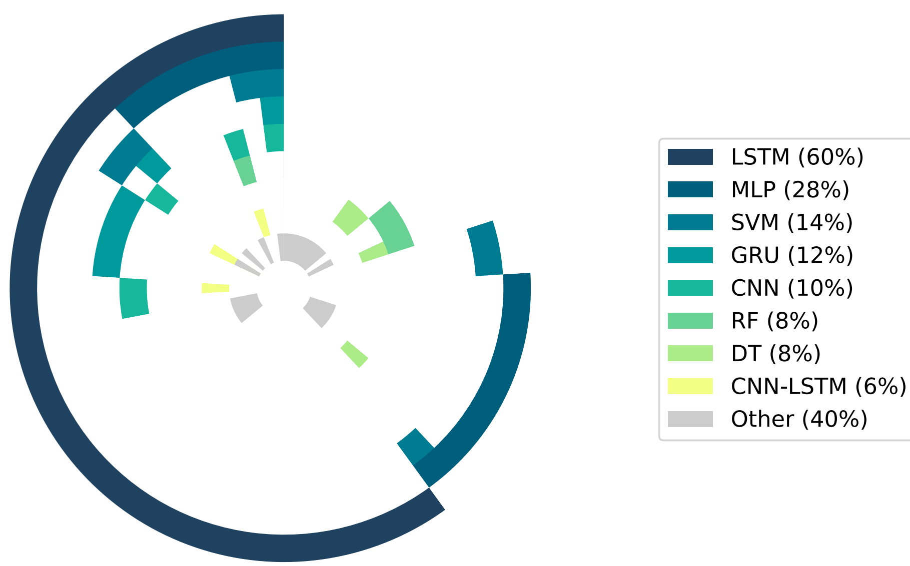

Fig. 7 Frequency of machine learning (ML) methods for hydrological modeling of flash floods

4 Main findings and open questions

The future of machine learning (ML) for flash flood forecasting must be discussed in light of the limitations observed in current research. First, the review revealed that only 10% of studies share their datasets, a barrier that limits reproducibility and hinders comparative assessments across different methods. Thus, advancing the field requires open data initiatives and standardized benchmarks that directly address this lack of transparency.

Second, although long short-term memory (LSTM) networks dominate recent studies (60%), no method consistently outperforms others. This highlights the need for comparative experiments using common datasets and evaluation metrics. Developing benchmarks would reduce this methodological fragmentation and foster fair performance comparisons.

Third, remote sensing data were rarely employed (14%), despite their potential to capture rainfall dynamics over ungauged basins. This limitation suggests that future research should prioritize integrating remote sensing products with ground-based measurements to overcome data scarcity in vulnerable regions.

Finally, many studies treat ML as "black-box" models, providing little insight into uncertainty or physical interpretability. Addressing this limitation requires physics-informed ML approaches, explainable AI techniques, and systematic uncertainty propagation studies.

Feature selection is the process of choosing the set of variables to be used as input in an algorithm. It is a widely adopted data preprocessing step in ML. In addition to enabling faster algorithms, it can also provide a better understanding of the underlying physical processes being modeled [118]. Feature selection has been applied

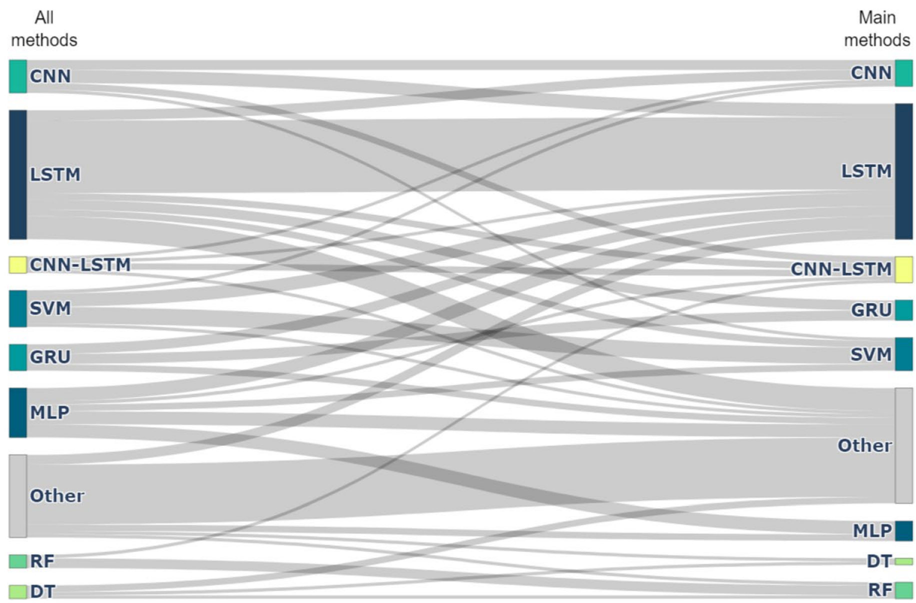

Fig. 8 Proportion of all mentioned machine learning (ML) methods (left) and the main methods of each selected paper (right). A connection (gray line) from a method on the right to a method on the left means that those methods were compared in the same paper - and the method from the right in this connection was one of those with the best performances in that paper

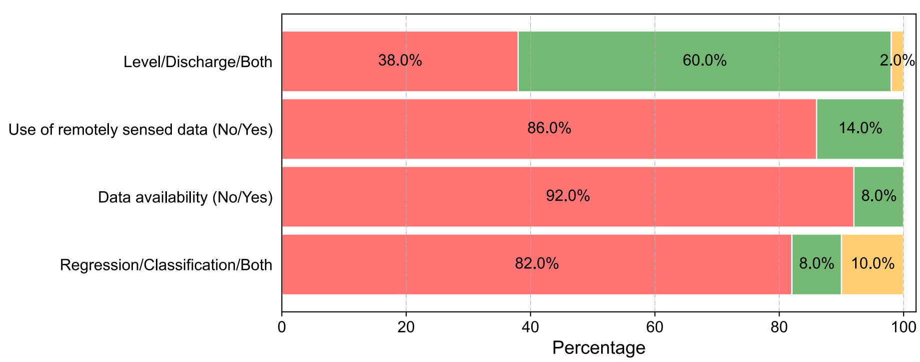

Fig. 9 Percentage ratio of the 50 articles that (i) presented water level, discharge, or both as output; (ii) used remotely sensed data; (iii) made data available; and (iv) applied regression, classification, or both

in various studies of streamflow forecasting. In [119], a comparison of eight filter-based feature selection methods is performed for monthly streamflow forecasting. In [120], within the context of daily streamflow forecasting, a comparison is made between the feature selection ability of a hydrologist and that of different model structures that select automatically. However, despite the work already performed, more comparative studies on the application of feature selection for hourly streamflow prediction still need to be conducted, which may be further explored.

The uncertainty analysis for hydrological models remains an important open question as well. The complex nature of modeling real-world hydrological processes, particularly flash floods, presents a persistent challenge. Understanding and quantifying the uncertainties associated with input and calibration data, model structural elements, and parameters is essential. These uncertainties not only affect the reliability of predictions but also influence decision-making processes for flash flood forecasting. A recent review of hydrological model uncertainties suggests that this issue is still in its early stages and requires further exploration and investigation [121]. Recent research has recognized the significance of this issue [122], but more is needed.

ML models could also help interpret flash flood events. For instance, [123–125] applied explainable ML methods (e.g., using SHAP - SHapley Additive exPlanations - values) through input data features as a scalable approach to identify flood events across different spatial and temporal scales.

Recently, new mesh-free approaches have emerged with the help of ML methods that integrate available observations and compute surrogate solutions for nonlinear partial differential equations (PDEs), such as the Saint-Venant equation related to hydraulic problems [126, 127]. For instance, [128] established a physics-informed ML (PIML) model to combine the predictive capabilities of ML algorithms with the understanding of hydrological processes in physics-based models. A physics-informed learning algorithm, such as physics-informed neural networks (PINN), can solve PDE using feed-forward neural network architectures and incorporate physical laws that represent spatial and temporal changes through computational methods for automatic differentiation [129]. Many challenges remain in ML algorithms for hydrology, including black box models and surrogate models, where the objective function is approximated by optimizing the model's hyperparameters to achieve optimal solutions. Therefore, there is a pressing need to generate mathematical and computational knowledge of substitute modeling related to physical phenomena and data observations, which may yield promising results as a support tool for hydrological studies in watersheds at various temporal and spatial resolutions.

In addition, two of the major challenges in real-time flash flood forecasting are the inherent trade-off between forecast lead time and accuracy, as well as robustness. In this regard, the reliability of early warnings can be compromised by systematic biases in rainfall forecasts, including the underestimation of extreme event intensity, errors in spatial placement, and temporal shifts in predicted rainfall. Several ML-based approaches can help mitigate these issues. Some of them are outlined below: (i) CNN can be used for downscaling numerical forecasts to obtain more accurate rainfall estimates with improved spatial resolution and to correct systematic error patterns; (ii) RNN and LSTM can be trained to learn corrections based on historical patterns of forecast errors, adjusting predicted rainfall to better match observations; (iii) MLPs, Random Forests, and XGBoost can be trained to estimate actual streamflow from biased rainfall forecasts, thereby reducing the impact of errors in the early detection of flash floods; and (iv) incorporating outputs from an ensemble of weather models into ML models can help reduce systematic bias and improve forecast reliability. Such integration of ML models into early warning systems offers a promising pathway to improving both the lead time and accuracy of flash flood alerts by mitigating biases in rainfall forecasts.

ML methods appear to be robust in predicting flash floods. Their data-oriented nature allows them to implicitly adapt to various input data sources, such as rain gauges, weather radar estimates of rainfall, water levels, or discharges. Additionally, ML methods offer a low processing cost for hydrological modeling. For example, a neural network may require a few hours for the training and validation phases, but once trained, the resulting model operates quickly enough to meet real-time demands within minutes or even seconds.

5 Getting evidence into practice

The application of ML approaches in flash flood forecasting is promising. However, to transform this theoretical potential into direct and practical products and applications and maximize its impact, a series of actions involving collective efforts must be undertaken. In this context, some recommendations are outlined below:

Integration of ML into early warning systems: Integrate ML models into early warning systems, as these models can be updated in real-time with hydrological, meteorological, and satellite data to identify patterns indicative of flood occurrences and issue alerts with a better balance between lead time and assertiveness. Close cooperation is essential among ML developers, specialists (e.g., meteorologists and hydrologists), and civil defense agents in monitored risk areas to ensure that alerts remain accurate and interpretable.

Development and dissemination of benchmarks: Establish standardized benchmarks derived from diverse datasets and realistic scenarios, providing them to the scientific community for (i) assessing the effectiveness of developed ML solutions, (ii) ensuring their reliability and practical applicability, and (iii) promoting fast innovations in the field.

Publications and reviews focused on case studies: Publications showcasing successful case studies provide valuable insights into the challenges encountered and the strategies employed to overcome them. This can bolster the confidence of other researchers and practitioners in ML approaches and offer practical guidance for implementing these solutions in their contexts.

Multidisciplinary collaboration and scientific events: Organizing events such as workshops, seminars, and scientific conferences that bring together experts in AI, hydrology, disaster management, and public policy encourages exchange and collaboration among these professionals. This is essential for developing and implementing integrated solutions that promote innovations in flood forecasting, aligned with social and environmental needs.

Lastly, the selection of keywords determines which papers are eligible for inclusion in the analysis. In this study, only papers containing the keywords “artificial intelligence”, “machine learning”, or “deep learning” were considered. This choice results in the exclusion of some relevant papers on flash flood forecasting that apply traditional statistical methods but were not associated with ML or AI by their authors, such as [130, 131]. Future systematic reviews on flash flood forecasting may explicitly consider statistical and physically based methods.

Abbreviations

AE	Autoencoders
GAN	Generative Adversarial Network
ANFIS	Artificial Neural Network and Fuzzy Inference System
GRU	Gated recurrent units
ARMA	Autoregressive–Moving-Average
k-NN	K-nearest neighbors algorithm
BMA	Bayesian Model Averaging
LSTM	Long Short-Term Memory
CGBR	Categorical Gradient Bosting Regression
MARS	Multivariate adaptive regression spline
CNN	Convolutional Neural Networks
MLP	Multilayer Perceptron
Conv-LSTM	Convolutional Long Short-Term Memory
OPENML	Open Machine Learning
DANN	Domain-Adversarial Neural Network
PSO	Particle swarm optimization
DNN	Deep Neural Network
RF	Random Forest
DSTGNN	Dynamic Spatiotemporal Graph Neural Network
RNN	Recurrent neural network
DT	Decision Tree
SVM	Support Vector Machine
ED	EnconderDecoder
XGBoost	Extreme Gradient Boosting
ELGBDT	Extreme Learning Machines and Gradient Boosted Decision Trees

Acknowledgements
Not applicable.

Authors' contributions
All authors contributed equally to this manuscript.

Funding
This study was financed by the CNPq Project 446053/2023-6 and by the São Paulo Research Foundation (FAPESP) grant 2024/02748-7. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Data availability
All data used in this work can be accessed at https://github.com/rogerionegri/iFAST.

Declarations

Ethics approval and consent to participate
Not applicable.

Competing interests
The authors declared no potential conflicts of interest with respect to the research.

Author details
[1]National Center for Monitoring and Early Warning of Natural Disasters, Cemaden, Estrada Dr. Altino Bondensan, 500, São José dos Campos 12247-016, São Paulo, Brazil. [2]National Institute for Space Research, INPE, Av. dos Astronautas, 1758, São José dos Campos 12227-010, São Paulo, Brazil. [3]Federal University of São Paulo, UNIFESP, Avenida Cesare Mansueto Giulio Lattes, 1201, São José dos Campos 12247-014, São Paulo, Brazil. [4]Federal Institute of São Paulo, IFSP, Rod. Presidente Dutra, km 145, São José dos Campos 12223-201, São Paulo, Brazil. [5]São Paulo State University, UNESP, Rod. Presidente Dutra, km 137.8, São José dos Campos 12247-004, São Paulo, Brazil. [6]Technological University of Uruguay, UTEC, Maciel sn/esq Morquio, Durazno, Durazno 97000, Durazno Department, Uruguay. [7]University of São Paulo, Av. Trabalhador São-carlense, 400, São Carlos 13566-590, São Paulo, Brazil. [8]Duke University, Science Dr., Durham 27710, North Carolina, USA.

Received: 28 July 2025 Revised: 16 September 2025 Accepted: 25 September 2025
Published online: 11 October 2025

References

1. WMO: (2021) WMO Atlas of Mortality and Economic Losses from Weather, Climate and Water Extremes (1970-2019). WMO-No. 1267, 90p., ISBN:978-92-63-11267-5, WMO, Geneva. https://library.wmo.int/idurl/4/57564481
2. EM-DAT U CRED (2024) The international disaster datase. https://public.emdat.be/data. Accessed 2 Sep 2024
3. Guerreiro SB, Dawson RJ, Kilsby C, Lewis E, Ford A (2018) Future heat-waves, droughts and floods in 571 European cities. Environ Res Lett 13(3):034009
4. Li Z, Gao S, Chen M, Gourley JJ, Liu C, Prein AF, Hong Y (2022) The conterminous United States are projected to become more prone to flash floods in a high-end emissions scenario. Commun Earth Environ 3(1):86
5. Gourley JJ, Hong Y, Flamig ZL, Arthur A, Clark R, Calianno M, Ruin I, Ortel T, Wieczorek ME, Kirstetter PE, Clark E, Krajewski WF (2013) A unified flash flood database across the United States. Bull Am Meteorol Soc 94(6):799–805. https://doi.org/10.1175/BAMS-D-12-00198.1
6. Georgakakos KP, Modrick TM, Shamir E, Campbell R, Cheng Z, Jubach R, Sperfslage JA, Spencer CR, Banks R (2022) The flash flood guidance system implementation worldwide: a successful multidecadal research-to-operations effort. Bull Am Meteorol Soc 103(3):E665–E679
7. Hapuarachchi H, Wang Q, Pagano T (2011) A review of advances in flash flood forecasting. Hydrol Process 25(18):2771–2784
8. Bucherie A, Werner M, van den Homberg M, Tembo S (2022) Flash flood warnings in context: combining local knowledge and large-scale hydro-meteorological patterns. Nat Hazard 22(2):461–480. https://doi.org/10.5194/nhess-22-461-2022
9. Yan H, Sun N, Wigmosta MS, Duan Z, Gutmann ED, Kruyt B, Arnold JR (2023) The role of snowmelt temporal pattern in flood estimation for a small snow-dominated basin in the Sierra Nevada. Water Resour Res 59(10):e2023WR034496
10. Yang Q, Guan M, Peng Y, Chen H (2020) Numerical investigation of flash flood dynamics due to cascading failures of natural landslide dams. Eng Geol 276:105765. https://doi.org/10.1016/j.enggeo.2020.105765
11. Li Z, Gao S, Chen M, Zhang J, Gourley JJ, Wen Y, Yang T, Hong Y (2023) Introducing flashiness-intensity-duration-frequency (f-idf): a new metric to quantify flash flood intensity. Geophys Res Lett 50(23):e2023GL104992. https://doi.org/10.1029/2023GL104992
12. Devia GK, Ganasri BP, Dwarakish GS (2015) A review on hydrological models. Aquat Procedia 4:1001–1007
13. Beven KJ (2012) Rainfall-runoff modelling: the primer. John Wiley & Sons
14. Clark MP, Bierkens MFP, Samaniego L, Woods RA, Uijlenhoet R, Bennett KE, Pauwels VRN, Cai X, Wood AW, Peters-Lidard CD (2017) The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism. Hydrol Earth Syst Sci 21(7):3427–3440. https://doi.org/10.5194/hess-21-3427-2017
15. Tripathy KP, Mishra AK (2024) Deep learning in hydrology and water resources disciplines: concepts, methods, applications, and research directions. J Hydrol 628:130458

16. Ng K, Huang Y, Koo C, Chong K, El-Shafie A, Ahmed AN (2023) A review of hybrid deep learning applications for streamflow forecasting. J Hydrol 625:130141
17. Soori M, Arezoo B, Dastres R (2023) Artificial intelligence, machine learning and deep learning in advanced robotics, a review. Cognitive Robotics 3:54–70
18. Mosavi A, Ozturk P, Kw C (2018) Flood prediction using machine learning models: literature review. Water 10(11):1536
19. Lyu HM, Yin ZY, Zhou A, Shen SL (2024) Sensitivity analysis of typhoon-induced floods in coastal cities using improved anp-gis. Int J Disaster Risk Reduct 104:104344
20. Lyu HM, Yin ZY, Zhou A, Shen SL (2023) Mcdm-based flood risk assessment of metro systems in smart city development: a review. Environ Impact Assess Rev 101:107154
21. Lyu HM, Yin ZY (2023) An improved mcdm combined with gis for risk assessment of multi-hazards in Hong Kong. Sustain Cities Soc 91:104427
22. Lyu H, Yin Z, Shen S, Chen X, Su D (2025) A novel framework incorporating machine learning into gis for flood susceptibility prediction of urban metro systems. Sci China Technol Sci 68(6):1–17
23. Deka PC et al (2014) Support vector machine applications in the field of hydrology: a review. Appl Soft Comput 19:372–386
24. Yaseen ZM, El-Shafie A, Jaafar O, Afan HA, Sayl KN (2015) Artificial intelligence based models for stream-flow forecasting: 2000–2015. J Hydrol 530:829–844
25. Zounemat-Kermani M, Batelaan O, Fadaee M, Hinkelmann R (2021) Ensemble machine learning paradigms in hydrology: a review. J Hydrol 598:126266
26. Mohammadi B (2021) A review on the applications of machine learning for runoff modeling. Sustain Water Resour Manag 7(6):98
27. Lange H, Sippel S (2020) Machine learning applications in hydrology. Forest-water interactions. Springer, Cham. 240. pp 233–257
28. Mosaffa H, Sadeghi M, Mallakpour I, Jahromi MN, Pourghasemi HR (2022) Application of machine learning algorithms in hydrology. In: Computers in earth and environmental sciences. Elsevier, pp 585–591
29. Mashala MJ, Dube T, Mudereri BT, Ayisi KK, Ramudzuli MR (2023) A systematic review on advancements in remote sensing for assessing and monitoring land use and land cover changes impacts on surface water resources in semi-arid tropical environments. Remote Sens 15(16):3926
30. Pati D, Lorusso LN (2018) How to write a systematic review of the literature. HERD Health Environ Res Des J 11(1):15–30
31. Ardabili S, Mosavi A, Dehghani M, Várkonyi-Kóczy AR (2020) Deep learning and machine learning in hydrological processes climate change and earth systems a systematic review. In: Engineering for Sustainable Future: Selected papers of the 18th International Conference on Global Research and Education Inter-Academia–2019 18. Springer, pp 52–62
32. Leitzke B, Adamatti D (2021) Multiagent system and rainfall-runoff model in hydrological problems: a systematic literature review. Water 13(24):3643
33. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, et al (2021) Prisma 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 372:1–36
34. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, et al (2021) The prisma 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372:1–9
35. Hu P, Zhang Q, Shi P, Chen B, Fang J (2018) Flood-induced mortality across the globe: spatiotemporal pattern and influencing factors. Sci Total Environ 643:171–182
36. Kundzewicz ZW, Su B, Wang Y, Xia J, Huang J, Jiang T (2019) Flood risk and its reduction in China. Adv Water Resour 130:37–45
37. Liu Y, Huang Y, Wan J, Yang Z, Zhang X (2020) Analysis of human activity impact on flash floods in China from 1950 to 2015. Sustainability 13(1):217
38. Zhao G, Liu R, Yang M, Tu T, Ma M, Hong Y, Wang X (2022) Large-scale flash flood warning in China using deep learning. J Hydrol 604(127):222
39. Ashley ST, Ashley WS (2008) Flood fatalities in the United States. J Appl Meteorol Climatol 47(3):805–818
40. Terti G, Ruin I, Anquetin S, Gourley JJ (2017) A situation-based analysis of flash flood fatalities in the United States. Bull Am Meteorol Soc 98(2):333–345
41. Khouakhi A, Villarini G, Vecchi GA (2017) Contribution of tropical cyclones to rainfall at the global scale. J Clim 30(1):359–372
42. Lee JH, Yuk GM, Moon HT, Moon YI (2020) Integrated flood forecasting and warning system against flash rainfall in the small-scaled urban stream. Atmosphere 11(9):971
43. Weng P, Tian Y, Liu Y, Zheng Y (2023) Time-series generative adversarial networks for flood forecasting. J Hydrol 622(129):702
44. Nevo S, Morin E, Gerzi Rosenthal A, Metzger A, Barshai C, Weitzner D, Voloshin D, Kratzert F, Elidan G, Dror G et al (2022) Flood forecasting with machine learning models in an operational framework. Hydrol Earth Syst Sci 26(15):4013–4032
45. Tang Y, Sun Y, Han Z, Wu Q, Tan B, Hu C et al (2023) Flood forecasting based on machine learning pattern recognition and dynamic migration of parameters. Journal of Hydrology Regional Studies 47(101):406
46. Hu C, Wu Q, Li H, Jian S, Li N, Lou Z (2018) Deep learning with a long short-term memory networks approach for rainfall-runoff simulation. Water (Basel) 10(11):1543. https://doi.org/10.3390/w10111543
47. Song T, Ding W, Wu J, Liu H, Zhou H, Chu J (2019) Flash flood forecasting based on long short-term memory networks. Water 12(1):109. https://doi.org/10.3390/w12010109
48. Yan L, Chen C, Hang T, Hu Y (2021) A stream prediction model based on attention-LSTM. Earth Sci Inform 14(2):723–733
49. Sanders W, Li D, Li W, Fang ZN (2022) Data-driven flood alert system (FAS) using extreme gradient boosting (XGBoost) to forecast flood stages. Water 14(5):747
50. Saint-Fleur BE, Allier S, Lassara E, Rivet A, Artigue G, Pistre S, Johannet A (2023) Towards a better consideration of rainfall and hydrological spatial features by a deep neural network model to improve flash floods forecasting: case study on the Gardon basin, France. Model Earth Syst Environ 9(3):3693–3708
51. Zhou F, Chen Y, Liu J (2023) Application of a new hybrid deep learning model that considers temporal and feature dependencies in rainfall-runoff simulation. Remote Sens 15(5):1395
52. Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. Science 313(5786):504–507
53. Devi G, Sharma M, Sarma P, Phukan M, Sarma KK (2022) Flood frequency modeling and prediction of beki and pagladia rivers using deep learning approach. Neural Process Lett 54(4):3263–3282
54. Jang JS (1993) Anfis: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23(3):665–685
55. Nayak P, Sudheer K, Rangan D, Ramasastri K (2005) Short-term flood forecasting with a neurofuzzy model. Water Resour Res. https://doi.org/10.1029/2004WR003562
56. Whittle P (1951) Hypothesis testing in time series analysis. Uppsala: Almquist Wiksell, 121p., Doctorate thesis, University of Uppsala
57. Toth E, Brath A, Montanari A (2000) Comparison of short-term rainfall prediction models for real-time flood forecasting. J Hydrol 239(1–4):132–147
58. Sun Y, Hong Y, Lee TH, Wang S, Zhang X (2021) Time-varying model averaging. J Econom 222(2):974–992
59. Zhou Y, Wu Z, Xu H, Wang H, Ma B, Lv H (2023) Integrated dynamic framework for predicting urban flooding and providing early warning. J Hydrol 618(129):205
60. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A (2018) CatBoost: unbiased boosting with categorical features. Adv Neural Inf Process Syst 31:1–11
61. Guo WD, Chen WB, Chang CH (2023) Prediction of hourly inflow for reservoirs at mountain catchments using residual error data and multiple-ahead correction technique. Hydrol Res 54(9):1072–1093
62. Guo WD, Chen WB, Chang CH (2023) Error-correction-based data-driven models for multiple-hour-ahead river stage predictions: a case study of the upstream region of the Cho-Shui River, Taiwan. J Hydrol Reg Stud 47(101):378
63. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
64. Chiacchiera A, Sai F, Salvetti A, Guariso G (2022) Neural structures to predict river stages in heavily urbanized catchments. Water 14(15):2330

65. Dehghani A, Moazam HMZH, Mortazavizadeh F, Ranjbar V, Mirzaei M, Mortezavi S, Ng JL, Dehghani A (2023) Comparative evaluation of lstm, cnn, and convlstm for hourly short-term streamflow forecasting using deep learning approaches. Eco Inform 75(102):119
66. Huang H, Lei X, Liao W, Liu D, Wang H (2023) A hydrodynamic-machine learning coupled (hmc) model of real-time urban flood in a seasonal river basin using mechanism-assisted temporal cross-correlation (mtc) for space decoupling. J Hydrol 624(129):826
67. Huang J, Li J, Oh J, Kang H (2023) Lstm with spatiotemporal attention for iot-based wireless sensor collected hydrological time-series forecasting. Int J Mach Learn Cybern 14(10):3337–3352
68. Shi X, Chen Z, Wang H, Yeung DY, Wong WK, Woo Wc (2015) Convolutional lstm network: a machine learning approach for precipitation nowcasting. Adv Neural Inf Process Syst 28:1–9
69. Sabour S, Frosst N, Hinton GE (2017) Dynamic routing between capsules. Adv Neural Inf Process Syst 30:1–11
70. Banihabib ME, Arabi A, Salha AA (2015) A dynamic artificial neural network for assessment of land-use change impact on warning lead-time of flood. Int J Hydrol Sci Technol 5(2):163–178
71. Robbins H, Monro S (1951) A stochastic approximation method. Ann Math Stat. https://doi.org/10.1214/aoms/1177729586
72. Diao Z, Wang X, Zhang D, Liu Y, Xie K, He S (2019) Dynamic spatial-temporal graph convolutional neural networks for traffic forecasting. Proceedings of the AAAI conference on artificial intelligence 33:890–897
73. Yang S, Zhang Y, Zhang Z (2023) Runoff prediction based on dynamic spatiotemporal graph neural network. Water 15(13):2463
74. Das Gupta S (1980) Discriminant analysis. In: Hinkley DV (ed) Fienberg SE. R.A. Fisher, An Appreciation, Springer, New York, New York, NY, pp 161–170
75. Erechtchoukova M, Khaiter P, Saffarpour S (2016) Short-term predictions of hydrological events on an urbanized watershed using supervised classification. Water Resour Manage 30:4329–4343
76. Adnan RM, Petroselli A, Heddam S, Santos CAG, Kisi O (2021) Comparison of different methodologies for rainfall-runoff modeling: machine learning vs conceptual approach. Nat Hazards 105:2987–3011
77. Belyakova P, Moreido V, Tsyplenkov A, Amerbaev A, Grechishnikova D, Kurochkina L, Filippov V, Makeev M (2022) Forecasting water levels in Krasnodar Krai rivers with the use of machine learning. Water Resour 49(1):10–22
78. Wang H, Xu S, Xu H, Wu Z, Wang T, Ma C (2023) Rapid prediction of urban flood based on disaster-breeding environment clustering and bayesian optimized deep learning model in the coastal city. Sustain Cities Soc 99(104):898. https://doi.org/10.1016/j.scs.2023.104898. Publisher: Elsevier Ltd
79. Liu J, Wu C (2017) A gradient-boosting decision-tree approach for firm failure prediction: an empirical model evaluation of Chinese listed companies. J Risk Model Validation. https://doi.org/10.21314/JRMV.2017.170
80. He S, Niu G, Sang X, Sun X, Yin J, Chen H (2023) Machine learning framework with feature importance interpretation for discharge estimation: a case study in Huitanggou sluice hydrological station, China. Water 15(10):1923
81. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2020) Generative adversarial networks. ACM New York, NY,
82. Chung J, Gulcehre C, Cho K, Bengio Y (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555
83. He S, Sang X, Yin J, Zheng Y, Chen H (2023) Short-term runoff prediction optimization method based on BGRU-BP and BLSTM-BP neural networks. Water Resour Manage 37(2):747–768
84. Zhang Y, Zhou Z, Van Griensven TJ, Yang SX, Gharabaghi B (2023) Flood forecasting using hybrid lstm and gru models with lag time preprocessing. Water 15(22):3982
85. MacQueen J et al (1967) Some methods for classification and analysis of multivariate observations. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability Oakland CA USA 1:281–297
86. Fix E (1985) Discriminatory analysis: nonparametric discrimination, consistency properties, vol 1. USAF School of Aviation Medicine
87. Graves A (2012) Long short-term memory. In: Supervised Sequence Labelling with Recurrent Neural Networks. Studies in Computational Intelligence. Springer, Berlin, Heidelberg. vol 385.
88. Li W, Kiaghadi A, Dawson C (2021) High temporal resolution rainfall-runoff modeling using long-short-term-memory (LSTM) networks. Neural Comput Appl 33(4):1261–1278
89. Han H, Morrison RR (2022) Data-driven approaches for runoff prediction using distributed data. Stoch Environ Res Risk Assess 36(8):2153–2171
90. Ho HV, Nguyen DH, Le XH, Lee G (2022) Multi-step-ahead water level forecasting for operating sluice gates in hai duong. Vietnam Environ Monit Assess 194(6):442. https://doi.org/10.1007/s10661-022-10115-7. Place: Cham Publisher: Cham: Springer International Publishing
91. Chen J, Li Y, Zhang C, Tian Y, Guo Z (2023) Urban flooding prediction method based on the combination of lstm neural network and numerical model. Int J Environ Res Public Health 20(2):1043
92. Cui Z, Guo S, Zhou Y, Wang J (2023) Exploration of dual-attention mechanism-based deep learning for multi-step-ahead flood probabilistic forecasting. Journal of hydrology (Amsterdam) 622(129):688. https://doi.org/10.1016/j.jhydrol.2023.129688. Publisher: Elsevier B.V
93. Dai Z, Zhang M, Nedjah N, Xu D, Ye F (2023) A hydrological data prediction model based on lstm with attention mechanism. Water 15(4):670
94. Kim D, Lee YO, Jun C, Kang S (2023) Understanding the way machines simulate hydrological processes—a case study of predicting fine-scale watershed response on a distributed framework. IEEE Trans Geosci Remote Sens 61:1–18
95. Koutsovili EI, Tzoraki O, Theodossiou N, Tsekouras GE (2023) Early flood monitoring and forecasting system using a hybrid machine learning-based approach. ISPRS Int J Geo-Inf 12(11):464
96. Le XH, Van LN, Nguyen GV, Nguyen DH, Jung S, Lee G (2023) Towards an efficient streamflow forecasting method for event-scales in ca river basin, Vietnam. J Hydrol Reg Stud 46:101328
97. Liu Y, Liu J, Li C, Liu L, Wang Y (2023) A wrf/wrf-hydro coupled forecasting system with real-time precipitation-runoff updating based on 3dvar data assimilation and deep learning. Water 15(9):1716
98. Moon H, Yoon S, Moon Y (2023) Urban flood forecasting using a hybrid modeling approach based on a deep learning technique. J Hydroinformatics 25(2):593–610
99. Tan WY, Lai SH, Pavitra K, Teo FY, El-Shafie A (2023) Deep learning model on rates of change for multi-step ahead streamflow forecasting. J Hydroinformatics 25(5):1667–1689
100. Zhang L, Qin H, Mao J, Cao X, Fu G (2023) High temporal resolution urban flood prediction using attention-based lstm models. J Hydrol 620:129499
101. Xu Y, Lin K, Hu C, Wang S, Wu Q, Zhang L, Ran G (2023) Deep transfer learning based on transformer for flood forecasting in data-sparse basins. Journal of hydrology (Amsterdam) 625:129956. https://doi.org/10.1016/j.jhydrol.2023.129956
102. Friedman JH (1991) Multivariate adaptive regression splines. Ann Stat 19(1):1–67
103. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323(6088):533–536
104. de Lima GR, Santos LB, de Carvalho TJ, Carvalho AR, Cortivo FD, Scofield GB, Negri RG (2016) An operational dynamical neuro-forecasting model for hydrological disasters. Modeling Earth Systems and Environment 2:1–9
105. Shirali E, Nikbakht Shahbazi A, Fathian H, Zohrabi N, Mobarak Hassan E (2020) Evaluation of WRF and artificial intelligence models in short-term rainfall, temperature and flood forecast (case study). J Earth Syst Sci 129(1):188
106. Lee EH (2023) Inflow prediction of centralized reservoir for the operation of pump station in urban drainage systems using improved multilayer perceptron using existing optimizers combined with metaheuristic optimization algorithms. Water 15(8):1543
107. Santos LB, Freitas CP, Bacelar L, Soares JA, Diniz MM, Lima GR, Stephany S (2023) A neural network-based hydrological model for very high-resolution forecasting using weather radar data. Eng 4(3):1787–1796
108. Miche Y, Sorjamaa A, Bas P, Simula O, Jutten C, Lendasse A (2009) Op-elm: optimally pruned extreme learning machine. IEEE Trans Neural Networks 21(1):158–162
109. Breiman L (2001) Random forests Machine learning 45:5–32
110. Muñoz P, Corzo G, Solomatine D, Feyen J, Célleri R (2023) Near-real-time satellite precipitation data ingestion into peak runoff forecasting models. Environmental Modelling & Software 160:105582

111. Amari SI (1972) Learning patterns and pattern sequences by self-organizing nets of threshold elements. IEEE Trans Comput 100(11):1197–1206
112. Wang Y, Wang W, Zang H, Xu D (2023) Is the lstm model better than rnn for flood forecasting tasks? A case study of Huayuankou station and Loude station in the lower Yellow River Basin. Water 15(22):3928
113. Cortes C, Vapnik V (1995) Support-vector networks. Machine learning 20:273–297
114. Wu J, Liu H, Wei G, Song T, Zhang C, Zhou H (2019) Flash flood forecasting using support vector regression model in a small mountainous catchment. Water 11(7):1327
115. Langhammer J (2023) Flood simulations using a sensor network and support vector machine model. Water 15(11):2004
116. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I (2017) Attention is all you need. Adv Neural Inf Process Syst 30:1–11
117. Munawar HS, Hammad AW, Waller ST (2022) Remote sensing methods for flood prediction: a review. Sensors 22(3):960
118. Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. J Mach Learn Res 3(Mar):1157–1182
119. Ren K, Fang W, Qu J, Zhang X, Shi X (2020) Comparison of eight filter-based feature selection methods for monthly streamflow forecasting-three case studies on camels data sets. J Hydrol 586:124897
120. Moreido V, Gartsman B, Solomatine DP, Suchilina Z (2021) How well can machine learning models perform without hydrologists? Application of rational feature selection to improve hydrological forecasting. Water 13(12):1696
121. Moges E, Demissie Y, Larsen L, Yassin F (2021) Review: sources of hydrological model uncertainties and advances in their analysis. Water 13(1). https://doi.org/10.3390/w13010028
122. Soares JAJP, Diniz MM, Bacelar L, Lima GRT, Soares AKS, Stephany S, Santos LBL (2024) Uncertainty propagation analysis for distributed hydrological forecasting using a neural network. Trans GIS 28(5):971–992. https://doi.org/10.1111/tgis.13169
123. Liu Z, Felton T, Mostafavi A (2023) Identifying flash flood hotspots with explainable machine learning using urban features. In: Proceedings of the 1st ACM SIGSPATIAL International Workshop on Advances in Urban-AI, , UrbanAI '23. Association for Computing Machinery, New York, pp 6–9. https://doi.org/10.1145/3615900.3628792
124. Hadi FAA, Mohd Sidek L, Ahmed Salih GH, Basri H, Sammen SS, Mohd Dom N, Muhamad Ali Z, Najah Ahmed A (2024) Machine learning techniques for flood forecasting. J Hydroinf 26(4):779–799. https://doi.org/10.2166/hydro.2024.208
125. Lyu HM, Yin ZY (2023) Flood susceptibility prediction using tree-based machine learning models in the gba. Sustain Cities Soc 97:104744. https://doi.org/10.1016/j.scs.2023.104744
126. Willard J, Jia X, Xu S, Steinbach M, Kumar V (2022) Integrating scientific knowledge with machine learning for engineering and environmental systems. ACM Comput Surv. https://doi.org/10.1145/3514228
127. Sirignano J, Spiliopoulos K (2018) Dgm: a deep learning algorithm for solving partial differential equations. J Comput Phys 375:1339–1364. https://doi.org/10.1016/j.jcp.2018.08.029
128. Bhasme P, Vagadiya J, Bhatia U (2022) Enhancing predictive skills in physically-consistent way: physics informed machine learning for hydrological processes. J Hydrol 615:128618. https://doi.org/10.1016/j.jhydrol.2022.128618
129. Raissi M, Perdikaris P, Karniadakis G (2019) Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys 378:686–707. https://doi.org/10.1016/j.jcp.2018.10.045
130. Prakash C, Barthwal A, Acharya D (2023) Retracted article: Floodalert: an internet of things based real-time flash flood tracking and prediction system. Multimedia Tools and Applications 82(28):43701–43727
131. Brito LA, Meneguette RI, De Grande RE, Ranieri CM, Ueyama J (2023) Floras: urban flash-flood prediction using a multivariate model. Appl Intell 53(12):16107–16125