JSCDM

JOURNAL OF SOFT COMPUTING AND DATA MINING
e-ISSN: 2716-621X

Vol. 7 No. 1 (2026) 152-184
https://publisher.uthm.edu.my/ojs/index.php/jscdm

# Advancing Flood Prediction and Classification: A Systematic Literature Review

**Jabir Abubakar Salisu[1], Hairulnizam Mahdin[1,4*], Salama A. Mostafa[2], Muhammad Aamir[3], Heru Nurwarsito[4**], Ammar Alazab[5]**

[1] *Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Parit Raja, MALAYSIA*
[2] *Department of Artificial Intelligence Engineering Techniques, College of Technical Engineering, Alnoor University, Mosul 41012, Nineveh, IRAQ*
[3] *Department of Computer Science, University of Oxford, Oxford, UNITED KINGDOM*
[4] *Faculty of Computer Science, University of Brawijaya, Malang, INDONESIA*
[5] *Centre for Artificial Intelligence Research and Optimisation (AIRO), Torrens University Australia, Adelaide, AUSTRALIA*

*Corresponding Author: *hairuln@uthm.edu.my, **heru@ub.ac.id
DOI: https://doi.org/10.30880/jscdm.2026.07.01.012

**Article Info**

Received: 30 October 2025
Accepted: 23 February 2026
Available online: 10 April 2026

**Keywords**

Flood prediction, flood classification, Machine Learning (ML), Deep Learning (DL), Explainable Artificial Intelligence (XAI), Geographic Information System (GIS).

**Abstract**

Floods are the most common and devastating natural disasters, aggravated by climate change and mass urbanization, and pose an urgent need for reliable techniques for prediction and classification to alleviate their situation. Although Machine Learning (ML) and Deep Learning (DL) methods show great potential, unlike existing literature that still mainly focuses on forecasting, prediction or susceptibility mapping, we combine prediction and classification, systematically review datasets and evaluation indicators and establish a model of prediction, method and performance taxonomy by carrying out a systematic literature review on 30 peer-reviewed articles from the period 2020 through 2025, extracted from primary academic databases with explicit search strings and inclusion criteria. The synthesis indicates that the existing ML based ML models, namely Random Forest (RF), Support Vector Machines (SVM) and Artificial Neural Networks (ANN), are functional in structured hydrological data sets. On the other hand, DL models such as Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN) have a strong performance in extracting spatiotemporal flood dynamics. A broader overview of some important research and methodological gaps — data availability in underrepresented geographic areas, uneven evaluation criteria, lack of Explainable Artificial Intelligence (XAI) implementation, limited interaction of socio-environmental variables — can be gleaned from the review, which highlights potential gaps in the existing literature. We build upon work in this field by mapping, systematically, datasets, methods, and performance metrics in the review. It provides pragmatic proposals for future research and policy-makers to improve flood risk modeling, disaster preparedness, and climate resilience.

## 1. Introduction

Environmental changes are significant events caused by natural processes that result in widespread destruction, often affecting both human and natural environments. Floods have significant impacts on community displacement, destruction of infrastructure, and long-term environmental damage [1], [2]. Floods have become more common globally by over 40% in the past two decades [3], [4], [5]. Floods happen when water levels rise above normal levels, flooding land areas. Flooding, caused by heavy rainfall, storm surge, or rapid snowmelt among others, has multiple forms such as flash flooding (fast floods) or moderate riverine inundations [5], [6]. Such phenomena severely impact communities across the globe and result in property damage, dislocation of population, deaths, and environmental problems [7]. Floods also contribute significantly to the structure of ecosystems and landscapes by depositing sediments and affecting the distribution of flora and fauna [8], [9].

Flooding is the most common natural disaster and affects all life forms including, but not limited to human beings as well as people engaged in agriculture, industry, and education [3], [10]. As noted by Khosravi et al [6], the percentage of human losses from natural disasters in Asia is approximately 90%, where floods are a major loss contributor. These problems indicate that preventing floods and minimizing the damages they cause are critical and great tactics to consider [11]-[13]. Between 2018 and 2022 worldwide, economic losses from floods are predicted to have totaled $299 billion U.S. dollars and insured losses represented around 15 per cent (around $44.5 billion) [14]. This reflects that 1.81 billion people — 23% of the world's population — are now directly touched by flood depths greater than 0.15 meters in a 1-in-100-year flood event, thereby putting them at high risk of losing their lives and livelihoods. Of these, 89% live in low and middle-income countries. Malaysia had the worst floods in several decades in 2021, killing more than 54 people and costing around RM6.1 billion to businesses around the world [15]-[17].

Recent reviews do not focus on recent advances in DL, nor specifically on their hybrid modeling for flood prediction, including classification of flood status or severity, which relies on the essential factors like meteorological, hydrological, or topographical data in addition to the land-use data [18], [19]. Major aspects cover precipitation, river, and statistical models that employ historical data to forecast future floods, along with hydraulic and hydrologic simulations that utilize the principles of physics [20], [21]. ML algorithms identify patterns in large datasets [22]. In addition, GIS is integrated with remote sensing for real-time monitoring and analysis [23]. Quantitatively predicting flood phenomena through statistical techniques is referred to as probabilistic modeling [24], [25]. Advantages of the probabilistic model, for example, include its ability to reflect the ambiguity and unpredictable nature of flood events, as well as its ability to offer a quantitative assessment of risk [26], [27]. In particular, its weaknesses concern reliance upon potentially faulty or missing data and sensitivity to analysis assumptions and techniques [28]. Utilizing satellite and aerial data, remote sensing, and Geographic Information System (GIS)-based modelling [29]-[31] is useful in data collection to map and assess flood events, which also allows researchers to aggregate various data sources and provide accurate information on the type of flooding as well as the impact of flooding [32], [33]. The downsides include the limited availability of qualitative data, which may be scarce or unavailable, and the computational complexity that makes it challenging to apply at large scale [34]. Artificial Intelligence (AI) constitutes the wider field; Machine Learning (ML) is a subfield; Deep Learning (DL) is a branch of machine learning which is utilized for hierarchical feature learning purposes [34], [35]. ML/DL are considered from this perspective in this review. ML looks at creating models and algorithms to analyze and predict from data [35]. ML seeks to make a model more powerful without training. It has three classes: supervised learning, unsupervised learning, and RL [36]. Supervised learning is a mechanism based on an algorithm trained on the labelled data to help generate predictions on new untrained data [37]. Unsupervised learning is an algorithm that is trained on a data set that is without labels but has been observed, patterned for identifying data patterns and relationships [38]. A number of applications, such as image recognition, natural language processing, recommendation systems, and predictive modelling, employ ML methods. The most commonly used ML approach in building flood status classification is Random Forest (RF) [39].

DL is a specific ML system inspired by the structure and operation of the human brain. It is an intelligent application that leverages multiple layers of connected neural networks to autonomously learn and create meaningful patterns, features, and representations from large and complex data sets [40], [41], [42]. This feature makes DL particularly suitable for big data analysis and pattern recognition problems. The best known application of DL includes traffic speed prediction, where a good model of traffic will accurately predict the traffic flow [43], [44]. Furthermore, DL methods have also been used in general applications such as transport systems, financial time-series analysis, and resource-booking platform usage [44]. Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) [45]-[47] and Deep Neural Networks (DNN) [48],[48] provide several specific popular DL algorithms that are widely utilized for flood prediction and status prediction and are being reported in recent years. Data-driven DL has various variants that cater for various applications.

The prediction and classification of floods have entered into the forefront of research due to their rising frequency and severity triggered by climate change, urbanization processes, and environmental degradation [49], [50]. The predictive models adopted for flood control are very important based on the ensemble of data including meteorological, hydrological, and geographical data [51], [52]. However, traditional statistical and hydrological models, while essential, often underperform at addressing the nonlinear and spatial differences in flood dynamics [53], [54]. ML/DL approaches have been identified as pioneering methods of analysis that can quickly adapt to huge and complex data sets while enhancing forecasting accuracy [55], [56]. A few reviews have been conducted on the different dimensions of AI-based flood modeling, however, mostly as a whole, most of them are focused on short-term forecasting whereas others are on susceptibility mapping or different DL architectures, and not the combination of the prediction and classification tasks, or systematically deal with datasets and criteria for evaluation. For the classification algorithms, the evaluation metrics are accuracy, precision, recall, and F1 Score, normally calculated from the confusion matrix. The evaluation metrics of the prediction algorithms, on the other hand, are RMSE, MAE, MSE, and $R^2$. In addition, many former reviews ignore regional inequalities, and methodological limitations (i.e., interpretability, uncertainty testing, and reproducibility).

This study fills these gaps through a systematic review on ML and DL approaches in flood prediction and classification between 2020 and 2025, which introduces a systematic classification that presents the datasets, the methods, and evaluation measures and identifies the regional/methodological limitations that hinder current understanding of the topic and we develop a summary that could provide guidance to researchers, disaster planning and managers, and policy-makers for more effective flood preparedness and mitigation.

The next section of this study is organised using the following categories: section 2 presents a literature review relating to this field, and section 3 explains methodologies employed in this research. Whereas, Section 4 introduces the new techniques like ML/DL approaches to the design of flood prediction and status classification models. The evaluation of the study dataset and eligibility criteria are presented in Section 5. This paper's analysis and discussion, which discusses ML/DL techniques in detail is provided in Section 6. The challenges and future directions of the study are addressed in Section 7. In Section 8, we summarise the salient issues facing us, and outline the research agenda for the coming years.

## 2. Related Reviews

Recent works of literature on review studies have explored the intersection between flood prediction, classification and new computational methods (e.g., ML and DL). Overall, the studies highlight the transformative power of data-driven processes, but also raise significant limitations, necessitating extensive system-wide investigation. For instance, Asif et al. [57] have shown a systematic review of ML applications for short-term flood forecasting. They reviewed 94 papers published from 2015 to 2022, classified forecasting models and discussed methodological strengths and weaknesses. While helpful, the review focused on short-term forecasting, not flood classification or evaluation metrics. Similarly, Anic et al. [58] described hybrid hydrological–ML models and displayed the efficiency achieved by including physical hydrology with data-driven methodology. Useful as a summary of the literature, however, the review was primarily narrative and did not provide a taxonomy or evaluation framework. In line with this trend towards susceptibility and mapping, Islam et al. [59] have summarized ML/DL models that are integrated with remote sensing for flood susceptibility mapping. They observed that a large adoption of convolutional and recurrent neural networks has been noted for their superior capacity for spatially-informed tasks. Instead, their scope was restricted to susceptibility studies; prediction and classification tasks were relatively unexplored. In a complementary manner, Byaruhanga et al. [60] identified Flood Early Warning Systems (FEWS) that highlight the increasing importance of data-driven notifications and communications methods. Their survey showed that there were inconsistent evaluation standards and minimal integration of ML in a manner that demanded developing overall technical approaches.

Other reviews have focused especially on DL. Liu et al. [61] performed a combined bibliometric and systematic review, demonstrating that CNNs, LSTMs and hybrid DL approaches remain the strongest contenders for real-time flood forecasting. However, data-set and comparative evaluation problems are still understudied. In analogous vein, Bentivoglio [62] supported the quick uptake of DL but offered scant coverage of datasets or metrics, discussing only algorithmic advance instead. In parallel, Munawar et al. [33] and Díez-Herrero & Garrote et al. [63] expanded on discussions of remote sensing and GIS enhanced flood risk assessment. Though both reviews validated geospatial tool importance, they both underestimated the role of predictive classification and evaluation frameworks. Finally, Ahmad et al. [64] provided a more holistic take on the phenomenon, which addressed AI-based environmental modelling in a variety of areas, and discussed floods as a subset. This diluted the nature of the insights which are directly relevant to flood prediction and classification.

These reviews indicate significant progress but also highlight three common gaps:

1. Little integration between prediction and classification. Reviews tend either to focus on forecasting or susceptibility mapping.
2. Poor handling of datasets and evaluation metrics, which are crucial for reproducibility and benchmarking, are often overlooked.
3. Fragmented scope, many reviews emphasize hydrological–ML hybrids whereas others concentrate on GIS, RS, or DL, but fail to establish a cohesive conceptual framework.

This work overshoots above and beyond this with a systematic literature review which integrates flood prediction and classification tasks while specifically noting the dataset acquisition, description, and evaluation criteria. Unlike previous reviews, this work not only describes studies by country and publication year, but also creates a structured taxonomy which pairs datasets, ML/DL techniques, and performance metrics. This comprehensive view, summarized in Table 1, allows the study to address the gaps, which contribute to a more general view of how up-coming computational approaches can contribute to advance in flood risk modelling and disaster preparedness.

**Table 1** *Comparative summary of existing review studies in flood prediction and classification*

| Ref. | Scope/Focus | Methods | Key Findings | Limitations | How this Study Differs |
|---|---|---|---|---|---|
| [57] | Short-term flood forecasting using ML | PRISMA, SWOT analysis | Classified ML forecasting approaches; identified gaps | Forecasting only; limited dataset analysis | Adds classification, datasets, and evaluation criteria |
| [58] | Hybrid hydrological–ML models | Narrative + case synthesis | Hybrid models improve accuracy through coupling | Limited generalizability; no structured taxonomy | Provides a dataset-driven taxonomy and evaluation framework |
| [59] | Flood susceptibility mapping (ML/DL + RS) | Systematic review | CNNs and RNNs are widely used with RS | Focus on susceptibility only | Broaden prediction and classification tasks |
| [60] | Flood Early Warning Systems (FEWS) | PRISMA-based review | FEWS is increasingly data-rich; it emphasizes communication | Inconsistent evaluation standards; weak ML integration | Adds AI-based classification; structured evaluation taxonomy |
| [61] | DL in flood forecasting/disaster | Bibliometric + systematic | LSTM, CNN, and hybrid DL models are dominant | Dataset aspects under-explored; DL only | Integrate ML/DL with datasets and evaluation metrics |
| [33] | Remote sensing–based flood modelling | Systematic review | RS is effective for flood extent mapping and monitoring | Focus on RS methods; less emphasis on predictive ML/DL | Provides balanced integration of datasets, metrics, and prediction |
| [63] | Flood risk assessment using AI + GIS | PRISMA + GIS synthesis | AI + GIS enhances spatial flood risk mapping | GIS-heavy; predictive classification not covered | Merges classification with datasets and evaluation frameworks |
| [62] | DL in flood modelling | Bibliometric + systematic | LSTM, CNN, and hybrid DL models are growing fast | Limited dataset focus; weak comparative evaluation | Introducing dataset categorization and evaluation taxonomy |
| [64] | AI-driven environmental modelling (incl. floods) | Mixed systematic approach | AI applied across domains; DL expansion | Flood treated as a subset; dataset/metrics secondary | Flood-centric review linking prediction, classification, datasets, and metrics |
| Our SLR | Flood prediction & classification (ML/DL) with dataset & evaluation focus | Systematic review + taxonomy | Synthesize prediction & classification; links datasets, metrics, ML/DL | -- | Provides a unified framework integrating prediction, classification, datasets, and metrics |

## 3. SLR Methodology

In this study, we investigate potential research gaps in flood prediction and classification while evaluating the state-of-the-art in ML/DL. In order to do this, a Systematic Literature Review (SLR) was performed using the PRISMA framework, which assisted in the retrieval, screening, and refinement of the most pertinent and latest studies in reputable journals. PRISMA had an organized process to guarantee transparency and rigour and initiated with an exhaustive search of the databases, followed by a screening/eligibility review process for duplicate and irrelevant literature, before limiting the search to articles of high quality. It allows us to critically assess state-of-the-art approaches for flood prediction and classification and to identify the shortcomings to which this work aims to contribute.

## 3.1 Academic Databases and Search Strategy

In this systematic review several databases were reviewed using ML/DL to find relevant studies used for flood prediction and classification. In these searches, Google Scholar was searched; GS-1 ("flood prediction" OR "flood forecasting" AND ("machine learning" OR "deep learning")), GS-2 ("flood classification" AND ("machine learning" OR "deep learning")). IEEE Xplore: ("flood prediction" OR "flood forecasting" OR "flood classification") AND ("machine learning" OR "deep learning"), ScienceDirect: ("flood forecasting" OR "flood classification") AND ("machine learning" OR "deep learning"), and Scopus: TITLE-ABS-KEY ("flood prediction" OR "flood forecasting") AND ("machine learning" OR "deep learning") AND PUBYEAR > 2019; and using customized search queries for each platform to concentrate on flood prediction models, algorithms and applications, use filters, such as date of publication (2020–2025), English language, and certain sort of document, research articles and conference papers, so as to ensure a sense of the relevancy and quality of the content. Significant research was identified through strict criteria (low-citation papers are excluded in Scopus). Searches were last conducted on 31 October 2025. Due that, this research selected the above database as the coverage database for studies in AI and engineering applications. Web of Science was omitted for its flood-AI coverage overlapped significantly with Scopus, and many relevant conference proceedings were better indexed in IEEE Xplore.

### 3.1.1 Arranging Studies by Country

The aim of this study was to explore the map of countries using MI or DL technologies to predict or classify flood models between 2020 and 2025. The geographical distribution for the selected papers on flood prediction models is shown in Figure 1. It illustrates the geographical concentration on flood prediction models, with one place in China taking the lead, with 11 papers and the next, with 4 papers from the USA. As a comparison, Taiwan and Iran each provided three papers, and Thailand received two. The steep cumulative percentage curve shows these regions make up a large part of the research contributions. The remaining locations (e.g. Portugal, Bangladesh, Indonesia, etc.) provide us only one paper and the research value is almost equal. This distribution illustrates that a portion of the world, mainly China, is still more than 60% of the world leading in the development of flood prediction researches. At the same time, many other regions are relatively underrepresented and may also demonstrate differences in worldwide literature or concerns of research resources.

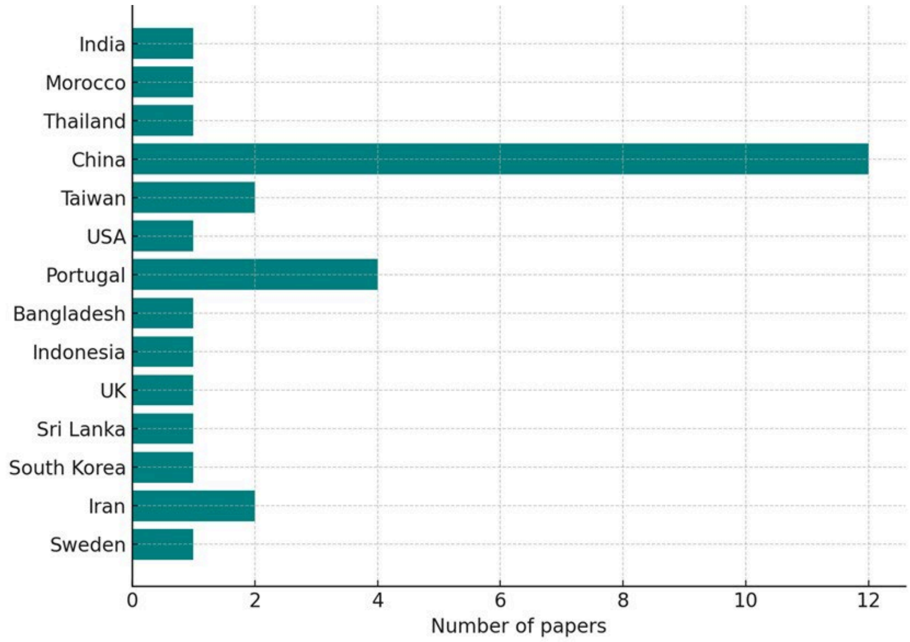


**Fig. 1** *Geographical distribution of selected papers on flood prediction models*

When looking at the composition by continent it can be observed that Asia has a significant representation of the table with a high percentage of the Table for countries in the Asia region and include China, Thailand, Indonesia, Sri Lanka and Bangladesh. That concentration reflects the region's importance in international events and reflects Asia's diversified array of countries. In addition to that, European countries like United Kingdom and Sweden are still representative on the world stage of why Europe is still significant in international affairs. And Australia's presence is a recognition of Oceania as a region that matters in global discussions. The table reflects the wide range of contributions and impacts across the regions — including Africa and South America — which are often overlooked by researchers in this area of studies, the table captures a glimpse of global dynamics.

### 3.1.2 Arranging Studies by Publishing Year

Figure 2 presents an overview in terms of the publication years and the amount of publications. It shows a spread of publications year over year in which there is an indication of any trend or an insight into the nature of the content in publication between years. In fact, the year of 2020 had the highest number of publications with 11 cases. A peak in publishing activity during that year, perhaps leading to a major event, a research breakthrough, or increased interest in the content. In addition, 2021 saw eight publications, further suggesting a period of notable activity in content creation.

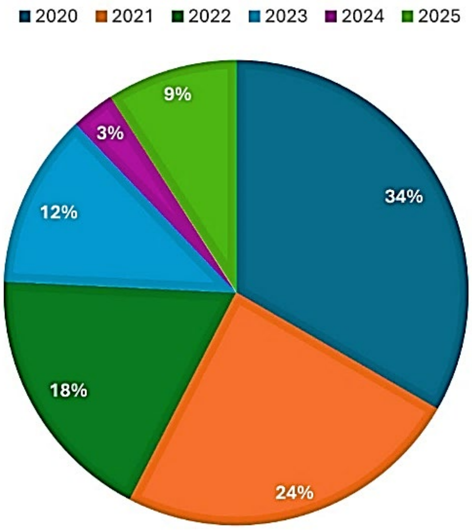


**Fig. 2** *Frequency of publications by year*

### 3.1.3 Arranging Studies by Publisher

Despite publishing in various media outlets, publications of the Journal of Hydrology (J1) among the institutions have surpassed others, accounting for nearly one-third of the studies, while the following ones are held by Water (MDPI) (J2) and IEEE Access (J3). Further publications which make significant contributions include Environmental Modelling & Software (J4) and Sustainable Cities & Society (J5); these contributions are distributed across other hydrology (J6), environmental science, and engineering journals (J7). This distribution reflects the disciplinary roots of the research in hydrological sciences as well as deepening interdisciplinary interaction with journals related to computational and urban systems.

## 3.2 Screening and Selection Process

The selection of studies followed the method and approach of PRISMA 2020. A total of 455 records were retrieved through database searches followed by the exclusion of 8 duplicates by Mendeley. The other 447 records were screened using Rayyan from the titles and abstracts of the review, and 396 studies were excluded from the review data that did not meet the inclusion criteria. A total of 13 records were excluded after abstract screening. After that, 38 full-text papers were evaluated for eligibility. Five of these were rejected on the basis of lack of methodological detail, little relevance to ML/DL-based flood prediction, or no empirical validation. In conclusion, 33 studies were selected, which satisfied each of the criteria and were included in the final synthesis. For this purpose, all included studies underwent a structured quality assessment along four dimensions: clarity of objectives, methodological transparency, data quality, and reproducibility.

## 3.3 Inclusion and Exclusion Criteria

The inclusive criteria are the required attributes of the paper for consideration in the review process. It includes document type, publication period, language used, and review scope.

- Document types: Peer-reviewed journal articles, high-quality conference papers, and technical reports.
- Publication Period: Studies published between 2020 and 2025
- Language: English only.

- Scope: Only articles applying ML/DL techniques in flood prediction or classification.

The exclusion criteria are the attributes of the paper that, if present, will be excluded from consideration in the review process. They are presented and described as follows:

- Duplicates identified during database searches.
- Papers not focused on ML/DL-based flood prediction, such as those emphasizing purely hydrological models without AI integration.
- Low-quality studies, e.g., insufficient methodological detail, non-peer-reviewed material.
- Non-English language documents to maintain accessibility and consistency in interpretation

## 3.4 Research Question

This paper includes specific research that pertains to the essential components of flood prediction and classification models and their role in effective flood management. The main research areas were the assessment of the efficiency of ML/DL models, choosing the best feature selection and model optimization approach, analyzing the effect of environmental, meteorological, and spatial variables on model accuracy, and exploring the implications for customized flood risk mitigation. These primary questions should be addressed in terms of improving predictive modelling approaches and reinforcing preparedness and resilience for the urban planners, researchers and policymakers in this review and might help push these key issues:

**RQ1:** How are ML and DL models designed and implemented for flood prediction and flood classification tasks?

**RQ2:** What types of data sources, feature representations, and hybrid or integrated modelling approaches have been shown to improve the performance of flood prediction and classification models?

**RQ3:** What major challenges, limitations, and research gaps exist in current flood prediction and classification studies, and what strategies have been proposed to overcome these issues and enhance real-world applicability?

## 3.5 SLR Process

The methodology that formed the structure of the systematic review was designed to achieve reliable, transparent and reproducible results. The steps to be taken to reply to research questions were reviewed by following the guidelines carefully. This chapter elucidates how the literature search, selection, analysis and quality evaluation were adopted.

The PRISMA process (as represented in figure 3) provides a systematic framework to promote transparency and rigor in systematic reviews by describing how studies in the database are retrieved, screened, and selected for inclusion. The whole process starts with the full identification of records from databases and other sources, and the omission of duplicate results. The rest of the records are then screened by title and abstract to remove studies that are not relevant. Full-text articles of potentially eligible studies are analyzed by predefined criteria, and exclusion reasons are recorded at each stage. The final selection of studies that satisfy all eligibility criteria is therefore considered as part of the review, giving a reproducible process by which the first batch of studies can be narrowed to the final selection.

This systematic study adhered to PRISMA recommendations, starting with identification of 455 articles, of which 440 were retrieved from electronic databases and 15 from manual searches. After removing 8 duplicates, 447 unique records were retained for screening. Title screening reduced the pool to 95 studies; subsequent abstract screening excluded 44 articles, leaving 51 eligible for full-text review. Of these, 38 articles were assessed in detail for eligibility, and 13 were excluded because they employed statistical approaches that did not meet the review criteria. Five articles were also excluded for lack of access to full text. Ultimately, 33 studies were included in the final review.

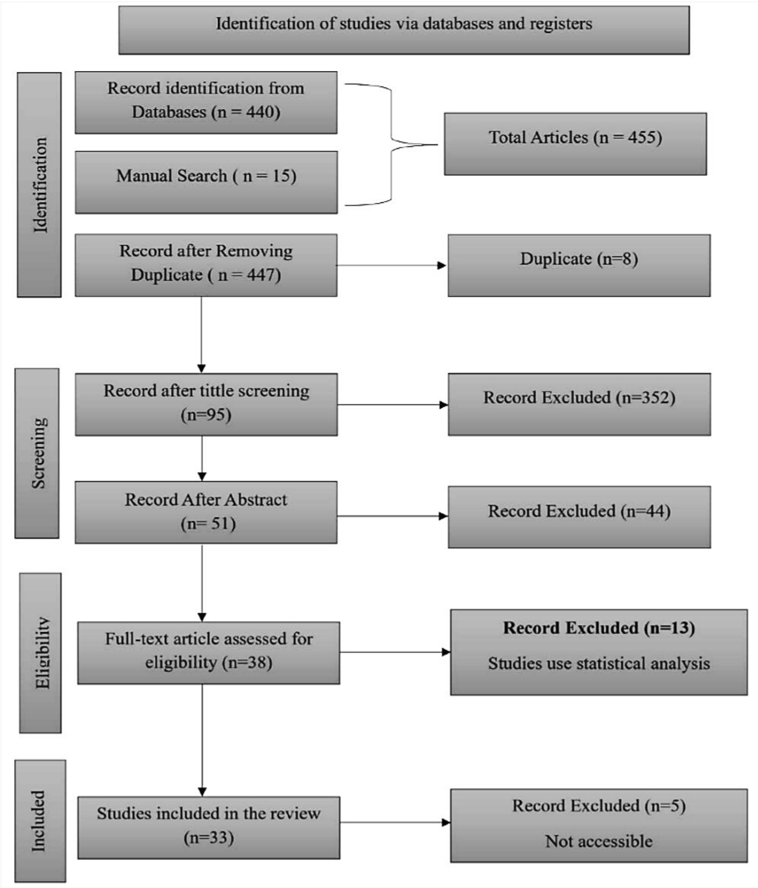


**Fig. 3** *A visual breakdown of publication selection based on PRISMA guidelines*

## 4. Emerging Techniques in Flood Prediction and Classification

New ML techniques have increased the effectiveness of flood prediction and status classification substantially. They are advanced methods using vast, varied datasets: past flood history, meteorological patterns, hydrological parameters, and geospatial information. Some hybrid ML models using physics-based simulations with a data-driven approach have been developed by researchers to improve flood forecasting with ML and classical hydrological models. DL architectures such as DNNs and RNNs have shown to be well adapted for the spatial and temporal modeling of dependencies that relate to the occurrence and severity of floods. Moreover, feature engineering is critical to adapt these models to incorporate inputs, including the trends of precipitation, soil moisture levels, land-use patterns, and topographical features for improved precision of the models. As ML becomes more advanced, the combination of ML with these real-time data sources and remote sensing tools will enhance the resilience of flood early warning systems to fluctuating environmental conditions.

### 4.1 Flood Prediction and Classification

Recent advancements in uncertainty quantification have led to more reliable and comprehensive probabilistic forecasts. Real-time monitoring for flood risk prediction using sensor technologies and remote sensing approaches can aid observing conditions and issuing flood warnings in a timely manner. Data collection, model validation, and improved prediction accuracy are also supported by citizen science and community participation efforts. Since climate change poses an ongoing burden to precipitation and weather extremes in the future, the need for real-time dynamic monitoring (i.e., rapid adaptation to changes in meteorological activities) and regular updates of predictive models with scenario analysis as an integral part will be paramount; adaptive flood prediction systems should be developed accordingly and continuously monitored and adjusted. Such initiatives cannot take place without the contributions of many disciplines and stakeholders to further the field and to build flood-resilient communities.

#### 4.1.1 ML in Flood Prediction and Classification

ML's contribution to flood prediction was studied by recent research: integrating meteorological data, hydrological data, and geospatial data. In these studies, ML models for such predictions have been found to be very good. In

Thailand, big and crowdsourced data could be included in this pipeline in an ML model proposed by S. Puttinaovarat [65] where methods such as DT, RF, Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Fuzzy Logic (FL) were used.Their ANN-based model was 97.83% accurate. But crowdsourced data have their issues too: data quality, consistency, and bias. Likewise, Q. Ke et al. [66] proposed an ensemble-based ML framework that quantifies rainfall thresholds for flood classification in China's Shenzhen. They used 14 classification algorithms with an overall accuracy of 96.5% but the false alarm rate was still 25%.

#### 4.1.2 Probabilistic and Data-Driven Approaches

A study conducted by Nguyen [67] Introduced a probabilistic flood forecasting model in Taiwan that integrates support vector regression (SVR), fuzzy inference models (FIM), and k-nearest neighbor (k-NN) algorithms. Using rainfall and river stage data from 2012 to 2022, their model attained high accuracy with an MSE of 0.07 and a correlation efficiency (CE) of 0.99. However, Zahura [68] Used RF as a surrogate model to predict floods in Norfolk, Virginia using physics-based methods. They illustrated that RF reduced computational time significantly but retained excellent predictive skill with RMSE values as low as 0.061. Likewise, [69] Implemented a neural network model (Elman Neural Network - ENN) + numerical flood simulation model (PCSWMM) for fast flood prediction in Tianjin, China. Their model showed high simulation accuracy such as RMSE of 2.49 and NSE of 0.88.

#### 4.1.3 Geospatial and GIS-Based Approaches

A work by [39] integrated ML and GIS methods to create a flood risk index for Lisbon, Portugal. In their RF model, high-risk flood regions were detected by analyzing historical flood data and meteorological conditions with 96% accuracy. The purpose of this study was to emphasize the role of geospatial data in flood risk assessment. Concurrently, [70] proposed the Flood Domain Adaptation Network (FloodDAN), an unsupervised approach which transfers knowledge across datasets to inform flood prediction. With very small labeled data, their model performed in a similar way to supervised models, highlighting the power of domain adaptation in flood prediction.

#### 4.1.4 Hybrid and Integrated Modelling Approaches

A combined ML–hydrodynamic modelling for compound flood prediction in Indonesia was proposed by [71]. 11 of 17 flooding occurrences with RF model were also correctly predicted and had NSE values greater than 0.8. In another investigation [72], DT, decision jungle (DJ) and RF algorithms were used to construct IoT based flood prediction models. Their model on Anglian River Basin Dataset was able to reach an accuracy of 85.72%, with DT being best with 93.22%. This study proposes the growing prevalence of hybrid solutions involving a combination of data sources and ML methods to enhance the predictive accuracy of floods.

#### 4.1.5 Enhancing Model Accuracy and Reliability

[70] developed the XAJ-MCQRNN model, which can apply conceptual hydrological models and ML-based regression techniques to multi-step-ahead flood prediction in China. The method obtained better coverage ratios and lower errors over classical regression models. In [73], we reported a model called cascaded ANFIS which combines fuzzy logic and rainfall-runoff modeling. Their model gave an RMSE of 0.66 and a Kling-Gupta efficiency (KGE) of 0.90, outperforming the RNNs and LSTM models. The need for model optimization and feature engineering to enhance forecasting reliability is highlighted in the literature and is the focus of these studies.

#### 4.1.6 Socioeconomic and Vulnerability-based Approaches

[12] examined flood susceptibility in China using tree-based ML models and probability certainty factor (PCF) based approaches. Their RF model achieved the best classification accuracy (94.3%) and showed the importance of integrating prior flood data and the environmental characteristics. [74] brings in socio-environmental aspect, such as socioeconomic susceptibility to flooding as part of its prediction by ANN model for flooding. They measured communities believed to be susceptible to flooding but had mixed performance for urban areas ($R^2$ = 0.231). These studies highlight the importance of inclusion of socioeconomic variables and environmental factors along with conventional ML-based flood forecasting methods.

#### 4.1.7 Effectiveness of ML in Flood Prediction and Classification

ML models significantly enhance flood prediction accuracy by assessing large-scale hydrological, meteorological, and geospatial data. Model types such as [65] and [66] leverage ML algorithms (DT, RF, SVM, ANN) to aggregate these input data from multiple datasets (e.g., crowdsourced dataset, real-time meteorological dataset) to estimate robust flood forecasts. ML and GIS based models [39], applied for the development of a flooding risk index, achieved a high degree of accuracy (96%) with respect to regions with vulnerability. [67] and [68] demonstrated that ML can make effective use of models in improving urban flood prediction methods based on probabilistic forecasting and surrogate modeling approaches, thereby reducing computational load without sacrificing prediction reliability. The promising characteristics of ML in varying environmental conditions (e.g., urban flooding [84] and rural flood risk assessment [87]) is evident. Unsupervised and ensemble-based learning models such as FloodDAN by [82] also tackle data constraints in flood prediction with domain adaptability. These trends also suggest increasing efficiency in flood prediction made possible by ML and early warning systems to support disaster mitigation.

### 4.1.8 Challenges to Consider in ML-Based Flood Prediction and Classification

Despite the powerful capabilities of ML models to predict flood, several challenges limit the scalability of the models. The first is quality and reliability in data, especially for crowdsourcing datasets such as [65], where poor predictions can occur due to inconsistencies, biases, and missing values. Also, a few models such as those of [66] and [70] were developed. They are limited to small to medium-sized data that may limit their ability to generalize well across many geographic areas. Another factor is the computationally costly nature of advanced ML models, including ensemble and hybrid models such as those described in [39]. According to research conducted by [72], real-time decision-making processing must consume high-power. In the same way, ML models frequently show interpretability shortcomings and it's difficult for policymakers and stakeholders to trust and act on predictions.

While [67] achieved high accuracy, the underlying probabilistic models being used make the model interpretable to have a lack of confidence that validity of the predictions will occur. Some ML model-based flood approaches were also in need of adaptation to mitigate their performance, while others include climate change or various land-use changes or infrastructure developments [68]. The problems to be tackled involve the incorporation of physics-based models, better techniques for data preprocessing and more open ML models to build model dependability and applicability. Descriptive results of flood prediction studies can be found in Table 2.

Summary of ML algorithms for flood prediction are presented in Table 2. This study considers the trends, evaluate the performance of each of their algorithms, and suggest possible future directions in research and application in this field. Increasing trends in algorithm selection for flood prediction reflect a general trend for advanced, traditional and hybrid approaches. Recent methods, such as TCN and XAJ-MCQRNN, have attracted more attention for solving the complex problems related to the predicting flood in a riverine environment with poor availability information, resulting in the performance enhancement of prediction [70]. At the same time, established techniques such as Support Vector Regression (SVR), RF, and ANN, for which reliable, versatile, interpretable methods, and have shown high capability across different flood types are heavily used [68], [75]. Hybrid approaches have found the momentum and are also catching on. This trend emphasizes the necessity of continuing to refine the methods of algorithm formulation so as to make an intelligent and effective approach for flood prediction.

**Table 2** *Summary of ML studies in flood prediction*

| Ref. | Problem domain | Algorithm | Evaluation | Limitation |
|---|---|---|---|---|
| [67] | Flash flood prediction with lead times of 1-3 hours in the Yilan River, Taiwan | SVR | MSE 0.07, CE 0.99 | The study provides strong results but lacks a detailed discussion of limitations, particularly model generalization and data variety. |
| [68] | Urban flood predictions for hourly water depths in the coastal city of Norfolk, Virginia, USA | RF | RMSE 0.2, MSE 0.06 | The study achieves reasonable accuracy but does not address potential limitations in data quality and model scalability. |
| [69] | Flood prediction of the depth of water accumulation in Tianjin Municipality, China. | ENN | MSE 0.096 | The high error suggests challenges in model accuracy, potentially due to insufficient data or inadequate model training. |
| [70] | The flood prediction model addresses limited historical data in Huangshan City, Anhui Province, China. | TCN | RMSE 7705, $R^2$ 45.58 | The exceptionally high RMSE and low $R^2$ indicate significant issues with model performance, possibly due to limited historical data. |
| [70] | Flood prediction based on a 3-hour streamflow series of the Jianxi River catchment in China. | XAJ-MCQRNN | RMSE 4.7, LNSE 9.0, $R^2$ 0.99. | Despite a high $R^2$, the unusual LNSE value suggests potential issues with data handling or modeling that require further exploration. |
| [73] | Flood prediction by modeling RR Variables. Sri Lanka | FL | RMSE 0.66, NSE 0.87 | The study achieves good accuracy; however, fuzzy logic may limit interpretability and scalability to other regions or datasets. |

| Ref. | Problem domain | Algorithm | Evaluation | Limitation |
|---|---|---|---|---|
| [74] | To improve the prediction accuracy and identify communities that are most affected by flooding | ANN | RSME 33.94, $R^2$ 0.231 | The high RMSE and low $R^2$ indicate significant accuracy issues, possibly due to complex flood dynamics or inadequate feature selection. |
| [76] | To address the challenge of data scarcity in developing a flood prediction model | RF-GB | $R^2$ 0.98, | The study relies on satellite data that can be obscured by cloud cover or dense vegetation, potentially underestimating small or fragmented water bodies due to the data's low resolution. |

The performance of two methods for flood prediction can be affected by the particular type of flooding considered in the present analysis. SVR and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) have achieved better accuracy and produced precise short-term forecasts of flash floods accurately by being able to predict flash floods effectively by accounting not only for the fast and chaotic dynamics of flash flood events [75]. In the area of urban flood prediction, RF and ANN models are valuable for simulating the complex relationship between water flow in an urban environment. They allow accurate prediction of water depth and support flood risk assessment and mitigation strategies [68]. In turn, state-of-the-art DL methods, such as Temporal Convolutional Networks (TCN) and extreme Attention-based Joint Multichannel Convolutional and Recurrent Neural Networks (XAJ-MCQRNN), are developing as robust tools in riverine flood prediction.

The innovative methodologies have been very well adapted to capture the complexities of river flow dynamics, particularly in data-poor settings, and are well suited to scenarios with few observable records [70]. These methods can thus be seen as mutually enhancing for flood predictions across different contexts, thus reinforcing the need to choose the best algorithm for each type of flood. Flood prediction models need sufficient data availability and quality to perform properly. A key challenge is data scarcity, and for DL models, which require many datasets to reach a successful solution. This restriction is stressed by [70] work as it relates to hydrological research as historical data is not available to fully predict. Also, quality of the information must be ensured, as poor values and incomplete data can result in biased predictions. Strong methods to clean and impute data are needed if the accuracy and precision of predictive models is to be preserved. In addition, data heterogeneity raises an additional layer of complexity, as flood prediction uses multiple data sources, including rainfall measurements, river gauge monitoring and satellite images.

Data integration techniques should be employed for harmonising these various inputs and improving reliability of the model. Flood prediction is a rapidly emerging area of research in which certain DL models (Temporal Convolutional Networks (TCN), XAJ-MCQRNN, etc.) are applied during riverine flood and for this data-poor environment. Common-place techniques like SVR, RF, ANN are still widely adopted across flood scenarios. Hybrid methodologies integrating these approaches also appear to be promising when accuracy is necessary. Yet, issues concerning data quality, lack of data and data heterogeneity remain as a challenge, indicating that the necessity for a coherent approach in managing data must be recognized. Overcoming these obstacles and utilizing multiple methodologies can greatly boost disaster preparedness and mitigation initiatives. Labeled data, on the other hand, is needed to classify flood status. These models adopt supervised learning as a means of training, where they divide the data into two independent sets (i.e. train and test) and use them at the time of training/test. Figure 4 allows the classification results using labeled data and unlabeled data, so such a semi-supervised learning approach can be generated.

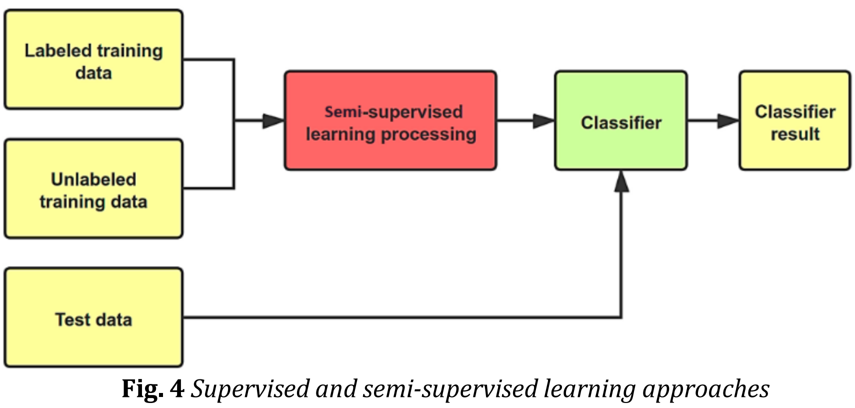


**Fig. 4** *Supervised and semi-supervised learning approaches*

Flood status classification is essential for disaster management, urban planning, and environmental monitoring. Various studies have explored different algorithms for flood status classification, yielding diverse accuracy results. This analysis examines the methodologies and outcomes of several research efforts in this domain, highlighting their strengths and limitations. Table 3 below presents the ML methods applied for flood status classification and the number of corresponding articles in the literature over the last half-decade. These tables were designed to tell readers which ML methods have become more popular among hydrologists.

**Table 3** *Summary of ML studies in flood classification*

| Ref. | Problem domain | Algorithm | Accuracy | Limitations |
|---|---|---|---|---|
| [39] | Classify the flood status as flood or no flood in Lisbon City, Portugal. | RF | 96.00% | The lower accuracy of GIS+RF suggests challenges in integrating geospatial data with ML models. |
| | | GIS and RF | 84.00% | |
| [77] | Flood status classification for Bangladesh's Sylhet and Sunamganj districts is crucial for environmental monitoring. It determines if these areas are currently experiencing flooding, classifying them as "yes" or "no" based on established criteria. | NB | 53.33% | The relatively low accuracy suggests that the chosen features or model configurations may not be optimal for the dataset. |
| | | LR | 56.67% | |
| | | SVM, | 60.00% | |
| | | ANN, | 48.89% | |
| | | GB | 60.00% | |
| | | DT | 53.33% | |
| | | KNN | 63.33% | |
| [71] | Flood status classification in the Kapuas River delta, Indonesia. | RF | 65.00% | The low overall accuracy indicates potential challenges with data quality or the complexity of flood status in the region. |
| | | MLR | 35.00% | |
| | | SVM | 59.00% | |
| [72] | Classification of 13 multivariate Flood statuses in parts of some cities in London. | DT | 93.22% | The variance in accuracy across algorithms suggests differing capabilities in handling the dataset's complexity. |
| | | RF | 86.57% | |
| | | DJ | 77.38 | |
| [65] | A systematic classification of flood status into three distinct categories helps to follow up and respond effectively. Different levels are yellow (which indicates caution), orange (highlighting an elevated risk level), and red (indicating critical and severe flooding conditions) for storm surge in southern Thailand, indicating it needs to be managed immediately. | DT | 95.33% | The study's high accuracy indicates robust models, but potential overfitting and the need for validation on more diverse datasets are not discussed. |
| | | RF | 96.67% | |
| | | SVM | 96.67% | |
| | | ANN (MLP) | 97.83% | |
| | | ANN (RBF) | 95.50% | |
| | | NB | 88.00% | |
| | | FL | 95.67% | |
| [66] | Flood status classification of flooded or non-flooded areas, the aim is for urban planning, emergency response, and resource allocation in these areas. This allows Shenzhen, China authorities to prioritize infrastructure, warnings, and rescue operations and thus achieve sustainable growth and improved safety. | DT | 95.60% | Very high accuracy suggests effective models, but the study may benefit from discussing the impact of data variety and real-world applicability. |
| | | DA | 95.90% | |
| | | SVM | 96.40% | |
| | | KNN | 95.90% | |
| | | Ensemble | 96.40% | |
| [12] | Multivariate flood status classification model prediction by classification into five major classes: very low, low, moderate, high, and very high in Quannan region of China. | RF | 91.50% | RF performed strongly in these data metrics, but the study did not address the limitations of the algorithm under different scenarios, or variations in data quality performance. |
| [78] | Urban flooding risk being increased due to climate change and rapid urbanization. | 2D-CNN–CapsNet–WOA | 89.42% 94.81% | The study is limited to historical flood data — no future climate-driven patterns are predicted — and depended on data quality and resolution. |
| [76] | The study aims to improve flood prediction in the Oum Er Rbia watershed, Morocco. | RF-GB | 98.00% | The study is based on satellite data that may be obscured by cloud cover or dense vegetation, the low resolution of the imagery could leave small and/or fragmented water bodies. |

Using a variety of algorithms, these studies report different methods used to classify flood status, achieving mixed accuracy [39]. A flood status report in Lisbon, Portugal reported high classification accuracy of 96% with RF. In contrast, an accuracy of 84% was achieved by GIS combined with RF, indicating difficulties in the coupling between the geospatial data and ML. Once again, [77] tested different algorithms in Sylhet and Sunamganj, Bangladesh, where K-Nearest Neighbours (KNN) attained the best performance at 63.33% accuracy and ANN the worst at 48.89% accuracy. With their relatively low accuracies, the results confirm that some aspects were disregarded and consequently the model is not yet well calibrated.

It analyzed flood status classification in Kapuas River delta, Indonesia, and the best accuracy (65%) was obtained from RF and lowest (35%) was found for MLR accuracy, which could be attributed mainly to the poor data quality or regional flood activity. In London [72] to mention a very high accuracy of DT (93.22%), RF (86.57%), and Decision Jungle (DJ) (77.38%), respectively, showing that the choice of algorithm can have an effect on the performance of different data source (complex data).

In southern Thailand, [65] obtained ANN (MLP) accuracy (97.83%), which causes over-fitting problem particularly when large datasets are not validated. In Shenzhen, China, [66] and Quannan, China, [12] performed well (e.g., 96.4% Linear SVM, 91.5% RF), but they did not give much useful views on data diversifications and actual applied values.

Despite the positive impact of data performance demonstrated, the drawbacks in these analyses are mainly down to data quality, integration difficulty and model robustness. A union of RF and GIS is provided in [39]. It is a testament of challenges of integrating geospatial data with ML techniques. We have obtained relatively low accuracy from [77]. [71] stress problems like feature picking, model settings or complexity of regions. The high accuracies [77] obtained in the latter indicate that the models are not only stable but also may be robust if different datasets are used in the validation of the models to prevent overfitting. Similarly, [78] and [27] stress to deal with diverse data and apply it beyond the lab.

These limitations can be partially alleviate with some enhancements. Diversified source : more data collection would ensure all our dataset(s) (real time and historical) is enriched and diverse enough for the best training of robust model. Strong validation techniques (e.g., “cross-validation”) can be used to ensure the successful deployment and adaptation of models in the new data. Fine-tuning hyperparameters, as well as looking at ensemble method, could also improve model performance. Comparative studies – across regions and with available data – can even reveal the best algorithms. If one can make models that succeed in real-world cases the real-world applicability of the model can be verified so that the model may be fitted and upgraded based on the feedback from users in real-life to boost the reliability of flood systems.

## 4.2 DL in Flood Prediction and Classification

DL is an architecture-based transformative technique by employing multiple layers to provide high-level features in raw data (Artificial Neural Networks [ANN]). Various types of it cover image recognition, speech recognition, natural language processing, etc., but it is more and more useful for analysis of data over time, as flood prediction to name only one case. Novel techniques include DNNs, RNNs, and LSTMs, using these neural architectures to solve the challenges in flood forecasting. DNNs handle spatial information collected from multiple sources and find significant parameters, e.g., land-cover change, elevations, spatial information obtained from satellite images and LiDAR. RNNs and LSTMs, however, excel at learning temporal dependencies, which is critically important for studies on the patterns of rainfall and river dynamics. The hybrid models of DL and classic hydrological models contribute to more accurate and reliable predictions with good accuracy and reliability; the Bayesian approaches and ensemble methods produced robust estimates that are sufficient with appropriate confidence intervals.

DL models are required in operational scenarios of real-time monitoring and early warning system, by detecting multiple forms of data (e.g weather forecasts, river gauge readings etc.) suitable for timely warning in time. They enhanced disaster preparedness and response by enabling communities to build resilience and minimize the consequences of flooding. Adaptive learned mechanism enables the improvement in performance of model for improved performance by providing constant feed-back into the modified model to make the model to adjust to environmental changes, so that it is more sustainable under climate change and changes in land use of activities. These advances highlight how DL could revolution the prediction of floods in a more accurate, reliable way and the speed with which disaster response plans are designed from inside and greater resilience is engrafted upon communities in a timely manner.

### 4.2.1 DL for Streamflow and Flood Prediction and Classification

DL technologies also extensively developed and thus substantially improved the reliability and accuracy of flood prediction models. As an introduction, [46] introduced a LSTM model and a model as the Encoder-Decoder LSTM (En-De-LSTM) model was applied as a model with EMD for Yangtze River streamflow prediction. Based on monthly streamflow data from 1952 to 2008, their model performed 30% less RMSE and 20% better in Legates-McCabe's Index (LMI) than alternative LSTM models. In addition, [79] compared and outperformed ANN, DNN, and LSTM models in the flood prediction and anomaly detection at Danshui River of Taiwan using a Conv-GRU neural network approach. In extreme events, such as Typhoon Soudelor in 2015, their model efficiently identified the abnormal water levels.

### 4.2.2 Interpretable and Hybrid DL Models

The demand for interpretable DL models has led to innovative approaches in flood forecasting. The STA-LSTM model was effectively employed to predict floods in three Chinese basins via an attention mechanism combined with LSTM [45]. Their model was a great improvement over other classic techniques such as DNNs and GCNs (Graph Convolutional Networks) with an RMSE of 97.03 and $R^2$ of 0.96. In a parallel way, [80] employed stochastic regularization for uncertainty estimation in flood prediction. Their LSTM-based model, trained on gauge height values for the Meramec River in Missouri, performed significantly better with an RMSE of 0.8722 and predictions were made every 15 minutes, which is an improvement over the 6-hour update rates provided by typical models.

### 4.2.3 Urban Flood Prediction and Hydrological Modelling

Urban flood forecasting has been enhanced by combining hydrological models with DL techniques. [75] built an LSTM model predicting urban flooding in Seoul based on three-hour Mean Areal Precipitation (MAP) predictions. Their model, tested across 24 heavy rainfall events, outperformed traditional logistic regression models. Likewise, [81] proposed a gradient-boosted Decision Tree (GBDT) model to predict urban flood depth. They took 16 rainfall stations and GIS-based surface factors and combined them into their study, with 88.48% accuracy and an $R^2$ of 0.96.

### 4.2.4 DNN and RNN Models for Spatial Flood Prediction and Classification

Integrating DNN and RNN has greatly improved spatial flood prediction. [48] studied DNN versus RNN models for flood hazard modeling in Seoul, South Korea. Their results showed a minor benefit of DNN, with an AUC of 84% and RMSE of 0.163. Similarly, [55] applied DNN and RNN models to predict flash floods in Iran's Golestan Province, where DNN outperformed RNN with an AUC of 0.832 and RMSE of 0.144 [82]. Again, the effectiveness of DNN in flood risk mapping for Iran was shown in this research with an RMSE of 0.094 and an $R^2$ of 0.098.

### 4.2.5 Hybrid Models for Flood Prediction and Classification

The fusion of different DL models has led to significant improvements in flood prediction. [47] combined RF, MLP, DNN along with LSTM in a hybrid FLUD model for predicting water level in the Sihu basin in China. This provided an effective accuracy, with an $R^2$ value of 0.9965 for predicting farmland waterlogging. Similarly, [83] developed a Conv-LSTM model for flood forecasting, using both DNN and LSTM for Xi County, China, which had an RMSE of 0.0200 and an $R^2$ of 0.9750. [84] further explored ensemble learning models that combine IoT data and weather forecasts to predict flood risk in Sweden, which achieved an MSE of 1.25.

### 4.2.6 Next-Generation Flood Prediction and Classification Models

Recent technologies have presented hybrid and attention modelologies for better accuracy of flood forecasting [85]. AGRS-LSTM-Transformer, a model proposed integrating LSTM, Transformer, and Adaptive Random Search Algorithm for flood forecasting in Jingle watershed, China. The high performance of their model with NSE and RMSE of 0.905 and 34.891 $m^3/s$ respectively, indicates significant power of hybrid models in forecasting in helping to minimize uncertainty and augment the early warning of these phenomena. Similarly, [86] proposed the U-Net River model that predicts flood depth and extent, and showed a 29% enhancement over standard hydraulic models. These researches show ongoing evolution of flood prediction technique incorporating DL, spatial analysis, and real-time processing as well.

### 4.2.7 Effectiveness of DL in Flood Prediction and Classification

By exploiting vast amounts of data, intricate hydrological structures and instantaneous meteorological data, DL models enable high accuracy in flood prediction. As mentioned in studies [46] and [45], for long-term streamflow forecasting, LSTM based models are better, resulting in RMSE reductions of up to 30% and improved accuracy in extreme flooding prediction. Attention mechanisms have been utilized to add to the temporal and spatial dependencies in flood forecasting, as exemplified in the STA-LSTM and Conv-LSTM models. Furthermore, urban flood models like those in [75] and [81] contribute rainfall, land use and real-time hydrological information to improve prediction performance. The DNN and hybrid models are proposed in [47] and [85]. DL models can also be applied to integrate spatial and temporal flood data towards a more accurate, multi-step forecasting. Such capacity to process high resolution data, automate feature extraction, and perform flood risk assessments in a timely manner by the means of beneficial tools used for disaster preparedness and water resource management.

### 4.2.8 Challenges to Consider in DL-based Flood Prediction and Classification

The challenges of forecasting flood models constructed using DL-based predictive technologies are clear such that even though flood prediction models are efficient, many challenges remain such as data quality, computational complexity, and interpretability. The dependence on a plethora of historical data is an issue as demonstrated in [46] and [82]. This can pose problem in areas with limited hydrological records. Furthermore, the performance of models like [84] depends on the accuracy of weather forecasts, which could create uncertainty beyond 72 hours. Models such as AGRS-LSTM-Transformer remain difficult because they are computationally expensive [85]. Training and real-time applications require high-performance computing. Moreover, black-box models like DNNs and LSTMs that offer high accuracy but poor decision transparency are also an issue for policymakers, as detailed in Table 4, as DL are notoriously difficult to interpret. There is also a limitation in models that will not generalize in other geographical regions as [80] pointed out where external effects (such as climate change and urban development) are not well integrated. To overcome these challenges, we have to combine physics-based models, explainable AI (XAI), and better data assimilation to improve model reliability and generalizability.

**Table 4** *Summary of DL studies in flood prediction and classification*

| Ref. | Problem domain | Algorithm | Evaluation | Limitation |
|---|---|---|---|---|
| [46] | Flood prediction by observing Monthly Streamflow. | En-De LSTM | RMSE 1171 | The high RMSE indicates significant prediction errors, possibly due to the complexities of temporal data. |
| [79] | Predict the water level flood phenomenon of a river in Taiwan | Conv-GRU | RMSE 0.77, MAPE 30.68, MAE 0.56 | A high MAPE indicates issues with percentage errors, suggesting potential data variability. |
| [45] | Improve the interpretability and accuracy of the Flood prediction model | LSTM | RMSE 97.03, $R^2$ 0.96, MAE 37.49, | High RMSE and MAE highlight substantial errors despite good $R^2$, indicating possible overfitting. |
| [80] | Flood prediction involves predicting gauge height and evaluating it. | LSTM | RMSE 0.8722, MAE 0.2595 | The study achieves good accuracy but lacks discussion on data variety and robustness. |
| [75] | Urban Flood prediction by generating a three-hour MAP in Seoul, South Korea. | LSTM | RSME 0.1, $R^2$ 0.95 | High accuracy, but requires validation across various urban settings. |
| [81] | Predicting four multivariate depths of water accumulation for urban flood in Zhengzhou city, China. | GBDT | R2 88.48%, $R^2$ 0.96 | High accuracy suggests strong model performance, but it requires a diverse validation dataset. |
| [48] | The study, which predicted flood hazard modeling in urban areas, was conducted in Seoul, South Korea. | RNN | RMSE 0.186, $R^2$ 82% | Both models perform well but need further exploration of their generalizability to other urban environments. |
| | | DNN | RMSE 0.163, $R^2$ 84% | |
| [55] | Predicting Flash Flood Probability in Iran | DNN | RMSE 0.144, $R^2$ 0.832 | High performance but necessitates validation with more diverse climatic data. |
| | | RNN | RMSE 0.181, $R^2$ 0.814 | |
| [47] | A flood prediction model, such as the FLUD model, is needed to predict street-level floods accurately. | RF | RMSE 0.01, $R^2$ 0.95 | Variation in accuracy among models suggests the need for hybrid approaches or feature enhancement. |
| | | MLP | RMSE 0.02, $R^2$ 0.93 | |
| | | DNN | RMSE 0.06, $R^2$ 0.78 | |
| | | LSTM | RMSE 0.04, $R^2$ 0.88 | |

| Ref. | Problem domain | Algorithm | Evaluation | Limitation |
|---|---|---|---|---|
| [84] | Predicting future flooding based on IoT data and weather forecasts in Kristianstad, Sweden | LSTM | MSE 1.25 | The study needs extensive validation to ensure robustness against different environmental data inputs. |
| [87] | A flood prediction model based on DNN that utilizes DL for complex feature extraction with two-dimensional | DNN | RMSE 6925.89, $R^2$ 0.8938 | Extremely high RMSE indicates significant prediction errors, underscoring the need for improved model tuning and feature selection. |
| [86] | Flood prediction model for predicting river flood depth and extent | CNN | MSE 0.00096, MAE 0.0077 | Excellent performance, but needs broader validation to confirm model generalization. |
| [88] | To predict floods by developing long lead-time streamflow prediction at the scales of various lead times. | LSTM | $R^2$ 0.86 | Both models exhibit substantial prediction errors, underscoring the need for more refined modeling approaches. |
| | | ANN | $R^2$ 0.76 | |
| [83] | Flood prediction model in the form of temporal features. | Conv-LSTM | RMSE 0.0200, R2 0.9750 | High accuracy but requires validation with varied temporal datasets to confirm robustness. |
| [82] | flood prediction area | DNN | MSE 0.008, RMSE 0.094, MAE 0.033, $R^2$ 0.098. | The study presents strong performance metrics, but it lacks a discussion of the model's generalization across different flood scenarios. |
| [85] | Flood prediction to accurately predict flood events | LSTM | NSE 0.905 and RMSE of 34.891 | The model's reliance on historical data from a single watershed limits its generalizability to regions with different hydrological conditions or data availability. |
| [84] | Predicting future flooding based on IoT data and weather forecasts in Kristianstad, Sweden | LSTM | MSE 1.25 | The study needs extensive validation to ensure robustness against different environmental data inputs. |
| [87] | A flood prediction model based on DNN that utilizes DL for complex feature extraction with two-dimensional | DNN | RMSE 6925.89, $R^2$ 0.8938 | Extremely high RMSE indicates significant prediction errors, underscoring the need for improved model tuning and feature selection. |

Flood prediction is an important factor in disaster forecasting of storm impacts and planning for an urban environment for disaster management and urban planning, providing timely warning and avoiding the impact of flooding. Different algorithms, such as flash floods, cities flooding, flooding and predicting water depths, have been used in various analyses of different floods. The approach and results, assessment and scope of several studies are reviewed in the review. The above studies use multiple algorithms for flood prediction which have a wide range of success or failures. [46] used Encoder Decoder LSTM (En-De LSTM) based model for monthly flow forecasting, obtained a high RMSE of 1171 showing serious prediction errors. [79] used Convolutional GRU (Conv-GRU) for water level prediction in Taiwan (RMSE = 0.77), with MAPE (30.68), reflecting variability challenges. Similarly, [45] used LSTM for improving model interpretability and accuracy, with an RMSE of 97.03 and an $R^2$ of 0.96.

However, one could observe a relatively high RMSE which indicates that prediction mistakes are still retained. The efficacy of LSTM and similar models for different flood prediction tasks is also shown in multiple researches. [80] used LSTM for gauge height prediction, which resulted in an RMSE of 0.8722, whilst [75] applied LSTM for urban flood prediction in Seoul, obtained an RMSE of 0.1 and an $R^2$ of 0.95. Using GBDT for urban flood prediction (Zhengzhou, China), it achieved an accuracy of 88.48% and an $R^2$ 0.96 [81]. In the work by [48], an analysis of RNN and CNN on flood hazard modeling in Seoul shows that DNN was able to overcome the RNN by providing an RMSE of 0.163 compared to 0.186 achieved with RNN. [55] DNN and RNN were also applied for predicting flash flood probability in Iran, where DNN outperformed RNN at RMSE 0.144 and AUC 0.832 [55]. [47] RF produced the good prediction results of streets for flood prediction (RMSE 0.01, $R^2$ 0.95). Other significant results were achieved by Hsu et al. also found by [84], which applied LSTM approach to flood prediction at Kristianstad, Sweden (MSE 1.25) and Hu et al. [149]. In this regard, they used LSTM-ROM for spatial-temporal flood prediction (RMSE 2.16). [87] used a CNN for the advanced feature extraction, nevertheless, showed high RMSE 6,925.89, indicating highly inaccurate predictions.

Though significant advances have been made in flood prediction models, there are still many challenges to overcome, especially with regard to data quality, model accuracy and generalization capabilities. However, although many papers gave robust results, the issues of model dimension, variety and scale of data have been under-investigated in studies. A high RMSE [46, 87] leads to high prediction errors, and indicates better data preprocessing and optimization of the model. Finally, models were developed by [79] and [45]. Even though they performed well, this highlights an opportunity for reducing data variability and preventing overfitting. Limitations for these approaches suggest that more advanced approaches, enhanced data integration procedures and more complex hybrid strategies are required for the robustness and reliability of flood prediction systems [88–90]. There are many improvements suggested on flood prediction models. This would also substantially improve the robustness of the proposed model if heterogeneous and high-quality data is used for data-collection, considering both real-time and historical data. Using strong validation techniques, such as cross-validation and bootstrapping, the generalization to unseen data can be improved. By altering hyperparameters and applying ensemble methods, the predictive performance may gain even more. Comparison of analysis to other databases then allows you to determine which algorithms perform best according to application — with different types of scenarios. In practice, therefore, it will also be, the case — and indeed the real possibility — that models will be applied for evaluation in addition to real-time calibration of their power and reliability as we endeavor to advance flood management.

## 5. Characteristics of the Flood Dataset

The rapid advancement of data-driven methods for flood prediction has generated a diverse body of literature with varied emphases on flood types, modelling strategies, data sources, scales, and applications. To unify these insights, this review synthesizes the selected studies into a coherent taxonomy that highlights both methodological trends and practical applications, as shown in Figure 5.

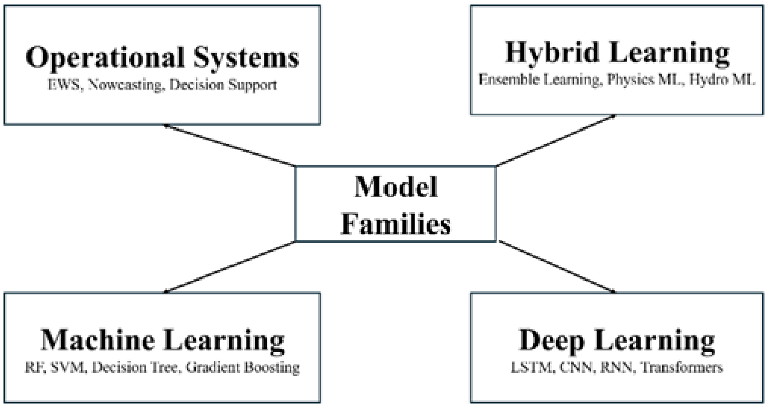


**Fig. 5** *Taxonomy of model families for flood prediction, highlighting ML, DL, hybrid, and operational systems*

### 5.1 Categories of flood

The literature reviewed covered multiple flood categories, indicative of the heterogeneous character of hydrological extremes. Most existing literature describes fluvial floods, especially riverine systems, and in which LSTM and ensemble neural networks were applied to project large catchment streamflow [90], [91]. Pluvial floods are well covered by ML classifiers and CNN-based inundation models exhibit strong predictive performance [66], [92], which are commonly due to heavy downpours in urban areas. Coastal flood events, such as tsunami and storm-surge events, are relatively less prevalent, but hybrid physics–data approaches have been developed to enhance the predictive accuracy [93]. Increasingly, it is reported that compound events, where concomitant factors such as river discharge and excessive rainfall worsen risk of flooding [94], [95], [96]. Urban drainage floods, an important subset of pluvial events, are modelled with high resolution data to account for the consequences of infrastructure bottlenecks [97]. Table 5 provides an overview of the distribution of flood types found in the studies reviewed.

**Table 5** *Distribution of flood types in the reviewed literature*

| Year | Studies | Flood Type | Notes |
|---|---|---|---|
| 2020 | [66] | Pluvial (Urban) | ML classifiers used for rainfall–runoff-based inundation in urban areas |
| 2021 | [69] | Urban drainage / Pluvial | CNN-based fast inundation model for drainage-limited urban floods |
| 2020 | [67] | Fluvial | Probabilistic ML for river discharge forecasting |
| 2019 | [98] | Flash/Pluvial | LSTM for rapid-onset flash floods |
| 2020 | [81] | Urban drainage | Predictive ML model for flood depth estimation |
| 2020 | [79] | Fluvial | DL model for river flooding and anomaly detection |
| 2022 | [71] | Compound (Fluvial + Pluvial) | Hybrid hydrodynamic-ML to capture interacting drivers |
| 2023 | [73] | Coastal/Fluvial | Comparative analysis for river–coastal interactions |
| 2022 | [72] | Fluvial / Susceptibility | Classification of flood-prone rural areas |
| 2022 | [84] | Pluvial/Urban | Sensor-driven ML prediction |
| 2023 | [99] | Fluvial | Regional-scale long-lead prediction |
| 2019 | [100] | Compound (Fluvial + Coastal) | Physics-guided ML for generalizability |
| 2022 | [70] | Fluvial/Pluvial | Unsupervised DL for cross-domain flood forecasting |

## 5.3 Model Families

The classification of model families represents a transformation from traditional ML to novel DL and hybrid models. ML techniques (e.g., RF, SVM, decision trees) can still be commonly employed for susceptibility mapping and classification tasks [101], [102]. The majority of streamflow and inundation forecasting tasks are based on deep learning (DL) architectures like LSTM, CNN, and spatio-temporal attention models [103], [104]. More recently, transformer-based models have been investigated to capture long-range dependencies in sequential hydrological data [105]. Hybrid frameworks combine physical models with ML to enhance interpretability and generalization [106]. Finally, operational systems, such as probabilistic early warning platforms and nowcasting models, demonstrate the translation of academic research into real-time decision-making [107]. Figure 5 illustrates the taxonomy of model families and their overlaps.

## 5.4 Model Data Sources

Predictive performance is rooted in data diversity. Hydrological and meteorological gauges are still foundational, however, especially when training ML classifiers. [108]. Remote sensing (RS) datasets, both SAR and optical, support susceptibility mapping at larger scales [109]. Increasingly reanalysis and forecast products are incorporated to extend lead times. [110]. Topographic and land-cover information (DEM/LC) facilitate urban flood models as they encode aspects of the terrain [111]. Social and crowdsourced based data, although developing in recent years, have worked well in complement to physical observations. [112]. Infrastructure data, including drainage networks and reservoir levels, are finally included in operational flood forecasting systems. [113]. A brief description of data sources and their usual uses is indicated in Table 6.

**Table 6** *Data sources utilized in flood prediction models across the reviewed studies*

| Year | Study | Data Sources | Applications |
|---|---|---|---|
| 2020 | [66] | Gauge rainfall, DEM, land use | Urban flood susceptibility mapping |
| 2020 | [65] | Gauges, crowdsourced reports, and social media | Real-time forecasting |
| 2021 | [69] | Rainfall, DEM, drainage network | Short-term inundation mapping |
| 2019 | [98] | Rainfall gauges, streamflow gauges | Flash flood early warning |
| 2020 | [81] | DEM, land cover, drainage data | Estimation of urban flood depth |
| 2020 | [79] | Streamflow gauges, hydrological observations | Long-lead River flood forecasting |

| Year | Study | Data Sources | Applications |
|---|---|---|---|
| 2022 | [71] | Hydrodynamic simulation outputs, precipitation forecasts | Hybrid flood forecasting |
| 2023 | [73] | Tide gauges, river flow, climate projections | Coastal and riverine risk assessment |
| 2022 | [72] | DEM, land cover, rainfall gauges | Flood susceptibility classification |
| 2022 | [84] | IoT sensor data (rainfall, water levels) | Real-time monitoring |
| 2023 | [99] | Climate reanalysis (ERA5), rainfall forecasts | Long-lead extreme event prediction |
| 2019 | [100] | River flow, tidal data, DEM | Physics-guided predictive modelling |
| 2022 | [70] | Cross-domain hydrological datasets | Transfer learning for flood prediction |

The reviewed studies vary in scale, similar to their decision contexts. Spatially, urban case studies are dominated by local-scale models, in which fine-scale prediction is crucial for disaster management [114]. Generalized flood risk maps are generated by regional and national scale models using satellite imagery and climate characteristics [115]. Timing allows the lead time on the models to vary from real-time nowcasting to more strategic prediction beyond 48 hours. The dominant intervals that are widely adopted are short-term horizons as they are used by the LSTM and CNN models which can surpass the statistical baselines [116], [117]. Hybrid hydrodynamic–ML architectures achieve medium-term windows (24–48 h) [118], whereas long-lead forecasting (>48 h) usually uses ensemble DL models [119], [120].

Flood prediction models are used in different contexts and as such, their applications show different emphases. The core encompasses forecasting and classification of flooding, from susceptibility mapping in rural Bangladesh [121] to estimated depths of inundation in large urban corridors [122]. Early Warning Systems (EWS) remain a fundamental operational tool and have been employed to improve disaster preparedness through real-time probabilistic models [123]. However despite the progress made in risk analysis by classical models, such as Monte Carlo, the persistence of susceptibility mapping in developing places is present, and spatial risk zoning is effected by ML classifiers [124]. Use cases for decision support include joining up policy and infrastructure frameworks. The research of [51] demonstrates not only the general applicability of these models but also their extended applications beyond their academic utility. The one-dimensional taxonomy reveals a strong methodological evolution toward ML classifiers for early susceptibility mapping, DL construct for time series prediction, and hybrid and operational framework to bridge science and practice. Diverse data sources and multiplicity of modelling contribute also to the generalist nature of large generic and implementable flood forecasting systems. Collectively, these three dimensions give a comprehensive perspective that can inform our understanding of the state of the art and lead, in the process, future studies.

## 5.5 Evaluation Practice & Datasets

The evaluation practices for flood prediction research are heterogeneous and vary across datasets, modeling techniques, application contexts detailed in the literature. Thus, the synthesis is of paramount importance for examining the manner in which studies maintain methodological rigor and generalizability. This section summarizes the primary approaches in four domains: datasets, preprocessing and feature engineering, evaluation metrics and validation strategies, reproducibility and openness.

The field of flood prediction and classification work requires a wide range of datasets and is broadly divided into public benchmark datasets and those collected locally or through private collection. For example, public datasets (e.g., ERA5 reanalysis, TRMM and GPM precipitation products, DEM repositories like SRTM) provide universal, large-scale standardized data that facilitate reproducibility and cross-study comparability. Some of the reviewed articles such as [55], [83], [125] used these global products to simulate rainfall–runoff dynamics or provide meteorological forcing for ML and DL models.

On the other hand, private or local datasets are still a great deal of data [126]-[128], mainly in studies carried out in small watersheds or urban basins. These involve streamflow measurements from local gauges, rainfall data from municipal meteorological agencies, and remote sensing data specific to the region of interest. Frequent limitation continues to be, however, the absence of standardized reporting and limited access to it, so reproducibility across the territory is restricted. Table 7 summarizes dataset types with a structured summary including their spatial/temporal resolution, accessibility and representative studies.

**Table 7 Representative datasets used in flood prediction research**

| Dataset Type | Sources | Resolution | Accessible | Studies |
|---|---|---|---|---|
| Gauge-based hydrology | Local rainfall & streamflow gauges | Point-scale, hourly–daily | Private | [46], [67], [77], [79], [85], [129], [130], [131], [132] |
| Remote sensing | TRMM, GPM, MODIS, Sentinel SAR | 0.1–0.25°, hourly–daily | Public | [55], [66], [78], [99], [133], [134], [135], [136] |
| Reanalysis/Forecasts | ERA5, NCEP, WRF outputs | 0.25°, hourly–6-hourly | Public | [73], [83], [88], [99], [131], [137], [138] |
| Terrain/Landcover | SRTM DEM, CORINE, NLCD | 30m–90m, static | Public | [39], [73], [81], [130], [139], [140] |
| Social/Crowdsourced | Tweets, volunteered rainfall reports | Irregular, event-based | Limited | [32], [136], [141] |

Preprocessing plays a central role in ensuring the quality of inputs to ML models. Typical techniques use temporal lags, signal decomposition (e.g., wavelet transforms and empirical mode decomposition) to extract dominant patterns, and data fusion to combine heterogeneous data sources (e.g., radar rainfall, satellite-derived indices, and ground observations).

For instance, studies such as [142], which combined LSTM with wavelet decomposition showed significantly improved short-term lead time forecasts, and feature selection strategies such as recursive feature elimination or principal component analysis (PCA) was applied to limit dimensions in large-scale datasets [143]. Also, some works used physics-informed feature engineering in which hydrodynamic model outputs, for example, water levels and runoff coefficients, were built into the analysis as predictors, contributing to the interpretability of models based on data.

Evaluation metrics are selected according to the prediction task. The most popular metrics for regression-based flood prediction are RMSE, MAE, Nash–Sutcliffe Efficiency (NSE) and $R^2$. Such classification-based studies (e.g., flood susceptibility mapping) use precision, recall, F1-score, and AUC to evaluate discrimination power. Validation techniques, however, were highly variable. Traditional k-fold cross-validation was typical in ML-based research, whereas DL was often based on spatial- and temporal-split approaches for assessing generalization to unseen situations. Some recent works have applied uncertainty quantification, e.g., probabilistic predictions based on Bayesian DL and quantile regression forests while acknowledging that extreme events are inherently stochastic events [144]-[146]. The metrics and validation schemes included in the reviewed studies are summarized in Table 8.

**Table 8** *Evaluation metrics and validation practices in reviewed studies*

| Task Type | Metrics Used | Validation Strategy | Studies |
|---|---|---|---|
| Flood prediction (regression) | RMSE, MAE, NSE, $R^2$ | k-fold CV, temporal splits (event-based train/test) | [46], [65], [70], [73], [79], [83], [85], [88], [99], [119], [129], [131], [132], [147] |
| Flood classification | Accuracy, Precision, Recall, F1, AUC | Stratified CV, spatial splits | [39], [55], [66], [77], [78], [130], [139], [140], [148] |
| Probabilistic forecasting | CRPS, Prediction Interval Coverage | Probabilistic calibration tests | [67], [80], [137], [141], [149] |
| Hybrid/Physics-informed | RMSE, NSE + hydrological fit indices | Leave-one-station-out | [69], [84], [100], [133], [150], [151], [152], [153], [154], [155], [156] |

A key methodological issue found in the reviewed documents is the inconsistent evaluation of metrics. Although RMSE, MAE, and $R^2$ are extensively used, essential elements such as uncertainty quantification, probabilistic accuracy, and skill scores are employed to a lesser extent, which may affect the robustness of the comparison between models. This discrepancy may add a significant challenge to reproducibility and to building consistent standard benchmarks from study to study. Also, poor regional coverage was observed in the reporting of evaluation metrics. Studies from Asia and North America usually included strong statistical validation. Research conducted in Africa and South America, however, often used a smaller number of basic error metrics due to data limitations. Such disparities prevent global comparison of results and highlight the importance of global evaluation protocols that balance methodological rigor and local data availability.

Reproducibility is still a major hurdle. As for the latter, a few recently published works adopted open science behavior, providing both datasets and source code [28], [157], [158]. A large volume of this research is done with proprietary or inaccessible local datasets. This impedes independent replication and benchmarking. One smaller, albeit growing trend is the publication of open repositories by scholars of pre-trained models, scripts, and pre-processed data to support reuse, particularly when applying to international projects such as the Global Flood Detection System (GFDS) and Copernicus Emergency Management Services. A fifth of the works considered, however, only explicitly reported a link to code or archive of data — a gap between methodological novelty and transparency.

## 6. Analysis and Discussion

The analysis of Tables 2-5 highlights key insights into the performance of ML, DL, and hybrid models for flood prediction and classification. As demonstrated in Table 2, the proposed RF and SVM models can be adopted in general with ML-based flood forecasting, RF, especially handling nonlinear relationships and large amounts of data, and its good performance especially well, were very good in flood prediction. However, their generalization limitation remains to be determined. As can be seen in Table 3, the ANN (primarily MLP) has the best accuracy (97.83%) for flood classification but is not interpretable. Thus, DT and RF aim to be more interpretable for real-time use. As suggested by Table 4, DL models, especially LSTM and Conv-LSTM, have robust effectiveness at capturing temporal dependencies and for the performance improvement of flood prediction, but they require substantial training data and computational resources. Table 5 shows that Reinforcement Learning (RL), including PPO (Proximal Policy Optimization) with LSTM, can be combined to ensure a balance in the performance of classification through informed decision-making and to address the problems arising from the large computational demands and overfitting. In summary, hybrid models of ML/DL and hydrological frameworks show great promise. However, they require additional efforts for interpretability, real-time adaptability, and scalability to floods. This review describes the performance and limitations of ML/DL techniques on flood prediction and classification, then addresses the following outcomes.

- ML-based Flood Prediction Models: Multiple studies that use such models illustrate different performances for flood prediction with ML models like RF, SVM and ANN. Generally, RF outperforms other models because the model is flexible enough to deal with different dataset types and nonlinear relationships. When data are limited, SVMs work well; they incur much less computational requirements compared with ANNs that depend on much more computational resources. However, due to dataset heterogeneity, the dataset has not been generalized across different geographical regions.
- ML-based Flood Classification Models: The highest accuracy (~97.83%) for flood classification is obtained in ANN models, especially in MLP-based architectures. However, despite their superior performance, DT and RF are explainable and robust despite their structural limitations, making them preferred for real-time applications. It has been demonstrated that integrating GIS data into classification models has mixed results due to difficulties integrating spatial information into ML techniques.
- DL-based Flood Prediction Models: Employing DL methods, including LSTM, Conv-LSTM, and transformer-based architectures, improves predictive accuracy through spatial and temporal dependency capture. Generally, LSTM-based models work well for streamflow prediction but require large datasets for training. Although DNNs perform well in image-based flood classification, they are insensitive to temporal aspects. However, hybrid approaches that meld DL with physical hydrological models achieve better accuracy and generalizability.
- DL-based Flood Classification Models: PPO combined with LSTM is promising for real-time flood classification using DRL techniques. However, training these models is very data-intensive and may overfit due to the scarcity of labelled data. The tension lies in striking a balance between prediction accuracy and interpretability.

Alternatively, the best prediction models across various application contexts tend to integrate the strengths of several modeling paradigms to enhance stability and extrapolation. DL and RL models are frequently cited as examples in discussions of flood prediction and classification, rather than as solutions to the problem. Modeling sequence-dependent processes, including LSTM, is widely applied because it is highly effective at capturing time-related dependencies, seasonality, and long-term relationships in hydrological and meteorological time-series data. GRU and temporal convolutional networks (TCNs) are also common sequential models that are often used in flood level prediction and event classification. But these methods are usually time-intensive for large-scale historical data and require significant training time, which can restrict their use in urgent situations.

Policy-gradient or value-based methods, including RL methods such as PPO and Q-learning, have become promising solutions for adaptive flood prediction and decision support in dynamic and uncertain environments. The techniques are highly appropriate for learning the best prediction or control policies under variable hydrological conditions, for reservoir operation, or for early-warning levels. Also, hybrid deep RL models that combine sequence models and RL agents yielded more positive outcomes. Recurrent models incorporate temporal memory and feature representations in such combinations, whereas the RL component enables adaptive learning and policy optimization. This collaboration has also been documented to enhance both flood classification and flood forecasting performance compared to single-model technologies.

These sophisticated models, although highly predictive, are limited by their high computational requirements, centralized data, and low interpretability, which makes it difficult to use them in real-time operations, flood forecasting, and early-warning systems. This means that, despite the fact that the current state of DL and RL models is the most accurate and flexible, future studies ought to be done on hybrid and lightweight models that combine data-driven learning with hydrological domain knowledge, physics-based modeling, and explainable AI methods to increase reliability, transparency, and applicability to the real world.

- Sequence-based DL models: Sequential hydrological and meteorological data models like LSTMs are considered to be the most effective DL methods to predict and classify floods since they can be used to model time-based high temporalities and dependencies. GRU and TCN, similar architectures, are also employed in flood forecasting. Although these models perform well, they require large datasets and substantial computational power to operate in real time due to their stringent requirements [46], [75].
- RL-based methods: PPO and Q-learning algorithms are methods of RL that are being actively studied in flood forecasting and optimization of decisions in changing and uncertain settings. Such techniques are applicable to adaptive flood management and early-warning cases because they trade off exploration and exploitation in acquiring prediction or control policies. But their cost of computation has limited their use in the real world [84].
- Hybrid Deep Reinforcement Learning (DRL) models: It has been demonstrated that a combination of sequence models (like LSTM) with reinforcement learning (such as PPO) is promising when it comes to predicting and classifying floods. Within these frameworks, LSTM adds temporal memory and feature learning, whereas PPO adds adaptive policy optimization, thereby increasing accuracy and robustness in prediction. However, the complexity of these hybrid models, centralized data requirements, and limited interpretability continue to pose challenges to their training and operational deployment [83], [85].

Critically, beyond model-level insights, two broader limitations emerge. First, regional gaps persist. Most reviewed studies are concentrated in Asia (notably China and India) and North America, while flood-prone regions such as Africa and South America remain underrepresented. This imbalance raises concerns about the transferability of AI-based models to areas with limited hydrological infrastructure and data availability. Second, methodological gaps constrain progress. The overwhelming emphasis on predictive accuracy overshadows equally critical aspects such as interpretability, reproducibility, and uncertainty quantification. Few studies employ XAI or probabilistic frameworks, despite their importance for gaining stakeholder trust in operational forecasting [159], [160]. Moreover, evaluation metrics remain inconsistently applied across studies, hindering systematic comparison.

## 7. Challenges and Future Directions

Several challenges continue to limit the effectiveness and real-world use of flood prediction models, as summarized in the open research directions. One persistent issue is data scarcity, since many flood-prone regions still lack sufficient and well-labeled datasets, which affects model reliability and generalization. Real-time adaptability is also weak, as many existing models cannot update predictions as environmental conditions change, reducing their value for early warning systems. Interpretability is a major issue, especially for deep learning models such as DNNs and LSTMs, which are black boxes and it can be hard for the user to interpret or trust their outputs. Moreover, high computational cost limits the application since high-level models are generally required to rely on hardware that is limited in low-resource environments. Hybrid strategies fusing ML/DL with classical hydrological models have the potential to enhance accuracy; however, they lack implementation power due to data fusion and model stability concerns. Addressing these gaps will demand additional efforts in, among other things, augmenting data, explainable AI, developing efficient models and building real-time processing frameworks for more feasible and reliable systems of flood prediction and classification. The remaining challenges which still exist are summarized in Table 9:

**Table 9** *Possible research directions*

| Research Direction | Description |
|---|---|
| Data Scarcity | Many regions lack extensive, labelled flood datasets, which impacts model generalization. |
| Real-time Adaptability | Current models struggle to adapt to real-time flood events, limiting practical deployment. |
| Model Interpretability | DL models lack transparency, particularly DNNs and LSTMs, making decision-making difficult. |
| Computational Demand | Advanced DL models require significant resources, hindering deployment in low-resource areas. |
| Hybrid Model Integration | Combining ML/DL with hydrological models could enhance predictions, but integration remains challenging. |
| Data Scarcity | Many regions lack extensive, labelled flood datasets, which impacts model generalization. |
| Real-time Adaptability | Current models struggle to adapt to real-time flood events, limiting practical deployment. |

The main limitation of this review is the generalizations made about data and models, as the studies examined are specific to regions with unique hydrological characteristics and, therefore, not applicable to a broader context. In addition, the evaluation of metrics in the studies is not homogeneous, as some report RMSE, while others report accuracy or NSE, preventing the establishment of common performance indicators. Moreover, there is no real-time implementation evaluation, which means the models will work well with historical data but are uncertain in real-time flood prediction. Moreover, non-English texts were excluded and, by extension, important work from countries such as India and Bangladesh, often affected by flooding, may have been ignored and overlooked. The review also highlights the limited data available and computational complexity associated with the DL models, but is not very insightful on the solutions that may be used to overcome the limitations of these models. Conclusively, hybrid models, combining ML/DL together with hydrological simulations, have potential; however, such models must consider real-world application, feasibility issues, and interpretability for use cases such as flood prediction to be robust and effective.

## 8. Conclusion

This systematic overview of the formulation and deployment of ML/DL models for flood prediction and classification answers the guiding research questions. In RQ1, the analysis highlights that traditional ML models like SVMs have obtained a good MSE of 0.07, are proving effective, and provide accurate (albeit short-term) flood predictions because they consider the very fast and complex forces that occur when flood events occur, with lead times of 1–3 hours in the Yilan River. Models like RF perform well for predicting complex patterns of surface water flow in the city. The RF model classified 13 multivariate flood statuses in some parts of a few central London cities: 86.57% accuracy is proved. Another study also demonstrated a 96% accuracy of predicting flood status for Lisbon, Portugal. In comparison, DL models (such as the LSTM models for predicting water-level under China's Sihu basin) show a remarkably good RMSE of 0.04, and are better on higher dimensional complex data. Regarding RQ2, the review points out that model performance may be substantially improved if diverse data types (e.g., hydrological, meteorological, spatial) are included in the model and hybrid models combining DL and RL for generating DRL or numerical hydrological simulations are developed. Not only do these strategies improve precision, but they also result in better real-time adaptability to the behavior of floods. In response to RQ3, the work focuses on recurring issues such as the lack of data, demand for computer resources and interpretability in advanced models that could accurately forecast and monitor the short and long-term floods over time. There still is a geographic disparity: most studies originated in Asia and North America and the flood-vulnerable areas such as Africa and South America are mostly missing, which limits global application as well. In order to fill this gap, this study presents three ways to advance the literature. It combines the flood prediction and classification problem in a single systematic literature review, offers to link datasets, methods and performance measurement; and presents to the public both regional and methodological shortcomings with which any future study needs to address. In sum, accuracy is not enough: operational flood forecasting must be model- interpretable, reproducible, regionally adaptable, and underpinned by solid open datasets. There are ways to address such gaps through the use of XAI frameworks, standardized open datasets, and hybrid models that combine predictive capability with transparency. Finally, this review underscores the importance for current scholarship to place an emphasis on enhancing interpretable, hybrid, and evidence-based flood models to assist policymakers, urban planners, and disaster managers in developing flood risk management strategies/applications that are resilient to flooding. The primary key contribution of this research is that this review presents the first integrated blend of ML and DL approaches for flood prediction and classification (2020–2025), proposing a structured taxonomy that connects datasets, methods, and evaluation metrics while highlighting urgent regional, methodological, and data gaps relevant to future research and practice.

## Acknowledgement

The authors would like to acknowledge that this research was supported by the Ministry of Higher Education (MOHE) through the Fundamental Research Grant Scheme (FRGS/1/2022/ICT02/UTHM/02/2). The authors also thank Alnoor University for supporting this project through the Research Supporting Fund ANUI/2025/ENG51.

## Conflict of Interest

The authors declare that they have no conflict of interest regarding the publication of this paper.

## Author Contribution

*The authors acknowledge their contributions to the paper as follows:* ***literature search, screening, data extraction and drafting the first draft of the manuscript:*** *Jabir Abubakar Salisu and Salama A. Mostafa;* ***conception and design of the study, such as formulation of review framework, taxonomy and research objectives:*** *Salama A. Mostafa and Hairulnizam Mahdin;* ***methodological approach particularly on machine learning and deep learning approaches classification and critically reviewing evaluation measures:*** *Muhammad Aamir and Heru Nurwarsito;* ***analysis, interpretation and refinement of the discussion on emerging trends and research gaps:*** *Ammar Alazab. The authors of this manuscript also all engaged in critical revision of the manuscript and contributed intellectually, as well as approved the final version for publication.*

## References

[1] Nogueira, Keiller, Penatti, Otávio A. B., dos Santos, José A., & Torres, Ricardo da S. (2018). Exploiting ConvNet diversity for flooding identification. IEEE Geoscience and Remote Sensing Letters, 15(9), 1446–1450. https://doi.org/10.1109/LGRS.2018.2845549

[2] Rosselló, Jaume, Becken, Susanne, & Santana-Gallego, María. (2020). The effects of natural disasters on international tourism: A global analysis. Tourism Management, 79, 104080. https://doi.org/10.1016/j.tourman.2020.104080

[3] Merz, Bruno, Kreibich, Heidi, Schwarze, Reimund, & Thieken, Annegret. (2021). Causes, impacts and patterns of disastrous river floods. Nature Reviews Earth and Environment, 2(9), 592–609. https://doi.org/10.1038/s43017-021-00195-3

[4] Merghadi, Abdelghani, Yunus, Ali P., Dou, Jianping, Whiteley, James, ThaiPham, Binh, Bui, Dieu Tien, & Yilmaz, Ismail. (2020). Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. Earth-Science Reviews, 207, 103225. https://doi.org/10.1016/j.earscirev.2020.103225

[5] Costache, Radu, Zaharia, Liliana, Prăvălie, Remus, & Arabameri, Alireza. (2020). Spatial prediction of flood potential areas using novel hybridizations of fuzzy decision-making, bivariate statistics, and machine learning. Journal of Hydrology, 585, 124808. https://doi.org/10.1016/j.jhydrol.2020.124808

[6] Khosravi, Kamran, Nohani, Ehsan, Maroufinia, Ehsan, & Pourghasemi, Hamid Reza. (2018). A comparative assessment of decision tree algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. Science of the Total Environment, 627, 744–755. https://doi.org/10.1016/j.scitotenv.2018.01.266

[7] Alabbad, Yousef, & Demir, Ibrahim. (2022). Comprehensive flood vulnerability analysis in urban communities: Iowa case study. International Journal of Disaster Risk Reduction, 74, 102955. https://doi.org/10.1016/j.ijdrr.2022.102955

[8] Chen, Liang, Wang, Hongtao, Zhang, Yi, Liu, Yanyan, & Wu, Qiang. (2024). Influence of flood events on the ecological characteristics of Bolboschoenus planiculmis: Implications for restoration of Grus leucogeranus habitats. Water, 16(24), 3672. https://doi.org/10.3390/w16243672

[9] Sun, Bo, Wu, Jian, He, Dongxiao, Wang, Mingyu, & Li, Wei. (2022). Experimental warming reduces ecosystem resistance and resilience to severe flooding in a wetland. Science Advances, 8(4), eabl9526. https://doi.org/10.1126/sciadv.abl9526

[10] Huang, Zhiwei, Pan, Bo, Zhao, Xinyu, Liu, Xiaowei, Liu, Xinxin, & Zhao, Guohua. (2023). Hydrological disturbances enhance stochastic assembly processes and decrease network stability of algae communities in a highland floodplain system. Science of the Total Environment, 903, 166207. https://doi.org/10.1016/j.scitotenv.2023.166207

[11] Hua, Peng, Qiu, Zhe, Wang, Jian, & Li, Ming. (2020). Evaluating the effect of urban flooding reduction strategies in response to design rainfall and low-impact development. Journal of Cleaner Production, 242, 118515. https://doi.org/10.1016/j.jclepro.2019.118515

[12] Chen, Wei, Li, Xinyi, Wang, Yichen, Liu, Xiaoping, & Hong, Haoyuan. (2020). Modeling flood susceptibility using data-driven approaches of naïve Bayes tree, alternating decision tree, and random forest methods. Science of the Total Environment, 701, 134979. https://doi.org/10.1016/j.scitotenv.2019.134979

[13] Tanoue, Masahiro, Taguchi, Ryosuke, Nakata, Shota, Watanabe, Shogo, Fujimori, Shinichiro, & Hirabayashi, Yukiko. (2020). Estimation of direct and indirect economic losses caused by a flood with long-lasting inundation: Application to the 2011 Thailand flood. Water Resources Research, 56(5), e2019WR026092. https://doi.org/10.1029/2019WR026092

[14] Rentschler, Jun, Salhab, Melda, & Jafino, Brandon A. (2022). Flood exposure and poverty in 188 countries. Nature Communications, 13(1), 3527. https://doi.org/10.1038/s41467-022-30727-4

[15] Muzamil, Syed A. H. B. S., Shafie, Azman, Hamzah, Aini, & Rahman, Zainal. (2022). Proposed framework for the flood disaster management cycle in Malaysia. Sustainability, 14(7), 4088. https://doi.org/10.3390/su14074088

[16] Ashikin, Nurul, Diana, Nor Izzati, Siwar, Chamhuri, Alam, Mahbub, & Yasar, Mahmud. (2021). Community preparation and vulnerability indices for floods in Pahang State of Malaysia. Land, 10(2), 198. https://doi.org/10.3390/land10020198

[17] Raihan, Abdul, Pereira, Joao J., Begum, Rokiah, & Rasiah, Rajah. (2023). The economic impact of water supply disruption from the Selangor River, Malaysia. Blue-Green Systems, 5(2), 102–120. https://doi.org/10.2166/bgs.2023.031

[18] Sun, Wei, Bocchini, Paolo, & Davison, Brian D. (2020). Applications of artificial intelligence for disaster management. Natural Hazards, 103(3), 2631–2689. https://doi.org/10.1007/s11069-020-04124-3

[19] Zhang, Jun, Zheng, Hao, Zhang, Xiaodong, Lin, Zhe, VanLooy, Joris, & Kennedy, Andrew. (2023). Strategic land acquisition for efficient and equitable flood risk reduction in the United States. Climate Risk Management, 42, 100554. https://doi.org/10.1016/j.crm.2023.100554

[20] Mishra, Anil, Pradhan, Biswajeet, Pourghasemi, Hamid Reza, & Bui, Dieu Tien. (2022). An overview of flood concepts, challenges, and future directions. Journal of Hydrologic Engineering, 27(6), 04022014. https://doi.org/10.1061/(ASCE)HE.1943-5584.0002164

[21] Donnelly, James, Abolfathi, Sam, Pearson, James, Chatrabgoun, Omid, & Daneshkhah, Alireza. (2022). Gaussian process emulation of spatio-temporal outputs of a two-dimensional inland flood model. Water Research, 225, 119100. https://doi.org/10.1016/j.watres.2022.119100

[22] Tang, Yu, Liu, Kai, Zhang, Fan, Li, Zhi, & Wang, Wei. (2023). Flood forecasting based on machine learning pattern recognition and dynamic migration of parameters. Journal of Hydrology: Regional Studies, 47, 101406. https://doi.org/10.1016/j.ejrh.2023.101406

[23] Zelentsov, Viktor A., Potryasaev, Sergey A., Pimanov, Ivan Y., & Ponomarenko, Mikhail R. (2019). Integrated use of GIS, remote sensing data, and a set of models for operational flood forecasting. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII(3/W8), 477–483. https://doi.org/10.5194/isprs-archives-XLII-3-W8-477-2019

[24] Lamb, Rob, Garside, Peter, Pant, Raghav, & Hall, Jim W. (2019). A probabilistic model of the economic risk to Britain's railway network from bridge scour during floods. Risk Analysis, 39(11), 2457–2478. https://doi.org/10.1111/risa.13370

[25] Kang, Beomseok, Feagin, Rusty A., Huff, Timothy, & Durán Vinent, Oscar. (2024). Stochastic properties of coastal flooding events: Part 2. Probabilistic analysis. Earth Surface Dynamics, 12(1), 105–115. https://doi.org/10.5194/esurf-12-105-2024

[26] Li, Xiaojing, & Willems, Patrick. (2020). A hybrid model for fast and probabilistic urban pluvial flood prediction. Water Resources Research, 56(6), e2019WR025128. https://doi.org/10.1029/2019WR025128

[27] Goenner, Andrew R., Franz, Kristie J., Gallus, William A., & Roberts, Benjamin. (2020). Evaluation of probabilistic quantitative precipitation forecasts for flood forecasting. Water, 12(10), 2860. https://doi.org/10.3390/w12102860

[28] Luo, Peng, Zhou, Ming, Kang, Shichang, & Li, Xinyi. (2022). Urban flood numerical simulation: Research, methods, and future perspectives. Environmental Modelling and Software, 156, 105478. https://doi.org/10.1016/j.envsoft.2022.105478

[29] Rosser, James F., Leibovici, Didier G., & Jackson, Matthew J. (2017). Rapid flood inundation mapping using social media, remote sensing, and topographic data. Natural Hazards, 87(1), 103–120. https://doi.org/10.1007/s11069-017-2755-0

[30] Farhadi, Hamid, & Najafzadeh, Mohammad. (2021). Flood risk mapping using remote sensing data and random forest technique. Water, 13(21), 3115. https://doi.org/10.3390/w13213115

[31] Shastry, Anish, Carter, Elizabeth, Coltin, Brian, Sleeter, Rachel, McMichael, Shane, & Eggleston, John. (2023). Mapping floods from remote sensing data and quantifying the effects of surface obstruction by clouds and vegetation. Remote Sensing of Environment, 291, 113556. https://doi.org/10.1016/j.rse.2023.113556

[32] Munawar, Hammad S., Hammad, Amin W. A., & Waller, Scott T. (2021). A review on flood management technologies related to image processing and machine learning. Automation in Construction, 132, 103916. https://doi.org/10.1016/j.autcon.2021.103916

[33] Munawar, Hammad S., Hammad, Amin W. A., & Waller, Scott T. (2022). Remote sensing methods for flood prediction: A review. Sensors, 22(3), 960. https://doi.org/10.3390/s22030960

[34] Farhadi, Hamid, Esmaeily, Ali, & Najafzadeh, Mohammad. (2022). Flood monitoring by integrating remote sensing techniques and multi-criteria decision-making methods. Computers and Geosciences, 160, 105045. https://doi.org/10.1016/j.cageo.2022.105045

[35] Antonopoulos, Ilias, Kitsikoudis, Vasileios, Glykas, Michael, & Ioannou, Nikolaos. (2020). Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review. Renewable and Sustainable Energy Reviews, 130, 109899. https://doi.org/10.1016/j.rser.2020.109899

[36] Saravanan, Rajesh, & Sujatha, Priya. (2018). State-of-the-art supervised machine learning techniques for data classification. In Proceedings of the 2nd International Conference on Intelligent Computing and Control Systems (pp. 945–949). IEEE. https://doi.org/10.1109/ICCONS.2018.8663155

[37] Jha, Kundan Kumar, Jha, Rakesh, Jha, Anil Kumar, Hassan, Mohammed A. M., Yadav, Shyam Kumar, & Mahesh, T. (2021). A comparative survey of machine learning algorithms across applications. In Proceedings of the 5th International Conference on Computational Systems and Information Technology for Sustainable Solutions (pp. 1–5). IEEE. https://doi.org/10.1109/CSITSS54238.2021.9683524

[38] Naeem, Saqib, Ali, Asif, Anam, Sana, & Ahmed, Muhammad M. (2023). Unsupervised machine learning algorithms: A comprehensive review. International Journal of Computing and Digital Systems, 13(1), 911–921. https://doi.org/10.12785/ijcds/130172

[39] Motta, Marco, de Castro Neto, Miguel, & Sarmento, Pedro. (2021). A mixed approach for urban flood prediction using machine learning and GIS. International Journal of Disaster Risk Reduction, 56, 102154. https://doi.org/10.1016/j.ijdrr.2021.102154

[40] Kumarasinghe, Kanishka, Kasabov, Nikola, & Taylor, Daniel. (2020). Deep learning and deep knowledge representation in spiking neural networks for brain-computer interfaces. Neural Networks, 121, 169–185. https://doi.org/10.1016/j.neunet.2019.08.029

[41] Livezey, Jesse A., & Glaser, Joshua I. (2021). Deep learning approaches for neural decoding across architectures and recording modalities. Briefings in Bioinformatics, 22(2), 1577–1591. https://doi.org/10.1093/bib/bbaa355

[42] Sejnowski, Terrence J. (2020). The unreasonable effectiveness of deep learning in artificial intelligence. Proceedings of the National Academy of Sciences, 117(48), 30033–30038. https://doi.org/10.1073/pnas.1907373117

[43] Zhang, Kun, Zheng, Lei, Liu, Zhaoyang, & Jia, Ning. (2020). A deep learning based multitask model for network-wide traffic speed prediction. Neurocomputing, 396, 438–450. https://doi.org/10.1016/j.neucom.2018.10.097

[44] Yu, Jeffrey Jingqiang Q., Markos, Christos, & Zhang, Shunqing. (2022). Long-term urban traffic speed prediction with deep learning on graphs. IEEE Transactions on Intelligent Transportation Systems, 23(7), 7359–7370. https://doi.org/10.1109/TITS.2021.3069234

[45] Ding, Yu, Zhu, Yu, Feng, Jian, Zhang, Peng, & Cheng, Zhen. (2020). Interpretable spatiotemporal attention LSTM model for flood forecasting. Neurocomputing, 403, 348–359. https://doi.org/10.1016/j.neucom.2020.04.110

[46] Liu, Dong, Jiang, Wei, Mu, Linlin, & Wang, Shaowen. (2020). Streamflow prediction using deep learning neural network: Case study of the Yangtze River. IEEE Access, 8, 90069–90086. https://doi.org/10.1109/ACCESS.2020.2993874

[47] Jiang, Zhenhua, Liu, Xiaoyu, Zhang, Yifan, Chen, Xuebin, & Wang, Haibo. (2022). Coupling machine learning and weather forecasts to predict farmland flood disasters: A case study in the Yangtze River basin. Environmental Modelling and Software, 155, 105436. https://doi.org/10.1016/j.envsoft.2022.105436

[48] Lei, Xiaohui, Chen, Jiahui, Zhang, Yujie, Wang, Rui, & Kim, Sunmin. (2021). Urban flood modeling using deep learning approaches in Seoul, South Korea. Journal of Hydrology, 601, 126684. https://doi.org/10.1016/j.jhydrol.2021.126684

[49] KC, Sudarshan, Shrestha, Shree, Ninsawat, Somporn, & Chonwattana, Supat. (2021). Predicting flood events in Kathmandu Metropolitan City under climate change and urbanisation. Journal of Environmental Management, 281, 111894. https://doi.org/10.1016/j.jenvman.2020.111894

[50] Wang, Li, Chen, Jun, Zhang, Wei, Liu, Peng, & Huang, Yifan. (2024). Climate change impacts on magnitude and frequency of urban floods under scenario and model uncertainties. Journal of Environmental Management, 366, 121679. https://doi.org/10.1016/j.jenvman.2024.121679

[51] Khoshkonesh, Amir, Nazari, Reza, Nikoo, Mohammad Reza, & Karimi, Mohsen. (2024). Enhancing flood risk assessment in urban areas by integrating hydrodynamic models and machine learning techniques. Science of the Total Environment, 952, 175859. https://doi.org/10.1016/j.scitotenv.2024.175859

[52] Saksena, Sanjeev, Dey, Sayan, Merwade, Venkatesh, & Singhofen, Patrick J. (2020). A computationally efficient and physically based approach for urban flood modeling using a flexible spatiotemporal structure. Water Resources Research, 56(1), e2019WR025769. https://doi.org/10.1029/2019WR025769

[53] Brunner, Manuela I., Viviroli, Daniel, Sikorska, Anna E., & Seibert, Jan. (2021). Flood spatial coherence, triggers, and performance in hydrological simulations: A large-sample evaluation of four streamflow-calibrated models. Hydrology and Earth System Sciences, 25(1), 105–119. https://doi.org/10.5194/hess-25-105-2021

[54] Alperen, Caglar I., Artigue, Guillaume, Kurtulus, Burak, Pistre, Stephane, & Johannet, Alain. (2021). A hydrological digital twin using artificial neural networks for flood simulation in the Gardon de Sainte-Croix basin, France. IOP Conference Series: Earth and Environmental Science, 906(1), 012112. https://doi.org/10.1088/1755-1315/906/1/012112

[55] Panahi, Mohammad, Gayen, Anindita, Pourghasemi, Hamid Reza, Rezaie, Fariborz, & Lee, Sang-Eun. (2021). Deep learning neural networks for spatially explicit prediction of flash flood probability. Geoscience Frontiers, 12(3), 101076. https://doi.org/10.1016/j.gsf.2020.09.007

[56] Vongkusolkit, Jirapon, Peng, Bowen, Wu, Mengdi, Huang, Qiang, & Andresen, Cameron G. (2023). Near real-time flood mapping with weakly supervised machine learning. Remote Sensing, 15(13), 3263. https://doi.org/10.3390/rs15133263

[57] Asif, Muhammad, Kuglitsch, Michael M., Pelivan, Ivan, & Albano, Roberto. (2025). Review and intercomparison of machine learning applications for short-term flood forecasting. Water Resources Management, 39(5), 1971–1991. https://doi.org/10.1007/s11269-025-04093-x

[58] Anik, Mohammad S. B. M., An, Chen, & Li, Sheng S. (2025). Evolution from physical process-based approaches to machine learning approaches for predicting urban floods: A literature review. Environmental Systems Research, 14(1), Article 409. https://doi.org/10.1186/s40068-025-00409-3

[59] Islam, Tanjina, Zeleke, Ebrahim B., Afroz, Meher, & Melesse, Assefa M. (2025). A systematic review of urban flood susceptibility mapping: Remote sensing, machine learning, and other modeling approaches. Remote Sensing, 17(3), 524. https://doi.org/10.3390/rs17030524

[60] Byaruhanga, Nicholas, Kibirige, Daniel, Gokool, Sharmila, & Mkhonta, Gideon. (2024). Evolution of flood prediction and forecasting models for flood early warning systems: A scoping review. Water, 16(13), 1763. https://doi.org/10.3390/w16131763

[61] Liu, Zhe, Coleman, Nicholas, Patrascu, Florin I., Yin, Ke, Li, Xin, & Mostafavi, Ali. (2025). Artificial intelligence for flood risk management: A comprehensive state-of-the-art review and future directions. International Journal of Disaster Risk Reduction, 117, 105110. https://doi.org/10.1016/j.ijdrr.2024.105110

[62] Bentivoglio, Roberto, Isufi, Elvin, Jonkman, Sebastiaan N., & Taormina, Riccardo. (2022). Deep learning methods for flood mapping: A review of existing applications and future research directions. Hydrology and Earth System Sciences, 26(16), 4345–4378. https://doi.org/10.5194/hess-26-4345-2022

[63] Díez-Herrero, Andrés, & Garrote, Javier. (2020). Flood risk analysis and assessment, applications and uncertainties: A bibliometric review. Water, 12(7), 2050. https://doi.org/10.3390/w12072050

[64] Ahmad, Hamza, Kenneth, Isaac I., Emmanuel, Andrew E., Aniekan, Abraham U., & Valentine, Isaac I. (2024). Artificial intelligence and machine learning in climate change research: A review of predictive models and environmental impact. World Journal of Advanced Research and Reviews, 21(1), 1999–2008. https://doi.org/10.30574/wjarr.2024.21.1.0257

[65] Puttinaovarat, Somsak, & Horkaew, Pichai. (2020). Flood forecasting system based on integrated big data and crowdsourced data using machine learning techniques. IEEE Access, 8, 5885–5905. https://doi.org/10.1109/ACCESS.2019.2963819

[66] Ke, Qiang, Zhang, Yifan, Chen, Jiahui, Wang, Rui, & Liu, Xiaoyu. (2020). Urban pluvial flooding prediction using machine learning approaches: A case study of Shenzhen City, China. Advances in Water Resources, 145, 103719. https://doi.org/10.1016/j.advwatres.2020.103719

[67] Nguyen, Duc T., & Chen, Shie-Yui. (2020). Real-time probabilistic flood forecasting using multiple machine learning methods. Water, 12(3), 787. https://doi.org/10.3390/w12030787

[68] Zahura, Fariha T., Goodall, Jonathan L., Sadler, Jeffrey M., Shen, Yaling, Morsy, Mahmoud M., & Behl, Madhur. (2020). Training machine learning surrogate models from a high-fidelity physics-based model: Application for real-time street-scale flood prediction in an urban coastal community. Water Resources Research, 56(10), e2019WR027038. https://doi.org/10.1029/2019WR027038

[69] Yan, Xiaojun, Xu, Kai, Feng, Wei, & Chen, Jie. (2021). A rapid prediction model of urban flood inundation in a high-risk area coupling machine learning and numerical simulation approaches. International Journal of Disaster Risk Science, 12(6), 903–918. https://doi.org/10.1007/s13753-021-00384-0

[70] Zhou, Rui, & Pan, Yuchen. (2022). FloodDAN: Unsupervised flood forecasting based on adversarial domain adaptation. In Proceedings of the 5th International Conference on Big Data and Artificial Intelligence (BDAI 2022) (pp. 6–12). IEEE. https://doi.org/10.1109/BDAI56143.2022.9862723

[71] Sampurno, Joko, Vallaeys, Vincent, Ardianto, Rachmat, & Hanert, Emmanuel. (2022). Integrated hydrodynamic and machine learning models for compound flooding prediction in a data-scarce estuarine delta. Nonlinear Processes in Geophysics, 29(3), 301–315. https://doi.org/10.5194/npg-29-301-2022

[72] Aswad, Firas M., Kareem, Ahmed N., Khudhur, Ahmed M., Khalaf, Bassam A., & Mostafa, Salama A. (2021). Tree-based machine learning algorithms in the Internet of Things environment for multivariate flood status prediction. Journal of Intelligent Systems, 31(1), 1–14. https://doi.org/10.1515/jisys-2021-0179

[73] Rathnayake, Nadeesha, Rathnayake, Upaka, Chathuranika, Ishara, Dang, Thi L., & Hoshino, Yuki. (2023). Projected water levels and identified future floods: A comparative analysis for the Mahaweli River, Sri Lanka. IEEE Access, 11, 8920–8937. https://doi.org/10.1109/ACCESS.2023.3238717

[74] Zhang, Rui, Kim, Hyungjun, Lien, Eric, Zheng, Dongmei, Band, Lawrence E., & Lakshmi, Venkat. (2021). Deep learning approach to predict peak floods and evaluate socioeconomic vulnerability to flood events: A case study in Baltimore, Maryland, USA. In Proceedings of the IEEE Systems and Information Engineering Design Symposium (SIEDS 2021) (pp. 1–6). IEEE. https://doi.org/10.1109/SIEDS52267.2021.9483782

[75] Nguyen, Duy H., & Bae, Dong-Hyuk. (2020). Correcting mean areal precipitation forecasts to improve urban flooding predictions using a long short-term memory network. Journal of Hydrology, 584, 124710. https://doi.org/10.1016/j.jhydrol.2020.124710

[76] Hajji, Soufiane, Benaabidate, Lahcen, El Mendili, Youssef, & Ouazar, Driss. (2025). Enhancing flood prediction through remote sensing, machine learning, and Google Earth Engine. Frontiers in Water, 7, 1514047. https://doi.org/10.3389/frwa.2025.1514047

[77] Ahmmed, Md. Rashed, Monir, Jasim, & Khushbu, Syeda A. (2022). Analysis of flood risk prediction using different machine learning classifiers: A study of rural areas in Bangladesh. In Proceedings of the 13th International Conference on Computing Communication and Networking Technologies (ICCCNT 2022) (pp. 1–6). IEEE. https://doi.org/10.1109/ICCCNT54827.2022.9984449

[78] Wang, Yu, Zhang, Peng, Xie, Yu, Chen, Lin, & Li, Yong. (2025). Toward explainable flood risk prediction: Integrating a novel hybrid machine learning model. Sustainable Cities and Society, 120, 106140. https://doi.org/10.1016/j.scs.2025.106140

[79] Miau, Shu-Yu, & Hung, Wei-Hsuan. (2020). River flooding forecasting and anomaly detection based on deep learning. IEEE Access, 8, 198384–198402. https://doi.org/10.1109/ACCESS.2020.3034875

[80] Gude, Vishal, Corns, Steven, & Long, Stephen. (2020). Flood prediction and uncertainty estimation using deep learning. Water, 12(3), 884. https://doi.org/10.3390/w12030884

[81] Wu, Zhen, Zhou, Yifan, Wang, Hongtao, & Jiang, Zhenhua. (2020). Depth prediction of urban flood under different rainfall return periods based on deep learning and data warehouse. Science of the Total Environment, 716, 137077. https://doi.org/10.1016/j.scitotenv.2020.137077

[82] Khosravi, Kamran, Pham, Binh Thai, Chapi, Kamran, Shirzadi, Amin, Shahabi, Himan, Revhaug, Ivar, & Pradhan, Biswajeet. (2020). Convolutional neural network approach for spatial prediction of flood hazard at the national scale of Iran. Journal of Hydrology, 591, 125552. https://doi.org/10.1016/j.jhydrol.2020.125552

[83] Chen, Chen, Jiang, Jianshi, Liao, Zongjian, Zhou, Yifan, Wang, Hongtao, & Pei, Qiang. (2022). A short-term flood prediction based on a spatial deep learning network: A case study for Xi County, China. Journal of Hydrology, 607, 127535. https://doi.org/10.1016/j.jhydrol.2022.127535

[84] Wang, Qiang. (2022). Machine learning model design for IoT-based flooding forecast. In Proceedings of the 2022 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC 2022) (pp. 97–103). IEEE. https://doi.org/10.1109/CyberC55534.2022.00025

[85] Li, Wei, Zhang, Hao, Chen, Xu, Liu, Jian, & Wang, Yifan. (2024). An interpretable hybrid deep learning model for flood forecasting based on Transformer and LSTM. Journal of Hydrology: Regional Studies, 54, 101873. https://doi.org/10.1016/j.ejrh.2024.101873

[86] Hosseiny, Hadi. (2021). A deep learning model for predicting river flood depth and extent. Environmental Modelling and Software, 145, 105186. https://doi.org/10.1016/j.envsoft.2021.105186

[87] Chen, Chen, Hui, Qiang, Xie, Wenjun, Wan, Shunming, Zhou, Yifan, & Pei, Qiang. (2021). Convolutional neural networks for forecasting flood processes in Internet-of-Things-enabled smart cities. Computer Networks, 186, 107744. https://doi.org/10.1016/j.comnet.2020.107744

[88] Cheng, Ming, Fang, Fangxing, Kinouchi, Tetsuya, Navon, Isaac M., & Pain, Christopher C. (2020). Long lead-time daily and monthly streamflow forecasting using machine learning methods. Journal of Hydrology, 590, 125376. https://doi.org/10.1016/j.jhydrol.2020.125376

[89] Dubey, Vinay, & Katarya, Rahul. (2025). FloodCNN-BiLSTM: Predicting flood events in urban environments. Engineering Analysis with Boundary Elements, 177, 106277. https://doi.org/10.1016/j.enganabound.2025.106277

[90] Zhong, Min, Zhang, Yifan, Li, Jun, Wang, Rui, & Kim, Sunmin. (2023). A hybrid model combining the CaMa-Flood model and deep learning methods for streamflow prediction. Water Resources Management, 37(12), 4841–4859. https://doi.org/10.1007/s11269-023-03583-0

[91] Solanki, Hitesh, Vegad, Uday, Kushwaha, Ankit, & Mishra, Vikram. (2025). Improving streamflow prediction using multiple hydrological models and machine learning methods. Water Resources Research, 61(1), e2024WR038192. https://doi.org/10.1029/2024WR038192

[92] Waseemullah, Hassan, Tariq, Majeed, Sajid, & Memon, Muhammad Qasim. (2024). Urban pluvial flood prediction using machine learning models. In Proceedings of the 4th International Conference on Innovations in Computer Science (ICONICS 2024) (pp. 1–6). IEEE. https://doi.org/10.1109/ICONICS64289.2024.10824627

[93] Xu, Lei, & Gao, Lei. (2024). A hybrid surrogate model for real-time coastal urban flood prediction: An application to Macao. Journal of Hydrology, 642, 131863. https://doi.org/10.1016/j.jhydrol.2024.131863

[94] Fereshtehpour, Mohsen, Najafi, Mohammad Reza, & Cannon, Alex J. (2025). Characterizing compound inland flooding mechanisms and risks in North America under climate change. Earth's Future, 13(2), e2024EF005353. https://doi.org/10.1029/2024EF005353

[95] Ming, Xiaoqing, Liang, Qiuhua, Dawson, Richard, Xia, Xiaoyu, & Hou, Jun. (2022). A quantitative multi-hazard risk assessment framework for compound flooding considering hazard interdependencies and interactions. Journal of Hydrology, 607, 127477. https://doi.org/10.1016/j.jhydrol.2022.127477

[96] Hsiao, Shih-Chieh, Hsu, Ting-Wei, Lin, Yi-Hsin, Wu, Wen-Cheng, & Chen, Wei-Hao. (2021). Flood risk influenced by the compound effect of storm surge and rainfall under climate change for low-lying coastal areas. Science of the Total Environment, 764, 144439. https://doi.org/10.1016/j.scitotenv.2020.144439

[97] Piadeh, Farhad, Behzadian, Kourosh, & Alani, Amir M. (2022). A critical review of real-time modeling of flood forecasting in urban drainage systems. Journal of Hydrology, 607, 127476. https://doi.org/10.1016/j.jhydrol.2022.127476

[98] Song, Tao, Ding, Wei, Wu, Jie, Liu, Hong, Zhou, Hao, & Chu, Jian. (2020). Flash flood forecasting based on long short-term memory networks. Water, 12(1), 109. https://doi.org/10.3390/w12010109

[99] Patil, Kaustubh R., Doi, Toshio, & Behera, Swadhin K. (2023). Predicting extreme floods and droughts in East Africa using a deep learning approach. npj Climate and Atmospheric Science, 6(1), Article 35. https://doi.org/10.1038/s41612-023-00435-x

[100] Qian, Kun, Mohamed, Ahmed, & Claudel, Christian. (2019). Physics-informed data-driven model for flood prediction: Application of deep learning in prediction of urban flood development. arXiv. http://arxiv.org/abs/1908.10312

[101] Prasad, Pradeep, Loveson, Victor J., Das, Bimal, & Kotha, Mahesh. (2022). Novel ensemble machine learning models in flood susceptibility mapping. Geocarto International, 37(16), 4571–4593. https://doi.org/10.1080/10106049.2021.1892209

[102] Ilia, Ioannis, Tsangaratos, Paraskevas, Tzampoglou, Panagiotis, Chen, Wei, & Hong, Haoyuan. (2022). Flash flood susceptibility mapping using stacking ensemble machine learning models. Geocarto International, 37(27), 15010–15036. https://doi.org/10.1080/10106049.2022.2093990

[103] Le, Xuan-Hien, Nguyen, Duc-Hieu, Jung, Sungmin, Yeon, Minjae, & Lee, Gi-Hoon. (2021). Comparison of deep learning techniques for river streamflow forecasting. IEEE Access, 9, 71805–71820. https://doi.org/10.1109/ACCESS.2021.3077703

[104] Rahimzad, Mohammad, Moghaddam Nia, Amir, Zolfonoon, Hamed, Soltani, Javad, Danandeh Mehr, Ali, & Kwon, Hyun-Han. (2021). Performance comparison of an LSTM-based deep learning model versus conventional machine learning algorithms for streamflow forecasting. Water Resources Management, 35(12), 4167–4187. https://doi.org/10.1007/s11269-021-02937-w

[105] Xu, Yifan, Zhang, Hao, Wang, Rui, Chen, Jiahui, & Kim, Sunmin. (2023). Deep transfer learning based on Transformer for flood forecasting in data-sparse basins. Journal of Hydrology, 625, 129956. https://doi.org/10.1016/j.jhydrol.2023.129956

[106] Abbasi, Amir, Kambali, Prashanth N., Shahidi, Parisa, & Nataraj, C. (2024). Physics-informed machine learning for modeling multidimensional dynamics. Nonlinear Dynamics, 112(24), 21565–21585. https://doi.org/10.1007/s11071-024-10163-3

[107] Imhoff, Rolf O., Brauer, Claudia C., van Heeringen, Koen J., Uijlenhoet, Remko, & Weerts, Albrecht H. (2022). Large-sample evaluation of radar rainfall nowcasting for flood early warning. Water Resources Research, 58(3), e2021WR031591. https://doi.org/10.1029/2021WR031591

[108] Kratzert, Frederik, Klotz, Daniel, Shalev, Gal, Klambauer, Günter, Hochreiter, Sepp, & Nearing, Grey. (2019). Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. Hydrology and Earth System Sciences, 23(12), 5089–5110. https://doi.org/10.5194/hess-23-5089-2019

[109] Hussain, Muhammad A., Chen, Zhi, Zheng, Yu, Zhou, Yang, & Daud, Hafiz. (2023). Deep learning and machine learning models for landslide susceptibility mapping with remote sensing data. Remote Sensing, 15(19), 4703. https://doi.org/10.3390/rs15194703

[110] Hamill, Thomas M., Whitaker, Jeffrey S., Bates, Gary T., & McLay, John G. (2022). The reanalysis for the Global Ensemble Forecast System, version 12. Monthly Weather Review, 150(1), 59–79. https://doi.org/10.1175/MWR-D-21-0023.1

[111] Zhou, Xiaolin, Mou, Lei, Ao, Tianyu, Huang, Xin, & Yang, Hailong. (2025). Effect of the spatial resolution of digital terrain data obtained by drone on urban fluvial flood modeling of mountainous regions. Hydrology and Earth System Sciences, 29(7), 1963–1980. https://doi.org/10.5194/hess-29-1963-2025

[112] Hwang, Jackelyn, & Naik, Nikhil. (2023). Systematic social observation at scale: Using crowdsourcing and computer vision to measure visible neighborhood conditions. Sociological Methodology, 53(2), 183–216. https://doi.org/10.1177/00811750231160781

[113] Ahmadi, Seyed Mohammad, Balahang, Sina, & Abolfathi, Sam. (2024). Predicting the hydraulic response of critical transport infrastructures during extreme flood events. Engineering Applications of Artificial Intelligence, 133, 108573. https://doi.org/10.1016/j.engappai.2024.108573

[114] Qi, Wenjing, Ma, Chao, Xu, Hong, Chen, Zhiqiang, Zhao, Kun, & Han, Haijun. (2021). A review on applications of urban flood models in flood mitigation strategies. Natural Hazards, 108(1), 31–62. https://doi.org/10.1007/s11069-021-04715-8

[115] Guerriero, Luca, Ruzza, Giorgio, Guadagno, Fausto M., & Revellino, Paola. (2020). Flood hazard mapping incorporating multiple probability models. Journal of Hydrology, 587, 125020. https://doi.org/10.1016/j.jhydrol.2020.125020

[116] Aderyani, Fatemeh R., Mousavi, Seyed Jamshid, & Jafari, Farzad. (2022). Short-term rainfall forecasting using machine learning-based approaches of PSO-SVR, LSTM, and CNN. Journal of Hydrology, 614, 128463. https://doi.org/10.1016/j.jhydrol.2022.128463

[117] Huang, Huan, Zhang, Hao, Liu, Jian, Chen, Xu, & Wang, Yifan. (2024). Improving the explainability of CNN-LSTM-based flood prediction by integrating SHAP technique. Ecological Informatics, 84, 102904. https://doi.org/10.1016/j.ecoinf.2024.102904

[118] Li, Li, & Jun, Kyung S. (2022). A hybrid approach to improve flood forecasting by combining a hydrodynamic flow model and artificial neural networks. Water, 14(9), 1393. https://doi.org/10.3390/w14091393

[119] Kordani, Mohammad, Nikoo, Mohammad Reza, Fooladi, Mohammad, Ahmadianfar, Iman, Nazari, Reza, & Gandomi, Amir H. (2024). Improving long-term flood forecasting accuracy using ensemble deep learning models and an attention mechanism. Journal of Hydrologic Engineering, 29(6), 04024019. https://doi.org/10.1061/JHYEFF.HEENG-6262

[120] Piadeh, Farhad, Behzadian, Kourosh, Chen, Albert S., Kapelan, Zoran, Rizzuto, Joseph P., & Campos, Luis C. (2023). Enhancing urban flood forecasting in drainage systems using dynamic ensemble-based data mining. Water Research, 247, 120791. https://doi.org/10.1016/j.watres.2023.120791

[121] Rahman, Md. Mostafizur, Jahan, Shamsul, Islam, Md. Rashedul, & Ahmed, Md. Shafiqul. (2019). Flood susceptibility assessment in Bangladesh using machine learning and multi-criteria decision analysis. Earth Systems and Environment, 3(3), 585–601. https://doi.org/10.1007/s41748-019-00123-y

[122] Xu, Chao, Li, Wei, Zhang, Hao, Chen, Xu, Liu, Jian, & Wang, Yifan. (2024). A hybrid model coupling process-driven and data-driven models for improved real-time flood forecasting. Journal of Hydrology, 638, 131494. https://doi.org/10.1016/j.jhydrol.2024.131494

[123] Zhao, Yifan, Liu, Jian, Chen, Xu, Zhang, Hao, Wang, Rui, & Li, Wei. (2025). A deep learning-based probabilistic approach to flash flood warnings in mountainous catchments. Journal of Hydrology, 652, 132677. https://doi.org/10.1016/j.jhydrol.2025.132677

[124] Al-Areeq, Ahmed M., Arabameri, Alireza, Pradhan, Biswajeet, Pourghasemi, Hamid Reza, & Lee, Sang-Eun. (2023). Flood subsidence susceptibility mapping using elastic-net classifier: A new approach. Water Resources Management, 37(13), 4985–5006. https://doi.org/10.1007/s11269-023-03591-0

[125] Bui, Dieu Tien, Tsangaratos, Paraskevas, Nguyen, Van Phong, Pradhan, Biswajeet, & Dou, Jianping. (2020). A novel deep learning neural network approach for predicting flash flood susceptibility: A case study in a high-frequency tropical storm area. Science of the Total Environment, 701, 134413. https://doi.org/10.1016/j.scitotenv.2019.134413

[126] Wang, Qiang, & Abdelrahman, Wael. (2023). High-precision AI-enabled flood prediction integrating local sensor data and third-party weather forecasts. Sensors, 23(6), 3065. https://doi.org/10.3390/s23063065

[127] Farooq, Muhammad S., Ahmad, Zahid, Khan, Asad, & Jeong, Jong-Myon. (2023). FFM: Flood forecasting model using federated learning. IEEE Access, 11, 24472–24483. https://doi.org/10.1109/ACCESS.2023.3252896

[128] Guo, Zhaoxiang, Leitão, João P., Simões, Nuno E., & Moosavi, Vahid. (2021). Data-driven flood emulation: Speeding up urban flood predictions by deep convolutional neural networks. Journal of Flood Risk Management, 14(1), e12684. https://doi.org/10.1111/jfr3.12684

[129] Zhou, Xiaolin, Huang, Xin, Jiang, Xinyu, & Jiang, Jie. (2025). Real-time error correction of multi-hour-ahead flash flood forecasting based on sliding runoff–rain data and deep learning models. Journal of Hydrology, 655, 132918. https://doi.org/10.1016/j.jhydrol.2025.132918

[130] Waleed, Muhammad, & Sajjad, Muhammad. (2025). Advancing flood susceptibility prediction: A comparative assessment and scalability analysis of machine learning algorithms in high-risk regions of Pakistan. Journal of Flood Risk Management, 18(1), e13047. https://doi.org/10.1111/jfr3.13047

[131] Palash, Wasif, Akanda, Ali S., & Islam, Sayed. (2024). A data-driven global flood forecasting system for medium to large rivers. Scientific Reports, 14(1), 8979. https://doi.org/10.1038/s41598-024-59145-w

[132] Wang, Shanshan, & Wang, Jian. (2022). Research on prediction model of mountain flood level in small watersheds based on deep learning. In Proceedings of the 4th International Conference on Intelligent Control, Measurement and Signal Processing (ICMSP 2022) (pp. 1024–1027). IEEE. https://doi.org/10.1109/ICMSP55950.2022.9859047

[133] Mohamadiazar, Narges, Ebrahimian, Abbas, & Hosseiny, Hadi. (2024). Integrating deep learning, satellite image processing, and spatiotemporal analysis for urban flood prediction. Journal of Hydrology, 639, 131508. https://doi.org/10.1016/j.jhydrol.2024.131508

[134] Wan, Han-Hui, Zhang, Hao, Liu, Jian, Chen, Xu, & Wang, Yifan. (2025). An interpretable and scalable model for rapid flood extent forecasting using satellite imagery and machine learning with rotated EOF analysis. Environmental Modelling and Software, 192, 106562. https://doi.org/10.1016/j.envsoft.2025.106562

[135] Maspo, Nur Afiqah, Harun, Ahmad N. Bin, Goto, Masaru, Cheros, Farah, Haron, Nur Aina, & Mohd Nawi, Mohd Nizam. (2020). Evaluation of machine learning approaches in flood prediction scenarios and input parameters: A systematic review. IOP Conference Series: Earth and Environmental Science, 479(1), 012038. https://doi.org/10.1088/1755-1315/479/1/012038

[136] Karyotis, Christos, Maniak, Thomas, Doctor, Faisal, Iqbal, Rami, Palade, Vasile, & Tang, Rong. (2019). Deep learning for flood forecasting and monitoring in urban environments. In Proceedings of the 18th IEEE International Conference on Machine Learning and Applications (ICMLA 2019) (pp. 1392–1397). IEEE. https://doi.org/10.1109/ICMLA.2019.00227

[137] Xiang, Xiaolong, Guo, Shenglian, Li, Chao, Sun, Bo, & Liang, Zhijun. (2025). Deep learning model for flood probabilistic forecasting considering spatiotemporal rainfall distribution and hydrologic uncertainty. Journal of Hydrology, 655, 132879. https://doi.org/10.1016/j.jhydrol.2025.132879

[138] Srinu, Nandamuri, & Bindu, Bhavani H. (2022). A review on machine learning and deep learning-based rainfall prediction methods. In Proceedings of the 3rd International Conference on Power, Energy, Control and Transmission Systems (ICPECTS 2022) (pp. 1–4). IEEE. https://doi.org/10.1109/ICPECTS56089.2022.10047554

[139] Adeyemi, Adebowale B., & Komolafe, Adewale A. (2025). Flood hazard zone prediction using a machine-learning-based geospatial approach in the lower Niger River basin, Nigeria. Natural Hazards Research, 5(2), 399–412. https://doi.org/10.1016/j.nhres.2025.01.002

[140] Costache, Radu, Arabameri, Alireza, Pourghasemi, Hamid Reza, Pradhan, Biswajeet, & Lee, Sang-Eun. (2022). Flash-flood hazard assessment using deep learning based on the H2O R package and fuzzy multi-criteria decision-making analysis. Journal of Hydrology, 609, 127747. https://doi.org/10.1016/j.jhydrol.2022.127747

[141] Tools, P. (2023). Know to predict, forecast to warn: A review of flood risk.

[142] Werner de Vargas, Victor, Schneider Aranda, José A., dos Santos Costa, Rodrigo, da Silva Pereira, Paulo R., & Victória Barbosa, Jorge L. (2023). Imbalanced data preprocessing techniques for machine learning: A systematic mapping study. Knowledge and Information Systems, 65(1), 31–57. https://doi.org/10.1007/s10115-022-01772-8

[143] Kunft, Alexander, Katsifodimos, Anastasios, Schelter, Sebastian, Breß, Sebastian, Rabl, Tilmann, & Markl, Volker. (2019). An intermediate representation for optimizing machine learning pipelines. Proceedings of the VLDB Endowment, 12(11), 1553–1567. https://doi.org/10.14778/3342263.3342633

[144] Tabbussum, Rida, & Dar, A. Q. (2021). Performance evaluation of artificial intelligence paradigms—Artificial neural networks, fuzzy logic, and adaptive neuro-fuzzy inference system for flood prediction. Environmental Science and Pollution Research, 28(20), 25265–25282. https://doi.org/10.1007/s11356-021-12410-1

[145] Mahdizadeh Gharakhanlou, Narges, & Perez, Laura. (2023). Flood susceptible prediction through the use of geospatial variables and machine learning methods. Journal of Hydrology, 617, 129121. https://doi.org/10.1016/j.jhydrol.2023.129121

[146] Arora, Aman, Pourghasemi, Hamid Reza, Pradhan, Biswajeet, & Lee, Sang-Eun. (2021). Optimization of state-of-the-art fuzzy–metaheuristic ANFIS-based machine learning models for flood susceptibility prediction mapping in the Middle Ganga Plain, India. Science of the Total Environment, 750, 141565. https://doi.org/10.1016/j.scitotenv.2020.141565

[147] Pianforini, Matteo, Dazzi, Stefano, Pilzer, Andrea, & Vacondio, Renato. (2024). Real-time flood map forecasting for dam-break scenarios using a transformer-based deep learning model. Journal of Hydrology, 635, 131169. https://doi.org/10.1016/j.jhydrol.2024.131169

[148] Mahadware, Amol, Saigiridhari, Aditya, Mishra, Anurag, Tupe, Amol, & Marathe, Nilesh. (2022). Rainfall prediction using different machine learning and deep learning algorithms. In Proceedings of the 2nd Asian Conference on Innovation and Technology (ASIANCON 2022) (pp. 1–8). IEEE. https://doi.org/10.1109/ASIANCON55314.2022.9908857

[149] Hu, Ruiming, Fang, Fangxing, Pain, Christopher C., & Navon, Isaac M. (2019). Rapid spatiotemporal flood prediction and uncertainty quantification using a deep learning method. Journal of Hydrology, 575, 911–920. https://doi.org/10.1016/j.jhydrol.2019.05.087

[150] Moradian, Shima, AghaKouchak, Amir, Gharbia, Sam, Broderick, Conor, & Olbert, A. I. (2024). Forecasting compound ocean–fluvial floods using machine learning. Journal of Environmental Management, 364, 121295. https://doi.org/10.1016/j.jenvman.2024.121295

[151] Chang, Li-Chiu, Yang, Ming-Tsang, & Chang, Fi-John. (2025). Flood resilience through hybrid deep learning: Advanced forecasting for Taipei's urban drainage system. Journal of Environmental Management, 379, 124835. https://doi.org/10.1016/j.jenvman.2025.124835

[152] Aderyani, Fatemeh R., Jafarzadegan, Kamran, & Moradkhani, Hamid. (2025). A surrogate machine learning modeling approach for enhancing the efficiency of urban flood modeling at metropolitan scales. Sustainable Cities and Society, 123, 106277. https://doi.org/10.1016/j.scs.2025.106277

[153] Yang, Fei, Ding, Wei, Zhao, Jun, Song, Lei, Yang, Dong, & Li, Xia. (2024). Rapid urban flood inundation forecasting using a physics-informed deep learning approach. Journal of Hydrology, 643, 131998. https://doi.org/10.1016/j.jhydrol.2024.131998

[154] Sit, Matthew, & Demir, Ibrahim. (2019). Decentralized flood forecasting using deep neural networks. arXiv. http://arxiv.org/abs/1902.02308

[155] Hu, Chao, Wu, Qiang, Li, Hong, Jian, Shuo, Li, Ning, & Lou, Zhen. (2018). Deep learning with long short-term memory networks for rainfall–runoff simulation. Water, 10(11), 1543. https://doi.org/10.3390/w10111543

[156] Kumar, Vinay, Azamathulla, H. M., Sharma, K. V., Mehta, D. J., & Maharaj, K. T. (2023). The state of the art in deep learning applications, challenges, and future prospects: A comprehensive review of flood forecasting and management. Sustainability, 15(13), 10543. https://doi.org/10.3390/su151310543

[157] Adnan, Md. Shahriar G., Pourghasemi, Hamid Reza, Pradhan, Biswajeet, & Lee, Sang-Eun. (2023). A novel framework for addressing uncertainties in machine learning-based geospatial approaches for flood prediction. Journal of Environmental Management, 326, 116813. https://doi.org/10.1016/j.jenvman.2022.116813

[158] Brunner, Manuela I., Slater, Louise, Tallaksen, Lena M., & Clark, Martyn. (2021). Challenges in modeling and predicting floods and droughts: A review. WIREs Water, 8(3), e1520. https://doi.org/10.1002/wat2.1520

[159] Alawad, M., & Ali, B. S. (2025). Leveraging ChatGPT in the loop for enhanced robustness in deep learning models. Information Technology and Cybersecurity, 2(2), 49–53.

[160] Abid, S. K., & Sulaiman, N. (2025). Revolutionizing the future: unleashing the potential of generative AI ChatGPT in disaster management. Environmental Science and Pollution Research, 1-16.