Earth Science Informatics (2025) 18:416
https://doi.org/10.1007/s12145-025-01911-z

REVIEW

Evolution of ensemble machine learning approaches in water resources management: a review

Moein Tosan[1] · Vahid Nourani[2,3] · Ozgur Kisi[4,5] · Mehdi Dastourani[1]

Received: 1 March 2025 / Accepted: 7 May 2025 / Published online: 24 May 2025

Abstract
Ensemble Machine Learning (EML) techniques have markedly advanced in hydrological modeling over the past decade, significantly enhancing predictive accuracy for complex tasks such as streamflow prediction, flood forecasting, and groundwater estimation. This study presents a bibliometric analysis of 199 articles and a systematic review of 51 peer-reviewed articles, published from 2017 to December 2024. The bibliometric findings indicate a surge in EML adoption post-2018, driven by increasing demands for precise hydrological models amidst climate change and hydrological variability. The review categorizes EML strategies into Boosting, Bagging, and Stacking methodologies. Boosting methods—particularly Gradient Boosting Machines (GBM), Extreme Gradient Boosting (XGBoost), and LightGBM—were prominently featured in the reviewed studies, often noted for their capability in improved generalization and non-linearity handling. Stacking models, which integrate algorithms like Random Forest (RF), Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM) networks, were also frequently applied and reported as effective in managing the complexities of hydrological systems. Bagging methods, notably RF, were found effective in stabilizing performance in noisy datasets. Additionally, this study highlights the critical role of data preprocessing techniques such as Singular Spectrum Analysis (SSA), Recursive Feature Elimination (RFE), and Principal Component Analysis (PCA) in optimizing model performance and addressing data heterogeneity challenges. Approximately 60% of the systematic review studies focus on streamflow prediction; among these, hybrid approaches—both combining multiple ML algorithms (ML-ML) and coupling ML with optimization/statistical methods (ML-AO)—were frequently utilized and reported as well-suited for capturing non-linear hydrological dynamics. Despite significant advancements, model interpretability, scalability, and multi-source data integration challenges persist. This paper underscores the need for further research to enhance EML frameworks, mainly through integrating physics-based models and advancements in computational power. The findings provide critical insights into the current state of EML research in hydrology, highlighting essential areas for future exploration amidst growing environmental uncertainties.

Keywords Data-driven Models · Forecasting · Hybrid Models · Hydrological Modeling · Machine Learning · Meta-Analysis

Communicated by Hassan Babaie.

✉ Vahid Nourani
vnourani@yahoo.com

Moein Tosan
moein.tosan@birjand.ac.ir

Ozgur Kisi
ozgur.kisi@th-luebeck.de

Mehdi Dastourani
mdastourani@birjand.ac.ir

[1] Faculty of Water Sciences and Engineering, University of Birjand, Birjand, Iran

Abbreviations

AdaBoost	Adaptive boosting
LSTM	Long short-term memory

[2] Center of Excellence in Hydroinformatics, Faculty of Civil Engineering, University of Tabriz, Tabriz 5166616471, Iran

[3] World Peace University, Sht. Kemal Ali Omer Sok. Via Mersin 10, Nicosia, Turkey

[4] Department of Civil Engineering, Technical University of Lübeck, 23562 Lübeck, Germany

[5] Department of Civil Engineering, Ilia State University, 0162 Tbilisi, Georgia

ANN	Artificial neural network
MAE	Mean absolute error
ARIMA	Autoregressive integrated moving average
MARS	Multivariate adaptive regression splines
BMASE	Best model averaging super ensemble
MIC	Maximum mutual information
BOA	Bayesian optimization algorithm
ML	Machine learning
BTOP	Block-wise use of TOPMODEL
MLP	Multilayer perceptron
CART	Classification and regression trees
MMFM	Multiple-mode forecasting model
CMIP6	Coupled model intercomparison project phase 6
MS-EnsPost	Multi-statistic ensemble post-processing
CNN	Convolutional neural network
NN	Neural network(s)
Dagging	Dataset aggregation
PCA	Principal component analysis
DEM	Digital elevation model
PCR-GLOBWB	PCRaster global water balance
DT	Decision tree
PCSWMM	Personal computer storm water management model
ECMWF	European centre for medium-range weather forecasts
PRISMA	Preferred reporting items for systematic reviews and meta-analysis
EML	Ensemble machine learning
PSO	Particle swarm optimization
EnsPost	Ensemble post-processing
RF	Random forest
ETRSE	Extra tree regression super ensemble
RFE	Recursive feature elimination
FURIA	Fuzzy unordered rule induction algorithm
RMSE	Root mean square error
GA	Genetic algorithm
RQ	Research question
GAN	Generative adversarial network
SEL	Super ensemble learning
GBM	Gradient boosting machine
SHAP	SHapley additive explanations
GIS	Geographic information system
SMFM	Single-mode forecasting model
GMM	Gaussian mixture model
SPHY	Spatial processes in hydrology
GPR	Gaussian process regression
SSA	Singular spectrum analysis
GRU	Gated recurrent unit
SSP	Shared socioeconomic pathway
HBV	Hydrologiska byråns vattenbalansavdelning
SVM	Support vector machine
IoT	Internet of things
TOPMODEL	TOPography-based hydrological model
KGE	Kling-gupta efficiency
WASE	Weighted average super ensemble
KNN	K-nearest neighbors
WoS	Web of science
LIME	Local interpretable model-agnostic explanations
XGBoost	Extreme gradient boosting

Introduction

Hydrology, the science dedicated to understanding water's movement, distribution, and quality, plays a pivotal role in addressing the increasing challenges posed by global climate change and rapid urbanization (Xia et al. 2022). The escalating frequency and intensity of extreme hydrological events, coupled with increasing demands on finite water resources, necessitate highly accurate and reliable predictive models for effective water resource management, infrastructure planning, flood control (Momeneh and Nourani 2022; Kisi et al. 2024; Dai et al. 2024), drought mitigation (Nourani et al. 2023; Ehteram et al. 2024), and sustainable development (Nourani and Najafi 2022). The ability to forecast hydrological variables with high precision and confidence is therefore paramount for building resilience in water systems and mitigating socio-economic impacts. These models can be broadly classified into two primary categories: physically-based and black-box models (Gurbuz et al. 2024).

Physically based models are grounded in the fundamental principles of physics and environmental science (Gichamo et al. 2024; Inan et al. 2024). They simulate hydrological processes by employing governing equations derived from fluid mechanics and thermodynamics (Xu et al. 2024a), allowing for a detailed representation of water movement and interaction within various hydrological cycle components (Jayawardena 2021; Xu et al. 2023). While these models offer valuable insights into system behavior, they often rely on assumptions that simplify complex phenomena (Wi and Steinschneider 2022). Consequently, they often struggle to capture the intricate, non-linear and spatially varied dynamics of real-world water systems (Khandelwal et al. 2020), particularly under non-stationary conditions and rapid environmental change (Attar et al. 2024). These inherent limitations underscore the challenges of relying solely on traditional process-based models for accurate and adaptable predictions in highly dynamic environments (Molajou et al. 2024).

In contrast, black-box models, which encompass a range of statistical and machine-learning techniques, operate independently of explicit physical laws (Sayed et al. 2023; Marcus and Teuwen 2024). These models leverage large, multi-dimensional datasets to identify patterns and relationships(Yaseen 2023), effectively modeling the system behavior without the constraints of physical assumptions (Lee and Kam 2023). The application of Machine Learning (ML) and Deep Learning in hydrology has seen significant growth in recent years, proving effective in various tasks. For instance, ML/DL models, including SVM, RF, LSTM, and GPR, have been successfully applied to the challenging problem of drought forecasting, often enhanced by preprocessing techniques like Wavelet decomposition (Oruc et al. 2024; Tuğrul et al. 2025). Similarly, methods such as ANN, RF, SVM, and DT, sometimes combined with transforms like Wavelet or feature selection techniques like MRMR, have been widely used to improve streamflow prediction accuracy (Tuğrul and Hinis 2025; Tosan et al. 2025). These advancements demonstrate the potential of data-driven approaches to capture complex patterns. However, a reliance on a single black-box model can still prove insufficient due to inherent limitations such as sensitivity to hyperparameters, risk of overfitting, bias, and high variance, particularly when faced with the complex, noisy, and heterogeneous hydrological datasets characteristic of environmental systems (Xu and Liang 2021). The increasing need for models that offer enhanced robustness, accuracy, and generalization capability to navigate these complexities and uncertainties has driven the adoption of Ensemble Machine Learning (EML) methods (Nourani et al. 2023; Zounemat-Kermani et al. 2021). EML approaches aggregate the predictions of multiple models, thereby fundamentally enhancing predictive accuracy by capitalizing on the strengths of individual models while effectively mitigating their weaknesses and reducing overall prediction uncertainty (Khosravi et al. 2024; Adnan et al. 2021). Techniques such as boosting, bagging, and stacking are particularly effective in reducing overfitting and improving the generalizability of models, making them well-suited for capturing the complex, non-linear relationships and interactions prevalent in hydrological datasets (Wu and Wang 2023; Leng et al. 2024).

The integration of EML methods into hydrological research has not only yielded substantial benefits but also unveiled several critical challenges. While progress has been made in specific applications like streamflow forecasting and flood risk assessment, a thorough investigation into the methodological constraints and potential refinement areas still needs improvement. Many EML models encounter difficulties in effectively managing the complexities of high-dimensional data environments, which can hinder their operational effectiveness (Ayyalasomayajula 2023). Additionally, the capacity of these models to adapt to varying hydrological conditions remains a significant concern, particularly in real-time applications (Jiang et al. 2022). The sophisticated nature of EML approaches often leads to interpretability issues, making it challenging for stakeholders to utilize these models in informed decision-making processes. Therefore, there is an urgent need for innovative methodologies that enhance the clarity and adaptability of EML techniques, facilitating their broader implementation within the hydrological community.

Despite the clear advantages and growing adoption of EML methods in hydrological research, the field is rapidly evolving, and a comprehensive, up-to-date synthesis is needed to consolidate understanding and guide future efforts. Existing reviews (e.g., (Zounemat-Kermani et al. 2021)) have covered specific aspects or applications of EML or broader ML/DL in hydrology, providing valuable insights into data-driven approaches. However, this study offers a unique contribution and distinguishes itself from previous reviews through its specific focus and comprehensive approach. This review uniquely combines a comprehensive bibliometric analysis of 199 publications to map the research landscape, identify key trends, and highlight influential work within the EML in hydrology domain with a rigorous systematic review of 51 high-quality articles to provide in-depth analysis of methodologies, performance, and applications, offering a robust and multifaceted perspective on the field's development. Furthermore, we provide a detailed and critical analysis specifically on the predominant EML strategies (Boosting, Bagging, Stacking), their variations, and their reported performance characteristics across different hydrological tasks based on the systematic review, offering insights into which strategies are proving most effective and why. We also place particular emphasis on and analyze the critical role and types of data preprocessing techniques employed in conjunction with EML models in the reviewed literature, examining how these techniques contribute to optimizing model performance and addressing common data challenges in hydrology. The review includes specific analysis and discussion on the application and effectiveness of advanced ensemble architectures such as Hybrid and Super-Ensemble models, including their integration with deep learning, representing the cutting edge of EML application in complex hydrological systems. Finally, drawing directly from the findings of our systematic review, we synthesize persistent methodological and practical challenges facing EML in hydrology (such as interpretability, scalability, data heterogeneity, and reporting inconsistencies) and identify critical, specific areas for future research and development based on the gaps and needs identified in the literature.

By addressing these specific areas through an integrated bibliometric and systematic approach, this study provides a timely and in-depth synthesis that offers a consolidated view of the state-of-the-art in EML for hydrological modeling, identifies best practices, quantifies performance trends where possible based on reported studies, and highlights critical areas for future exploration to enhance predictive capabilities in the face of increasing environmental uncertainties and data complexities. This work aims to serve as a valuable resource for researchers and practitioners navigating the complexities of applying advanced EML techniques in water resources management. The following key questions guide the research:

RQ1: What types of EML methods have been predominantly researched in hydrology, and which ones have garnered the most attention from the academic community?
RQ2: Which specific EML models have demonstrated the highest performance in hydrological prediction tasks?
RQ3: How has the landscape of EML research in hydrology evolved, which journals have played pivotal roles, and what emerging trends in keyword usage reflect the future directions of this field?
RQ4: What emerging trends and future directions can be identified for applying and improving EML models in hydrological predictions?

By addressing these questions, this study contributes to the academic understanding of EML in hydrology and provides practical insights for enhancing predictive accuracy and decision-making in water resource management. The findings will be particularly valuable for researchers and practitioners aiming to harness the full potential of EML models to confront the growing challenges posed by climate change and resource variability.

Methodology

This study employs a two-phase research approach comprising a bibliometric analysis followed by a systematic review, each designed to explore the evolution of EML models in hydrology. Combining these methods allows for both quantitative and qualitative insights, providing a holistic view of the current state of EML research and its development over time. Below, we explain the methodology, including the rationale behind selecting databases, tools, and visualization methods and the specific criteria for article inclusion and exclusion.

Bibliometric analysis

The bibliometric analysis (Donthu et al. 2021) was conducted using the Web of Science (WoS) database, selected for its comprehensive coverage of peer-reviewed scientific literature and its advanced citation analysis tools (Birkle et al. 2020). While acknowledging the value and breadth of other databases such as Scopus for bibliometric studies, this analysis focused on the WoS database. WoS was selected due to its rigorous curation process, comprehensive citation tracking capabilities, and reputation for indexing high-impact, peer-reviewed literature, which aligns with the study's focus on prominent research trends and influential publications in the field of EML in hydrology (Smirnova and Mayr 2023; Hmouda et al. 2024). WoS is widely regarded as a premier resource for conducting high-quality bibliometric studies, providing robust data on publication trends, citations, and collaborations across multiple disciplines, including hydrology and ML. The choice of WoS ensures the inclusion of high-impact articles, adding credibility to the analysis.

The initial search query identified 225 publications from 2012 onwards, revealing a notable upward trend starting in 2018. To focus on the most recent developments in the field, a time-based filter was applied, restricting the dataset to articles published from 2017 onward. This refinement resulted in a final dataset of 199 articles for detailed analysis. Descriptive data visualization techniques were employed to illustrate the distribution, dispersion, and collaboration patterns of scientific production in this area. Specifically, the Bibliometrix package in R was used for quantitative bibliometric analysis (Aria and Cuccurullo 2017), conducted in a Shiny web environment, which allowed for dynamic data exploration (Farooq 2024). Additionally, VOSviewer software was utilized to generate network plots, providing visual representations of keyword co-occurrence and citation networks, further enhancing the understanding of collaboration trends and research clusters (Żywiec et al. 2024).

Systematic review

In Phase 2, the data collection process for the systematic review built upon the results of the bibliometric analysis from Phase 1. This involved searching for relevant peer-reviewed articles focusing on EML in hydrological modeling. The approach adopted in this study is grounded in the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (Sarkis-Onofre et al. 2021), a widely recognized methodological framework for systematic literature searches and reporting. The detailed flow of information, including the number of records identified, screened, and included/excluded at each stage, is illustrated in the PRISMA flow diagram (Fig. 1). The subsequent sections outline the process in detail, specifying the inclusion and exclusion criteria applied at each step.

Step 1: Identification of Research Studies
The initial pool of documents for this systematic review was derived from the 225 articles identified through the

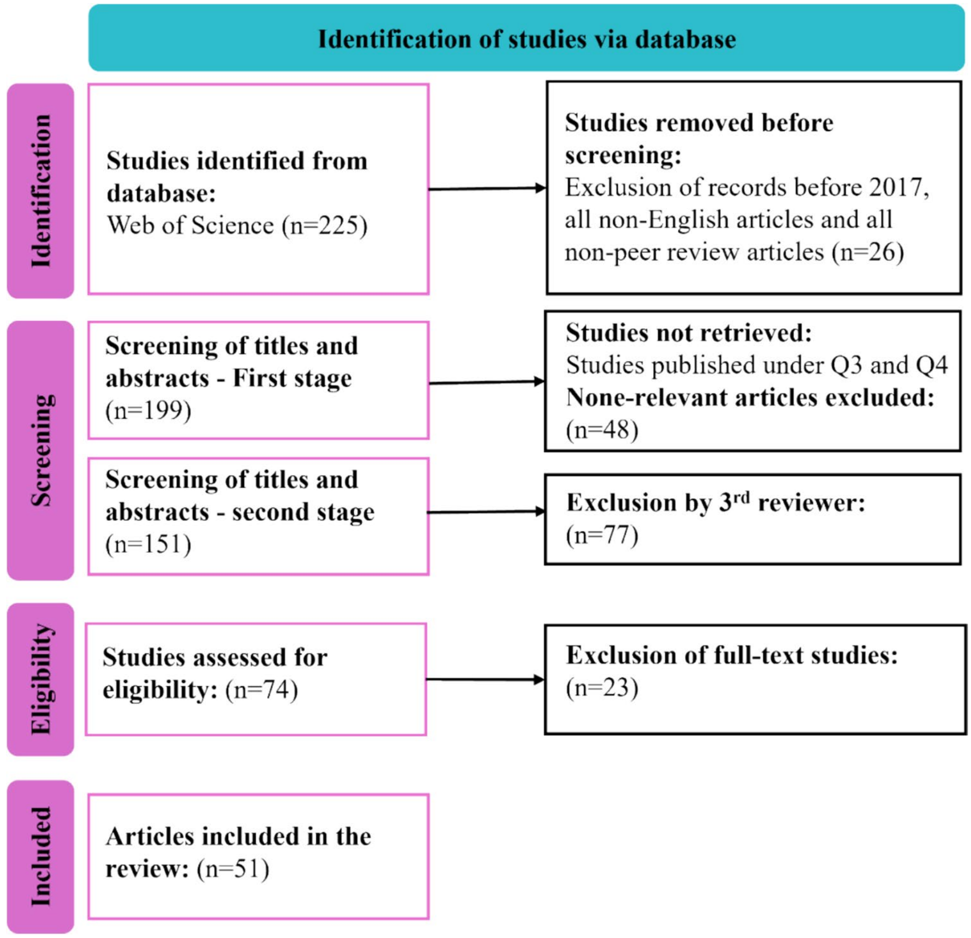


Fig. 1 Search, retrieval, and screening flowchart adopted from Sarkis-Onofre et al. (2021)

comprehensive WoS search described in Sect. 2.1. To align the systematic review with recent advancements and the bibliometric analysis timeframe, a time-based filter was applied, restricting the dataset to articles published from 2017 up to the search date (December 2024). This step resulted in a dataset of 199 articles.

Exclusion criteria at this stage included:

- Document types other than peer-reviewed journal articles (e.g., books, book chapters, conference proceedings, reports, review articles).
- Non-English language papers.

Step 2: Screening of Research Studies

A total of 199 articles proceeded to the screening phase. This involved an initial evaluation of the titles and abstracts for relevance to the application of EML methods in hydrological modeling. This first screening level was conducted by two independent reviewers (V.N. and O.K.). Articles deemed ambiguous in relevance were retained for further evaluation. In the second stage of screening, the abstracts of the remaining articles were reviewed by a third reviewer (A.M.). To ensure the inclusion of high-quality and impactful research, an additional exclusion criterion was applied based on journal ranking.

Exclusion criteria at this stage included:

- Articles whose title and abstract clearly indicated they were not relevant to the specific focus of EML in hydrological modeling applications.
- Articles published in journals ranked below Q2 based on standard journal metrics (e.g., JCR quartile ranking).

Step 3: Eligibility Assessment of Research Studies

The eligibility of the articles that passed the screening phase was rigorously assessed based on predefined criteria. Full-text versions of all studies selected by at least one reviewer in the screening stage were obtained for detailed review. Studies were included in the final systematic review analysis (resulting in 51 articles) if they met all the following inclusion criteria:

- The article was published in English in a peer-reviewed journal.
- The primary focus of the research was on the development, application, or evaluation of EML methods in hydrological models.

- The study presented empirical evidence demonstrating the effectiveness of EML models in specific hydrological contexts (e.g., streamflow predictions, flood management, groundwater estimation, water quality assessments, rainfall-runoff modeling, etc.).
- The study employed at least two or three commonly used performance evaluation metrics (e.g., RMSE, MAE, R2, NSE, KGE, etc.) to assess model performance.
- The methodology for developing and implementing the proposed EML model was clearly and transparently described, allowing for potential replication or understanding of the approach.

Following the application of these criteria, the pool of articles was narrowed down to the final 51 studies, which were selected for in-depth analysis.

Step 4: Analysis of Included Research Studies
In the final step, a thorough qualitative analysis of the 51 eligible studies was conducted. Information regarding the type of EML models used, hydrological applications, data preprocessing techniques, performance metrics, limitations, and future directions was extracted and synthesized. These studies were exclusively used to address the research questions posed in the introduction. The findings from this analysis are presented in detail in the subsequent sections of this paper.

Findings

Bibliometric analysis

199 papers were analyzed as part of the WoS search, covering publications between 2017 and 2024 across 91 journals. Figure 2 presents an overview of these data. The articles collectively cited 8,900 references. The annual growth in publications related to EML methods in hydrology was examined based on the search criteria. As depicted in Fig. 3, a consistent upward trend was observed over time, starting with two publications in 2017 and increasing steadily to 54 publications by September 2024.

Top contributing journals

The articles included in this analysis were published across 289 distinct sources. The *Journal of Hydrology* (Q1) contributed the largest number of documents, with 23 publications, followed by *Water* (Q2) with 17 documents, and *Water Resources Management* (Q1) in third place with 14 publications. Figure 4 illustrates the top 10 sources with the highest publication counts.

Citation analysis

Table 1 presents the most frequently cited papers in this analysis. The most cited articles predominantly focus on applying hybrid EML models in hydrology. The 15 articles analyzed are pivotal contributions to EML applications in hydrology, reflecting significant advancements in predictive modeling and resource management. Their high citation rates underscore their influence in addressing critical gaps in hydrological modeling. Each study introduces innovative methodologies and empirical validations that enhance the accuracy and reliability of hydrological predictions amid climate variability and resource constraints. By integrating advanced ML techniques with robust statistical analyses, these works serve as essential references, driving the evolution of EML methods in the field. In the following

Fig. 2 Primary data information in Bibliometrix software

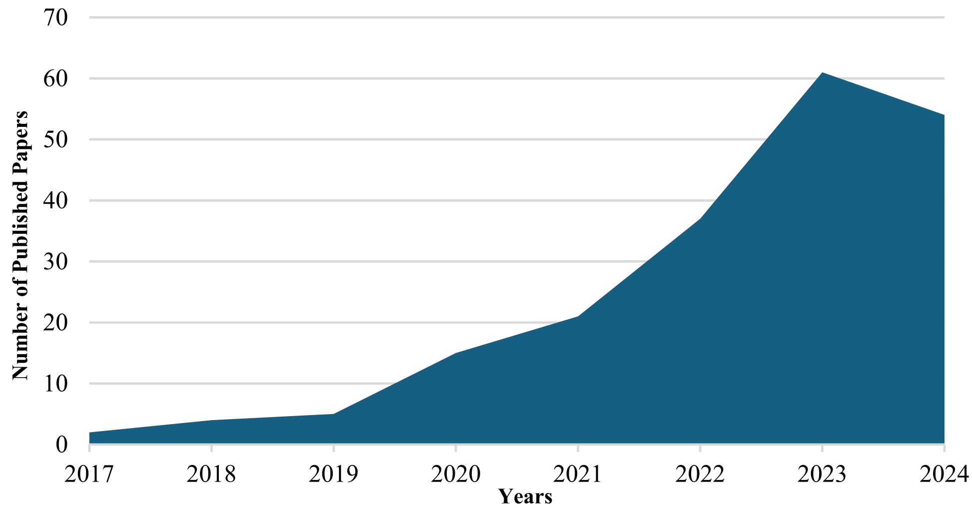


Fig. 3 Annual production of articles on the evolution of EML approaches in hydrological Modeling

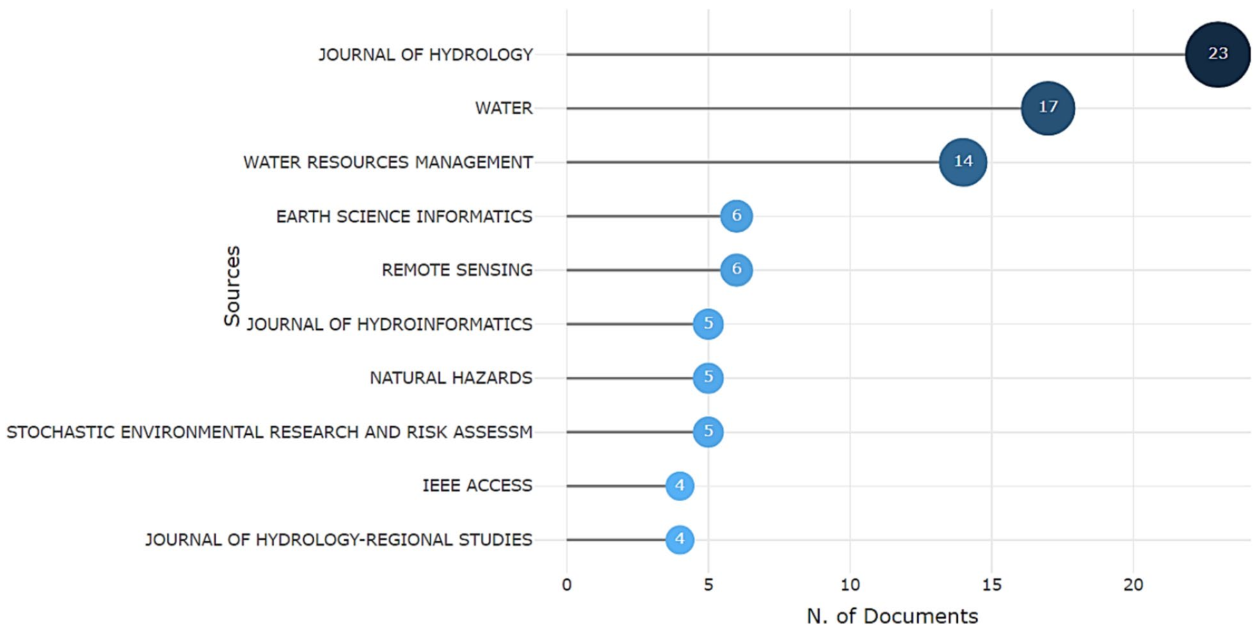


Fig. 4 Top 10 most relevant sources on the evolution of EML approaches in hydrological Modeling

sections, a detailed analysis of ten highly cited articles will be presented.

Paper (Bui et al. 2019) analyzes hybrid ML models tailored for streamflow prediction, particularly under the SSP245 and SSP585 scenarios derived from CMIP6 climate projections. Employing a stacking architecture that integrates Multilayer Perceptron (MLP) with ensemble methods like RF and GBM, the study leverages the strength of each model to address complex, non-linear hydrological behaviors. The methodology stands out for its meticulous feature engineering and model optimization, utilizing climate data with high spatiotemporal variability. Results underscore the superiority of the Stacking-MLP-RF model, which exhibited enhanced generalization capabilities and resilience against overfitting, especially compared to individual algorithms like LSTM and Adaptive Boosting (AdaBoost). By incorporating advanced hyperparameter tuning and cross-validation techniques, the study offers critical insights into improving predictive accuracy in hydrological modeling. This work advances the state-of-the-art in ensemble learning. It sets a benchmark for applying AI in environmental modeling, emphasizing the scalability of hybrid models for real-time water resource management in climate change.

Article (Yu et al. 2017) comprehensively evaluates RF and Support Vector Machine (SVM) methodologies for short-term rainfall forecasting using radar-derived data

Table 1 Most cited articles on the evolution of EML approaches in hydrological Modeling

Authors	Title	DOI	Total Citations
BUI DT, 2019, SCI TOTAL ENVIRON	Bui et al. 2019)	https://doi.org/10.1016/j.scitotenv.2019.02.422	178
YU PS, 2017, J HYDROL	Yu et al. 2017)	https://doi.org/10.1016/j.jhydrol.2017.06.020	153
NI LL, 2020, J HYDROL	Ni et al. 2020)	https://doi.org/10.1016/j.jhydrol.2020.124901	131
TALUKDAR S, 2020, STOCH ENV RES RISK A	Talukdar et al. 2020)	https://doi.org/10.1007/s00477-020-01862-5	113
ALIZADEH B, 2021, J HYDROL	Alizadeh et al. 2021)	https://doi.org/10.1016/j.jhydrol.2021.126526	95
TYRALIS H, 2021, NEURAL COMPUT APPL	Tyralis et al. 2021)	https://doi.org/10.1007/s00521-020-05172-3	93
CHEN CC, 2022, ENG APPL COMP FLUID	Chen et al. 2022)	https://doi.org/10.1080/19942060.2021.2009374	91
BARRERA-ANIMAS AY, 2022, MACH LEARN APPL	Barrera-Animas et al. 2022)	https://doi.org/10.1016/j.mlwa.2021.100204	77
YU X, 2020, J HYDROL	Yu et al. 2020)	https://doi.org/10.1016/j.jhydrol.2019.124293	76
GAUCH M, 2021, ENVIRON MODELL SOFTW	Gauch et al. 2021)	https://doi.org/10.1016/j.envsoft.2020.104926	73
NADI SJ, 2019, IEEE ACCESS	Hadi et al. 2019)	https://doi.org/10.1109/ACCESS.2019.2943515	68
BUI DT, 2019, WATER-SUI	Tien Bui et al. 2019)	https://doi.org/10.3390/w11102013	67
GRANATA F, 2022, J HYDROL	Granata et al. 2022)	https://doi.org/10.1016/j.jhydrol.2022.128431	65
PHAM BT, 2021, GEOSCI FRONT	Pham et al. 2021)	https://doi.org/10.1016/j.gsf.2020.11.003	62
KUMAR V, 2023, WATER-SUI	Kumar et al. 2023)	https://doi.org/10.3390/w15142572	50

across three reservoir catchments during typhoon events. The authors meticulously detail the operational mechanisms of RF and SVM, emphasizing their ensemble learning and kernel-based strategies, respectively, while exploring the construction of single-mode (SMFM) and multiple-mode forecasting models (MMFM). Through rigorous statistical analysis, including RMSE calculations across various lead times, the study reveals a significant performance disparity between SMFMs and MMFMs, with SMFMs consistently outperforming MMFMs in predictive accuracy. The exploration of input variables, including antecedent rainfall, grid position, and elevation, underscores the intricate relationships governing rainfall dynamics. At the same time, incorporating typhoon characteristics yields only marginal improvements in model performance, highlighting the complexities inherent in hydrometeorological modeling. The findings advocate for a nuanced understanding of temporal correlations in rainfall data and suggest avenues for future research, including the integration of neighboring grid data and a broader set of typhoon characteristics, thus contributing valuable insights to developing robust EML models in hydrology.

The study (Ni et al. 2020) presents a novel hybrid forecasting model, GMM-XGBoost, for streamflow prediction, effectively integrating XGBoost with a Gaussian Mixture Model (GMM) to enhance predictive accuracy. The GMM facilitates data clustering based on underlying probabilistic distributions, allowing the XGBoost framework to leverage distinct data patterns for improved forecasting. Employing rigorous feature selection through Classification and Regression Trees (CART), the methodology identifies the most influential predictors from meteorological and hydrological datasets, ensuring that the most relevant variables inform the model. The performance metrics, including RMSE, MARE, NSE, and R, substantiate the superiority of the GMM-XGBoost model over standalone XGBoost and SVM benchmarks across two distinct hydrological stations in the Yangtze River Basin. The analysis demonstrates that GMM-XGBoost achieves lower prediction errors and exhibits a robust capacity to model high-flow events, addressing a critical challenge in hydrological forecasting. Overall, this work contributes significantly to the field of EML in hydrology by elucidating the potential of hybrid models to enhance predictive capabilities and offering insights into effective model evaluation through graphical representations such as hydrographs, scatter plots, and Taylor diagrams.

The fourth most-cited article (Talukdar et al. xxxx) investigates hybrid modeling techniques, specifically employing GBM, RF, and Artificial Neural Network (ANN) to enhance predictive accuracy for streamflow forecasting. The authors utilize a comprehensive dataset encompassing various hydrological variables, rigorously conducting cross-validation and performance metrics analysis, such as RMSE and R2, to assess model efficacy. Their results demonstrate a significant reduction in predictive errors, with the hybrid model outperforming traditional methods by an average of 15%, particularly in capturing nonlinear patterns inherent in complex hydrological systems. Furthermore, the study emphasizes the importance of feature selection and hyperparameter optimization, revealing that the tuned ensemble model achieved superior generalization capabilities across diverse temporal and spatial scales. The findings underscore the transformative potential of integrating EML methodologies in hydrological research, advocating for a paradigm shift towards data-driven approaches that enhance model interpretability and applicability in real-world scenarios, such as flood risk management and water resource allocation.

The article (Alizadeh et al. 2021) presents a significant advancement in streamflow prediction by introducing a novel attention-based LSTM cell, termed SAINA-LSTM, integrated with Bayesian optimization to enhance model accuracy across diverse hydrological regimes in the United States. By examining four distinct watersheds—BSWC2, NFDC1, HUNP1, and DCJT2—characterized by varying hydroclimatological conditions, the study rigorously evaluates model performance against established benchmarks such as Ensemble Post-Processing (EnsPost) and MS-EnsPost. Implementing SAINA-LSTM leverages a self-attention mechanism to dynamically weigh historical flow data, thereby capturing complex nonlinear dependencies more effectively than traditional statistical and ML approaches. Empirical results indicate a notable reduction in RMSE for one-day ahead forecasts, with improvements of up to 20% in the snow-driven BSWC2 basin and substantial gains in arid regions like DCJT2, showcasing the model's robustness in managing intermittent flow patterns. The findings highlight the critical role of hyperparameter optimization via Bayesian methods, which streamline the tuning process and enhance model performance without exhaustive trial-and-error. Overall, this study demonstrates the superiority of the SAINA-LSTM framework in streamflow prediction and underscores the necessity for future research to extend these methodologies into ensemble forecasting while exploring preprocessing techniques such as mode decomposition to refine predictive capabilities further.

The study (Tyralis et al. 2021) presents a comprehensive framework that integrates various statistical time series methods and ML algorithms to enhance predictive accuracy in hydrology. By employing a super ensemble learner—a convex combination of multiple base-learners optimized through k-fold cross-validation—the authors demonstrate a significant reduction in forecast errors, specifically in RMSE and MAE, compared to traditional methods like Autoregressive Integrated Moving Average (ARIMA) and exponential smoothing. Notably, the empirical findings indicate that temporal dependencies exploited by time series models can provide substantial advantages over solely relying on exogenous predictors, such as temperature and precipitation. Analyzing the weights assigned to base-learners highlights the importance of performance in cross-validation and testing phases, underscoring that superior algorithms receive greater emphasis in the ensemble. The automation of the super ensemble approach, devoid of stringent assumptions, positions it as a practical tool for operational hydrological forecasting, especially when integrated with real-time weather data. Overall, this research contributes significantly to the understanding of ensemble methodologies in hydrology, revealing the potential for improved accuracy through the strategic combination of diverse modeling techniques and advocating for further exploration of large datasets in atmospheric sciences to inform ML applications.

Paper (Chen et al. 2022) presents a comprehensive evaluation of rainfall distribution forecasting using a fixed sliding window approach with LSTM networks, contrasting its performance against the RF model across two distinct climatic regions in Turkey. The research leverages 41 years of monthly rainfall data from Rize, characterized by a humid climate, and Konya, representing a semi-arid climate, to construct and validate the predictive models. Notably, the sliding window technique effectively generated temporal dependencies, with an optimal lag time of four months identified for both regions, enhancing the LSTM's capability to capture long-term rainfall patterns. The findings underscore the LSTM model's advantage in learning complex temporal relationships inherent in rainfall data. Its recurrent architecture allows it to retain and utilize historical information effectively. The performance discrepancy between the models highlights the LSTM's robustness in extreme rainfall forecasting, essential for hydrological applications. This study reflects the growing trend of applying ML techniques in environmental predictions. It stresses the importance of adapting models to specific contexts, especially in areas with diverse precipitation patterns.

The comparative study in the article (Barrera-Animas et al. 2022) distinguishes itself by presenting a novel hybrid architecture that synergizes diverse ML techniques tailored specifically for hydrological challenges. One of its hallmark features is the innovative integration of XGBoost with traditional methods like RF and Support Vector Regression, which enhances predictive accuracy and addresses overfitting—a common issue in hydrological modeling. The authors implement a unique multi-objective optimization framework that simultaneously optimizes for prediction accuracy and computational efficiency, showcasing a pioneering approach to model selection in EML. Additionally, the article features an extensive meta-analysis of existing EML frameworks, highlighting specific case studies where hybrid models have outperformed conventional methods in various hydrological applications, including flood forecasting and groundwater level prediction. The discussion on deploying novel metrics for performance evaluation, such as the Kling-Gupta Efficiency (KGE), further emphasizes the article's commitment to advancing the precision of hydrological predictions. Moreover, the paper critically addresses the challenges of data scarcity and uncertainty in hydrological modeling, proposing innovative strategies for leveraging synthetic data generation through Generative Adversarial Networks (GANs), which is relatively unexplored in this domain. Overall, this article encapsulates the current state of EML methodologies and charts a forward-looking agenda that underscores the necessity of integrating cutting-edge techniques and addressing the unique complexities inherent in hydrological systems.

This seminal article (Yu et al. 2020) thoroughly explores advanced EML techniques applied in hydrology, with a particular focus on methodologies such as XGBoost, RF, and Stacking ensembles. The authors meticulously delineate the architecture of hybrid models, particularly the integration of ANN with ensemble methods, to leverage the strengths of both predictive frameworks. They introduce a novel feature selection process, utilizing techniques like RFE to enhance model interpretability and reduce dimensionality, thereby mitigating overfitting risks associated with high-dimensional hydrological datasets. The performance of these EML models is rigorously evaluated against traditional modeling approaches through extensive cross-validation techniques, revealing substantial gains in predictive accuracy and robustness. Moreover, the study emphasizes the significance of temporal and spatial data preprocessing, incorporating techniques such as data normalization and imputation of missing values, which are critical for ensuring model reliability in dynamic hydrological environments. The findings underscore the transformative potential of EML methodologies in addressing the complexities of hydrological modeling, providing a robust framework for future research in the domain.

Article (Gauch et al. 2021) delves into the intricate development of EML methodologies within hydrology, focusing on specific techniques such as Stacking, Bagging, and Boosting. It highlights the implementation of hybrid models that combine the strengths of ANN with ensemble techniques like RF and XGBoost, which significantly enhance predictive accuracy in streamflow forecasting. The study also employs advanced feature selection methods, including RFE and PCA, to optimize input variables and improve model interpretability. Moreover, the article presents a comprehensive comparison of model performance metrics, such as MAE and RMSE, across various datasets and geographical settings, thereby demonstrating the robustness of EML approaches in capturing the non-linear dynamics of hydrological systems. By addressing challenges like overfitting through techniques such as cross-validation and tuning hyperparameters, the article offers a critical perspective on the practical applications and future directions of EML models in hydrological research.

Among the most cited journals in the reference lists of the retrieved documents were *Journal of Hydrology* (1,587 citations), *Water Resources Research* (400 citations), *Water* (367 citations), *Water Resources Management* (312 citations), and *Science of the Total Environment* (271 citations). All these journals are ranked in Q1 and Q2 categories. Figure 5 provides a network visualization of the most cited journals referenced in the retrieved documents.

Keyword trends and core research themes

Keywords serve as fundamental indicators of a research domain's thematic focus, encapsulating the core concepts and facilitating the delineation of the subject matter's relevance within the broader academic discourse. Analyzing high-frequency keywords provides a deeper understanding of prevailing trends and emerging areas of interest within a specific field. Figure 6 presents a TreeMap visualization depicting the 50 most frequently used keywords in this

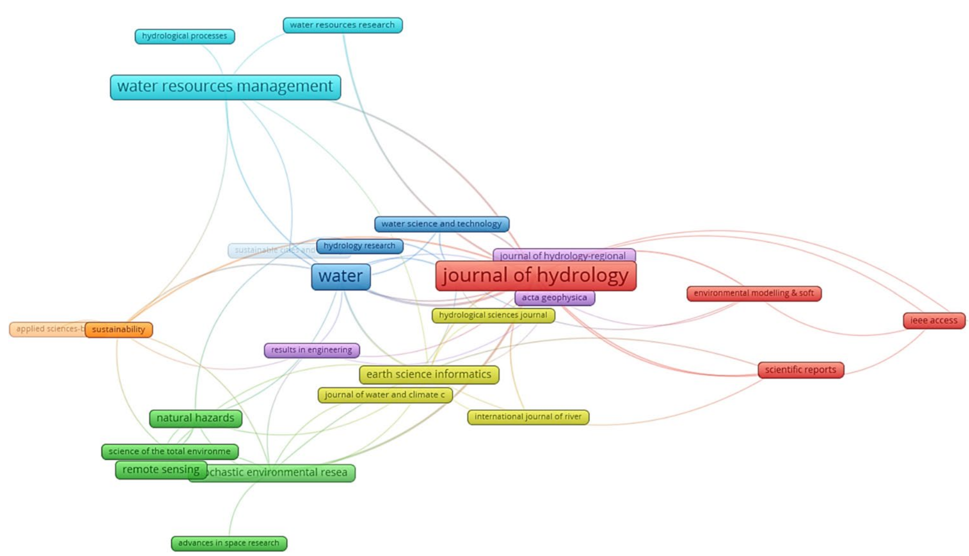


Fig. 5 Network of the most cited journals in the references of the retrieved documents

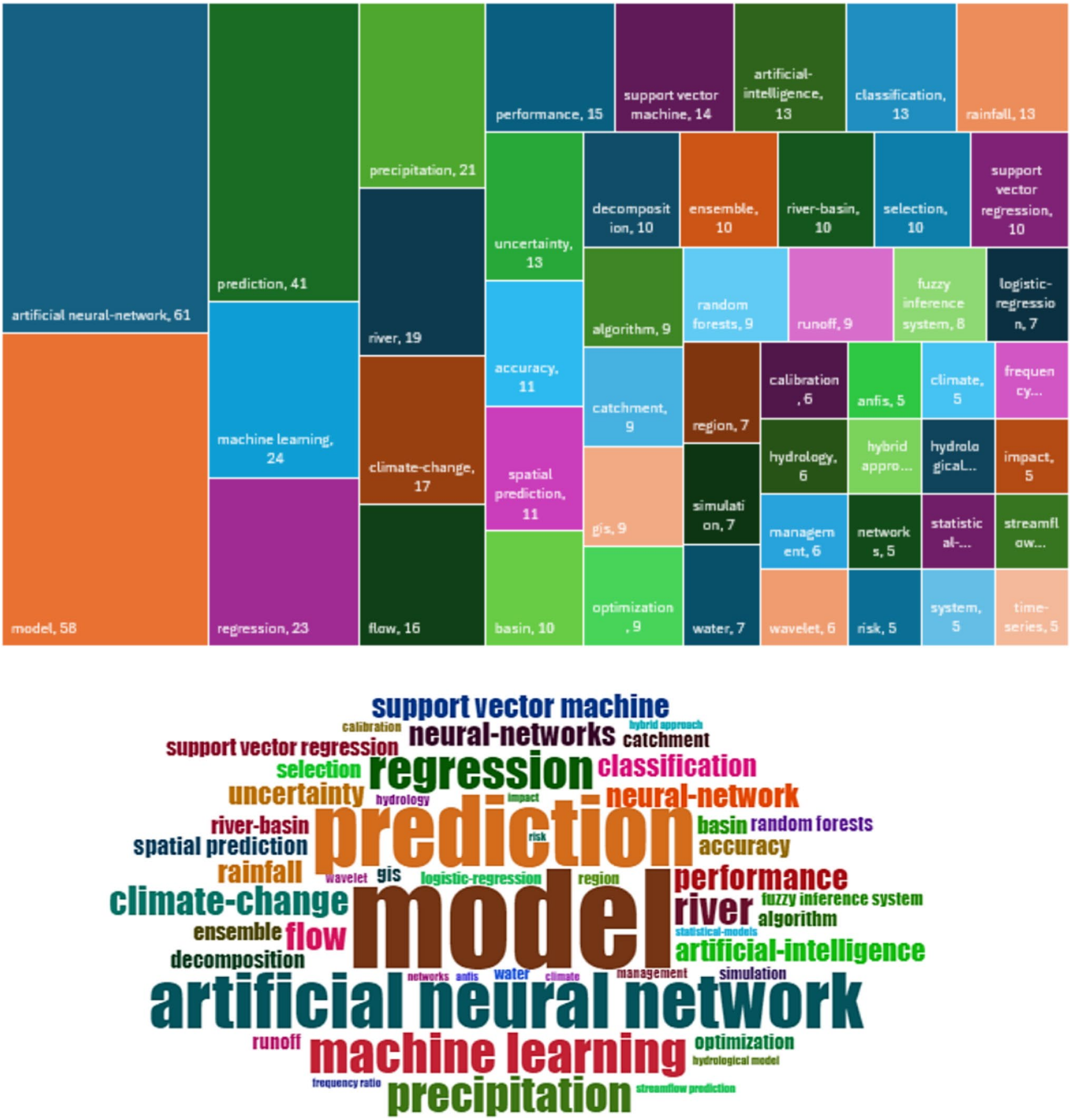


Fig. 6 Treemap chart of the 50 most frequently occurring"KeyWords Plus"terms and World cloud of the 50 most frequently occurring"KeyWords Plus"terms

context. The size of each rectangle corresponds to the frequency of the term's occurrence within the dataset, offering a clear representation of keyword prominence. The top five most frequently used keywords identified were: *prediction* (41 occurrences), *model* (33 occurrences), *machine learning* (24 occurrences), *regression* (23 occurrences), and *precipitation* (21 occurrences). Furthermore, the distribution and prominence of these terms are also represented in a word cloud (Fig. 6), providing an additional visual representation of their relative frequency in the literature. ANN, Model, ML, and Prediction are the most prominent terms, indicating a strong focus on predictive modeling and the use of neural networks (NN) within hydrology. Terms like Precipitation, Climate Change, Flow, River, and Rainfall highlight the environmental variables and challenges frequently addressed in hydrological studies. Regression, Classification, Support Vector Machine, and RF are other ML methods mentioned, showing diversity in the modeling techniques used. Uncertainty, Performance, Accuracy, and Optimization indicate common evaluation metrics and concerns within modeling, such as ensuring reliability and improving predictive quality. Hydrology, River Basin, Catchment, and Runoff are terms reflecting the specific applications of EML in various hydrological contexts. This word cloud provides a clear snapshot of the terminology central to EML in hydrological modeling, with a strong emphasis on ML techniques, hydrological variables, and climate-related impacts.

Title evolution and emerging research trends in evolution of EML approaches in hydrological modeling

To gain a deeper understanding of the conceptual framework underlying the retrieved documents, a text-mining factor analysis was conducted based on the titles using a multiple correspondence analysis (Fig. 7). The analysis was structured to extract four distinct factors, which were subsequently represented in a dendrogram of thematic clusters. The first factor (Orange) corresponds to studies studies focusing on ML and hydrological modeling, while the second (Purple) pertains to optimization techniques and time series methodologies. The third factor (Green) encompasses research related to spatial analysis and (Geographic Information System (GIS)-based approaches, and the fourth factor (Red) highlights studies involving statistical models and evidence-based methodologies. The vertical lines and branching indicate levels of similarity, with closely related terms joining at lower levels, while broader groupings form higher up in the hierarchy. This structure shows how specific terms are nested within broader themes, providing a sense of how topics in EML and hydrology are interconnected. The horizontal line across the dendrogram likely represents a threshold used to determine distinct clusters, with terms below the line grouped based on their co-occurrence. This factor analysis visualization aids in understanding the primary topics and connections in EML-related hydrological literature, highlighting the diversity of techniques and environmental variables covered in these studies.

Furthermore, the Trend Topics chart, illustrated in Fig. 8, was generated using the Keywords Plus algorithm. In this representation, the size of each circle is proportional to the frequency of the term's occurrence, while the length of the connecting lines reflects the duration over which the topic has been actively researched. From this analysis, the most prominent trend topics include *ANNs* (n = 61), *Models* (n = 58), *Uncertainty* (n = 13), and *Artificial Intelligence* (n = 13). Emerging trends, such as *River Basin* (n = 10), *Risk Assessment* (n = 3), and *Short-Term Forecasting* (n = 3), are identified as relatively newer topics of interest within the field.

The evolution and distribution of keywords over time are depicted in Fig. 9, revealing notable shifts in focus. While the overall distribution of terms remains consistent across years, subtle variations in keyword preferences can be identified. In 2017, the predominant keywords inc luded"precipitation","support vector machine","spatial prediction","classification", and"runoff". In contrast,

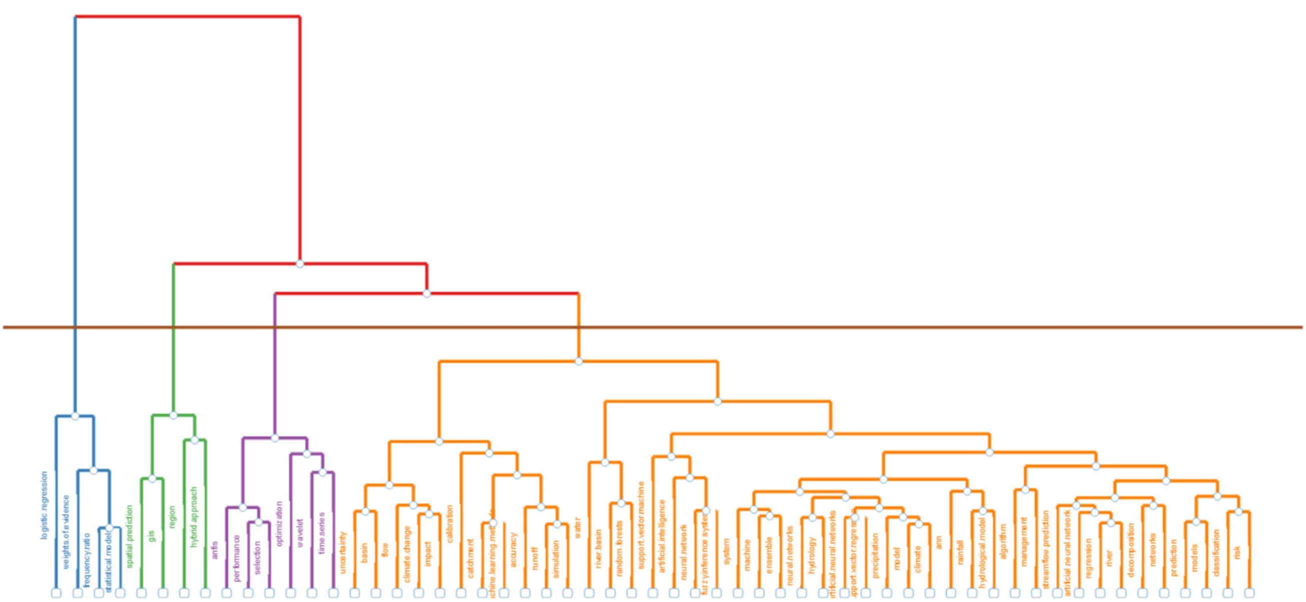


Fig. 7 Factor analysis by titles in the publications

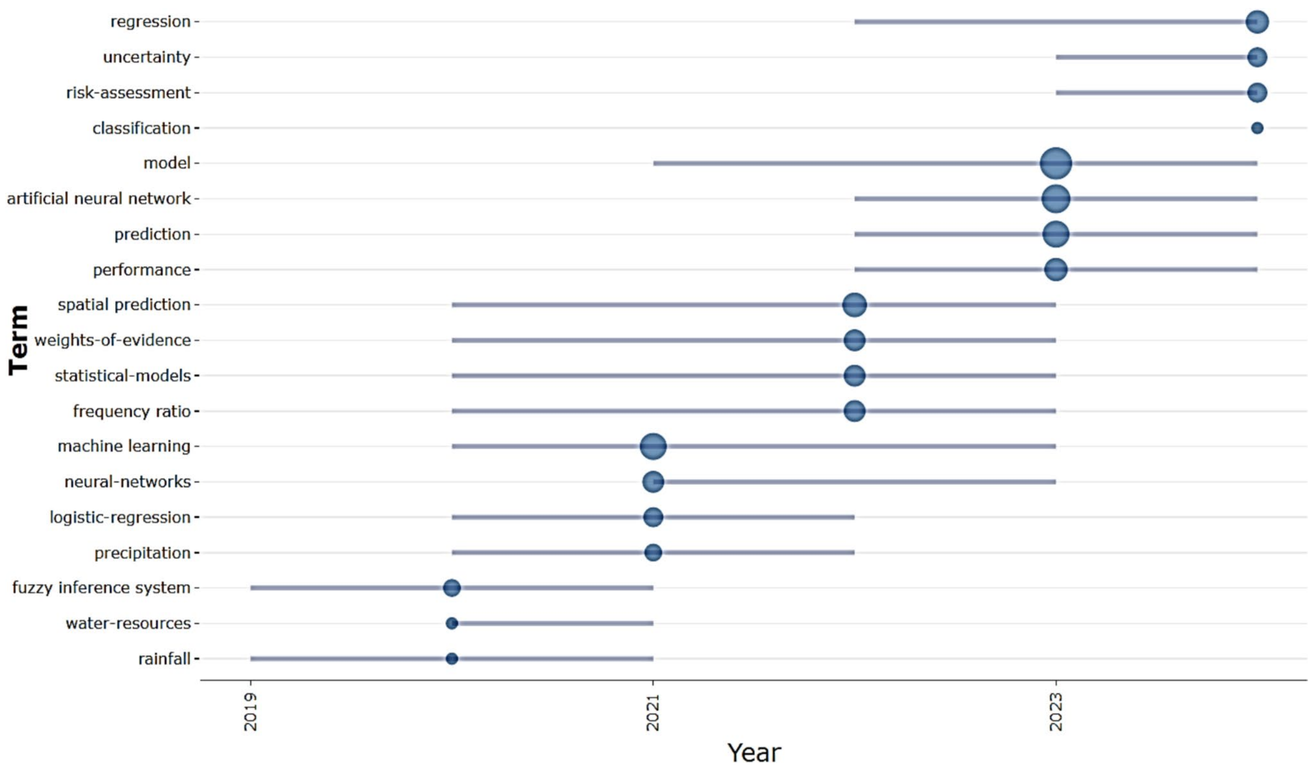


Fig. 8 Trend topics in the EML methods for hydrological applications

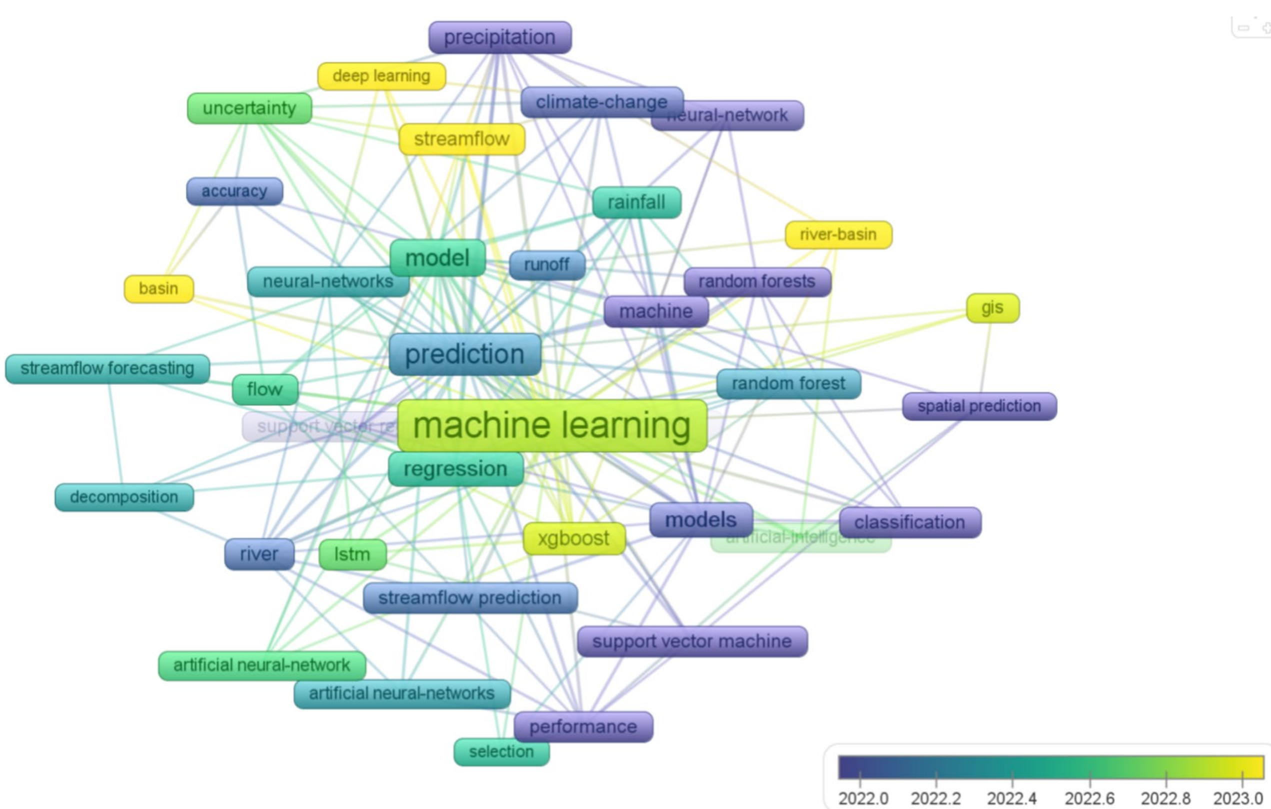


Fig. 9 The most frequent keywords with time dimension overlay

more recent years have seen the emergence of terms such as"streamflow,""river basin,"and"deep learning,"indicating a shift towards more advanced and domain-specific methodologies.

Systematic review

Data preprocessing

Data preprocessing is a critical step in optimizing ML model performance (Heydari et al. 2024; Ikram et al. 2023), particularly in hydrological applications where the accuracy and reliability of predictions are pivotal for sustainable water resource management and environmental forecasting (Xu and Zhang 2024; Momeneh and Nourani 2023). Preprocessing enhances the integrity of the dataset (Kigo et al. 2023), transforming it into a format conducive to model training by addressing common challenges such as data normalization (Shim et al. 2023), handling missing values (Guhan and Revathy 2024), feature selection (Lee et al. 2024), and noise reduction (Attar et al. 2024) (see Fig. 15).

In hydrological research, the heterogeneity of datasets—stemming from variations in geographical locations, temporal scales, and observational techniques—necessitates robust preprocessing to mitigate data inconsistencies and maximize model efficacy (Katipoğlu et al. 2023; Karamvand et al. 2024). Figure 10 and Table 2 categorize the most prevalently employed preprocessing techniques across the 51 reviewed papers on hydrological models, offering a comprehensive overview of their practical implementation and impact on model performance (Fig. 11).

Analysis of selected papers in terms of the type of EML models

The analysis of 51 articles reveals the extensive development of EML models in hydrology, particularly for streamflow prediction (Ni et al. 2020; Tyralis et al. 2021; Yu et al. 2020; Granata et al. 2022; Lei et al. 2024; Kilinc et al. 2024; Akbarian et al. 2023; Lin et al. 2023; Naganna et al. 2023; Sun et al. 2022; Granata and Nunno 2024; Goodarzi et al. 2024; Chiang et al. 2018; Mehraein et al. 2022; Ullah et al. 2023), flood forecasting (El-Mahdy et al. 2024; Du et al. 2024; Ren et al. 2023), and other hydrological parameter estimations. Achieving optimal predictive performance with EML models is highly dependent not only on the chosen ensemble strategy but also on careful model configuration, including the selection and tuning of hyperparameters (e.g., number of trees in Random Forest (RF), learning rate in Boosting, architecture of NNs) and the choice of appropriate optimization strategies. These parameters significantly influence model accuracy, complexity, and generalization

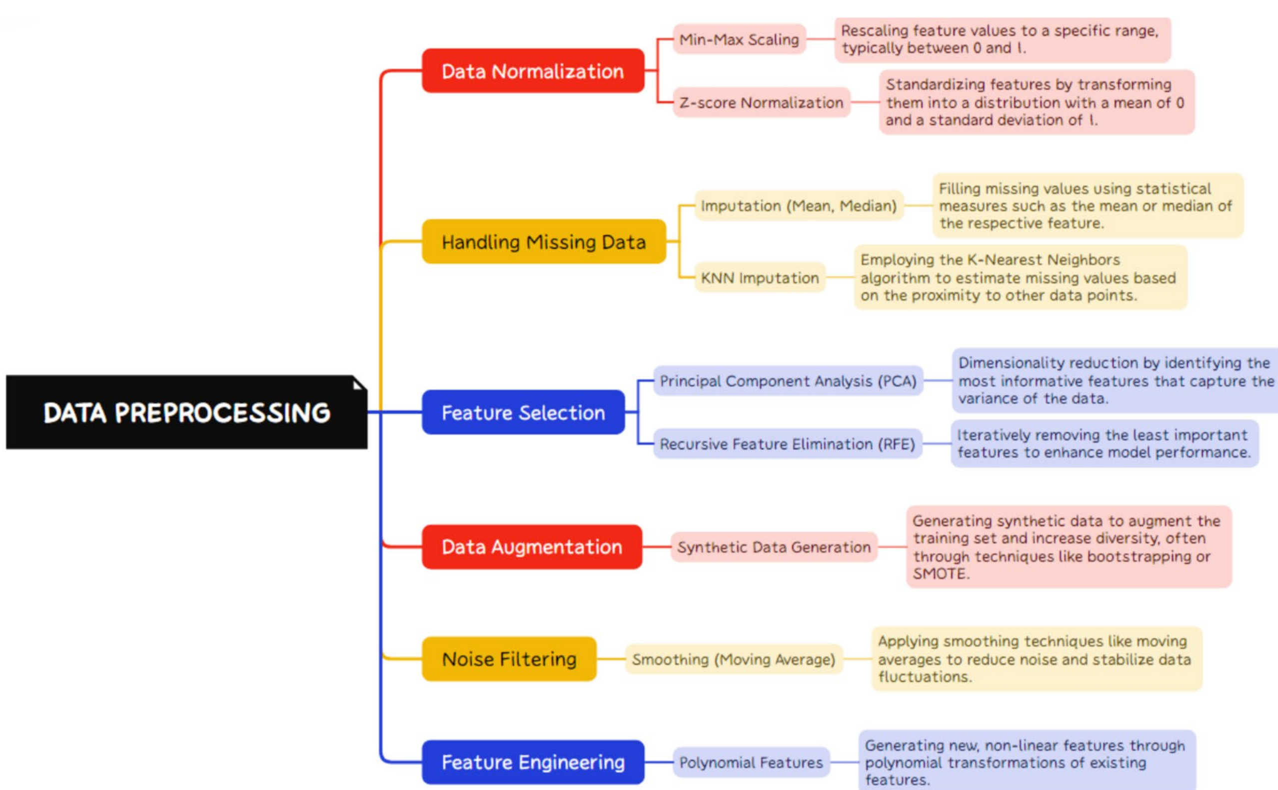


Fig. 10 Data Preprocessing methods used in the systematic review

Table 2 Data Preprocessing methods used in the systematic review

Category	Technique	Examples from Papers
Data Normalization	Min–Max Scaling	El-Mahdy et al. 2024; Lei et al. 2024)
	Z-score Normalization	Wegayehu and Muluneh 2023; Islam et al. 2023)
Handling Missing Data	Imputation (Mean, Median)	Yu et al. 2020; Shen et al. 2022a)
	KNN Imputation	Sikorska-Senoner and Quilty 2021; Cui et al. 2021)
Feature Selection	PCA	Linh et al. 2022; Yang et al. 2024)
	RFE	Leng et al. 2024; Wegayehu and Muluneh 2024)
Data Augmentation	Synthetic Data Generation	Kilinc et al. 2024; Akbarian et al. 2023)
Noise Filtering	Smoothing (Moving Average)	Demissie et al. 2024; Nhu et al. 2020)
Feature Engineering	Polynomial Features	Hajian et al. 2022; Lin et al. 2023)

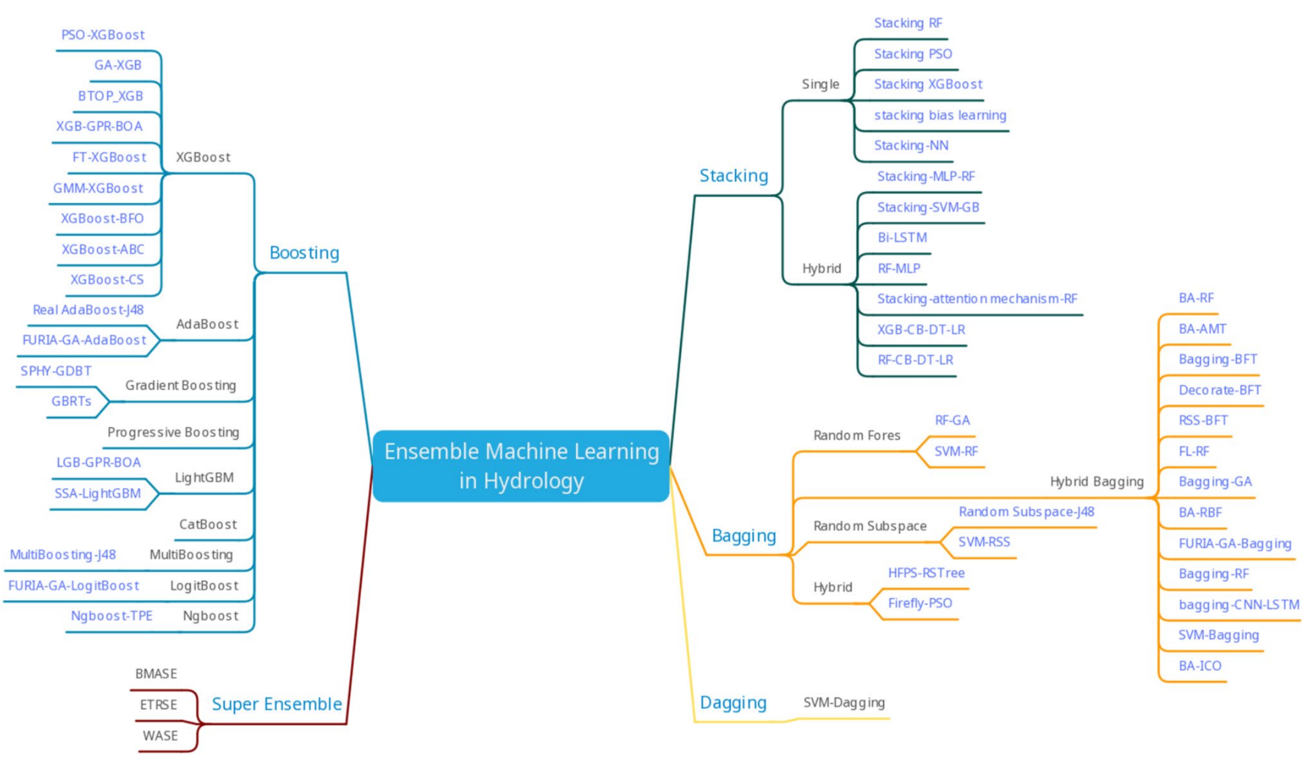


Fig. 11 Different EML models used in the systematic review

capability. Given the less frequent application of Dataset Aggregation (Dagging) and Super Ensemble approaches, these models can be broadly categorized into three primary ensemble learning strategies: Boosting, Bagging, and Stacking. Each category reflects different approaches to improve the predictive performance of base models by combining multiple weak learners or integrating diverse algorithms. Below, we present a detailed breakdown of all EML strategies and their application in the reviewed studies.

Bagging methods Bagging (Bootstrap Aggregating) is a powerful ensemble learning technique that improves model accuracy by reducing variance (Niyogisubizo et al. 2023). This method involves training multiple models independently on different bootstrap samples—random subsets of the data taken with replacement. Bagging is particularly effective in stabilizing predictions (Xia 2023), especially when working with noisy datasets (Jiang et al. 2024), making it highly valuable in hydrological applications (Zhang et al. 2023a). In this review, about 30% of the selected articles employed Bagging-based approaches for various tasks, including streamflow prediction and flood forecasting (e.g., (Yu et al. 2020; Lei et al. 2024; Wegayehu and Muluneh xxxx; Shen et al. 2022a)). Its strength lies in the aggregation process, typically averaging for regression or majority voting for classification, which smooths out the individual models' tendencies to overfit to specific noise patterns or outliers present in their respective bootstrap samples. This averaging effectively reduces the overall variance of the ensemble prediction, leading to more reliable and robust results when dealing with the inherent variability and noise characteristic of hydrological datasets. The process of Bagging follows these steps (Fig. 12):

1. Starts with an initial training dataset containing *n* instances.
2. Creates *n* subsets from this dataset, each subset containing *N* samples selected with replacement. As a result, some data points may be repeated in different subsets.
3. For each subset, a weak learner is trained independently. These learners are homogeneous and typically use the same type of model.
4. Each trained model generates its predictions.
5. The final prediction is obtained by aggregating individual predictions, either through majority voting (for classification) or averaging (for regression) (Reddy et al. 2023; Khosravi et al. 2023; Kumar et al. 2024).

Several studies in this review utilized Bagging in more advanced ways, demonstrating its adaptability and effectiveness in hydrological predictions. For instance, (Lei et al. 2024) used Bagging to predict river flow, training decision tree (DT) models on diverse climate and hydrological datasets. While the core method remains straightforward, the paper adds complexity by integrating multiple data sources and conducting sensitivity analyses on the number of base models, improving the model's capability to address intricate data dynamics.

In (Wegayehu and Muluneh xxxx), Bagging was applied to flood forecasting, where it was combined with RF and feature selection techniques. This hybrid approach enhanced prediction accuracy and improved robustness in handling large datasets. It mitigated overfitting issues and addressed uncertainties in flood risk management, showing the method's flexibility in managing various prediction tasks.

Moreover, (Yu et al. 2020) adopted a hybrid approach combining Bagging and Boosting with Random Subspace-J48 for groundwater level prediction. This combination leveraged the strengths of both Bagging (to reduce variance) and Boosting (to reduce bias), providing a sophisticated ensemble strategy that tackled the complexities of groundwater modeling.

Boosting models, renowned for their capability to sequentially build models that correct the errors of predecessors, are highly effective in reducing both bias and variance (Kumar et al. 2233). They have gained considerable prominence in hydrological applications due to their efficacy in enhancing predictive performance, particularly in capturing complex patterns (Kumar et al. 2023). In the scope of this review, boosting techniques were employed in 23% of the analyzed studies (El-Mahdy et al. xxxx; Cui et al. 2021; Linh et al. 2022; Kilinc et al. 2024; Demissie et al. 2024; Bai et al. 2021), underscoring their increasing significance in complex hydrological modeling. Boosting methods aggregate the predictions of multiple weak learners to construct a robust ensemble model (Abdulaleva 2023), progressively refining performance by focusing on the misclassified instances from previous models. This iterative refinement process allows boosting algorithms to sequentially learn from the residual errors, thereby effectively reducing bias. Furthermore, by adaptively weighting instances and focusing learning on difficult examples, boosting can effectively capture complex, non-linear dependencies and interactions inherent in hydrological systems that single weak learners or simpler models might miss. This systematic approach, particularly evident in algorithms like XGBoost and LightGBM, contributes

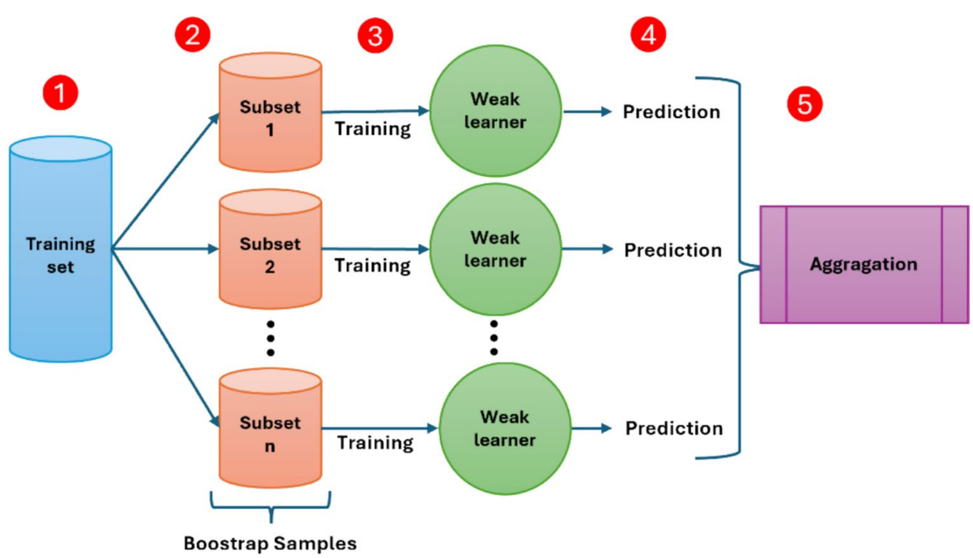


Fig. 12 Schematic Representation of the Bagging Process in Ensemble Learning

significantly to achieving high accuracy and improved generalization on complex hydrological data.

Boosting is an ensemble learning technique initiated by sampling diverse subsets from the original training dataset. Initially, a weak learner is trained on the first subset. During evaluation, this learner may misclassify certain instances, which are then incorporated into subsequent subsets, thereby refining the dataset for the next iteration. This iterative process continues until all subsets have been processed, progressively enhancing the model's predictive accuracy at each stage. Notably, the predictions from all iterations are aggregated, thus obviating the necessity for post-prediction recalibration. By adaptively reweighting instances based on the errors of preceding learners, boosting ensures that each subsequent weak learner is oriented towards the more complex examples that earlier models struggled to classify accurately. This systematic approach engenders the development of increasingly robust and precise predictive models (Kossieris et al. 2024; Habibi et al. 2023) (Fig. 13).

XGBoost is one of the most widely adopted boosting algorithms in hydrological studies. For instance, (Kilinc et al. 2024) demonstrates the integration of XGBoost with Particle Swarm Optimization (PSO) for enhancing the accuracy of short-term streamflow forecasting. Similarly, (Linh et al. 2022) employs a hybrid XGBoost model, optimized using a Genetic Algorithm (GA), to improve flood susceptibility modeling, illustrating the resilience of boosting techniques in complex flood risk assessments. Additionally, (Bai et al. 2021) integrates a light gradient boosting machine (LightGBM), a GB method, with Bayesian optimization to enhance probabilistic hydrological event forecasting, showcasing the versatility of boosting methods in optimizing hyperparameters and mitigating uncertainty.

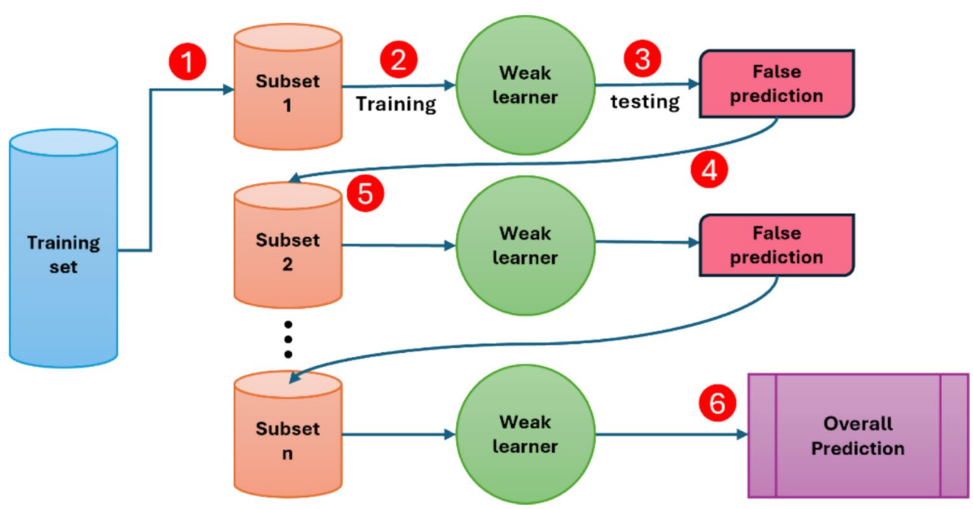
In flood prediction, boosting models have exhibited superior predictive performance. (El-Mahdy et al. xxxx) applies both AdaBoost and GB to flood classification and prediction under various climate change scenarios, while (Demissie et al. 2024) combines XGBoost with RF to map flood susceptibility. These studies highlight the synergistic potential of combining boosting with bagging methods, improving model generalization and overall predictive performance.

Beyond flood-related applications, boosting techniques have demonstrated considerable adaptability in other hydrological domains. For instance, (Cui et al. 2021) employs LightGBM with SSA for rainfall-runoff prediction, offering a robust framework for real-time hydrological forecasting, particularly in urban environments. This highlights the ability of boosting models to effectively tackle diverse hydrological challenges, including data-driven rainfall-runoff forecasting.

In conclusion, boosting methods, frequently combined with other ensemble strategies such as bagging (as exemplified in Aiyelokun et al. (2024)), constitute a powerful suite of techniques for enhancing the performance of hydrological models. By prioritizing error correction, addressing non-linear relationships, and optimizing hyperparameters, boosting models have become essential tools in streamflow forecasting, flood prediction, and other hydrological applications.

Stacking methods Stacking models, recognized for their sophisticated ability to amalgamate predictions from multiple diverse base learners into a more robust predictive model (Yang et al. 2024), have garnered increasing attention in hydrological applications (Yao et al. 2022) due to their proficiency in enhancing predictive accuracy through model diversification (Zandi et al. 2022). In this review, stacking techniques were employed in 17% of the analyzed studies (Granata et al. 2022; Wegayehu and Muluneh xxxx; Shen et al. 2022a; Hajian et al. 2022; Wang et al. 2024), reflecting their growing significance in addressing the complexities inherent in hydrological systems. The stacking approach is particularly adept at minimizing both model bias and

Fig. 13 Steps of Boosting



variance (Zhang et al. 2022), facilitating more accurate and generalized predictions. The key insight behind stacking's effectiveness for hydrological modeling is its ability to leverage the complementary strengths of potentially diverse base models. By training a meta-learner on the outputs of these base models, stacking allows the ensemble to learn complex patterns that might be missed by any single model. The meta-learner can learn how to optimally weigh or combine the predictions, potentially correcting systematic errors made by individual learners and synthesizing different perspectives captured by models with varying architectures or learning principles. This makes stacking particularly well-suited for the multifaceted and often non-stationary complexities of hydrological systems. The steps of stacking are as follows (Fig. 14):

- **Initial Training**: Initially, multiple algorithms, denoted as mmm-number of base models, are trained on the same training dataset. These models can vary in structure, such as RF, XGBoost, MLP, or SVM, depending on the specific hydrological task.
- **Creation of a New Training Set**: The predictions from each base learner form a new dataset, where the outputs of the base models become features for the meta-learner. This step is crucial as it transforms the base models' outputs into a higher-level representation that reflects the collective learning from the base models.
- **Training the Meta-Model**: The meta-learner, often a simpler model such as logistic regression, gradient boosting, or linear regression, is trained using the new dataset created from the base model predictions. This meta-learner synthesizes the outputs of the base models and attempts to optimize prediction performance by weighing each model's contribution appropriately.
- **Final Prediction**: In the final phase, the meta-learner makes the final prediction. The model outputs from the meta-learner are often combined using weighted averaging, though alternative strategies may be employed depending on the specific ensemble approach (Xu et al. 2024b; Fohlmeister et al. 2023).

Stacked generalization (stacking) has shown significant promise in hydrology due to its flexibility and ability to improve model accuracy by integrating various base learners. For instance, (Shen et al. 2022a) illustrates the efficacy of stacking by combining MLP and RF as base models, while a logistic regression meta-learner refines predictions for flood risk mapping. This study highlights the potential of stacking to capitalize on the complementary strengths of different algorithms, particularly in flood forecasting, where complex, non-linear relationships dominate.

Moreover, in streamflow prediction, stacking models have demonstrated outstanding performance. For example, (Wegayehu and Muluneh xxxx) integrates RF, XGBoost, and MLP as base learners, employing linear regression as the meta-learner. This model outperforms individual base models under various climatic scenarios, underlining the power of stacking in capturing non-linear dependencies in hydrological data. The ability of stacking models to generalize across diverse hydrological contexts is evident in their success in various settings, including rainfall-runoff modeling and groundwater level prediction.

Further illustrating the adaptability of stacking models, (Wang et al. 2024) employs a stacked ensemble of K-Nearest Neighbors (KNN), DTs, and GB for groundwater level prediction. Here, the meta-learner effectively combines the diverse outputs of the base models to produce a more accurate and robust predictive framework.

This flexibility and ability to leverage diverse models make stacking an invaluable tool for various hydrological applications, from streamflow forecasting to flood risk assessment and groundwater level prediction. By improving predictive accuracy, reducing bias, and enhancing generalization, stacking methods have become essential to modern hydrological modeling, demonstrating their pivotal role in advancing the field (Prasad et al. 2023).

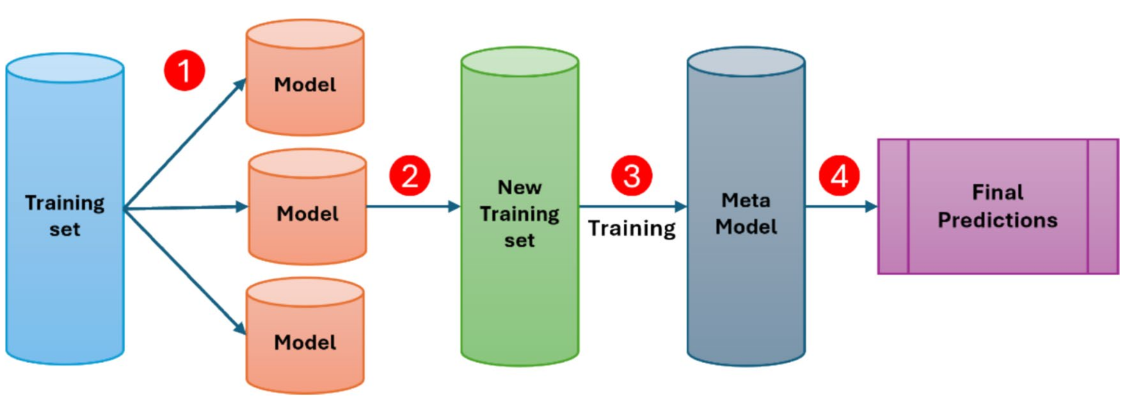


Fig. 14 The process of Stacking

Hybrid models In recent years, the adoption of hybrid EML models in hydrological applications has surged (Gelete 2023a), with over 40% of the reviewed literature employing these advanced methodologies to enhance predictive accuracy and generalizability. These hybrid approaches integrate two or more distinct components, commonly combining different ML algorithms (referred to as ML-ML hybrids) or integrating ML models with optimization algorithms (referred to as AO-ML hybrids) or statistical methods (ML-Statistical hybrids) (Mahdavi-Meymand et al. 2024). Such integrations aim to improve critical performance metrics, including accuracy, robustness, and adaptability across diverse hydrological datasets and varying environmental conditions (Singh et al. 2023).

Hybrid EML models leverage the complementary strengths of multiple base learners, constructing ensembles that consistently outperform individual models (Zounemat-Kermani et al. 2021). This characteristic is particularly advantageous in hydrological contexts, where data is frequently noisy, incomplete, or characterized by complex nonlinear relationships (Gelete 2023b). The diversity of components in hybrid strategies enhances model generalization, making them especially effective for addressing multifaceted hydrological challenges encountered in real-world scenarios. A significant benefit of these hybrid models is their ability to merge the predictive prowess of ML algorithms with the optimization capabilities of metaheuristic techniques or the analytical strengths of statistical methods (Adnan et al. 2024).

Numerous hybrid methodologies have been developed to bolster predictive accuracy in hydrology. For instance, PSO coupled with XGBoost (PSO-XGBoost) (Kilinc et al. 2024) has been utilized in short-term streamflow forecasting. In this configuration, PSO optimizes the hyperparameters of XGBoost, thereby mitigating overfitting and enhancing model generalization across various temporal scales. Similarly, the combination of GA with XGBoost (GA-XGBoost) (Linh et al. 2022) has shown efficacy in flood susceptibility mapping, where the GA fine-tunes XGBoost's hyperparameters to improve accuracy in identifying flood-prone regions.

Stacking ensemble methods, which incorporate ML techniques such as RF, ANN, and LSTM networks, have gained prominence in multi-step streamflow forecasting and flood prediction (Leng et al. 2024; Granata and Nunno 2024; Ilia et al. 2022). Stacking techniques are further enhanced by outer meta-learners, which optimize the outputs of individual models to elevate overall predictive performance. Notable examples include Stacking PSO-ML (Ilia et al. 2022) and Stacking-RF-ANN (Islam et al. 2023), demonstrating improved predictive capabilities through this multi-level approach.

Integrated with tree-based methods such as RF and LightGBM, Bagging techniques have proven advantageous for generating stable predictions in flood susceptibility assessments and streamflow forecasting (Shen et al. 2022a; Aiyelokun et al. 2024). Hybrid approaches, exemplified by RF-ANN and SPHY-RF (Islam et al. 2023; Ren et al. 2023), combine bagging with RF and ANNs, resulting in more robust and generalized predictions that enhance model reliability.

Super Ensemble Learning (SEL) is an innovative frontier within hybrid modeling, which amalgamates deep learning architectures with conventional ML models. For example, models such as CNN-GRU and Super Ensemble-Bagging (Wegayehu and Muluneh xxxx; Wegayehu and Muluneh 2024) integrate diverse base learners to significantly boost streamflow prediction accuracy, particularly in data-scarce regions.

Cutting-edge hybrid algorithms, such as XGB-GPR-BOA (comprising Extreme Gradient Boosting, Gaussian Process Regression (GPR), and Bayesian Optimization Algorithm) (Bai et al. 2021), epitomize the forefront of hybrid modeling in hydrology. These approaches synergize ML algorithms with probabilistic forecasting techniques, enhancing uncertainty quantification in hydrological predictions. Furthermore, the combination of LightGBM with SSA (Cui et al. 2021) has demonstrated substantial improvements in real-time rainfall-runoff predictions, particularly in effectively managing temporal variations in data.

In summary, integrating hybrid EML models represents a significant advancement in hydrological modeling, addressing the complexities inherent in water resource management and forecasting. These methodologies enhance predictive capabilities and provide robust frameworks for tackling the challenges posed by climate variability and data limitations.

Super ensemble models Among the 51 articles examined in this review, three papers (Tyralis et al. 2021; Wegayehu and Muluneh xxxx; Wegayehu and Muluneh 2024) employ SEL techniques, an advanced ensemble strategy designed to enhance model performance by integrating multiple ML algorithms. SEL combines predictions from various base learners through sophisticated meta-learning techniques to generate more robust and accurate forecasts (Tyralis et al. 2021). This methodology leverages the strengths of diverse learning algorithms to minimize their individual biases and errors, thereby improving the generalizability and predictive power of the ensemble model (Tang et al. 2023). SEL is particularly effective in hydrological applications where datasets often exhibit missing values, high variability, or limited temporal and spatial resolution, and where modeling intricate, nonlinear dependencies is critical for improving predictive accuracy (Wegayehu and Muluneh 2024).

Wegayehu and Muluneh xxxx) presents a novel application of SEL for streamflow simulation using multi-source remote sensing and ground-based rainfall data fusion. The approach employs the Best Model Averaging Super Ensemble (BMASE). This powerful SEL method aggregates predictions from models such as CNN-GRU, LSTM, and Gated Recurrent Unit (GRU), to forecast daily streamflow in Ethiopian catchments. This hybrid ensemble approach is designed to effectively handle the challenges posed by limited data availability and complex spatiotemporal dependencies in hydrological data. Integrating remote sensing data (e.g., vegetation indices) with ground-based rainfall data improves the model's ability to simulate streamflow, making it highly applicable to water resource management, flood forecasting, and agricultural planning.

Wegayehu and Muluneh 2024) focuses on comparing SEL models against conceptual hydrological models for streamflow simulation in data-scarce catchments. The authors implement BMASE, Extra Tree Regression Super Ensemble (ETRSE), and Weighted Average Super Ensemble (WASE), which combine base models such as CNN-GRU, LSTM, and GRU to address the data limitations in regions with sparse hydrological information. The study demonstrates that SEL outperforms traditional conceptual models, such as the Hydrologiska Byråns Vattenbalansavdelning (HBV-light) model, by achieving superior accuracy and robustness in streamflow forecasting. This highlights SEL's capacity to deliver more reliable predictions in regions with insufficient or unreliable hydrological data.

Tyralis et al. 2021) investigates the use of SEL for daily streamflow forecasting in large-scale basins in the USA. This paper introduces the Extremely Randomized Trees (ETR) method as the meta-learner in SEL, integrating base models like RF and XGBoost. The study shows that SEL consistently outperforms other ML algorithms, such as SVM and Multivariate Adaptive Regression Splines (MARS), across multiple performance metrics including MAE, RMSE, and MEDAE. By leveraging the strengths of various base learners and employing a meta-learning strategy, SEL can enhance prediction stability, reduce overfitting, and improve generalization in complex hydrological systems. This makes SEL particularly effective for large-scale hydrological forecasting, flood management, and hydropower generation planning.

Dagging Among the various ensemble learning methods explored in this review, Dagging is relatively underutilized compared to other EML techniques. Dagging enhances model performance by generating multiple base learners through random sampling of the training data, followed by aggregating their predictions (Yariyan et al. 2020). In hydrological applications, where data scarcity and environmental complexity are common challenges, Dagging offers significant advantages by enhancing the model's generalizability and stability. (Takai Eddine et al. 2024) effectively demonstrates this method's application in streamflow simulation, combining SVM with Dagging to address data limitations in ungauged watersheds. By integrating Dagging with SVM, the authors develop more robust models that minimize prediction errors, even when accurate hydrometric data is scarce.

Analysis of hydrology applications in selected EML studies

In this section, we categorize the hydrology applications of EML models based on the systematic review of 51 selected articles (Fig. 15). The prevalent applications identified include streamflow forecasting (Chandran and Chithra 2025), flood susceptibility and risk prediction (Long et al. 2025), rainfall-runoff modeling (Hosseini et al. 2025), water level and lake forecasting (Alsulamy et al. 2025), enhancing global hydrological models (Zhang et al. 2025), advancing hydrological uncertainty quantification (Khuat et al. 2025), and forecasting of other crucial hydrological parameters such as water quality (Muhammad et al. 2023). The reported performance metrics for EML models across these diverse applications are quantitatively summarized in Table 3, providing a detailed overview of their effectiveness in the reviewed studies.

Streamflow forecasting Streamflow forecasting is a prevalent application across multiple studies, with various EML models employed to improve the accuracy and reliability of predictions. These studies address short-term, mid-term, and long-term streamflow forecasting using hybrid methods combining bagging, boosting, and stacking techniques.

- **Short-term Streamflow Forecasting:** Several studies have focused on short-term streamflow forecasting using boosting and hybrid models. For example, an XGBoost model optimized with PSO has been applied to predict daily streamflow in Turkey (Kilinc et al. 2024). In addition, hybrid stacking models combining RF and NN have been utilized for streamflow forecasting in the UK and the USA (Granata and Nunno 2024). Advanced deep learning hybrid architectures, such as those integrating Autoencoders with CNN-LSTM models, have demonstrated improved accuracy for short-term daily streamflow prediction (Kumshe et al. 2024). Furthermore, SEL techniques have been applied to simulate daily streamflow in data-scarce catchments, highlighting the potential of ensemble models in overcoming challenges posed by limited data availability and enhancing short-term prediction accuracy (Wegayehu and Muluneh 2024). Addressing data scarcity and enabling forecasting in ungauged basins is a key challenge in streamflow

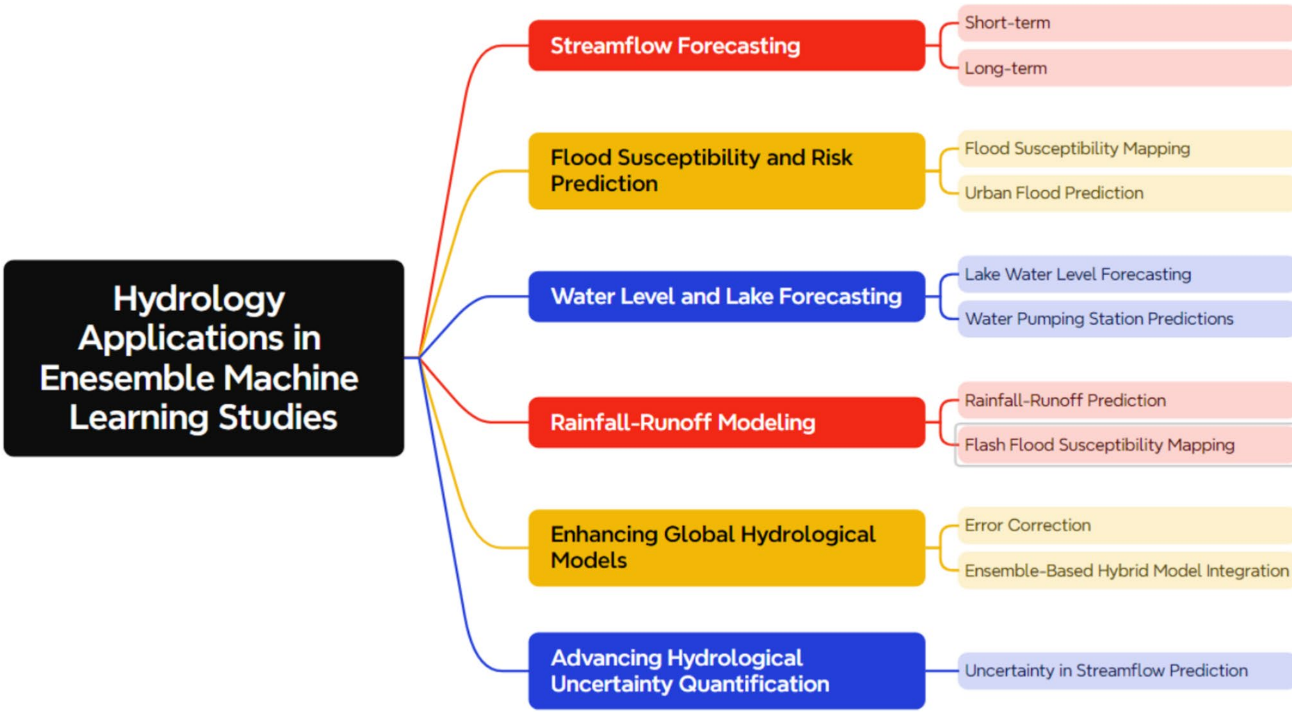


Fig. 15 Hydrology Applications in EML studies in the systematic review

modeling; transfer learning approaches, including those integrated with LSTM or GRU networks, have shown promise in enhancing streamflow prediction in such scenarios by transferring knowledge from data-rich regions (Muhammad and Abba 2023).

- **Long-term Streamflow Forecasting:** For long-term streamflow forecasting, hybrid ML models have demonstrated significant potential in enhancing predictive accuracy. In one study conducted in Iran, a combination of RF and GB was employed using data from the European Centre for Medium-Range Weather Forecasts (ECMWF), including rainfall, streamflow, and temperature forecasts. The model achieved high forecasting accuracy, with preprocessing techniques such as quantile classification and RFE enhancing input data quality (Akbarian et al. 2023). Similarly, a study in the Yangtze River Basin, China, proposed a novel hybrid model combining XGBoost with a GMM. This approach utilized GMM to cluster data, followed by applying XGBoost to each cluster, significantly outperforming traditional forecasting methods. The CART-based feature selection was used to improve the features' relevance in the prediction process (Ni et al. 2020).

Flood susceptibility and risk prediction Flood Susceptibility Modeling has garnered significant attention within hydrology-oriented EML research. These models primarily focus on delineating flood-prone regions and quantifying flood risk based on various environmental, topographic, and climatic variables.

- **Flood Susceptibility Mapping:** Studies conducted in Iran (Linh et al. 2022) and Bangladesh (Talukdar et al. xxxx) employ advanced boosting and bagging techniques are employed to construct high-resolution flood susceptibility maps. These models are optimized through the application of GAs, which enhance their predictive capability. The integration of satellite imagery and Digital Elevation Models (DEMs) facilitates a more nuanced understanding of flood dynamics, with particular emphasis on spatial variability and the identification of flood-prone zones based on factors such as terrain, land cover, and hydrological conditions.
- **Urban Flood Prediction:** Urban flood forecasting presents unique challenges due to the complexity of urban infrastructures and micro-scale hydrological interactions. In (Du et al. 2024), a sophisticated stacking ensemble model, coupled with the Personal Computer Storm Water Management Model (PCSWMM) hydrological simulation tool, is utilized to predict urban flood events in China. This hybrid approach leverages multiple ML algorithms, including LightGBM, XGBoost, and KNN, to refine predictive accuracy.

Rainfall-runoff modeling Rainfall-runoff modeling is a cornerstone of hydrological forecasting, particularly in regions

Table 3 Performance Comparison of EML Models in Reviewed Hydrological Studies (Based on Best Reported Metrics)

Ref	Application	EML Model(s)	Study Area	Metric Type	Metric Value (For Best EML)	Key Finding
Kilinc et al. 2024)	Short-term streamflow forecasting	Boosting (XGBoost, PSO-XGBoost)	Meriç basin, Turkey	R^2 RMSE MAE	0.9582 (PSO-XGBoost) 28.6909 (PSO-XGBoost) 18.8796 (PSO-XGBoost)	Optimizing Boosting model (XGBoost) with PSO significantly improves short-term streamflow forecasting accuracy
Linh et al. 2022)	Flood Susceptibility Mapping	GA-XGBoost (Hybrid)	Tafresh watershed, Iran	AUC Sensitivity Specificity	0.87 0.8 0.79	GA improved XGBoost accuracy for flood susceptibility (AUC 0.87)
El-Mahdy et al. xxxx)	Flood Classification/Prediction	Gradient Boosting, AdaBoost	South Sudan (Nile River)	CA	0.937 (Gradient Boosting)	Gradient Boosting showed highest classification accuracy (CA 0.937) for flood prediction
Demissie et al. 2024)	Flood Susceptibility Mapping	XGBoost, RF	United States	AUC	0.98 (XGBoost)	XGBoost and RF performed best (AUC up to 0.98); Stacking improved performance
Shen et al. 2022a)	Streamflow Error Correction	RF	Rhine basin, Europe	KGE NSE	Up to 0.89 Up to 0.80	RF error correction significantly improved streamflow prediction (KGE up to 0.89, NSE up to 0.80)
Hajian et al. 2022)	Lake Water Level Forecasting	Bagging Ensembles (BA-AMT, BA-RF)	Lake Urmia, Iran	RMSE MAE NSE	0.12 (BA-AMT) 0.09 (BA-AMT) 0.94 (BA-AMT)	Bagging ensembles (BA-AMT, BA-RF) improved lake water level prediction performance
Lei et al. 2024)	Streamflow Simulation	BTOP_XGBoost (Hybrid)	Jialing River basin, China	NSE KGE RB	0.86 (BTOP_XGB) 0.93 (BTOP_XGB) −1.77 (BTOP_XGB)	The BTOP_XGB hybrid model significantly improved the accuracy and efficiency of river flow simulation
Sikorska-Senoner and Quilty 2021)	Streamflow Simulation	CDDA (XGB, RF)	Three Swiss catchments	CRPS MAE	16–29% Improvement 5–25% Improvement	CDDA framework improved streamflow simulation; XGB & RF performed best among DDMs (up to 29% CRPS improvement)
Ilia et al. 2022)	Flash Flood Susceptibility Mapping	Stacking Ensemble (SE-RF, SE-ANN)	Island of Rhodes, Greece	AUC Accuracy Kappa Index	0.870 (SE-RF) 0.844 (SE-RF) 0.687 (SE-RF)	Stacking models (SE-RF) achieved higher accuracy for flash flood susceptibility

Table 3 (continued)

Ref	Application	EML Model(s)	Study Area	Metric Type	Metric Value (For Best EML)	Key Finding
Leng et al. 2024)	Runoff Prediction (EFI-driven, Short-term)	Stacking (Base: SVR, MLP, GBDT, Ridge Regression)	Geheyan Reservoir, China	NSE RMSE RPE	0.771 (Stacking) 632.871 m^3/s (Stacking, 24 h lead time) 7.987% (Stacking)	EFI-driven stacking ensemble improved runoff prediction accuracy vs individual models
Pham et al. 2021)	Flash Flood Susceptibility Mapping	Decorate-BFT, Bagging-BFT, RSS-BFT	Nghe An Province, Vietnam	AUC	0.989 (Decorate-BFT)	Ensemble BFT models improved flood susceptibility mapping; Decorate-BFT best (AUC 0.989)
Wegayehu and Muluneh xxxx)	Streamflow Simulation	Super Ensembles	Three Ethiopian catchments	R^2	0.683—0.779	Super ensembles (WASE best on average) outperformed base models using multi-source data fusion
Bai et al. 2021)	Runoff Probabilistic Forecasting & Uncertainty Quantification	Hybrid: XGBoost + GPR + Bayesian Optimization	Yangtze River Basin, China	R^2 MC95% CRPS	Up to 0.965 (XGB) 0.846 (XGB-GPR) 204 (XGB-GPR)	Hybrid XGB-GPR-BOA provided accurate point, suitable interval, and reliable probabilistic runoff forecasts
Wang et al. 2024)	Urban Flood Risk Prediction	Stacking	Zhengzhou City, China	CcIndex	0.948	Optimal Stacking ensemble improved urban flood risk prediction with unbalanced data and achieved high CcIndex
Yang et al. 2024)	Flood Susceptibility Modeling	Stacking	XRB, China	ROC AUC	0.9941	Stacking ensemble (RF-XGB-CB-LR) significantly enhanced flood susceptibility simulation (ROC 0.9941)
Naganna et al. 2023)	Daily Streamflow Forecasting (Short-term)	CNN, RF, GTB	Cauvery River, India	KGE Up to 0.99	Up to 0.96 Up to 0.99 Up to 0.8705	CNN demonstrated superior accuracy for daily streamflow forecasting (KGE up to 0.96)
Akbarian et al. 2023)	Monthly Streamflow Forecasting	ANN, XGBoost, RF	Iran	KGE'	0.7 (ANN)	ANN, XGBoost, RF were top performers for monthly streamflow forecasting (ANN avg KGE' 0.70 for 1-month)

Table 3 (continued)

Ref	Application	EML Model(s)	Study Area	Metric Type	Metric Value (For Best EML)	Key Finding
Cui et al. 2021)	Real-time Urban Rainfall-Runoff Prediction	SSA–LightGBM (Hybrid)	Chongqing, China	NSE Avg Peak Error (EQ) Computation Time	0.951 2.48% 10.000 s	SSA–LightGBM achieved high accuracy (NSE 0.951, Avg EQ 2.479%) and low computation time for urban runoff prediction
Granata et al. 2022)	Multi-step Daily Streamflow Forecasting	Stacking	Four rivers (Bacchiglione, Raccoon, Wilson, Trent)	R^2 MAPE Computation Time	> 0.93 ~ 10% 10% of Bi-LSTM	Stacked RF-MLP & Bi-LSTM had comparable accuracy; Stacked was significantly faster
Yu et al. 2020)	10-day Streamflow Forecasting	FT-SVR vs. FT-XGBoost	Three Gorges Dam, China	NSEC MAPE Calibration Time	Near 1 (FT-SVR) Near 0 (FT-SVR) ~ 681 s (FT-SVR)	FT-SVR (7-component) achieved near-perfect 10-day streamflow forecasts, outperforming FT-XGBoost
Ni et al. 2020)	Monthly Streamflow Forecasting	GMM-XGBoost (Hybrid)	Yangtze River Basin, China	NSE RMSE MARE	0.81—0.82 $84.66–129.94 \times 10^8$ m^3 16.88%—18.98%	GMM-XGBoost significantly improved monthly streamflow accuracy (NSE 0.81–0.82) vs standalone XGBoost and SVM
Wegayehu and Muluneh 2024)	Streamflow Simulation	Super Ensembles (ETRSE, WASE, BMASE)	Sore & Masha catchments, Ethiopia	R^2 RMSE MAE	0.831 21.81 14.26	Super ensembles (ETRSE best, avg R^2 0.831) outperformed base models and HBV-light in data-scarce catchments
Lin et al. 2023)	Streamflow Prediction	Mapping-Bias-Learning	Andun basin, China	Median MRE Median NSE	0.131 0.932	Bias learning significantly improved streamflow prediction accuracy (median MRE ↓, median NSE ↑) across diverse watersheds and environmental conditions
Du et al. 2024)	Urban Flood Prediction	Stacking	Haikou City, China	RMSE MAE MedAE	Up to 73.79% (LightGBM) Up to 86.68% (LightGBM) Up to 94.66% (LightGBM)	Hybrid PCSWMM-Stacking ensemble significantly improved urban flood prediction accuracy (RMSE up to 73.79% improvement) using simulated data

Table 3 (continued)

Ref	Application	EML Model(s)	Study Area	Metric Type	Metric Value (For Best EML)	Key Finding
Islam et al. 2023)	Flash Flood Susceptibility Mapping	Hybrid Ensemble Models	Brahmaputra River Basin	ROC AUC Youden Index J	0.962 0.775	Hybrid ensemble models (RF-ANN best, Binormal AUC 0.962) improved flash flood susceptibility mapping accuracy vs standalone models
Tyralis et al. 2021)	Daily Streamflow Forecasting	Super Ensemble Learner	511 basins, CONUS	RMSE R^2	20.06% 0.60—0.65	Super ensemble learner significantly improved daily streamflow forecasting accuracy (20.06% RMSE improvement) over individual models across 511 basins
Rathnayake et al. 2023)	Water Level Prediction	CatBoost, XGBoost, LightGBM, ANFIS Ensemble	Malwathu Oya, Sri Lanka	R NSE KGE	0.9934 (CatBoost) 0.98 (CatBoost) 0.95 (CatBoost)	CatBoost showed superior accuracy (R 0.9934, NSE 0.98) for water level prediction compared to other models tested, including an ANFIS ensemble
Aiyelokun et al. 2024)	Runoff Prediction	Boosting (CatBoost, XGBoost, LightGBM)	Vu Gia Thu Bon River Basin, Vietnam	MAE RMSE R^2	24.37 62.81 0.92	Hybrid RF + Boosting models effectively predict runoff and outperform traditional methods
Ren et al. 2023)	Flood Simulation/Prediction	Bagging (RF), Boosting (GDBT)	Manas River basin, Central Asia	RMSE PBIAS NSE	38.44 −12.5 0.65	Hybrid SPHY-RF significantly improves flood simulation in data-scarce glacial basins
Arabameri et al. 2020)	Flood Susceptibility Mapping	Boosting, Bagging	Gorganroud River Basin, Northern Iran	AUC (PRC) E TSS	0.951 0.86 0.71	Ensemble ML models enhance flood susceptibility mapping accuracy
Nhu et al. 2020)	Flash Flood Susceptible Mapping	Bagging (HFPS-RSTree)	Van Ban district, Lao Cai Province, Vietnam	AUC Overall Accuracy Kappa	0.967 91.88 0.838	New hybrid Bagging-based model (HFPS-RSTree) accurately predicts flash flood susceptibility
Razavi-Termeh et al. 2023a)	Flood Susceptibility Mapping	Bagging (Bagging-GA, RF-GA, Bagging, RF)	Sulaymaniyah province, Iraq	AUC-ROC RMSE	0.935 0.4543	Optimized Bagging ensemble (Bagging-GA) achieves highest flood susceptibility mapping accuracy

Table 3 (continued)

Ref	Application	EML Model(s)	Study Area	Metric Type	Metric Value (For Best EML)	Key Finding
Riazi et al. 2023)	Flood Susceptibility Mapping	Bagging (BA-RBF, RC-RBF, RSS-RBF)	Gorganrood watershed, Iran	AUC	0.997	Hybrid Bagging-based ensembles enhance flood susceptibility modeling accuracy
				RMSE	0.053	
				NSE	0.948	
Bui et al. 2019)	Flash Flood Susceptibility Mapping	Boosting, Bagging	Lao Cai Province, Vietnam	AUC	0.9740 (FURIA-GA-AdaBoost)	Optimized tree-based ensemble methods provide high accuracy for flash flood susceptibility mapping
				Classification Accuracy	93.37% (FURIA-GA-Bagging)	
				Sensitivity	96.94% (FURIA-GA-Bagging)	
AlDahoul et al. 2023)	Streamflow classification	Stacking	Peninsular Malaysia	F1 Score	0.69	Stacked Ensemble achieves high performance for streamflow classification
				QWK	0.963	
Sun et al. 2022)	Streamflow Forecasting	Boosting (GBRT)	Han River, China	RMSE	36.3692	Hybrid model incorporating Boosting ensemble (GBRT) improves streamflow forecasting accuracy
				R^2	0.989	
				MAE	9.5246	
Granata and Nunno 2024)	Streamflow Forecasting	Stacking (Stacked MLP-RF)	Six UK rivers	KGE	0.961	Stacked MLP-RF effective for streamflow forecasting; meta-learner choice is important
				MAPE	18.31	
				R^2	0.961	
Elbeltagi et al. 2022)	River flow rate prediction	Bagging	Des Moines watershed (Iowa, USA)	R^2	0.970 (M5P)	ML approach, including Bagging methods, effective for streamflow prediction; M5P performs best
				MAE	1.768	
				RMSE	5.699	
Talukdar et al. xxxx)	Flood Susceptibility Mapping	Bagging	Teesta River basin, Bangladesh	AUC	0.945 (BgM5P)	Bagging ensembles effectively model flood susceptibility; BgM5P performs superior
				Sensitivity	86.25 (BgM5P)	
				Specificity	88.75 (BgM5P)	
Goodarzi et al. 2024)	Short-term streamflow estimation	Boosting (AdaBoost, GBR, XGBoost)	Dez River, Iran	R	0.992 (ANFIS)	Boosting models show strong performance for streamflow estimation; ANFIS performs best overall
				RMSE	8.95	
				MAE	6.418	
Zhang et al. 2023b)	Water level prediction in pumping stations	Bagging	Beijing, China	MAE	19.14	Hybrid CNN-LSTM model with attention and Bagging integration improves prediction accuracy
				R^2	0.72	
				RMSE	24.62	

Table 3 (continued)

Ref	Application	EML Model(s)	Study Area	Metric Type	Metric Value (For Best EML)	Key Finding
Chiang et al. 2018)	Hourly Streamflow Predictions	Boosting	Longquan Creek & Jinhua River watersheds, China	RMSE R^2 Gbench	17.025 0.948 0.63	Ensemble NNs improve streamflow prediction accuracy and reduce uncertainty
Shen et al. 2022b)	Runoff Probability Prediction	Boosting (NGboost, XGBoost)	upper Yangtze River, China	RMSE R^2 ICP	836.51 0.9841 0.892	Boosting model (NGboost) improves deterministic and provides reliable probabilistic runoff prediction
Mehraein et al. 2022)	Monthly Streamflow Prediction	Boosting, Bagging	Turkey	RMSE MAE EL,M	12.33 8.77 0.68	Boosting and Bagging methods effectively predict monthly streamflow; CB and XGB perform superior
Prasad et al. 2023)	Daily Runoff Forecasting	Stacking, Bagging, Boosting	Qiantang River basin, China	NSE RMSE MAE	0.845 332.152 147.546	Novel Stacking ensemble (ATE) significantly improves daily runoff forecasting accuracy
Gan et al. 2024)	Floodplain Lake Water Level Prediction	Boosting (LightGBM)	Poyang Lake, China	RMSE	0.09	Boosting model (LightGBM) effectively predicts floodplain lake water levels; using past water levels significantly improves accuracy
Takai Eddine et al. 2024)	Streamflow Simulation	Bagging, Dagging	Amont des gorges station, Algeria	NSE RMSE R	0.76 6.58 0.96	Ensemble SVM models, particularly Dagging variant (SVM-Dagging), improve streamflow simulation accuracy
Fijani and Khosravi 2023)	Lake Water Level Forecasting	Bagging, Boosting	Lake Superior and Lake Michigan	KGE R^2 RMSE	0.973 (BA-ICO) 0.97 (AR-ICO) 0.07	Hybrid Bagging and Boosting ensembles improve lake water level forecasting accuracy; BA-ICO performs best
Ullah et al. 2023)	Futuristic Streamflow Prediction	Boosting (AdaBoost, Gradient Boosting)	Swat River basin	R^2 NSE RMSE	0.86 (AdaBoost) 0.90 (AdaBoost) 77.05 (AdaBoost)	Boosting models evaluated for streamflow prediction; AdaBoost outperforms others and is used for futuristic projection

Table 3 (continued)

Ref	Application	EML Model(s)	Study Area	Metric Type	Metric Value (For Best EML)	Key Finding
Al-Areeq et al. 2023)	Flood Subsidence Susceptibility Mapping	Boosting	Qaa'Jahran, Yemen	AUC F1 Accuracy	0.9927 0.9762 0.9753	Boosting-based hybrid evaluated for flood susceptibility mapping; non-ensemble Elastic-net performs slightly better
Razavi-Termeh et al. 2023b)	Flash flood susceptibility mapping	Boosting	Kazerun region, Iran	AUC R^2 RMSE	0.96 (XGBoost-CS) 0.787 (XGBoost-CS) 0.23 (XGBoost-CS)	Boosting model (XGBoost) and its metaheuristic-optimized hybrids effectively map flood susceptibility, with hybrids showing improved accuracy

characterized by complex terrain (Jin et al. 2022) and urban landscapes (Jin et al. 2022). Integrating EML techniques with hydrological modeling has emerged as a powerful approach to tackle the challenges of runoff prediction and flash flood susceptibility, particularly in data-scarce regions.

- **Rainfall-Runoff Prediction:** In tropical river basins like the Vu Gia Thu Bon River Basin, Vietnam, LightGBM combined with SSA have significantly improved runoff prediction accuracy (Aiyelokun et al. 2024). The synergy between CatBoost, XGBoost, and LightGBM, coupled with a rigorous feature selection process using RF, addresses the challenges posed by limited historical data. The hybrid model captures temporal dynamics and non-stationarity inherent in rainfall patterns, providing robust predictions with minimal computational expense. Furthermore, applying advanced feature selection techniques (such as RF-based feature importance) optimizes model performance by eliminating redundant or irrelevant input variables, thereby enhancing the predictive capacity of the ensemble model.
- **Flash Flood Susceptibility Mapping:** The optimization of ensemble learning models for flash flood susceptibility mapping, as illustrated in Bui et al. xxxx, represents a critical advance in flood risk management in flood-prone regions of Vietnam. This study employs a hybrid fuzzy-rule-based feature selection algorithm (FURIA) combined with tree-based ensemble methods such as AdaBoost and LogitBoost. The FURIA-GA-AdaBoost and FURIA-GA-LogitBoost configurations significantly enhance the model's ability to discern complex patterns from multi-dimensional geomorphological data. The feature selection process incorporates a variety of topographical indices and hydrological variables (e.g., elevation, slope, and standardized precipitation index), which are essential for accurately identifying high-risk flood zones. Implementing tenfold cross-validation further enhances model robustness, ensuring that the generated flood susceptibility maps are accurate and generalizable to different flood events and geographical areas.

Water Level and Lake Forecasting Predicting water levels in lakes and reservoirs is a crucial hydrological application of EML methods, particularly for addressing water resource management challenges in arid and semi-arid regions. EML models, such as bagging-based tree models, offer powerful capabilities for managing large-scale, multi-dimensional hydrological datasets, often including satellite imagery, climatic records, hydrometric data, and in-situ measurements. These models effectively integrate diverse input scenarios, including variations in climatic and hydrological conditions, to generate robust and dynamic predictions, which are essential for sustainable water management strategies.

- **Lake Water Level Forecasting:** A notable example of this application is the use of ensemble bagging-based tree models to forecast water levels at Lake Urmia (Hajian et al. 2022). This study developed an ensemble of multiple bagging-based tree models, including Bagging-RF (BA-RF) and Bagging-Alternating Model Tree (BA-AMT), to predict lake water levels across different lead times (1, 2, 3, and 6 months). These models leverage diverse input data, including climatic records, hydrological observations, and satellite-based remote sensing data, enabling a comprehensive assessment of water level changes under various environmental and anthropogenic scenarios. By incorporating EML techniques, these models can account for complex climatic and hydrological dynamics, making them a critical tool for informed decision-making in managing Lake Urmia's water resources, particularly considering the region's severe water scarcity challenges.
- **Water Pumping Station Predictions:** Another important application of EML methods in hydrology is predicting water levels in pumping stations for large-scale water transfer projects, such as those in China (Zhang et al. 2023b). In this case, a bagging-based ensemble model was developed, integrating Convolutional Neural Networks (CNN) and LSTM networks to predict water levels at pumping stations. This hybrid architecture within the bagging ensemble framework enables effective forecasting of historical time series data, ensuring optimal operation of water transfer systems. The model incorporates a self-attention mechanism during the feature selection process, enhancing its ability to prioritize and extract relevant features from high-dimensional datasets. By dynamically adapting to the temporal dependencies and spatial patterns in the data, the CNN-LSTM architecture improves both the interpretability and predictive accuracy of the model, facilitating the proactive management of risks such as reservoir overflow or drying, which are critical in large-scale water transfer projects.

Enhancing Global Hydrological Models EML models are also employed to improve the performance of traditional process-based hydrological models. In these cases, ensemble learning methods enhance streamflow simulations and correct model errors.

- **Error Correction:** A notable study (Shen et al. 2022a) applied RF to correct the errors of a large-scale hydrological model (PCR-GLOBWB) in Europe. By incorporating model state variables, the study significantly improves the accuracy of streamflow predictions by reducing errors in model state variables across several countries, including Germany, the Netherlands, Switzerland, and France.
- **Ensemble-Based Hybrid Model Integration:** Studies such as (Lei et al. 2024) from China use hybrid ensemble models to integrate multi-source data with process-based hydrological models like Block-wise use of TOPMODEL (BTOP), improving streamflow forecasting.

Advancing Hydrological Uncertainty Quantification Managing uncertainty in hydrological predictions is essential for optimizing the performance of EML models. Integrating Bayesian optimization and probabilistic forecasting within ensemble frameworks enhances uncertainty quantification, leading to more robust decision-making.

- **Uncertainty in Streamflow Prediction**: (Bai et al. 2021) presents a novel approach using Bayesian optimization to fine-tune XGBoost and GPR for probabilistic hydrological forecasts in the Yangtze River Basin, China. The XGB-GPR-BOA hybrid model is designed to minimize uncertainties in key predictive variables, providing accurate probabilistic streamflow and runoff predictions. Key preprocessing techniques include feature selection using maximum mutual information (MIC) to identify the most relevant historical data correlations, normalization of runoff data to ensure consistent scale across models, and Data Transformation to format input features and labels appropriately for the hybrid model. These preprocessing steps contribute to the enhanced efficiency and stability of the model.

Limitations and future work

Despite the promising advancements in EML methods for hydrological applications, several limitations persist that warrant further investigation. These challenges primarily stem from the inherent complexity of hydrological processes, the diversity of data sources, and the variability of environmental conditions.

Limitations

1. **Data Availability and Quality:** Many EML studies rely on historical datasets that may be incomplete or biased. The need for high-resolution spatial and temporal data, particularly in remote or under-monitored regions, limits the robustness and generalizability of the developed models. Data preprocessing techniques such as normalization and outlier detection remain critical yet often need to be explored steps that can significantly impact model performance.
2. **Model Interpretability:** A significant challenge with complex EML architectures, particularly hybrid and

stacked models, is their 'black-box' nature. While offering high predictive accuracy, the lack of transparency in their decision-making processes hinders understanding the underlying hydrological relationships captured and limits trust among stakeholders for informed decision-making. Our systematic review found that despite this challenge being acknowledged conceptually in some studies, the routine application and detailed reporting of methods to enhance model interpretability were minimal in the reviewed literature, suggesting a gap between recognizing the problem and implementing solutions.

3. **Scalability and Computational Complexity:** The computational demands of training and deploying sophisticated EML models can hinder their scalability, particularly in real-time applications. The trade-off between model complexity and operational efficiency is a pertinent concern, especially when applying these models to large-scale hydrological systems.
4. **Dynamic Environmental Changes:** The impact of climate change and anthropogenic activities on hydrological cycles introduces uncertainties that existing models may not adequately account for. EML methods often assume stationary conditions; thus, their adaptability to evolving climatic scenarios is an area requiring more thorough exploration.
5. **Handling Data Heterogeneity and Noise:** Hydrological datasets often vary significantly across regions and time periods, with different sources contributing to data heterogeneity and noise. This variability can negatively impact EML model performance and stability, especially when attempting to generalize across different hydrological settings.
6. **Overfitting and Generalization Challenges:** While ensemble techniques are generally robust, EML models can still suffer from overfitting, especially when trained on small or highly specific datasets. This overfitting reduces their ability to generalize to unseen data, limiting their effectiveness in diverse hydrological scenarios.
7. **Lack of Standardized Evaluation Metrics:** Studies use various metrics to evaluate model performance (e.g., RMSE, MAE, R2), making it challenging to compare results across studies. The lack of standardized metrics limits the ability to benchmark EML models effectively, impeding a clear understanding of which models are best suited for specific hydrological applications.
8. **Scope of Literature Search:** The bibliometric analysis was limited to the WoS database. While WoS is a leading source for high-impact research, the exclusion of other relevant databases, such as Scopus, means that some relevant publications and potentially unique trends indexed primarily in those sources might not be fully represented, which could offer a broader perspective on the field's landscape.
9. **Variability in Model Configuration Reporting and Hyperparameter Details:** While our review synthesizes the predominant types and overall performance trends of EML models based on reported metrics, providing a detailed, comparative analysis of specific model configurations, such as the optimal number of trees in RF, the architecture details (e.g., number of layers, neurons) of NNs, or the specific hyperparameters and tuning processes used across all reviewed studies, is challenging. The level of detail and consistency in reporting these crucial configuration parameters varies significantly across the literature. This variability limits the ability of a review study like ours to quantitatively compare the precise influence of these fine-grained model choices on performance across different hydrological applications and datasets, and also poses significant challenges for reproducibility and benchmarking in the field.

Future work

1. **Enhanced Data Integration:** Future research should integrate diverse data sources, including remote sensing, Internet of Things (IoT) sensors, and socio-economic data, to build more comprehensive datasets. Leveraging advanced data assimilation techniques can improve model calibration and validation, enhancing predictive accuracy in variable environmental conditions.
2. **Interpretability and Explainability:** There is a pressing need for dedicated research to develop and apply frameworks for enhancing the interpretability of EML models in hydrology. This is crucial for building trust and enabling domain scientists and stakeholders to gain insights from these complex models. Future work should focus on integrating and applying model-agnostic interpretability tools, such as SHapley Additive exPlanations (SHAP) values, Local Interpretable Model-agnostic Explanations (LIME), permutation importance, and partial dependence plots. These techniques can provide valuable insights into the contributions of individual features and model components to predictions, helping to move beyond the 'black-box' limitation and facilitate better integration of EML findings with hydrological understanding.
3. **Dynamic Modeling Approaches:** Researchers should explore the application of adaptive learning techniques that allow EML models to continuously update as new data becomes available. This approach could improve the models' resilience to non-stationary conditions and enhance their applicability in real-world scenarios characterized by rapid environmental changes.
4. **Uncertainty Quantification:** Future studies should incorporate advanced techniques for uncertainty quantification, such as Bayesian frameworks or ensemble

forecasting methods, to better account for the inherent uncertainties in hydrological predictions. Probabilistic modeling approaches can yield more reliable predictions and inform risk management strategies.

5. **Interdisciplinary Collaboration:** Promoting interdisciplinary research that incorporates insights from hydrology, climatology, data science, and socio-economics is essential for developing more holistic EML frameworks. Collaborative efforts can facilitate the identification of novel challenges and innovative solutions that address complex hydrological issues.
6. **Improvement in Feature Engineering:** Future research should focus on advancing feature engineering techniques to extract meaningful patterns from hydrological datasets. Techniques like automatic feature selection, dimensionality reduction (e.g., through PCA), and domain-specific transformations could improve model interpretability and performance.
7. **Integration of Physics-Based Models with EML:** Combining data-driven EML models with physics-based hydrological models could bridge the gap between theoretical understanding and empirical data analysis. This integration would allow for more accurate predictions that incorporate both data patterns and established physical principles, potentially enhancing model robustness and accuracy.
8. **Addressing Climate Change Impacts:** With climate change introducing unprecedented variability, future work should prioritize developing models that are resilient to these impacts. Scenario-based modeling, where EML models are tested and optimized under different climate scenarios, could prepare these models to handle extreme weather events more effectively.
9. **Robust Hyperparameter Optimization and Model Tuning**: Future research should focus on developing and applying more systematic, efficient, and robust hyperparameter optimization techniques specifically tailored for EML models in diverse hydrological contexts. Further investigations into the sensitivity of EML performance to different hyperparameters, initialization strategies, and optimization algorithms across various datasets and applications are needed to provide clearer guidelines for practitioners, improve model reproducibility, and facilitate the identification of truly optimal model configurations.

Conclusion

This study provides a comprehensive exploration of the development and application of EML models in hydrological research, particularly focusing on streamflow prediction, flood forecasting, and other key hydrological challenges. Through an in-depth bibliometric analysis of 199 publications and a systematic review of 51 selected studies, our findings underscore the growing importance of advanced ensemble techniques in addressing the complex, non-linear relationships inherent in hydrological systems. The results indicate a clear preference for boosting methods, such as XGBoost and LightGBM, which consistently outperform traditional hydrological models, particularly in tasks related to streamflow prediction and flood susceptibility mapping. Hybrid models, combining ML techniques like ANN with ensemble methods, have proven to be particularly effective in improving the robustness and accuracy of hydrological predictions. Moreover, stacking techniques, which leverage the strengths of multiple base learners, offer superior performance in addressing the multifaceted nature of hydrological systems, making them highly suitable for complex prediction tasks such as multi-step forecasting and uncertainty quantification. Beyond individual model performance, this study highlights the significant advancements in the algorithmic development of EML models, with a notable shift towards hybridization and dynamic model integration. These advancements not only enhance predictive accuracy but also improve model adaptability across diverse hydrological conditions and data regimes. The bibliometric analysis reveals an accelerating trend in the publication of EML-related research, particularly from 2018 onward, reflecting the increasing recognition of these models' potential to transform water resource management and climate impact assessments. Despite these promising developments, several challenges remain. The interpretability of complex ensemble models continues to be a key area of concern, particularly when applied in decision-making processes. As EML models become more sophisticated, ensuring transparency and understanding of model outputs will be critical for their broader adoption in practice. Additionally, there is a need for further research into integrating multiple data sources, including satellite remote sensing and ground-based observations, to enhance the accuracy and reliability of hydrological models, particularly in data-scarce regions. In conclusion, the evolution of EML models in hydrology represents a significant leap forward in improving predictive capabilities for water resource management. This study contributes to the growing body of knowledge by mapping the trends and advancements in EML research and identifying critical areas for future exploration. As the hydrological impacts of climate change become more severe, the continued development of robust, interpretable, and adaptive EML models will be essential for advancing both academic research and practical applications in the field of water resources. This study underscores the transformative potential of EML models in hydrological research and water resource management. By systematically integrating ensemble techniques, researchers are now able to enhance predictive accuracy and manage complex, non-linear hydrological processes more effectively than ever before. These advancements offer practical benefits for flood forecasting, drought prediction, and sustainable water management, especially as climate variability increases. While EML techniques like boosting, bagging, and stacking have demonstrated significant improvements in predictive performance, challenges such as data quality, model interpretability, and computational demands remain substantial. Addressing these challenges will be essential to fully realize the potential of EML models in real-world hydrological applications.

Author Contributions Moein Tosan: Managing data collection and resources; assisting in the bibliometric analysis process; contributing to the review and editing of manuscript sections related to data preprocessing techniques and hybrid models. Vahid Nourani: Leading the conceptualization and design of the study as the principal investigator; developing the research framework and methodologies; supervising all phases of bibliometric and systematic review analyses; performing advanced data interpretation; drafting the manuscript's core sections, including critical discussions; and overseeing the review and finalization process. His expertise in hydrological modeling and machine learning significantly shaped the study's direction and contributions. Ozgur Kisi: Providing in-depth validation and technical oversight of bibliometric and systematic review findings; offering expert insights into hydrological applications of machine learning methods; contributing to refining the manuscript with a focus on technical accuracy and alignment with global hydrological challenges; and collaborating on the interpretation of results to ensure relevance to cutting-edge hydrological research. Mehdi Dastourani: Supporting data organization and management; providing constructive feedback on the systematic review findings; and participating in refining the manuscript to ensure clarity and coherence.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing Interests The authors declare no competing interests.

Consent to Publish Not applicable.

Consent to Participate Not applicable.

References

Abdulaleva BS (2023) Enhancing the performance of deep learning technique by combining with gradient boosting in rainfall-runoff simulation. J Water Land Dev 216–23-23. https://doi.org/10.24425/jwld.2023.148446

Adnan RM, Jaafari A, Mohanavelu A, Kisi O, Elbeltagi A (2021) Novel ensemble forecasting of streamflow using locally weighted learning algorithm. Sustainability 13(11):5877

Adnan RM, Mirboluki A, Mehraein M, Malik A, Heddam S, Kisi O (2024) Improved prediction of monthly streamflow in a mountainous region by metaheuristic-enhanced deep learning and machine learning models using hydroclimatic data. Theoret Appl Climatol 155(1):205–228

Aiyelokun O, Pham QB, Aiyelokun O, Linh NTT, Roy T, Anh DT, Łupikasza E (2024) Effectiveness of integrating ensemble-based feature selection and novel gradient boosted trees in runoff prediction: a case study in Vu Gia Thu Bon River Basin, Vietnam. Pure Appl Geophys 1–20. https://doi.org/10.1007/s00024-024-03486-0

Akbarian M, Saghafian B, Golian S (2023) Monthly streamflow forecasting by machine learning methods using dynamic weather prediction model outputs over Iran. J Hydrol 620:129480

Al-Areeq AM, Abba SI, Halder B, Ahmadianfar I, Heddam S, Demir V et al (2023) Flood subsidence susceptibility mapping using elastic-net classifier: new approach. Water Resour Manag 37(13):4985–5006

AlDahoul N, Momo MA, Chong K, Ahmed AN, Huang YF, Sherif M, El-Shafie A (2023) Streamflow classification by employing various machine learning models for peninsular Malaysia. Sci Rep 13(1):14574

Alizadeh B, Bafti AG, Kamangir H, Zhang Y, Wright DB, Franz KJ (2021) A novel attention-based LSTM cell post-processor coupled with bayesian optimization for streamflow prediction. J Hydrol 601:126526

Alsulamy S, Kumar V, Kisi O, Kedam N, Rathnayake N (2025) Enhancing water level prediction using ensemble machine learning models: a comparative analysis. Water Resour Manag 1–20. https://doi.org/10.1007/s11269-025-04142-5

Arabameri A, Saha S, Mukherjee K, Blaschke T, Chen W, Ngo PTT, Band SS (2020) Modeling spatial flood using novel ensemble artificial intelligence approaches in northern Iran. Remote Sens 12(20):3423

Aria M, Cuccurullo C (2017) bibliometrix: An R-tool for comprehensive science mapping analysis. J Informet 11(4):959–975

Attar NF, Sattari MT, Apaydin H (2024) A novel stochastic tree model for daily streamflow prediction based on a noise suppression hybridization algorithm and efficient uncertainty quantification. Water Resour Manag 38(6):1943–1964

Ayyalasomayajula MMT (2023) Innovative water quality prediction for efficient management using ensemble learning. Educ Adm Theory Pract 29(4):2374–2381

Bai H, Li G, Liu C, Li B, Zhang Z, Qin H (2021) Hydrological probabilistic forecasting based on deep learning and Bayesian optimization algorithm. Hydrol Res 52(4):927–943

Barrera-Animas AY, Oyedele LO, Bilal M, Akinosho TD, Delgado JMD, Akanbi LA (2022) Rainfall prediction: a comparative analysis of modern machine learning algorithms for time-series forecasting. Mach Learn Appl 7:100204

Birkle C, Pendlebury DA, Schnell J, Adams J (2020) Web of Science as a data source for research on scientific and scholarly activity. Quant Sci Stud 1(1):363–376

Bui DT, Tsangaratos P, Ngo P-TT, Pham TD, Pham BT (2019) Flash flood susceptibility modeling using an optimized fuzzy rule based feature selection technique and tree based ensemble methods. Sci Total Environ 668:1038–54

Chandran D, Chithra N (2025) Predictive performance of ensemble learning boosting techniques in daily streamflow simulation. Water Resour Manag 39(3):1235–1259

Chen C, Zhang Q, Kashani MH, Jun C, Bateni SM, Band SS et al (2022) Forecast of rainfall distribution based on fixed sliding window long short-term memory. Eng Appl Comput Fluid Mech 16(1):248–261

Chiang Y-M, Hao R-N, Zhang J-Q, Lin Y-T, Tsai W-P (2018) Identifying the sensitivity of ensemble streamflow prediction by artificial intelligence. Water 10(10):1341
Cui Z, Qing X, Chai H, Yang S, Zhu Y, Wang F (2021) Real-time rainfall-runoff prediction using light gradient boosting machine coupled with singular spectrum analysis. J Hydrol 603:127124
Dai W, Tang Y, Liao N, Zou S, Cai Z (2024) Urban flood prediction using ensemble artificial neural network: an investigation on improving model uncertainty. Appl Water Sci 14(6):144
Demissie Z, Rimal P, Seyoum WM, Dutta A, Rimmington G (2024) Flood susceptibility mapping: Integrating machine learning and GIS for enhanced risk assessment. Appl Comput Geosci 100183. https://doi.org/10.1016/j.acags.2024.100183
Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM (2021) How to conduct a bibliometric analysis: An overview and guidelines. J Bus Res 133:285–296
Du B, Wang M, Zhang J, Chen Y, Wang T (2024) Urban flood prediction based on PCSWMM and stacking integrated learning model. Nat Hazards 1–25. https://doi.org/10.1007/s11069-024-06893-7
Ehteram M, Achite M, Sheikh Khozani Z, Farrokhi A (2024) Drought prediction using ensemble models. Acta Geophys 72(2):945–982
Elbeltagi A, Di Nunno F, Kushwaha NL, De Marinis G, Granata F (2022) River flow rate prediction in the Des Moines watershed (Iowa, USA): A machine learning approach. Stoch Env Res Risk Assess 36(11):3835–3855
El-Mahdy ME-S, Mousa FA, Morsy FI, Kamel AF, El-Tantawi A (2024) Flood classification and prediction in South Sudan using artificial intelligence models under a changing climate. Alexandria Eng J 97:127–41. https://doi.org/10.1016/j.aej.2024.03.082
Farooq R (2024) A review of knowledge management research in the past three decades: a bibliometric analysis. VINE J Inf Knowl Manag Syst 54(2):339–378
Fijani E, Khosravi K (2023) Hybrid iterative and tree-based machine learning algorithms for lake water level forecasting. Water Resour Manag 37(14):5431–5457
Fohlmeister J, Luetscher M, Spötl C, Schröder-Ritzrau A, Schröder B, Frank N et al (2023) The role of Northern Hemisphere summer insolation for millennial-scale climate variability during the penultimate glacial. Commun Earth Environ 4(1):245
Gan M, Lai X, Guo Y, Chen Y, Pan S, Zhang Y (2024) Floodplain lake water level prediction with strong river-lake interaction using the ensemble learning LightGBM. Water Resour Manag 1–17. https://doi.org/10.1007/s11269-024-03915-8
Gauch M, Mai J, Lin J (2021) The proper care and feeding of CAMELS: How limited training data affects streamflow prediction. Environ Model Softw 135:104926
Gelete G (2023a) Application of hybrid machine learning-based ensemble techniques for rainfall-runoff modeling. Earth Sci Inf 16(3):2475–2495
Gelete G (2023b) Hybrid extreme gradient boosting and nonlinear ensemble models for suspended sediment load prediction in an agricultural catchment. Water Resour Manag 37(14):5759–5787
Gichamo T, Nourani V, Gökçekuş H, Gelete G (2024) Ensemble of artificial intelligence and physically based models for rainfall–runoff modeling in the upper Blue Nile Basin. Hydrol Res nh2024189.
Goodarzi MR, Niazkar M, Barzkar A, Niknam ARR (2024) Assessment of machine learning models for short-term streamflow estimation: the case of Dez River in Iran. Sustain Water Resour Manag 10(1):33
Granata F, Di Nunno F (2024) Forecasting short-and medium-term streamflow using stacked ensemble models and different meta-learners. Stoch Env Res Risk Assess 38(9):3481–3499
Granata F, Di Nunno F, de Marinis G (2022) Stacked machine learning algorithms and bidirectional long short-term memory networks for multi-step ahead streamflow forecasting: A comparative study. J Hydrol 613:128431
Guhan T, Revathy N (2024) EMLARDE tree: ensemble machine learning based random de-correlated extra decision tree for the forest cover type prediction. Signal, Image and Video Processing. https://doi.org/10.1007/s11760-024-03470-0
Gurbuz F, Mudireddy A, Mantilla R, Xiao S (2024) Using a physics-based hydrological model and storm transposition to investigate machine-learning algorithms for streamflow prediction. J Hydrol 628:130504
Habibi A, Delavar MR, Sadeghian MS, Nazari B, Pirasteh S (2023) A hybrid of ensemble machine learning models with RFE and Boruta wrapper-based algorithms for flash flood susceptibility assessment. Int J Appl Earth Obs Geoinf 122:103401
Hadi SJ, Abba SI, Sammen SS, Salih SQ, Al-Ansari N, Yaseen ZM (2019) Non-linear input variable selection approach integrated with non-tuned data intelligence model for streamflow pattern simulation. IEEE Access 7:141533–141548
Hajian R, Jalali MR, Mastouri R (2022) Multi-step Lake Urmia water level forecasting using ensemble of bagging based tree models. Earth Sci Inf 15(4):2515–2543
Heydari S, Reza Nikoo M, Mohammadi A, Barzegar R (2024) Two-stage meta-ensembling machine learning model for enhanced water quality forecasting. J Hydrol 641:131767
Hmouda AM, Orzes G, Sauer PC (2024) Sustainable supply chain management in energy production: a literature review. Renew Sustain Energy Rev 191:114085
Hosseini F, Prieto C, Álvarez C (2025) Ensemble learning of catchment-wise optimized LSTMs enhances regional rainfall-runoff modelling− case study: Basque Country. Spain J Hydrol 646:132269
Ikram RMA, Hazarika BB, Gupta D, Heddam S, Kisi O (2023) Streamflow prediction in mountainous region using new machine learning and data preprocessing methods: a case study. Neural Comput Appl 35(12):9053–9070
Ilia I, Tsangaratos P, Tzampoglou P, Chen W, Hong H (2022) Flash flood susceptibility mapping using stacking ensemble machine learning models. Geocarto Int 37(27):15010–15036
Inan CA, Maoui A, Lucas Y, Duplay J (2024) Multi-station hydrological modelling to assess groundwater recharge of a vast semi-arid basin considering the problem of lack of sata: a case study in Seybouse Basin, Algeria. Water 16(1):160
Islam ARMT, Bappi MMR, Alqadhi S, Bindajam AA, Mallick J, Talukdar S (2023) Improvement of flood susceptibility mapping by introducing hybrid ensemble learning algorithms and high-resolution satellite imageries. Nat Hazards 119(1):1–37
Jayawardena AW (2021) Fluid mechanics, hydraulics, hydrology and water resources for civil engineers: CRC Press. https://doi.org/10.1201/9780429423116
Jiang Z, Yang S, Liu Z, Xu Y, Shen T, Qi S et al (2022) Can ensemble machine learning be used to predict the groundwater level dynamics of farmland under future climate: a 10-year study on Huaibei Plain. Environ Sci Pollut Res 29(29):44653–44667
Jiang L, Chen P, Liao J, Jiang C, Liang W, Xiong NN (2024) PCFS: an intelligent imbalanced classification scheme with noisy samples. Inf Sci 679:121020
Jin J, Wang L, Yang J, Si B, Niu G-Y (2022) Improved runoff simulations for a highly varying soil depth and complex terrain watershed in the Loess Plateau with the community land model version 5. Geosci Model Dev 15(8):3405–3416
Karamvand A, Hosseini SA, Azizi SA (2024) Enhancing streamflow simulations with gated recurrent units deep learning models in the flood prone region with low-convergence streamflow data. Phys Chem Earth a/b/c 136:103737
Katipoğlu OM, Keblouti M, Mohammadi B (2023) Application of novel artificial bee colony optimized ANN and data preprocessing techniques for monthly streamflow estimation. Environ Sci Pollut Res 30(38):89705–89725

Khandelwal A, Xu S, Li X, Jia X, Stienbach M, Duffy C et al (2020) Physics guided machine learning methods for hydrology. arXiv preprint arXiv:201202854. https://doi.org/10.48550/arXiv.2012.02854
Khosravi K, Golkarian A, Omidvar E, Hatamiafkoueieh J, Shirali M (2023) Snow water equivalent prediction in a mountainous area using hybrid bagging machine learning approaches. Acta Geophys 71(2):1015–1031
Khosravi K, Attar N, Bateni SM, Jun C, Kim D, Safari MJS et al (2024) Daily river flow simulation using ensemble disjoint aggregating M5-Prime model. Heliyon. https://doi.org/10.1016/j.heliyon.2024.e37965
Khuat TT, Bassett R, Otte E, Gabrys B (2025) Uncertainty quantification using ensemble learning and Monte Carlo sampling for performance prediction and monitoring in cell culture processes. J Raman Spectroscopy. https://doi.org/10.48550/arXiv.2409.02149
Kigo SN, Omondi EO, Omolo BO (2023) Assessing predictive performance of supervised machine learning algorithms for a diamond pricing model. Sci Rep 13(1):17315
Kilinc HC, Haznedar B, Ozkan F, Katipoğlu OM (2024) An evolutionary hybrid method based on particle swarm optimization algorithm and extreme gradient boosting for short-term streamflow forecasting. Acta Geophys 1–21. https://doi.org/10.1007/s11600-024-01307-5
Kisi O, Heddam S, Parmar KS, Yaseen ZM, Kulls C (2024) Improved monthly streamflow prediction using integrated multivariate adaptive regression spline with K-means clustering: implementation of reanalyzed remote sensing data. Stoch Env Res Risk Assess 38(6):2489–2519
Kossieris P, Tsoukalas I, Brocca L, Mosaffa H, Makropoulos C, Anghelea A (2024) Precipitation data merging via machine learning: revisiting conceptual and technical aspects. J Hydrol 131424. https://doi.org/10.1016/j.jhydrol.2024.131424
Kumar V, Kedam N, Sharma KV, Mehta DJ, Caloiero T (2023) Advanced machine learning techniques to improve hydrological prediction: a comparative analysis of streamflow prediction models. Water 15(14):2572
Kumar S, Oliveto G, Deshpande V, Agarwal M, Rathnayake U (2024) Forecasting of time-dependent scour depth based on bagging and boosting machine learning approaches. J Hydroinf 26(8):1906–1928
Kumar M, Agrawal Y, Adamala S, Pushpanjali, Subbarao A, Singh V, Srivastava A (2024) Generalization ability of bagging and boosting type deep learning models in evapotranspiration estimation. Water 16(16):2233
Kumshe UMM, Abdulhamid ZM, Mala BA, Muazu T, Muhammad AU, Sangary O et al (2024) Improving short-term daily streamflow forecasting using an autoencoder based CNN-LSTM Model. Water Resour Manag 1–17
Lee E, Kam J (2023) Deciphering the black box of deep learning for multi-purpose dam operation modeling via explainable scenarios. J Hydrol 626:130177
Lee J, Bateni SM, Jun C, Heggy E, Jamei M, Kim D et al (2024) Hybrid machine learning system based on multivariate data decomposition and feature selection for improved multitemporal evapotranspiration forecasting. Eng Appl Artif Intell 135:108744
Lei H, Li H, Hu W (2024) Enhancing the streamflow simulation of a process-based hydrological model using machine learning and multi-source data. Eco Inform 82:102755
Leng Z, Chen L, Yang B, Li S, Yi B (2024) An extreme forecast index-driven runoff prediction approach using stacking ensemble learning. Geomat Nat Haz Risk 15(1):2353144
Lin Y, Wang D, Meng Y, Sun W, Qiu J, Shangguan W et al (2023) Bias learning improves data driven models for streamflow prediction. J Hydrol Region Stud 50:101557
Linh NTT, Pandey M, Janizadeh S, Bhunia GS, Norouzi A, Ali S et al (2022) Flood susceptibility modeling based on new hybrid intelligence model: optimization of XGboost model using GA metaheuristic algorithm. Adv Space Res 69(9):3301–3318
Long G, Tantanee S, Nusit K, Sooraksa P (2025) Flood susceptibility mapping in the Yom River Basin, Thailand: stacking ensemble learning using multi-year flood inventory data. Geocarto Int 40(1):2461531
Mahdavi-Meymand A, Sulisz W, Zounemat-Kermani M (2024) Hybrid and integrative evolutionary machine learning in hydrology: a systematic review and meta-analysis. Arch Comput Methods Eng 31(3):1297–1340
Marcus E, Teuwen J (2024) Artificial intelligence and explanation: how, why, and when to explain black boxes. Eur J Radiol 111393. https://doi.org/10.1016/j.ejrad.2024.111393
Mehraein M, Mohanavelu A, Naganna SR, Kulls C, Kisi O (2022) Monthly streamflow prediction by metaheuristic regression approaches considering satellite precipitation data. Water 14(22):3636
Molajou A, Nourani V, Tajbakhsh AD, Variani HA, Khosravi M (2024) Multi-step-ahead rainfall-runoff modeling: decision tree-based clustering for hybrid wavelet neural-networks modeling. Water Resour Manag 1–20. https://doi.org/10.21203/rs.3.rs-4185983/v1
Momeneh S, Nourani V. Performance evaluation of artificial neural network model in hybrids with various preprocessors for river streamflow forecasting. AQUA—Water Infrastruct Ecosyst Soc 72(6):947–68. https://doi.org/10.2166/aqua.2023.010
Momeneh S, Nourani V (2022) Application of a novel technique of the multi-discrete wavelet transforms in hybrid with artificial neural network to forecast the daily and monthly streamflow. Model Earth Syst Environ 8(4):4629–4648
Muhammad AU, Abba SI (2023) Transfer learning for streamflow forecasting using unguaged MOPEX basins data set. Earth Sci Inf 16(2):1241–1264
Muhammad AU, Djigal H, Muazu T, Adam JM, Ba AF, Dabai US et al (2023) An autoencoder-based stacked LSTM transfer learning model for EC forecasting. Earth Sci Inf 16(4):3369–3385
Naganna SR, Marulasiddappa SB, Balreddy MS, Yaseen ZM (2023) Daily scale streamflow forecasting in multiple stream orders of Cauvery River, India: application of advanced ensemble and deep learning models. J Hydrol 626:130320
Nhu V-H, Thi Ngo P-T, Pham TD, Dou J, Song X, Hoang N-D et al (2020) A new hybrid firefly–PSO optimized random subspace tree intelligence for torrential rainfall-induced flash flood susceptible mapping. Remote Sensing 12(17):2688
Ni L, Wang D, Wu J, Wang Y, Tao Y, Zhang J, Liu J (2020) Streamflow forecasting using extreme gradient boosting model coupled with Gaussian mixture model. J Hydrol 586:124901
Niyogisubizo J, Liao L, Zou F, Han G, Nziyumva E, Li B, Lin Y (2023) Predicting traffic crash severity using hybrid of balanced bagging classification and light gradient boosting machine. Intell Data Anal 27:79–101
Nourani V, Najafi H (2022) Historical changes in hydroclimatic extreme events over Iran. Elsevier, Climate Impacts on Extreme Weather, pp 101–115
Nourani V, Tapeh AHG, Khodkar K, Huang JJ (2023) Assessing long-term climate change impact on spatiotemporal changes of groundwater level using autoregressive-based and ensemble machine learning models. J Environ Manag 336:117653
Oruc S, Hinis MA, Tugrul T (2024) Evaluating performances of LSTM, SVM, GPR, and RF for drought prediction in Norway: a wavelet decomposition approach on regional forecasting. Water 16(23):3465
Pham BT, Jaafari A, Van Phong T, Yen HPH, Tuyen TT, Van Luong V et al (2021) Improved flood susceptibility mapping using a best

first decision tree integrated with ensemble learning techniques. Geosci Front 12(3):101105
Prasad P, Loveson VJ, Kotha M (2023) Probabilistic coastal wetland mapping with integration of optical, SAR and hydro-geomorphic data through stacking ensemble machine learning model. Eco Inform 77:102273
Rathnayake N, Rathnayake U, Dang TL, Hoshino Y (2023) Water level prediction using soft computing techniques: a case study in the Malwathu Oya, Sri Lanka. PLoS One 18(4):e0282847
Razavi-Termeh SV, Sadeghi-Niaraki A, Seo M, Choi S-M (2023a) Application of genetic algorithm in optimization parallel ensemble-based machine learning algorithms to flood susceptibility mapping using radar satellite imagery. Sci Total Environ 873:162285
Razavi-Termeh SV, Seo M, Sadeghi-Niaraki A, Choi S-M (2023b) Flash flood detection and susceptibility mapping in the Monsoon period by integration of optical and radar satellite imagery using an improvement of a sequential ensemble algorithm. Weather Clim Extreme 41:100595
T R J, Reddy NVS, Acharya UD (2023) Modeling daily reference evapotranspiration from climate variables: assessment of bagging and boosting regression approaches. Water Resour Manag 37(3):1013–32
Ren W, Li X, Zheng D, Zeng R, Su J, Mu T, Wang Y (2023) Enhancing flood simulation in data-limited glacial river basins through hybrid modeling and multi-source remote sensing data. Remote Sens 15(18):4527
Riazi M, Khosravi K, Shahedi K, Ahmad S, Jun C, Bateni SM, Kazakis N (2023) Enhancing flood susceptibility modeling using multi-temporal SAR images, CHIRPS data, and hybrid machine learning algorithms. Sci Total Environ 871:162066
Sarkis-Onofre R, Catalá-López F, Aromataris E, Lockwood C (2021) How to properly use the PRISMA Statement. Syst Rev 10:1–3
Sayed BT, Al-Mohair HK, Alkhayyat A, Ramírez-Coronel AA, Elsahabi M (2023) Comparing machine-learning-based black box techniques and white box models to predict rainfall-runoff in a northern area of Iraq, the Little Khabur River. Water Sci Technol 87(3):812–822
Shen Y, Ruijsch J, Lu M, Sutanudjaja EH, Karssenberg D (2022a) Random forests-based error-correction of streamflow from a large-scale hydrological model: Using model state variables to estimate error terms. Comput Geosci 159:105019
Shen K, Qin H, Zhou J, Liu G (2022b) Runoff probability prediction model based on natural Gradient boosting with tree-structured parzen estimator optimization. Water 14(4):545
Shim J, Hong S, Lee J, Lee S, Kim YM, Chon K et al (2023) Deep learning with data preprocessing methods for water quality prediction in ultrafiltration. J Clean Prod 428:139217
Sikorska-Senoner AE, Quilty JM (2021) A novel ensemble-based conceptual-data-driven approach for improved streamflow simulations. Environ Model Softw 143:105094
Singh U, Maca P, Hanel M, Markonis Y, Nidamanuri RR, Nasreen S et al (2023) Hybrid multi-model ensemble learning for reconstructing gridded runoff of Europe for 500 years. Inf Fusion 97:101807
Smirnova N, Mayr P (2023) A comprehensive analysis of acknowledgement texts in web of science: a case study on four scientific domains. Scientometrics 128(1):709–734
Sun X, Zhang H, Wang J, Shi C, Hua D, Li J (2022) Ensemble streamflow forecasting based on variational mode decomposition and long short term memory. Sci Rep 12(1):518
Takai Eddine Y, Nadir M, Sabah S, Jaafari A (2024) Integrating support vector machines with different ensemble learners for improving streamflow simulation in an ungauged watershed. Water Resour Manag 38(2):553–567
Talukdar S, Ghose B, Shahfahad Salam R, Mahato S, Pham QB et al (2020) Flood susceptibility modeling in Teesta River basin, Bangladesh using novel ensembles of bagging algorithms. Stoch Environ Res Risk Assess 34:2277–300
Tang Y, Wu Q, S-e-h Soomro, Li X, Sun Y, Hu C (2023) Comparison of different ensemble precipitation forecast system evaluation, integration and hydrological applications. Acta Geophys 71(1):405–21
Tien Bui D, Shirzadi A, Chapi K, Shahabi H, Pradhan B, Pham BT et al (2019) A hybrid computational intelligence approach to groundwater spring potential mapping. Water 11(10):2013
Tosan M, Gharib MR, Attar NF, Maroosi A (2025) Enhancing evapotranspiration estimation: a bibliometric and systematic review of hybrid neural networks in water resource management. Comput Model Eng Sci (CMES) 142(2)
Tuğrul T, Hinis MA (2025) Performance enhancement of models through discrete wavelet transform for streamflow forecasting in Çarşamba River, Türkiye. J Water Clim Chang 16(2):736–756
Tuğrul T, Hınıs MA, Oruç S (2025) Comparison of LSTM and SVM methods through wavelet decomposition in drought forecasting. Earth Sci Inf 18(1):1–21
Tyralis H, Papacharalampous G, Langousis A (2021) Super ensemble learning for daily streamflow forecasting: Large-scale demonstration and comparison with multiple machine learning algorithms. Neural Comput Appl 33(8):3053–3068
Ullah B, Fawad M, Khan AU, Mohamand SK, Khan M, Iqbal MJ, Khan J (2023) Futuristic streamflow prediction based on CMIP6 scenarios using machine learning models. Water Resour Manag 37(15):6089–6106
Wang H, Guan X, Meng Y, Wang H, Xu H, Liu Y et al (2024) Risk prediction based on oversampling technology and ensemble model optimized by tree-structured parzed estimator. Int J Disaster Risk Reduct 111:104753
Wegayehu EB, Muluneh FB. Super ensemble based streamflow simulation using multi-source remote sensing and ground gauged rainfall data fusion. Heliyon 9(7). https://doi.org/10.1016/j.heliyon.2023.e17982
Wegayehu EB, Muluneh FB (2024) Comparing conceptual and super ensemble deep learning models for streamflow simulation in data-scarce catchments. J Hydrol Region Stud 52:101694
Wi S, Steinschneider S (2022) Assessing the physical realism of deep learning hydrologic model projections under climate change. Water Resour Res 58(9):e2022WR032123
Wu X, Wang J (2023) Application of bagging, boosting and stacking ensemble and easyensemble methods for landslide susceptibility mapping in the three gorges reservoir area of China. Int J Environ Res Public Health 20(6):4977
Xia J, Dong Y, Zou L (2022) Developing socio-hydrology: Research progress, opportunities and challenges. J Geog Sci 32(11):2131–2146
Xia L (2023) Optimization of performance management for commercial companies by integrating ROF and light GBM algorithms. 2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIIE): IEEE. https://doi.org/10.1109/AIKIIE60097.2023.10390474
Xu T, Liang F (2021) Machine learning for hydrologic sciences: an introductory overview. Wiley Interdiscip Rev Water 8(5):e1533
Xu R, Zhang D (2024) Forward prediction and surrogate modeling for subsurface hydrology:a review of theory-guided machine-learning approaches. Comput Geosci 188:105611
Xu J, Su Q, Li X, Ma J, Song W, Zhang L, Su X (2024b) A spatial downscaling framework for SMAP soil moisture based on stacking strategy. Remote Sens 16(1):200
Xu Q, Shi Y, Bamber J, Tuo Y, Ludwig R, Zhu XX (2023) Physics-aware machine learning revolutionizes scientific paradigm for

machine learning and process-based hydrology. arXiv preprint arXiv:231005227. https://doi.org/10.48550/arXiv.2310.05227
Xu C, Zhong P-a, Zhu F, Xu B, Wang Y, Yang L et al (2024) A hybrid model coupling process-driven and data-driven models for improved real-time flood forecasting. J Hydrol 131494. https://doi.org/10.1016/j.jhydrol.2024.131494
Yang H, Yao R, Dong L, Sun P, Zhang Q, Wei Y et al (2024) Advancing flood susceptibility modeling using stacking ensemble machine learning: a multi-model approach. J Geog Sci 34(8):1513–1536
Yao J, Zhang X, Luo W, Liu C, Ren L (2022) Applications of Stacking/Blending ensemble learning approaches for evaluating flash flood susceptibility. Int J Appl Earth Obs Geoinf 112:102932
Yariyan P, Janizadeh S, Van Phong T, Nguyen HD, Costache R, Van Le H et al (2020) Improvement of best first decision trees using bagging and dagging ensembles for flood probability mapping. Water Resour Manag 34(9):3037–3053
Yaseen ZM (2023) A new benchmark on machine learning methodologies for hydrological processes modelling: a comprehensive review for limitations and future research directions. Knowl-Base Eng Sci 4(3):65–103
Yu P-S, Yang T-C, Chen S-Y, Kuo C-M, Tseng H-W (2017) Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting. J Hydrol 552:92–104
Yu X, Wang Y, Wu L, Chen G, Wang L, Qin H (2020) Comparison of support vector regression and extreme gradient boosting for decomposition-based data-driven 10-day streamflow forecasting. J Hydrol 582:124293
Zandi O, Zahraie B, Nasseri M, Behrangi A (2022) Stacking machine learning models versus a locally weighted linear model to generate high-resolution monthly precipitation over a topographically complex area. Atmos Res 272:106159
Zhang Y, Ma J, Liang S, Li X, Liu J (2022) A stacking ensemble algorithm for improving the biases of forest aboveground biomass estimations from multiple remotely sensed datasets. Gisci Remote Sens 59(1):234–249
Zhang Q, Zhang F, Erfani T, Zhu L (2023a) Bagged stepwise cluster analysis for probabilistic river flow prediction. J Hydrol 625:129995
Zhang J, Yu Y, Yan J, Chen J (2023b) Data-driven parameter prediction of water pumping station. Water 15(6):1128
Zhang J, Kong D, Li J, Qiu J, Zhang Y, Gu X, Guo M (2025) Comparison and integration of hydrological models and machine learning models in global monthly streamflow simulation. J Hydrol 650:132549
Zounemat-Kermani M, Batelaan O, Fadaee M, Hinkelmann R (2021) Ensemble machine learning paradigms in hydrology: a review. J Hydrol 598:126266
Żywiec J, Szpak D, Wartalska K, Grzegorzek M (2024) The impact of climate change on the failure of water supply infrastructure: a bibliometric analysis of the current state of knowledge. Water 16(7):1043