Green Technologies and Sustainability 2 (2024) 100104

Contents lists available at ScienceDirect

Green Technologies and Sustainability

journal homepage: https://www.keaipublishing.com/en/journals/green-technologies-and-sustainability/

Review article

A comprehensive review towards resilient rainfall forecasting models using artificial intelligence techniques

Md. Abu Saleh *, H.M. Rasel, Briti Ray

Department of Civil Engineering, Rajshahi University of Engineering and Technology, Rajshahi 6204, Bangladesh

ARTICLE INFO

Keywords:
Rainfall
Forecasting
Artificial intelligence
Resilience

ABSTRACT

Rainfall is one of the remarkable hydrologic variables that is directly connected to the sustainable environment for any region over the globe. The present study aims to review different research papers on rainfall forecasting using artificial intelligence (AI) models including a bibliographic assessment of the most popular AI models and a comparison of the results based on the accuracy parameters. 39 journal papers, published in renowned international journals from 2000 to 2023, were studied extensively to categorize modeling techniques, best models, characteristics of input data, the period for the input variables, data division, and so forth. Although certain drawbacks still exist, the results of reviewed studies suggest that AI models may help simulate rainfall in various geographic locations. In some cases, the data splitting mechanism was delivered to the model itself so that the model accuracy gets improved. The recommendations from the reviewed papers will help future researchers fill the research gaps, especially tuning the hyperparameters while building the training models. Hybrid models were advised in some cases to minimize the gap between the simulated and the observed data. All recommendations from reviewed papers aimed to achieve a resilient rainfall forecasting model in the era of climate change.

1. Introduction

In the realm of rainfall modeling, conventional practice has relied upon conceptual or physical-based models as the primary analytical tool [1–14]. However, the models lack the requirement for substantial data sets and input features [15–17]. In several instances, there is a scarcity of data, but the paramount importance lies in acquiring precise predictions rather than comprehending the underlying mechanics [18–21]. Consequently, employing black-box artificial intelligence models can serve as a viable substitute [22,23]. Various approaches exist for modeling and predicting rainfall, including conceptual, physical, numerical, statistical, and others [24].

AI techniques are rapidly gaining traction due to their user-centric design and consistently impressive performance [25,26]. Numerous studies have explored the effectiveness of AI models in rainfall modeling across various regions worldwide. This study presents a comprehensive evaluation of scholarly works that have employed artificial intelligence (AI) techniques for the purpose of modeling and forecasting rainfall patterns. Naturally, these methodologies exhibit certain limitations, including but not limited to issues such as overfitting, limited applicability to broader contexts, potential reliance on irrelevant data, improper modeling using unsuitable techniques, and similar concerns [27–32]. Nevertheless, the ease of use, fast execution, and satisfactory precision of these methods, despite their lack of consideration for the complexities of physics, have prompted numerous researchers to employ them. It is critical to recognize that the qualities or constraints of AI models may prohibit them from successfully forecasting comparable time series, even if they were initially built for that reason.

The fundamental benefit of AI models, however, is their capacity to simulate nonlinear and complicated events without requiring a thorough comprehension of their internal mechanics [33]. Rainfall modeling using AI methodologies has witnessed a steady increase in popularity and attracted substantial attention from researchers around the world. To find out the probable scope of research based on the research gaps, the investigation regarding AI models in rainfall forecasting is needed for researchers to learn what literature and methodologies have already conducted in this regard. Though, AI models in hydrological parameters were studied in different studies, extensive literature on rainfall forecasting using different AI models is in dire need.

The present study reviewed the features and capabilities of different applied AI models for rainfall forecasting based on 39 papers. From these papers, the key information such as input variables, applied models, best output models, time span for input data, data splitting for training, testing, and validating the models were extracted. Moreover, these papers were collected from renowned publishing authorities like Elsevier and Springer using relevant key words. Fig. 1 illustrates the frequency of publications in different years from 2000 to 2023. With

* Corresponding author.
E-mail addresses: saleh110308@acc.edu.bd (M.A. Saleh), hmrasel@ce.ruet.ac.bd (H.M. Rasel).

https://doi.org/10.1016/j.grets.2024.100104
Received 12 January 2024; Received in revised form 26 May 2024; Accepted 26 May 2024
Available online 27 May 2024

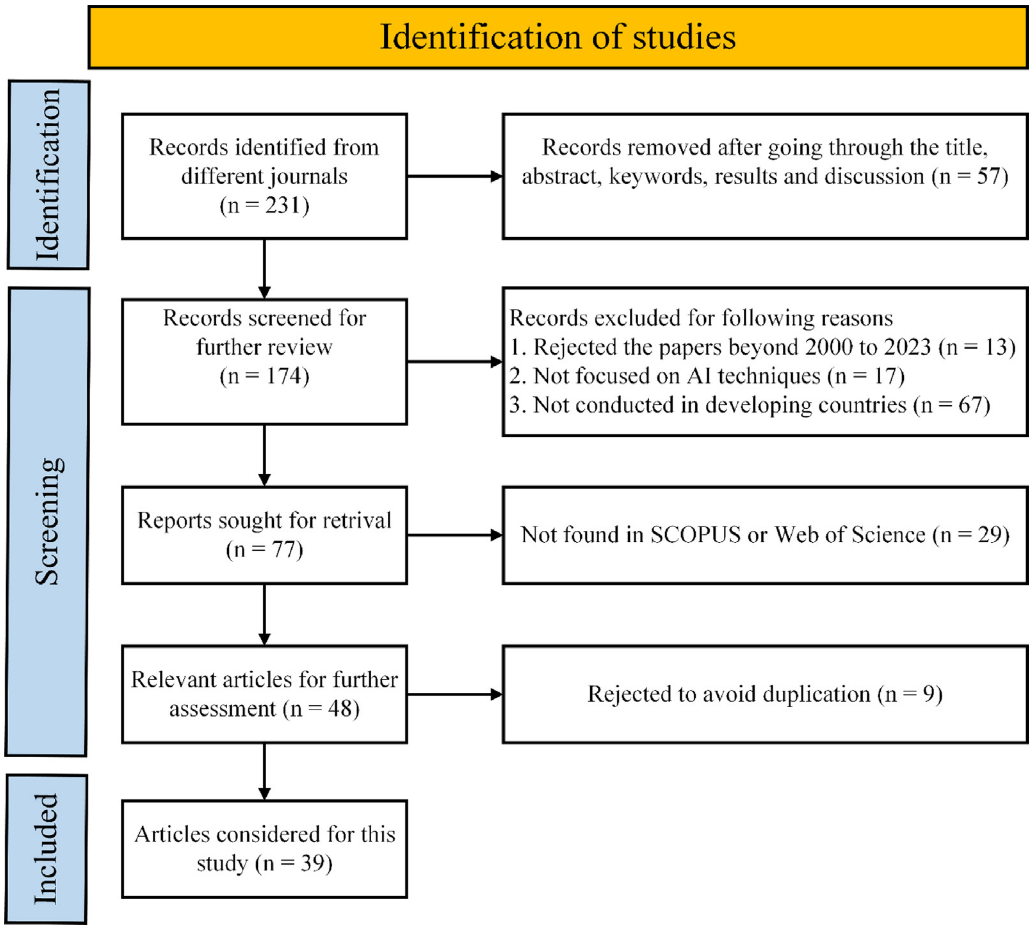

Fig. 1. PRISMA flow chart while selecting the papers.

the passage of time, the AI model quality was improved and the number of researches also. ANN, SVR, LSTM, NN, ANFIS and SVM were mostly found among the studies. Moreover, the other models include MLP, RF, ELM, XGBoost, RBF methods. An exhaustive breakdown of reviewed papers is listed in Table 1, covering details like authors and publication year, journal names with corresponding impact factors (based on 2022), the country of study, implemented AI models, top-performing models, utilized climatic input parameters, data time span, and the complete data range.

2. Importance of this study

Advances in rainfall predictions using AI modeling help farmers make informed planting, irrigation, and crop management decisions, improving agricultural production and food security [73,74]. Reliable rainfall predictions help water resource managers improve water allocation, mitigate drought, and prevent flooding, promoting efficient and sustainable water management. Rainfall forecasting in sustainability frameworks supports the global climate action agenda. Accurate predictions are needed to monitor and adapt to changing weather patterns and develop resilient solutions in the face of environmental unpredictability as climate change increases. In addition, water is extensively used in energy production for cooling, processing, and generation. Water supply and distribution, on the other hand, frequently rely on energy-intensive procedures. The use of energy sources and technologies can affect the availability and quality of water, and vice versa. Thus, this extensive literature review on rainfall forecasting by AI models is notably involved in the sustainable energy-water-environment nexus now and in the future.

Rainfall forecasting has long played a crucial role in mitigating the risks associated with floods, droughts, and landslides. However, the advent of Artificial Intelligence (AI) is ushering in a new era of precision and effectiveness in this domain. This technology holds immense potential to transform how we prepare for and respond to extreme weather events, ultimately saving lives and livelihoods. One of the most significant implications of AI-powered rainfall forecasting lies in its ability to create enhanced early warning systems. AI algorithms can analyze vast amounts of historical and real-time data, leading to far more accurate forecasts compared to traditional methods. This translates to earlier warnings, giving communities precious time to prepare for impending floods, droughts, or landslides. The ability to predict the severity and location of these events with greater accuracy allows authorities to take decisive action, minimizing casualties and property damage.

Furthermore, AI models can be trained on high-resolution data, enabling the creation of spatially localized forecasts. Rainfall patterns can vary significantly even within short distances. AI forecasts that account for these local variations empower communities to make more informed decisions based on their specific vulnerabilities. For instance, a village located near a riverbank might need to initiate evacuation procedures much sooner than one situated on higher ground during a flood warning. AI does not stop at point predictions. It can also generate probabilistic forecasts, indicating the likelihood of different rainfall amounts. This empowers communities to make risk-based decisions. Imagine farmers equipped with forecasts that not only predict a potential drought but also indicate the probability of varying degrees of severity. This allows them to take proactive measures such as planting drought-resistant crops or adjusting irrigation practices, potentially minimizing losses.

Table 1
Key observations found in the reviewed papers.

Sl. No.	Author & year	Journal with impact factor (IM/Cite score)	Study area	Used AI models	Input	Time steps	Span of datasets	Model parameters	Research Insights
1.	Luk et al. [34]	Journal of Hydrology (6.4)	Australia	ANN	Rainfall	15-Min	1991–1996	NMSE	Forecasted the spatial distribution of rainfall for an urban catchment.
2.	Luk et al. [35]	Mathematical and Computer Modelling	Australia	MLFN, PRNN, TDNN	Rainfall	15-Min	1991–1996	NMSE	Investigated the effect of temporal and spatial information on short-term rainfall forecasting using different orders of lag and different numbers of spatial inputs.
3.	Valverde Ramírez et al. [36]	Journal of Hydrology (6.4)	Brazil	ANN, MLR, ETA	Rainfall	Monthly	1992–2002	MAE, RMSE	ANN forecasts were superior to traditional linear regression models.
4.	Partal and Kişi [37]	Journal of Hydrology (6.4)	Turkey	Wavelet-Neuro-Fuzzy, Neuro-Fuzzy, ANN, MLR	P	Daily	1987–2001	RMSE, R^2	The wavelet-neuro-fuzzy performed better than the classical neuro-fuzzy model.
5.	Hong [38]	Applied Mathematics and Computation (4.0)	Taiwan	RSVRCPSO, SVRCPSO	Rainfall	Not Mentioned	Not Mentioned	NMSE, CE, CC	RSVRCPSO performed better among the other hybrid models.
6.	Bilgili and Sahin [39]	Energy Sources, Part A: Recovery, Utilization, and Environmental Effects (2.9)	Turkey	ANN	Rainfall, T	Monthly	1975–2006	MAE, R	Different neuron values were used for input (4) and output (2) layers in the model to find out the most reliable forecast.
7.	Mandal and Jothiprakash [40]	ISH Journal of Hydraulic Engineering (4.5 Cite Score)	India	ANN, MT	Rainfall	Daily	1961–2007	RMSE, R, CE	TLRN used least nodes and hidden layers among the other models during computational steps.
8.	Abbot and Marohasy [41]	Atmospheric Research (5.5)	Australia	ANN	Rainfall	Monthly	1997–2010	RMSE, R, MAE	Built a successful ANN prototype superior to the existing POAMA-1.5 GCM models in Australia.
9.	Nastos et al. [42]	Atmospheric Research (5.5)	Greece	ANN	P	Daily	1891–2009	RMSE, MBE, IA, R^2	Extreme rainfall values were overestimated and underestimated (in a considerable value) frequently.
10.	Farajzadeh et al. [43]	Water Resources and Industry (5.1)	Iran	NN, ARIMA	P	Daily	1973–2011	RMSE, MAE	Fed the rainfall values in the model as input and as output, they fed time lagged values using two hidden layers with 10 sigmoid hidden neurons.
11.	Abbot and Marohasy [44]	Atmospheric Research (5.5)	Australia	ANN	Rainfall	Monthly	1997–2011	RMSE, R, MAE	Introduced lagged relationships in the forecast model.
12.	Lin and Jhong [45]	Journal of Hydrology (6.4)	Taiwan	SVM, MGSVM	Rainfall	Hourly	1996–2009	MAE, EC	Optimized the input dimensions by identifying combinations of features that minimize redundant or irrelevant information.
13.	Yu et al. [46]	Journal of Hydrology (6.4)	Taiwan	RF, SVM	Rainfall	1-hour, 3-hour	2012–2015	RMSE	Single-mode models outperform multi-mode models. SVM performed better than RF in single-mode operation.
14.	Cramer et al. [47]	Expert Systems with Applications (8.5)	USA, Europe	MCRP, GP, SVR, RBF, KNN	Rainfall	Daily	1990–2009	RMSE	Forecasting with accumulated rainfall is more accurate than using daily values.
15.	Dash et al. [48]	Computers and Electrical Engineering (4.3)	India	KNN, ANN, ELM	Rainfall	Monthly	1871–2016	MAE, RMSE, MASE, PP	For the ELM model used here, the number of hidden nodes significantly impacts accuracy. The best results were obtained with 8 hidden nodes (ELM (8-15-1)).
16.	Xiang et al. [49]	Applied Soft Computing (8.7)	China	SVR, E-ANN, ANN, E-SVR, E-ANN-SVR	Rainfall	Daily	1951–2015	RMSE, R, MAE	Input data was decomposed by IMFs and then ELM-SVR was applied to compare the error results.
17.	Akbari Asanjan et al. [50]	Journal of Geophysical Research: Atmospheres (4.4)	USA	RNN-PER, Persist-PER, LSTM-PER, Farne-PER, RNN-PER	Rainfall	30-min	2011–2013	RMSE, CC, POD, FAR, CSI	Different number of nodes and hidden layers were tested to optimize the hyperparameter tuning.
18.	Alotaibi et al. [51]	Water (3.4)	Saudi Arabia	ANN, ANFIS, GCM	Rainfall, T	Daily, Weekly, Monthly	1978–2015	AIC, R^2	Feed-forward networks (ANN) performed better in long-term forecasting.

(*continued on next page*)

Table 1 (*continued*).

Sl. No.	Author & year	Journal with impact factor (IM/Cite score)	Study area	Used AI models	Input	Time steps	Span of datasets	Model parameters	Research Insights
19.	Pour et al. [52]	Atmospheric Research (5.5)	Bangladesh	SVM-MOS	Rainfall, CMIP5	Monthly	1961–2005	MBE, NSE, Nash Sutcliffe	Using SVM-based MOS downscaling with GCMs offers a reliable method for predicting future rainfall changes.
20.	Zeynoddin et al. [53]	Journal of Environmental Management (8.7)	Malaysia	ARMA, ELM	Rainfall	Monthly	2000–2014	MAE, RMSE, MARE, RMSRE	Decomposing data using techniques like Empirical Mode Decomposition (EMD), Vibrational Mode Decomposition, and Wavelet Transform can improve the accuracy of forecasting models.
21.	Kumar et al. [54]	Civil Engineering Journal (4.1)	India	ENN, ANN	Rainfall, Discharge	Not Mentioned	1986–2014	R, R^2, Nash Sutcliffe, RMSE, RPE	Compared ENN and ANN for rainfall-runoff modeling in a flood-prone area, finding that ENN outperforms ANN in simulating discharge based on performance metrics.
22.	Liu et al. [55]	Applied Soft Computing Journal (8.7)	India	RBF-NN, RBF-GA, RBF-HPSOGA	Rainfall	Monthly	1949–2011	AARE, RMSE, CC	Neural Networks have the superior ability to handle the complex, non-linear nature of rainfall patterns.
23.	Poornima and Pushpalatha [56]	Atmosphere (2.9)	India	LSTM	Rainfall	Not Mentioned	1980–2014	RMSE, Loss	Intensified LSTM model for rainfall prediction showed superior performance to traditional methods like ARIMA and LSTMs.
24.	Aguasca-Colomo et al. [57]	Applied Sciences (2.7)	Canary Islands	RF, GBM, XGBoost	Rainfall	Monthly	1976–2016	Kappa-Coefficient	ML models predict effectively in similar geographical regions.
25.	Hossain et al. [58]	Meteorology and Atmospheric Physics (2.0)	Australia	ANN	ENSO, IOD	Monthly	1957–2013	RMSE, R, MAE, d	They suggested that for seasonal rainfall prediction in this region, considering non-linear relationships between past climate factors (ENSO and IOD) and rainfall patterns was more accurate.
26.	Hossain et al. [59]	International Journal of Water (0.36)	Australia	ANN	ENSO, IOD	Monthly	1957–2014	RMSE, R, MAE, d	The non-linear ANN models, trained on past climate data (ENSO and IOD), outperformed MLR models.
27.	Suparta and Samah [60]	Geodesy and Geodynamics (2.4)	Indonesia	ANFIS	Rainfall	Daily, Monthly	2009–2018	MAPE, MSE, RMSE	This suggests ANFIS could be a valuable tool for flood risk management.
28.	Ni et al. [61]	Journal of Hydrology (6.4)	China	MLP, LSTM	Rainfall	Rainfall, Streamflow	1951–2017	RMSE, NSE, MARE	WLSTM and CLSTM outperformed traditional methods (MLP and LSTM) for predicting monthly streamflow and rainfall.
29.	Mohammed et al. [62]	International Journal of Scientific and Technology Research	India	MLR, SVR, Lasso Regression	Rainfall	Monthly	1901–2015	MAE, R	Provided non-experts with an accessible way to understand and compare these techniques for improved prediction accuracy.
30.	Velasco et al. [63]	ProcediaComputerScience (4.0)	Philippines	MLPNN, ANN	T, WS, H, Rainfall, V, Days, Months, Years	Daily	2006–2018	RMSE, MAE, MAPE	MLPNN model was successful with minimal errors, suggesting it could be useful for planning purposes.
31.	Ridwan et al. [64]	Ain Shams Engineering Journal (6.0)	Malaysia	BLR, BDTR, DFR, NNR	Rainfall	Daily, Weekly, 10-day, monthly	2010–2019	MAE, RMSE, RAE, R	Boosted Decision Tree Regression (BDTR) achieved the highest accuracy, especially for daily and monthly forecasts using an Autocorrelation Function (ACF) approach.
32.	Barrera-Animas et al. [65]	Machine Learning with Applications	UK	S-LSTM, B-LSTM, XGBoost, GBR, LSVR	Rainfall	Hourly	2000–2020	Loss, RMSE, MAE	Stacked-LSTM and Bidirectional-LSTM networks performed best, suggesting these approaches can be effective for budget-conscious rainfall forecasting applications.
33.	Elbeltagi et al. [66]	Arabian Journal of Geosciences (2.3)	Egypt	DL-NN	Rainfall	Not Mentioned	Not Mentioned		They proposed a hybrid deep learning approach with an adaptive hyperparameter optimization algorithm.
34.	Di Nunno et al. [67]	Sustainability (3.9)	Bangladesh	M5P, SVR, M5P-SVR	T, H, CC, SSH, P	Monthly	1956–2013	MAE, RAE, RMSE, R^2	The hybrid model, M5P-SVR, achieved the best results with accuracy (R^2) exceeding 0.87.
35.	Billah et al. [68]	Open Computer Science (1.5)	Bangladesh	KNN, Logistic Regression, SVM, RF, NB, NN, LSTM	Rainfall	Daily	2012–2018	RMSE, SD, CC, PP	LSTM outperformed traditional techniques by achieving 97.14% accuracy in predicting rainfall.
36.	Ojo and Ogunjo [69]	Scientific African (2.9)	Nigeria	ANN, ANFIS, SVM, MLR	Rainfall	Daily	1983–2013	RMSE, AIC, R^2	Different ANFIS algorithms performed best for different months, suggesting a need to tailor the model based on the target prediction period.

(*continued on next page*)

Table 1 (*continued*).

Sl. No.	Author & year	Journal with impact factor (IM/Cite score)	Study area	Used AI models	Input	Time steps	Span of datasets	Model parameters	Research Insights
37.	Hunasigi et al. [70]	Acta Ecologica Sinica (4.1 Cite Score)	India	MLP-ANN	Rainfall	Daily	1983–2016	CC, RMSE	AI approach is a valuable tool for future weather parameter prediction in India.
38.	Abebe and Endalie [71]	Journal of Big Data (8.1)	Ethiopia	ANN, ANFIS	Rainfall	Not Mentioned	2011–2021	RMSE, E, MAE, MAPE, R^2	ANFIS was found more accurate than ANN in predicting rainfall across various locations in Ethiopia.
39.	Nithyashri et al. [72]	Measurement: Sensors (0.9 Cite Score)	India	LSTM	Rainfall	Monthly	Not Mentioned	MAE, RMSE, Loss	LSTM predicted rainfall with an average accuracy of 89%, and required less data for training, and works efficiently.

The benefits extend beyond early warnings. AI-powered forecasts can significantly improve preparedness and response efforts. Earlier and more accurate flood warnings allow authorities to plan and execute evacuations efficiently, minimizing casualties. Similarly, drought forecasts can trigger proactive measures like water conservation or preemptive water distribution to drought-stricken areas, mitigating the impact on agriculture and domestic water supplies. The ultimate goal is to build community resilience. AI-powered forecasting systems can be integrated with public awareness campaigns, educating communities about potential risks and appropriate preparedness actions. Farmers can utilize forecasts to optimize agricultural practices, leading to increased productivity and reduced vulnerability to droughts. More accurate forecasts can also facilitate the development of better risk assessment tools for insurance companies, potentially leading to more affordable flood and drought insurance options for communities. This financial safety net can play a crucial role in helping communities recover from disasters. However, it is important to acknowledge the challenges. The effectiveness of AI models hinges on the quality and quantity of data used for training. Gaps in historical data or unreliable real-time measurements can limit forecast accuracy. Additionally, ensuring effective communication and accessibility of forecasts to all community members, especially those with limited access to technology, is crucial for maximizing their impact. Finally, communities should not become overly reliant on AI forecasts. Traditional knowledge and local weather observations should still be considered for a comprehensive understanding of weather patterns.

In summary, AI-powered rainfall forecasting offers a game-changing solution for mitigating the risks associated with extreme weather events. By enabling more accurate, localized, and probabilistic forecasts, AI empowers communities to prepare more effectively and build resilience in the face of potential disasters. As we move forward, addressing data quality challenges and ensuring equitable access to forecasts will be essential to fully unlock the transformative power of this technology.

3. Article selection

We conducted a systematic review to identify relevant studies on rainfall forecasting using Artificial Intelligence (AI). Our goal was to determine the sample size of the literature, profile its characteristics, identify tools used for synthesis, and uncover existing research gaps. We meticulously followed the 2020 Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines, but with a specialized structure tailored to analyze AI adoption and research gaps within the rainfall forecasting domain. A PRISMA flow diagram (Fig. 1) showcases the systematic review process. It visually depicts the flow of studies at each stage, including the initial number identified, the number included after applying selection criteria, and the number excluded with justifications provided for each exclusion.

For article acquisition, we cast a wide net, encompassing renowned publishing authorities like Elsevier, Springer, Taylor & Francis, and more. We crafted a targeted keyword list to find out the most relevant works. This list included terms like "Rainfall", "Forecasting", "Artificial Intelligence", "Neural Network", "Deep Learning", and "Hybrid AI model", optimized for each database's search engine. At least one of these keywords had to be present in the target articles. Furthermore, the publication timeframe was restricted to articles published between 2000 and 2023.

Prior to diving into the selection process, we established clear inclusion and exclusion criteria. Articles had to be published in English, feature a readily available full-text for review, and emphasize the utilization of AI techniques. Additionally, the research could originate from any developing country, alongside developed countries categorized by the United Nations. This study emphasized to collect studies representing all the continents excluding the Antarctica. Finally, review articles and case studies were not considered. Through this rigorous selection process, we aimed to construct a comprehensive and focused sample that illuminates the development landscape of AI-powered rainfall forecasting research from 2000 to 2023. Analyzing this sample do not only reveal the volume of existing research but also exposes areas where further investigation is crucial for advancing the field of rainfall forecasting.

The initial search identified 231 studies potentially relevant to rainfall forecasting using AI techniques. These studies underwent a multi-stage screening process (Fig. 1). The first stage involved examining titles, abstracts, keywords, results, and discussion sections. Studies lacking a clear focus on AI-based rainfall forecasting were excluded at this stage (n=57). Additional exclusion criteria were applied: publication date outside the defined timeframe (2000–2023) (n=13), non-utilization of AI techniques (n=17), and research not conducted in developing countries [75] (n=67) – the latter aimed to achieve a geographically balanced sample.

Following these exclusions, reports for 77 potentially relevant studies were retrieved. However, 29 reports faced further exclusion due to unavailability in chosen databases (e.g., SCOPUS, Web of Science) and some papers were excluded to avoid duplication (n=9). The final selection process yielded a sample of 39 studies deemed highly relevant and suitable for inclusion in the review paper.

To gain deeper insights from this sample, the authors undertook a qualitative coding and analysis process. This involved a meticulous breakdown of the data, categorizing it into key themes and patterns. Each study was systematically reviewed, with a specific reading sequence: abstract, conclusion, results and discussions, methodology, and finally, the introduction. This method facilitated the identification of potential themes and categories based on the information presented in each paper. The identified themes and categories were then extracted and tabulated for further analysis. A comprehensive table was created in Excel to capture key variables, including author names, publication year, journal with impact factor (or CiteScore), study area, utilized AI models, input parameters, time steps, dataset span, and performance evaluation parameters. This multi-stage screening process and subsequent in-depth analysis ensured a focused and informative sample size for the review paper. The extracted and tabulated data provided a solid foundation for further investigation of the research landscape surrounding AI-powered rainfall forecasting.

4. Non-comparative studies

4.1. Bibliographic review

Luk et al. [34] paved the way for advancements in short-term rainfall forecasting within urban environments by proposing the use of ANNs. Their work was further complemented by Luk et al. [35] who explored the potential of ANFIS method. Valverde Ramírez et al. [36] proposed an ANN-based method for Sao Paulo, Brazil. Their approach involved establishing a non-linear relationship between regional ETA model output data and surface rainfall data. Moreover, Bilgili and Sahin [39] successfully employed ANN architecture to model monthly rainfall and temperature in the context of Turkey. Another study conducted by Abbot and Marohasy [41] further contributed to this field by exploring the feasibility of using climatic indices for rainfall prediction. Nastos et al. [42] investigated the potential of ANNs in estimating the likelihood of predicting maximum precipitation on a daily basis for the following year. In contrast, Abbot and Marohasy [44] utilized an ANN model combined with genetic optimization to identify relationships between long-term data series lag values and rainfall at geographically dispersed sites in Queensland. Kumar et al. [54] addressed ENN and ANN techniques together for rainfall-runoff modeling in Sone Command, Bihar, aiming to address the challenges posed by frequent flooding in this region. Suparta and Samah [60] built three different ANN models for rainfall forecasting, evaluating their respective performance and potential for further development. Table 1 illustrates the key findings from the reviewed papers.

4.2. Results

This study analyzed 9 papers (24%) that explored various improvements to the ANN method for rainfall forecasting, with a strong emphasis on minimizing error terms. The findings of these studies are presented below:

1. ENN, an enhanced ANN model, outperformed than the conventional ANN models in runoff-rainfall modeling concluded by the accuracy parameters [54].
2. Daily precipitation time series enhanced ANN's ability to forecast such extreme events, reducing the negative effects on urbanized areas' infrastructure and even preventing fatalities.
3. Extensive optimization of both parameters and hyperparameters was conducted to significantly reduce the residual error.
4. A comprehensive investigation into the consequences of climate change was conducted, encompassing various environmental parameters.

5. Comparative studies

5.1. Bibliographic review

Seasonal rainfall prediction in Kerala, India, was investigated using three machine learning algorithms: ANN, and extreme learning machine (ELM), and KNN [48]. Partal and Kişi [37] decomposed daily precipitations into subseries using discrete wavelet transforms, which are fed to neuro-fuzzy models to anticipate daily precipitation. Typhoon rainfall prediction was investigated using a novel method combining RNN and support vector regression, where RSVRCPSO served as the driving optimization algorithm [38]. Mandal and Jothiprakash [40] used ANNs and model trees (MT) to estimate future rainfall using observed data. The effectiveness of a neural network and an ARIMA technique was assessed while predicting monthly rainfall within the Urmia Lake basin [43].

Lin and Jhong [45] focused on the well-known MLP and ANN models, while [47] explored a broader range of algorithms, including KNN, GP, SVR, RBF, M5 Rules, and M5 Model Trees. Addressing the challenges of monthly streamflow and rainfall forecasting, [76] proposed two hybrid models, WLSTM and CLSTM, based on LSTM networks. Alotaibi et al. [51] employed a combination of diverse techniques, including ANNs, ANFISs, and GCMs, to project rainfall and temperature patterns in the Qassim region. Fan et al. [77] proposed a hybrid method combining SVR and ANNs for rainfall prediction.

Both [52,53] focused on advancing rainfall prediction methods. Pour et al. [52] applied a downscaling method for Bangladesh, while [53] studied the efficiency of hybrid models, specifically a fusion of a non-linear ELM and a linear stochastic model, for predicting rainfall in tropical regions. Short-term Quantitative Precipitation Forecasting was conducted by Deep neural networks by improving Cloud-Top Brightness Temperature (CTBT) [50]. Random Forest and Extreme Gradient Boosting was used by Aguasca-Colomo et al. [57]. Long-term forecasting was explored by Poornima and Pushpalatha [56] using both RNN and LSTM, while ANNs were implemented with diverse algorithms [55].

MLP and ANN were applied in the Western Australia [58]. In that study, different LSTM techniques and an ensemble model combining gradient boosting regression, simple SVR, and extra-tree regression were utilized. Lasso Regression and SVM proved an effective forecasting model using regression [62]. Ni et al. [61] conducted a comparative study on the performance of SVM and random forests for real-time rainfall forecasting. Meanwhile, Ridwan et al. [64] explored regression technique using decision forest, Bayesian linear, neural network, and boosted decision tree, for their forecasting model. In another study by Ojo and Ogunjo [69], fourteen machine learning algorithms were evaluated Nigeria's monthly and annual rainfall.

The available models encompass two polynomial regression methods, three ANNs with varied configurations, four accommodative neuro-fuzzy inference with different fuzzy memberships (Fuzzy C-Means, Gaussian, Generalized Bell, Subtractive Clustering), and five support vector machine kernels. For rainfall prediction, LSTM [68], M5P-SVR [67], optimization of hyperparameters [66]. While [65] focused on comparing simplified models built with traditional and deep learning approaches, considering their downstream application suitability, [71] investigated on ANFIS and ANN using topographical and cyclic component data collected between 2011 and 2021. The temporal relationships between coastal rainfall data were captured using LSTM networks [72].

5.2. Results

72% of the reviewed papers (28 papers) concentrated on comparing AI models for a better understanding of rainfall patterns and training the models for better accuracy. Luk et al. [35] highlighted the role of short-term memory in rainfall data, suggesting that it may explain the similar performance achieved by various neural network models, including TDNN, MLFN, and Elman networks, in rainfall prediction tasks. ANN and Model Trees (MT) models perform equally well, suggesting they are promising short-term rainfall forecast methods [40]. The suggested RSVRCPSO model has a satisfactory NMSE accuracy index. The RSVRCPSO model exhibits superior performance in capturing the average rainfall anomalies. This indicates its ability to accurately predict rainfall depth values and their correlation with actual values [42]. MGSVM outperforms SVM for most events, notably long-term forecasting [45].

Yu et al. [76] concluded that the RF-SMFM tends to underestimate rainfall compared to the SVM-SMFM, which delivers superior performance in most cases. Additionally, their analysis revealed that SVR, GP, and Radial Basis Functions were the top performing algorithms, with Radial Basis Functions statistically outperforming the conventional Markov chain extension [47]. Additionally, the LSTM/LSTM-PER demonstrated a key advantage in its sustained performance, exhibiting a slower decline in accuracy compared to other methods examined [49].

Adjusting the invisible layers and their underlying nodes directly impacts the accuracy of predictions, and the ELM design surpasses other approaches in terms of accuracy [48]. Liu et al. [55] found that the hybrid particle swarm approach holds significant promise in generating highly efficient architectures for radial basis neural networks. Combining wavelet-transform and convolutional layers with LSTM models for rainfall forecasting can lead to significant accuracy improvements, especially for longer prediction horizons. Furthermore, CLSTM and WLSTM models demonstrate superior performance over standard LSTM models when forecasting rainfall for longer time steps ahead, as informed by Ni et al. [61]. The tweaked SVR simulation predicted best, showed by Mohammed et al. [62]. According to Barrera-Animas et al. [65], LSTM neural networks that possess a substantial quantity of hidden layers may not be as effective in acquiring the intricacies of meteorological time-series data for the purpose of predicting hourly rainfall volume values. Based on the accuracy parameters, neural networks (76.9%), KNN (76.8%), and LSTM (97.14%) were the best [68]. The hybrid deep learning model achieved an impressive 62% improvement in RMSE compared to previous methods [66]. In tropical monsoon-climate zones, ML algorithms, particularly the hybrid model M5P-SVR, should predict precipitation reliably [67]. ANFIS outperformed ANN across all evaluation metrics and testing stations. This is highlighted by the higher Nash–Sutcliffe efficiency scores, with ANFIS reaching 0.995 compared to 0.935 for ANN [71].

After careful data preparation, model training, and performance evaluation, two MLPNN models achieved promising results [78]. These models, SCG-Tangent and SCG-Sigmoid, had Mean Absolute Error (MAE) values of 0.01297 and 0.1388, respectively, and Root Mean Squared Error (RMSE) values of 0.01512 and 0.01557, respectively. These low error values suggest that the MLPNN approach has the potential to provide valuable lead-time for organizations and individuals to

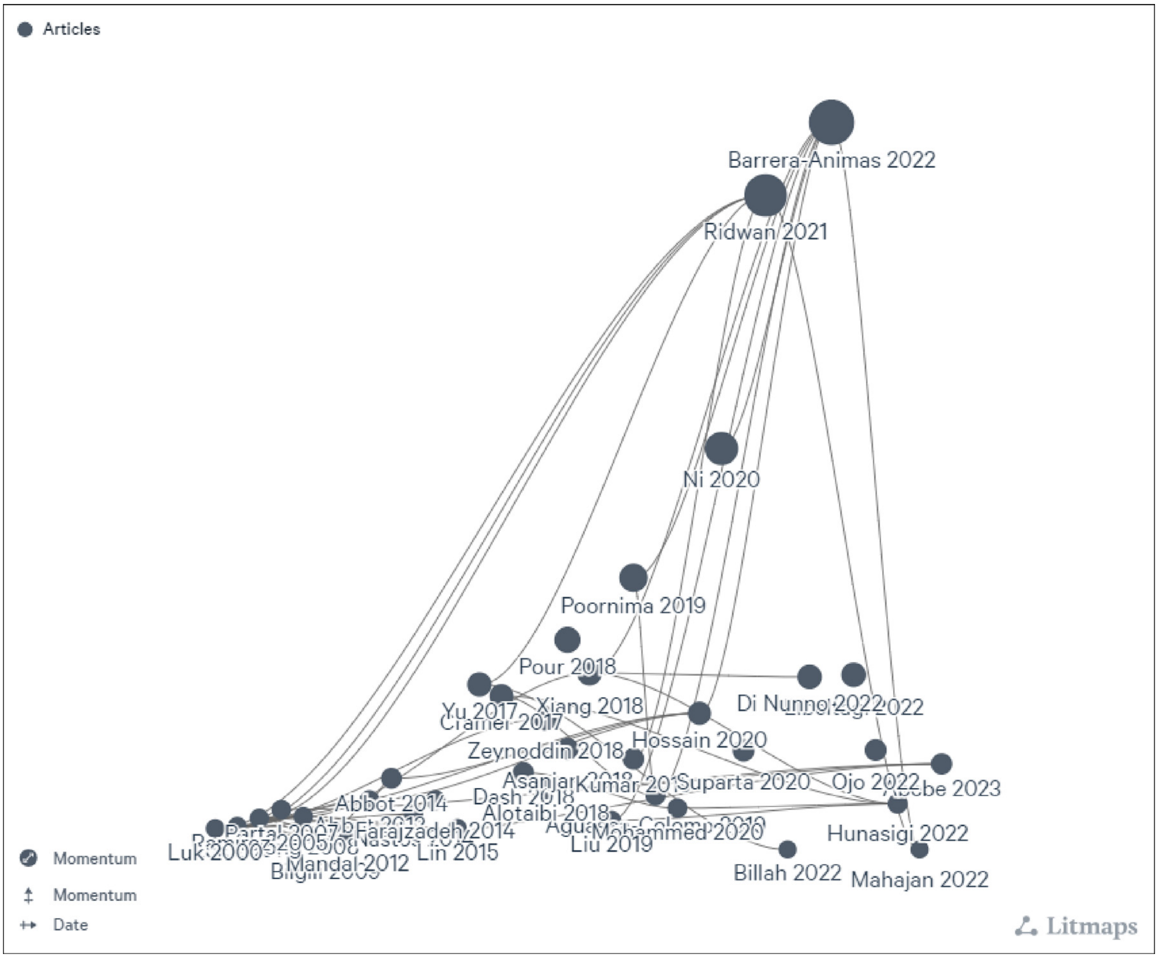

Fig. 2. Momentum of the reviewed papers (based on citations and recency).

plan and take action in response to upcoming rainfall events. Moreover, Another MLPNN with the same number of hidden neurons and Sigmoid activation function, but using the Resilient Propagation training algorithm, had the lowest Mean Absolute Error (MAE) of 0.0209 [63]. These results emphasize the importance of performance analysis when using MLPNNs for rainfall forecasting, as different configurations can lead to variations in accuracy metrics. On the other study, the optimized SVRM model achieved a Mean Square Error (MSE) of 3.461315, indicating a close match between predicted and actual rainfall values [79]. These results suggest that SVRM has the potential to be a valuable tool for rainfall forecasting, provided careful data preparation, appropriate kernel function selection, optimal parameter tuning, and consideration of lag variables are implemented.

6. Generalized results and discussions

6.1. Interrelationships among the reviewed papers

The analysis of the selected 39 papers from various reputable journals involved examining the study profile, citations, impact factors, or Cite Score, of the journals in which they were published, the year of publication, and the significant research outcomes. Several characteristics were taken into consideration in order to determine the most suitable research methodologies utilizing artificial intelligence (AI). Fig. 2 illustrates the title similarity among the reviewed papers, and the following illustration, Fig. 3, mentions the relationship among the papers based on the momentum of the individual papers.

The calculation of momentum is derived from the number of citations a publication receives, which is then adjusted to account for the passage of time. The significance of time adjustment lies in its ability to account for the inherent tendency of older publications to accumulate a greater number of citations. Hence, through the consideration of publication recency, it becomes evident that current articles possess greater potential compared to previously published studies. The process of AI forecasting encompasses a series of essential stages aimed at generating precise predictions by leveraging past data [80]. After gathering the necessary data, it goes through a preprocessing stage to ensure its accuracy and suitability for further analysis. Feature selection and engineering techniques aid in the extraction of variables that possess the highest degree of information.

The process of model selection and training entails the careful selection of a suitable algorithm and subsequently training it using provided training data. The process of fine-tuning hyperparameters is fundamental to maximizing a model's performance. This involves carefully adjusting various parameters within the model to optimize its accuracy. Evaluating the model's performance against the independent test set provides a crucial measurement of its effectiveness in real-world scenarios [81]. The act of deploying a model in a practical setting enables the generation of predictions for previously unseen data. The model's continual relevance and effectiveness in producing predictions are ensured through continuous monitoring, maintenance, documentation, and feedback loops. A generalized concept on building AI models for forecasting is illustrated in the Fig. 3 [82].

Momentum, often measured by citation count, offers an indirect indicator of a paper's influence within the research community (Fig. 1). Studies with high momentum suggest they have resonated with other researchers, potentially shaping the direction of future investigations. By analyzing the momentum of research papers in AI-based rainfall

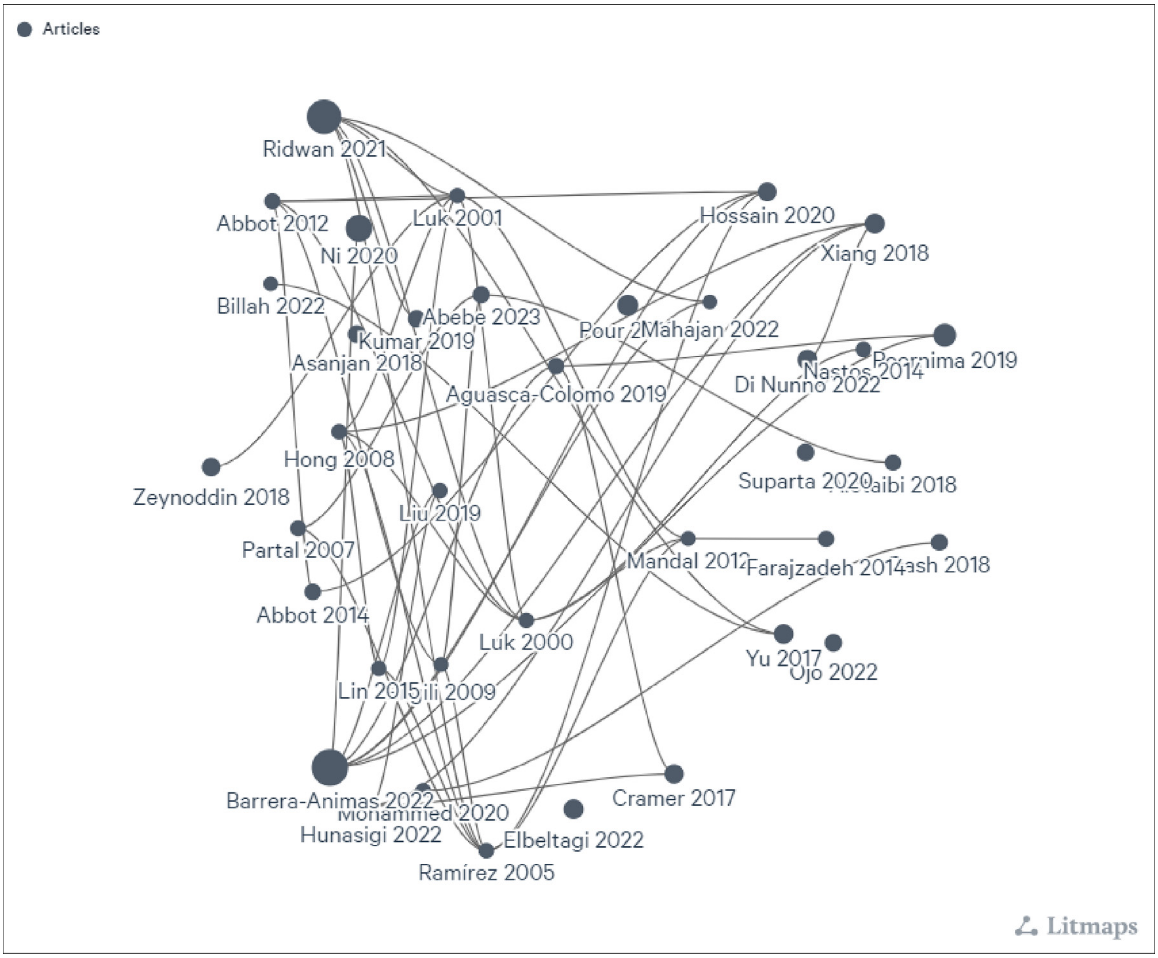

Fig. 3. The similarity in the methodology among the papers.

forecasting, we can identify emerging trends and areas of focus within the field. These highly cited papers often provide foundational knowledge, propose groundbreaking methodologies, or present significant advancements in rainfall forecasting using AI techniques. They serve as valuable starting points for further research and development. The very process of analyzing research momentum can also highlight gaps in existing research. Areas with relatively low momentum might indicate a lack of exploration or suggest opportunities to revisit older techniques with a fresh perspective using more advanced AI models or data sources.

The reviewed literature exhibited a consistent methodology for AI-powered forecasting techniques (Fig. 4). The majority of studies (over 93%) opted for a manual partition of the data for training and testing purposes. Cross-validation, a statistically robust technique for mitigating overfitting, was employed by a minority of studies (less than 7%). This suggests a potential bias towards a specific data division strategy, which could limit the generalizability of the obtained forecasting models. Furthermore, the complexity of the employed AI models varied considerably across the reviewed studies. This heterogeneity can be attributed to several factors. Firstly, Different AI models possess distinct theoretical foundations, influencing their suitability for capturing specific data patterns. For instance, recurrent neural networks excel at modeling sequential data, while support vector machines might be preferred for tasks involving high-dimensional feature spaces. Secondly, the characteristics of the input data significantly impact model selection.

The number of features, their data types (categorical, numerical), and the presence of non-linearities all influence the choice of an appropriate AI model architecture. Lastly, the performance of an AI model is highly sensitive to its hyperparameters, which are settings that control the learning process. The complexity of the hyperparameter tuning procedure can vary depending on the chosen model architecture.

Future research could benefit from exploring the development of more intricate hybrid models. These models could combine the strengths of different AI paradigms to achieve superior accuracy with minimal errors. However, constructing such hybrid models presents a significant challenge. For example, increasing model complexity can enhance the model's capacity to learn intricate data patterns. However, excessively complex models are prone to overfitting, leading to poor performance on unseen data. Striking a balance between model complexity and generalizability is crucial for achieving robust forecasting models. Moreover, certain AI models, particularly deep neural networks, can be challenging to interpret. While they might deliver exceptional accuracy, understanding their inner workings can be difficult. Future research should strive for the development of interpretable hybrid models that provide insights alongside accurate forecasts. By addressing these challenges, future research can pave the way for the creation of novel and powerful hybrid models that deliver superior forecasting performance with minimal errors.

6.2. Study area

This review paper presented papers from 06 continents except Antarctica. Papers were collected from India (8 papers), Australia (7 papers), Bangladesh (3 papers), Taiwan (3 papers), China (2 papers), Malaysia (2 papers). One paper from Brazil, Canary Islands, Indonesia, Iran, UK, Turkey, Saudi Arabia, Ethiopia, Nigeria, and Greece were studied for different AI models. A geographically-based map in Fig. 5

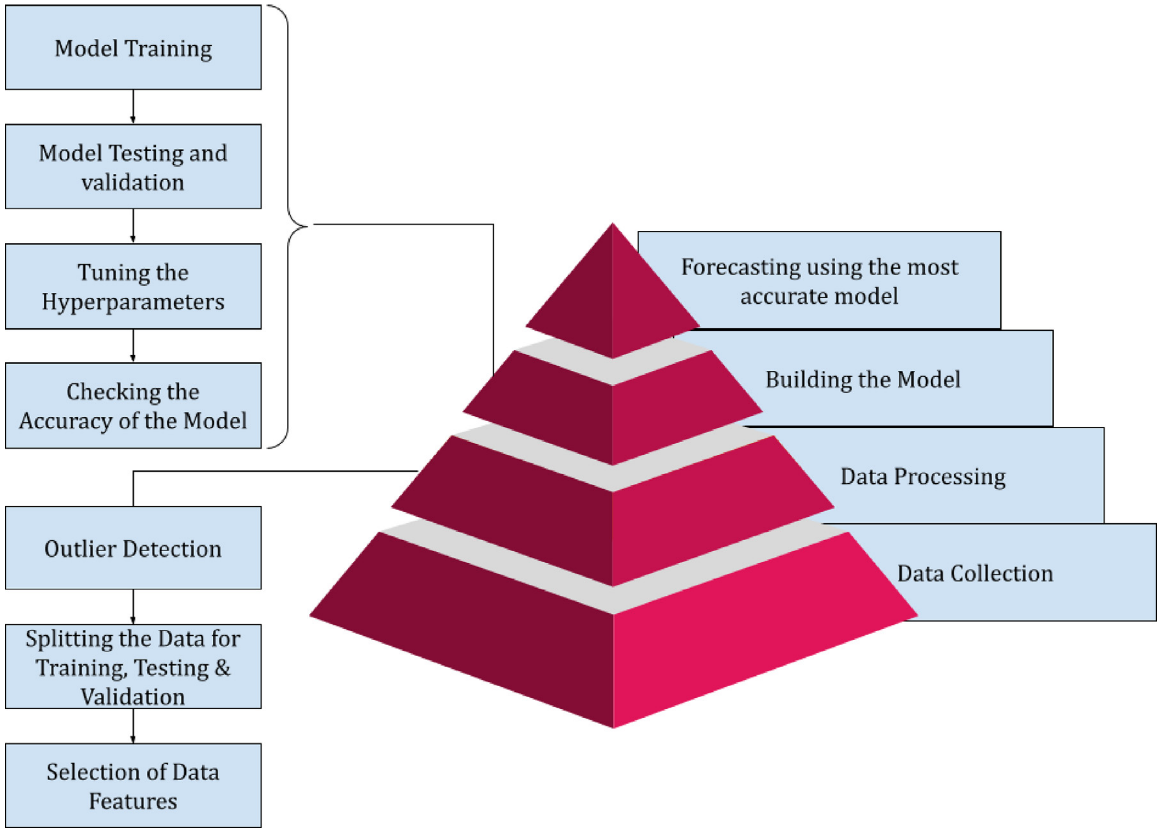

Fig. 4. Generalized steps in AI model-based forecasting.

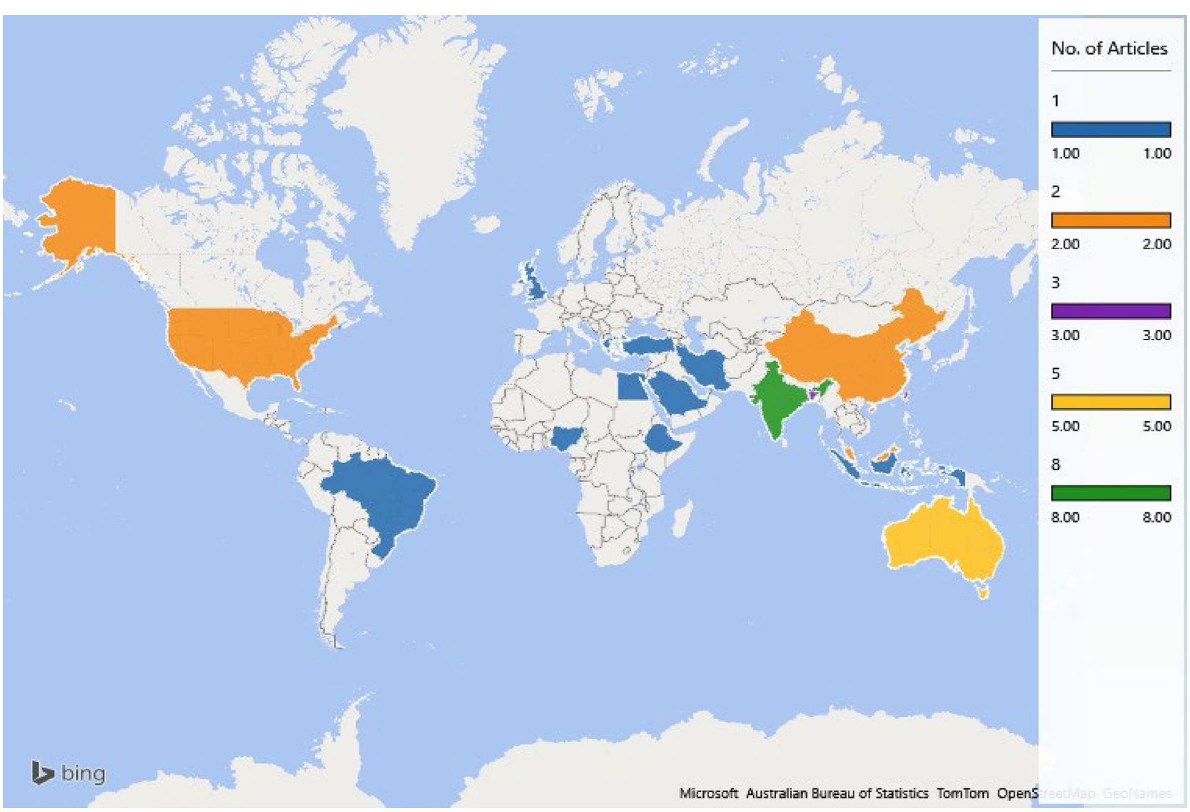

Fig. 5. Study regions of the reviewed papers.

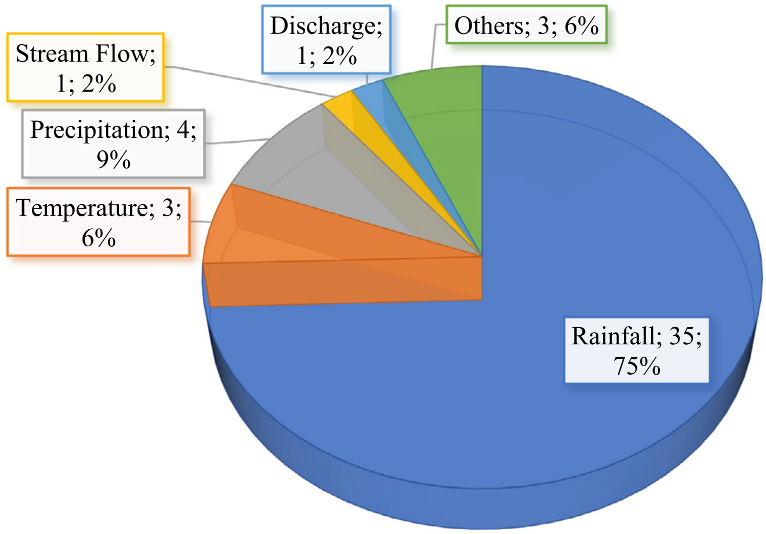

Fig. 6. Input parameter statistics for rainfall forecasting (no. of reviewed papers, proportion).

illustrates the distribution of countries based on conducted research papers.

Environmental parameters, including rainfall, exhibit significant spatial variability across geographical regions. Consequently, the techniques employed for rainfall forecasting must adapt to these regional variations. This review study serves a twofold purpose. Firstly, by mentioning a selection of countries (Fig. 4), it provides a concise overview of forecasting techniques utilized in diverse locations, which can be a valuable resource for future researchers. This knowledge can guide them in selecting appropriate techniques for their specific research objectives. Secondly, for future studies conducted in geographically similar regions with comparable seasonal patterns to the countries mentioned (Fig. 12), this review offers a starting point for exploring relevant forecasting techniques (as potential first attempts). Based on the reviewed literature on rainfall characteristics within these countries, the following overview is presented.

Monsoon Dominance in South Asia

India and Bangladesh experience a dominant wet season driven by the monsoon. From June to September, these countries receive the majority of their annual rainfall. Conversely, winter months bring drier conditions. Similarly, Taiwan experiences a rainy season influenced by typhoons, typically lasting from May to August. This contrasts with drier winters in the region.

Regional Variations Across Continents:

Australia showcases a continent-wide variation in rainfall patterns. The northern regions experience a wet season during their summer months (December–February), while the south receives winter rain (June–August). China exhibits a similar disparity. Eastern China experiences significant summer rainfall (June–August) due to the East Asian monsoon, while arid conditions dominate the northwestern regions.

Tropical Rainforests and Arid Climates:

Malaysia boasts a tropical rainforest climate with abundant rainfall year-round. However, slightly drier periods may occur between May and September. Indonesia, another Southeast Asian nation, also enjoys consistent rainfall throughout the year due to its tropical climate. Eastern regions tend to receive more precipitation due to prevailing winds. In contrast, the Canary Islands and Saudi Arabia represent arid climates with very little rainfall throughout the year. Iran falls largely within the arid or semi-arid category, experiencing limited rainfall with occasional spring showers in specific regions.

Temperate and Mediterranean Climates:

The United Kingdom exemplifies a temperate climate with year-round rainfall, with the heaviest occurring during winter (December-February). Turkey and Greece, on the other hand, demonstrate a Mediterranean climate with wet winters and dry summers. However, southeastern regions in Turkey have a more arid climate, while island regions in Greece tend to be drier.

Sub-Saharan Africa:

Ethiopia showcases regional variations in rainfall. The highlands experience a wet season from June to September, while the lowlands are drier. Nigeria experiences a tropical climate with a rainy season from April to October. The north of the country is drier with a short rainy season in summer.

This review analyzed 39 research articles, focusing on the input parameters (Fig. 5), data frequency (Fig. 6), and data range (Fig. 7) employed in each study. These considerations were evaluated based on the specific requirements of the chosen models and the data availability reported by the authors. The reviewed studies employing machine learning for forecasting did not explicitly outline a standardized protocol for data selection.

6.3. Input parameters

Fig. 6 shows the input parameters while forecasting the rainfall according to the studied papers. Rainfall data was mentioned in 35 papers. Precipitation and Temperature were used in 4 and 3 papers respectively. Moreover, stream flow and discharge data were fed in one paper each. Other regional variables were mentioned in 3 papers.

6.4. Time span of data

Based on the reviewed papers, the mostly used time span was monthly for rainfall forecasting. In 18 papers, monthly data was used while 12 papers came with daily data. Among the others, 15-min, hourly, and weekly data were used in 2 papers each. 30-min and 10-day data were processed in 1 paper each. In 6 papers, the time span was not mentioned. Fig. 7 illustrates the time span against the no. of publications.

6.5. Data size

In the reviewed 39 papers, the data size was chosen based on the availability and there was no mention of adapting particular data sets. The span of years started from 3 to 146 among the papers. The 20

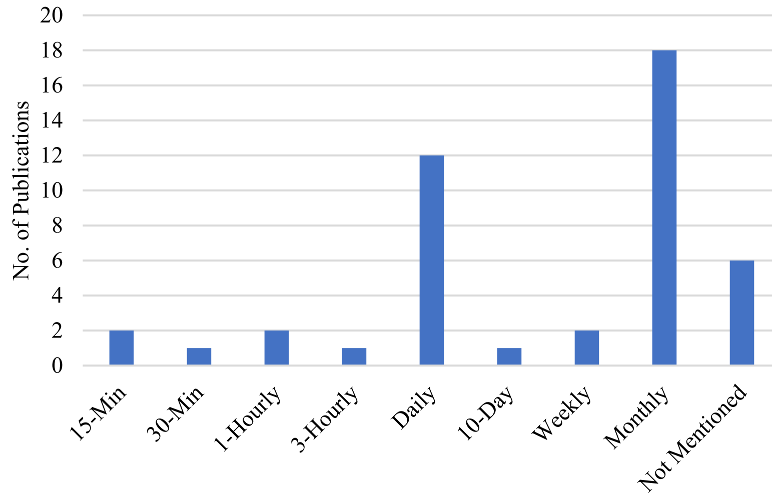

Fig. 7. Data characteristics in the reviewed papers.

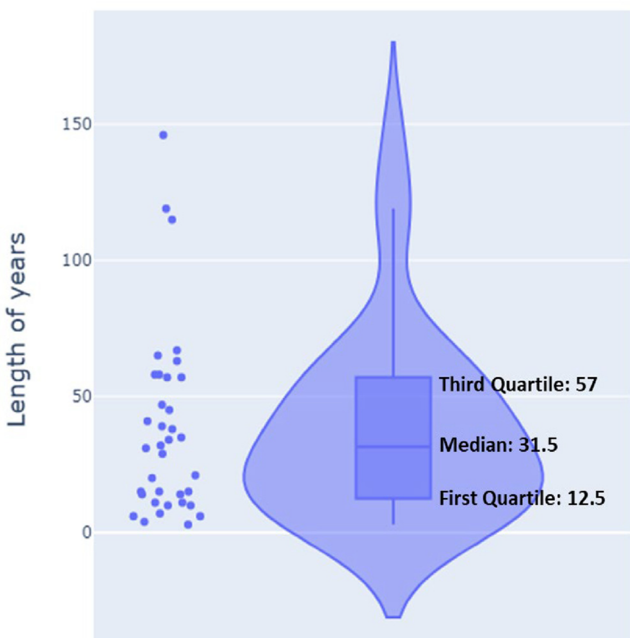

Fig. 8. The violin plot for the quartiles of the data range (in years).

papers span from 11 to 49.5 years (Second and third quartile). The 8 papers (First quartile) are between 3 to 11 years and the other 5 papers (Fourth quartile) remain from 49.5 to 67 years. The 3 papers containing the data range of 146, 119, and 115 years, were detected as outliers. While calculating all the quartiles in the violin plot, illustrated in Fig. 8, these three papers were excluded. AI models could handle a different span of data, but, no specific recommendations or remarks were mentioned in the reviewed papers regarding the span of the datasets. 3 papers did not mention the range of data.

6.6. Data splitting

Rainfall forecasting using Artificial Intelligence (AI) techniques relies heavily on the proper utilization of data for model development and evaluation. In this context, the practice of dividing the available data into training, testing, and validation sets becomes an absolute necessity. Firstly, the training dataset forms the backbone of the AI model. It consists of historical rainfall data paired with relevant environmental factors (temperature, humidity, wind speed, etc.). The AI model is trained on this data, essentially learning the relationships between the input features and the corresponding rainfall patterns. Secondly, testing dataset serves as an independent benchmark for assessing the model's generalizability. Once trained on the training set, the model is evaluated on the testing set. If the model performs well on unseen data, it indicates its ability to learn from the training data and apply that knowledge to predict future rainfall events. Finally, the validation Set (Optional but Highly Recommended) plays a crucial role in optimizing the model's hyperparameters.

These are settings within the AI model that influence the learning process (e.g., number of hidden layers in a neural network). The validation set allows us to fine-tune the hyperparameters to achieve the best possible performance on the training data without overfitting. Overfitting occurs when the model becomes too specific to the training data and performs poorly on unseen data. Without this data partitioning approach, AI models for rainfall forecasting would suffer from several shortcomings like overfitting, biased estimates, difficulty in Hyperparameter Tuning and so forth.

If the model is trained and evaluated on the same data, it might simply memorize patterns in that specific data. This leads to poor performance when presented with new, unseen data, rendering the forecasts unreliable. Without a dedicated testing set, it is impossible to objectively assess the model's accuracy. The model's performance on the training data might be misleadingly high if the training data does not represent the full range of potential rainfall scenarios. Without a validation set, it becomes challenging to optimize the model's hyperparameters. Picking hyperparameters blindly could lead to suboptimal performance, limiting the model's forecasting potential. The practice of using training, testing, and validation sets is not just a good practice, but an absolute necessity for developing reliable and accurate AI-powered rainfall forecasting models. By partitioning the data effectively, researchers can build models that learn from the past, generalize to the future, and ultimately provide valuable insights for water resource management, flood prediction, and agricultural planning.

In the reviewed papers, any data splitting technique was not observed in most cases. Several papers employed cross-validation techniques as a means of improving model performance. Dataset was split into training, testing and validation in 9 papers. 3 papers used cross validation process and the ratio of training, testing and validation was not mentioned. No specific information about data division was found in 5 papers. Data was only divided into training and testing in 18 papers. One paper used training and validation, while another paper used testing and validation only. The ratio of data division is

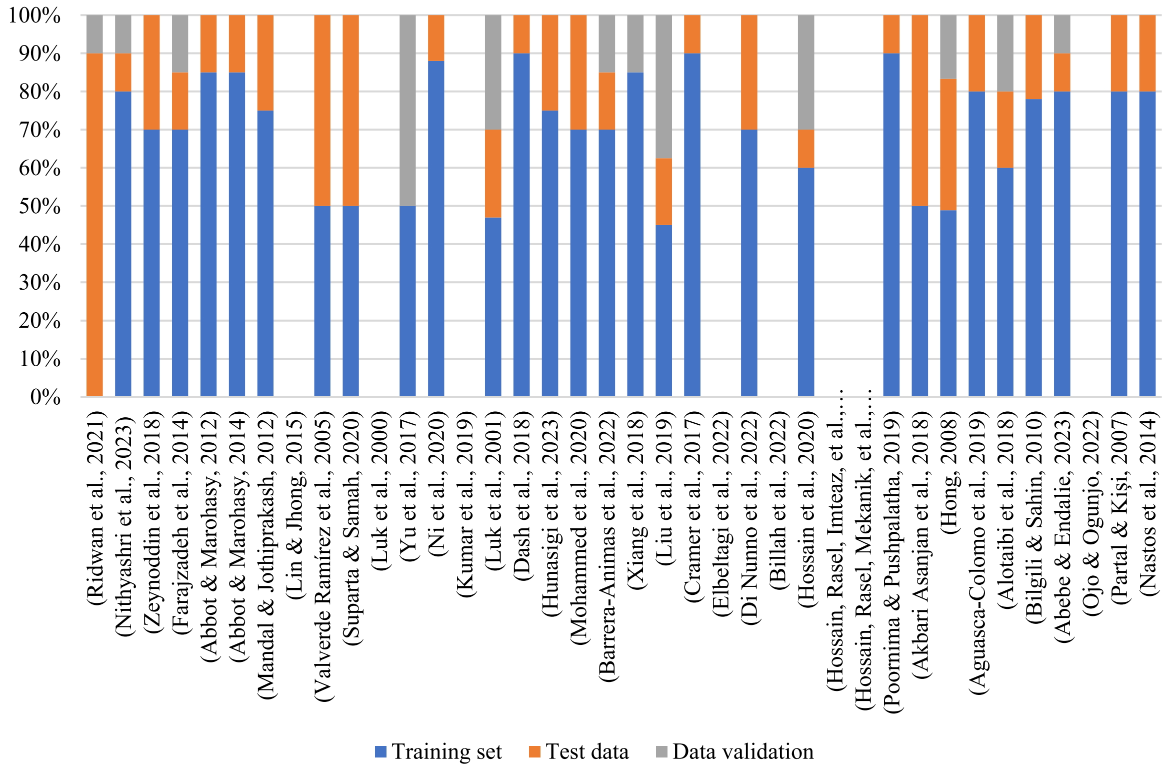

Fig. 9. Data division in the reviewed papers.

illustrated in the Fig. 9. To determine the optimal size of training and test sets, two studies implemented a trial-and-error methodology, iteratively adjusting the split until achieving the lowest error rate.

6.7. Software used

83% of studies (31 papers) were conducted without mentioning the software programs. The software selection for AI modeling varies across these studies, with MATLAB being the most popular choice with 3 studies, followed by Python with 2. This trend reflects the current popularity of MATLAB in engineering and scientific research, while Python's versatility and growing AI libraries are also attracting researchers. Notably, the use of R in one study highlights its potential for specific research domains, particularly those involving data analysis and statistical modeling. The architecture and steps in the model building, data division, accuracy parameters also diverse for the different programming environment. Fig. 10 illustrates the methods used in the reviewed papers and the corresponding software programs used. AI techniques are well known as a black-box that barely involved in the physical process, rather, accuracy in forecasting is the main goal of these models where the background mechanism does not concern much.

6.8. Model accuracy parameters

Different accuracy parameters were analyzed to compare the model performance. RMSE was used in 65% of the total 39 papers whereas MAE was the second highest parameter used in the 13 papers which is 41% of the total reviewed papers. R-squared (19%), R (22%), CC (11%), CE (8%), NMSE (8%), d (8%) was found comparatively in higher quantity. Loss, RAE, MARE, PP, AIC, Nash Sutcliffe, MBE were observed in 5% papers each and 3% papers contained NSE, RPE, AARE, AD, Kappa-Coefficient, MAPE, IA, RMSE. These parameters were used combinedly or separately in the paper. Fig. 11 describes the parameters and the corresponding number of publications. Moreover, the best models were mentioned in Fig. 12 describing the model names against the frequency of the models in publications.

6.9. Best models for specific regions

AI models produce output from the trained data. Similar trends can be seen within particular locations according to the climate conditions. This suggests that the models may be able to identify the region based solely on the data, potentially leading to better results without requiring comparisons with other models. This served as the impetus for conducting an initial analysis of 39 papers with a focus on the study area (the country) and the most promising models identified within each study. ANN showed better accuracy in Australia (in 5 papers), Brazil, Saudi Arabia, Turkey, and Greece. Hybrid ANN models were found to be significant in India and Egypt. India is a country of extensive geographical diversities. Thus, LSTM, MT, ENN, SVR, and RBF-HPSOGA were found to be influential over the other models in different locations. Moreover, rainfall in Bangladesh was forecast more accurately when SVM, M5P-SVR, and LSTM were practiced. Studies in other countries such as Malaysia (BDTR, ELM), Taiwan (MGSVM, MLFN), China (LSTM, E-ANN-SVR), the USA (SVR, GP, RBF, LSTM-PER), the Canary Islands (XGBoost), the UK (LSTM), and Turkey (Wavelet-Fuzzy-Neuro) showed better accuracy in the models mentioned in the parenthesis. However, ANFIS proved to be the best model in Indonesia, Ethiopia, and Nigeria. Fig. 13 illustrates all the best models against the countries.

Identifying these gaps paves the way for novel research directions and fosters innovation in AI-based rainfall forecasting (Fig. 10). The field of AI-based rainfall forecasting stands to gain significant advantages by adopting similar research methodologies across various studies. These commonalities pave the way for several positive outcomes that will ultimately lead to more accurate and reliable rainfall predictions. One key benefit lies in the ability to conduct meaningful comparisons and benchmarking. When studies share a methodological foundation, researchers can directly compare their findings. This comparative analysis reveals which AI techniques perform best under specific conditions, such as short-term versus long-term forecasting or across different geographical regions. By identifying the strengths

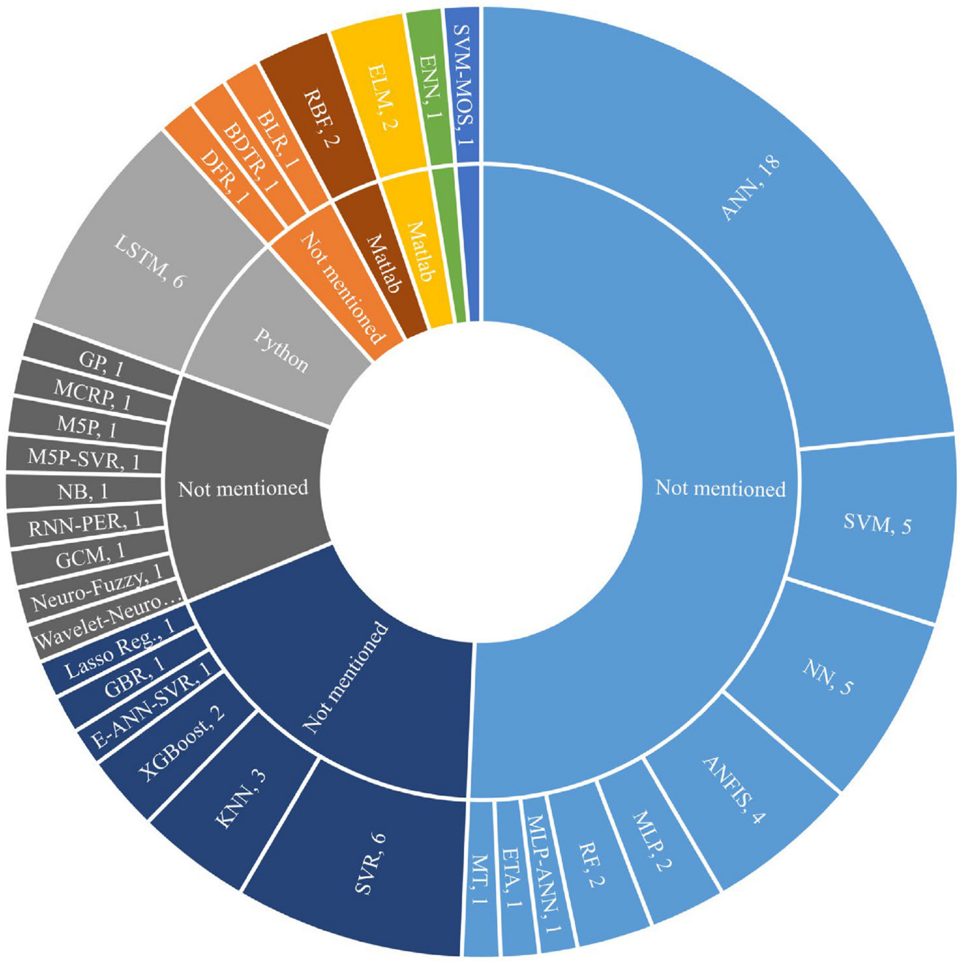

Fig. 10. Different AI methodology used in different software programs and quantity of papers.

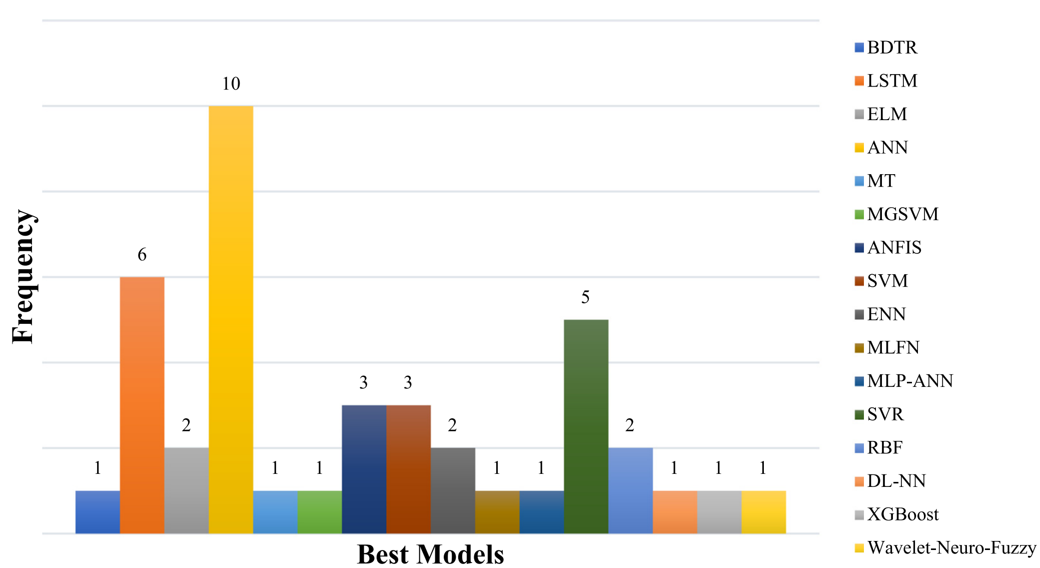

Fig. 11. The best models found in studies.

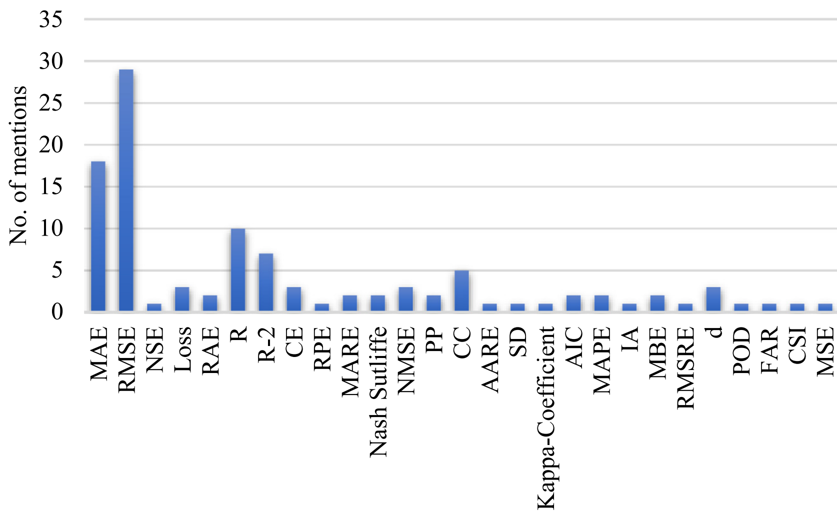

Fig. 12. Accuracy parameters used in the reviewed papers.

and weaknesses of various AI models, this analysis helps guide future research directions and resource allocation.

Furthermore, similar methodologies facilitate the process of replication and validation. Researchers can readily verify the original findings of previous studies by replicating their methods with different datasets. Consistent results across these replications strengthen the overall validity of the AI approach for rainfall forecasting. This not only builds trust in the field but also allows for a broader understanding of the model's generalizability. Standardization is another crucial benefit that emerges from shared methodologies. A common ground allows researchers to establish standardized approaches for AI-based rainfall forecasting. This encompasses factors like data pre-processing techniques, model evaluation metrics, and reporting formats. Standardized practices lead to more consistent and reliable results, ultimately improving the overall quality of future research. Researchers can spend less time reinventing the wheel and more time focusing on pushing the boundaries of the field.

The ability to build upon existing knowledge becomes significantly easier when methodologies are similar. New studies can leverage the groundwork laid by previous research. This facilitates a cumulative approach where researchers can refine existing models, explore new AI architectures, or incorporate additional data sources. This collaborative knowledge-building process accelerates advancements in AI-based rainfall forecasting, leading to more sophisticated and accurate models over time. Finally, similar methodologies open the door for meta-analysis. This powerful statistical technique combines data from multiple studies with shared methodologies. By analyzing this combined dataset, researchers can draw broader conclusions about the effectiveness of AI models for rainfall forecasting. Meta-analysis provides a more robust understanding of the overall field and helps identify promising areas for future research endeavors.

7. Conclusions

39 research articles were reviewed in this study to find out the overall scenario of rainfall forecasting models based on AI from 2000 to 2023 in different study areas over the globe. Based on the reviewed papers in this study, the rainfall forecasting using different AI models produced their apparent best performance while compared to another AI models. ANN proved to be one of the promising models that showed the highest accuracy compared to other models, but it is also true that, in some papers, there is another best model than ANN, say, Wavelet-Neuro-Fuzzy, which is a hybrid model, as example. A complex interplay between regional variables, the historical dataset, the cross-validation technique, and the model itself determines the model's overall performance. Thus, no model is supreme to be concluded. The best model based on the accuracy parameters is only can be relative for any specific region and for against some specific AI models. Some recommendations were proposed in different papers as mentioned below:

1. Forecasting models should be built in possible different hybrid AI models so that the best one can be chosen. Hybrid models are proved to consider the extreme and lowest values while forecasting for which the overall accuracy gets better [83].
2. Rainfall prediction models, despite their good accuracy, face limitations due to the stochastic nature of rainfall and dependence on various factors like geography, climate, and ocean currents. Achieving 100% accuracy remains a challenge [62].
3. The incorporation of diverse environmental parameters will be central to the continued exploration and understanding of climate variability and extreme weather events [70].
4. Integrating prediction models with advanced deep learning methods like stacked autoencoders could be promising for enhancing the performance [55].
5. Future research should focus on addressing model outliers exceeding the training set range to ensure consistent predictive power across all climates [66].

Rainfall forecasting presents a significant challenge due to the inherent variability and complexity of precipitation patterns. These patterns exhibit non-linearity, fluctuate across space and time, and possess an element of randomness. Choosing the most effective forecasting technique hinges on several factors, including the available data, the desired forecast lead time (short-term versus long-term), and the specific research goals. This section delves into the applicability of stand-alone machine learning (ML), stand-alone statistical models, and hybrid approaches in rainfall prediction. Based on the reviewed papers, conclusions were drawn as follows:

Stand-alone Machine Learning (ML) for Rainfall Forecasting

Machine learning algorithms offer a data-driven approach to rainfall forecasting. Their strength lies in their ability to learn intricate, non-linear relationships between multiple input variables, such as temperature, humidity, and pressure, and the resulting rainfall, without

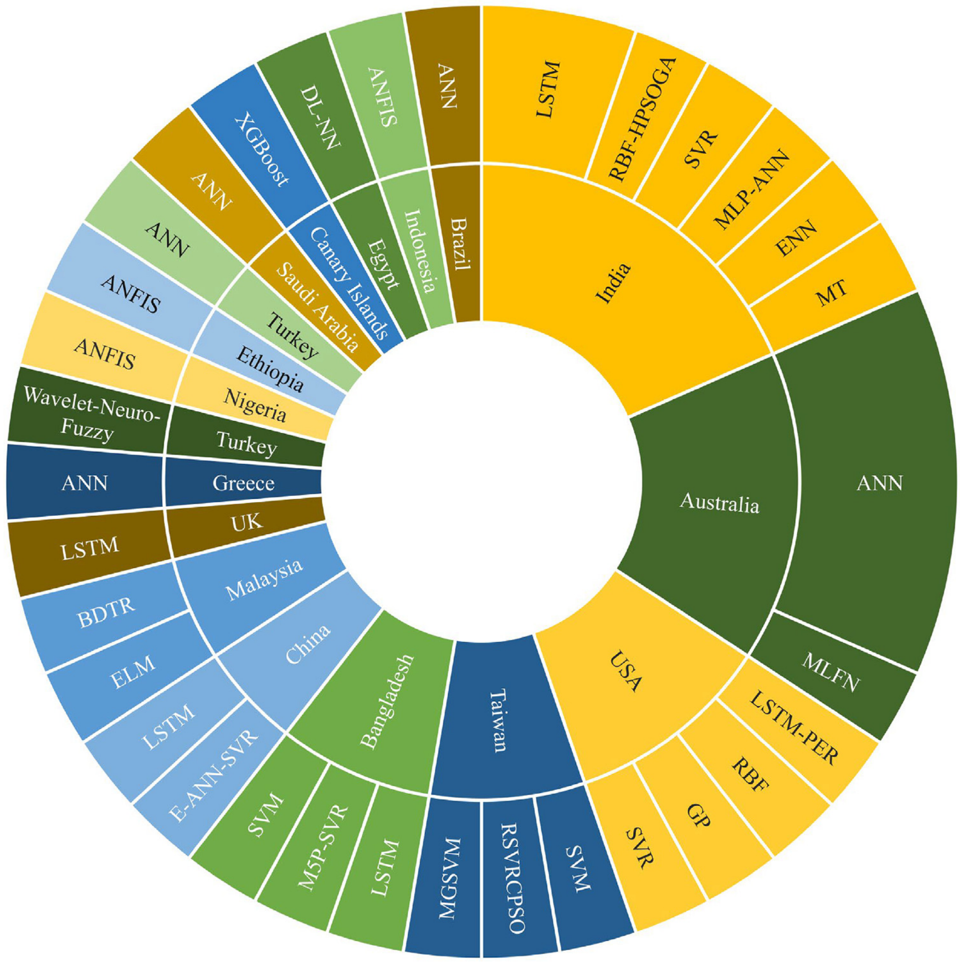

Fig. 13. The country-wise best models.

relying on explicitly defined physical relationships. This adaptability makes them particularly valuable in the context of climate change, where rainfall patterns may exhibit non-stationary behavior. Additionally, modern ML algorithms can effectively handle high-dimensional datasets, a common occurrence in environmental studies, incorporating diverse data sources like satellite imagery or radar data.

However, stand-alone ML models also have limitations. Their performance heavily relies on the quality and quantity of training data. Limited or poor-quality data can lead to overfitting, where the model performs well on the training data but fails to generalize to unseen data, resulting in inaccurate forecasts. Another drawback concerns interpretability. The inner workings of complex ML models can be opaque, making it difficult to understand how they arrive at their predictions. This lack of interpretability can limit their adoption in practical applications where understanding the reasoning behind the forecast is crucial.

Stand-alone Statistical Models for Rainfall Forecasting:

Statistical models offer an alternative approach, drawing upon established physical relationships between atmospheric variables and rainfall. This foundation in physical principles provides valuable insights into the underlying mechanisms driving precipitation patterns. Additionally, statistical models are generally easier to interpret compared to complex ML algorithms. This allows for a better understanding of the factors influencing rainfall predictions and fosters trust in the forecast. Furthermore, statistical models can be computationally less expensive to train and run compared to some ML algorithms, making them suitable for real-time forecasting applications where speed is a critical factor.

However, traditional statistical models also have limitations. They may struggle to capture the full complexity of non-linear relationships present in rainfall data, potentially leading to less accurate forecasts, particularly for complex weather patterns. Additionally, statistical models often assume a degree of stationarity in the underlying data, meaning the patterns remain relatively constant over time. This assumption might not hold true for long-term forecasts, especially in the context of climate change where rainfall patterns may exhibit significant shifts.

Hybrid Approaches: Combining Strengths for Enhanced Performance:

Hybrid approaches aim to capitalize on the complementary strengths of both ML and statistical models. Statistical models can provide a foundational framework incorporating established physical relationships, while ML techniques can learn complex non-linear patterns from the data. This combination has the potential to outperform stand-alone approaches by leveraging the power of data-driven learning alongside the established knowledge of physical processes.

Choosing the Right Approach for Rainfall Forecasting:

The optimal forecasting technique depends on the specific application and research objectives. When data availability is limited or interpretability is a top priority, statistical models might be preferred. For complex, non-linear relationships and situations with abundant data, ML approaches might be more suitable. Hybrid models offer a promising avenue for leveraging the strengths of both methodologies, potentially leading to more accurate and robust rainfall forecasts that can inform critical decisions in various sectors.

8. Limitations

This study examined a selection of 39 research papers from reputable journals that explore various artificial intelligence (AI) techniques in different countries worldwide. While the current study emphasizes the potential of hybrid models with confidence levels, expanding the scope of reviewed papers could unlock significant advances in algorithm development, reveal new research gaps, and provide specific directions for future research. This, in turn, would enable a more comprehensive literature survey to uncover new avenues for research in the domain of rainfall forecasting, ultimately contributing to the sustainable management of water resources on a global scale.

9. Interpretation of key findings

Artificial intelligence (AI) is undergoing continuous development, including increasingly sophisticated elements to devise novel approaches that offer enhanced precision in issue prediction. Recent research has focused on hybrid models rather than singular models. Furthermore, there is a strong correlation between geographic orientation and meteorological data, which significantly impacts AI models, particularly during the occurrence of extreme occurrences such as natural disasters, floods, heavy snowfall, and similar phenomena. The integration of multiple artificial intelligence (AI) models has demonstrated enhanced reliability in accurately capturing the peak, trough, and average values. The study establishes the dominance of hybrid models by integrating a predetermined confidence level, highlighting their advantage over traditional statistical or singular AI models. Furthermore, it was noticed in some reviewed works on rainfall forecasting that there is a propensity to incorporate the selection of the training and test dataset ratio into the models themselves in order to conduct a grid search for identifying the optimal model with the lowest error values. The application of AI approaches has led to significant improvements in the robustness of rainfall forecasting systems, facilitating their direct contribution to the understanding and management of climate change impacts and mitigation strategies all over the world.

Rainfall forecasting remains a complex task due to the inherent non-linearity, spatiotemporal variability, and randomness of precipitation patterns. While both stand-alone statistical models and machine learning (ML) models offer valuable tools, they each have limitations. Statistical models, while interpretable, may struggle with complex, non-linear relationships. Conversely, ML models can handle complexity but lack interpretability and may not account for established physical relationships. Hybrid models offer a compelling solution by strategically combining these approaches to optimize predictive performance in rainfall forecasting. This optimization is achieved through several key strategies:

Leveraging Strengths:

Hybrid models capitalize on the complementary strengths of both techniques. The statistical component incorporates established physical relationships between atmospheric variables and rainfall, providing a foundational framework and interpretability to the forecasts. The ML component, embedded within this framework, learns complex, non-linear patterns from the data, capturing the intricate dynamics of rainfall events. This data-driven learning enhances the model's ability to adapt to changing weather patterns and potentially outperform stand-alone models.

Addressing Data Stationarity:

Rainfall patterns, particularly over long-time scales, may exhibit non-stationary behavior due to climate change. Hybrid models address this challenge in two ways. Statistical techniques can be employed to detrend or remove non-stationary components from the data before feeding it to the ML model. Alternatively, certain hybrid models can incorporate time-varying parameters within the statistical framework, allowing for dynamic adaptation to changes in the underlying data distribution over time. Both strategies improve forecast accuracy, especially for long-term predictions [84].

Residual Considerations:

Residuals, the difference between observed and predicted rainfall values, play a crucial role in model evaluation and improvement. Hybrid models leverage residuals for further optimization. Analyzing the distribution and patterns of residuals can reveal potential biases or limitations in the model, informing the refinement of the statistical components or guiding the training process of the ML component. Additionally, ensemble learning techniques can be applied within the hybrid framework. These techniques combine predictions from multiple statistical and/or ML models, with models exhibiting lower residuals contributing more significantly to the final forecast, potentially leading to improved overall accuracy.

Data Linearity Assumptions:

Statistical models often rely on the assumption of linear relationships between variables. While hybrid models incorporate ML components that can handle non-linearity, it is important to consider the underlying assumptions of the statistical components. Transformation techniques can be employed within the hybrid framework to convert non-linear relationships into linear ones, allowing the statistical component to function effectively. Furthermore, feature engineering techniques can be used to create new features derived from the original data. These new features may capture the non-linear relationships more effectively, improving the overall performance of the hybrid model.

On the other hand, Stand-alone Machine learning (ML) algorithms excel at identifying complex patterns in data without the need for explicit programming of the underlying relationships. This makes them well-suited for rainfall forecasting, where the connections between various environmental factors and rainfall patterns can be intricate and non-linear. ML algorithms learn directly from historical rainfall data and associated environmental factors. This allows them to capture complex relationships that might be difficult to model with traditional statistical methods. However, a wide range of ML models are available, each with its own strengths and weaknesses. This allows researchers to choose the model that best suits the specific characteristics of the rainfall data and the desired forecasting horizon (short-term, medium-term, long-term). In addition, many ML models can effectively handle high-dimensional datasets, incorporating a multitude of environmental factors beyond just temperature and humidity. This comprehensive picture can lead to more accurate forecasts. On the other hand, some ML models, particularly deep neural networks, can be opaque in their decision-making processes. While they may deliver accurate forecasts, understanding the "why" behind the prediction can be challenging. Finally, the performance of ML models heavily relies on the quality and quantity of training data. Insufficient or noisy data can lead to inaccurate or unreliable forecasts.

Stand-alone statistical models are typically easier to understand than complex ML models. This allows researchers to pinpoint the specific factors driving the forecasts and assess the model's assumptions. Statistical models generally require less computational power compared to some ML algorithms. This can be advantageous for real-time forecasting applications with resource constraints. Moreover, statistical methods for rainfall forecasting have been well established and refined over decades, providing a solid foundation for reliable predictions. However, stand-alone statistical models might struggle with complex, non-linear relationships often present in rainfall data.

Hybrid approaches combine the strengths of both standalone ML and statistical models, aiming to achieve superior forecasting accuracy. Hybrid models can utilize the data-driven learning capabilities of ML models while maintaining the interpretability of statistical models. This allows for accurate forecasts with insights into the driving factors. By combining different modeling techniques, hybrid approaches can potentially capture a wider range of rainfall patterns, leading to more generalizable forecasts. On the contrary, building and tuning a hybrid model is often more complex than using standalone approaches. This requires expertise in both statistical modeling and machine learning. Combining models can sometimes lead to overfitting, especially if the individual models already have a high capacity. Careful model selection and hyperparameter tuning are crucial.

Stand-alone machine learning, statistical models, and hybridization all offer viable approaches for rainfall forecasting using AI techniques. Each approach has its own advantages and limitations. The choice of approach depends on the specific forecasting task, available data, computational resources, and the desired level of interpretability. In many cases, hybridization can be a powerful way to leverage the strengths of both machine learning and statistical models, leading to more accurate and informative rainfall forecasts.

10. Future scope of research

The utilization of AI models in rainfall forecasting holds potential benefits not just for meteorologists, scientists, and researchers, but also for diverse industries and society at large. The potential for utilizing AI models in rainfall forecasting for sustainable water resource management is considerable, including various areas such as reservoir management, groundwater recharge planning, drought mitigation, flood prediction, and management, among others. The growing concern about water scarcity on a global scale has prompted the exploration of AI-driven rainfall forecasting as a crucial tool for enhancing the efficient distribution and preservation of water resources. This technology offers hope for a sustainable future by holding potential for advancements in mitigating climate change, adapting to its impacts, and addressing growing water needs. This research highlights the importance of international cooperation in developing AI-powered rainfall forecasting for water resource sustainability, through the exchange of knowledge, data, and successful strategies.

This comprehensive review has shed light on the potential and limitations of various machine learning (ML) models used in rainfall forecasting. While these models offer significant advantages, further research is crucial to unlock their full potential and ensure their effectiveness in real-world applications. Here, we propose specific and actionable recommendations to guide future research efforts.

Enhancing Interpretability and Understanding:

One key area for exploration involves Explainable Artificial Intelligence (XAI) techniques. Integrating XAI within ML models can significantly enhance their interpretability. By understanding the factors driving rainfall predictions, researchers and users can gain valuable insights into the model's decision-making process. Furthermore, evaluating the effectiveness of different XAI methods in explaining complex, non-linear relationships learned by ML models is crucial. This will allow researchers to choose the most appropriate XAI technique for a specific forecasting scenario.

Leveraging Physical Knowledge for Improved Accuracy:

Another promising avenue for research lies in the development of hybrid approaches. These models combine the strengths of ML with established physical relationships governing atmospheric dynamics. By incorporating this knowledge within the ML framework, researchers can potentially improve the accuracy and robustness of the forecasts, particularly for long-term predictions. Evaluating the performance of hybrid models compared to standalone ML models will provide valuable insights into the effectiveness of this approach.

Quantifying Uncertainty for Informed Decision-Making:

Rainfall forecasts are inherently uncertain due to the complex nature of weather patterns. To address this challenge, future research should focus on developing frameworks for uncertainty quantification within ML models. These frameworks can equip the models with the ability to quantify and communicate the inherent uncertainties associated with the predicted rainfall amounts. Evaluating the effectiveness of different uncertainty quantification methods is essential. This will ensure that users receive clear information about the confidence level associated with the forecasts, allowing them to make more informed decisions.

Addressing Data Limitations through Augmentation:

The performance of ML models heavily relies on the quality and quantity of training data. Sparse or noisy real-world rainfall data can hinder model performance. Investigating the efficacy of data augmentation techniques is a promising approach to address these limitations. Data augmentation involves techniques like generating synthetic data to artificially expand the training dataset. Assessing the impact of such techniques on the performance and generalizability of ML models will be crucial in determining their effectiveness for rainfall forecasting.

Incorporating High-Resolution Data for Improved Representation:

Rainfall patterns exhibit significant spatial and temporal variability. To capture these nuances more effectively, future research should explore the feasibility of incorporating high-resolution data sources into ML models. Examples of such data sources include satellite imagery and radar data. These data sources offer a more detailed picture of atmospheric conditions, potentially leading to improved forecasts. Evaluating the computational efficiency and scalability of ML models trained on high-resolution data is vital. This will ensure that the benefits of increased detail do not come at the cost of impractical computational demands.

By pursuing these research directions within a defined timeframe, the field of ML-based rainfall forecasting can move forward significantly. The proposed recommendations offer a roadmap for researchers and developers, focusing on key areas that can contribute to the development of more robust, interpretable, and reliable models for accurate rainfall predictions.

11. List of abbreviations

ANN	Artificial Neural Network
SVR	Support Vector Regression
LSTM	Long Short-Term Memory
NN	Neural Network
ANFIS	Adaptive Neuro-Fuzzy Inference System
SVM	Support Vector Machine
MLP	Multilayer Perceptron
MLPNN	Multilayer Perceptron Neural Network
RF	Random Forest
ELM	Extreme Learning Machine
XGBoost	Extreme Gradient Boosting
RBF	Radial Basis Function
ETA	Estimated Time of Arrival
ENN	Ensemble Neural Network
KNN	K-Nearest Neighbors
RNN	Recurrent Neural Network
RSVRCPSO	Recurrent SVR with Chaotic Particle Swarm Optimization
MT	Model Trees
ARIMA	Autoregressive Integrated Moving Average
GP	Genetic Programming
M5	Model Tree
WLSTM	Wavelet Long Short-Term Memory
CLSTM	Convolutional Long Short-Term Memory
GCM	Global Climate Model
MLFN	Multilayer Feedforward Network
PRNN	Probabilistic Recurrent Neural Network
TDNN	Time Delay Neural Network
SVRCPSO	SVR with Chaotic Particle Swarm Optimization
MGSVM	Multi-Group Support Vector Machine

MCRP	Markov Chain Recurrent Plot
E-ANN	Ensemble Artificial Neural Network
E-SVR	Ensemble Support Vector Regression
E-ANN-SVR	Ensemble of ANN and SVR
RNN-PER	Recurrent Neural Network for Periodicity Detection
Persist-PER	Persistence-Based Periodicity Detection
Farne-PER	Farneback-Based Periodicity Detection
LSTM-PER	LSTM-Based Periodicity Detection
SVM-MOS	Support Vector Machine for Mean Opinion Score Prediction
RBF-NN	Radial Basis Function Neural Network
RBF-GA	Radial Basis Function with Genetic Algorithm
RBF-HPSOGA	Radial Basis Function with Hybrid Particle Swarm Optimization and Genetic Algorithm
BLR	Bayesian Linear Regression
DFR	Dynamic Factor Regression
NNR	Neural Network Regression
BDTR	Boosted Decision Tree Regression
S-LSTM	Stacked Long Short-Term Memory
B-LSTM	Bidirectional Long Short-Term Memory
GBR	Gradient Boosting Regression
LSVR	Least Squares Support Vector Regression
DL-NN	Deep Learning Neural Network
M5P	Model Tree with Pruning
M5P-SVR	Model Tree with Pruning combined with Support Vector Regression
NB	Naïve Bayes
MLPNN	Multilayer Perceptron Neural Network
MLP-ANN	Multilayer Perceptron Artificial Neural Network
POAMA-1.5 GCM	Predictive Ocean Atmosphere Model for Australia –1.5 general circulation model
TLRN	Time Lagged Recurrent Networks
IMFs	Intrinsic Mode Functions

Funding information

This research was undertaken without the benefit of prior external funding.

CRediT authorship contribution statement

Md. Abu Saleh: Writing – review & editing, Writing – original draft, Visualization, Software, Resources, Methodology, Investigation, Formal analysis. **H.M. Rasel:** Writing – review & editing, Writing – original draft, Supervision, Formal analysis, Conceptualization. **Briti Ray:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The analysis in this publication is based on the reviewed research papers and presented in this paper.

Acknowledgments

The present study expresses sincere gratitude to the authors of the examined research papers for their contributions in terms of concepts, methodology, algorithm developments, innovative problem-solving approaches, and potential recommendations for future research directions.

References

[1] R. Ahmed, Probabilistic estimates of rainfall extremes in Bangladesh during the premonsoon season, Indian Geogr. J. 64 (1989) 39–53, https://cir.nii.ac.jp/crid/1570291225229159424.

[2] D. Angchok, Traditional method of rainfall prediction through Almanacs in Ladakh, Indian J. Traditional Knowl. 5 (2006) 145–150.

[3] M. Balehegn, S. Balehey, C. Fu, W. Liang, Indigenous weather and climate forecasting knowledge among Afar pastoralists of north eastern Ethiopia: Role in adaptation to weather and climate variability, Pastoralism 9 (1) (2019) 8, http://dx.doi.org/10.1186/s13570-019-0143-y.

[4] M. Dastorani, M. Mirzavand, M.T. Dastorani, S.J. Sadatinejad, Comparative study among different time series models applied to monthly rainfall forecasting in semi-arid climate condition, Nat. Hazards 81 (3) (2016) 1811–1827, http://dx.doi.org/10.1007/s11069-016-2163-x.

[5] I. Hossain, H.M. Rasel, M.A. Imteaz, F. Mekanik, Long-term seasonal rainfall forecasting: Efficiency of linear modelling technique, Environ. Earth Sci. 77 (7) (2018) 280, http://dx.doi.org/10.1007/s12665-018-7444-0.

[6] F. Islam, M. Imteaz, V. Boulomytis, H.M. Rasel, Combined regression modelling of autumn rainfall in Western Australia using potential climate indices, in: 37th Hydrology & Water Resources Symposium 2016: Water, Infrastructure and the Environment, 2016.

[7] F. Islam, M. Imteaz, H.M. Rasel, Analysing the effect of lagged climate indices on rainfall predictability for Western Australia's North Coast Region, in: 22nd International Congress on Modelling and Simulation, Hobart, Tasmania, Australia, 2017, pp. 1600–1606.

[8] Y. Lai, D.A. Dzombak, Use of the Autoregressive Integrated Moving Average (ARIMA) model to forecast near-term regional temperature and precipitation, Weather Forecast. 35 (3) (2020) 959–976, http://dx.doi.org/10.1175/WAF-D-19-0158.1.

[9] P. Narayanan, A. Basistha, S. Sarkar, S. Kamna, Trend analysis and ARIMA modelling of pre-monsoon rainfall data for Western India, C. R. Geosci. 345 (1) (2013) 22–27, http://dx.doi.org/10.1016/j.crte.2012.12.001.

[10] T. Olatayo, A. Taiwo, Statistical modelling and prediction of rainfall time series data statistical modelling and prediction of rainfall time series data, 2019.

[11] M.A. Rahman, L. Yunsheng, N. Sultana, Analysis and prediction of rainfall trends over Bangladesh using Mann–Kendall, Spearman's rho tests and ARIMA model, Meteorol. Atmos. Phys. 129 (4) (2017) 409–424, http://dx.doi.org/10.1007/s00703-016-0479-4.

[12] H.M. Rasel, M.A. Imteaz, F. Mekanik, Evaluating the effects of lagged ENSO and SAM as potential predictors for long-term rainfall forecasting, in: Proceedings of the International Conference on Water Resources and Environment (WRE 2015), Beijing, China, 2015, pp. 125–129, http://dx.doi.org/10.1201/b19079-23.

[13] H.M. Rasel, M.A. Imteaz, F. Mekanik, Multiple regression modelling approach for rainfall prediction using large-scale climate indices as potential predictors, Int. J. Water 11 (3) (2017) 209–225, http://dx.doi.org/10.1504/IJW.2017.085879.

[14] S. Swain, S. Nandi, P. Patel, Development of an ARIMA model for monthly rainfall forecasting over Khordha district, in: P.K. Sa, S. Bakshi, I.K. Hatzilygeroudis, M.N. Sahoo (Eds.), Odisha, India BT - Recent Findings in Intelligent Computing Techniques, Springer, Singapore, 2018, pp. 325–331.

[15] J.G. De Gooijer, The role of time series analysis in forecasting: A personal view, Int. J. Forecast. 6 (4) (1990) 449–451, http://dx.doi.org/10.1016/0169-2070(90)90020-C.

[16] J.G. De Gooijer, R.J. Hyndman, 25 Years of time series forecasting, Int. J. Forecast. 22 (3) (2006) 443–473, http://dx.doi.org/10.1016/j.ijforecast.2006.01.001.

[17] F. Joutz, R. Trost, Using stochastic simulation to test the effect of seasonal adjustment on forecast standard errors of motor gasoline demand, Int. J. Forecast. 8 (2) (1992) 219–231, http://dx.doi.org/10.1016/0169-2070(92)90120.

[18] I.-C. Baciu, Stochastic models for forecasting inflation rate, empirical evidence from Romania, Procedia Econ. Finance 20 (2015) 44–52, http://dx.doi.org/10.1016/S2212-5671(15)00045-3.

[19] C. Buelens, Inflation forecasting and the crisis: Assessing the impact on the performance of different forecasting models and methods, 2012, Directorate General Economic and Financial Affairs (DG ECFIN).

[20] J.C.C. Chan, Moving average stochastic volatility models with application to inflation forecast, J. Econometrics 176 (2) (2013) 162–172, http://dx.doi.org/10.1016/j.jeconom.2013.05.003.

[21] Nugent W.R., 2 Regression-discontinuity and arima models, in: W. Nugent (Ed.), Analyzing Single System Design Data, Oxford University Press, 2009, http://dx.doi.org/10.1093/acprof:oso/9780195369625.003.0002.

[22] D. Hain, R. Jurowetzki, S. Lee, Y. Zhou, Machine learning and artificial intelligence for science, technology, innovation mapping and forecasting: review, synthesis, and applications, Scientometrics 128 (3) (2023) 1465–1472, http://dx.doi.org/10.1007/s11192-022-04628-8.

[23] I.H. Sarker, Machine learning: Algorithms, real-world applications and research directions, SN Comput. Sci. 2 (3) (2021) 160, http://dx.doi.org/10.1007/s42979-021-00592-x.

[24] H.S. Barjouei, H. Ghorbani, N. Mohamadian, D.A. Wood, S. Davoodi, J. Moghadasi, H. Saberi, Prediction performance advantages of deep machine learning algorithms for two-phase flow rates through wellhead chokes, J. Petrol. Explor. Prod. 11 (3) (2021) 1233–1261, http://dx.doi.org/10.1007/s13202-021-01087-4.

[25] T. Götze, M. Gürtler, E. Witowski, Forecasting accuracy of machine learning and linear regression: evidence from the secondary CAT bond market, J. Bus. Econ. (2023) http://dx.doi.org/10.1007/s11573-023-01138-8.

[26] H.M. Rasel, Long term rainfall forecast modelling for South Australia using large scale climate predictors (Issue April), 2018, https://researchbank.swinburne.edu.au/file/5cea66e5-fb87-493e-8909-f0cc7c543265/1/h_rasel_thesis.pdf.

[27] V. Barcala, J. Rozemeijer, K. Ouwerkerk, L. Gerner, L. Osté, Value and limitations of machine learning in high-frequency nutrient data for gap-filling, forecasting, and transport process interpretation, Environ. Monit. Assess. 195 (7) (2023) 892, http://dx.doi.org/10.1007/s10661-023-11519-9.

[28] M.R. Carbone, When not to use machine learning: A perspective on potential and limitations, MRS Bull. 47 (9) (2022) 968–974, http://dx.doi.org/10.1557/s43577-022-00417-z.

[29] F. Chen, Z. Cao, E.M. Grais, F. Zhao, Contributions and limitations of using machine learning to predict noise-induced hearing loss, Int. Arch. Occup. Environ. Health 94 (5) (2021) 1097–1111, http://dx.doi.org/10.1007/s00420-020-01648-w.

[30] M.R. Dobbelaere, P.P. Plehiers, R. Van de Vijver, C.V. Stevens, K.M. Van Geem, Machine learning in chemical engineering: Strengths, weaknesses, opportunities, and threats, Engineering 7 (9) (2021) 1201–1211, http://dx.doi.org/10.1016/j.eng.2021.03.019.

[31] M.K. Gajendran, I.F.S.A. Kabir, S. Purohit, E.Y.K. Ng, On the Limitations of Machine Learning (ML) Methodologies in Predicting the Wake Characteristics of Wind Turbines BT - Renewable Energy Systems in Smart Grid, Springer Nature, Singapore, 2022, pp. 15–23,

[32] E.M. Onyema, K.K. Almuzaini, F.U. Onu, D. Verma, U.S. Gregory, M. Puttaramaiah, R.K. Afriyie, Prospects and challenges of using machine learning for academic forecasting, Comput. Intell. Neurosci. 2022 (2022) 5624475, http://dx.doi.org/10.1155/2022/5624475.

[33] T. Rajaee, A. Boroumand, Forecasting of chlorophyll-a concentrations in South San Francisco Bay using five different models, Appl. Ocean Res. 53 (2015) 208–217, http://dx.doi.org/10.1016/j.apor.2015.09.001.

[34] K.C. Luk, J.E. Ball, A. Sharma, A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting, J. Hydrol. 227 (1–4) (2000) 56–65, http://dx.doi.org/10.1016/S0022-1694(99)00165-1.

[35] K.C. Luk, J.E. Ball, A. Sharma, An application of artificial neural networks for rainfall forecasting, Math. Comput. Model. 33 (6–7) (2001) 683–693, http://dx.doi.org/10.1016/S0895-7177(00)00272-7.

[36] M.C. Valverde Ramírez, H.F. de Campos Velho, N.J. Ferreira, Artificial neural network technique for rainfall forecasting applied to the São Paulo region, J. Hydrol. 301 (1) (2005) 146–162, http://dx.doi.org/10.1016/j.jhydrol.2004.06.028.

[37] T. Partal, Ö. Kişi, Wavelet and neuro-fuzzy conjunction model for precipitation forecasting, J. Hydrol. 342 (1–2) (2007) 199–212, http://dx.doi.org/10.1016/j.jhydrol.2007.05.026.

[38] W.-C. Hong, Rainfall forecasting by technological machine learning models, Appl. Math. Comput. 200 (1) (2008) 41–57, http://dx.doi.org/10.1016/j.amc.2007.10.046.

[39] M. Bilgili, B. Sahin, Prediction of long-term monthly temperature and rainfall in Turkey, Energy Sour. A 32 (1) (2010) 60–71, http://dx.doi.org/10.1080/15567030802467522.

[40] T. Mandal, V. Jothiprakash, Short-term rainfall prediction using ANN and MT techniques, ISH J. Hydraul. Eng. 18 (1) (2012) 20–26, http://dx.doi.org/10.1080/09715010.2012.661629.

[41] J. Abbot, J. Marohasy, Application of artificial neural networks to rainfall forecasting in Queensland, Australia, Adv. Atmos. Sci. 29 (2012) http://dx.doi.org/10.1007/s00376-012-1259-9.

[42] P.T. Nastos, A.G. Paliatsos, K.V. Koukouletsos, I.K. Larissi, K.P. Moustris, Artificial neural networks modeling for forecasting the maximum daily total precipitation at Athens, Greece, Atmos. Res. 144 (2014) 141–150, http://dx.doi.org/10.1016/j.atmosres.2013.11.013.

[43] J. Farajzadeh, A. Fakheri Fard, S. Lotfi, Modeling of monthly rainfall and runoff of Urmia lake basin using feed-forward neural network and time series analysis model, Water Resour. Ind. 7 (8) (2014) 38–48, http://dx.doi.org/10.1016/j.wri.2014.10.003, 2014.

[44] J. Abbot, J. Marohasy, Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks, Atmos. Res. 138 (July) (2014) 166–178, http://dx.doi.org/10.1016/j.atmosres.2013.11.002.

[45] G.F. Lin, B.C. Jhong, A real-time forecasting model for the spatial distribution of typhoon rainfall, J. Hydrol. 521 (2015) 302–313, http://dx.doi.org/10.1016/j.jhydrol.2014.12.009.

[46] P. Yu, T. Yang, S. Chen, C. Kuo, Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting, 2017.

[47] S. Cramer, M. Kampouridis, A.A. Freitas, A.K. Alexandridis, An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives, Expert Syst. Appl. 85 (2017) 169–181, http://dx.doi.org/10.1016/j.eswa.2017.05.029.

[48] Y. Dash, S.K. Mishra, B.K. Panigrahi, Rainfall prediction for the Kerala state of India using artificial intelligence approaches, Comput. Electr. Eng. 70 (May) (2018) 66–73, http://dx.doi.org/10.1016/j.compeleceng.2018.06.004.

[49] Y. Xiang, L. Gou, L. He, S. Xia, W. Wang, A SVR–ANN combined model based on ensemble EMD for rainfall prediction, Appl. Soft Comput. 73 (2018) 874–883, http://dx.doi.org/10.1016/j.asoc.2018.09.018.

[50] A. Akbari Asanjan, T. Yang, K. Hsu, S. Sorooshian, J. Lin, Q. Peng, Short-term precipitation forecast based on the PERSIANN system and LSTM recurrent neural networks, J. Geophys. Res.: Atmos. 123 (22) (2018) 12, http://dx.doi.org/10.1029/2018JD028375, 512-543, 563.

[51] K. Alotaibi, A.R. Ghumman, H. Haider, Y.M. Ghazaw, M. Shafiquzzaman, Future predictions of rainfall and temperature using GCM and ANN for arid regions: A case study for the Qassim region, Saudi Arabia, Water 10 (9) (2018) http://dx.doi.org/10.3390/w10091260.

[52] S.H. Pour, S. Shahid, E.-S. Chung, X.-J. Wang, Model output statistics downscaling using support vector machine for the projection of spatial and temporal changes in rainfall of Bangladesh, Atmos. Res. 213 (2018) 149–162, http://dx.doi.org/10.1016/j.atmosres.2018.06.006.

[53] M. Zeynoddin, H. Bonakdari, A. Azari, I. Ebtehaj, B. Gharabaghi, H. Riahi Madavar, Novel hybrid linear stochastic with non-linear extreme learning machine methods for forecasting monthly rainfall a tropical climate, J. Environ. Manag. 222 (February) (2018) 190–206, http://dx.doi.org/10.1016/j.jenvman.2018.05.072.

[54] S. Kumar, T. Roshni, D. Himayoun, A comparison of Emotional Neural Network (ENN) and Artificial Neural Network (ANN) approach for rainfall-runoff modelling, Civ. Eng. J. (Iran) 5 (10) (2019) 2120–2130, http://dx.doi.org/10.28991/cej-2019-03091398.

[55] Q. Liu, Y. Zou, X. Liu, N. Linge, A survey on rainfall forecasting using artificial neural network, Int. J. Embed. Syst. 11 (2) (2019) 240–249, http://dx.doi.org/10.1504/IJES.2019.098300.

[56] S. Poornima, M. Pushpalatha, Prediction of rainfall using intensified LSTM based recurrent neural network with weighted linear units, Atmosphere 10 (11) (2019) http://dx.doi.org/10.3390/atmos10110668.

[57] R. Aguasca-Colomo, D. Castellanos-Nieves, M. Méndez, Comparative analysis of rainfall prediction models using machine learning in Islands with complex orography: Tenerife Island, Appl. Sci. 9 (22) (2019) http://dx.doi.org/10.3390/app9224931.

[58] I. Hossain, H.M. Rasel, M.A. Imteaz, F. Mekanik, Long-term seasonal rainfall forecasting using linear and non-linear modelling approaches: A case study for Western Australia, Meteorol. Atmos. Phys. 132 (1) (2019) 131–141, http://dx.doi.org/10.1007/s00703-019-00679-4.

[59] I. Hossain, H.M. Rasel, F. Mekanik, M.A. Imteaz, Artificial neural network modelling technique in predicting Western Australian seasonal rainfall, Int. J. Water 14 (1) (2020) 14–28, http://dx.doi.org/10.1504/IJW.2020.112711.

[60] W. Suparta, A.A. Samah, Rainfall prediction by using ANFIS times series technique in South Tangerang, Indonesia, Geodesy Geodynam. 11 (6) (2020) 411–417, http://dx.doi.org/10.1016/j.geog.2020.08.001.

[61] L. Ni, D. Wang, V.P. Singh, J. Wu, Y. Wang, Y. Tao, J. Zhang, Streamflow and rainfall forecasting by two long short-term memory-based models, J. Hydrol. 583 (2020) 124296, http://dx.doi.org/10.1016/j.jhydrol.2019.124296.

[62] M. Mohammed, R. Kolapalli, N. Golla, S.S. Maturi, Prediction of rainfall using machine learning techniques, Int. J. Sci. Technol. Res. 9 (01) (2020) 3236–3240.

[63] L.C.P. Velasco, R.P. Serquiña, M.S.A. Abdul Zamad, B.F. Juanico, J.C. Lomocso, Week-ahead rainfall forecasting using multilayer perceptron neural network, Procedia Comput. Sci. 161 (2019) 386–397, http://dx.doi.org/10.1016/j.procs.2019.11.137.

[64] W.M. Ridwan, M. Sapitang, A. Aziz, K.F. Kushiar, A.N. Ahmed, A. El-Shafie, Rainfall forecasting model using machine learning methods: Case study Terengganu, Malaysia, Ain Shams Eng. J. 12 (2) (2021) 1651–1663, http://dx.doi.org/10.1016/j.asej.2020.09.011.

[65] A.Y. Barrera-Animas, L.O. Oyedele, M. Bilal, T.D. Akinosho, J.M.D. Delgado, L.A. Akanbi, Rainfall prediction: A comparative analysis of modern machine learning algorithms for time-series forecasting, Mach. Learn. Appl. 7 (2021) (2022) 100204, http://dx.doi.org/10.1016/j.mlwa.2021.100204.

[66] A. Elbeltagi, Z. Bilel, N. Bailek, K. Bouchouicha, D.C. Pande, C. Santos, A.R.M.T. Islam, N. Al-Ansari, E.-S. El-kenawy, Optimizing hyperparameters of deep hybrid learning for rainfall prediction: A case study of a mediterranean basin, Arab. J. Geosci. 15 (10) (2022) http://dx.doi.org/10.1007/s12517-022-10098-2.

[67] F. Di Nunno, F. Granata, Q.B. Pham, G. de Marinis, Precipitation forecasting in northern Bangladesh using a hybrid machine learning model, Sustainability 14 (5) (2022) http://dx.doi.org/10.3390/su14052663.

[68] M. Billah, M.N. Adnan, M.R. Akhond, R.R. Ema, M.A. Hossain, S. Md. Galib, Rainfall prediction system for Bangladesh using long short-term memory, Open Comput. Sci. 12 (1) (2022) 323–331, http://dx.doi.org/10.1515/comp-2022-0254.

[69] O.S. Ojo, S.T. Ogunjo, Machine learning models for prediction of rainfall over Nigeria, Sci. Afr. 16 (2022) e01246, http://dx.doi.org/10.1016/j.sciaf.2022.e01246.

[70] P. Hunasigi, S. Jedhe, M. Mane, V. Patil-Shinde, Multilayer perceptron neural network based models for prediction of the rainfall and reference crop evapotranspiration for sub-humid climate of Dapoli, Ratnagiri district, India, Acta Ecol. Sinica 43 (1) (2023) 154–201, http://dx.doi.org/10.1016/j.chnaes.2022.09.004.

[71] W.T. Abebe, D. Endalie, Artificial intelligence models for prediction of monthly rainfall without climatic data for meteorological stations in Ethiopia, J. Big Data 10 (1) (2023) 2, http://dx.doi.org/10.1186/s40537-022-00683-3.

[72] J. Nithyashri, R.K. Poluru, S. Balakrishnan, M. Ashok Kumar, P. Prabu, S. Nandhini, IOT based prediction of rainfall forecast in coastal regions using deep reinforcement model, Measurement: Sensors 29 (August) (2023) http://dx.doi.org/10.1016/j.measen.2023.100877.

[73] C.S. Bojer, Understanding machine learning-based forecasting methods: A decomposition framework and research opportunities, Int. J. Forecast. 38 (4) (2022) 1555–1561, http://dx.doi.org/10.1016/j.ijforecast.2021.11.003.

[74] K. Prathibha, G.Rithvik. Reddy, H. Kosre, K.Lohith. Kumar, A. Rajak, R. Tripathi, in: D.S. Sisodia, L. Garg, R.B. Pachori, M. Tanveer (Eds.), Rainfall Prediction Using Machine Learning BT - Machine Intelligence Techniques for Data Analysis and Signal Processing, Springer Nature, Singapore, 2023, pp. 457–468.

[75] Affairs, U.N.D. of E. and S., World Economic Situation and Prospects 2022, 2022nd ed., United Nations, 2022, https://www.un-ilibrary.org/content/books/9789210011839.

[76] P.-S. Yu, T.-C. Yang, S.-Y. Chen, C.-M. Kuo, H.-W. Tseng, Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting, J. Hydrol. 552 (2017) 92–104, http://dx.doi.org/10.1016/j.jhydrol.2017.06.020.

[77] J. Fan, W. Yue, L. Wu, F. Zhang, H. Cai, X. Wang, X. Lu, Y. Xiang, Evaluation of SVM, ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in different climates of China, Agric. Forest Meteorol. 263 (2018) 225–241, http://dx.doi.org/10.1016/j.agrformet.2018.08.019.

[78] L.C.P. Velasco, R.P. Serquiña, M.S.A.A. Zamad, B.F. Juanico, J.C. Lomocso, Performance analysis of multilayer perceptron neural network models in week-ahead rainfall forecasting, Int. J. Adv. Comput. Sci. Appl. 10 (3) (2019) 578–588, http://dx.doi.org/10.14569/IJACSA.2019.0100374.

[79] L.C. Velasco, J.M. Aca-ac, J.J. Cajes, N.J. Lactuan, S.C. Chit, Rainfall forecasting using support vector regression machines, Int. J. Adv. Comput. Sci. Appl. 13 (3) (2022) 231–237, http://dx.doi.org/10.14569/IJACSA.2022.0130329.

[80] G. Bontempi, S. Ben Taieb, Y.-A. Le Borgne, in: M.-A. Aufaure, E. Zimányi (Eds.), Machine Learning Strategies for Time Series Forecasting BT - Business Intelligence: Second European Summer School, eBISS 2012, Brussels, Belgium, July (2012) 15-21, in: Tutorial Lectures, Springer Berlin Heidelberg, 2013, pp. 62–77, http://dx.doi.org/10.1007/978-3-642-36318-4_3.

[81] S. Kolassa, Why the best point forecast depends on the error or accuracy measure, Int. J. Forecast. 36 (1) (2020) 208–211, http://dx.doi.org/10.1016/j.ijforecast.2019.02.017.

[82] I. Kalinina, A. Gozhyj, in: S. Babichev, V. Lytvynenko (Eds.), Methodology for Solving Forecasting Problems Based on Machine Learning Methods BT - Lecture Notes in Data Engineering, in: Computational Intelligence, and Decision Making, Springer International Publishing, 2023, pp. 105–125.

[83] L. Parviz, K. Rasouli, A. Torabi Haghighi, Improving hybrid models for precipitation forecasting by combining nonlinear machine learning methods, Water Resour. Manag. 37 (10) (2023) 3833–3855, http://dx.doi.org/10.1007/s11269-023-03528-7.

[84] L.C.P. Velasco, D.L.L. Polestico, G.P.O. Macasieb, M.B.V. Reyes, F.B. Vasquez, Load forecasting using autoregressive integrated moving average and artificial neural network, Int. J. Adv. Comput. Sci. Appl. 9 (7) (2018) 14–22, http://dx.doi.org/10.14569/IJACSA.2019.0101103.