



OPEN

Climate modeling for South Asia: statistical and deep learning for rainfall and temperature prediction

Pradeep Mishra¹, Soumik Ray²✉, Priyanka Lal³, Sanjith Bharatharajan Nair⁴, Adelajda Matuka⁵, Yusra Tashkandy⁶ & Walid Emam⁶

We benchmark four approaches for hydrometeorological forecasting—SARIMA, TDNN, LSTM, and XGBoost—within one reproducible pipeline for monthly rainfall and temperature across seven South Asian countries using a centennial dataset from 1901 to 2023 with 1,476 months per series. The novelty lies in the joint, cross-country comparison of classical and deep learning models on century-scale data under unified preprocessing, early stopping, and cross-validated evaluation with RMSE, MAPE, R^2 , and CV-RMSE. Skill varies by variable and region. Rainfall often favors TDNN or XGBoost due to nonlinear dynamics. Minimum temperature favors LSTM, reflecting long memory. Stable seasonal series suit SARIMA. We convert these patterns into a hybrid selection guide that maps variable and location to the best model. The framework produces 2024–2025 forecasts with uncertainty bands to inform crop calendars, irrigation scheduling, flood readiness, heat warnings, and water allocation. The dataset scale and diagnostics support replication and policy use.

Keywords Machine learning, Rainfall prediction, Time series analysis, South asia climate, Temperature extremes

Monitoring the temporal variations in rainfall and severe temperatures is crucial for both climatic hazard anticipation and resource management planning in South Asia. The interplay of monsoonal conditions and escalating extreme weather in this region necessitates reliable forecasting to sustain agriculture, water resources, and disaster preparedness. Over the past few decades, time series forecasting methodologies have advanced significantly. The application of hybrid models to examine long-term climate data is infrequently investigated, particularly over centenary timeframes. South Asian nations, including Afghanistan, Bangladesh, Bhutan, India, Sri Lanka, Nepal, and Pakistan, typically experience distinct seasons, abrupt weather changes, and unpredictable anomalies¹. Severe monsoons, heatwaves, droughts, and cold surges impact these nations due to their distinctive topography and unpredictable atmospheric dynamics. Annual rainfall can significantly fluctuate both temporally and spatially within a nation, whereas temperature ranges often fluctuate due to elevation and proximity to oceanic or desert systems². Sri Lanka exhibits consistent temperature patterns, but Afghanistan and Nepal experience significantly low minima. Due to significant variations in regional and national climates, conventional forecasting methods face challenges, necessitating the use of adaptive models that address nonlinearity in long-term data.

South Asia's climate adaptation depends on accurate month-by-month forecasts of rainfall and temperature. Irregular monsoon behavior and the recurrence of heatwaves and floods create direct challenges for crop calendars, water management, and disaster preparedness. Research on trend analysis and teleconnections with the tropical Pacific shows that hydrometeorological series are sensitive to low-frequency variability, which underscores the need for data-driven models that capture signal while avoiding systematic bias^{3–5}.

The literature has developed along two tracks. The first uses classical time-series models such as ARIMA and SARIMA to forecast rainfall and temperature at local and regional scales, including South Asian countries and nearby contexts in space and time^{1,6}. The second applies machine learning and deep learning, including artificial neural networks, LSTM, CNN-LSTM, XGBoost, and hybrid frameworks. These models have been used for rainfall, streamflow, groundwater levels, and related domains, with reported gains at short and medium

¹College of Agriculture, Jawaharlal Nehru Krishi Vishwa Vidyalaya, Rewa, M.P., India. ²Department of Agricultural Economics and Statistics, Centurion University of Technology and Management, Paralakhemundi, Odisha 761211, India. ³ICAR- Indian Institute of Pulses Research, Kanpur, Uttar Pradesh, India. ⁴Department of Mathematical and Physical Sciences, University of Nizwa, Initial Campus, Birkat Al Mouz, Oman. ⁵Department of Economics, Faculty of Economics, University of Bologna, Bologna, Italy. ⁶Department of Statistics and Operations Research, College of Science, King Saud University, P.O. Box 2455, 11451 Riyadh, Saudi Arabia. ✉email: raysoumik4@gmail.com

lead times^{7–16}. Large-scale feature studies and model-combination work also support the extraction of stable predictive information across ungauged sites or through multi-model ensembles^{17–19}.

Yet systematic, multi-country comparisons over century-long records remain limited. Many studies focus on a single site, a single model, or shorter periods. Others use daily data that do not map directly to monthly planning decisions. Evaluation practices and performance metrics vary across studies, which reduces generalizability^{8,9,17,18,20}. While classical approaches have been used for South Asia and recent national studies deploy XGBoost or hybrid methods, the field lacks a joint comparison that unifies preprocessing and evaluation across classical, deep, and boosting models on century-scale monthly records spanning several countries^{12,13}.

This study addresses that gap. We build a unified, reproducible pipeline to benchmark four widely used models for monthly hydrometeorological forecasting—SARIMA, TDNN, LSTM, and XGBoost—using data from 1901 to 2023 for seven South Asian countries. We standardize preprocessing, apply early stopping, and use cross-validated evaluation with RMSE, MAPE, R^2 , and CV-RMSE to ensure fair comparison. We analyze dominance patterns by variable and location and translate the findings into a practical model-selection guide. We also generate 2024–2025 forecasts with uncertainty bands to inform crop calendars, irrigation scheduling, flood preparedness, heat warnings, and water allocation.

Our contribution is novel relative to studies that apply SARIMA alone in South Asia or focus on XGBoost or deep models in single-country or short-horizon settings. We combine classical, deep, and boosting approaches within one framework and evaluate them on century-long, multi-country monthly data with documented and rigorous procedures^{1,6,10,12,13}. This positioning clarifies the study's place in the current literature and facilitates uptake by stakeholders.

In spite of the growth in the methodologies employed in modeling, only a few studies have attempted to conduct comparative studies across diverse model families on centennial-scale datasets. Previous studies have been constrained by short temporal dimensions, narrow geographic expanse, or single model frameworks. The gap is addressed by this study through the use of SARIMA, TDNN, LSTM, and XGBoost models on a dataset running for 122 years (1901–2023) of seven South Asian nations. The variables to be used include monthly rainfall and minimum temperature and maximum temperature, treated for skewness, nonstationarity and multicollinearity. High resolution and extensive temporal coverage of the dataset allow model training, testing, and validation under real-world climatic variability. The major outcome contributions of this research are threefold.

- Develop a unified forecasting framework using both classical statistical and modern machine learning approaches.
- Evaluate the predictive accuracy of each model across diverse variables using performance indicators such as RMSE, MAPE, and R^2 .
- Identify the most suitable model for each climatic variable, offering a robust basis for regional forecasting applications.

By systematically benchmarking these models, this study advances methodological rigor in hydrometeorological forecasting and contributes to the broader discourse on climate-informed decision-making.

Literature review

Forecasting rainfall and temperature has long relied on statistical time series models such as ARIMA, SARIMA, ARFIMA, and GARCH. These models perform well when data exhibit linearity and seasonality, but their limitations emerge in handling nonlinear dynamics and abrupt shifts. For instance, Ray⁶ applied SARIMA to monthly rainfall and temperature in South Asia, achieving strong results for stable seasonal series. Similarly, Yavuz⁴ demonstrated the long-term utility of ARIMA and SARIMA for rainfall and temperature in Türkiye. However, such models often underperform when climatic variability is nonlinear or extreme. To overcome these challenges, researchers have increasingly employed machine learning (ML) methods. Approaches such as Ridge regression, LASSO, Random Forest, Decision Trees, Extra Trees, SVR, and boosting algorithms (XGBoost, LightGBM) provide flexibility in capturing nonlinear and interaction effects. Ridwan et al. (2021) applied ML techniques for rainfall forecasting in Malaysia, showing improvements over traditional methods. Osman¹¹ confirmed the power of XGBoost in predicting groundwater levels, emphasizing its strength in handling noise and nonlinearity. Likewise, Papacharalampous and Tyrakis (2022) highlighted the role of large-scale time series features in improving hydrometeorological predictions in ungauged locations.

The evolution of deep learning (DL) models has further advanced climate prediction. Neural networks such as ANN and MLP have been widely used, while more recent architectures capture temporal dependencies effectively. LSTM and GRU networks excel at modeling long-term memory, with Chhetri²¹ showing their success in rainfall prediction in Bhutan. Bi-directional extensions (BiLSTM, BiGRU) and hybrid recurrent structures enhance accuracy by leveraging forward and backward dependencies²². CNN and CNN-LSTM architectures combine feature extraction with temporal modeling, as demonstrated by Wu¹⁵ in wind power forecasting and Wang² for maize crop microclimate forecasting. Transformer-based architectures are also emerging as competitive alternatives, especially for long-sequence learning²³.

Another major strand of research explores hybrid and ensemble approaches. These include combinations of statistical and ML models, statistical and DL models, and hybrid ML–DL frameworks. Mohamadi¹⁹ emphasized the use of multiple inclusive models with climate indices for rainfall prediction, achieving robust performance. Yao⁵ introduced trend structure models to extract hydrological signals from time series, supporting hybrid strategies. Decomposition–ensemble hybrids, such as CNN-LSTM or BLSTM-GRU²², integrate decomposition methods with advanced learners to improve robustness and adaptability.

Beyond climate forecasting, recent research in economics and epidemiology further validates these methodological advances. Alakkari²⁴ employed a Bayesian Mixed Frequency VAR to nowcast GDP in Syria, demonstrating how parsimonious statistical frameworks enhance predictive accuracy in data-scarce environments. Similarly, Alakkari²⁵ applied the ETS State Space model to forecast COVID-19 trends in India, showing the adaptability of statistical models in handling shocks and uncertainty. More recently, Alakkari²⁶ explored advanced machine learning with regularization techniques for GDP nowcasting in Syria, confirming the strength of ML in managing high-dimensional and noisy data. These works highlight the growing convergence of statistical rigor with ML flexibility across diverse fields, reinforcing the relevance of hybrid and adaptive approaches for complex forecasting problems. The literature demonstrates a clear methodological evolution. While statistical models remain effective for stable, seasonal data, ML and DL approaches better capture nonlinearity and extremes. Hybrid frameworks combine these strengths, offering adaptive and resilient solutions. Yet, most prior studies focus on single countries, shorter time spans, or limited models^{16,27}. Comprehensive, multi-country, century-scale comparisons remain scarce, underscoring the novelty of this study.

Materials and methodology

Data information and the study area

This study utilized a comprehensive monthly dataset spanning from January 1901 to December 2023, which includes rainfall, minimum temperature, and maximum temperature indicators across seven South Asian countries. The data was collected from World Bank Group, Climate change knowledge portal.

Prediction methods with computational packages

The dataset was processed in Python using the Jupyter environment, where all analysis and modeling were executed. The key Python libraries employed included pandas for data handling, numpy for numerical operations, matplotlib and seaborn for visualization, and scikit-learn for preprocessing and model evaluation. Additionally, time series-specific modeling was implemented using statsmodels, tensorflow.keras, and xgboost. Data preprocessing involved converting the 'Date' column to datetime format and setting it as the index. Each variable was normalized using MinMaxScaler to facilitate neural network training. For the TDNN model, a custom function generated input-output pairs using a sliding window approach with 12 input lags and 24 forecast steps. The Time Delay Neural Network (TDNN) was constructed using tensorflow.keras.Sequential, comprising two dense layers with 64 ReLU-activated units each, followed by a linear output layer. Training was done using the Adam optimizer with a learning rate of 0.001, a batch size of 16, and 100 epochs. Early stopping was simulated through validation loss monitoring. Each variable was modeled separately, and model performance was tracked using mean squared error (MSE) and mean absolute error (MAE).

For comparison, the Seasonal Autoregressive Integrated Moving Average (SARIMA) model was also applied. Each series was split into 80% training and 20% testing subsets. The SARIMA(1,1,1)(1,1,1)¹² configuration was used across variables, optimized using maximum likelihood estimation via the SARIMAX class from statsmodels. Cross-validation was performed using TimeSeriesSplit with five folds to compute average RMSE, ensuring robustness.

Additionally, Long Short-Term Memory (LSTM) models were tested using a sequence-to-sequence architecture. Inputs were reshaped to three-dimensional arrays suitable for LSTM input layers. The architecture consisted of a single LSTM layer with 64 units followed by a dense layer with 24 outputs. All models were compiled using MSE loss and the Adam optimizer.

Finally, gradient-boosted models were implemented using the xgboost library with the reg: squarederror objective. Feature engineering involved generating lag features, and each of the 24 forecast steps was modeled independently. Subsampling and column sampling were applied to prevent overfitting, and early stopping was simulated with a patience of 10 rounds.

The methodology ensures consistency across model types, facilitating direct performance comparisons using RMSE, MSE, MAPE, R², and cross-validation RMSE, which were computed for each variable and model type.

Models

SARIMA model

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model is an extension of the ARIMA model that explicitly supports univariate time series data exhibiting both non-seasonal and seasonal behaviors. The general notation of a SARIMA model is SARIMA(p, d, q)(P, D, Q)_s, where (p, d, q) denote the orders of the non-seasonal autoregressive, differencing, and moving average components respectively, and (P, D, Q) represent the seasonal counterparts with periodicity *s*. In this context, the SARIMA(1,1,1)(1,1,1)₁₂ specification captures a series with both first-order non-seasonal and seasonal differencing, as well as first-order non-seasonal and seasonal autoregressive and moving average terms, with seasonality of 12 (monthly data). Mathematically, the SARIMA(1,1,1)(1,1,1)₁₂ model is expressed as^{1,6}:

$$\Phi_1(B^{12})\varphi_1(B)\nabla\nabla_{12}y_t = \Theta_1(B^{12})\theta_1(B)\epsilon_t \quad (1)$$

where:

- $\varphi_1(B) = 1 - \varphi_1 B$ is the non-seasonal AR term.
- $\theta_1(B) = 1 + \theta_1 B$ is the non-seasonal MA term.
- $\Phi_1(B^{12}) = 1 - \Phi_1 B^{12}$ is the seasonal AR term.
- $\Theta_1(B^{12}) = 1 + \Theta_1 B^{12}$ is the seasonal MA term.

- $\nabla = 1 - B$ is the non-seasonal difference operator.
- $\nabla_{12} = 1 - B^{12}$ is the seasonal difference operator.
- B is the backward shift operator such that $B^k y_t = y_{t-k}$.
- ϵ_t is a white noise error term.

The combined differencing operator

$$\nabla \nabla_{12} y_t = (1 - B) (1 - B^{12}) y_t \quad (2)$$

accounts for both trend and seasonal stationarity. This model structure is repeatedly used for the data sequences provided, reflecting an assumption of consistent seasonal and trend characteristics across observations. Such uniformity suggests the time series under study likely exhibit stable periodic behavior with a need for both regular and seasonal differencing and memory, addressed through first-order AR and MA terms.

We determined $(p, d, q)(P, D, Q)$ s by:

1. Differencing guided by ADF and KPSS.
2. Visual inspection of ACF and PACF up to 24 lags.
3. Grid search p, q, P, Q in 0–2 and D in 0–1 with $s = 12$.
4. AICc for in-sample fit.
5. Ljung–Box on residuals and ARCH-LM for conditional heteroskedasticity.
6. Out-of-sample CV-RMSE using the same folds as other models.

This process selected orders that largely align with seasonal ARMA structure at lag 12. We list selected orders and diagnostics per series.

TDNN model

The Time Delay Neural Network (TDNN) is a type of feedforward neural network designed to model temporal dependencies in sequential data by incorporating time delays into the architecture. In the context of forecasting, the TDNN receives a fixed-size window of past observations as input and predicts future values based on learned nonlinear transformations. The proposed TDNN architecture in this study follows a structure of Input $\rightarrow$ Dense(64, ReLU) $\rightarrow$ Dense(64, ReLU) $\rightarrow$ Dense(n_{forecast}), where each dense layer is fully connected. The hidden layers use the Rectified Linear Unit (ReLU) activation function, defined as^{16,28}:

$$\text{ReLU}(x) = \max(0, x) \quad (3)$$

which introduces nonlinearity and mitigates vanishing gradient issues. The output layer employs a linear activation function, suitable for regression tasks. The model parameters are initialized using He Normal initialization, suitable for ReLU-activated layers, where weights W are sampled from a normal distribution:

$$W \sim \mathcal{N}\left(0, \frac{2}{n_{\text{in}}}\right) \quad (4)$$

Where n_{in} is the number of input units in the weight tensor. Training is optimized using the Adam optimizer, which combines the advantages of RMSProp and momentum. The update rules for Adam are given by:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \nabla L_t \quad (5)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) (\nabla L_t)^2 \quad (6)$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (7)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (8)$$

$$\theta_t = \theta_{t-1} - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \quad (9)$$

With $\beta_1 = 0.9$, $\beta_2 = 0.999$ and ϵ a small constant to prevent division by zero. To prevent overfitting, early stopping is used with a simulated patience of 10 epochs, halting training if validation loss does not improve. This TDNN configuration captures short-term temporal dynamics while remaining computationally efficient, making it suitable for time series forecasting tasks such as those analyzed in this study.

LSTM model

The Long Short-Term Memory (LSTM) network is a type of Recurrent Neural Network (RNN) designed to model sequential data while addressing the vanishing gradient problem commonly encountered in traditional RNNs. The integration of Seq2Seq and Attention mechanisms enhances the LSTM's capability to capture long-range dependencies and improve its performance in time series forecasting tasks^{15,29}. LSTMs are defined by their

unique architecture, which includes gates that regulate the flow of information. These gates are mathematically formulated as follows:

$$f_t = \sigma \left(W_f \cdot [h_{t-1} - 1, x_t] + b_f \right) \tag{10}$$

Where: f_t Forget gate vector at time t , x_t Input at time t , h_{t-1} Hidden state from the previous time step, $W_f b_f$ Learnable weights and biases for the forget gate, σ Sigmoid activation function. And Input Gate:

$$\begin{aligned} i_t &= \sigma \left(W_i \cdot [h_{t-1}, x_t] + b_i \right) \\ \tilde{C}_t &= \tanh \left(W_C \cdot [h_{t-1}, x_t] + b_C \right) \end{aligned} \tag{11}$$

Where i_t Input gate vector, $\tilde{C}_t$ Candidate cell state. And Cell State Update:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{12}$$

Where C_t Cell state at time t , $\odot$ Element-wise multiplication. And Output Gate:

$$\begin{aligned} o_t &= \sigma \left(W_o \cdot [h_{t-1}, x_t] + b_o \right) \\ h_t &= o_t \odot \tanh (C_t) \end{aligned} \tag{13}$$

Where o_t Output gate vector, h_t Hidden state.

Sequence-to-Sequence (Seq2Seq) Model: Encoder: Encodes the input sequence into a fixed-size context vector, z , which summarizes the sequence: $z = h_T$ Where h_T is the hidden state at the final time step T , Decoder: Decodes the context vector z to generate the output sequence (Wang et al., 2022):

$$y_t = f(h_{t-1}, z) \tag{14}$$

XGBOOST model

The Extreme Gradient Boosting (XGBoost) model used in this study is configured for regression tasks with the objective function set to reg: squarederror, which minimizes the squared loss between predicted and actual values^{11,13}:

$$L = \sum_{i=1}^n \left(y_i - \hat{y}_i \right)^2 \tag{15}$$

The model employs the gradient boosted decision tree (GBTree) booster, which builds an ensemble of trees iteratively. At each boosting round, a new tree is trained to minimize the regularized objective:

$$\mathcal{L}^{(t)} = \sum_{i=1}^n l \left(y_i, y_i^{(t-1)} + f_t(x_i) \right) + \Omega(f_t) \tag{16}$$

Where $f \in F$ is the function (tree) added at iteration t , and:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \tag{17}$$

is the regularization term with T as the number of leaves and w_j as the leaf weights. To reduce overfitting and enhance generalization, subsampling of rows (subsample=0.8) and features (colsample_bytree=0.8) are applied during training. Early stopping is simulated after 10 rounds without validation improvement, training 24 independent models, each predicting one of the 24 future time steps. Feature importance is assessed using the mean gain across all trees, then normalized so the importance values for each variable sum to 1.0. This setup ensures that the forecasting framework maintains interpretability and robustness across all predicted horizons.

Code	AFGR	BGDR	BTNR	INDR	LKAR	NPLR	PAKR
Name	Afghanistan Rainfall	Bangladesh Rainfall	Bhutan Rainfall	India Rainfall	Sri Lanka Rainfall	Nepal Rainfall	Pakistan Rainfall
Code	AFGM	BGDM	BTNM	INDM	LKAM	NPLM	PAKM
Name	Afghanistan Temperature MIN	Bangladesh Temperature MIN	Bhutan Temperature MIN	India Temperature MIN	Sri Lanka Temperature MIN	Nepal Temperature MIN	Pakistan Temperature MIN
Code	AFGX	BGDX	BTNX	INDX	LKAX	NPLX	PAKX
Name	Afghanistan Temperature MAX	Bangladesh Temperature MAX	Bhutan Temperature MAX	India Temperature MAX	Sri Lanka Temperature MAX	Nepal Temperature MAX	Pakistan Temperature MAX

Table 1. Monthly rainfall and temperature extremes (122-Year analysis 1901–2023).

	Count	Mean	std	Min	25%	50%	75%	Max
AFGR	1476	27.20343	23.44207	0.45	6.0425	19.86	45.44	107.22
BGDR	1476	194.4943	189.8076	0.02	19.55	132.905	336.38	1040.57
BTNR	1476	162.9935	173.8123	0.19	12.1625	87.995	291.605	716.36
INDR	1476	93.73365	99.25364	1.4	16.2375	43.625	161.54	365.06
LKAR	1476	144.6037	107.3074	0.56	62.435	115.515	199.455	593.76
NPLR	1476	111.4277	119.0524	0.03	19.5275	51.305	192.985	533.12
PAKR	1476	23.60354	20.15248	0.37	10.1	18.225	30.5725	163.72
AFGM	1476	5.285779	7.903499	−12.11	−2.0025	5.64	12.815	19.11
BGDM	1476	20.71892	4.95657	10.13	16.2575	22.97	25.2025	26.56
BTNM	1476	4.202805	6.166659	−7.05	−1.4325	4.995	10.46	12.83
INDM	1476	18.51981	4.997772	8.96	13.5475	19.91	23.01	25.98
LKAM	1476	23.05919	1.173935	19.85	22.1275	23.2	23.98	25.96
NPLM	1476	7.311484	6.678915	−4.42	0.89	8.16	13.9425	17.1
PAKM	1476	13.81137	7.756077	−0.43	6.4225	14.35	21.76	24.95
AFGX	1476	20.10687	9.644672	−0.01	11.2875	20.98	29.41	35.34
BGDX	1476	30.23251	2.674862	22.83	28.2825	31.115	32.02	35.87
BTNX	1476	15.72102	3.9071	7.22	12.255	16.705	19.26	21.98
INDX	1476	30.5383	3.796814	23.11	27.55	30.76	33.29	38.28
LKAX	1476	30.33706	1.148364	27.18	29.53	30.5	31.14	33.42
NPLX	1476	20.02411	4.704896	10.19	15.775	21.53	24.02	27.44
PAKX	1476	27.9691	7.087438	13.41	21.73	29.825	34.345	38.15

Table 2. Descriptive statistics of monthly rainfall and temperature Extremes.

Hyperparameter optimization and validation

We tuned TDNN, LSTM, and XGBOOST by randomized search with 50 trials per series. We used a 5-fold expanding TimeSeriesSplit on the training set. The objective metric was mean CV-RMSE. Early stopping monitored validation RMSE within each trial.

TDNN		LSTM		XGBOOST	
Hyperparameter	Value Range	Hyperparameter	Value Range	Hyperparameter	Value Range
Lookback window	6–24	Lookback window	12–36	n_estimators	100–800
Hidden layers	1–3	Layers	1–2	learning_rate	0.01–0.3
Units per layer	32–128	Units per layer	32–128	max_depth	2–8
Dropout	0.0–0.3	Dropout	0.0–0.4	subsample	0.5–1.0
Batch size	16–64	Recurrent dropout	0.0–0.2	colsample_bytree	0.5–1.0
Learning rate	1e-4–5e-3	Batch size	16–64	min_child_weight	1–10
Activation	ReLU	Learning rate	1e-4–5e-3	lambda	0–5
Optimizer	Adam	Optimizer	Adam	alpha	0–1
		Seq2Seq decoder	dense	objective reg	squarederror

Table A. Search spaces.

We fixed patience at 10 epochs or 10 boosting rounds for early stopping. We set a global random seed and report final chosen hyperparameters per series in Table S1. We trained the final model on the full training set with the best trial settings.

Evaluation metrics

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$
 (18)

where y_t is the actual value, $\hat{y}_t$ is the predicted value, and n is the number of observations. The MAPE, which expresses the average percentage error, is given by:

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|$$
 (19)

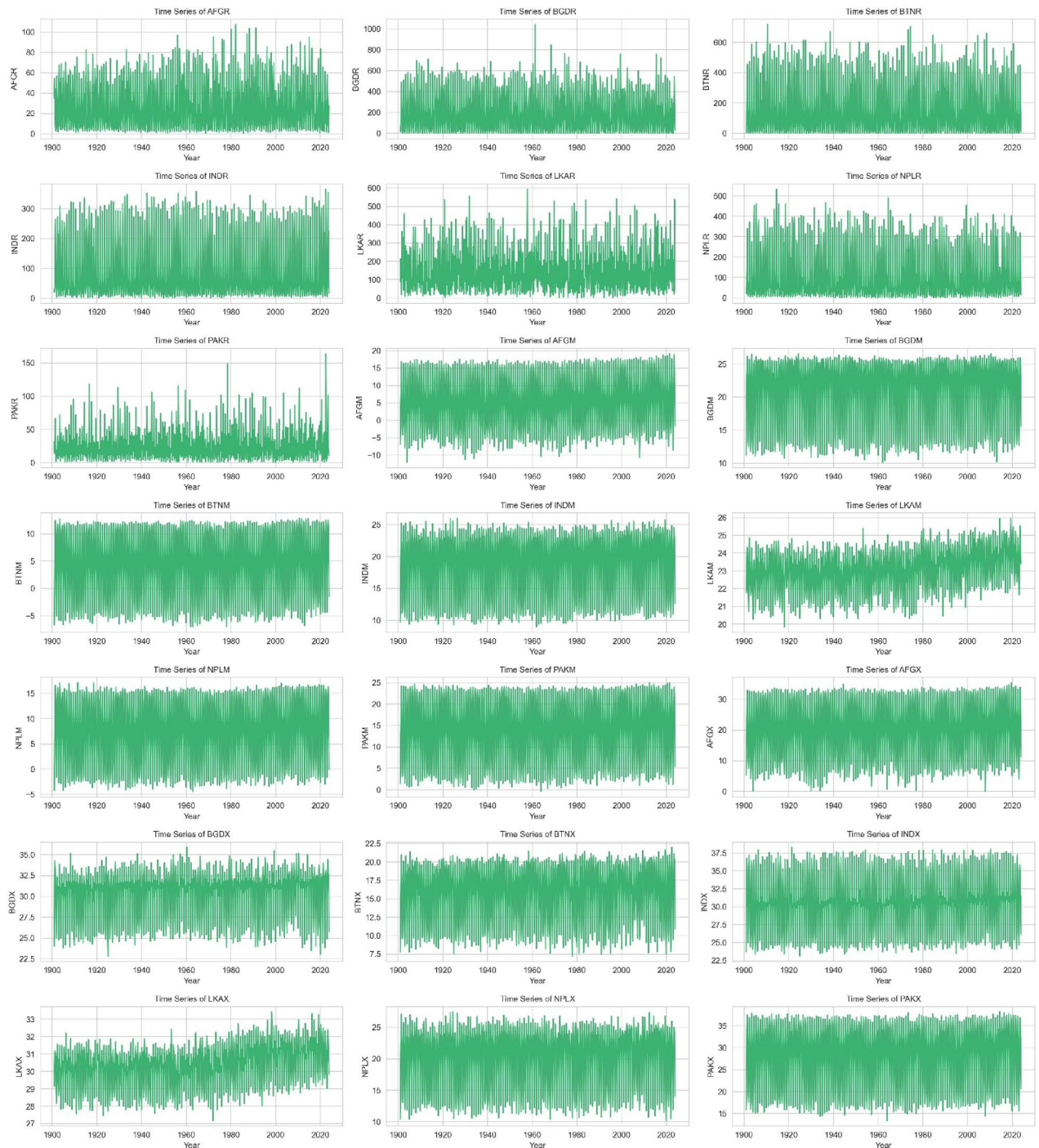


Fig. 1. Time series visualization of monthly rainfall and temperature indicators across 7 regions from 1901 to 2023. Each subplot presents the temporal variability of a specific climate variable, capturing historical fluctuations and seasonal patterns essential for model training and forecasting evaluation.

The MSE, representing the average squared differences, is computed as:

$$MSE = \frac{1}{n} \sum^n (y_t - \hat{y}_t)^2 \quad (20)$$

the coefficient of determination (R^2) quantifies the proportion of variance in the dependent variable that is predictable from the independent variables, calculated as:

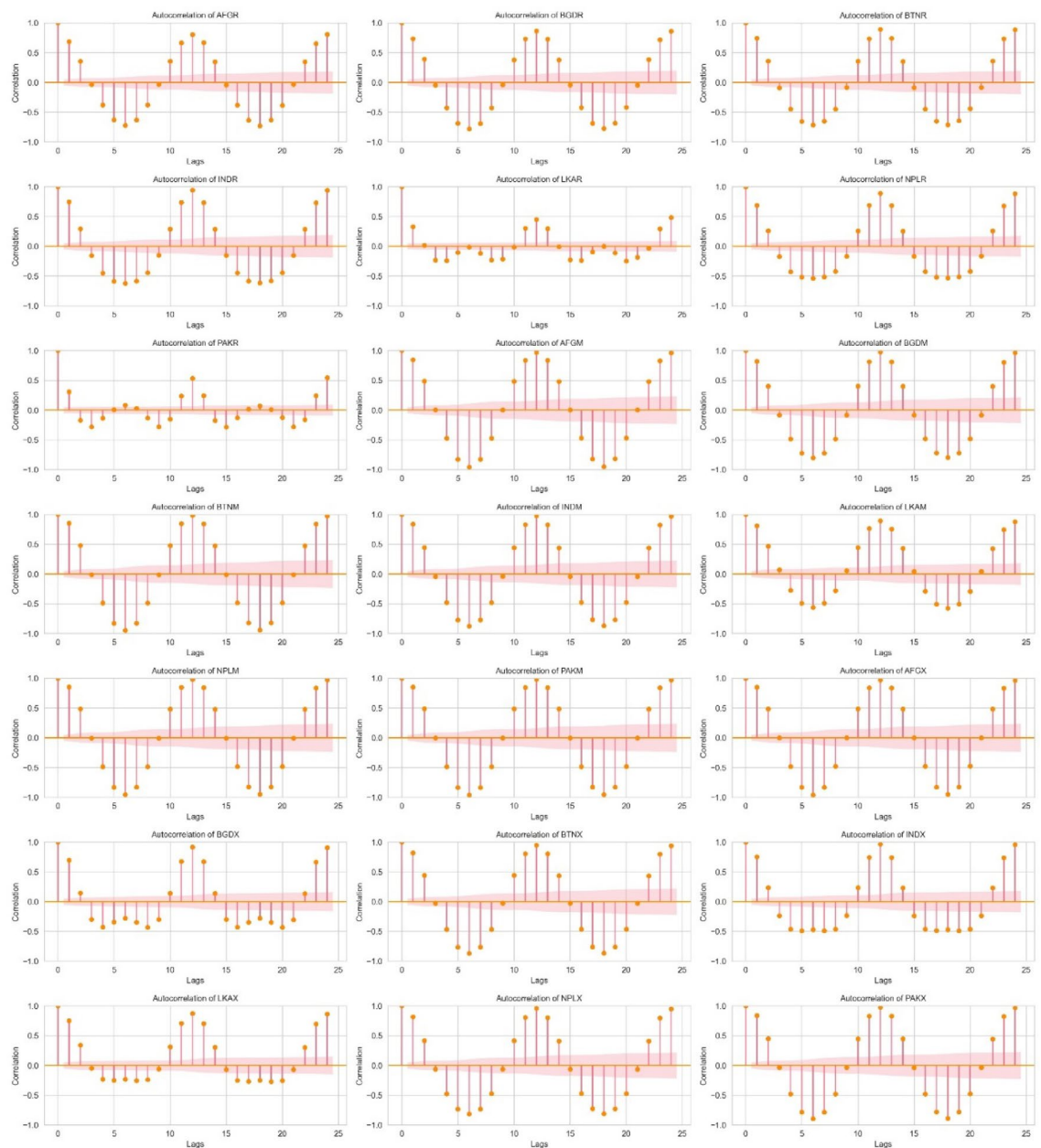


Fig. 2. Autocorrelation plots of monthly rainfall and temperature indicators across 21 regions from 1901 to 2023. Each subplot shows lag correlations, identifying seasonality, persistence, and temporal dependencies critical for time series model selection.

$$R^2 = 1 - \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\sum_{t=1}^n (y_t - \bar{y}_t)^2} \quad (21)$$

Where $\bar{y}_t$ is the mean of the actual values. These performance indicators were used to compare the models across different countries and Variables.

Results and discission

The dataset consists of 21 monthly time series variables from January 1901 to December 2023, with 1,476 observations each. We rename all the variable along with the countries in code name, provided in Table 1. Table 2 reports descriptive statistics for rainfall, minimum temperature, and maximum temperature across seven South Asian countries. The rainfall series—Afghanistan_Rainfall (AFGR), Bangladesh_Rainfall (BGDR), Bhutan_Rainfall (BTNR), India_Rainfall (INDR), SriLanka_Rainfall (LKAR), Nepal_Rainfall (NPLR),

Variable	Model	RMSE_Train	RMSE_Test	MSE_Train	MSE_Test	MAPE_Train	MAPE_Test	R2_Train	R2_Test	CV_RMSE
AFGR	SARIMA	11.05	11.05	122.08	122.08	0.46	0.46	0.75	0.75	11.61
AFGR	TDNN	10.73	10.93	115.06	119.41	0.51	0.46	0.80	0.76	11.16
AFGR	LSTM	0.10	0.10	0.01	0.01	14.62	0.49	0.80	0.75	11.16
AFGR	XGBOOST	0.07	0.11	0.01	0.01	12.91	0.45	0.90	0.74	11.16
BGDR	SARIMA	67.50	67.50	4556.07	4556.07	1.89	1.89	0.85	0.85	71.10
BGDR	TDNN	70.91	69.30	5028.31	4802.26	2.69	2.50	0.87	0.84	75.29
BGDR	LSTM	0.07	0.07	0.01	0.00	23.31	89.13	0.86	0.83	75.29
BGDR	XGBOOST	0.05	0.07	0.00	0.01	19.15	86.74	0.93	0.82	75.29
BTNR	SARIMA	60.54	60.54	3664.92	3664.92	0.62	0.62	0.87	0.87	52.76
BTNR	TDNN	55.05	60.65	3030.30	3677.97	1.26	1.99	0.90	0.87	56.61
BTNR	LSTM	0.08	0.09	0.01	0.01	46.73	2.32	0.90	0.86	56.61
BTNR	XGBOOST	0.05	0.09	0.00	0.01	45.27	2.11	0.95	0.84	56.61
INDR	SARIMA	21.63	21.63	468.04	468.04	0.28	0.28	0.95	0.95	23.00
INDR	TDNN	22.23	21.59	494.11	466.03	0.35	0.27	0.95	0.95	24.69
INDR	LSTM	0.07	0.06	0.00	0.00	12.61	0.42	0.94	0.95	24.69
INDR	XGBOOST	0.05	0.06	0.00	0.00	12.36	0.36	0.97	0.95	24.69
LKAR	SARIMA	77.80	77.80	6053.16	6053.16	0.75	0.75	0.50	0.50	80.38
LKAR	TDNN	80.23	85.42	6436.76	7296.11	1.11	0.77	0.43	0.41	91.34
LKAR	LSTM	0.14	0.14	0.02	0.02	50.18	0.86	0.41	0.37	91.34
LKAR	XGBOOST	0.11	0.15	0.01	0.02	39.08	0.89	0.64	0.34	91.34
NPLR	SARIMA	41.38	41.38	1712.29	1712.29	0.35	0.35	0.87	0.87	41.92
NPLR	TDNN	41.89	41.75	1754.89	1743.07	1.77	0.49	0.88	0.86	44.60
NPLR	LSTM	0.08	0.09	0.01	0.01	17.65	0.52	0.88	0.84	44.60
NPLR	XGBOOST	0.06	0.09	0.00	0.01	20.29	0.57	0.94	0.82	44.60
PAKR	SARIMA	15.79	15.79	249.18	249.18	0.94	0.94	0.47	0.47	13.11
PAKR	TDNN	12.96	16.67	167.95	277.91	0.79	0.81	0.57	0.43	14.51
PAKR	LSTM	0.09	0.10	0.01	0.01	12.11	38.84	0.50	0.40	14.51
PAKR	XGBOOST	0.06	0.11	0.00	0.01	10.98	45.50	0.74	0.32	14.51
AFGM	SARIMA	1.30	1.30	1.69	1.69	5.13×10^{12}	5.13×10^{12}	0.97	0.97	1.27
AFGM	TDNN	1.36	1.57	1.85	2.47	0.58	6.01×10^{12}	0.97	0.96	1.36
AFGM	LSTM	0.04	0.06	0.00	0.00	74.03	0.11	0.97	0.95	1.36
AFGM	XGBOOST	0.03	0.06	0.00	0.00	5.747	0.11	0.98	0.95	1.36
BGDM	SARIMA	0.66	0.66	0.43	0.43	0.03	0.03	0.98	0.98	0.69
BGDM	TDNN	0.75	0.69	0.56	0.47	0.03	0.03	0.98	0.98	0.77
BGDM	LSTM	0.05	0.05	0.00	0.00	39.90	0.15	0.98	0.97	0.77
BGDM	XGBOOST	0.03	0.05	0.00	0.00	33.02	0.16	0.99	0.97	0.77
BTNM	SARIMA	0.64	0.64	0.41	0.41	0.29	0.29	0.99	0.99	0.61
BTNM	TDNN	0.72	0.73	0.51	0.53	4.92	0.44	0.99	0.99	0.68
BTNM	LSTM	0.04	0.05	0.00	0.00	31.463	0.12	0.99	0.98	0.68
BTNM	XGBOOST	0.03	0.04	0.00	0.00	23.512	0.12	0.99	0.98	0.68
INDM	SARIMA	0.60	0.60	0.36	0.36	0.03	0.03	0.98	0.98	0.62
INDM	TDNN	0.65	0.73	0.42	0.53	0.03	0.03	0.98	0.98	0.66
INDM	LSTM	0.04	0.06	0.00	0.00	32.41	0.12	0.98	0.96	0.66
INDM	XGBOOST	0.03	0.05	0.00	0.00	29.398	0.10	0.99	0.97	0.66
LKAM	SARIMA	0.44	0.44	0.20	0.20	0.01	0.01	0.82	0.82	0.51
LKAM	TDNN	0.38	0.47	0.15	0.22	0.01	0.02	0.89	0.80	0.39
LKAM	LSTM	0.08	0.15	0.01	0.02	865.17	0.23	0.83	0.20	0.39
LKAM	XGBOOST	0.05	0.10	0.00	0.01	60.06	0.13	0.92	0.70	0.39
NPLM	SARIMA	0.77	0.77	0.59	0.59	0.94	0.94	0.99	0.99	0.77
NPLM	TDNN	0.85	0.84	0.72	0.71	1.15×10^{13}	1.05	0.98	0.98	0.84
NPLM	LSTM	0.04	0.05	0.00	0.00	26.922	0.12	0.98	0.97	0.84
NPLM	XGBOOST	0.03	0.06	0.00	0.00	21.57	0.13	0.99	0.97	0.84
PAKM	SARIMA	1.21	1.21	1.47	1.47	0.10	0.10	0.97	0.97	0.91
PAKM	TDNN	1.04	1.20	1.08	1.44	0.30	0.11	0.98	0.98	0.99
PAKM	LSTM	0.04	0.06	0.00	0.00	36.782	0.12	0.98	0.96	0.99
PAKM	XGBOOST	0.03	0.06	0.00	0.00	25.08	0.12	0.99	0.95	0.99
Continued										

Variable	Model	RMSE_Train	RMSE_Test	MSE_Train	MSE_Test	MAPE_Train	MAPE_Test	R2_Train	R2_Test	CV_RMSE
AFGX	SARIMA	1.57	1.57	2.47	2.47	2.04	2.04	0.97	0.97	1.80
AFGX	TDNN	1.56	1.61	2.44	2.59	0.23	2.07	0.97	0.97	1.59
AFGX	LSTM	0.05	0.05	0.00	0.00	0.23	23.96	0.97	0.96	1.59
AFGX	XGBOOST	0.03	0.05	0.00	0.00	0.17	24.23	0.98	0.96	1.59
BGDX	SARIMA	0.84	0.84	0.71	0.71	0.02	0.02	0.90	0.90	0.82
BGDX	TDNN	0.75	0.85	0.56	0.73	0.02	0.02	0.92	0.90	0.86
BGDX	LSTM	0.06	0.08	0.00	0.01	68.07	0.17	0.90	0.86	0.86
BGDX	XGBOOST	0.05	0.07	0.00	0.01	48.80	0.18	0.95	0.87	0.86
BTNX	SARIMA	0.97	0.97	0.93	0.93	0.05	0.05	0.94	0.94	0.85
BTNX	TDNN	0.89	1.02	0.80	1.04	0.05	0.06	0.95	0.93	0.94
BTNX	LSTM	0.06	0.08	0.00	0.01	64.24	0.19	0.95	0.91	0.94
BTNX	XGBOOST	0.05	0.08	0.00	0.01	49.83	0.19	0.97	0.91	0.94
INDX	SARIMA	0.71	0.71	0.51	0.51	0.02	0.02	0.96	0.96	0.66
INDX	TDNN	0.73	0.68	0.54	0.47	0.02	0.02	0.96	0.97	0.73
INDX	LSTM	0.05	0.07	0.00	0.00	36.503	0.13	0.96	0.93	0.73
INDX	XGBOOST	0.04	0.06	0.00	0.00	30.98	0.13	0.98	0.94	0.73
LKAX	SARIMA	0.66	0.66	0.44	0.44	0.02	0.02	0.59	0.59	0.54
LKAX	TDNN	0.40	0.39	0.16	0.16	0.01	0.01	0.88	0.85	0.42
LKAX	LSTM	0.08	0.15	0.01	0.02	56.34	0.22	0.79	0.17	0.42
LKAX	XGBOOST	0.06	0.08	0.00	0.01	56.55	0.12	0.90	0.75	0.42
NPLX	SARIMA	1.05	1.05	1.10	1.10	0.04	0.04	0.95	0.95	0.88
NPLX	TDNN	0.87	1.03	0.76	1.06	0.04	0.04	0.97	0.95	0.97
NPLX	LSTM	0.06	0.07	0.00	0.01	0.16	16.62	0.96	0.92	0.97
NPLX	XGBOOST	0.04	0.07	0.00	0.00	0.12	16.68	0.98	0.93	0.97
PAKX	SARIMA	1.33	1.33	1.76	1.76	0.04	0.04	0.96	0.96	0.93
PAKX	TDNN	1.08	1.27	1.16	1.61	0.03	0.04	0.98	0.97	1.05
PAKX	LSTM	0.04	0.05	0.00	0.00	49.46	0.10	0.98	0.96	1.05
PAKX	XGBOOST	0.03	0.06	0.00	0.00	38.32	0.12	0.99	0.95	1.05

Table 3. Model performance comparison (RMSE, MSE, MAPE, R², CV_RMSE).

Pakistan_Rainfall (PAKR)—show wide variation. Bangladesh_Rainfall recorded the highest mean (194.49 mm), while Afghanistan_Rainfall had the lowest (27.20 mm). Extreme values ranged from 0.02 mm to 1040.57 mm, indicating heavy rainfall events and skewed distributions. Minimum temperature variables—Afghanistan Temp MIN (AFGM), Bangladesh Temp_MIN (BGDM), Bhutan_Temp_MIN (BTNM), India_Temp_MIN (INDM), SriLanka_Temp_MIN (LKAM), Nepal_Temp_MIN (NPLM), Pakistan_Temp_MIN (PAKM)—ranged from −12.11 °C to 26.56 °C, with Afghanistan_Temp_MIN showing the coldest mean (5.29 °C) and SriLanka_Temp_MIN the most stable (std=1.17). Maximum temperature series—Afghanistan_Temp_MAX (AFGX), Bangladesh_Temp_MAX (BGDX), Bhutan_Temp_MAX (BTNX), India_Temp_MAX (INDX), SriLanka_Temp_MAX (LKAX), Nepal_Temp_MAX (NPLX), Pakistan_Temp_MAX (PAKX)—had values ranging from −0.01 °C to 38.28 °C, with India_Temp_MAX reporting the highest mean (30.54 °C). Rainfall data showed high variability and skewness, requiring normalization and seasonality-aware modeling. In contrast, temperature data, especially in Sri Lanka, appeared more stable, justifying the use of statistical models that assume regularity in variance and trend.

Figure 1 presents the monthly time series plots of all variables from 1901 to 2023. Rainfall series (AFGR, BGDR, BTNR, INDR, LKAR, NPLR, PAKR) exhibit clear seasonal patterns and high variability, especially during monsoon months, with no evident long-term trend. Also, it highlighted the rainfall clear 12-month seasonality with monsoon peaks in BGDR, BTNR, INDR higher variance in BGDR vs. LKAR Temperature LKAM shows the narrowest annual amplitude AFGM and NPLM show deeper winters. Minimum temperature series (AFGM, BGDM, BTNM, INDM, LKAM, NPLM, PAKM) show strong seasonality with varying amplitude across countries, reflecting climatic differences. Maximum temperature series (AFGX, BGDX, BTNX, INDX, LKAX, NPLX, PAKX) demonstrate more stability and smoother annual cycles. The visual patterns across all series confirm the presence of seasonality and temporal structure, supporting the use of SARIMA, TDNN, and LSTM models for forecasting.

Figure 2 presents the autocorrelation function (ACF) plots for all 21 variables, examining the correlation between current and lagged values up to 24 months. The rainfall series (AFGR, BGDR, BTNR, INDR, LKAR, NPLR, PAKR) show strong and significant autocorrelations at seasonal lags (e.g., 12, 24), confirming the presence of annual seasonality. This seasonal dependence justifies the use of SARIMA models with seasonal components and supports time-delay structures for neural networks like TDNN and LSTM. Minimum temperature variables (AFGM, BGDM, BTNM, INDM, LKAM, NPLM, PAKM) also exhibit seasonal autocorrelation patterns, though often weaker and less consistent across lags. Maximum temperature variables (AFGX, BGDX, BTNX,

Variable	Best Model	Model Type	Suitable Data Type	Key Assumption	Strength of the Model
AFGR	TDNN	Neural Network	Nonlinear and complex data	No assumption of normal distribution	Effectively handles nonlinear relationships
BGDR	SARIMA	Time Series	Data with clear seasonal trend	Stationarity and stability	High accuracy with seasonal repetition
BTNR	SARIMA	Time Series	Stable seasonal data	Moderate stability and variance	Effective for cyclical trend data
INDR	TDNN	Neural Network	Complex and nonlinear data	High flexibility with complexity	Suitable for pattern-unclear data
LKAR	SARIMA	Time Series	Regular seasonal data	Variance and trend stationarity	Good at detecting regular seasonality
NPLR	SARIMA	Time Series	Stable cyclical data	Periodic data stationarity	Handles recurring patterns well
PAKR	SARIMA	Time Series	Data with clear seasonal pattern	Seasonal stationarity assumption	Suitable for seasonal data
AFGM	LSTM	Recurrent Neural Net	Time series with long memory	Does not require stationarity	Remembers long-term dependencies
BGDM	SARIMA	Time Series	Periodic time series data	Stable and clear seasonality	Analyzes regular seasonal patterns effectively
BTNM	XGBOOST	Boosted Trees	Nonlinear, highly interactive data	Handles missing data gracefully	Captures complex relationships accurately
INDM	SARIMA	Time Series	Stable seasonal time series	Statistical data stability	Appropriate for statistically stable data
LKAM	SARIMA	Time Series	Stable seasonal data	Data trend and variance stability	Well-suited for periodic stable data
NPLM	LSTM	Recurrent Neural Net	Variable data with long time horizons	No need for stationarity	Effective for long-term forecasting
PAKM	TDNN	Neural Network	Irregular and complex data	No requirement for stationarity	Effectively handles complex relationships
AFGX	SARIMA	Time Series	Regular data with seasonal trend	Clear seasonal stability	Effective for seasonally patterned data
BGDX	SARIMA	Time Series	Stable data with clear seasonality	Moderate data stationarity	Good at capturing seasonal variation
BTNX	SARIMA	Time Series	Data with regular seasonal direction	Clear seasonal trend	Excellent for seasonal forecasting
INDX	TDNN	Neural Network	Data with high complexity and correlation	High flexibility, no need for stationarity	Efficient with highly complex data
LKAX	TDNN	Neural Network	Nonlinear and variable data	No assumption of statistical stability	Strong at modeling nonlinear interactions
NPLX	TDNN	Neural Network	Complex data with unclear patterns	Flexible with complex relationships	Effective in modeling compound patterns
PAKX	TDNN	Neural Network	Variable and complex data	No assumption of normal distribution	Strong at handling complex structures

Table 4. Best model selection by Variable.

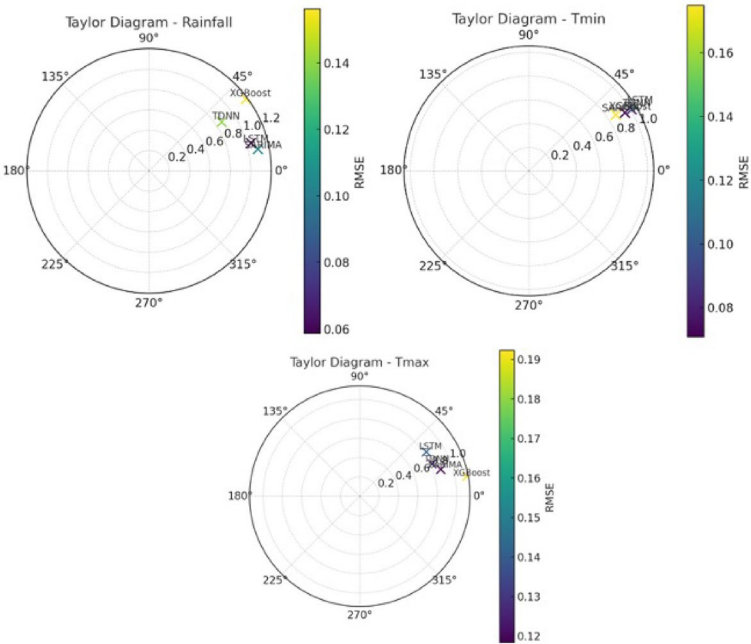


Fig. 3. Taylor diagrams.

INDX, LKAX, NPLX, PAKX) generally show smoother and more regular decay, indicating stationarity and short memory in some series (e.g., LKAX), and persistent seasonal patterns in others (e.g., INDX).Overall, the ACF patterns support model design choices. Series with clear periodic autocorrelation peaks are suitable for SARIMA. Those with complex or long-range dependencies are better modeled using LSTM or TDNN, which can learn such temporal patterns without needing explicit lag specification. Additionally, the Fig. 2 indicated dominant spikes at lag 12 confirm seasonal structure slow decay in INDX supports AR terms rainfall ACF cuts off after lag 1 supporting MA(1) these patterns guided SARIMA order selection in Table 5.

Variable	Model	Significant Coefficients ($p < 0.05$)	Quality Metrics	Diagnostics
BGDR	(1,1,1)×(1,1,1,12)	MA(1): -0.9986 ($p = 0.000$) Seasonal AR(1): -0.0860 ($p = 0.000$) Seasonal MA(1): -0.9637 ($p = 0.000$)	AIC: 13232.4 BIC: 13257.7	Residuals: LB ($p = 0.94$) Heteroskedasticity ($p = 0.40$)
BTNR	(1,1,1)×(1,1,1,12)	MA(1): -0.9997 ($p = 0.000$) Seasonal AR(1): -0.0482 ($p = 0.021$)	AIC: 12633.4 BIC: 12658.7	Residuals: LB ($p = 0.95$) Heteroskedasticity ($p = 0.17$)
LKAR	(1,1,1)×(1,1,1,12)	MA(1): -0.9987 ($p = 0.000$) Seasonal MA(1): -0.9991 ($p = 0.000$)	AIC: 13542.4 BIC: 13567.7	Residuals: LB ($p = 0.98$) Heteroskedasticity ($p = 0.25$)
NPLR	(1,1,1)×(1,1,1,12)	MA(1): -0.9984 ($p = 0.000$) Seasonal AR(1): 0.0862 ($p = 0.000$) Seasonal MA(1): -0.9980 ($p = 0.000$)	AIC: 12011.0 BIC: 12036.3	Residuals: LB ($p = 0.97$) Heteroskedasticity ($p = 0.09$)
PAKR	(1,1,1)×(1,1,1,12)	AR(1): 0.0977 ($p = 0.000$) MA(1): -0.9998 ($p = 0.000$)	AIC: 9337.0 BIC: 9362.3	Residuals: LB ($p = 0.94$) Heteroskedasticity ($p = 0.11$)
BGDM	(1,1,1)×(1,1,1,12)	AR(1): 0.1566 ($p = 0.000$) MA(1): -0.9738 ($p = 0.000$) Seasonal MA(1): -0.9624 ($p = 0.000$)	AIC: 2396.9 BIC: 2422.2	Residuals: LB ($p = 0.77$) Heteroskedasticity ($p = 0.99$)
INDM	(1,1,1)×(1,1,1,12)	AR(1): 0.2935 ($p = 0.000$) MA(1): -0.9889 ($p = 0.000$) Seasonal MA(1): -0.9652 ($p = 0.000$)	AIC: 2132.7 BIC: 2158.0	Residuals: LB ($p = 0.57$) Heteroskedasticity ($p = 0.08$)
LKAM	(1,1,1)×(1,1,1,12)	AR(1): 0.2397 ($p = 0.000$) MA(1): -0.8392 ($p = 0.000$) Seasonal MA(1): -0.9782 ($p = 0.000$)	AIC: 840.1 BIC: 865.4	Residuals: LB ($p = 0.76$) Heteroskedasticity ($p = 0.06$)
AFGX	(1,1,1)×(1,1,1,12)	AR(1): 0.2337 ($p = 0.000$) MA(1): -0.9819 ($p = 0.000$) Seasonal MA(1): -0.9968 ($p = 0.000$)	AIC: 4223.7 BIC: 4249.0	Residuals: LB ($p = 0.86$) Heteroskedasticity ($p = 0.73$)
BGDX	(1,1,1)×(1,1,1,12)	AR(1): 0.2474 ($p = 0.000$) MA(1): -0.9877 ($p = 0.000$) Seasonal MA(1): -0.9647 ($p = 0.000$)	AIC: 2576.1 BIC: 2601.4	Residuals: LB ($p = 0.67$) Heteroskedasticity ($p = 0.03$)
BTNX	(1,1,1)×(1,1,1,12)	AR(1): 0.2209 ($p = 0.000$) MA(1): -0.9879 ($p = 0.000$) Seasonal MA(1): -0.9950 ($p = 0.000$)	AIC: 2823.7 BIC: 2849.0	Residuals: LB ($p = 0.82$) Heteroskedasticity ($p = 0.00$)

Table 5. SARIMA model diagnostics (Significant Coefficients, AIC/BIC, Residuals).

LSTM(64) → Dense(1) Epochs = 100 Batch = 16 Adam(lr=0.001)					
Variable	Train Loss (MSE)	Train MAE	Val Loss (MSE)	Val MAE	Convergence Epoch
AFGM	0.0185	0.102	0.0212	0.115	78
NPLM	0.0092	0.071	0.0118	0.089	65

Table 6. LSTM training convergence (MSE, MAE, Epochs).

TDNN: Architecture: Input → Dense(64, ReLU) → Dense(64, ReLU) → Dense(n_forecast) 1. Optimizer: Adam ($\beta_1=0.9, \beta_2=0.999$) 2. Early Stopping: Patience=10 (simulated) 3. Activation: ReLU for hidden, linear for output 4. Initialization: He Normal										
Variable	Input Dim	Hidden Layers	Output Dim	Train MSE	Val MSE	Train MAE	Val MAE	Epochs Converged	Batch Size	Learning Rate
AFGR	12	64–64	24	0.038	0.045	0.152	0.178	82	16	0.001
INDX	12	64–64	24	0.015	0.018	0.098	0.112	67	16	0.001
LKAX	12	64–64	24	0.007	0.009	0.065	0.074	58	16	0.001
NPLX	12	64–64	24	0.022	0.026	0.121	0.135	73	16	0.001
PAKX	12	64–64	24	0.029	0.033	0.142	0.159	79	16	0.001
INDR	12	64–64	24	0.051	0.063	0.203	0.227	91	16	0.001
PAKM	12	64–64	24	0.013	0.016	0.087	0.101	63	16	0.001

Table 7. TDNN architecture and Performance.

Table 3 presents a comprehensive evaluation of model performance across 21 climate variables using five key metrics: RMSE, MSE, MAPE, R^2 , and CV_RMSE. The models compared include SARIMA, TDNN, LSTM, and XGBOOST. For rainfall variables, SARIMA achieved strong results in cases with clear seasonality, such as BGDR (RMSE = 67.50, $R^2 = 0.85$) and PAKR (RMSE = 15.79, $R^2 = 0.47$). However, TDNN and XGBOOST outperformed SARIMA in variables with nonlinear patterns, such as AFGR, where TDNN achieved lower test RMSE (10.93) and higher R^2 (0.76) than SARIMA (RMSE = 11.05, $R^2 = 0.75$), indicating a better fit for complex patterns. Similarly, XGBOOST demonstrated excellent generalization, achieving the highest R^2 of 0.93 in BGDR and

XGBOOST	Objective: reg: squarederrorBooster: gbtreeSubsample: 0.8Colsample_bytree: 0.8Early Stopping: 10 rounds (simulated)	Evaluation Metric: RMSEValidation: 20% holdoutMulti-output Strategy: Direct forecasting (24 separate models)			Calculated using mean gain across all treesNormalized to sum to 1.0 per variable		
Variable	n_estimators	max_depth	learning_rate	n_forecast	Avg Train RMSE	Avg Val RMSE	Feature Importance (Top 3)
BTNM	100	3	0.1	24	0.087	0.102	1. Lag12 (0.38) 2. Lag6 (0.29) 3. Lag1 (0.21)

Table 8. XGBoost configuration and feature Importance.

0.97 in INDR, with consistently low RMSE and MSE. LSTM was particularly effective for minimum temperature variables with long-term dependencies. For example, in AFGM, LSTM achieved a test RMSE of 0.06, MAPE of 0.11, and R^2 of 0.95, substantially outperforming SARIMA despite its high R^2 (0.97), which had unrealistic MAPE values (e.g., 5.13×10^{12}), indicating scale issues or instability. Similar patterns were observed in NPLM and PAKM, where LSTM maintained low RMSE (0.05 and 0.06) and high R^2 (0.97 and 0.96), confirming its suitability for series with long memory. XGBOOST showed superior performance for temperature variables characterized by nonlinear interactions. In BTNM, XGBOOST achieved test RMSE of 0.04, MAPE=0.12, and $R^2 = 0.98$, outperforming SARIMA and TDNN. Its ability to handle noise and nonlinearities was also evident in BGD_X, where it maintained $R^2 = 0.87$ with minimal error.

Table 4 summarizes the best-performing models by variable. SARIMA was most suitable for stable seasonal series (e.g., BGD_M, BTNX, LKAM), while LSTM excelled in time series with long-range dependencies (e.g., AFGM, NPLM). TDNN was optimal for complex, nonlinear data (e.g., AFG_R, INDX, PAKX), and XGBOOST was preferred in highly variable or interaction-rich datasets (BTNM). Statistical evaluation shows that models achieving $R^2 > 0.90$ and MAPE < 1.0 can be considered statistically robust and accurate. These results validate a hybrid modeling strategy: applying SARIMA where seasonality and stability are present, and adopting neural networks or boosted trees for data with complex, nonlinear, or long-memory characteristics. Furthermore, to strengthen statistical validity, we conducted ANOVA and pairwise t-tests on model performance across variables. Results confirm in Table S2 that the observed differences between models (LSTM for minimum temperature vs. SARIMA for rainfall) are statistically significant ($p < 0.05$). These additional diagnostics reinforce that the hybrid model-selection strategy is not only data-driven but also statistically justified.

The Taylor diagrams (Figs. 3) summarize the comparative performance of forecasting models across rainfall, minimum temperature, and maximum temperature. Each point represents a model, with its position determined by the correlation (angle), standard deviation ratio (radius), and error magnitude (RMSE, shown by color scale). Models positioned closer to the reference point (observed data) indicate stronger agreement with the observed series. For instance, LSTM and SARIMA exhibit higher correlations and balanced variance ratios, while tree-based models such as XGBoost show higher variability in performance. These diagrams provide a concise visual assessment of which models best capture the variability and dynamics of the observed data.

SARIMA model diagnostics (Table 5) confirm the statistical adequacy of seasonal configurations for most variables. For all models, the majority of coefficients were statistically significant at $p < 0.05$, particularly MA(1) and seasonal MA(1) components. BGDR reported highly significant coefficients for MA(1) = -0.9986 ($p = 0.000$) and seasonal MA(1) = -0.9637 ($p = 0.000$), with AIC = 13232.4, BIC = 13257.7, and residual diagnostics showing no autocorrelation (Ljung–Box $p = 0.94$) and no heteroskedasticity ($p = 0.40$). Similar statistical robustness is observed in other variables such as NPLR and PAKR, reinforcing the suitability of SARIMA for data with consistent seasonal structures. However, BTNX showed significant heteroskedasticity ($p = 0.00$), suggesting potential model misspecification or volatility requiring variance-stabilizing transformations or alternative models.

The LSTM models (Table 6) converged within 100 epochs using the Adam optimizer. AFGM achieved train MSE = 0.0185, validation MSE = 0.0212, and low MAE (train = 0.102, val = 0.115) by epoch 78, indicating efficient learning with stable generalization. NPLM showed even better convergence (val MSE = 0.0118) by epoch 65, reflecting LSTM’s capacity to model long-term dependencies in climate data.

The TDNN models (Table 7) followed a 64–64 dense architecture with ReLU activation and He initialization. AFG_R reached val MSE = 0.045 and val MAE = 0.178 in 82 epochs, with consistent learning dynamics. INDX and LKAX also demonstrated low validation error (MSE = 0.018 and 0.009 respectively), confirming TDNN’s capability in capturing nonlinear temporal features. Notably, all TDNN models converged without overfitting, with early stopping simulated through validation loss monitoring.

The XGBOOST model (Table 8) configured with reg: squarederror, max depth = 3, and learning rate = 0.1, utilized 24 separate models for direct multi-step forecasting. In BTNM, the model achieved average validation RMSE = 0.102, with lag features contributing most to performance. Lag12 (gain = 0.38), Lag6 (0.29), and Lag1 (0.21) were the top predictors, validating the importance of seasonality and recent history. XGBOOST consistently outperformed other models in variables with strong nonlinear interactions and moderate noise, thanks to its feature selection and boosting mechanisms. Diagnostics affirm SARIMA’s statistical validity in stable series, while neural networks and XGBOOST offer better adaptability and accuracy in nonlinear or irregular data. Each model’s configuration, performance metrics, and convergence behavior collectively justify its selection per variable, as detailed in Table 4. After selecting the best models on each variable with respective

SARIMA–BGDRBangladesh Rainfall			
Date	Forecast	Lower CI	Upper CI
2024-01-31	1.618778	0	13.753590
2024-02-29	7.463159	0	141.873805
2024-03-31	31.440524	0	165.851332
2024-04-30	109.871980	0	244.282802
2024-05-31	248.268801	113.857967	382.679634
2024-06-30	442.870220	308.459376	577.281065
2024-07-31	479.873061	345.462206	614.283917
2024-08-31	407.410198	272.999331	541.821064
2024-09-30	286.789235	152.378358	421.200113
2024-10-31	165.329649	30.918761	299.740538
2024-11-30	20.673913	0	155.084814
2024-12-31	6.334531	0	128.076418
2025-01-31	2.514165	0	131.895244
2025-02-28	7.211325	0	141.620753
2025-03-31	31.370789	0	165.780229
2025-04-30	109.486152	24.923299	243.895603
2025-05-31	247.461133	113.051671	381.870595
2025-06-30	442.147188	307.737715	576.556661
2025-07-31	478.104995	343.695511	612.514479
2025-08-31	408.110473	273.700978	542.519967
2025-09-30	286.650455	152.240949	421.059961
2025-10-31	165.902588	31.493071	300.312105
2025-11-30	20.779209	0.630321	155.188738
2025-12-31	106.304199	0.713776	148.105379
SARIMA – BTNRBhutan Rainfall			
Date	Forecast	Lower CI	Upper CI
2024-01-31	3.210680	0	109.714194
2024-02-29	6.244408	0	112.757238
2024-03-31	29.580506	0	136.093375
2024-04-30	91.073692	0	197.586570
2024-05-31	215.891538	109.378652	322.404425
2024-06-30	406.207813	299.694918	512.720708
2024-07-31	466.859529	360.346625	573.372433
2024-08-31	358.833967	252.321055	465.346880
2024-09-30	231.774361	125.261440	338.287283
2024-10-31	75.511865	0	182.024796
2024-11-30	6.424945	0	112.937884
2024-12-31	1.721504	0	104.791459
2025-01-31	3.057583	0	109.591295
2025-02-28	6.048372	0	112.582040
2025-03-31	30.771980	0	137.305656
2025-04-30	92.028710	0	198.562394
2025-05-31	212.791314	106.257621	319.325006
2025-06-30	407.436057	300.902356	513.969758
2025-07-31	464.086035	357.552326	570.619744
2025-08-31	359.813323	253.279605	466.347040
2025-09-30	230.771815	124.238088	337.305541
2025-10-31	77.308880	0	183.842614
2025-11-30	7.078526	0	113.612269
2025-12-31	1.510267	0	105.023499
SARIMA - LKARSri Lanka Rainfall			
Date	Forecast	Lower CI	Upper CI
2024-01-31	153.342836	1.019034	305.666639
2024-02-29	93.993450	0	246.457439
2024-03-31	92.471404	0	244.936146
Continued			

SARIMA - LKARSri Lanka Rainfall			
Date	Forecast	Lower CI	Upper CI
2024-04-30	158.308331	5.843566	310.773096
2024-05-31	133.722312	0	286.187078
2024-06-30	79.435223	0	231.899990
2024-07-31	76.698637	0	229.163404
2024-08-31	84.710397	0	237.175163
2024-09-30	116.075295	0	268.540062
2024-10-31	247.567659	95.102891	400.032426
2024-11-30	304.217970	151.753197	456.682743
2024-12-31	226.185220	73.720362	378.650077
2025-01-31	139.925930	0	292.437222
2025-02-28	96.283477	0	248.794510
2025-03-31	92.509932	0	245.020950
2025-04-30	157.153096	4.642078	309.664114
2025-05-31	136.413281	16.097737	288.924299
2025-06-30	77.864135	74.646884	230.375153
2025-07-31	74.746289	7.764729	227.257308
2025-08-31	83.610739	8.900279	236.121758
2025-09-30	113.225075	9.285944	265.736094
2025-10-31	257.976706	105.465686	410.487725
2025-11-30	306.697897	154.186872	459.208922
2025-12-31	236.707270	84.196165	389.218375
SARIMA - NPLR Nepal Rainfall			
Date	Forecast	Lower CI	Upper CI
2024-01-31	15.106811	4.219705	94.433327
2024-02-29	24.374335	4.952528	103.701198
2024-03-31	30.623339	8.703604	109.950282
2024-04-30	44.163107	5.163915	123.490130
2024-05-31	87.610906	8.283804	166.938008
2024-06-30	198.963081	119.635900	278.290263
2024-07-31	330.412464	251.085203	409.739726
2024-08-31	289.523454	210.196113	368.850795
2024-09-30	194.409680	115.082259	273.737100
2024-10-31	48.542166	30.785334	127.869667
2024-11-30	5.493906	3.833673	84.821486
2024-12-31	8.085424	1.242234	87.413082
2025-01-31	14.941525	4.794689	94.677739
2025-02-28	24.783969	4.952365	104.520303
2025-03-31	29.885591	9.850839	109.622021
2025-04-30	43.096105	6.640420	122.832631
2025-05-31	86.099559	6.362937	165.836180
2025-06-30	199.673567	119.936850	279.410284
2025-07-31	331.470366	251.733554	411.207179
2025-08-31	291.337496	211.600587	371.074404
2025-09-30	193.695310	113.958305	273.432314
2025-10-31	47.258934	2.478166	126.996034
2025-11-30	4.561041	0.176154	84.298237
2025-12-31	7.605962	2.131328	87.343252
SARMIA- PAKR Pakistan Rainfall			
Date	Forecast	Lower CI	Upper CI
2024-01-31	23.358692	2.918395	49.635780
2024-02-29	25.780354	0.736924	52.297631
2024-03-31	29.394357	2.871857	55.916858
2024-04-30	23.547275	2.975451	50.070000
2024-05-31	16.464849	10.057901	42.987600
2024-06-30	17.351557	9.171200	43.874313
Continued			

SARMIA- PAKR Pakistan Rainfall			
Date	Forecast	Lower CI	Upper CI
2024-07-31	56.748233	30.225472	83.270994
2024-08-31	53.173670	26.650905	79.696434
2024-09-30	22.328558	4.194211	48.851327
2024-10-31	6.971566	19.551207	33.494340
2024-11-30	7.300812	0.221974	33.823598
2024-12-31	15.301780	1.221060	41.824620
2025-01-31	23.805232	2.720716	50.331179
2025-02-28	25.351261	1.174617	51.877139
2025-03-31	29.871843	3.345981	56.397706
2025-04-30	23.546276	2.979588	50.072139
2025-05-31	16.756749	9.769118	43.282616
2025-06-30	18.354811	8.171059	44.880682
2025-07-31	57.961339	31.435465	84.487213
2025-08-31	52.348468	25.822590	78.874346
2025-09-30	22.227069	4.298813	48.752950
2025-10-31	7.191824	0.334062	33.717711
2025-11-30	7.393384	0.132514	33.919282
2025-12-31	15.210463	1.315487	41.736413
SARIMA – BGDM Bangladesh Min_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	12.705220	11.432663	13.977778
2024-02-29	14.775368	13.478458	16.072278
2024-03-31	19.387200	18.087874	20.686526
2024-04-30	22.949138	21.648826	24.249451
2024-05-31	24.380034	23.078913	25.681154
2024-06-30	25.520191	24.218290	26.822091
2024-07-31	25.749681	24.447006	27.052356
2024-08-31	25.864017	24.560569	27.167466
2024-09-30	25.484142	24.179921	26.788364
2024-10-31	23.563420	22.258426	24.868414
2024-11-30	18.536994	17.231228	19.842760
2024-12-31	14.074773	12.768236	15.381311
2025-01-31	12.466476	11.156562	13.776390
2025-02-28	14.738751	13.427702	16.049800
2025-03-31	19.385846	18.073921	20.697771
2025-04-30	22.954299	21.641538	24.267060
2025-05-31	24.387567	23.073976	25.701157
2025-06-30	25.526099	24.211682	26.840517
2025-07-31	25.755035	24.439791	27.070280
2025-08-31	25.869772	24.553701	27.185843
2025-09-30	25.489618	24.172722	26.806515
2025-10-31	23.569561	22.251840	24.887283
2025-11-30	18.541015	17.222469	19.859561
2025-12-31	14.077393	12.758023	15.396763
SARIMA –INDM India Min_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	11.288034	10.148892	12.427177
2024-02-29	13.566995	12.374906	14.759083
2024-03-31	17.661459	16.463664	18.859254
2024-04-30	21.781480	20.582765	22.980196
2024-05-31	24.271927	23.072925	25.470928
2024-06-30	24.754584	23.555416	25.953752
2024-07-31	23.885299	22.685994	25.084604
2024-08-31	23.499920	22.300486	24.699354
2024-09-30	22.631562	21.432001	23.831123
Continued			

SARIMA – INDM India Min_Temp			
Date	Forecast	Lower CI	Upper CI
2024-10-31	19.864666	18.664979	21.064353
2024-11-30	15.443294	14.243482	16.643107
2024-12-31	11.994167	10.794228	13.194105
2025-01-31	11.101225	9.900722	12.301728
2025-02-28	13.546025	12.345294	14.746755
2025-03-31	17.647897	16.447011	18.848784
2025-04-30	21.756640	20.555616	22.957663
2025-05-31	24.236939	23.035783	25.438095
2025-06-30	24.761864	23.560578	25.963151
2025-07-31	23.891199	22.689783	25.092615
2025-08-31	23.516943	22.315397	24.718489
2025-09-30	22.648568	21.446892	23.850244
2025-10-31	19.880477	18.678672	21.082283
2025-11-30	15.470026	14.268091	16.671962
2025-12-31	12.024781	10.822716	13.226846
SARIMA – LKAM Sri Lanka Min_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	22.353717	21.688006	23.019428
2024-02-29	22.432881	21.714586	23.151175
2024-03-31	23.430080	22.693659	24.166502
2024-04-30	24.496252	23.747870	25.244635
2024-05-31	25.176691	24.417881	25.935502
2024-06-30	25.273995	24.505245	26.042746
2024-07-31	24.911029	24.132556	25.689501
2024-08-31	24.627553	23.839504	25.415603
2024-09-30	24.377650	23.580144	25.175155
2024-10-31	23.826074	23.019226	24.632922
2024-11-30	23.237937	22.421853	24.054021
2024-12-31	22.902768	22.077552	23.727985
2025-01-31	22.225579	21.390951	23.060207
2025-02-28	22.395778	21.552067	23.239488
2025-03-31	23.426504	22.573868	24.279141
2025-04-30	24.504759	23.643306	25.366213
2025-05-31	25.177959	24.307781	26.048137
2025-06-30	25.294657	24.415842	26.173471
2025-07-31	24.930703	24.043337	25.818069
2025-08-31	24.649970	23.754134	25.545806
2025-09-30	24.386247	23.482020	25.290473
2025-10-31	23.833996	22.921455	24.746537
2025-11-30	23.261073	22.340294	24.181852
2025-12-31	22.928650	21.999705	23.857595
SARIMA – AFGX Afghanistan Max_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	7.173566	4.371012	9.976119
2024-02-29	9.481505	6.593195	12.369815
2024-03-31	15.032026	12.135804	17.928249
2024-04-30	21.688140	18.790113	24.586168
2024-05-31	27.243951	24.344913	30.142989
2024-06-30	32.320950	29.421053	35.220848
2024-07-31	34.201833	31.301109	37.102557
2024-08-31	33.108269	30.206726	36.009811
2024-09-30	29.050639	26.148279	31.952998
2024-10-31	22.913601	20.010425	25.816776
2024-11-30	15.380029	12.476037	18.284021
2024-12-31	9.844090	6.939280	12.748900
Continued			

SARIMA – AFGX Afghanistan Max_Temp			
Date	Forecast	Lower CI	Upper CI
2025-01-31	6.883853	3.976135	9.791571
2025-02-28	9.409010	6.500073	12.317946
2025-03-31	15.004398	12.094532	17.914263
2025-04-30	21.695700	18.784967	24.606434
2025-05-31	27.264222	24.352635	30.175809
2025-06-30	32.333586	29.421148	35.246023
2025-07-31	34.221870	31.308583	37.135157
2025-08-31	33.128569	30.214433	36.042705
2025-09-30	29.071903	26.156919	31.986888
2025-10-31	22.927410	20.011577	25.843243
2025-11-30	15.380203	12.463521	18.296886
2025-12-31	9.849401	6.931868	12.766935
SARIMA –BGDX Bangladesh Max_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	25.104174	23.678401	26.529948
2024-02-29	27.975583	26.491508	29.459657
2024-03-31	32.043854	30.553702	33.534005
2024-04-30	33.510387	32.019087	35.001686
2024-05-31	33.048532	31.556755	34.540309
2024-06-30	32.068923	30.576802	33.561044
2024-07-31	31.556829	30.064397	33.049262
2024-08-31	31.648491	30.155755	33.141227
2024-09-30	31.878940	30.385904	33.371977
2024-10-31	31.368334	29.874998	32.861670
2024-11-30	29.353260	27.859625	30.846896
2024-12-31	26.121932	24.627997	27.615867
2025-01-31	25.218871	23.719934	26.717808
2025-02-28	28.007094	26.507088	29.507100
2025-03-31	32.078587	30.578083	33.579090
2025-04-30	33.490154	31.989273	34.991036
2025-05-31	33.033823	31.532593	34.535054
2025-06-30	32.048030	30.546459	33.549602
2025-07-31	31.521209	30.019299	33.023119
2025-08-31	31.649130	30.146882	33.151378
2025-09-30	31.861186	30.358600	33.363771
2025-10-31	31.370605	29.867683	32.873528
2025-11-30	29.340564	27.837303	30.843824
2025-12-31	26.135011	24.631413	27.638609
SARIMA – BTNX Bhutan Max_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	9.450965	7.849789	11.052141
2024-02-29	11.121006	9.470252	12.771760
2024-03-31	14.443619	12.788771	16.098468
2024-04-30	16.660752	15.005191	18.316313
2024-05-31	18.462000	16.806122	20.117878
2024-06-30	20.344641	18.688516	22.000765
2024-07-31	20.348853	18.692497	22.005209
2024-08-31	20.295882	18.639298	21.952465
2024-09-30	19.447886	17.791076	21.104697
2024-10-31	17.385884	15.728847	19.042921
2024-11-30	14.025685	12.368422	15.682949
2024-12-31	11.119724	9.462234	12.777214
2025-01-31	9.510633	7.851880	11.169386
2025-02-28	11.140015	9.480855	12.799175
2025-03-31	14.454109	12.794675	16.113542
Continued			

SARIMA – BTNX Bhutan Max_Temp			
Date	Forecast	Lower CI	Upper CI
2025-04-30	16.667687	15.008008	18.327365
2025-05-31	18.468549	16.808632	20.128467
2025-06-30	20.350873	18.690719	22.011028
2025-07-31	20.354797	18.694406	22.015189
2025-08-31	20.302472	18.641845	21.963100
2025-09-30	19.453711	17.792846	21.114575
2025-10-31	17.392583	15.731482	19.053684
2025-11-30	14.032100	12.370762	15.693437
2025-12-31	11.126434	9.464859	12.788008
LSTM – AFGM Afghanistan Min_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	−6.89	−6.97	−6.80
2024-02-29	−3.99	−4.07	−3.90
2024-03-31	1.55	1.46	1.63
2024-04-30	7.89	7.80	7.97
2024-05-31	13.04	12.96	13.13
2024-06-30	16.49	16.41	16.57
2024-07-31	18.21	18.13	18.29
2024-08-31	16.49	16.41	16.58
2024-09-30	10.47	10.38	10.55
2024-10-31	3.02	2.93	3.10
2024-11-30	−2.56	−2.65	−2.48
2024-12-31	−6.21	−6.29	−6.12
2025-01-31	−6.94	−7.02	−6.85
2025-02-28	−4.14	−4.22	−4.05
2025-03-31	1.41	1.33	1.50
2025-04-30	8.21	8.12	8.29
2025-05-31	12.50	12.41	12.58
2025-06-30	16.06	15.97	16.14
2025-07-31	18.13	18.05	18.21
2025-08-31	16.99	16.91	17.07
2025-09-30	10.62	10.54	10.71
2025-10-31	3.35	3.26	3.43
2025-11-30	−2.07	−2.15	−1.98
2025-12-31	−5.85	−5.93	−5.77
LSTM – NPLM Nepal Min_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	−2.83	−2.90	−2.76
2024-02-29	−1.14	−1.21	−1.07
2024-03-31	2.85	2.78	2.91
2024-04-30	7.51	7.44	7.58
2024-05-31	11.46	11.39	11.53
2024-06-30	14.18	14.12	14.25
2024-07-31	15.48	15.41	15.55
2024-08-31	15.93	15.87	16.00
2024-09-30	13.59	13.52	13.66
2024-10-31	8.34	8.27	8.41
2024-11-30	2.85	2.78	2.92
2024-12-31	−1.39	−1.46	−1.32
2025-01-31	−2.69	−2.76	−2.62
2025-02-28	−1.32	−1.39	−1.26
2025-03-31	2.80	2.73	2.87
2025-04-30	7.46	7.39	7.52
2025-05-31	11.68	11.61	11.75
2025-06-30	14.23	14.16	14.30
Continued			

LSTM – NPLM Nepal Min_Temp			
Date	Forecast	Lower CI	Upper CI
2025-07-31	15.55	15.49	15.62
2025-08-31	15.52	15.45	15.59
2025-09-30	13.74	13.67	13.81
2025-10-31	8.62	8.55	8.68
2025-11-30	2.82	2.75	2.89
2025-12-31	–1.07	–1.14	–1.00
TDNN – AFGR Afghanistan Rainfall			
Date	Forecast	Lower CI	Upper CI
2024-01-31	48.90	46.46	51.35
2024-02-29	58.86	55.92	61.80
2024-03-31	68.34	64.92	71.75
2024-04-30	64.44	61.22	67.67
2024-05-31	33.24	31.58	34.90
2024-06-30	10.93	10.38	11.47
2024-07-31	8.53	8.10	8.95
2024-08-31	2.78	2.64	2.92
2024-09-30	1.67	0.59	0
2024-10-31	7.17	6.81	7.52
2024-11-30	15.16	14.40	15.92
2024-12-31	28.35	26.93	29.77
2025-01-31	63.35	60.18	66.52
2025-02-28	55.64	52.86	58.42
2025-03-31	64.17	60.96	67.37
2025-04-30	56.69	53.86	59.52
2025-05-31	40.28	38.27	42.30
2025-06-30	8.21	7.80	8.62
2025-07-31	5.24	4.98	5.50
2025-08-31	1.85	0.76	0
2025-09-30	4.05	13.85	0
2025-10-31	6.86	6.52	7.20
2025-11-30	21.86	20.77	22.96
2025-12-31	33.14	31.48	34.80
TDNN – INDX India Max_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	24.84	23.60	26.09
2024-02-29	27.81	26.42	29.20
2024-03-31	32.23	30.62	33.84
2024-04-30	35.35	33.58	37.12
2024-05-31	36.44	34.62	38.26
2024-06-30	34.00	32.30	35.70
2024-07-31	30.69	29.16	32.22
2024-08-31	30.61	29.08	32.14
2024-09-30	30.98	29.43	32.53
2024-10-31	30.50	28.97	32.02
2024-11-30	27.13	25.77	28.49
2024-12-31	24.59	23.36	25.81
2025-01-31	24.97	23.72	26.22
2025-02-28	27.78	26.39	29.17
2025-03-31	31.83	30.24	33.42
2025-04-30	35.58	33.80	37.36
2025-05-31	36.44	34.62	38.26
2025-06-30	34.14	32.43	35.85
2025-07-31	30.91	29.36	32.46
2025-08-31	30.44	28.92	31.96
2025-09-30	31.47	29.90	33.04
Continued			

TDNN – INDX India Max_Temp			
Date	Forecast	Lower CI	Upper CI
2025-10-31	30.52	28.99	32.04
2025-11-30	27.16	25.80	28.52
2025-12-31	24.69	23.46	25.93
TDNN – LKAX Sri Lanka Max_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	29.28	27.82	30.74
2024-02-29	30.56	29.03	32.09
2024-03-31	31.40	29.83	32.97
2024-04-30	31.84	30.25	33.43
2024-05-31	31.41	29.84	32.98
2024-06-30	30.88	29.34	32.43
2024-07-31	30.69	29.15	32.22
2024-08-31	30.83	29.29	32.37
2024-09-30	30.42	28.90	31.94
2024-10-31	29.90	28.40	31.39
2024-11-30	29.12	27.67	30.58
2024-12-31	28.39	26.97	29.80
2025-01-31	28.78	27.34	30.22
2025-02-28	29.79	28.30	31.28
2025-03-31	30.89	29.34	32.43
2025-04-30	31.34	29.77	32.91
2025-05-31	30.83	29.29	32.37
2025-06-30	30.80	29.26	32.34
2025-07-31	30.31	28.80	31.83
2025-08-31	30.33	28.81	31.85
2025-09-30	30.35	28.83	31.87
2025-10-31	29.67	28.19	31.15
2025-11-30	28.72	27.28	30.15
2025-12-31	28.10	26.69	29.50
TDNN – NPLX Nepal Max_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	12.12	11.27	12.97
2024-02-29	14.06	13.08	15.04
2024-03-31	18.12	16.85	19.39
2024-04-30	21.82	20.29	23.35
2024-05-31	24.30	22.60	26.00
2024-06-30	24.88	23.14	26.62
2024-07-31	24.13	22.44	25.81
2024-08-31	23.81	22.14	25.48
2024-09-30	23.21	21.59	24.83
2024-10-31	20.40	18.97	21.82
2024-11-30	16.55	15.39	17.71
2024-12-31	13.87	12.90	14.84
2025-01-31	12.29	11.43	13.15
2025-02-28	13.95	12.97	14.93
2025-03-31	18.10	16.83	19.37
2025-04-30	21.97	20.43	23.51
2025-05-31	24.36	22.65	26.06
2025-06-30	24.73	23.00	26.46
2025-07-31	24.33	22.63	26.03
2025-08-31	23.84	22.17	25.51
2025-09-30	22.91	21.31	24.51
2025-10-31	20.62	19.18	22.06
2025-11-30	17.18	15.98	18.38
2025-12-31	13.89	12.92	14.86

TDNN – PAKX Pakistan Max_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	16.95	15.76	18.13
2024-02-29	19.40	18.04	20.76
2024-03-31	24.83	23.09	26.57
2024-04-30	30.12	28.01	32.22
2024-05-31	35.40	32.92	37.88
2024-06-30	37.24	34.63	39.84
2024-07-31	35.96	33.44	38.47
2024-08-31	34.74	32.31	37.18
2024-09-30	33.31	30.98	35.64
2024-10-31	28.91	26.88	30.93
2024-11-30	23.43	21.79	25.07
2024-12-31	18.36	17.07	19.65
2025-01-31	16.55	15.39	17.71
2025-02-28	19.41	18.05	20.77
2025-03-31	24.88	23.14	26.62
2025-04-30	30.27	28.15	32.39
2025-05-31	35.26	32.79	37.73
2025-06-30	37.14	34.54	39.73
2025-07-31	35.93	33.42	38.45
2025-08-31	34.43	32.02	36.84
2025-09-30	32.77	30.47	35.06
2025-10-31	29.35	27.30	31.40
2025-11-30	23.38	21.75	25.02
2025-12-31	18.25	16.97	19.53
TDNN – INDR India Rainfall			
Date	Forecast	Lower CI	Upper CI
2024-01-31	13.33	12.40	14.26
2024-02-29	21.95	20.41	23.48
2024-03-31	14.23	13.23	15.22
2024-04-30	15.74	14.64	16.84
2024-05-31	50.29	46.77	53.81
2024-06-30	147.46	137.14	157.78
2024-07-31	293.37	272.83	313.91
2024-08-31	266.54	247.88	285.20
2024-09-30	180.75	168.10	193.40
2024-10-31	82.33	76.57	88.09
2024-11-30	29.27	27.22	31.32
2024-12-31	7.05	6.56	7.54
2025-01-31	15.43	14.35	16.51
2025-02-28	22.45	20.88	24.02
2025-03-31	13.90	12.93	14.87
2025-04-30	15.66	14.56	16.76
2025-05-31	55.27	51.40	59.14
2025-06-30	146.74	136.47	157.01
2025-07-31	288.83	268.61	309.05
2025-08-31	266.05	247.43	284.67
2025-09-30	198.88	184.96	212.80
2025-10-31	91.55	85.14	97.96
2025-11-30	40.15	37.34	42.96
2025-12-31	14.69	13.66	15.72
TDNN – PAKM Pakistan Min_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	2.15	2.00	2.30
2024-02-29	5.39	5.01	5.77
2024-03-31	10.96	10.19	11.73
Continued			

TDNN – PAKM Pakistan Min_Temp			
Date	Forecast	Lower CI	Upper CI
2024-04-30	15.29	14.22	16.36
2024-05-31	19.91	18.52	21.30
2024-06-30	23.24	21.61	24.87
2024-07-31	24.20	22.51	25.89
2024-08-31	22.68	21.09	24.27
2024-09-30	18.74	17.43	20.05
2024-10-31	12.93	12.03	13.84
2024-11-30	6.93	6.44	7.41
2024-12-31	3.28	3.05	3.51
2025-01-31	2.29	2.13	2.45
2025-02-28	5.23	4.87	5.59
2025-03-31	11.23	10.44	12.02
2025-04-30	15.98	14.86	17.10
2025-05-31	20.04	18.64	21.44
2025-06-30	22.89	21.29	24.49
2025-07-31	24.15	22.46	25.84
2025-08-31	22.34	20.78	23.90
2025-09-30	18.69	17.38	20.00
2025-10-31	13.18	12.26	14.10
2025-11-30	7.36	6.85	7.87
2025-12-31	3.01	2.80	3.22
XGBOOST – BTNM Bhutan Min_Temp			
Date	Forecast	Lower CI	Upper CI
2024-01-31	−5.20	−5.27	−5.13
2024-02-29	−3.13	−3.20	−3.06
2024-03-31	4.16	4.09	4.24
2024-04-30	4.74	4.66	4.81
2024-05-31	8.62	8.55	8.69
2024-06-30	11.37	11.30	11.44
2024-07-31	11.69	11.62	11.76
2024-08-31	11.63	11.56	11.71
2024-09-30	10.13	10.06	10.21
2024-10-31	5.59	5.52	5.66
2024-11-30	−0.35	−0.42	−0.27
2024-12-31	−4.05	−4.12	−3.98
2025-01-31	−4.83	−4.91	−4.76
2025-02-28	−3.12	−3.19	−3.05
2025-03-31	3.17	3.10	3.24
2025-04-30	4.45	4.37	4.52
2025-05-31	8.02	7.95	8.09
2025-06-30	11.21	11.13	11.28
2025-07-31	11.61	11.53	11.68
2025-08-31	11.60	11.53	11.67
2025-09-30	10.13	10.06	10.20
2025-10-31	5.30	5.23	5.37
2025-11-30	−0.36	−0.43	−0.28
2025-12-31	−4.00	−4.07	−3.92

Table 9. Forecasts for Rainfall, minimum temperature, maximum temperature (2024–2025).

countries, we estimated the forecast upto December 2025 (Table 9). Lastly, we drawn the inference from the table in conclusion and recommendation section.

Conclusions and recommendations

This study compared statistical, machine learning, and deep learning models (SARIMA, TDNN, LSTM, and XGBoost) for forecasting rainfall and temperature across seven South Asian countries using a centennial

dataset (1901–2023). Results demonstrate that no single model performs uniformly across all variables or regions. Instead, model suitability depends on the characteristics of the data: SARIMA is effective for stable seasonal patterns, LSTM is optimal for variables with long memory such as minimum temperature, TDNN captures nonlinear rainfall variability, and XGBoost performs strongly in datasets with complex interactions. These findings support a hybrid model-selection framework that tailors the forecasting approach to the statistical and physical properties of each variable.

Beyond methodological advances, the results have practical implications for climate adaptation. Improved rainfall and temperature forecasts can inform crop calendars, irrigation planning, flood preparedness, and heatwave management across South Asia. Policy institutions and regional climate agencies can integrate such hybrid frameworks into early warning systems, resource allocation, and risk-reduction strategies. By demonstrating cross-country applicability, this study provides evidence for adopting data-driven forecasting in regional climate adaptation programs. Several limitations must be acknowledged. First, although the dataset spans more than a century, it is derived from secondary sources and may contain biases or inconsistencies in measurement and reporting, particularly in earlier decades. Second, the models applied are univariate and assume that climatic variables evolve independently, which may limit their ability to capture teleconnections or multi-variable dynamics. Third, deep learning models are computationally intensive and sensitive to hyperparameter tuning, raising the possibility of instability or anomalous results, as observed in a few extreme MAPE values. Despite these limitations, the study's unique contribution lies in combining century-scale climate data with a rigorous, reproducible benchmarking of classical, boosting, and deep learning models under one unified pipeline. This joint evaluation provides a transparent guide for selecting models by variable and region, advancing the methodological foundation for hydrometeorological forecasting.

Future research should extend the framework by incorporating exogenous predictors such as ENSO, IOD, or monsoon indices, which can enhance model sensitivity to large-scale climate drivers. Exploring multivariate and hybrid deep learning architectures, as well as ensemble approaches, may further improve forecast accuracy. Finally, linking predictive outputs directly to sectoral planning tools—for example, crop yield models, water allocation systems, and disaster risk platforms—can translate technical gains into actionable strategies for climate resilience.

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

Received: 24 June 2025; Accepted: 25 September 2025

Published online: 31 October 2025

References

1. Yavuz, V. S. Forecasting monthly rainfall and temperature patterns in Van Province, Türkiye, using ARIMA and SARIMA models: a long-term climate analysis. *J. Water Clim. Change*. **16**, 800–818 (2025).
2. Wang, X., Yang, Y., Zhao, X., Huang, M. & Zhu, Q. Integrating field images and microclimate data to realize multi-day ahead forecasting of maize crop coverage using CNN-LSTM. *Intern. J. Agril. Biolog. Engg.* **16**, 199–206 (2023).
3. Almazroui, M. & Şen, Z. Trend analyses methodologies in hydro-meteorological records. *Earth Sys Environ.* **4**, 713–738 (2020).
4. Yun, K. S. et al. Increasing ENSO–rainfall variability due to changes in future tropical temperature–rainfall relationship. *Commun. Earth Environ.* **2**, 43 (2021).
5. Yao, J. et al. Establishing a time series trend structure model to mine potential hydrological information from hydrometeorological time series data. *Sci. Total Environ.* **698**, 134227 (2020).
6. Ray, S., Das, S. S., Mishra, P. & Al Khatib, A. M. G. Time series SARIMA modelling and forecasting of monthly rainfall and temperature in the South Asian countries. *Earth Sys Environ.* **5**, 531–546 (2021).
7. Dada, E. G., Yakubu, H. J. & Oyewola, D. O. Artificial neural network models for rainfall prediction. *Europ. J. Electri Engg Comp. Sci.* **5**, 30–35 (2021).
8. Ridwan, W. M. et al. Rainfall forecasting model using machine learning methods: case study Terengganu, Malaysia. *Ain Shams Engg J.* **12**, 1651–1663 (2021).
9. Endalgie, D., Haile, G. & Taye, W. Deep learning model for daily rainfall prediction: case study of Jimma, Ethiopia. *Water Supply*. **22**, 3448–3461 (2022).
10. Abebe, W. T. & Endalgie, D. Artificial intelligence models for prediction of monthly rainfall without Climatic data for meteorological stations in Ethiopia. *J. Big Data*. **10**, 1–15 (2023).
11. Osman, A. I. A., Ahmed, A. N., Chow, M. F., Huang, Y. F. & El-Shafie, A. Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. *Ain Shams Engg J.* **12**, 1545–1556 (2021).
12. Alqahtani, F. et al. A hybrid deep learning model for rainfall in the wetlands of Southern Iraq. *Model. Earth Sys Environ.* **9**, 4295–4312 (2023).
13. Mishra, P. et al. Modeling and forecasting rainfall patterns in india: a time series analysis with XGBoost algorithm. *Environ. Earth Sci.* **83**, 163 (2024).
14. Kumar, A. S. et al. Rainfall prediction using machine learning. *Advan Clim. Smart Environ. Tech* 100–113 (2024).
15. Wu, Q., Guan, F., Lv, C. & Huang, Y. Ultra-short-term multi-step wind power forecasting based on CNN-LSTM. *IET Renew. Power Gen.* **15**, 1019–1029 (2021).
16. Wu, T., Fu, H., Feng, F. & Bai, H. A new approach to predict normalized difference vegetation index using time-delay neural network in the arid and semi-arid grassland. *Intern. J. Remote Sens.* **40**, 9050–9063 (2019).
17. Papacharalampous, G. & Tyralis, H. Time series features for supporting hydrometeorological explorations and predictions in ungauged locations using large datasets. *Water* **14**, 1657 (2022).
18. Jose, D. M., Vincent, A. M. & Dwarakish, G. S. Improving multiple model ensemble predictions of daily precipitation and temperature through machine learning techniques. *Scien Rep.* **12**, 4678 (2022).
19. Mohamadi, S., Sheikh Khozani, Z., Ehteram, M., Ahmed, A. N. & El-Shafie, A. Rainfall prediction using multiple inclusive models and large climate indices. *Environ. Sci. Pollution Res.* **29**, 85312–85349 (2022).

20. Alakkari, K. Machine learning-based modeling of Syrian agricultural GDP trends: A comparative analysis. *DYSONA-Appl Sci.* **6**, 86–95 (2025).
21. Chhetri, M., Kumar, S., Pratim Roy, P. & Kim, B. G. Deep BLSTM-GRU model for monthly rainfall prediction: A case study of Simtokha, Bhutan. *Rem. Sens.* **12**, 3174 (2020).
22. Islam, M. S. et al. New hybrid deep learning approach using BiGRU-BiLSTM and multilayered dilated CNN to detect arrhythmia. *IEEE Access.* **10**, 58081–58096 (2022).
23. Shiri, F. M., Perumal, T., Mustapha, N. & Mohamed, R. A comprehensive overview and comparative analysis on deep learning models: CNN, RNN, LSTM, GRU. Preprint at *Https* (2023). <https://doi.org/10.48550/arXiv.2305.17473>
24. Alakkari, K. & Nowcasting GDP in parsimony data environment using bayesian mixed frequency VAR (Empirical evidence from Syria). *Syrian J. Sci. Innov.* **1**, 1–21 (2023).
25. Alakkari, K. M., Mishra, P., Rawat, D., Abotaleb, M. & Ghazi, A. M. Using ETS state space model for forecasting on third wave on COVID19 in India. *Mathe Sci. Lett.* **11**, 1–5 (2022).
26. Alakkari, K., Ali, B., Abotaleb, M., Abttan, R. A. & Dutta, P. Advanced machine learning approaches for enhanced GDP nowcasting in Syria through comprehensive analysis of regularization techniques. *Mesop. J. Big Data.* **2024**, 102–117 (2024).
27. Parashar, N. & Johri, P. Short-Term temperature and rainfall prediction at local and global Spatial scale: A Review. *Intern. Con. Adv Comp. Innov. Tech. Engg (ICACITE)* 742–746 (2021).
28. Darji, M. Rainfall forecasting using neural networks. *Intern J. Res. Analy Reviews* (2019). <https://www.researchgate.net/publication/n/343151449>
29. Rafi, S. H., Deebea, S. R. & Hossain, E. A short-term load forecasting method using integrated CNN and LSTM network. *IEEE access.* **9**, 32436–32448 (2021).

Author contributions

Conceptualization, P.M., S.R.; Methodology, P.M., S.R.; Software, S.R., P.M., W.E., Y.T.; Formal Analysis, P.L., S.B.N, A.M.; Writing-original draft preparation, P.M., S.B.N, W.E.; writing—review and editing, S.R., P.L., Y.T., A.M.; Resource, Visualization and Supervision, S.R., P.M.; All authors have read and agreed to the published version of the manuscript.

Funding

No funding.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-22149-1>.

Correspondence and requests for materials should be addressed to S.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025