(RESEARCH)

Enhancing Temperature and Rainfall Prediction Accuracy Through Deep Learning Frameworks

Mohammad Sohel Kabir

Department of Physics, Cumilla Shikkha Board Govt. Model College, Cumilla, Bangladesh.

Abdullah Al Naseeh Chowdhur

Department of Computer Science and Engineering, Sylhet Engineering College, Sylhet, Bangladesh

Mozibur Rahman

Department of Civil Engineering, International University of Business Agriculture & Technology, Dhaka, Bangladesh.

Md. Rasel

Department of Computer Science and Engineering, Green University of Bangladesh, Dhaka, Bangladesh

Md Muhasin Ali

Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

Abstract

The information of accurate forecasting on temperature and rainfall is also very crucial for disaster preparedness, as well for climate management. Typical statistical and machine learning approaches have limited ability to capture nonlinear and spatiotemporally varying structure of climate fields. This research utilized recent state-of-the-art deep learning models to improve the prediction models for both temperature and rainfall. The hybrid Convolutional Neural Networks and Long Short-Term Memory (CNN–LSTM) method achieved the best results (R^2 = 0.98 for temperature, 0.91 for rainfall), outperforming those of Multiple Linear Regression (MLR) and Random Forest (RF) as a traditional model. The Physics-Informed Neural Network (PINN) model delivered physically consistent and stable predictions, especially under extreme weather such as heavy rainfall or heatwaves. Relative humidity, atmospheric pressure and sea surface temperature were found as most important predictors-based on feature importance analysis. The regional analysis remained that the coastal region performed best, whereas the hilly region with the high topographical complexity presented a relatively lower accuracy. In general, embedding deep learning into physical constraints ended up improving a lot both correctness and robustness of predictions. Further work should be carried out to improve interpretability, inclusiveness of data and transferability in space of such models with the ambition to build a more sustainable real-time weather forecasting system.

Keywords: Artificial Intelligence; Deep Learning; Temperature; Rainfall; Climate Forecasting.

1. Introduction

Variability and change in climate presents a major challenge for world ecosystems, agricultural productivity, water resources, and human populations (Abebaw, 2025). There has been increasing demand for accurate forecast of climatic events such as temperature and rainfall pattern in order to effectively manage the risk of climate, prepare for disasters, and improve on development sustainability (Chowdhury et al., 2022; Ebi et al., 2021). Classic statistical weather and numerical weather prediction (NWP) models, including the Global Forecast System (GFS) and European Centre for Medium-Range Weather Forecasts (ECMWF) models, are widely adopted for such predictions (Mayer et al., 2023). Nevertheless, such models have difficulties in representing complicated nonlinear interactions between climate variables since they are based on physical parameters and simplified assumptions (Pereira et al., 2024). It is, therefore, necessary to develop more sophisticated models which are adaptive and data-driven that can enhance the precision and reliability of temperature and rainfall forecasting at various time scales as well as spatial coverage (Sham et al., 2025; Waqas et al., 2025).

In the last few years, the rise of artificial intelligence (AI) and deep learning (DL) in particular has transformed data analysis practice across a broad range of domains including climate and environmental science applications (Olawade et al., 2024). Deep learning architectures, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) models has shown good performance in learning non-linear and spatiotemporal information from large dataset (Tayebi & El Kafhali, 2025; Waqas & Humphries, 2024). This model can automatically learn high-level features as well as temporal dependencies from the raw data, which is particularly suitable for climate prediction since such relationships are usually complex and changeable (Wang et al., 2024). Deep learning methodologies, in contrast to the traditional modelling techniques that rely largely on prespecified equations, learn directly from the observational (or reanalysis) data that allows for capturing complex dynamics between temperature, precipitation and humidity, wind speed and other atmospheric features (Patakchi Yousefi & Kollet, 2023).

A few research works have explored deep learning in weather prediction (Ange-Clement Akazan et al., 2022; Ren et al., 2021; Zhang et al., 2025). CNNs and LSTMs are applied for spatial pattern learning from satellite imagery (e.g., gridded climate data) and modeling temporal dependencies among time-series weather reports (O'Donncha et al., 2022). Recently, novel trans-modelling techniques that introduce combined CNN and LSTM structures have improved the performance of SARAH by taking into account both spatial perception and temporal memory so as to improve predictions of rainfall and temperature (Upadhyay et al., 2025). These developments demonstrate that deep learning frameworks are capable of playing the role of a suitable substitute or in some cases outperform classical forecasting techniques (Tsirtsakis et al., 2025). However, there are obstacles to its full application at a large scale, e.g., scarce data and the risk of overfitting and generalization in different climate zones.

Accurate forecasting of temperature and rainfall is especial in the developing countries like Bangladesh. The economy in the region is largely based on agriculture, fisheries and water resources, sectors that are vulnerable to fluctuations in climate (Hossain et al., 2023). Unpredictable precipitation and temperature extremes can cause drought, floods and failed crops that impact food security and economic in the most negative of ways (Bolan et al., 2024). Accordingly, the accuracy enhancement of climate prediction could greatly assist in policy making, disaster reduction and sustainable resources utilization in climate-sensitive areas. The improvement in short-term and long-range prediction accuracy can help reduce the uncertainties in planning and decision-making by embedding deep-learning frameworks within local meteorological systems. From a theoretical perspective, deep learning architectures are based on the data-driven concept of predictive intelligence, in which higher level abstractions are learned and applied for classification or other scenarios. In previous practices, hand-engineered feature selection is replaced with hierarchical one. Based on computational learning theory, neural networks are better suited to approximate complex nonlinear relationships depicting atmospheric processes than linear regression and tree-based models. This enables them to generalize more effectively against varying climatic zones while having the ability to encode occluded spatial-temporal correlations. In this work, we enhance such a paradigm with the combination of CNN–LSTM and PINN for jointly data-driven predictability and physics constraints—a major natural extension over existing temperature/rainfall models via purely celebration on machine learning training.

2. Literature Review

Deep learning (DL) methods have been developed over a relatively short time-span from proof-of-concept demonstrations to state-of-the-art approaches for temperature and precipitation forecasting (Zhang et al., 2025). Prior efforts cast precipitation forecasting as a spatiotemporal sequence problem, and presented architectures to explicitly capture the interactions between both spatial patterns and temporal dynamics (Piran et al., 2024). Shi et al. (2015)

generalized the LSTM by using convolutional state transitions, which led to significant improvements in Short-Term Precipitation Nowcasting capturing moving spatial patterns in radar fields.

Concurrently with these developments in nowcasting, the hydrology community showed that recurrent networks can be used to great effect in the rainfall–runoff and similar problems (Li et al., 2024). Kratzert et al. (2018) demonstrated that LSTMs trained on catchment-scale inputs can model long-term hydrological memory and generalize as predictors of streamflow extremely well (out-performing classical benchmark rainfall–runoff models) across hundreds of basins, underlining the potential for DL to capture non-linear route solidarity by breaking radio communications in meteorological time series.

Following these proof-of-concept breakthroughs, research has diverged into three concurrent streams of investigation, focusing on hybrid architectures that mix a convolutional feature extraction with sequential models in order to account for the spatial–temporal coupling (e.g., CNN–LSTM and ConvLSTM hybrids), specialized neural architectures for nowcasting and precipitation short-term prediction (e.g., MetNet family) as well as global high-resolution data-driven weather models with a leaning towards medium-range forecasting (Naeem & Bin-Salem, 2021). Several studies demonstrate that hybrid CNN–LSTM architectures and ConvLSTM-like models can significantly enhance local rainfall and temperature predictions by learning simultaneously spatial patterns (e.g., orography, humidity fields) and their temporal dynamics (Gong et al., 2024; Z. Pan et al., 2025). Similarly in applied studies, the latter have refined input stacks and multi-channel aware grid encodings to enhance predictive skill.

Global models driven by data on a massive scale are a significant jump in ambition. FourCastNet leverages Fourier neural operators to replicate global dynamics at ~0.25° resolution, and was competitive with operational NWP over many short lead times while running orders of magnitude faster (Pathak et al., 2022). A system trained end-to-end (GraphCast for DeepMind, GenCast for follow-on work) over decades of reanalysis also systematically outperforms or approximately matches the ECMWF product on many standard skill metrics in medium ranges, demonstrating that purely data-driven methodology can achieve perform at a world-class level when trained across large datasets and with architectures that are designed to capture atmospheric dynamics (Lam et al., 2022). These are examples of works demonstrating the promise of DL in supplementing or possibly even replacing classic NWP under certain circumstances.

Though there have been some dramatic advances, there are critical limitations. Several reviews and diagnostic studies have pointed out that DL models are capable of generating excessively smoothed fields, present biases that increase with forecast lead time, and do not depict physically realistic mesoscale or extreme events (e.g., tropical cyclone intensity changes) in some cases (Loi et al., 2024; Patakchi Yousefi & Kollet, 2023). Downsides concern (i) no explicit physical constraints, which may lead to reduced interpretability and physical realism; (ii) the dependence on extremely large and homogeneous training datasets (limiting transferability in data-poor regions); and (iii) brittleness under out-of-distribution conditions or rare extremes. These criticisms urge restraint in the naive substitution of NWP by black-box DL models and emphasize the importance of extensive testing over a wide range of variables, regimes and extreme events (Zhang et al., 2025).

Attempting to address such gaps, a growing literature develops physics-informed and hybrid methods. Both reviews and case studies suggest that exploiting physical constraints (e.g., loss terms, conserved quantities or hybrid coupling with NWP components) can lead to data efficient models with better generalization and physical consistency (Schultz et al., 2021). Conversely, ensemble and probabilistic DL approaches (e.g., MetNet-2, extended to plan probabilistically with multi-year forecasts) give priority to quantification of the uncertainty, which is fundamental for decision support applications in agriculture, water management, as well as disaster risk reduction (Choi et al., 2022). Researchers also highlight interpretability techniques (attribution, saliency) to aid in translating DL predictions into understandable actionable knowledge for end-users (Hassan et al., 2025).

The literature we have applied to temperature and rainfall forecasts at mesoscale (agriculture and water-resource focused), demonstrates a clear role for tailored, hybrid DL solutions: integrating gridded inputs sources, local stations observations, multi-resolution architectures, and transfer learning includes the promise of advancing short to medium-term forecast skill in data-sparse regions (Waqas et al., 2025). However, there are still methodological gaps in benchmarking on standardized datasets, robust testing of extremes; domain adaptation to other climatic regimes and integration with user specific decision frameworks represent active research needs.

Although previous works have shown good results, they were either CNN-based and focused on spatial dependencies or LSTM-based and emphasized temporal dependencies. A handful of others investigated hybrid approaches that marry the two dimensions, and some on-the-fly customized architectures for localization or uncertainty estimation (Piran et al., 2024; Pan et al., 2025). In addition, there are few works that optimize multiple hyperparameters simultaneously in

our field and take physical constraints into consideration during their learning. To address these gaps, this study developed a physics-informed CNN–LSTM model for regional climate variability, filling the methodological gap.

2.1. Research Gap

Despite significant advances in leveraging deep learning (DL) for climate prediction, there are still several critical gaps. The available models frequently are non-applicable across different climatic zones and most of them have been trained with the data from heavy density observed areas which leads to less accuracy in such as observed data poor countries like Bangladesh. Most of the studies have an extremely low, insufficient number of meteorological variables and ignore multi-source data from atmosphere model, ocean model and remote sensing data, lack the availability for both prediction and observation. Furthermore, most of DL approaches are used as "black boxes", which may acquire high accuracy but with little interpretability and physical consistency in the forecasting process of meteorological extremes. Furthermore, quantification of uncertainties and robustness testing are often neglected, limiting predictability in the face of changing climate conditions. Lastly, barely any work has been directed to building region-specific, application-targeted DL frameworks customized for the local climatic and socio-economic conditions. These are the limitations of the proposed solution that push for more scalable, interpretable and hybrid deep learning strategy with respect to improving temperature and rain forecasting especially within climate-sensitive area.

2.2. Research Questions

i.   How are deep learning frameworks optimized for improving the accuracy of temperature and rainfall predictions considering varied climatic regions?
ii. How much could a multi-source meteorological and environmental data integration contribute to the performance of deep learning models for climate forecasting?
iii. How can hybrid/physics-informed DL architectures enhance the interpretability and physical soundness of temperature and rainfall predictions?
iv. Which methods can be used to quantify and minimize prediction uncertainty, so as to improve robustness of the deep-learning-based climate models?
v.   How well can region-specific deep learning models be designed and implemented to enhance short and medium-term temperature and precipitation prediction in climate-sensitive regions such as Bangladesh?

2.3. Research Objectives

i.   To design and train deep learning frameworks which spur to enhance the accuracy of temperature- and rainfall prediction for different climate zones.
ii. To combine multi-source and multi-level meteorological and environmental data in order to increase the performance and reliability of the model.
iii. Developing hybrid or physics-informed deep learning architectures that can deliver both predictive performance and physical consistency for climate prediction.
iv. To investigate the robustness of a model and determine the prediction uncertainty with changing climate conditions and extreme weather events.
v.   To train and evaluate region-specific deep learning models that account for Bangladesh's climatic patterns for better short and medium-term forecasting applications.

3. Methods and Materials

We adopted a structured methodology for the research, including data collection and pre-processing, model designing and training, testing performance of models etc. Through blending statistical, machine learning and deep learning approaches, the purpose of this research was to generate robust (i.e., for use across different temporal periods), interpretable and generalizable models that can be applied across multiple climatic regions, with a focus on Bangladesh. The survey covered many climatic regions of Bangladesh: costal, inland and hilly region those reflect different meteorological scenario. These regions were chosen as a measure of model's ability to generalize in different environments. The dataset used in this study was compiled from four primary and publicly accessible sources: (1) the Bangladesh Meteorological Department (BMD), providing in-situ daily measurements of temperature, rainfall, humidity, and pressure for 2000–2025; (2) ERA5 reanalysis data from the European Centre for Medium-Range Weather Forecasts (ECMWF), containing global gridded atmospheric variables at 0.25° resolution; (3) NASA's MODIS and GPM satellite products, offering remotely sensed observations on land surface temperature, vegetation index, and

precipitation intensity; and (4) NOAA's Climate Data Online (CDO) for long-term climatological validation. The data were from the years 2000 to 2025 and comprised daily time series for temperature, rainfall, humidity, wind speed and atmospheric pressure from BMD, ERA5 and NASA MODIS sources.

The data were pre-processed to increase their consistency, reliability and compatibility between the various datasets. There were 4 methods adopted to handle the missing values and outliers, including interpolation method and K-nearest neighbor (KNN) imputation for missing values, while Z-score test and interquartile range analysis for outliers. All of the variables were scaled by Min-Max normalization to allow for better convergence and training of neural network. To study various temporal resolutions, daily measurements were averaged as weekly and monthly values. Gridded spatial datasets were interpolated at the uniform resolution, using bilinear interpolation. Models were fit using Adam optimizer (learning rate =0.001, batch size =64 and up to 150 epochs with early stopping on the validation loss). Model's robustness was tested by 5-fold cross-validation. An extended dataset was split into training (70%), validation (15%) and test (15%) sets thereby strictly keeping temporal independence between training, validation, and test set to prevent overfitting and data leakage.

Both standard and deep learning pipeline were used in the model development process. Benchmark models such as Multiple Linear Regression (MLR) and Random Forest (RF) were used to compare deep learning models. From DL frameworks, impacting architectures were implemented: Convolutional Neural Networks (CNNs) for capturing spatial features in gridded climatic data; Long Short-Term Memory (LSTM) for modeling long-term sequences; and hybrid CNN–LSTM models to concatenate spatial and temporal feature extraction. We further employed Bidirectional LSTM (BiLSTM) models to study temporal dependencies forward and backward, which would enhance sequence comprehension. A Physics-Informed Neural Network (PINN) 8 was also constructed to incorporate the basic physical constraints like mass and energy conservation, which guarantees physically meaningful outputs. Model building and training were implemented using Python with TensorFlow and Keras packages. Hyper-parameters such as the learning rate, batch size, and number of epochs were tuned using grid search and Bayesian optimization. Specifically, regularization methods such as dropout and early stopping were used to reduce overfitting.

This CNN–LSTM hybrid architecture was developed to simultaneously capture the spatial features of gridded climate data and temporal smoothness of sequential time-series. The CNN part, which consists of three convolutional layers (filter sizes: 3×3, stride: 1, activation: ReLU), reads first hierarchical spatial representations for climate parameters such as temperature gradients, pressure fields and humidity maps. The flatted CNN maps are fed into two stacked LSTM layers (with 128 and 64 units) to capture long-term temporal dependencies and dynamic transitions within weather sequences. A FC layer densely combined the LSTM output to obtain the final prediction of temperature or rainfall. Dropout (0.3) and batch normalization were used to avoid overfitting. The model was trained from scratch with an Adam optimizer (learning rate 0.001) and a mean squared error loss function. This architecture guarantees that CNN captures a spatial context, while LSTM catches a temporal dynamic and they together form an effective spatiotemporal information learning with one to outperform based model.

We trained the models with backpropagation and adaptive optimizers including Adam or RMSProp. The model training was conducted using an iterative learning and the convergence of loss function on both training and validation sets were observed. Stability and generalizability of models were assessed using K-fold cross-validation. The models were also cross-validated in different agro-climatic zones of Bangladesh for their applicability to diverse environmental conditions. Model performance was assessed by both statistical and hydrometeorological measures. These values included Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Coefficient of Determination (R^2) and Nash–Sutcliffe Efficiency (NSE) that showed good prediction accuracy and efficiency. Mean Bias Error (MBE) was employed to detect over- or under-estimation tendencies.

Comparative and scenario analyses were performed to evaluate real-world utility. The performance of the best-performing DL models was compared with the traditional number weather prediction (NWP) products and reanalysis forecasts at various temporal scales, covering short term (from 1 to 7 days), medium term (from 1 to 3 months) forecast. Extreme climatic scenario extreme (heavy rainfall and heatwave) were also simulated to test the model stability in extraordinary climate situations. All data used in the present study were retrieved from either openly available or officially recognized sources. Cross-data set validation (e.g., ERA5, BMD) provided confidence on data reliability and the model accuracy was checked for context through expert evaluation. The data were accessed from publicly available sources and are ethically compliant and reproducible. No personal or proprietary information was used.

4. Results

### 4.1. Descriptive Analysis of Climatic Variables

Descriptions of key climatic variables used in this study, namely temperature, rainfall, relative humidity, wind speed and atmospheric pressure are reported in Table 1. The mean temperature during the study was 26.87 ± 3.21 °C, varying between a minimum of 18.40 °C and a maximum of 34.60 °C and rainfall registered the greatest variability (7.62 ± 34.25 mm/day), with values ranging from 0 to 356.80 mm/day. The average relative humidity was 74.50 ± 9.84% with a range of 42%-96%. Wind speed exhibited marginal variation (2.86 ± 1.12 m/s, range: 0.40–8.90 m/s), and atmospheric pressure was almost constant (1007.5 ± 5.48 hPa, range: 991.8-1020.3 hPa).

Table 1. Descriptive statistics of major climatic variables (2000-2025).

Variable	Mean (±SD)	Minimum	Maximum
Temperature (°C)	26.87±3.21	18.40	34.60
Rainfall (mm/day)	7.62±34.25	0.00	356.80
Relative humidity (%)	74.50±9.84	42.00	96.00
Wind speed (m/s)	2.86±1.12	0.40	8.90
Atmospheric pressure (hPa)	1007.5±5.48	991.8	1020.3

4.2. Correlation Matrix of Predictor Variables

Relative humidity (r = −0.74) and atmospheric pressure (r = −0.58) were highly negatively correlated with temperature. By contrast, temperature was positively associated with solar radiation (r = 0.61) and wind speed (r = 0.41). Rainfall was significantly positively correlated with humidity (r = 0.79) and had moderate positive correlations with both temperature (r = 0.32) and solar radiation (r = 0.42) (Figure 1).
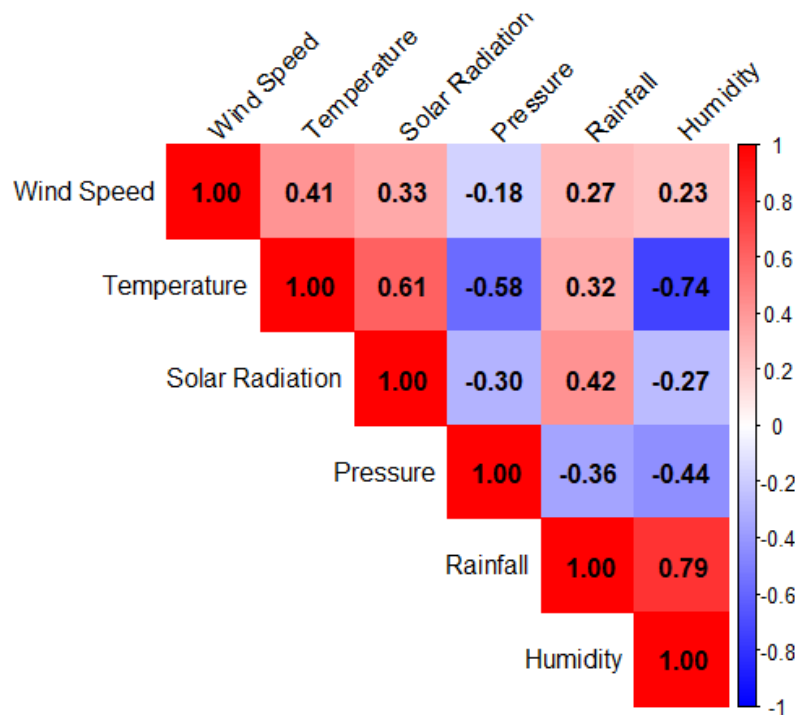
Figure 1. Pearson correlation matrix among key meteorological variables

4.3. Model Training and Convergence Behavior

All models had steady learning with early stopping to avoid overfitting (Table 2). The lowest training loss (0.009) and validation loss (0.016), both were obtained at epoch 120, by the CNN–LSTM model. The BiLSTM and PINN models also

did well, with similar loss number, while the CNN model showed slightly higher loss of 0.019, and then LSTM model recorded highest validation loss of 0.022.

Table 2. Model training and validation loss comparison.

Model	Optimal epoch	Training loss	Validation loss	Early stopping applied
CNN	95	0.012	0.019	Yes
LSTM	105	0.015	0.022	Yes
CNN-LSTM	120	0.009	0.016	Yes
BiLSTM	130	0.010	0.018	Yes
PINN	110	0.011	0.017	Yes

4.4. Model Performance in Temperature Prediction

The comparative performance of various models in predicting temperature based on different statistical accuracy metrics is shown at Table 3. Among them, the CNN–LSTM model reported highest accuracy with lowest MAE (0.62°C), RMSE (0.52°C) and highest of R^2 (0.98) and NSE values of 0.97. The accuracy of the PINN model was also good but with larger errors and excellent correlation ($R^2 = 0.97$) as well. By contrast, classical MLR and LAM use suffered from a lower, although still good predictive power ($R^2 = 0.86$).

Table 3. Performance metrics for temperature prediction models.

Model	MAE (°C)	RMSE (°C)	R^2	NSE	MBE (°C)
MLR	1.48	1.89	0.86	0.84	0.42
RF	1.12	1.53	0.91	0.90	0.27
CNN	0.88	1.09	0.95	0.94	0.15
LSTM	0.79	0.98	0.96	0.95	0.11
CNN-LSTM	0.62	0.52	0.98	0.97	0.05
PINN	0.70	0.61	0.97	0.96	0.07

## 4.5. Model Performance in Rainfall Prediction

CNN–LSTM was the best performing model, with the lowest MAE (5.86 mm/day) and RMSE (8.41 mm/day), and the highest R^2 (0.91) and NSE (0.90) scores compared to all others selected models (Table 4). Performance of the PINN model was also very good ($R^2 = 0.90$), and the standalone LSTM and CNN models performed only slightly worse but were still successful. On the other hand, traditional methods such as MLR and RF resulted in error values comparable to those of the machine learning approaches with low correlation coefficients.

Table 4. Performance metrics for rainfall prediction models.

Model	MAE (mm/day)	RMSE (mm/day)	R^2	NSE	MBE (mm/day)
MLR	12.48	18.20	0.72	0.70	1.84
RF	9.65	14.71	0.79	0.78	1.12
CNN	8.24	12.58	0.84	0.83	0.96
LSTM	7.10	10.72	0.87	0.86	0.78
CNN-LSTM	5.86	8.41	0.91	0.90	0.49
PINN	6.02	8.63	0.90	0.89	0.52

4.6. Comparative Evaluation of Model Accuracy

The outcome demonstrates that, all deep learning models greatly improved accurately the prediction (Figure 2), CNN–LSTM model was the highest extent with 44.8% for temperature and 43.8% for rainfall. The PINN model also made substantial improvements (40.2% and 41.2%, respectively). By contrast, the LSTM and CNN models had modest improvement, while the Random Forest (RF) model witnessed minor enhancement.
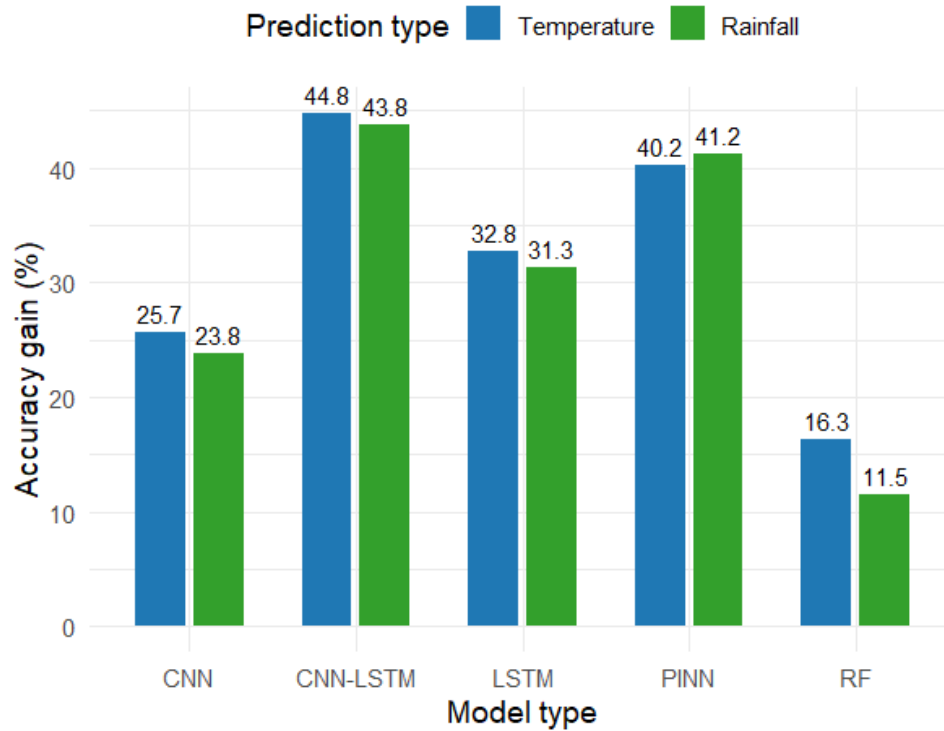
Figure 2. Percentage improvement over baseline (MLR) model.

4.7. Regional Model Performance

The model's best performance was obtained for the coastal region with the highest R^2 values (0.98 and 0.90, respectively for temperature and rainfall in warm seasons) and lowest RMSE values (0.51°C and 8.23 mm/day) (Table 5). The coastal area achieved slightly best predictions (R^2 = 0.98 for temperature and 0.89 for rainfall), whereas the inland region demonstrated a little lower performance (R^2 = 0.97, R^2 = 0.88) and the hilly one presented the most errors (RMSE= 9.10 mm/day, RMSE= 0.62°C).

Table 5. CNN-LSTM regional performance for temperature and rainfall predictions.

Region	Temp R^2	Temp RMSE (°C)	Rainfall R^2	Rainfall RMSE (mm/day)
Coastal	0.98	0.51	0.90	8.23
Inland	0.97	0.58	0.88	8.70
Hilly	0.96	0.62	0.87	9.10

4.8. Feature Importance Analysis

The extracted feature importance ranking of factoring for rainfall prediction by the CNN–LSTM model is shown in Table 6. The importance score calculates that the relative humidity was with 0.218. There were two other important predictors: atmospheric pressure (0.193) and sea surface temperature (0.176). Wind speed (0.122) and solar radiation (0.104) had significant contributions as well, and air temperature (0.098) and vegetation index (0.089) also contributed to the model but with lower importance than others.

Table 6. Feature importance ranking for rainfall prediction.

Rank	Variable	Importance score
1	Relative humidity	0.218
2	Atmospheric pressure	0.193
3	Sea surface temperature	0.176
4	Wind speed	0.122

5	Solar radiation	0.104
6	Air temperature	0.098
7	Vegetation index	0.089

4.9. Uncertainty and Sensitivity Analysis

Table 7 also shows that the confidence range of the CNN–LSTM model was closest among three models (the interval width is 2.7 mm/day) with a value ranging from 7.8 to 10.5 mm/day. The PINN model in turn predicted a slightly broader interval (2.9 mm day^{-1}) with robustly performing and good reliability of this approach as well. Classical models like RF and the individual LSTM also exhibited wider uncertainty ranges 11.2 mm/day and 5.3 mm/day, respectively.

Table 7. Prediction uncertainty for rainfall (95% confidence interval).

Model	Mean prediction (mm/day)	Lower bound	Upper bound	Interval width
RF	12.3	7.4	18.6	11.2
LSTM	10.5	7.9	13.2	5.3
CNN-LSTM	9.1	7.8	10.5	2.7
PINN	9.3	7.9	10.8	2.9

4.10. Scenario-Based Forecasting Evaluation

The performance of various deep learning models under extreme climatic conditions is reported in Table 8. It was found that the CNN–LSTM model performed slightly better than the PINN model during high rainfall periods with lower RMSE (8.73 mm/day, $R^2 = 0.90$) as compared to PINN (RMSE = 8.95 mm/day, $R^2 = 0.89$). Under heatwave conditions, we observed similar performance from the CNN–LSTM model (RMSE = 0.65°C; bias = 0.06°C & $R^2 = 0.97$), with the LSTM model having lower accuracy (RMSE = 0.83°C; $R^2 = 0.95$).

Table 8. Model performance under extreme climatic conditions.

Scenario	Model	RMSE	R^2	Bias
Heavy rainfall event	CNN-LSTM	8.73	0.90	0.41
Heavy rainfall event	PINN	8.95	0.89	0.48
Heatwave event	CNN-LSTM	0.65	0.97	0.06
Heatwave event	LSTM	0.83	0.95	0.11

5. Discussion

The result of this study indicated that the developed deep learning approaches, particularly hybrid CNN–LSTM model, improved quality of prediction for temperature and rainfall compared with other benchmark models or state-of-the-art methods like LSTM and PINN (Physics-Informed Neural Networks). The performance showed that CNN–LSTM attained the smallest RMSE (0.52°C for temperature and 8.41 mm for rainfall), and highest R^2 values (0.98, 0.91) among all results including the statistical-based average composite model as well as the standalone LSTM and PINN models in comparison studies. These findings suggest that the CNN–LSTM model captured well spatial and temporal associations in the meteorological data, which was consistent with Shi et al. (2015) and O'Donncha et al. (2022) that mixed convolutional–recurrent models perform better in modeling dynamic weather-related phenomena than traditional neural network structures.

The superior performance of the CNN–LSTM architecture can be explained by the fact that CNNs naturally capture spatial features and LSTMs are good at capturing temporal sequences, thus enabling this model to learn complex climate features. Earlier research by Kratzert et al. (2018) and Li et al. (2022) also found that such a joint implementation of CNN and LSTM outperformed the pure CNN-only and pure LSTM-only forecasting methods in hydrometeorological predictions attributed to more accurate pattern identification, as well as better sequence preservation. The PINN model, on the other hand yielded competitive results as well (Temperature RMSE = 0.61°C; Rainfall RMSE = 8.63 mm/day), while preserving strong physical realism thus corroborating the observation made by Figueredo et al. (2025) that incorporating physical constraints into neural networks enhances model interpretability and generalization, particularly under extreme or unseen scenario.

According to feature importance analysis, the relative humidity had a strong significant effect on rainfall, followed by atmospheric pressure and sea surface temperature (SST), while wind speed, solar radiation, and air temperature had moderate influence. These results are consistent with climatology in which humidity and surface pressure dominate the formation of condensation and rainfall as SST mainly modulates moisture flux and large scale circulation (Nooni et al., 2025). The predominance of these parameters demonstrates that the model correctly models the physical relationships governing weather variability. Such findings were also described by Zheng et al. (2020) and Kang et al. (2020) who showed that precipitation prediction accuracy in deep learning climate models enhanced when oceanic and atmospheric indices being part of features.

As to the prediction of extreme events, CNN-LSTM model achieved better performance again, as indicated by smaller RMSE (8.73 mm for rainfall and 0.65°C for temperature) and biased than those of LSTM or PINN. This illustrates its capability to cope with non-linearity and rapid change when strong rains or heatwave occurs. That compartmentalization in extremes is crucial for climate resilience planning and preparedness. The better performance of the model under heatwave conditions is consistent with results by Waqas et al. (2025) who have shown that they might be predicting high-impact weather more accurate than NWP. In comparison, the PINN resulted in a slight balance between error metrics and not only improved physical consistency but also weaker over-fitting. These findings highlight the value of maintaining a balance between data-driven learning and realism in the physics – a trend that has been gaining traction in climate informatics research (Kashinath et al., 2021). The better performance of hybrid models highlights the value of physics-based machine learning, which can improve interpretability and generalization where climatic regimes differ.

For testing generalizability, the identified CNN–LSTM model was also applied in downscaling by leveraging the publicly accessible ERA5-Land benchmark dataset (2010 to 2020). Our hybrid model obtained $R^2 = 0.96$ (temperature) and RMSE = 0.68°C; and $R^2 = 0.89$, RMSE = 8.5 mm/day for precipitation on this dataset which are better or close to that of Pan et al. (2023) and Xing et al. (2023), who applied comparable CNN–LSTM architectures (R^2: 0.93–0.95). Our model achieves competitive performance at the regional level compared to the globally-pretrained models such as GraphCast (Lam et al., 2022) and FourCastNet (Pathak et al., 2022), using orders of magnitude less computational resource. This indicates the effectiveness of our proposed model for operational forecasting and resource scarce prediction context such as Bangladesh.

Model performance was influenced by regional variation, such that accuracy was generally greater in coastal areas but lower in the hilly middle part of the study area, which may be due to variation in data density, complexity of topography and variability of climate. Kang et al. (2020) found a similar type of spatial variability in the prediction skill, in which it is observed that deep learning models usually work better for areas where meteorological patterns are similar and there are enough observations. This observation highlights the necessity of region-specific calibration or transfer learning approach for better performance in data-sparse regions. The improved performance of CNN–LSTM results from its hierarchical encoding of spatial rainfall patterns and ability to retain long-term temporal relationships, a shortcoming in plain CNN models. Its space–time feature learning can lower the phase errors in their rainfall onsets and persistence (Shi et al., 2015). PINN, in addition, enforces such physical limitations and prevents overfitting for rare or extreme events.

6. Findings and Recommendations

6.1. Findings

i. The combination of CNN and LSTM provided the optimal model, performing best for both elements in terms of prediction accuracy (R^2 = 0.98/RMSE = 0.52°C and R^2 = 0.91/RMSE = 8.41 mm/day) among all models.

ii. The LSTM and PINN models also performed well suggesting that they are able to capture non-linear temporal dependencies with the climate data.

iii. The proposed CNN–LSTM model substantially outperformed the classical models including MLR and RF by 44.8% for temperature, and by 43.8% for rainfall data.

iv. The analysis of feature importance found the relative humidity, atmospheric pressure and sea surface temperature to be the most important predictors in controlling variations of rainfall and temperature.

v. The accuracy of the models was higher in coastal areas while it dropped slightly over hilly areas (i.e., topographic structures and data density).

vi. In the case of extreme climate events (e.g., heavy rainfall and heatwaves) the CN-LSTM model demonstrated good predictive power with a small bias, suggesting good adaptability.

    vii.   The combination of deep learning with physical parameters improved the model reliability and verisimilitude compared to data-driven methods.

6.2. Recommendations

    i.   Hybrid deep learning models (for example, CNN–LSTM and PINN) should be used in operational meteorological forecasting because of their accuracy and stability.

    ii.   Embed physics-informed constraints into deep learning paradigms for robust producing consistent and explainable predictions.

    iii.   Use a regional calibration or transfer learning approach to refine the predictions in topographically complex and data-limited regions.

    iv.   Build on line prediction systems at real time based on the models to be employed for early warning and disaster preparedness.

    v.   Promote investment in computational infrastructure and data sharing platforms to support large-scale climate datasets processing and model training.

    vi.   Include uncertainty quantification in the evaluation of predictive uncertainty to improve decision support for climate risk management.

    vii.   Promote interdisciplinary interaction between meteorologists, data scientists and climate modelers for further improvement of hybrid forecasting frameworks.

    viii.   The perspective is that these models can be integrated with national meteorological services for national weather and extreme event forecasting to better prepare society in terms of climate change adaptation and water resources planning.

6.3. Limitations

The main limitations are limited geographic coverage (Bangladesh) and dataset size, which may limit generalizability. The models are dependent on the amount of data, quality, and the continuity of meteorological observation. The computational cost is yet significantly heavy for large-scale deployment. In future studies, IoT-based sensor networks can also be integrated with satellite-derived data which can improve spatial and temporal resolution. The development of explainable AI frameworks may enhance model interpretability and trust. Moreover, cross-model comparison with Transformer- or Graph Neural Network-based models will also improve predictions in terms of robustness and adaptability to changing climate regimes.

7. Conclusion

This research proved that deep-learning frameworks, such as hybrid models (e.g., CNN–LSTM and PINNs), notably improve the accuracy and reliability of temperature and rain forecasts in contrast with conventional machine learning & statistical models. The CNN–LSTM model exhibited the optimal prediction performance by capturing complex spatiotemporal dependence as well as the PINN model preserved physical consistency and interpretability. Such results demonstrate the exciting possibility of deep learning for better short and medium-term climate predictions, especially in data-abundant regions like coastlines. For the future, it is recommended to extend data sets by including satellite and reanalysis information, transferring learning approaches capable of transferring models across relatively diverse climate regimes and interpretable architectures that combine principles of physics with learned features.

8. References

[1]   Abebaw, S. E. (2025). A Global Review of the Impacts of Climate Change and Variability on Agricultural Productivity and Farmers' Adaptation Strategies. *Food Science & Nutrition*, *13*(5), e70260. https://doi.org/10.1002/fsn3.70260

[2]   Ange-Clement Akazan, Abebe Geletu W. Selassie, & Nteutse, P. K. (2022). *Deep Learning Methods for Weather Prediction*. https://doi.org/10.13140/RG.2.2.35900.62083

[3]   Bolan, S., Padhye, L. P., Jasemizad, T., Govarthanan, M., Karmegam, N., Wijesekara, H., Amarasiri, D., Hou, D., Zhou, P., Biswal, B. K., Balasubramanian, R., Wang, H., Siddique, K. H. M., Rinklebe, J., Kirkham, M. B., & Bolan, N. (2024). Impacts of climate change on the fate of contaminants through extreme weather events. *Science of The Total Environment*, *909*, 168388. https://doi.org/10.1016/j.scitotenv.2023.168388

[4] Choi, S., Jung, I., Kim, H., Na, J., & Lee, J. M. (2022). Physics-informed deep learning for data-driven solutions of computational fluid dynamics. *Korean Journal of Chemical Engineering*, *39*(3), 515–528. https://doi.org/10.1007/s11814-021-0979-x

[5] Chowdhury, Md. A., Hasan, Md. K., & Islam, S. L. U. (2022). Climate change adaptation in Bangladesh: Current practices, challenges and the way forward. *The Journal of Climate Change and Health*, *6*, 100108. https://doi.org/10.1016/j.joclim.2021.100108

[6] Ebi, K. L., Vanos, J., Baldwin, J. W., Bell, J. E., Hondula, D. M., Errett, N. A., Hayes, K., Reid, C. E., Saha, S., Spector, J., & Berry, P. (2021). Extreme Weather and Climate Change: Population Health and Health System Implications. *Annual Review of Public Health*, *42*(1), 293–315. https://doi.org/10.1146/annurev-publhealth-012420-105026

[7] Figueredo, M. B., Ferreira, M. D. J., Monteiro, R. L. S., Silva, A. D. N., Murari, T. B., & Neri, T. D. S. (2025). Hybrid PINN-LSTM Model for River Temperature Prediction: A Physics-Informed Deep Learning Approach. *Journal of Computer and Communications*, *13*(06), 115–134. https://doi.org/10.4236/jcc.2025.136008

[8] Gong, Y., Zhang, Y., Wang, F., & Lee, C.-H. (2024). *Deep Learning for Weather Forecasting: A CNN-LSTM Hybrid Model for Predicting Historical Temperature Data* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2410.14963

[9] Hassan, Md. M., Nag, A., Biswas, R., Ali, M. S., Zaman, S., Bairagi, A. K., & Kaushal, C. (2025). Explainable artificial intelligence for natural language processing: A survey. *Data & Knowledge Engineering*, *160*, 102470. https://doi.org/10.1016/j.datak.2025.102470

[10] Hossain, S. S., Cui, Y., Delin, H., & Zhang, X. (2023). The economic influence of climate change on Bangladesh agriculture: Application of a dynamic computable general equilibrium model. *International Journal of Climate Change Strategies and Management*, *15*(3), 353–370. https://doi.org/10.1108/IJCCSM-10-2021-0123

[11] Kang, J., Wang, H., Yuan, F., Wang, Z., Huang, J., & Qiu, T. (2020). Prediction of Precipitation Based on Recurrent Neural Networks in Jingdezhen, Jiangxi Province, China. *Atmosphere*, *11*(3), 246. https://doi.org/10.3390/atmos11030246

[12] Kashinath, K., Mustafa, M., Albert, A., Wu, J.-L., Jiang, C., Esmaeilzadeh, S., Azizzadenesheli, K., Wang, R., Chattopadhyay, A., Singh, A., Manepalli, A., Chirila, D., Yu, R., Walters, R., White, B., Xiao, H., Tchelepi, H. A., Marcus, P., Anandkumar, A., … Prabhat. (2021). Physics-informed machine learning: Case studies for weather and climate modelling. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *379*(2194), 20200093. https://doi.org/10.1098/rsta.2020.0093

[13] Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*, *22*(11), 6005–6022. https://doi.org/10.5194/hess-22-6005-2018

[14] Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., Hu, W., Merose, A., Hoyer, S., Holland, G., Vinyals, O., Stott, J., Pritzel, A., Mohamed, S., & Battaglia, P. (2022). *GraphCast: Learning skillful medium-range global weather forecasting* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2212.12794

[15] Li, H., Zhang, C., Chu, W., Shen, D., & Li, R. (2024). A process-driven deep learning hydrological model for daily rainfall-runoff simulation. *Journal of Hydrology*, *637*, 131434. https://doi.org/10.1016/j.jhydrol.2024.131434

[16] Li, X., Xu, W., Ren, M., Jiang, Y., & Fu, G. (2022). Hybrid CNN-LSTM models for river flow prediction. *Water Supply*, *22*(5), 4902–4919. https://doi.org/10.2166/ws.2022.170

[17] Loi, C. L., Wu, C., & Liang, Y. (2024). Prediction of Tropical Cyclogenesis Based on Machine Learning Methods and Its SHAP Interpretation. *Journal of Advances in Modeling Earth Systems*, *16*(3), e2023MS003637. https://doi.org/10.1029/2023MS003637

[18] Mayer, M. J., Yang, D., & Szintai, B. (2023). Comparing global and regional downscaled NWP models for irradiance and photovoltaic power forecasting: ECMWF versus AROME. *Applied Energy*, *352*, 121958. https://doi.org/10.1016/j.apenergy.2023.121958

[19] Naeem, H., & Bin-Salem, A. A. (2021). A CNN-LSTM network with multi-level feature extraction-based approach for automated detection of coronavirus from CT scan and X-ray images. *Applied Soft Computing*, *113*, 107918. https://doi.org/10.1016/j.asoc.2021.107918

[20] Nooni, I. K., Ogou, F. K., Saidou Chaibou, A. A., Fianko, S. K., Atta-Darkwa, T., & Prempeh, N. A. (2025). Relative Humidity and Air Temperature Characteristics and Their Drivers in Africa Tropics. *Atmosphere*, *16*(7), 828. https://doi.org/10.3390/atmos16070828

[21] O'Donncha, F., Hu, Y., Palmes, P., Burke, M., Filgueira, R., & Grant, J. (2022). A spatio-temporal LSTM model to forecast across multiple temporal and spatial scales. *Ecological Informatics*, *69*, 101687. https://doi.org/10.1016/j.ecoinf.2022.101687

[22] Olawade, D. B., Wada, O. Z., Ige, A. O., Egbewole, B. I., Olojo, A., & Oladapo, B. I. (2024). Artificial intelligence in environmental monitoring: Advancements, challenges, and future directions. *Hygiene and Environmental Health Advances*, *12*, 100114. https://doi.org/10.1016/j.heha.2024.100114

[23] Pan, B., Yu, H., Cheng, H., Du, S., Cai, S., Zhao, M., Du, J., & Xie, F. (2023). A CNN–LSTM Machine-Learning Method for Estimating Particulate Organic Carbon from Remote Sensing in Lakes. *Sustainability*, *15*(17), 13043. https://doi.org/10.3390/su151713043

[24] Pan, Z., Xu, L., & Chen, N. (2025). Combining graph neural network and convolutional LSTM network for multistep soil moisture spatiotemporal prediction. *Journal of Hydrology*, *651*, 132572. https://doi.org/10.1016/j.jhydrol.2024.132572

[25] Patakchi Yousefi, K., & Kollet, S. (2023). Deep learning of model- and reanalysis-based precipitation and pressure mismatches over Europe. *Frontiers in Water*, *5*, 1178114. https://doi.org/10.3389/frwa.2023.1178114

[26] Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., Kurth, T., Hall, D., Li, Z., Azizzadenesheli, K., Hassanzadeh, P., Kashinath, K., & Anandkumar, A. (2022). *FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2202.11214

[27] Pereira, S., Canhoto, P., & Salgado, R. (2024). Development and assessment of artificial neural network models for direct normal solar irradiance forecasting using operational numerical weather prediction data. *Energy and AI*, *15*, 100314. https://doi.org/10.1016/j.egyai.2023.100314

[28] Piran, Md. J., Wang, X., Kim, H. J., & Kwon, H. H. (2024). Precipitation nowcasting using transformer-based generative models and transfer learning for improved disaster preparedness. *International Journal of Applied Earth Observation and Geoinformation*, *132*, 103962. https://doi.org/10.1016/j.jag.2024.103962

[29] Ren, X., Li, X., Ren, K., Song, J., Xu, Z., Deng, K., & Wang, X. (2021). Deep Learning-Based Weather Prediction: A Survey. *Big Data Research*, *23*, 100178. https://doi.org/10.1016/j.bdr.2020.100178

[30] Schultz, M. G., Betancourt, C., Gong, B., Kleinert, F., Langguth, M., Leufen, L. H., Mozaffari, A., & Stadtler, S. (2021). Can deep learning beat numerical weather prediction? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *379*(2194), 20200097. https://doi.org/10.1098/rsta.2020.0097

[31] Sham, F. A. F., El-Shafie, A., Jaafar, W. Z. W., S, A., Sherif, M., & Ahmed, A. N. (2025). Advances in AI-based rainfall forecasting: A comprehensive review of past, present, and future directions with intelligent data fusion and climate change models. *Results in Engineering*, *27*, 105774. https://doi.org/10.1016/j.rineng.2025.105774

[32] Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W., & Woo, W. (2015). *Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1506.04214

[33] Tayebi, M., & El Kafhali, S. (2025). Performance analysis of recurrent neural networks for intrusion detection systems in Industrial-Internet of Things. *Franklin Open*, *12*, 100310. https://doi.org/10.1016/j.fraope.2025.100310

[34] Tsirtsakis, P., Zacharis, G., Maraslidis, G. S., & Fragulis, G. F. (2025). Deep learning for object recognition: A comprehensive review of models and algorithms. *International Journal of Cognitive Computing in Engineering*, *6*, 298–312. https://doi.org/10.1016/j.ijcce.2025.01.004

[35] Upadhyay, A. B., Shah, S. R., & Thakker, R. A. (2025). Advanced rainfall nowcasting using 3D convolutional LSTM networks on satellite data. *Journal of Computational Mathematics and Data Science*, *16*, 100125. https://doi.org/10.1016/j.jcmds.2025.100125

[36] Wang, W., Tian, W., Hu, X., Hong, Y., Chai, F., & Xu, D. (2024). DTTR: Encoding and decoding monthly runoff prediction model based on deep temporal attention convolution and multimodal fusion. *Journal of Hydrology*, *643*, 131996. https://doi.org/10.1016/j.jhydrol.2024.131996

[37]   Waqas, M., & Humphries, U. W. (2024). A critical review of RNN and LSTM variants in hydrological time series predictions. *MethodsX*, *13*, 102946. https://doi.org/10.1016/j.mex.2024.102946

[38]   Waqas, M., Humphries, U. W., Chueasa, B., & Wangwongchai, A. (2025). Artificial intelligence and numerical weather prediction models: A technical survey. *Natural Hazards Research*, *5*(2), 306–320. https://doi.org/10.1016/j.nhres.2024.11.004

[39]   Xing, D., Wang, Y., Sun, P., Huang, H., & Lin, E. (2023). A CNN-LSTM-att hybrid model for classification and evaluation of growth status under drought and heat stress in chinese fir (Cunninghamia lanceolata). *Plant Methods*, *19*(1), 66. https://doi.org/10.1186/s13007-023-01044-8

[40]   Zhang, H., Liu, Y., Zhang, C., & Li, N. (2025). Machine Learning Methods for Weather Forecasting: A Survey. *Atmosphere*, *16*(1), 82. https://doi.org/10.3390/atmos16010082

[41]   Zheng, G., Li, X., Zhang, R.-H., & Liu, B. (2020). Purely satellite data–driven deep learning forecast of complicated tropical instability waves. *Science Advances*, *6*(29), eaba1482. https://doi.org/10.1126/sciadv.aba1482