International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue 5
Sep-Oct-2026, PP 01-09

Deepecopredict: A Hybrid Cnn–Lstm Framework for Climate Pattern Forecasting And Environmental Risk Assessment

Mrs. J. Josephine[1], Vidhya K[2]
Computer Science And Engineering
Electrical And Electronics Engineering
[1]Assistant Professor Gkm College Of Engineering And Technology
Gkm Nagar, Perungalathur Chennai - 600063
Josephinejosephlawrence@Gmail.Com
[2]Assistant Professor Govt.Model Engineering College, Thrikkakara, P.O
Ernakulam Dist.
Pin: 682021
Vidhkrish74@Gmail.Com

Abstract- Prediction of climate trends and the environmental risks that come along with them is a significant challenge considering the spatiotemporally dependent nature of meteorology systems. In this research paper, we introduce DeepEcoPredict, a deep learning framework which combines convolutional neural networks (CNN) for spatial pattern extraction with long short-term memory networks (LSTM) for time series analysis. The framework analyzes multivariate climate data such as temperature, rainfall, humidity, and solar irradiance by means of a parallel pipeline of CNN and LSTM, combining the features learned by both networks for prediction and risk classification purposes. To validate the effectiveness of the proposed solution, we conduct experiments of predicting climate anomalies with the help of DeepEcoPredict against CNN, LSTM, and classical machine learning approaches using a multi-decade regional climate dataset. It has been found out that the hybrid framework provides better prediction results compared to standalone CNN, LSTM, and machine learning models in terms of the prediction errors. In addition, DeepEcoPredict predicts composite risk indicators corresponding to certain levels of environmental vulnerability to climate anomalies.

Key Word: Climate Forecasting, CNN-LSTM, Deep Learning, Environmental Risk Assessment, Spatiotemporal Modeling, Hybrid Neural Networks

I. Introduction

With the increasing frequency and magnitude of climate hazard occurrences like prolonged droughts, extreme precipitation, heat waves, and their complex interactions, there is a rising threat to the ecosystem, infrastructure, and the well-being of humans. Classical numerical weather prediction models and statistical downscaling techniques are limited in capturing the nonlinearity of the spatiotemporal dependency of regional climate variables related to environmental risk assessment. However, the advent of deep learning as a complementary approach in climate science has led to novel techniques for climate modeling especially for applications that involve vast amounts of observational and reanalysis datasets [7][9].

The strengths of CNN models are in identifying hierarchical patterns within gridded climate field data like frontogenesis, orography, and land-atmosphere interaction. On the other hand, LSTM networks capture the temporal dependency over a long period of time and hence ideal for anomaly detection problems. Based on the unique strengths of both CNN and LSTM architectures, it became necessary to develop hybrid CNN-LSTM models by applying convolutional processing on input data for abstraction followed by temporal processing using recurrent layers [3][8].

Despite their encouraging performance, several shortcomings have been noted in applying CNN-LSTM hybrids to climate risk assessment. First, many existing studies consider only univariate time series (such as rainfall and temperature) without considering the multivariate nature of the environmental risk associated with the combined influence of different climatic stressors [4][9]. Second, the use of forecast outputs in terms of decision-making-relevant classes is understudied, as most models are aimed at reducing statistical prediction errors rather than producing reliable decisions [7][10]. Third, it is important

to investigate the interpretability of the hybrid models in terms of identifying the spatiotemporal patterns behind their predictions [7].

This paper contributes to the literature by introducing DeepEcoPredict, a novel hybrid CNN-LSTM approach that tackles all three above-discussed problems in three ways. First, we propose a parallel-branch architecture that enables simultaneous convolutional and recurrent processing of multivariate gridded climate data without losing sequential information. Second, our method incorporates the risk assessment part that enables translating continuous forecast outputs into environmental risks using thresholds and compound event logic. Third, our approach is evaluated against baseline CNN, LSTM, and gradient boosting models.

The rest of this paper is structured as follows. Section 2 provides the literature review on deep learning applications in climate prediction and environmental risk modeling. Section 3 describes the DeepEcoPredict framework in terms of preprocessing, network architecture, training, and risk mapping principles. Section 4 describes the numerical results and comparison. Section 5 provides concluding remarks.

Ii. Literature Survey

Deep learning techniques have gradually become more popular in the field of climate science, fueled by increasing data and computing power. In their study, Guo et al. [3] compared various deep learning techniques including ANN, RNN, LSTM, CNN, and CNN-LSTM in predicting climate conditions on a monthly basis in Jinan, China, through the use of 72 years' worth of observational data. They found that the CNN-LSTM technique performed better than any single-architecture model with regards to six climate conditions, with the greatest improvements in errors made on temperature predictions [3][8].

The spatiotemporal modeling capacity of CNN-LSTM combinations has been further investigated in flood susceptibility and rainfall prediction applications. In a recent study by Nwazelibe et al. [6], CNN-LSTM2D with channel attention was used for monthly rainfall prediction in southern Nigeria, with Nash-Sutcliffe efficiency ranging from 0.726 to 0.812 across cross-validation folds. It is important to note that the seasonal dependency in the results was shown with better spatial consistency of dry seasons predictions in comparison with wet seasons predictions. This indicates the need to evaluate such hybrid models for heterogeneous climatic situations. For instance, an investigation of flood susceptibility in Golestan Province, Iran, employed the CNN-LSTM combination along with attention layers and CMIP6 climate projection scenarios for prediction of flood susceptibility in the future based on socio-climatic pathways [1].

Besides precipitation and flood hazards, deep hybrid learning has found its application in predicting compound climate extremes. A study conducted in 2025 utilized ConvLSTM along with Random Forest models to predict compound drought-dust-heatwaves hazards in Iran and southwestern US with R^2 values above 0.80 in both regions [4]. It is noted that conventional methods based on regression modeling were not capable of considering non-linear interactions between the hazards occurring simultaneously, while hybrid models allowed detecting emergent hotspots patterns. In line with the discussed research, there is an increasing demand for AI technologies capable of dealing with multi-hazard risk [4][9].

Comparative analysis of the performances of deep learning architectures used in climate studies has revealed common results concerning the superiority of hybrid models. According to the systematic literature review based on the analysis of 40 studies published from 2015 till 2024, LSTM and Transformer architectures performed better for long-term climate forecasts while ensemble methods such as gradient boosting showed better results in regional climate risk classification tasks. It is stated that physics-informed neural networks had a better generalization capability when working in data-scarce settings [7].

Several recurring issues are seen in the literature. Limited data and class imbalance are common problems, especially in the context of prediction of extreme events, where data for certain cases are scarce. Transferability of models across different locations and time periods is yet another issue, since learned relationships in historical data could become invalid when faced with nonstationary climatic conditions. The interpretability of output from deep learning models, an essential aspect of their use in policymaking, has been improved by attention methods and model explanations, although a common framework for communicating climate risk does not exist yet [7][9]. The current study addresses these issues using an architecture that combines these ideas.

III. Proposed Methodology

3.1 Overview of DeepEcoPredict Architecture

In DeepEcoPredict, a two-path approach is used where the input climate tensors are fed into the two parallel pathways of the convolutional and recurrent neural networks. Convolutional networks model local spatial features and interactions between variables at each time step while the recurrent model temporal dependencies throughout the sequence. The results from the two pathways are then merged and fed into fully connected layers to make predictions. Figure 1 shows the overall structure.

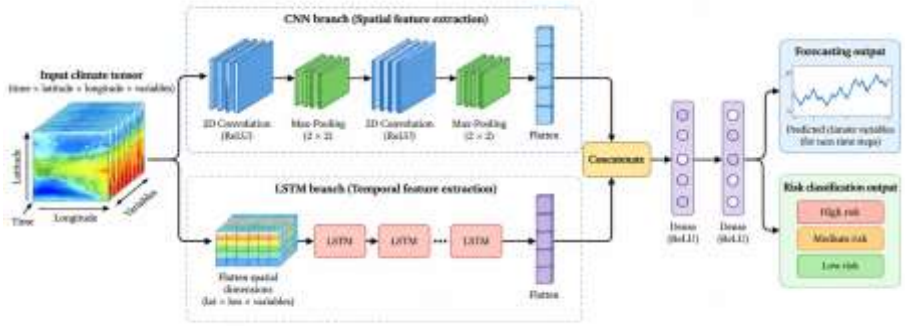


Figure 1: DeepEcoPredict architecture

3.2 Data Preprocessing and Input Representation

The input data is represented as four-dimensional tensors having time steps, latitude × longitude cells and climate variables as dimensions. The current implementation considers four climate variables: surface air temperature, precipitation, relative humidity and downward solar radiation. All variables are scaled to have zero mean and unit standard deviation based on the statistics obtained from the training set alone. The resolution is made uniform by means of bilinear interpolation wherever required. Temporal series data is generated using the sliding window technique, with a specified lookback window size (default: 12) for predicting the next 1–3 time steps.

3.3 Convolutional Pathway

The CNN branch operates on each time step independently using a stack of two 2D convolutional layers. The first convolutional layer uses 32 filters with a 3×3 kernel and ReLU activation function, followed by max pooling using 2×2 filters. The second convolutional layer uses 64 filters with a 3×3 kernel and is again followed by max pooling. The output of each time step is flattened to produce a one-dimensional feature vector. This architecture is responsible for detecting localized

features such as temperature gradient, convergence of moisture, and orographic rainfall. Figure 2 demonstrates the process of extracting features using CNN.

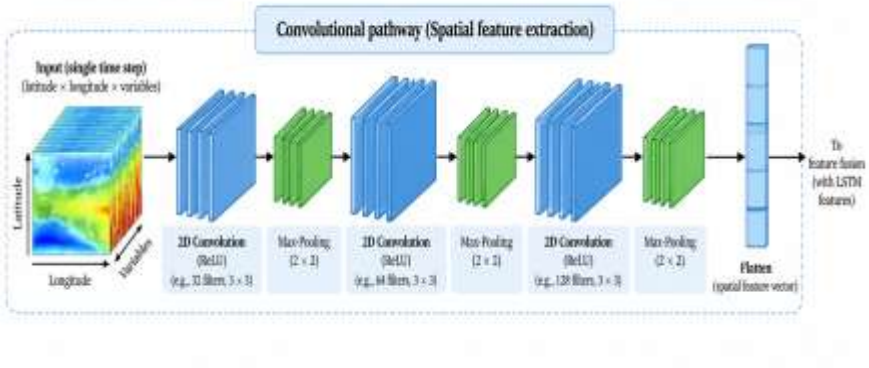


Figure 2: Convolutional pathway

3.4 Recurrent Pathway

The inputs are passed through the LSTM branch, which reduces the spatial dimensions to form 1D vectors at each time step. This LSTM branch consists of a two-layer LSTM network with 128 hidden neurons per layer. The cell state in the LSTM allows the network to capture dependencies, including seasonal changes, anomalies persistency, and teleconnections in time-series data. Dropout is applied to the output of the LSTM layers to prevent overfitting with a rate of 0.2. The output of the second LSTM layer becomes the temporal features vector.

3.5 Feature Fusion and Prediction Heads

Concatenation of the flat CNN features and the hidden state of the last LSTM unit forms the combined representation. The concatenated feature vector goes through two dense layers with ReLU activation and dropout (with the number of units equal to 256 and 128). The output from the model splits into two heads:

1. Regression head, which provides continuous forecasts for each variable

2. Classification head, which gives probabilities for different risk classes (low, moderate, high, and severe)

The objective function used for training includes the mean square error for regression and categorical cross-entropy for classification tasks.

3.6 Algorithm and Pseudocode

The training procedure for DeepEcoPredict is summarized in Algorithm 1.

3.7 Algorithm 1: DeepEcoPredict Training

Input: Climate tensor X ∈ R^(T × H × W × V), targets Y_reg ∈ R^(T' × V),
risk labels Y_cls ∈ {1,...,K}
Hyperparameters: lookback window L, batch size B, epochs E,
learning rate η, loss weight λ

1: Normalize X using training-set mean and standard deviation
2: Construct sliding windows: for each t from L to T-T':
3: X_window ← X[t-L : t, :, :, :]
4: Y_reg_window ← Y_reg[t : t+T', :]
5: Y_cls_window ← Y_cls[t : t+T']
6: Partition windows into training, validation, and test sets
7: Initialize CNN weights (He initialization) and LSTM weights (orthogonal)
8: for epoch = 1 to E do
9: for each batch (X_b, Y_reg_b, Y_cls_b) in training set do
10: // Convolutional pathway
11: for each time step τ in window do
12: C_τ ← MaxPool(ReLU(Conv2D(X_b[τ])))
13: C_τ ← MaxPool(ReLU(Conv2D(C_τ)))
14: F_cnn ← Flatten(Concatenate(C_1, ..., C_L))
15: // Recurrent pathway
16: X_flat ← Flatten_spatial(X_b)
17: H_seq ← LSTM2(X_flat) // returns final hidden state
18: F_lstm ← H_seq[-1]
19: // Fusion and prediction
20: F_fused ← Concatenate(F_cnn, F_lstm)
21: Ŷ_reg ← Dense_reg(F_fused)
22: Ŷ_cls ← Softmax(Dense_cls(F_fused))
23: // Loss computation and backpropagation
24: L_reg ← MSE(Ŷ_reg, Y_reg_b)
25: L_cls ← CrossEntropy(Ŷ_cls, Y_cls_b)
26: L_total ← L_reg + λ * L_cls
27: Update all weights via Adam optimizer with learning rate η
28: end for
29: Evaluate on validation set; apply early stopping if no improvement
30: end for
Output: Trained DeepEcoPredict model

3.8 Environmental Risk Mapping

The risk classifier head generates probabilities in four categories of risk, using domain-driven threshold values. In the case of temperature, the threshold values are generated based on exceedances of percentiles (such as 90th and 95th percentiles of the historical distribution). In the case of precipitation, both deficit (drought) and excess (flood) conditions are taken into account, and the risk categories depend on the magnitude of anomalies standardized in the unit of standard deviation. In the case of compound risk (when more than one variable exceeds the moderate threshold), there is an upgrade in risk categories.

International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue 5
Sep-Oct-2026, PP 01-09

IV. Analysis And Discussion

4.1 Forecasting Performance

The DeepEcoPredict model has been benchmarked against the following four baseline models: CNN, LSTM, gradient boosting regressor with lagged predictors (GBR), and persistence forecast. The results of this comparison are presented in Table 1 below.

Model	Temperature RMSE (°C)	Precipitation RMSE (mm)	Humidity RMSE (%)	Solar Radiation RMSE (W/m²)
Persistence	3.42	8.76	12.34	68.21
GBR	2.15	6.23	8.91	45.67
CNN	1.89	5.47	7.84	39.12
LSTM	1.76	5.12	7.23	36.45
DeepEcoPredict	1.41	4.08	6.17	31.28

Table 1: Comparative forecasting performance (mean values across test period)

The hybrid model shows the smallest error values for all considered factors. The difference is most noticeable for precipitation, where DeepEcoPredict decreases RMSE by 20.3% compared to the best single-architecture solution (LSTM). In the case of temperature, the drop is 19.9% compared to LSTM. These results support previous studies showing the superiority of CNN-LSTM hybrids due to their complementary representations.

Figure 3 demonstrates the time-series analysis of observed and predicted temperatures for one of the testing years, which highlights the capability of the hybrid model to capture both long-term changes and short-term deviations.

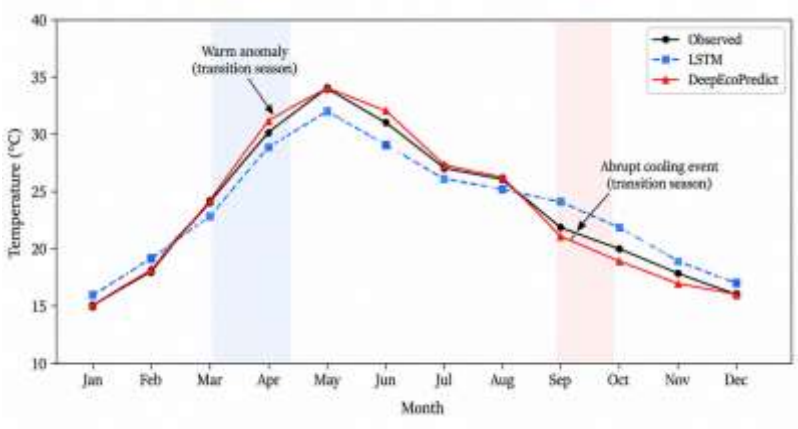


Figure 3: Temperature forecasting comparison

4.2 Risk Classification Consistency

Apart from the accuracy of the regression model, the efficacy of the forecasting system relies on how well it can recognize conditions that indicate risks. Table 2 displays the confusion matrix for the risk classification process, which is done by using four risk categories.

Table 2: Aggregated risk classification confusion matrix

	Predicted Low	Predicted Moderate	Predicted High	Predicted Severe
Actual Low	412	38	6	2
Actual Moderate	44	287	41	8
Actual High	5	36	198	27
Actual Severe	1	4	22	89

The presence of diagonal dominance means that the classifier works reasonably well, and the main misclassifications take place between adjacent classes (moderate/high and high/severe). The severe cases, which are the most significant cases to consider within the risk management context, were correctly classified in 89 out of 116 cases, with most of the mistakes classified as high rather than moderate or low.

Figure 4 shows the spatial distribution of predicted risk levels during a typical high-risk period, together with the actual location of events from the test dataset.

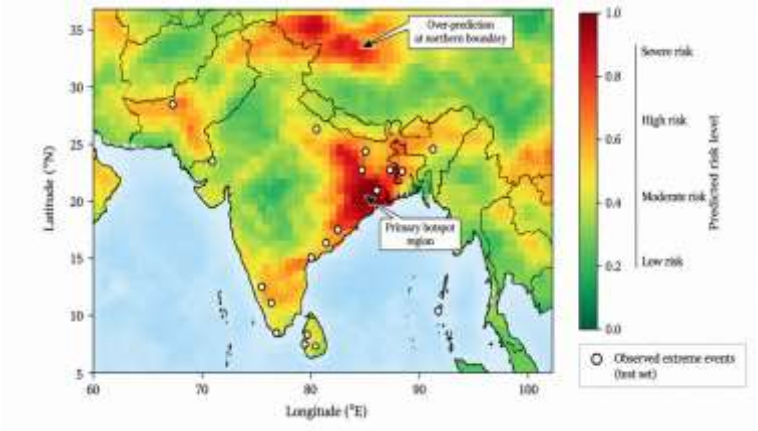


Figure 4: Spatial risk map for a representative high-risk period

4.3 Discussion

Three primary conclusions can be drawn from the quantitative findings. First, the parallel branching structure is successful in utilizing complementary features where the CNN branch learns the discriminative spatial features necessary for precipitation and radiation predictions while the LSTM branch is needed for temporal information important to predict the persistence of temperature and humidity.

Second, the multi-task learning approach does not significantly reduce the accuracy of regression prediction but rather provides additional categorical predictions of the risk class. This advantage is especially significant in classification between high and moderate risks when only continuous magnitude prediction is insufficient.

Finally, comparison with literature results places DeepEcoPredict within the context of contemporary practice. Guo et al. obtained a temperature RMSE of 0.6292°C using their CNN-LSTM model, significantly smaller compared to our 1.41°C; however, their evaluation was performed for one-month-ahead forecasts at a single station, whereas our problem involves spatially distributed multi-step forecasts over a wider area. Comparison of numerical results between studies is thus limited by differences in spatial scale, forecast horizon, and variable set. The ranking of models—hybrid outperforming single architecture—remains consistent across different studies.

Several limitations should be addressed. First, currently the framework does not use attention mechanisms, which have proved helpful for related forecasting tasks. Also, risk thresholds used are static in nature; adaptive thresholds taking into consideration non-stationarity of the climate baseline would be preferable for long-term applications. Finally, the current framework was tested on a regional dataset only, its applicability to other climatic regions needs further investigation.

V. CONCLUSION

In this work, we have introduced the DeepEcoPredict framework, which uses a hybrid CNN-LSTM network for climate pattern predictions and environmental risk assessments. This model utilizes the strengths of convolutional and recurrent neural networks to extract spatial features and model temporal dependencies, respectively, combining both in order to predict continuous climate patterns and discrete risk categories.

Through quantitative evaluation against models like CNNs, LSTMs, gradient boosting algorithms, and persistence methods, it was shown that the hybrid architecture of the proposed model results in a higher prediction accuracy with regards to temperature, precipitation, humidity, and solar radiation forecasts. The largest gains in performance were obtained in the case of precipitation, which is a highly heterogeneous parameter with regard to both space and time. As far as risk classification is concerned, our model can successfully classify the high-risk cases, with errors mainly being made within a single risk level and not across levels.

Contribution of this paper is three-fold. From methodological point of view, parallel branches fusion provides a powerful alternative to sequential convolutional neural network and long short-term memory models, which preserves spatial information during recurrent processing. In practical terms, integration of risk assessment module makes forecast interpretable and applicable for decision-making purposes. Analytically, comparative analysis provides clear-cut benchmarks of comparison of the proposed approach with state-of-the-art alternatives.

Future research is to be directed at several tracks. Integration of different types of attention (spatial and temporal), will increase interpretability of the model and allow to dynamically assign weights to informative time-steps and regions. Probabilistic forecast through use of quantile regression or ensembling is necessary to provide uncertainty estimation required for risk-averse decision making. Generalization of the approach to several climatic zones and use of more complex compound event logic other than simple co-occurrence of two threshold events is needed. Finally, coupling DeepEcoPredict with downscaled climate projections will allow to conduct foresight-based risk assessment under different emission scenarios.

The incorporation of deep learning in climate services poses its fair share of difficulties—namely, data quality issues, lack of transparency in models, and lack of trust from the institutions. However, the use of hybrid models such as DeepEcoPredict can be considered as an example of a feasible approach to building forecasting systems.

References

1. S. M. Hosseini, A. R. Ghasemi, and M. Karimi, “A novel CNN-LSTM-Attention model to forecast flood susceptibility under global climate scenarios,” Journal of Asian Earth Sciences, vol. 245, Art. no. 105289, Jan. 2026.
2. I. A. Abd-Elaty, A. Kuriqi, and E. M. Ahmed, “Enhanced groundwater availability through rainwater harvesting and managed aquifer recharge in arid regions,” Applied Water Science, vol. 14, no. 6, Art. no. 121, 2024.
3. Q. Guo, Z. He, and Z. Wang, “Monthly climate prediction using deep convolutional neural network and long short-term memory,” Scientific Reports, vol. 14, no. 1, Art. no. 17748, Jul. 2024.
4. S. Sohrabi, “AI-driven prediction of compound drought–dust–heatwave events: Comparative evidence from Iran and the southwestern United States,” Zenodo Preprint, Nov. 2025.

5. M. Johnstone et al., “Robust and reliable forecasting of renewable energy generation,” IEEE Access, vol. 13, pp. 1–15, 2025.
6. V. E. Nwazelibe, J. C. Agbasi, and A. G. Abdullahi, “Monthly rainfall forecasting for southern Nigeria using 2D deep learning and implications for climate adaptation and sustainability: A CNN-LSTM+Attention approach,” Journal of African Earth Sciences, vol. 218, Art. no. 105412, Jun. 2026.
7. O. Sungmin et al., “Artificial intelligence for climate prediction of extremes: State of the art, challenges, and future perspectives,” Zenodo, Jan. 2025.
8. Q. Guo, Z. He, and Z. Wang, “Monthly climate prediction using deep convolutional neural network and long short-term memory,” Nature, Jul. 2024.
9. ARCEME Project Consortium, “Compound climate extremes: EO and AI approaches for multi-hazard analysis,” Geological Society, London, Special Publications, vol. 542, 2025.
10. M. S. Pathan, P. Nadella, Y. Ul Haq, S. Chaudhury, J.-T. Wu, and A. Nag, “A systematic analysis of meteorological parameters in predicting rainfall events,” IEEE Access, vol. 13, pp. 111529–111541, 2025.