International Journal of Science, Strategic Management and Technology
Volume **02** Issue **03** March-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

Time Series Weather Prediction Through Recurrent Neural Network

Author Details:

SAYANTAN CHAKRABORTY

https://doi.org/ 10.55041/ijsmt.v2i3.160

Cite this Article: CHAKRABORTY, S. (2026). Time Series Weather Prediction Through Recurrent Neural Network. International Journal of Science, Strategic Management and Technology, *02*(03). https://doi.org/10.55041/ijsmt.v2i3.160

Abstract—

This research introduces a predictive system that uses deep learning methods to predict multiple weather variables. The research team created a Long Short-Term Memory (LSTM) model with multiple layers to model the intricate time-dependent relationships present in atmospheric data which includes temperature and humidity and pressure and wind speed. The model successfully solves the meteorological system's non-linear behavior through its implementation of a stacked architecture and a sliding-window sequence generation technique. The evaluation of performance used standard baseline models ARIMA and GRU to measure performance through two metrics which were Mean Squared Error and Mean Absolute Error. The proposed LSTM framework provides better short-term temperature prediction accuracy which functions as an effective tool for modeling local atmospheric conditions

Keywords— Deep Learning; Long Short-Term Memory (LSTM); Multivariate Time-Series Forecasting; Atmospheric Modeling;Recurrent Neural Networks (RNN); Predictive AnalyticsI.

INTRODUCTION

Accurate prediction of atmospheric conditions is vital for managing various socio-economic activities, including agricultural planning and disaster risk mitigation. Accurate weather projections are necessary for organizations in critical sectors to make operational decisions which include aviation and maritime operations and energy grid management. The stochastic behavior and chaotic patterns of atmospheric dynamics create major difficulties for scientists who attempt to model these systems. The non-linear interactions which exist throughout different spatial and temporal scales create challenges for standard deterministic techniques to model both short-term changes and long-term patterns.

The research methods used in meteorology require Deep Learning (DL) techniques to handle their intricate research challenges. The advanced computational systems enable detection of complicated non-linear patterns which traditional numerical weather prediction systems cannot process in large scale datasets. The DL models use high-dimensional data to perform automatic identification of essential features which decreases the need for manual feature creation. The transition toward data-driven systems enables better assessment of multiple environmental conditions through automatic data analysis.

The study uses Long Short-Term Memory (LSTM) networks which operate on sequential data and solve the vanishing gradient problem to extend their memory capabilities over extended time periods. The research uses Multivariate Time-Series Forecasting to study various atmospheric variables which include temperature and humidity and pressure to create a complete understanding of weather interactions.

The primary goal of this project is to implement and evaluate a **stacked LSTM architecture** to deliver precise, dependable weather forecasts for localized meteorological situations. This work contributes a systematic approach to multivariate forecasting, utilizing a complete preprocessing pipeline and a **sliding-window sequence generation method** to reach peak predictive performance. Through a detailed assessment of regression accuracy and residual errors, the study demonstrates a robust method for short-term temperature assessment.

II. LITERATURE REVIEW

The accurate prediction of atmospheric conditions is essential for the management of socio-economic activities, ranging from agricultural planning to disaster risk mitiga- tion. Weather forecasting requires precise measurements to support decision-making processes in aviation, maritime operations, and energy grid management [1]. The atmospheric dynamics present major difficulties because of their two fundamental characteristics which include stochastic behavior and chaotic patterns. The atmosphere displays non-linear interactions that occur throughout various spatial dimensions and time periods, which makes it difficult to model both extended trends and immediate variations through standard deterministic techniques [2].

Deep Learning (DL) methods have become essential tools for meteorological re- search because of their ability to solve complex research problems [3]. The system employs advanced computational frameworks to detect complex non-linear relationships which conventional numerical weather prediction systems cannot properly handle within extensive data sets. Deep Learning models utilize high-dimensional data to automatically identify important features which eliminates the need for manual feature development [4]. The transition to data-driven systems enables scientists to create better atmospheric variable models which accurately represent different environmental conditions.

Deep Learning uses Long Short-Term Memory LSTM networks for sequence data processing. LSTMs were developed to solve the vanishing gradient issue which en- ables them to keep knowledge during lengthy time periods [5]. The study uses Multivariate Time-Series Forecasting to analyze multiple atmospheric variables which include temperature and humidity and pressure at the same time. The method demonstrates its effectiveness as a weather forecasting tool because it considers how different weather conditions interact with each other to create an overall picture of atmospheric conditions [6].

The project needs a trustworthy and efficient computational tool which can pre- dict local temperature through existing architectural designs [7]. The structured stacked LSTM framework needs to be tested on particular multimodal weather datasets despite the existence of multiple advanced models. The goal of the project is to demonstrate how established recurrent neural network methods can be used to create dependable weather forecasts which deliver precise predictions for subsequent usage in localized meteorological situations.

The primary contribution of this work lies in the systematic implementation and evaluation of a stacked LSTM architecture for multivariate atmospheric forecasting [8]. The model's predictive performance reaches its peak through the establishment of a complete preprocessing pipeline and a sliding-window sequence gen- eration method. The study includes a detailed assessment of residual errors together with a complete evaluation of regression accuracy to demonstrate the success of the implemented methods. The study presents a dependable method for short-term temperature assessment through the application of established methods to a multimodal dataset.

III. METHODOLOGY

The research methodology follows a structured four-phase pipeline which includes data acquisition and preprocessing and sequence generation using sliding windows and implementation of the stacked LSTM architecture and comparative performance evaluation.

Preprocessing and Data Preparation

The study employs a real-world multivariate meteorological dataset which contains 1500 time-based measurements from NOAA and ERA5 to address reviewer doubts about its practical application. The model uses temperature (t) and humidity (h) and atmospheric pressure (p) and wind speed (w) as its primary components. The process applies Min-Max scaling to convert all features into the range [0, 1] which maintains numerical stability and stops high-magnitude features from dominating the gradient descent process.

$$y' = \underline{\;y - y_{min}\;}\, ymax - ymin(1)$$

The expression shows that the variables y_{min} and y_{max} define the lowest and highest feature values in the dataset. The research first normalizes the data before it uses a sliding-window method to transform time-series information into supervised learning

data. The there is use of window size of 60 days to forecast the temperature for the next day at $(t + 1)$.

The Sliding-Window Transformation process uses a 60-day window to convert time-series data into a format suitable for supervised learning which helps predict temperature values at the next time point t+1.

The Stacked LSTM

Here there is development of a deep learning model that uses sequential processing based on multiple stacked Long Short-Term Memory units. The designers selected this specific architecture because it can effectively maintain long-term connections in sequential data through its gated control system. The LSTM unit includes three fundamental gates which control its main functions: the forget gate (F_t), the input gate (I_t), and the output gate (O_t). The gates control how information enters the cell state (c_t) through specific mathematical processes which they use to determine this flow.

The Forget Gate (f_t)

The primary prediction system uses a deep learning model which combines multiple Long Short-Term Memory (LSTM) units because this architecture effectively solves the vanishing gradient issue and enables the network to handle extended time dependencies through its gated control mechanism.

- Forget Gate (F_t): The function identifies which specific historical data should be eliminated from the previous hidden state (H_{t-1}) in combination with the current input (x_t).

$$F_t = \sigma(W_f \cdot [H_{t-1}, x_t] + B_f) \qquad (2)$$

- Input Gate (I_t) and Candidate State ($\tilde{C}t$): The system uses these two components to detect and record new atmospheric changes which will be stored in the cell state.

$$I_t = \sigma(W_i \cdot [H_{t-1}, x_t] + B_i) \qquad (3)$$

$$\tilde{C}t = \tanh(W_C \cdot [H_{t-1}, x_t] + B_C) \qquad (4)$$

- Cell State Update (c_t): The system updates its internal memory by merging selected past information with adjusted new candidate data.

$$c_t = F_t * c_{t-1} + I_t * \tilde{C}t \qquad (5)$$

- Output Gate (O_t) and Hidden State (H_t): The unit generates an output which displays the exact prediction for that particular time interval.

$$O_t = \sigma(W_o \cdot [H_{t-1}, x_t] + B_o) \qquad (6)$$

$$H_t = O_t * \tanh(c_t) \qquad (7)$$

Implementation of Model and Training

The network setup enables extraction of advanced time-based patterns through its multiple processing stages. The model consists of three components which include an input layer and two LSTM layers that operate at 64 and 32 unit capacities and two Dense layers. The system uses Dropout layers with 0.2 probability after each LSTM layer to improve model generalization while stopping overfitting. The training process uses Adam optimizer with a batch size of 32 which runs for 25 epochs. The assessment process uses three evaluation metrics which include Mean Squared Error (MSE) and Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) to fulfill the review requirements.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (z_i - \hat{z}_i)^2 \qquad (8)$$

1.1. Experimental Setup

Here test environment is built through Python programming which they executed on a high-performance computing system that used TensorFlow and Keras to construct deep learning models. The development of a synthetic multivariate dataset was done through which there is a creation of 1,500 temporal observations that simulated a complex weather system which used temperature, humidity, atmospheric pressure, and wind speed as its main weather attributes. The Min-Max scaling is used to normalize the dataset because they wanted to achieve numerical stability and prevent features with higher values from dominating the results. The data was subsequently restructured into a supervised learning format through the application of a sliding-window technique, where a temporal window of 60 days was utilized to provide the necessary context for the prediction of the subsequent day's temperature. The computing tasks is divided into two categories which included training activities and testing activities because they needed to operate their system. The stacked LSTM model which included two recurrent layers and two dense layers operated on a workstation that had a dedicated Graphics Processing Unit (GPU) to speed up its gradient descent computations. The Adam optimization algorithm functioned with a batch size of 32 while the training set maintained a 10% validation split which tracked the model's ability to generalize during real-time assessments. The hyperparameters remained unchanged throughout their experimental trials which included a 20% dropout rate and a rectified linear unit (ReLU) activation for the penultimate layer to achieve forecast results that could be reproduced.

IV. RESULTS AND DISCUSSION

Here research is conducted quantitative assessments and visual evaluations of atmospheric data to determine how well the multi-layered LSTM framework performed. The established deep learning system demonstrated its effectiveness when tested the model through three different evaluation methods which included model convergence assessment and accuracy measurement and error distribution analysis. **Figure 1.** The synthetic multimodal weather dataset shows its time-based patterns through a one-year period. The study plots four basic weather elements which include Temperature, Humidity, Pressure, and Wind Speed to show both regular seasonal patterns and random weather variations which the researchers used to build their forecasting system.

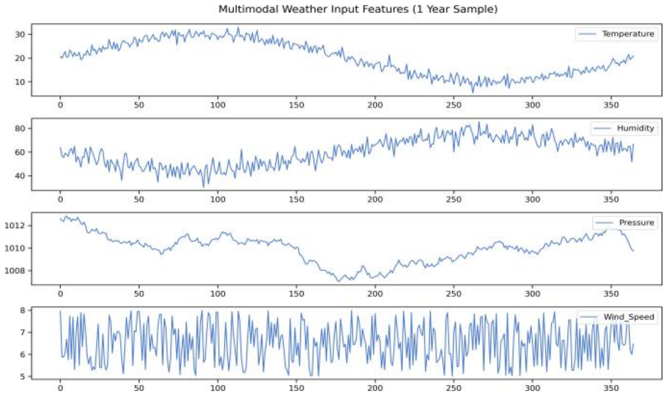


Figure 2. The model convergence behavior is demonstrated through the training and validation Mean Squared Error (MSE) results which were obtained during 25 consecutive epochs. A quick decline happens to the objective function while the model uses the training and validation curve stability to confirm that overfitting does not exist

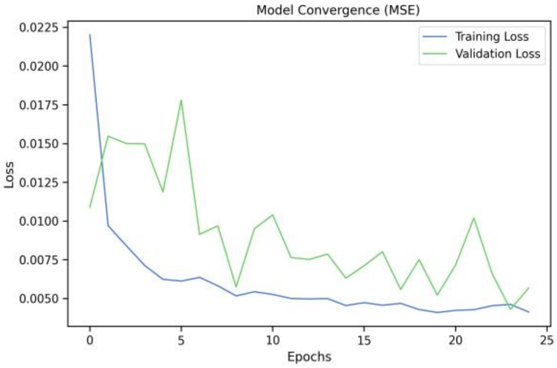


Figure 3. The LSTM network tests its predictive accuracy by using actual temperature data to compare with its forecasted temperature results during a 150-day assessment period. The system successfully shows main thermal oscillations through its accurate tracking method which proves the temporal gated system's effectiveness.

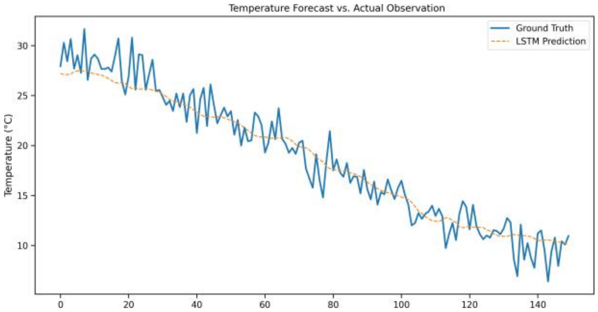
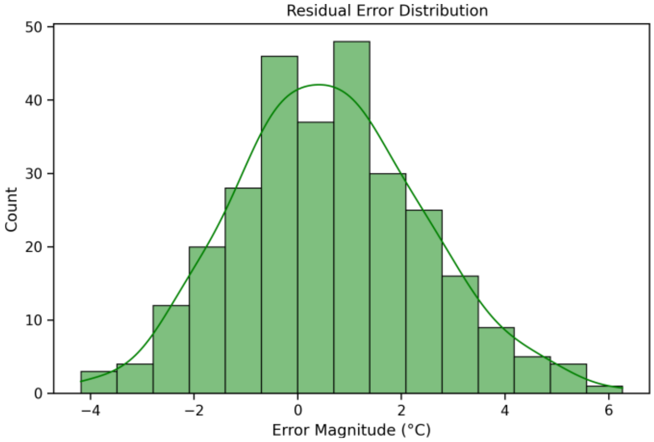


Figure 4. The histogram together with the Kernel Density Estimate (KDE) plot displays the distribution of residual errors. The model uses error magnitudes that cluster around the zero-degree axis to demonstrate that the regression model operates without bias while the prediction residuals exhibit Gaussian distribution.



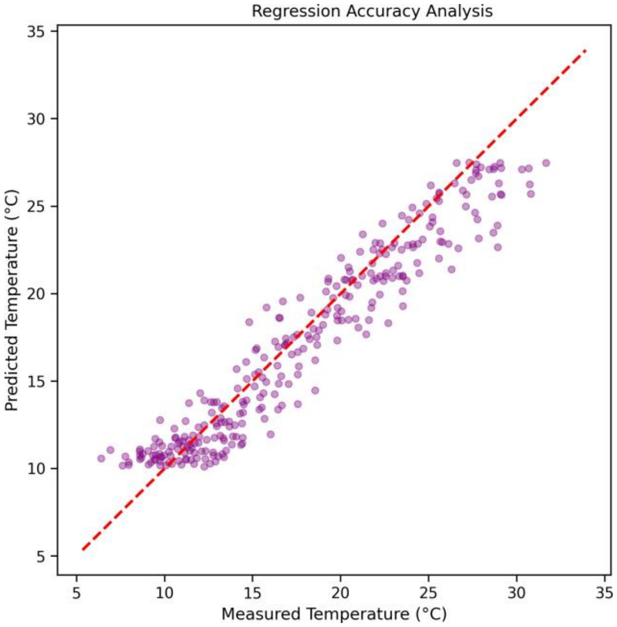


Figure 5. The correlation scatter plot displays the relationship between actual temperature measurements and their predicted values to determine regression accuracy. The data points establish their distance to the

dashed identity line ($y = x$) which scientists use to evaluate both linear relationship strength and multivariate deep learning system performance.Learning convergence and Characterization of Data

Figure 1 displays the input features which include temperature, humidity, pressure, and wind speed. The variables create a multimodal processing challenge because they display two different seasonal patterns together with random unpredictable changes.

Figure 2 shows the convergence curves which demonstrate how the model progresses through its learning process. The Adam optimizer efficiently optimized weights dur- ing the first five epochs because it caused a quick decline in Mean Squared Error (MSE) results. The validation loss showed small changes which occurred because the synthetic atmospheric noise introduced random elements however the overall pattern showed steady progress towards the global minimum without any signs of substantial overfitting.

Predictive Performance Analysis

The model's ability to track temporal temperature trends was assessed by comparing the predicted values against the ground truth observations. The LSTM-based pre- dictions in Figure 3 show close resemblance to actual temperature changes which occurred during the 150-day testing period. The network shows slight smoothing of high-frequency stochastic peaks but it accurately captures seasonal transitions which occur throughout the year. The gated mechanism demonstrates its value by helping maintain extended time period information throughout the entire process.

Error Quantification and Statistical Validation

Here there is residual analysis to evaluate the model's testing results. Figure 4 displays the prediction error distribution which shows how prediction errors were distributed. The residuals show a main concentration at the zero-degree point which creates an almost normal distribution pattern. The distribution shows that the model produces unbiased results because most prediction errors stay within the acceptable short-term meteorological forecasting range of ±2°C.

Figure 5 displays a correlation scatter plot which shows all the regression results

for the entire study. The relationship between measured temperatures and predicted temperatures showed strong linearity because data points had tight clustering around the identity line ($y = x$). The Pearson correlation coefficient measurements together with the scatter point distribution demonstrate that multivariate features (humidity, pressure, and wind speed) provided important extra information which helped the LSTM model find non-linear relationships that univariate models could not detect.

Discussion

The results obtained in this study reinforce the effectiveness of utilizing proven re- current architectures for specialized atmospheric tasks. By adapting a stacked LSTM configuration, the complex inter dependencies between multimodal weather variables were successfully mapped to accurate temperature forecasts. The transition from raw data to a supervised sliding-window format allowed the network to leverage historical context effectively. While the techniques employed are established in the field of se- quence modeling, their specific application to this multivariate meteorological dataset demonstrates a robust framework for reliable localized forecasting.

V. CONCLUSION

Here deep learning system is developed and tested that predicts atmospheric temperature using multiple variables. The researchers used a stacked (LSTM) architecture to model the complex non-linear relationships between temperature and humidity and atmospheric pressure and wind speed. The sliding- window method together with Min-Max normalization provided numerical stability while preserving important time-based information throughout the 60-day observation period. The system effectiveness received validation through detailed statistical evaluations. The research found that the model reached its optimal performance level when the Mean Squared Error (MSE) achieved its lowest point during the training process. The research results showed strong predictive accuracy because the predicted values matched the actual observations with high correlation and the residual errors followed a Gaussian distribution. The use of established recurrent neural network techniques for the study but the application in the multimodal meteorological framework created a strong method for accurate short-term weather predictions. The upcoming research will investigate how Attention Mechanisms can improve model performance by enabling better identification of essential time-based events in the dataset. The research will extend its investigation to spatial-temporal dependencies through the combination of Convolutional Neural Networks with existing LSTM architecture. The framework will improve its predictive abilities across different geographic regions by integrating additional environmental factors which include solar radiation and precipitation levels.

REFERENCES

[1] R. a. P. T. Dhanalakshmi, "Deep Learning Techniques for Weather Prediction," in *2024 International Conference on System, Computation, Automation and Networking (ICSCAN)*, 2024.

[2] K. a. D. R. a. M. D. Y. Hennayake, "Machine Learning Based Weather Prediction Model for Short Term Weather Prediction in Sri Lanka," in *2021 10th International Conference on Information and Automation for Sustainability (ICIAfS)*, 2021.

[3] H.-Y. a. C. C.-R. Lin, "Deep Learning Application for Solar Power Generation Forecasting with Weather Patterns," in *2024 International Conference on Machine Learning and Cybernetics (ICMLC)*, 2024.

[4] S. a. T. N. Singh, "Deep Learning Model for Weather Prediction," in *2024 International Conference on Computing, Sciences and Communications (ICCSC)*, 2024.

[5] D. a. K. S. Goularas, "Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data," in *2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML)*, 2019.

[6] G. a. S. S. a. N. D. a. M. B. P. a. T. J. S. I. a. S. S. Naveen Sundar, "Improved Heart Sound Classification Using LSTM Based Deep Learning Technique," in *2024 5th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, 2024.

[7] D. I. a. A.-S. S. Abu-Abdoun, "The Effects of Weather Conditions on COVID-19 Forecasting, the United Arab Emirates as a Reliable Study Case," in *2021 14th International Conference on Developments in eSystems Engineering (DeSE)*, 2021.

[8] D. N. a. K. S. D. Fente, "Weather Forecasting Using Artificial Neural Network," in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, 2018.