E-ISSN 2586-9027
Homepage : https://tci-thaijo.org/index.php/SciTechAsia Science & Technology Asia
Vol.31 No.1 January - March 2026 Page: [260-281]
*Original research article*

# A Comparative Framework of VAR and LSTM Networks for Soil Moisture Prediction Using Atmospheric Data

Tawsif Mahmud[1,2], Preetom Nag[1,2], Goutam Saha[3,*], Jiaul Haque Saboj[1,2], Suvash C. Saha[4]

[1]*Department of Mathematics & Physics, North South University, Dhaka 1229, Bangladesh*
[2]*Center of Applied & Computational Sciences, North South University, Dhaka 1229, Bangladesh*
[3]*Department of Mathematics, International University of Business Agriculture and Technology, Dhaka 1230, Bangladesh*
[4]*School of Mechanical and Mechatronic Engineering, University of Technology Sydney, NSW 2007, Australia*

Received 30 July 2025; Received in revised form 6 March 2026
Accepted 11 March 2026; Available online 27 March 2026

**ABSTRACT**

Accurate prediction of soil moisture is vital for optimizing water resource management and agricultural productivity, yet it remains a challenging problem due to the complex, non-linear relationships between soil dynamics and atmospheric conditions. To address this, this study proposes a comprehensive forecasting framework comparing conventional statistical methods with advanced deep learning techniques. The key contribution of this work is a robust methodological pipeline that integrates a comprehensive set of ten meteorological variables, including temperature, relative humidity, wind speed, radiation, and vapor pressure deficit, following rigorous time-series preprocessing and stationarity checks. Specifically, we formulate and compare a Vector Auto-regression (VAR) model against a Long Short-Term Memory (LSTM) network for one-day ahead forecasting of soil moisture during cropping seasons. To evaluate the efficacy of the proposed techniques, we utilized historical data spanning from 1972 to 2024. The evaluation results demonstrate that the LSTM network significantly outperforms the conventional VAR model, achieving exceptional accuracy with an $R^2$ exceeding 0.99 and notably lower RMSE values. These findings validate the reliability of the proposed LSTM framework and establish its superiority over traditional statistical approaches for precise soil moisture estimation.

**Keywords:** Long short-term memory; Prediction; Soil moisture; Vector auto-regression; Weather data

***Corresponding author:** gsahamath@du.ac.bd **doi:** xx.xxxxx/scitechasia.xxxx.xx

## 1. Introduction

In recent years, the rising need for sustainable farming and effective water resource management has highlighted the value of precise soil moisture predictions. Soil moisture, an essential factor in determining the water content within the soil, profoundly influences agricultural productivity, water conservation efforts, and ecosystem health. Currently, machine learning (ML) stands as a state-of-the-art technique for predicting unknown values. The adoption of artificial intelligence tools and techniques in agriculture and climatology has led to the increasing popularity of ML approaches for estimating and forecasting soil moisture.

Recently, research in hydrology has progressively adopted data-driven approaches based on deep learning, such as convolutional neural networks (CNNs) [1], recurrent neural network (RNN) [2], and long short-term memory (LSTM) [3–5]. These models exhibit superior predictive performance with fewer parameters compared to traditional physically based models [6–8]. Baruah et al. [9] focused on utilizing ML for the prediction of soil moisture. Han et al. [10] proposed a novel framework called ”pyramid” to achieve accurate and high-resolution soil moisture estimation. They utilized the CART algorithm to address challenges commonly associated with global fitting approaches. Zhang et al. [11] explored how initializing soil moisture affects the accuracy of Weather Research and Forecasting (WRF) model predictions in Xinjiang, China. The research focused on a series of simulations using both the single-column model (SCM) and the WRF model for a real weather scenario in that locality. Fereres et al. [12] established that, notwithstanding a potential reduction in yield due to decreased water application, such a measure may enhance profitability in the cultivation of horticultural crops. Ballester et al. [13] conducted an inquiry into the impacts of varying irrigation deficits upon both the quality and the yield of cotton. Datta and Faroughi [14] proposed a multihead LSTM model that exhibited robust forecasting capabilities for soil moisture, delivering accurate predictions up to one month in advance with a coefficient of determination ($R^2$) of 95.04%. This outcome underscores the model's accuracy and reliability in predictive performance.

Cai et al. [15] introduced a Deep Learning Regression Network (DNNR) designed to forecast soil moisture in the Beijing region, with the objective of addressing the accuracy constraints inherent in existing predictive models. Their research validated the efficacy of the DNNR model in soil moisture forecasting, emphasizing its superior data-fitting capabilities and its precision in predicting soil moisture trends and values. Moghadas and Badorreck [16] explored machine learning techniques, particularly artificial neural networks, for enhancing the understanding of the relationship between electrical conductivity and soil water content. Their findings highlight the superiority of ANN in capturing complex and nonlinear variations, offering a valuable tool for soil moisture estimation from geophysical data. Greifeneder et al. [17] presented a machine learning-based methodology for mapping soil moisture at high spatial resolution, responding to the need for enhanced detail in applications including agriculture and environmental monitoring. Wai et al. [18] performed a comprehensive review of multiple deep learning models applied to time series forecasting, including Recurrent Neural Networks (RNN), Long Short-Term Memory

(LSTM) networks, Convolutional Neural Networks (CNN), Gated Recurrent Units (GRU), and Temporal Convolutional Networks (TCN). In their analysis, LSTM proved effective in capturing long-term patterns. Jiang et al. [19] addressed the challenges of time series characteristics of soil moisture changes; the study integrates Principal Component Analysis (PCA) with LSTM neural networks to predict soil moisture levels. They found that the proposed PCA-LSTM model outperforms RBF NN, RNN, and traditional LSTM models in terms of prediction accuracy and efficiency. Singh et al. [20] focused on precision agriculture and the introduction of a soil moisture forecasting model, particularly employing LSTM, addresses the need for advanced techniques in farming practices. Their investigation advances the field of precision agriculture by training a model for the accurate prediction of soil moisture.

Huang et al. [21] proposed a soil moisture forecasting model utilizing a genetic neural network to tackle the intricate and nonlinear characteristics of soil moisture prediction. This model was applied to forecast soil moisture levels at Hongxing Farm in Heilongjiang Province. Zuo et al. [22] addressed the challenge of precisely estimating the distribution of soil moisture content across large scales. Their study introduced an innovative hybrid model integrating Particle Swarm Optimization (PSO) with an LSTM network, designated as PSO-LSTM. The model aims to enhance prediction accuracy across different soil depths using Sentinel-1A data. Han et al. [23] concentrated on the significance of employing deep learning, particularly the LSTM model with a sequence-to-sequence structure, for accurate runoff prediction. Filipovi´c et al. [24] created a robust soil moisture prediction system that utilizes meteorological data to forecast soil moisture levels three days in advance, supporting irrigation scheduling and agricultural planning. They observed that the LSTM network yielded the lowest errors compared to the statistical and classical machine learning models, demonstrating superior prediction accuracy and generalization properties.

Mok et al. [25] advanced the use of deep learning, particularly LSTM models, for predicting dam inflows, a vital component of water resource management. Their results demonstrated that the LSTM model attained high accuracy in inflow forecasts when leveraging historical datasets comprising precipitation and inflow records. Bai et al. [26] highlighted the comparative performance of advanced machine learning (LSTM network) and traditional hydrologic models in predicting runoff under changing climatic conditions. The findings emphasize the LSTM network's consistent performance across different climatic conditions during calibration.

As we know, VAR models are widely recognized as a conventional statistical approach for capturing linear interdependencies among multiple time series. For instance, Lu et al. [27] successfully employed a Bayesian hierarchical VAR model to forecast medical and psychological conditions which demonstrated the method's capability to handle multivariate inputs. However, despite its utility in linear domains, the conventional VAR approach possesses lots of limitations that hinder its effectiveness in solving complex environmental problems. Specifically, VAR relies on the assumption of linearity and struggles to model the nonlinear, chaotic dynamics inherent in soil moisture and atmospheric interactions. Furthermore, VAR models often fail to capture long-term temporal dependencies and are sensitive to noise in high-frequency weather

data. To address these limitations, our study proposes a Deep Learning-based approach using LSTM networks which are inherently designed to model non-linear relationships and possess gating mechanisms that allow them to filter noise and retain long-term memory of seasonal patterns. Our study aims to demonstrate that the proposed LSTM framework overcomes the weaknesses of traditional statistical models, and hence provide a more robust and accurate solution for soil moisture prediction under dynamic atmospheric conditions.

Therefore, our study aimed to examine the relationship and forecasting potential between surface and soil moisture, taking into account their interdependence as shaped by variable weather conditions in Dhaka, Bangladesh. The novelty of this research lies in its comprehensive approach to soil moisture prediction, integrating advanced machine learning techniques with practical agricultural applications. This investigation focuses on Soil Moisture depending on weather factors, including Temperature (°C), Relative Humidity (%), Wind Gust ($ms^1$), Wind Speed ($ms^1$), Shortwave Radiation ($Wm^2$), UV Radiation ($Wm^{-2}$), Mean Sea Level Pressure (hPa), Vapor Pressure Deficit (hPa) and Soil Temperature (°C). This study breaks new ground by incorporating a comprehensive analysis of local atmospheric conditions, which has not been extensively covered in previous research. By using advanced machine learning algorithms, we aim to provide a more accurate understanding of soil moisture dynamics in this region. The objectives of this study were to assess the predictability of soil moisture under varying weather conditions and to formulate VAR and LSTM models for soil moisture forecasting, with a specific focus on the cropping season, followed by a comparison of their predictive accuracy. The insights derived from this research are intended to offer valuable guidance to decision-makers in the fields of water resource management, agriculture, and environmental planning, thereby enhancing the understanding and prediction of soil moisture dynamics. The current study is informed by the research conducted by Basir et al. [28], and the modeling concept is derived from their work.

## 2. Methodology and Materials

### 2.1 Research area and dataset

Agricultural land covers a diverse range of land uses, including arable land for crop cultivation, such as paddy rice, vegetables, and other crops, as well as silvopasture, which integrates forest and pasture. Additionally, it includes mangrove ecosystems, freshwater bodies like rivers, lakes, beels, and haors, and land used for aquaculture, tea plantations, and salt pans. These land types may be further categorized into seasonal cropland, orchard land, wet meadow or pasture land, and fallow land, which is temporarily not in use. In the Dhaka division, cropland has been largest in Bangladesh, though its extent has been lessening over time. Cropland area measured 2,313,751 hectares in 1976, 2,312,961 hectares in 2000, and 2,161,650 hectares in 2010. This research investigates the relationship between surface and soil moisture content at a specific site in Dhaka, Bangladesh, with geographical coordinates corresponding to a latitude of 23.75 and a longitude of 90.5, respectively. For a visual representation of the study area's geographical location, Fig. 1 illustrates Dhaka's position on the map. Data acquisition for this study leveraged the comprehensive historical weather dataset available through Meteoblue (https://www.meteoblue.com/en/historyplus).

The dataset utilized in this study encompasses a temporal span ranging from January 1, 1972, to February 15, 2024, encompassing a total of 19,039 individual observations, all recorded at a depth of 0 to 7 cm. The soil surveyed in this study exhibits a diverse range of characteristics, encompassing various types such as Non-calcareous Alluvium, Calcareous Alluvium, Peat, Non-calcareous Grey Floodplain soils, Grey Piedmont soil, and Acid Basin clay. This extensive dataset offers a valuable resource for examining the influence of diverse weather elements on soil moisture dynamics. The specific weather parameters retrieved include Temperature (°C), Relative Humidity (%), Wind Gust ($ms^1$), Wind Speed ($ms^1$), Short-wave Radiation ($Wm^2$), UV Radiation ($Wm^2$), Mean Sea Level Pressure (hPa), Vapor Pressure Deficit (hPa), Soil Temperature (°C), and Soil Moisture ($m^3m^3$). The dataset comprises daily records, presenting the daily average of hourly measurements, offering a valuable resource for examining the influence of diverse weather elements on soil moisture.

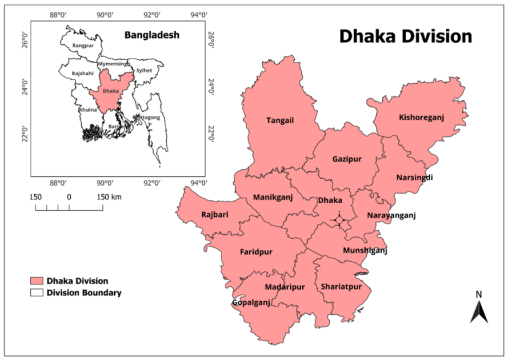


**Fig. 1.** Data Collection site (latitude, longitude = 23.75, 90.5) Dhaka Division.

## 2.2 Missing value handling

The dataset includes numerous missing values due to delays and gaps in data reception and execution by the weather station. An established methodology involves systematically removing rows containing any instances of null values, commonly called dropping rows. This paper handles missing values using the technique mentioned earlier, ensuring the reliability and validity of subsequent analyses.

## 2.3 Time series analysis

To develop forecasting models for a time series variable based on additional time series variables, it is imperative to examine their forecastability and the relationships linking the dependent variable with the independent input variables, typically known as predictors. The dependent variable in focus is the soil moisture, and its forecast is based upon input variables such as Temperature, Relative Humidity, Wind Gust, Wind Speed, Shortwave Radiation, UV Radiation, Mean Sea Level Pressure, Vapor Pressure Deficit, and Soil Temperature. The results of the time series analysis prompted the inclusion of soil moisture data from previous days as an additional input variable. Multiple statistical investigations were performed to examine the properties of the time series variables, confirming their autocorrelation, stationarity, cointegration with the soil moisture series, and mutual interdependence. The respective tests employed for this purpose are outlined in Table 1.

## 2.4 Formulation of soil moisture forecasting models

The findings from the time series analysis indicated notable autocorrelation and cointegration within the soil moisture dataset. This study concentrates

**Table 1.** Time series analysis statistical tests [28].

| Test & Objective | Used Library |
|---|---|
| Durbin-Watson test (Autocorrelation) | statmodels-stattools |
| Augmented Dickey-Fuller test (Stationarity) | statmodels-stattools |
| Johansen's cointegration test (Cointegration) | statmodels-stattools |
| Granger causality test (Interdependency of Variables) | statmodels-stattools |
| Seasonal decompose (Seasonality) | statmodels-tsa-seasonal_decompose |

on constructing predictive models for soil moisture forecasting, utilizing two distinct methodologies: Vector Autoregressive (VAR) models and Long Short-Term Memory (LSTM) networks. The VAR model, a conventional tool in time-series analysis, is applied to identify and model the interrelationships and lagged effects among various factors affecting soil moisture dynamics. In contrast, the LSTM network, a sophisticated artificial intelligence method, is employed owing to its capacity to model complex, non-linear relationships and to discern long-term dependencies within the soil moisture time series data. Fig. 2 presents a concise outline of the model development process. The historical weather data underwent thorough cleaning procedures, including systematically removing rows containing missing data. The dataset was divided into two distinct subsets: a training subset and a testing subset. Subsequently, the models were trained utilizing the training subset, and their accuracy was assessed using the testing subset.

## 3. Evaluation Metrics

Considering the variability in the magnitude of soil moisture levels, multiple evaluation metrics are employed to provide a thorough assessment of the model's performance. These metrics include the mean average error (MAE), mean squared error (MSE), Root mean squared error (RMSE), $R^2$, and Overall Performance Index (OI).

MAE is calculated as the average of the absolute differences between the predicted and recorded soil moisture values. It is represented by Eq. (3.1) [41, 42]. Also, MSE measures the average of the squared differences between the predicted and recorded values, given in Eq. (3.2) [15, 41]. And, RMSE is the square root of the MSE and provides a measure of the average magnitude of the errors, presented in Eq. (3.3) [15, 41]. In addition, $R^2$ assesses the proportion of the variance in the soil moisture values that is predictable from the independent variables. An $R^2$ value of 1 signifies a perfect fit, indicating that the model accurately explains all variability in the data. Conversely, an $R^2$ value of 0 suggests that the model fails to explain any variability in the data [43]. It is computed using Eq. (3.4). And OI measures the effectiveness of the model relative to the variability in the observed data. It is calculated using Eq. (3.5) [50].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{3.1}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \tag{3.2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{3.3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}, \tag{3.4}$$

$$OI = \frac{MSE}{\text{Variance of Observed Data}}. \tag{3.5}$$

Here, $n$ represents the number of samples, $y_i$ is the recorded value, $\hat{y}_i$ signifies the predicted value and $\bar{y}_i$ indicates the mean of the recorded values.

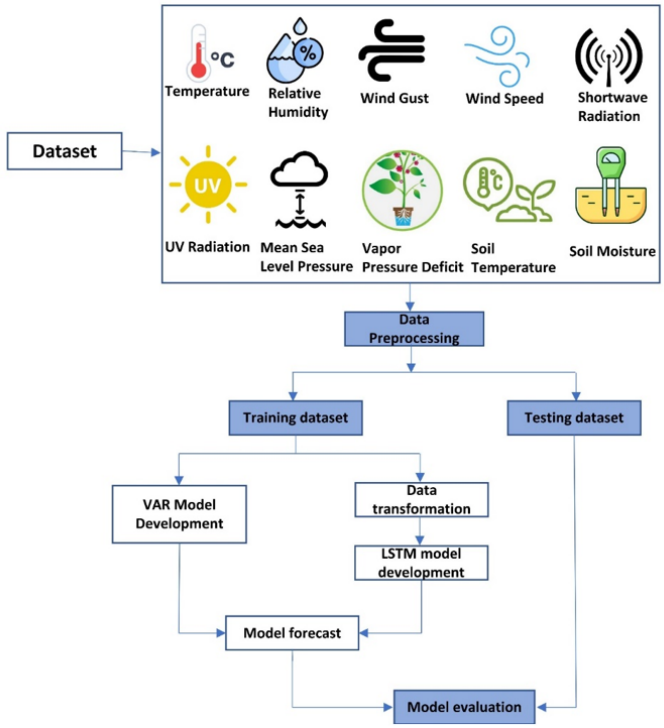


**Fig. 2.** Flow diagram of Soil Moisture forecast model development and evaluation.

### 3.1 VAR model

The development of a VAR model represents a pivotal aspect of our research, allowing us to gain valuable insights into the dynamic interactions among essential environmental variables. The VAR model serves as a flexible and robust instrument for examining the relationships and interactions across multiple time series variables. In this study, we utilized historical climate and meteorological data from 1972 to 2023, carefully preparing the dataset for the training and forecasting of our VAR model. Using the VAR model, we aimed to capture the relationships among temperature, humidity, wind characteristics, radiation, and other relevant factors that influence soil moisture dynamics.

#### *3.1.1 Selection of lag order*

An appropriate lag order is critical in time series analysis, particularly for VAR models. In our study, the lag order selection process involved identifying the optimal number of lagged variables that capture the temporal dependencies in the data while avoiding overfitting [31]. For lag order selection, the Akaike Information Criterion (AIC) method [32, 33] was employed in our study, as detailed in Table 2. After evaluating various lag orders, it was determined that the 31st lag order yielded the minimum AIC value. This selection was based on minimizing the AIC value, indicating a balance between model complexity and goodness of fit. By choosing the 31st lag order, we aim to capture the most relevant temporal dependencies in the soil moisture data while ensuring the model's effectiveness in forecasting.

**Table 2.** Selection of Order for VAR (*representing the minimum).

| Lag Order | AIC | Lag Order | AIC |
|---|---|---|---|
| 0 | -0.6447 | 17 | -9.112 |
| 1 | -8.713 | 18 | -9.114 |
| 2 | -8.927 | 19 | -9.114 |
| 3 | -8.997 | 20 | -9.114 |
| 4 | -9.026 | 21 | -9.116 |
| 5 | -9.050 | 22 | -9.117 |
| 6 | -9.063 | 23 | -9.118 |
| 7 | -9.076 | 24 | -9.120 |
| 8 | -9.083 | 25 | -9.120 |
| 9 | -9.086 | 26 | -9.121 |
| 10 | -9.093 | 27 | -9.121 |
| 11 | -9.099 | 28 | -9.120 |
| 12 | -9.102 | 29 | -9.121 |
| 13 | -9.106 | 30 | -9.121 |
| 14 | -9.109 | 31 | -9.122* |
| 15 | -9.110 | 32 | -9.122 |
| 16 | -9.111 | | |

#### *3.1.2 Construction of model equation*

The construction of the model equation in our study follows the VAR framework, as depicted in Eq. (3.6) [34]. This equation represents the relationship between the output vector $Y_t$ at time $t$, which, in our case, represents the soil moisture and its lagged values. The model includes a constant term $A_0$, the vector intercept, and coefficient matrices $A_{v,l}$ for each variable and lag combination. These coefficients capture the impact of each variable on the current soil moisture content, considering its historical values up to the lag order. Additionally, the equation incorporates the vector of exogenous variables $Y_{t-l}$ and the residual vector et at time $t$. This model equation provides a framework for analyzing the relationships between soil moisture and other relevant variables over time. From Table 2, it is evident that the 31st lag order emerged as the optimal choice, corresponding to the minimum AIC value.

$$Y_t = A_0 + \sum_{l=1}^{31} \sum_{v=1}^{10} A_{v,1} Y_{t-1} = e_t. \quad (3.6)$$

#### *3.1.3 VAR model evaluation*

Evaluating the VAR model is paramount to ensuring its reliability and efficacy in capturing the intricate dynamics of environmental variables. The VAR model is evaluated for the cropping season in Bangladesh from May to November [35] for the years 2022 and 2023. A range of statistical metrics was utilized to evaluate the model's performance, including the Durbin-Watson statistic, which offered insights into the existence of autocorrelation within the residuals. The Durbin-Watson statistic spans from 0 to 4, with a value approximating 2 signifying the absence of autocorrelation, values below 2 indicating positive autocorrelation, and values exceeding 2 suggesting negative autocorrelation. Furthermore, the evaluation matrices were computed to determine the precision and goodness of fit of the model's forecasts.

### 3.2 LSTM model

The development of the LSTM model marks a pivotal phase in our research, aiming to harness the power of neural networks for enhanced soil moisture prediction. Through an extensive training process on the designated dataset, the LSTM model learned the complex patterns and nonlinear relationships, enabling it to make accurate soil moisture-level predictions. The outcomes of this research underscore the efficacy of LSTM-based approaches in enhancing our capability to forecast and understand the dynamics of soil moisture, contributing valuable insights to sustainable water resource

management. In the training phase, a min-max scaling technique is applied to the training dataset [36], ensuring all features possess consistent ranges. This training dataset comprised daily weather records spanning 50 years, from January 1st, 1972, to April 30th, 2022. Additionally, the model is tested for the cropping season in Bangladesh, which occurs from May to November [35], for the years 2022 and 2023.

### *3.2.1 LSTM model architecture*

The architecture of the LSTM model was designed to effectively predict soil moisture based on historical data. The selection of hyperparameters was guided by the need to capture complex, non-linear atmospheric patterns while preventing overfitting [36].

The Tanh activation function was implemented across all LSTM layers. This choice is standard for LSTMs as it centers the data, aiding in the optimization process and mitigating the vanishing gradient problem over long time lags. Following the LSTM layers, Dense layers with ReLU activation were added to introduce further non-linearity and speed up convergence [37].

**Table 3.** Architecture of LSTM network.

| Layer (Type) | Nodes | Parameter | Activation Function |
|---|---|---|---|
| LSTM 1 | 128 | 70,656 | Tanh |
| LSTM 2 | 64 | 49,408 | Tanh |
| LSTM 3 | 32 | 12,416 | Tanh |
| Dense | 16 | 516 | ReLU |
| Dense (Output) | 1 | 17 | Linear |

The network consists of multiple LSTM layers with varying numbers of units (as detailed in Table 3). The specific number of layers and units was determined through iterative testing, ensuring the model was sufficiently deep to learn long-term dependencies from the 10 input variables without becoming computationally prohibitive. The final layer uses a linear activation function to output the continuous soil moisture value. Details are presented in Fig. 3.

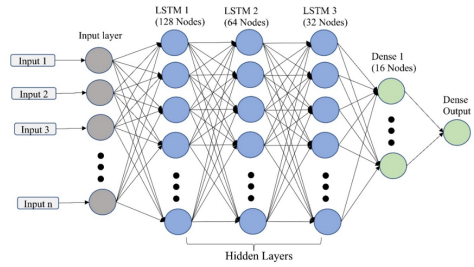


**Fig. 3.** Architecture of the LSTM model.

### *3.2.2 Training the LSTM model*

To ensure effective prediction, the LSTM model was trained on a comprehensive historical dataset spanning from January 1st, 1972, to April 30th, 2022.

The model was trained for 10 epochs. This specific number of epochs was selected by monitoring the validation loss during preliminary trials. Due to the large volume of daily records (over 19,000 observations), the model converged rapidly. We observed that the validation loss minimized within the first few epochs; extending training to a higher number of epochs resulted in over-fitting, where the model performed well on training data but poorly on unseen test data.

The model was compiled using the Mean Squared Error (MSE) loss function to penalize large errors in soil moisture prediction.

### *3.2.3 Testing the LSTM model*

The performance of the model was assessed using a one-step-ahead (1-day ahead) recursive forecasting strategy. In this setup, the model predicts soil moisture for the current day using the meteorological observations and soil moisture values from the preceding days. This process is repeated iteratively to generate the contin-

uous time-series predictions for the testing period. The testing phase encompassed a two-year period, spanning the cropping seasons from 1st May 2022 to 30th November 2022 and 1st May 2023 to 30th November 2023. Various statistical error measures were used to quantify the error between predicted and recorded values to assess the model's effectiveness. Key metrics included the coefficient of determination ($R^2$), MAE, MSE, RMSE, and Overall Performance Index (OI). These metrics provided a deeper understanding of the model's forecasting accuracy and its ability to capture the underlying dynamics of soil moisture variability. Additionally, a residual analysis was also performed to check whether the residuals were normally distributed and to verify the assumption of zero mean error. This analysis further confirmed the robustness and appropriateness of the developed model for predicting soil moisture.

## 4. Algorithm

Proposed Framework for Soil Moisture Prediction

Input: Historical time-series dataset D containing atmospheric variables $\{V_1, V_2, ..., V_n\}$ and Soil Moisture $S$.

Output: Predicted Soil Moisture vector $\hat{s}$.

i. Data Preprocessing:

a. Load dataset D.

b. Remove rows with null values.

c. Split D into Training set and Testing set.

d. Normalize features in Training set using Min-Max Scaling.

ii. Stationarity Analysis for VAR:

a. Perform ADF test on $S$.

b. If non-stationary, apply differencing until stationary.

c. SCalculate optimal lag order using Akaike Information Criterion (AIC).

d. Train VAR (p) model on Training set.

iii. Deep Learning Model Construction for LSTM):

a. Initialize LSTM network architecture.

b. Define Input layer dimension based on number of atmospheric variables n.

c. Add LSTM layers with L units and Tanh activation.

d. Add dense layer with ReLU activation.

e. Compile model with optimizer and loss function (MSE).

iv. Training and Prediction:

a. Train LSTM model on Training set for $E$ epochs.

b. Predict Soil Moisture on Testing set using $VAR(p) \longrightarrow \hat{S}_{VAR}$.

c. Predict Soil Moisture on Testing set using $LSTM \longrightarrow \hat{S}_{LSTM}$.

v. Evaluation:

a. Calculate Error Metrics (MAE, RMSE, $R^2$) for $\hat{S}_{VAR}$ and $\hat{S}_{LSTM}$ against actual $S$.

b. Select model with lowest error and highest $R^2$.

## 5. Results and Discussions

### 5.1 Correlation analysis

In this study, we utilize the Pearson correlation [29, 30] to examine relationships between dataset features. The coefficient for two variables is computed using the following formula:

$$\rho_{X,Y} = \frac{N\sum XY - \sum X \sum Y}{\sqrt{N\sum X^2 - (\sum X)^2}\sqrt{N\sum Y^2 - (\sum Y)^2}}. \quad (5.1)$$

The correlation matrix in Fig. 4 depicts the strength and direction of linear relationships between soil moisture and several other environmental factors. Soil moisture has a strong positive correlation (values closer to 1) with soil temperature (0.57), temperature (0.58), and relative humidity (0.80). As soil temperature, temperature, and relative humidity increase, soil moisture also increases. Conversely, soil moisture has a strong negative correlation (values closer to -1) with vapor pressure deficit (-0.53) and mean sea level pressure (-0.70). So, higher vapor pressure deficit and mean sea level pressure are likely to correspond with lower soil moisture.

### 5.2 Analysis of time series variables

#### *5.2.1 Data analysis on historical soil moisture*

Soil moisture data collected between 1972 and 2024 reveal an average moisture level of 30.09%. Fig. 5 illustrates the monthly distribution of soil moisture throughout a typical year, highlighting that moisture levels are above average from April to October and decrease between November and March.

#### *5.2.2 Seasonality tests of soil moisture*

The seasonal decomposition presented in Fig. 6(a) displays the measured soil moisture content at each point in the time series. Whereas Fig. 6(b) isolates the long-term trend in soil moisture content. In this case, the trend appears to fluctuate over time, which means the soil moisture content varies based on month throughout a year. Fig. 6(c) shows a repeating pattern over time, likely reflecting seasonal variations in precipitation and temperature. The seasonal pattern seems cyclical, with one peak and one trough per year. This suggests that soil moisture content is generally higher in one specific period. After accounting for the trend and seasonality, residuals represent the unexplained variations in the soil moisture data. Ideally, the residuals should be random and have no discernible pattern. The residuals might have some structure, which could indicate short-term events or measurement errors that are not captured by the trend and seasonal components.

#### *5.2.3 Autocorrelation test of variables*

Fig. 7 exhibits the temporal dependence of soil moisture content through the analysis of autocorrelation and partial autocorrelation functions. Fig. 7(a) depicts that the rapid decrease in the autocorrelation factor indicates that the soil moisture data is not random, but rather exhibits some serial dependence. This means that the current soil moisture content is related to the previous time lags (days). Fig. 7(b) indicates a distinct pattern in the partial autocorrelation, where successive lags alternate in a positive-negative-positive sequence. This function reveals a significant positive correlation for the first lag. Following this, there's a significant negative correlation at the second lag, and then potentially another positive correlation at the third lag.

The alternating positive-negative-positive correlation pattern observed in the

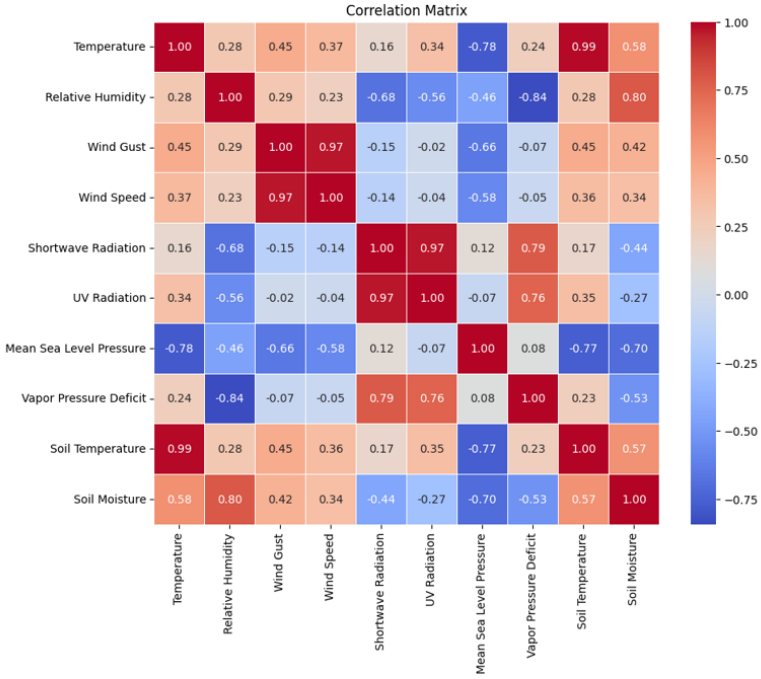


**Fig. 4.** Correlation coefficient of variables.

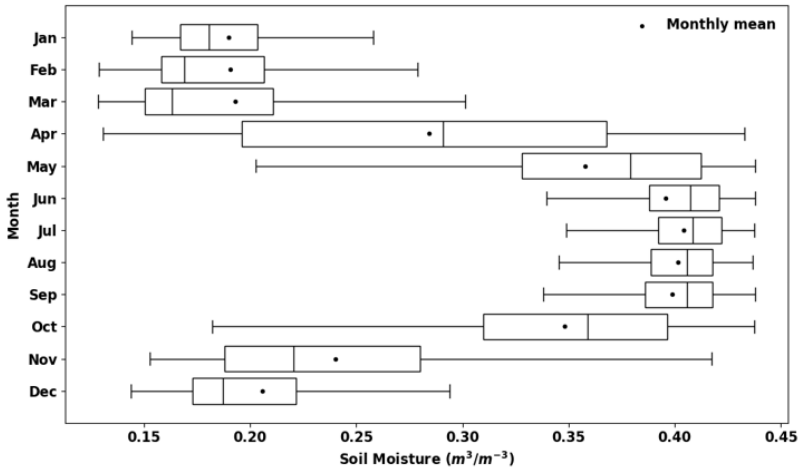


**Fig. 5.** Soil moisture over 52 years span by month (1972-2024).

partial autocorrelation function indicates cyclical or oscillatory dynamics between the current soil moisture content ($n^{th}$ day) and its past values ($(n - 1)^{th}$ days).

Table 4 displays the outcomes of the Durbin-Watson test conducted for various variables in our study. The Durbin-Watson statistic assesses the existence of autocorrelation in the residuals of a regression analysis. A value approximating 2 indicates the absence of autocorrelation, whereas values substantially deviating from 2 suggest the presence of autocorrelation [44, 45]. For the variables Temperature, Rel-

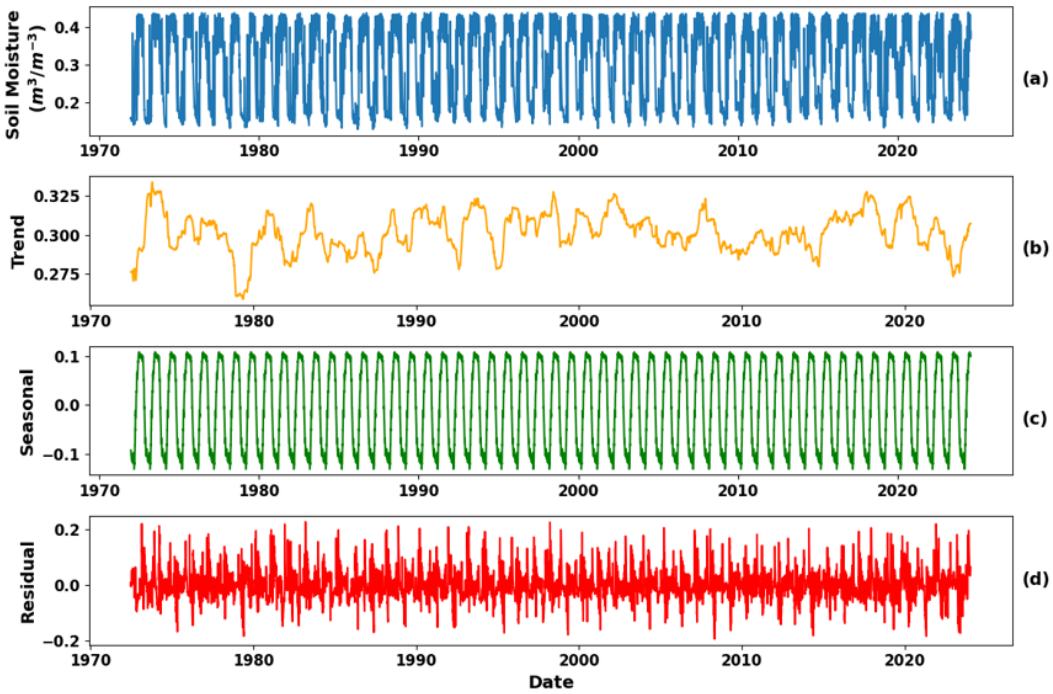


**Fig. 6.** Seasonal Decomposition of soil moisture time series: (a) value of soil moisture, (b) trend, (c) seasonality, (d) Residual.

**Table 4.** Durbin-Watson test.

| Variables | Durbin Watson Score |
|---|---|
| Temperature | 0.001505 |
| Relative Humidity | 0.00289 |
| Wind Gust | 0.0452818 |
| Wind Speed | 0.06187 |
| Shortwave Radiation | 0.041708 |
| UV Radiation | 0.037921 |
| Mean Sea Level Pressure | 1.6436130 |
| Vapor Pressure Deficit | 0.048163 |
| Soil Temperature | 0.000845 |
| Soil Moisture | 0.0038337 |

ative Humidity, Wind Gust, Wind Speed, Shortwave Radiation, UV Radiation, Vapor Pressure Deficit, and Soil Moisture, the Durbin-Watson scores range from 0.000845 to 0.06187, indicating very low autocorrelation. However, the Mean Sea Level Pressure variable exhibits a Durbin-Watson score of 1.6436130, indicating the existence of some autocorrelation in the residuals. This insight is vital for assessing the regression model's reliability and interpreting the coefficients' significance.

#### *5.2.4 Stationarity test*

The Augmented Dickey-Fuller (ADF) test [46] is known as a statistical method used to identify if a time series data is stationary. In our analysis, all variables are considered including both the input and target variables. The ADF scores for all variables are negative, and the results are displayed in Table 5. A negative ADF score indicates a greater likelihood of stationarity in the data. Additionally, a more negative ADF score indicates a stronger indication of stationarity for the variables. In a stationary pattern of data, the ADF score is lower than the critical value for a certain confidence interval (such as 1%, 5%, or 10%). The ADF score

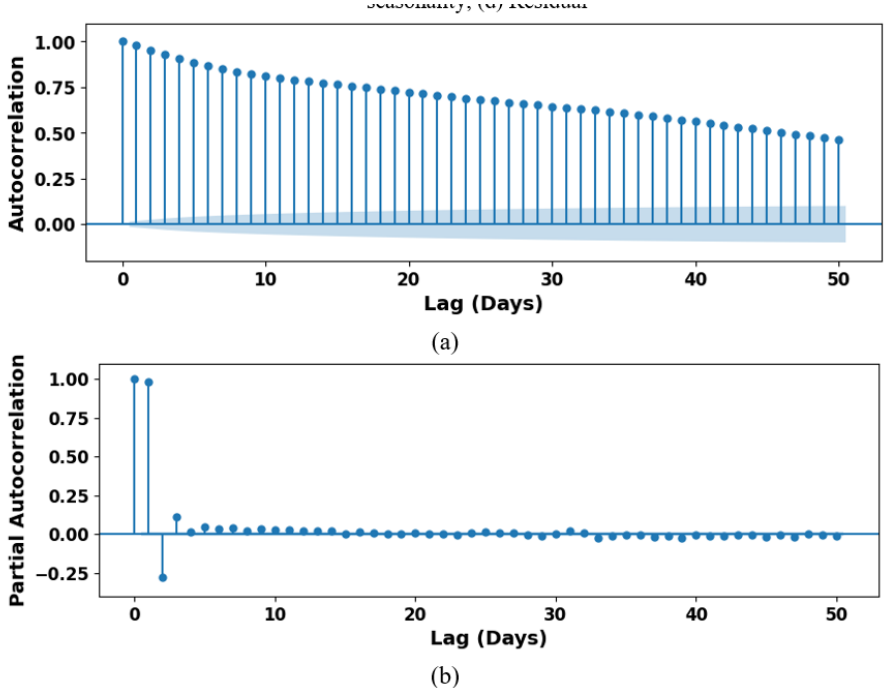


**Fig. 7.** (a) Autocorrelation (b) Partial autocorrelation in soil moisture time series.

for all variables was less than the critical values at a 1%, 5%, and 10% confidence interval. This indicates that all variables are significantly stationary. Furthermore, the *p*-values displayed in the table also support the conclusion that the time series data for each variable is stationary. Specifically, all *p*-values were less than 0.01 for every variable.

### *5.2.5 Cointegration test*

A cointegration analysis was performed to assess long-term equilibrium relationships between potential influencing factors and soil moisture. The findings are compiled in Table 6, which lists the test statistic for each weather variable along with the critical value at a 95% confidence level. This critical value serves as a benchmark for statistical significance [47–49]. Variables with test statistic values exceeding the critical value are considered to have a cointegrated relationship with soil moisture. In simpler terms, this indicates that these variables exhibit a long-term equilibrium relationship with soil moisture. The table allows us to identify key weather factors that significantly influence soil moisture in the long term. For instance, variables like temperature and relative humidity might exhibit cointegration, suggesting that they share a stable, long-term association with soil moisture. These provide valuable insights into the influence of weather patterns on soil moisture dynamics.

### *5.2.6 Forecastability test*

The results of the Granger causality test, as presented in Table 7, suggest that all the input variables could be valuable for predicting soil moisture. These include temperature (T), relative humidity (RH), wind gust (WG), wind speed (WS), shortwave radiation (SR), UV radiation (UVR),

**Table 5.** Augmented Dickey-Fuller (ADF) test.

| Variables | ADF Score | $p$-Value | Critical Values | | |
|---|---|---|---|---|---|
| | | | 1% | 5% | 10% |
| Temperature | -14.9422 | $1.32 \times 10^{-27}$ | -3.431 | -2.862 | -2.567 |
| Relative Humidity | -11.1498 | $2.96 \times 10^{-20}$ | -3.431 | -2.862 | -2.567 |
| Wind Gust | -10.0161 | $1.72 \times 10^{-17}$ | -3.431 | -2.862 | -2.567 |
| Wind Speed | -11.0307 | $5.68 \times 10^{-20}$ | -3.431 | -2.862 | -2.567 |
| Shortwave Radiation | -12.5404 | $2.32 \times 10^{-23}$ | -3.431 | -2.862 | -2.567 |
| UV Radiation | -12.6489 | $1.38 \times 10^{-23}$ | -3.431 | -2.862 | -2.567 |
| Mean Sea Level Pressure | -10.2711 | $3.99 \times 10^{-18}$ | -3.431 | -2.862 | -2.567 |
| Vapor Pressure Deficit | -15.6877 | $1.45 \times 10^{-28}$ | -3.431 | -2.862 | -2.567 |
| Soil Temperature | -14.9388 | $1.33 \times 10^{-27}$ | -3.431 | -2.862 | -2.567 |
| Soil Moisture | -10.7398 | $2.84 \times 10^{-19}$ | -3.431 | -2.862 | -2.567 |

**Table 6.** Cointegration test.

| Variables | Test Statistic | Critical Value at 95% Confidence Level |
|---|---|---|
| Temperature | 26337.8404 | 197.3772 |
| Relative Humidity | 20259.5438 | 159.5290 |
| Wind Gust | 15719.1690 | 125.6185 |
| Wind Speed | 11398.2819 | 95.7542 |
| Shortwave Radiation | 7612.4764 | 69.8189 |
| UV Radiation | 4362.9965 | 47.8545 |
| Mean Sea Level Pressure | 2385.1978 | 29.7961 |
| Vapor Pressure Deficit | 973.6506 | 15.4943 |
| Soil Temperature | 139.4316 | 3.8415 |

mean sea level pressure (MSLP), vapor pressure deficit (VPD), soil temperature (ST), and soil moisture (SM). This test confirms the potential of weather-related inputs to forecast changes in soil moisture levels. The significance of the findings is determined by the p-values, which are below 0.0001. This suggests a significant statistical correlation between the soil moisture time series and meteorological inputs. In other words, historical data on weather variables significantly predict soil moisture content [46]. The statistical tests and analyses conducted indicate a strong relationship between soil moisture and weather conditions. Moreover, they suggest that historical weather data, along with past soil moisture, can be utilized to forecast future soil moisture levels effectively.

## 5.3 Evaluation of soil moisture forecast

### *5.3.1 Evaluation of the VAR model*

Predicting soil moisture levels for the cropping seasons of 2022 and 2023 was used to assess the VAR model's performance. Fig. 8(a) displays the forecasted and observed soil moisture values from May 1, 2022, to November 30, 2022, while Fig. 8(b) illustrates the corresponding data for the period from May 1, 2023, to November 30, 2023. The model's evaluation metrics are presented in Table 8.

For the two forecasted years, 2022 and 2023, the error values (MAE, MSE, and RMSE) are significantly small. However, the $R^2$ values are -0.0318 for the 2022 crop season and 0.2729 for the 2023 crop season. These $R^2$ values indicate that while the VAR model yields low error metrics, it does not effectively capture day-to-day variations in the data. The negative $R^2$ value for 2022 suggests that the model's performance was suboptimal for that year, and the $R^2$ value for 2023, though positive, still indicates limited explanatory power.

**Table 7.** Gragner causality test.

| Variables | T | RH | WG | WS | SR | UVR | MSLP | VPD | ST | SM |
|---|---|---|---|---|---|---|---|---|---|---|
| T | 1 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| RH | <0.0001 | 1 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| WG | <0.0001 | <0.0001 | 1 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| WS | <0.0001 | <0.0001 | <0.0001 | 1 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| SR | <0.0001 | <0.0001 | <0.0001 | <0.0001 | 1 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| UVR | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | 1 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| MSLP | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | 1 | <0.0001 | <0.0001 | <0.0001 |
| VPD | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | 1 | <0.0001 | <0.0001 |
| ST | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | 1 | <0.0001 |
| SM | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | 1 |

This observation is consistent with Fig. 8, which demonstrates that the VAR model struggles with forecasting short-term fluctuations. The VAR model struggles to forecast day-to-day variations in daily weather data primarily because it assumes linear relationships between variables and focuses on long-term trends rather than short-term fluctuations. This linear assumption makes it difficult for the model to capture the non-linear patterns often present in weather data, such as sudden temperature spikes or drops. Additionally, the model's reliance on past values (lags) means it may not effectively capture rapid weather changes that occur from one day to the next. Furthermore, the VAR model isn't well-equipped to handle sudden external factors like unexpected storms or shifts in wind patterns, which can significantly influence daily weather outcomes. Its tendency towards mean-reversion also leads to underestimating the magnitude of daily weather variations, particularly during extreme or volatile conditions. These limitations make the VAR model less effective for accurately forecasting day-to-day weather variations.

### *5.3.2 Evaluation of the LSTM network model*

Fig. 9 presents a visual comparison of recorded soil moisture value with the predictions made by the LSTM model for the cropping seasons of 2022 and 2023. This LSTM model successfully captures the general seasonal trends in soil moisture content for both cropping seasons. This is evident from the similarity between the recorded

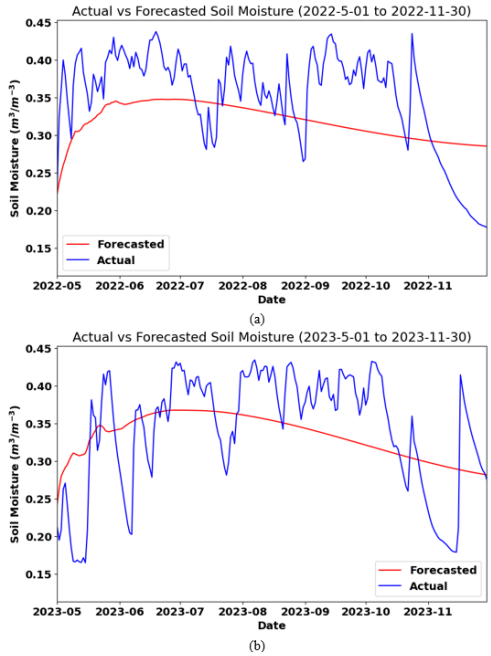


**Fig. 8.** Cross-sectional Scale-up 1.5× with Fixed Froude Number: (a) Downstream 1 (200 mm after junction); (b) Downstream 2 (200 mm after junction); (c) Outlet 1; (d) Outlet 2.

**Table 8.** Performance evaluation of VAR model for soil moisture.

| Parameters | Test for Cropping Season 2022 | Test for Cropping Season 2023 |
|---|---|---|
| MAE | 0.0571 | 0.0559 |
| MSE | 0.0043 | 0.0044 |
| RMSE | 0.0653 | 0.0665 |
| $R^2$ | -0.0318 | 0.2729 |
| Durbin-Watson Statistics | 2.0000 | 2.0000 |

and predicted values in both years. While there are a few instances where the model predicts higher soil moisture than what occurred, the differences between the predicted and recorded values are relatively slight overall. This suggests that the LSTM model accurately forecasts soil moisture values.

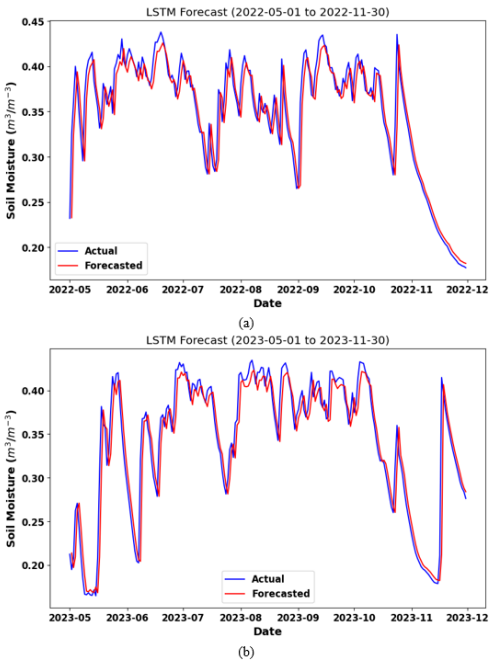


**Fig. 9.** Recorded and LSTM model forecasted soil moisture for (a) 2022 and (b) 2023 cropping seasons.

Fig. 10 depicts an analysis of the LSTM model's performance in predicting soil moisture. The distribution of data points relative to the diagonal line, representing a scenario of perfect prediction, provides key insights into the model's performance. Figs. 10(a)-(b) reveal a tight clustering of data points around the diagonal line, suggesting strong concordance between the recorded soil moisture and the model's predictions. Table 9 provides an assessment of the performance and predictive accuracy of the model applied to soil moisture prediction during the cropping seasons of 2022 and 2023. These metrics serve as vital indicators of the model's effectiveness in identifying the inherent patterns and variability in the observed data. The model fit is $R^2$ = 0.9971 and $R^2$ = 0.9977 for the cropping seasons of 2022 and 2023, respectively. Furthermore, the low values of MAE, MSE, and RMSE indicate minimal errors in the model's predictions [51], confirming its reliability for practical applications in agriculture. Overall Performance Index (OI) offers a comprehensive evaluation of the model's predictive performance by considering both the variance captured by the model and the magnitude of errors in the predictions [50]. The OI values of 0.9972 for the 2022 cropping season and 0.9977 for the 2023 cropping season indicate strong agreement between predicted and recorded soil moisture values. This suggests that the LSTM model accurately captures underlying data patterns, enhancing confidence in its predictive capabilities. This evaluation metric indicates that this model can be more viable than conventional empirical equations in soil moisture forecasting.

Fig. 11 illustrates the distribution of errors between the predicted and recorded soil moisture content for the cropping seasons of 2022 and 2023. The residuals appear relatively generally distributed for both cropping seasons, as evidenced by the close resemblance between the bars and the fitted standard distribution curve. This implies that the discrepancies between the predicted and recorded soil moisture content are random and independent, with no substantial bias evident in the model's forecasts.

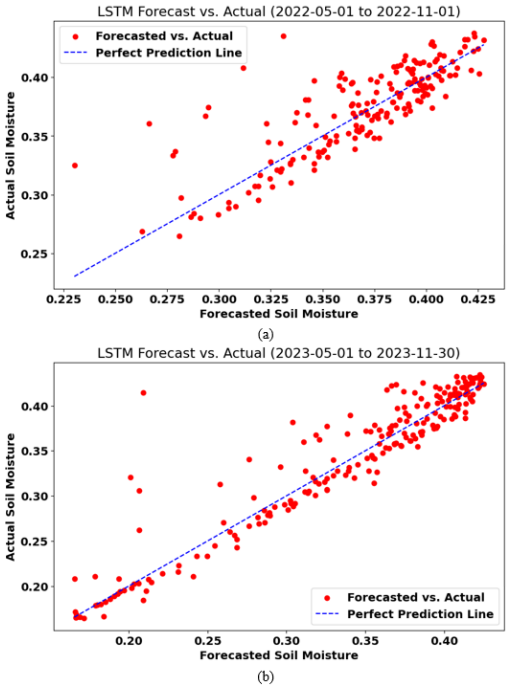


**Fig. 10.** LSTM model fit for soil moisture (a) 2022 and (b) 2023 cropping seasons.

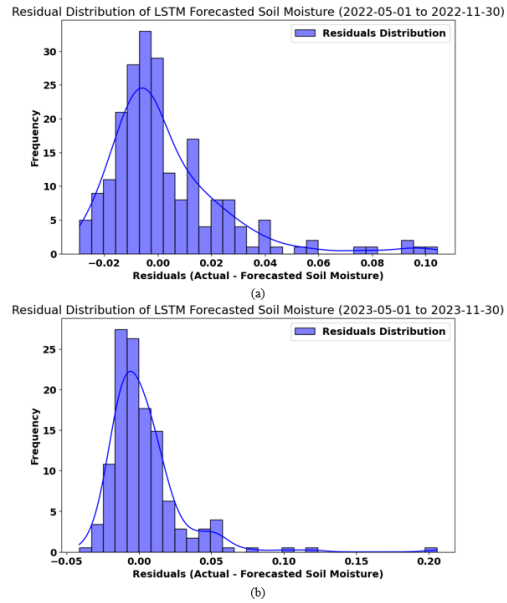


**Fig. 11.** Distribution of residuals for LSTM-predicted soil moisture during the (a) 2022 and (b) 2023 cropping seasons.

### *5.3.3 Iterative 1-day ahead forecasting for future periods*

Fig. 12 illustrates an LSTM model's prediction of soil moisture from December 1, 2023, to May 31, 2024. The blue line represents recorded soil moisture measurements up to early December 2023, while the red line shows the model's prediction. The predicted soil moisture values range from approximately 0.15 to 0.40, exhibiting cyclical patterns with notable fluctuations. The model suggests lower soil moisture levels during winter months and higher levels in spring, potentially reflecting seasonal precipitation patterns. Short-term variations in the prediction may indicate the model's attempt to capture individual weather events. The smooth transition between recorded and predicted values suggests the LSTM model has identified underlying patterns in the historical data. However, without uncertainty bounds, it's challenging to assess the model's confidence in its longer-term predictions. This forecast could be valuable for agricultural planning and water resource management.

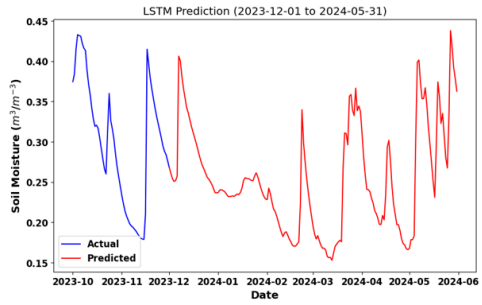


**Fig. 12.** Predicted soil moisture result from December 1, 2023 to May 31, 2024.

## 5.4 Comparison of LSTM and VAR model

An evaluation of the VAR model and LSTM networks was conducted to assess their effectiveness in forecasting soil mois-

ture for the 2022 and 2023 cropping seasons. Both models showed relatively low MAE, MSE, and RMSE values, indicating minimal differences between predicted and recorded soil moisture. However, the VAR model exhibited a negative $R^2$ value for 2022, signaling poor performance in capturing day-to-day variations. The LSTM model consistently outperformed the VAR model, achieving higher $R^2$ values across both years. This superior performance is attributed to the LSTM's ability to handle non-linear relationships and complex temporal dependencies more effectively than the VAR model. The enhanced accuracy and robustness of the LSTM in forecasting soil moisture underscore its advantages and superiority over traditional forecasting methods. These findings demonstrate the LSTM model's potential for providing more reliable and precise predictions in soil moisture forecasting and agricultural decision-making.

## 6. Conclusion

The study employed time series analysis to examine the relationship between soil moisture and a comprehensive set of weather variables. Although this study utilized data from Dhaka Division, the proposed methodological framework is region-agnostic. The mathematical models and algorithmic workflow presented herein are applicable to any geographic region provided the requisite atmospheric input variables are available. Researchers in diverse climatic zones can adapt this framework to predict local soil moisture dynamics by simply training the models on region-specific historical datasets. The present study evaluated the efficacy of two established forecasting models: VAR and LSTM networks. While both models demonstrated statistically sound performance, the LSTM network outperformed the VAR model in terms of accuracy and trend-tracking capabilities. Specifically, for the 2022 cropping season, the VAR model achieved an MAE of 0.0571, MSE of 0.0043, RMSE of 0.0653, and an $R^2$ value of -0.0318. In the 2023 cropping season, the VAR model showed a slight improvement with an MAE of 0.0559, MSE of 0.0044, RMSE of 0.0665, and an $R^2$ of 0.2729. In contrast, the LSTM model demonstrated superior performance with an MAE of 0.0027, MSE of 0.000031, RMSE of 0.0034, and an impressive $R^2$ of 0.9971 for the 2022 season. For the 2023 season, the LSTM model maintained high accuracy with an MAE of 0.003, MSE of 0.000014, RMSE of 0.0037, and an $R^2$ of 0.9977. Additionally, the LSTM model has been successfully used to predict soil moisture from December 1, 2023, to May 31, 2024. These findings emphasize the remarkable reliability and precision of the LSTM model in soil moisture forecasting, underscoring its potential as a robust tool for the agricultural industry. The outstanding performance of the LSTM network highlights its significance in enhancing soil moisture management and meeting the practical demands of agriculture.

Furthermore, this research highlights that while deep learning models require rigorous hyper parameter tuning and a substantial historical dataset, the gain in predictive reliability justifies the computational investment. Future implementations should focus on optimizing this architecture for real-time deployment which suggests a hybrid approach could further bridge the gap between computational efficiency and the high accuracy achieved by LSTM in this study.

## Acknowledgments

First and Second author acknowledge the following grants: NSU CTRGC Grant No.: CTRG-23-SEPS-31. The Ministry of Science and Technology (MOST) (SRG-246481), Government of the People's Republic of Bangladesh.

## References

[1] LeCun Y, Haffner P, Bottou L, Bengio Y, van Leeuwen J, Hartmanis J, et al. Object recognition with gradient-based learning. In: Lecture Notes in Computer Science. Germany: Springer Berlin/Heidelberg; 1999. p. 319-45.

[2] Rumelhart DE. Learning representations by backpropagating errors. Nature. 1986;323:533-6.

[3] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735-80.

[4] Feng D, Fang K, Shen C. Enhancing streamflow forecast and extracting insights using long short-term memory networks with data integration at continental scales. Water Resour Res. 2020;56(9).

[5] Sheikhi F, Kowsari Z. Time series forecasting of COVID-19 infections and deaths in Alpha and Delta variants using LSTM networks. PLoS One. 2023;18(10).

[6] Mosavi A, Ozturk P, Chau KW. Flood prediction using machine learning models: Literature review. Water. 2018;10(11).

[7] Solomatine DP, Ostfeld A. Data-driven modelling: some past experiences and new approaches. J Hydroinform. 2008;10(1):3-22.

[8] Usmani S, Shamsi JA. LSTM based stock prediction using weighted and categorized financial news. PLoS One. 2023;18(3).

[9] Baruah RD, Roy S, Bhagat RM, Sethi LN. Use of data mining technique for prediction of tea yield in the face of climate change of Assam, India. In: 2016 International Conference on Information Technology (ICIT). IEEE; 2016. p. 265-9.

[10] Han J, Mao K, Xu T, Guo J, Zuo Z, Gao C. A soil moisture estimation framework based on the CART algorithm and its application in China. J Hydrol. 2018;563:65-75.

[11] Zhang H, Liu J, Li H, Meng X, Ablikim A. The impacts of soil moisture initialization on the forecasts of weather research and forecasting model: A case study in Xinjiang, China. Water (Basel). 2020;12(7).

[12] Fereres E, Goldhamer DA, Parsons LR. Irrigation water management of horticultural crops. HortScience. 2003;38(5):1036-42.

[13] Ballester C, Brinkhoff J, Quayle WC, Hornbuckle J. Monitoring the effects of water stress in cotton using the green red vegetation index and red edge ratio. Remote Sens. 2019;11(7).

[14] Datta P, Faroughi SA. A multihead LSTM technique for prognostic prediction of soil moisture. Geoderma. 2023;433.

[15] Cai Y, Zheng W, Zhang X, Zhangzhong L, Xue X. Research on soil moisture prediction model based on deep learning. PLoS One. 2019;14(4).

[16] Moghadas D, Badorreck A. Machine learning to estimate soil moisture from geophysical measurements of electrical conductivity. Near Surf Geophys. 2019;17(2):181-95.

[17] Greifeneder F, Notarnicola C, Wagner W. A machine learning-based approach for surface soil moisture estimations with Google Earth Engine. Remote Sens. 2021;13(11).

[18] Wai KP, Chia MY, Koo CH, Huang YF, Chong WC. Applications of deep learning in water quality management: A state-of-the-art review. J Hydrol. 2022;613.

[19] Jiang S, Chen G, Chen D, Chen T. Application and evaluation of an improved LSTM model in the soil moisture prediction of southeast Chinese tobacco-producing areas. J Indian Soc Remote Sens. 2023;51(9):1843-53.

[20] Singh S, Kaur S, Kumar P, Singh SN, Panigrahi BK, Kothari DP, et al. Forecasting soil moisture based on evaluation of time series analysis. In: Advances in Power and Control Engineering. Singapore: Springer; 2019. p. 145-56.

[21] Huang C, Li L, Ren S, Zhou Z, Liu Y, Chen Y, et al. Research of soil moisture content forecast model based on genetic algorithm BP neural network. In: Computer and Computing Technologies in Agriculture IV. Berlin: Springer; 2011. p. 309-16.

[22] Wu Z, Cui N, Zhang W, Liu C, Jin X, Gong D, et al. Estimating soil moisture content in citrus orchards using multi-temporal Sentinel-1A data-based LSTM and PSO-LSTM models. J Hydrol. 2024;637.

[23] Han H, Choi C, Jung J, Kim HS. Deep learning with long short-term memory based sequence-to-sequence model for rainfall-runoff simulation. Water (Basel). 2021;13(4).

[24] Filipović N, Brdar S, Mimić G, Marko O, Crnojević V. Regional soil moisture prediction system based on long short-term memory network. Biosyst Eng. 2022;213:30-8.

[25] Mok JY, Choi JH, Moon YI. Prediction of multipurpose dam inflow using deep learning. J Korea Water Resour Assoc. 2020;53(2):97-105.

[26] Bai P, Liu X, Xie J. Simulating runoff under changing climatic conditions: A comparison of the long short-term memory network with two conceptual hydrologic models. J Hydrol. 2021;592.

[27] Lu F, Zheng Y, Cleveland H, Burton C, Madigan D. Bayesian hierarchical vector autoregressive models for patient-level predictive modeling. PLoS One. 2018;13(12).

[28] Basir MS, Noel S, Buckmaster D, Ashik-E-Rabbani M. Enhancing subsurface soil moisture forecasting: A long short-term memory network model using weather data. Agriculture (Basel). 2024;14(3).

[29] Yan J, Hu L, Zhen Z, Wang F, Qiu G, Li Y, et al. Frequency-domain decomposition and deep learning based solar PV power ultra-short-term forecasting model. IEEE Trans Ind Appl. 2021;57(4).

[30] Hutson AD. A robust Pearson correlation test for a general point null using a surrogate bootstrap distribution. PLoS One. 2019;14(5).

[31] Ozcicek O, Douglas MW. Lag length selection in vector autoregressive models: Symmetric and asymmetric lags. Appl Econ. 1999;31(4):517-24.

[32] Li H, Yang Z, Yan W. An improved AIC onset-time picking method based on regression convolutional neural network. Mech Syst Signal Process. 2022;171.

[33] Rajakumar G, Du KL, Rocha Á. Intelligent communication technologies and virtual mobile networks: Proceedings of ICICV 2023. Singapore: Springer; 2023.

[34] Holtz-Eakin D, Newey W, Rosen HS. Estimating vector autoregressions with panel data. Econometrica. 1988;56(6):1371-95.

[35] Mohsenipour M, Shahid S, Chung ES, Wang XJ. Changing pattern of droughts during cropping seasons of Bangladesh.

Water Resour Manag. 2018;32(5):1555-68.

[36] Pires IM, Hussain F, Garcia NM, Lameski P, Zdravevski E. Homogeneous data normalization and deep learning: A case study in human activity classification. Future Internet. 2020;12(11).

[37] Namin AH, Leboeuf K, Muscedere R, Wu H, Ahmadi M. Efficient hardware implementation of the hyperbolic tangent sigmoid function. In: 2009 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE; 2009.

[38] Agarap AF. Deep learning using rectified linear units (ReLU). arXiv. 2018.

[39] Reyad M, Sarhan AM, Arafa M. A modified Adam algorithm for deep neural network optimization. Neural Comput Appl. 2023;35(23):17095-112.

[40] Singh S, Kumar S, Singh DS, Satapathy SC, Biswal BN, Udgata SK. Modified mean square error algorithm with reduced cost of training and simulation time for character recognition in back propagation neural network. In: FICTA 2013. Cham: Springer; 2014. p. 137-45.

[41] Zhang B, Zhang Y, Jiang X. Feature selection for global tropospheric ozone prediction based on the BO-XGBoost-RFE algorithm. Sci Rep. 2022;12(1).

[42] Biazar SM, Fard AF, Singh VP, Dinpashoh Y, Majnooni-Heris A. Estimation of evaporation from saline-water with more efficient input variables. Pure Appl Geophys. 2020;177(11):5599-619.

[43] Chicco D, Warrens MJ, Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Comput Sci. 2021;7.

[44] Dufour JM, Dagenais MG. Durbin-Watson tests for serial correlation in regressions with missing observations. J Econometrics. 1985;27(3):371-81.

[45] King ML. The alternative Durbin-Watson test: An assessment of Durbin and Watson's choice of test statistic. J Econometrics. 1981;17(1):51-66.

[46] Jalil A, Rao NH, Ozcan B, Ozturk I. Time series analysis (stationarity, cointegration, and causality). In: Environmental Kuznets Curve (EKC): A Manual. London: Elsevier; 2019. p. 85-99.

[47] Franses PH. A multivariate approach to modeling univariate seasonal time series. J Econometrics. 1994;63(1):133-51.

[48] Campos J, Ericsson NR, Hendry DF. Cointegration tests in the presence of structural breaks. J Econometrics. 1996;70(1):187-220.

[49] Ozturk I, Al-Mulali U. Investigating the validity of the environmental Kuznets curve hypothesis in Cambodia. Ecol Indic. 2015;57:324-30.

[50] Salaeh N, Ditthakit P, Pinthong S, Hasan MA, Islam S, Mohammadi B, et al. Long short-term memory technique for monthly rainfall prediction in Thale Sap Songkhla River Basin, Thailand. Symmetry (Basel). 2022;14(8).

[51] Behroozi-Khazaei N, Nasirahmadi A. A neural network based model to analyze rice parboiling process with small dataset. J Food Sci Technol. 2017;54(8):2562-9.