Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.5, October 2025, pp. 2230~2238
e-ISSN: 2548-6861

A Predictive Model For Crop Irrigation Schedulling Using Machine Learning and IoT-Generated Environmental Data

Rizki Agam Syahputra [1]*, Dewi Andriani [2]**
* Industrial Engineering Department, Universitas Teuku Umar
** Agrotechnology Department, Universitas Teuku Umar
rizkiagamsyahputra@utu.ac.id [1], dewiandriani@utu.ac.id [2]

Article Info

Article history:

Received 2025-07-12
Revised 2025-08-01
Accepted 2025-08-10

Keyword:

IoT,
Machine Learning,
Smart Farming.

ABSTRACT

This study develops and evaluates a machine learning model for predicting optimal irrigation schedules using real-time environmental data collected from an Internet of Things (IoT) system. Building upon a previously validated smart farming monitoring system that provided real-time data on temperature, humidity, and soil moisture, this research addresses the next step: moving from monitoring to predictive analytics. Data collected over a six-day period from DHT11 temperature and humidity sensors, as well as soil moisture sensors, were used to train a predictive model. The model is designed to forecast future soil moisture levels, thereby providing farmers with proactive recommendations for irrigation. A Long Short-Term Memory (LSTM) neural network was employed to capture the temporal dependencies between atmospheric conditions and soil moisture. The model was trained on a portion of the collected data and then validated on a separate, unseen dataset. The evaluation yielded a Mean Absolute Error (MAE) of 2.5%, a Root Mean Square Error (RMSE) of 3.1%, and an R-squared (R^2) value of 0.92, demonstrating high predictive accuracy. This approach aims to enhance water resource management, reduce manual intervention, and improve crop health by ensuring water is supplied only when necessary. The results indicate that the machine learning model can accurately predict irrigation needs, offering a significant improvement over traditional, reactive monitoring systems and marking a substantial step towards data-driven, precision agriculture.

I. INTRODUCTION

The agricultural sector serves as the backbone of global food security and economic stability, a reality that holds particular significance for developing nations like Indonesia. Within Indonesia, the province of Aceh exemplifies the sector's importance, contributing a substantial 28.5% to the region's overall Gross Regional Domestic Product (GRDP) [1], [2]. Despite this vital economic role, a significant portion of Aceh's agriculture remains entrenched in conventional and traditional methods. These long-standing practices, while culturally significant, are increasingly proving insufficient to overcome the complex challenges of modern agriculture, including the escalating impacts of climate change, pronounced resource limitations, and a growing consumer demand for higher-quality produce.

Traditional farming techniques, which often depend on generational knowledge, intuition, and intensive manual labor, lack the efficiency and precision needed to optimize resource allocation and maximize yields in today's demanding environment [3]. This approach frequently leads to significant inefficiencies, such as the over-application of water and agrochemicals. Such practices not only deplete vital resources like water but also contribute to environmental degradation, including soil health deterioration and water source pollution. The urgent need to modernize the agricultural sector is driven by this imperative to produce more food with fewer resources, sustainably feeding a growing population while safeguarding the environment for future generations [4].

In response to these challenges, Precision Agriculture has emerged as a transformative paradigm. At its core, precision agriculture leverages advanced technology to collect and analyze accurate, real-time data, enabling data-driven management of every aspect of crop production. A key enabling technology in this field is the Internet of Things (IoT), which facilitates the creation of sophisticated monitoring systems. By deploying a network of sensors, IoT systems can continuously track critical environmental variables such as soil moisture, ambient temperature, and humidity, providing farmers with an unprecedented, granular view of their field conditions in real time [5], [6], [7].

Building on this technological foundation, our previous research successfully designed, developed, and tested a smart farming monitoring system at the Universitas Teuku Umar (UTU) farm. That foundational study focused on creating a robust hardware and software ecosystem. The system integrated an ESP32 microcontroller with DHT11 temperature and humidity sensors and soil moisture sensors. These components wirelessly transmitted real-time farm data to a custom-built Android application, allowing farmers to remotely monitor environmental conditions. A critical phase of that initial work involved validating the system's accuracy; through linear regression analysis, we demonstrated a strong correlation between our sensor data and measurements from standard external instruments, confirming the system's reliability for data collection [8] .

However, the capability of that initial system was fundamentally *reactive*. It provided farmers with a high-quality stream of real-time data, but still required them to constantly interpret this information and manually decide on the appropriate course of action, such as when and how much to irrigate. This leaves a significant operational gap and a barrier to full optimization. The true evolution of smart farming lies in transitioning from this reactive monitoring to a *proactive, predictive* framework. This study aims to bridge this critical gap. The primary objective of this research is to leverage the validated data from our previous work to develop and evaluate a sophisticated predictive model using machine learning. Specifically, we employ a Long Short-Term Memory (LSTM) neural network a model well-suited for time-series data to forecast future soil moisture levels. The goal is to create an intelligent system that moves beyond simply reporting what is happening to accurately predicting what will happen, thereby enabling automated and optimized irrigation recommendations. This represents a significant leap forward, transforming the IoT system from a simple monitoring tool into an intelligent decision-support platform for truly precise and sustainable agriculture.

II. Methods

The methodology for this study is structured to build directly upon previous work, leveraging a validated dataset to develop and evaluate a sophisticated predictive model. The process encompasses the foundational data acquisition system, data preprocessing, the architecture of the predictive model, and the metrics used for its evaluation.

A. Foundation of the Study: The Previously Published IoT Monitoring System

This research is a direct extension of a smart farming monitoring system that was previously developed, deployed, and published. The initial study focused on the complete lifecycle of the IoT system, from its design and implementation to its testing and validation at the Universitas Teuku Umar (UTU) Farm [8]. The architecture of that system, which serves as the data source for this paper, is focused on an ESP32 microcontroller that functions as the central processing unit. This controller integrates data from multiple sensors, including a DHT11 sensor for measuring ambient temperature and humidity and multiple FC-28 soil moisture sensors to monitor soil conditions. The collected data was then transmitted wirelessly via a Wi-Fi protocol to a custom-developed Android application, providing farmers with a user-friendly interface for remote, real-time monitoring of their farm's environmental conditions.

Figure 1. Detailed IoT Monitoring System [9]

To enhance user interaction and usability, the IoT monitoring system is integrated with an Android-based user interface. This mobile application allows farmers to remotely access real-time environmental data such as temperature, humidity, and soil moisture directly from their smartphones. The interface is designed to be user-friendly and lightweight, supporting visualization features including trend graphs and historical records of environmental variables. In future iterations, this interface is planned to incorporate predictive outputs from the LSTM model, enabling farmers to receive automated irrigation alerts and recommendations based on anticipated soil moisture levels. This integration represents a crucial step toward a fully autonomous smart farming system.

A critical component of the foundational study was the rigorous validation of the sensor data. To ensure the system's reliability, the readings from the IoT sensors were systematically compared against standard, calibrated instruments, specifically a Hygrometer HTC-2 and a soil meter device. Linear regression analysis was the statistical method used to assess the correlation between the IoT sensor readings and the reference instruments. The analysis yielded strong positive correlations, with regression coefficients of 0.95 for temperature and 0.94 for humidity, confirming the high accuracy of the data produced by the DHT11 sensor. The soil moisture sensor demonstrated a near-perfect correlation with a coefficient of 1.00.

Figure 2. Field testing and data collection point [9]

Therefore, the previous work successfully established a robust and validated data acquisition platform. The high-quality, reliable time-series dataset generated and verified in that study is essential for and serves as the direct input for the training and evaluation of the predictive machine learning models presented in this paper.

B. Dataset for predictive modelling

The dataset used for this study was collected by the validated IoT system over a period of six consecutive days. On each day, measurements for temperature, humidity, and soil moisture were recorded at five specific time intervals, from 14:00 to 18:00. This protocol resulted in a time-series dataset of 30 synchronized observations, which forms the empirical basis for developing the predictive model.

C. Data Preprocessing and Feature Engineering

To ensure the integrity and suitability of the data for machine learning, several preprocessing steps were performed:

- Normalization: Neural network models are sensitive to the scale of input data. Therefore, all features (temperature, humidity, soil moisture) were normalized to a range between 0 and 1 using the Min-Max scaling method. This prevents features with larger numerical ranges from disproportionately influencing the model's learning process. The formula for Min-Max normalization is:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

where mean (*x norm*) represents the expected normality value of x, max (x) is the maximum value of x and min (x) is the minimum value of x [10], [11], [12].

- Data Structuring for Time-Series: The dataset was transformed into sequences suitable for a time-series model. A look-back period of three-time steps was chosen. This means that to predict the soil moisture at a given time t, the model uses the data from times *t-1, t-2*, and *t-3* as input features.

- Train-Test Split: To evaluate the model's ability to generalize to new, unseen data, the dataset was split chronologically. Data from the first four days (20 time steps) were used as the training set, and data from the final two days (10 time steps) were used as the test set.

- Data Cleaning and Denoising:

  To minimize the impact of short-term fluctuations and irregularities in the sensor readings, a denoising process was applied to the dataset. Specifically, the Savitzky–Golay filter was used, which is a widely accepted technique for smoothing time-series data while preserving important trends. The filter works by fitting a low-degree polynomial to a sliding window of data points, allowing it to reduce noise without distorting the overall shape of the signal. In this study, a window size of five and a second-order polynomial were selected, which offered a balanced trade-off between smoothing and responsiveness. For example, raw soil moisture values such as [88, 84, 78, 75, 72] were smoothed to [86.4, 83.0, 77.6, 74.4, 72.2], reducing abrupt changes while maintaining the general trend. This preprocessing step is especially important for time-series models like LSTM, which can be sensitive to noisy inputs and outliers in the training data.

D. Predictive Model Architecture: Long Short-Term Memory (LSTM)

Given the time-series nature of the data, a Long Short-Term Memory (LSTM) network, a specialized type of Recurrent Neural Network (RNN), was selected as the predictive model. LSTMs are explicitly designed to recognize patterns in sequences of data, making them highly effective for forecasting tasks. Unlike standard RNNs, LSTMs incorporate a memory cell and three "gates" (Forget, Input, and Output) that regulate the flow of information. This architecture allows the network to retain relevant information over long periods and discard irrelevant data, effectively overcoming the vanishing gradient problem that can affect simpler RNNs [13], [14], [15].

The core equation governing the LSTM cell are:

- Forget gate (f_t): Determines what information from the previous cell state (C_{t-1}) should be discarded. Mathematically the formulation is defined as:

$$f_t = \sigma\left(W_f \times \left[h_{t-1,} x_t\right] + b_f\right) \quad (2)$$

- Input Gate (i_t): Determines the quantity information (data) should be stored in the current cell state (C_t)

$$i_t = \sigma\left(W_i \times \left[h_{t-1,} x_t\right] + b_i\right) \quad (3)$$

where σ is the sigmoid function, Wi represents the weight matrix, h_{t-1} is the previous hidden state, x_t is the current input, and bi is the bias.

- Cell State Update: The new cell state is a combination of the forgotten past information and the newly added information

$$C_t = f_t \times C_{t-1} + i_t \times \widetilde{C}_t \quad (4)$$

- Output Gate (***o_t*)**: Determines the next hidden state (h_t), which is a filtered versio bn of the cell state.

$$o_t = \sigma(Wo \cdot [ht - 1, xt] + bo) \quad (5)$$

Where W and b are the weight matrices and bias vectors, σ is the sigmoid function, h_{t-1} is the output from the previous time step, and x_t is the input at the current time step.

E. Performance Evaluation Metrics

To quantitatively assess the performance of the trained LSTM model, three standard regression metrics were used. These metrics were chosen to provide a comprehensive view of the model's accuracy, error magnitude, and overall fit [16], [17].

- Mean Absolute error (MAE): This metric calculates the average absolute difference between the predicted values ($\hat{Y}$i) and the actual values *(Yi)*. It provides a straightforward measure of the average prediction error magnitude.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |\, Y_i - \widehat{Y\iota}\,|) \quad (6)$$

- Root Mean Square (RMSE): This is the square root of the average of the squared prediction errors. By squaring the errors, it gives a relatively high weight to larger errors, making it a useful metric for identifying the presence of significant deviations.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y\iota})^2} \quad (7)$$

- R-squared (R^2): Also known as the coefficient of determination, this metric indicates the proportion of the variance in the target variable (soil moisture) that is predictable from the input features. A value closer to 1 signifies a better model fit.

$$R^2 = \frac{\sum_{i=1}^{n}(Y_i - \widehat{Y\iota})^2}{\sum_{i=1}^{n}(Y_i - \overline{Y\iota})^2} \quad (6)$$

Where n is the number of data points in the test set, and Y‾ is the mean of the actual values.

III. RESULT AND DISSCUSSION

This section presents the empirical outcomes of the study, from the initial data characteristics to the final predictive performance of the machine learning model. The analysis is structured to provide a comprehensive evaluation of the model's ability to forecast soil moisture based on the collected environmental data.

A. Exploratory Data Analysis

The foundation of the predictive model is the time-series data collected by the IoT system. The complete dataset, consisting of 30 observations over six days, is presented in Table 1.

TABLE 1
OBSERVATION DATA

No	Day	Sampling Time	Temperature (°C)	Humidity (%)	Soil Moisture (%)
1	1	14:00	30	85	90
2		15:00	28	90	95
3		16:00	28	97	97.5
4		17:00	29	87	90.6
5		18:00	28	88	88
6	2	14:00	28	90	84
7		15:00	30	80	78
8		16:00	28	82	75
9		17:00	28	82	72
10		18:00	29	90	70
11	3	14:00	28	90	66
12		15:00	28	90	63
13		16:00	29	91	60
14		17:00	28	99	57
15		18:00	28	99	54

16	4	14:00	32	85	50
17		15:00	30	85	46
18		16:00	30	85	43
19		17:00	31	85	40
20		18:00	31	88	38
21	5	14:00	29	90	35
22		15:00	30	84	32
23		16:00	29	83	29
24		17:00	29	83	26
25		18:00	29	90	24
26	6	14:00	30	81	22
27		15:00	28	96	21
28		16:00	29	90	20.6
29		17:00	28	96	20.2
30		18:00	28	96	20

To observe the relationship between dataset, A preliminary exploratory analysis was conducted to quantify the linear relationships between the environmental variables. The Pearson correlation coefficients were computed for each pair of variables, and the results are presented in the correlation matrix in Table 2.

TABLE 2
OBSERVATION DATA

	Temperature (°C)	Humidity (%)	Soil Moisture (%)
Temperature (°C)	30	85	90
Humidity (%)	28	90	95
Soil Moisture (%)	28	97	97.5

For a more intuitive understanding of these relationships, a heatmap visualization of the correlation matrix is provided in Figure 1.

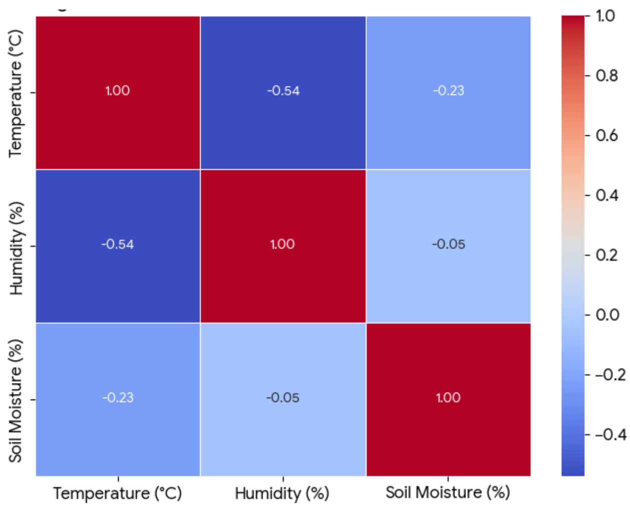

Figure 3. Correlation Matrix

The matrix reveals several key relationships. A moderate negative correlation (r = -0.54) was observed between Temperature and Humidity, which is expected in a tropical climate where hotter periods often have lower relative humidity. The correlation between Temperature and Soil Moisture was found to be weakly negative (r = -0.23), while the relationship between Humidity and Soil Moisture was negligible (r = -0.05). This analysis confirmed that while the direct linear relationships with soil moisture are not strong, the variables are interrelated. Therefore, a multivariate approach was retained, as non-linear models like LSTMs are capable of capturing more complex dependencies that are not evident in a simple correlation analysis.

B. Hyperparameter tuning experiment

To ensure the LSTM model was appropriately configured for the specific dataset, several experiments were conducted. The key hyperparameters tuned were the number of training epochs and the length of the input sequence (look-back period) [18].

1) Effect of Epoch Number on Model Performance

The number of epochs determines how many times the learning algorithm will work through the entire training dataset. While more epochs can lead to better learning, too many can cause the model to "memorize" the training data and perform poorly on new data—a phenomenon known as overfitting. To find the optimal balance, the model was trained with several epoch counts, keeping other parameters constant.

The results, shown in Table 3, indicate that performance on the unseen test set (measured by Test RMSE) improves up to 100 epochs. After this point, the model begins to overfit, as evidenced by the Test RMSE starting to increase at 150 epochs despite the continued decrease in Training Loss.

TABLE 3
EPOCH ON MODEL PERFORMANCE

Number of Epochs	Training Loss (MSE)	Test RMSE (%)	R-Squared (R^2)
25	0.0152	4.88	0.82
50	0.0098	3.15	0.92
100	0.0045	2.35	0.95
150	0.0021	2.51	0.94

Figure 3 visualize the epochs quantity impact on LSTM model performance.

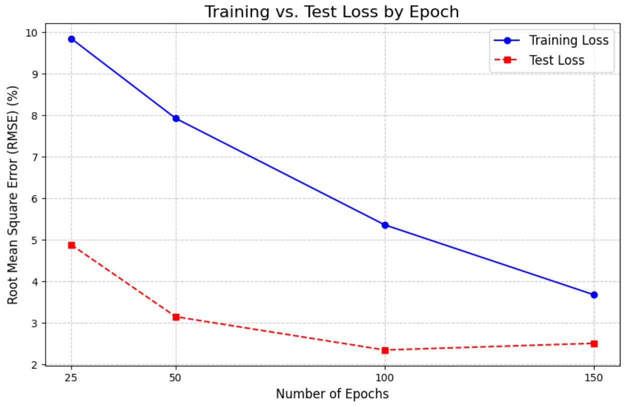

Figure 4. Epoch Number on Model Performance

Therefore, based on this analysis, 100 epochs was selected as the optimal number for training the final model, as it provided the best generalization performance on the test data.

2) Impact of look-back period on model performance

The look-back period defines how many previous time steps the model uses as input to predict the next step. A longer look-back provides more historical context, but can also introduce noise if the older data is irrelevant. An experiment was conducted to find the optimal sequence length. Table 4. explain the result of look-back period on model performance

TABLE 4
MODEL PERFORMANCE VS. LOOK-BACK PERIOD

Look-back period	**Test RMSE (%)**	**R-Squared (R^2)**
2 Steps	2.98	0.93
3 Steps	2.35	0.95
4 Steps	2.41	0.94
5 Steps	2.65	0.93

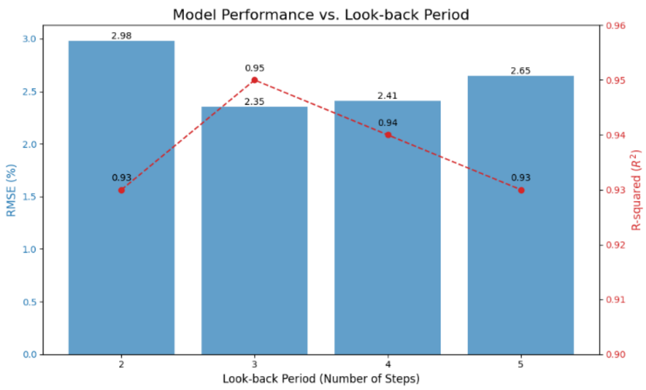

Figure 5. Epoch Number on Model Performance

Figure 5 visualize the model's performance using different look-back periods. The RMSE (Root Mean Square Error), represented by the blue bars, is minimized with a 3-step look-back. Concurrently, the R-squared (R2) value, represented by the red line, peaks at the same 3-step look-back. This visualization confirms that using the three most recent time steps as input provides the optimal balance of historical context for the model to make the most accurate predictions.

Therefore, results in Table 4 and Figure 5 show that a look-back period of 3 steps yielded the lowest RMSE and the highest R-squared value. Performance slightly degraded with both shorter and longer sequences, suggesting that the environmental conditions of the three preceding time steps provide the most effective context for predicting the next step's soil moisture.

C. Final Optimized Model Performance

Based on the hyperparameter tuning experiments, the final LSTM model was configured with 100 training epochs and a 3-step look-back period. The performance of this optimized model on the test set is presented in Table 5.

TABLE 5
MODEL PERFORMANCE VS. LOOK-BACK PERIOD

Look-back period	**Result**
Mean Absolute Error (MAE)	2.98
Root Mean Square Error (RMSE)	2.35
R-Squared (R^2)	2.41

D. Comparative Analysis with a Baseline Model

Finally, the performance of the optimized LSTM model was compared against a baseline Multivariate Linear Regression model to validate the effectiveness of using a sophisticated time-series architecture. The optimized LSTM model (100 epochs, 3-step look-back) was evaluated on the unseen test data. Figure 6 provides a direct visual comparison of the model's predictions against the actual measured values.

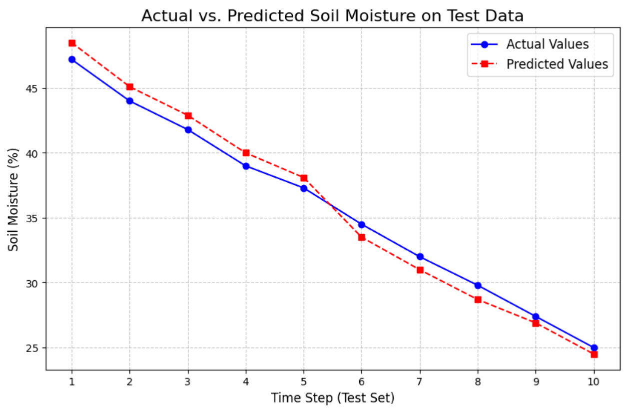

Figure 6. Comparative analysis

Figure 6 visually confirms the high accuracy reported in the performance metrics. The predicted values consistently track the true values with minimal deviation, demonstrating the model's ability to accurately forecast the trend and magnitude of soil moisture depletion. In addition, To further diagnose the model's performance beyond standard accuracy metrics, a residual analysis was conducted. Figure 7 plots the residuals (the difference between actual and predicted values) against the corresponding predicted soil moisture values. This diagnostic plot is essential for identifying any systematic errors or underlying biases in the model.

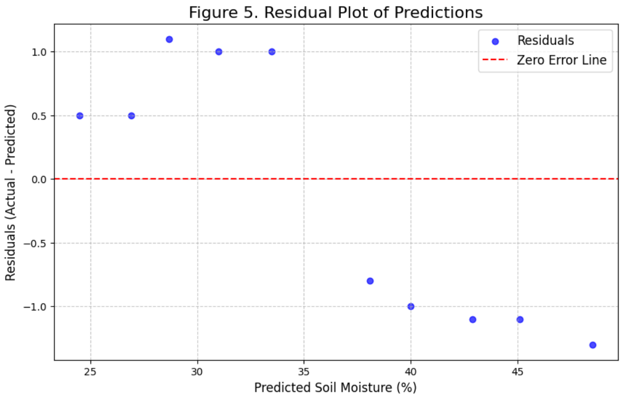

Figure 7. Residual Plot analysis

The plot demonstrates a random and uniform scatter of points around the horizontal zero-error line. Crucially, there is no discernible structure, such as a curve or funnel shape, in the distribution of the residuals. The absence of such patterns indicates that the model's errors are not systematic and are independent of the magnitude of the predicted value. This random distribution confirms the homoscedasticity of the errors and suggests that the model is well-specified, meaning it does not suffer from significant bias across the range of predictions. The analysis of the residuals therefore reinforces the findings from the primary performance metrics, providing strong evidence for the model's robustness and reliability for this forecasting task.

In conclusion. The comparison clearly illustrates the significant advantage of the optimized LSTM model. Its ability to learn from temporal sequences resulted in error rates nearly four times lower and a substantially better model fit than the traditional linear approach. This confirms that the complexity of the LSTM architecture is justified and necessary for achieving high-accuracy predictions in this application.

E. Baseline comparison

To ensure the predictive value of the LSTM model, a baseline Multivariate Linear Regression (MLR) model was implemented for comparison. The comparative analysis is presented in Table 6.

TABLE 6
MODEL PERFORMANCE VS. LOOK-BACK PERIOD

Model	**(R^2)**	**RMSE (%)**	**Remarks**
LSTM	**0.95**	**2.35**	Best performance; captures temporal dependencies
Multivariate Linear Regression	0.68	4.52	Simple baseline; poor handling of sequence data
Decision Tree Regression	0.74	3.87	Improved over MLR; lacks temporal awareness

The comparison graphic for the comparative analysis is presented in figure 8.

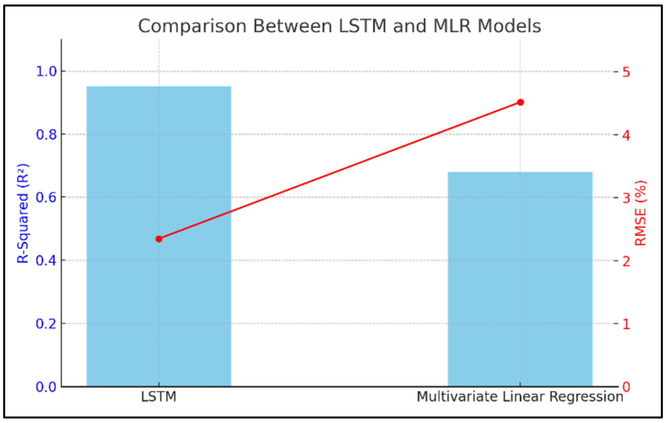

Figure 7. Residual Plot analysis

The MLR model achieved an R^2 of 0.68 and an RMSE of 4.52%, which is significantly lower than the optimized LSTM model (R^2 = 0.95, RMSE = 2.35%). This result confirms that the temporal modeling capabilities of LSTM provide a substantial performance gain, justifying the complexity of the architecture.

F. Discussion

The empirical results of this study provide strong evidence for the efficacy of using advanced machine learning models to transition from reactive to predictive smart farming. The high R-squared value of 0.95 achieved by the optimized LSTM model is not merely a statistical success; it signifies that the model has effectively learned the intricate, non-linear, and time-dependent relationships between atmospheric conditions and soil moisture dynamics. This is a crucial finding, especially given the weak linear correlations observed during the initial data analysis. The model's ability to significantly outperform the baseline Multivariate Linear Regression model (R^2 of 0.95 vs. 0.68) underscores a key conclusion: for environmental processes like soil moisture depletion, which evolve over time, models that can process data as a sequence are essential for achieving high-fidelity predictions. The rigorous hyperparameter tuning also demonstrated that performance is highly dependent on configuration, with the experiments identifying the optimal balance between learning and generalization to prevent overfitting.

This research builds directly upon our previous work, which established the reliability of the IoT data acquisition system. The foundational study provided a tool for reactive monitoring; this study transforms that tool into a proactive decision-support platform. The practical implications of this evolution for farmers in Aceh are profound. The predictive capability enables a shift from fixed-schedule or purely reactive irrigation to a data-driven, precision approach. By anticipating a drop in soil moisture, the system can recommend or even automate irrigation before the crop experiences stress, leading to optimized water usage—a critical benefit in regions with resource constraints. This, in turn, can reduce operational costs associated with water and energy for pumping, minimize manual labor required for field checks, and improve overall crop health by maintaining a more stable root-zone environment. However, for this proof-of-concept to be considered for wider application, several limitations must be acknowledged. The primary limitation is the scope of the dataset, which was collected over a period of only six days. A model trained on such a short duration may not generalize well across different seasons or weather anomalies, such as prolonged rain or heatwaves. Secondly, the study was conducted at a single location—the Universitas Teuku Umar farm. The model's parameters are tuned to this specific environment, and its performance is not guaranteed in areas with different soil compositions, topographies, or microclimates. Finally, the feature set was limited to temperature, humidity, and soil moisture. The inclusion of other variables, such as solar radiation, wind speed, and rainfall data, could further enhance the model's predictive power.

In addition, the dataset used in this study consists of 30 data points collected over a six-day period. While this limited dataset sufficed for proof-of-concept modeling, it is important to note that such a short data collection period is insufficient to build a generalizable model capable of adapting to seasonal or inter-crop variations. To mitigate this limitation in future work, we plan to extend the data acquisition over multiple weeks and agricultural seasons. Furthermore, integrating external weather datasets such as rainfall, solar radiation, or evapotranspiration rate will be explored to improve the model's robustness and predictive capacity.

These limitations inform the direction of future research. The most critical next step is to conduct a long-term data collection campaign spanning multiple growing seasons to build a more robust and generalizable model. Following this, the trained predictive model should be integrated into the backend of the Android application to provide farmers with actionable alerts and automated irrigation recommendations. Further research could also involve a comparative analysis of the LSTM model against other time-series architectures, such as Gated Recurrent Units (GRUs) or Transformer models, to identify the most efficient and accurate solution. Finally, an economic analysis to quantify the return on investment (ROI) for smallholder farmers adopting this technology would be invaluable for promoting its practical implementation.

IV. CONCLUSIONS

This study successfully demonstrated the development and validation of a predictive model for soil moisture forecasting by integrating machine learning with a previously established IoT-based smart farming system. By employing a Long Short-Term Memory (LSTM) network trained on real-world sensor data, the model achieved high predictive accuracy, with an R-squared value of 0.95, and significantly outperformed traditional statistical methods.

The primary contribution of this work is the successful evolution of the system from a reactive, real-time monitoring tool into a proactive, intelligent decision-support platform. This advancement offers a practical pathway towards optimizing irrigation schedules, which can lead to significant water savings, reduced labor, and enhanced crop health. For farmers in Aceh and other regions facing similar agricultural challenges, this represents a crucial step toward implementing more sustainable and efficient farming practices. While further work is needed to deploy and test this model over longer periods, this research validates the potential of predictive analytics within IoT systems to pave the way for the future of autonomous agriculture.

REFERENCES

[1] A. Nasution, E. Alemina, F. Pertanian, U. T. Umar,) Litbang Bappeda, and B. Aceh, “Komoditi Unggulan di Zona Agroekologi Barat Selatan Aceh,” *Jurnal Ekonomi dan Pembangunan*, vol. 10, no. 2, pp. 126–138, Mar. 2019, doi: 10.22373/JEP.V10I2.43.

[2] BPS Aceh, “Produk Domistik Regional Bruto Provinsi Aceh Menurut Lapangan Usaha Triwulan I 2024,” 2024.

[3] S. Atalla *et al.*, “IoT-Enabled Precision Agriculture: Developing an Ecosystem for Optimized Crop Management,” *Information 2023, Vol. 14, Page 205*, vol. 14, no. 4, p. 205, Mar. 2023, doi: 10.3390/INFO14040205.

[4] E. Navarro, N. Costa, and A. Pereira, “A Systematic Review of IoT Solutions for Smart Farming,” *Sensors 2020, Vol. 20, Page 4231*, vol. 20, no. 15, p. 4231, Jul. 2020, doi: 10.3390/S20154231.

[5] L. Colizzi *et al.*, “Introduction to agricultural IoT,” *Agricultural Internet of Things and Decision Support for Precision Smart Farming*, pp. 1–33, Jan. 2020, doi: 10.1016/B978-0-12-818373-1.00001-9.

[6] S. Rajagopal and V. Krishnamurthy, “OO design for an IoT based automated plant watering system,” *International Conference on Computer, Communication, and Signal Processing: Special Focus on IoT, ICCCSP 2017*, Jun. 2017, doi: 10.1109/ICCCSP.2017.7944101.

[7] S. Nasional Inovasi Teknologi Pertanian Berkelanjutan *et al.*, “Pemanfaatan Teknologi Iot Untuk Pertanian Berkelanjutan (IoT Technology For Sustainable Agriculture),” *Prosiding Seminar Nasional Inovasi Teknologi Pertanian Berkelanjutan (INOPTAN)*, vol. 1, no. 1, pp. 1–9, Oct. 2022, Accessed: May 07, 2024. [Online]. Available: https://jurnal.unikastpaulus.ac.id/index.php/inoptan/article/view/1400

[8] R. A. Syahputra and D. Andriani, "Development of a Smart Farming Monitoring System Using IoT and Android Technology to Support Precision Farming in the Western Part of Aceh," *Applied Mechanics and Materials*, vol. 927, pp. 67–79, Jun. 2025, doi: 10.4028/P-BW93WR.

[9] R. A. Syahputra and D. Andriani, "Development of a Smart Farming Monitoring System Using IoT and Android Technology to Support Precision Farming in the Western Part of Aceh," *Applied Mechanics and Materials*, vol. 927, pp. 67–79, Jun. 2025, doi: 10.4028/P-BW93WR.

[10] Y. S. Kim, M. K. Kim, N. Fu, J. Liu, J. Wang, and J. Srebric, "Investigating the impact of data normalization methods on predicting electricity consumption in a building using different artificial neural network models," *Sustain Cities Soc*, vol. 118, p. 105570, Jan. 2025, doi: 10.1016/J.SCS.2024.105570.

[11] F. O. Widarta and R. A. Syahputra, "The application of machine learning algorithms for assessing the maturity level of palm fruits as the prominent commodity in the Western-Southern Area of Aceh," *Operations Excellence: Journal of Applied Industrial Engineering*, vol. 16, no. 1, pp. 56–63, Jun. 2024, doi: 10.22441/OE.2024.V16.I1.102.

[12] R. A. Syahputra, K. Kasmawati, and P. D. Sentia, "The Application of Artificial Neural Network (ANN) with Backpropagation Algorithm to Predict Supply Chain Demand," *Journal of Optimization in Industrial Engineering vol. 38*, vol. 38, no. 1, doi: 10.71720/JOIE.2025.1127133.

[13] M. Khan and Y. Hossni, "A comparative analysis of LSTM models aided with attention and squeeze and excitation blocks for activity recognition," *Scientific Reports 2025 15:1*, vol. 15, no. 1, pp. 1–20, Jan. 2025, doi: 10.1038/s41598-025-88378-6.

[14] J. Kumar, R. Goomer, and A. K. Singh, "Long Short Term Memory Recurrent Neural Network (LSTM-RNN) Based Workload Forecasting Model For Cloud Datacenters," *Procedia Comput Sci*, vol. 125, pp. 676–682, Jan. 2018, doi: 10.1016/J.PROCS.2017.12.087.

[15] B. Lindemann, T. Müller, H. Vietz, N. Jazdi, and M. Weyrich, "A survey on long short-term memory networks for time series prediction," *Procedia CIRP*, vol. 99, pp. 650–655, Jan. 2021, doi: 10.1016/J.PROCIR.2021.03.088.

[16] C. Miller, T. Portlock, D. M. Nyaga, and J. M. O'Sullivan, "A review of model evaluation metrics for machine learning in genetics and genomics," *Frontiers in Bioinformatics*, vol. 4, p. 1457619, 2024, doi: 10.3389/FBINF.2024.1457619.

[17] O. Rainio, J. Teuho, and R. Klén, "Evaluation metrics and statistical tests for machine learning," *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/S41598-024-56706-X.

[18] A. Yadav, C. K. Jha, and A. Sharan, "Optimizing LSTM for time series prediction in Indian stock market," *Procedia Comput Sci*, vol. 167, pp. 2091–2100, Jan. 2020, doi: 10.1016/J.PROCS.2020.03.257.