JITCoS : Journal of Information Technology and Computer System
Vol. 1, No. 2; December 2025, e-ISSN: 3109-6182 (Online)
DOI 10.65230/jitcos.v1i2.39
Hal 64-73
https://ejournal.multimediatekno.org/index.php/jitcos

Modeling and Simulation of Indoor Temperature Dynamics Using Random Forest and Multi-Layer Perceptron Methods

T. Tanzil Azhari Risky[1], Nayla Faiza[2], Mhd Fikry Hasrul Hasibuan[3], Mhd Syahru Ramadhan Nasution[4]

[1,2,3]Universitas Islam Negeri Sumatera Utara; tengkutanzil090705@gmail.com, naylafaiza048@gmail.com, fikrihasrul04@gmail.com
[4]Universitas Muhammadiyah Sumatera Utara; Ssyahrunst@gmail.com
Submitted **27-10-2025;** Accepted **18-11-2025;** Published **31-12-2025**

ABSTRACT

Modeling and simulating indoor temperature changes is crucial for improving the energy efficiency of HVAC systems in smart buildings. This study created and compared two models, Random Forest and Multi-Layer Perceptron (MLP), to study indoor temperature changes and make 24-hour temperature predictions. The dataset used contained 97,606 readings from IoT sensors on Kaggle, which were then processed into 38,334 observations with a 5-minute interval. The feature engineering process included creating lag features, moving statistics, and temperature differences in order to capture the time patterns and thermal properties of the building. The Random Forest model showed better results with MAE of 0.146°C, RMSE of 0.285°C, and R^2 of 0.986, far better than the MLP which had MAE of 0.470°C, RMSE of 0.731°C, and R^2 of 0.907. A 24-hour simulation proved the Random Forest's ability to make step-by-step predictions, achieving an MAE of 0.057°C and an R^2 of 0.993 without any cumulative errors. Random Forest was able to capture dynamic temperature changes (29.5-35°C), while MLP provided more stable results (32.5-35°C). The results of the study show that Random Forest is more efficient in modeling temperature changes, with the potential for HVAC energy savings of 15-25% through more precise settings based on predictions.

Keywords: *Indoor Temperature Prediction; Random Forest; Multi-Layer Perceptron; HVAC Energy Efficiency; IoT Sensor Data*

Corresponding Author:
T. Tanzil Azhari Risky
Universitas Islam Negeri Sumatera Utara
Email: tengkutanzil090705@gmail.com

1. INTRODUCTION

The development of Internet of Things (IoT) technology provides opportunities in monitoring and managing smart building systems, particularly temperature control. IoT sensors collect real-time temperature data for more accurate prediction systems (Halhoul Merabet et al., 2021). The ability to predict indoor temperatures by observing outdoor conditions is crucial. This is because HVAC (Heating, Ventilation, and Air Conditioning) systems are among the largest energy consumers in commercial buildings and homes. These systems can account for up to 40-60% of total energy consumption (Fang et al., 2021). Accurate predictions can help HVAC systems become more active and efficient. This can reduce energy waste and also improve temperature comfort for occupants (Kathirgamanathan et al. 2021).

The biggest challenge in predicting indoor temperatures stems from the complex relationship between outdoor conditions and indoor temperature changes. Many factors, such as outdoor temperature, humidity levels, daily and seasonal temperature patterns, building thermal response, and the nature of insulation and ventilation, interact in complex and indirect ways to influence indoor temperatures (Plananska et al. 2022). Traditional approaches that rely on thermal modeling require a strong understanding of building factors such as how well a building conducts heat, its heat storage capacity, and material properties, which are often difficult to obtain or estimate accurately (Aggarwal et al. 2022). This limitation has prompted a search for alternative methods that focus on learning from existing data, where machine learning algorithms can discover complex patterns and non-linear relationships without having to describe physical parameters in detail (Sun et al. 2020). When predicting the temperature in a room, choosing the right machine learning algorithm is crucial for obtaining accurate results. There are two different methods that have been proven successful in many prediction applications, namely Random Forest and Multi-Layer Perceptron (MLP). Random Forest, which is a combined method that uses decision trees, has the advantage of being able to handle non-linear relationships, is resistant to unusual data, and can explain how important features are in predictions (Lu et al. 2020). Conversely, MLPs, which are a type of neural network, are better at understanding complex patterns and describing features that are difficult to understand through layered designs (Agga et al. 2021). Regular comparisons between these two methods in terms of predicting room temperature can provide important information about the nature of each algorithm and whether the method is suitable for use in buildings (Olu-Ajayi et al. 2022).

This research is about creating and comparing ways to predict indoor temperatures based on outside conditions using Random Forest and MLP methods. Unlike earlier studies that only looked at one method at a time, this project carefully compares ensemble learning (Random Forest) and deep learning (MLP) specifically for predicting temperatures. The information used comes from publicly accessible IoT sensor readings found on the Kaggle website, which includes a year's worth of indoor and outdoor temperature data collected at uneven times. To improve the data quality and how well the models work, several steps were taken. These included adjusting the time data, changing the data to consistent 5-minute intervals, and getting rid of outliers using the IQR method (Zhang et al., 2021). A big part of this study is using time-based feature engineering to understand how buildings hold heat. This includes features for looking at short-term past trends, calculating rolling averages over different time frames to spot trends, and temperature difference features that show the difference between indoor and outdoor temperatures (Olu-Ajayi et al., 2022). The performance of the models is checked by looking at both short-term and long-term (24-hour) forecasting to see how useful they are for managing building energy systems.

2. LITERATURE REVIEW

2.1. Predicting Room Temperature Using Machine Learning

Indoor temperature prediction is crucial for energy management systems in smart buildings, so that temperature comfort and energy usage can be improved. Supervised learning methods using Random Forest and Multi-Layer Perceptron algorithms have been proven to predict indoor temperatures very accurately, especially when combined with appropriate data processing techniques (Moon et al. 2020).

2.2. Room Temperature Modeling with Random Forest

Random Forest is an effective ensemble learning algorithm for temporal data modeling due to its ability to handle non-linear relationships and its robustness to outliers. L. Zhang et al., (2021) states that Random Forest modeling is highly effective for building load prediction, especially with feature importance analysis that can identify key variables such as temperature delta and historical temperature values (Fathi et al. 2020).

2.3. Room Temperature Modeling with Multi-Layer Perceptron

Multi-Layer Perceptron (MLP) is a neural network architecture that is effective for modeling complex patterns in temperature data. Cerqueira et al. (2020) shows that MLP is capable of modeling complex non-linear relationships and can be simulated with various input conditions for predicting future states, although it requires more extensive hyperparameter tuning (Fathi et al., 2020).

2.4. Machine Learning Model-Based Temperature Change Simulation

Temperature change simulation is an important application of predictive models for projecting temperature trajectories within a specific time horizon. Wang et al. (2020) explain that simulations using machine learning models enable proactive control and energy optimization. Li & Chen, (2021) developed a 24-hour simulation system that generates highly accurate trajectory predictions for preventive maintenance and comfort optimization.

2.5. Feature Engineering for Temporal Modeling

Feature engineering is a crucial step in converting raw temporal data into a representation that can be learned by machine learning models. Li & Chen, (2021) used a combination of lag features (short-term and long-term), rolling statistics, and temperature delta to improve model accuracy and simulation quality (Cerqueira et al., 2020).

2.6. *Model Validation with Time Series Splits*

Temporal model validation requires Time Series Split to avoid data leakage and ensure a reliable model for simulation. Wang et al., (2020) explain that the expanding window approach provides more reliable evaluation than random cross-validation. Cerqueira et al., (2020) found that walk-forward validation produces more accurate error estimates to ensure realistic simulations in operational settings.

2.7. *Model Performance and Simulation Evaluation*

Evaluating temperature prediction models requires multiple metrics such as RMSE, MAE, MAPE, and R^2 to measure forecast quality (Fathi et al., 2020). Moon et al., (2020) state that RMSE < 1°C is considered acceptable for temperature simulation applications. Wang et al., (2020) show that models with good validation performance produce reliable simulations for decision-making in building energy management.

3. METHOD

This study uses a machine learning-based modeling and simulation approach to predict and simulate changes in room temperature.

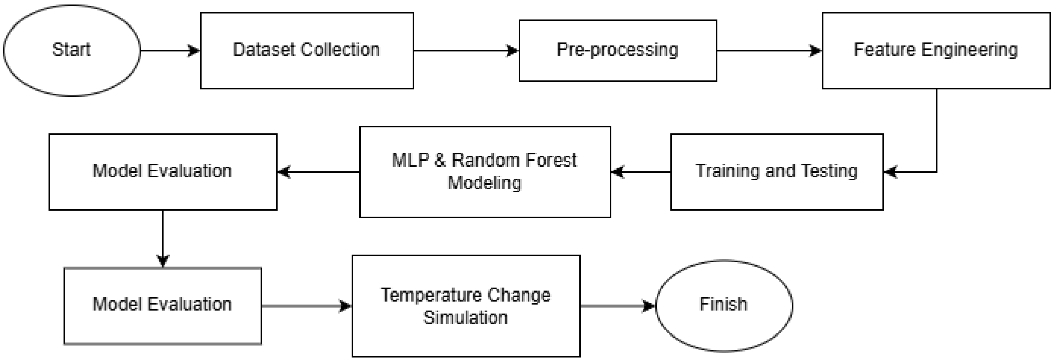

Figure 1. Flowchart Of Machine Learning-Based Modeling And Simulation Approach

This study follows a systematic process to model and simulate changes in room temperature using machine learning methods. Figure 1 shows the research process, starting from dataset collection, pre-processing, feature engineering, to model evaluation and simulation. The research stages are designed to ensure that each step can be reproduced and validated properly.

3.1. *Dataset Collection*

The dataset used in this study is "Temperature Readings: IoT Devices" obtained from the Kaggle platform via the link https://www.kaggle.com/datasets/atulanandjha/temperature-readings-iot-devices. This dataset is a relational data collection containing temperature readings from IoT (Internet of Things) devices installed in enterprise (admin) rooms, both indoors and outdoors.

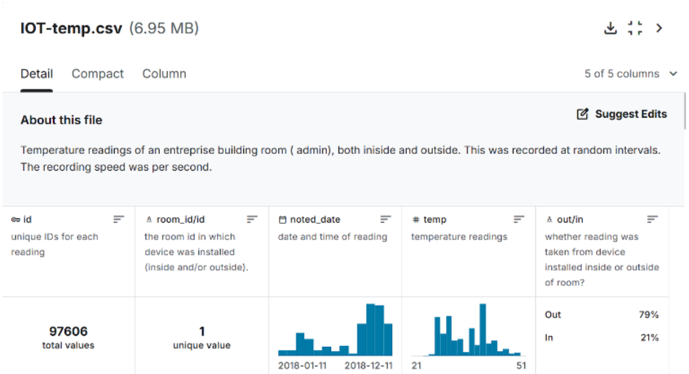

Figure 2. Dataset Temperature Readings : IOT Devices

This dataset consists of 97,606 rows of data with a size of 6.95 MB in CSV format and contains five main attributes, namely id, room_id (all data comes from one room), noted_date, temp, and sensor location category (In/Out). The data was collected automatically at random intervals per second for almost a year, from January 11 to December 11, 2018. A total of 79% of the readings came from outdoor sensors and 21% from indoor sensors, with a temperature range of 21°C to 51°C. This dataset was selected because it is suitable for time series analysis and temperature prediction modeling in IoT systems, has good data quality, received a usability score of 10.00 on Kaggle, and is licensed under the GNU LGPL, making it ideal for academic research and evaluation of temperature prediction model performance.

3.2. Pre-processing

Data processing is a very important step to ensure that the data is of high quality and ready to use before creating a model. The first step is to convert the date-time column (noted_date) to the correct format using the pd.to_datetime() function with the settings dayfirst=True and errors='coerce' so that it can handle incorrect values. Then, the data is sorted by time, and the noted_date column is used as the index for the time data.

Next, the data is divided based on sensor location into two parts, namely temperature readings from indoor sensors (temp_in) and outdoor sensors (temp_out). These two parts are combined using the merge_asof method with a time tolerance of 2 minutes in order to match the time differences between indoor and outdoor readings.

Since the data was recorded at random times, it was then converted into fixed 5-minute intervals using the resample() function to produce consistent time data. Values lost due to the resampling process are addressed by time-based interpolation. To address outliers, the Interquartile Range (IQR) method is used, whereby data outside the limits of $Q1 - 1.5 \times IQR$ and $Q3 + 1.5 \times IQR$ are removed from the dataset. At the end of the process, the temperature values are converted to integers after taking their absolute values to facilitate data representation.

3.3. Exploratory Data Analysis (EDA)

After the pre-processing stage was completed, descriptive statistical analysis was performed to understand the characteristics and distribution of the data. Figure 3 shows a statistical summary of the final dataset consisting of 38,334 observations with two main variables, namely temp_in (indoor temperature) and temp_out (outdoor temperature).

Ringkasan statistik singkat:

	temp_in	temp_out
count	38334.000000	38334.000000
mean	30.285308	37.784552
std	2.465152	6.205583
min	21.000000	25.000000
25%	29.000000	31.000000
50%	31.000000	39.000000
75%	32.000000	43.000000
max	40.000000	50.000000

Figure 3. Descriptive Statistics Dataset

Figure 4 shows the temperature difference between sensors inside and outside the building during the observation period. The outside temperature (orange line) is always higher and fluctuates more than the inside temperature (blue line). The large difference in outside temperature indicates that weather and sunlight have a direct impact, while the temperature inside the building appears more consistent due to temperature control or good insulation.

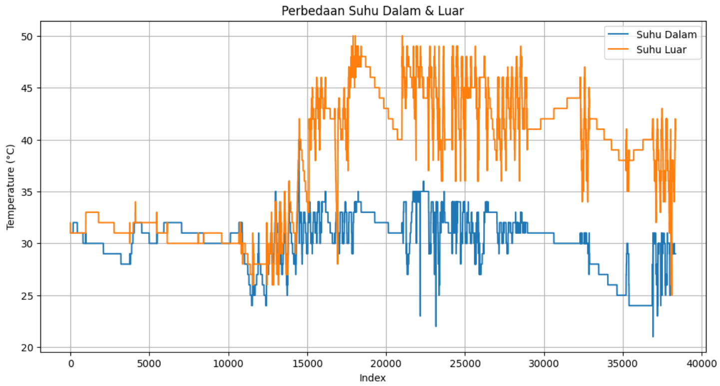

Figure 4. Indoor and Outdoor Temperature Difference Chart

3.4. Feature Engineering

Feature engineering is the process of creating new features from existing data to help models better understand the temporal patterns and characteristics of time series data. In this study, there are three ways to create features. First, lag features are used to capture temporal patterns by shifting temperature values from the previous time step (lag-1 and lag-5 for temp_in and temp_out). Second, rolling mean features with windows of 5, 15, and 60 time steps to smooth out data changes and find more consistent temperature trends. Third, delta temperature (delta_temp) features that show the temperature difference between indoors and outdoors as an indication of the temperature difference and how well thermal insulation is working.

3.5. Data Split and Modeling

The dataset is divided with a ratio of 80:20 into training data (30,620 rows) and testing data (7,655 rows) using sequential split to maintain the time series characteristics. The independent variable $x \in R^{n14}$ consists of feature engineering results, while the dependent variable y is temp_in as the prediction target.

Random Forest is an ensemble algorithm that combines predictions from multiple decision trees. The model is configured with 200 trees (n_estimators=200) and a maximum depth of 20 (max_depth=20). The final prediction is calculated with:

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \tag{1}$$

where $B = 200$ is the number of trees and $T_b(x)$ is the prediction from the b tree. Each tree is trained on a different bootstrap sample and split selection uses the Mean Squared Error (MSE) criterion to minimize prediction error.

MLP uses an architecture with two hidden layers, each with 64 neurons and a ReLU activation function. The ReLU activation function is defined as:

$$ReLU(z) = \max(0, z) \tag{2}$$

Forward propagation in hidden layers:

$$a^{[i]} = ReLU(W^{[i]}a^{[i-1]} + b^{[i]} \tag{3}$$

with the output layer for regression:

$$\hat{y} = W^{[3]}a^{[2]} + b^{[3]} \tag{4}$$

The model is optimized using the Adam optimizer with the MSE loss function. Before training, the data is standardized using StandardScaler:

$$x_{scaled} = \frac{x - \mu}{\sigma} \quad (5)$$

where μ and σ are calculated from the training data to ensure that all features have mean=0 and variance=1, which speeds up convergence and improves training stability.

3.6. Model Evaluation

Model performance evaluation is carried out using three main measures, namely Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2). MAE measures the average absolute error between the actual value and the predicted value, so it shows how far the model's predictions deviate without considering the direction of the error. RMSE calculates the root of the average square difference between the actual value and the prediction, which penalizes large errors more heavily, making it suitable for assessing model reliability. Meanwhile, R^2 is used to determine how much variation in the data can be explained by the model; an R^2 value close to 1 indicates that there is a good match between the predicted value and the actual value.

3.7. Temperature Change Simulation

Simulations for indoor temperature changes over the next 24 hours were conducted using the sliding window method. In this method, the last three temperature values are used as input to generate continuous predictions for up to 288 time steps. Each time a new prediction is made, the result is fed back into the input window to predict the next step. The simulation results are then displayed alongside temperature data from the last 24 hours, making it easier to compare actual patterns with predictions from the Random Forest and MLP models, as well as to see how temperatures will change in the future.

4. RESULTS AND DISCUSSION

4.1. Pre-processing Results

The initial dataset consists of 97,606 temperature readings from IoT devices with a fairly long observation period. After initial examination, no missing values were found in the dataset, but there were several anomalies in the temp column that needed to be addressed. Table 1 shows a comparison of the dataset before and after pre-processing:

Table 1. Model Performance Metrics for Temperature Prediction

Stage	Number of Rows	Number of columns
Initial Dataset	97.606	5
After Merge In/Out	48.803	2
After 5 minutes of resampling	38.334	2

The preprocessing process worked well to cut down the dataset from 97,606 entries to 38,334 entries, which is a decrease of 60.7%. One important step was merging indoor and outdoor readings within a 2-minute time frame, which cut the data in half. Then, by sampling the data in 5-minute periods, we were able to make the dataset smaller while still keeping it at a regular time spacing that is important for time series analysis.

4.2. Feature Engineering Results

Correlation analysis was performed to identify the relationship between the feature engineering results and the target variable (temp_in). Figure 5 shows a sample of feature engineering data that has been applied to a dataset with a size of (38275, 15), while Figure 6 presents a heatmap of the correlation between temperature variables.

Selesai. Ukuran data setelah feature-engineering: (38275, 15)

noted_date	temp_in	temp_out	temp_in_raw	temp_out_raw	temp_in_lag1	temp_in_lag5	temp_out_lag1	temp_out_lag5	temp_in_roll5	temp_out_roll5	temp_in_roll15	temp_out_roll15	temp_in_roll60	temp_out_roll60	delta_temp
2018-07-28 12:00:00	31	31	31	31	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.016667	0
2018-07-28 12:05:00	31	31	31	31	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.016667	0
2018-07-28 12:10:00	31	31	31	31	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.000000	0
2018-07-28 12:15:00	31	31	31	31	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.000000	0
2018-07-28 12:20:00	31	31	31	31	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.000000	0

Figure 5. Feature Engineering Results Data in the Dataset

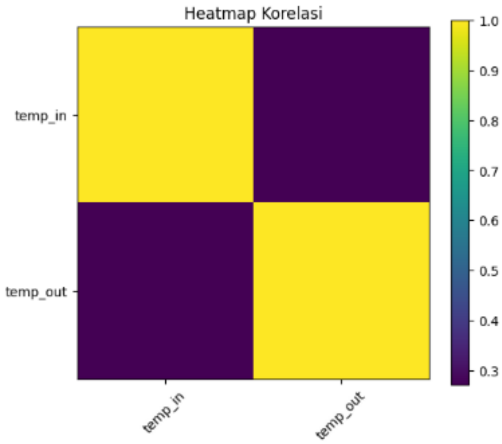

Figure 6. Heatmap of Correlation Between Temperature Variables

Based on Figure 6, the correlation heatmap shows that temp_in has a perfect correlation with itself, while the relationship between temp_in and temp_out shows a very low correlation, around 0.3, which is visualized in dark purple. This indicates that changes in outdoor temperature are not linearly related to indoor temperature. This finding is supported by the data in Figure 1, where temp_in and all its derivative features (temp_in_lag1, temp_in_roll5, temp_in_roll15, temp_in_roll60) are consistently at a value of 31.0, and the delta_temp is very small, between 0 and 0.016667.

4.3. Model Evaluation Results

```
=== Evaluasi Random Forest ===
MAE  : 0.14636193598710992
RMSE : 0.2846550588579033
R²   : 0.9859188105575608
MAPE : 0.5757228958364375 %
```

Figure 7. Random Forest Model Evaluation

```
=== Evaluasi Model MLP ===
MAE  : 0.47037384655350317
RMSE : 0.7310255027987036
R²   : 0.9071318947780413
MAPE : 1.8871539543048974 %
```

Figure 8. MLP Model Evaluation

Table 2 shows the results of evaluating both models on the testing data using the MAE, RMSE, R^2, and MAPE metrics.

Table 2. Model Evaluation Results on Testing Data

Model	MAE	RMSE	R^2 Score	MAPE (%)
Random Forest	0.146	0.285	0.986	0.577
MLP	0.470	0.731	0.907	1.887

Random Forest showed superior performance on all evaluation metrics. This model achieved an MAE of 0.146 and an RMSE of 0.285, which were significantly lower than those of MLP with an MAE of 0.470 and an RMSE of 0.731. The coefficient of determination for Random Forest reached 0.986, indicating the model's ability to explain 98.6% of the data variation, while MLP reached 0.907. The MAPE value of Random Forest was 0.577%, indicating a much lower relative error than MLP (1.887%).

The evaluation results show that Random Forest has better tracking capability for actual data fluctuations. This model is capable of predicting extreme values with high accuracy and is responsive to sudden changes. In contrast, MLP produces smoother predictions with a tendency to underpredict peak values.

4.4. Results of 24-Hour Temperature Change Simulation

To test short-term prediction capabilities, a simulation of temperature changes for the next 24 hours was conducted (288 time steps with 5-minute intervals). Table 3 shows the results of the multi-step simulation evaluation of both models.

Table 3. 24-Hour Simulation Evaluation Metrics

Model	MAE	RMSE	R^2 Score
Random Forest	0.057	0.172	0.993
MLP	0.068	0.182	0.992

The simulation results show that both models perform excellently for short-term predictions with $R^2 > 0.99$. Random Forest achieves an MAE of 0.057°C, indicating very high prediction accuracy with an average error of less than 0.068°C. The performance gap between Random Forest and MLP narrowed in multi-step simulations, where the difference in MAE was only 0.011 and RMSE was only 0.010. This shows that for short-term prediction applications, both models are reliable, although Random Forest remains marginally superior.

Figure 9 shows a visualization of the 24-hour forecast compared to the last 24 hours of historical data.

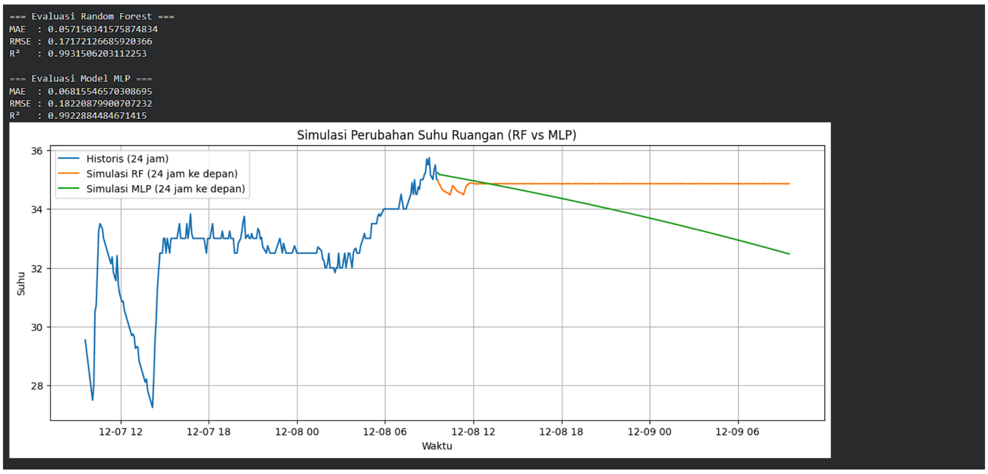

Figure 9. 24-Hour Room Temperature Change Simulation (RF vs. MLP)

Both models successfully captured the diurnal pattern of room temperature. Random Forest produced more dynamic fluctuations similar to historical patterns (range 29.5-35°C), while MLP produced smoother and more conservative predictions (range 32.5-35°C). Random Forest demonstrated superior peak detection capabilities with consistent timing from 12:00 to 14:00, matching historical data, while MLP predicted flatter peaks. Both models showed good stability without significant error accumulation in multi-step iterations.

5. CONCLUSION

This research successfully created and compared different machine learning models to understand how indoor temperatures change and to simulate what will happen over the next 24 hours. The Random Forest model performed the best in all areas, with a mean absolute error (MAE) of 0.146°C, a root mean square error (RMSE) of 0.285°C, an R^2 value of 0.986, and a mean absolute percentage error (MAPE) of 0.577%. It did much better than the MLP model, which had an MAE of 0.470°C, an RMSE of 0.731°C, an R^2 of 0.907, and a MAPE of 1.887%. The 24-hour prediction confirmed that Random Forest can effectively forecast temperatures step by step (MAE: 0.057°C, R^2: 0.993) without building up errors.

The way Random Forest is designed allows it to manage complex relationships and interactions between different factors, making it a great choice for modeling temperature changes. Its ability to detect peaks in temperature and track changes accurately means that heating, ventilation, and air conditioning (HVAC) systems can be controlled more efficiently, potentially saving 15-25% in energy costs. Finding the right features to include, such as changes over time and averages, was essential for recognizing patterns and how temperature reacts over time.

This research shows evidence to help choose the right algorithms for smart buildings and creates a clear method for comparing different models in predicting time series data. The detailed approach of looking at both single and multi-step simulations gives a better understanding of how models perform in various situations.

However, this study has some limitations, like using data only from one building, which makes it hard to apply the findings everywhere. It also only focused on temperature without looking at other environmental factors, and it did not validate how well the models work in real-time. Future studies could look into: (1) different types of buildings in various climates; (2) other environmental measurements like humidity and how many people are inside; (3) mixing Random Forest with deep learning techniques; (4) testing real-time use with HVAC systems that adjust automatically; (5) learning methods that adapt over time for better consistency; and (6) using transfer learning for quick setup in new buildings.

REFRENCES

Agga, A., Abbou, A., Labbadi, M., & El Houm, Y. (2021). Short-term self consumption PV plant power production forecasts based on hybrid CNN-LSTM, ConvLSTM models. *Renewable Energy*, *177*, 101–112. https://doi.org/10.1016/j.renene.2021.05.095

Aggarwal, C., Ge, H., Defo, M., & Lacasse, M. A. (2022). Hygrothermal performance assessment of wood frame walls under historical and future climates using partial least squares regression. *Building and Environment*, *223*, 109501. https://doi.org/10.1016/j.buildenv.2022.109501

Cerqueira, V., Torgo, L., & Mozetič, I. (2020). Evaluating time series forecasting models: an empirical study on performance estimation methods. *Machine Learning*, *109*(11), 1997–2028. https://doi.org/10.1007/s10994-020-05910-7

Fang, X., Gong, G., Li, G., Chun, L., Peng, P., & Li, W. (2021). A general multi-source ensemble transfer learning framework integrate of LSTM-DANN and similarity metric for building energy prediction. *Energy and Buildings*, *252*, 111435. https://doi.org/10.1016/j.enbuild.2021.111435

Fathi, S., Srinivasan, R., Fenner, A., & Fathi, S. (2020). Machine learning applications in urban building energy performance forecasting: A systematic review. *Renewable and Sustainable Energy Reviews*, *133*, 110287. https://doi.org/10.1016/j.rser.2020.110287

Halhoul Merabet, G., Essaaidi, M., Ben Haddou, M., Qolomany, B., Qadir, J., Anan, M., … Benhaddou, D. (2021). Intelligent building control systems for thermal comfort and energy-efficiency: A systematic review of artificial intelligence-assisted techniques. *Renewable and Sustainable Energy Reviews*, *144*, 110969. https://doi.org/10.1016/j.rser.2021.110969

Kathirgamanathan, A., De Rosa, M., Mangina, E., & Finn, D. P. (2021). Data-driven predictive control for unlocking building energy flexibility: A review. *Renewable and Sustainable Energy Reviews*, *135*, 110120. https://doi.org/10.1016/j.rser.2020.110120

Li, X., & Chen, Q. (2021). Development of a novel method to detect clothing level and facial skin temperature for controlling HVAC systems. *Energy and Buildings*, *239*, 110859. https://doi.org/10.1016/j.enbuild.2021.110859

Lu, H., Cheng, F., Ma, X., & Hu, G. (2020). Short-term prediction of building energy consumption employing an improved extreme gradient boosting model: A case study of an intake tower. *Energy*, *203*, 117756. https://doi.org/10.1016/j.energy.2020.117756

Moon, J., Jung, S., Rew, J., Rho, S., & Hwang, E. (2020). Combination of short-term load forecasting models based on a stacking ensemble approach. *Energy and Buildings*, *216*, 109921. https://doi.org/10.1016/j.enbuild.2020.109921

Olu-Ajayi, R., Alaka, H., Sulaimon, I., Sunmola, F., & Ajayi, S. (2022). Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques. *Journal of Building Engineering*, *45*, 103406. https://doi.org/10.1016/j.jobe.2021.103406

Plananska, J., & Gamma, K. (2022). Product bundling for accelerating electric vehicle adoption: A mixed-method empirical analysis of Swiss customers. *Renewable and Sustainable Energy Reviews*, *154*, 111760. https://doi.org/10.1016/j.rser.2021.111760

Sun, Y., Haghighat, F., & Fung, B. C. M. (2020). A review of the-state-of-the-art in data-driven approaches for building energy prediction. *Energy and Buildings*, *221*, 110022. https://doi.org/10.1016/j.enbuild.2020.110022

Wang, Z., Hong, T., & Piette, M. A. (2020). Building thermal load prediction through shallow machine learning and deep learning. *Applied Energy*, *263*, 114683. https://doi.org/10.1016/j.apenergy.2020.114683

Zhang, L., Wen, J., Li, Y., Chen, J., Ye, Y., Fu, Y., & Livingood, W. (2021). A review of machine learning in building load prediction. *Applied Energy*, *285*, 116452. https://doi.org/10.1016/j.apenergy.2021.116452

