

Electrical Load Prediction Model for the SG 06 Feeder of ULP Meureudu Using a Gated Recurrent Unit Architecture

**Analdi Muttaqin, Khairun Saddami, Ramdhan Halid Siregar,
Mohd Syaryadhi, and Mahdi Syukri***

Department of Electrical and Computer Engineering,

Universitas Syiah Kuala, Banda Aceh, Indonesia

Corresponding Author: Mahdi Syukri* | mahdisyukri@usk.ac.id

Abstract

The increasing electricity demand and the growing complexity of distribution networks require PLN to employ accurate load forecasting to ensure system reliability. This challenge is evident in the SG 06 Feeder of ULP Meureudu, which exhibits highly fluctuating load patterns, significant disparities between average and peak loads, and a relatively low load factor. These characteristics reduce the effectiveness of conventional forecasting methods and highlight the need for a model capable of accurately capturing temporal patterns. This study develops a short-term load forecasting model based on a Gated Recurrent Unit (GRU) architecture, utilizing time-based features as the sole input variables. The research process includes historical data collection, preprocessing, time-series windowing, model training, and performance evaluation using RMSE, MAE, and MAPE. The results show that the GRU model is able to generate stable and accurate predictions for both daily and weekly load patterns of the SG 06 Feeder, outperforming the simple statistical and time-based reference models used for comparison. The proposed model provides a reliable solution to support operational decision-making and enhance the reliability of the 20 kV distribution network managed by ULP Meureudu.

Keywords: GRU, Short-Term Load Forecasting, Jaringan distribusi 20 kV, Penyulang SG 06.

I. INTRODUCTION

The growing demand for electrical energy across various regions requires PLN to manage distribution systems more effectively and adaptively. The dynamics of electricity consumption—driven by community activity patterns, weather conditions, and customer growth cause energy needs to fluctuate over time. These conditions increasingly necessitate more accurate load forecasting capabilities to ensure that energy distribution within the network, particularly the 20 kV distribution system, operates optimally [1].

Inaccurate load forecasting can lead to various operational consequences, including supply imbalances, increased power losses, feeder overloading, and reduced service reliability [2]. Within the operational area of PLN ULP Meureudu, one of the feeders exhibiting highly dynamic load characteristics is the SG 06 Feeder. The load fluctuations on this feeder are strongly influenced by customers' daily consumption patterns, thereby necessitating a forecasting approach capable of capturing these temporal variations with high precision [3].

Various studies on electrical load forecasting has advanced significantly in line with the growing demand for more efficient and reliable energy management. Various studies have employed conventional methods based on fuzzy logic and statistical approaches. Syahputra [4] applied fuzzy logic for short-term load forecasting in Banda Aceh and demonstrated that while this method can capture fundamental load patterns, it is less adaptive to highly dynamic load variations. Statistical approaches such as the Seasonal

Autoregressive Integrated Moving Average (SARIMA) combined with an Artificial Neural Network (ANN) have also been examined by Silfiani [5]; however, their performance declines when confronted with nonlinear consumption patterns.

With the increasing complexity of power systems, research has increasingly shifted toward the use of machine learning techniques. Asshiddiq [6] and Syukri [7] demonstrated that machine learning algorithms and Artificial Neural Networks (ANN) can enhance prediction accuracy for distribution transformers as well as real-time. have also been widely applied for energy forecasting and customer load estimation, as shown in the works of Binoto and Kristiawan [8] and Wibowo et al. [9]. Other studies have utilized hybrid approaches, such as combining fuzzy logic with a Multilayer Perceptron (MLP) for forecasting loads during holiday periods [10]. Additionally, research by Susanti [11], Ardyansyah and Kartini [12], and Sugiarto et al. [13] further underscores the important role of ANN and Backpropagation Neural Networks (BPNN) in electrical load forecasting.

Further developments in load forecasting research are marked by the adoption of deep learning architectures. Ningrum et al. [14] employed a Long Short-Term Memory (LSTM) network for transformer lifetime prediction and found that the model was capable of capturing long-term temporal patterns. Improvements to MLP-based ANN models for load estimation have also been explored in several studies [15]– [17]. Septian and Kartini [18] specifically utilized LSTM for electrical load modeling and

demonstrated that it provides higher prediction accuracy compared to conventional Backpropagation Neural Networks (BPNN).

Many of the forecasting methods discussed above often exhibit limitations in capturing complex and nonlinear temporal patterns. With advances in artificial intelligence, modern neural network architectures such as the Gated Recurrent Unit (GRU) offer improved capabilities in learning time-series behaviors, including daily and seasonal variations in electrical load patterns [19]. Moreover, the incorporation of temporal features (such as hour, day, and other calendar-based categories) can further enrich the representation of consumption patterns and enhance forecasting accuracy [20].

Although previous studies have demonstrated the advantages of deep learning-based models, there remains a gap in research specifically examining the use of GRU architectures with time-based features as the primary input at the 20 kV distribution network level, particularly within PLN's operational areas. This gap is significant because RNN-based models, especially GRU, possess strong capabilities in capturing highly fluctuating temporal patterns on feeders with mixed customer characteristics, such as the SG 06 Feeder in ULP Meureudu. This research gap forms the basis for the present study, which aims to develop a more adaptive and accurate GRU-based short-term load forecasting model.

Based on these conditions, this study develops a GRU-based short-term load forecasting model that utilizes time-based features as input variables. The model is expected to provide more accurate load predictions for the SG 06 Feeder, thereby supporting operational decision-making, improving network reliability, and assisting PLN in optimizing daily load management.

The main contribution of this article lies in the development of a GRU-based short-term load forecasting model specifically tailored for the SG 06 Feeder of ULP Meureudu, which is characterized by highly fluctuating load behavior and a relatively low load factor. By utilizing time-based features as the sole input variables, the proposed model demonstrates that temporal information alone is sufficient to capture complex daily and weekly consumption patterns in a 20 kV distribution network. This study also presents a comprehensive methodological framework encompassing historical data acquisition, preprocessing, time-series windowing, model training, and performance evaluation using RMSE, MAE, and MAPE. The experimental results show that the GRU architecture outperforms baseline forecasting methods, producing more accurate and stable predictions for the SG 06 Feeder. Furthermore, the proposed model contributes to PLN's operational needs by offering a reliable forecasting tool that can support decision-making,

enhance feeder reliability, and optimize daily load management within the distribution system.

II. RELATED WORK

Previous studies have demonstrated that deep learning architectures, particularly the Gated Recurrent Unit (GRU), demonstrate strong capabilities in learning temporal patterns within time-series data and are widely applied in short-term electrical load forecasting. GRU is often considered more efficient than the Long Short-Term Memory (LSTM) network due to its simpler structure while maintaining the ability to capture temporal dependencies effectively [19]. Several studies have also integrated GRU with external variables such as weather parameters or other operational indicators. For instance, the work of Silfiani [5] and other related research shows that multimodal inputs can enhance forecasting accuracy; however, their effectiveness largely depends on the availability of complementary datasets, which are not always accessible at the feeder level.

In addition, a number of studies have proposed more sophisticated GRU-based forecasting models, including hybrid PCA-GRU frameworks, evolution-algorithm-based optimization, and GRU architectures augmented with attention mechanisms [20]–[22]. Although these approaches can yield improved prediction performance, they typically involve higher computational complexity, making them less suitable for deployment within PLN's operational units, which require lightweight, fast, and easily implementable forecasting systems.

Despite the growing body of work on GRU-based forecasting, research specifically examining the use of pure time-based features (e.g., hour, day, and week indicators) as the sole input variables in a 20 kV distribution network remains limited. In fact, previous studies on distribution systems have shown that temporal dimensions alone can sufficiently represent daily and weekly load fluctuation patterns without requiring additional external variables that may be difficult to obtain [20], [23]. This gap is particularly relevant for operational areas such as ULP Meureudu, where data structures are relatively simple and depend primarily on hourly historical load records.

Therefore, this study addresses this research gap by developing a GRU-based short-term load forecasting model that uses time-based features as the only input. The proposed model is designed to remain accurate, stable, and computationally efficient, allowing it to be directly implemented on the SG 06 Feeder, which exhibits highly fluctuating load characteristics and a low load factor. This approach aligns with PLN's operational needs for a forecasting model that is both simple and reliable, thereby supporting

improved daily load management within the 20 kV distribution network.

III. METHODS

This research was conducted through several stages, as illustrated in Figure 1, including the collection of electrical load data, data preprocessing, which involved constructing time-series sequences, the design of the Gated Recurrent Unit (GRU) model architecture, model training, and performance evaluation using statistical metrics.

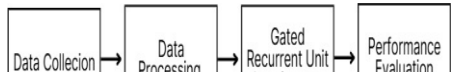


Figure 1. The Proposed Framework

3.1 Data Collection

The data used in this study consist of historical daily electrical load measurements from the SG 06 Feeder of ULP Meureudu, recorded at an hourly resolution (24 data points per day) over two years from 2023 to 2024 [24]. A detailed summary of the dataset is presented in Table 1. The dataset consists of electrical load data from the SG 06 Feeder at ULP Meureudu, recorded at an hourly resolution over the period from January 1, 2023, to December 31, 2024, resulting in a total of 17,544 data entries. The target variable is the kVA value, while the input features are derived from timestamp extraction, namely hour of the day (0–23), day of the week, and week of the year. Using time-based features as the sole input aims to evaluate the GRU’s ability to learn temporal load patterns without relying on external variables.

The data were obtained in CSV format and underwent several preprocessing steps before model training, including data cleaning (handling missing values and outliers), time transformation, and normalization to meet the requirements of the GRU neural network.

Figure 2 illustrates the daily load profile of the SG 06 Feeder at ULP Meureudu. Based on the initial data analysis, the SG 06 Feeder exhibits a load factor of less than 0.7, indicating that the distribution system is not operating efficiently. Additionally, Figure 2 illustrates a substantial difference between the average load and the peak load. This gap introduces significant challenges for load forecasting. The forecasting model must be capable of recognizing sudden load spikes at specific hours while maintaining accurate predictions during low-load periods.

Moreover, the relatively low load factor increases the risk of imbalance between power supply and consumer demand. If the load forecasting model is inaccurate, system operators may face conditions of either surplus power generation or insufficient supply during peak-load hours.

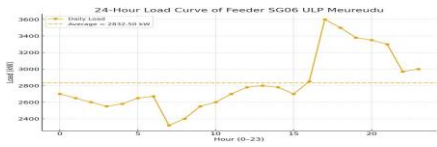


Figure 2. Daily Load Curve of the SG 06 Feeder

Table 1. Summary of the dataset	
Nama	Description
Dataset	Historical hourly electrical load data over two years
Format of the dataset	CSV or Excel format, containing timestamp and electrical load (kWh) columns
Data Source	ULP Meureudu, PLN UP3 Sigli Region

3.2 Data Pre-processing

Next, the collected data proceeds to the Feature Engineering stage, in which the timestamp is transformed into time-based features such as hour of the day (0–23), day of the week (Monday–Sunday), and week of the year. These temporal features are extracted to capture cyclical consumption patterns and periodic behaviors that commonly occur in electrical load profiles. Transforming the timestamp into structured temporal attributes enables the GRU model to better learn daily and weekly load dynamics without requiring external variables. The Feature Engineering process is carried out through several steps. First, the timestamp column is converted into a standard datetime format to ensure consistency. Next, numerical representations of time are generated by decomposing the datetime object using built-in functions such as *datetime.hour*, *datetime.weekday*, and *datetime.isocalendar() week*. These new features are then appended to the dataset as independent input variables. In addition, categorical time features (such as day names) are encoded into numerical form to ensure compatibility with the GRU network, which requires numerical inputs. The data preprocessing stage involves cleaning missing values in the kVA column, applying Min–Max normalization to scale all features within the 0–1 range, and partitioning the dataset into training, validation, and testing subsets while preserving the chronological order. Normalization is essential because the GRU model is sensitive to variations in feature scales. The load profiles before and after normalization are illustrated in Figures 3 and 4.

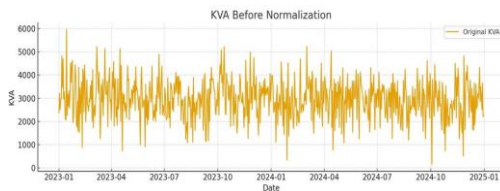


Figure 3. kVA Load Profile Before Normalization

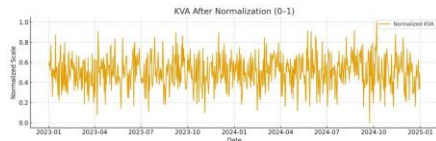


Figure 4. kVA Load Profile After Normalization

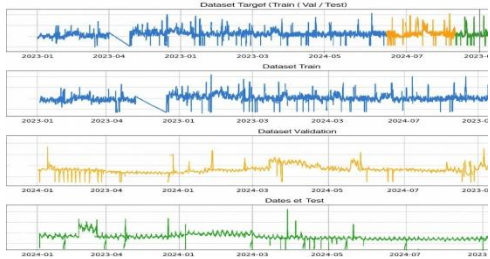


Figure 5. Dataset Partitioning

The dataset was partitioned using a time-series split method to prevent information leakage across periods. The distribution of the dataset is illustrated in Figure 5. Afterward, the dataset was partitioned into separate subsets for training, validation, and testing to ensure proper model training, tuning, and evaluation. The data were then divided into two main subsets: a training set and a testing set. The training set comprised 80% of the total samples, while the remaining 20% were allocated to the testing set.

3.3 Gated Recurrent Unit architecture

Gated Recurrent Unit (GRU) is a variant of the Recurrent Neural Network (RNN) designed to process sequential data and address the limitations of conventional RNNs in preserving long-term dependencies, particularly those caused by the vanishing gradient problem. The GRU architecture is a simpler alternative to the Long Short-Term Memory (LSTM) network. Despite its more compact structure, utilizing only two main gates, the *update gate* and the *reset gate*, GRU remains highly effective in capturing temporal dynamics and retaining essential information over extended time intervals.

One of the primary advantages of GRU lies in its computational efficiency. With fewer parameters than LSTM, GRU enables faster training while maintaining strong predictive accuracy. This makes it particularly suitable for electrical load forecasting, which involves time-series data characterized by daily, weekly, and seasonal variations as well as nonlinear fluctuations. GRU also performs reliably on small to medium-sized datasets, which are commonly encountered in electric power distribution systems.

In this study, the use of GRU offers several benefits, including the ability to learn daily and weekly load fluctuations, maintain stable training performance, and produce accurate short-term load predictions. Its lightweight gating mechanism makes GRU an optimal choice for deep learning–

based load forecasting applications that require high accuracy with efficient computation.

Based on the internal structure shown in Figure 6, the GRU consists of two main gates: the reset gate (r_t) and the update gate (z_t). The reset gate controls the extent to which the memory from the previous hidden state ($\hat{h}_{t-1}$) contributes to the formation of the new candidate memory ($\hat{h}_t$). When r_t approaches zero, the network effectively “forgets” most of the memory. Conversely, the update gate determines the proportion of old memory and new memory that will be used to generate the current hidden state (h_t). With these two gates, the GRU can adaptively regulate how much past information is retained and how much new information is incorporated. Mathematically, the GRU hidden state is computed using the following equations [22]:

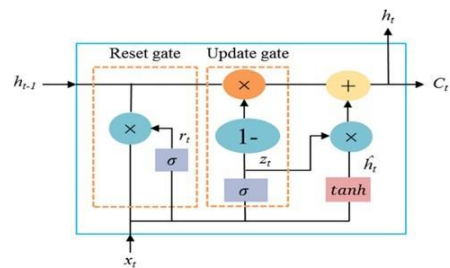


Figure 6. Gated Recurrent Unit (GRU) Architecture [9]

$$\begin{aligned} Z_t &= \sigma(W_z x_t + U_z h_{t-1} + b_z) \\ r_t &= \sigma(W_r x_t + U_r h_{t-1} + b_r) \\ \hat{h}_t &= \tanh(W_h x_t + U_h (r_t \odot h_{t-1}) + b_h) \\ h_t &= (1 - z_t) \odot h_{t-1} + z_t \odot \hat{h}_t \end{aligned}$$

Where :

x_t is the input at time step t , h_{t-1} is the input at time step t , z_t is the update gate, r_t is the reset gate, $\hat{h}_t$ is the candidate hidden state, $\odot$ denotes element-wise multiplication, σ is the sigmoid activation function, and W , U , and b are the weight matrices and bias vectors associated with each gate.

The sigmoid (σ) and hyperbolic tangent ($\tanh$) functions are used to regulate the level of activation at each stage, enabling the GRU to handle data with recurring patterns such as hours, days, or weeks in the context of load forecasting. Unlike LSTM, the GRU does not have a separate cell state, which makes the learning process lighter and faster, making it highly suitable for real-time applications and systems with limited computational resources.

In the context of electrical load forecasting, GRU is widely used to model temporal energy consumption data. Its ability to learn from time-based inputs, such as hours and days, allows the network to efficiently identify seasonal and cyclical patterns without requiring additional external variables [1], [22]. Consequently, GRU has become one of the preferred architectures for developing short-term load forecasting models that are simple, accurate, and computationally efficient. The next stage involves determining the hyperparameters used for testing, as presented in Table 2.

Table 2. GRU Model Hyperparameter		
No	Hyperparameter	Value
1	Number of GRU Layers	2 Layers
2	Number of Neurons	64 and 32 units
3	Optimizer	Adam
4	Regularisasi	L2 = 0,001
5	Optimizer	Adam
6	Learning Rate	10^{-3} dan 10^{-5}
7	Function Loss	Mean Squared Error
8	Model Training Evaluation Metrics	Mean Absolute Error
9	Model Test Evaluation Metrics	MAE dan MAPE
10	Input Sequence Length	24 Jam dan 168 jam
11	Epoch	100 dan 200
12	Batch Size	32

3.4 Performance Evaluation

The next stage involves training and validation, where the model is trained using the training dataset and tested on the validation dataset. The results of this stage are evaluated using statistical metrics such as Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). MAPE is one of the most commonly used metrics because it is relative and can represent the accuracy level in percentage terms. The MAPE formula is given as follows [26]:

IV. RESULTS AND DISCUSSION

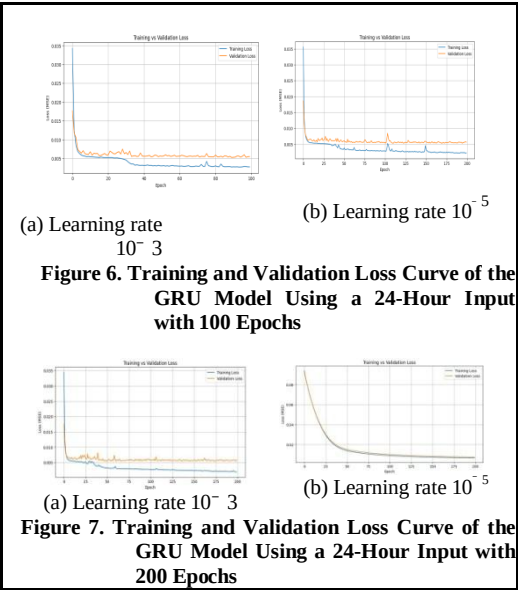
4.1 Training and Validation Results of the GRU Model

The GRU model was trained using two configurations of input sequence length, namely 24 hours and 168 hours, representing daily and weekly load patterns, respectively. These variations were designed to address the second and third research questions concerning the model’s ability to learn different temporal structures and the influence of sequence length on forecasting accuracy. In addition, two learning rates (10^3 and 10^5) and two epoch settings (100 and 200 epochs) were employed to evaluate the effects of hyperparameter selection on training stability and model performance.

4.1.1 GRU Model with a 24-Hour Input Window

The model with a 24-hour input sequence length is designed to capture daily fluctuation patterns. Figure 6 presents the training and validation loss

curves for the GRU model with a 24-hour input window trained for 100 epochs using two learning rate values, namely 10^3 and 10^5 in Figures 6(a) and 6(b), respectively



In Figure 6(a), the loss decreases rapidly during the first 10 epochs, where the training loss continues to decline until it reaches a stable value, while the validation loss follows a similar pattern, even though with slightly more fluctuation. This behavior indicates that the model successfully learns the daily load patterns and does not exhibit any significant signs of overfitting, as the gap between the training and validation loss remains consistent. In Figure 6(b), the loss reduction appears slower but smoother due to the use of a smaller learning rate. The training loss decreases steadily, and the validation loss remains within a controlled range. This demonstrates that smaller weight updates lead to a more precise training process with strong generalization stability.

Subsequently, Figure 7 presents the training results of the GRU model with a 24-hour input window using a longer training duration of 200 epochs. In Figure 7(a), with a learning rate of 10^3 , the training loss decreases until it reaches a stable low value, while the validation loss shows an initial downward trend before stabilizing, despite slight fluctuations after the 100th epoch. This pattern indicates that increasing the number of epochs provides the model with a longer learning period without introducing significant overfitting. Meanwhile, in Figure 7(b), using a learning rate of 10^5 , both the training and validation loss curves exhibit a very smooth and consistent decline. The extended training duration enables the model to reach a stable convergence point, and the close alignment of the curves suggests strong generalization capability. Overall, across all four graphs, it can be concluded that the GRU model with a 24-hour input window effectively learns daily fluctuation patterns, with the combination of a smaller learning rate and a

longer training duration providing the best training stability.

4.1.2 GRU Model with a 168-Hour Input Window

This section discusses the performance of the GRU model trained using a 168-hour input sequence, representing a full week, to capture the weekly temporal patterns of the SG 06 Feeder load. Using a 168-hour input window provides the model with a broader contextual representation compared to the 24-hour configuration, enabling it to identify weekly trends, gradual load variations, and recurring consumption patterns that occur weekly.

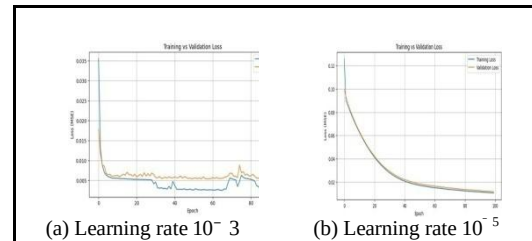
Figure 8 and Figure 9 illustrate the training and validation loss curves of the GRU model using a 168-hour input window, evaluated under two learning rates 10^3 and 10^5 and two training durations (100 and 200 epochs). These plots provide insights into the model's convergence behavior, learning stability, and generalization performance when trained to capture weekly load patterns.

Figure 8(a), which presents the results for a learning rate of 10^3 with 100 training epochs, shows that the training loss decreases rapidly during the initial epochs and stabilizes afterward. The validation loss follows a similar trajectory but exhibits more noticeable fluctuations. This behavior indicates that the model can learn the weekly load patterns; however, the relatively larger learning rate contributes to instability in the validation curve, suggesting that the weight updates may be too aggressive for this longer input sequence.

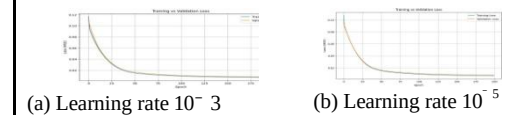
Figure 8(b), with a learning rate of 10^5 , displays a smoother and more gradual reduction in both training and validation loss. The two curves remain closely aligned throughout the training period, indicating strong generalization and stable learning dynamics. The smaller learning rate results in slower convergence but offers significantly more stable performance with minimal overfitting. In Figure 9(a), where the model is trained for 200 epochs using a learning rate of 10^3 , both loss curves continue to decline and reach lower values compared to the 100-epoch configuration.

Despite slight fluctuations in the validation loss, the curves remain reasonably aligned. The extended training duration allows the model to refine its learning further, although the relatively high learning rate still introduces small instabilities. Finally, Figure 9(b) demonstrates the most stable convergence pattern among all configurations. With a learning rate of 10^5 and 200 training epochs, both the training and validation loss curves decrease smoothly and remain consistently close. This indicates that the model achieves highly stable convergence and exhibits excellent generalization capability. The

combination of a longer training duration and a small learning rate proves to be the most effective for learning weekly temporal dependencies. Overall, the analysis of Figures 8 and 9 shows that the GRU model with a 168-hour input window is capable of capturing weekly load fluctuations effectively. While a higher learning rate accelerates loss reduction, it introduces more fluctuations, especially for longer input sequences. In contrast, a smaller learning rate paired with extended training yields the most stable and reliable learning behavior.



(a) Learning rate 10^{-3} (b) Learning rate 10^{-5}
Figure 8. Training and Validation Loss Curve of the GRU Model Using a 168-Hour Input with 100 Epochs



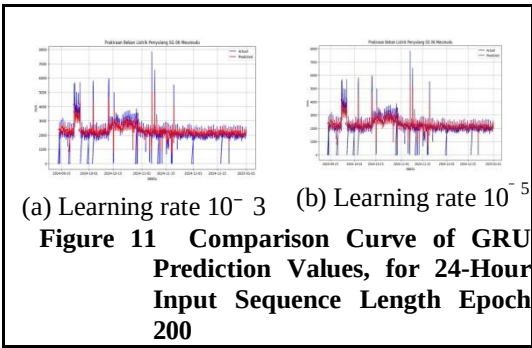
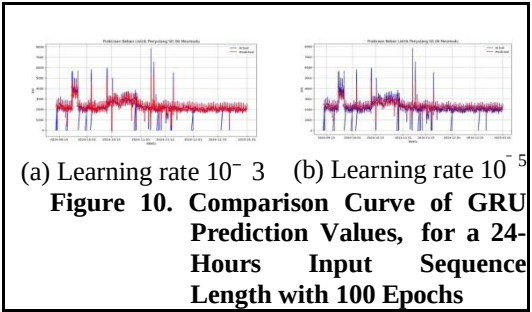
(a) Learning rate 10^{-3} (b) Learning rate 10^{-5}
Figure 9. Training and Validation Loss Curve of the GRU Model Using a 168-Hour Input with 200 Epochs

4.2 Evaluation of the GRU Model

4.2.1 Evaluation of the GRU Model with a 24-Hour Input Length

Figure 10 presents a comparison between the actual load values and the predicted values generated by the GRU model using a 24-hour input sequence at epoch 100, with learning rate 10^3 shown in Figure 10(a) and learning rate 10^5 in Figure 10(b). In both graphs, the predicted curve (red) is able to follow the fundamental pattern of the daily load as represented by the actual curve (blue), particularly within the regular load range of approximately 2,000–4,000 kVA. The GRU model demonstrates a solid ability to capture recurring daily fluctuations, especially during periods of relatively stable demand. However, when extreme load spikes occur exceeding 6,000 kVA, the model tends to underestimate the load, as the predicted values fail to reach the actual peak. This indicates that the 24-hour input window does not provide sufficient temporal context for the model to learn non-periodic anomalies or irregular peak events. A comparison of the two learning rates shows that the learning rate of 10^5 produces slightly smoother and more stable predictions, although the difference is not particularly significant at epoch 100. Figure 11 shows the prediction results at epoch 200. With a longer training duration, the prediction curves better follow the general daily

trend compared to those at epoch 100, particularly within the regular load range. Nevertheless, the model continues to exhibit underestimation during peak load events and steep declines, suggesting that the limitation arises not from training duration but from the limited temporal context inherent in the 24-hour input window. The learning rate of 10^3 tends to produce predictions that are more responsive to rapid changes, though this also results in greater fluctuations. Conversely, the learning rate of 10^5 generates smoother and more consistent prediction curves, reflecting more stable weight updates throughout training.



4.2.2 Evaluation of the GRU Model with a 168-Hour Input Length

Figures 12 and 13 present the comparison between the actual load values and the GRU model predictions using a 168-hour input window across two learning rates and two training durations. At epoch 100, the model with a learning rate of 10^3 can follow the general weekly load pattern, particularly within the regular load range of 2,000–4,000 kVA, although it still exhibits noticeable underestimation during extreme peaks exceeding 6,000 kVA. The use of a smaller learning rate (10^5) produces smoother and more stable predictions, as weight updates occur more gradually, yet the model continues to struggle in capturing sudden spikes and abrupt drops in load. When the number of epochs is increased to 200, the prediction performance improves slightly for both learning rates, with the model showing better alignment with the underlying weekly trend and reduced fluctuations in the predicted curve. However, the persistent underestimation of extreme load events indicates that these irregular peaks cannot be fully learned even with extended training. Overall, the 168-hour input length enables

the GRU to understand weekly temporal patterns effectively, while the influence of learning rate appears primarily on prediction smoothness rather than on the model’s ability to capture high-amplitude anomalies.

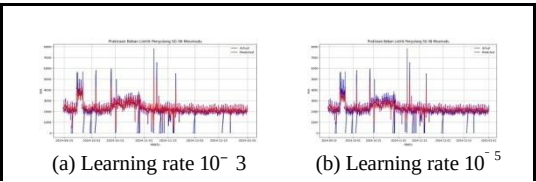


Figure 12. Comparison Curve of GRU Prediction Values , Input Sequence Length 168 Hours, Epoch 100

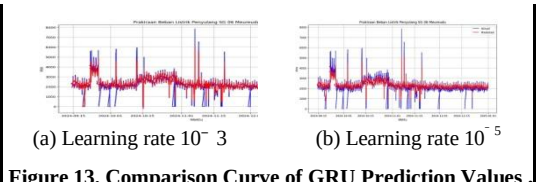


Figure 13. Comparison Curve of GRU Prediction Values , Input Sequence Length 168 Hours, Epoch 200

4.3. Discussion

The findings of this study demonstrate that the GRU model is capable of performing short-term load forecasting using time-based features alone, although with certain limitations. For regular daily load patterns, the GRU produces stable and accurate predictions. However, for rare and extreme load fluctuations, the model exhibits clear constraints. This directly addresses the first research question: while the model can generate time-based predictions, its performance is strongly dependent on the load characteristics. The results show that both daily and weekly patterns can be effectively captured. Models trained with longer input sequences (168 hours) achieve more stable and accurate performance; thus, the input sequence length has a significant impact on prediction accuracy. A prediction model that relies solely on time-based features tends to struggle when facing extreme, non-periodic spikes. Nevertheless, the GRU model still performs well in forecasting regular, recurring patterns that appear consistently on a daily and weekly basis.

Several limitations were identified in this study. First, the model does not incorporate external variables such as temperature or solar irradiation, even though these factors are available in the dataset and could potentially enhance prediction accuracy. Second, extreme load events are not sufficiently represented within the 24-hour input window, limiting the model’s ability to capture rare but critical fluctuations. Third, short-term forecasts remain sensitive to non-periodic anomalies that cannot be inferred from time-based features alone. Despite these limitations, the GRU model built solely on temporal features demonstrates considerable potential for supporting day-to-day operational decision-making in 20 kV distribution networks, particularly for feeders

characterized by relatively stable and repetitive consumption patterns.

4.4. Comparative Analysis

This section compares the performance of the proposed GRU model with one of the conventional forecasting methods commonly used in electrical load studies, namely ARIMA. The ARIMA model is selected as a benchmark because it has long served as a baseline in time-series forecasting and remains widely applied in power system analysis, despite its limitations in capturing nonlinear patterns and highly dynamic load fluctuations. The comparative analysis is conducted to evaluate the effectiveness of the proposed GRU model in forecasting the electrical load of the SG 06 Feeder by contrasting its performance against the conventional ARIMA approach.

The comparison results presented in Figure 14 show that the GRU model consistently follows load variation patterns more adaptively than ARIMA, particularly in regions with significant daily and weekly fluctuations. The GRU produces prediction curves that are more flexible and responsive to dynamic load behavior, whereas ARIMA exhibits a more rigid and lagging response to sudden changes. This occurs because ARIMA is built upon assumptions of linearity and stationarity, making it unable to capture the temporal complexity and nonlinear characteristics inherent to the SG 06 Feeder load profile. In contrast, the gating mechanism within the GRU allows the model to retain long-term memory while dynamically adjusting information updates at each time step.

The evaluation results indicate that GRU achieves substantially higher accuracy compared to ARIMA. The optimal GRU configuration using a 168-hour input window, a learning rate of 10^3 , and 200 epochs achieves an MAE of 188.13 kVA and a MAPE of 10.27%, reflecting relatively low prediction error and stable forecasting performance. ARIMA, on the other hand, exhibits larger deviations, especially during peak-load periods above 6,000 kVA. In these conditions, ARIMA fails to capture the dynamics of extreme spikes and produces flatter, less adaptive predictions. This demonstrates ARIMA’s limited capability when confronted with non-stationary and anomaly-prone energy consumption patterns such as those found in SG 06. These limitations become particularly evident when the load changes drastically from hour to hour, where ARIMA can only track short-term trends without capturing the full daily and weekly cyclic patterns.

Furthermore, the GRU model proves to have stronger generalization capabilities for weekly temporal patterns, as evidenced by the stability of

its predictions when using the 168-hour input window. The longer temporal window enriches contextual information, enabling the model to recognize recurring weekly consumption rhythms, including days with consistently higher or lower loads. Although ARIMA can incorporate seasonality through SARIMA extensions, it still struggles to represent weekly patterns that are not fully stationary. This indicates that the GRU’s RNN-based structure aligns more closely with the characteristics of 20 kV distribution loads, which exhibit recurrent but irregular consumption variations.

From an operational perspective, the advantages of GRU in capturing short- and medium-term load patterns have direct implications for feeder management. GRU’s ability to closely follow load trends makes it suitable for supporting operational decision-making, including switching scheduling, transformer loading management, and overload risk mitigation. ARIMA may still serve as a baseline model or as an initial approach when data availability is limited; however, its performance is insufficient for distribution systems with rapid and irregular dynamics such as SG 06. Overall, the comparative analysis confirms that the GRU model significantly outperforms ARIMA in terms of adaptability, accuracy, and prediction stability, making it the more appropriate core model for short-term load forecasting in 20 kV distribution networks.

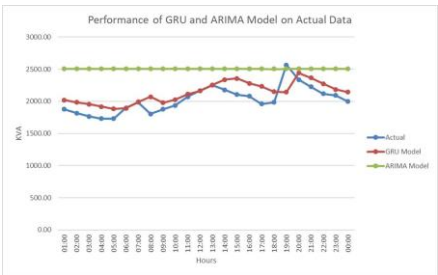


Figure 14. Comparison of the GRU Model with the ARIMA Model

Table 3. Performance Comparison of GRU and ARIMA Models Based on MAE and RMSE

Model	MAE (kVA)	RMSE (kVA)
GRU	231,74	437,36
ARIMA	548,29	783,80

The comparison Table 3 of MAE and RMSE values shows a significant performance difference between the GRU and ARIMA models in forecasting the electrical load of the SG 06 Feeder. The GRU model achieves an MAE of 231.74 kVA and an RMSE of 437.36 kVA, indicating that both the average prediction error and the deviation from actual values are relatively low. This reflects the

GRU's ability to follow load variations with higher accuracy and stability. In contrast, the ARIMA model produces an MAE of 548.29 kVA and an RMSE of 783.80 kVA, demonstrating that its prediction errors are more than twice those of the GRU model. The substantially higher RMSE value further highlights ARIMA's inability to capture sudden load spikes or extreme fluctuations in daily consumption patterns. Overall, the table clearly illustrates that the GRU model possesses far superior generalization and adaptability compared to ARIMA, making it more suitable for short-term load forecasting in 20 kV distribution networks characterized by fluctuating and nonlinear load behavior.

V. CONCLUSIONS

This study proposed a short-term load forecasting model for the SG 06 Feeder of ULP Meureudu based on a Gated Recurrent Unit (GRU) architecture that relies solely on time-based features. This approach was motivated by the research gap identified in the introduction, namely the lack of studies that evaluate GRU performance using pure temporal inputs at the 20 kV distribution-feeder level, where external variables are often unavailable or difficult to obtain. The results of this research confirm the feasibility and effectiveness of this modeling strategy.

The proposed model demonstrates that temporal features such as hour of the day, day of the week, and week of the year are sufficient for capturing the recurring daily and weekly consumption patterns of the SG 06 Feeder. The GRU architecture effectively learns these temporal dependencies, validating the premise stated in the introduction that modern RNN-based models can provide superior forecasting capability compared to conventional statistical methods. Furthermore, the study shows that the input sequence length significantly influences predictive accuracy: the 168-hour (weekly) input window consistently outperforms the 24-hour configuration, producing more stable convergence, stronger generalization, and the lowest error metrics. The best-performing configuration GRU with a 168-hour input, a learning rate of 10^{-3} , and 200 epochs achieved a MAPE of 10.27%, demonstrating reliable prediction performance suitable for feeder-level operational planning. However, the model still exhibits limitations when handling non-periodic extreme load events, which remain difficult to predict even with an extended temporal context. The overall results show that the proposed GRU-based approach provides an accurate, computationally efficient, and practical tool for supporting short-term operational decision-making in PLN's 20 kV distribution network.

Future studies may integrate external variables such as weather factors, holidays, and

operational events to improve the model's ability to estimate irregular load spikes. The incorporation of customer-specific or system-level features may also enhance prediction robustness in feeders with heterogeneous consumption profiles. Exploring more advanced architectures such as Attention-based GRU, Transformers, or hybrid CNN RNN models may yield better performance for complex or non-stationary load patterns. Additionally, implementing online or adaptive learning methods could allow the model to dynamically update as consumption behavior evolves.

REFERENCES

- [1] PT. PLN. (Persero), 2024, *Statistik PLN 2024 Data Operasional*.
- [2] L. Assaffat, 2017, *Prediksi Energi Industri Menggunakan Support Vector Machine dan Kernel*, Prosiding Seminar Nasional Teknologi dan Informatika, pp. 140–145.
- [3] PLN Wilayah Banda Aceh, 2015, *Kajian Kelayakan Proyek*.
- [4] R. Syahputra, 2012, *Prakiraan Beban Listrik Jangka Pendek Kota Banda Aceh Berbasis Logika Fuzzy*, Jurnal Rekayasa Elektriika, vol. 4, no. 2, pp. 72–78.
- [5] M. Silfiani, 2023, *Model Gabungan SARIMA dan ANN untuk Peramalan Beban Listrik*, Variance: Jurnal Ilmiah Matematika dan Statistik, vol. 5, no. 1, pp. 49–57.
- [6] H. K. Asshiddiq, 2022, *Prediksi Beban pada Transformator 150kV Menggunakan Algoritma Machine Learning*, Universitas Muhammadiyah Yogyakarta, Tesis.
- [7] M. Syukri, 2022, *Monitoring dan Identifikasi Pemakaian Beban Listrik Secara Real-Time Berbasis Artificial Neural Network*, Universitas Hasanuddin, Tesis.
- [8] M. Binoto and Y. Kristiawan, 2015, *Peramalan Energi Listrik Terjual dan Daya Listrik Tersambung di Solo Menggunakan Model Jaringan Syaraf Tiruan*, Prosiding Seminar Nasional Teknologi dan Informatika, pp. 230–236.
- [9] H. Wibowo, Y. Mulyadi, and A. G. Abdullah, 2012, *Prediksi Beban Terklasifikasi dengan Metode ARIMA*, Electrans, vol. 10, no. 1, pp. 15–22.
- [10] A. T. Rahman, A. Riza, and A. Pambudi, 2014, *Prediksi Beban Listrik pada Hari Libur Menggunakan Kombinasi Fuzzy Subtractive Clustering dan MLP*, Reka Elkomika, vol. 2, no. 2, pp. 44–51.
- [11] A. Susanti, 2015, *Implementasi Algoritma Hybrid Firefly dan MLP untuk Prediksi Konsumsi Listrik*, Universitas Airlangga, Tesis.

- [12] D. P. Ardyansyah and U. T. Kartini, 2020, *Peramalan Beban Listrik Konsumen Menggunakan Metode CBL dan Backpropagation Neural Network*, Jurnal Teknik Elektro, vol. 9, no. 1, pp. 37–44.
- [13] N. Sugiarto, U. T. Kartini, and S. I. Haryudo, 2021, *Prediksi Beban Listrik Menggunakan C-BPNN (Combined-Backpropagation Neural Network)*, Jurnal Teknik Elektro, vol. 10, no. 2, pp. 71–78.
- [14] A. A. Ningrum, I. Syarif, and A. I. Gunawan, 2021, *Prediksi Umur Transformator Menggunakan Deep Learning-LSTM*, Jurnal Ilmu Informasi dan Teknologi, vol. 5, no. 1, pp. 31–37.
- [15] R. Saputra and S. Sunardiyo, 2023, *Analisis Arsitektur JST-MLP untuk Estimasi Beban Energi Listrik Area Pelayanan PLN Salatiga*, Elkomika: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika, vol. 11, no. 1, pp. 102–109.
- [16] N. P. Sakinah, I. Cholissodin, and A. W. Widodo, 2018, *Prediksi Permintaan Menggunakan Jaringan Syaraf Tiruan Backpropagation*, J-PTIIK, vol. 4, no. 2, pp. 90–96.
- [17] H. Dermawan and S. I. Alkadri, 2025, *Prediksi Beban Listrik Harian Menggunakan ANN Backpropagation*, Journal of Engineering and Science, vol. 6, no. 1, pp. 44–50.
- [18] B. A. Septian and U. T. Kartini, 2023, *Pemodelan Prediksi Beban Listrik Menggunakan LSTM Neural Network*, Jurnal Teknik Elektro, vol. 11, no. 1, pp. 18–25.
- [19] Z. Liu and H. Bao, 2018, *Short-term load forecasting with multi-source data using GRU*, in 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, China, May, pp. 2787–2792.
- [20] A. Hasanat, A. Anwar, M. Aslam, and M. A. Khan, 2024, *Enhancing short-term load forecasting with a CNN-GRU hybrid model: Insights from single-step and multi-step forecasting on the AEP dataset*, Energy Reports, vol. 12, pp. 111–124, Jan. 2024
- [21] Y. Hua, Y. Gao, and X. Wang, 2024, *A short-term power load forecasting method using CNN-GRU with an attention mechanism*, Sustainability, vol. 16, no. 5, p. 2064, Feb.
- [22] X. Xia, Y. Zhou, J. Li, and Y. Liang, 2023, *A stacked GRU-RNN-based approach for predicting renewable energy and electricity load*, Sustainability, vol. 15, no. 3, p. 2705, Feb.
- [23] N. Nurmahaludin, 2014, *Perbandingan Metode PSO dan Regresi Linear untuk Prediksi Beban Puncak Listrik*, Poros Teknik, vol. 6, no. 2, pp. 75–83.
- [24] *Single Line Diagram PLN ULP Meureudu*, 2025.
- [25] H. K. Ramadhan, 2021, *Simulasi Prediksi Beban Energi Listrik Area Pelayanan Jaringan Kendal dengan Algoritma MLP*, Universitas Semarang, Tesis.
- [26] M. Haddin and A. Marwanto, 2022, *Peramalan Potensi PLTS dan PLTB Menggunakan Jaringan Syaraf Tiruan MLP*, Jurnal ELKOM, vol 14, no. 2, pp. 22–29.
- [27] M. P. Damayanti, 2022, *Klasifikasi Beban Listrik Berdasarkan Citra CT-scan Menggunakan K-Means Clustering*, Skripsi, UIN Walisongo Semarang.
- [28] V. Apriana and R. I. Handayani, 2017, *Analisis Algoritma Prediksi untuk Beban Listrik Jangka Pendek Menggunakan Backpropagation Neural Network*, Jurnal Ilmiah Teknologi Informasi, vol. 3, no. 2, pp. 104–109, 2017.