ENTHUSIASTIC
INTERNATIONAL JOURNAL OF APPLIED STATISTICS AND DATA SCIENCE
Volume 6, Issue 1, April 2026, pp. 57-68

Household Electricity Demand Forecasting in Batam from 2023 to 2047 Using Multilayer Perceptron Neural Network

Tiffani Giofanta Ginting [a,1], Hermansah [a,2,*], Yudhi Hanggara [a,3]

[a] Department of Mathematics Education, University of Riau Kepulauan, Batam, Indonesia
[1] giofantagintingt@gmail.com; [2] hermansah@fkip.unrika.ac.id*; [3] yudhi@fkip.unrika.ac.id
* Corresponding author

ARTICLE INFO

History
Submitted: January 14, 2026
Revised: April 22, 2026
Accepted: April 25, 2026

Keywords
Energy Demand Forecasting
Artificial Neural Network
Multilayer Perceptron (MLP)
Household Electricity Demand
Batam

ABSTRACT

The rapid growth of electricity demand in Batam, driven by increasing household and industrial consumption, necessitates accurate long-term energy forecasting. This study aimed to forecast household electricity demand in Batam from 2023 to 2047 using the multilayer perceptron (MLP) artificial neural network (ANN) model. Secondary data from PT PLN Batam (2013-2022), including customer numbers, electricity sales volume, and revenue, were analyzed. A total of 200 MLP models were trained, varying the number of hidden layers and nodes, with algorithms including BACKPROP, RPROP+, RPROP−, SAG, and SLR. The partial autocorrelation function (PACF) was used to determine the number of input layer nodes. The optimal model, using the smallest learning rate (SLR) algorithm with four hidden layers and ten nodes, achieved the best performance with the lowest mean squared error (MSE) of 35.93 and mean absolute percentage error (MAPE) of 0.47%. The projection results show a consistent increase in electricity demand, with a peak forecast of 2,114 GWh by 2047. These findings provide valuable insights for long-term energy planning and policy-making, ensuring adequate electricity supply and infrastructure development in Batam.

1. Introduction

The demand for electricity is a critical indicator of urban growth, economic activity, and quality of life, especially in rapidly developing regions. Batam, as a strategic hub for industry, trade, and residential settlements, has seen a significant increase in electricity consumption, particularly in the household sector, driven by population growth, urbanization, and rising living standards. In this context, electricity provision must not only focus on meeting current demand but also anticipate long-term needs with higher accuracy. The primary challenge lies in developing reliable demand forecasting models that can support electricity generation capacity planning, energy distribution, and policy-making [1]–[3]. As such, there is a pressing need for forecasting approaches that are both accurate and efficient, enabling more precise and sustainable energy planning.

Although numerous studies have addressed electricity demand forecasting using both traditional statistical methods and machine learning techniques such as autoregressive integrated moving average (ARIMA), exponential smoothing, artificial neural networks (ANN), and deep learning, significant gaps remain in the literature. First, many studies focus on aggregate electricity consumption forecasting, neglecting the household sector, which exhibits distinct consumption

p-ISSN 2798-253X
e-ISSN 2798-3153

patterns and growth trajectories. Second, there is limited research using electricity data specifically from Batam, which may differ from other regions in Indonesia in terms of growth dynamics and energy consumption characteristics. Additionally, few studies have systematically compared different neural network architectures and learning algorithms to determine the most effective configuration for accurate forecasting [4]–[6]. These gaps suggest the need for further research to refine existing methods, making them more contextually relevant and methodologically robust for long-term energy planning.

This research aimed to develop an effective forecasting model for household electricity demand in Batam, focusing on three key variables: the number of customers, electricity sales volume (measured in GWh), and revenue from electricity sales (measured in million Rupiah). Specifically, the study intended to optimize the multilayer perceptron (MLP) neural network model, which was selected for its ability to handle complex, nonlinear relationships within time series data [7]–[9]. By evaluating various network configurations—including different hidden layer counts, node numbers, and learning algorithms—the study sought to identify the optimal model for forecasting electricity demand from 2023 to 2047. The goal is to provide accurate long-term forecasts that not only reflect historical trends but also contribute to the strategic planning of energy infrastructure and policy in Batam.

This study contributes significantly to the field of energy forecasting, particularly in the context of electricity demand in developing urban regions. The primary contribution lies in the application of the MLP model, which has been adapted to capture the nonlinearities inherent in electricity consumption patterns. Additionally, the study offers a comprehensive comparison of various network architectures and learning algorithms, which has not been systematically addressed in previous literature. By selecting the most accurate configuration based on rigorous model evaluation [10]–[14], the research contributes a robust methodological framework that can be used for energy demand forecasting in other regions or sectors. Furthermore, the findings of this study provide valuable insights for energy planners, policymakers, and utility providers, offering a data-driven approach to long-term energy capacity planning and demand management.

To achieve the research objectives, this study employed the MLP model, utilizing time series analysis, data normalization, lag identification via partial autocorrelation function (PACF), and iterative training and testing procedures. The MLP architecture was optimized through a series of experiments involving varying numbers of hidden layers and nodes, along with different learning algorithms, such as BACKPROP, RPROP+, RPROP−, SAG, and SLR. These algorithms were selected to refine the model's ability to generalize from historical data, ensuring that the predictions for future demand are as accurate as possible. The model's performance was evaluated using standard error metrics, including mean squared error (MSE) and mean absolute percentage error (MAPE), with the best-performing model chosen based on these quantitative evaluations. This methodology was designed to yield a forecasting model that was not only accurate but also capable of adapting to complex demand patterns over time [15]–[17].

Despite its robust methodology and promising results, this study is subject to several limitations. The dataset used in this research is confined to secondary data from PT PLN Batam, covering the period from 2013 to 2022, which may limit the generalizability of the findings to other regions or time periods. Moreover, the study focuses exclusively on three key variables related to household electricity consumption and does not account for external factors such as economic fluctuations, technological advancements, or policy changes that may influence electricity demand. The model's reliance on annual data also means that it cannot capture short-term variations or seasonal trends, which might affect electricity usage [18], [19]. These limitations should be considered when applying the findings to broader contexts or longer-term predictions.

Future research should seek to address these limitations by expanding the dataset to include a broader range of variables, incorporating external influences such as economic growth, changes in electricity pricing, and technological developments in energy efficiency. Furthermore, extending the forecast horizon and utilizing higher-frequency data (e.g., monthly or quarterly data) would allow

p-ISSN 2798-253X
e-ISSN 2798-3153

for a more detailed understanding of short-term fluctuations in electricity demand. Exploring hybrid models that combine statistical approaches with advanced machine learning techniques, such as long short-term memory (LSTM) networks or hybrid ARIMA-ANN models, could also improve forecasting accuracy and robustness [20]–[24]. Lastly, testing the model's applicability in other regions or sectors would provide further validation of its effectiveness in diverse settings and contribute to the development of a more universally applicable forecasting methodology.

2. Method

2.1. Data Description

This study utilized secondary data obtained from the annual statistical reports of PT PLN Batam, covering the period from 2013 to 2022. The focus was on household electricity demand in Batam. The key variables analyzed included the number of electricity customers, electricity sales volume, and the revenue generated from electricity sales. These data were used to develop a reliable model for forecasting electricity demand in the coming decades.

2.2. Model Description

The study applied a MLP ANN model for forecasting electricity demand. The MLP model was selected due to its proven ability in handling complex, nonlinear relationships between input features and output predictions. The network architecture consisted of an input layer, one or more hidden layers, and an output layer. The MLP model's performance was determined by evaluating different configurations, including varying the number of hidden layers and nodes per layer [4]–[6].

The MLP model used in this study was a form of artificial neural network structured with three layers: input, hidden, and output. The configuration of the model involved multiple combinations of input parameters and varying numbers of neurons across hidden layers. The primary objective was to predict electricity consumption metrics, including the number of customers, electricity volume sold, and the revenue generated from electricity sales. The network architecture was designed as input, hidden, and output layers. The input layer comprised significant historical lag values from the PACF analysis. The hidden layers consisted of 1 to 5 layers, with neurons varying between 2, 3, 5, 10, 15, 20, and 25. Meanwhile, output layer generated predictions for three key energy demand variables. The activation function used in the hidden layers was the sigmoid function, which is defined as:

$$f(x) = \frac{1}{1+e^{-x}}. \quad (1)$$

The learning process in this study was executed using multiple algorithms, including BACKPROP, RPROP+, RPROP−, SAG, and SLR, to identify the most effective learning technique for the dataset. The general formulation of the MLP model is as follows:

$$Y = f(W \cdot X + b) \quad (2)$$

where X is the input lag values, W is the weights between layers, f is the activation function, b is the bias term.

2.3. Data Processing

The data underwent preprocessing, which included normalization using min-max scaling, to ensure that all input variables were within a comparable range, thereby optimizing the training process. Additionally, the number of nodes in the input layer was determined using the PACF method. PACF helped identify the most relevant lags that influence future predictions, ensuring that only the most pertinent historical data points were used for training [10]–[14].

Data normalization is an essential preprocessing step in ensuring the stability and efficiency of the model training process. The process is necessary to scale all input variables to a comparable range, enabling faster convergence and more stable learning. In this study, min–max scaling was employed, which normalized the data into a specific range, typically between 0 and 1. The normalization formula used is shown in (3).

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (3)$$

where X is the original data, X_{min} is the minimum value in the dataset, X_{max} is the maximum value in the dataset. This step ensured that the neural network treated all features equally, without bias towards any particular range of data values [7]–[9].

The identification of time series dependency patterns is crucial for constructing an effective MLP model. The model's predictive accuracy heavily depends on how well past values influence future predictions. To determine the optimal number of lags, the PACF was used. This method identified significant lags that were most directly associated with the prediction of subsequent values. The PACF was analyzed, and the significant peaks were used to define the lag structure for the MLP model, ensuring that only the most relevant historical data points were used for forecasting [20]–[22].

2.4. Combination of Parameters and Learning Algorithms

The model was trained using five learning algorithms: BACKPROP, RPROP+, RPROP−, SAG, and SLR. For each algorithm, multiple configurations were tested, varying the number of hidden layers (1 to 5 layers) and the number of nodes per layer (2, 3, 4, 5, 10, 15, 20, 25 nodes). A total of 200 distinct MLP configurations were evaluated. Hyperparameter tuning was conducted following established practices from [25] to refine the model's performance.

2.5. Model Training

During the training phase, 200 MLP models were evaluated based on their performance using the MSE and MAPE metrics. The model that demonstrated the best performance, in terms of the lowest MSE and MAPE values, was selected for further testing. The training results for all model configurations are presented in the Results and Discussion section, particularly in the Training subsection, which includes the MSE and MAPE values for each learning algorithm and network architecture.

2.6. Model Testing

For model validation, the five best-performing models from the training phase were tested using data from 2022. This testing phase aimed to evaluate the accuracy and stability of the models. The MSE and MAPE metrics were again used to assess model performance. The testing results are presented in the Results and Discussion section, particularly in the Testing subsection, highlighting the performance of each model based on the selected learning algorithms and configurations. Among the five models, the one using the SLR algorithm with four hidden layers and ten nodes was identified as the most optimal model.

The selection of the best model was based on several evaluation metrics that assessed the predictive performance and stability of the model. The primary criteria for model evaluation included MSE, MAPE, Prediction stability, and consistency across multiple variables. MSE is measure of the average squared difference between predicted and actual values, while MAPE is metric that evaluates the accuracy of forecasts by calculating the percentage error between predicted and actual values. Prediction stability is the consistency of predictions when tested on hold-out data, specifically the data from the year 2022. Meanwhile, consistency across multiple variables is the model's ability to produce stable and accurate forecasts across the three key variables. These metrics ensured that the selected model was not only accurate but also robust and consistent in its performance.

The training and validation of the MLP model were conducted in two distinct phases: training and testing phases. During the training phase, 200 combinations of MLP architectures were tested. For each combination, the MSE and MAPE were calculated. Based on these error metrics, the five best models from each learning algorithm were selected for further evaluation. In the testing phase, the top models from phase 1 were validated using the 2022 data. The model with the highest

p-ISSN 2798-253X
e-ISSN 2798-3153

prediction stability and lowest error metrics was then selected for future use in projections. This two-phase process ensured that the final model was not only accurate but also robust against overfitting, as it was validated on a separate testing dataset.

Once the best model was selected, it was utilized for projecting energy demand in the residential sector up to 2047. The projections were performed iteratively, with the output of each year used as the input for predicting the following year's demand (recursive forecasting). This methodology allowed for the extrapolation of future energy demand based on historical data, providing a reliable forecast for long-term planning and resource allocation [26]–[31].

2.7. Energy Demand Projection

The selected model, with an optimal configuration (smallest learning rate (SLR) algorithm, four hidden layers, and ten nodes), was used to project household electricity demand in Batam from 2023 to 2047. The model performance was validated by comparing the projected results with actual electricity consumption data from PT PLN Batam for the period 2013 to 2022. The forecasting results are presented in the Results and Discussion section, particularly in the Projection sub-section, which shows a steady increase in electricity demand, indicating the need for future energy planning and infrastructure development.

3. Results and Discussion

This study utilized secondary data obtained from the statistical reports of PT PLN Batam, covering the period from 2013 to 2022, with a particular focus on the household sector in Batam. The study was conducted at PT PLN Batam, with the primary object of analysis being household electricity demand, which included data on the number of electricity customers, the volume of electricity sold, and revenue from electricity sales. In the initial stage, the architecture of a MLP ANN was analyzed to examine the effectiveness of this method in projecting electricity demand from 2023 to 2047. Data processing was carried out using the R software environment through the neural network function. The network architecture was tested by varying the number of hidden layers from 1 to 5 and the number of nodes in each layer at 2, 3, 4, 5, 10, 15, 20, and 25. The network training methods included the BACKPROP, RPROP+, RPROP−, SAG, and SLR algorithms. In addition, the number of nodes in the input layer was determined using the PACF approach. The research findings are presented in three main stages, namely training, testing, and long-term electricity demand projection.

3.1. Training

In this study, the parameter combinations consisted of five learning algorithms, five variations in the number of hidden layers, and eight different numbers of nodes in each hidden layer, resulting in a total of 200 ANN models being analyzed. Meanwhile, the other hyperparameter tuning settings were determined based on the approach proposed in [25]. The complete training results for the various model configurations are presented in Table 1 and Table 2.

Table 1. MAPE of Training Performance Across Network Configurations

Hidden Layer	Hidden Layer Nodes	Learning Algorithm				
		BACKPROP	*RPROP+*	*RPROP-*	*SAG*	*SLR*
1	2	0.02077	0.02030	0.01770	0.02003	0.01920
	3	0.02072	0.01517	0.01592	0.01541	0.01494
	4	0.01680	0.01513	0.01491	0.01320	0.01553
	5	0.01600	0.01535	0.01481	0.01380	0.01503
	10	0.02027	0.01519	0.01517	0.01491	0.01494
	15	0.02043	0.01503	0.01488	0.01510	0.01486
	20	0.02130	0.01501	0.01135	0.00855	0.00954
	25	0.02166	0.01507	0.01095	0.01456	0.01164

Based on the results presented in Table 1 and Table 2, it can be concluded that each learning algorithm used in this study exhibited different performance characteristics, where each algorithm achieved its optimal performance under a specific network architecture configuration. For the BACKPROP algorithm, the best model was obtained using two hidden layers with five nodes in each layer, resulting in a prediction error of MSE of 363.867 and a MAPE of 1.532%. Meanwhile, the RPROP+ algorithm demonstrated its best performance when using a network architecture consisting of four hidden layers with fifteen nodes, achieving an MSE value of 37.9211 and a MAPE of 0.448%. These results indicate that increasing network complexity in certain algorithms can improve predictive accuracy, particularly in adaptive gradient-based learning methods. In addition, differences in network structures across algorithms suggest that there is no single optimal configuration applicable to all methods, as each algorithm exhibits different sensitivity to the number of layers and nodes used during the training process.

Table 2. MAPE of Training Performance Across Network Configurations

Hidden Layer	Hidden Layer Nodes	Learning Algorithm				
		BACKPROP	***RPROP+***	***RPROP-***	***SAG***	***SLR***
2	2	0.01659	0.01565	0.01764	0.01747	0.01894
	3	0.01544	0.01111	0.01813	0.01111	0.01342
	4	0.01534	0.01104	0.01168	0.00917	0.01378
	5	0.01532	0.01085	0.01333	0.01065	0.00675
	10	0.02109	0.01225	0.01180	0.00708	0.01064
	15	0.02103	0.01420	0.01259	0.01443	0.01374
	20	0.02105	0.01400	0.01502	0.01400	0.00572
	25	0.02108	0.01456	0.01487	0.01445	0.01398
3	2	0.02107	0.02084	0.02101	0.02101	0.02099
	3	0.02120	0.01184	0.01789	0.01924	0.02085
	4	0.02109	0.01053	0.01384	0.00964	0.01704
	5	0.02105	0.00718	0.00922	0.01029	0.00847
	10	0.02082	0.00528	0.00778	0.00529	0.00506
	15	0.02109	0.00493	0.00725	0.00999	0.00479
	20	0.02109	0.01358	0.01399	0.01412	0.00464
	25	0.02108	0.00487	0.01164	0.01475	0.00457
4	2	0.02108	0.02104	0.02106	0.02108	0.02109
	3	0.02104	0.02097	0.02106	0.02107	0.02105
	4	0.02109	0.01736	0.02090	0.02106	0.02094
	5	0.02108	0.01580	0.02095	0.02065	0.01757
	10	0.02109	0.00473	0.01635	0.00612	0.00445
	15	0.02108	0.00448	0.00473	0.00484	0.00493
	20	0.02108	0.00475	0.00446	0.00759	0.00457
	25	0.02109	0.00476	0.00468	0.00572	0.00499
5	2	0.02110	0.02108	0.02108	0.02109	0.02108
	3	0.02110	0.02107	0.02108	0.02107	0.02108
	4	0.02109	0.02104	0.02106	0.02106	0.02100
	5	0.02108	0.02094	0.02098	0.02106	0.02106
	10	0.02108	0.00581	0.02057	0.00696	0.00936
	15	0.02109	0.00463	0.00481	0.00696	0.00546
	20	0.02109	0.00463	0.00493	0.00497	0.00464
	25	0.02109	0.00457	0.00464	0.00511	0.00461

Furthermore, for the RPROP− algorithm, the best network configuration was found with four hidden layers and twenty nodes, yielding an MSE of 38.5580 and a MAPE of 0.446%. The stochastic average gradient (SAG) algorithm also achieved optimal performance with a similar architecture, namely four hidden layers with fifteen nodes, resulting in an MSE of 42.4889 and a MAPE of

0.484%. Meanwhile, the SLR algorithm produced its best performance using a network architecture consisting of four hidden layers with ten nodes, with an MSE of 35.9261 and a MAPE of 0.445%. Based on these results, configurations with four hidden layers tended to dominate in producing optimal performance across most algorithms, although the optimal number of nodes varied among methods. The five best-performing network configurations from each learning algorithm were subsequently selected as the final models for the testing stage. These findings highlight that the process of selecting neural network architecture plays a crucial role in improving prediction accuracy, and therefore the selection of hidden layers and nodes must be conducted systematically based on empirical evaluation rather than purely theoretical assumptions.

3.2. Testing

The testing stage was conducted to evaluate the five best network models obtained from the training process in order to identify the single most optimal model for projecting electricity demand. The MSE and MAPE values are presented in Table 3. The testing results of the five best network variations obtained from each learning algorithm in the training stage showed considerable variation in MSE and MAPE values. Nevertheless, all tested models produced MAPE values below 10%, which, according to forecasting accuracy criteria, may be classified as very good. The average MSE and MAPE values across the five networks were 177.894 and 0.803%, respectively. Among the five models, the network trained using the SLR learning algorithm with four hidden layers and ten nodes was proven to be the most optimal, as it produced the lowest MSE value of 39.3399 and a MAPE of 0.470%. Based on this superior performance, the SLR network with this configuration was selected for projecting household electricity demand in Batam through 2047. This selection was based on considerations of both accuracy and model stability in predicting data during the testing period.

Table 3. Error Values and Percentages from the Testing Results

Learning Algorithm	Hidden Layer	Hidden Layer Nodes	MSE	MAPE
BACKPROP	2	5	723.707	0.02111
RPROP+	4	15	40.8718	0.00473
RPROP-	4	20	42.4532	0.00489
SAG	4	15	43.1020	0.00475
SLR	4	10	39.3399	0.00470

3.3. Projection

The electricity demand projection process was carried out using a single artificial neural network model that had passed both the training and testing stages. The selected model was the network trained using the SLR learning algorithm, consisting of four hidden layers and ten nodes in each layer. This network was chosen because it demonstrated the best performance with the lowest prediction error during the testing stage. The selection of this model was intended to ensure that the projection results would have a high level of accuracy and closely approximate actual data, as demonstrated in the previous evaluation stage. In other words, this model exhibited a strong predictive capability for capturing the historical pattern of electricity demand, making it reliable for forecasting future trends. The implementation of this best-performing model was then used to project household electricity demand in Batam through 2047. The projected results were compared with the actual data reported by PT PLN Batam and are presented separately in Tables 4 and 5 to illustrate the consistency and predictive accuracy of the model.

The projection results for PT PLN Batam's electricity sales over the 2023–2047 period indicated a consistent upward trend over the next twenty-five years. The average electricity sales during this period were projected to reach 1,550 GWh per year, with a cumulative total of 38,741 GWh. The highest projected sales were expected to occur in 2047 at 2,114 GWh, while the lowest were projected for 2023 at 991 GWh. A similar increasing pattern was also evident in the number of customers and revenue from electricity sales. Based on the data presented in Table 5, the gradual increase in electricity demand from year to year required adequate support in terms of energy supply. Therefore, a policy strategy governing long-term electricity availability is needed, including efforts to plan additional generating capacity. This is important as an anticipatory measure against potential

p-ISSN 2798-253X
e-ISSN 2798-3153

operational disruptions in power generation, whether due to maintenance activities or technical failures.

Table 4. Actual Household Electricity Demand of PT PLN Batam from 2013 to 2022

Year	Household Customer Group		
	Number of Customers	*Electricity Sales (GWh)*	*Electricity Sales Revenue (Million IDR)*
2013	192,580	539	473,029
2014	191,766	583	562,654
2015	190,667	604	616,904
2016	196,127	637	650,816
2017	203,687	621	779,089
2018	273,012	689	997,738
2019	279,676	816	1,170,561
2020	286,181	876	1,250,686
2021	294,762	892	1,282,520
2022	301,207	942	1,357,562

Table 5. Projected Household Electricity Demand of PT PLN Batam from 2023 to 2047

Year	Household Customer Group		
	Number of Customers	*Electricity Sales (GWh)*	*Electricity Sales Revenue (Million IDR)*
2023	313,400	991	1,381,351
2024	329,526	1,034	1,497,118
2025	339,979	1,063	1,722,041
2026	350,372	1,111	1,892,536
2027	360,764	1,158	1,993,717
2028	372,132	1,207	2,072,979
2029	381,539	1,256	2,126,395
2030	391,943	1,308	2,340,733
2031	402,335	1,360	2,490,824
2032	412,728	1,413	2,608,728
2033	423,121	1,466	2,677,435
2034	435,303	1,511	2,751,411
2035	447,478	1,555	2,958,659
2036	459,659	1,599	3,111,224
2037	471,858	1,653	3,208,268
2038	484,063	1,699	3,263,436
2039	496,267	1,748	3,383,755
2040	508,471	1,791	3,565,797
2041	520,668	1,832	3,720,380
2042	532,865	1,875	3,799,271
2043	545,062	1,929	3,862,173
2044	557,259	1,975	4,000,863
2045	569,456	2,022	4,184,814
2046	581,653	2,069	4,327,802
2047	593,850	2,114	4,400,668

3.4. Discussion

The results from both the training and testing phases of the model demonstrate that the selected MLP model with the SLR algorithm provides superior performance in predicting electricity demand. The model’s accuracy is particularly evident in the low values of MSE and MAPE observed during testing, which are crucial indicators of a model’s precision in real-world applications. A low MSE indicates that the predictions are closely aligned with actual values, minimizing forecasting error, while the low MAPE reflects the model’s ability to maintain accuracy across different demand values. These findings suggest that the MLP model is highly effective for reliable long-term forecasting of household electricity demand. The ability to predict future demand with such accuracy is particularly important for energy planners, as it provides them with a solid foundation for making informed decisions regarding energy capacity planning, infrastructure investment, and policy development.

p-ISSN 2798-253X
e-ISSN 2798-3153

Compared with previous studies in energy demand forecasting, the proposed model demonstrates competitive predictive performance. Previous studies have applied various forecasting approaches, including ARMA-based statistical models and LSTM-based machine learning techniques, to predict energy consumption patterns [11], [15], [24]. However, accurate electricity demand forecasting remains challenging due to nonlinear characteristics, dynamic variations, and complex relationships within consumption data [23], [24], [27]. The optimized MLP model developed in this study improves prediction capability by reducing forecasting errors and generating long-term household electricity demand projections over a 25-year horizon. This improvement is supported by the optimization of key network parameters, particularly hidden layer configuration and neuron numbers, which significantly affect MLP performance [4], [6], [20]. The proposed approach provides a robust computational framework for supporting long-term electricity demand forecasting and energy planning.

This study makes significant contributions to the field of electricity demand forecasting, particularly through the introduction of an optimized MLP model. The model's performance was enhanced by carefully tuning critical parameters, such as the number of hidden layers and the number of neurons in each layer, which allowed it to better capture the complex relationships inherent in the electricity demand data. Furthermore, the application of the SLR algorithm as the learning method improved the model's stability and accuracy, ensuring that it could provide reliable predictions even in the presence of noisy or uncertain data. These improvements make the model an invaluable tool for energy planners, as it can be used to forecast electricity demand more accurately and with greater confidence. Additionally, this research adds to the growing body of knowledge on the application of ANN in energy forecasting, offering new insights and methodologies that can be applied to other regions or sectors facing similar challenges in energy demand forecasting.

The forecasting results of this study highlight a steady increase in electricity demand over the next 25 years, which carries significant implications for energy policy and planning. The projected rise in electricity consumption underscores the urgent need for proactive measures in expanding power generation capacity to meet future demand. This trend emphasizes the importance of long-term strategic energy planning, which must take into account projected population growth, urbanization, and economic development. As electricity demand increases, the development of sustainable and reliable energy sources becomes even more critical. The findings suggest that energy planners must not only focus on current demand but also anticipate future energy needs, ensuring that infrastructure investments are made well in advance to avoid shortages or supply disruptions. Additionally, the results of this study provide evidence that can inform policy decisions regarding energy efficiency, renewable energy integration, and grid modernization.

Despite the promising results of this study, several limitations should be acknowledged. One significant limitation is the reliance on secondary data from PT PLN Batam, which, while useful, may contain inherent biases or limitations in scope. The dataset used in this research covers the period from 2013 to 2022, and its generalizability to other regions or future periods may be restricted. Furthermore, the study primarily focused on three key variables—number of customers, electricity sales volume, and revenue from electricity sales—without incorporating external factors that could influence electricity demand, such as economic shifts, technological advancements, or policy changes. The model's inability to account for such external variables may limit its applicability in real-world scenarios where these factors play a significant role in shaping energy consumption patterns. These limitations should be considered when interpreting the results, particularly in the context of making long-term policy or investment decisions.

Given these limitations, future research should aim to incorporate more diverse datasets, including additional variables that may affect electricity demand, such as economic indicators, technological developments, and government policies. Integrating these external factors would enhance the model's robustness, allowing it to better capture the complexities of electricity demand in a dynamic environment. Moreover, exploring alternative forecasting techniques, such as gated recurrent units (GRU), could provide valuable insights into refining predictions, especially for time series data with complex dependencies. By considering a wider range of influencing factors and adopting different modeling approaches, future studies could improve the accuracy and generalizability of the predictions, making them more applicable to diverse contexts and helping policymakers better anticipate future energy needs.

Further research could benefit from extending the forecast horizon beyond 2047 to assess the long-term sustainability of electricity demand growth. While this study provides valuable insights for the next 25 years, extending the model's forecast period could help identify potential shifts in energy demand as a result of changing socioeconomic factors, technological advancements, or environmental considerations. Additionally, exploring hybrid models that combine traditional statistical methods with advanced machine learning techniques could yield even more accurate results. For example, combining ARIMA with deep learning techniques like LSTM or GRU could leverage the strengths of both methods and improve forecasting accuracy. Finally, applying this methodology to other regions or sectors could help validate the effectiveness and adaptability of the model across different contexts, further establishing its utility as a tool for long-term energy planning.

4. Conclusion

Based on the findings of this study, it can be concluded that the forecasting model using MLP ANN provides highly accurate predictions for household electricity demand in Batam from 2023 to 2047. The projections indicate a consistent increase in electricity demand, with the highest forecasted value reaching 2,114 GWh in 2047, reflecting a steady rise in both customer numbers and electricity sales. The model, particularly the SLR algorithm with four hidden layers and ten nodes, demonstrated the lowest error rates, making it the optimal choice for long-term projections. These findings highlight the critical need for robust energy planning and the expansion of electricity generation capacity in Batam to meet future demand. However, the study's limitations, such as the reliance on secondary data and the exclusion of external factors that may influence electricity consumption, suggest that further research incorporating a wider range of variables is needed. Overall, this research significantly contributes to the field of energy demand forecasting and provides valuable insights for policymakers and energy planners in Batam and similar regions.

References

[1] M.A. Saputra, and R. Rachmawati. "Perkembangan kawasan industri dan permukiman di Kota Batam tahun 1997-2007," *J. Bumi Indones.*, vol. 4, no. 1, pp. 409–417, 2015, Art. no. 222889, 2015.

[2] I. Saputra, D. Istardi, M. Ansori, C.B. Nugroho, and A. Maskarai. "Kajian proyeksi pemenuhan kebutuhan energi Provinsi Kepulauan Riau," *J. Integr.*, vol. 11, no. 2, pp. 125–129, Oct. 2019, doi: 10.30871/ji.v11i2.1116.

[3] N. Azizah, "Peramalan konsumsi beban listrik triple-SARIMA berdasarkan periodik seasonal (S)," B.S. thesis, Dept. Math. Inf. Technol., Institut Teknologi Kalimantan, Balikpapan, Indonesia, 2023.

[4] R. Saputra, S. Sunardiyo, A. Nugroho, and S. Subiyanto, "Analisis arsitektur jaringan syaraf tiruan-multilayer perceptron untuk efektivitas estimasi beban energi listrik PT. PLN (Persero) UP3 Salatiga," *ELKOMIKA*, vol. 11, No. 3, pp. 664–676, Jul. 2023, doi: 10.26760/elkomika.v11i3.664.

[5] H. Hermansah, D. Rosadi, A. Abdurakhman, and H. Utami, "Automatic time series forecasting using nonlinear autoregressive neural network model with exogenous input," *Bull. Electr. Eng. Inform.*, vol. 10, no. 5, pp. 2836–2844, Oct. 2021, doi: 10.11591/eei.v10i5.2862.

[6] R. Saputra, S. Sunardiyo, A. Nugroho, and S. Subiyanto, "Implementasi multilayer perceptron artificial neural network untuk prediksi konsumsi energi listrik PT PLN (Persero) UP3 Salatiga," *Elektrika*, vol. 15, no. 2, pp. 60–68, Oct. 2023, doi: 10.26623/elektrika.v15i2.6411.

[7] M.W. Purnama, S.I. Haryudo, W. Ariwibowo, and U.T. Kartini, "Peramalan kebutuhan energi listrik jangka panjang sektor rumah tangga UID Jawa Timur menggunakan metode analysis time series: Proyeksi tren quadratic dan regresi linear berbasis software Minitab V19," *J. Tek. Elektr.*, vol. 10, no. 2, pp. 485–495, May 2021, doi: 10.26740/jte.v10n2.p485-495.

[8] A.S. Caessar, I.M. Nrartha, and R. Rosmaliati, "Perbandingan metode koefisien dengan jaringan syaraf tiruan pada peramalan beban listrik jangka pendek," *Pros. Sem. Nas. Sains Teknol.*, vol. 6, 2024, pp. 176–183.

 p-ISSN 2798-253X
e-ISSN 2798-3153

[9] P.A. Nugroho, “Implementasi jaringan syaraf tiruan multi-layer perceptron untuk prediksi penyinaran matahari Kota Bandung,” *Komputa, J. Ilm. Komput. Inform.*, vol. 12, no. 1, pp. 83–90, May 2023, doi: 10.34010/komputa.v12i1.9419.

[10] A. Fathurrozi, G.D. Kalandro, A.R. Chaidir, S. Prasetyono, and M. Gozali, “Peramalan beban jangka panjang pada Gardu Induk Bangil dengan metode generalized regression neural network,” *Techné, J. Ilm. Elektrotek.*, vol. 23, no. 2, pp. 185–198, Nov. 2024, doi: 10.31358/techne.v23i2.461.

[11] H. Hermansah, M. Muhajir, and P.C. Rodrigues, “Indonesian inflation forecasting with recurrent neural network long short-term memory (RNN-LSTM),” *Enthusiastic, Int. J. Appl. Stat. Data Sci.*, vol. 4, no. 2, pp. 132–142, Oct. 2024, doi: 10.20885/enthusiastic.vol4.iss2.art5.

[12] M. Muhajir, H. Hermansah, A.P. Wicaksono, and L.A. Pratiwi, “A new approach for selecting optimal neuron number in input layer of generalized regression neural network model,” *Baghdad Sci. J.*, vol. 22, no. 8, pp. 2738–2751, Aug. 2025, doi: 10.21123/2411-7986.5034.

[13] R. Yotenka, M. Muhajir, H. Hermansah, and P.C. Rodrigues, “Comparative analysis of activation functions in recurrent neural network: An application to Indonesian inflation forecasting,” *Math. Modell. Eng. Probl.*, vol. 12, no. 3, pp. 754–762, Mar. 2025, doi: 10.18280/mmep.120302.

[14] H. Hermansah, “Strategy for enhancing GRU-RNN performance through parameter optimization,” *Math. J. Modell. Forecast.*, Vol. 3, No. 1, pp. 16–24, Jun. 2025, doi: 10.24036/mjmf.v3i1.41.

[15] H. Hermansah, “Hybrid MODWT-ARMA model for Indonesia stock exchange LQ45 index forecasting,” *Enthusiastic, Int. J. Appl. Stat. Data Sci.*, vol. 4, no. 1, pp. 51–57, Apr. 2024, doi: 10.20885/enthusiastic.vol4.iss1.art5

[16] H. Hermansah and M. Muhajir, “Comparison of accuracy between neural network and regression models in forecasting,” *MAP J.*, vol. 5, no. 1, pp. 61–69, 2023, doi: 10.15548/map.v5i1.5949.

[17] T.S. Azzahra, J.J. Cerelia, F.A.L. Nugraha, and A.A. Pravitasari, “MRI-based brain tumor classification using inception ResNet V2,” *Enthusiastic, Int. J. Appl. Stat. Data Sci*, vol. 3, no. 2, pp. 163-175, Oct. 2023, doi: 10.20885/enthusiastic.vol3.iss2.art4.

[18] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, “The M4 competition: Results, findings, conclusion and way forward,” *Int. J. Forecast.*, vol. 34, no. 4, pp. 802–808, Oct.–Dec. 2018, doi: 10.1016/j.ijforecast.2018.06.001.

[19] P. Gunoto and S. Sofyan, “Perancangan pembangkit listrik tenaga surya 100 Wp untuk penerangan lampu di ruang selasar Fakultas Teknik Universitas Riau Kepulauan,” *Sigma Tek.*, vol. 3, no. 2, pp. 96–106, Nov. 2020, doi: 10.33373/sigma.v3i2.2754.

[20] S. Zahara, Y.N. Sukmaningtyas, R.M. Akbar, and M.Z. Abidin, “Pengaruh jumlah hidden layer dan neuron pada model multilayer perceptron untuk prediksi emas,” *J. Ilm. ILKOMINFO*, vol. 8, no. 2, pp. 269–275, Jul. 2025.

[21] E. Enung, H. Kasyanto, and R.R. Sari, “Penerapan algoritma multilayer perceptron (MLP) untuk memprediksi debit di Sungai Citarum bagian hulu (Pos Pengukuran Majalaya), Kab. Bandung, Jawa Barat,” *Potensi, J. Sipil Politek.*, vol. 25, no. 1, pp. 1–8, Apr. 2023, doi: 10.35313/potensi.v25i1.4513.

[22] M. Tanaka, R. Takashima, S. Mori, and T. Oyama, “Special issue on developing sustainable energy and environmental systems in Japan: Energy crisis and challenges,” *J. Energy Eng.*, vol. 143, no. 3, 2017, Art. no. Art. no. F2017001, doi: 10.1061/(ASCE)EY.1943-7897.0000457.

[23] H.S. Hippert, C.E. Pedreira, and R.C. Souza, “Neural networks for short-term load forecasting: A review and evaluation,” *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 44–55, Feb. 2001, doi: 10.1109/59.910780.

[24] C. Deb, F. Zhang, J. Yang, S.E. Lee, and K.W. Shah, “A review on time series forecasting techniques for building energy consumption,” *Renew. Sustain. Energy Rev.*, vol. 74, pp. 902–924, Jul. 2017, doi: 10.1016/j.rser.2017.02.085.

[25] H. Hermansah, D. Rosadi, A. Abdurakhman, and H. Utami, “Selection of input variables of nonlinear autoregressive neural network model for time series data forecasting,” *Media Stat.*, vol. 13, no. 2, pp. 116–124, Dec. 2020, doi: 10.14710/medstat.13.2.116-124.

[26] W. Triyuly, S. Triyadi, and S. Wonorahardjo, “Synergising the thermal behaviour of water bodies within thermal environment of wetland settlements,” *Int. J. Energy Environ. Eng.*, vol. 12, pp. 55–68, 2021, doi: 10.1007/s40095-020-00355-z.

[27] S. Makridakis, E. Spiliotis, and V. Assimakopoulos. “Statistical and machine learning forecasting methods: Concerns and ways forward,” *PloS One*, vol. 13, no. 3, 2018, Art. no. e0194889, doi: 10.1371/journal.pone.0194889.

[28] Y. Rahmawati, and P.C. Taylor, Eds. *Empowering Science and Mathematics for Global Competitiveness*. Leide, The Nederlands: CRC Press/Balkema, 2019.

[29] T. Hong and S. Fan, “Probabilistic electric load forecasting: A tutorial review,” *Int. J. Forecast.*, vol. 32, no. 3, pp. 914–938, Jul.–Sep. 2016, doi: 10.1016/j.ijforecast.2015.11.011.

[30] S. Fan and R.J. Hyndman, “Short-term load forecasting based on a semi-parametric additive model,” *IEEE Trans. Power Syst*, vol. 27, no. 1, pp. 134–141, Feb. 2012, doi: 10.1109/TPWRS.2011.2162082.

[31] H. Yu, D. Liu, G. Chen, B. Wan, S. Wang, and B. Yang, “A neural network ensemble method for precision fertilization modeling,” *Math. Comput. Modell.*, vol. 51, no. 11–12, pp. 1375–1382, Jun. 2010, doi: 10.1016/j.mcm.2009.10.028.

 p-ISSN 2798-253X
e-ISSN 2798-3153