Energy Engineering

Tech Science Press

Doi:10.32604/ee.2025.071545

ARTICLE

Comparison of Objective Forecasting Method Fit with Electrical Consumption Characteristics in Timor-Leste

Ricardo Dominico Da Silva[1,2], **Jangkung Raharjo**[1,3,*] **and Sudarmono Sasmono**[1,3]

[1]School of Electrical Engineering, Telkom University, Bandung, 40257, Indonesia
[2]Vocational Technical School of Don Bosco, Fatumaca, Baucau, 43014, Timor-Leste
[3]Center of Excellence of Sustainable Energy and Climate Change, Telkom University, Bandung, 40257, Indonesia

*Corresponding Author: Jangkung Raharjo. Email: jangkungraharjo@telkomuniversity.ac.id

Received: 07 August 2025; Accepted: 30 September 2025; Published: 27 November 2025

ABSTRACT: The rapid development of technology has led to an ever-increasing demand for electrical energy. In the context of Timor-Leste, which still relies on fossil energy sources with high operational costs and significant environmental impacts, electricity load forecasting is a strategic measure to support the energy transition towards the Net Zero Emission (NZE) target by 2050. This study aims to utilize historical electricity load data for the period 2013–2024, as well as data on external factors affecting electricity consumption, to forecast electricity load in Timor-Leste in the next 10 years (2025–2035). The forecasting results are expected to support efforts in energy distribution efficiency, reduce operational costs, and inform decisions related to the sustainable energy transition. The method used in this study consists of two main approaches: the causality method, represented by the econometric Principal Component Analysis (PCA) model, which involves external factors in the data processing process, and the time series method, utilizing the LSTM, XGBoost, and hybrid (LSTM+XGBoost) models. In the time series method, data processing is combined with two approaches: the sliding window and the rolling recursive forecast. The performance of each model is evaluated using the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The model with the lowest MAPE (<10%) is considered the best-performing model, indicating the highest accuracy. Additionally, a Monte Carlo simulation with 50,000 iterations was used to process the data and measure the prediction uncertainty, as well as test the calibration of the electricity load projection data. The results showed that the hybrid model (LSTM+XGBoost) with a rolling forecast recursive approach is the best-performing model in predicting electricity load in Timor-Leste. This model yields an RMSE of 75.76 MW, an MAE of 55.76 MW, and an MAPE of 5.27%, indicating a high level of accuracy. In addition, the model is also indicated as one that fits the characteristics of electricity load in Timor-Leste, as it produces the lowest percentage of forecasting error in predicting electricity load. The integration of the best model with Monte Carlo Simulation, which yields a *p*-value of 0.565, suggests that the results of electricity load projections for the period 2025–2035 are well-calibrated, reliable, accurate, and unbiased.

KEYWORDS: Load forecasting; econometric PCA; LSTM; XGBoost; Monte Carlo; sliding window; rolling forecast; recursive; retraining; Timor-Leste

1 Introduction

In recent decades, electricity demand has increased including Timor-Leste. This rise is due to factors such as economic growth, population growth, climate change, and rapid technological development within the country [1]. According to the International Energy Agency (IEA), global electricity demand is expected to continue to increase at a faster rate, growing 3.4% per year through 2026 [2]. Meanwhile, according to Our World in Data, the growth in electrical energy demand in 2023 is 183,230 TWh, with fossil energy accounting for 140,231 TWh of the total use [3].

With the increasing consumption of electrical energy, forecasting electrical energy plays a crucial role, as it enables the utilization of historical data to predict future energy demand. The results of this forecasting can assist utilities and governments in planning energy transitions and deploying efficient energy distribution strategies, as well as ensuring that the supply of electrical energy remains in line with the demand for electrical loads in a developing country. If correlated with Timor-Leste, which is still dependent on fossil fuels as its primary energy source, it results in relatively high operating costs and harms the environment. According to a government report [4], the allocation of funds for the electricity sector is the highest compared to other sectors. This fact suggests that electricity load forecasting plays a crucial role. In this study, the results of electricity load forecasting can help utilities and the government to plan for more optimal and efficient electricity distribution, thereby reducing or limiting government funding allocations for the electricity sector. With a high potential for renewable energy sources [5], the results of electricity load forecasting can provide a parameter for the government and Electricidade De Timor-Leste (EDTL) as a utility in supporting Timor-Leste's energy transition roadmap, which targets 50% clean energy use by 2030 [6]. Furthermore, this effort will contribute to the acceleration of achieving the NZE roadmap globally by 2050 [7].

Electricity load forecasting plays an important role in planning the distribution of electrical energy, both for the short-term [8], medium-term, and long-term. In the context of Timor-Leste, which relies heavily on fossil-based energy sources such as diesel, electricity load forecasting with a daily [9], monthly, or annual time frequency is also very important because it can help EDTL to plan efficient and optimal distribution of electrical energy while reducing the cost of operating very expensive diesel power plants. As Timor-Leste consists of several districts and enclaves with highly fluctuating economic growth, spatial electricity load forecasting for each region [10] is important. It is possible to assist the utility in planning the allocation of electrical energy effectively in each region, so that energy distribution can be done optimally and efficiently.

Given the urgency of electricity load forecasting, which is crucial for Timor-Leste, the development and novelty of this research are notable. It is the first time that LSTM, XGBoost, and hybrid (LSTM+XGBoost) models, as well as econometric PCA models, are used to consider various external variables and predict electrical energy consumption in Timor-Leste. With a long-term time horizon of the next 10 years. Additionally, using Monte Carlo can generate a probabilistic forecast with a 5% confidence threshold for the lower bound and a 95% confidence threshold for the upper bound. It can offer various scenarios to utilities and governments as a parameter to plan efficient energy distribution, ensuring the supply of electrical energy remains in accordance with load demand in the country's sustainable development. The purpose of integrating various models in predicting electricity load in this research is to provide accurate forecasting results for use in real applications, such as energy distribution planning and considerations for energy transition. It can produce a forecasting model that is suitable for the characteristics of the electricity load in Timor-Leste. Additionally, PCA can identify factors that influence changes in electrical energy consumption in Timor-Leste.

Various studies have been conducted in predicting electrical energy consumption using different methods with applications in various contexts and using various dataset dimensions, so the following Table 1 is added to compare the methods, contributions, and limitations of these studies to define innovation in the research to be carried out:

Table 1: Current status of literature research

Literature	Methods	Contribution	Limitation
Zhang and Jánošík [11]	CatBoost and XGBoost	Integrating various machine learning models to predict short-term electricity load	Exclude external factors and refrain from performing in a country context, poor generalization.
Almuhaini and Sultana [12]	ARIMA, ARIMAX, SVR, NARX with BOA optimization	Integrating various load forecasting models with an optimization algorithm to predict long-term electricity load in the country context	Include external factors that affect the increase in electricity demand, but do not perform the PCA to extract the data, and do not investigate other models.
Rofiqi et al. [13]	Trend (linear and quadratic) and Monte Carlo	Integrating a forecasting model to predict electricity load in a small context	Exclude external factors and refrain from performing in a country context, as this leads to poor generalization.
Semmelmann et al. [14]	LSTM-XGBoost	Introducing the integration of forecasting models to predict electricity load using a dataset from a smart meter	Exclude external factors, avoid datasets from conventional electrical systems, and refrain from performing in a country context.
Edoka et al. [15]	LSTM	Developing a forecasting model from a deep learning algorithm to predict electricity in a city with short-term load forecasting	Does not investigate other models, the frequency dataset used is relatively small, and it excludes the external factors that affect the increase in electricity demand.
Sasmono et al. [16]	Qualitative and quantitative approaches	Introduced a forecasting method by integrating a quantitative and qualitative approach to predict electricity load growth spatially by identifying various customer sectors with each different variables that contribute to the increase in electricity load usage.	Does not perform well in a country context, has poor generalizability, and fails to investigate and integrate with other models.

(Continued)

Table 1 (continued)

Literature	Methods	Contribution	Limitation
This study	Econometric PCA, LSTM, XGBoost, hybrid (LSTM+XGBoost), and Monte Carlo	Integrating various forecasting models of causality method and time series method to predict long-term electricity load in the context of the country (Timor-Leste)	The complexity of integrating various forecasting models with multiple stages of data processing (discussed later)

Referring to the relevant research in the previous table, which describes the limitations of each study, such as the model used in study [11] with a short-term forecasting time horizon, it is inadequate for long-term planning, such as the transition of electrical energy in contributing to NZE. The lack of involvement of external factors that are considered to affect electric energy consumption, such as economic growth, population growth, and others, in the forecasting process, as in the study [11–15], cannot determine the causes of load changes. Load forecasting is done in a smaller context [11,13,14], meaning not in the context of a country, which allows for a narrower generalization of model application. Using data characteristic of modern electricity systems, as in the study [14], will produce data with high resolution and more detail. This system is not used in Timor-Leste, which is a developing country whose electricity system still relies on a conventional system that can only produce data with low resolution.

Due to the lack of uncertainty analysis as in the study [11–16], it is not possible to ensure the reliability of load projection with fluctuating energy consumption demand. From the limitations and weaknesses of each study described in the literature research table, the placement of this research is to integrate various models, such as econometric PCA, LSTM, XGBoost, and hybrid (LSTM+XGBoost), to predict electricity load in the context of a country with a long-term time horizon and consider external factors that affect electricity consumption. In addition, the use of Monte Carlo with a 5% confidence threshold for the lower bound and 95% for the upper bound in this study can ensure accurate load projection and provide various scenarios to the utility for long-term energy distribution planning.

The contributions from this research fill the identified gaps both in practice and academia:

(a) It offers direct solutions to real-world problems, such as better energy distribution planning, improved plant cost efficiency, and guidance on energy transition decisions.
(b) It creates a new forecasting model and analytical framework to enhance scientific literature.
(c) It gives new insight into integrating different forecasting models, including sliding window and rolling forecast recursive approaches, to predict electricity load for a country over the long term.

2 Methods

In this study, we aim to predict electrical energy consumption in Timor-Leste using an objective, quantitative approach that employs two methods: causality and the time series method. The research utilizes quantitative data, including historical electricity load data and external factors data that influence the electrical energy consumption, covering the period from 2013 to 2024. Historical data are analyzed with causality and time series methods, each represented by forecasting models such as econometric PCA, LSTM, XGBoost, and a hybrid (LSTM+XGBoost). The study's proposed prediction model is illustrated in the research block diagram shown in Fig. 1 below:

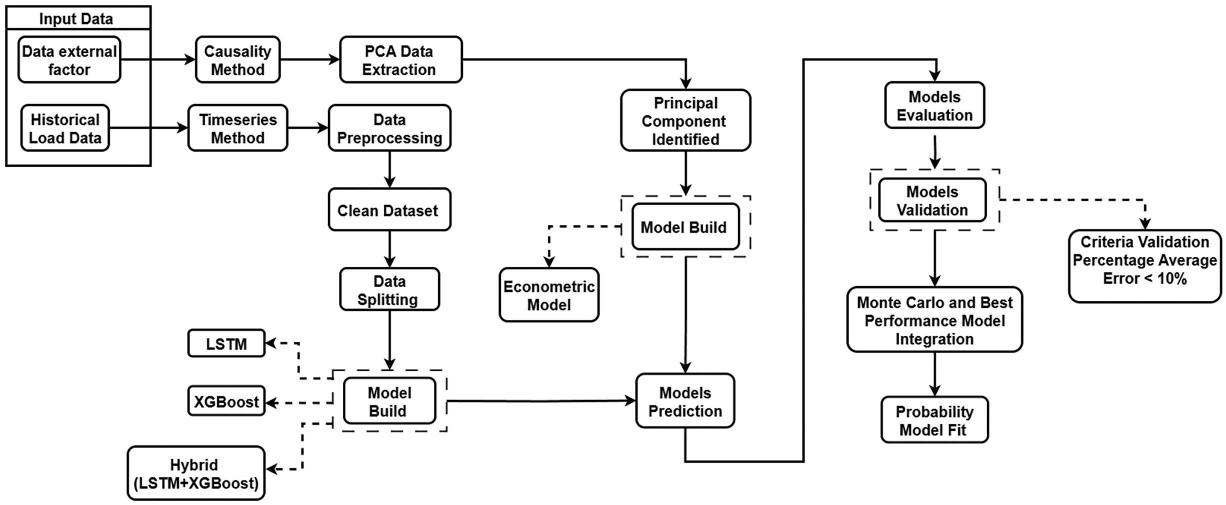

Figure 1: Research block diagram

The historical data on electrical energy consumption used in this study were obtained directly from the utility and came from existing plants. External data factors comprise 8 variables considered to influence electrical energy consumption in Timor-Leste. These variables are: GDP per capita, GDP growth, population growth, electricity access, number of tourists, temperature, and import and export activities. The data for each variable is sourced from international official websites related to Timor-Leste, such as the World Bank, government reports, and other sources [17–19].

This study forecasts electricity load using various models that represent causality and the time series method. For the causality method, data preprocessing starts with extracting 8 external variables using PCA. This step identifies the factors that contribute most to changes in electricity consumption in Timor-Leste. For the time series method, data preprocessing involves cleaning and normalizing data to ensure a uniform scale and facilitate processing.

Identified external factors with the highest percentage contribution coefficient will be integrated with an econometric model. This model uses multiple linear regression to predict electricity load with the following equation:

$$Y' = \beta_0 + (\beta_1.X_1) + (\beta_2.X_2) + (\beta_3.X_3)\ldots(\beta_n.X_n) \tag{1}$$

For the timeseries method, the preprocessed data will be further processed by forming a sequence according to the predetermined window size of 365 data points. In this research, only one window size is used, as the primary objective is to produce load forecasting with a long-term time horizon. Then, divide the data into training data (80%) and testing data (20%) to train the model and learn from the data. The model training process in data processing refers to the model architecture and feature engineering for each model approach, as illustrated in Table 2.

Table 2 below shows the feature engineering of LSTM, XGBoost, and hybrid (LSTM+XGBoost) models for each approach, as below:

Table 2: Feature engineering parameters

Feature engineering	Sliding window recursive			Rolling forecast recursive		
	LSTM	XGBoost	Hybrid	LSTM	XGBoost	Hybrid
Number of layers	2 × 64	–	2 × 64	2 × 64	–	2 × 64
Activation function	Tanh	–	ReLu	Tanh	–	ReLu
Batch size	32	–	32	32	–	32
Epoch	100	–	100	100	–	100
Dropout	0.1	–	0.1	0.1	–	0.2
N estimator	–	500	100	–	500	100
Col sample	–	0.8	0.8	–	0.8	0.8
Sub sample	–	0.8	0.8	–	0.8	0.8
Max depth	–	5	5	–	5	5
GPU tree method	–	–	–	–	√	√

From the feature engineering for each forecasting model that represents the timeseries method in the previous table. For the LSTM forecasting model, the architecture consists of two layers, each with 64 neurons. The regularization parameters of the LSTM model for each layer use dropouts of 0.1 and 0.2. As for forecasting models that employ a rolling forecast approach, particularly for XGBoost and hybrid (LSTM+XGBoost) models in this study, the GPU tree method is utilized to enhance the computation time of the XGBoost model during the data processing process.

Hyperparameter tuning is performed on each forecasting model for each approach, with the aim of increasing the model's accuracy in predicting electricity loads, thereby providing more robust and accurate results. The forecasting model, with its engineering features as outlined in the feature engineering table for both approaches, is a feature engineering parameter that yields the most optimal forecasting results for each model. For the hybrid model (LSTM+XGBoost) itself in this study, the data processing mechanism is residual-based, meaning that the input timestep data will be processed first using the LSTM model, then the residual prediction from the LSTM will be improved and further processed by the XGBoost model to produce the final prediction results.

The results of data processing from each forecasting model will be evaluated based on the model's performance in predicting electrical energy consumption in Timor-Leste. The performance metrics used to evaluate the forecasting models in this research are based on the following three metrics:

$$RMSE = \sqrt{\frac{1}{n}\sum(y_i - y_i')^2} \tag{2}$$

$$MAE = \frac{1}{n}\sum(y_i - y_i') \tag{3}$$

$$MAPE = \frac{100}{n}\sum\frac{(y_i - y_i')}{y_i} \tag{4}$$

The evaluation results of each forecasting model will be validated based on the criteria established in this study, as illustrated in Table 3 [10,16,20,21]. The performance of a model in predicting electricity load that produces a percentage error <10% indicates that the model is accurate, and the model with the best performance in forecasting electricity load. The parameter criteria to validate the model performance are as follows:

Table 3: Validation parameters

Mean average percentage error	Meaning of value
<10%	Excellent forecasting model capability
10%–20%	Good forecasting model capability
20%–50%	Reasonable forecasting model capability
>50%	Poor forecasting model capability

We will evaluate electricity load predictions for Timor-Leste based on the performance of each model in handling data. Each model's results will be validated using the criteria in the validation parameters table. The best-performing and most accurate model will be selected to project Timor-Leste's electricity load for the next 10 years.

To predict electrical energy consumption over the next 10 years, this study employs the causality method, focusing on external factors that contribute the most to the increase in energy use in Timor-Leste. PCA results define the contribution coefficient of each factor. After defining the external factor data, it is integrated with the econometric model using the formula in Eq. (1) to project the future electricity load. The highest-contribution coefficient external factor values used for prediction were obtained from the Timor-Leste government projections [18,22].

Meanwhile, the time series forecasting model, consisting of LSTM, XGBoost, and a hybrid model (LSTM+XGBoost), is used in this study to predict electrical energy consumption in Timor-Leste over the next 10 years. The three models are combined using a sliding window and a rolling forecast recursive approach. The electricity load prediction process is performed gradually using a sliding window or rolling forward approach. The prediction results are then used again as input to perform the next prediction, and the process repeats until the forecasting time horizon is complete.

This research also involves data retraining scenarios, with two data retraining scenarios for each approach. For the sliding window approach, the data retraining process is performed only once, by retraining all historical data (2013–2024). Then, the data is used according to the size of the window, which is 365 data points, to perform future forecasting until the forecasting time horizon is complete, namely until 2035. Meanwhile, the rolling forecast approach re-trains the data repeatedly for each period. For example, the model uses 365 data points according to the size of the window, then predicts the first period of 2025. After the prediction process for the 2025 period is complete, the model will retrain the data, incorporating the latest prediction results, and then predict for the next period. This process is repeated until the forecasting time horizon is complete, namely, 2035.

The data retraining scenario in the rolling forecast approach involves repeated retraining; therefore, the rolling forecast model utilizes the GPU tree method to increase computation time in data processing. This retraining scenario is carried out in research, especially for time series forecasting models, with the aim of keeping the model consistent in learning data patterns and remaining adaptive to changes in data trends, so that the forecasting results obtained are also more robust.

The results of electricity load projections for the next 10 years from each forecasting model, integrated with Monte Carlo, refer to the validated model performance based on the criteria in Table 3. The results of the identified best performance model integrated with Monte Carlo will be processed using 50,000 iterations with a confidence threshold of 5% for the lower bound and 95% for the upper bound. The objective in using Monte Carlo is to obtain a probabilistic forecast resulting from a simulation of 50,000 iterations with a 90% confidence level. It can offer various scenarios for utilities, energy policy makers, and the Timor-Leste government in planning a more efficient and optimal distribution of electrical energy. In addition to ensuring that the supply of electrical energy continues to meet the country's needs in the context of sustainable development. It can also be used as a parameter to transition electrical energy from fossil-based to clean energy in accordance with the roadmap for the use of electrical energy in Timor-Leste, which is to use 50% clean energy by 2030 [6].

3 Result and Discussion

3.1 Data Description

The historical electricity load data used in this study covers the period January 2013 to December 2024. For a description of the historical data of electrical energy consumption in Timor-Leste from 2013 to 2024, see Table 4 below:

Table 4: Description of historical electricity load data

Parameter	Value
Minimum	345.408 MW
Mean	388.518 MW
Maximum	485.958 MW
Standard deviation	45.438 MW

Based on the data of electrical energy consumption in Timor-Leste for 12 years (2013–2024), the average annual increase in electricity load is 1.65%. Fig. 2 below illustrates how electrical energy consumption changes in Timor-Leste from 2013 to 2024:

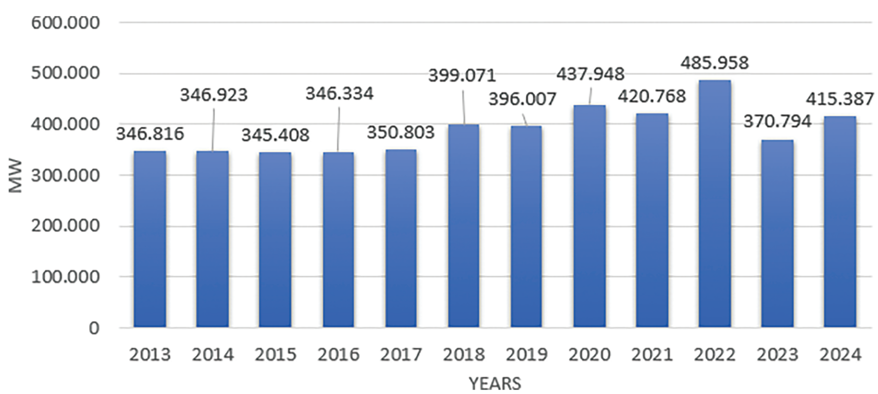

Figure 2: Historical data of electrical load

The performance evaluation results of each forecasting model used in this study to predict electrical energy consumption in Timor-Leste are illustrated in Table 5 below:

Table 5: Forecasting model evaluation result

No.	Forecasting model	Evaluation parameter			Training time (second)
		RMSE (MW)	MAE (MW)	MAPE (%)	
1	Econometric PCA	123.18	93.47	8.46	–
2	LSTM sliding window recursive	80.80	59.50	5.54	112.60
3	XGBoost sliding window recursive	90.62	68.49	6.33	10.24
4	Hybrid (LSTM+XGBoost) sliding window recursive	78.14	57.94	5.46	514.20
5	LSTM rolling forecast recursive	80.84	59.42	5.57	107.08
6	XGBoost rolling forecast recursive	89.34	66.52	6.16	3.0
7	Hybrid (LSTM+XGBoost) rolling forecast recursive	75.76	55.76	5.27	450.10

3.2 Causality Method

In this research, the causality method in the data processing process involves 8 variables as external factors that are considered to affect the consumption of electrical energy in Timor-Leste, which are then extracted using PCA to identify the percentage contribution of each external variable to the consumption of electrical energy in Timor-Leste. The extraction results obtained are as shown in Fig. 3 below:

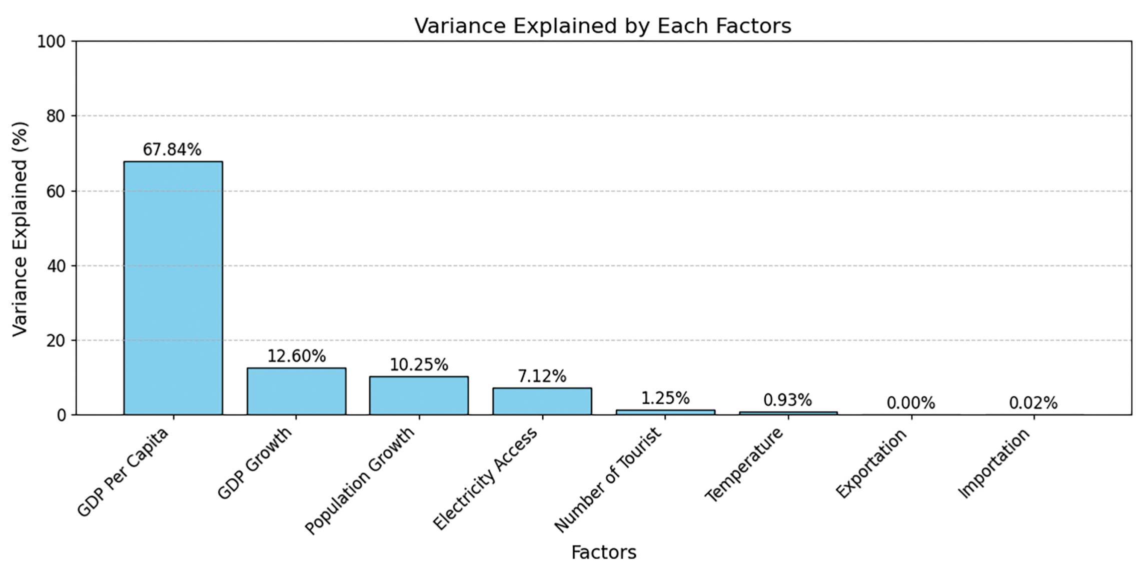

Figure 3: Variance of external factors

Based on the results of data extraction of 8 external factors using PCA, the factors with the highest contribution to changes in electrical energy consumption in Timor-Leste were identified, as illustrated in previous figure, so this study decided to involve only three external factors with the highest percentage contribution to the increase in consumption of electrical energy in Timor-Leste, namely: GDP per capita, GDP growth and population growth. The underlying reason for focusing on only these three factors is that they are important parameters as indicators of a country's development.

Involving these three important factors in the data processing process, it can provide a basis or parameter for the utility to plan the optimal distribution of electrical energy in accordance with the load demand, thereby supporting the country's sustainable development, as indicated by these three factors. The three factors identified as having the highest contribution will be integrated with an econometric model that uses multiple linear regression to predict electricity load. The value for each external factor involved in the data processing process in predicting the electricity load is obtained from the projections of the Timor-Leste government report [18,22]. For the econometric model that represents the causality method, which involves three important factors such as GDP per capita, GDP growth, and population growth in projecting electricity load, the projection results are as illustrated in Fig. 4 below:

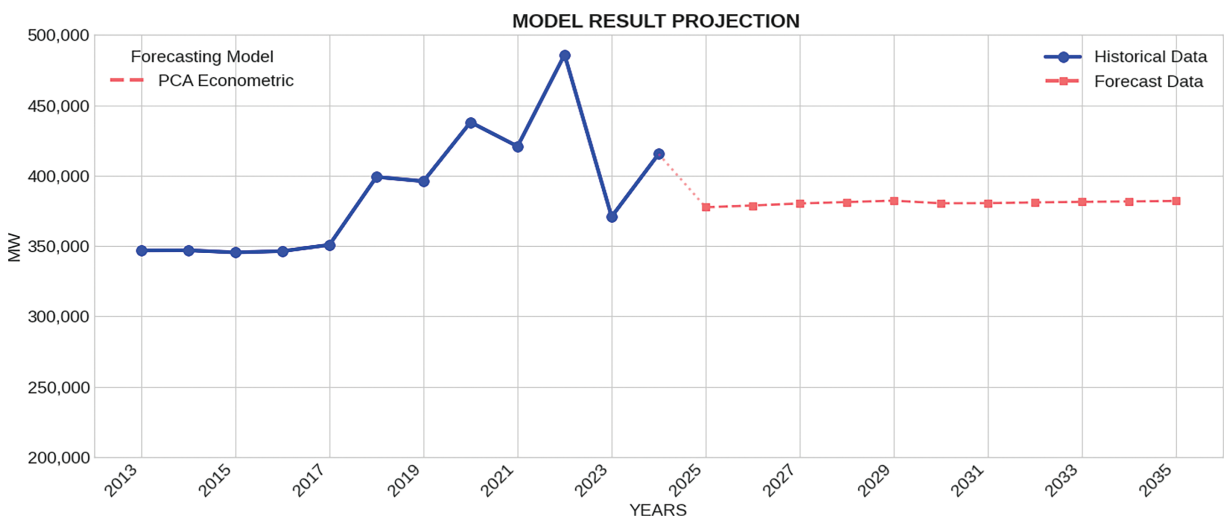

Figure 4: Projection results from the econometric PCA model

The electricity load projection for Timor-Leste over the next 10 years, generated from the PCA econometric model, considers three main external factors that most significantly influence electricity consumption: GDP per capita, GDP growth, and population growth. The projection results shown in Fig. 4 indicate an increase in electricity consumption during the forecast period, although the growth is fluctuating and not very significant. This pattern suggests that future electricity load growth will be relatively more consistent than historical patterns, although a downward trend in electricity consumption is still expected during certain periods.

3.3 Timeseries Method

For time series methods represented by various models, such as LSTM, XGBoost, and hybrid (LSTM+XGBoost) models, combined with a sliding window approach and rolling recursive forecast, to process data and predict electricity load in Timor-Leste. By using the engineering features of each approach presented in Table 2, and based on the performance evaluation of each forecasting model in handling data, as shown in Table 5, the load projection results of each forecasting model from the time series method are as illustrated in Fig. 5 below:

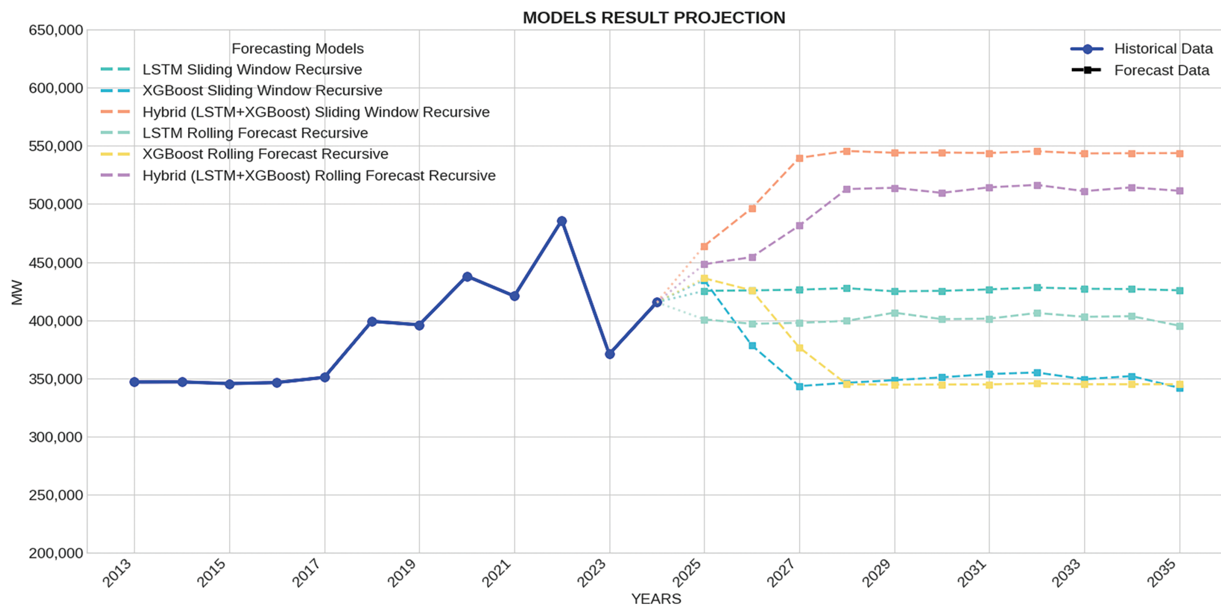

Figure 5: Projection results of all models representing the time series method

The results of electricity load projections for the next 10 years, as projected by each forecasting model represented the time series method, are illustrated in Fig. 5. For both the sliding window and rolling recursive forecasting approaches, it is shown that models like LSTM and hybrid (LSTM+XGBoost) capture the general pattern of electricity load growth, but their long-term projections vary significantly. However, the XGBoost model generates distinct electricity load projections for both approaches, as illustrated in Fig. 5. These results differ from those of other models. XGBoost struggles to learn the electricity load data patterns from Timor-Leste. Table 5 confirms this with higher error percentages compared to the LSTM and hybrid (LSTM+XGBoost) models.

Electricity load projections for Timor-Leste over the next 10 years, derived from causality and time series methods using their respective forecasting models, including econometric PCA, LSTM, XGBoost, and hybrid (LSTM+XGBoost). The projection results are as illustrated in Fig. 6 below:

The combined graph in Fig. 6 shows that each model provides different projection results for the next 10-year horizon. The econometric PCA model exhibits a moderate and relatively stable growth pattern, with lower projected values. A similar pattern is also produced by LSTM, both with the *sliding window recursive* and *rolling forecast recursive* approaches, although the LSTM projection values are higher than those of the PCA econometric model.

The XGBoost model generates projections with greater variability and demonstrates lower consistency with historical trends. The hybrid (LSTM+XGBoost) model, which employs the recursive sliding window approach, produces a relatively higher pattern and tends to stagnate at a certain level, making it less adaptable to historical changes. In contrast, the hybrid model employing the recursive rolling forecast approach yields the most stable and consistent results relative to historical growth patterns and produces realistic projection levels. These findings align with the quantitative evaluation results in Table 5, which identify the recursive rolling forecast hybrid model as the most effective approach.

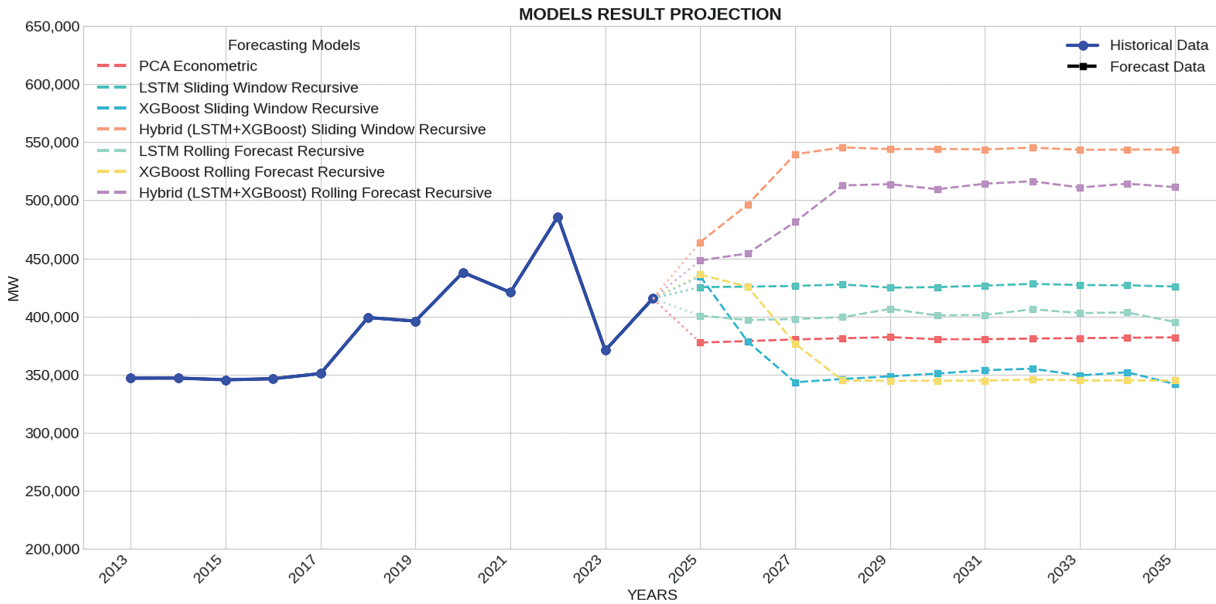

Figure 6: Projection results of all models representing causality and time series methods

These results of the best performance model projection, namely the hybrid model (LSTM+XGBoost) with a rolling forecast approach, show a continuing upward trend in electricity growth, indicating that Timor-Leste's development prospects are progressing positively. This is supported by government plans to develop infrastructure, such as upgrading airport capacity [23], building a FIFA-standard stadium [24], and facilitating foreign investment in natural resources [25], all of which are expected to encourage economic growth and future electricity consumption. These initiatives also reflect the government's preparation for becoming a full member of ASEAN [26].

3.4 Monte Carlo Method

Based on the research block diagram in Fig. 1, the forecasting model integrated with Monte Carlo was selected because it demonstrated the best performance among all models evaluated. This selection was supported by its validation results, with a forecasting error of less than 10%. As the top-performing model, it was used to project Timor-Leste's electricity load for the next 10 years, and those projections were reprocessed using Monte Carlo simulation with 50,000 iterations to provide a probabilistic forecast at the 5% lower and 95% upper confidence bounds.

The probabilistic forecast from the best-performing model and Monte Carlo integration presents various scenarios. These scenarios can be used for efficient and optimal planning of electrical energy distribution. They are also useful for planning the energy transition to accelerate achieving the roadmap goal of using 50% clean energy by 2030 in Timor-Leste [6]. The evaluation results of each forecasting model are presented in Table 5. Among these, the hybrid model (LSTM+XGBoost) with a rolling forecast approach is selected as the best-performing model because it produces the lowest average error percentage in predicting electricity load in Timor-Leste.

With the best-performing forecasting model identified, its projected electricity load data will be processed using a Monte Carlo simulation. The simulation will run 50,000 iterations, with a 5% confidence threshold for the lower limit and a 95% confidence threshold for the upper limit. The results of the 10-year electricity load projection for Timor-Leste using these Monte Carlo simulations are shown in Table 6 below:

Table 6: Monte Carlo projection results

Years	Percentile 5% (MW)	Median 50% (MW)	Percentile 95% (MW)
2025	412,324	433,245	478,755
2026	417,406	454,009	491,184
2027	441,750	467,006	521,088
2028	469,991	489,135	555,743
2029	470,373	517,029	557,539
2030	465,775	529,123	553,427
2031	469,501	531,886	559,205
2032	470,802	536,468	562,112
2033	465,258	540,002	556,871
2034	467,645	536,645	561,051
2035	464,330	538,081	558,424

Electricity load projection data generated from a Monte Carlo simulation of 50,000 iterations shows that the type of distribution that best fits the projection data is the beta distribution, with the lowest value of β = 0.678 and an α value of 1.305. Meanwhile, the p-value is 0.565 with the Kolmogorov-Smirnov (KS) statistical test. For the type of distribution shaped is right-skewed, the annual growth rate of the Monte Carlo projection results is 2.19%. The Monte Carlo projection results in Table 6 show that the growth of electrical energy consumption in Timor-Leste in the next 10 years is significant. With a stable growth trend. Although there is a decrease in certain periods, such as the 5% and 95% quartiles, there is a notable decline in 2030, 2033, and 2035. While the median quartile or 50% growth consistently increases, it only decreases in 2034.

Based on the forecasting models offered in this study to predict electrical energy consumption in Timor-Leste, and considering the evaluation results in Table 5 as well as the Monte Carlo projection in Table 6, the following interpretations can be made:

(a) The time series method, represented by its forecasting model, outperforms the causality method by producing the lowest average forecasting error percentage.
(b) Using the causality method with the PCA model, three external factors show the highest contribution to changes in electrical energy consumption in Timor-Leste: GDP per capita, GDP growth, and population growth.
(c) The three factors that are important indicators in the development of a country are involved in the data processing process, with the aim that the forecasting results obtained provide a parameter or foundation for the utility in planning the distribution of electrical energy that is efficient, optimal, and still answers the demand for electricity in Timor-Leste in the sustainable development of the country. In addition, by considering the three main factors, the utility can better identify the causes of changes in electrical energy consumption.
(d) The use of time series methods in this study, represented by LSTM, XGBoost, and hybrid (LSTM+XGBoost) models combined with sliding window and rolling forecast recursive approaches, based on the model evaluation results obtained, indicates that the forecasting model with a rolling forecast approach outperforms the model with a sliding window approach, by performing proper hyperparameter tuning as illustrated in Table 2.
(e) The hyperparameter tuning results for each approach, illustrated in Table 2, show that the model with aggressive regularization using 0.2 dropout for each layer in LSTM, as done in the hybrid model (LSTM+XGBoost), rolling recursive forecasting. The results are better than those of the other models,

indicating that the input data processed by the model contains high noise and complex data patterns that require aggressive model regularization and a complex model architecture, as well as hybrid models, to process the data effectively. This enables the model to provide robust and accurate forecasting results. High data noise and complex temporal data patterns of electricity load in Timor-Leste are caused by the uneven and inefficient distribution of electrical energy in the country. With these problems, it is expected that the results of electricity load forecasting from this research will provide insight and parameters for utilities in planning a more efficient distribution of electrical energy in Timor-Leste.

(f) Referring to the evaluation results of each forecasting model from the timeseries method, the model that performs best in handling data is identified as a hybrid model (LSTM+XGBoost) with a rolling forecast approach that produces the lowest evaluation parameter metrics compared to other models. These results also indicate that the model is the most accurate in predicting electricity load and is suitable for the characteristics of electricity load in Timor-Leste.

(g) The best performance model projections are processed using Monte Carlo with 50,000 iterations and a 5% to 95% confidence threshold, as shown in Table 6. This approach produces probabilistic forecasting results with various scenarios. These results help utilities and governments make informed decisions for efficient electrical energy distribution planning, ensuring that supply meets demand in the country's sustainable development.

(h) The Monte Carlo method processes electricity load projections from the best performance model with 50,000 iterations. It produces a p-value of 0.565, which is greater than the null hypothesis value of 0.05 (H0). This means the Monte Carlo results fail to reject hypothesis 0, indicating the projection data are unbiased and probabilistically valid.

(i) Monte Carlo by defining the beta distribution type as the distribution that best fits the data and produces the lowest value. Where the value is still within the range of 0 to 1, namely 0.678, it can be interpreted that the Monte Carlo projection data generated from the 50,000-iteration simulation is well-calibrated and can be validated that the projection results are reliable and accurate.

(j) Monte Carlo, combined with the best performance model (LSTM+XGBoost) rolling forecast, indicates that the projected electricity load in Timor-Leste is expected to grow by an average of 2.19% per year from 2025 to 2035. This is a significant increase compared to the historical 1.65% growth rate from 2013 to 2024.

The use of PCA in extracting data is quite effective in assisting data processing, as it can reduce dimensions by retaining only the most relevant variables. Through this approach, factors that influence changes in electrical energy consumption can be identified more efficiently [27]. Furthermore, the identified factors can be integrated with other forecasting models to predict electricity load. As shown in Fig. 2, historical electricity load data indicate that the COVID-19 pandemic is a significant phenomenon contributing to increased consumption in Timor-Leste. During the COVID-19 transition period, electrical energy consumption rose significantly in both Timor-Leste and Indonesia [10].

The results of this study indicate that the hybrid (LSTM+XGBoost) model, combined with a recursive rolling forecast approach, yields the best model performance, capable of producing high accuracy with an MAPE value of 5.27% and an RMSE of 75.76 MW for a 10-year long-term horizon. When compared to previous studies, the (LSTM+XGBoost) [14] and PCA–LSTM [28] models only achieved MAPE of 16.81% and 17.25%, respectively, in the short-term horizon, while studies using CNN–LSTM [29] and LSTM standalone [15] did achieve MAPE < 10% but were limited to short-term forecasting, making them less relevant for long-term energy planning needs. Meanwhile, other studies using ARIMA, ARIMAX, SVR, and NARX with BOA [12], as well as ARIMA, regression, and hybrid CST–LSTM [30] in the context of countries show higher RMSE values than the model proposed in this study. These differences confirm that the integration of econometric PCA, LSTM, XGBoost, hybrid (LSTM+XGBoost), and Monte Carlo simulation in this study not only yields a higher level of accuracy but also provides calibrated and unbiased projections.

Thus, these results are more relevant to the context of Timor-Leste's electricity system, particularly in supporting EDTL as a utility, energy policy maker, and the government in efficiently planning electricity distribution, reducing dependence on expensive fossil-based power plants [4], and considering the transition to 50% clean energy by 2030 [6], and Net Zero Emissions by 2050 [7]. Efforts in energy transition are one of the concrete actions to utilize and maximize renewable energy sources in Timor-Leste, whose potential is very high [5]. Referring to the objectives of Timor-Leste's energy roadmap, the results of electricity load forecasting from this research can provide a parameter to determine the capacity of renewable energy plants, such as solar power plants, that will be built [6]. In addition, it can provide a basis for forecasting the solar energy that will be generated by the solar power plant to adjust to the electricity demand [31]. So that the distribution of electrical energy produced still meets the needs of the electricity load in Timor-Leste, supporting the country's sustainable development.

In terms of academic aspects, the implications of this research can provide a method with a new analytical framework that can be used to conduct long-term electricity load forecasting within the country's context. It can enrich the existing literature and provide a reference base for researchers in the energy field to predict electrical energy consumption using a new approach. Additionally, it can produce a model that integrates various forecasting models to predict electrical energy consumption, providing accurate forecasting results. In discussing the process of forecasting electrical energy consumption in this study, it was identified that it can still be further developed by considering other external factors, such as government regulations or utility-related electricity tariffs for each customer sector. So that it can identify the contribution of electrical energy use from each customer sector more explicitly, such as industry, households, offices, and others.

4 Conclusion

Based on the results of predicting electricity load in Timor-Leste using forecasting models for causality and time series methods, the following conclusions can be drawn:

(a) Among the methods tested, the most effective for predicting future electrical energy consumption in Timor-Leste is the time series method, which is represented by the hybrid (LSTM+XGBoost) rolling forecast model.
(b) It was identified that the forecasting model that is suitable for the characteristics of electricity load in Timor-Leste is a hybrid model (LSTM+XGBoost) with a rolling forecast approach. The model is also identified as the best-performing model because it produces the lowest average error percentage and is the most accurate in predicting electricity load, as determined by the validation criteria established in this study.
(c) By using the causality method that involves various external factor variables in the data processing. This study identifies that the external factors contributing the most to the increase in electrical energy consumption in Timor-Leste comprise three main variables: GDP per capita, GDP growth, and population growth, which are key parameters of Timor-Leste's indicators.

For further research, the forecasting model and approach used in this study can be further developed to perform electricity load forecasting in a smaller context. It can be further developed by using various window sizes for short-term electricity load forecasting. Additionally, it can integrate the model with optimization algorithms to predict electricity load, providing more robust results.

Acknowledgement: The authors would like to thank the Postgraduate Program of Telkom University for supporting and facilitating the implementation of this research.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: Ricardo Dominico Da Silva: Conceptualization, methodology, writing—original draft, editing and translating, Jangkung Raharjo: Supervision, formal analysis and validation, Sudarmono Sasmono: Conceptualization, methodology, and validation. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The data used in this study are confidential and have limited access, so they cannot be included directly in this publication.

Ethics Approval: The data used in this study is limited and protected by a confidentiality agreement with the data provider. Therefore, the data cannot be published openly.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

Nomenclature

Abbreviation

NZE	Net Zero Emission
IEA	International Energy Agency
EDTL	*Electricidade De Timor-Leste*
GDP	Gross Domestic Product
ASEAN	Association of Southeast Asian Nations
ARIMA	Autoregressive Integrated Moving Average
BOA	Bayesian Optimization Algorithm
NARX	Nonlinear Autoregressive Exogenous
SVR	Support Vector Regressor
PCA	Principal Component Analysis
LSTM	Long-Short Term Memory
XGBoost	eXtreme Gradient Boosting
CatBoost	Categorical Boosting
RMSE	Root Mean Square Error
MAE	Mean Average Error
MAPE	Mean Average Percentage Error
GPU	Graphics Processing Unit

Symbol

$\boldsymbol{Y'}$	Prediction value
$\boldsymbol{\beta_0}$	Intercept
$\boldsymbol{\beta}$	Coefficient variable
$\boldsymbol{X}$	Value of the external factor
$\boldsymbol{y_i}$	Actual value
$\boldsymbol{y_i}$	Prediction value
$\boldsymbol{n}$	Total number of samples

References

1. Mir AA, Alghassab M, Ullah K, Khan ZA, Lu Y, Imran M. A review of electricity demand forecasting in low and middle income countries: the demand determinants and horizons. Sustainability. 2020;12(15):5931. doi:10.3390/su12155931.

2. IEA. Electricity 2024: analysis and forecast to 2026. 2024 [cited 2024 Dec 21]. Available from: https://www.iea.org/reports/electricity-2024.
3. Ritchie H, Rosado P, Roser M. Energy production and consumption. 2024 [cited 2025 Jan 20]. Available from: https://ourworldindata.org/energy-production-consumption#article-citation.
4. Informativos E, Complementares E. State budget 2024. 2024 [cited 2025 Sep 29]. Available from: https://www.mof.gov.tl/publicationdetails/orcamento-geral-de-estado-2024.
5. International Renewable Energy Agency. Timor-leste country indicators and SDGS; 2023 [cited 2025 Sep 29]. Available from: https://www.irena.org/-/media/Files/IRENA/Agency/Statistics/Statistical_Profiles/Asia/Timor-Leste_Asia_RE_SP.pdf.
6. Admin EDTL. The President of the Executive Commission of EDTL, E.P. participated in the 15th Meeting of the Expert Working Group on Energy Connectivity (EWG-EC) in Bangkok-Thailand. 2025 [cited 2025 Mar 12]. Available from: https://edtl-ep.tl/news-detail/en/2025/03/2979994931093398342981/the-president-of-the-executive-commission-of-edtl-e-p-participated-in-the-15th-meeting-of-the-expert-working-group-on-energy-connectivity-ewg-ec-in-bangkok-thailand/.
7. International Energy Agency. Net zero by 2050: a roadmap for the global energy sector [Internet]. 2021 [cited 2025 Sep 29]. Available from: https://www.energy.gov/sites/default/files/2023-06/Ex%20E%20IEA%20Net%20Zero%20by%202050.pdf.
8. Raharjo A, Budiharjo S. Short-term electric load forecasting using auto regressive-candidates area shifting technique. Int J Eng Appl. 2021;9(5):286–94. doi:10.15866/irea.v9i5.20668.
9. Aditya I, Simaremare A, Raharjo J, Suyanto, Wijayanto I. Daily power plant load prediction using grammatical evolution. In: Proceedings of the 2022 International Conference on Electrical Engineering, Computer and Information Technology; 2022 Nov 22–23; Jember, Indonesia. p. 122–6. doi:10.1109/ICEECIT55908.2022.10030558.
10. Indra Wijaya IGPO, Kristi MS, Nabila Sri Maharani JR. Data adjustment for ARIMA method to investigate the impact of pandemic on electricity consumption prediction in Indonesia. Emit J Tek Elektro. 2025;25(1):47–53.
11. Zhang L, Jánošík D. Enhanced short-term load forecasting with hybrid machine learning models: catBoost and XGBoost approaches. Expert Syst Appl. 2024;241(4):122686. doi:10.1016/j.eswa.2023.122686.
12. Almuhaini SH, Sultana N. Forecasting long-term electricity consumption in Saudi Arabia based on statistical and machine learning algorithms to enhance electric power supply management. Energies. 2023;16(4):2035. doi:10.3390/en16042035.
13. Rofiqi NZ, Bagaskoro MC, Gumilar L, Aripriharta A. Long-term electricity power forecasting at PT PLN UP3 Kediri using trend and monte carlo simulation methods. J E-Komtek. 2023;7(1):21–8. doi:10.37339/e-komtek.v7i1.1131.
14. Semmelmann L, Henni S, Weinhardt C. Load forecasting for energy communities: a novel LSTM-XGBoost hybrid model based on smart meter data. Energy Inform. 2022;5(Suppl 1):24. doi:10.1186/s42162-022-00212-9.
15. Edoka EO, Abanihi VK, Amhenrior HE, Evbogbai EMJ, Bello LO, Oisamoje V. Time series forecasting of electrical energy consumption using deep learning algorithm. Niger J Technol Dev. 2023;20(3):163–75. doi:10.4314/njtd.v20i3.1424.
16. Sasmono S, Sinisuka NI, Atmopawiro MW, Darwanto D. Macro demand spatial approach (MDSA) at spatial demand forecasting for transmission system planning. Int J Electr Eng Inform. 2015;7(2):193–206. doi:10.15676/ijeei.2015.7.2.3.
17. World Bank Group. Timor leste data [Internet]. 2025 [cited 2024 Nov 8]. Available from: https://data.worldbank.org/country/timor-leste.
18. República Democrática De Timor-Leste. Investment in strategic infrastructure, strengthening the economy and improving citizens' welfare [Internet]. 2025 [cited 2025 Jun 30]. Available from: https://www.laohamutuk.org/econ/OGE25/Prop/RelatorioEn.pdf.
19. Data W. East timor 2015. 2015 [cited 2024 Nov 8]. Available from: https://www.worlddata.info/asia/east-timor/index.php.
20. Lewis CD. Industrial and business forecasting methods: a practical guide to exponential smoothing and curve fitting. London, UK: Butterworth Scientific; 1982.

21. Norvaiša E, Bobinaitė V, Konstantinavičiūtė I, Miškinis V. Energy intensity forecasting models for manufacturing industries of catching up economies: lithuanian case. Energies. 2024;17(12):2860. doi:10.3390/en17122860.
22. Timor-Leste National Institute of Statistics. Timor-leste population and housing census 2022, thematic report—population projection 2024 [Internet]. 2024 [cited 2025 May 6]. Available from: https://inetl-ip.gov.tl/2024/05/14/census-2022-thematic-report-population-projection/.
23. Febiola A. Waskita karya develops facilities at timor-leste airport [Internet]. 2025 [cited 2025 Jul 13]. Available from: https://www.tempo.co/ekonomi/waskita-karya-garap-pembangunan-fasilitas-di-bandara-timor-leste-1513932.
24. Moniz A. FIFA plans to build an international-standard stadium in timor-leste [Internet]. 2023 [cited 2025 Jul 18]. Available from: https://id.tatoli.tl/2023/09/13/fifa-ingin-bangun-stadion-berstandar-internasional-di-timor-leste/.
25. Estrella Resources Ltd. Company announcement: ASX Release. Perth, Australia: estrella resources. 2024 [cited 2025 Sep 29]. Available from: https://company-announcements.afr.com/asx/esr/37813af3-0b69-11ef-9cd8-fa0ad8c92e08.pdf.
26. Darmawan AP, Dzulfaroh AN. Timor-leste will become a full ASEAN member in october 2025 [Internet]. 2025 [cited 2025 Jul 20]. Available from: https://www.kompas.com/tren/read/2025/06/20/161500265/timor-leste-bakal-jadi-anggota-penuh-asean-pada-oktober-2025.
27. Niyozov NN, Rikhsitillaev BK, Jalilova DA, Kuatbaev PM. Forecasting of electricity consumption by industrial enterprises with a continuous nature of production based on the principal component method (PCA). E3S Web Conf. 2023;384(56):01031. doi:10.1051/e3sconf/202338401031.
28. Fang Q, Zhong Y, Xie C, Zhang H, Li S. Research on PCA-LSTM-based short-term load forecasting method. IOP Conf Ser Earth Environ Sci. 2020;495(1):012015. doi:10.1088/1755-1315/495/1/012015.
29. Özdemir S, Demir Y, Yildirim Ö. The effect of input length on prediction accuracy in short-term multi-step electricity load forecasting: a CNN-LSTM approach. IEEE Access. 2025;13:28419–32. doi:10.1109/ACCESS.2025.3540636.
30. González Grandón T, Schwenzer J, Steens T, Breuing J. Electricity demand forecasting with hybrid classical statistical and machine learning algorithms: case study of Ukraine. Appl Energy. 2024;355(12):122249. doi:10.1016/j.apenergy.2023.122249.
31. Marhraoui S, Saad B, Silkan H, Laasri S, Hannani AE. Forecasting solar energy production across multiple sites using deep learning. Energy Eng. 2025;122(7):2653–72. doi:10.32604/ee.2025.064498.