Electric Load Demand Forecasting using Neural Networks (LSTM), Autoregressive Methods (ARMA), and Statistical Methods (Monte Carlo)

GALO JIMÉNEZ[1], ALEX ROCHA[1], DIEGO L. JIMÉNEZ[1], PABLO MUÑOZ[2], ROBERTO SALAZAR-ACHIG[1]
[1]Facultad de Ciencias de la Ingeniería y Aplicadas,
Universidad Técnica de Cotopaxi,
Latacunga,
ECUADOR

[2]Departamento de Ciencias Exactas,
Universidad de las Fuerzas Armadas - ESPE,
Latacunga,
ECUADOR

Abstract: - This article presents a comparative study of three short-term electricity demand forecasting approaches: the Monte Carlo statistical method, the autoregressive moving average (ARMA) model, and the long short-term memory (LSTM) neural network. Historical electricity demand data from the San Rafael substation in Cotopaxi, Ecuador, for the period 2020-2023 were used. The Monte Carlo method generated a synthetic series with an accuracy greater than 70% in terms of R2, but it exhibited deviations throughout the predictions. The ARMA model, while showing a prediction accuracy greater than 80% in terms of R2 During weekdays, deteriorated over longer prediction periods. The LSTM network demonstrated superior predictive capacity, reaching values below 10% for MAPE and NMSE, an RMSE below 100 kW, and a coefficient of determination. R2 above 90%, across all prediction horizons (24, 48, and 72 hours). Finally, the LSTM network offers better performance for the planning and energy management of electric power systems.

Key-Words: - Electric load demand forecasting, LSTM network, ARMA model, Monte Carlo method, Time series analysis, Deep learning.

Received: June 27, 2025. Revised: October 24, 2025. Accepted: January 21, 2026. Published: June 9, 2026.

1 Introduction

Electricity is an essential element of our current society, fundamental to industrial development and technological progress. However, this situation has increased due to recent population expansion, as well as the socioeconomic, technological, and industrial progress of each country, [1]. This is reflected in the Electricity Master Plan 2023–2032, which was prepared by the Ministry of Energy and Mines of Ecuador, which estimates that the demand for electricity will have an average annual growth of 4.4%, reaching 35.88 GWh in 2032, [2]. Therefore, proper management of electrical energy is essential for the independent and sustainable development of the country, [3].

Since electrical energy cannot be stored efficiently, it must be generated constantly to meet the demand required by end consumers, [4]. Therefore, it is necessary to adequately predict electricity demand, since it directly influences the planning processes of the national energy system, allowing for anticipating its future behavior and providing the required energy supply, [5]. Through efficient prediction, transmission system operation can be optimized, facilitating electricity market management, reducing energy waste, and maintaining the stability of the grid's energy management system, [6]. This allows the system to continuously adjust to variations in demand, thus ensuring an uninterrupted supply of electricity to different consumers, [7].

Demand forecasting can be classified into three-time horizons based on its behavior: short, medium, and long term. Short-term forecasts range from a few hours to several days, medium-term forecasts extend to weeks or months, while long-term forecasts cover periods of several years, [8]. Time series data can be defined as a chronological sequence of observations about a target variable, such as electricity demand. In this context, demand-related time series often exhibit complex behavior, as they tend to be time-varying and do not follow a defined linear pattern, [4], [9].

Electricity demand prediction can be classified into three main approaches: statistical analysis, machine learning, and deep learning. Statistical models, such as Markov chains (MC), exponential smoothing, Monte Carlo, and autoregressive moving average (ARMA) models, are suitable for linear data but have limitations when dealing with nonlinear time series. Machine learning models, including decision trees (DT), support vector machines (SVM), and artificial neural networks (ANN), improve prediction capabilities in nonlinear environments, although their performance declines over long-term horizons. Deep learning models, on the other hand, are designed to capture highly nonlinear relationships through multiple layers, offering better results in complex time series. Among them, recurrent neural networks (RNNs) stand out for their ability to process sequences and retain historical information, although they struggle to maintain long-term dependencies due to gradient vanishing or explosion. To overcome this limitation, long short-term memory (LSTM) networks were introduced, which incorporate a cellular memory capable of preserving relevant information, consolidating themselves as one of the most effective techniques for predicting time series, [4], [9], [10].

In this work, we propose to compare three models for predicting electricity demand: two statistical models and a deep learning model, such as the Monte Carlo method, the autoregressive moving average (ARMA) model, and finally a deep learning-based model, the LSTM network. The main objective of this research is to make predictions over three evaluation horizons, in order to evaluate and compare the performance of each model in these contexts. To this end, validation techniques such as the error metrics normalized mean squared error (NMSE), root mean squared error (RMSE), mean absolute percentage error (MAPE), and coefficient of determination (R^2) will be used, which will allow quantifying the accuracy of the predictions and determining the most appropriate model according to the time horizon analyzed.

2 Theoretical Framework

In this section, the operation and equations that characterize each of the three prediction models used in this study will be described.

2.1 Monte Carlo Method

This technique was initially developed by the Polish mathematician Stanislaw Ulam, who participated in the Manhattan Project. After the war, Ulam used to play numerous games of solitaire, interested in analyzing the distribution of an experiment, [11].

With the Monte Carlo Method, synthetic time series can be generated through random sampling. Although there are various ways to apply the Monte Carlo method, it generally follows this pattern: first, determine the statistical distributions for the values to be generated. Second, repeatedly draw random samples from the determined distributions to obtain the desired result, [12]. For this study, the normal probability distribution was used. This distribution is fully defined by only two parameters, mu (μ) and sigma (σ). The mean (μ) represents the expected value or central tendency of the distribution, indicating where the data clusters, the standard deviation (σ) measures the dispersion or spread of the data around the mean, [11], [13]. Mathematically, it is represented in equation (1), through which random numbers are generated from the mean and standard deviation. In addition, the Equations for determining its parameters are shown in Table 1.

$$f\ (X; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2s^2}} \tag{1}$$

Table 1. Equations for Determining the Parameters of the Normal Distribution

Parameter	Equation	
Mean	$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$	(2)
Variance	$s^2 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})}{N-1}$	(3)
Standard Deviation	$s = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})}{N-1}}$	(4)

Source: created by the authors.

where:
x: is the value of the sample.
$\mu, \bar{x}$: is the average.
σ^2: is the variance.
σ: is the standard deviation.
N: represents the number of measurements or data

2.2 Autoregressive Moving Average Model.

ARMA models are dynamic, discrete, linear models in time in which the output at each time point depends on past outputs (AR) and past inputs (MA), [14]. It consists of two parts: an autoregressive term of order p and a moving average term of order q, designated as ARMA (p, q). This model combines both terms to forecast future values of univariate time series, considering the linear dependence between past observations (AR) and historical residual errors (MA), [15], [16]. Mathematically, it is expressed in equation (5).

$$F_t = c + \varepsilon_t + \sum_{i=1}^{p} \phi_i F_{t-i} + \sum_{j=1}^{q} \theta_j F_{t-j} \qquad (1)$$

where:
F_t: Value of the time series at time t.
c: Constant (mean term), explains a base level around which the series fluctuates.
ε_t: White noise term at time t, with zero mean and constant variance.
ϕ_i: Autoregressive (AR) coefficients for lags.$i = 1, \ldots., p$
θ_j: Moving average (MA) coefficients for lags.$j = 1, \ldots., q$
p: Moving average (MA) coefficients for lags.
q: Order of the moving average component.

The autoregressive moving average (ARMA) model treats data formed over time as a random time series. The value at the time t of the sequence is not only related to the value at time. $(t - i)$, but also the amount of random interference at the time is related $(t - j)$, thus establishing a model for predicting future values, [17].

The autoregressive component $\sum_{i=1}^{p} \phi_i F_{t-i}$ models the dependence of the current value on its p past values, while the moving average component $\sum_{j=1}^{q} \theta_j F_{t-j}$ captures the impact of the past errors. For the model to be valid, two conditions must be satisfied: Stationarity and invertibility. These conditions ensure that the process is stable and that the impact of past errors diminishes over time, [17].

2.3 Long Short-Term Memory Network

The LSTM neural networks are an advanced variant of recurrent neural networks (RNNs), specifically designed to capture long-term temporal dependencies in data sequences. Unlike traditional RNNs, LSTMs incorporate a specialized internal structure that allows them to retain relevant information, [18]. This structure includes a memory cell and three types of gates: the input gate, the forget gate, and the output gate. These gates dynamically regulate the input, retention, and output of information, allowing the network to learn complex patterns in time series with greater accuracy and stability, [5], [19]. Its architecture is shown in Figure 1.

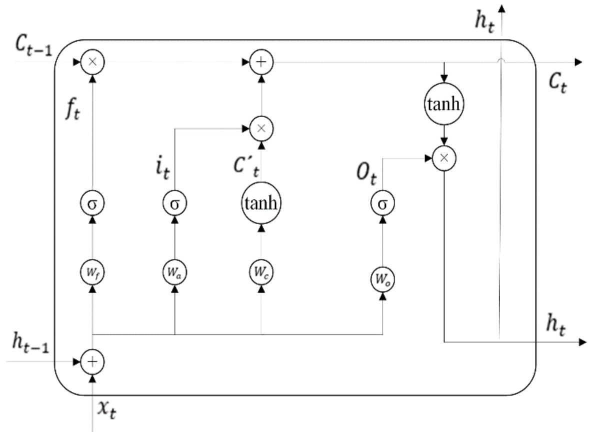

Fig. 1: Long Short-Term Memory Neural Network
Source: created by the authors

where:
x_t: Input value
h_{t-1} *and* h_t: Output value at time $t - 1$ and t
C_{t-1} *and* C_t: Cell state at time $t - 1$ and t
$b = \{b_a, b_f, b_c, b_o\}$:Biases for the input gate, forget gate, cell state, and output gate.
$\overrightarrow{W_1} = \{W_a, W_f, W_c, W_o\}$:Weight matrices for the input gate, forget gate, cell state, and output gate.
$\overrightarrow{W_2} = \{W_{ha}, W_{hf}, W_{hc}, W_{ho}\}$:Recurrent weight matrices.
$\vec{a} = \{i_t, f_t, C_t, O_t\}$:The output activations for the input gate, forget gate, candidate cell state, and output gate.

The forgetting gate f_t uses x_t and h_{t-1} as input to calculate the information to be retained in C_{t-1}, using a sigmoid activation. While the input gate i_t uses x_t and h_{t-1} to calculate C_t, using a sigmoid activation. Furthermore, the output gate O_t regulates the output of the LSTM cell by considering C_t, applying sigmoid and tanh layers. Mathematically, forward learning for an LSTM network is given by the equations (6)-(10).

$$i_t = \sigma(W_a x_t + W_{ha} h_{t-1} + b_a) \qquad (2)$$

$$f_t = \sigma(W_f x_t + W_{hf} h_{t-1} + b_f) \qquad (3)$$

$$C_t = f_t \times C_{t-1} + i_t \times tanh(W_c x_t + W_{hc}(h_{t-1} + b_c)) \qquad (4)$$

$$O_t = \sigma(W_o x_t + W_{ho} h_{t-1} + b_o) \qquad (5)$$

$$h_t = O_t \times tanh(C_t) \qquad (6)$$

where σ and $tanh$ are activation functions, and $\times$ represents vector multiplication. The sigmoid function generates values between 0 and 1, followed by process multiplication, which helps determine whether past information should be remembered or forgotten, [19], [20]. This is represented by equation (11).

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{7}$$

On the other hand, the tanh function is used to regulate the magnitude and sign of the information entering or leaving the cell. Its range between -1 and 1 allows for representing positive and negative contributions, preventing values from skyrocketing and keeping the information within controlled limits, [19]. This equation is defined in (12).

$$tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{8}$$

In general, the LSTM network learns by computing the output from equations (6) – (10). Subsequently, the error between the obtained values and the expected data is determined, and this error is back-propagated through the input gate, the memory cell, and the forget gate, [21].

3 Methodology

To identify the most optimal model for predicting electricity demand, the methodology shown in Figure 2 was used. This consists of evaluating electricity demand data, developing each of the three proposed models, and validating their results to select the one that shows the best performance based on the lowest prediction error. The steps of the methodology are described below.

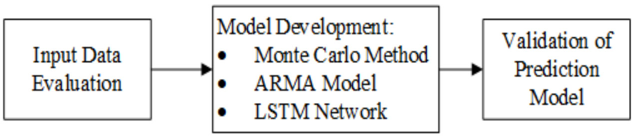

Fig. 2: Proposed methodology for electricity demand forecasting
Source: created by the authors

3.1 Input Data Evaluation

The electrical power data were obtained from the Cotopaxi Provincial Electric Company (ELEPCO) records from 2020 to 2023, belonging to the San Rafael substation in the Latacunga canton, Cotopaxi province. The information was released in hourly resolution, identifying daily and annual patterns in electrical demand, allowing its behavior to be characterized as a basis for use in the various prediction models.

Figure 3 shows the evolution of average annual power, reflecting growth in the average demand for electrical power in the area registered by ELEPCO during the 2020–2023 period. This behavior may be related to the growth in energy consumption driven by population growth, industrial development, and the increased use of electronic devices in daily life. This is reflected in the average growth in electricity demand for each year, reaching 1,442 MW for 2020, 1,542 MW for 2021, 1,651 MW for 2022, and 1.7 MW for 2023, reflecting a progressive increase in demand over time.

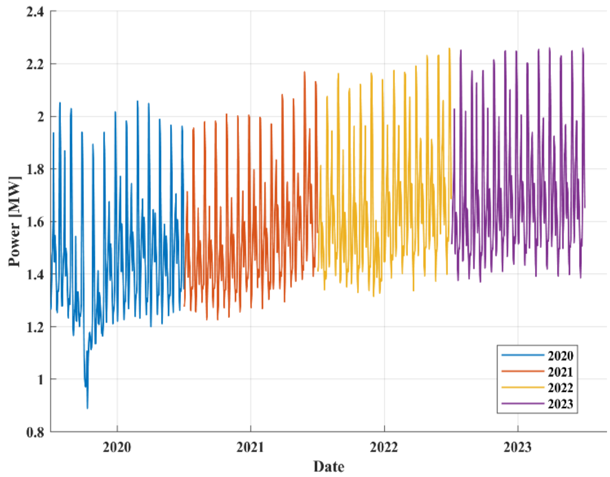

Fig. 3: Annual behavior of electricity demand for the period 2020-2023
Source: created by the authors

Additionally, it can be observed that 2020 shows greater variability, with higher dispersion and more pronounced drops in demand. On the other hand, 2021 and 2022 exhibit more stable behavior and a slight upward trend. Finally, 2023 presents a less pronounced growth compared to the previous years.

3.2 Model Development

This section details the development and application of three different methodologies for electricity demand forecasting: the Monte Carlo method, the ARMA model, and the LSTM network. For each method, the flowchart is described, beginning with data preparation and analysis, followed by the specific configuration of each model, and ending with the forecast generation process and the assessment of their accuracy using defined error metrics.

3.2.1 Monte Carlo Method

Figure 4 represents the development of the Monte Carlo method, where, first, the electricity demand database was imported. Next, the mean and standard deviation were calculated for each of the 24 hours of the day. With these parameters, random numbers were generated for each hour from the normal probability distribution, representing possible consumption scenarios for a full day. Finally, the Normalized Mean Squared Error (NMSE) of less than 30% was incorporated as a validation parameter, with the aim of comparing the synthetic series generated with the real data. This condition allows filtering and selecting those synthetic days that present a greater similarity in trend and magnitude with respect to the behavior of the electrical demand

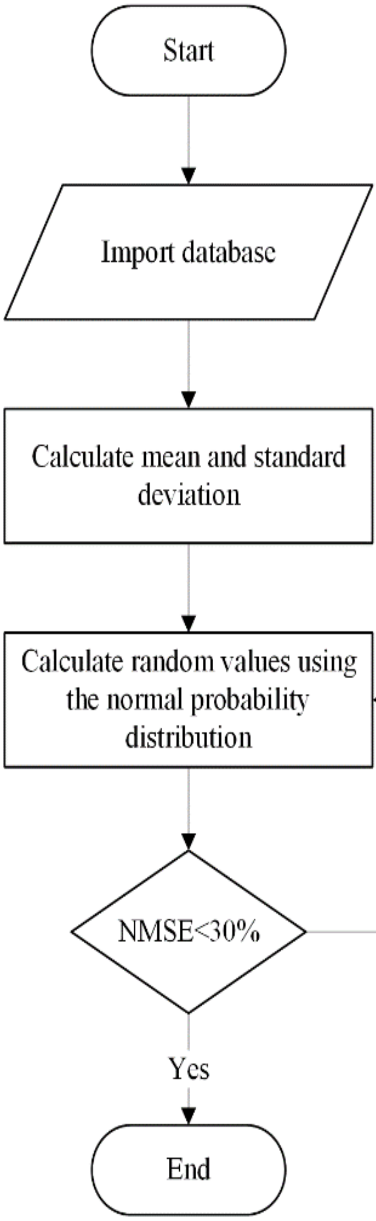

Fig. 4: Flowchart of the Monte Carlo method
Source: created by the authors

3.2.2 ARMA Model

In Figure 5 the development of the ARMA model is shown, where it starts by importing historical data from the time series, these data were subjected to the Augmented Dickey-Fuller (ADF) test to check their seasonality, if the data are not stationary they must be subjected to differentiation or logarithmic transformation for the predictions to be stable and not depend on time, once the data are stationary they must be adjusted to an ARMA model using the Akaike Corrected Information Criterion (AIC), where the minimum value of the AIC is identified, which allows selecting the order of p and q associated with the best fitting model, [16].

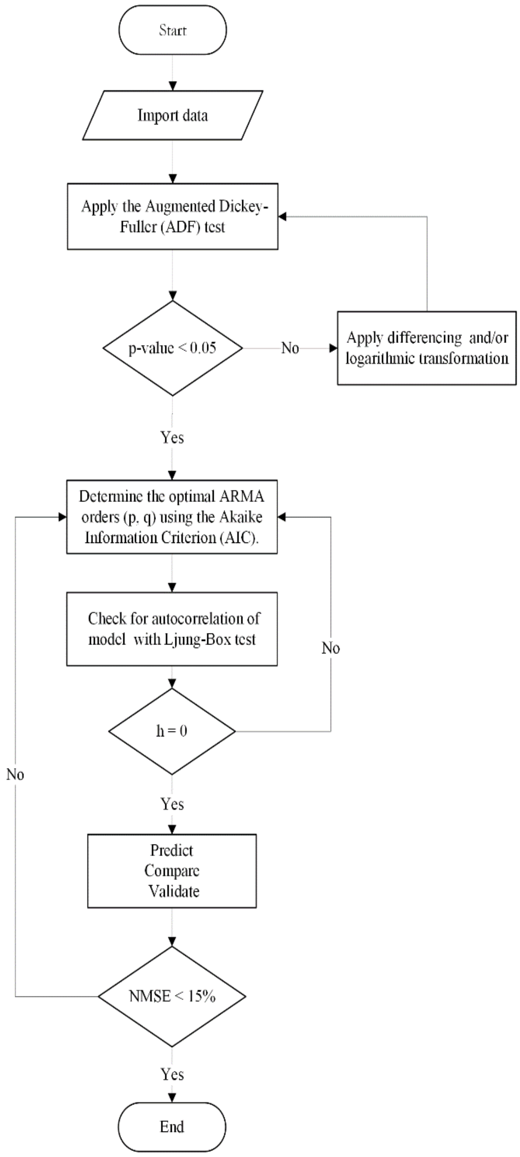

Fig. 5: Flowchart of the ARMA model
Source: created by the authors

The data were then subjected to a Ljung-Box test to verify whether the ARMA model residuals behaved like white noise; if so, the model was statistically adequate to represent the time series. Finally, a prediction is made based on the values of p and q and the performance of the model is evaluated by comparing the predicted values with the actual values of the electrical demand using the NMSE error metric, if the conditions based on the AIC criterion, the white noise test and the prediction error less than 15% are met, the combination of parameters p and q is selected as optimal. It should be noted that several tests must be carried out, since there may be a variety of parameters of p and q that may meet the aforementioned conditions, however, the model that best captures the temporal trends in demand and has a lower NMSE error will be selected when comparing the predictions with the actual data.

3.2.3 LSTM Network

Figure 6 shows the development of the neural network, where, in the first step, the database is imported corresponding to the electricity demand. These data are normalized to ensure a common measurement scale that facilitates improved data quality and consistency, [22].

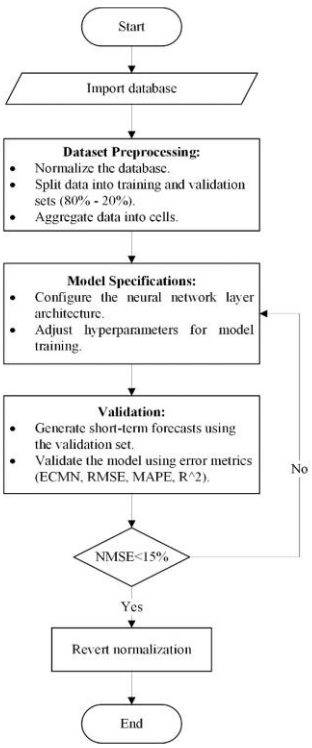

Fig. 6: Flowchart of the LSTM model
Source: created by the authors

For this purpose, equation (13) was used.

$$datanormalized = \frac{data - \mu}{s} \quad (13)$$

Once the data is normalized, the data set is divided into 80% for training and 20% for validation. The data intended for training must be organized into cells so that the network can more effectively identify and learn temporal trends in electricity demand. Once the dataset is ready for training, the network architecture and its hyperparameters must be defined. These hyperparameters were determined through trial and error, as each dataset and prediction problem has unique characteristics. Therefore, there is no single combination of hyperparameters that guarantees the best performance in all cases. Values such as the optimizer, training epochs, learning time, and other parameters are defined depending on the model's performance.

Finally, a short-term prediction is performed to evaluate the model's predictive capability. To do this, a comparison is made between the prediction and the actual data, using an NMSE error of less than 15% as a stopping criterion. Once the most suitable model has been identified, the initially applied normalization must be reversed to obtain the prediction values on their original scale.

3.3 Validation of Prediction Models

This section presents the main metrics used to evaluate the performance of predictive models applied to electricity demand. Relative and absolute error indicators are included.

3.3.1 Normalized Mean Squared Error (NMSE)

It is a dimensionless metric variant of the MSE, which is normalized by dividing it by the variance of the real values, giving a metric that expresses the error in relative terms, which facilitates the comparison between models and allows for the identification of those that have a more stable and consistent performance in different data sets, [23]. As shown in equation (14).

$$NMSE = \frac{1}{\sigma^2 N}\sum_{i=1}^{N}(x_i - y_i)^2 \quad (9)$$

where x_i and y_i represent the actual values and predicted values of electrical demand, respectively.

3.3.2 Root Mean Squared Error (RMSE)

It is a common metric used to compare the predictive performance of various models, where a low RMSE value means better predictive ability in terms of its absolute deviation, [24]. This is expressed in the original units of the data, providing an intuitive measure of predictive accuracy. Defined in equation (15).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2} \quad (10)$$

3.3.3 Mean Absolute Porcentaje Error (MAPE)

It is a relative metric that quantifies the prediction error as an average percentage with respect to the actual values, ideal for evaluating performance in contexts with variable magnitudes, [24]. Represented by equation (16).

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - y_i}{x_i}\right| \times 100 \quad (11)$$

3.3.4 Coefficient of determination - R^2

This metric is often used to estimate model performance. It represents the fraction of predicted values that are closest to the actual data line, [25]. Unlike the other metrics presented, whose ideal values tend toward zero, coefficient of determination values close to one (100%) indicate a higher level of efficiency in the model. This is represented in equation (17).

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(x_i - y_i)^2}{\sum_{i=1}^{N}(x_i - \bar{x})^2} \quad (12)$$

4 Results and Discussion

To verify the accuracy of each model, an individual prediction was made for each method, taking Tuesday, March 23, 2023, as a reference. Each model will be evaluated based on the validation parameters presented in Section III. In addition, a regression graph is presented that compares the actual values and the predicted values, allowing to visualize the degree of fit of each model and demonstrating its predictive capacity. Next, different evaluation horizons with the applied methodologies are presented.

4.1 Monte Carlo Method

The statistical parameters used as input for the Monte Carlo method were: mean=1.5837, variance=0.1164, standard deviation=0.3412, skewness=0.2804, kurtosis=2.4227, max value=2.6752, and min value=0.6988. The skewness and kurtosis values indicate that the data distribution approximates a normal probability distribution, making it suitable for generating synthetic series using the Monte Carlo method.

4.2 ARMA Model

The stationarity of the electrical demand series was evaluated using the ADF test, where the results (p=1, h=1, and t=-114664, which is more negative than the critical value of -10416) confirmed that the series is stationary and suitable for the ARMA model.

The Akaike Information (AIC) was then used to determine the optimal parameters, selecting the model with p=48 and 1=9, which achieved the lowest AIC value (-5.5933), indicating the best fit. Finally, the Ljung-Box test was applied to verify the independence of the residuals, obtaining a p-value of 0.3775 (>0.05), which confirms that the residuals behave as white noise and that the model adequately captures the temporal structure of the data without autocorrelation.

4.3 LSTM Network

The LSTM network was designed with three layers of 128 hidden units, using dropout rates of 0.4 and 0.25 to prevent overfitting and trained with the Adam optimizer at a learning rate of 0.0009 for 200 epochs. Based on the computational complexity analysis O(nd^2) and memory complexity O(nd) [26], with an input sequence of 720 steps and 128 hidden units, the model requires approximately 1.18x10^7 operations per sequence and 9.22x10^4 memory units, indicating a high computational cost.

4.4 Comparison of Prediction Methods

To evaluate the models' predictive capability, a comparison was conducted between the three prediction methods, considering prediction horizons of 24, 48, and 72 hours. Sunday, April 2, 2023, was selected as the starting point, given that it behaves differently from weekdays. The graphs are generally composed of a vertical line representing the start of the prediction (purple line), actual electricity demand (dashed black line), the synthetic series generated by the Monte Carlo method (yellow line), the ARMA model prediction (blue line), and the LSTM network prediction (orange line).

Figure 7 presents the 24-hour Sunday forecast. It shows that the LSTM network's prediction is almost identical to actual demand, especially during the midday peak and afternoon decline, while the ARMA model and Monte Carlo method are noticeably more erratic. It is clearly observed how they deviate from the actual curve between 6:00 and 18:00, forming peaks and valleys where none should exist, which visually confirms their difficulty in tracking actual demand on weekends.

Figure 8 shows the 48-hour forecast covering Sunday and Monday. A change in the ARMA model's behavior can be seen on the second day, as its prediction is closer to the actual demand trend. The LSTM network, meanwhile, continues to adjust accurately over the two days, although slight differences can be seen in the morning peaks and the afternoon peaks, where the prediction is slightly different from actual demand.

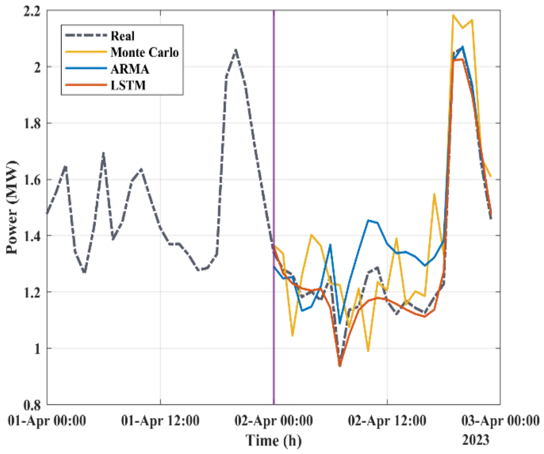

Fig. 7: 24-Hour Electric Load Demand Forecasting Using Monte Carlo, ARMA, and LSTM Methods vs. Real Data

Source: created by the authors

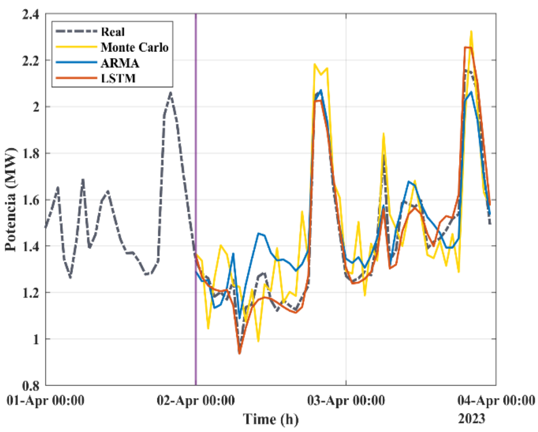

Fig. 8: 48-Hour Electric Load Demand Forecasting Using Monte Carlo, ARMA, and LSTM Methods vs. Real Data

Source: created by the authors

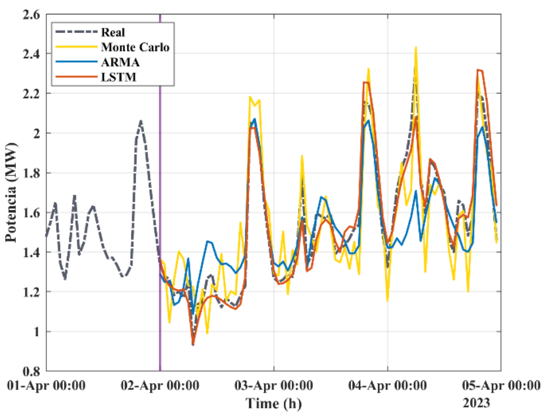

Fig. 9: 72-Hour Electric Load Demand Forecasting Using Monte Carlo, ARMA, and LSTM Methods vs. Actual Data

Source: created by the authors

Figure 9 shows the 72-hour forecast covering Sunday, Monday, and Tuesday. The main observation is the deterioration of the ARMA model on the third day, where it again shows pronounced deviations from the actual curve, while the Monte Carlo method remains variable across the actual electricity demand, as it does not deviate abruptly, but neither does it achieve accuracy. Likewise, the LSTM network demonstrates robustness in its prediction, since no matter how long the evaluation horizon is extended, the prediction gets closer to the actual data, with the only visible and consistent deviations occurring during the consumption peaks of each day.

Finally, Table 2 shows the prediction performance results of the Monte Carlo Method, the ARMA Model, and the LSTM Network for the three evaluation horizons (24, 48, and 72 hours). Comparing the performance of the three models across different time periods, it is observed that the LSTM is clearly the most reliable. It shows lower errors, and the coefficient of determination (R^2) value is above 93% in all cases. This means that it adapts very well to changes in electricity demand over any time horizon. Regarding the ARMA model, although it improves its prediction for 48 hours, with an NMSE that is reduced by 4.5%, this improvement is not sustainable over time, since for 72 hours its errors increase significantly, with an NMSE that reaches 26.02%, demonstrating that it has difficulty developing a long-term prediction. Finally, the Monte Carlo method shows an oscillatory evolution. Its NMSE and MAPE errors improve over time, but its RMSE error remains high, around 146 kW.

Table 2. Electrical Load Demand Forecasting Results

Time	Model	NMSE (%)	RMSE (kW)	MAPE (%)	R^2 (%)
24 hours	Monte Carlo	29.81	162.91	10.24	68.89
	ARMA	17.92	126.31	8.78	81.29
	LSTM	2.71	49.18	2.96	97.16
48 hours	Monte Carlo	23.96	146.99	8.7	75.53
	ARMA	13.41	110.00	6.85	86.29
	LSTM	4.88	66.36	3.41	95.013
72 hours	Monte Carlo	19.98	145.12	7.91	79.73
	ARMA	26.01	165.60	7.90	73.61
	LSTM	6.17	80.68	3.81	93.73

Source: created by the authors.

5 Conclusion

In this research, a comparative study was carried out between three electricity demand prediction models:

Monte Carlo Method, ARMA Model and LSTM Network, based on the hourly records of the San Rafael substation, in Cotopaxi, Ecuador, corresponding to the period 2020–2023. For this purpose, the error metrics NMSE, RMSE, MAPE and R^2 were used, with the purpose of evaluating the predictive capacity of the models in three evaluation horizons (24, 48 and 72 hours). In this way, it was possible to determine the performance of each model, in addition to its limitations and strengths depending on the time period analyzed, to establish which is most appropriate to optimize the planning and operation of the Ecuadorian electrical system.

The results showed that the Monte Carlo method is capable of generating synthetic series with a trend, although it presents fluctuations in electricity demand during most of the prediction period, which limits its accuracy. However, as the prediction horizon increases, its performance remains stable, with an NMSE error close to 20%. The ARMA model showed acceptable performance in the 48-hour prediction, with an NMSE value below 15%, but its predictive capacity deteriorated significantly as the prediction horizon increased due to the linear nature of the model, which limits its ability to capture non-linear patterns in the long term. In contrast, the LSTM network was consolidated as the most robust model, maintaining NMSE and MAPE values below 10%, along with a coefficient of determination greater than 93% in all case studies, confirming its ability to model complex patterns of long-term electricity demand.

As future work, we propose incorporating exogenous variables such as temperature and seasonality into the LSTM network for analysis in electricity demand forecasting. We also propose analyzing socioeconomic factors to evaluate forecasting behavior in rural environments, and integrating advanced deep learning architectures such as Bi-LSTM, GRU, and hybrid models to compare different forecasting models. Finally, we encourage evaluating the entire national electricity demand and identifying its annual behavior to achieve a forecast that is more realistic and serves as a benchmark for the Ecuadorian Electricity Regulation and Control Agency (ARCONEL).

Acknowledgement:
The authors express their gratitude to the provincial electricity company Cotopaxi ELEPCO SA for providing the electricity demand data for the San Rafael substation for the period 2020-2023.

References:

[1] M. Bilgili, N. Arslan, A. Şekertekin, y A. Yaşar, «Application of long short-term memory (LSTM) neural network based on deeplearning for electricity energy consumption forecasting», Turk. J. Electr. Eng. Comput. Sci., vol. 30, n.o 1, pp. 140-157, ene. 2022, doi: 10.3906/elk-2011-14.

[2] *(text in Spanish) Electricity Master Plan – Ministry of Energy and Mines (*«Plan Maestro de Electricidad – Ministerio de Energia y Minas»), [Online] https://www.recursosyenergia.gob.ec/plan-maestro-de-electricidad/ (Accessed Date: January 10, 2026).

[3] J. Bedi y D. Toshniwal, «Deep learning framework to forecast electricity demand», Appl. Energy, vol. 238, pp. 1312-1326, mar. 2019, doi: 10.1016/j.apenergy.2019.01.113.

[4] J. F. Torres, F. Martínez-Álvarez, y A. Troncoso, «A deep LSTM network for the Spanish electricity consumption forecasting», Neural Comput. Appl., vol. 34, n.o 13, pp. 10533-10545, jul. 2022, doi: 10.1007/s00521-021-06773-2.

[5] H. Abbasimehr, M. Shabani, y M. Yousefi, «An optimized model using LSTM network for demand forecasting», Comput. Ind. Eng., vol. 143, p. 106435, May 2020, doi: 10.1016/j.cie.2020.106435.

[6] N. Jin et al., «Highly accurate energy consumption forecasting model based on parallel LSTM neural networks», Adv. Eng. Inform., vol. 51, p. 101442, ene. 2022, doi: 10.1016/j.aei.2021.101442.

[7] A. Manowska, «Using the LSTM Network to Forecast the Demand for Electricity in Poland», Appl. Sci., vol. 10, n.o 23, Art. n.o 23, ene. 2020, doi: 10.3390/app10238455.

[8] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, y Y. Zhang, «Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network», IEEE Trans. Smart Grid, vol. 10, n.o 1, pp. 841-851, ene. 2019, doi: 10.1109/TSG.2017.2753802.

[9] G. Memarzadeh y F. Keynia, «Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm», Electr. Power Syst. Res., vol. 192, p. 106995, mar. 2021, doi: 10.1016/j.epsr.2020.106995.

[10] S. Mahjoub, L. Chrifi-Alaoui, B. Marhic, y L. Delahoche, «Predicting Energy Consumption Using LSTM, Multi-Layer GRU and Drop-

GRU Neural Networks», Sensors, vol. 22, n.o 11, Art. n.o 11, ene. 2022, doi: 10.3390/s22114062.

[11] «Normal distribution», en Basic Statistics and Epidemiology, 5.a ed., London: CRC Press, 2022, pp. 29-31. doi: 10.1201/9781003148111-11.

[12] M. Turowski et al., «Generating synthetic energy time series: A review», Renew. Sustain. Energy Rev., vol. 206, p. 114842, dic. 2024, doi: 10.1016/j.rser.2024.114842.

[13] «Applications of the Univariate Normal Distribution», en The Multivariate Normal Distribution, WORLD SCIENTIFIC, 2021, pp. 81-124. doi: 10.1142/9789811235290_0005.

[14] P. O. Ugbehe, O. E. Diemuodeke, y D. O. Aikhuele, «Electricity demand forecasting methodologies and applications: a review», Sustain. Energy Res., vol. 12, n.o 1, p. 19, abr. 2025, doi: 10.1186/s40807-025-00149-z.

[15] M. Bilgili y E. Pinar, «Gross electricity consumption forecasting using LSTM and SARIMA approaches: A case study of Türkiye», Energy, vol. 284, p. 128575, dic. 2023, doi: 10.1016/j.energy.2023.128575.

[16] N. B. Hassine, R. Milocco, y P. Minet, «ARMA based popularity prediction for caching in Content Delivery Networks», en 2017 Wireless Days, Porto, Portugal: IEEE, mar. 2017, pp. 113-120. doi: 10.1109/WD.2017.7918125.

[17] J. Wang, Q. Zhou, y X. Zhang, «Wind power forecasting based on time series ARMA model», IOP Conf. Ser. Earth Environ. Sci., vol. 199, p. 022015, dic. 2018, doi: 10.1088/1755-1315/199/2/022015.

[18] I. A. Ibrahim y M. J. Hossain, «LSTM Neural Network Model for Ultra-short-term Distribution Zone Substation Peak Demand Prediction», en 2020 IEEE Power & Energy Society General Meeting (PESGM), ago. 2020, pp. 1-5. doi: 10.1109/PESGM41954.2020.9281973.

[19] A. R. Abdellah y A. Koucheryavy, «VANET Traffic Prediction Using LSTM with Deep Neural Network Learning», en Internet of Things, Smart Spaces, and Next Generation Networks and Systems, vol. 12525, O. Galinina, S. Andreev, S. Balandin, y Y. Koucheryavy, Eds., en Lecture Notes in Computer Science, vol. 12525. , Cham: Springer International Publishing, 2020, pp. 281-294. doi: 10.1007/978-3-030-65726-0_25.

[20] J. Q. Wang, Y. Du, y J. Wang, «LSTM based long-term energy consumption prediction with periodicity», Energy, vol. 197, p. 117197, abr. 2020, doi: 10.1016/j.energy.2020.117197.

[21] S. Muzaffar y A. Afshari, «Short-Term Load Forecasts Using LSTM Networks», Energy Procedia, vol. 158, pp. 2922-2927, feb. 2019, doi: 10.1016/j.egypro.2019.01.952.

[22] H. Fraihat, A. A. Almbaideen, A. Al-Odienat, B. Al-Naami, R. De Fazio, y P. Visconti, «Solar Radiation Forecasting by Pearson Correlation Using LSTM Neural Network and ANFIS Method: Application in the West-Central Jordan», Future Internet, vol. 14, n.o 3, p. 79, mar. 2022, doi: 10.3390/fi14030079.

[23] T. O. Hodson, T. M. Over, y S. S. Foks, «Mean Squared Error, Deconstructed», J. Adv. Model. Earth Syst., vol. 13, n.o 12, p. e2021MS002681, dic. 2021, doi: 10.1029/2021MS002681.

[24] R. Bareth, A. Yadav, S. Gupta, y M. Pazoki, «Daily average load demand forecasting using LSTM model based on historical load trends», IET Gener. Transm. Distrib., vol. 18, n.o 5, pp. 952-962, mar. 2024, doi: 10.1049/gtd2.13132.

[25] M. Despotovic, V. Nedic, D. Despotovic, y S. Cvetanovic, «Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation», Renew. Sustain. Energy Rev., vol. 56, pp. 246-260, abr. 2016, doi: 10.1016/j.rser.2015.11.058.

[26] M. Larrat y C. Sales, «Classification of Flying Drones Using Millimeter-Wave Radar: Comparative Analysis of Algorithms Under Noisy Conditions», Sensors, vol. 25, n.º 3, p. 721, ene. 2025, doi: 10.3390/s25030721.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

- Galo Jimenez: Investigation, Methodology, Validation, Software.
- Alex Rocha: Investigation, Methodology, Data curation, Software.
- Diego L. Jimenez: Writing – review & editing, Conceptualization.
- Roberto Salazar-Achig: Formal analysis, Project administration.
- Pablo Muñoz: Supervision, Visualization.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare.