Comparison of Conventional Techniques for House Electricity Consumption Forecasting

Sandra Pajares Centeno, Hugo Alatrista-Salas
Escuela De Posgrado Newman, Tacna, Peru

***Abstract*—Electricity consumption monitoring is the automated process of recording, processing, and analyzing electricity usage in real time to make informed decisions. This research aims to implement an artificial intelligence- and deep learning-based methodology to forecast monthly electricity consumption in Tacna, Peru, and generate decision-making indicators. To this end, we used electricity consumption records from Electrosur S.A., the company responsible for electricity distribution and marketing in the departments of Tacna and Moquegua, from February 2015 to December 2022 (a total of 95 months). We compared three artificial intelligence models in this context: (i) eXtreme Gradient Boosting (XGBoost), (ii) Light Gradient Boosting (LGBM), and (iii) Prophet. While all models effectively forecasted electricity consumption, the Prophet model demonstrated superior performance, achieving a mean absolute percentage error (MAPE) of 0.7% compared to actual consumption values. Additionally, the study discusses the potential of recurrent neural networks to further enhance predictive accuracy.**

Keywords*—Electricity consumption; forecasting; recurrent neural networks; deep learning

I. Introduction

According to a report by the United Nations Department of Economic and Social Affairs, cities account for up to 80% of global energy consumption and 75% of carbon emissions, despite covering only 3% of the Earth's surface. This problem is prevalent in various countries, including in Peru, which emitted over 91.4 million tons of greenhouse gases (GHGs) in 2018. Peru also has one of the highest per capita GHG emissions in South America, at 2.8 tons. The indiscriminate use of electricity has resulted in our cities having one of the highest carbon footprints in Latin America. Additionally, international organisations such as the UN and the OAS are prioritising issues related to carbon emissions and climate change mitigation.

An increasing number of companies and government agencies are using artificial intelligence to improve the decision-making process [1]. In this sense, extracting information from data representing complex dynamics has been used to predict future events and make informed decisions in several areas, as discussed in [2], [3], [4], [5].

Electricity distribution and marketing companies, in particular, have processes that require continuous monitoring because the results of these processes directly impact the environment [6]. Therefore, it is essential to forecast energy consumption in the short, medium, and long term for environmental impact studies and capacity expansion. Power or energy consumption monitoring involves the automated recording, processing, and analysis of power network data in real time. The purpose of monitoring is to enable informed decision-making. Several algorithms have been implemented to predict electricity consumption [7], [8], [9], [10].

Energy consumption analysis has been applied to various fields, including the analysis of traffic conditions [11], the energy consumption of IoT devices [12], [13], the estimation of electric vehicle consumption [14], [15], and the use of aerial imagery to predict energy consumption [16].

This paper focuses on forecasting time series data. Apart from the classification proposed in [17], machine learning and deep learning techniques have become very relevant in recent years for various prediction tasks [18], [19]. Many works have been proposed for time series prediction tasks, particularly for energy consumption estimation For example, the authors in [20] used deep neural networks, specifically long short-term memory (LSTM) neural networks. In their paper, the authors compare two ways of representing data: classical LSTMs with time series data and sequences using the sequence-to-sequence (S2S) method. Similarly, LSTM models have improved thanks to algorithms based on the bidirectional long-short-term memory (Bi-LSTM) model [21]. In this context, this paper compares three well-known techniques for predicting energy consumption and generating indicators to facilitate better energy management and decision-making in Tacna, Peru.

After the introduction, Section II summarizes the existing literature on the topic. Section III presents the definitions and conceptual framework. Section IV describes the methodology used in this study, and Section V details the dataset and results. The work concludes with the conclusions and suggestions for future work.

II. Related Works

Various authors have used machine learning and deep learning techniques in their literature to predict electricity usage in residential, commercial, and public buildings. For instance, [8] use algorithms such as Support Vector Machine (SVM) and decision trees to predict household-level electricity consumption in the US. Similarly, [10] use support vector regression to estimate electricity consumption in a Turkish city. In [22], the authors use the Random Forests algorithm to predict short-term energy consumption based on five datasets collected over a year from various buildings. Similar studies have been conducted on houses in various countries, including China [23], Malaysia [24], Brazil and Spain [25], France [26], Albania [27] and several countries simultaneously [28], [29].

Additionally, some authors have used a variety of techniques to estimate electrical energy consumption. For example, in [7], the authors proposed a new fuzzy wavelet neural network model to analyse hourly electricity consumption. The model was validated using a database of electricity consumption from the Greek island of Crete. Similarly, [8] propose a method based on neural network back propagation and the improved particle swarm optimization (IPSO) algorithm for predicting electricity consumption and measuring the impact of changes and imbalances in energy consumption caused by the pandemic.

Other papers focus on more advanced techniques, such as deep learning. For instance, [30] the authors use *Long Short-Term Memory* In this text, the authors discuss the use of LSTM neural networks to predict household electricity consumption and their potential to reduce carbon emissions. The authors also discuss the potential of these networks to reduce carbon emissions. Additionally, they reference a study by [31] that proposes a model for predicting electric vehicle energy consumption using transfer learning. The proposed method utilizes data from the target domain as well as data or a model from a similar domain. The authors used data on electric vehicle trips, including the battery charge state, departure and arrival times at rest areas, and GPS data similar to the predicted training data.

In [32], three energy consumption prediction algorithms were compared: Asynchronous Advantage Actor-Critic (A3C), Deep Deterministic Policy Gradient (DDPG), and Recurrent Deterministic Policy Gradient (RDPG). The algorithms were tested on data collected from an office building. In this context, [33] predict monthly consumption using three deep learning models: Deep Fully Connected, Convolutional Neural Network, and Long Short-Term Memory Neural Networks. The authors also employed pre-processing techniques, such as normalization, to homogenize the input domain. Finally, they applied the algorithms to over 9 million sample datasets. In [34] study, data from 100 civil public buildings in China was used to predict short and medium-term electrical energy consumption. The authors proposed an algorithm based on a deep neural network and compared it with other algorithms such as the back-propagation algorithm, the Elman neural network, and the fuzzy neural network. In another study, [35] proposed integrating two deep learning models, including a long short-term memory (LSTM) model. They claim that combining LSTM and reinforcement learning agents (reinforced learning) is more efficient for predicting energy consumption at the University of Wollongong in Australia.

Additionally, several efforts have been made to propose efficient algorithms in terms of both prediction quality and ease of interpretation, particularly when the results are used to make informed decisions. For example, in [36], the authors use three interpretable models based on long short-term memory (LSTM) with self-attention to predict electrical energy consumption in a dataset related to an office building. Similarly, in [37], the authors use the SHAP tool to analyze the electrical energy consumption of an electric arc furnace used to produce stainless steel after an artificial neural network (ANN) makes a prediction. Through this study, the authors aim to provide practical examples of how SHAP can be used to analyse the contributions of each input variable to the consumption predictions generated by black box machine learning models when the data are associated with steel processes.

Furthermore, the authors of [38] propose a four-component methodology to predict and explain electrical energy consumption over approximately two million minutes collected from 2006 to 2010. These components include a power encoder, an auxiliary encoder for a two-dimensional, interpretable latent space, a deep-learning-based prediction model, and an explainer that identifies the most significant input features for predicting future demand. Finally, in [39], the authors focus on improving the efficiency of existing algorithms - implemented in the Darts library - for predicting the daily, weekly and monthly electricity consumption of 43 customers using the Neural Basis Expansion Analysis for Interpretable Time Series (N-BEATS) technique.

This research underscores the practical importance of adopting automated forecasting techniques for predicting electricity consumption, especially in contexts where current practices still depend on manual projections and basic trend extrapolations. Unlike complex models such as ARIMA, recurrent neural networks (RNNs), or Transformer-based architectures, which often require extensive parameter tuning and substantial computational resources, our findings demonstrate that simpler, well-established models like Prophet, XGBoost, and LGBM can deliver highly accurate results when applied to robust, well-structured datasets. This suggests that advanced forecasting performance does not always require sophisticated models, especially when the data is consistent in quality and time-based patterns. Our objective is not to conduct an extensive comparison of forecasting techniques, propose novel methods, or introduce new evaluation metrics. Instead, we aim to present a practical, evidence-based report reviewing existing algorithms and demonstrating the effectiveness of three simple models with promising results[1].

III. Basic Definitions

A. Electric Energy Consumption: Electrosur Tacna, Peru

The electricity balance illustrates the flow of electricity from generation to final consumption. Electricity travels through transmission and distribution networks. In 2022, Peru generated $59,713$ GWh of electricity and imported 32 GWh from Ecuador. The electrical energy balance, including production, consumption, and losses, is listed below:

- Of the total energy produced in Peru, $57,814$ GWh (97%) came from companies in the electricity market, while $1,898$ GWh (3%) came from industrial companies that produce electricity for their own use.
- Of the total electrical energy available, which includes national production plus imports, 1.4% is used for auxiliary services in power stations. Meanwhile, 11.1% is lost in transmission and distribution, leaving 87.4% to reach the final consumer.
- Of the energy produced by enterprises in the electricity market, 1.2% was used for auxiliary services in power stations, 11.3% was lost and 87.5% reached final consumers.
- Of the energy produced by industrial enterprises for their own use, 7.2% was used for auxiliary services

[1] The code and data are available at https://github.com/huvaso/electricityForecasting

of their plants, 5.78% was lost and 87.7% was used for their operations.

- Of the $48,092$ GWh of electrical energy sold by generation and distribution companies to final consumers, 38% ($18,410$ GWh) was sold on the regulated market, while 62% ($29,682$ GWh) was sold on the free market.

Electrosur S.A.'s main activity is distributing and marketing electrical energy in its concession area, which encompasses the Tacna and Moquegua departments. The company purchases the energy it distributes and markets from nine generation companies and the National Interconnected System of Peru.

In 2022, energy sales amounted to $399,786$ MWh, an increase of 2.3% compared to 2021. This increase is due to the progressive and sustained rise in consumption by both regulated and free customers, resulting from the economic reactivation of commercial activities in Tacna, Moquegua, and Ilo. ENEL Generación Perú was the company's main energy supplier in 2022.

In Peru, the forecast for electrical energy purchases (demand) is currently based on historical data projections, considering new user demand and population growth rates. The electricity distribution company's energy efficiency can be estimated by dividing its energy sales and purchases into periods. Therefore, its energy efficiency is estimated to be 89.97% in the 2022 financial year. One of the main activities undertaken in 2022 to reduce technical losses was the installation of two reactive compensation banks, one in Tacna with a total capacity of 300 kWh and one in Ilo with a total capacity of 300 kWh, for a total capacity of 600 kWh. This reduced technical energy losses to $36,500$ kWh per year.

Another important activity in reducing technical losses has been installing LED lights. A total of $14,084$ units were installed in Tacna ($7,452$), Moquegua ($2,863$), and Ilo ($3,769$). Due to the replacement of LED lights and the program to control losses, change drivers, recover energy, prevent illegal connections, and provide exceptional services, Electrosur S.A. is now among the electricity distribution companies in Peru with the lowest losses, below the average of 87.46%. Additionally, the loss of electrical energy was less than 10%, corresponding to $44,536$ MWh, which equates to approximately 15 million soles in economic terms.

Therefore, using historical data from a specific period and machine learning algorithms will enable more accurate purchase forecasts for annual planning and more confident economic and operational decisions. In terms of energy efficiency, correctly forecasting electricity demand will enable households to reduce their electricity bills.

B. Conventional Classifiers

1) XGBoost (eXtreme Gradient Boosting): This is an open-source library for performing machine learning tasks, especially regression and classification problems. It is based on the boosting algorithm, which combines several weaker models to form a stronger one. XGBoost uses decision trees and employs advanced techniques such as regularisation and gradient optimisation to improve model accuracy and performance. It is also widely used in data science competitions due to its effectiveness.

2) LGBM (LightGBM): This is a decision tree-based machine learning library designed to be efficient and scalable. Unlike XGBoost, LightGBM uses a leaf-wise rather than level-wise growth approach, which significantly improves training speed. It is designed to handle large datasets and high dimensionality.

3) Prophet: This is an open-source library developed by Facebook for time series forecasting. It is designed to handle time series data that may have trends, seasonality, holidays and also allows the inclusion of special events. Prophet can handle missing data and outliers, and it is suitable for short- and medium-term time series forecasting.

C. Evaluation Metrics

After training the model, it is essential to evaluate its performance on unseen data using evaluation metrics. The choice of metrics depends on the problem's specific context and preferences regarding error interpretation. The objective is to minimize these metrics' values to enhance the prediction model's accuracy.

1) MAE (Mean Absolute Error): measures the average size of absolute errors between predictions and actuals in a time series. It is useful for understanding the average size of errors regardless of their direction.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{1}$$

Where: n is the number of observations; y_i is the actual value at time i; and $\hat{y}_i$ is the prediction at time i.

2) MSE (Mean Squared Error): measures the average squared error between predictions and actual values in a time series. The squared term emphasizes larger errors, making it more sensitive to outliers.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2}$$

Where: n is the number of observations; y_i is the actual value at time i; and $\hat{y}_i$ is the prediction at time i.

3) MAPE (Mean Absolute Percentage Error): measures the average percentage of absolute errors with respect to the actual values. It is useful to evaluate the relative error in percentage terms.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{|y_i - \hat{y}_i|}{|y_i|}\right) \times 100 \tag{3}$$

Where: n is the number of observations; y_i is the actual value at time i; and $\hat{y}_i$ is the prediction at time i.

IV. Methodology

Predictive techniques based on temporal data are now fundamental to data science and artificial intelligence because they can model and predict patterns, trends, and behaviors in time series data.

In this context, and after analysing the work presented in the literature [40], the methodology described in Figure 1 was proposed in this paper. The dataset contains the periodic electricity consumption of each subscriber in kWh. The dataset was preprocessed before building a time series. This process can include various steps, such as filling in missing data, detecting outliers, and normalizing. This step depends on the dataset. Next, the time feature is preprocessed. Depending on the algorithm, the dataset must meet certain format specifications. Then, the time series is built according to the desired temporal granularity (daily, monthly, etc.). After building the time series, three forecasting algorithms were compared using the quality metrics described in the previous section. Finally, the predictions were visualized to better understand consumption behavior patterns.

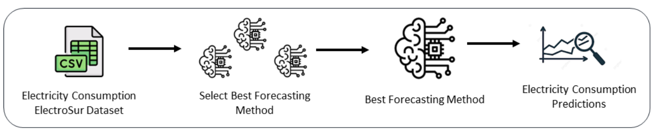


Fig. 1. Proposed methodology.

V. Experiments and Results

This section describes the available dataset. Additionally, the main findigs of our approach were detailed and discussed.

A. Dataset

The required data for the study was provided by Electrosur S.A., Tacna's electricity distribution company in Peru. The data, which corresponded to the electricity consumption data for the district, province, and department of Tacna, was obtained from the commercial system database. It is worth noting that the district's users are subject to a residential tax.

The research will examine data from February 2015 to December 2022. This includes the supply number and active energy consumption in kilowatt-hours (kWh) per user for each month and year within this period. The supply number is a code that identifies the service user. It is also known as the customer number that appears on the electricity consumption voucher. Active energy consumption (AEC) is measured in kWh and corresponds to a user's energy consumption as recorded by an electricity meter, also known as a wattmeter, which is calibrated in billing units. AEC is calculated by subtracting the electric meter reading from the previous month from the current month's reading, resulting in the billed consumption for the month.

It is important to note that there are $8,171$ users, each with 88 months of billing information. These records correspond to active energy consumption from February 2015 to December 2022, resulting in 719,048 total records of electrical energy consumption.

Table I presents the main statistical characteristics of the data by year.

As expected, the number of services increases each year due to the progressive growth of urban centers. However, contrary to urban growth, the average electricity consumption expenditure in Tacna is decreasing. Figure 2 shows a slight increase in 2019; however, the average consumption decreased during the period of the 2020-2021 (Covid-19) pandemic. Furthermore, the maximum annual electricity consumption value in Tacna increases in 2020 and 2021, during the period of the pandemic. Note that the quarantine period began in March 2020.

B. Results

First, any missing, null, or duplicate data in the dataset was addressed. This cleaning process is essential to ensure that the models are based on high-quality data. Next, a data transformation process was performed. This process includes converting data into appropriate formats, creating new features from existing ones, and rescaling variables to facilitate modeling. Since the number of houses has increased over the years, our particular dataset focuses on only those houses with records for each of the 95 months.

Additionally, to explore which model works best with our dataset, the average monthly consumption of all houses was calculated. This produced one time series representing the average consumption of houses in the Tacna districts over the 95 months (see Figure 3).

The training set is used to train the model, while the validation set is used to evaluate its performance (see figure 4). When working with time series, the data set must be divided according to chronological order, not randomly. In this sense, our data sets correspond to:

- Training set: Consumption corresponding to the dates from February 2012 to May 2021 (76 months).
- Validation set: Consumption corresponding to the period from June 2021 to December 2022 (19 months).

Identifying the features (independent variables) that will be used to predict the target variable (dependent variable) is essential for training the model. Then a machine learning algorithm is used to fit the model to the training data. The model learns the patterns and relationships in the data that can be used for predictions.

Moreover, three learning algorithms were chosen: eXtreme Gradient Boosting (XGBoost), Light Gradient Boosting (LGBM), and Prophet. The Mean Absolute Error, Mean Squared Error, and Mean Absolute Percentage Error were used to evaluate performance.

To train the model with XGBoost and LightGBM, new features were generated from the data: Year, Month and Quarter. An example of these new features is shown in Table II.

The Prophet algorithm generates features related to trends and seasonality. These features relate to different components of the model and their prediction intervals: (i) The trend component captures the general direction of the time series.

TABLE I. STATISTICAL FEATURES OF THE DATA SET BY YEAR

Features	2015	2016	2017	2018	2019	2020	2021	2022
Count	6004.00	6156.00	6346.00	6676.00	7437.00	7506.00	8013.00	8247.00
Mean	123.49	123.43	122.18	116.98	117.91	117.23	110.68	106.30
Std.Deviation	175.54	185.99	185.04	172.37	190.08	183.41	197.29	181.57
Minumum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25%	38.63	38.58	39.16	36.39	34.33	34.75	28.75	27.33
50%	91.13	89.00	89.33	84.83	82.41	84.75	77.83	73.33
75%	153.13	150.83	149.02	144.33	142.08	144.33	136.83	129.66
Maximum	4156.45	4312.58	4483.83	3706.41	4052.41	4327.09	9000.00	5075.00

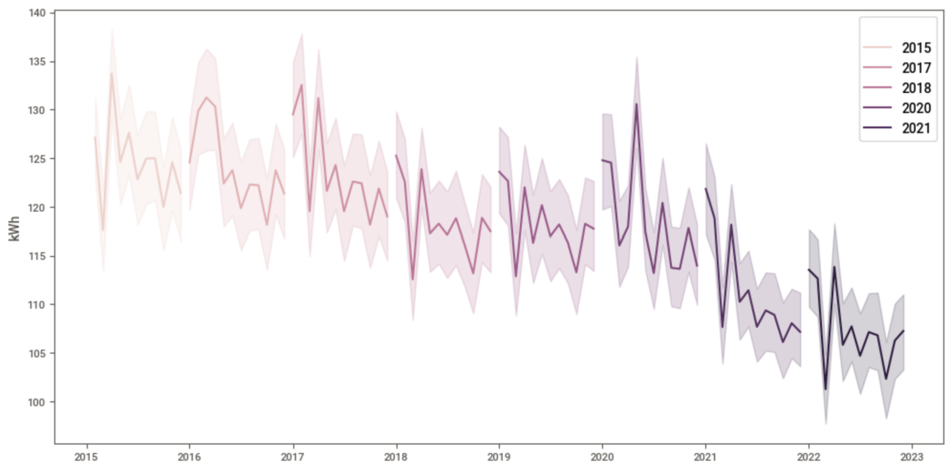


Fig. 2. Electricity consumption per year.

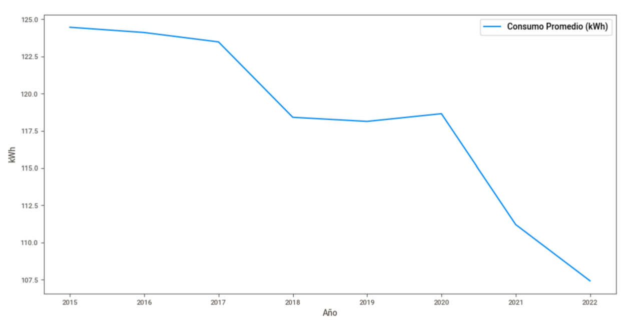


Fig. 3. Time series representing the mean consumption in Tacna.

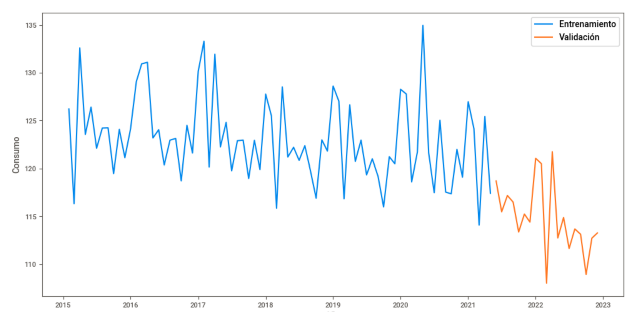


Fig. 4. Splitting the dataset.

It represents the long-term growth or decline trajectory in the data. (ii) Variables representing the additive components of the model, including seasonality and holiday effects. (iii) Variables representing the annual component of the model, which captures seasonal patterns over the years; and (iv) variables representing the multiplicative components of the model.

Table III shows the values obtained for each method according to the evaluated metrics. Based on these results, we can conclude that the Prophet model is the best. Furthermore, Figure 5 shows that the Prophet model's predictions are closer to the actual consumption values than those of the other methods.

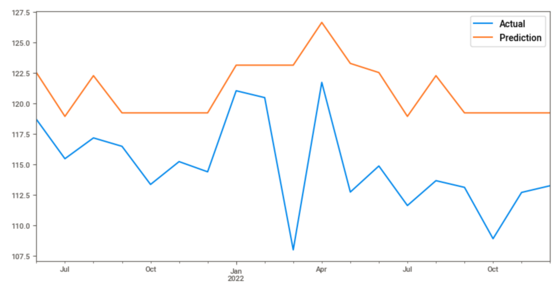
TABLE II. NEW ADDITIONAL FEATURES

Period	Consumptions	**Four-month Period**	**Month**	**Year**
2015-02-01	126.208909	1	2	2015
2015-03-01	116.323349	1	3	2015
2015-04-01	132.613204	2	4	2015
2015-05-01	123.545468	2	5	2015
2015-06-01	126.405208	2	6	2015

C. Discussion

The current methodology for forecasting electricity demand in Peru relies primarily on extrapolating historical data while

XGBoost



Light GBM

Prophet

Fig. 5. Forecasting of electricity consumption in validation dataset.

TABLE III. Metrics Evaluated for the Three Algorithms

Algorithm	MAE	MSE	MAPE
XGBoost	3.026	12.264	0.027
LGBM	6.196	48.351	0.055
Prophet	0.771	1.056	0.007

considering demographic growth and the addition of new users. While this traditional approach provides a baseline estimation, it often lacks the granularity and adaptability required for dynamic energy consumption environments. In this context, this study research proposes a data-driven methodology that uses machine learning and statistical forecasting models, namely eXtreme Gradient Boosting (XGBoost), Light Gradient Boosting (LGBM), and Prophet, to improve the accuracy of electricity consumption forecasts in Tacna, Peru.

The results indicate that the Prophet model outperforms the other two algorithms across all evaluated metrics (MAPE, MAE, and MSE), with a particularly low MAPE (see Table III), demonstrating its strong suitability for capturing seasonal and trend-based patterns in time series data. The proposed methodology benefits from several strengths. First, the dataset is extensive and consists of monthly active energy consumption records from $8,171$ users over an 8-year period, resulting in over 700,000 observations. This rich dataset allows for robust model training and evaluation. Second, the inclusion of real operational data from Electrosur S.A., a key electricity provider, ensures practical relevance and reflects actual consumption patterns in the region.

However, some limitations must be acknowledged. While the proposed models are effective in forecasting, they do not yet incorporate exogenous factors, such as weather patterns, economic activity, or special events, that could affect consumption behavior. Additionally, while Prophet shows excellent performance, its black-box nature may limit its usefulness to decision-makers who prefer transparent models. From a policy and operational perspective, improving forecasting accuracy could support Electrosur S.A.'s ongoing efforts to reduce technical losses, such as installing reactive compensation banks and LED lighting. Enhanced forecasting could also support targeted interventions for high-consumption users, energy-saving initiatives, and more efficient demand planning.

VI. Conclusion

This study presented a data-driven approach to forecasting monthly electricity consumption in Tacna, Peru. The study used nearly eight years of historical data from the Peruvian Electricity Administration Office. Implementing and comparing three machine learning models demonstrated the efficacy of AI-based forecasting techniques in the energy sector. Of the evaluated models, Prophet demonstrated the highest predictive accuracy, achieving a mean absolute percentage error (MAPE) of 0.7%, highlighting its robustness for time series forecasting in this domain.

Future work will focus on exploring deep learning approaches, particularly recurrent neural networks (RNNs) and transformers, to better model temporal consumption patterns, as discussed in [41], [42]. Comparative analyses will be conducted to assess the performance of these models against traditional ensemble methods such as XGBoost and LGBM. Finally, a real-time alert system to identify users with potentially high electricity consumption. This system will allow us to provide targeted energy-saving recommendations and support demand management efforts.

References

[1] M. Havey, *Essential business process modeling.* " O'Reilly Media, Inc.", 2005.

[2] J. Bascompte and R. V. Solé, "Rethinking complexity: modelling spatiotemporal dynamics in ecology," *Trends in Ecology & Evolution*, vol. 10, no. 9, pp. 361–366, 1995.

[3] A. McGovern, D. John Gagne, N. Troutman, R. A. Brown, J. Basara, and J. K. Williams, "Using spatiotemporal relational random forests to improve our understanding of severe weather processes," *Statistical Analysis and Data Mining: The ASA Data Science Journal*, vol. 4, no. 4, pp. 407–429, 2011.

[4] K. S. Oberoi and G. Del Mondo, "Graph-based pattern detection in spatio-temporal phenomena," in *16th Spatial analysis and geomatics conference (SAGEO 2021)*, 2021.

[5] K. V. Rao, A. Govardhan, and K. C. Rao, "Spatiotemporal data mining: Issues, tasks and applications," *International Journal of Computer Science and Engineering Survey*, vol. 3, no. 1, p. 39, 2012.

[6] J. W. Lee, "The contribution of foreign direct investment to clean energy use, carbon emissions and economic growth," *Energy policy*, vol. 55, pp. 483–489, 2013.

[7] M. Amina, V. Kodogiannis, I. Petrounias, and D. Tomtsis, "A hybrid intelligent approach for the prediction of electricity consumption," *International Journal of Electrical Power & Energy Systems*, vol. 43, no. 1, pp. 99–108, 2012.

[8] S. Nallathambi and K. Ramasamy, "Prediction of electricity consumption based on dt and rf: An application on usa country power consumption," in *2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE)*. IEEE, 2017, pp. 1–7.

[9] X. Li, Y. Wang, G. Ma, X. Chen, J. Fan, and B. Yang, "Prediction of electricity consumption during epidemic period based on improved particle swarm optimization algorithm," *Energy Reports*, vol. 8, pp. 437–446, 2022.

[10] K. Kavaklioglu, "Modeling and prediction of turkey's electricity consumption using support vector regression," *Applied Energy*, vol. 88, no. 1, pp. 368–375, 2011.

[11] P. Li, Y. Zhang, K. Zhang, and M. Jiang, "The effects of dynamic traffic conditions, route characteristics and environmental conditions on trip-based electricity consumption prediction of electric bus," *Energy*, vol. 218, p. 119437, 2021.

[12] G. Bedi, G. K. Venayagamoorthy, and R. Singh, "Development of an iot-driven building environment for prediction of electric energy consumption," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 4912–4921, 2020.

[13] A. Chaudhari, B. Rodrigues, and S. More, "Automated iot based system for home automation and prediction of electricity usage and comparative analysis of various electricity providers: Smartplug," in *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*. IEEE, 2016, pp. 390–392.

[14] C. Beckers, I. Besselink, J. Frints, and H. Nijmeijer, "Energy consumption prediction for electric city buses," in *Proceedings of the 13th ITS European Congress, Brainport, The Netherlands*, 2019, pp. 3–6.

[15] Z. Yi and P. H. Bauer, "Adaptive multiresolution energy consumption prediction for electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10 515–10 525, 2017.

[16] M. Rosenfelder, M. Wussow, G. Gust, R. Cremades, and D. Neumann, "Predicting residential electricity consumption using aerial and street view images," *Applied Energy*, vol. 301, p. 117407, 2021.

[17] M. Bourdeau, X. qiang Zhai, E. Nefzaoui, X. Guo, and P. Chatellier, "Modeling and forecasting building energy consumption: A review of data-driven techniques," *Sustainable Cities and Society*, vol. 48, p. 101533, 2019.

[18] A. G. Salman, B. Kanigoro, and Y. Heryadi, "Weather forecasting using deep learning techniques," in *2015 international conference on advanced computer science and information systems (ICACSIS)*. Ieee, 2015, pp. 281–285.

[19] Y. Lin, N. Mago, Y. Gao, Y. Li, Y.-Y. Chiang, C. Shahabi, and J. L. Ambite, "Exploiting spatiotemporal patterns for accurate air quality forecasting using deep learning," in *Proceedings of the 26th ACM SIGSPATIAL international conference on advances in geographic information systems*, 2018, pp. 359–368.

[20] D. L. Marino, K. Amarasinghe, and M. Manic, "Building energy load forecasting using deep neural networks," in *IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2016, pp. 7046–7051.

[21] J. Huang, C. Pang, W. Yang, X. Zeng, J. Zhang, and C. Huang, "A deep learning neural network for the residential energy consumption prediction," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 17, no. 4, pp. 575–582, 2022.

[22] A.-D. Pham, N.-T. Ngo, T. T. H. Truong, N.-T. Huynh, and N.-S. Truong, "Predicting energy consumption in multiple buildings using machine learning for improving energy efficiency and sustainability," *Journal of Cleaner Production*, vol. 260, p. 121082, 2020.

[23] Y. Liu, H. Chen, L. Zhang, X. Wu, and X.-j. Wang, "Energy consumption prediction and diagnosis of public buildings based on support vector machine learning: A case study in china," *Journal of Cleaner Production*, vol. 272, p. 122542, 2020.

[24] M. K. M. Shapi, N. A. Ramli, and L. J. Awalin, "Energy consumption prediction by using machine learning for smart building: Case study in malaysia," *Developments in the Built Environment*, vol. 5, p. 100037, 2021.

[25] T. C. Carneiro, P. A. Rocha, P. C. Carvalho, and L. M. Fernández-Ramírez, "Ridge regression ensemble of machine learning models applied to solar and wind forecasting in brazil and spain," *Applied Energy*, vol. 314, p. 118936, 2022.

[26] A. J. del Real, F. Dorado, and J. Durán, "Energy demand forecasting using deep learning: Applications for the french grid," *Energies*, vol. 13, no. 9, p. 2242, 2020.

[27] A. Qosja, E. G. Dhamo, D. Georges, L. Nikolla, and A. Cela, "Comparative analysis of neural network and radial basis functions approaches for electricity consumption forecasting in five different regions in albania," in *2024 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*. IEEE, 2024, pp. 1–6.

[28] D. Sen, K. M. Tunç, and M. E. Günay, "Forecasting electricity consumption of oecd countries: A global machine learning modeling approach," *Utilities Policy*, vol. 70, p. 101222, 2021.

[29] U. Şahin, S. Ballı, and Y. Chen, "Forecasting seasonal electricity generation in european countries under covid-19-induced lockdown using fractional grey prediction models and machine learning methods," *Applied Energy*, vol. 302, p. 117540, 2021.

[30] Z. Zheng, H. Chen, and X. Luo, "Spatial granularity analysis on electricity consumption prediction using lstm recurrent neural network," *Energy Procedia*, vol. 158, pp. 2713–2718, 2019.

[31] A. Fukushima, T. Yano, S. Imahara, H. Aisu, Y. Shimokawa, and Y. Shibata, "Prediction of energy consumption for new electric vehicle models by machine learning," *IET Intelligent Transport Systems*, vol. 12, no. 9, pp. 1174–1180, 2018.

[32] T. Liu, Z. Tan, C. Xu, H. Chen, and Z. Li, "Study on deep reinforcement learning techniques for building energy consumption forecasting," *Energy and Buildings*, vol. 208, p. 109675, 2020.

[33] R. F. Berriel, A. T. Lopes, A. Rodrigues, F. M. Varejao, and T. Oliveira-Santos, "Monthly energy consumption forecast: A deep learning approach," in *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2017, pp. 4283–4290.

[34] L. Lei, W. Chen, B. Wu, C. Chen, and W. Liu, "A building energy consumption prediction model based on rough set theory and deep learning algorithms," *Energy and Buildings*, vol. 240, p. 110886, 2021.

[35] X. Zhou, W. Lin, R. Kumar, P. Cui, and Z. Ma, "A data-driven strategy using long short term memory models and reinforcement learning to predict building electricity consumption," *Applied Energy*, vol. 306, p. 118078, 2022.

[36] Y. Gao and Y. Ruan, "Interpretable deep learning model for building energy consumption prediction based on attention mechanism," *Energy and Buildings*, vol. 252, p. 111379, 2021.

[37] L. S. Carlsson, P. B. Samuelsson, and P. G. Jönsson, "Interpretable machine learning—tools to interpret the predictions of a machine learning model predicting the electrical energy consumption of an electric arc furnace," *Steel research international*, vol. 91, no. 11, p. 2000053, 2020.

[38] J.-Y. Kim and S.-B. Cho, "Explainable prediction of electric energy demand using a deep autoencoder with interpretable latent space," *Expert Systems with Applications*, vol. 186, p. 115842, 2021.

[39] A. K. Shaikh, A. Nazir, I. Khan, and A. S. Shah, "Short term energy consumption forecasting using neural basis expansion analysis for interpretable time series," *Scientific reports*, vol. 12, no. 1, p. 22562, 2022.

[40] D. Mariano-Hernández, L. Hernández-Callejo, M. Solís, A. Zorita-Lamadrid, O. Duque-Pérez, L. Gonzalez-Morales, F. S. García, A. Jaramillo-Duque, A. Ospino-Castro, V. Alonso-Gómez *et al.*, "Analysis of the integration of drift detection methods in learning algorithms for electrical consumption forecasting in smart buildings," *Sustainability*, vol. 14, no. 10, p. 5857, 2022.

[41] R. Kumar, B. Shakila, and M. Prakash, "A novel hybrid deterministic temporal fusion transformer and dynamic recurrent neural network architecture for multi-horizon multi-variate electric vehicle energy consumption forecasting: mitigation pathway," *International Journal of Information Technology*, pp. 1–19, 2025.

[42] J. He, Y. Li, X. Xu, and D. Wu, "Energy consumption forecasting for oil and coal in china based on hybrid deep learning," *PloS one*, vol. 20, no. 1, p. e0313856, 2025.