A new interval prediction methodology for short-term electric load forecasting based on pattern recognition

Xavier Serrano-Guerrero[1], Guillermo Escrivá-Escrivá[2], Marco Briceño-León[2,3], Jean-Michel Clairand[3,*]

[1] *Grupo de Investigación en Energías, Universidad Politécnica Salesiana, Cuenca 010103, Ecuador*
[2] *Institute for Energy Engineering, Universitat Politècnica de València, edificio 8E, escalera F, 2ª planta, 46022 Valencia, Spain*
[3] *Facultad de Ingeniería y Ciencias Aplicadas, Universidad de las Américas, Quito 170122, Ecuador*
**Corresponding author. E-mail address: jean.clairand@udla.edu.ec*

Abstract
Demand prediction is an essential tool for electricity management, and is fundamental to the corresponding decision-making. Short-term forecasting has crucial utility for generating dispatching commands, managing the spot market, and detecting anomalies. The techniques associated with machine learning are those currently preferred by researchers for making predictions. However, there are concerns regarding limiting the uncertainty of the obtained results. In this work, a statistical methodology with a simple implementation is presented for obtaining a prediction interval with a time horizon of seven days (15-min time steps), thereby limiting the uncertainty. The methodology is based on pattern recognition and inferential statistics. The predictions made differ from those from a classical approach which predicts point values by trying to minimize the error. In this study, 96 intervals of absorbed active power are predicted for each day, one for every 15 min, along with a previously defined probability associated with the real values being within each obtained interval. To validate the effectiveness of the predictions, the results are compared with those from techniques with the best recent results, such as artificial neural network (ANN) long short-term memory (LSTM) models. A case study in Ecuador is analyzed, resulting in a prediction interval coverage probability (PICP) of 81.1% and prediction interval normalized average width (PINAW) of 10.13%, with a confidence interval of 80%.

Keywords: electricity demand; pattern recognition; prediction intervals; short-term forecasting

1. Introduction
Demand prediction is an essential tool for electricity management, and is critical to the corresponding decision-making. Forecasts based on time horizons can be divided into short-, medium-, and long-term forecasts. Short-term electricity forecasts can range from one hour to a week, and are vital for generation dispatching, managing spot markets, and detecting anomalies. A medium-term forecast comprises a period between one week and one year and is used for planning, negotiating contracts, and operating electricity systems. Finally, in long-term forecasts, periods of more than one year are used. These find their most significant applications in planning distribution and transmission networks and new generation plants [1].

Interest in predicting and forecasting the electricity demand has increased considerably, especially in the last decade. In particular, forecasting the demand for electricity for new electricity grids is not a trivial task, as it depends on various factors, including climatic, social, economic, and labor factors.

Various tools have been developed and used to forecast electrical power demands. The first work on electricity prediction dates from 1955 [2], and was based on interpolation and extrapolation techniques from historical data. In 1978, the prediction approach was improved by using multivariate regression methods derived from economic theories [3]. In 1980 [4], the use of discriminative Bayesian models was introduced to improve electricity prediction, based on using the mean square error as a metric for evaluating the quality of a method. In 1988 [5], the study of non-stochastic adaptive methods was incorporated into the prediction of electricity demand with the least mean square algorithm.

In the 1990s, machine learning began to be used to predict electricity demand. Since 1996, artificial neural networks (ANNs) have been used to predict electricity demand [6], and have provided good results. Over time, the application of ANNs has evolved, and new proposals have been made [7]–[9]. In addition to ANNs, other tools based on machine learning have been used for predictions of consumption, such as support vector machines [10] and decision trees [11].

Advancement in ANN algorithms, big data, and in the processing capacities of computers have given rise to what is known today as "deep learning." It's very recent use in prediction and forecasting has presented good results [12]–[13]. The results reported in the literature highlight the use of long short-term memory (LSTM) neural networks, owing to their good performance [14].

Table 1 summarizes the most widely used electricity demand forecasting methods. Deterministic, non-deterministic, and hybrid tools can be distinguished as follows. Deterministic tools are based on using mathematical equations to model physical phenomena. These techniques are also called white or transparent boxes, as the relationships between the electricity demand and other variables are known. Non-deterministic tools can be divided into two types: those that use statistical methodologies, and those that use machine learning. Machine learning is a derivation of artificial intelligence that allows machines to learn on their own. These methodologies are called black boxes, as the relationships between the predicted electricity consumption and other variables are not known. Finally, hybrid or "gray box" prediction tools combine the features of white box and black box tools.

Table 1. Electricity demand forecasting methods

Prediction tools	Methodologies	Advantages	Disadvantages
Deterministic	Mathematical-physical models (engineering)	They model any physical or energy system.	They are complex, requiring detailed knowledge of physical or energy systems. High errors when the models do not conform to reality. High computational cost.
Non-deterministic - Statistics	Linear regression models	Simple to apply. The prediction model is described with a simple equation.	Data for one or more variables that have a significant correlation with the dependent variable are required. Limited results when the independent variables have non-linear relationships with the output variable. Difficulty in managing multicollinearity.
	Holt and Winters	Simple application models. They can be modified to suit new conditions.	Poor prediction owing to the multiple seasonality of the electricity demand data.
	Box Jenkins (Auto-regressive moving average (ARMA), auto-regressive integrated moving	They admit non-stationary data series [15].	Efficacy decreases when the time series is dominated by the seasonal component [16]. Complex application.

	average (ARIMA)		
	Bayesian Models	Simple to apply. Good predictions when the data fit a normal probability distribution.	Data do not always fit a normal distribution.
Non-deterministic-Machine learning	Neural networks	They work for linear or nonlinear problems. Collinearity between variables in the training data is not a problem. Anomalies in the data do not significantly affect the prediction results.	They use many indeterminate parameters adjusted without determined rules [17]. For training they require a complete and extensive database of electricity consumption and its related variables, consistent over time. It is difficult to limit the uncertainty of the results obtained and to interpret them physically, especially in stochastic variables.
	Vector Support Machine	It is capable of working with heterogeneous and incomplete databases. Good prediction results are obtained as they use an optimization algorithm.	There are few clues to selecting the best kernel function, its corresponding parameters, and two additional constants. It is difficult to quantify the uncertainty of the results obtained and to interpret them physically.
	Decision Tress	Simple models of easy interpretation. It is not affected by outliers.	Limited prediction for continuous variables. The reliability of the results depends on the accuracy of the training values. A small change in the input can cause large changes in the tree.
	Deep Learning	High adaptability to the data.	High computer costs. Establishing the network´s structure, i.e. number of neurons, layers, optimization algorithm, etcetera, is not a simple task and can require considerable time. It is difficult to limit the uncertainty of the obtained results and to interpret them physically.
	Genetic algorithms	They can solve nonlinear problems. High adaptability to data.	High computational cost. The results are not always optimal as an adequate adjustment of the algorithm depends on the number of data in the population, iterations, properties of the chromosomes and a correct definition of the fitness function, in addition the processing time can be high.
Hybrids	Combine deterministic and non-deterministic models	A good criterion in the selection of the models improves the predictions. They allow to maintain physical interpretations without the need for a very detailed and complex mathematical model.	An expert is required to select the parameters of the non-deterministic models. Implementation can be complex.

An important consideration regarding electricity demand prediction is that although many variables can be associated with electricity consumption, it is impractical to use all of them in a real-time surveillance and monitoring system [18]. A multi-variable prediction system increases the computational requirement and complexity, and simultaneously introduces greater uncertainty, by requiring the predictions of other variables. It has been shown that, for short-term prediction models, a univariate model is sufficient, as in these cases, the external variables may have little influence [16]. For this reason, the contributions in this study are mostly based on univariate models, combined with data segmentation criteria.

The literature review shows that machine learning is one of the methodologies preferred by researchers today. However, for this type of methodology, certain weaknesses have been identified, as follows.

- Machine learning approaches have difficulties in limiting the uncertainty of the obtained results.
- Establishing the structure of the network is not an easy task; in addition, these methodologies use many adjusted parameters, without specific rules.
- The selection of the training data requires considerable effort.
- Artificial networks are trained in specific conditions. The models are adjusted to a certain facility; therefore, the same model cannot be used for other facilities.

Recently, a few studies have investigated probabilistic load forecasting, and its use in quantifying uncertainty through prediction intervals (PIs). A PI is a prediction of an interval in which a real value in the future will drop, with a certain probability called the confidence level. PIs are a current topic in energy management. The corresponding works use different techniques; for instance, the authors of [19]–[21] used Gaussian processes, autoregressive integrated moving average models, log-normal processes, and kernel encoders to produce probabilistic forecasts for electricity consumption. However, Gaussian processes require considerable time to learn the hyperparameters of the covariance functions. Moreover, they eventually obtain negative PIs, which is inconsistent with energy consumption. In [22], a combination of a kernel-based support vector quantile regression model and Copula theory was proposed for forecasting a short-term power load. Then, in [22], the same authors used a Yeo-Johnson transformation quantile regression and Gaussian kernel function for short-term power load forecasting. Similarly, Zhang et al. [24] studied load forecasting using quantile regression forest, quantile determination, and gradient boosting machine approaches. He et al. [25] considered a variational mode decomposition-based quantile regression forest method and Bayesian optimization algorithm for short-term load forecasting. In these works, the results were evaluated based on the PI coverage probability (PICP) and the PI normalized average width (PINAW). Although these works demonstrate satisfactory performance, they use complex methods since electricity demand prediction has seasonal components and responds to a stochastic process. This may result in high computational resources. Therefore, it is necessary to explore computationally efficient methods to forecast electricity demand while limiting the uncertainty.

The innovative contributions of this paper are highlighted as follows:

(1) A new interval prediction methodology of electricity demand based on pattern recognition (IPMPR) is proposed, which can be used for any consumer type. The predictions made differ from the classical approach, which predicts point values while minimizing the forecasting error.

(2) This methodology limits forecast uncertainty through robust statistical analysis considering the data's seasonality using low computational complexity.

(3) To demonstrate the effectiveness of the methodology, the results are compared with a new ANN LSTM structure. The obtained results show that the prediction's sharpness is better than LSTM networks and other complex methods, especially on holidays.

This article is organized as follows. Section 2 presents the new interval prediction methodology. Section 3 presents the evaluation metrics used to quantify the goodness of the results. Section 4 presents the results obtained when applying the methodology to the national consumption of Ecuador. Section 5 provides an analysis of the results and discussion. Finally, Section 6 presents the main conclusions.

2. Methodology

The IPMPR for demand prediction has three stages based on the "statistical assessment for identifying changes in consumption" method [26], as depicted in Figure 1.

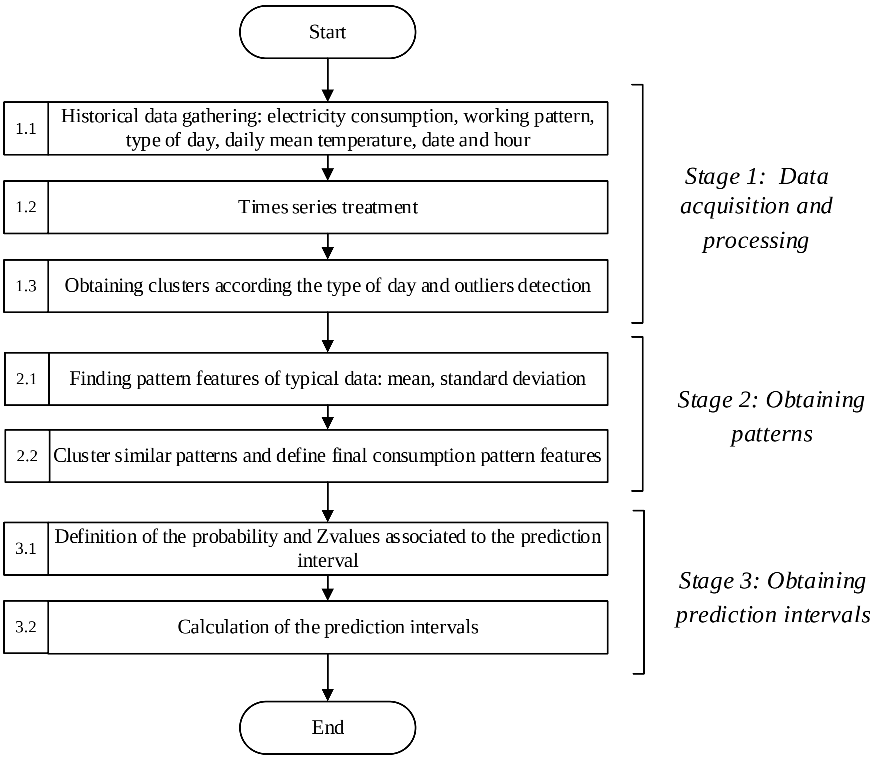


Figure 1. Interval prediction methodology of electricity demand based on pattern recognition (IPMPR)

2.1 Stage 1: Data acquisition and processing

In the first stage of the proposed methodology, the data of the daily load profiles (DLPs) are acquired and processed in such a way that they conform to a normal probability distribution, allowing the anomalies to be separated. Figure 2 shows how data is used in the proposed methodology, at the end of Stage 1 are obtained 14 disagregated matrices, which are utilized in the next step. Each of the sub-stages is described below.

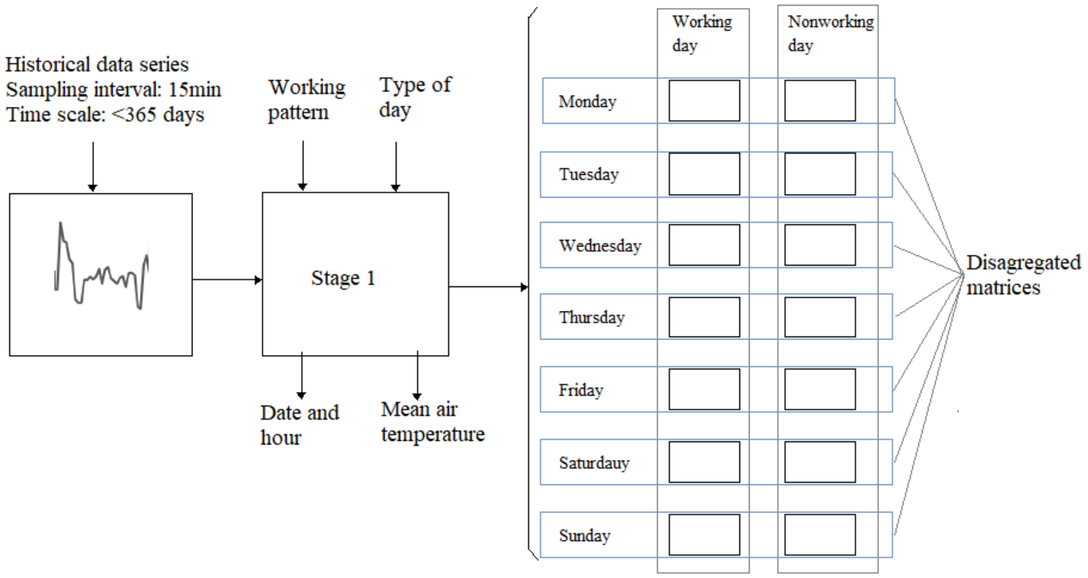


Figure 2. Data acquisition and processing in the proposed methodology

2.1.1 Historical data gathering

Three initial parameters are defined. The first is the analysis period (the number of days for which the electricity consumption data is collected). Each day is characterized by 96 values of quarter-hour measurements, thereby forming a DLP. In addition, the type of analysis day is defined by selecting the day of the week (Monday, Tuesday, Wednesday, etc.) and defining whether the day is working or non-working.

The DLP data is organized into a matrix of size M × A. M indicates the number of days in the selected period, and A indicates the number of characteristics considered for each day. The characteristics considered to represent the electricity consumption of a day are the 96 data points for the quarter-hour average power measurements and four variables used for data segmentation, i.e., the date and time, day of the week, working hours, and temperature. Therefore, the total number of characteristics is 100 (A = 100).

2.1.2 Times series treatment

The electricity demand is considered as a time series with seasonal variations owing to changes in work processes, ambient temperature, vacation periods, etc. [27]. In this type of data, it is necessary to evaluate whether a transformation or treatment of the data is required [28]. The data transformation in this sub-stage is conducted by applying the seasonal analysis of energy consumption (SAEC) method [29], which eliminates the trends and seasonal components of the data in such a way that the data instances fit better to the normal probability distribution, facilitating the interpretation of the results.

2.1.3 Obtaining clusters

If the data have been processed, they are reorganized in the M × A size matrix described in Section 2.1.1. In the opposite case, the original matrix is taken. Next, three columns are added to the data matrix to represent the maximum, average, and minimum power corresponding to each day, respectively. The data matrix then has dimensions of M × N, where N = 103.

The data in the matrix are segmented based on three segmentation variables, which are in turn based on two criteria. The first criterion considers two temporary categorical variables (the fourth-hour interval and day of the week), whereas the second considers the working pattern (working or non-working day). Then, the M × N matrix is broken down into 14 matrices, two for each day of the week. One matrix contains the working days, and the other contains the non-working days. Once data segmentation has been conducted, the segmentation variables are eliminated. The og matrices then have 100 columns or characteristics (N´ = 100), i.e., the 96 data points of the quarter-hour average power load and mean air temperature in the local area where the facility is located, in addition to the maximum, average, and minimum power load corresponding to each row.

The disaggregated matrices are standardized, making the mean zero (μ = 0) and standard deviation one (σ = 1), according to the procedure described in [26]. Standardization is performed for each column of the matrix. The value of Z corresponding to each row r and each column c is calculated as follows:

$$Z_{rc} = \frac{x_{rc} - u_c}{\sigma_c}, \tag{1}$$

In the above x_{rc} is the value of variable X in the disaggregated matrix of row *r* and column *c*, u_c is the mean, and σ_c is the standard deviation of variable X in column *c*. The maximum values of Z for each column are stored in a vector Z_{max}. The minimum values of Z for each column are stored in a vector Z_{min}.

Subsequently, anomalous data are detected and eliminated; a DLP is considered anomalous when at least one of the N values of a day of analysis is outside the normal distribution's 95% confidence interval. This procedure is performed for each column of each disaggregated matrix, including the mean air temperature.

The probability that a Z value corresponding to column c is within the confidence interval can be expressed as follows:

$$P\left(Z_{\alpha_1,c} < Z_c < Z_{\alpha_2,c}\right) = 1 - \alpha_{1,c} - \alpha_{2,c} = 1 - \alpha \quad \textbf{(2)}$$

Here, $Z_{\alpha 1,c}$ and $Z_{\alpha 2,c}$ are the lower and upper limits of the confidence interval for each column c, respectively. $\alpha_{1,c}$ and $\alpha_{2,c}$ represent the left and right tail areas of the standard normal distribution, respectively.

At the end of this stage, all of the data corresponding to the atypical days are eliminated from the disaggregated matrices, leaving only the data for typical electricity demand days. Each matrix's size is $M_d \times 96$, where M_d is the number of days comprising each pattern matrix, and 96 represents the hourly quarter-power values for each day.

2.2 Stage 2: Obtaining patterns

2.2.1 Finding pattern features

The stochastic consumption pattern for each day of the week is obtained from the pattern matrices presented in the previous section. This pattern is represented by four vectors representing its characteristics: the mean, standard deviation, and values of Z_{min} and Z_{max} for each column corresponding to each pattern matrix.

2.2.2 Clustering of similar patterns

Similar patterns are grouped in this stage. Clustering is beneficial, as statistical methods work better when more data are available. As a grouping metric, the Euclidean distance is used, whose equation is expressed as follows:

$$Distance(A,B) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2} \quad \textbf{(3)}$$

In the above, A and B are vectors with the same number of elements. In this case, n = 96, as the values a_i and b_i represent the normalized pattern characteristics, which were defined in the previous section. Normalization in this case simply refers to a scale change by dividing each value a_i and b_i for the maximum value of the two vectors.

The grouping process begins by calculating the Euclidean distance between the normalized vectors of the mean and standard deviation for each pattern matrix (seven matrices, one for each day of the week). If the calculated distance is less than an arbitrary threshold value (in this work 0.2), the patterns are considered similar, and then the corresponding pattern matrices are grouped.

The final consumption pattern is represented by a final pattern matrix formed by the data from the days of demand that are similar to the day to be predicted. From this matrix, the characteristics expressed in the vectors are obtained, i.e., the mean, standard deviation, Z_{min}, and Z_{max}, and these variables are computed column-by-column in each final pattern matrix.

The information in each final pattern matrix is best appreciated when represented with box-and-whisker plots; these can also be called stochastic patterns. Using the proposed SAEC method, less variable stochastic consumption patterns are obtained, improving the precision in detecting the anomalies in DLPs.

2.3 Stage 3: Obtaining prediction intervals

Before calculating the PIs, an adjustment is made to the mean, provided that the data transformation using the SAEC method has been conducted. The adjustment procedure is as follows.

1. A base load value is subtracted from the obtained mean values, i.e., 1% of all of the observations.

2. The obtained values from the previous step are multiplied by the value of the seasonality index corresponding to the most recent week, according to the procedure detailed in [29].

3. The values of the base load are added to the values obtained from the previous step.

4. The probability $(1-\alpha)$ is defined and associated with the PI. For this study, probabilities of 60%, 80%, and 95% are selected.

5. The values $z_{\alpha/2}$ y $z_{-\alpha/2}$ of the standard normal distribution are defined to meet a condition based on the probability, as follows:

$$\int_{z_{-\alpha/2}}^{z_{\alpha/2}} n(z;0,1)dz = 1-\alpha \tag{4}$$

6. A PI is calculated from the mean and standard deviation. Equation 5 shows the expression used.

$$\mu - z_{\alpha/2}\sigma\sqrt{1+\frac{1}{n}} < X_0 < \mu + z_{\alpha/2}\sigma\sqrt{1+\frac{1}{n}} \tag{5}$$

In the above, μ is the median, $Z_{\alpha/2}$ is the value of Z that leaves a value of $\alpha/2$ under the normal curve, X_0 represents the random variable to predict, and n denotes the number of data. In this case, it represents the number of DLPs within the final pattern matrix. The upper and lower bounds of the PI are defined by Equations 6 and 7, respectively, as follows:

$$U = \mu + z_{\alpha/2}\sigma\sqrt{1+\frac{1}{n}} \tag{6}$$

$$L = \mu - z_{\alpha/2}\sigma\sqrt{1+\frac{1}{n}} \tag{7}$$

If the data transformation is not performed, the mean values of the final pattern do not change; that is, steps 1 to 3 are not performed, and the procedure continues from step 4.

3. Evaluation metrics

This section defines evaluation metrics commonly used in these studies, and specifically in this work, for quantifying the goodness of predictions.

3.1 Mean absolute percentage error (MAPE)

The MAPE is one of the most extended measures for evaluating errors in power load forecasting. Being a relative error, it allows for a comparison of results regardless of the magnitude of the values, as follows:

$$MAPE = \frac{1}{p}\sum_{i=1}^{p}\frac{\left|x_{o_i} - x_i\right|}{x_i} \times 100\ \% \tag{8}$$

Here, p is the number of predictions, x_o is the prediction value, and x is the real value of the observation.

3.2 Prediction Interval Coverage Probability (PICP)

The PICP assesses the accuracy of the PI. The PICP is defined as the cardinal feature of the PIs, and demonstrates the percentage of targets that will be covered by the upper and lower bounds [22]. The PICP is defined as [31] follows:

$$PICP = \frac{1}{p}\sum_{i=1}^{p} C_i \tag{9}$$

In the above, C_i is defined as follows:

$$C_i = \begin{cases} 1\ if\ x_i \in [L_i, U_i] \\ 0\ if\ x_i \notin [L_i, U_i] \end{cases} \tag{10}$$

Here, L_i and $U_i$are the lower and upper bouds of the PI, respectively. However, the PICP alone is an insufficient metric for indicating whether the PI is adequate. A large interval ensures a high value of the PICP, but it will be useless. Consequently, the PINAW is required as a complementary metric.

3.3 Prediction interval normalized average width (PINAW)

The PINAW evaluates the width of the PI. It is desirable to have as small of a PI width as possible. The PINAW is defined as follows [32]:

$$\mathrm{PINAW} = \frac{1}{pR}\sum_{i=1}^{p}(U_i - L_i) \tag{11}$$

In the above, R is the difference between the maximum and minimum values of the bounds of the PI.

4. Application of interval prediction methodology of electricity demand based on pattern recognition (IPMPR) to the national electricity demand of Ecuador

The data of the quarter-hour demand for electricity for the entire Republic of Ecuador corresponding to the dates between January 1, 2017 and December 9, 2018 were obtained through the National Electricity Operator [33]. Ecuador is one of the Latin American countries for which the demand for electric power has grown the most; for example, in the period from 1990 to 2016, the electricity consumption grew by 378% [34]. The supply of electricity for the year 2019 had a contribution of 76.3% from hydroelectricity [35]; thus, it is considered that Ecuador's generation matrix is relatively clean.

The air temperature data were obtained from the National Institute of Meteorology and Hydrology [36]. As the obtained prediction of the electricity demand is for the entire country, the acquired temperature data comprised the weighted averages for the three largest cities in Ecuador: Guayaquil, Quito, and Cuenca. The weights used for this calculation took values corresponding to the total population in each city.

Sunday, October 14, 2018 was selected as the final day for obtaining consumption patterns. The data analysis period was fixed at 364, 182, and 119 days, because these values contained an integer number of weeks. In each of the three cases, the PIs and percentage of variation of the intervals relative to their means were obtained. The evaluation of the results allowed for selecting the most appropriate analysis period.

In this case of the PIs with a data analysis period of 364 days, the data from October 16, 2017 to October 14, 2018 were analyzed. The seasonality in such consumption was low, as indicated by the seasonality index defined in [27] (Figure 3).

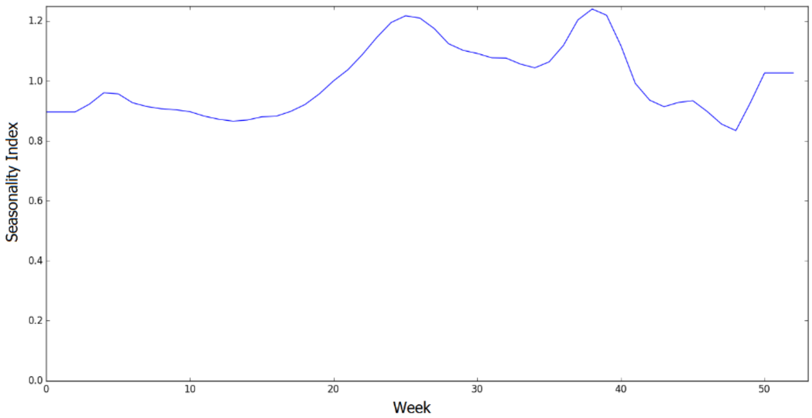


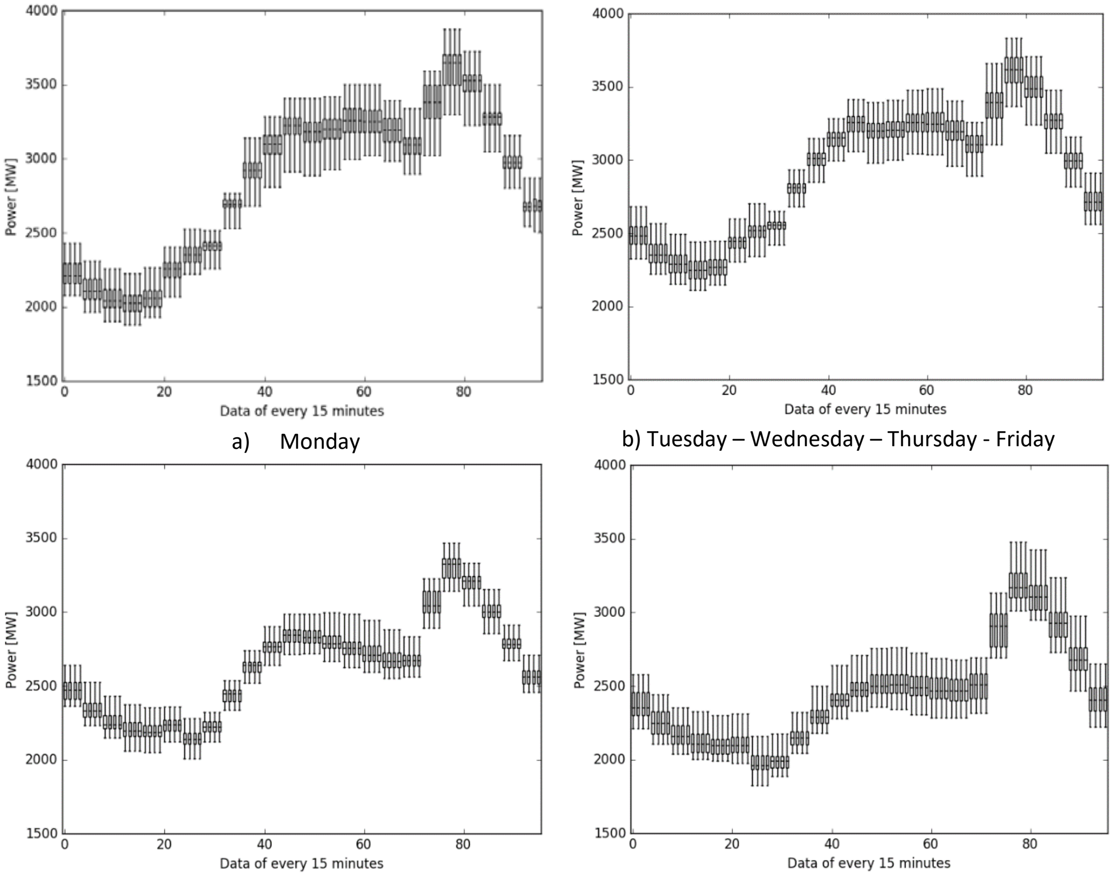
Figure 3. Seasonality index for a period of 364 days

The consumption patterns conformed to the values in Table 2 and could be represented with box plots, as shown in Figure 4.

Table 2. Formation of electricity consumption patterns for the 364-day period

Day of the week	Working pattern	Days with which it is grouped	Number of typical daily load profiles (DLPs)	Number of atypical DLPs	Total number of DLPs in the final pattern
Monday	Working day	-	42	6	42
Tuesday	Working day	Wednesday-Thursday-Friday	33	17	140

Wednesday	Working day	Tuesday-Thursday-Friday	33	19	140
Thursday	Working day	Tuesday-Wednesday-Friday	39	12	140
Friday	Working day	Tuesday-Wednesday-Thursday	35	13	140
Saturday	Working day	-	33	12	33
Sunday	Non-working day	-	42	10	42



a) Monday

b) Tuesday – Wednesday – Thursday - Friday

b) Saturday

d) Sunday

Figure 4. Electricity consumption patterns considering data from October 16, 2017 to October 14, 2018 (period of obtaining patterns: 364 days)

Figure 5 shows the PIs obtained from October 15–21, 2018, and the actual electricity demand.

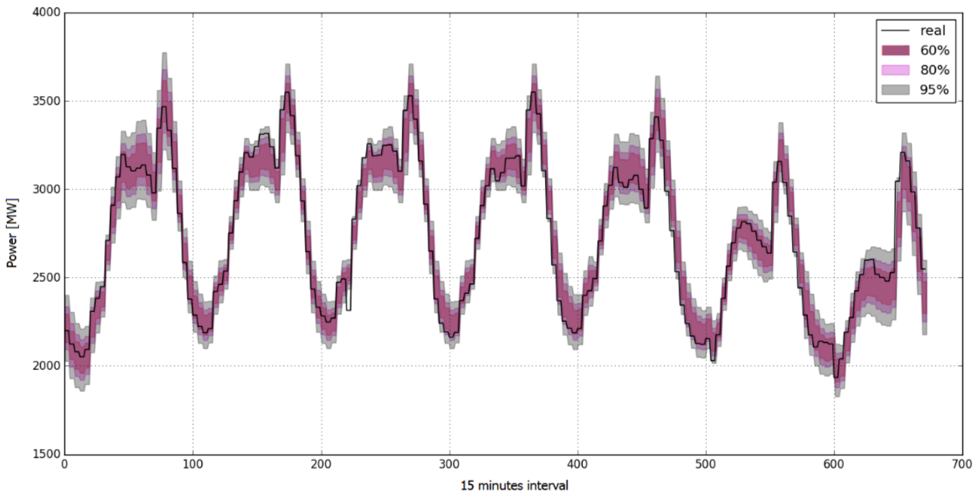


Figure 5. Prediction intervals with 60%, 80%, and 95% probability from October 15, 2018 to October 21, 2018 (period of obtaining patterns 364 days)

Table 3 shows the PICP values obtained for each PI. For example, in one week, the PICP values are 83.9% and 99.4% for the intervals associated with 80% and 95% probability, respectively. In contrast, the PINAW values are 9.39% and 13.33% for the same intervals, respectively.

Table 3. Summary of prediction intervals from October 15 to 21, 2018 (period for obtaining patterns 364 days)

Probability	Values within the range	Prediction interval coverage probability (PICP) [%]	Prediction interval normalized average width (PINAW) [%]
60%	416	61.90	6.43
80%	564	83.93	9.39
95%	668	99.40	13.33

The same procedure was applied to obtain the PIs for the defined periods of 182 and 119 days, and for the following 3 weeks. Table 4 shows a summary of the results from October 15 to November 11, 2018.

Table 4. Summary of the application of the prediction intervals from October 15 to November 11, 2018

No. days for analysis	Probability	PICP [%]	PINAW [%]
	60%	60.9	7.12
	80%	81.1	10.13
364	95%	96.7	14.13
	60%	60.0	6.21
	80%	77.1	8.79
182	95%	91.7	12.19
	60%	56.0	5.11
	80%	76.3	7.59
119	95%	91.8	10.88

The results indicate that the 364-day pattern collection period is the only period providing PIs in which the PICP is greater than the associated probability.

Figure 6 shows the PICPs for each predicted day in a scatter plot. The squares correspond to a 60% probability of occurrence, the circles to 80%, and the rhombuses to 95%. The PICPs obtained with patterns of 364 and 182 days are represented on the vertical and horizontal axes, respectively. It can be observed that in the patterns formed from 364 days of analysis, the PICP values are substantially higher than those obtained by patterns with 182 days. Specifically, for the case of the 95% probability, it can be seen that all points are above the diagonal line. This indicates that the PIs obtained with 182-day patterns never achieved better results than those obtained with 364 days. The results obtained with pattern periods of 119 days are discarded, owing to the low value of the PICPs.

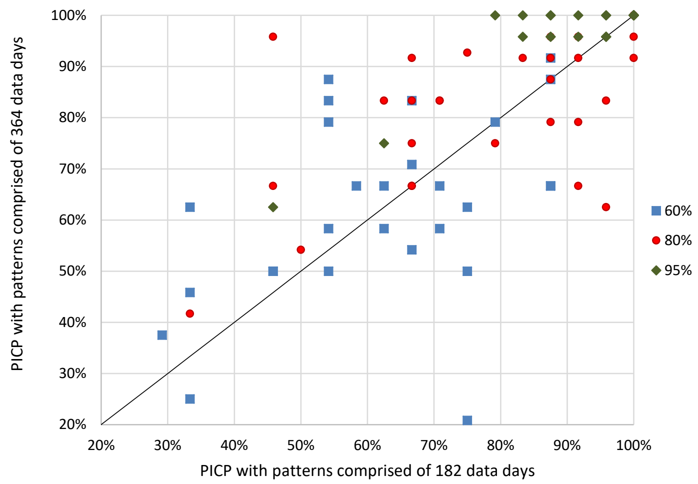


Figure 6. Prediction interval coverage probability (PICP) for 364 and 182 day pattern collection periods

5. Analysis of results and discussion

The results obtained for the demand from continental Ecuador indicate that the developed methodology is useful for generating PIs. For the case analyzed in Ecuador, where the electricity demand data have a weak seasonal component, the number of days prior to the week to be predicted can simply be selected and entered into the model. In addition, the interval with 80% probability appears to be the most adequate for predicting the electricity demand.

In this section, the results from an ANN LSTM model are compared with those from the proposed methodology (IPMPR). LSTM networks are characterized by achieving very good predictions for time series of electrical energy data.

5.1 Structure of the long short-term memory (LSTM) Network

The proposed ANN LSTM was structured according to the parameters in Table 5. The proposed network obtained the best results after multiple tests. The mean square error was evaluated for the different numbers of intermediate layers and neurons in each layer to find the configuration in which the least prediction error was obtained. This procedure was justified, as there is no analytical method for determining the optimal number of intermediate layers and neurons in an

ANN. The training optimization algorithm used in this study was Adam. Adam was proposed in 2014 by Kingma and Ba [37], and has become one of the most widely used algorithms in deep learning. Its major advantages are its easy implementation, low computational cost, low memory requirements, and appropriate use when the data have weak or very strong gradients.

Table 5. ANN LSTM parameters

Parameters	Configuration
Optimization algorithm for training	Adam
Number of layers	3
Number of neurons in the input layer	96
Number of neurons in the intermediate	100
Numbers of neurons in the last layer	96
Look back	7 days

The input and output layers had 96 neurons each. Tests have shown that when predicting electricity demands, the error decreases when this configuration is applied. Although LSTM networks are widely used in systems that vary with time, it has been demonstrated that when all of the data of the time series enter through the same neuron, the errors are considerably greater. This problem is solved by introducing a data entry matrix into a neuronal network of 96 × M dimensions, where each row represents a fourth hourly data of the DLP, and M is the number of days entering the neuronal network. As shown in Figure 7, each neuron in the input layer receives the fourth hourly data with 24-hour periodicity. Owing to this configuration, the network obtains lower errors, as the energy consumption at a specific moment of the day does not differ much from consumption on another day at the same time.

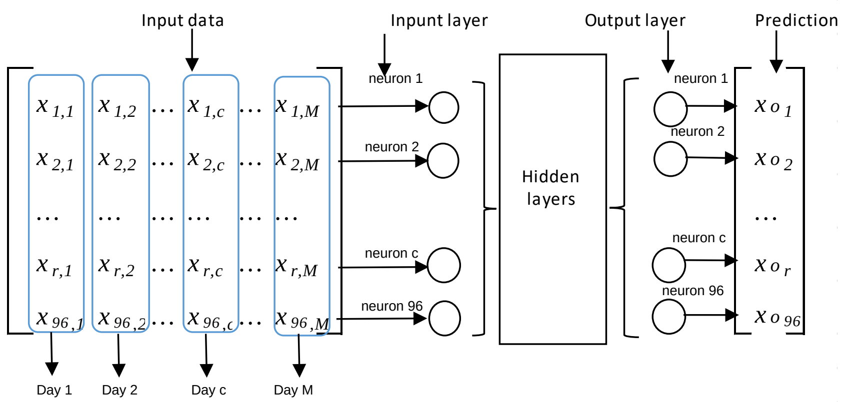


Figure 7. Artificial neural network (ANN) long short-term memory (LSTM) used for prediction

5.2 Results for the case study in Ecuador

The analyzed case study was the same as that described in Section 4.1. The available data from January 1, 2017 to September 12, 2018 (67970 records) were segmented in such a way that 70% of the initial data were used for training, and the remaining 30% (91 days) were used for evaluation of the predictions.

The results of the predictions are presented from September 10, 2018 to December 9, 2018 (13 weeks), and their goodness is compared with the predictions made by the proposed IPMPR approach. For this purpose, the MAPE is used, it is the most commonly used error metric for evaluating predictions.

In the literature, some studies indicate that the use of MAPE brings with it some issues, such as (i) MAPE cannot be used if there are zero values. Still, the power load never reaches zero in some instances, such as in many industries, commercial buildings, power grid substations, or the entire country's power demand. (ii) "MAPE is based on assumptions that (a) accurate forecasting of small loads is important and (b) one large error is not more significant than an equally large sum of small absolute errors" [30]. Nevertheless, all types of errors incur this issue in a certain way because they are based on an arithmetic average of the difference between the actual values and the predictions. (iii) "MAPE also penalizes over-forecasts (where forecast load is greater than realized load) more than under–forecasts" [30]. However, when predictions are accurate, this penalization is small. In this methodology, the use of patterns for prediction allows the over-forecast and the under-forecast, which are not significant so that the MAPE is between 2% and 3%.

In the 13 weeks evaluated, an average MAPE value of 2.73% is evident (Table 6). The obtained error is acceptable relative to other studies. In contrast, there are days in which the prediction of the ANN fails considerably. For example, in the prediction weeks 4, 5, and 8 in Table 6, the values are considerably higher than the total mean value. A closer look at week 8, when the highest errors in the prediction occur, reveals a weakness in LSTM networks. As indicated above, LSTM networks have the capacity to memorize patterns. In this case, however, this characteristic becomes a disadvantage, because when there are holidays (from November 1st to 4th), the network makes errors in its predictions. As can be seen in Figure 8, Thursday is the first holiday to behave like a Saturday, and Friday behaves like a Sunday. The network then predicts consumption on Thursday as a common working day, but because the consumption on that day is very similar to that on a Saturday, the next day is predicted as a Sunday when it is actually a Friday. The neuronal network then predicts the following day (Saturday) as if it were a Monday. In this way, errors are accumulated, resulting in a deficient prediction.

Table 6. MAPE in predicting electricity demand in Ecuador for thirteen weeks using the ANN LSTM and interval prediction methodology of electricity demand based on pattern recognition (IPMPR) (October 9, 2018 to September 12, 2018)

Week	Start date	End date	Methodology	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Mean
1	10/09/2018	16/09/2018	ANN LSTM	2.10%	1.34%	2.12%	1.91%	1.85%	2.60%	1.49%	1.92%
			IPMPR	1.17%	1.28%	1.39%	1.62%	2.30%	1.24%	1.66%	1.52%
2	17/09/2018	23/09/2018	ANN LSTM	2.03%	1.79%	2.32%	1.66%	2.49%	2.23%	1.77%	2.04%
			IPMPR	2.54%	1.06%	1.05%	0.97%	1.37%	2.16%	3.17%	1.76%
3	24/09/2018	30/09/2018	ANN LSTM	2.34%	1.07%	1.52%	2.52%	1.09%	2.22%	2.33%	1.87%
			IPMPR	1.39%	0.93%	1.71%	2.73%	1.52%	3.13%	3.07%	2.07%
4	01/10/2018	07/10/2018	ANN LSTM	3.37%	5.97%	3.17%	3.65%	2.02%	2.12%	2.62%	3.28%
			IPMPR	2.60%	2.26%	1.42%	2.35%	2.71%	1.71%	2.51%	2.22%
5	08/10/2018	14/10/2018	ANN LSTM	8.74%	8.10%	2.95%	1.90%	2.35%	4.48%	2.81%	4.48%
			IPMPR	6.83%	2.35%	2.00%	2.08%	1.96%	2.55%	1.47%	2.75%
6	15/10/2018	21/10/2018	ANN LSTM	2.15%	2.81%	2.25%	1.43%	1.68%	2.84%	2.16%	2.19%
			IPMPR	1.57%	1.69%	1.94%	2.17%	2.16%	2.24%	3.00%	2.11%
7	22/10/2018	28/10/2018	ANN LSTM	3.24%	2.45%	1.19%	1.33%	1.32%	2.68%	1.15%	1.91%

			IPMPR	2.82%	2.77%	2.18%	2.09%	1.84%	1.54%	1.74%	2.14%
8	29/10/2018	04/11/2018	ANN LSTM	1.79%	1.32%	2.17%	8.06%	2.96%	11.57%	12.91%	5.82%
			IPMPR	1.95%	2.20%	1.72%	1.79%	1.82%	3.23%	5.37%	2.58%
9	05/11/2018	11/11/2018	ANN LSTM	2.53%	2.24%	2.23%	2.92%	3.11%	3.36%	3.12%	2.79%
			IPMPR	2.57%	2.99%	1.70%	2.37%	2.40%	3.91%	1.88%	2.55%
10	12/11/2018	18/11/2018	ANN LSTM	3.63%	1.18%	1.40%	1.56%	2.04%	3.33%	2.53%	2.24%
			IPMPR	2.01%	1.02%	1.38%	2.01%	2.05%	2.78%	2.15%	1.91%
11	19/11/2018	25/11/2018	ANN LSTM	2.22%	2.55%	2.50%	1.75%	3.44%	4.41%	2.25%	2.73%
			IPMPR	1.40%	2.10%	1.84%	1.43%	2.20%	4.94%	5.05%	2.71%
12	26/11/2018	02/12/2018	ANN LSTM	2.57%	1.87%	1.62%	2.19%	1.42%	2.25%	2.34%	2.04%
			IPMPR	3.05%	1.71%	2.51%	1.37%	1.15%	2.26%	3.09%	2.16%
13	03/12/2018	09/12/2018	ANN LSTM	2.99%	1.78%	1.15%	2.21%	1.80%	2.65%	2.69%	2.18%
			IPMPR	1.33%	0.73%	1.53%	2.11%	2.64%	3.52%	5.92%	2.54%
Media total			ANN LSTM	3.05%	2.65%	2.04%	2.54%	2.12%	3.60%	3.09%	2.73%
			IPMPR	2.40%	1.77%	1.72%	1.93%	2.01%	2.71%	3.08%	2.23%

The following is a summary of the results obtained by the IPMPR approach for the PIs in the same period as the ANN LSTM. The electricity demand forecasting intervals were obtained from September 10 to December 9, 2018. As indicated above, interval prediction is a different approach from the traditional one. In particular, ANNs obtain point prediction values, and their goodness is evaluated by calculating an error. In contrast, the proposed method obtains minimum and maximum values and the probability associated with the real value falling within that interval; thus, comparing the goodness of each method is not a trivial task. Accordingly, it is proposed to obtain the mean of the intervals, and to consider this value as the prediction, so as to later calculate an MAPE value that can be compared with that obtained by the ANN LSTM network. The mean value of the MAPE for the 13 weeks evaluated is 2.23%, i.e., 20% less than the error obtained by the ANN LSTM. These error values are listed in Table 6. Other types of error such as mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE) were also analyzed in this study. The results for the 13 weeks evaluated are described in Table 7.

Table 7. Evaluation of different errors in forecasting results

Methodology	MAPE [%]	MAE [MW]	MSE [MW^2]	RMSE [MW]
ANN LSTM	2,73	53,49	5373,95	68,12
IPMPR	2,23	65,48	6579,87	83,31

When evaluating the 91-day predictions, it is evident that on 55 occasions, the IPMPR obtains a lower error than the LSTM ANN. In prediction week 8, the greatest decrease in the MAPE is achieved, from 5.82% to 2.58%. The reduction is achieved owing to the fact that the proposed methodology segments the data in such a way that the pattern of a holiday is known a priori, and is generally grouped with consumption of Saturdays and/or Sundays. Figure 7 shows the actual demand for electricity, prediction made by the neural network, and 80% PI from October 29th to November 4th, 2018.

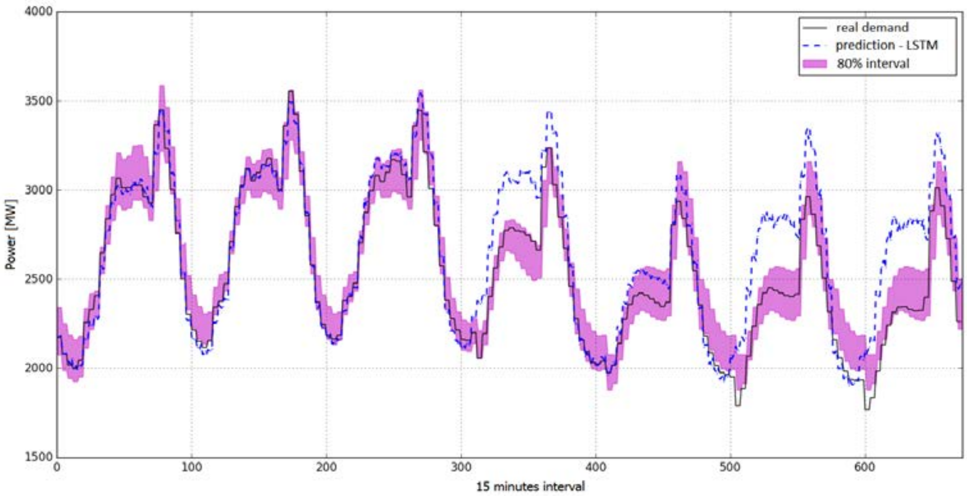


Figure 8. Forecast from October 29, 2018 to November 4, 2018 for Ecuador using ANN LSTM and IPMPR

The proposed methodology presents additional information regarding the predictions made. For example, for Friday, September 28, 2018, the method provides different PIs. As an example, the intervals associated with probabilities of occurrence of 60%, 80%, and 95% are chosen, where the PICP values are 66.7%, 95.8%, and 100%, respectively (Figure 9). Similarly, the PINAW values are 5.7%, 8.22%, and 11.6% for the same probability of occurrence. In addition, the percentage of daily variation of each PI with respect to its mean is known (Table 8). With this information, it is possible to know the estimated fourth hour variation. For example, for Friday between 9:30 and 9:45, the variations in the PIs of 60%, 80%, and 95% are ± 1.82%, ± 2.77%, and ± 4.23%, respectively.

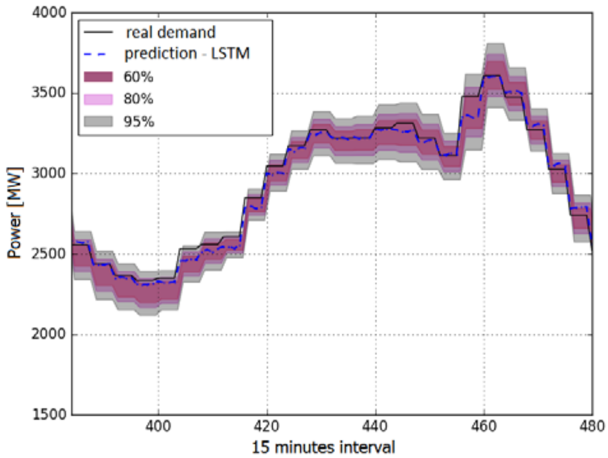


Figure 9. Forecast for Friday, September 28, 2018 using the ANN LSTM and IPMPR

Table 8. Percentage variation of the prediction intervals with respect to their mean (Friday, September 28, 2018)

Average range with respect to the pattern average [±]			Maximum range with respect to the pattern average [±]			Minimum range with respect to the pattern average [±]		
Probability			Probability			Probability		
60%	80%	95%	60%	80%	95%	60%	80%	95%
2.13%	3.24%	4.96%	2.99%	4.56%	6.97%	1.36%	2.08%	3.18%

5.3 Discussion

The results from the analyzed case studies allow for the identification of the advantages from obtaining the PIs using the proposed methodology, relative to the predictions made with the ANN LSTM. The most noteworthy advantages are as follows.

- The PIs present information for limiting the uncertainty of the prediction.
- The proposed methodology is able to establish the PI and PICP, in contrast to the ANNs.
- A percentage variation of the interval with respect to its mean is defined for each defined PI and PINAW.

The mean of the PI can be defined as a point value of prediction, so it is possible to calculate the MAPE. The error values are similar to those obtained when making predictions with the ANN LSTM.

Moreover, the usefulness of the PIs is compared with other methods found in the literature. The obtained results show that the sharpness of the predictions are similar. Table 9 presents the obtained values for the PICP and PINAW from different works.

Table 9. PICP and PINAW in different works

Work	Used techniques	Case study	Time horizon	Confidence interval [%]	PICP [%]	PINAW [%]
Proposed IPMPR	Pattern recognition and Statistics	Electricity demand of Ecuador	7 days (96 points/day)	80	81.10	10.13
[20]	Gaussian processes (GPs) log-normal process (LP) kernels encodes	residential customers in Sydney, Australia	N/D	N/D	82.00	13.00
[25]	Variational mode decomposition (VMD). Quantile Regression Forest (QRF) Bayesian optimization algorithm	Power grid load in Henan Province, China.	1 day (24 points/day)	80	91.67	16.64
[23]	Yeo-Johnson transformation quantile regression Probability density forecasting Gaussian kernel function Quantile function crossing	Electricity load in winter of Ottawa, Canada	7 days (24 points/day)	N/D	99.40	22.50
[24]	Quantile regression forest Gradient boosting machine	GEFCom 2014 load track	1 day	80	89.13	N/D
[19]	Gaussian Processes ARIMA model	Residential customers in Sydney, Australia	30 minutes	80	81.22–91.12	9.69–15.46
[22]	kernel-based Support vector quantile regression Copula theory	Electricity demand of Singapore	10 days (48 points per day)	N/D	94.79	37.04
[38]	Multi-objective particle swarm optimization algorithm (MOPSO) and least squares support vector regression LSSVR	Electricity demand of Singapore	N/D	90	94.17	3.00

* N/D No data

This work utilizes nationwide power consumption as input data. In the presented case study, the use of temperature variable's as the weighted average of the most populated areas has resulted in more accurate results. Note that this paper presents a general methodology that can be used for any type of consumer or facility. Generally, a consumer is in a specific zone, so it would not be necessary to perform the temperature's weighted average since it will be introduced directly.

The IPMPR approach provides predictions with a range of values with a defined probability that the real value is within the range. This has an important application for the detection of anomalies in electrical consumption; it can easily identify situations where there are inefficiencies in the consumption of a facility. If an actual value is very far from the PIs, it can be assumed with great certainty that the consumption has been abnormal. In addition, this approach employs a simple methodology and it is easily programmable in an electrical consumption management system.

6. Conclusions

The application of machine learning in the prediction of electricity demand revealed certain limitations, as follows.

- Difficulty in quantifying the uncertainty of the results and interpreting them physically, as electricity demand is a continuous stochastic variable.
- High computer expenditures (computing and time).
- Considerable expertise and time is required to establish the structure and configuration of the network, that is, number of neurons, layers, optimization algorithm, etc.

This study presents an interval prediction methodology of electricity demand based on pattern recognition (IPMPR) that solves the limitations of the above-mentioned machine learning methods by obtaining a PI with a time horizon of seven days (15 min time step). In addition, it presents the percentage of variation of the interval with respect to its mean, thus limiting the uncertainty of the prediction. The case study analyzed in Ecuador obtained a PICP of 81.1% and PINAW of 10.13% with a confidence interval of 80%.

The predictions made by the IPMPR are generally better than those obtained with the ANN LSTM. Specifically, the average MAPE value for the evaluated weeks is 2.23%, which is 20% less than the error achieved by the ANN LSTM. The errors in conventional methods are substantially higher when there are atypical consumption patterns, such as those occurring on holidays.

It can further be concluded that the use of the IPMPR approach provides predictions with a range of values, and has a defined probability that the real value is within the range. This has a very important application for the detection of anomalies in consumption; it can easily identify situations where there are inefficiencies in the consumption of an installation. If an actual value is very far from the PIs, it can be assured with great certainty that the consumption has been abnormal. In addition, because this approach employs a simple methodology, it is easily programmable in a consumption management system.

In addition, the prediction based on intervals can very effectively characterize the variability of the analyzed consumptions, as for a certain probability, a wider interval indicates that there is more dispersion in the initial data.

7. References

[1] V. M. Rueda, V. Henao, J. David, F. Cardona, and C. Jaime, “Recent Advances in Load Forecasting Using Nonlinear Models,” *Dyna*, vol. 78, no. 167, pp. 36–43, 2011.

[2] R. G. HOOKE, “Forecasting the Demand ror Electricity,” *Electr. Eng. 75(2), 132-132.*, 1956.

[3] H. G. Lakhani and B. Bumb, “Forecasting Demand for Electricity in Maryland: An Econometric Approach.”

[4] O. Derek W Bunn Department of Engineering Science, University of Oxford, Parks Road

and U. OXI 3PJ, “Experimental study of a Bayesian method for daily electricity load forecasting,” 1979.

[5] R. F. Engle, S. J. Brown, G. Stern, and J. Wiley, “A Comparison of Adaptive Structural Forecasting Methods for Electricity Sales,” *J. Forecast.*, vol. 7, pp. 149–172, 1988.

[6] J. T. Connor, “A Robust Neural Network Filter for Electricity Demand Prediction,” *J. Forecast.*, vol. 15, pp. 437–458, 1996.

[7] J. Ahmad, H. Larijani, R. Emmanuel, M. Mannion, A. Javed, and M. Phillipson, “Energy demand prediction through novel random neural network predictor for large non-domestic buildings,” in *2017 Annual IEEE International Systems Conference (SysCon)*, Apr. 2017, pp. 1–6, doi: 10.1109/SYSCON.2017.7934803.

[8] G. Escrivá-Escrivá, C. Álvarez-Bel, C. Roldán-Blay, and M. Alcázar-Ortega, “New artificial neural network prediction method for electrical consumption forecasting based on building end-uses,” *Energy Build.*, vol. 43, no. 11, pp. 3112–3119, 2011, doi: 10.1016/j.enbuild.2011.08.008.

[9] C. Roldán-Blay, G. Escrivá-Escrivá, C. Álvarez-Bel, C. Roldán-Porta, and J. Rodríguez-García, “Upgrade of an artificial neural network prediction method for electrical consumption forecasting using an hourly temperature curve model,” *Energy Build.*, vol. 60, pp. 38–46, 2013, doi: 10.1016/j.enbuild.2012.12.009.

[10] X. Wang, D. Luo, X. Zhao, and Z. Sun, “Estimates of energy consumption in China using a self-adaptive multi-verse optimizer-based support vector machine with rolling cross-validation,” *Energy*, vol. 152, pp. 539–548, Jun. 2018, doi: 10.1016/J.ENERGY.2018.03.120.

[11] M. W. Ahmad, M. Mourshed, and Y. Rezgui, “Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption,” *Energy Build.*, vol. 147, pp. 77–89, Jul. 2017, doi: 10.1016/J.ENBUILD.2017.04.038.

[12] S. Ryu, J. Noh, and H. Kim, “Deep Neural Network Based Demand Side Short Term Load Forecasting,” *Energies*, vol. 10, no. 1, p. 3, Dec. 2016, doi: 10.3390/en10010003.

[13] E. Mocanu, P. Nguyen, M. Gibescu, and W. L. Kling, “Deep learning for estimating building energy consumption,” *Sustain. Energy, Grids Networks*, 2016, doi: 10.1016/j.segan.2016.02.005.

[14] A. Rahman, V. Srikumar, and A. D. Smith, “Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks,” 2017, doi: 10.1016/j.apenergy.2017.12.051.

[15] X. Serrano-Guerrero, L. Siavichay, J. Clairand, and G. Escrivá-Escrivá, “Forecasting Building Electric Consumption Patterns Through Statistical Methods,” in *Advances in Intelligent Systems and Computing*, 2020, vol. 1067, doi: 10.1007/978-3-030-32033-1_16.

[16] J. W. Taylor, “Short-term electricity demand forecasting using double seasonal exponential smoothing,” *J. Oper. Res. Soc.*, vol. 54, no. 8, pp. 799–805, 2003, doi: 10.1057/palgrave.jors.2601589.

[17] A. Foucquier, S. Robert, F. Suard, L. Stéphan, and A. Jay, “State of the art in building modelling and energy performances prediction: A review,” *Renew. Sustain. Energy Rev.*, vol. 23, pp. 272–288, 2013, doi: 10.1016/j.rser.2013.03.004.

[18] X. Serrano-Guerrero, R. Prieto-Galarza, E. Huilcatanda, J. Cabrera-Zeas, and G. Escriva-Escriva, “Election of variables and short-term forecasting of electricity demand based on backpropagation artificial neural networks,” *2017 IEEE Int. Autumn Meet. Power, Electron. Comput.*, pp. 1–5, Nov. 2017, doi: 10.1109/ROPEC.2017.8261630.

[19] D. W. van der Meer, M. Shepero, A. Svensson, J. Widén, and J. Munkhammar, “Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using Gaussian Processes,” *Appl. Energy*, vol. 213, no. October 2017, pp. 195–207, 2018, doi: 10.1016/j.apenergy.2017.12.104.

[20] M. Shepero, D. van der Meer, J. Munkhammar, and J. Widén, “Residential probabilistic load forecasting: A method using Gaussian process designed for electric load data,” *Appl. Energy*, vol. 218, no. March, pp. 159–172, 2018, doi: 10.1016/j.apenergy.2018.02.165.

[21] D. W. van der Meer, J. Munkhammar, and J. Widén, “Probabilistic forecasting of solar power, electricity consumption and net load: Investigating the effect of seasons, aggregation and penetration on prediction intervals,” *Sol. Energy*, vol. 171, no. March, pp. 397–413, 2018, doi: 10.1016/j.solener.2018.06.103.

[22] Y. He, R. Liu, H. Li, S. Wang, and X. Lu, “Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory,” *Appl. Energy*, vol. 185, pp. 254–266, 2017, doi: 10.1016/j.apenergy.2016.10.079.

[23] Y. He and Y. Zheng, “Short-term power load probability density forecasting based on Yeo-Johnson transformation quantile regression and Gaussian kernel function,” *Energy*, vol. 154, pp. 143–156, 2018, doi: 10.1016/j.energy.2018.04.072.

[24] W. Zhang, H. Quan, and D. Srinivasan, “Parallel and reliable probabilistic load forecasting via quantile regression forest and quantile determination,” *Energy*, vol. 160, pp. 810–819, 2018, doi: 10.1016/j.energy.2018.07.019.

[25] F. He, J. Zhou, L. Mo, K. Feng, G. Liu, and Z. He, “Day-ahead short-term load probability density forecasting method with a decomposition-based quantile regression forest,” *Appl. Energy*, vol. 262, no. December 2019, 2020, doi: 10.1016/j.apenergy.2019.114396.

[26] X. Serrano-Guerrero, G. Escrivá-Escrivá, and C. Roldán-Blay, “Statistical Methodology to Assess Changes in the Electrical Consumption Profile of Buildings,” *Energy Build.*, vol. 164, pp. 99–108, 2018, doi: https://doi.org/10.1016/j.enbuild.2017.12.059.

[27] J. X. Serrano Guerrero, “Caracterización de la demanda de energía mediante patrones estocásticos en las Redes Eléctricas Inteligentes,” Universitat Politècnica de València, 2020.

[28] J. E. Seem, “Pattern recognition algorithm for determining days of the week with similar energy consumption profiles,” *Energy Build.*, vol. 37, no. 2, pp. 127–139, 2005, doi: 10.1016/j.enbuild.2004.04.004.

[29] X. Serrano-Guerrero, G. Escrivá-Escrivá, S. Luna-Romero, and J.-M. Clairand, “A Time-Series Treatment Method to Obtain Electrical Consumption Patterns for Anomalies Detection,” *Energies*, 2020, doi: 10.3390/en13051046.

[30] P. Koponen, J. Ikäheimo, J. Koskela, C. Brester, and H. Niska, “Assessing and comparing

short term load forecasting performance," *Energies*, vol. 13, no. 8, 2020, doi: 10.3390/en13082054.

[31] A. Khosravi, S. Nahavandi, and D. Creighton, "Construction of optimal prediction intervals for load forecasting problems," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1496–1503, 2010, doi: 10.1109/TPWRS.2010.2042309.

[32] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Comprehensive review of neural network-based prediction intervals and new advances," *IEEE Trans. Neural Networks*, vol. 22, no. 9, pp. 1341–1356, 2011, doi: 10.1109/TNN.2011.2162110.

[33] Operador Nacional de Electricidad, "CENACE," 2018. http://www.cenace.org.ec/ (accessed Nov. 05, 2018).

[34] International Energy Agency, "Statistics data browser." https://www.iea.org/statistics (accessed Mar. 04, 2019).

[35] I. de I. G. y E.- IIGE, "Balance energético nacional 2019." p. 174, 2020, [Online]. Available: http://bit.ly/3nvFB5l.

[36] "Instituto Nacional de Meteorología e Hidrología," "No Title," *RED DE ESTACIONES METEOROLÓGICAS E HIDROLÓGICAS*, 2019. http://www.serviciometeorologico.gob.ec/ (accessed Nov. 02, 2019).

[37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv Prepr. arXiv1412.6980*, 2014.

[38] P. Jiang, R. Li, N. Liu, and Y. Gao, "A novel composite electricity demand forecasting framework by data processing and optimized support vector machine," *Appl. Energy*, vol. 260, no. November 2019, p. 114243, 2020, doi: 10.1016/j.apenergy.2019.114243.