

# A COMPUTATIONAL APPROACH TO FORECASTING AND MINIMIZING ELECTRICITY COSTS IN THE SHORT-TERM MARKET FOR DISTRIBUTORS IN BRAZIL

Francisco E. Bezerra<sup>1</sup> – Luis C. Santos Jr<sup>2</sup> – Cleber G. Dias<sup>2</sup> – Fabio H. Pereira<sup>1,2\*</sup>

<sup>1</sup>Industrial Engineering Graduate Program, Universidade Nove de Julho, Rua Vergueiro, 235/249, 01504-001, Liberdade, São Paulo – Brazil

<sup>2</sup>Informatics and knowledge management Graduate Program, Universidade Nove de Julho, Rua Vergueiro, 235/249, 01504-001, Liberdade, São Paulo – Brazil

---

## ARTICLE INFO

### Article history:

Received: 17.10.2019.

Received in revised form: 26.12.2019.

Accepted: 27.01.2020.

### Keywords:

Power distributor

Commercialization of Electric Energy

Forecasting electricity

Multilayer Network

DOI: <https://doi.org/10.30765/er.1529>

## Abstract:

*In Brazil, the electric power distributors must buy electricity on auctions one, three and five years ahead. If there is inefficiency in the contracting of electric energy, the chamber of Commercialization of Electric Energy, which enables the commercialization, can apply penalties. Thus, this paper proposes a computational approach to forecasting electricity by the class of the consumer using a multi-layer perceptron artificial neural network with a backpropagation algorithm and a prediction using time series techniques through the Bayesian and Akaike selection criteria. The forecast of electricity consumption can serve as support in the purchase of electricity in auctions in the regulated contracting environment and in the process of settlement of differences and for energy management, customer service, and distributor billing. The results show that a multilayer network with a backpropagation algorithm is able to learn the behavior of the data that influences the electric energy consumed by consumption class and can be used to follow the evolution in the demand of each class of consumption and, consequently, to help distributors in the process of contracting of electricity, reduce losses like fines, and reduce the costs of the energy distributor.*

---

## 1 Introduction

Since the creation of the new commercial model for electric power in Brazil, Decree No. 5,163 of 2004, electric power distributors have been obliged to purchase electric power through auctions conducted in the Regulated Contracting Environment (RCE), to attend to its customers belonging to various consumption classes, such as commercial, own consumption, public illumination, industrial, public authorities, domestic, rural, and public service, among others [1].

The new model also obliges each distributor to meet 100% of its demand, in such a way that if there is more consumption than what is contracted for, the distributor will be penalized. On the other hand, it is not allowed to contract for energy over 5% of its annual demand, because above this value, the distributor will also suffer penalties applied by national electric power agency ANEEL [1].

Each country adopts a different system of marketing of electric energy. In China, for example, the entire electric sector is controlled by state-owned enterprises: there is no competition in the services of generation, transmission and distribution of electric energy. The marketing is done through the wholesale market, however, only the electric companies can buy electrical energy on the short-term market. The electric companies can carry out monthly or annually contracts involving the generation [2].

---

\* Corresponding author

E-mail address: [fabiohp@uni9.pro.br](mailto:fabiohp@uni9.pro.br)

The energy marketing in Germany works through a wholesale market that works in two environments: short-term market, in which energy for the short term is marketed to cover the risk of variations of prices, or through negotiation directly between the agents of the market [2].

The state of California in the United States works with a wholesale energy market in which the operator of the system of distribution, California Independent System Operator (CASIO), alone administers nearly 80% of the electric network, with the other 20% being the responsibility of public enterprises. Great quantities of energy are marketed in the market for the next day or next hour, in which the prices for the purchase and sale of electricity for the next day are determined from hour to hour [2].

Hong and Fan, in [3] provide a tutorial review of the probabilistic forecasting of electricity prices, which is considered a new branch of the load prediction problem, with the main objective of offering ideas and assisting in the development of models and methods useful for researchers and professionals in the field of load forecasting. According to the authors, there are five new problems that introduce greater uncertainties for the electric power industry: climate variability, electric vehicles, wind and solar energy generation, energy efficiency, and demand response.

With this, electric energy forecasting has become vital for those who work in the electricity market, mainly for the enterprise distributors that need to annually report electricity consumption to the National Electric Power Agency (ANEEL). Besides, forecasting provides conditions to solve several decision-making problems, such as: price planning, tariffs, distribution network operations, energy production planning, tariff management, customer care and collections [4].

The forecast of electricity demand is the main task in planning for electricity production, as it can determine the resources required for the operation of hydroelectric and thermal power plants of the country, such as the volume of water required to generate the power level requested or consumption of fuels [5].

As each country has different types of electric plant, different factors affect the demand for electricity. Therefore, it is necessary to adopt a special forecasting method, since no method can be generalized to all demand patterns [5].

In relation to energy prices, the forecast is necessary because it supports those involved in the generation, distribution, and trading of energy to set their bidding strategies, increase revenues and reduce possible losses in the short-term market [6]. Point out that price forecasts for energy producers and buyers are important in the short- and long-term markets because they offer better conditions to develop strategies in the bidding process [7]. Reports which forecasts are important for the efficient implementation of energy policies, for accurate predictions of energy production, are crucial in increasing demand conditions, as is the case of Turkey, for example [8].

### 1.1 Related Work

Several techniques have been used to assist distributors in this process of buying and selling electricity. Forecasting electric power demands to determine the planning and operation of the energy system of the region of Al Batinah region in Oman, was presented in [9]. Those authors proposed the use of artificial neural networks (ANN), in which they used historical data monthly from 2007, 2008 and 2009 on consumption, temperature, and humidity to forecast the consumption for 2010.

In [5] a method was demonstrated to integrate different forecasting models to analyze the pattern of electrical consumption and forecast load demand. The three-year historical data were used, and Kuwait consumption was forecasted using a moving average for seven days, which had a mean error of 30.55 MW with ASM and 0.0384 prediction for 30 days with an average error of 174.47 MW.

In [10] the authors developed load forecasting for the PJM energy company and the British energy market using historical load data with a time delay of 1 hour and 1 day in an ANN with the selection of intelligent chaotic characteristics, which is based on the reconstructed phase space theorem, and thus, a load forecast for the following days of the year 2006: 11-17 February, 8-14 June, 9-15 August and 6-12 November.

The work developed in [11] was based upon the optimization of the power acquisition on the Short-Term Market (STM) in the auctions for one year ahead (A-1) and tuning, using a genetic algorithm, seeking to minimize the cost related to the contracting of electric power, the penalties for under-contracting, and the Price for Differences Settlement (PDS). The optimal contracting was calculated using a Monte Carlo simulation to generate several consumption scenarios for a five-year period.

While [4] compared the development of the Holt-Winters method and the ANN in short and long-term forecasting of electric power demands for the State of the Philippines in the United States, where the results were compared through the MAPE, Absolute Medium Error (AME) and the root mean square error (RMSE) development indicators.

In [12] the ARIMA and ANN techniques are used to forecast the electricity consumption of Hebei Province in China, where the historical data from 1980 to 2008 were used, in which efficiency in prediction was compared using AME, RMSE, and MAPE.

A study on the scope of some forecasting techniques of time series was accomplished in [13], such as ARIMA, ANN, fuzzy, neural, wavelet, evolutionary algorithms, and hybrid algorithms that function with two or more techniques to predict consumer energy in buildings. The results showed that the use of the hybrid approaches is more efficient because if a technique does not provide everything efficiently, another technique can supply this deficiency.

In [14] an electricity forecasting method for Turkey was developed, which used historical data 1975-2013: population, Gross Domestic Product (GDP) per capita, percentage of inflation, percentage of unemployment, the average temperature in the summer, and the average temperature in the winter. They used multiple linear regressions and a multilayer network (MLP) and compared the results from 2007 to 2013 with other results in the literature.

A way to forecast by minute and hour the load for a residence was developed in [15]. For this, they used deep neural networks with two architectures of the algorithm LSTM (long-short-term memory), LSTM standard and a direct sequence-to-sequence (S2S) architecture. There are a set of samples from December 2006 to November 2010 per minute and hour, in which the first three years were used to train the model and the last year was used to test. The results showed good forecasts for the first 60 hours with a standard LSTM network just for a forecast per minute.

An ANN model was also used in [16] for pattern recognition, classification, clustering and time series forecasting, producing results with a high degree of accuracy in several areas.

In this way, this work proposes a computational approach to predicting energy consumption by class for distributors in Brazil to minimize electricity costs in the short-term market. Historical data of the income from the electricity supply were used, as well as number of units of consumption, and average tariff of power supply. So, the main contribution of this approach is its ability to accurately forecast energy consumption using only publicly available data independent of distributors' strategy.

The approach comprises a multi-layer perceptron (MLP) artificial neural network (ANN) with a backpropagation algorithm, while time series techniques use only historical consumption data from each consumer class.

The efficiency of the forecast is tested by means of the mean absolute percentage error (MAPE) for all classes of consumption of a distributor Southeast Brazil and were compared with the results of other studies with significant accuracy.

The remainder of this paper is organized as follows: Section 2 presents the rules relating to the marketing of electricity in Brazil, Section 3 presents the techniques used for prediction: artificial neural networks and techniques of time series and criteria for selecting models; Section 4 gives the methodology used in the research. The results and conclusions are presented in Sections 5 and 6, respectively.

## 2 Rules of electric power commercialization in Brazil

In order to purchase electric power in the regulated hiring environment (ACR), distributors are obliged to take part in auctions three to five years ahead (A-3, A-5) of time, call new energy auctions, where the contracts of this category may be done with a minimum term of fifteen and a maximum term of thirty-five years. There is the possibility of purchasing power in the existing energy auctions one year ahead, called A-1, and in the adjustment auctions with immediate delivery, where the contracts may be for a minimum of three and a maximum of fifteen years [17].

The procedures for fare regulation (PRORET) are used to regulate the fare processed and were approved by the Normative Resolution number 435/2011. They are organized in 12 modules, subdivided into submodules where the submodule 4.3 displays the over contracting of energy and exposure to the short-term market, detailing the mechanisms involved in verifying of the exposure of the distributor [1].

Accounting of the energy is made by the Commercialization of Electric Power Board (CEPB) as a way

of extinguishing the differences between contraction and consumption of electric power. It depends on the energy balance of the distributor (EBD), which is composed of the difference between the total consumption (TRC) and contracts signed by the distributors (PCL), all in Mega-Watt-Hour (MWh), as observed in equation 1 [18].

$$EBD = -TRC + PCL \quad (1)$$

The section Rules of energy commercialization describes the way the accounting of the energy volume traded in the short-term market is done, which depends on the product of the energy balance EBD in MWh with the price of the PDS in Brazilian Real, calculated on a weekly basis with the standard charges for each submarket (region of the country) as we may observe in equation 2 [18, 19].

$$STM = \pm EBD \times PDS \quad (2)$$

The consolidation of the total monthly result in the short-term market, TM(STM), or financial exposure of the distributor in the STM, as shown in Equation 3. This is also described in the Rules of commercialization: consolidation of the results of the CEPB, showing that the TM(STM), which is calculated monthly for each distributor, depends on the result of the sum in the STM, where the TM(STM) results may be positive, when the distributor will have more contracted energy than consumed, or negative, when there is more consumption than energy contracts. The indicators  $r$  and  $s$  represent charge and weekly standards respectively [18].

$$TM(STM) = \pm \sum_{r,s}^{n} (STM) \quad (3)$$

The CEPB counts the amount of electric power contracted, consumed and sold in the STM for each distributor, and the ascertainment of the energy bought and sold on the STM (VSTM and CSTM) is done on a monthly basis, which was exceeded in the contracting of electric energy of the distributor in the year. When the volume of energy sold in the SSTM is larger than the VSTM (SSTM > VSTM), there is an over contracting of the distributor; when the energy volume purchased in contracts is larger than the volume sold in the STM (VSTM > SSTM), there is an exposure of the distributor [20].

The energy demanded in the year by the distributor,  $ED$  (year), is booked by the sum of the monthly demanded energy, and the limit above contracting in the year,  $LAC$  (year), is calculated as 5% of the annual demanded energy, as displayed in equation 4 [20].

$$LAC(year) = 1.05 \times ED(year) \quad (4)$$

When the contracted energy by the distributor is not enough to attend to its market, the distributor may be penalized for the under contracting, purchase of energy in the short term, and by the error in the energy contracting, where the real cost of the energy not passed depends upon the contracted volume, PDS, and the annual reference value.

### 3 Techniques employed in the prediction

#### 3.1 Artificial neural networks

In recent years, artificial neural networks (ANNs) have been applied in several areas, and constitute an Artificial Intelligence (AI) tool that is incorporated in the modeling of human intelligence, that works in parallel, and whose knowledge is acquired through learning, where weighted synapses are used for knowledge. Its main characteristic is to store experimental knowledge to be used [21].

A multilayer perceptron (MLP) is an artificial neural network that can work with several non-linear decision surfaces, it uses several learning rules, and has a high degree of connectivity through the synaptic weights [21].

The Figure 1 shows the structure of the MLP network, which consists of an input layer, two hidden layers, and an output layer. The input signal propagates in a single direction, layer by layer. The input layer will receive input data from the network. The first hidden layer is fed from the input layer, which is in turn constituted by the source nodes, the outputs resulting from the first hidden layer being applied to as input to the second hidden layer. The output layer then has the function of processing the information [14].

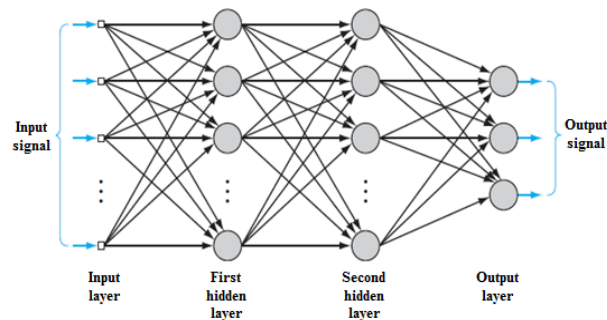


Figure 1. MLP network with a structure of two hidden layers Source: Haykin, 2008, p. 124.

A function widely used to determine the values of change of each weight is the logistic function, which defines the output of a neuron in terms of  $v$ : when  $v$  is greater than or equal to 0,  $\phi(v)$  will be equal to 1; if  $v$  is less than 0, then  $\phi(v)$  will be 0 and  $v$  represents the sum of the inputs plus the activation threshold. The output signal is represented by the sum of the product of the input signals, by the synapse-toxic weight, plus the bias.

To minimize the error of training of the network, a backpropagation algorithm can be implemented, which works in two phases: the feed-forward phase, in which the synaptic weights are fixed and the input signal is propagated through the network, layer by layer up to the output, and the backward phase, in which an error signal is produced, comparing an output of the network with the desired response. This has propagated back through the network, layer by layer, resulting in continuous adjustments of the synaptic weights [21].

### 3.2 Time series techniques and criterion for selection of models

A time series consists of any set of observations organized in time. The essential characteristic of a time series is that observations near each other are dependent. The analysis of this dependence requires the use of specific techniques, for this, there is a need to develop models [22] [23]. Stochastic models include both stationary and non-stationary models. In stationary models, it is assumed that the process remains in statistical equilibrium with probabilistic properties that do not change over time, varying around a fixed constant mean level and with constant variance, while in non-stationary models, such characteristics of the data change over time [22]. A forecast is related to the prediction of some events and is used in areas such as business, industry, government, economics, environmental sciences, medicine, social sciences, politics, and finance, among others. Forecasting problems can be classified into short-term, that involve forecasts of just a few periods of time (days, weeks, months) in the future, forecasts in the medium term, which extend to a few months, and long-term forecasts, which are related to forecasts over one year [24].

There are several methods used to make predictions: some work with trend, seasonality, as well as the autocorrelation function (ACF), which is a way to measure and explain the association between observations in a time series. There are methods that seek to explore the autocorrelation structure of the prediction residuals of the current period with those occurring in previous periods, among other characteristics [24]. Time series techniques are chosen by the model selection criteria that seek to select models that are closer to reality.

The Akaike Information Criterion (AIC) was created by the Japanese statistician Hirotugu Akaike in 1973 [25], based on information theory. It works by minimizing the Kullback-Leibler distance (K-L), which is a measure of the distance between the true model and a selected model. It is used for the selection of a model. According to [23], the smaller the AIC value the better the model fit.

Equation 5 illustrates the calculation of the AIC, which depends on the maximum likelihood estimator  $\hat{\sigma}_p^2$  (compatible with the available information) and  $p$ , the number of parameters:

$$AIC = -2\log\hat{\sigma}_p^2 + 2p \tag{5}$$

In 1978, the Bayesian information criterion (BIC) was developed [26], which maximizes the probability of identifying the true model among the evaluated models and functions similarly to the AIC: the lower the value the better the model fit. Equation 6 shows the development of this model, using the maximum likelihood estimator, in which  $p$  the number of parameters to be estimated and  $m$  is the number of sample points:

$$BIC = -2\log\hat{\sigma}_k^2 + p\log m \tag{6}$$

The best models are selected by lower values of the AIC and BIC. These criteria add a penalty that increases with the number of regressors. As the penalty is higher for the BIC, this criterion tends to favor more parsimonious models than AIC. The idea is to use these criteria is to find a parsimonious model that describes the dynamic process of a variable [27].

#### 4 Methodology

This work proposes the use of simple and publicly available data such as the historical power supply revenue data (RFE), number of consumer units (UC) and average tariff of power supply (TMF) in the MLP with a backpropagation algorithm to make a forecast of demand for consumer class.

Since the efficiency of predictions RNA and time series techniques for commercial, own consumption, rural, public lighting, public power, residential, industrial, water, sewage and sanitation, and Electric traction consumption classes are made by the MAPE performance index.

As reported in [5], in developing countries the electricity demand grows with a dynamic and high growth rate. Therefore, each country is indifferent to factors that affect the demand for electricity. On the other hand, [6] describes that the effect of temperature and other climatic variables are incorporated in the demand. Thus, we chose to use financial variables (income and energy supply tariff) and number of consumer units to forecast electricity.

Thus, this work follows two approaches for predicting energy consumption by consumer class as can be seen in Figure 2. The first uses time series techniques and the second is based on artificial neural network. The first approach has used a history of electricity consumption from January 2003 to June 2016 to provide 12 months for each consumer class. AR, MA, ARMA, ARCH, ARIS and GARCH.

The data is transformed into stationary and then is made to forecast electricity consumption by class for the next 12 months through the AIC and BIC selection criteria.

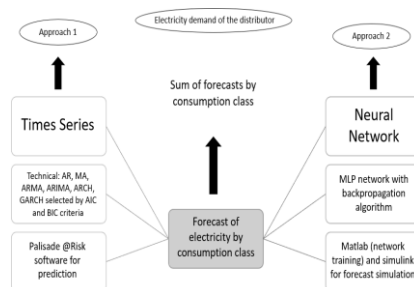


Figure 2. Approach to energy consumption forecast.

In [7] is described that the selection of the type and number of input variables is an important issue to find the best ANN architecture. So, if some variables that have no impact on the variable output are included into the model, it can suffer significant loss performance, increase training time, increase network complexity, and reduce network runtime.

The ANN used in this work used the same network architecture proposed by [28], with one input layer with three neurons, two hidden layers and the output layer with one neuron.

In this work we have trained five networks for each consumer class, with the input layer with the variables income, number of consumer units and average tariff of electricity supply. The first hidden layer has 20, 30, 40, 50, 60 neurons in the second hidden layer has 10, 20, 30, 40, 50 neurons.

The sigmoid function has been selected to process the signal generated by combining the inputs and the synaptic weights, thus generate the neuron output signal. The forecast electricity consumption is performed with the input variables at time  $t$  as shown by equation 6, where  $ic$  corresponds to energy supply revenue,  $nc$ , number of consuming units,  $mt$  average energy supply tariff, and  $d(t)$ , energy demand by consumption class.

$$d(t) = f((ic_t, nc_t, mt_t, \dots, ic_{tn}, nc_t, mt_{tn})) \tag{6}$$

Matlab toolbox 2016 was used for network training and the capabilities of Simulink were used for construction and simulation of the dynamic systems. Through the ‘gensim’ command (net-1), ANNs were created, customized for each class of consumption, and its Simulink model was constructed, as shown in Figure 3.

Note that the input variables were entered on a monthly basis in the model, and before sending the data to the ANN, it was necessary to standardize the input data to work with values between 0 and 1. With these normalized values, the already trained network receives the input data and generalizes from parameters that have been previously described. Denormalization is necessary to return the expected demand values.

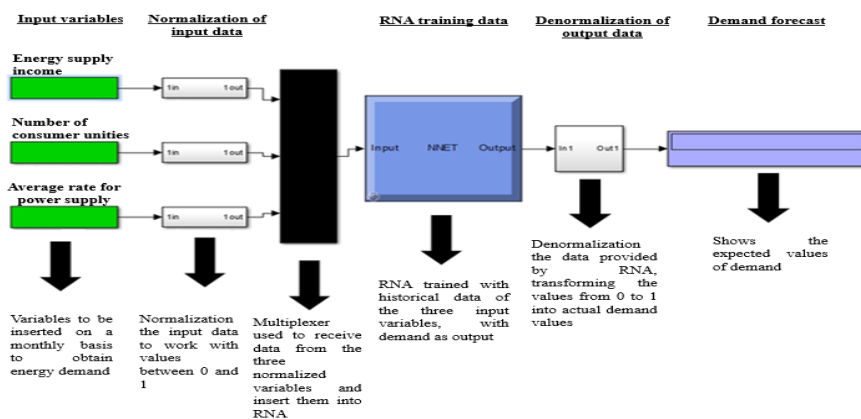


Figure 3. Simulink consumption classes electricity forecasting structure.

## 5 Results and discussion

This section shows all results obtained in the training and forecasting of electricity of all consumption classes of the distributor *AES Eletropaulo*, which distributes energy in the city of São Paulo, Brazil. It is possible to see the result of the network training (MSE and training time of the network), the performance of the electric energy forecast for each set of neurons used in the work. Subsequently, it shows the results obtained with techniques of the time series, which have been selected with the criterion AIC BIC. The efficiency of forecasting with time series techniques is demonstrated by the MAPE performance index.

Table 1 shows the performance of the ANN obtained by the training set, represented by the MSE and the forecast of the five network configurations for all classes of consumption of the distributor, represented by the MAPE index. The highlighted values represent the smallest values of the MAPE and MSE for each case.

Table 2 shows the five neuronal architectures used in the ANNs. One can observe that with an increasing number of neurons, there is an increase in computational time. Because of the different values for different classes of consumption, the different consumption classes require different training times.

In accordance with the best MAPE scores in Table 1 the best neural network architecture is different for each consumption class. It means there is not a general neural model for all classes. This represent an important result since the different architectures regarding the number of neurons converged in a different

time, as can be seen in Table 2.

Regarding the accuracy of the electricity forecast, ANN presented the best performance in terms of the MAPE for the nine consumer classes. Figure 4 presents a comparison of current consumption and expected monthly performance index with the MAPE from July 2016 to June 2017, in which it is possible to see that the MAPE index does not exceed 1.24%.

Table 1. Development and training of the ANN.

Consumption Class / N° of neurons	Performance index of training & forecast of ANN									
	20   10   1		30   20   1		40   30   1		50   40   1		60   50   1	
	MSE	MAPE (%)	MSE	MAPE (%)	MSE	MAPE (%)	MSE	MAPE (%)	MSE	MAPE (%)
Commercial	1.63E-06	0.188	5.57E-07	0.098	<b>9.98e-10</b>	<b>0.0282</b>	4.18E-07	0.095	4.98E-08	0.121
Own consumption	9.44E-06	34.793	7.54E-06	40.506	3.23E-06	33.381	1.60E-06	<b>5.85</b>	<b>5.81e-7</b>	26.830
Rural	2.10E-05	12.210	2.62E-05	<b>11.29</b>	5.50E-06	342.170	<b>3.97e-6</b>	417.356	3.25E-05	126.627
Public Lighting	4.37E-05	<b>0.50</b>	7.57E-06	2.732	1.89E-05	8.278	<b>6.37E-6</b>	0,906	1.86E-05	5.088
Public Power	1.15E-06	<b>1.57</b>	<b>6.82e-8</b>	2.339	1.15E-08	2.009	4.51E-06	8,77	1.75E-06	8.677
Residential	5.24E-06	13.415	7.48E-08	<b>0.028</b>	7.55E-08	0.108	4.20E-08	0,33	<b>4.23e-8</b>	0.260
Industrial	6.32E-08	6.555	1.40E-08	13.494	2.23E-08	<b>6.45</b>	8.31E-09	12,175	<b>9.99e-10</b>	11.864
Water, sewage and sanitation	3.25E-05	<b>0.47</b>	<b>2.85e-10</b>	0.693	4.54E-05	1.476	1.59E-08	9,005	1.19E-06	8.363
Electric traction	4.25E-09	<b>0.25</b>	9.88E-10	0.466	9.96E-10	2.339	7.31E-07	26,85	<b>9.9e-10</b>	35.544

Table 2. Training time for different ANN architecture.

Classes of consumption	Time for training the ANN				
	20   10   1	30   20   1	40   30   1	50   40   1	60   50   1
Commercial	13s	57s	3 min 2s	11 min 52s	20min 52s
Own consumption	12s	51s	3 min 44s	9 min 26s	22min 24s
Rural	12s	50s	3 min 26s	9 min 5s	21min 31s
Public Lighting	12s	52s	3 min 20s	17 min 44s	22min 57s
Public Power	13s	37s	2 min 48s	5 min 44s	28min 4s
Residential	12s	49s	2 min 55s	8 min 34s	34min 25s
Industrial	12s	53s	3 min 1s	8 min 25s	22min 58s
Water, sewage and sanitation	12s	57s	3 min 57s	9 min 45s	27min 2s
Electric traction	14s	1 min 2s	2 min 42s	10 min 2s	23min 4s

The use of revenue variables, the number of consumer units and average tariff of power supply were enough for the artificial neural network to learn the behavior of the data and generalize from them. This solves the problem mentioned by [7] regarding the use of temperature and its quick changes as the input variables in the neural network.

Corresponding forecast results for time series techniques selected by the AIC criterion are presented in Figure 5. MAPE index reaches more than 10% in some cases, in the months of October 2016 and May and June 2017.

So, ANN produces better results regarding time series techniques in the sense of MAPE index, in contrast to [26] in which the authors present an application of ARMA technical process with better accuracy in demand for load forecasting in Greece compared with ANN.



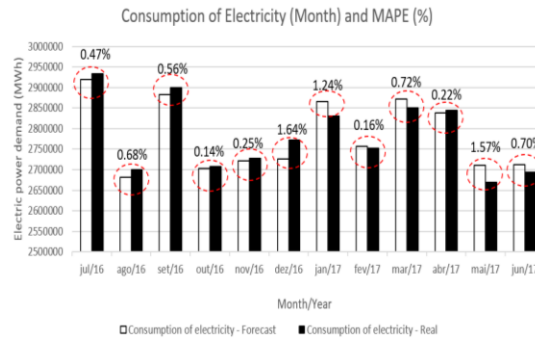


Figure 4. Power consumed target and forecast by ANN and MAPE index.

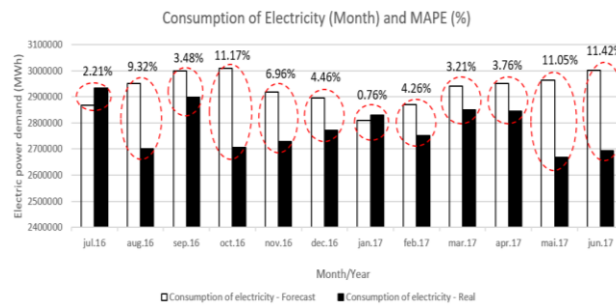


Figure 5. Power consumed target and forecast with time series (MWh) and MAPE index.

In the development of the experiments, it was observed that by increasing the amount of network neurons, there is an increase in the computational time and there is no guarantee that the increase in the number of neurons improves efficiency in forecasting power, as noted in Tables 1 and 2, and summarized in Table 3, through the MAPE index and computational time.

Regarding the time series techniques, it was observed that the techniques chosen by the AIC and BIC criteria have not changed substantially in the choice of technique since the number of samples used for the prediction was 162 samples. The number of samples affects the choice of the best technical and BIC model is more efficient than the AIC when there is a considerable number of samples (over 500 samples) [29].

Table 3. Best results regarding MSE, MAPLE and ANN architecture.

Classes of consumption	MSE (ANN training)	MAPE (%)	ANN training time	Best ANN architecture (Neuron)
Commercial	9.98e-10	0.02	3m2s	40   30   1
Own consumption	5.81e-7	5.85	9m26s	50   40   1
Rural	3.97e-6	11.29	50s	30   20   1
Public Lighting	6.37E-6	0.50	12s	20   10   1
Public Power	6.82e-8	1.57	13s	20   10   1
Residential	4.23e-8	0.02	49s	30   20   1
Industrial	9.99e-10	6.45	3m1s	40   30   1
Water, sewage and sanitation	2.85e-10	0.47	12s	20   10   1
Electric traction	9.9e-10	0.25	14s	20   10   1

## 6 Conclusion

This work demonstrated a computational approach to predict and minimize electricity costs in the short-term market for electricity distributors in Brazil, which used artificial neural networks and time series techniques. The results show that an electricity distributor can use historical data of the number of consumer units, revenue and electricity supply tariff to predict its demand.

This makes the proposed computational approach used to forecast electricity of great value to the

electricity distributors in Brazil, since these variables are in the public domain and are not subject to large variations that can make accurate prediction difficult. The efficiency of monthly forecasting produced an average error of less than 0.7% regarding the annual energy demand contracted, which would result in no penalty for this distributor. Thus, it is expected that the distribution decreases the cost of electricity trading in the spot market, reduces the losses to subcontracting or due to over contracting, and gets healthier recipes.

## 7 References

- [1] ANEEL: Normative Resolution No. 703, from March 28, 2016. Structure of the Taxes Regulation Procedures – PRORET. Available at at: <[http://www2.aneel.gov.br/cedoc/aren2016703\\_Proret\\_Submod\\_4\\_3\\_V0.pdf](http://www2.aneel.gov.br/cedoc/aren2016703_Proret_Submod_4_3_V0.pdf)>. Accessed Feb. 2, 2016.
- [2] SRE - REPORT IV: Conceptual aspects of economic regulation, tariff model and training facility of rates in 25 countries studied. Available at <[https://www.cpfl.com.br/energiassustentaveis/inovacao/projetos/Documents/PB302/aspectos\\_conceituais.pdf](https://www.cpfl.com.br/energiassustentaveis/inovacao/projetos/Documents/PB302/aspectos_conceituais.pdf)>. Accessed Dec. 11, 2016.
- [3] Hong, T., Fan, S.: *Probabilistic electric load forecasting: A tutorial review*. Int. J. Forecasting, 32 (2016), 3, 914–938.
- [4] Bantugon, M. J. T., Gallano, R. J. C.: *Short- and long-term electricity load forecasting using classical and neural network-based approach: A case study for the Philippines*, IEEE Region 10 Conference (TENCON), Singapore, 2016, 3822-3825.
- [5] Almeshai, E., Soltan, H.: *A methodology for Electric Power Load Forecasting*. Alexandria Engineering Journal 50 (2011), 137–144.
- [6] Catalao, J., Mariano, S., Mendes, V., Ferreira, L.: *Short-term electricity prices forecasting in a competitive market: A neural network approach*, Electric Power Systems Research, 77 (2007), 1297-1304.
- [7] Sarada, K., Bapiraju, V.: *Comparison of day-ahead price forecasting in energy market using Neural Network and Genetic Algorithm*, Int. Conf. on Smart Electric Grid (ISEG), Guntur, 2014, 1-5.
- [8] Kankal, M., Uzlu, E.: *Neural network approach with teaching–learning-based optimization for modeling and forecasting long-term electric energy demand in Turkey*. Neural Computing and Applications, 28 (2017), 1, 737–747.
- [9] Swaroop, R.; Hussein, A.: *Load forecasting for power system planning and operation using artificial neural network at al batinah region oman*. Journal of Engineering Science and Technology, 7 (2012), 4, 498-504.
- [10] Kouhi, S., Keynia, F., Ravadanegh, S.: *A new short-term load forecast method based on neuro-evolutionary algorithm and chaotic feature selection*. Int. J. Electr. Power Energy Syst., 62 (2014), 862–867.
- [11] Lazo, J. G. L., Figureueiredo, K., Lima, D. A.: *Analysis of an intelligent module of electric power contracting in short term for distributors*. Automation & Control (Rio de Janeiro), 23 (2012), 6, 711-725.
- [12] Wang, L., Zhang, Z., Chen, J.: *Short-Term Electricity Price Forecasting with Stacked Denoising Autoencoders*. IEEE Transactions on Power Systems. 32 (2016), 2673-2681.
- [13] Deb, C., Zhang, F., Yang, J., Lee, S., Shah, K.: *A review on time series forecasting techniques for building energy consumption*. Renewable and Sustainable Energy Reviews, 74 (2017), 902-924.
- [14] Gunay, M.E.: *Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: case of Turkey*. Energy Policy, 90 (2016), 92–101.
- [15] Marino, D., Amarasinghe, K., Manic, M.: *Building Energy Load Forecasting using Deep Neural Networks*. 42nd Annual Conference of the IEEE Industrial Electronics Society, Florence, 2016, 23-26.
- [16] Khashei, M., Bijari, M.: *An artificial neural network (p, d, q) model for time-series forecasting*. Expert Syst. Appl., 37 (2010), 479-489.
- [17] Electric Energy Commercialization Rules – CEPB(c), (2016) version. Contracts. Available at:

- <<http://www.CEPB.org.br>>. Accessed on June 20, 2016.
- [18] Electric Energy Commercialization Rules – CEPB(a), (2012) version. Energy Balance. Available at: <<http://www.CEPB.org.br>>. Accessed on June 20, 2016.
  - [19] Electric Energy Commercialization Rules – CEPB(d), (2016) version. Price of the difference liquidation. Available at: <<http://www.CEPB.org.br>>. Accessed on June 20, 2016.
  - [20] Electric Energy Commercialization Rules – CEPB(b), (2016) version. Results Consolidation. Available at: <<http://www.CEPB.org.br>>. Accessed on June 20, 2016.
  - [21] Haykin, S.: *Neural Networks: A Comprehensive Foundation*. Pearson, Toronto, 1999.
  - [22] Box, G., Jenkins, G.: *Time Series Analysis: Forecasting and Control*, Holden-Day, New Jersey, 2015.
  - [23] Morettin, P. A., Toloí, C. M. C.: *Time Series Analysis* (in Portuguese). 2<sup>nd</sup> ed. Edgar Blücher, São Paulo, 2006.
  - [24] Montgomery, D. C., Jennings, C. L., Kulahci, M.: *Introduction to Time Series Analysis and Forecasting*, 2<sup>nd</sup> ed. John Wiley & Sons Inc, New York, 2015.
  - [25] Akaike, H.: *A new look at the statistical model identification*. IEEE Transactions on Automatic Control 19 (1974), 6, 716-723.
  - [26] Schwarz, G.: *Estimating the dimension of a model*. Annals of Statistics, 6 (1978), 2, 461-464.
  - [27] Verbeek, M.: *A Guide to Modern Econometrics*. 2nd edition, John Wiley & Sons Inc., 31, 285–8, 2004.
  - [28] Osman, Z., Awad, M., Mahmoud T.: *Neural network based approach for short-term load forecasting*. IEEE/PES Power Systems Conference and Exposition, Seattle, 2009, 15-18.
  - [29] Ding, J., Tarokh, V., Yang, Y.: *Bridging AIC and BIC: A New Criterion for Autoregression*, IEEE Transactions on Information Theory, 64 (2018), 6, 4024 - 4043.