Available online at www.sciencedirect.com
ScienceDirect
Energy Reports xxx (xxxx) xxx
www.elsevier.com/locate/egyr

6th International Conference on Energy and Environment Research, ICEER 2019, 22–25 July, University of Aveiro, Portugal

Forecasting residential electric power consumption for Bogotá Colombia using regression models

Carlos Peña-Guzmán*, Juliana Rey

Environmental Engineering Program, Santo Tomas University, Cr 9 # 51-11, Bogotá, 110231, Colombia

Received 21 July 2019; accepted 12 September 2019
Available online xxxx

Abstract

This study presents three models of multiple linear regressions for forecasting energy demand. The first is a simple multiple linear regression, the second model has an economic interpretation of coefficients (econometric), while the third model is developed in the form of double logarithm economic regression. The article was developed based on the six socio-economic strata in Bogotá City. The second model is shown to be superior to the multiple linear regression model with a climatic approach and the econometric model of double logarithm in terms of precision in the calculation of the electric energy demand, as evidenced in the model evaluation tools used, such as the coefficient of determination, with values higher than 0.9 except in stratum 5.

Peer-review under responsibility of the scientific committee of the 6th International Conference on Energy and Environment Research, ICEER 2019.

Keywords: Energy demand; Forecasting; Multiple linear regressions

1. Introduction

The consumption of electrical energy has grown exponentially over the years, becoming the public service with the greatest coverage at a global level [1]. This is as a consequence of the changes to which the world is constantly exposed, such as population growth, climatic conditions and economic development [2].

Currently, the world's energy supply systems are at a turning point, which is why different countries are investing resources in energy studies, estates and tools, to make sound decisions in this field [3], with the prognosis of energy consumption representing one of the most used tools. This is associated with the fact that these models allow for the prediction and identification of trends in consumption behaviors, thus generating proposals for possible future scenarios to different trend situations, because they are based on data recorded in the past [4]. The use of these models has been carried out in different parts of the world, using regression models, multivariate models [4], economics models [5], or based on the use of machine learning [6,7].

This article presents the behavior of three energy consumption forecast models in the city of Bogotá, using three models of multiple regressions. The article was developed based on the six socio-economic strata existing in the city.

* Corresponding author.
E-mail address: carlos.pena@usantotomas.edu.co (C. Peña-Guzmán).

https://doi.org/10.1016/j.egyr.2019.09.026

Peer-review under responsibility of the scientific committee of the 6th International Conference on Energy and Environment Research, ICEER 2019.

2. Material and methods

2.1. Geographic location

Bogotá is the capital of Colombia and has had an approximate population growth rate of 1.31% in recent years. As a capital city, it has the largest gross domestic product (GDP) in the country, contributing most of the total national GDP (24.5%) for 2016, and has been characterized by having the highest levels of electricity service coverage in Colombia, with percentages greater than 97% in the last 15 years, reaching in 2011 at 100% coverage.

Within the city, the residential sector has the highest consumption in terms of electricity, given that in 2015 there was a consumption of 2.82 GWh, which represents ∼42% of the total consumption of electricity. The residential sector is divided into six strata (Stratum 1: low–low, Stratum 2: low, Stratum 3: medium–low, Stratum 4: medium, Stratum 5: medium–high and Stratum 6: high).

2.2. Models

Three types of regression models were selected, namely, a multiple regression model, a multiple econometric regression model and a linear regression model of double logarithm. For the econometric model, the interpretation of the coefficients changes, converting them into elasticity indicators with respect to energy demand. Therefore, this estimate is now partially effective, converting the coefficients into the elasticities of these variables with respect to the response variable. That is to say that when X varies in a unit, the remaining constant variables, Y varies 1 in average in β units. For the double-regression model, the formulation of this was based on a linear-type specification in logarithms.

For the validation of the models, the coefficient of determination (R^2) was used as a performance metric and a significance of 0.05 was used for the evaluated coefficients.

2.3. Variables

The following variables were selected for the models by stratum: electricity consumption, end users of the residential sector, total energy billing, total natural gas billing, revenue per user, GDP and surface temperature. These variables were obtained for a record of 11 years, from the month of January 2005 to the month of December 2016. The Bogota's annual average temperature is 13 °C, and does not have a significant impact on energy demand.

3. Results

Regarding energy consumption, Fig. 1 shows the historical multi-year consumption of electricity by stratum between the years 2005 and 2016, where we can observe that for strata 1–4, there is an annual growth in consumption. In contrast, for stratum 5 since 2012, there has been a reduction in consumption. As for stratum 6, this reduction occurs in 2013. Consumption for stratum 3 presents the maximum consumption values, with values between 2.37E10 and 5.63E10 kWh. While for the period between 2005 and 2011, stratum 1 presents the lowest consumption values, which are in a range between 1.45E7 and 1.74E7 kWh.

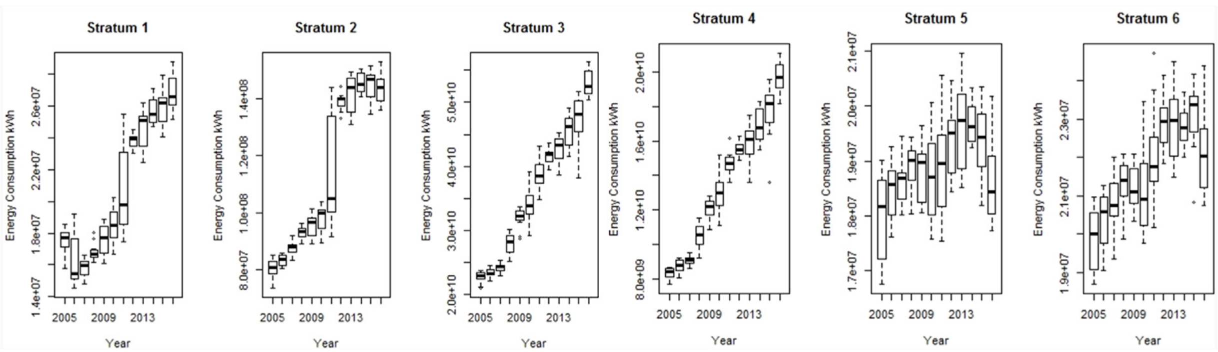

Fig. 1. Energy consumption per stratum.

The number of users presents an accelerated growth in strata 1–3 as of 2010, whereas after this year, the growth decreases in proportion. However, strata 2 and 3 have the highest number of users, as shown in Fig. 2. This is due to the fact that they represent the largest distribution of the population in Bogotá with values of 41.3% and 36%. In contrast, stratum 6 is the one with the lowest number of users with a minimum of 51,616 end users.

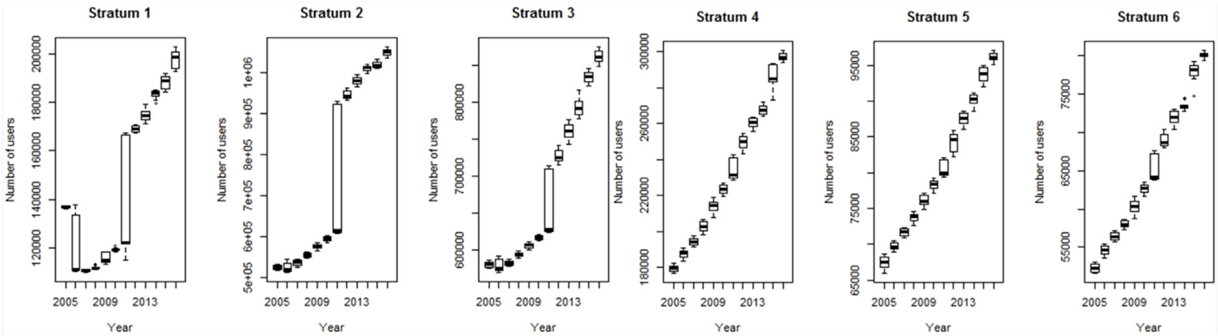

Fig. 2. Number of users per stratum.

For the payment of the service, in Fig. 3, the final prices are shown without taking subsidies into account. The behavior of this variable increases and the same level of growth is observed for all strata. Strata 5 and 6 have a higher turnover than the other strata, given that they do not have a subsidy and their purchasing power is higher, therefore higher rates are applied.

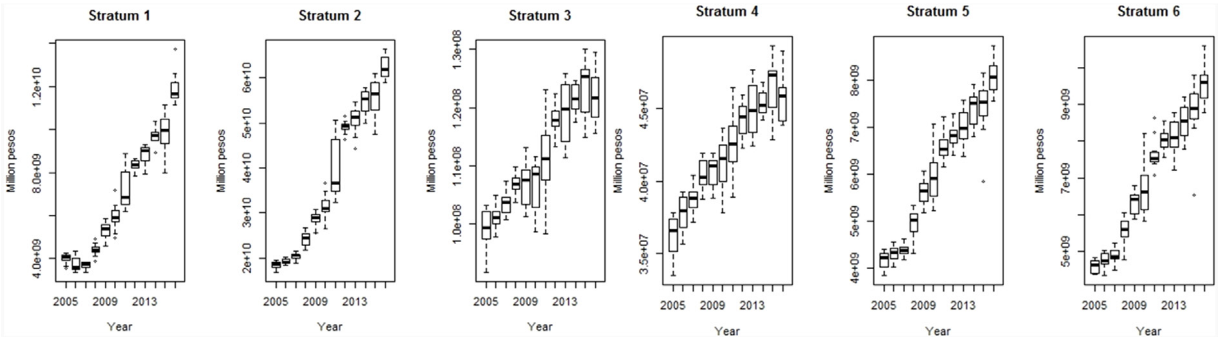

Fig. 3. Total billed per stratum.

The multi-year monthly GDP from 2005 to 2016 is shown in Fig. 4a. This variable increases over time due to the economic growth that the city has. Fig. 4b shows the evolution of per capita income per end user, which

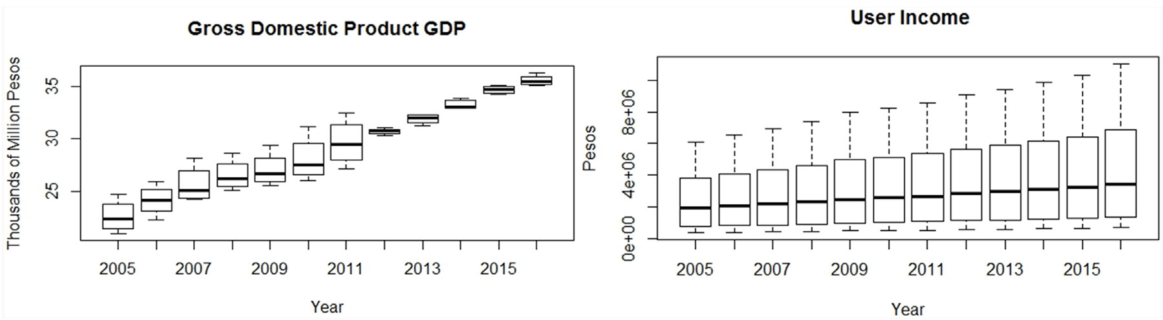

Fig. 4. (a) GDP; (b) user income.

maintains a constant average with a marginal increase over time. The values above the median correspond to the users of strata 5 and 6 who present between 10 and 12 minimum wages in the past due year.

3.1. Forecast models

The first model developed was the multiple regressions by stratum. For this, a stepwise regression model was proposed, in order to determine the variables that best explain the energy consumption. From this process, the variables found were the final energy users, the surface temperature and the price of electric power. The equation is shown below and the values of the coefficients and their significance are shown in Table 1.

$$Energy\,consumption = B_o + B_1 Users + B_2 Temperature + B_3 Price \tag{1}$$

Table 1. Coefficients and evaluation measures by variable of the multiple regression model.

Stratum	Coefficients				R^2	*p*-value	Significance			
	B_0	B_1	B_2	B_3			B_0	B_1	B_2	B_3
1	5.87E+06	5.55E+01	−3.22E+04	1.07E−03	0.95	<0.05	<0.05	<0.05	<0.05	<0.05
2	3.71E+07	5.57E+01	−3.77E+05	1.06E−03	0.97	<0.05	<0.05	<0.05	<0.05	<0.05
3	−6.72E+10	7.53E+04	4.82E+08	4.37E+10	0.88	<0.05	<0.05	<0.05	<0.05	<0.05
4	−1.64E+10	8.66E+04	6.36E+07	2.24E+07	0.95	<0.05	<0.05	<0.05	<0.05	<0.05
5	2.03E+07	−7.28E+01	−5.00E+04	8.19E−04	0.74	<0.05	<0.05	<0.05	<0.05	<0.05
6	205E+07	−842E+01	−9.21E+04	1.10E+00	0.93	<0.05	<0.05	<0.05	<0.05	<0.05

For the price variable, the lower weights of these coefficients are for strata 1, 2, 5 and 6. For strata 1 and 2, as they have government subsidies, the price of energy it is not that important, and for strata 5 and 6, their purchasing power explains that the value of energy is indifferent to its consumption. In the same way, in the analysis of the values of the coefficients, we can deduce some behaviors related to the strata, for example, in strata 3 and 4, the variable of the price of energy has a high value, which means that for this socioeconomic group the consumption of energy may depend on the price of it. This is because stratum 4 does not have any subsidy on its invoice, nor does it make contributions to the other strata. Likewise, stratum 3 is the stratum that has the lowest subsidy, so the price of this service influences consumption. For strata 5 and 6, the price of energy is indifferent, the same happens for strata 1 and 2.

The second model was based on multiple econometric regressions. For the calculation of the regressions, the following were taken into account as economic explanatory variables: the price of electricity, the gross domestic product and end users. The results are presented below:

$$Energy\,consumption = B_o + B_1 GDP + B_2 Price + B_3 Users \tag{2}$$

$$\beta_{1=} EQc = Elasticity\,GDP = \frac{(1+\frac{\Delta Q}{Q})}{(1+\frac{\Delta GDP}{GDP})} \tag{3}$$

$$\beta_{2=} EQc = Elasticity\,Price = \frac{(1+\frac{\Delta Q}{Q})}{(1+\frac{\Delta P}{P})} \tag{4}$$

$$\beta_{3=} EQc = Elasticity\,Users = \frac{(1+\frac{\Delta Q}{Q})}{(1+\frac{\Delta U}{U})} \tag{5}$$

where $\frac{\Delta Q}{Q}$ represents the variation of energy demand, Q is the initial value of the demand, $\frac{\Delta GDP}{GDP}$ represents the variation of income measured in capital by GDP and $\frac{\Delta P}{P}$ represents the variation of the price with respect to the demand for energy, with P being the initial price.

Table 2 shows the elasticity values that result from the model, reiterating that for strata 1, 2, 5 and 6, the demand for the energy price is perfectly inelastic, since its values are very close to zero. That is, an increase in the price leaves the quantity demanded without any change. In contrast, in strata 3 and 4, if a decrease in demand were to occur, the value of the price of energy would increase. This is due to climatic phenomena, which modified the amount of energy offered, that is, the demand was greater than the supply, that is why there is an increase in price.

Table 2. Coefficients and evaluation measures by variable of the econometric regression model.

Stratum	Coefficients				R^2	p-value	Significance			
	B_0	B_1	B_2	B_3			B_0	B_1	B_2	B_3
1	5.20E+06	5.23E+03	1.00E−03	5.87E+01	0.95	<0.05	<0.05	<0.05	<0.05	<0.05
2	1.77E+07	5.58E+05	7.00E−03	7.32E+01	0.97	<0.05	<0.05	<0.05	<0.05	<0.05
3	4.67E+10	1.08E+09	1.59E+02	5.07E+04	0.88	<0.05	<0.05	<0.05	<0.05	<0.05
4	−1.51E+10	1.04E+07	1.04E+07	1.04E+07	0.95	<0.05	<0.05	0.84	<0.05	<0.05
5	1.86E+07	2.03E+05	4.00E−03	−1.03E+02	0.76	<0.05	<0.05	<0.05	<0.05	<0.05
6	1.67E+07	276E+05	6.00E−03	1.15E+02	0.80	<0.05	<0.05	<0.05	<0.05	<0.05

Alternatively, the regression gives values that indicate the utility of the model, given that it has higher coefficients of determination of 0.9, except stratum 5, as in the previous model. These values of R^2 are close to 1, indicating that the relationships of the economic variables with respect to energy demand are important. For strata 1 and 2, the purchasing power variable, such as GDP, is not significant; on the contrary, the other variables show significance.

As for the economic model of double logarithms for the calculation of mathematical parameters, explanatory variables were taken into account, such as the price of electricity, the per capita income of users, the price of natural gas and end users. The results are described below:

$$Energy\, consumption = \beta_0 * Price_1^{\beta 1} * Income_2^{\beta 2} * Price\, Natural\, Gas_3^{\beta 3} * Users_4^{\beta 4} * \exp(\varepsilon) \tag{6}$$

$$\beta_{1=} EQc = Elasticity\, Price = \frac{(1+\frac{\Delta Q}{Q})}{(1+\frac{\Delta P}{P})} \tag{7}$$

$$\beta_{2=} EQc = Elasticity\, Income = \frac{(1+\frac{\Delta Q}{Q})}{(1+\frac{\Delta Ing}{Ing})} \tag{8}$$

$$\beta_{3=} EQc = Elasticity\, Natural\, Gas = \frac{(1+\frac{\Delta Q}{Q})}{(1+\frac{\Delta Pg}{Pg})} \tag{9}$$

$$\beta_{4=} EQc = Elasticity\, Users = \frac{(1+\frac{\Delta Q}{Q})}{(1+\frac{\Delta U}{U})} \tag{10}$$

Table 3 shows how the cross elasticity that was calculated based on a natural gas, has a value less than 1, meaning it is inelastic. It can also be identified, as the price behaves inelastic for all strata, this happens in luxury goods or necessary goods, such as energy service, which is an indifferent change in demand if there is an increase in the price. In the elasticity of income for demand, it stands out that the only stratum that presented a positive elasticity was stratum 3, which associates demand with the level of income. Since an increase in this variable will cause an increase in the energy demand for this layer, also showing a spring value in the relationship of user demand, this indicates that this layer is the one with the highest ratio between these two variables. This is because it is the stratum with the highest percentage of occupation, meaning that the number of users positively affects its demand.

Table 3. Coefficients and evaluation measures by variable of the logarithmic regression model.

Stratum	Coefficients					R^2	p-value	Significance				
	B_0	B_1	B_2	B_3	B_4			B_0	B_1	B_2	B_3	B_4
1	2.69	0.39	−0.250	0.031	0.376	0.95	<0.05	<0.05	<0.05	<0.05	0.284	<0.05
2	2.77	0.232	−0.74	0.107	0.478	0.97	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05
3	−4.28	0633	1.154	0.04	0.353	0.94	<0.05	<0.05	<0.05	<0.05	0.435	<0.05
4	−3.76	0.658	0.289	0.023	1.263	0.95	<0.05	<0.05	<0.05	<0.05	0.586	<0.05
5	5.91	0.182	−0.525	0.031	0.035	0.74	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05
6	6.41	0.211	−0.671	0.049	0.515	0.79	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05

Additionally, it can be seen how the R^2 values are found to be higher than 0.9 in all strata except for strata 5 and 6, where the independent variables explain the model in 70%. This analysis of significance helps to see how the price of natural gas does not influence the consumption of electricity, given that for this variable, the probability

Please cite this article as: C. Peña-Guzmán and J. Rey, Forecasting residential electric power consumption for Bogotá Colombia using regression models. Energy Reports (2019), https://doi.org/10.1016/j.egyr.2019.09.026.

values were higher than 0.05. This is how it is verified once again that it is a complementary good, but not a substitute.

4. Conclusions

The users variable explains the demand for electric power, since they are the main players in it. This is why the strata with the highest occupancy have the highest energy consumption.

All regression models have a good accuracy in the calculation of electric power demand, as evidenced by the model evaluation tools used, such as the coefficient of determination, with values higher than 0.9 except in stratum 5.

It is important to emphasize that the econometric model is preferred since it has quantitative variables that can partially explain qualitative variables, such as the gross domestic product, which thanks to the analysis of the purchasing power of the users can analyze their consumption hypothetically based on the models of household appliances that they have in their homes. As well as the inelasticity of the demand of energy with respect to the price of it in strata where there are subsidies or high purchasing power, such as the group of strata 1–2 and 5–6. Consequently, it is important to mention the importance of analyzing the behavior of demand by grouping the data by socioeconomic strata, showing how subsidies and contributions make this demand inelastic over time.

References

[1] World Bank (2015) Sustainable Energy for All 2015: Progress Toward Sustainable Energy. The World Bank, Washington DC.
[2] Pablo-Romero Mdel P, De Jesús J. Economic growth and energy consumption: The energy-environmental kuznets curve for latin america and the Caribbean. Renew Sustain Energy Rev 2016;60:1343–50.
[3] Goldemberg J. World energy assessment: energy and the challenge of sustainability. New York: United Nations Development Programme; 2000.
[4] Suganthi L, Samuel A. Energy models for demand forecasting—A review. Renew Sustain Energy Rev 2012;16:1223–40.
[5] Parajuli R, Østergaard P, Dalgaard T, Pokharel G. Energy consumption projection of nepal: An econometric approach. Renew Energy 2014;63:432–44.
[6] Ahmad A, Hassan M, Abdullah M, Rahman H, Hussin F, Abdullah H, Saidur R. A review on applications of ANN and SVM for building electrical energy consumption forecasting. Renew Sustain Energy Rev 2014;33:102–9.
[7] Deb C, Zhang F, Yang J, Lee SE, Shah KW. A review on time series forecasting techniques for building energy consumption. Renew Sustain Energy Rev 2017;74:902–24.

Please cite this article as: C. Peña-Guzmán and J. Rey, Forecasting residential electric power consumption for Bogotá Colombia using regression models. Energy Reports (2019), https://doi.org/10.1016/j.egyr.2019.09.026.