



Multiple Linear Regression Model Applied to the Projection of Electricity Demand in Colombia

Jesús García-Guiliany¹, Emiro De-la-hoz-Franco², Andrés-David Rodríguez-Toscano³, Juan-David De-la-Hoz-Hernández^{4*}, Hugo G Hernandez-Palma⁵

¹School of Business Administration, Universidad Simón Bolívar, Barranquilla, Colombia, ²Department Software Engineering and Networks, Universidad de la Costa - CUC, Barranquilla, Colombia, ³Department of Energy, Universidad de la Costa, Barranquilla, Colombia, ⁴Department Software Engineering and Networks, Corporación Universitaria Latinoamericana, Barranquilla, Colombia, ⁵School of Business Administration, Universidad del Atlántico, Barranquilla, Colombia. *Email: jdelaohzh@ul.edu.co

Received: 09 March 2019

Accepted: 10 October 2019

DOI: <https://doi.org/10.32479/ijeep.7813>

ABSTRACT

The exigencies as soon as to competitiveness and productivity have influenced in the energetic consumption and the demand of electrical energy in Colombia, reason why at the present time it is of much interest and utility to have access to tools or valid models to reach greater knowledge in which related to the possible future projections. Next, the results of a quantitative study are presented that through the analysis of data collected between 2007 and 2017 that made possible the construction of a multiple linear regression model to estimate the demand of electric energy. These types of instruments currently originate as alternatives to promote management strategies in the energy field in the country. The final results allow to visualize an estimated figure for the next periods which will serve to contrast with the official results and to generate from this information possible lines of intervention in different organisms.

Keywords: Energy Consumption, Electric Demand, Multiple Linear Regression Model

JEL Classifications: D24, Q43, M31

1. INTRODUCTION

There is no doubt that observing demand in terms of electricity consumption is a very important topic in Colombia (Holmberg and Erdemir, 2017), having in mind that the high exigency of competitiveness for all economic sectors is influencing the demand of this important component (Palma, 2017). There are several factors that influence the demand for electric energy, which is why projective methods or alternatives take on great significance in the current moment the country is going through (GPO, 2016), since they would allow control organizations as well as organizations that promote decision-making for future plans to have valid instruments to support their different investigations (Andrews-Speed et al., 2014).

As the demand for electricity fluctuates, the costs also vary in the electricity market (Fabra and Reguant, 2014). In addition

to the mentioned, it is also important to take into account for a correct approach to the energy component, aspects such as the environment, the behavior of the power generation system and the regulation rules that the country has been imparting to all actors or stakeholders (Ardila and Cardona, 2017). Looking at the behavior of electricity demand, taking as a reference official reports or derived from the energy sector, it can be seen that some sectors may impact more than others on the high and low consumption in different seasons during a year or period (Ñustes and Riviera, 2017).

For the foregoing, reaching models or systems that facilitate the projection of electricity consumption demand is now a valuable option for stimulating the design of intervention plans, as well as the implementation of strategies that derive in competitiveness but are also aligned with global energy

consumption parameters or trends (Nejat et al., 2015). Making short and long term projections can provide the necessary bases for control and monitoring bodies to make timely decisions and to direct alternatives closer to market requirements (Pukšec et al., 2014).

For the proposed goals, various tools are available from statistics that can be converted into valuable options to make models or observation systems. In this way, this document proposes the multiple linear regression model as an estimation technique (Kaytez et al., 2015). This model facilitates appreciating future values of a dependent variable (Y) based on independent variables ($X_1, X_2, X_3 \dots X_n$) with the condition that these are continuous, in other words, numeric. Multiple linear regression is then based on the adjustment of real data to a model that allows projecting according to an equation posed as (Montgomery et al., 2012):

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 \dots$$

Where:

Y = Dependent variable

a = Constant

b_n = Coefficients

X_n = Independent variables.

For the case that occupies the present article, it was investigated in the National Energy Plan Colombia: Energy Ideology 2050, carried out by the Mining and Energy Planning Unit (UPME) (2015) with the purpose of identifying the sectors of greater final consumption according to the exposed scenarios; this orientation, allowed to clarify the potential variables that would affect the demand of energy. By virtue of the foregoing, multiple regression was used to project the use of electricity in Colombia based on the data reported by the Operation Report of the National Interconnected System and Market Administration carried out by the non-governmental organization XN, S.A. corresponding to the series of data for the period 2007-2017 (SIN, 2017).

Taking into account the series of data provided by the entity, indicated in Table 1, we proceeded to perform the multiple linear regression model, based on variables reported by the World

Bank (2017), such as: Agriculture, value added (% of GDP), Industry, value added (% of annual growth), Industrialization, value added (% of annual growth), Urban population (% of total), Urban population growth (% annual), Expenditure on research and development (% of GDP), Imports of goods and services (% of GDP), Population density (people per km²), Gini index, GDP growth (annual %), Investment in energy with private participation (US\$ at current prices) and GDP per capita, PPP (US\$ at current international prices).

Using stepwise the inclusion method provided by the SPSS program, the data of the mentioned variables were incorporated to the model data, which allowed evaluating their initial correlation coefficients (Pearson's R) with the statistical significance measured by the p-value; as a result of this procedure, only those that added significantly to the multiple regression equation, discarding those that obtained $P > 0.05$, were added to the multiple regression equation. Next, Table 1 shows the independent variable (energy demand) and the independent variables that are relevant for the analysis.

2. THEORETICAL FOUNDATIONS OF THE MULTIPLE LINEAR REGRESSION MODEL

To facilitate the interpretation of results it is necessary to make a brief explanation of the elements that compose the multiple linear regression model. In the first place, the data provided by the SPSS is the goodness of adjustment, which is indicated by the value R^2 . In simple terms, this indicator allows us to identify if there is a correlation between the dependent variable (GWh) and the independent variables (GDP per capita, added value of the industry), which can be interpreted according to the following criteria (Montgomery et al., 2012):

- Values close to 1 would indicate a positive correlation between the variables (as the independent variable grows, so does the dependent variable).
- Values close to or equal to 0 show that there is no correlation between the variables.
- Values close to -1 show a negative relationship (as the independent variable grows, the dependent variable decreases).

Table 1: Annual energy demand in Colombia

Year	Y	X_1	X_2
	Annual energy demand in Colombia (GWh)	GDP per capita, PPP (US\$ at current international prices)	Industry, value added (% of GDP)
2007	52,853	97,109,176	3063
2008	53,870	101,322,657	3242
2009	54,679	102,602,267	3171
2010	56,148	10,680,003	3202
2011	57,155	114,964,777	3427
2012	59,370	120,583,438	3457
2013	60,890	127,250,437	3416
2014	63,571	133,955,162	3271
2015	66,174	138,276,728	3054
2016	66,318	141,654,512	2972
2017	66,893	145,520,086	2925

Operation report of the national interconnected system (SIN) (2017) and market administration, World Bank (2017)

On the other hand, the ANOVA variance analysis indicates whether it is feasible to construct a regression model from the selected predictor variables. The results will show the value of the so-called F test and its statistical significance value which will be interpreted according to the following criteria's:

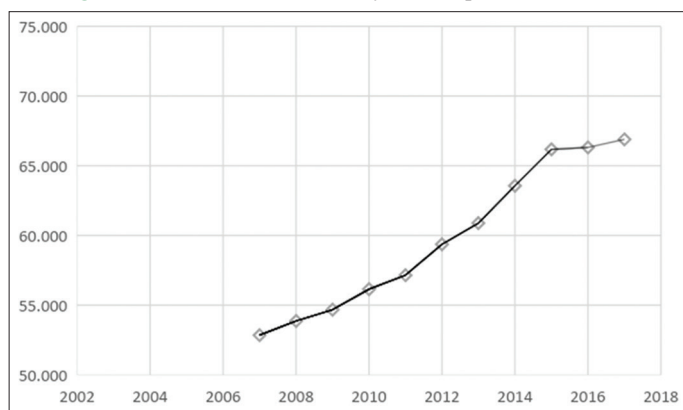
- a. If the significance value $P > 0.05$, the regression model is not viable and its execution is not recommended.
- b. If the significance value $P < 0.05$, the regression model is viable and its execution is recommended.

Finally, the third component of the results shown by the SPSS is the coefficient table, which will expose the variables required for the construction of the equation indicated at the beginning of the section.

3. METHODS

Data analysis was done with the support of SPSS version 24 software, using the multiple linear regression procedure accepted for the purposes expressed in the initial part (Sánchez-Villegas, 2014). The variables shown in Table 1 were incorporated into the program database and the dialog chart corresponding to the regression was executed, followed by the introduction of

Figure 1: Behaviour of electricity consumption in Colombia



Source: Operation Report of the National Interconnected System (SIN) (2017) and Market Administration

Table 2: Summary of the model

Model	R	R square	R square adjusted
1	0.993 ^a	0.986	0.984
2	0.998 ^b	0.996	0.995

^aPredictors: (Constant), GDP per capita, PPP (\$ at current international prices).

^bPredictors: (Constant), GDP per capita, PPP (\$ at current international prices), Industry, value added (% of GDP)

Table 3: ANOVA of the model

Model		Summation of squares	gl	Root mean square	F	Following
1	Regression	132,788	1	132,788	420,649	0,000 ^b
	Residue	1,894	6	0,316		
	Total	134,682	7			
2	Regression	134,187	2	67,093	677,019	0,000 ^c
	Residue	0,496	5	0,099		
	Total	134,682	7			

^aDependent variable: Annual energy demand in Colombia (GWh), ^bPredictors: (Constant), GDP per capita, PPP (\$ at current international prices). ^cPredictors: (Constant), GDP per capita, PPP (\$ at current international prices), Industry, value added (% of GDP)

the dependent variable (GWh) and the independent variables (Stephanidis, 2018), obtaining the results of the model indicated below.

4. RESULTS

Initially, the graphical level behavior of energy demand in Colombia for the period 2007 - 2017 is shown in order to observe the growth that has been reported during these last years and to be able to make an adequate projection of the data; Figure 1 shows this trend.

It is clearly observed a positive trend in energy consumption during recent years in Colombia, thus, it is relevant to make a projection based on data using the multiple regression model, taking into account the predictor variables, Table 2 shows the summary of the model.

As explained in previous sections, the step-by-step method allows the variables to be progressively incorporated into the equation and thus determine those that really contribute significantly. Table 2 shows model 1 (first step) where only per capita GDP is taken into account with an adjusted $R^2 = 0.984$; then, model 2 clearly has the best adjustment evidenced in the adjusted $R^2 = 0.995$, also considering the added value of the industry as a predictor. On the other hand, Table 3 of the ANOVA points out the relevance of the regression using these variables.

This result would indicate the relevance of whether or not to continue with regression analysis based on two fundamental data: the F test and the sig. value, or commonly referred to as the P value. Based on the theoretical criteria explained in previous sections, it is possible to confirm that the regression model is viable and its execution is recommended, since the $P < 0.05$ in step 2. Continuing with the procedure, the coefficients obtained from the multiple linear regression are shown in Table 4.

The table summarizes the coefficients in column B necessary for the construction of the model, both the constant and the independent variables; additionally, the $P < 0.05$ would indicate that these data can be used, since they are statistically significant. With this, we proceed to the configuration of the following regression equation:

$$Y = 34,050 + 0,003 (X_1) + (-0,279(X_2))$$

Table 4: Multiple linear regression coefficients

Modelo		Standardized coefficients		
		B	t	P value
1	(Constant)	24,307	14,272	0,000
	GDP per capita, PPP (\$ at current international prices)	0,993	20,510	0,000
2	(Constant)	34,050	12,322	0,000
	GDP per capita, PPP (\$ at current international prices)	0,003	35,413	0,000
	Industry, value added (% of GDP)	-0,279	-3,757	0,013

*Dependent variable: Annual energy demand in Colombia (GWh)

5. CONCLUSIONS

The data exposed and analyzed through the regression model, allow us to observe that in Colombia the existing trend between 2007 and 2017, shows a marked growth in the demand of electric energy. Thus things the premise that suggests that despite constant campaigns for savings in consumption and the insertion of renewable energy, electricity continues to be the most employed nationwide. The proposed model indicates in a similar way that due to the behavior of consumption in the Colombian territory this instrument can be taken to make diverse projections and reflections, giving place to the generation of strategies that allow to visualize the way to attenuate the increase of the demand and to receive certain measures for the formulation of plans consistent with the national and international energetic plans and oriented to incorporate new energetic alternatives to fulfill the demands as for environment and economic factors associated to the long term sustainability.

REFERENCES

- Andrews-Speed, P., Liao, X., Dannreuther, R. (2014), *The Strategic Implications of China's Energy Needs*. London: Routledge.
- Ardila, L.M.C., Cardona, C.J.F. (2017), Structure and current state of the wholesale electricity markets. *IEEE Latin America Transactions*, 15(4), 669-674.
- Banco Mundial. (2017), Sección Indicadores. Available from: <https://www.datos.bancomundial.org/indicador>.
- Fabra, N., Reguant, M. (2014), Pass-through of emissions costs in electricity markets. *American Economic Review*, 104(9), 2872-2899.
- Government Publications Office. editor. (GPO). (2016), *International Energy Outlook 2016: With Projections to 2040*. Government Printing Office.
- Holmberg, K., Erdemir, A. (2017), Influence of tribology on global energy consumption, costs and emissions. *Friction*, 5(3), 263-284.
- Informe de Operación del Sistema Interconectado Nacional (SIN). (2017), *Demanda de Energía Nacional*. Available from: <http://www.informes anuales.xm.com.co/2017/SitePages/operacion/4-1-Demanda-de-energia-nacional.aspx>.
- Kaytez, F., Taplamacioglu, M.C., Cam, E., Hardalac, F. (2015), Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. *International Journal of Electrical Power and Energy Systems*, 67, 431-438.
- Montgomery, D., Peck, E.A., Vining, G. (2012), *Introduction to Linear Regression Analysis*. Vol. 821. New Jersey: John Wiley and Sons.
- Nejat, P., Jomehzadeh, F., Taheri, M.M., Gohari, M., Majid, M.Z.A. (2015), A global review of energy consumption, CO₂ emissions and policy in the residential sector (with an overview of the top ten CO₂ emitting countries). *Renewable and Sustainable Energy Reviews*, 43, 843-862.
- Ñustes, W., Riviera, S. (2017), Colombia: territorio de inversión en fuentes no convencionales de energía renovable para la generación eléctrica. *Revista Ingeniería, Investigación y Desarrollo*, 17, 37-48.
- Palma, H.H. (2017), *Direccionamiento estratégico para la dinamización del sector salud en el departamento del Atlántico*. *BIOCIENCIAS*, 12(1), 79-84.
- Pukšec, T., Mathiesen, B.V., Novosel, T., Duić, N. (2014), Assessing the impact of energy saving measures on the future energy demand and related GHG (greenhouse gas) emission reduction of Croatia. *Energy*, 76, 198-209.
- Sánchez-Villegas, A. (2014), In: Martínez-González, M.A., Faulín, F.J., editors. *Bioestadística Amigable*. Barcelona: Elsevier.
- Stephanidis, C. editor. (2018), *HCI International 2018 Posters' Extended Abstracts: 20th International Conference*. Vol. 852. HCI International 2018, Las Vegas, NV, USA, Proceedings. Springer.
- Unidad de Planeación Minera y Energética (UPME). (2015), *Plan Energetico Nacional Colombia: Ideario Energético 2050*. Available from: http://www1.upme.gov.co/Documents/PEN_IdearioEnergetico2050.pdf.