



## ANÁLISIS ESTADÍSTICO DE LAS EMISIONES DE CO<sub>2</sub> POR EL CONSUMO DE COMBUSTIBLE GASEOSO EN ECUADOR MEDIANTE LA APLICACIÓN DE LA REGRESIÓN LINEAL SIMPLE

**Luis Alberto De Lucas Coloma**

Economista, Master en Administración de Empresas  
Docente de la Universidad Regional Autónoma de Los Andes  
[deluccaec@yahoo.com](mailto:deluccaec@yahoo.com)

Para citar este artículo puede utilizar el siguiente formato:

Luis Alberto De Lucas Coloma (2018): "Análisis estadístico de las emisiones de CO<sub>2</sub> por el consumo de combustible gaseoso en Ecuador mediante la aplicación de la regresión lineal simple", Revista Caribeña de Ciencias Sociales (julio 2018). En línea:  
[//www.eumed.net/rev/caribe/2018/07/combustible-gaseoso-ecuador.html](http://www.eumed.net/rev/caribe/2018/07/combustible-gaseoso-ecuador.html)

### Resumen

Este estudio parte de la problemática ambiental que influye las emisiones de CO<sub>2</sub> por el consumo de combustible gaseoso en Ecuador, como también el análisis de los factores sociales y económicos que influyen a las emisiones, por lo tanto, se determinó como objetivo desarrollar un análisis estadístico mediante la aplicación de la regresión lineal, en la cual correlaciona una variable independiente como el número anual de habitantes y las emisiones de CO<sub>2</sub> anuales específicamente el consumo de combustible gaseoso. En consecuencia los resultados obtenidos presentan que si existe una relación estadística entre dichas variables, es decir el aumento en el número de habitantes infiere directamente en el aumento de emisiones de CO<sub>2</sub>.

Palabras claves: Regresión lineal simple, combustible gaseoso, Población, factores socioeconómicos, CO<sub>2</sub>.

### Abstract

This study is based on the environmental issues that influence CO<sub>2</sub> emissions from gaseous fuel consumption in Ecuador, as well as the analysis of the social and economic factors that influence emissions, therefore, the objective was to develop a statistical analysis through the application of linear regression, in which an independent variable such as the annual number of inhabitants and annual CO<sub>2</sub> emissions, specifically the consumption of gaseous fuel, correlates. Consequently, the results obtained show that if there is a statistical relationship between these

variables, that is, the increase in the number of inhabitants directly influences the increase in CO<sub>2</sub> emissions.

Key words: Simple linear regression, gaseous fuel, Population, socioeconomic factors, CO<sub>2</sub>.

## INTRODUCCIÓN

En los últimos años el planeta ha sufrido cambios significativos en torno al medio ambiente y estos cambios se ven representados en los cambios climáticos, aumento en la temperatura global, sequías, inundaciones, etc. Sin embargo, uno de los factores imperantes de dichos cambios es el uso indiscriminado de combustibles fósiles, como el petróleo, carbón y gas (Rentería, Toledo, Bravo, & Ochoa, 2016). Estos derivados su insumos importantes no solo en la economía de un país como Ecuador, también se lo considera a nivel mundial un elemento esencial para activación productiva a nivel mundial, como en la actividad industrial, comercial y residencial, por tanto, son los causantes importantes de efectos negativos en los ciclos de la atmosfera y por generar gran cantidad de gases de efecto invernadero fomentando el calentamiento global, afectado de forma negativos en los seres vivos con miles de millones de afectados y provocando grandes pérdidas económicas (Barros, 2006).

Según Ortega y Abad (2014) determinan que de acuerdo a los resultados presentados en la investigación y dada la realidad del Ecuador, es necesaria la sugerencia de recomendaciones de política que permitan alcanzar un equilibrio entre lo económico y ambiental, lo cual se sobrepone a la idea de “esperar y crecer”.

La emisión de CO<sub>2</sub> dentro del Ecuador por el consumo de combustible gaseoso es controlada por el Ministerio del Ambiente, el cual posee datos desde 1979 hasta 2013. Los mismos datos son los que serán usados en este proyecto el cual tiene como método de análisis la regresión lineal simple. Mediante este método se comprobara si durante los próximos años se reducirá o incrementará el nivel de CO<sub>2</sub> en el Ecuador. (Banco Mundial, 2017)

Para el desarrollo del análisis de regresión lineal se utiliza el software Gretl, el cual es de fácil manejo. La regresión lineal es un método útil para encontrar tendencias de valor dispersos pero que siguen un orden. El método de regresión lineal fue encontrado por el “príncipe de las matemáticas” Carl Gauss quien aplico métodos algebraicos para encontrar una relación de varias observaciones.

El método de regresión líneas sigue los supuestos de Gauss-Markov formulados por Carl Gauss y Andrei Markov. Los cuales son: correcta especificación, muestreo aleatorio simple, esperanza condicionada, correcta identificación y homocedasticidad. Dicho supuestos limitan mucho el campo de la regresión lineal pero sirven como la base de modelos más complejos.

## 2. FUNDAMENTACIÓN TEÓRICA

En el presente estudio analiza la relación existente entre factores de crecimiento poblacional y emisiones de CO<sub>2</sub>, considerando también el impacto socioeconómico del país. Según Zilio (2010) señala que desde la perspectiva de la validez de la Curva de Kuznets de Carbono (CKC) aplicada para la región de América Latina y el Caribe, existe la relación empírica entre el producto y las emisiones de CO<sub>2</sub> para una muestra conformada por veintitrés países de la región a través de un enfoque no paramétrico basado en la estimación kernel de polinomios suavizados. Espinoza (2013), se analizó la relación entre crecimiento económico y la contaminación ambiental en el Ecuador durante las últimas cinco décadas. Utilizó pruebas de cointegración con el enfoque de testear una relación a largo plazo entre dichas variables, sin embargo fue imperante considerar dentro del análisis la densidad poblacional como variable explicativa (Rentería, Toledo, Bravo, & Ochoa, 2016).

La regresión lineal es un método muy utilizado para realizar análisis, ya que con este método se pueden realizar modelos donde se muestra la dependencia de una variable con respecto a otra. Donde según la nomenclatura, "X" es la variable independiente o variable explicativa y "Y" es la variable dependiente o variable respuesta (Carrasquilla-Batista, 2016).

La estructura de la regresión lineal es la siguiente:

$$Y = \beta_0 + \beta_1 X + e$$

En la expresión anterior hay dos factores o causas que influyen en Y, también está presente el e, el error aleatorio, que provoca la dependencia entre las dos variables, variable dependiente e independiente no se perfecta, ya que existe incertidumbre.

$\beta_0$  es la ordenada al origen del modelo y  $\beta_1$  es la pendiente, que representa el incremento de variable dependiente por el incremento una unidad de la variable independiente.

Se establecen algunas hipótesis:

La varianza de e es constante, para cualquier valor de x.

$$Var\left(\frac{e}{x} = x\right) = \sigma^2$$

La distribución de e es normal, la media 0 y la desviación  $\sigma$ .

Los errores asociados a los valores de Y son independientes.

### **Método de mínimos cuadrados**

Es un procedimiento de minimización para la estimación de parámetros. El método de mínimos cuadrados consiste en minimizar los errores de la siguiente forma (Walpole, Myers, & Myers, 1999):

$$\text{Min}_{\beta_1, \beta_2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$$

Se deriva parcialmente:

$$\frac{dS}{d\beta_0} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\frac{dS}{d\beta_1} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) x_i = 0$$

Se obtienen las siguientes expresiones:

$$\sum_{i=1}^n y_i = \left( n\beta_0 + \beta_1 \sum_{i=1}^n x_i \right)$$

$$\sum_{i=1}^n y_i x_i = \left( \beta_0 \sum_{i=1}^n x_i + \beta_1 \sum_{i=1}^n x_i^2 \right)$$

Y aplicando procedimientos matemáticos obtenemos los parámetros:

$$\beta_1 = \frac{\sum_{i=1}^n y_i x_i - \bar{y} \sum_{i=1}^n x_i}{\sum_{i=1}^n x_i^2 - \bar{x} \sum_{i=1}^n x_i}$$

$$\beta_0 = \bar{y} - \beta_1 \bar{x}$$

Para medir la significancia de la regresión se usa la bondad de ajuste por medio del  $R^2$  que se define de la siguiente forma:

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

### ***Planteamiento de las variables***

Para esta investigación se obtuvo que la variable independiente es la población total y que la variable dependiente son las emisiones de CO<sub>2</sub> por el consumo de combustible gaseoso.

**Tabla 1 - 2 – Variables**

<b>Población total</b>	<b>Emisiones de CO2</b>
7,976445	73,34
8,183194	132,012
8,39294	172,349
8,606213	201,685
8,823751	194,351
9,045979	187,017
9,272906	190,684
9,504129	95,342
9,739176	190,684
9,977377	363,033
10,218091	509,713
10,46099	601,388
10,705667	465,709
10,951202	535,382
11,196479	484,044
11,440583	465,709
11,683479	495,045
11,924993	451,041
12,163885	462,042
12,398691	432,706
12,628596	473,043
12,852755	480,377
13,07206	476,71
13,289601	931,418
13,509647	1136,77
13,735233	894,748
13,96748	762,736
14,205453	891,081
14,447562	777,404
14,691275	876,413
14,93469	971,755
15,177355	858,078
15,419666	1169,773
15,661547	1320,12

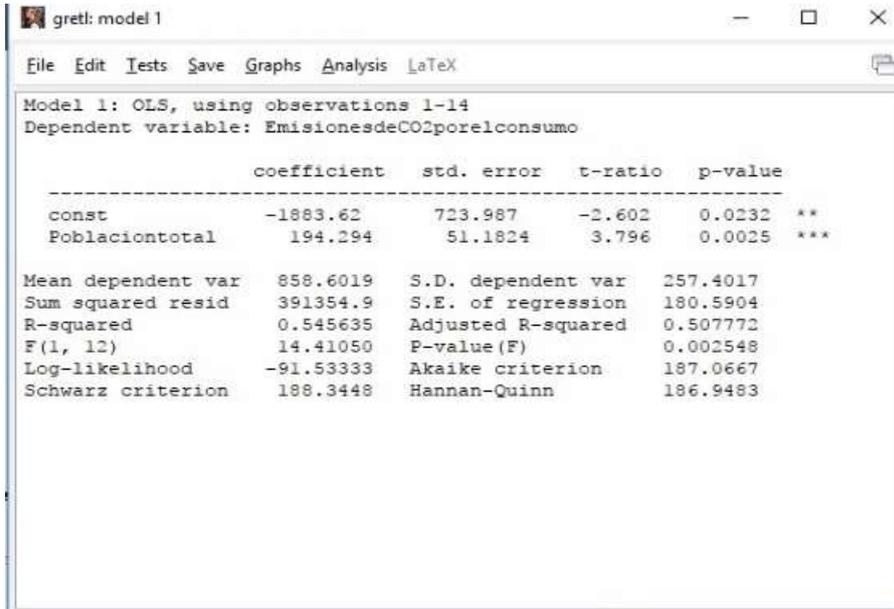
**Fuente:** (Banco Mundial, 2018)

**Elaborado por:** El autor

### 3. CONSTRUCCIÓN DEL MODELO E INTERPRETACIÓN

#### *Construcción del modelo*

Las variables serán analizadas por el software “Gretl”, con el cual se obtuvieron los siguientes resultados.



The screenshot shows the Gretl software interface with the following regression results:

	coefficient	std. error	t-ratio	p-value	
const	-1883.62	723.987	-2.602	0.0232	**
Poblaciontotal	194.294	51.1824	3.796	0.0025	***

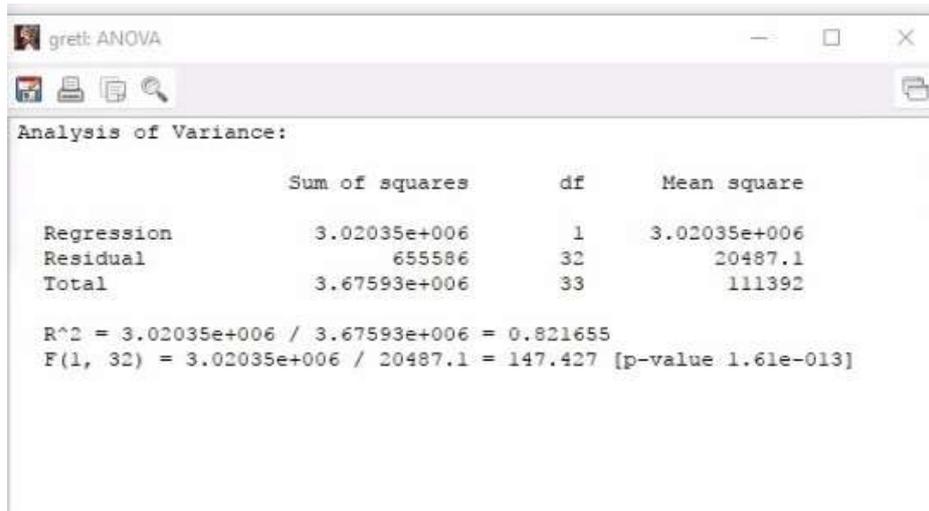
  

Mean dependent var	858.6019	S.D. dependent var	257.4017
Sum squared resid	391354.9	S.E. of regression	180.5904
R-squared	0.545635	Adjusted R-squared	0.507772
F(1, 12)	14.41050	P-value(F)	0.002548
Log-likelihood	-91.53333	Akaike criterion	187.0667
Schwarz criterion	188.3448	Hannan-Quinn	186.9483

Figura 1 – III: Modelo 1

Fuente: Software gretl

Elaborado por: el autor



The screenshot shows the ANOVA results in Gretl:

	Sum of squares	df	Mean square
Regression	3.02035e+006	1	3.02035e+006
Residual	655586	32	20487.1
Total	3.67593e+006	33	111392

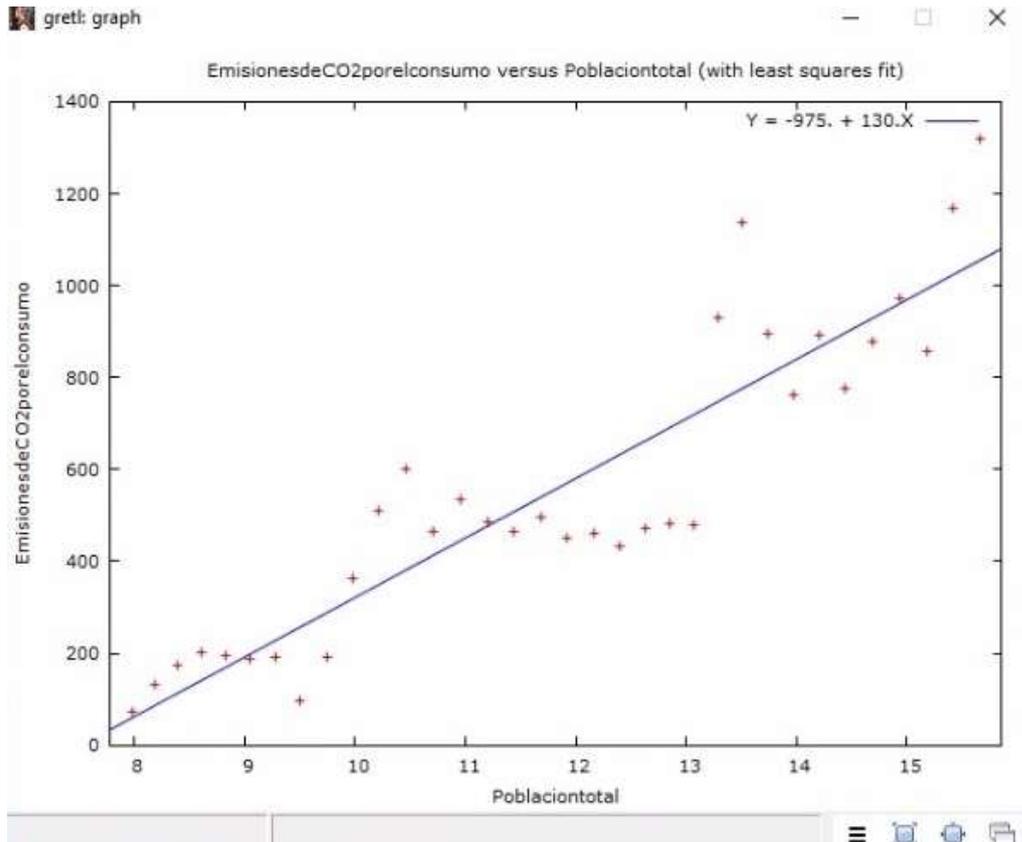
  

$R^2 = 3.02035e+006 / 3.67593e+006 = 0.821655$   
 $F(1, 32) = 3.02035e+006 / 20487.1 = 147.427$  [p-value 1.61e-013]

Figura 2 – III: Análisis de Varianza

Fuente: Software gretl

Elaborado por: el autor



**Figura 3 – III:** Gráfica de la regresión

**Fuente:** Software gretl

**Elaborado por:** el autor

### ***Interpretación de coeficientes y $R^2$***

*El coeficiente de determinación ( $R^2$ )*

El coeficiente de determinación muestral, indicará cuán eficiente es el modelo de regresión lineal simple para este estudio. El cual está dado por:

$$R^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}}$$

Haciendo uso de programa de Gretl, se obtuvo un  $R^2$  ajustado de = 0,8161

De esta manera una correlación del 80% de la variación total de los valores de Y en la muestra se explica mediante una relación lineal con los valores de X.

*Aplicando el Estadístico de prueba "f"*

Tendiendo como hipótesis:

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

$$F_0 = \frac{SCR/1}{SSE/n-2}$$

Donde:

$$SCR = 147,715736$$

$$SSE = 32,0626497$$

La toma de decisión estaría dada por:

La  $H_0$  se rechaza si:  $F_0 > F_\alpha(1, n-2)$

El nivel de significancia a usar en este modelo es de 5% con  $n = 34$

Entonces, si el  $F_0$  se excede se concluye que hay una cantidad significativa de variación en la respuesta por el modelo que se postula.

El valor de  $F_0$  es: 147,427103

Y el  $F_\alpha$  es: 4.17

Se rechaza la hipótesis nula. Se concluye que el modelo de regresión es significativo.

*Aplicando el Estadístico de prueba "t"*

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

Para determinar la validez de las hipótesis, se aplicaría un Estadístico t -

\* Con un nivel de significancia del 5%

La toma de decisión estaría dada por:

Se rechaza  $H_0$  si:  $t_0 > t(\alpha/2, n-2)$

\*La región crítica estaría dada por  $t > 2.042$ . El cual es un valor de tabla.

\*Entonces se determina el Estadístico t:

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

Donde el coeficiente de correlación muestral "r" sería = 0,9065

$$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

El valor de la estadística T que tiene una distribución t con  $n - 2$  grados de libertad.

Dando paso al valor " $t_0$ " de = **12,1419563**

Decisión:

Se rechaza la hipótesis nula, El valor  $t_0$  resulto ser mayor al t

Con respecto al parámetro  $\beta_0$  (Intersección al eje y)

$H_0: \beta_0 = 0$

$H_1: \beta_0 \neq 0$

Para determinar la valides de la hipótesis  $\beta_0$ , se aplica un Estadístico t

Se rechaza  $H_0$  si:  $t_0 > t(\alpha/2, n-2)$

Dando paso al valor " $t_0$ " de = - 12,1419563. El cual es un valor de tabla.

Entonces se determina el Estadístico t:

$$t_0 = \frac{\beta}{\sqrt{CME \left[ \frac{1}{n} + \frac{X^2}{S_{xx}} \right]}}$$

Dando paso el valor  $t_0$  de = -5625,109

### **Análisis de resultados**

Se rechaza la hipótesis nula, El valor  $t_0$  resulto ser menor a la t de la tabla.

## **4. CONCLUSIONES**

Una vez realizado el análisis de regresión mediante el software Gretl y con los resultados que se obtuvo, se rechaza las hipótesis nulas de ambos análisis (de regresión y de varianza) se concluye que la ecuación de regresión simple, es significativa, por lo que se puede afirmar que el modelo es importante para preceder la "Emisiones de CO2 por el consumo de combustibles gaseosos"

Sin embargo, estos resultados muestran que desde el punto de vista socioeconómico el país debe seguir con proyectos que generen energía eléctrica limpia como las hidroeléctricas de tal forma el transporte público y particular dependan en mediano plazo en la totalidad de este tipo de energía y así, acabar con la mayor parte de emisiones de gases tóxicos en el país.

## **5. GLOSARIO DE TÉRMINOS**

CO2

Sustancia gaseosa que por inhalación prolongada produce diferentes efectos y consecuencias en la salud del ser humano, desde pérdida de conocimiento, hasta efectos que al no ser atendidos pueden producir la muerte (Centro Regional de Información sobre desastres, 2017).

#### Combustible gaseoso

Combustible gaseoso. Están formados principalmente por hidrocarburos, es decir, compuestos moleculares de carbono e hidrógeno. Las propiedades de los diferentes gases dependen del número y disposición de los átomos de carbono e hidrógeno de sus moléculas. Todos estos gases son inodoros en estado puro, igual que ocurre con el Monóxido de Carbono (tóxico) que a veces contienen. Por eso es corriente añadir compuestos de azufre al gas comercial; estos compuestos, que a veces están presentes de forma natural en el gas, tienen un olor desagradable y sirven para advertir un escape en las tuberías o en los aparatos de gas. Además de sus componentes combustibles, la mayoría de los combustibles gaseosos contienen cantidades variables de Nitrógeno y agua (ECURED, 2017).

#### Densidad de población

Se llama densidad de población al número total de habitantes dividido entre el número de km<sup>2</sup> de una zona, región o país determinados (Instituto Nacional de Estadística y Censos · INEC, 2017).

#### Factores socioeconómicos

Factores económicos y sociales que caracterizan a un individuo o un grupo dentro de una estructura social (Centro Regional de Información sobre desastres, 2017).

#### Regresión lineal simple

El objetivo de un modelo de regresión es tratar de explicar la relación que existe entre una variable dependiente (variable respuesta) Y un conjunto de variables independientes (variables explicativas) X<sub>1</sub>,..., X<sub>n</sub>. En un modelo de regresión lineal simple tratamos de explicar la relación que existe entre la variable respuesta Y y una única variable explicativa X (Universidad de Santiago de Compostela, 2012).

## 6. BIBLIOGRAFÍA

Banco Mundial. (Julio de 2017). <http://www.bancomundial.org/es/country/ecuador>.

Banco Mundial. (5 de Mayo de 2018). [datos.bancomundial.org](https://datos.bancomundial.org/indicador/EN.ATM.CO2E.GF.KT?end=2014&locations=EC&start=1978). Obtenido de <https://datos.bancomundial.org/indicador/EN.ATM.CO2E.GF.KT?end=2014&locations=EC&start=1978>

Barros, V. (2006). *Cambio climático global*. Libros del Zorzal.

- Carrasquilla-Batista, A. C.-R.-M.-E.-B. (2016). Regresión lineal simple y múltiple: aplicación en la predicción de variables naturales relacionadas con el crecimiento microalgal. *Revista Tecnología en Marcha*, 29(8), 33-45.
- Centro Regional de Información sobre desastres. (2017). Vocabulario controlado sobre desastres. San José: CRID; 2000. (CRID).
- ECURED. (2017). [www.ecured.cu](http://www.ecured.cu). Obtenido de [https://www.ecured.cu/Combustible\\_gaseoso](https://www.ecured.cu/Combustible_gaseoso)
- Espinosa, J. (2013). Estimación de la curva de Kuznets medioambiental en el Ecuador durante el período 1961-2010. *Universidad de Cuenca*.
- Instituto Nacional de Estadística y Censos · INEC. (2017). [ecuadorencifras.gob.ec](http://ecuadorencifras.gob.ec). Obtenido de <http://www.ecuadorencifras.gob.ec/documentos/web-inec/Bibliotecas/Libros/Panorama%20Laboral%202017.pdf>
- Ortega, J., & Abad, R. (2014). Medición de los niveles de contaminación de CO y CO<sub>2</sub>, a través de un sistema electrónico basado en PLC's, para el monitoreo de la calidad del aire en la Universidad Católica de Cuenca, Sede Azogues. *Revista Interamericana de Ambiente y Turismo*, Vol. 10, Nº 2, P. 217-224. doi: <http://dx.doi.org/10.4067/riatvol10iss2pp217-224%250718-235X>
- Rentería, V., Toledo, E., Bravo, D., & Ochoa, D. (2016). Relación entre Emisiones Contaminantes, Crecimiento Económico y Consumo de Energía. El caso de Ecuador 1971-2010. *Revista Politécnica-Septiembre*, 38(1).
- Universidad de Santiago de Compostela. (2012). <http://eio.usc.es>. Obtenido de [http://eio.usc.es/eipc1/BASE/BASEMASTER/FORMULARIOS-PHP-DPTO/MATERIALES/Mat\\_50140116\\_Regr\\_%20simple\\_2011\\_12.pdf](http://eio.usc.es/eipc1/BASE/BASEMASTER/FORMULARIOS-PHP-DPTO/MATERIALES/Mat_50140116_Regr_%20simple_2011_12.pdf)
- Walpole, Myers, & Myers. (1999). *Probabilidad y Estadística para ingenieros*. México: Prentice Hall.
- Zilio, M. (2010). LA CURVA DE KUZNETS AMBIENTAL:EVIDENCIA PARA AMÉRICA LATINA Y EL CARIBE . *UNIVERSIDAD NACIONAL DEL SUR* , 112.