

# Estimación del consumo eléctrico colombiano en el corto y largo plazo empleando regresión multivariable y series temporales

## Forecast of the colombian electricity consumption in the short and long term using multivariable regression and time series

---

Daniel Orlando Garzón Medina<sup>1</sup>, Geovanny Alberto Marulanda García<sup>2</sup>

<sup>1</sup>Universidad de la Salle, [danior12@hotmail.com](mailto:danior12@hotmail.com)

<sup>2</sup>Universidad de la Salle, [gemarulanda@unisalle.edu.co](mailto:gemarulanda@unisalle.edu.co)

DOI: <https://doi.org/10.18041/1794-4953/avances.1.1294>

Fecha de recepción: 26/02/2017

Fecha de aceptación del artículo: 14/11/2017

---

### Cómo citar:

GARZÓN MEDINA, Daniel Orlando; MARULANDA GARCÍA, Geovanny Alberto. Estimación del consumo eléctrico colombiano en el corto y largo plazo empleando regresión multivariable y series temporales. **AVANCES <BR>Investigación en Ingeniería**, [S.l.], v. 14, p. 155-168, dec. 2017. ISSN 2619-6581. Disponible en: <http://revistas.unilibre.edu.co/index.php/avances/article/view/1294>. Fecha de acceso: 19 feb. 2018 doi: <https://doi.org/10.18041/1794-4953/avances.1.1294>.

### Resumen

La previsión de consumo de energía eléctrica constituye un pilar importante para desarrollar proyectos de expansión en generación, transmisión y distribución. En este trabajo se propone realizar una proyección de demanda para el consumo de energía eléctrica en el sector residencial colombiano por medio de una serie temporal y una regresión multivariable que relacione el crecimiento económico del país con su consumo eléctrico. Para validar la metodología propuesta, se compararán los resultados obtenidos con la información oficial suministrada por la Unidad de Planeación Minero-Energética (UPME).

**Palabras claves:** Estimación, consumo eléctrico, Regresión multivariable, series temporales.

### Abstract

The forecast electricity consumption constitutes an important step towards the development of new technologies to meet the energy consumption in the coming decades. This paper intends to make a projection by means of a time series and multivariate regression which allows to connect the country's economic growth with its electric consumption. To validate the proposed methodology, the error is calculated in respect to official information provided by the UPME.

**Keywords:** Forecasting, Electrical consumption, multivariable regression, time series.

## 1. Introducción

El constante desarrollo de los mercados de energía eléctrica genera escenarios cada vez más competitivos, por lo que es de crucial importancia poseer sistemas de transmisión y distribución con planes de manejo bien estructurados [1] y [2]. Entre los planes de manejo, se debe contar con modelos de pronóstico de la demanda de energía adecuados y es por esto por lo que actualmente se realizan investigaciones en torno a esta temática. El pronóstico de la demanda de electricidad es una herramienta fundamental para la toma de decisiones operativas y estratégicas en las empresas de energía, cuya falta de precisión puede generar altos costos económicos [3]. Dicha predicción puede determinar de forma previa la carencia de energía en el sistema, o si pudiera ser conveniente, la consideración en la construcción de nuevas subestaciones para suplir la demanda necesaria. Con la información suministrada por el estudio de pronóstico de demanda de energía eléctrica a corto y largo plazo, los operadores pueden tomar decisiones en torno al despacho óptimo de centrales de generación, y realizar análisis de confiabilidad y seguridad en la operación del sistema.

Los métodos para la predicción se dividen en dos grandes grupos: Los métodos basados en inteligencia artificial, y los métodos paramétricos.

El primer grupo está compuesto por modelos basados en redes neuronales. Tienen la característica que no necesitan un modelo matemático para representar el consumo de energía eléctrica, pero si poseen parámetros importantes tales como entradas, funciones de transferencia, algoritmo de entrenamiento y el

número de neuronas. Este método utiliza diferentes algoritmos de aprendizaje y actualmente es muy utilizado debido a su amplio uso en el área de la ingeniería [5].

El segundo método de este primer grupo se basa en lógica difusa, y considera la imprecisión de los datos a través de reglas de extracción de variables de entrada y de modelos estadísticos [5]. En general, esta técnica es usada para el tratamiento de datos que al combinarse con otra técnica permite lograr una predicción completa y satisfactoria [6].

El segundo grupo está compuesto por modelos paramétricos, y se basan en el modelado matemático de la demanda. Los parámetros de los modelos son estimados usando técnicas estadísticas sobre históricos de datos, o posibles factores que puedan afectar estos datos [1] y [5].

En [1] se explica la importancia de la regresión multivariable en el proceso de pronóstico, pues son aquellas variables las que darán precisión a la respuesta que se obtendrá. La autora desarrolla el proceso de tratamiento de datos estadísticos para obtener un proceso “suave” para la línea de tendencia deseada. Por su parte, en [4], se propone un modelo de previsión de demanda de electricidad de largo plazo por medio de Modelos Auto-Regresivos Integrados por Media Móvil (ARIMA por sus siglas en inglés). Esta herramienta de análisis desarrollada por la Escuela Técnica Superior de Ingeniería (ICAI) es capaz de crear distintos modelos de predicción de series temporales, así como modelos de decisión. Se argumenta que los factores que influyen una predicción de largo plazo son las nuevas variables que aparecen en el horizonte temporal (1-10 años) [4]. Por otro lado, el planeamiento integra-

do de recursos energéticos propuesto en [8], muestra la forma en que los recursos pueden ser aprovechados de forma tal que no se genere contaminación ni grandes emisiones de CO<sub>2</sub>. Se realiza un recuento de las formas de obtención de energía, su aporte para el desarrollo sostenible y las estrategias para la conservación mundial. El estudio desarrollado por el autor se torna más amplio dándole soluciones a los problemas de demanda futura que parte de la proyección demográfica a largo plazo e incluye un análisis de desarrollo sostenible enfocado al sector energético (residencial, industrial, comercial y rural).

Por su parte, la Agencia Nacional de Energía Eléctrica de Brasil (ANEEL), expone las directrices para el planeamiento de expansión del sistema de distribución, estableciendo requisitos mínimos e información para realizar la planeación. La ANEEL hace énfasis en que los modelos de previsión son de libre elección por parte de las distribuidoras, y que los resultados de los modelos de previsión de demanda están sujetos a validación por parte de la ANEEL [7].

En algunos de los trabajos realizados sobre previsión de la demanda de energía eléctrica se establecen una serie de falencias en los modelos propuestos. Algunas de estas son [11]:

- No existen razones empíricas, metodológicas o teóricas para preferir un modelo específico entre los diferentes tipos de modelos.
- Falta información sobre los procedimientos más adecuados para la estimación y la prueba de cada modelo particular.

- Es confusa la metodología empleada.
- Es difícil incorporar información cualitativa en los pronósticos, pues modifica completamente el desarrollo del modelo.

Por medio del desarrollo de las series temporales y la regresión multivariable propuestas en este trabajo, se pretende establecer una metodología clara y capaz de explicar el tratamiento de datos que permitan llegar a un pronóstico acertado de la demanda de energía eléctrica gracias a la recolección de históricos de consumo de energía eléctrica, el PIB, y el crecimiento poblacional colombiano. Se diseñarán los modelos para calcular la demanda a corto y largo plazo del sector residencial, teniendo así una estimación acertada para verificar la calidad de la respuesta en la metodología propuesta. De esta forma se busca establecer una nueva metodología aplicada a un problema de pronóstico de energía eléctrica implementando una variable poblacional que posiblemente mejorará la precisión respecto a los modelos econométricos dinámicos que tienen en cuenta los efectos climatológicos utilizados por la UPME. Se compararán los resultados con los pronósticos realizados por la UPME para así poder garantizar un servicio confiable y de calidad para los futuros usuarios que hagan parte del sector en Colombia [9].

Este trabajo se encuentra ordenado de la siguiente forma: La sección II presenta los modelos matemáticos empleados para la previsión de corto y largo plazo, la sección III presenta los resultados obtenidos y los datos reales por parte de la UPME, y la sección IV presenta las conclusiones y recomendaciones para ser desarrollados en futuros trabajos.

## 2. Modelos Matemáticos

### A Regresión multivariable

Más conocida como Multivariate Adaptive Regression Splines (MARS por sus siglas en inglés), es un método estadístico no paramétrico que fue propuesto con el fin de aproximar, de manera eficiente la relación de más de dos variables para la obtención de una variable predictora [10].

Se utiliza cuando se estudia la posible relación entre varias variables independientes (predictoras o explicativas) y otra variable dependiente (respuesta). Los usos que con mayor frecuencia toma este modelo además de la predicción de respuesta a partir de variables explicativas son:

- Identificación de variables explicativas: Ayuda a crear modelos donde se seleccionen las variables que pueden influir en la respuesta, descartando aquellas que no aporten información.
- Detección de interacciones: Relación entre variables que afectan a la variable respuesta. Esto permite descartarlas o incluirlas en el modelo para que la respuesta sea más precisa.

Para el desarrollo del modelo multivariable, se parte de una función lineal suponiendo que el impacto de cada variable independiente a considerarse afectará de manera proporcional la variable dependiente que se desea analizar:

$$Y_i = a_0 + a_1 x_1 + a_2 x_2 + \dots + \varepsilon_i \tag{1}$$

Dónde:

|                 |   |
|-----------------|---|
| $Y_i$           | Variable dependiente, donde muestra el valor inicial de la función. |
| $a_0$           | Punto de corte con la ordenada.                                     |
| $a_1, a_2$      | Constantes de proporción.   |
| $x_1, x_2$      | Variables independientes.   |
| $\varepsilon_i$ | Residuos y posible error de medición.                               |

Para un conjunto de  $n$  variables, donde  $i \in Z^+$  se tiene que

$$Y = XA + E \tag{2}$$

Dónde:

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1m} \\ 1 & x_{21} & x_{22} & \dots & x_{2m} \\ 1 & x_{31} & x_{32} & \dots & x_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nm} \end{pmatrix}$$

$$A = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}, \quad E = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Siendo  $n$  la cantidad de variables y  $m$  la cantidad de muestras.

Para determinar la función que mejor representa el conjunto de datos con los cuales se realizará la predicción, se aplica a (2) el método de los mínimos cuadrados [12] y [14]. Esto es

$$E^T E = (U - XA)^T (U - XA) \tag{3}$$

Donde  $U$  es el valor histórico que se tiene de la variable dependiente que se desea predecir y relacionar con las variables independientes.

Se aplica la condición de primer orden para optimizar y obtener un conjunto apropiado de respuestas llegando a

$$\frac{dE^T E}{dA} = -2X^T U + 2X^T X = 0 \quad (4)$$

Que conlleva a la siguiente expresión general para  $i \geq 1$ , con  $n$  variables y  $m$  cantidad de muestras.

$$A = (X^T X)^{-1} X^T U \quad (5)$$

Note que la expresión (5) corresponde a la relación existente entre los valores históricos de las variables independientes y los valores históricos de la variable dependiente. Las constantes de proporción almacenadas en el vector  $A$  hacen posible realizar la predicción del conjunto de variables dependientes almacenados en  $U$ .

Los resultados obtenidos por medio de la regresión multivariable son anuales, mientras que para las series temporales se tendrá una estimación mensual de consumo eléctrico. Esto debido a que el modelo matricial del primer método relaciona las variables de una forma general y más directa.

### a. Series Temporales

Los modelos de series temporales son algunos de los modelos más empleados para predecir el consumo de energía eléctrica [11]. El objetivo del análisis de series temporales es identificar patrones no aleatorios en la serie temporal de una variable de interés, la observación de este comportamiento pasado permite realizar previsiones sobre el futuro, orientado a

la toma de decisiones [13]. Dos modelos clásicos han sido empleados para desarrollar una serie temporal: el modelo aditivo y el multiplicativo. Las ecuaciones (6) y (7) presentan el modelo aditivo y multiplicativo respectivamente:

$$\hat{y} = T + S + C + R \quad (6)$$

$$\hat{y} = T \times S \times C \times R \quad (7)$$

Dónde:

- $T$  Tendencia.
- $C$  Variaciones cíclicas.
- $S$  Variaciones estacionales o estacionalidad.
- $R$  Variaciones irregulares.
- $Y$  Variable dependiente.

Según el modelo clásico, todas las series temporales se componen de cuatro patrones a saber [17]: La tendencia, las variaciones cíclicas, estacionales e irregulares. La tendencia se define como el comportamiento a largo plazo de la serie. Ésta puede ser causada por el crecimiento demográfico, cambios graduales en el hábito de consumo, o cualquier otro aspecto que afecte la variable de interés a largo plazo. Cuando una serie de datos presenta muchas irregularidades, es común "alisarla" a través de medias móviles, de forma tal, que pueda extraerse información relevante de la serie. Por su parte, las variaciones cíclicas son fluctuaciones en los valores de la variable dependiente con duración superior a un año, y que se repiten con cierta periodicidad. Son resultado de variaciones económicas como periodos de crecimiento o recesión. Las variaciones estacionales son fluctuaciones en los valores de la variable con duración inferior a un año y se repiten todos los años. Generalmente en función

de las estaciones del año, o feriados. Finalmente, las variaciones irregulares son fluctuaciones inexplicables, resultado de hechos fortuitos e inesperados como catástrofes naturales, atentados terroristas, o decisiones del gobierno.

Para el pronóstico desarrollado en este trabajo, se optó por el modelo aditivo que posee una metodología más clara y precisa al momento de calcular las componentes de la serie [17]. Por otro lado, el cálculo de la tendencia y las variaciones estacionales se realizó por medio de medias móviles con ajuste exponencial simple ( $\alpha = 0.75$  y  $k=4$ ), debido a las ventajas que tiene el método para realizar previsiones de corto y largo plazo y a la facilidad de acomodar la serie de datos de manera trimestral. El cálculo de la tendencia se detalla en la siguiente sección. Por último, las variaciones irregulares no fueron consideradas en el modelo, pues son resultado de datos fortuitos teóricamente imprevisibles [11].

#### *Obtención de la tendencia por medio de medias móviles*

Las medias móviles son una forma de calcular la tendencia en una serie temporal. Para identificar la tendencia, se calcula la suma de los primeros periodos o elementos de la serie y se almacena de forma centrada en una columna llamada total móvil. Posteriormente, se calcula el total móvil para el siguiente periodo, esto quiere decir que se elimina el primer término de la serie utilizado y se adiciona el siguiente cuyo resultado ira centrado al igual que en el primer caso. El proceso se repite hasta finalizar con los datos de toda la serie. Posteriormente, se calcula un total móvil para 2 periodos con los resultados obtenidos anteriormente. Esto quiere decir que en una columna diferente

se calcularan los nuevos totales móviles centrados de acuerdo con la metodología anteriormente explicada. Luego de ello, y para finalizar, se dividen los totales móviles de 2 periodos en la cantidad de conjuntos por periodo utilizados, es decir que si el periodo es 4 se debe dividir en 8 porque se agruparon 2 conjuntos de 4 periodos para calcular los índices.

Para una serie de datos mayor, se grafica el periodo y las medias móviles resultantes. Esto dará como resultado la línea de tendencia de los datos suavizada de acuerdo con la media móvil escogida.

Existen tres tipos de medias móviles aplicadas al pronóstico de energía eléctrica: promedios móviles, promedio móvil doble y promedio móvil ponderado. De acuerdo a [1], es recomendable trabajar con medias móviles con ajuste exponencial debido a que los datos de la serie presentan tendencia y estacionalidad, y algunos de los periodos no son estables.

#### *Obtención de la tendencia por medio de medias móviles con ajuste exponencial*

El ajuste exponencial simple es otra forma de obtener la tendencia de una serie temporal. Presenta algunas ventajas respecto a las medias móviles:

- Permite realizar previsiones de corto plazo (para el periodo siguiente de la serie), lo que no es posible por medias móviles.
- Tiene en cuenta todos los valores previamente observados y no solamente los más cercanos, como ocurre con las medias móviles.

Existen tres tipos de ajuste exponencial aplicados a los modelos de previsión en



series temporales: La atenuación exponencial ajustada a la tendencia o modelo de Holt, la atenuación exponencial ajustada a la tendencia y a la variación estacional o modelo de Winter, y el ajuste exponencial simple.

El ajuste exponencial simple es el más popular y el más utilizado debido a que responde de la mejor manera ante cambios fuertes en la demanda [1], brindando mejor precisión a los resultados obtenidos. Se trabajó con este modelo para el cálculo de la tendencia.

Por otro lado, el ajuste exponencial simple suministra una media móvil exponencialmente ponderada a lo largo de la serie temporal, es decir, cada previsión o valor ajustado depende de todos los valores previos. Para realizar el ajuste exponencial simple basta aplicar la ecuación (8) para un periodo de tiempo cualquiera [17]

$$F_i = W \cdot D_i + (1 - W) \cdot F_{(i-1)} \quad (8)$$

Dónde:

- $i$  Periodo de tiempo de la serie de datos
- $D_i$  Valor de la serie original en el periodo  $i$
- $F_i$  Valor de la serie exponencialmente ajustado en el periodo  $i$
- $F_{(i-1)}$  Valor de la serie exponencialmente ajustado en el periodo  $i - 1$  (periodo anterior)
- $W$  Constante de regularización o coeficiente de ajuste ( $0 < W < 1$ )

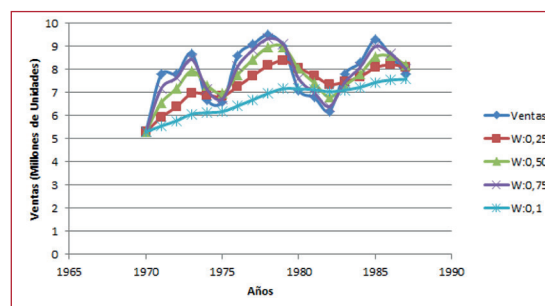
Se considera que el primer valor de la serie original será igual al primer valor ajustado ( $F_0 = D_1$ ), esto significa que el

ajuste realmente comienza a partir del segundo periodo de la serie.

La selección de la constante de regularización  $W$  es crucial para el ajuste exponencial y se selecciona de forma heurística, es decir, según el conocimiento de la persona que realiza la predicción. No obstante, es posible establecer una regla general de selección [17]

- Si el interés es simplemente obtener la tendencia, eliminando el efecto de las otras componentes, el valor deberá ser cercano a cero.
- Si hay interés en realizar una estimación con la serie, es recomendable que el valor de  $W$  sea más cercano a uno, a manera de reflejar mejor el comportamiento de la serie a corto plazo.

Considere las ventas de un determinado producto, por año. La representación gráfica de los resultados obtenidos se presenta en la Figura 1. El gráfico que presenta la tendencia respecto a los diferentes valores de  $W$  es:



**Figura # 1.** Ajuste exponencial con varios valores de  $W$

Cuanto menor es el valor de  $W$  más 'suavizada' es la serie temporal, de esta forma es posible prever el comportamiento a largo plazo de la serie siendo este creciente, decreciente o estacionario.

### *Obtención de la estacionalidad de una serie temporal*

Hay varios métodos para la obtención de las variaciones estacionales, entre ellos el método de la razón para la media móvil (o método de la media móvil porcentual) que consiste en:

1. Obtener medias móviles de orden igual al número de periodos estacionales (4 si la serie es trimestral, 12 si es mensual)
2. Obtener medias móviles de 2 periodos, centradas, a partir de las medias móviles calculadas en el paso uno.
3. Obtener los índices estacionales para cada periodo estacional, es decir que, de acuerdo con el modelo aditivo, sustrayendo de los valores originales de la serie las medias móviles centradas calculadas en el paso dos, [17].
4. Organizar los valores estacionales de los periodos similares, es decir, todos los índices de enero, febrero, marzo, etc.
5. Calcular la media aritmética de los valores correspondientes al periodo estacional (media de los índices obtenidos en todos los meses enero de la serie, por ejemplo)
6. Hacer las correcciones necesarias para que la suma de los índices sea igual a cero.

Los pasos 1 y 2 son virtualmente idénticos al procedimiento para la obtención de la tendencia por medias móviles anteriormente comentado.

En el modelo aditivo, si todos los índices estuvieran próximos o exactamente iguales a cero, las componentes estacionales parecerían no ejercer gran efecto sobre la serie. Por otra parte, si los índices fueran diferentes de cero, ya sea positivo o negativo, el valor de la tendencia sería modificado por ellos, indicando influencia de las componentes estacionales en la serie. En algunos casos, se sugiere realizar pequeñas correcciones para garantizar tal comportamiento dependiendo de la serie de datos con la que se esté trabajando [17].

### *Obtención de las variaciones cíclicas e irregulares*

Es importante resaltar que algunos autores no mencionan las variaciones cíclicas porque en ciertos casos la serie temporal necesita abarcar décadas para que sea posible identificar el comportamiento cíclico [18], y especialmente en series socio económicas los datos más antiguos pueden ser anticuados y contribuir para la construcción de un modelo irreal. Sin embargo, se optó por tenerlas en cuenta esperando que los datos tomados presenten variaciones cíclicas en cortos periodos de tiempo y de esta forma obtener un modelo completo.

Las variaciones cíclicas e irregulares son obtenidas a través de la remoción de las componentes tendencia y estacionalidad de la siguiente forma:

En el modelo aditivo:

$$C = Y - T - S \quad (9)$$

En el modelo multiplicativo:

$$C = \frac{Y}{(T \cdot S)} \quad (10)$$



Dónde:

- $Y$  Valor original de la serie
- $T$  Tendencia
- $S$  Componente estacional

Es costumbre construir un gráfico de variaciones cíclicas en donde se pueda identificar si los ciclos realmente influyen la serie, cuál es su periodicidad, y si el efecto de tales variaciones es muy grande.

Otro aspecto importante en las series temporales es que la previsión de tendencia para largo plazo exige una ecuación (como la de la recta) donde se substituyen los valores de los periodos para los cuales se desea hacer la estimación y se adicionan únicamente los índices sazonales, y las variaciones cíclicas (Si existen en la serie) [17] y [9].

Las variables con las cuales se trabajaron los modelos son:

- Histórico del PIB suministrado por el banco de la república.
- Histórico de crecimiento poblacional colombiano suministrado por el DANE.
- Histórico de consumo residencial en Colombia suministrado por los reportes de la UPME.

Para la regresión multivariable es necesario trabajar con las tres variables anteriormente mencionadas, mientras que para las series temporales es posible calcular la tendencia, los índices estacionales y las variaciones cíclicas únicamente con el histórico de consumo residencial colombiano.

### 3. Resultado de los Modelos

Es importante resaltar que el histórico de datos consultados de población colombiana, producto interno bruto y consumo de energía eléctrica residencial es de libre acceso y cualquier persona puede hacer uso de esta información. Fueron consultadas las bases de datos del banco de la república para obtener el histórico de PIB colombiano, el Departamento Nacional de Estadística (DANE) para consultar el crecimiento poblacional, y XM (empresa que opera y administra el mercado eléctrico colombiano) para obtener el histórico de consumo de energía eléctrica residencial. Todos los datos se obtuvieron para los años 2005 a 2013 [15] y [16]. Los datos empleados en este trabajo pueden ser consultados en el anexo 1.

La información mencionada anteriormente se empleó para analizar cuatro casos de estudio donde se evaluó el desempeño de la regresión multivariable y las series temporales para la predicción del consumo de energía eléctrica en Colombia en el corto y largo plazo.

#### Caso 1: Regresión Multivariable para Corto Plazo

Después de haber modelado la ecuación (5) en Matlab, siendo el vector columna con los datos de consumo de energía eléctrica residencial en GWh (2005-2013) y la matriz que relaciona la población y el PIB per cápita (2005-2013) se obtiene el vector columna que relaciona las 3 variables. Se procede a crear la expresión lineal (ec. 11) que dará como resultado la previsión para el año 2014.

$$Y = a_0 + a_1 * \text{año} + a_2 * \text{Población} + a_3 \quad (11)$$

Dónde:

$$U = \begin{bmatrix} 4069,58 \\ 4218,66 \\ 4372,25 \\ 4540,5 \\ 4693,41 \\ 4861,3 \\ 5034,41 \\ 5221,66 \\ 5393 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & 2005 & 42888590 & 7931153 \\ 1 & 2006 & 43405960 & 8361479 \\ 1 & 2007 & 43926930 & 8832464 \\ 1 & 2008 & 44451150 & 9037877 \\ 1 & 2009 & 44978830 & 9079360 \\ 1 & 2010 & 45509580 & 9329881 \\ 1 & 2011 & 46044600 & 9829122 \\ 1 & 2012 & 46581820 & 10108664 \\ 1 & 2013 & 47121090 & 10487056 \end{bmatrix}$$

$$A = \begin{bmatrix} 1,0e + 0,6 \\ 1,0773 \\ -0,0006 \\ 0,0000 \\ 0,0000 \end{bmatrix}$$

$$Y = 5454.7\text{GWh}$$

Para validar la precisión del modelo, se realizó una comparación del resultado obtenido con las previsiones que la UPME expide en sus reportes. En la Tabla 1 se muestran los resultados estimados y el valor real en GWh para el año 2014.

**Tabla # 1.** Regresión multivariable- corto plazo

| Año  | UPME | Calculada | Valor real 2014 | ε1    | ε2    | ε3    |
|------|------|-----------|-----------------|-------|-------|-------|
| 2014 | 5577 | 5454,7    | 5690            | 1,98% | 4,13% | 2,19% |

La UPME estimo el consumo de energía eléctrica residencial para el año 2014 en 5577GWh, cuando el valor real fue de

5690GWh. El primer error calculado (ε1) es la previsión de la UPME respecto al valor real, el segundo error (ε2) es el valor obtenido con la metodología propuesta respecto al valor real y el último (ε3) hace referencia al valor obtenido con la metodología propuesta respecto al valor de la UPME.

La metodología usada para el corto plazo presento una precisión esperada como lo sugería [14], y fue de 4,13% respecto al valor real. Hay que tener en cuenta que la metodología usada por la UPME presenta ajustes de corrección mensual que reduce la incertidumbre en los resultados y los hace más precisos.

**Caso 2: Regresión Multivariable para Largo Plazo**

El contraste entre crecimiento económico, la demanda de energía y la población colombiana evidencia una pérdida de correlación entre las tres variables para el largo plazo. Esto significa que la variable de población crece de manera abrupta respecto a la variable de PIB lo que puede generar resultados atípicos [19].

El valor calculado obtenido para la estimación a largo plazo registrado en la Tabla 2 se obtuvo con la misma expresión lineal (11) utilizada para el corto plazo y con los datos del año 2030, tales como población colombiana que será de 53.175.000 habitantes, y PIB colombiano igual a 14,611.270 pesos constantes [15] y [16].

**Tabla # 2.** Regresión multivariable- Largo plazo (GWh)

| Año  | UPME | Calculada | ε1     |
|------|------|-----------|--------|
| 2030 | 8958 | 4036      | 54,94% |

**Caso contrario sucedió con la previsión a largo plazo. El modelo utilizado no fue**

**útil, pues se trabajó bajo la premisa de PIB y población estimados del año 2030 que generaron un error ( $\epsilon_1$ ) aproximado del 55% respecto a la previsión de la UPME.**

### Caso 3: Series Temporales para Corto Plazo

Se procede a organizar el histórico de consumo de energía eléctrica en trimestres dando como resultado 36 periodos y trabajando con una media móvil igual a 4. A partir de ello se calcularon las respectivas componentes de la ecuación (6) de acuerdo con lo sugerido en [17].

Teniendo el histórico anual y el consumo de energía eléctrica se debe realizar una regresión exponencial (debido al comportamiento que presentan los datos) para cada mes del año y de esta forma conocer la ecuación característica que modelara la demanda de energía eléctrica. Luego de ello, se procede a adicionar los índices de estacionalidad, y variaciones cíclicas.

El resultado de la regresión para corto plazo se muestra en la Tabla 3:

**Tabla # 3.** Series Temporales- Corto plazo

| Mes        | GWh  |           |           |
|------------|------|-----------|-----------|
|            | UPME | Calculada | Error (%) |
| Enero      | 5402 | 5897,11   | 9,17      |
| Febrero    | 5210 | 5792,7    | 11,18     |
| Marzo      | 5630 | 5376,82   | 4,50      |
| Abril      | 5509 | 6163,72   | 11,88     |
| Mayo       | 5589 | 4269,82   | 23,60     |
| Junio      | 5390 | 5422,99   | 0,61      |
| Julio      | 5670 | 6914,92   | 21,96     |
| Agosto     | 5690 | 5714,54   | 0,43      |
| Septiembre | 5702 | 6944,71   | 21,79     |
| Octubre    | 5640 | 5503,35   | 2,42      |
| Noviembre  | 5680 | 6036,96   | 6,28      |
| Diciembre  | 5810 | 3646,7    | 37,23     |

**Tabla # 4.** Series Temporales- Corto plazo

| Año  | UPME | Calculada | Valor real 2014 | $\epsilon_1$ | $\epsilon_2$ | $\epsilon_3$ |
|------|------|-----------|-----------------|--------------|--------------|--------------|
| 2014 | 5577 | 5640,37   | 5690            | 1,98%        | 0,87%        | 1,13%        |

Se observa que el comportamiento de consumo eléctrico no sigue un modelo matemático del todo exponencial por lo que hay meses del año que presentan errores demasiado altos, como lo es el caso de diciembre, mayo y julio. Sin embargo, se consideran buenos los resultados porque al adicionarse las componentes temporales hace que el pronóstico a corto plazo sea mucho más preciso.

Se calcula el promedio obtenido de los 12 meses para conocer cuál es el valor anual de consumo para el año 2014 como se muestra en la Tabla 4.

El primer error calculado ( $\epsilon_1$ ) es el error obtenido por la previsión de la UPME respecto al valor real, el segundo error ( $\epsilon_2$ ) es el valor obtenido con la metodología propuesta respecto al valor real y el último ( $\epsilon_3$ ) hace referencia al valor obtenido con la metodología propuesta respecto al valor de la UPME.

La metodología usada para el corto plazo presento una precisión esperada como lo sugería [12], y fue de 0,87% respecto al valor real. Las series de tiempo permitieron tener una mayor precisión respecto al resultado esperado, pues el método de medias móviles con ajuste exponencial con el cual se calculó la tendencia tenía en cuenta la totalidad de los periodos analizados.

### Caso 4: Series Temporales para Largo Plazo

Para la obtención de las proyecciones de demanda de energía eléctrica, la UPME

emplea una combinación de modelos econométricos y series temporales a fin de obtener la mejor aproximación para el horizonte de tiempo [9].

Para realizar la previsión de consumo a largo plazo, se procede a utilizar las expresiones características de los 12 meses y los mismos índices que se calcularon para el corto plazo. La covarianza de los 12 meses es de aproximadamente 0,98 lo que representa la buena relación existente entre las variables. La previsión realizada será para el año 2030 y se observa en la Tabla 5:

**Tabla # 5.** Series Temporales-Nueva metodología: Largo plazo (GWh)

|            |          |
|------------|----------|
| Enero      | 10317,37 |
| Febrero    | 10468,99 |
| Marzo      | 9337,94  |
| Abril      | 10583,98 |
| Mayo       | 7352,08  |
| Junio      | 9415,2   |
| Julio      | 11936,14 |
| Agosto     | 9915,44  |
| Septiembre | 12101,91 |
| Octubre    | 9801,02  |
| Noviembre  | 10829,92 |
| Diciembre  | 6588,59  |
| Promedio   | 9887,39  |

**Se procede a calcular el promedio de los 12 meses para así saber el valor exacto anual de consumo que se tendrá para el año 2030 como se observa en la Tabla 6.**

**Tabla # 6.** Series Temporales-Nueva metodología: Largo plazo

| Año  | GWh  |           | ε1     |
|------|------|-----------|--------|
|      | UPME | Calculada |        |
| 2030 | 8958 | 9887,39   | 10,37% |

Para el largo plazo se obtuvo un comportamiento poco deseado debido a que el consumo de algunos meses del año se comportaba de forma lineal, como el caso de julio. Se logró una precisión con el modelo de 10,37%, Mucho mejor que la obtenida con el modelo de regresión multivariable. De cualquier forma, es posible que el error disminuya con respecto al consumo real que se de en este periodo.

## 4. Conclusiones

Cabe resaltar que los eventos de gran consumo energético en Colombia en el año 2014 marcaron diferencia en la curva de estimación para los dos métodos utilizados. Esta variación permitió que los resultados obtenidos en las diferentes previsiones tuvieran errores considerables. Para evitar ello, los pronósticos realizados por la UPME cuentan con seguimientos mensuales y correcciones de los modelos para que sean tenidos en cuenta eventos atípicos en las variables y de esta forma lograr una precisión mayor. Los modelos aplicados por esta entidad cuentan con variables de incertidumbre tales como consumo eléctrico en días feriados y días atípicos de consumo que minimizan el error gradualmente.

Algunas de las falencias observadas en los dos modelos parten en la poca correlación que tienen sus variables a largo plazo. Es decir, que según las fuentes

consultadas la población aumenta a un ritmo mayor que el PIB para los años 2020 y 2030. Adicional a ello, los modelos no tienen en cuenta variaciones de consumo debido a días festivos, o factores climatológicos que han de ser tenidos en cuenta.

Se observó además que la precisión del modelo por regresión multivariable dependía directamente de la cantidad de variables que fueran utilizadas. Para nuestro caso, las tres variables permitieron obtener errores bajos en los pronósticos a corto plazo, mientras que para el largo plazo el modelo presentaba un error aproximado del 55% respecto a la estimación dada por la UPME debido a la inconsistencia entre las variables que se tuvieron en cuenta para el análisis de largo plazo. Sin embargo, este error puede disminuir con respecto al consumo real. Se recomienda a los futuros trabajos e investigaciones en el área de estimación del consumo eléctrico a corto y largo plazo trabajar con modelos de pronóstico que incluyan una variable de aleatoriedad tal como el consumo eléctrico en días festivos con el fin de reducir el error debido a situaciones imprevistas, así como trabajar con series temporales diarias que permitan aprovechar al 100% la información de los históricos disponibles.

## 5. Referencias

- [1] Ariza, A. M, "Métodos utilizados para el pronóstico de la demanda de energía eléctrica en sistemas de distribución," Tesis de pregrado, Pereira, Colombia, 2013.
- [2] Barrientos, A. F., & Olaya, J, "Un modelo "spline" para el pronóstico de la demanda de energía eléctrica," Escuela de Ingeniería Industrial (Cali), Colombia, pp. 187-202, 2007.
- [3] Gómez de la Calle, M, "Modelo de previsión de demanda de electricidad de largo plazo," Madrid, España, 2010.
- [4] Cobo berbera, I, "Modelo de previsión de la demanda de corto plazo," Madrid, España, 2013.
- [5] Salas, J. S., Desarrollo con Matlab de una Red Neuronal para Estimar la Demanda de Energía Eléctrica. Valladolid: Universidad de Valladolid, España, 2016.
- [6] Chen, Toly; and Wang, Yu-Cheng. Long-Term Load Forecasting by a Collaborative FuzzyNeural Approach. International Journal of Electrical Power & Energy Systems, 12, 2012, vol. 43, no. 1. pp. 454-464. ISSN 0142-0615.
- [7] Agência nacional de energia elétrica (ANEEL), 2014. [Online], disponible: [www.aneel.gov.br](http://www.aneel.gov.br)
- [8] Morales Udaeta, M. E, "Planejamento integrado de recursos energéticos-PIR-para o setor elétrico (pensando o desenvolvimento sustentável)," São Paulo, Brasil, 1997.
- [9] Unidad de Planeación Minero Energética (UPME), "Proyecciones regionales de demanda de energía eléctrica y potencia máxima en Colombia," Bogotá, Colombia, 2009.
- [10] Rodríguez Betancourt, R., & Arrieta Gallardo, M, "Econometría Moderna," México D.F, México, 2011.
- [11] Mejía, V. M, "Predicción del consumo de Energía Eléctrica en Colombia utilizando modelos no lineales," Universidad Nacional, Medellín, Colombia, 2011.
- [12] J.J Gómez-Cadenas, "El método de los mínimos cuadrados," Curso de estadística, Bogotá, Colombia, 2005.

- [13] De la Fuente Fernández S., "Estadística teórica I: Series temporales," Universidad Autoónoma de Madrid, España, 2014.
- [14] J.M. Rojo Abuin, "Regresión lineal múltiple," Instituto de Economía y Geografía, Madrid, España, 2007.
- [15] Departamento Nacional de Estadística (DANE), "Censo poblacional Colombiano," Colombia, 2014.
- [16] Banco de la República de Colombia, "Producto interno bruto Colombiano a largo plazo (PIB)," Colombia, 2014.
- [17] M. Meneses Reis., "Análise de series temporais," Universidade Federal de Santa Catarina (UFSC), Florianópolis, Brasil, 2013.
- [18] F. J. Barón López, F. Téllez Montiel, "Apuntes de bioestadística," Capítulo 6, pp. 35-40, Málaga, España, 2014.
- [19] Cárdenas, M., & Junguito, R, "Nueva introducción a la economía colombiana," Texto básico sobre economía colombiana (Bogotá), Colombia, pp. 2-12, 2013.