



Pontificia Universidad Javeriana

Facultad De Ciencias Económicas Y Administrativas

Trabajo para optar por el título de Magister en Economía

Título:

**Estimación de los Costos de las Plantas de Tratamiento de la Empresa de Acueducto
y Alcantarillado De Bogotá -ESP**

Estudiante:

RODRIGO ALFONSO MILLÁN SERNA

Asesor:

FLAVIO HERNADO JACOME LIEVANO

Bogotá D.C.

noviembre 2022



Contenido

Introducción	2
Contexto: Funcionamiento de las Plantas de Agua	5
Marco Teórico	6
Datos y Estadística Descriptiva	9
Metodología	11
Modelo de Regresión Lineal y Estimación por Mínimos Cuadrados	11
Series de Tiempo	12
Modelos con Errores ARIMA	14
Resultados	16
Conclusiones	19
Referencias	21
Anexos	24

Introducción

Este trabajo de grado estudia la eficiencia económica de las plantas de tratamiento de la Empresa de Acueducto y Alcantarillado de Bogotá. Para ello, se estiman los costos agregados de los dos Sistemas de las plantas de tratamiento, norte; compuesto por las plantas de Wiesner y Tibitoc y, sur; compuesto por las plantas Vitelma, Dorado y Yomasa, en los que están agrupadas las plantas de tratamiento de agua para consumo humano de la Empresa de Acueducto y Alcantarillado de Bogotá.

Las estimaciones de los costos son una herramienta adecuada por cuanto permite aproximar costos mínimos de producir cantidades específicas de un bien (agua), a los precios de los factores e insumos, y con información tecnológica dada.

Las estimaciones de los costos tienen una justificación adicional que consiste en que son insumo para la fijación de tarifas para los consumidores de agua. En las metodologías usuales de tarifas con base en costos, éstas se igualan a los Costos Medios, es decir a los Costos Totales Óptimos divididos por las cantidades producidas. De aquí la importancia de contar con las estimaciones de estos Costos Mínimos.

En términos generales, los costos de potabilización del agua y tratamiento de aguas residuales se pueden clasificar en costos de inversión y de funcionamiento. Los primeros, a su vez, se subdividen en costos de inversión inicial, que incluyen aspectos como los estudios preliminares, el desarrollo de la inversión, la interventoría, la conservación y manejo de la cuenca y, el desarrollo o fortalecimiento institucional. Adicionalmente, los costos de inversión futura se deben ejecutar para poder asegurar el suministro del agua tratada a todos los habitantes de la ciudad. Además de los costos de inversión anteriormente descritos, también se destacan los costos de reposición, de depreciación, de financiación y los ambientales.

Por su parte, los costos de funcionamiento se subdividen en costos de operación y mantenimiento y los costos de administración. Los primeros se asocian con la mano de obra, la adquisición de insumos químicos, el uso de energía eléctrica y el desarrollo de actividades de control de la calidad del agua y del proceso de tratamiento. En cuanto al costo de administración, se incluye el mantenimiento de los equipos, el personal administrativo y los gastos generales (Aristizábal & Galvis, 2012).

Las inversiones en el sector de agua potable y saneamiento básico son cuantiosas, especialmente en las zonas rurales y en los pequeños y medianos municipios el problema es mayor, la estimación de los costos totales comprende la inversión inicial y futura, costos de reposición, administración, operación y mantenimiento de las alternativas de potabilización de agua, para poblaciones del sector urbano, industrial y rural tanto pequeño como mediano.

Estimación de los Costos de las Plantas de Tratamiento de la Empresa de Acueducto y Alcantarillado de Bogotá -ESP

Resumen

Este trabajo estudia la eficiencia económica de las plantas de tratamiento de la Empresa de Acueducto y Alcantarillado de Bogotá. Para ello, se estiman los costos de los dos Sistemas de las plantas de tratamiento: Norte (plantas de Wiesner y Tibitoc) y Sur (plantas de Vitelma, Dorado y Yomasa). Al ajustar modelos de regresión con errores ARIMA, y al analizar las elasticidades obtenidas para cada costo en cada sistema, se encuentra que el sistema sur es más eficiente en el uso de los insumos, mientras que el sistema norte es más eficiente administrativamente. Por tanto, y teniendo en cuenta los pesos relativos de cada uno de los factores, se encuentra que el sistema sur es más eficiente que el sistema norte.

Palabras Clave: Planta de tratamiento de Agua, Modelos Arima, Costos

Clasificación JEL: C22

Abstract

This paper studies the economic efficiency of the treatment plants of Empresa de Acueducto y Alcantarillado de Bogotá. To do this, we will estimate the cost of the two treatment plants systems, North (Wiesner and Tibitoc plants) and South (Vitelma, Dorado and Yomasa plants). By adjusting regression models with ARIMA errors, and analyzing the elasticities obtained for each cost in each system, we found that the southern system is more efficient in the use of inputs, while the northern system has a more efficient administrative structure. Therefore, and considering the relative weights of each factor, we find that the southern system is more efficient than the northern system.

Key Words: Water treatment plants, Arima Models, Cost

JEL Classification: C22

Contexto: Funcionamiento de las Plantas de Agua

Son tres las principales tecnologías de producción de agua potable: la combinación de barreras múltiples, con la cual se logran bajas condiciones de riesgo; el tratamiento integrado de dichas tecnologías, que busca producir un efecto esperado y, las tecnologías de tratamiento por objetivo en el cual cada una de las etapas siguen una meta específica (Sedimentación, coagulación, floculación, sedimentación, filtración y desinfección).

En cuanto a la clasificación de las plantas de tratamiento de agua potable (PTAP), estas pueden catalogarse como de tecnología convencional, las cuales se caracterizan por incorporar los procesos de coagulación, floculación, decantación y filtración. Por otra parte, las plantas de filtración directa incluyen proceso de coagulación-decantación y de filtración rápida, aunque algunas incorporan en ocasiones una etapa de floculación. Por último, las PTAP de filtración en múltiples etapas o FIME, incorporan etapas de filtración gruesa dinámica, filtración gruesa ascendente y filtración lenta en arena (Romero, 2008).

Las etapas de coagulación y floculación se utilizan para eliminar los sólidos de suspensión. Por un lado, en la coagulación se agrega cierta sustancia al agua para hacer cambiar el comportamiento de las partículas en suspensión, haciendo que estas se atraigan y conformen un cúmulo agregado. Por otro lado, en la floculación, proceso siguiente al de coagulación, se agita de manera suave y lenta el agua, generando un mayor contacto entre las partículas y la unión de estas, las cuales se eliminan posteriormente por medio de la sedimentación o filtración (López, 1997).

Ahora bien, durante la sedimentación algunos sólidos en suspensión y algunos patógenos se depositan en el fondo del contenedor de agua, esto como producto del efecto de gravedad y las diferencias sustanciales asociadas a la densidad del material contaminante. La filtración incluye

un tamizado mecánico, a través de filtros de arena lentos. Dichos filtros extraen los sólidos en suspensión más pequeños, los patógenos y determinados productos químicos, como sabores y olores (Pizzi, 2007).

En el caso explícito de la ciudad de Bogotá D.C, la Empresa de Acueducto y Alcantarillado de Bogotá (EAAB-ESP) cuenta con cinco plantas de tratamiento de agua potable, las cuales poseen la capacidad de tratar el agua para Bogotá y 11 municipios aledaños.

Marco Teórico

Desde la teoría económica, los análisis de Costos son una herramienta que permite caracterizar el costo de producción de una empresa dadas unas combinaciones de costos de los factores y niveles de producción (Nicholson, 2005). Así, la aproximación por costos evita tener que construir Funciones de Producción (y demandas de factores) como pasos previos a la descripción del comportamiento económico de las empresas y garantiza consistencia con un comportamiento racional de minimización de costos. En adición, la Función de Costos permite obtener resultados de estática comparativa sin requerir supuestos sobre la divisibilidad de los bienes o sobre convexidad de los conjuntos de posibilidades de producción (McFadden, 1978).

Es así como el análisis de costos permite considerar tres tipos de funciones que pueden ser implementadas a nivel econométrico, las cuales son: funcionales lineales, cuadráticas o polinómicas (Greer, 2010), estas permiten incluir una mayor cantidad de factores de producción, así mismo su linealización puede ser lograda mediante una transformación logarítmica como lo es el método de mínimos cuadrados.

Adicionalmente, y tal como se referenció previamente, los parámetros pueden ser interpretados directamente como elasticidades de producción, por lo que no se necesitan hacer transformaciones.

De hecho, la aproximación a este problema mediante costos ya ha sido estudiado por Revollo y Londoño (2010), quienes plantean aproximaciones por costos linealizados verificando la presencia de economías de escala para los sistemas de acueducto y alcantarillado en Colombia.

Ahora bien, abordando el tema de las PTAP, se debe mencionar que existen distintas tecnologías que llevan a la potabilización del agua, no obstante, no todas cumplen los mismos principios y esto puede afectar enormemente la estructura de costos y a su vez la función de costos, por consiguiente, cada planta de potabilización tiene una función de costos que debe ser analizada por separado.

De esta manera, los costos de producción o tratamiento se pueden clasificar en una función de costos que incluya factores de inversión y de funcionamiento. Por ente, los factores de inversión deben incluir costos asociados a los factores de inversión inicial que incluyen aspectos como los estudios preliminares, el desarrollo de la inversión, la interventoría, la conservación y manejo de la cuenca y, el desarrollo o fortalecimiento institucional, seguidamente los costos de inversión futura se deben ejecutar para poder asegurar el suministro del agua tratada a todos los habitantes de la ciudad.

Además de los costos de inversión anteriormente descritos, también se destacan los costos de reposición, de depreciación, de financiación y ambientales.

Por otra parte, los costos de funcionamiento se subdividen en costos de operación y mantenimiento y, los costos de administración. Los primeros se asocian con la mano de obra, la adquisición de insumos químicos, el uso de energía eléctrica y el desarrollo de actividades de control de la calidad del agua y, del proceso de tratamiento. En cuanto al costo de administración, se incluye el mantenimiento de los equipos, el personal administrativo y los gastos generales (Aristizabal & Galvis, 2012).

Desde la teoría económica, la Función de Costos Totales es una herramienta que permite caracterizar el costo de producción de una empresa dadas unas combinaciones de costos de los factores, niveles de producción y los costos mínimos de existencia de la empresa (Nicholson, 2005), por lo que se convierte en una buena aproximación para abordar el problema del costeo en las PTAP.

Frente a las funciones de costo, se les otorga un papel preponderante a las funciones lineales tipo Cobb-Douglas (Cobb and Douglas, 1928), donde se asume que el tipo de rendimientos a escala, elasticidad constante de producción y que los cambios tecnológicos son neutrales en el sentido de Hicks (Redondo, 2011), lo que facilita su análisis e interpretación. Una vez definida la forma funcional adoptada para la función de costos, calcularemos la función de costos medios, que nos permite conocer los costos de cada unidad, al dividir la función de costos totales entre el número de unidades producidas (Nicholson, 2005).

En efecto, esta metodología tiene como ventaja que permite entender cómo se comporta el costo para diferentes unidades de producción y facilita la comparación entre diferentes estructuras de producción (por ejemplo, en este caso diferentes plantas de agua), pero tiene como desventaja que, al ser un método de costeo agregado, no permite saber el impacto de cada unidad producida en los diferentes componentes del costo. En adición, en el caso de que los supuestos asumidos sobre la fórmula funcional no se cumplan, podríamos incurrir en problemas de especificación y estimación del modelo. No obstante, la comprobación de estos supuestos no se incluía dentro de los alcances de esta tesis y a la fecha no existe un compendio bibliográfico que permita contradecir o no, los supuestos detrás de la función Cobb-Douglas para el análisis de costos de las PTAP de Colombia. Sin embargo, futuros estudios podrían profundizar sobre este aspecto.

Datos y Estadística Descriptiva

Los datos fueron obtenidos mediante una solicitud expresa a la Dirección de abastecimiento de la EAAB-ESP, la cual remitió información completa para el periodo de enero de 2006 a diciembre de 2018. A continuación, las variables de costos fueron deflactadas a valores de diciembre de 2018 con el fin de asegurar su comparación y, de evitar que la información reflejara cambios en los precios y no en las cantidades.

De esta forma, se obtuvieron los datos de costo de energía (CE), costos de los químicos para el tratamiento del agua (CQ), la Tasa de Uso del Agua (TU) y costos administrativos (CA) mensuales de todas las plantas de la red de acueducto de Bogotá, para el periodo comprendido entre enero de 2006 y diciembre de 2018. Se debe señalar que en los costos administrativos se incluye lo referente a personal, casino y vigilancia.

En valores corrientes, el costo con mayor peso es el costo de los químicos, seguido por el costo de la energía, el costo administrativo y finalmente el costo de la tasa de uso del agua. Sin embargo, este análisis incluye los efectos de los precios, por lo que a continuación se analiza el comportamiento de cada una de estas variables descontando el efecto de los precios, por lo que todos estos valores son deflactados mediante la variación del IPC con año base diciembre de 2018, para tener en cuenta el crecimiento real de los costos y no los efectos de la inflación, se expresaron en millones de pesos para facilitar su lectura y su comprensión.

Igualmente, se tiene en cuenta la cantidad suministrada de agua (AS) de cada planta, con el fin de medir la producción real de los dos sistemas; norte y sur, dada su estructura actual de costos, la cual se encuentra medida en millones de metros cúbicos. Adicionalmente, para medir el impacto de la financiación de la construcción de las plantas se utiliza como variable proxy la

tasa de los Depósitos a Término Fijo (DTF), que es la tasa de captación de referencia del mercado financiero colombiano.

En la Gráfica 1 se observan los costos totales de todo el sistema de las plantas de tratamiento de la EAAB-ESP. Se evidencia que estos han presentado gran variabilidad durante todo el periodo de estudio, la cual se ha incrementado a partir de 2014, junto con una tendencia al alza en el costo. En lo que respecta a los costos administrativos (Gráfica 2), estos han tenido una media constante a lo largo del tiempo, con excepción del último año, en donde se observa un leve incremento de los costos. El costo de la energía (Gráfica 3) ha tenido una gran variabilidad, manteniendo una tendencia al alza durante el periodo, la cual se ha acelerado desde 2014. En la Gráfica 4, se observa como la línea de tendencia del costo medio de químicos disminuye levemente desde enero de 2006 hasta 2018. En la Gráfica 5, se observa como la línea de tendencia de la Tasa de los Depósitos a Término Fijo (DTF) disminuye de aproximadamente 7% en 2006 a un 5% aproximadamente en 2018. En la Gráfica 6 se observa como la línea de tendencia del agua suministrada presenta un aumento de dos millones de metros cúbicos de 2006 a 2018. En la Gráfica 7 se observa como la línea de tendencia de la tasa de uso de agua aumenta de un valor negativo en 2006 a un valor de \$200 millones en 2018.

En la Tabla 1, se encuentra la estadística descriptiva por cada una de las variables. En promedio, durante todo el periodo analizado, el costo promedio total sistemas norte y sur de la EAAB-ESP fue de cerca de 6,6 miles de millones de pesos por mes, con una desviación de 2.3 miles de millones.

Frente a los demás costos, se evidencia que el costo de la energía fue el que más contribuyó al costo total, con un valor promedio de 1,4 miles de millones de pesos, seguido por el costo de los químicos con 1.1 miles de millones de pesos promedio, el costo administrativo

con cerca de 0.9 miles de millones de pesos y la tasa de uso de última con un valor promedio de 0.1 miles de millones de pesos.

El sistema total de las plantas de tratamiento de agua de la Empresa de Acueducto y Alcantarillado de Bogotá se encuentra subdividido a su vez en el sistema norte (compuesto por las plantas Wiesner, Tibitoc) y el sistema sur (Vitelma, Dorado, Yomasa). A partir de este punto analizaremos cuál de estos dos sistemas es más eficiente en sus costos.

Para facilitar la comparación entre ellos, se analizarán los datos de cada sistema en términos de sus costos medios al dividir los costos totales entre el agua suministrada, de manera que podamos comprender cada costo de acuerdo con las unidades producidas por planta. A continuación, se presenta el mismo análisis por sistema, normalizado por el agua suministrada en cada sistema.

Se evidencia que el mayor peso ponderado por unidad se encuentra en los costos de la energía, seguido por los costos administrativos, los costos químicos y por último los costos de la tasa de uso del agua.

A continuación, se explicará la metodología bajo la cual se estimarán los costos de tratamiento de la EAAB-ESP, así como la estrategia empírica que se utilizará para su estimación.

Metodología

Modelo de Regresión Lineal y Estimación por Mínimos Cuadrados

Este modelo parte del supuesto de que una variable Y puede ser explicada mediante una función lineal que depende de unas variables X y un término aleatorio ε , el cual se distribuye normal con media cero, varianza constante y sin correlación serial (Greene, 2002; Wooldrige, 2002) que en forma matricial se representa como:

$$Y = X\beta + \varepsilon \text{ Ec. 2}$$

El objetivo es estimar el vector de parámetros β que determinan el grado de correlación o dependencia (según si es posible o no asumir la causalidad) entre las variables X y la Y . Para esto, el método de estimación más utilizado es el Mínimos Cuadrados, que busca minimizar el cuadrado de la distancia entre cada una de las observaciones y la recta dada por Ec. 3 (Wooldrige, 2002) y se obtiene que los valores de estimados de los parámetros y sus varianzas corresponden a:

$$b = (X'X)^{-1}X'Y \text{ Ec. 3}$$

$$V(b) = s^2(X'X)^{-1} \text{ Ec. 4}$$

con $s^2 = \frac{e'e}{n-p}$ con $e = Y - Xb$, n el número de datos y p el número de variables independientes. De acuerdo con lo demostrado por Gauss-Markov (Rao, 1968; Hayashi, 2000), si todos los supuestos del modelo, que se mencionaron previamente, se cumplen entonces el estimador MCO es el mejor estimador lineal insesgado puesto que va a ser el más eficiente (con menor varianza) entre todos los estimadores lineales insesgados.

Sin embargo, en el caso en el que esto no se pueda garantizar, es necesario recurrir a otro tipo de estimadores para encontrar estimaciones para β que sean eficientes e insesgadas.

Series de Tiempo

Una serie de tiempo es un tipo de datos que surge de la medición repetida, cronológica y generalmente con igual espaciamiento temporal, en un intervalo de tiempo de una o varias características de un individuo (Montgomery, et al., 2015).

Las series de tiempo tienen una serie de componentes, los cuales son importantes porque marcarán su comportamiento. En general, se pueden identificar cuatro componentes principales:

la tendencia, que muestra el comportamiento de largo plazo; el componente estacional, que es corresponde a un comportamiento repetitivo y periódico para un momento específico, el ciclo, que muestra las oscilaciones de los datos durante periodos cortos de tiempo; y el componente aleatorio que no puede ser explicado (Brockwell & Davis, 2002).

La mayoría de los modelos que utilizan series de tiempo tienen como principal supuesto que los errores se comportan como ruido blanco (Hamilton, 1994). Es decir, que el componente aleatorio se distribuya normal, con media cero, varianza constante en el tiempo y no existe correlación en el tiempo, supuesto que no se cumple cuando la serie no es estacionaria, es decir que tiene tendencia, correlación serial o varianza no constante. Debido a que la mayoría de las series no cumplen con estos requisitos, es necesario recurrir a una serie de transformaciones que permiten que la serie obtenga este comportamiento deseado, tales como diferenciaciones (diferencias absolutas o variaciones porcentuales de las series) o a filtros que permitan descomponer las series, tales como el filtro de Hodrick y Prescott (1997) o el filtro STL (Cleveland, et al. 1990).

En general, cuando las series no son estacionarias se dice que son integradas, y el número de diferenciaciones necesarias para convertirla en estacionarias, indicará el orden de integración.

Para el modelamiento de las series se puede recurrir a aproximaciones univariadas (Granger and Newbold, 2014) o multivariadas (Lütkepohl, 2013). Sin embargo, esta tesis está enfocada en análisis de regresión que tienen en cuenta las estructuras de modelamiento del error bajo aproximaciones de series de tiempo univariadas, con el fin de mantener las principales propiedades de las funciones de producción tipo Cobb-Douglas.

Modelos con Errores ARIMA

Dentro del modelamiento univariado de las series, uno de los modelos más utilizado para abordar son los modelos Autorregresivos (Akaike, 1969), denotados como AR(p), que en los que el valor actual de la serie y_t puede explicarse por p valores pasados de la serie tal que:

$$y_t = \alpha + \rho_1 y_{t-1} + \dots + \rho_p y_{t-p} + \epsilon_t \quad Ec. 5$$

Donde ϵ_t es un proceso de ruido blanco y ρ_1, \dots, ρ_p son los parámetros de autocorrelación que indica indican el tipo de autocorrelación de los datos, tal que si $\rho_t = 0$ no existe correlación serial de los datos para el tiempo t, $|\rho_t| < 1$ existe una correlación serial (positiva o negativa) con el tiempo t que se diluye (es decir, que disminuye con el paso de los periodos) y si $|\rho_t| > 1$ la serie es explosiva, que nunca se estabilizará y siempre aumentará o disminuirá en el tiempo. Por lo anterior, se hace necesario imponer una restricción de manera que se consideren únicamente modelos con coeficientes de autocorrelaciones con valores absolutos menores a uno.

En contraposición a estos modelos surgieron los modelamientos de Medias Móviles (Durbin, 1959), denotados como MA(q). Estos modelos parten del supuesto de que la serie está influenciada por observaciones pasadas del error (Kotu and Deshpande, 2018), tal que:

$$y_t = \alpha + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} \quad Ec. 6$$

con ϵ_t un proceso de ruido blanco y $\theta_1, \dots, \theta_q$ son los parámetros del modelo. Por construcción, los parámetros del modelo son siempre estacionarios en media, varianza y covarianza, por lo que no es necesario imponer restricciones al parámetro. Sin embargo, por la condición de invertibilidad (Bentarzi and Hallin, 1994), la ecuación solo tiene solución cuando $|\theta_t| < 1$.

En enfoques más recientes surgieron los modelos Autorregresivos de Medias Móviles (Whittle, 1951; Box et al., 2011), denotados como ARMA(p,q) que corresponden a la unión de los dos modelos anteriores tal que la serie puede ser representada por (Hamilton, 1994):

$$y_t = \alpha + \rho_1 y_{t-1} + \dots + \rho_p y_{t-p} + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} + \epsilon_t \quad Ec. 7$$

Estos modelos heredan las mismas características que los dos modelos anteriores, y puede ser estimado mediante el método de Box-Jenkins.

Finalmente, los modelos más generales en esta familia son los llamados Autorregresivos Integrados de Medias Móviles (Kotu and Deshpande, 2018), denotados como ARIMA(p,d,q), en el cual se puede incluir la operación de diferenciación de la serie, las d veces que sea necesarias, hasta obtener una serie estacionaria:

$$y_t^d = \alpha + \rho_1 y_{t-1}^d + \dots + \rho_m y_{t-m}^d + \theta_1 \epsilon_{t-1}^d + \dots + \theta_q \epsilon_{t-q}^d + \epsilon_t^d \quad Ec. 8$$

donde y_t^d indica la serie con d diferenciaciones. Si bien, múltiple literatura ha señalado que estos modelos tienen un alto poder predictivo (Iqbal and Naveed, 2016; Namin and Namin, 2018). Estos últimos modelos son complejos de interpretar y no permiten incluir variables explicativas.

Como solución a este problema, surgieron los llamados modelos de regresión con errores ARIMA (Pankratz, 1991) en los que se busca explicar la serie en función de otra variable independiente, pero se asume que los errores siguen una especificación ARIMA, tal que

$$y_t = \beta x_t + n_t^d$$

$$\rho_1 n_{t-1}^d + \dots + \rho_p n_{t-p}^d + \theta_1 \epsilon_{t-1}^d + \dots + \theta_q \epsilon_{t-q}^d + \epsilon_t^d \quad \text{Ec. 9}$$

En este caso β corresponde a los parámetros del modelo que muestran la relación entre x_t y y_t y se interpretan igual que en los modelos de regresión lineal. Para la estimación de estos modelos se debe recurrir a la estimación por máxima verosimilitud (Van den Bossche et al., 2004).

Es pertinente resaltar que a todos estos modelos se les debe realizar pruebas de validación de supuestos y bondad de ajuste, antes de realizar cualquier tipo de interpretación.

Bajo esta aproximación se pueden controlar las violaciones a los supuestos del modelo lineal estimado por mínimos cuadrados, sin perder la facilidad de interpretación, especialmente en términos de eficiencias económicas, y los supuestos subyacentes de las funciones de producción tipo Cobb-Douglas.

Resultados

Para la estimación de los costos de los sistemas de tratamiento de agua de la Empresa de Acueducto y Alcantarillado de Bogotá, se planteó la siguiente función:

$$C = f(C_A, C_E, C_q, C_u, r) \quad \text{Ec. 10}$$

En donde C el estimativo del costo medio, y C_A , el costo medio administrativo, que incluye el costo del personal; C_e , el costo medio de la energía; C_q , el costo medio de los insumos para el tratamiento del agua, incluidos los productos químicos; C_u el costo medio de la tasa de uso del agua y r , tasa de interés, que refleja el costo del capital. A continuación, se obtuvieron los logaritmos, con el fin de asegurar una estimación lineal (Ec. 11):

$$\ln(C) = (A) + \alpha \ln(C_A) + \beta \ln(C_E) + \gamma \ln(C_q) + \delta \ln(C_u) + \theta \ln(r) + \varepsilon \quad \text{Ec. 11}$$

Donde los valores de $\alpha, \beta, \gamma, \delta, \theta$ corresponden a las respectivas elasticidades del costo medio a cada uno de los insumos de producción.

Resultados de la Estimación

A partir de la base de datos anterior, se estimó un modelo base de costos, mediante mínimos cuadrados, todo el sistema y los sistemas norte y sur, cuyos resultados se presentan en la Tabla 2. A primera vista, el modelo base para todo el sistema de la EAAB-ESP tiene un buen ajuste, pues su R^2 corresponde al 86% y la prueba de normalidad de Shapiro y Wilks (1965) da normalidad de los errores del modelo. En el caso del sistema norte, encontramos que el ajuste del modelo medido por R^2 corresponde a aproximadamente el 89% y para el sistema sur al 88%. Los errores pasan la prueba de normalidad para el sistema norte, mientras que para el sistema sur se debe rechazar la hipótesis nula de normalidad.

Sin embargo, al realizar la prueba de autocorrelación de Durbin y Watson (1951), encontramos que el valor asociado de la prueba es 0.87 lo que indica que con un nivel de significancia del 1% se rechaza la hipótesis nula de no autocorrelación. Adicionalmente, las pruebas de heterocedasticidad de Breush Pagan (1979) indican la posible existencia de heteroscedasticidad en los datos (con un estadístico de prueba asociado del 27.82 que es significativo con un nivel de significancia del 0.01). Para el sistema norte, existe evidencia de heterocedasticidad (con un estadístico BP de 15.81, que es significativo al 1%) y autocorrelación (con un estadístico DW de 1.11 significativo al 1%), mientras que en sistema sur no existe evidencia de heteroscedasticidad (con un BP de 4.04 y un p valor asociado de 0.54), pero sí de autocorrelación (con un DW de 0.68 que es significativo al 1%).

Adicionalmente, el análisis de los residuos indica que no existe un buen ajuste del modelo, puesto que aún se perciben patrones en el gráfico de residuos y se observa la existencia de posibles datos atípicos (Figura 13).

Por consiguiente, se realizará un segundo modelo, en donde se ajusta un modelo de regresión con errores ARIMA de acuerdo con lo especificado en la sección de metodología. Para este ejercicio se ajustan varios modelos y se selecciona aquel con el criterio de información AIC más bajo (Akaike, 1974).

En la Tabla 3 se ajustan los modelos con errores ARIMA para el sistema total, en donde para cada uno de los datos se seleccionaron los modelos con el criterio de información más bajo, más parsimoniosos y con mejor capacidad de pronóstico respectivamente.

Para el modelo que contiene la información de todo el sistema de la Empresa de Acueducto y Alcantarillado de Bogotá, encontramos que la prueba de normalidad de Shapiro y Wilks (1965) indica que los errores son normales y las pruebas de Ljung - Box (1978) sobre los errores indica la no existencia de problemas de correlación serial del modelo y al aplicar modelos sobre los errores, encontramos que estos se comportan como ruido blanco, por lo que estos resultados ya son interpretables.

En el caso de los sistemas norte y sur, se evidencia que en ambos casos se rechaza la hipótesis nula de errores normales, aparentemente por la presencia de datos atípicos. Sin embargo, para ambos modelos no es posible rechazar la hipótesis nula de no correlación serial bajo las pruebas de Ljung - Box (1978) y en ambos casos los errores se comportan como ruido blanco.

A partir de esto, se encuentra que la elasticidad de CA con respecto a CT es de 0.25, lo que quiere decir que un incremento en un 1% en los costos administrativos medios repercute en

un incremento de 0.25% en los costos medios, mientras que un incremento en un 1% en los costos medios de la energía se reflejan en un incremento de 0.4% en los costos medios, un aumento de un 1% en los costos medios de los químicos implica un alza de 0.25% en los costos medios y un aumento de un 1% en la tasa de uso se refleja en un aumento de 0,03% en los costos medios. No se observan efectos significativos de la DTF ni del suministro de agua en los costos.

A partir de estos modelos, y analizando las elasticidades obtenidas para cada uno de los factores de producción para cada sistema, se encuentra que el incremento de los costos administrativos medios impacta en menor medida en el sistema norte que en el sistema sur, pero el costo medio del uso de la energía y de los químicos tienen un menor impacto en el incremento de los costos medios en el sistema sur.

Debido a su bajo peso por unidad, la tasa de uso no tiene impacto estadístico en ninguna de los dos sistemas. Por tanto, se podría concluir que el sistema sur es más eficiente en el uso de los insumos, mientras que el sistema norte tiene una estructura administrativa más eficiente. Por tanto, en términos generales, y teniendo en cuenta los pesos relativos de cada uno de los factores, se encontró que el sistema sur es más eficiente que el sistema norte.

Ahora bien, desde una visión estratégica empresarial encaminada a la mejora de la EAAB-ESP, se recomendaría que los dos sistemas compartan sus experiencias sobre el uso de factores, para que ambas redes generen potabilización de agua de la manera más eficiente.

Conclusiones

A lo largo de esta investigación se realizó un análisis de los costos medios asociados al proceso de tratamiento de aguas desde la captación, tratamiento y distribución de los sistemas norte y sur de las plantas de tratamiento de agua de la Empresa de Acueducto y Alcantarillado de

Bogotá con el fin de determinar cuál de los dos sistemas es más eficiente en el uso de los factores de producción por unidad producida.

Para esto, se recurrió a un análisis de series de tiempo, teniendo en cuenta que se disponía de la información de los costos medios (usando como unidad de estandarización la cantidad de agua suministrada por cada sistema) del costo de energía (CE), costos de los químicos para el tratamiento del agua (CQ), costos de la Tasa de Uso del Agua (TU) y costos administrativos (CA) mensuales de todas las plantas de los sistemas de la Empresa de Acueducto y Alcantarilla de Bogotá, para el periodo comprendido entre enero de 2006 y diciembre de 2018.

En definitiva, se determinó ajustar modelos de regresión con errores ARIMA, teniendo en cuenta que este tipo de modelos corrigen los errores de estimación propios de los datos tipo serie de tiempo, sin perder la facilidad de interpretación de las variables.

A partir de estos modelos, y analizando las elasticidades obtenidas para cada uno de los factores de producción para cada sistema, se encontró que el incremento de los costos administrativos medios impacta en menor medida en el sistema norte que en el sistema sur, pero el costo medio del uso de la energía y de los químicos tienen un menor impacto en el incremento de los costos medios en el sistema sur. Debido a su bajo peso por unidad, la tasa de uso no tiene impacto estadístico en ninguno de los dos sistemas. Por consiguiente, se podría concluir que el sistema sur es más eficiente en el uso de los insumos, mientras que el sistema norte tiene una estructura administrativa más eficiente. Finalmente, en términos generales y teniendo en cuenta los pesos relativos de cada uno de los factores, encontramos que el sistema sur es más eficiente que el sistema norte.

Referencias

- Akaike, H. (1969). Fitting autoregressive models for prediction. *Annals of the institute of Statistical Mathematics*, 21(1), 243-247.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.
- Aristizábal, H., & Galvis, A. (2012). *Análisis de costos en sistemas de potabilización del agua*. Cali: Universidad del Valle.
- Bentarzi, M., & Hallin, M. (1994). On the invertibility of periodic moving-average models. *Journal of time series analysis*, 15(3), 263-268.
- Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2011). *Time series analysis: forecasting and control*, 734. John Wiley & Sons.
- Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, (pp. 1287-1294).
- Brockwell, P. J. & Davis, R. A. (2002). *Introduction to Time series and Forecasting*. New York. Wpringer.
- Cleveland, R. B., Cleveland, W. S., McRae, J.E. & Terpenning, I. (1990) STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics*, 6, 3–73.
- Durbin, J. (1959). Efficient estimation of parameters in moving-average models. *Biometrika*, 46(3/4), 306-316.
- Durbin, J., & Watson, G. S. (1951). Testing for serial correlation in least squares regression. *II*. *Biometrika*, 38(1/2), 159-177.

- Granger, C. W. J., and Newbold, P. (1974). Spurious Regressions in Econometrics. *Journal of Econometrics* 2 (2), 111-120.
- Granger, C. W. J., & Newbold, P. (2014). *Forecasting economic time series*. Academic Press.
- Greer, M. (2010). *Electricity cost modeling calculations*. Academic Press.
- Greene, W. (2002). *Econometric Analysis*. New Jersey. Prentice-Hall.
- Hamilton, J., (1994). *Time Series Analysis*. New Jersey. Princeton University Press.
- Hayashi, F., (2000). *Econometrics*. s.l. Princeton University Press.
- Hodrick, R. J. & Prescott E. C. (1997). Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking* 29, 1-16.
- Iqbal M, Naveed A (2016). Forecasting inflation: Autoregressive Integrated Moving Average Model. *European Scientific Journal*, 12(1), 83-92.
- Kotu, V., & Deshpande, B. (2018). *Data science: concepts and practice*. Morgan Kaufmann.
- Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297-303.
- López, P. (1997). *El agua: Tecnología de su distribución y uso*. Promotora General de Estudios, S.A.
- López, N., & Polo, G. (2009). *Análisis del sistema actual de abastecimiento de Bogotá*. Universidad de La Salle.
- Lütkepohl, H. (2013). *Introduction to multiple time series analysis*. Springer Science & Business Media.
- Matiz, F. (2008). *Optimización del proceso de floculación de la planta de tratamiento de agua potable El Dorado de Bogotá*. Universidad de los Andes.

- McFADDEN, D. (1978). Cost, revenue, and profit functions. Production economics: A dual approach to theory and applications. *The theory of production, 1*
- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to time series analysis and forecasting*. John Wiley & Sons.
- Namin SS, Namin A (2018). Forecasting economic and financial time series: ARIMA vrs. LSTM. *Machine Learning*, pp. 1-19.
- Nicholson, W. (2005). *Teoría microeconómica: principios básicos y ampliaciones*. Paraninfo.
- Pankratz, A. (1991). *Forecasting with Dynamic Regression Models*. John Wiley & Sons.
- Pizzi, N. (2007). *Pre-Treatment Field Guide: American Water Works Association*. AWWA.
- Rao, C. R. (1968). A note on a previous lemma in the theory of least squares and some further results. *Sankhyā: The Indian Journal of Statistics, Series A*, 259-266.
- Revollo, D. & Londoño, G. (2010). Análisis de las economías de escala y alcance en los servicios de acueducto y alcantarillado en Colombia. *Desarrollo y Sociedad*, 66.
- Romero, M. (2008). *Tratamientos utilizados en potabilización de agua*. Facultad de Ingeniería - Universidad Rafael Landívar.
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591-611.
- Usa Global Market S.A. (2016). Planta de Tratamiento Francisco Wiesner.
<http://www.usagmsa.com/index.php/2012-09-24-02-35-25>
- Van den Bossche, F., Wets, G., & Brijs, T. (2004). *A regression model with ARIMA errors to investigate the frequency and severity of road traffic accidents* (No. RA-2004-35). LUC.
- Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis*, 1951. Almquist and Wiksell, Upsala.

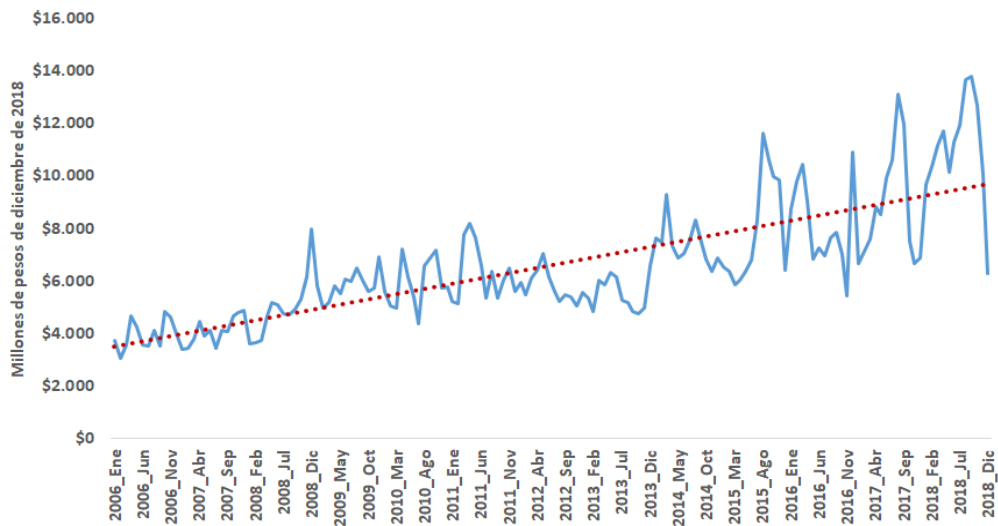
Wooldridge, J. H. (2002). *Econometric Analysis of Cross Section and Panel Data* (2nd Ed)

Books.

Anexos

Gráfica 1

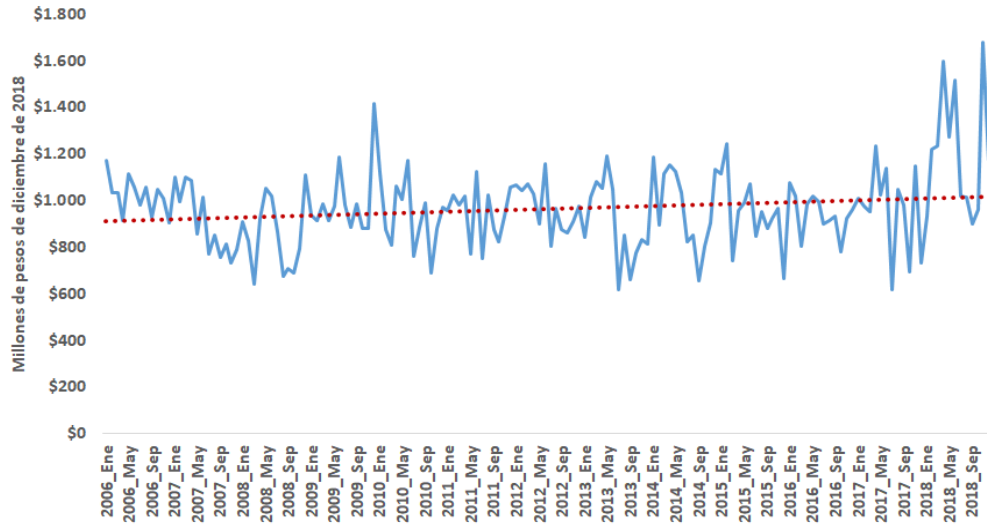
Costos totales de los sistemas norte y sur de la EAAB-ESP



Fuente: Cálculos del autor con base en los datos de la Empresa de Acueducto y Alcantarillado de Bogotá.

Gráfica 2

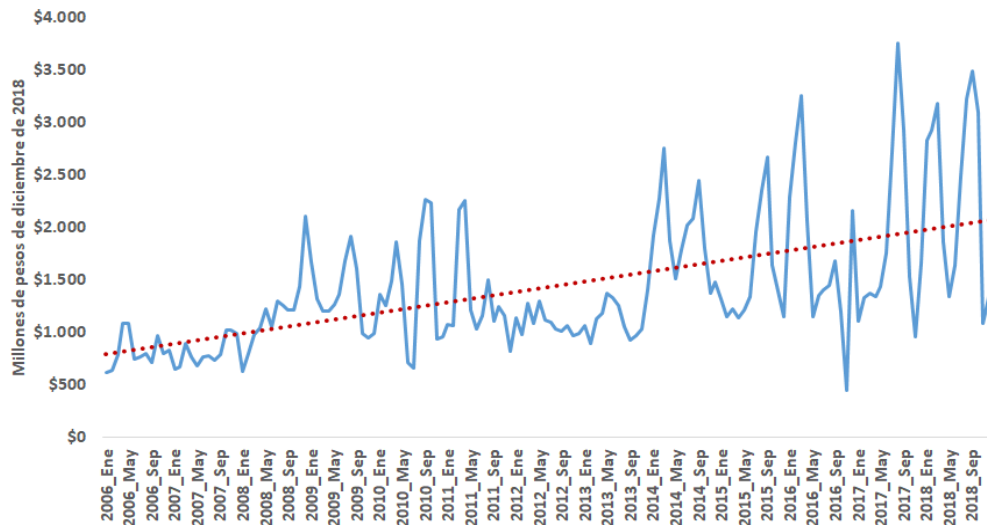
Costos administrativos de los sistemas norte y sur de la EAAB-ESP



Fuente: Cálculos del autor con base en los datos de la Empresa de Acueducto y Alcantarillado de Bogotá.

Gráfica 3

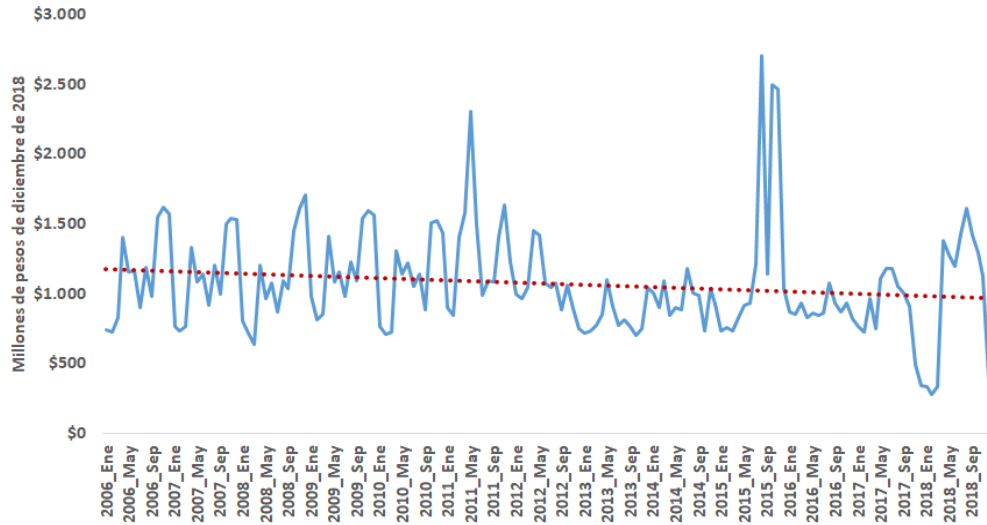
Costos de la energía de los sistemas norte y sur de la EAAB-ESP



Fuente: Cálculos del autor con base en los datos de la Empresa de Acueducto y Alcantarillado de Bogotá.

Gráfica 4

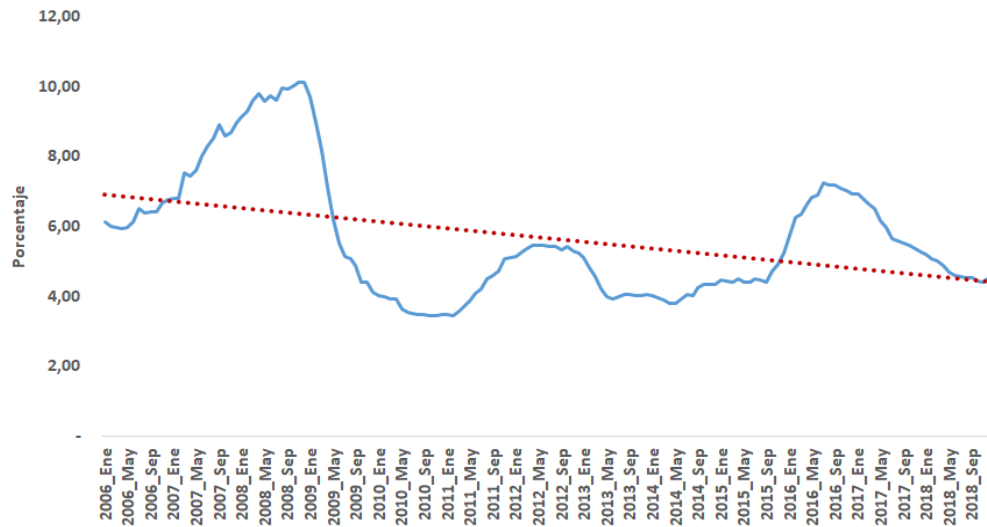
Costos de los químicos de los sistemas norte y sur de la EAAB-ESP.



Fuente: Cálculos del autor con base en los datos de la Empresa de Acueducto y Alcantarillado de Bogotá.

Gráfica 5

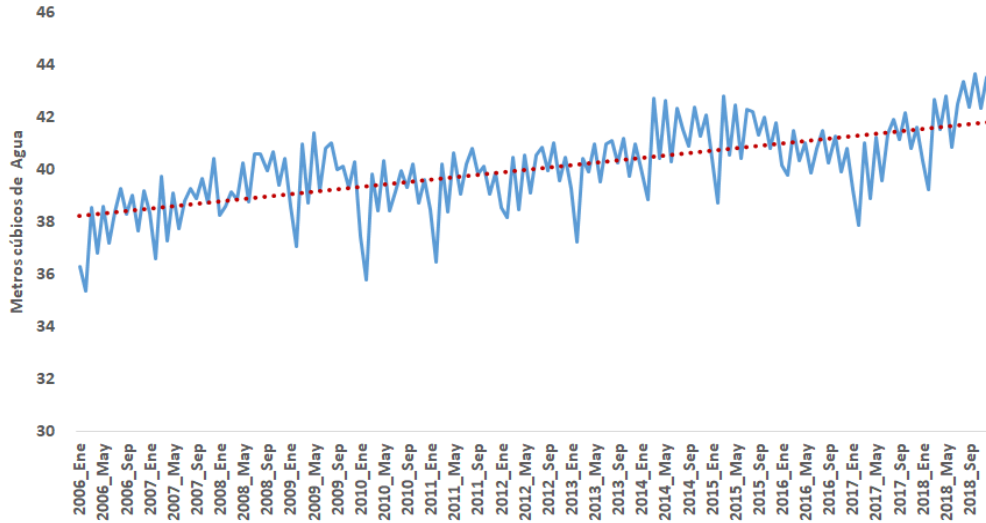
Tasa de los Depósitos a Término Fijo



Fuente: Cálculos del autor con base en los datos del Banco de la República de Colombia.

Gráfica 6

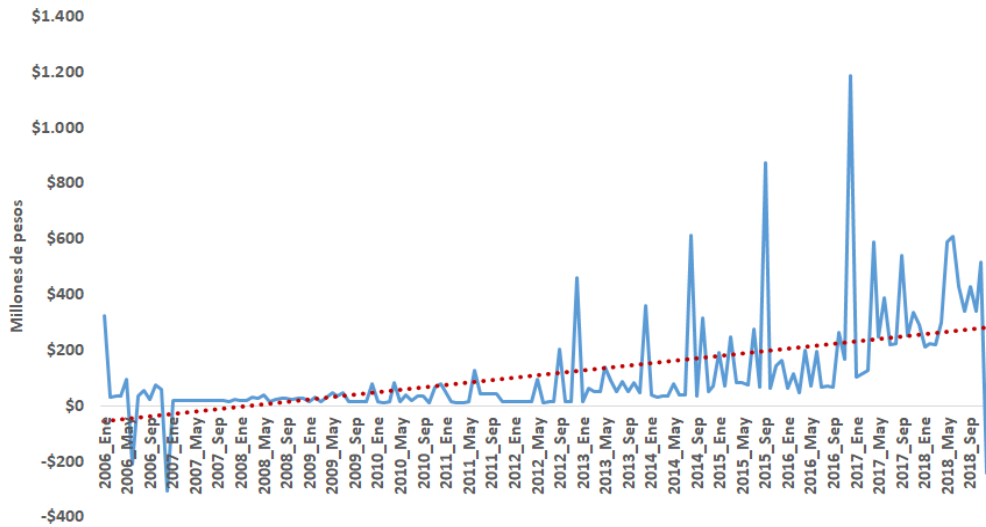
Cantidad Suministrada de Agua por la EAAB-ESP en millones de metros cúbicos



Fuente: Cálculos del autor con base en los datos de la Empresa de Acueducto y Alcantarillado de Bogotá.

Gráfica 7

Costos de la Tasa de Uso de los sistemas norte y sur de la EAAB-ESP



Fuente: Cálculos del autor con base en los datos de la Empresa de Acueducto y Alcantarillado de Bogotá.

Tabla 1

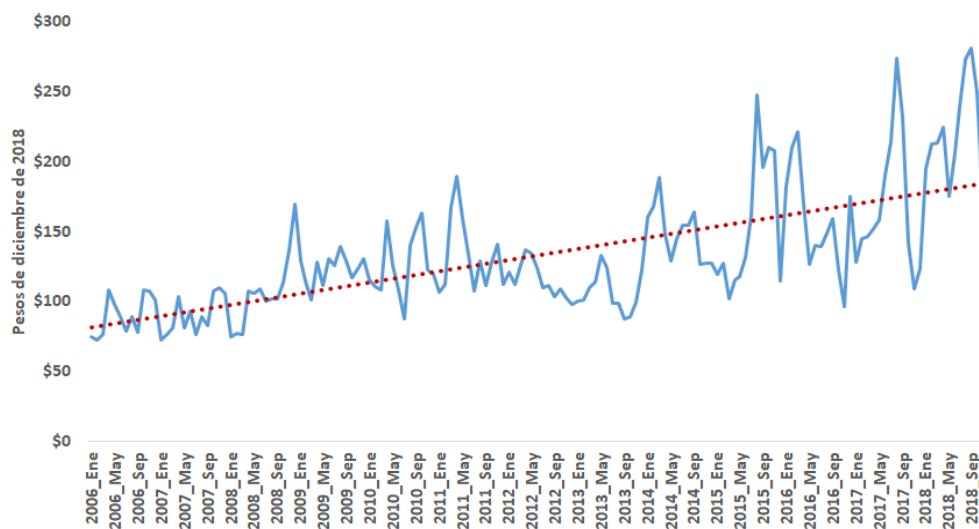
Estadística descriptiva de las variables

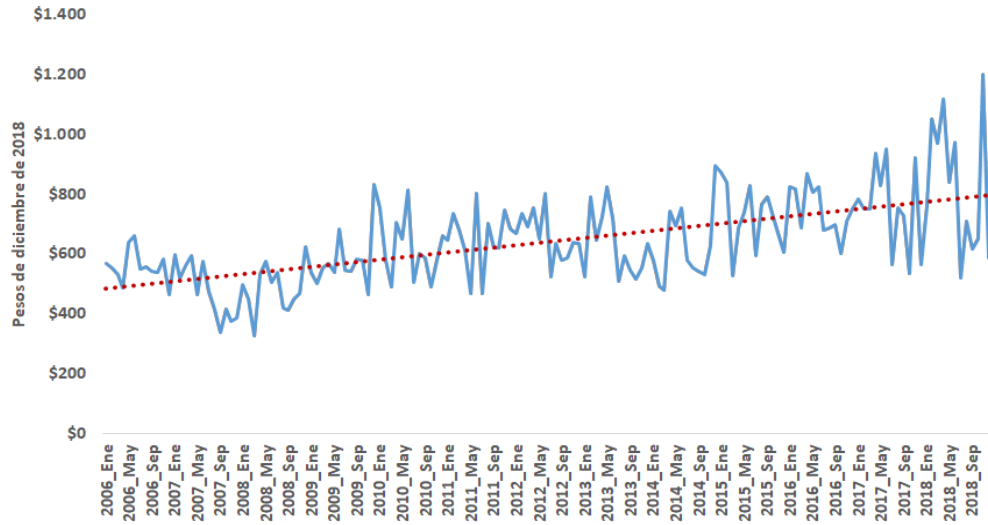
	CT	CA	CE	CQ	DTF	AS	TU
Media	\$6.597	\$967	\$1.431	\$1.076	5,67	40,02	\$113
Mediana	\$6.133	\$971	\$1.236	\$1.005	5,22	40,20	\$46
Desviación estándar	\$2.301	\$172	\$654	\$373	1,81	1,58	\$181
Rango	\$10.774	\$1.062	\$3.307	\$2.424	6,69	8,28	\$1.492
Mínimo	\$3.052	\$618	\$448	\$278	3,44	35,39	-\$306
Máximo	\$13.827	\$1.680	\$3.755	\$2.703	10,13	43,67	\$1.186

Fuente: Cálculos del autor con base en los datos de la empresa de Acueducto y Alcantarillado de Bogotá.

Gráfica 8

Costos medios por sistema de la EAAB-ESP

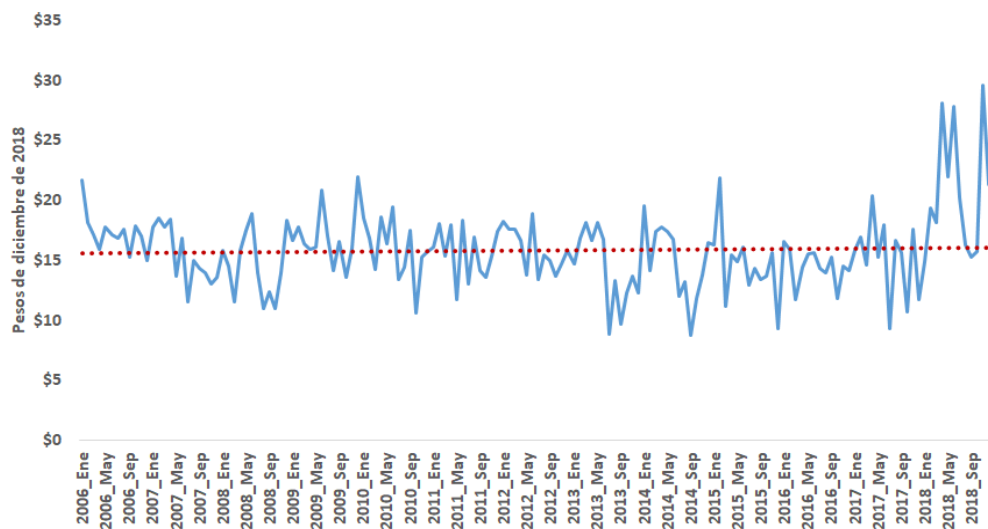


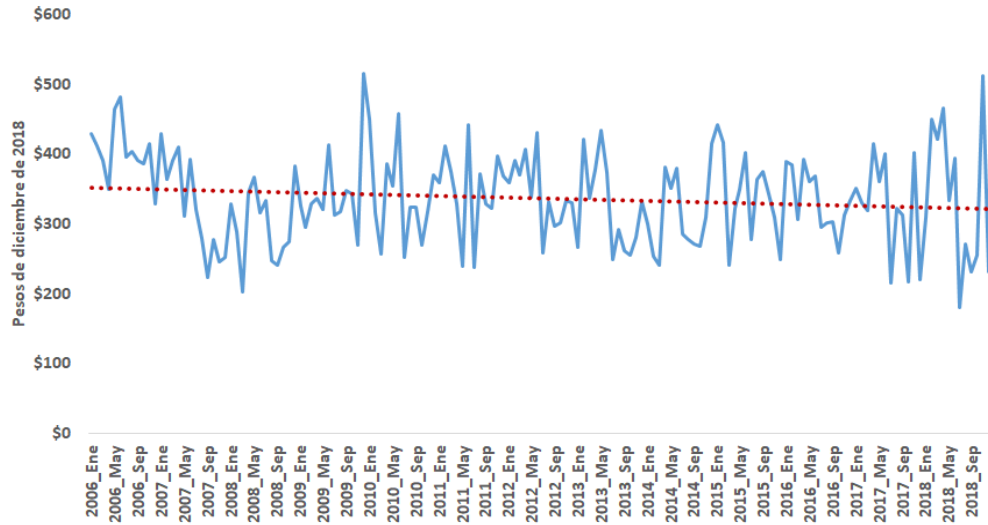


Fuente: Cálculos del autor con base en los datos de la Empresa de Acueducto y Alcantarillado de Bogotá. El primer gráfico hace referencia al sistema norte y el segundo al sistema sur.

Gráfica 9

Costos medios administrativos por sistema de la EAAB-ESP

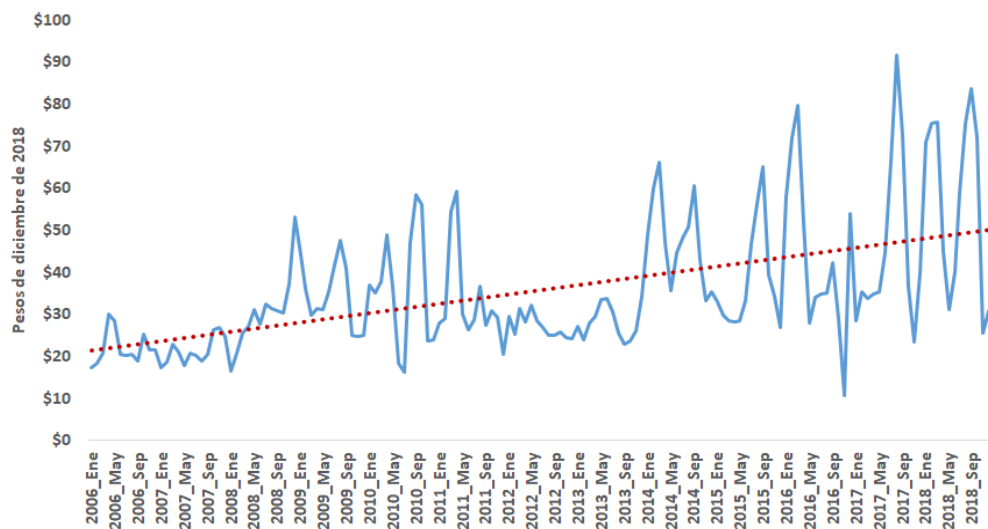


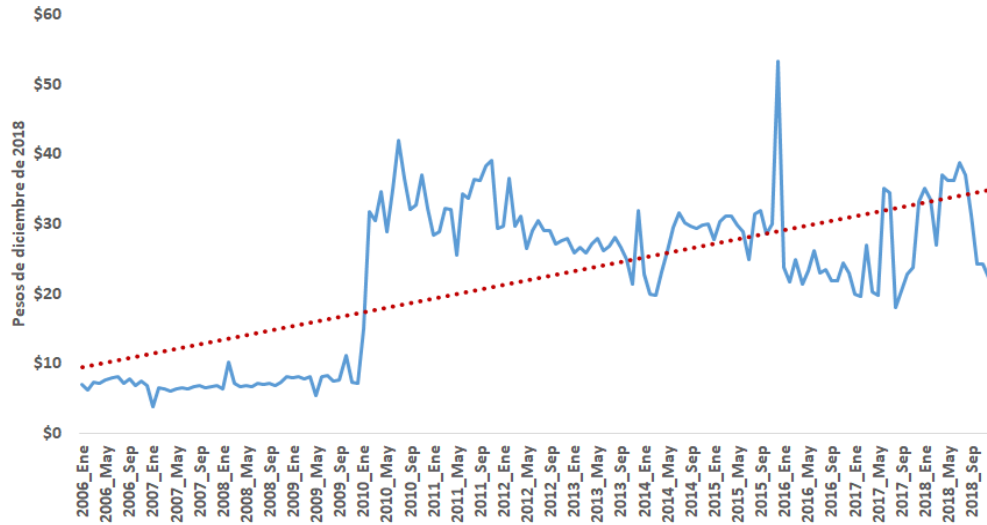


Fuente: Cálculos del autor con base en los datos de la empresa de Acueducto y Alcantarillado de Bogotá. El primer gráfico hace referencia al sistema norte y el segundo al sistema sur.

Gráfica 10

Costos medios de la energía por sistema de la EAAB-ESP

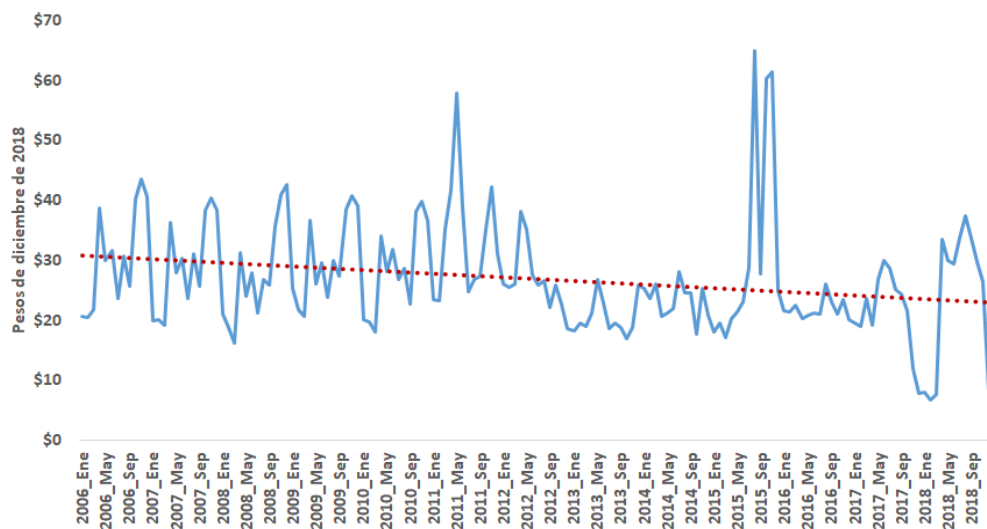


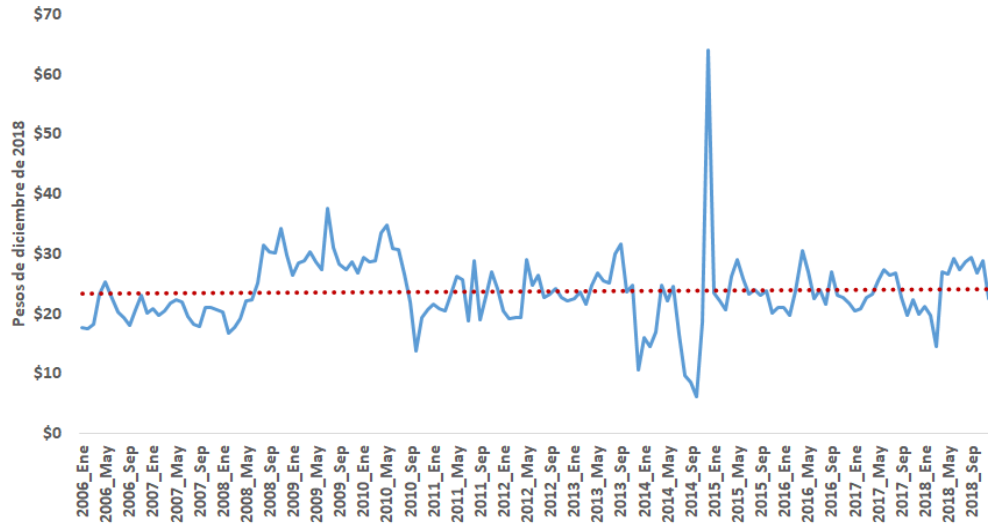


Fuente: Cálculos del autor con base en los datos de la empresa de Acueducto y Alcantarillado de Bogotá. El primer gráfico hace referencia al sistema norte y el segundo al sistema sur.

Gráfica 11

Costos medios de los químicos por sistema de la EAAB-ESP

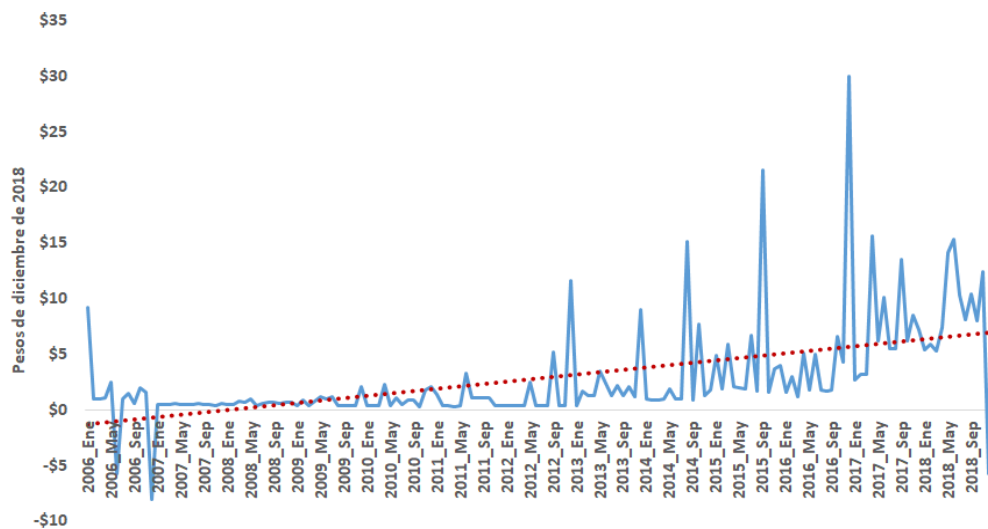


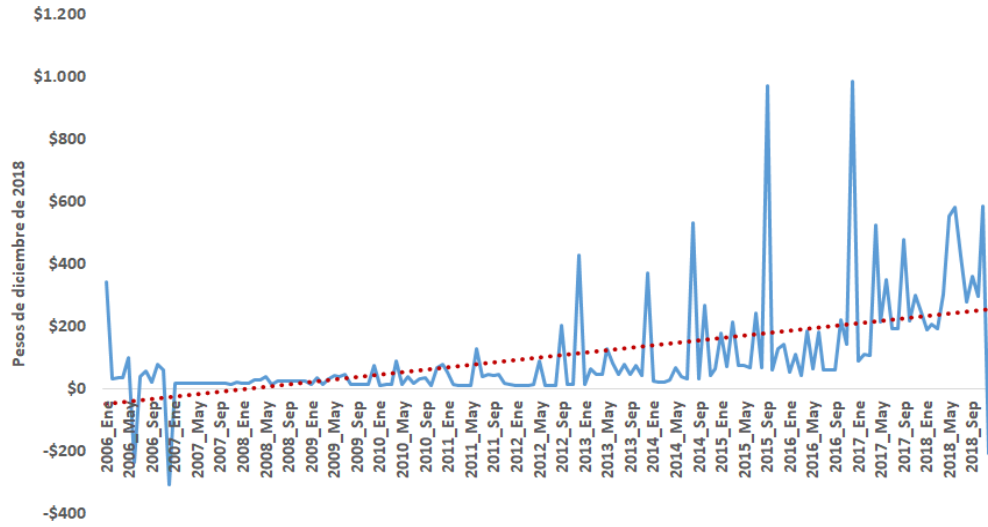


Fuente: Cálculos del autor con base en los datos de la empresa de Acueducto y Alcantarillado de Bogotá. El primer gráfico hace referencia al sistema norte y el segundo al sistema sur.

Gráfica 12

Costos de la Tasa de Uso del Agua por Sistema de la EAAB-ESP





Fuente: Cálculos del autor con base en los datos de la empresa de Acueducto y Alcantarillado de Bogotá. El primer gráfico hace referencia al sistema norte y el segundo al sistema sur.

Tabla 2

Resultados del modelo Base

	Modelo Total	Modelo Norte	Modelo Sur
(Intercept)	1.959*** (0.236)	1.503*** (0.178)	0.457* (0.227)
CA	0.218*** (0.052)	0.195*** (0.042)	0.796*** (0.033)
CE	0.555*** (0.026)	0.613*** (0.023)	0.210*** (0.015)
CQ	0.171*** (0.028)	0.229*** (0.024)	0.115*** (0.025)
DTF	-0.073* (0.033)	0.064*** (0.008)	0.044*** (0.006)
TU	0.094*** (0.009)	-0.059* (0.029)	0.128*** (0.032)
R-squared	0.858	0.889	0.883
N	156	156	156

Nota: desviaciones estándar en paréntesis y * indica que el coeficiente es significativo con un nivel de confianza del 95%.

Gráfica 13

Residuos y residuos vs predictor lineal para el modelo base

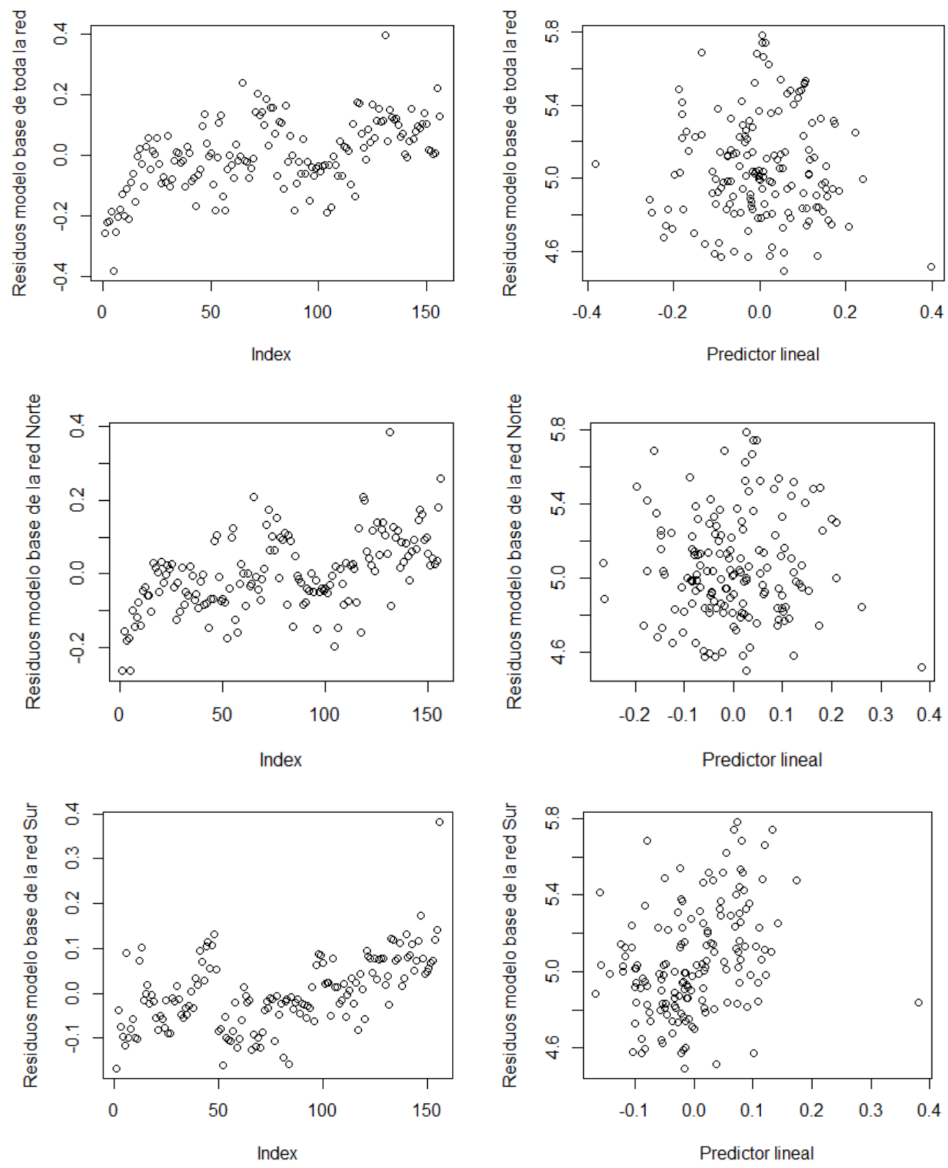


Tabla 3

Resultados del modelo con errores ARIMA

	Modelo total	Modelo Norte	Modelo sur
ar1	0.30 ** (0.10)	0.69 *** (0.11)	-0.41 *** (0.08)
ma1	-0.87 *** (0.05)	-0.73 *** (0.08)	
sar1	0.01 (0.10)		0.48 *** (0.08)
drift	0.00 *** (0.00)		
CA	0.25 *** (0.02)	0.20 *** (0.01)	0.86 *** (0.00)
CE	0.40 *** (0.01)	0.45 *** (0.01)	0.06 *** (0.00)
CQ	0.25 *** (0.01)	0.33 *** (0.01)	0.05 *** (0.00)
DTF	0.04 (0.03)	-0.01 (0.00)	0.00 (0.00)
TU	0.03 *** (0.00)	-0.02 (0.05)	0.02 (0.02)
ar2		-0.30 *** (0.09)	-0.17 (0.09)
AIC	-489.26	-551.18	-932.81
AICc	-487.74	-549.94	-931.57
BIC	-458.83	-523.79	-905.42
Log Likelihood	254.63	284.59	475.41
Num. obs.	155	155	155

Nota: desviaciones estándar en paréntesis y * indica que el coeficiente es significativo con un nivel de confianza del 95%.