

# *Modelos ARIMA y estructural de la serie de precios promedio de los contratos en el Mercado Mayorista de Energía Eléctrica en Colombia*

*Recibido para evaluación: 28 de Abril de 2005  
Aceptación: 26 de Sept de 2005  
Entrega de versión final: 01 de Dic de 2005*

Olver Hernández N.<sup>1</sup>  
Juan David Velásquez<sup>2</sup>  
Isaac Dyner<sup>3</sup>

## **RESUMEN**

Este artículo considera técnicas de modelamiento para la serie de precios de los contratos del mercado eléctrico colombiano. Las técnicas consideradas incluyen los modelos ARIMA y los modelos estructurados de series de tiempo. Los resultados del modelo estructural son mejores en el proceso de calibración mientras los resultados del modelo ARIMA son mejores en el periodo de validación.

**PALABRAS CLAVES:** Precio de Contratos, tendencia cúbica, estacionalidad, ARIMA

## **ABSTRACT**

This paper considers techniques for modeling the time series contract prices in the Colombian Electricity Market. Considered techniques include the ARIMA and Structural time series models. Results of the structural model are better in the calibration process while results of the ARIMA model are better in the validation process.

**KEYWORDS:** Contract Prices, Cubical Trend, Seasonality, ARIMA models.

<sup>1, 2, 3</sup> Universidad  
Nacional de Colombia,  
Sede Medellín.

<sup>1</sup>oohernan@unbalmed.edu.co  
<sup>2</sup>javelasq@unalmed.edu.co  
<sup>3</sup>idynner@unbalmed.edu.co

## 1. INTRODUCCIÓN

La estructura de los mercados de electricidad que han sufrido procesos de desregulación, liberación y privatización difiere, en gran medida respecto de otros tipos de mercados, debido a las características propias de la electricidad. De ahí que sus series de precios presentan características propias usualmente no encontradas en otras series económicas tales como la alta volatilidad de los precios de bolsa que se presentan en los mercados liberalizados o desregulados.

La consecuencia en Colombia de esta tendencia mundial, se materializó mediante la creación del Mercado Mayorista de Energía Eléctrica, donde la electricidad se negocia en un mercado de subastas de corto plazo, o a través de contratos bilaterales entre agentes. Estos dos tipos de mercados poseen sus propias series de precios, las cuales muestran las características propias de cada forma de negociación. Mientras la teoría económica de los mercados eléctricos esboza los principios fundamentales que gobiernan la formación del precio en los mercados de electricidad, las particularidades propias de cada mercado hacen que no se pueda especificar directamente un modelo econométrico que permita explicar la dinámica de la formación de los precios de electricidad; es así entonces como el modelamiento empírico de series temporales cumple un rol fundamental en el entendimiento de las particularidades propias de cada serie de precios.

Particularmente en el caso de Colombia, ya se tienen algunas experiencias en el modelamiento de los precios de bolsa; no obstante, en nuestro conocimiento, no existen trabajos publicados que expliquen desde un punto de vista econométrico las características de las series de precios de contratos.

Motivados por la importancia que tiene para los distintos agentes comprender la dinámica que gobierna los precios de electricidad negociada en contratos de largo plazo, este trabajo presenta dos modelos univariados que representan la dinámica de dicha serie, buscando establecer algunas de sus propiedades, así como de comparar su desempeño en la predicción de su comportamiento futuro. En primer lugar se desarrollará un modelo ARIMA siguiendo la metodología de Box-Jenkins (Box and Jenkins, 1970), para posteriormente, desarrollar un modelo de regresión donde los errores son modelados usando la misma metodología de Box-Jenkins. Para finalizar se realiza un análisis comparativo de la calidad de ambos modelos en la representación de las propiedades de esta serie.

En el siguiente capítulo se muestra la serie de precios y se plantea la metodología a utilizar, primero se muestra el modelo Box-Jenkins y luego el modelo regresivo. En el capítulo siguiente se muestran los resultados de la estimación de los modelos, luego se analizan los resultados para llegar por último a las conclusiones.

## 2. INFORMACIÓN UTILIZADA Y METODOLOGÍA

### 2.1 Formación del Precio de Contratos

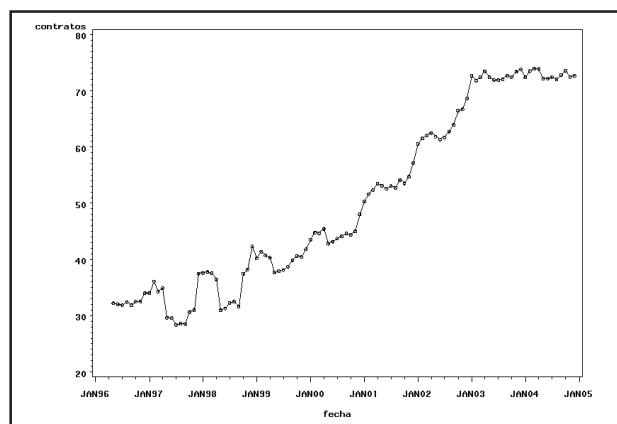
La Energía puede ser negociada (ISA, 2002) a través de contratos bilaterales de largo plazo entre agentes que pactan libremente su volumen y su precio, o mediante un proceso de subasta en un mercado de corto plazo, donde los excedentes de energía son vendidos a bolsa y esta a su vez se los vende a aquellos agentes cuyo volumen de contratación no cubre su demanda real. La decisión de contratación de compras y ventas es dificultada, por la gran cantidad de factores tanto en la oferta y la demanda que pueden modificar el comportamiento de los precios, aunque pareciera que desde el lado de la oferta, la cantidad del recurso hídrico y las expectativas de su disponibilidad futura fueran los factores más

importantes en la explicación de los movimientos de los precios (Mesa et al., 1994; Poveda et al., 2002); los cuales están relacionados con los fenómenos hidrológicos extremos de El Niño y La Niña. Como una consecuencia de lo anterior, el establecimiento de las metas de contratación de la energía generada realizado por los agentes generadores requiere de un proceso de análisis complejo en donde deben considerarse factores como la cantidad de energía contratada en el largo plazo, la evolución de su demanda, el precio de los contratos y el precio de Bolsa; estos últimos son influenciados por factores físicos tales como la hidrología, embalse ofertable, demanda total del sistema, etc. (Montoya, 2003). En el caso de los comercializadores, adicionalmente a los factores anteriores deben tenerse en cuenta otros aspectos como los riesgos asociados a las compras de energía destinadas a suplir la energía que ya fue vendida. De esta forma el precio promedio de los contratos despachados resulta de la confluencia de los factores previamente numerados, así como también de los diferentes tipos de contratos que se han realizado entre los agentes. De ellos los más representativos son:

- Pague lo contratado (PC): El comercializador que compra se compromete a pagar toda la energía contratada, independiente que esta sea consumida o no. Si se contrató una cantidad mayor a la requerida, el comprador debe venderla en la Bolsa de Energía. Este es el único caso en que los comercializadores venden energía en la Bolsa.
- Pague lo demandado con tope (PD con tope): el agente comprador sólo paga su consumo, siempre y cuando la cantidad consumida sea inferior o igual a la energía contratada. Si el consumo es superior al tope, la diferencia se paga al precio de bolsa, a menos que el comprador tenga otros contratos y supla su déficit con ellos.

### 2.2 Información Utilizada

Para realizar el presente estudio los datos del precio promedio mensual de los contratos en el Mercado Eléctrico Colombiano (MEM) fueron obtenidos del Sistema de Información NEON. Su serie temporal es presentada en la Figura 1.



**Figura 1.** Precio promedio mensual de la Electricidad de contratos despachados y de la bolsa de energía en \$/kWh

La inspección visual de la serie muestra la presencia aparente de tres tramos con características estructurales diferentes; en el primero de ellos se aprecia una tendencia leve y una estacionalidad muy marcada, que cubre las observaciones entre 1996:5 y 1999:12; el segundo tramo tienen una tendencia con una pendiente positiva superior a la del tramo anterior, la cual se mantiene hasta principios

de 2003, finalmente el último tramo se caracteriza por mantener un nivel aproximadamente horizontal alrededor del cual oscilan los precios de los contratos.

Una característica de toda la serie es la presencia de una componente estacional de un año de periodicidad la cual puede ser explicada por los ciclos de invierno-verano, en donde los meses de la estación de lluvia presentan precios más bajos que el resto del año, debido a la abundancia del recurso hídrico y la alta componente de plantas hidráulicas del parque de generación.

### 2.3 Modelos Usados

#### 2.3.1 Modelos de Box-Jenkins

En la familia de modelos ARIMA (Box and Jenkins, 1970), se pretende estimar un modelo de la forma

$$\phi_p(L)(1-L^s)(1-L)^d y_t = \theta_q(L)e_t \quad (1)$$

Donde  $L$  simboliza el operador de rezago, los polinomios  $\phi_p(L)$  representan la componente autoregresiva y los polinomios  $\theta_q(L)$  representan la componentes de medias móviles. Los factores

$(1-L)^d(1-L^s)$  corresponden a los operadores de diferenciación simple de orden  $d$ , y de diferenciación estacional de periodo  $s$ ; el proceso de modelado se basa en determinar el orden de integración de la serie y la presencia de componentes estacionales para especificar los valores de  $d$  y  $s$ ; posteriormente se realiza un proceso de identificación donde se postulan diferentes polinomios para el modelamiento de las componentes autoregresivas y de medias móviles, los cuales son identificados a través de los correlogramas simples y parciales de la serie diferenciada. El modelo final es seleccionado con base en un criterio de información que mide conjuntamente el error de ajuste a los datos y el error de varianza inducido por los grados de libertad del modelo. Una selección común es el criterio de Información de Akaike (AIC) o el criterio de Hanan-Quinn (Hanhan and Quinn, 1979).

Uno de los problemas fundamentales de la metodología es la determinación del orden de integración de los datos, ya que se hace difícil distinguir entre un proceso autorregresivo de orden uno con un coeficiente cercano a la unidad y un proceso que sigue un paseo aleatorio puro. Para resolver este problema Dicky-Fuller proponen una prueba estadística para determinar cuando el coeficiente del modelo autorregresivo de orden uno es significativamente diferente de la unidad. La prueba se basa en realizar una regresión.

$$y_t = \mu + \beta * t + \rho * y_{t-1} + e_t \quad (2)$$

Donde  $\mu$  es el término de media constante,  $\beta$  es la pendiente,  $\rho$  es parámetro autorregresivo y  $e_t$  es ruido blanco.  $y_t$  sigue un proceso integrado si  $\rho = 1$  (camino aleatorio con constante). La prueba asume  $H_0: \rho = 1$  (hipótesis nula) contra la alternativa  $H_1: \rho < 1$  (hipótesis alterna). (Dickey and Fuller, 1979). Un tratamiento detallado sobre este tipo de pruebas es presentado por Hamilton (1994) y Mills (1993).

#### 2.3.2. Modelo Estructural

Esta clase de modelos representa explícitamente las componentes de tendencia ( $\mu_t$ ) y de estacionalidad ( $\gamma_t$ ), tal que la variable dependiente es obtenida como la suma de las dos anteriores, es decir: (Diebold, 2001):

$$y_t = \mu_t + \gamma_t + \epsilon_t \quad (3)$$

Donde el término  $\epsilon_t$  representa los residuos del modelo, los cuales pueden contener relaciones no capturadas en la ecuación anterior, por lo que usualmente se incluye una segunda ecuación dentro del modelo que permite capturar dichas relaciones; la dinámica que sigue la serie de residuos es comúnmente modelada a través de la aproximación de Box-Jenkins:

$$\epsilon_t = \frac{\theta_q(L)}{\phi_p(L)} \eta_t \quad (4)$$

donde  $\eta_t$  es un ruido blanco.

La componente estacional  $\gamma_t$ , en el caso de series mensuales es modelada a través de doce variables indicadoras, una para cada mes, tal que:

$$\gamma_t = \sum_{j=1}^{12} \delta_j I_{j,t} \quad (5)$$

donde,  $I_{j,t}$  representan las variables indicadoras de cada uno de los meses del año:

$$I_{j,t} = \begin{cases} 1 & t = j, j+12, j+24, \dots \\ 0 & \text{En otro caso} \end{cases} \quad (6)$$

y son los parámetros de la estacionalidad que representan el efecto del  $j$ -ésimo mes del año. La tendencia es modelada a través de un polinomio cúbico tal que  $\gamma_t = at + bt^2 + ct^3$ . Para evitar la multicolinealidad no se incluye un intercepto en el modelo, ya que se asume que los coeficientes lo incluyen; de esta forma la sumatoria de los factores estacionales sobre los doce meses consecutivos es diferente de cero, a diferencia de otras aproximaciones que incluyen implícitamente esta condición. Así el modelo final puede ser escrito como:

$$y_t = at + bt^2 + ct^3 + \sum_{j=1}^{12} \delta_j I_{j,t} + \epsilon_t \quad (7)$$

Donde  $a, b, c$  y  $\delta_j$  son los parámetros que hay que estimar.

### 3. Resultados empíricos

#### 3.1. Modelos de Box-Jenkins

##### 3.1.1. Evidencias de Integración de la Serie

Un problema fundamental dentro de la modelación de Box-Jenkins es reducir la serie a un estado homogéneo en el sentido estadístico. En este sentido es necesario diferenciar procesos integrados de procesos autorregresivos de orden uno, con coeficiente positivo cercano a la unidad. Dickey y Fuller analizaron este problema encontrando que el modelo de regresión tradicional con el que se estiman los modelos autorregresivos se encuentra fuertemente sesgado cuando la serie es integrada, puesto que la distribución que siguen los coeficientes no pertenecen a ningún modelo paramétrico tradicional, por lo que los valores críticos para determinar si el coeficiente de  $y_{t-1}$  es significativamente distinto de uno, deben ser estimados numéricamente. Se distinguen tres pruebas: la primera asume que la media es cero y que no hay tendencia (Zero Mean), la segunda asume que la media del proceso es distinta de cero, pero sin tendencia (Single Mean) y la última que hay tendencia con intercepto diferente de cero (TREND). Los resultados de la aplicación de estas tres variantes se pueden apreciar en la Tabla 1

Como los valores-p de la prueba son mayores que 0.05, con media cero en ambos casos, se concluye no hay evidencias suficientes para rechazar la hipótesis nula de que la serie tiene una raíz unitaria para el rezago uno y doce.

### 3.1.2. Modelo obtenido

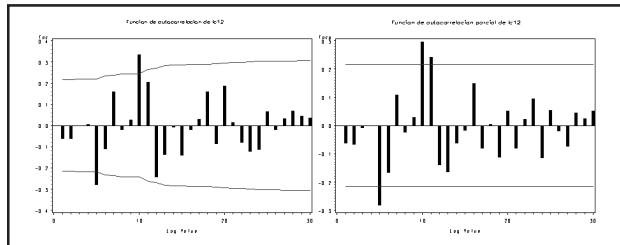
A partir del análisis de los Autocorrelogramas simple y parcial mostrados en la Figura 2, se postularon diferentes modelos alternativos para describir la dinámica de la serie de precios de contratos; usando el criterio de Información de Akaike se obtuvo el siguiente modelo como el más adecuado:

$$(1 - 0.48 * B^{10})(1 - B)(1 - B^{12})y_t = (1 + 0.37 * B^{11} - 0.24 * B^{12})a_t$$

**Tabla 1.** Prueba de Dickey-Fuller para probar Raíces Unitarias

	Rezago 1		Rezago 12*	
	Estadístico	Valor-p	Estadístico	Valor-p
Zero Mean	0.2055	0.729	-5.02	0.1204
Single Mean	-0.7647	0.9049	-3162.9	0.0001
Trend	-19.8348	0.0572	-113.628	0.0001

\*La prueba se aplica luego de aplicar diferenciación de orden uno



**Figura 2.** Autocorrelogramas simple y parcial de la serie  $(1 - B)(1 - B^{12})y_t$

Sus parámetros fueron estimados a partir de las observaciones entre Mayo de 1996 y Junio de 2004, reservándose un periodo de seis meses para analizar la capacidad de predicción del modelo.

Para determinar la normalidad de los residuos, se utilizó la prueba de Jarque-Bera, obteniéndose un valor del estadístico de 10.5 y un valor crítico de 0.005, lo que indica que los residuos no siguen una distribución normal. Los resultados de la prueba de Ljung-Box indican que no existen correlaciones seriales consecutivas en los primeros 18 rezagos del modelo. Los resultados de la prueba de McLeod-Lee indican que no existen no linealidades remanentes en

**Tabla 2.** Prueba de Ljung Box para los residuos del modelo diferenciado

To Lag	Chi Sq	DF	Valor-p	Autocorrelaciones					
6	6.59	6	0.36	-0.16	0.04	-0.02	0.10	-0.15	-0.09
12	8.26	12	0.76	0.093	0.01	0.06	-0.05	-0.01	-0.03
18	14.98	18	0.66	-0.13	0.04	-0.11	0.09	-0.03	0.139

**Tabla 3.** Prueba de Ljung Box para los residuos al cuadrado del modelo diferenciado

To Lag	Chi Sq	DF	Valor-p	Autocorrelaciones					
6	1.81	6	0.93	-0.06	0.08	0.04	0.03	0.02	-0.07
12	3.59	12	0.99	0.100	0.004	0.06	0.03	0.008	0.038
18	5.78	18	0.99	0.104	-0.03	0.02	-0.02	0.089	-0.01

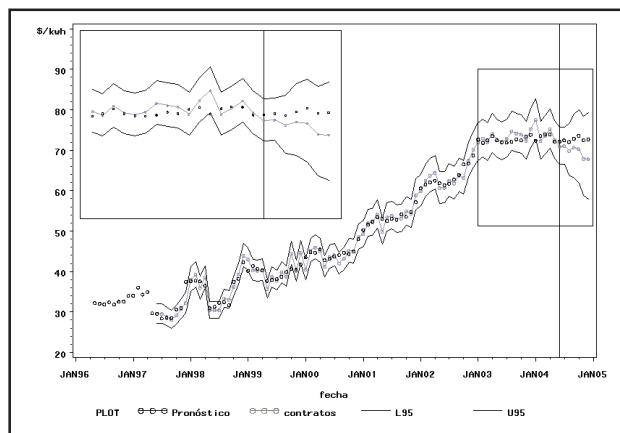
los residuos del modelo. Los resultados numéricos de estas pruebas son mostrados en las tablas 2 y 3.

Los pronósticos del modelo un paso adelante durante el proceso de estimación de parámetros son mostrados en la Figura 3.

### 3.2. Modelo Estructural

#### 3.2.1. Estacionalidad y Tendencia

Al estimar el modelo presentado en la Ecuación 7, se obtuvieron los valores mostrados en la tabla 4, junto con sus valores críticos.



**Figura 3.** Resultados del modelamiento de la serie temporal usando diferenciación y un modelo ARMA

**Tabla 4.** Parámetros del modelo estructural

Variable	Valor Estimado	Valor-p
a	-0.008244	<.0001
b	0.000399	<.0001
c	-2.37E-06	<.0001
$E_{ene}(\delta_1)$	3.5723	<.0001
$E_{feb}(\delta_2)$	3.5844	<.0001
$E_{mar}(\delta_3)$	3.571	<.0001
$E_{abr}(\delta_4)$	3.5671	<.0001
$E_{may}(\delta_5)$	3.4931	<.0001
$E_{jun}(\delta_6)$	3.4852	<.0001
$E_{jul}(\delta_7)$	3.4806	<.0001
$E_{ago}(\delta_8)$	3.4797	<.0001
$E_{sep}(\delta_9)$	3.4778	<.0001
$E_{oct}(\delta_{10})$	3.5069	<.0001
$E_{nov}(\delta_{11})$	3.5083	<.0001
$E_{dic}(\delta_{12})$	3.5641	<.0001

En la Figura 4 se presenta el ciclo estacional de la serie temporal estimados a través de los coeficientes  $\delta_j$  en la regresión. A partir de ella puede corroborarse que los precios entre Mayo y Octubre (Estación de invierno) tienden a ser más bajos que en los correspondiente para la estación de verano.

Al analizar los residuales del modelo se encontró que estos poseen una estructura determinística (véase la Figura 5) por lo que pueden ser representados mediante un modelo ARIMA. En consecuencia se procedió a su modelación usando la metodología de Box-Jenkins previamente descrita, obteniéndose el siguiente modelo que describe su dinámica:

$$(1 - 0.63 * B)(1 + 0.42 * B^5)e_t = (1 - 0.40 * B^{13} - 0.19 * B^{14} - 0.34 * B^{19})a_t$$

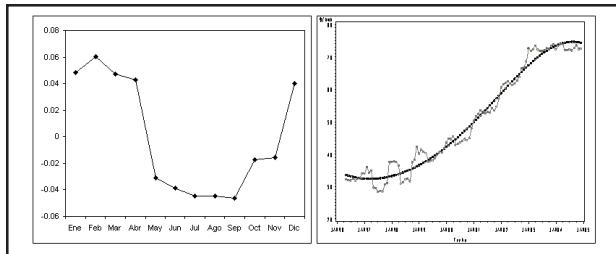


Figura 4. Estacionalidad y tendencia de la serie

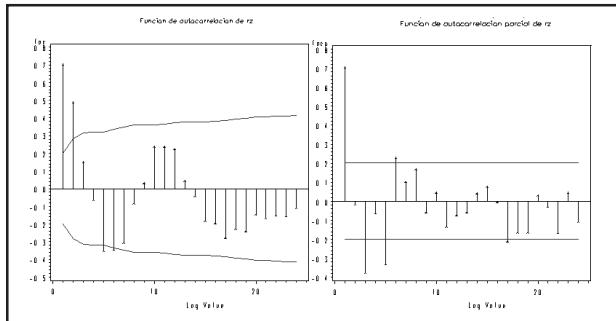


Figura 5. Fac y Facp de los residuales del modelamiento de la tendencia y la estacionalidad

Se verificaron los residuos de este último modelo encontrando que ellos siguen una distribución normal (de acuerdo con el estadístico de Jarque-Bera  $F=0.3428$ , valor crítico: 0.843) y que hay incorrelación entre los residuos (según la prueba de Ljung-Box mostrada en la tabla 5). La prueba de McLeod-Lee indica que para el séptimo rezago existe posiblemente no linealidades remanentes en los residuos que el modelo no es capaz de capturar (véase la tabla 6).

Los pronósticos del modelo un paso adelante durante el proceso de estimación de parámetros son mostrados en la Figura 6.

Tabla 5. Prueba de Ljung Box para los residuos del modelo estructural

To Lag	Chi Sq	DF	Valor-p	Autocorrelaciones					
6	3.63	6	<b>0.73</b>	0.06	0.131	0.073	0.051	-0.05	0.089
12	8.72	12	<b>0.77</b>	-0.06	0.113	-0.02	0.171	-0.01	0.001
18	20.67	18	<b>0.30</b>	0.012	-0.05	-0.20	-0.06	-0.22	-0.07

Tabla 6. Prueba de Ljung Box para los residuos al cuadrado del modelo estructural

To Lag	Chi Sq	DF	Valor-p	Autocorrelaciones					
				0.054	0.020	0.127	-0.06	0.098	0.058
6	3.72	6	<b>0.71</b>	0.054	0.020	0.127	-0.06	0.098	0.058
12	24.40	12	<b>0.02</b>	0.264	-0.08	0.022	0.333	-0.02	-0.01
18	34.80	18	<b>0.01</b>	-0.05	-0.03	0.007	0.015	0.277	-0.07

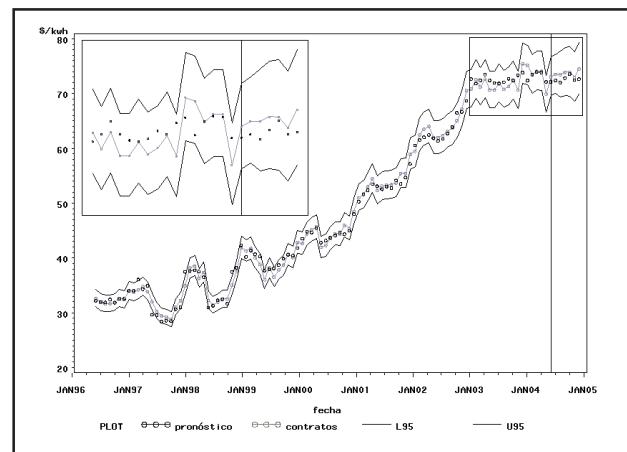


Figura 6. Resultados del modelamiento de la serie temporal usando el modelo estructural

### 2.3. Evaluación de la Capacidad de Predicción

Para los dos modelos analizados se evaluó su capacidad de predicción realizando el pronóstico ciego seis meses adelante a partir del último mes usado para la estimación de los parámetros. El término pronóstico ciego se refiere a que la predicción en cada periodo fue realizada usando los datos reales hasta junio de 2004 y a partir de ahí los pronósticos previos entregados por el modelo. Para efectos comparativos, los estadísticos tanto de los períodos de entrenamiento como de validación son mostrados en la tabla 7.

Tabla 7. Estadísticos de la fase de entrenamiento

Estadístico	ARIMA	Estructural
mape	1.8240	2.5042
msimp	2.5955	2.5961
u-theil	0.5415	0.7255
rmse	1.0816	1.5630
r <sup>2</sup>	0.9313	0.9008

Tabla 8. Estadísticos de la fase de validación

Estadístico	ARIMA	Estructural
mape	1.4292	4.0814
msimp	0.7994	0.7995
u-theil	1.7426	4.8124
rmse	1.1594	3.2113
r <sup>2</sup>	0.9263	0.7961

Analizando los resultados la tabla de estadísticos, podemos observar que los modelos en la etapa de entrenamiento mejoran el modelo simple de usar el valor anterior como predicción del mes en cuestión, debido a que el mape es menor que el msimp.

Debido al posible cambio estructural de la serie en la etapa de validación los dos modelos son peores que el modelo simple de usar el valor en el periodo anterior, esto se puede concluir al comparar el mape y el msimp.

#### 4. CONCLUSIONES

Se ha realizado el análisis de la serie de precios promedio de contratos despachados en el Mercado Eléctrico Colombiano usando la metodología de Box-Jenkins y la técnica de modelamiento estructural con errores siguiendo una estructura temporal.

De acuerdo con la información disponible, se ha demostrado que la serie de precios es y que presenta una componente estacional asociada a los ciclos de invierno-verano propios de la hidrología. Los resultados del modelo de Box-Jenkins indican que sus residuos no siguen una distribución normal aunque son incorrelacionados. Esto indica la necesidad de evaluar posteriormente modelos heterocedásticos.

Los residuales del modelo estructural aunque siguen una distribución normal y son incorrelacionados, presentan evidencias de una posible relación no-lineal existente que el modelo no es capaz de capturar. Para la tarea de predicción a la luz de la información disponible, el modelo estructural captura de una forma más completa las características de la serie de precios de los contratos.

#### BIBLIOGRAFÍA

BOX, G. E. P. and JENKINS, G. M. Time Series Analysis: Forecasting and Control. Holden—Day Inc. 1970.

DICKEY, D. A. and FULLER, W. A.. Distribution of the estimators for autoregressive time series with a unit root. En: Journal of the American Statistical Association, vol 74 (1979), p. 427-431.

DIEBOLD, F. Elementos de Pronóstico. International Thomson. 2001.

HAMILTON, J. D. Time Series Analysis. Princeton University Press. 1994.

HANNAN, E. AND QUINN, B. The determination of the order of an autoregression. Journal of Royal Statistical Society Series B, Vol 41(1979), p. 190—195.

ISA. Informe Mercado de Energía Mayorista. ISA. 2002.

MESA, O., CARVAJAL, F., SALAZAR, L., AND POVEDA, G. Efecto del fenómeno del niño- oscilación del sur en la hidrología de Colombia. In Memorias del XVI Congreso Latinoamericano de Hidráulica e Hidrología, Vol. 2, (1994), p. 373-384, Santiago de Chile. IAHR.

MILLS, T. C. The econometric modelling of financial time series. Cambridge Press University. 1993

Montoya, R. D. Herramienta para el apoyo de toma de decisiones

de empresas comercializadoras y distribuidoras de energía. Master's thesis, Universidad Nacional de Colombia. Sede Medellín. Posgrado en Aprovechamiento de Recursos Hídricos. 2003.

POVEDA, G., VELEZ, J., HOYOS, C., MEJIA, J., BARCO, J., AND CORREA, P. Influencia de fenómenos macroclimáticos sobre el ciclo anual de la hidrología colombiana: Cuantificación lineal, no lineal y percentiles probabilísticas. En: Memorias VI Congreso Colombiano de Meteorología, Vol. 2, (2002). Bogotá. Socolmet.