

---

# ANÁLISIS DE SERIES DE TIEMPO PARA LA PREDICCIÓN DE LOS PRECIOS DE LA ENERGÍA EN LA BOLSA DE COLOMBIA

---

Sergio Botero Botero\*  
Jovan Alfonso Cano Cano

## Resumen

Botero, Sergio y Cano, Jovan Alfonso. "Análisis de series de tiempo para la predicción de los precios de la energía en la bolsa de Colombia", Cuadernos de Economía, v. XXVII, n. 48, Bogotá, 2008, páginas 173-208

Debido a la reestructuración del sector eléctrico colombiano, durante las dos últimas décadas, el comportamiento del precio de la energía eléctrica ha incrementado su volatilidad, reflejando el riesgo existente para los diferentes agentes que intervienen en el mercado. El objetivo de este artículo es presentar una metodología para la implementación de modelos de regresión, sobre la serie histórica de precios de bolsa de energía en Colombia. A medida que la cantidad de datos aumente, podrán desarrollarse modelos más amplios, que describan de forma adecuada comportamientos del mercado, que empleando las técnicas y la información disponible actualmente, no es posible identificar.

**Palabras clave:** mercado de energía, spot, series de tiempo, intervención del mercado. **JEL:** C13, C15, C53, Q49.

---

\* Sergio Botero Botero es Doctor of Science, Engineering Management, actualmente se desempeña como Profesor Asociado de Facultad de Minas de la Universidad Nacional de Colombia (Bogotá, Colombia). E-mail: sbotero@unal.edu.co.

Jovan Alfonso Cano Cano es estudiante de Doctorado en Ingeniería y Magíster en Ciencias Económicas de los Recursos Energéticos y Naturales. E-mail: jcano@isa.com.co.

Este artículo fue **recibido** el 26 de octubre de 2007 y su **publicación aprobada** el 20 de febrero de 2008.

**Abstract**

**Botero, Sergio y Cano, Jovan Alfonso. "Time series analysis for prediction of energy prices on the Colombian stock exchange," Cuadernos de Economía, v. XXVII, n. 48, Bogota, 2008, pages 173-208**

Because of the restructuring of the Colombian electricity sector over the last two decades, the behavior of the price of electrical energy has shown increased volatility, reflecting the risk that exists for the different agents who intervene in this market. The purpose of this article is to present a methodology for the implementation of regression models on the historical series of stock market prices of energy in Colombia. As the quantity of data increases, broader models can be developed to adequately describe market behaviors that are impossible to identify using currently available techniques and information.

**Keywords:** energy market, spot market, time series, market intervention. **JEL:** C13, C15, C53, Q49.

**Résumé**

**Botero, Sergio y Cano, Jovan Alfonso. « Analyse de séries temporelles pour la prévision des prix en bourse de l'énergie en Colombie », Cuadernos de Economía, v. XXVII, n. 48, Bogota, 2008, pages 173-208**

Comme conséquence des réformes structurelles du secteur énergétique colombien, pendant les deux dernières décennies, le comportement des prix dans ce secteur a augmenté sa volatilité. Cette volatilité reflète le risque existant pour les différents agents qui interviennent sur le marché. L'objectif de cet article est de présenter une méthodologie pour l'emploi de modèles de régression, sur la série historique de prix en bourse de l'énergie en Colombie. Dans la mesure où la quantité de données disponibles augmente, on pourra développer des modèles plus vastes, qui décrivent d'une façon adéquate certains comportements du marché, dont l'identification n'est pas possible avec les techniques et l'information disponibles actuellement.

**Mot clés :** marché de l'énergie, spot, séries temporelles, régulation du marché. **JEL :** C13, C15, C53, Q49.

Los mercados competitivos de este bien son muy recientes en la historia, dado que desde sus comienzos esta industria fue un monopolio natural. En algunos casos de carácter estatal y en otros de carácter privado con una alta regulación por parte del Estado.

A finales de los años 80 y el decenio de los 90, muchos sistemas energéticos en el mundo y, en especial en los países latinoamericanos, tuvieron reformas muy relevantes, las cuales han mostrado gran diversidad en cuanto a su profundidad, alcance y velocidad de ejecución (OLADE, CEPAL, GTZ 2003). Entre las principales razones para esta reestructuración es posible destacar la búsqueda de: eficiencia, libre competencia en un sector tradicionalmente estatal y transparencia en el proceso de formación de los precios (Millan 1999). De igual forma, la introducción de nuevas tecnologías en generación permitió aprovechar las economías de escala, y los avances en computación permitieron determinar, en forma más exacta, las interacciones en los sistemas eléctricos.

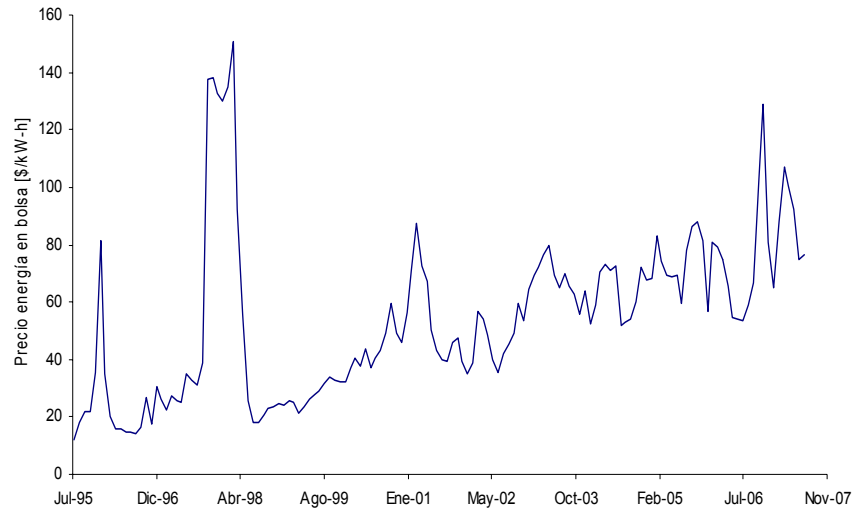
La introducción de la competencia en la generación de electricidad contribuyó a que los precios comenzaran a reflejar el costo real de la producción de la energía y a responder a variaciones de los determinantes de este precio, por ejemplo, cambios en la disponibilidad del agua, precios internacionales de los hidrocarburos, ciclos económicos y condiciones particulares del mercado (restricciones en la red, posición dominante, entre otras).

La desregularización de los mercados de energía y la introducción de la competencia han conducido a cambios que afectan la precisión en los pronósticos y, por tanto, no es posible garantizar los beneficios para los agentes del mercado. El riesgo de mercado en las negociaciones es considerable, por ejemplo, la volatilidad de los precios en el mercado spot para Colombia ha presentado valores superiores al 350% en una

escala mensual<sup>1</sup>. A pesar de ello la predicción de los precios de la Energía en Bolsa presenta patrones de comportamiento –por ejemplo, estacionalidades en diferentes escalas de tiempo–, los cuales pueden ser incorporados en modelos de regresión (Weron 2005).

La serie con los precios históricos de la Energía en bolsa mensual en Colombia, en \$/kW-h se presenta en la Gráfica 1<sup>2</sup>.

GRÁFICA 1  
PRECIOS DE BOLSA HISTÓRICOS MENSUAL [\$/KW-H] (\$ CORRIENTES)



Fuente: Base de datos NEON.

Es importante recordar que la bolsa de energía en Colombia nació a mediados de 1995, por tanto, sólo se dispone de una serie de datos de 12 años (144 meses). En este orden de ideas, el análisis que se haga puede orientar la evolución de esta metodología, la cual puede ser refinada y llegar a estudios concluyentes cuando se disponga de un mayor número de datos para el análisis.

<sup>1</sup> Valor publicado en el Boletín N° 16 Expertos en Mercados (XM), 14 de diciembre de 2005.

<sup>2</sup> Base de datos NEON, disponible en el sitio web: [www.xm.com.co](http://www.xm.com.co).

Estudiar los factores que afectan la formación de los precios de bolsa e identificar nuevas metodologías para la predicción de los mismos, tiene un objetivo final (generalmente, intangible para muchos): el empleo de este tipo de herramientas debe reflejarse sobre el precio del kW-h pagado por los usuarios. En otros términos, proyectar adecuadamente el futuro, podría mejorar el desempeño económico de los agentes y aumentar la eficiencia del mercado; en este sentido, el precio final del producto debería reflejar cada vez más el valor real de la energía.

### **MERCADO DE ENERGÍA MAYORISTA**

Durante los años ochenta, el sector eléctrico colombiano entró en crisis, al igual que en otros de países de América Latina. El deterioro en su desempeño se debió, especialmente, al subsidio de tarifas y a la politización de las empresas estatales. De igual forma, los grandes proyectos de generación, llevados a cabo con sobrecostos y atrasos considerables, condujeron a que el sector se convirtiera en una gran carga para el Estado.

Por otra parte, en el mundo comenzó a ponerse en duda la eficacia de los monopolios estatales para la prestación de los servicios públicos, iniciándose grandes reformas en países como: Nueva Zelanda, Inglaterra, los miembros de la Unión Europea, Australia y Argentina (Rudnick 1999).

Dentro de los principales motivos para la revisión del esquema de funcionamiento del sector eléctrico en Colombia, se encuentran:

- El racionamiento de energía en 1992.
- El elevado porcentaje de la deuda del sector eléctrico dentro de la deuda externa nacional.
- Los altos costos de la energía.
- La ineficiencia en la administración de muchas empresas.
- No prioridad en el gasto público para mejorar los servicios públicos.
- Un régimen tarifario inadecuado que no permitía recuperar la inversión.

- La baja cobertura (sobre todo en el sector rural).
- La calidad deficiente en la prestación de los servicios (no se cumplía con estándares internacionales).
- La inexistencia de políticas de protección al usuario.
- La concentración de funciones: operación, regulación y control.

A lo anterior se suma que la Constitución Política de 1991 planteó lineamientos como:

- La cobertura y la calidad de servicios públicos como indicadores del nivel de vida de una nación.
- Los servicios públicos son inherentes a la finalidad social del Estado.
- El acceso masivo a los servicios públicos como un factor fundamentado en el proceso económico y social del país.
- La competencia mejora la productividad.
- El fortalecimiento de la inversión en lo social.
- El Estado no debe ser empresario ejecutante sino regulador-controlador.

En particular, el cuestionamiento sobre la eficiencia del papel del Estado como administrador de los servicios públicos, llevó al planteamiento de propuestas como (Kirschen 2006):

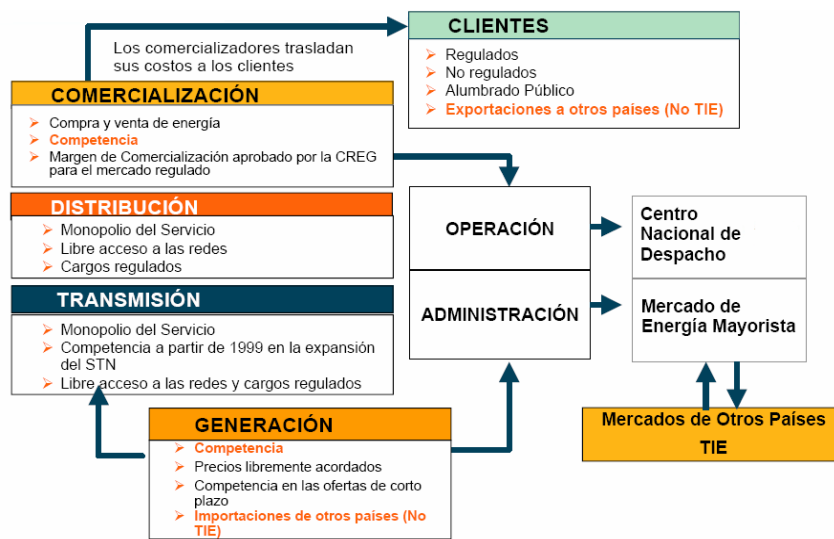
- Introducción de competencia en el sector eléctrico.
- Penetración de la inversión privada, incluso privatización de compañías estatales.
- Eliminación de la integración vertical, separando los negocios de generación, transmisión, distribución y comercialización.
- Consolidación del papel del Estado como ente regulador.

En este contexto, a principio de los años 90, en Colombia, se vio la necesidad de modernizar el sector eléctrico, abriéndolo a la participación privada y siguiendo un esquema similar al de los países pioneros en este desarrollo, en especial el Reino Unido. Esta reestructuración se materializó con la promulgación de las Leyes 142 y 143 de 1994, las cuales definieron el marco legal que permitió el desarrollo de un

esquema de mercado en competencia. Estas leyes crearon el Mercado Mayorista de Energía Eléctrica (Raigoza 1999).

La reglamentación de este mercado fue desarrollada por la Comisión de Regulación de Energía y Gas (CREG). Para este propósito, la Comisión se asesoró de consultores nacionales e internacionales y con el apoyo de las empresas del mismo sector, promulgó las reglamentaciones básicas y puso en funcionamiento el nuevo esquema a partir del 20 de julio de 1995. La estructura actual del Mercado Mayorista se presenta en la Figura 1, en ella se encuentran definidos los agentes que intervienen en el mercado y las características básicas de cada negocio (comercialización, distribución, transmisión y generación).

FIGURA 1  
ESTRUCTURA DEL MERCADO DE ENERGÍA MAYORISTA



Fuente: Compañía de Expertos en Mercados (XM).

### FORMACIÓN DE LOS PRECIOS DE LA ENERGÍA EN BOLSA EN COLOMBIA

La Bolsa de Energía es una figura comercial tipo subasta en sobre cerrado, la cual toma las ofertas de precio y declaración de disponibilidad

hora a hora de cada recurso de generación y establece un programa diario de generaciones, igualmente, hora a hora.

Al inicio de todos los días, en forma confidencial, cada agente generador hace una oferta de precio (en \$/MWh) y la declaración de disponibilidad (en MW) de cada uno de sus recursos de generación. Se presenta una única oferta para las 24 horas del día siguiente.

El registro histórico de estas ofertas y la declaración de disponibilidad se encuentran para consulta pública, de esta manera, los agentes pueden conocer la información de todos los oferentes en el mercado. El análisis de las ofertas de los agentes, y de la intervención de ofertas y disponibilidades, es muy valioso puesto que, los demás agentes pueden conocer no sólo el precio sino la disponibilidad histórica de generación de sus competidores.

El ejercicio de la Bolsa de Energía ha sido construido sobre la hipótesis de competencia perfecta, es decir, información perfecta, numerosos oferentes y demandantes, inexistencia de posiciones dominantes, neutralidad del dinero, no intervención del Estado, inexistencia de barreras para las entradas y salidas, entre otros. En los modelos de competencia perfecta se obtienen soluciones de tipo Paretiano que maximizan el bienestar del consumidor y concluyen que la decisión de precios que más beneficio puede traerle a los competidores es ofertar un precio igual a su costo marginal (Jaramillo 2000).

Tal como se observa en la Figura 2, el Centro Nacional de Despacho (CND<sup>3</sup>) recibe las ofertas de los generadores e inicia el predespacho de energía basado en las proyecciones de demanda construidas a partir de: datos históricos, modelos de planeación y estudios dados por la Unidad de Planeación Minero Energética (UPME<sup>4</sup>), sin considerar las restricciones del sistema; es decir, que no se toman en consideración los límites operativos de los elementos asociados al Sistema Interconectado Nacional (SIN) ni las indisponibilidades de los mismos no declaradas

---

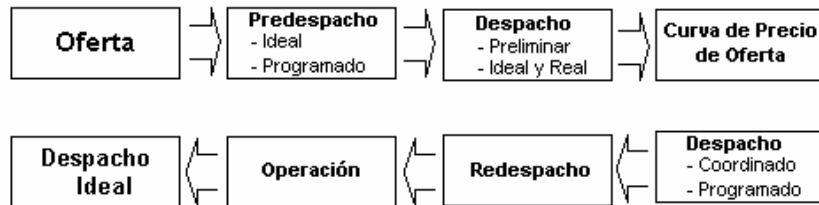
<sup>3</sup> Es una de las gerencias de la compañía XM (Compañía de Expertos en Mercados), encargada de operar, administrar y desarrollar los mecanismos del Mercado de Energía Mayorista establecidos por el regulador (CREG).

<sup>4</sup> Unidad administrativa encargada de realizar la planeación del desarrollo sostenible de los sectores de minas y energía en Colombia.



por los agentes (por ejemplo, voladuras de torres). Posteriormente, se realiza el despacho de energía considerando las restricciones del sistema y se incorporan las curvas de oferta de los generadores y las inflexibilidades de los mismos, por lo tanto, se tienen en cuenta las curvas de operación de cada generador dependiendo de su estado anterior (frio, tibio, caliente<sup>5</sup>), tanto para el arranque como para su parada.

FIGURA 2  
 PROCESO PARA EL CÁLCULO DEL DESPACHO IDEAL



Fuente: Compañía de Expertos en Mercados (XM).

Simultáneamente a la operación del sistema y debido a las contingencias que pueden presentarse, se realiza el redespacho de energía, el cual tiene como fin orientar la operación de la forma más económica posible. Finalmente, la información obtenida de la operación, del redespacho y de las inflexibilidades de los generadores, es empleada para realizar el cálculo del Precio de la Energía en Bolsa (Despacho Ideal), el cual realiza el cubrimiento de la demanda, despachando las generaciones obligadas: plantas inflexibles, en prueba, no despachadas centralmente, filo de agua y cogeneradores<sup>6</sup> (su valor de oferta no afecta el Precio de Bolsa).

Posteriormente, se despacha los generadores con los precios de oferta más bajos acorde con la disponibilidad comercial (cantidad de MW), hasta cubrir la demanda total. De esta forma, el Precio de Bolsa, es el precio de oferta del último recurso de generación no-inflexible requerido para atender la demanda real hora a hora (Gráfica 2); y es con este

<sup>5</sup> El estado de un generador térmico depende de la cantidad de horas que ha estado apagado.

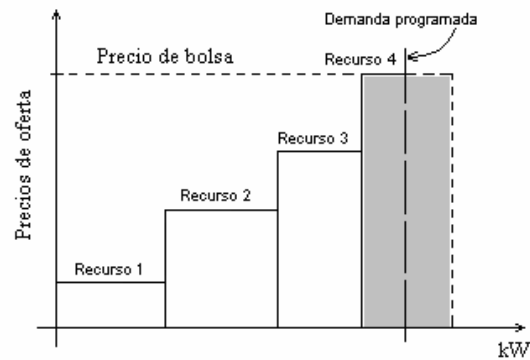
<sup>6</sup> Cogenerador es aquella persona natural o jurídica que produce energía utilizando un proceso de Cogeneración. Entendiendo como Cogeneración, el proceso de producción de energía eléctrica y energía térmica, que hace parte integrante de una actividad productiva, destinadas ambas al consumo propio o de terceros y destinadas a procesos industriales o comerciales.

precio que se realiza la liquidación de las transacciones realizadas por los agentes (Cano 2004).

Teniendo clara la dinámica empleada para el cálculo del Precio de Bolsa, es importante resaltar que a pesar de considerar las limitantes del sistema durante todo el proceso, éstas finalmente no afectan el Precio de Bolsa. En consecuencia el valor calculado obedece, básicamente, al punto de equilibrio entre la demanda real y las ofertas de los agentes generadores, evaluada hora a hora.

El despacho ideal se simula para determinar un balance sin las restricciones eléctricas que vician el proceso comercial. Este despacho considera que las redes tienen capacidad de transporte ilimitada, convirtiéndose así en un juego comercial de oferta y demanda.

GRÁFICA 2  
DESPACHO DE LOS RECURSOS DE GENERACIÓN PARA LA  
FORMACIÓN DEL PRECIO DE BOLSA



Fuente: elaboración propia.

A partir de la Gráfica 2, se podría inferir que los generadores con menores costos serían los que ofertarían a menor precio y, por ende, serían despachados completamente (caso de los recursos 1 y 2), mientras que, un generador que tenga mayores costos ofertaría a un precio mayor (caso de los recursos 3 y 4). Como se observa el recurso 4 no generaría toda su capacidad, puesto que, proveería energía hasta el punto en el cual se cubriera la demanda programada, determinando el precio de

bolsa o marginal en ese momento, este tipo de generador se denomina *generador marginal*.

En un mercado libre los agentes pueden ofertar a un precio determinado con libre albedrío; sin embargo, el sistema regulatorio ha establecido algunos parámetros para orientar lo que se espera sea la oferta de cada generador. La regulación CREG 055 de 1994, emitida cuando se estaba comenzando el nuevo mercado de energía en Colombia, es un claro ejemplo de las señales dadas a los agentes que tomen la decisión de participar en el mercado.

De acuerdo con la resolución CREG 055 de 1994 Artículo 6 – Ofertas de precio en la bolsa de energía:

“Los precios a los cuales las empresas generadoras ofrecen diariamente al CND, deben reflejar los costos variables de generación en que esperan incurrir.

Para plantas Termoeléctricas:

Precio de oferta = CSC + CTC + COM + CAP + OCV + Utilidades

(CSC) Costo de Suministro de Combustible [\$/MWh]

(CTC) Costo de Transporte de Combustible [\$/MWh]

(COM) Costo de Operación y Mantenimiento [\$/MWh]. Depende en gran medida del tipo de tecnología empleada gas natural, carbón y otros combustibles.

(CAP) Costo de Arranque-Parada [\$/MWh]. Es el costo asociado a cada arranque-parada de cada unidad de generación.

(OCV) Otros Costos Variables [\$/MWh]. Corresponden a los siguientes Costos Variables calculados por el ASIC, como: CEE (CERE), FAZNI, Aportes Ley 99 de 1993 y Costo Unitario por Servicio de AGC proporcional a la Generación<sup>7</sup>.

Para plantas hidroeléctricas:

---

<sup>7</sup> ASIC: Administrador del sistema de intercambios comerciales (dependencia del CND); CEE: Costo equivalente de energía; CERE: Costo equivalente real en energía del cargo por capacidad; FAZNI: Fondo de apoyo financiero para la energización de zonas no interconectadas; AGC: Regulación secundaria de frecuencia.

Precio de oferta = costo de oportunidad (valor de agua) + COM + Utilidades

El costo de oportunidad hace referencia a la decisión de generar en el momento de la oferta, teniendo en cuenta la operación económica a mediano y largo plazo del Sistema Interconectado Nacional.

Las utilidades en ambos casos pueden estar incluidas dentro de los costos, proporcional al capital de cada uno de estos”.

### **SERIES DE TIEMPO**

Los pronósticos se realizan con el fin de orientar las decisiones en muchas áreas del mundo como los mercados, el transporte, la identificación de fallas, el clima, entre otros. Para pronosticar una variable se debe construir un modelo y estimar sus parámetros usando datos históricos, es decir, logrando una caracterización estadística de los enlaces entre el presente y el pasado. En forma simple se usa el modelo estimado del pronóstico para extrapolar los datos observados (Diebold 2001).

En el caso de la energía, algunos modelos ARIMA y ARMAX han sido probados para la serie de tiempo del Mercado de California, empleando como variable exógena la carga del sistema (Weron 2005). Otro tipo de modelos y comparación de resultados como: modelos lineales, regresión dinámica, función de transferencia y redes neuronales han sido evaluados para el Mercado de PJM (Guang 2005).

La estimación de precios con los modelos desarrollados en este trabajo se realiza para un horizonte mensual, el cual será de gran utilidad para un agente comercializador al momento de tomar decisiones de contratos a corto o largo plazo, comprar en bolsa, pujar para obtener un mejor precio, entre otras. De igual manera, sería de utilidad para un agente generador que decida acumular sus recursos en busca de un mejor precio.

En otros términos, conocer con certeza un valor o el intervalo de valores posibles del precio de la energía en bolsa para meses futuros, implica para los agentes del mercado la reformulación de muchas de sus estrategias de compra o venta de energía.

## Definiciones

$Z_t$ : Variable aleatoria definida en el tiempo  $t$ .

$z_t$ : Valor tomado por la variable aleatoria  $Z_t$  en el tiempo  $t$  o realización de  $Z_t$  en el período  $t$ .

$z_1, z_2, z_3, \dots, z_n$ : Serie de tiempo que tiene  $n$  observaciones o una realización del proceso  $Z_t$  cuando  $t$  varía entre 1 y  $n$ .

$n$ : Origen de los pronósticos. Entero que indica la posición de la última observación disponible, el presente o el punto a partir del cual se empieza a pronosticar.

Como  $z_t$  es un valor que toma la variable aleatoria  $Z_t$  en el tiempo  $t$ , si  $t < n$ ,  $z_t$  es una observación del pasado; si  $t = n$ ,  $z_t$  es la observación presente; y si  $t > n$ ,  $z_t$  es un valor que tomará la variable  $Z_t$ . Cabe anotar que los valores de  $z_t$  para  $t > n$  son desconocidos y es el objeto de estudio del presente trabajo.

## Modelos de regresión

El objetivo de regresión es estimar los parámetros asociados con una relación funcional  $z_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} \dots + \alpha_n x_{t-n} + \varepsilon_t$ , que se supone existe entre el valor esperado de la variable dependiente y las variables independientes, minimizando la suma de los cuadrados de los errores de una muestra de  $n$  observaciones de la forma  $(Z_i, x_{i1}, x_{i2}, x_{i3}, \dots, x_{in})$  que, teóricamente, satisfacen la relación funcional.

Para inferir sobre los parámetros estimados es necesario establecer algunos supuestos sobre las variables independientes y sobre los errores que deben validarse antes de utilizar los resultados obtenidos de la estimación del modelo.

Los supuestos de los residuales en el modelo de regresión son:

- Las variables independientes del modelo son determinísticas.
- Los errores son variables aleatorias que cumplen:
  - $E[\varepsilon_j] = 0$  para  $j = 1, 2, \dots, n$  la media es constante e igual a cero.
  - $Var[\varepsilon_j] = \sigma^2$  para  $j = 1, 2, \dots, n$  la varianza es constante e igual  $\sigma^2$ .

- $Cov(\varepsilon_i, \varepsilon_j) = 0$   $\varepsilon_i$  y  $\varepsilon_j$  incorrelacionados para todo  $i \neq j, i, j = 1, 2, \dots, n$ .
- $\varepsilon_j \sim N(0, \sigma^2)$  donde  $j = 1, 2, \dots, n$ .

La media cero indica que la función de autocorrelación coincide con su autocovarianza y considerar que la varianza del proceso es constante, indica un proceso estacionario. Adicionalmente, el ruido blanco en sentido estricto presenta un valor en cada instante del tiempo que no depende de cual haya sido su valor en los instantes precedentes y que no ejerce ninguna influencia en sus valores futuros, es decir, que no es posible obtener más información de la serie de tiempo que aporte a la estimación de valores futuros a partir de un modelo de regresión (Diebold 2001).

#### *Modelo autoregresivo*

Para implementar el modelo de estimación de precios de la energía en bolsa, se partió de la serie histórica de datos entre enero de 1996 y junio de 2007, tal como se observa en la Gráfica 1 y de la cual se realiza un análisis para establecer las siguientes características: tendencias pronunciadas o ciclos, existencia de estabilidad o de periodicidad en las observaciones, variancia de las observaciones (constante o variable en el tiempo), existencia de observaciones inusuales y de puntos extremos, cambios en la estructura de la serie, entre otras.

De esta manera se encontraron las siguientes observaciones sobre la serie de tiempo:

- Muestra una irregularidad para los precios comprendidos entre septiembre de 1997 y abril de 1998, período durante el cual se presentó el fenómeno del Niño, sumado a la reciente implementación de la Bolsa de Energía. Estos dos factores desencadenaron un alza en los precios presentándose la intervención del agente regulador. Este punto es clave para determinar la porción de los datos que puede ser despreciada al desarrollar el modelo.
- Durante los meses de mayo a septiembre de 2000, el precio de bolsa presenta una tendencia creciente que puede asociarse con los incrementos del precio de oferta de los agentes generadores térmicos. En septiembre, la CREG expidió la Resolución CREG 063 como una medida para reducir la volatilidad de las

ofertas, lo cual se reflejó de manera transitoria en el precio de bolsa. A partir de noviembre de 2000 se observa un alza del precio que combina factores como los atentados que ocasionaron atrapamientos<sup>8</sup> de energía hidráulica y el inicio del período de verano. Luego de la expedición de las resoluciones CREG 026 de 2001 y CREG 034 de 2001, el precio de bolsa disminuyó de una manera significativa, coincidiendo con una reducción de los atentados a la infraestructura (UPME 2002).

- Entre septiembre y octubre de 2006 se observa un incremento súbito en los precios, relacionado con una alerta dada por el Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia (IDEAM) respecto a un fenómeno del Niño que se presentaría en el corto plazo y que en realidad no ocurrió. Ambos valores serán considerados como *outliers*<sup>9</sup> y, por tanto, se calcularán valores promedio históricos para el análisis de los modelos.
- Se han identificado los períodos en los cuales el agente regulador ha decidido intervenir el mercado, al considerar que existen irregularidades de las cuales puede sacarse provecho como son las cargas atrapadas, restricciones del sistema, entre otros. Estos períodos y fechas de intervención serán analizados e ingresados al modelo de estimación.

Ahora es necesario determinar si un modelo autoregresivo es apropiado para explicar el comportamiento futuro de los Precios de la Energía en Bolsa; es decir, se debe identificar una estructura de correlación entre las variables aleatorias del proceso estocástico que genera la serie.

Para ello es necesario estudiar la Función de Autocorrelación (FAC), conociendo que un proceso estocástico  $Z_t / t \in I \subset Z$ , es estacionario en covarianza o débilmente estacionario de orden 2 si:

*i.* Para todo  $t \in I$ , la media y varianza existen y son constantes ( $E[Z_t] = \mu$  y  $Var[Z_t] = \sigma_z^2$ ) y

---

<sup>8</sup> Término empleado para indicar que una planta de generación no puede suministrar energía eléctrica al sistema debido a restricciones del mismo.

<sup>9</sup> Observación numéricamente distante de los demás datos (promedio y/o tendencia). Un análisis más detallado de este tipo de datos permiten obtener una estimación significativamente mejor (Contreras 2005).

ii. Si para cualquier  $i, j \in I$ ,  $Cov(z_i, z_j)$  depende únicamente del número de períodos que separa las variables, es decir, de  $j - i$ , que se denota  $k$ : número de rezagos (Correa 2000).

A partir de la FAC obtenida en SAS se observa que la variable “precios de Bolsa de Energía (MEM)” presenta una alta correlación con sus rezagos hasta el orden 3. Esto quiere decir que un valor determinado de la serie en cualquier instante de tiempo, depende de los valores anteriores que la serie hubiese tomado hasta 3 períodos de tiempo atrás.

Adicionalmente, con los resultados de la Función de Autocorrelación Parcial (FAP) se infiere que el orden que debe ser aplicado en el modelo de regresión y que garantice que sus parámetros sean significativos, es dos. Por lo tanto, el modelo inicial sería un modelo autoregresivo de orden 2 AR(2), es decir, un modelo con la siguiente forma:

$$z_t = \alpha_0 + \alpha_1 z_{t-1} + \alpha_2 z_{t-2} + \varepsilon_t \quad [1]$$

donde:

$z_t$ : variable aleatoria

$z_{t-1}$ : variable aleatoria rezagada un período de tiempo

$z_{t-2}$ : variable aleatoria rezagada dos períodos de tiempo

$\alpha_0$ : intercepto (tendencia)

$\alpha_1 \alpha_2$ : parámetros del modelo

$\varepsilon_t$ : ruido blanco

Después de identificar el orden del modelo de regresión, es necesario determinar sus parámetros y verificar la significancia de los mismos. Algunos de los resultados obtenidos en SAS se observan en el Cuadro 1.

A partir del cálculo del Total R-Square, se tiene que el modelo representa la serie de precios en un 74,62%; los valores de criterio de información de Akaike (AIC) y del criterio de información bayesiano de Schwarz (SBC) son empleados para comparar los modelos y determinar cuál permite una mejor estimación.



CUADRO 1  
RESULTADOS MODELO AUTOREGRESIVO DE ORDEN 2

The AUTOREG Procedure					
Model	AR2				
Dependent Variable	mem				
Ordinary Least Squares Estimates					
SSE	29626,7578	DFE	139		
MSE	213,14214	Root MSE	14,59939		
SBC	1176,2121	AIC	1167,34462		
Regress R-Square	0,7462	Total R-Square	0,7462		
Durbin-Watson Statistics					
Order	DW	Pr < DW	Pr > DW		
1	1,9751	0,4066	0,5934		
2	2,1019	0,7391	0,2609		
3	2,0227	0,6034	0,3966		
Standard					
Variable	DF	Estimate	Error	t Value	Approx Pr >  t
Intercept	1	9,5016	2,6859	3,54	0,0005
lrmem	1	1,0093	0,0835	12,09	<,0001
lrmem2	1	-0,1775	0,0830	-2,14	0,0342

Fuente: elaboración propia.

De los estadísticos generalizados de Durbin-Watson y su respectivo valor  $p$  se tiene que a un nivel de significancia de 0,05, la prueba de hipótesis indica que  $\phi_1 = 0$ ,  $\phi_2 = 0$  y  $\phi_3 = 0$ , dado que todos son mayores a 0,05. Esto nos lleva a concluir que el supuesto de errores no correlacionados se debe cumplir.

Los parámetros estimados para el intercepto y los rezagos 1 y 2 presentan además el cálculo de las desviaciones estándar y la información necesaria para probar su significancia, dado que en este caso todos los valores de  $p$  son menores que 0,05, los parámetros son significativos.

Para evaluar si la serie de tiempo de los residuales obtenida, luego de aplicar el modelo autoregresivo es ruido blanco, se debe determinar el cumplimiento de las condiciones descritas anteriormente. Los resultados de la evaluación de los residuales se presentan en el Cuadro 2.

CUADRO 2  
PRUEBA SOBRE LA SERIE DE RESIDUALES

Modelo	Media	Var.	Des	Estadísticos de prueba				
				Mediana	Asim.	Curtosis	Wnormal	Vpnorm
Auto-regresivo AR(2)	0	210,12	14,5	-1,46	1,72	13,78	0,81	0,00

Fuente: elaboración propia.

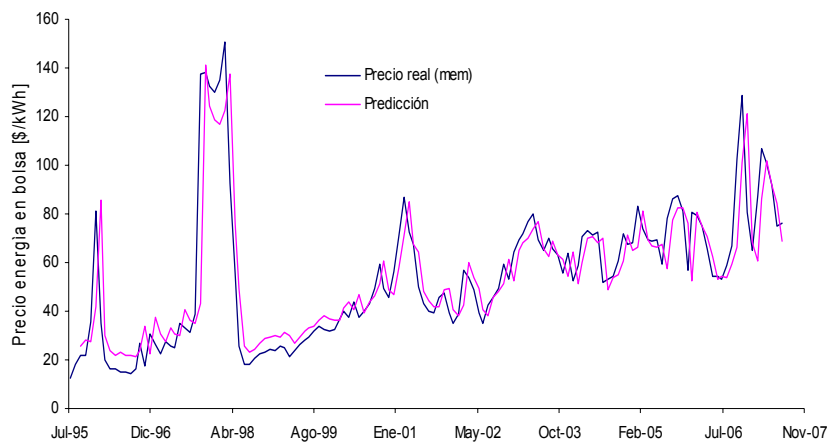
Con los resultados del Cuadro 2 se observa que la serie de tiempo formada por los residuales del modelo autoregresivo, **no cumple** las condiciones del supuesto de normalidad de los errores. El valor  $p$  de la prueba de Shapiro-Wilk es  $v_{pnor} = 0,00$  (menor que el valor de significancia de 0,05), luego se rechaza  $H_0$ , es decir, los errores NO son normales.

La Gráfica 3 presenta el comportamiento del modelo comparado con los datos históricos de la serie, a pesar de que la serie de residuales no hubiese cumplido con los supuestos de ruido blanco.

#### *Modelo de estacionalidades*

Una serie de tiempo tiene componente estacional de período  $s$ , si presenta patrones que se repiten con una frecuencia constante ( $s$ ), produciendo en ella un efecto periódico (Diebold 2001). Para realizar la evaluación sobre la serie de tiempo estudiada, se procederá de la siguiente forma:

GRÁFICA 3  
MODELO AR(2) -VS- PRECIOS DE BOLSA HISTÓRICOS



Fuente: elaboración propia.

- i. Se caracteriza gráficamente cada una de las formas funcionales, mediante procesos estacionales simulados, mostrando cada una de las componentes que determinan la serie.

- ii. Luego se caracterizan las series estacionales puras, para buscar funciones de tiempo que permitan generarlas, lo cual lleva al conjunto de  $s$  variables indicadoras (funciones de tiempo) o por  $s$  funciones trigonométricas de período  $s$ .
- iii. Conocidas las formas funcionales que generan la parte estacional de la serie, se emplea el método de regresión para estimar los parámetros, que acompañan las  $s$  funciones de tiempo que definen la componente estacional.
- iv. Finalmente, se emplean los resultados obtenidos y se combinan con los modelos estudiados anteriormente, con el fin de obtener un mejor modelo de estimación.

Definición 1: La variable indicadora de la estación  $j$  ( $1 \leq j \leq s$ ) de una serie estacional pura ( $S_t$ ), de período estacional  $s$ , notada  $I_{jt}$ , es una serie de longitud igual a  $S_t$ , está definida por la siguiente expresión:

$$I_{jt}: \text{variable indicadora de la estación } j = \begin{cases} 1 & \text{si } t \text{ está en la estación } j \\ 0 & \text{si } t \text{ no está en la estación } j \end{cases}$$

De esta forma se obtienen  $I_{1t}, I_{2t}, \dots, \text{ y } I_{st}$ ,  $s$  variables indicadoras que forman una base.

La representación funcional del modelo para una serie con componente estacional, tendencia y componente aleatoria (en este caso con errores autoregresivo) es:

$$Z_t = \beta_1 t + \sum_{i=1}^n \delta_i \text{mes}_{it} + U_t \quad \text{y} \quad U_t = \sum_{j=1}^k \phi_j U_{t-j} + \varepsilon_t \quad [2]$$

donde:

$Z_t$ : variable aleatoria

$\text{mes}_{it}$ : variable indicadora de estación

$\delta_i$ : parámetro de la estación  $i$

$\beta_1$ : pendientes de la tendencia

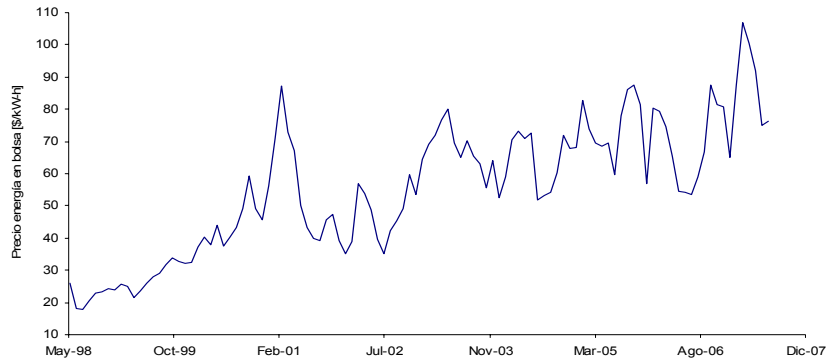
$U_t$ : residuales

$\varepsilon_t$ : ruido blanco

Según el análisis gráfico, se identificó una irregularidad en los precios entre septiembre de 1997 y abril de 1998. Por tal motivo, se procedió a analizar la serie de tiempo desde abril de 1998 hasta junio de 2007, considerando los

valores de septiembre y octubre de 2006 como *outliers*. La Gráfica 4 presenta la nueva serie de tiempo analizada para el modelo de estacionalidades.

GRÁFICA 4  
PRECIO DE BOLSA HISTÓRICO MENSUAL [\$/KWH]



Fuente: elaboración propia.

Sobre la nueva serie de datos se aplica la metodología descrita, en busca de la tendencia, las estacionalidades y los errores autorregresivos que sobre la misma puedan existir.

Se evaluaron estacionalidades entre 2 y 12 meses identificando una estacionalidad anual con errores autorregresivos de orden 1, en la cual todos los parámetros son significativos, como se observa en el Cuadro 3.

El parámetro estimado para el rezago de orden uno es significativo  $t_{\text{value}} = |10.66|$ ; la comparación se realiza con el valor 1,96, dado que para las pruebas de hipótesis, se emplea la distribución t con 47 grados de libertad que se aproxima a una normal.

El modelo identificado con las estacionalidades y la tendencia representa los precios en un 85,06%, mientras que el modelo completo incluyendo la parte autoregresiva en los errores representa la serie de precios en un 98,67%.

De los estadísticos generalizados de Durbin-Watson y su respectivo valor  $p$  se tiene que a un nivel de significancia de 0,05, la prueba de hipótesis indica que  $\phi_1 = 0$ ,  $\phi_2 = 0$  y  $\phi_3 = 0$ , dado que todos son mayores a 0,05. Esto nos lleva a concluir que el supuesto de errores no correlacionados se debe cumplir.

CUADRO 3  
RESULTADOS DEL MODELO CON TENDENCIA, ESTACIONALIDAD  
ANUAL Y ERRORES AUTORREGRESIVOS DE ORDEN 1

Dependent Variable		mem			
Yule-Walker Estimates					
Estimates of Autoregressive Parameters					
Lag	Coefficient	Standard Error		t Value	
1	-0,736299	0,069061		-10,66	
SSE	5107,51644	DFE	96		
MSE	53,20330	Root MSE	7,29406		
SBC	800,933264	AIC	763,126539		
Regress R-Square	0,8506	Total R-Square	0,9867		
Durbin-Watson Statistics					
Order	DW	Pr < DW	Pr > DW		
1	1,9552	0,4145	0,5855		
2	1,8982	0,3409	0,6591		
3	2,1255	0,7913	0,2087		
Standard					
Approx Variable	DF	Estimate	Error	t Value	Pr >  t
t	1	0,5206	0,0772	6,74	<,0001
in1	1	22,8736	5,5586	4,11	<,0001
in2	1	17,4798	5,6128	3,11	0,0024
in3	1	21,1145	5,6257	3,75	0,0003
in4	1	23,4980	5,6375	4,17	<,0001
in5	1	28,5773	5,6520	5,06	<,0001
in6	1	27,3289	5,6694	4,82	<,0001
in7	1	22,4453	5,6883	3,95	0,0002
in8	1	23,4922	5,7058	4,12	<,0001
in9	1	31,0645	5,7181	5,43	<,0001
in10	1	36,9088	5,7198	6,45	<,0001
in11	1	32,3781	5,7029	5,68	<,0001
in12	1	29,4817	5,6557	5,21	<,0001

Fuente: elaboración propia.

Los parámetros estimados para las estacionalidades (in1 a in12) y la tendencia (t) presentan además el cálculo de las desviaciones estándar y la información necesaria para probar su significancia. En este caso todos los valores de  $p$  son menores que 0,05; por tanto, los parámetros son significativos.

Hasta el momento, todos los parámetros estimados de este modelo (parte estacional y errores autorregresivos) son significativos. La validación sobre los supuestos de los errores se presenta en los Cuadros 4 y 5.

CUADRO 4  
PRUEBA DE NORMALIDAD SOBRE LA SERIE DE RESIDUALES

Modelo	Estadísticos de prueba							
	Media	Var.	Des	Mediana	Asim.	Curtosis	Wnormal	Vpnorm
Estacionariedades	-0,02	46,86	6,85	-0,8	0,16	0,22	0,98	0,22

Fuente: elaboración propia.

El supuesto de normalidad de los errores se satisface, puesto que el valor  $p$  de la prueba de Shapiro-Wilk es  $v_{pnor} = 0,22$ ; se acepta  $H_0$  lo que quiere decir que los errores son normales.

CUADRO 5  
PRUEBA DE IGUALDAD DE VARIANZA SOBRE LA SERIE DE RESIDUALES

		The TTEST Procedure							
		Statistics				Upper CL			
Variable	Class	Lower CL	Upper CL	Lower CL	Upper CL	Lower CL	Upper CL	Lower CL	Upper CL
Std Err		N	Mean	Mean	Mean	Std Dev	Std Dev	Std Dev	Std Dev
res32	A	55	-1,602	-0,178	1,2463	4,4345	5,2674	6,4887	
0,7103									
res32	B	55	-2,069	0,1402	2,3495	6,8799	8,1722	10,067	
1,1019									
res32	Diff (1-2)		-2,917	-0,318	2,2807	6,0677	6,875	7,9321	
1,311									
		T-Tests							
Variable	Method	Variances		DF	t Value	Pr >  t			
res32	Pooled	Equal		108	-0,24	0,8089			
res32	Satterthwaite	Unequal		92,3	-0,24	0,8089			
Equality of Variances		Method	Num DF	Den DF	F Value	Pr > F			
res32	Folded F		54	54	2,41	0,0016			

Fuente: elaboración propia.

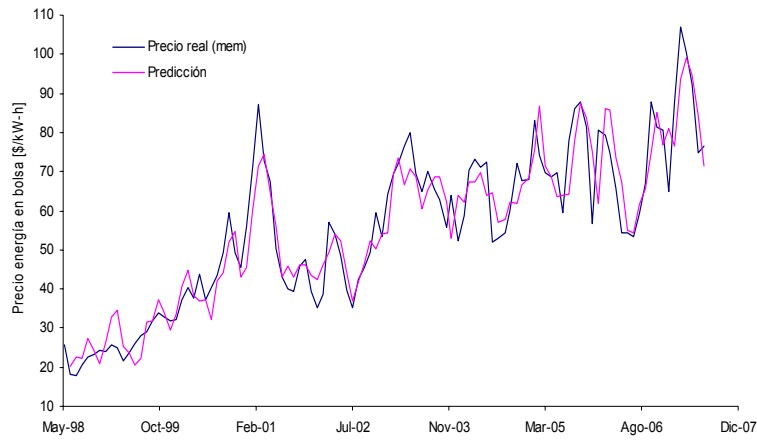
De los resultados anteriores se observa que no se cumple la igualdad de la varianza,  $Pr > F = ,0016$  es menor que el valor de significancia 0,05, dado lo anterior no se continúa con las demás sobre los residuales.

La Gráfica 5 presenta el comportamiento del modelo comparado con los datos históricos de la serie.

#### *Modelo de intervención*

Un evento externo, exactamente definido en el tiempo, que afecta el comportamiento normal de una serie o de un proceso estocástico, se denomina intervención.

GRÁFICA 5  
 MODELO DE ESTACIONALIDADES Y ERRORES AUTORREGRESIVOS  
 [US\$/KW-H]



Fuente: elaboración propia.

Las intervenciones se modelan por medio de variables indicadoras de los intervalos o de los períodos, en los cuales el evento afecta la serie o a través de funciones de tiempo que explican la variación que producen.

La representación funcional de este tipo de modelos puede tomar, en cada etapa de intervención, formas funcionales similares a las estudiadas e incluso tener intercepto global o independiente, de la siguiente forma:

$$\begin{aligned}
 Z_t = & \beta_{00} I_0 + \beta_{01} P_0 + \sum_{i=0}^n \delta_{0i} I_{0\_mei} + \beta_{10} I_1 + \beta_{11} P_1 + \sum_{i=0}^n \delta_{1i} I_{1\_mei} + \beta_{20} I_2 + \beta_{21} P_2 \\
 & \dots + \sum_{i=0}^n \delta_{ji} I_{j\_mei} + \beta_{j0} I_j + \beta_{j1} P_j + U_t \quad y \quad U_t = \phi U_{t-1} + \varepsilon_t
 \end{aligned}
 \tag{3}$$

donde:

$Z_t$ : variable aleatoria

$\delta_{ji}$ : parámetros para cada estación identificada y en cada etapa de intervención

$I_{i\_mei}$ : estaciones identificadas en cada etapa de intervención

$\beta_{ij}$ : pendientes de cada etapa de intervención

$U_t$ : variable aleatoria de residuales

$\varepsilon_t$ : ruido blanco

Para la serie formada por los precios de bolsa, se tiene que las intervenciones realizadas sobre el mercado más representativas, es decir aquellas que afectan directamente la formación de precios, fueron realizadas entre mayo y septiembre de 2000. En este período el precio de bolsa presenta una tendencia creciente que puede asociarse con los incrementos del precio de oferta de los agentes generadores térmicos. Posteriormente, en septiembre, la CREG expidió la Resolución CREG 063 como una medida para reducir la volatilidad de las ofertas, lo cual se reflejó de manera transitoria en el precio de bolsa.

A partir de noviembre de 2000 se observa un alza del precio de bolsa que combina factores como los atentados que ocasionaron atrapamientos de energía hidráulica y el inicio del período de verano.

Luego de que se expidieron las resoluciones CREG 026 de 2001 y CREG 034 de 2001, el precio de bolsa disminuye de una manera significativa, coincidiendo con una reducción de los atentados a la infraestructura. Sin embargo, a final de 2001 y principio de 2002, ante una nueva oleada de atentados y una expectativa de ocurrencia del fenómeno del pacífico, se observa un repunte en el precio de bolsa, sin llegar a los niveles de precios observados previamente a la expedición de las resoluciones mencionadas.

Considerando las intervenciones descritas, se realizan dos (2) particiones en la serie de tiempo, quedando tres períodos de análisis:

1. Desde mayo de 1998 a julio de 2000, se evaluará un modelo de estacionalidades con errores autoregresivo (ver Cuadro 6).
2. Entre agosto de 2000 y mayo de 2001 (período con diferentes intervenciones del mercado) las observaciones son insuficientes para implementar un modelo. Por lo tanto, no se hará un análisis adicional para estos datos.
3. Desde junio de 2001 hasta junio de 2007 se evaluará un modelo similar al de primer período (estacionalidades con errores autoregresivo) eliminando los "outliers" identificados en el modelo de estacionalidad (ver Cuadro 9).



### Análisis del período mayo de 1998 a julio de 2000

El parámetro estimado para el rezago de orden uno se encuentra en el límite para la toma de decisiones,  $t_{\text{value}} = |1.97|$ , dado lo anterior, la comparación se realiza con el valor 1.96, considerándose el parámetro significativo.

El modelo identificado con las estacionalidades y la tendencia, representa los precios en un 99,16%, mientras que el modelo completo incluyendo la parte autoregresiva en los errores representa la serie de precios en un 99,77%.

En este caso los valores  $p$  para las respectivas pruebas indican que  $\phi_1 = 0$  y  $\phi_2 = 0$ , puesto que son mayores que 0,05. Esto nos lleva a inferir que el supuesto de errores no correlacionados se debe cumplir.

CUADRO 6  
RESULTADOS MODELO ESTACIONAL ANUAL CON ERRORES  
AUTORREGRESIVOS (MAYO 1998-JULIO 2000)

Dependent Variable		mem			
Yule-Walker Estimates					
Estimates of Autoregressive Parameters					
		Standard			
Lag	Coefficient	Error	t Value		
1	-0,493708	0,251040	-1,97		
SSE	55,8028766	DFE	12		
MSE	4,65024	Root MSE	2,15644		
SBC	139,550194	AIC	121,936842		
Regress R-Square	0,9916	Total R-Square	0,9977		
Durbin-Watson Statistics					
Order	DW	Pr < DW	Pr > DW		
1	1,8288	0,3598	0,6402		
2	1,8671	0,4637	0,5363		
3	1,9423	0,5838	0,4162		
Standard					
Approx					
Variable	DF	Estimate	Error	t Value	Pr >  t
t	1	1,0006	0,1046	9,57	<,0001
in1	1	14,8614	2,0562	7,23	<,0001
in2	1	15,1241	2,1318	7,09	<,0001
in3	1	16,4874	2,1715	7,59	<,0001
in4	1	17,4201	2,1845	7,97	<,0001
in5	1	16,1715	2,2104	7,32	<,0001
in6	1	15,2212	2,2511	6,76	<,0001
in7	1	14,1374	2,3017	6,14	<,0001
in8	1	16,3737	2,3565	6,95	<,0001
in9	1	16,6611	2,4089	6,92	<,0001
in10	1	12,7332	2,4488	5,20	0,0002
in11	1	15,9209	2,4552	6,48	<,0001
in12	1	13,1019	2,3770	5,51	0,0001

Fuente: elaboración propia.

Los parámetros estimados para las estacionalidades (in1 a in12) y la tendencia (t) presenta además el cálculo de las desviaciones estándar y la información necesaria para probar su significancia. En este caso todos los valores de  $p$  son menores que 0,05 y, por lo tanto, los parámetros son significativos.

Hasta el momento todos los parámetros estimados de este modelo (parte estacional y errores autorregresivos) son significativos. La validación sobre los supuestos de los errores se presenta en el Cuadro 7.

De los resultados anteriores se concluye que la media es igual a cero (0), el supuesto de normalidad de los errores se satisface. Dado que el valor  $p$  de la prueba de Shapiro-Wilk es  $v_{pnor} = 0,97$ , no se rechaza  $H_0$ , considerándose que los errores son normales. En este orden de ideas, se procede a realizar la verificación de la igualdad de varianzas, la probabilidad de los estadísticos de Liung Box superior a  $\alpha$ , la autocorrelación y autocorrelación parcial.

CUADRO 7  
PRUEBA DE NORMALIDAD SOBRE LA SERIE DE RESIDUALES

Modelo	Estadísticos de prueba							
	Media	Var.	Des	Mediana	Asim.	Curtosis	Wnormal	Vpnorm
Intervención 1	-0,04	2,23	1,49	-0,25	-0,12	-0,21	0,99	0,97

Fuente: elaboración propia.

Se observa que los residuales satisfacen la prueba de verificación de igualdad de varianzas  $Pr > F = 0,9739$ , al ser superior que el valor de significancia 0,05. La probabilidad de los estadísticos de Liung Box es superior (0,9405) a  $\alpha$  (0,05), con lo cual se reafirma la no correlación de los errores, además la gráfica de autocorrelación ratifica lo indicado por los demás estadísticos de prueba. Los valores de autocorrelación dados por los asteriscos (\*) se encuentran dentro del intervalo de confianza definido por puntos (·).

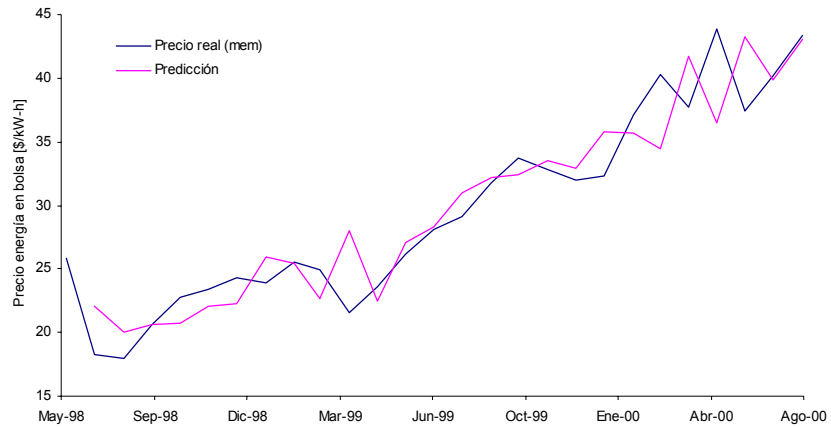
La Gráfica 6 presenta el comportamiento de la serie de datos comparada con el modelo, el cual ha cumplido con los estadísticos de prueba para la significancia de los parámetros y cuyos residuos se comportan como ruido blanco.

**CUADRO 8  
PRUEBAS SOBRE LA SERIE DE RESIDUALES**

Name of Variable = res32																								
Mean of Working Series -0.03479																								
Standard Deviation 1.462081																								
Number of Observations 26																								
Autocorrelations																								
Lag	Covariance	Correlation	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1	Std Error
0	2.145037	1.00000												*****										0
1	0.104268	0.04861												*										0.196116
2	0.040760	0.01900												.										0.196579
3	-0.182291	-0.08498												**										0.196650
4	-0.402593	-0.18769												****										0.198057
5	-0.156445	-0.07293												*										0.204784
6	0.107605	0.05016												.										0.205780
7	0.175583	0.08186												**										0.206250
Autocorrelation Check for White Noise																								
To	Chi-Square	DF	Pr > ChiSq	-----Autocorrelations-----																				
Lag 6	1.75	6	0.9412	0.049	0.019	-0.085	-0.188	-0.073	0.050															
The TTEST Procedure																								
Statistics																								
Variable	Class	N	Lower CL	Mean	Upper CL	Lower CL	Mean	Upper CL	Std Dev	Std Dev	Std Dev	Std Err												
res32	A	13	-1.03	-0.116	0.7974	1.0843	1.5121	2.496	0.4194															
res32	B	13	-0.876	0.0468	0.9694	1.0948	1.5267	2.5202	0.4234															
res32	Diff (1-2)		-1.393	-0.163	1.0669	1.1864	1.5194	2.1137	0.596															
T-Tests																								
Variable	Method	Variations	DF	t Value	Pr >  t																			
res32	Pooled	Equal	24	-0.27	0.7867																			
res32	Satterthwaite	Unequal	24	-0.27	0.7867																			
Equality of Variances																								
Variable	Method	Num DF	Den DF	F Value	Pr > F																			
res32	Folded F	12	12	1.02	0.9739																			

Fuente: elaboración propia.

**GRÁFICA 6  
MODELO DE ESTACIONALIDADES Y ERRORES AUTORREGRESIVOS  
PRIMER PERÍODO DE LA SERIE**



Fuente: elaboración propia.

## Análisis del período junio de 2001 a junio de 2007

El parámetro estimado para el rezago de orden uno es significativo  $t_{\text{value}} = |6.35|$ ; la comparación se realiza con 1,96 dado que para las pruebas de hipótesis se emplea la distribución t con 47 grados de libertad que se aproxima a una normal.

El modelo identificado con las estacionalidades representa los precios en un 91,94%, mientras que el modelo completo incluyendo la parte autorregresiva en los errores representa la serie de precios en un 98,79%.

En este caso los valores  $p$  para las respectivas pruebas indican que  $\phi_1 = 0$ ,  $\phi_2 = 0$  y  $\phi_3 = 0$ , dado que todos son mayores que 0,05. Esto nos lleva a concluir que el supuesto de errores incorrelacionados se debe cumplir.

CUADRO 9  
RESULTADOS DEL MODELO ESTACIONAL CON  
ERRORES AUTORREGRESIVOS (JUNIO 2001 A JUNIO 2007)

Dependent Variable		mem			
Yule-Walker Estimates					
Estimates of Autoregressive Parameters					
Lag	Coefficient	Standard Error	t Value		
1	-0,637035	0,100354	-6,35		
SSE	3878,04218	DFE	59		
MSE	65,72953	Root MSE	8,10737		
SBC	557,756276	AIC	525,689844		
Regress R-Square	0,9194	Total R-Square	0,9879		
Durbin-Watson Statistics					
Order	DW	Pr < DW	Pr > DW		
1	1,9837	0,4749	0,5251		
2	1,8996	0,3863	0,6137		
3	2,0184	0,6129	0,3871		
Variable	DF	Estimate	Standard Error	t Value	Approx Pr >  t
t	1	0,5072	0,1164	4,36	<,0001
in1	1	35,7084	5,9595	5,99	<,0001
in2	1	40,6609	5,9901	6,79	<,0001
in3	1	42,5660	6,0093	7,08	<,0001
in4	1	48,0028	6,0370	7,95	<,0001
in5	1	48,1267	6,0762	7,92	<,0001
in6	1	41,6431	6,1244	6,80	<,0001
in7	1	41,7458	6,1768	6,76	<,0001
in8	1	49,8392	6,2279	8,00	<,0001
in9	1	55,7977	6,2706	8,90	<,0001
in10	1	52,7164	6,2938	8,38	<,0001
in11	1	48,2902	6,2793	7,69	<,0001
in12	1	42,3116	6,1922	6,83	<,0001

Fuente: elaboración propia.

Los parámetros estimados para las estacionalidades (in1 a in12) y la tendencia (t) presenta además el cálculo de las desviaciones estándar y la información necesaria para probar su significancia. En este caso, todos los valores de  $p$  son menores que 0,05, de forma que los parámetros son significativos.

Hasta el momento todos los parámetros estimados de este modelo (parte estacional y errores autorregresivos) son significativos. La validación sobre los supuestos de los errores se presenta en los Cuadros 10 y 11.

CUADRO 10  
PRUEBA DE NORMALIDAD SOBRE LA SERIE DE RESIDUALES

Modelo	Estadísticos de prueba							
	Media	Var.	Des	Mediana	Asim.	Curtosis	Wnormal	Vpnorm
Intervención 2	-0,08	53,86	7,34	-1,18	0,16	-0,31	0,99	0,84

Fuente: elaboración propia.

CUADRO 11  
PRUEBAS SOBRE LA SERIE DE RESIDUALES

Name of Variable = res32		Mean of Working Series		Standard Deviation		Number of Observations	
			-0.07763		7.288192		73

Lag	Covariance	Correlation	Autocorrelations										Std Error										
			-1	9	8	7	6	5	4	3	2	1		0	1	2	3	4	5	6	7	8	9
0	53.117740	1.00000																					0
1	0.013850	0.00026																					0.117041
2	1.385441	0.02608										*											0.117041
3	-1.931016	-0.03635										*											0.117121
4	0.231216	0.00435										*											0.117275
5	7.143734	0.13449										***											0.117277
6	-4.974829	-0.09366										**											0.119371
7	-5.689726	-0.10712										**											0.120374
8	3.568987	0.06719										*											0.121673
9	0.488595	0.00920										*											0.122180
10	-12.625509	-0.23769										*****											0.122189
11	-0.866537	-0.01631										*											0.128367
12	-13.215500	-0.24880										*****											0.128395
13	4.470046	0.08415										**											0.134838
14	5.731550	0.10790										**											0.135555
15	-10.671986	-0.20091										*****											0.136727
16	-0.886582	-0.01669										*											0.140713
17	-2.084656	-0.03925										*											0.140740
18	2.428087	0.04571										*											0.140890

Autocorrelation Check for White Noise										
To	Chi-Square	DF	Pr > ChiSq	-----Autocorrelations-----						Pr >
Lag										
6	2.33	6	0.8869	0.000	0.026	-0.036	0.004	0.134	-0.094	
12	14.16	12	0.2907	-0.107	0.067	0.009	-0.238	-0.016	-0.249	
18	20.08	18	0.3283	0.084	0.108	-0.201	-0.017	-0.039	0.046	

The TTTEST Procedure Statistics									
Variable	Class	N	Lower CL Mean	Upper CL Mean	Lower CL Std Dev	Upper CL Std Dev	Std Dev	Std Dev	Std Err
res32	A	37	-1.884	0.1547	2.1939	4.9737	6.1161	7.9446	1.0055
res32	B	36	-3.191	-0.316	2.5585	6.8917	8.4969	11.084	1.4161
res32	Diff (1-2)		-2.977	0.4712	3.919	6.3459	7.3863	8.8379	1.7292

T-Tests					
Variable	Method	Variations	DF	t Value	Pr >  t
res32	Pooled	Equal	71	0.27	0.7860
res32	Satterthwaite	Unequal	63.5	0.27	0.7870

Equality of Variances					
Variable	Method	Num DF	Den DF	F Value	Pr > F
res32	Folded F	35	36	1.93	0.0533

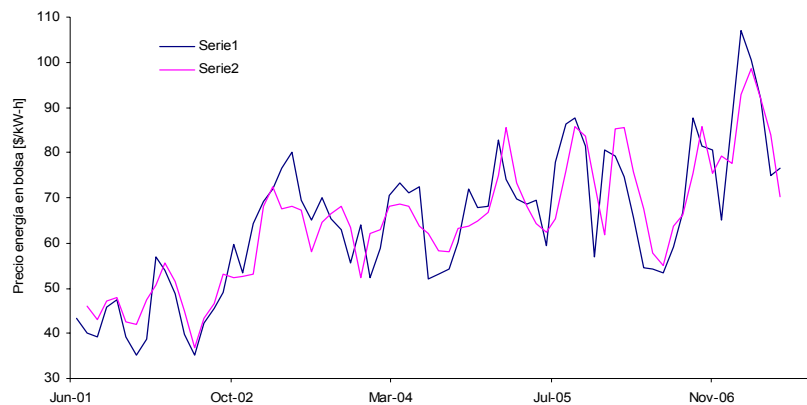
Fuente: elaboración propia.

De los resultados anteriores se concluye que la media es igual a cero (0), el supuesto de normalidad de los errores se satisface. Dado que el valor  $p$  de la prueba de Shapiro-Wilk es  $vpnor = 0,84$ , no se rechaza  $H_0$ , considerándose que los errores son normales. En este orden de ideas, se procede a realizar la verificación de la igualdad de varianza, la probabilidad de los estadísticos de Liung Box superior a  $\alpha$  y la autocorrelación.

Se observa que los residuales satisfacen la prueba de verificación de igualdad de varianza  $Pr > F = 0,0533$ , al ser superior que el valor de significancia 0,05. La probabilidad de los estadísticos de Liung Box es superior a  $\alpha$ , con lo cual se reafirma la no correlación de los errores, además la gráfica de autocorrelación ratifica lo indicado por los demás estadísticos.

La Gráfica 7 presenta el comportamiento de la serie de datos comparada con el modelo, el cual ha cumplido con los estadísticos de prueba para la significancia de los parámetros y cuyos residuos se comportan como ruido blanco.

GRÁFICA 7  
MODELO DE ESTACIONALIDADES, ERRORES AUTORREGRESIVOS  
ÚLTIMO PERÍODO DE LA SERIE



Fuente: elaboración propia.

Ahora lo que resta es medir que tan apto es el modelo para pronosticar valores futuros de los precios de la energía en bolsa, empleando para

ello algunos estadísticos básicos y otros construidos a partir del error relativo de pronóstico.

Dentro de los estadísticos básicos encontramos:

- Error promedio de pronóstico:

$$ME = \frac{1}{m} \sum_{i=0}^{m-1} e_{n+i}(1) = \frac{1}{m} \sum_{i=0}^{m-1} (z_{n+i+1} - \hat{z}_{n+i}(1))$$

- Error promedio absoluto de pronóstico:  $MAE = \frac{1}{m} \sum_{i=0}^{m-1} |e_{n+i}(1)|$

- Suma de cuadrados de los errores de pronóstico:

$$SSE = \sum_{i=0}^{m-1} e_{n+i}^2(1)$$

- Error cuadrático medio del pronóstico:  $MSE = \frac{1}{m} \sum_{i=0}^{m-1} e_{n+i}^2(1)$

Dentro de los estadísticos basados en el error relativo de pronóstico se encuentran:

- Porcentaje de error al estimar:

$$Z_{n+i} = 100 \left[ \frac{z_{n+i} - \hat{z}_{n+i-1}(1)}{z_{n+i}} \right] = 100 \left[ \frac{e_{n+i-1}(1)}{z_{n+i}} \right]$$

- Porcentaje medio de error:

$$MPE = \frac{100}{m} \sum_{i=0}^{m-1} \frac{e_{n+i}(1)}{z_{n+i+1}} = \frac{100}{m} \sum_{i=0}^{m-1} \frac{z_{n+i+1} - \hat{z}_{n+i}(1)}{z_{n+i+1}}$$

- Porcentaje medio absoluto del error:

$$MAPE = \frac{100}{m} \sum_{i=0}^{m-1} \left| \frac{z_{n+i+1} - \hat{z}_{n+i}(1)}{z_{n+i+1}} \right|$$

El  $MPE$  y el  $MAPE$  son estadísticos intuitivos y fáciles de interpretar puesto que representan el porcentaje promedio del error o el porcentaje

absoluto promedio del error de los pronósticos y suministran una información más concreta sobre la precisión del pronóstico que el MSE.

Para hallar el incremento en la precisión de los pronósticos, al aplicar un método, debe existir un modelo que sirva de referencia para establecer las comparaciones. Makridakis-Wheelwright-Mcgee proponen el siguiente método:

- Método de pronóstico simple o ingenuo (SIMP). Éste considera que el pronóstico estimado en  $t=n$  para la variable  $Z_{n+1}$  es el valor observado en  $t=n$ , es decir, que pronostica la observación futura con la observación presente. Por tanto:

$$\hat{z}_{n+1} = \hat{z}_n(1) = z_n \quad y \quad e_n(1) = z_{n+1} - z_n$$

Con lo anterior, el *MAPE* del método de referencia *SIMP*, para las últimas  $m$  observaciones de una serie de tiempo, será:

$$MSIMP = \frac{100}{m} \sum_{i=0}^{m-1} \left| \frac{z_{n+i+1} - z_{n+i}}{z_{n+i+1}} \right|$$

### **Criterio de comparación con el modelo de referencia simple**

Si el *MAPE* de un modelo bajo estudio es menor que el *MSIMP*, significa que este método debe implementarse porque mejora la calidad del pronóstico, luego se justifica implementar el método.

Si el *MAPE* de un modelo bajo estudio es mayor o igual al *MSIMP* significa que este método no debe implementarse porque no se mejora la calidad del pronóstico, pero sí requiere más trabajo.

En concordancia con lo anterior, el método de referencia sólo sería aplicable en casos donde el horizonte de pronóstico es uno o cuando la serie que se modela tiene tendencia constante. Situación poco común en aplicaciones reales.

Las medidas relativas estudiadas dan igual peso a todos los errores en contraste con el *MSE* que al elevar al cuadrado los errores pondera los más altos. Podría ser útil tener una medida que considere estos aspectos:



pondere los errores más altos, tenga en cuenta los efectos de un cambio de escala y permita comparaciones con los sistemas de pronósticos simples. Una medida con estas características es el *Theil*.

$$Theil = \sqrt{\frac{\sum_{i=0}^{m-1} \left[ \frac{\hat{z}_{n+i+1} - z_{n+i+1} (I)}{z_{n+i+1}} \right]^2}{\sum_{i=0}^{m-1} \left[ \frac{z_{n+i} - z_{n+i}}{z_{n+i+1}} \right]^2}}$$

El *Theil* es igual 0, si el numerador es cero; esto implica que el método de pronósticos empleado es exacto. El *Theil* es igual a 1, si el numerador y el denominador son iguales, esto indica que la precisión del método de pronóstico bajo estudio es igual a la del método; en conclusión no se justifica implementar el método bajo estudio. Si el *Theil* es mayor que 1, el sistema de pronósticos del método bajo estudio tiene un error relativo de pronóstico mayor que le del método simple, en consecuencia no se justifica implementar el método. Si  $0 < Theil < 1$  el método de pronóstico bajo estudio es mejor que el método simple.

Los resultados luego de implementar la metodología anterior para el modelo se presenta en el Cuadro 12.

CUADRO 12  
RESULTADOS ESTADÍSTICOS PARA MEDIR EL PODER DE PREDICCIÓN

Mês de prueba	mem	predicción
Julio 2007	79,81	80,52907
Agosto 2007	84,11	81,63564
The MEANS Procedure		
Variable		Mean
ffffffffffffffffffffffffffff		
me		0,878
mae		1,597
mpe		1,020
mape		1,921
mse		3,320
msimp		5,112
ffffffffffffffffffffffffffff		
theil		0,58

Fuente: elaboración propia.

El estadístico  $MAPE = 1,92\%$  es inferior al  $MSIMP = 5,11\%$ , lo que indica que el modelo disminuye el porcentaje de error relativo en la predicción en un  $3,19\%$  con respecto al método simple, es decir, que se justifica implementar el modelo de predicción. Esto es ratificado por el estadístico Theil que presenta un valor inferior a 1 para la prueba realizada con los últimos dos valores del precio de bolsa que NO fueron empleados para la identificación de los parámetros del modelo (julio y agosto de 2007).

En conclusión se he identificado un modelo que cumple con los estadísticos de prueba en la significancia de los parámetros, con las condiciones de ruido blanco sobre los residuales y que presenta una buena capacidad de predicción.

### CONCLUSIONES

Con los datos disponibles hasta la fecha se presenta una metodología para la estimación de los precios en bolsa de la energía, como una herramienta adicional a las existentes en la literatura revisada para el caso colombiano y que cumple con las validaciones estadísticas apropiadas.

Los resultados obtenidos a lo largo del artículo confirman lo observado en el análisis gráfico inicial, en el cual fueron identificados algunos períodos de intervención por parte del regulador. Sólo hasta el momento en que fueron incorporadas dichas intervenciones a las técnicas de estimación, se logró obtener un modelo que cumpliera todos los estadísticos de prueba tanto en la significancia de los parámetros como en el análisis de residuales.

Este hallazgo tiene implicaciones fuertes en la construcción de modelos para estimaciones en el largo plazo, dado que deben incorporar los estudios que el regulador tenga bajo desarrollo y que posiblemente afectarán la formación de los precios en la Bolsa de Energía de Colombia, tal es el caso del Cargo por Confiabilidad (CxC), el Sistema Electrónico de contratos (SEC), el Mercado Organizado Regulado (MOR), las Transacciones Internacionales de Energía (TIES), entre otros.

Los modelos con tendencia, estacionalidad anual y errores autorregresivos, presentan grandes ventajas en su forma de estimación,

debido al alcance que puede obtenerse con la estacionalidad anual y la tendencia identificada sobre la serie de tiempo.

Los modelos identificados a lo largo del artículo presentan un horizonte de estimación que, generalmente, depende de la estacionalidad y el orden de la parte autoregresiva, lo que obliga a realimentar de forma permanente el modelo, en aras de obtener una buena estimación. En concordancia con lo anterior, la predicción de los Precios de la Energía en Bolsa en el largo plazo (años) requiere modelos más elaborados que incluyan variables como la operación del sistema de transmisión nacional, la estructura del mercado, los mecanismos de contratación, la simulación del despacho económico, como es el caso de los modelos de Análisis de Equilibrio (Guang 2005).

En otras palabras se requiere construir un modelo que simule de forma más precisa el proceso completo de formación de los precios de bolsa, para luego realizar sensibilidad (determinación estadística de la significancia de los parámetros) de las variables más representativas y llegar así a un modelo más parsimonioso, que permita a los agentes del sistema emplear herramientas de predicción rápidas y fáciles de implementar que ayuden a la toma de decisiones frente a las transacciones de energía.

Dado que la metodología desarrollada se concentró en la bolsa de energía de Colombia, si se pretendiera hacer un análisis similar para otro mercado o energético se debe realizar un proceso completamente nuevo, siguiendo los mismos lineamientos usados en el presente trabajo.

## REFERENCIAS BIBLIOGRÁFICAS

- Cano, Jovan (2004). Tesis - Estimación de los Precios de la Energía en la Bolsa de Colombia. Universidad Nacional de Colombia, Facultad de Ciencias Humanas, Medellín. p. 1-110. Para optar al título de Magíster en Ciencias Económicas.
- Compañía de Expertos en Mercados XM. Boletín del Sector Eléctrico Colombiano, números 1 al 178, Medellín.
- Conejo, Antonio; Plazas, Miguel; Espinola, Rosario y Molina Ana (2005). "Day-Ahead Electricity Price Forecasting Using the Wavelet Transform and ARIMA Models". *IEEE transactions on power systems*, 20(2): 1035-1042.

- Contreras, Javier; Espinola, Rosario; Nogales, Francisco y Conejo, Antonio (2003). "ARIMA Models to Predict Next-Day Electricity Prices". *IEEE transactions on power systems*, 18(3): 1014 - 1020.
- CREG (Comisión de Regulación de Energía y Gas) (1994). Resolución 055-1994. Santafé de Bogotá.
- CREG (Comisión de Regulación de Energía y Gas) (2000). Resolución 063-2000. Bogotá.
- CREG (Comisión de Regulación de Energía y Gas) (2001a). Resolución 026-2001. Bogotá.
- CREG (Comisión de Regulación de Energía y Gas) (2001b). Resolución 034-2001. Bogotá.
- Correa, Emilia (2000). *Series de Tiempo*. Medellín: Centro de publicaciones Universidad Nacional de Colombia.
- Diebold, Francis (2001). *Elements of forecasting*, segunda edición. Australia: South Western College Publishing.
- Guang Li, Chen-Ching Liu; Lawarree, Jacques; Gallanti, Massimo y Venturini Andrea (2005). State-of-Art of Electricity Prices Forecasting. CIGRE/IEEE PES, International Symposium.
- Jaramillo, Andrés (2000) Tesis - La Bolsa de Energía a la Luz de la Teoría de Juegos. Universidad Nacional de Colombia, Facultad de Ciencias Humanas. Medellín. p. 1 - 60. Para optar al título de Magíster en Ciencias Económicas.
- Kirschen, Daniel (2006). "Fundamentals of Power System Economics", The University of Manchester, U.K.
- Millan, Jaime y Von Der Fehr, Nils-Henrik (2003). *Keeping the Lights On: Power Sector Reform in Latin America*. Washington. D.C.: Inter-American Development Bank.
- OLADE-CEPAL-GTZ (2003). *Energía y Desarrollo Sustentable en América Latina y el Caribe: Guía para la Formulación de Políticas Energéticas*. Santiago de Chile: CEPAL.
- Raigoza, Diana (1999). Tesis - Toma de Decisiones para la Compra de Energía Eléctrica. Universidad Nacional de Colombia, Facultad de Minas, Medellín. p. 17 - 81. Para optar al título de Ingeniera Electricista.
- Rudnick, Hugh y Zolezzi, Juan Manuel (1999). "Expansión de los sistemas de transmisión en mercados competitivos de generación", Actas I Reunión Internacional G&T CIER 99, Optimización de los Procesos de Generación y Transmisión en Ambientes Competitivos, Bogotá, Colombia, Noviembre 3-5, 1999.
- UPME (Unidad de Planeación Minero Energética). *Plan de Expansión de Referencia 2002 - 2011*. Bogotá: Ministerio de Minas y Energía. 18 - 77.
- Weron, Rafal y Misiorek, Adam (2005). *Forecasting Spot Electricity Prices With Time Series Models. International Conference*. "The European Electricity Market EEM-05", mayo 10-12, Lodz (Polonia): 133-141.