

ANÁLISIS DE UN PROCESO DE DIFUSIÓN POR SALTOS CON REVERSIÓN A LA MEDIA APLICADO AL MERCADO COLOMBIANO DE ELECTRICIDAD

Trabajo final de grado presentado por Mauricio Aguiar Gil como requisito para optar al título de Maestría en Ingeniería Administrativa

Director:

Sergio Botero Botero, D.Sc., M.Sc. I.M.

Profesor asociado a la Escuela de Ingeniería de la Organización

Facultad de Minas, Universidad Nacional de Colombia, sede Medellín

INTRODUCCIÓN.....	2
MARCO TEÓRICO.....	4
CONVERSIÓN DE LA SERIE DE PRECIOS DE PESOS CORRIENTES A PESOS CONSTANTES.....	7
IDENTIFICACIÓN DEL CARÁCTER DE LA SERIE ¿ADITIVA O MULTIPLICATIVA?.....	7
TRATAMIENTO PARA LA IDENTIFICACIÓN DE LA COMPONENTE DE TENDENCIA EN LA SERIE.....	10
CARACTERIZACIÓN DE LAS COMPONENTES ESTACIONALES DE LARGO Y CORTO PLAZO.....	12
IDENTIFICACIÓN Y REMOCIÓN DE PICOS EN LA SERIE PRODUCIDOS POR SALTOS, REPETICIÓN DEL CÁLCULO DE LOS PARÁMETROS ESTACIONALES	16
IDENTIFICACIÓN DE LA EXISTENCIA DE REVERSIÓN A LA MEDIA Y CARACTERIZACIÓN.....	19
CARACTERIZACIÓN DE LA VOLATILIDAD.....	24
CARACTERIZACIÓN DE LOS PARÁMETROS DEL PROCESO DE SALTOS.....	29
POSIBILIDAD DE APLICACIONES.....	34
CONCLUSIONES.....	34
BIBLIOGRAFÍA	35

INTRODUCCIÓN

Dado el contexto actual en el que está inmersa la República de Colombia, clasificada como un “mercado de frontera” por unos [1] y como “mercado emergente” por otros [2], se hace necesario que el sector público, el sector privado, la academia y la sociedad redoblen esfuerzos para construir dinamismo, estabilidad y prosperidad nacional. Y es evidente que un pilar del progreso (que contempla el dinamismo, la estabilidad y la prosperidad nacional) es la generación, comercialización y transmisión de la energía eléctrica [3]. Las potencialidades de Colombia como generador y exportador a nivel regional son amplias y requieren ser analizadas, pero este no es el objetivo del presente trabajo.

El mercado local de electricidad ha experimentado cambios motivados en momentos políticos que ha vivido el país y en momentos de dificultades energéticas tales como el racionamiento de energía sucedido a principios de la década de 1990. La liberalización (parcial) del mercado de electricidad colombiano se emprendió trazando reglas tales que impidieran el desabastecimiento de energía causado por múltiples factores (el clima, problemas de orden público, fallas en el sistema de transmisión), que permitan la formación transparente de precios y la libre competencia. Dicha liberalización contribuyó a que los precios reflejaran de manera más fidedigna los costos de producción de la energía eléctrica [4]

Estos costos están ligados a varias variables: La disponibilidad de recurso hídrico en los embalses, los precios del gas natural y del carbón, el clima, los ciclos económicos, la situación de orden público y la operatividad del sistema de generación y transmisión; pero es importante mencionar que la liberalización de los mercados conlleva agregar a la formación de los precios un elemento adicional: La puja que produce la oferta y la demanda, implicando un aumento de la incertidumbre o la *volatilidad* en los precios [4], [5] porque a pesar de que la formación de los precios en bolsa de la electricidad en Colombia viene ligada de manera indirecta a un sistema de subasta basado en un derivado financiero de cobertura conocido como *cargo por confiabilidad* [6], las variables mencionadas al inicio de este párrafo no permiten la eliminación total del riesgo de mercado, ligado a la *volatilidad*.

Por citar un ejemplo, en Colombia la disponibilidad de energía eléctrica depende fuertemente del estado del tiempo nacional ya que alrededor del 70% del parque energético está conformado por hidroeléctricas [6], [7] y de manera particular, el fenómeno del niño impacta la producción en las centrales hidroeléctricas reduciendo la oferta de electricidad se reduce, creando alteraciones inesperadas en la formación del precio en bolsa. Las consecuencias en el consumo residencial, industrial, comercial y la economía en general son negativas.

En otros mercados de electricidad del mundo se han identificado algunos hechos estilizados que podrían considerarse comunes a los mercados liberalizados. Uno de ellos es la naturaleza aleatoria de los precios, que no es más que producto de la interacción entre la oferta y la demanda que, desde el punto de vista de los procesos estocásticos se conoce como *proceso de difusión*. Otro hecho estilizado es la formación repentina de precios atípicos o proceso de *saltos* de los precios previamente formados a causa de problemas de elasticidad en la

demanda y además está la oscilación alrededor de un valor base o proceso de *reversión a la media* en el que las mismas fuerzas del mercado o las reglas de regulación en este retornan los precios a un nivel comparativamente aceptable (media). El objetivo central de este trabajo es identificar y analizar la presencia de estos hechos estilizados en una serie histórica de precios en bolsa para la electricidad en Colombia.

Se tomó una serie de precios diarios partiendo de enero 1 de 1997 hasta febrero 6 de 2013, lo que comprende un total de 5881 observaciones. No se hizo distinción previa de días feriados o fines de semana.

MARCO TEÓRICO

Los mercados de energía eléctrica se caracterizan por contar con un regulador que interviene en ellos para garantizar que el precio de la energía llevado al consumidor no sea manipulado por los generadores ni que sea elevado en exceso así como para evitar la aparición de oligopolios que distorsionen la dinámica de precios. Por esto y de manera general los precios de la energía eléctrica se revierten a un régimen dado o media [8], [9]. Las desviaciones de dicha media son ocasionadas por diversas razones: Ruptura del sistema eléctrico, aumentos en la demanda, oferta inelástica, cambios en precios de las materias primas necesarias para la generación entre otras [10].

Estas desviaciones se conocen como “picos” y para su modelado se ha propuesto diversos enfoques en el que de manera común se considera la serie temporal de precios como un proceso estocástico; la diferencia entre un enfoque y otro es en el tipo de proceso estocástico que se propone. Algunos autores [9] proponen estudiar la serie como un proceso de Markov, en el que el precio en un tiempo t no está ligado al precio en un tiempo $t - \Delta t$ o el precio en un tiempo $t + \Delta t$, pero los parámetros de dicho proceso cambian en el tiempo para retornar a un estado original. Este hecho podría explicar el fenómeno de reversión a la media.

Otros autores [11], [7], [10] proponen estudiar la serie como un proceso de difusión por saltos con reversión a la media. Estos procesos tienen una componente de tendencia o *drift* el cuál puede definirse en el mediano plazo como el régimen de precios, una componente aleatoria a largo plazo o volatilidad, una componente aleatoria que asigna la probabilidad de un salto y una componente a largo plazo que en el caso de un salto, hace que el precio retorne a la media (régimen) [11]. Los cambios de nivel o saltos son importantes para la gestión del riesgo.

En finanzas, la forma más básica de un proceso estocástico para modelado puede esquematizarse con la relación:

$$dP = \mu P dt + \sigma P dz \quad (1)$$

En donde se indica que un cambio en los precios de un activo (dP) puede modelarse con la combinación de dos componentes: Una componente de tendencia ($\mu P dt$) donde (dt) representa cambios en el tiempo y una componente aleatoria ($\sigma P dz$) cuya componente (dz) o “Proceso de Wiener” representa *shocks* aleatorios y basados en un movimiento browniano geométrico (ruido aleatorio basado en una distribución lognormal); esto último para no permitir la existencia de precios negativos [11]; sin embargo, este proceso no da cuenta de ciertos hechos que se presentan en los mercados energéticos:

- Los retornos de los precios de la energía, definidos como $\ln\left(\frac{P_t}{P_{t-1}}\right)$ no se distribuyen de manera normal [7], [9].
- Los precios de la energía presentan desviaciones bruscas de la componente de tendencia que se dan en un lapso corto de tiempo (saltos). El proceso de Wiener no da cuenta de esos saltos [11].

- Debido a las políticas regulatorias y a otras características de los mercados de energía eléctrica, después de un salto hay un retorno al régimen de precios. Este hecho tampoco está tenido en cuenta en el proceso de Wiener [9], [11].

Para tener en cuenta estos hechos, se ha propuesto otros procesos estocásticos para modelar el comportamiento de las series temporales de precios. Por ejemplo:

$$dP_t = \mu dt + \sigma_t P_t dW_t + Jdq \quad (2)$$

Indica que las variaciones en el precio del activo P_t se pueden modelar combinando tres componentes: μdt como componente de *drift*, $\sigma_t dW_t$ como componente aleatoria (volatilidad y proceso de Wiener) y Jdq es la componente de saltos. En la literatura se menciona frecuentemente que esta componente Jdq se modela con un proceso de Poisson compuesto con parámetro λ y comportamiento lognormal de J [12], [13]. Tener en cuenta el proceso de reversión a la media implica usar un proceso del tipo [14]:

$$dP_t = \alpha(\mu - P_t)dt + \sigma_t dW_t + Jdq \quad (3)$$

Donde α representa la velocidad con la que los precios retornan a su régimen después de un salto (velocidad de reversión a la media), μ es ahora el nivel de precios en el largo plazo y los demás términos siguen siendo similares a los anteriormente mencionados [11]. Calibrar un proceso de este tipo para el modelado consiste en definir los parámetros: Velocidad de reversión a la media α , nivel de precios μ , volatilidad σ_t y parámetro para el proceso de saltos λ así como un carácter lognormal para la magnitud de los saltos J .

Desde el punto de vista del análisis de series temporales, una serie de precios puede definirse como un esquema de componentes que se suman, llamado esquema aditivo:

$$P_t = T_t + S_t + s_t + C_t + K_t \quad (4)$$

O como un esquema de componentes que se multiplican llamado esquema multiplicativo:

$$P_t = T_t S_t s_t C_t K_t \quad (5)$$

Dónde:

P_t = Precio en el tiempo t

S_t = Componente que refleja el comportamiento estacional a largo plazo

s_t = Componente que refleja el comportamiento estacional a corto plazo

C_t = Componente que refleja el comportamiento cíclico a lo largo de t

K_t = Componente con el comportamiento aleatorio, de picos y reversión de la serie temporal

En línea con el planteamiento general de este trabajo, cabe suponer que en la componente Z_t (aleatoria) puede encontrarse un término de difusión y un término de saltos mientras que en

el resto de las componentes de la serie temporal se encuentra enmascarado el término de reversión a la media. Se abordará ambos enfoques: Procesos estocásticos y series temporales, con el objetivo de separar y modelar por aparte las componentes determinísticas para lograr un análisis más preciso de los términos en la componente aleatoria: Difusión y saltos.

CONVERSIÓN DE LA SERIE DE PRECIOS DE PESOS CORRIENTES A PESOS CONSTANTES

Para entender mejor la dinámica de precios, lo ideal es hacer que todos los datos de precios estén en valores equivalentes a pesos de último periodo temporal de la serie (febrero de 2013). Para ello se utiliza la base de datos en el portal del Banco de la República para obtener los datos mensuales del índice de precios del productor (IPP). La operación necesaria para llevar un valor P_t^α a precio corriente del año α hasta un valor P_t^β a precio constante con base en el año β es la siguiente:

$$IPP_A^\beta = 100 \left(\frac{IPP_t^\alpha}{IPP_A^\alpha} \right) \quad (6)$$

Dónde:

IPP_A^β = Índice de precios con base en el periodo β para el año A

IPP_A^α = Índice de precios con base en el periodo α para el año A

IPP_t^α = Índice de precios con base en el periodo α para un año t

Dado que el conjunto de datos de IPP que proporciona el portal del Banco de la República contiene valores con base en el periodo “junio de 1999” se requiere reformular cada valor basado en el periodo “febrero de 2013” usando (6) Con esto se puede lograr un nuevo conjunto de datos de IPP necesario para llevar los precios corrientes a precios constantes a través de la relación:

$$P_t^{\text{Constante}} = \frac{IPP_t^\beta \times P_t^{\text{Corriente}}}{100} \quad (7)$$

En adelante, se llevará a cabo el trabajo con precios constantes que se notarán como P_t en lugar de $P_t^{\text{Constante}}$

IDENTIFICACIÓN DEL CARÁCTER DE LA SERIE ¿ADITIVA O MULTIPLICATIVA?

Para continuar con el tratamiento a la serie temporal de precios en bolsa de la electricidad en Colombia, es necesario identificar su naturaleza. Se parte del hecho de que una serie de tiempo puede expresarse como un esquema aditivo o un esquema multiplicativo. Para estudiar la dinámica de la serie temporal se hace necesario identificar el esquema que se le ajusta. Existen tres métodos para ello:

- Inspección visual de la serie temporal. Una serie con niveles (límites) constantes se considera un esquema aditivo mientras que una serie con niveles (límites) crecientes o decrecientes se considera un esquema multiplicativo [15], [16], [17].

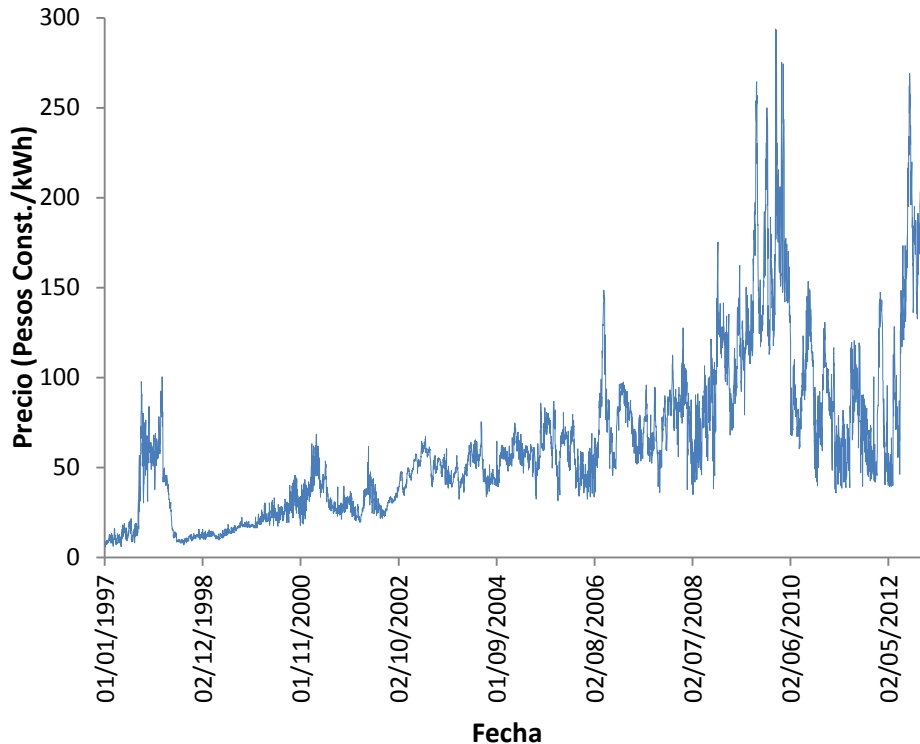


Figura 1 Serie temporal P_t de precios constantes en bolsa de la energía eléctrica en Colombia. Fuente: Creación propia basada en datos alojados en xm.com.co

De la Figura 1 se observa que la serie de datos presenta cambios de nivel pronunciados a medida que transcurre el tiempo y de acuerdo a esto se podría concluir que la serie de tiempo es multiplicativa. No obstante se emprenderá el uso del segundo método:

- Trazado de un gráfico de dispersión. Consistente en trazar un gráfico que plasme los valores ordenados de los promedios de precios para un mes y un año dado. Cada punto en el gráfico es el promedio de cada mes para cada año contra su respectiva desviación estándar. Si este gráfico no presenta una tendencia definida, se considera una serie con esquema aditivo; si presenta una tendencia definida se considera una serie con esquema multiplicativo.

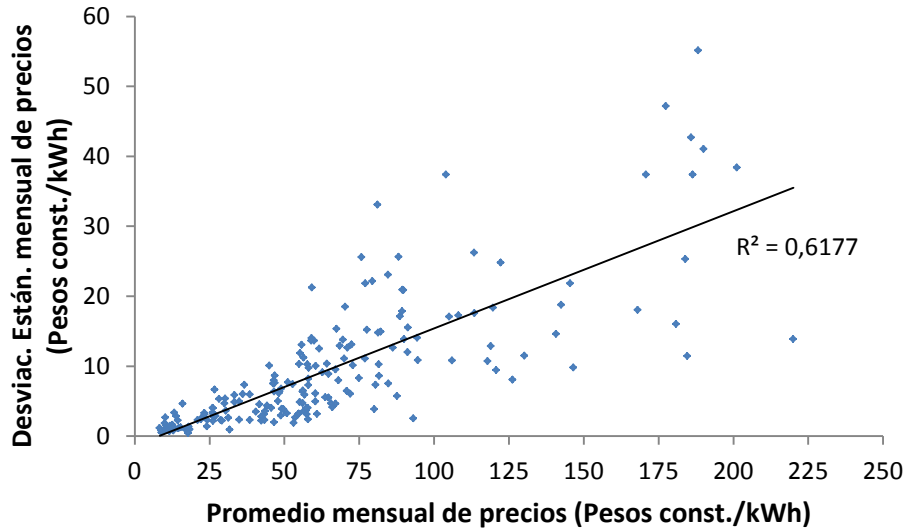


Figura 2 Gráfico de dispersión para la serie temporal de precios. Fuente: Creación propia basada en datos alojados en xm.com.co

La tendencia observada en la Figura 2 no es concluyente, por ello se ensaya otro enfoque:

- Método de las diferencias y cocientes estacionales [18]. Se calculan las diferencias estacionales

$$d_t = P_t - P_{t-\tau} \quad (8)$$

Y los cocientes estacionales

$$c_t = \frac{P_t}{P_{t-\tau}} \quad (9)$$

Donde τ es una estación de tiempo que para el caso se escogió igual a 7 y 365 días (semanal y anual) Después se calcula el promedio de todas las diferencias estacionales \bar{d}_t y su respectiva desviación estándar σ_{d_t} así como el promedio y desviación estándar de todos los cocientes estacionales (\bar{c}_t y σ_{c_t}). En este proceso se pierde siempre una cantidad τ de datos. Se calculan los coeficientes de variación definidos como:

$$C_{vd} = \frac{\sigma_{d_t}}{\bar{d}_t} \quad (10)$$

$$C_{vc} = \frac{\sigma_{c_t}}{\bar{c}_t} \quad (11)$$

Y de acuerdo al resultado obtenido se concluirá que:

Si $C_{vd} > C_{vc}$ se asume un esquema multiplicativo

Si $C_{vd} < C_{vc}$ se asume un esquema aditivo

En las Tablas se muestra el resultado de calcular y analizar los datos en la serie temporal de precios de la electricidad en Colombia. Basado en los resultados obtenidos se asume que esta serie sigue un esquema multiplicativo.

Tabla 1 Resultados del análisis de las diferencias y cocientes estacionales usando promedios móviles a 7 días

ITEM	DIFERENCIAS ESTACIONALES	COCIENTES ESTACIONALES
Promedio a 7 días	0,251	1,023
Desviación estándar	$\sigma_{dt} = 15,616$	$\sigma_{ct} = 0,205$
Coefficiente de variación (C_v)	$C_{vd} = 62,217$	$C_{vc} = 0,200$

Tabla 2 Resultados del análisis de las diferencias y cocientes estacionales usando promedios móviles a 365 días

ITEM	DIFERENCIAS ESTACIONALES	COCIENTES ESTACIONALES
Promedio a 365 días	6,860	1,341
Desviación estándar	$\sigma_{dt} = 41,865$	$\sigma_{ct} = 0,931$
Coefficiente de variación	$C_{vd} = 6,103$	$C_{vc} = 0,694$

TRATAMIENTO PARA LA IDENTIFICACIÓN DE LA COMPONENTE DE TENDENCIA EN LA SERIE

Como se concluyó, la serie de datos es una serie de esquema multiplicativo. Para abordar en el esquema de trabajo propuesto por [19] se procede a convertir la serie original de precios en otra que siga un esquema aditivo calculando logaritmo natural para cada dato de la serie original. Básicamente esto implica tener lo siguiente:

$$P'_t = \ln P_t \quad (12)$$

Para la cual se tiene que definir componentes estacionales (de largo y corto plazo) y aleatoria, ya que es evidente que serán diferentes para P'_t que para P_t . Adicionalmente, usar la serie de logaritmos en lugar de la serie original permite suavizar los datos. Dicha serie se muestra en la Figura 4.

No se debe perder de vista que la hipótesis sobre la que se fundamenta este trabajo contempla la existencia de saltos de precios (picos) que deben ser estudiados y que enrarecen la naturaleza global de la serie temporal para la cuál es recomendable entonces removerlos con el fin de desentrañar mejor las componentes. Una vez removidos, se espera determinar con mayor exactitud de la nueva serie temporal las componentes propias del proceso estocástico que sigue (componentes estacionales, componente de reversión a la media, componente de volatilidad)

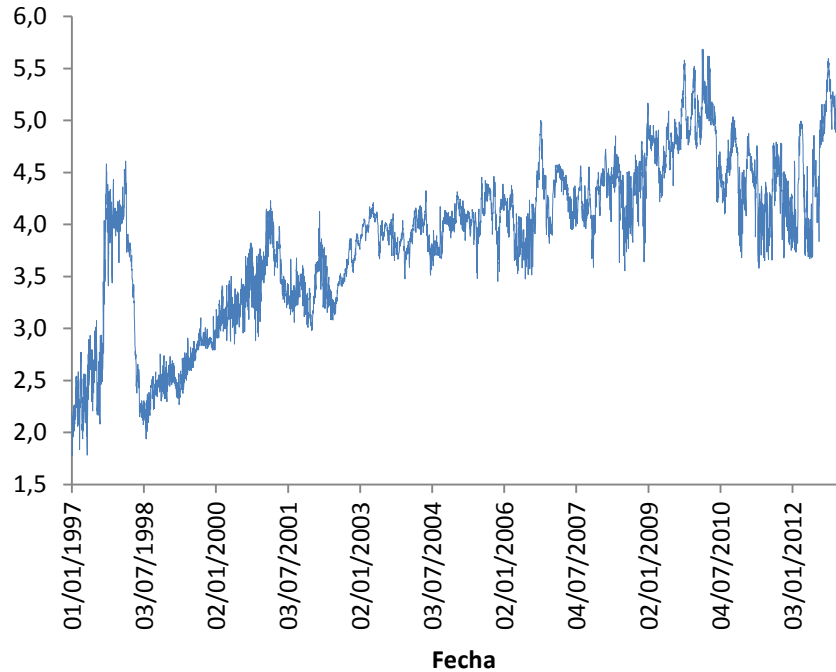


Figura 3 Serie de tiempo de $P'_t = \ln P_t$

Si se desea remover los picos de precios generados por el proceso de saltos, se debe en primera medida identificarlos. El proceso de identificación de picos exige que sea sustraída toda componente estacional y toda componente de tendencia. La determinación de la componente de tendencia es directa y se puede realizar usando Microsoft Excel® de la siguiente manera: Se ajusta una línea de tendencia a la serie de logaritmos de precios y se utiliza la ecuación de ajuste como modelo para la componente de tendencia; luego, esta componente se resta de la serie de logaritmos.

Se ajustó una tendencia polinómica de orden dos para la cual Microsoft Excel® entrega la siguiente ecuación:

$$T_t = (-4 \times 10^{-8})t^2 + (6 \times 10^{-4})t + 2,5763 \quad (13)$$

Con $t =$ tiempo. Se filtra la componente de tendencia restando T_t de la serie de datos P'_t y se procede a filtrar las componentes estacionales de largo y corto plazo. El resultado de este filtrado se observa en la Figura 5.

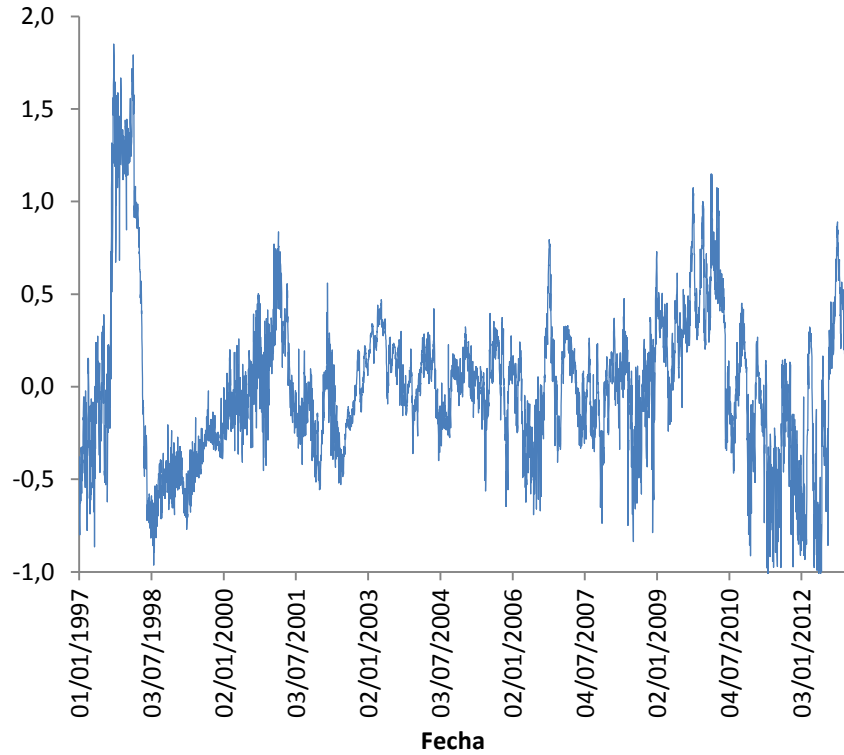


Figura 4 Serie de tiempo de $P'_t - T_t$. Acá se removió la componente de tendencia

CARACTERIZACIÓN DE LAS COMPONENTES ESTACIONALES DE LARGO Y CORTO PLAZO

Para la determinación de la componente estacional de largo plazo con picos sobre el logaritmo de la serie se siguió este procedimiento: Apoyados en [19], se propone la siguiente relación para la componente estacional a largo plazo:

$$S_t = a_1 \operatorname{sen} \left[2\pi \left(\frac{t}{365} + a_2 \right) \right] + a_3 + a_4 \operatorname{EWMA}_t^{0,975} \quad (14)$$

Donde

$$\operatorname{EWMA}_t^{0,975} = (1 - 0,975) \ln P_t + 0,975 \operatorname{EWMA}_{t-1}^{0,975} \quad (15)$$

Con $\operatorname{EWMA}_0^{0,975} = \ln P_0$. Es necesario aclarar que $\operatorname{EWMA}_t^{0,975}$ es una media móvil ponderada exponencial con un factor de decaimiento de 0,975. Para la estimación de los coeficientes a_1, \dots, a_4 se lleva a cabo un ajuste por mínimos cuadrados no lineales sobre la serie de datos

usando el software R. También es necesario indicar valores iniciales para dichos coeficientes $(a_{1i}, a_{2i}, a_{3i}, a_{4i})$ y para ello se partirá de estos supuestos:

- La componente de tendencia T_t está combinada con la parte estacional. En $t = 0$ estas dos componentes no tienen efecto alguno sobre el precio, por lo que el precio inicial va a estar definido como:

$$S_0 = a_{1i} \operatorname{sen} \left[2\pi \left(\frac{0}{365} + a_{2i} \right) \right] + a_{3i} + a_{4i} \operatorname{EWMA}_0^{0,975} \quad (16)$$

- Por definición, $\operatorname{EWMA}_0^{0,975}$ es distinto de cero en $t = 0$, pero al tratarse de una media móvil ponderada basada en observaciones anteriores, en $t = 0$ ella no aporta a la componente estacional, y para ello se definirá $a_{4i} = 0$
- Se supondrá que en $t = 0$ las expectativas del precio por parte del mercado fueron altas, por eso se maximizará la ecuación (14) y esto se da cuando $\operatorname{sen}(2\pi a_{2i})$ sea máxima es decir, cuando sea igual a 1. Por tanto se tendrá que:

$$2\pi a_{2i} = (2n + 1) \frac{\pi}{2}, n = 0, 1, 2, \dots$$

$$a_{2i} = \frac{(2n + 1)}{4} \quad (17)$$

- Como la función seno es armónica, se puede usar cualquier valor de n . Se usa $n = 0$, por lo que $a_{2i} = \frac{1}{4}$.

Para determinar los valores iniciales a_{1i} y a_{3i} se requiere contar con dos ecuaciones. Los supuestos se extienden a $t = 1$ sin incurrir en suposiciones excesivas. Partiendo de la ecuación (14) se utiliza MatLab® que resuelva el sistema de ecuaciones:

$$a_{1i} \operatorname{sen} \left[2\pi \left(\frac{1}{4} \right) \right] + a_{3i} \cong -0,38 \quad (18) \text{ para } t = 0$$

$$a_{1i} \operatorname{sen} \left[2\pi \left(\frac{1}{365} + \frac{1}{4} \right) \right] + a_{3i} \cong -0,38 \quad (19) \text{ para } t = 1$$

Teniendo en cuenta los siguientes hechos:

- La ecuación (18) corresponde a $t = 0$, mientras que la ecuación (19) corresponde a $t = 1$. Los valores iniciales a_{1i} y a_{3i} son calculados por MatLab® y se obtiene que:

$$a_{1i} \cong -7,26$$

$$a_{3i} \cong 6,89$$

Se usarán estos valores iniciales en el cálculo de mínimos cuadrados no lineales. Usando el software R, se obtienen los siguientes resultados:

Para los coeficientes estimados a_1, a_2, a_4 , se obtienen valores P menores a 0,05 con un nivel de significancia del 0% (representado por el código de asteriscos generado por el software R) mientras que para la estimación de a_3 se logra un nivel de significancia de aproximadamente 5% Entonces para todos ellos se acepta la hipótesis de que son diferentes a cero y por lo tanto los coeficientes que definen la componente estacional de la serie temporal con saltos de precios en bolsa de la electricidad son:

```
Parameters:
  Estimate Std. Error t value Pr(>|t|)
a1 0.071412  0.004124  17.315  <2e-16 ***
a2 0.397402  0.009337  42.563  <2e-16 ***
a3 0.005709  0.002907   1.964  0.0496 *
a4 1.003834  0.008466 118.571  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figura 5 Resultado generado en R del ajuste por mínimos cuadrados no lineales para el modelo de la componente estacional aplicable a la serie de tiempo filtrada de tendencia

Tabla 3 Coeficientes obtenidos por regresión no lineal para el modelo estacional S_t estimados en ausencia de tendencia

COEFICIENTE	VALOR
a_1	0,0714
a_2	0,3974
a_3	0,0057
a_4	1,0038

Con estos coeficientes se construye un modelo para la estacionalidad a largo plazo, definida en (14) y se sustrae de la serie temporal de precios. Así, el modelo aditivo para dicha serie queda:

$$P_t'' = P_t' - T_t - S_t \quad (20)$$

Tal como se observa en la Figura 7.

En T_t está contenida la tendencia y en S_t la estacionalidad de largo plazo, es necesario tener en cuenta que sustraer T_t y S_t de la serie de precios no elimina del todo su componente estacional, ya que se podría tener presente una componente estacional de corto plazo s_t [20]. También es necesario tener en cuenta que en la remoción de esta nueva componente se corre el riesgo de remover también parte de la componente aleatoria de la serie [18] y para evitar esto se procederá a hacer la remoción de s_t usando índices estacionales de corto plazo para cada día de la semana. El proceso es así:

- De la serie temporal P_t'' se calculan medias móviles de siete en siete datos (una semana) que representen el valor promedio de una semana y con ellas la estacionalidad semana a semana (corto plazo): Se calcula una primera media móvil partiendo de la observación P_1'' hasta la observación P_7'' y se continúa hacia adelante. Como se tratará de medias con cantidad impar de datos, estarán centradas en el cuarto

día y no será necesario un proceso adicional de centrado. Por efecto del cálculo de las medias móviles se perderán los tres primeros datos y los tres últimos de la serie.

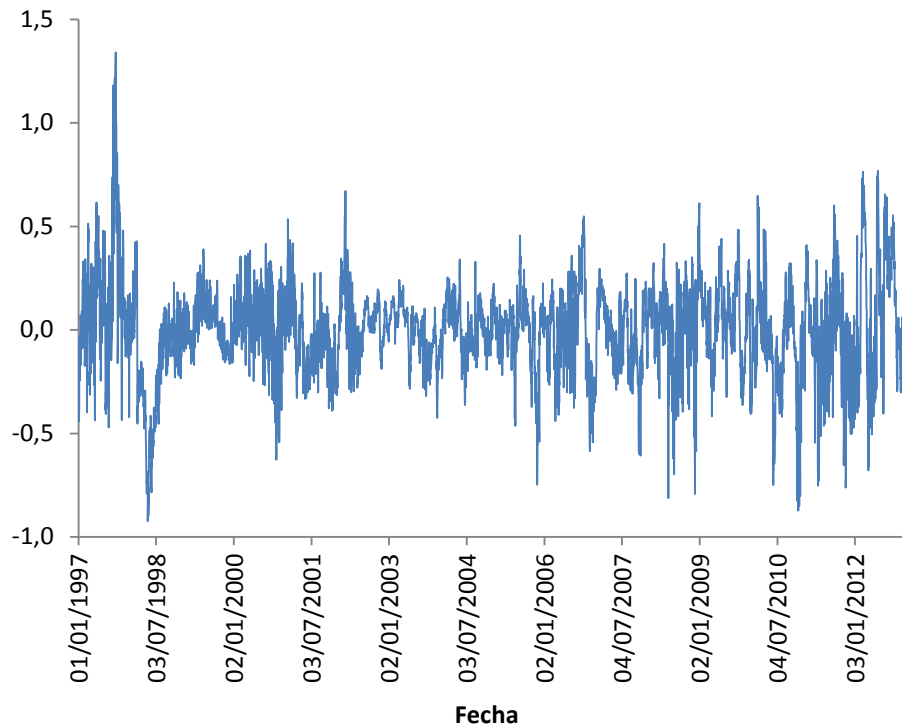


Figura 6 Serie de tiempo de P'_t

- Como la serie de tiempo original $P'_t = \ln P_t$ manifiesta un esquema aditivo, el resultado de cada media móvil se *restará* del correspondiente valor P'_t . Cada uno de los resultados anteriores corresponderá al día de la semana en el que esté ese valor P'_t . Llámese a este resultado V_t .
- Se calculan siete promedios, cada uno correspondiente a un día de la semana; es decir, se promedian todos los V_t del lunes, todos los V_t del martes, todos los V_t del miércoles etcétera, incluyendo días feriados. Luego se calcula el promedio de todos estos resultados (un promedio de promedios) así:

V_l = promedio de los valores V_t de todos los lunes

V_m = promedio de los V_t de todos los martes

V_c = promedio de los V_t de todos los miércoles

⋮
⋮
⋮

$$\bar{V} = \text{promedio de los promedios} = \frac{\sum_{\text{lunes}}^{\text{domingo}} V_d}{7} \quad (21)$$

El valor \bar{V} es un factor de ajuste que se usa para calcular los índices estacionales de corto plazo para cada día:

$$\begin{aligned} s_l &= V_l - \bar{V} \text{ (índice estacional de los lunes)} \\ s_m &= V_m - \bar{V} \text{ (índice estacional de los martes)} \\ s_c &= V_c - \bar{V} \text{ (índice estacional de los miércoles)} \\ &\vdots \\ &\vdots \\ &\vdots \end{aligned}$$

Dado el esquema aditivo de la serie original P'_t , estos índices estacionales s_t se restan a P'_t de manera selectiva:

$$\begin{aligned} P''_t - s_l &\text{ si } t \text{ corresponde a un día lunes} \\ P''_t - s_m &\text{ si } t \text{ corresponde a un día martes} \\ P''_t - s_c &\text{ si } t \text{ corresponde a un día miércoles} \\ &\vdots \\ &\vdots \\ &\vdots \end{aligned}$$

Es decir, a cada P''_t que caiga un lunes se le resta s_l , a cada P''_t que caiga un martes se resta s_m y así para el resto de días en la serie temporal P''_t . El resultado obtenido será

$$\begin{aligned} K_t &= P''_t - s_t \\ K_t &= (P'_t - T_t - S_t) - s_t \quad (22) \end{aligned}$$

Esta es la serie de precios filtrada (sin tendencia ni estacionalidades de largo y corto plazo), se muestra en la Figura 8 y contiene la componente de saltos junto con la componente de reversión a la media y la componente aleatoria.

IDENTIFICACIÓN Y REMOCIÓN DE PICOS EN LA SERIE PRODUCIDOS POR SALTOS, REPETICIÓN DEL CÁLCULO DE LOS PARÁMETROS ESTACIONALES

Se procede entonces a identificar los picos en esta nueva serie K_t y filtrarlos. El filtrado es para que el modelado de la componente de volatilidad del proceso estocástico esté libres de estos picos pues este proceso de picos enmascara los procesos determinístico y estocástico.

Las fuentes [9], [21], [22] proponen el uso de un “filtro recursivo” consistente en llamar “picos” a todos los valores que sobrepasen un cierto número de desviaciones estándar para los datos de la serie, mientras que [19] muestra que para los mercados de electricidad alemán (EEX) y australiano (New South Wales) se logran mejores resultados cuando se llaman “picos”

a los valores que se encuentren dentro del 10% superior (inferior) o dentro del 2,5% superior (inferior); sin embargo en este estudio no se abordará este último enfoque.

En este estudio se abordará el enfoque de “filtro recursivo”, en particular se llamará picos a todos los valores que se encuentren dentro fuera de un rango de ± 2 desviaciones estándar, similar a la propuesta de [21] que propone usar un rango de $\pm 2,3$ desviaciones estándar.

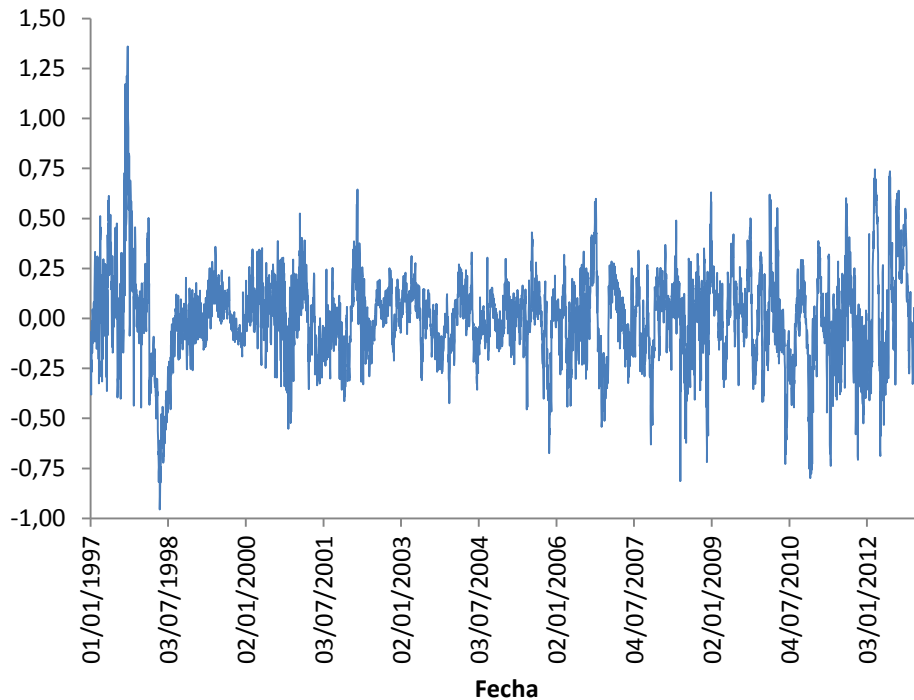


Figura 7 Serie de tiempo de Z_t

Con la identificación de los picos, sigue la remoción de ellos. El filtrado recursivo debe hacerse tantas veces como sea necesario hasta no encontrar ningún pico [21], [22]. Cada uno de los valores que se clasifique como pico debe ser remplazado. En [19] se propone remplazar cada uno de ellos por el promedio simple de la serie desestacionalizada; sin embargo se usará la propuesta de [23] consistente en reemplazar cada valor pico de un día dado por la mediana de los valores del mismo día de la semana dentro del mismo mes.

Con este filtrado, cuyo resultado puede visualizarse en la Figura 8 se tiene entonces una base para reconstruir una serie temporal de precios sin picos ni tendencia ni componente estacional de corto plazo. Esto se logra reintegrando (sumando) la componente estacional de largo plazo que fue restada previamente. Después de esto se procede nuevamente a estudiar el comportamiento de dicha componente S_t usando mínimos cuadrados no lineales. Siguiendo la misma línea de razonamiento aplicada en la primera regresión cuadrática, se definen las siguientes constantes como valores iniciales para la regresión:

$$a_{2i} = \frac{1}{4}$$

$$a_{1i} = -52,36$$

$$a_{3i} = 51,95$$

$$a_{4i} = 0$$

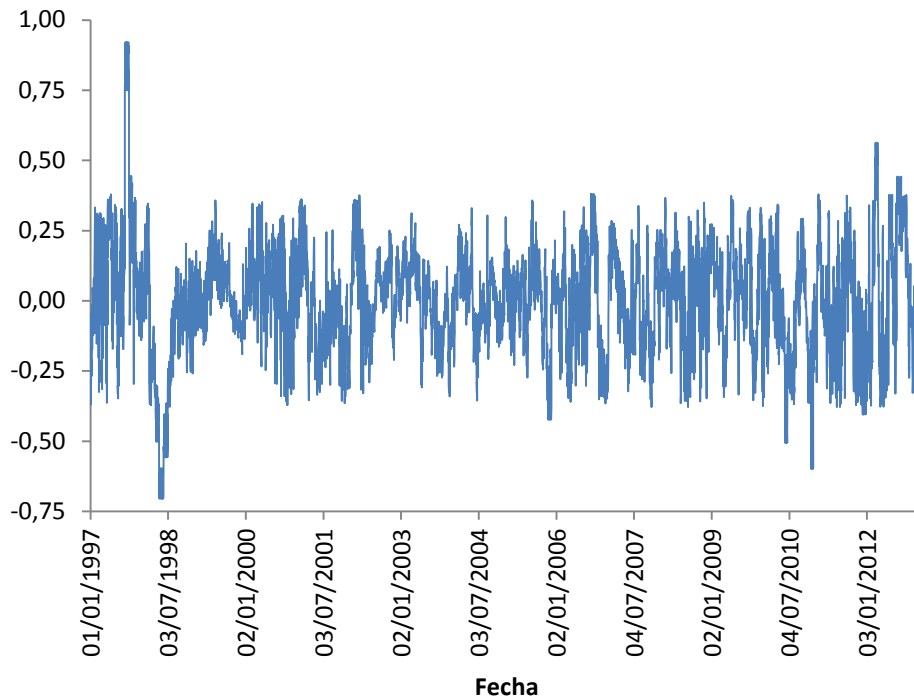


Figura 8 Serie de tiempo de L_t consistente en la serie Z_t libre de picos

Usando el software R para el procedimiento de mínimos cuadrados no lineales se obtienen los nuevos parámetros para la ecuación de componente de tendencia y estacionalidad a largo plazo (14)

```
Parameters:
  Estimate Std. Error t value Pr(>|t|)
a5 0.072820  0.003688  19.745  <2e-16 ***
a6 0.399166  0.008199  48.687  <2e-16 ***
a7 0.005852  0.002600   2.251  0.0244 *
a8 1.001750  0.007689 130.291  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figura 9 Resultado generado en R del ajuste por mínimos cuadrados no lineales para el modelo de la componente estacional para la serie de tiempo filtrada de saltos, tendencia y estacionalidad de corto plazo

Para los nuevos coeficientes estimados a_1, a_2, a_4 , se obtienen valores P menores a 0,05 con un nivel de significancia del 0% (representado por el código de asteriscos generado por el

software R) mientras que para la estimación de a_3 se logra un nivel de significancia de aproximadamente 1% Entonces para todos ellos se acepta la hipótesis de que son diferentes a cero y por lo tanto los coeficientes que definen la componente estacional de la serie temporal con saltos de precios en bolsa de la electricidad son los mostrados en la Tabla 4.

Tabla 4 Coeficientes obtenidos por regresión no lineal para el modelo estacional S_t estimados en ausencia de tendencia (T_t), picos y de estacionalidad de corto plazo s_t

COEFICIENTE	VALOR
a_1	0,0728
a_2	0,3992
a_3	0,0058
a_4	1,0018

Si se comparan los resultados en la Figura 9 con los resultados en la Figura 5 se puede concluir que:

- El error estándar para cada coeficiente estimado disminuye levemente, haciendo que el valor t para el ajuste por mínimos cuadrados no lineales sea mayor. Esto significa que los valores obtenidos para los coeficientes a_1, \dots, a_4 son significativos estadísticamente.
- El nivel de significancia para los coeficientes calculados sigue validando la hipótesis alterna (los coeficientes son distintos de cero)

Hasta el momento se cuenta con una serie temporal L_t originada en la serie P'_t , a la que se le ha removido su componente de tendencia T_t , sus componentes estacionales de largo y corto plazo S_t, s_t y los picos, que se modelan con un proceso estocástico de saltos. Queda entonces en la serie L_t la componente de reversión a la media y la componente aleatoria. Con el fin de contar con valores positivos para el cálculo de retornos y estimar correctamente modelos para la componente de reversión a la media y para la volatilidad se suman los picos removidos y una cantidad constante (se escogió sumar 1) a dicha L_t obteniendo:

$$K_t = L_t + Jdq + 1 \quad (23)$$

IDENTIFICACIÓN DE LA EXISTENCIA DE REVERSIÓN A LA MEDIA Y CARACTERIZACIÓN

A través de un análisis de rango reescalado sobre la serie (23) se puede determinar si el comportamiento de esta es netamente aleatorio o persistente o de reversión a la media [24]. Dicho análisis consiste en determinar un valor H conocido como *exponente de Hurst* para el que se puede decir que:

$H = \frac{1}{2}$ implica un proceso de ruido blanco

$H > \frac{1}{2}$ implica un proceso persistente o de tendencias

$H < \frac{1}{2}$ implica un proceso antipersistente o de reversión a la media

Esto es, si la serie de precios de la energía presenta un exponente de Hurst mayor a $1/2$ exhibe un comportamiento de correlación positiva entre incrementos (los aumentos en los precios van seguidos de más aumentos y las disminuciones en los precios van seguidas de más disminuciones) Si la serie de precios presenta un exponente de Hurst menor a $1/2$ exhibe un comportamiento de correlación negativa entre incrementos (los aumentos en los precios vienen seguidos de una disminución o reversión) Si la serie de precios presenta un exponente de Hurst igual a $1/2$ se tendrá un proceso completamente aleatorio, lo cual es imposible para un mercado de electricidad regulado.

La identificación se realiza sobre la serie de retornos de los precios constantes. Los pasos llevados a cabo para la identificación de dicho coeficiente [9] son:

- Se crearon siete copias de la serie de retornos sobre K_t (r_{ti}) con $i = 1, 2, \dots, 7$, que constarán de 5880 datos.
 - La primera copia ($i = 1$) se usó tal como es y por lo tanto tendrá un tamaño de $n_1 = 5880$ datos.
 - la segunda copia ($i = 2$) se dividió en $j = 2$ subseries de igual tamaño:
$$n_2 = \frac{5880}{2} = 2940 \text{ datos}$$
 - la tercera copia ($i = 3$) fue dividida en $j = 4$ subseries de igual tamaño:
$$n_3 = \frac{5880}{4} = 1470 \text{ datos}$$
 - la cuarta copia ($i = 4$) fue dividida en $j = 8$ subseries de igual tamaño:
$$n_4 = \frac{5880}{8} = 735 \text{ datos}$$
 - la quinta ($i = 5$) fue dividida en $j = 16$ subseries y se procedió así hasta la séptima copia que fue subdividida en $j = 64$ subseries. Como la cantidad total de datos es de 5880, no se tuvieron subseries de igual tamaño a partir de la quinta copia; sin embargo en todas estas copias se dejó recortada la última subserie.
- Para cada subserie en cada copia se calculó su media y su desviación estándar E_{ij} y S_{ij} respectivamente, con $i = 1, 2, \dots, 7$ representando la copia de r_t y $j = 1, 2, \dots, 64$ representando la subserie en dicha copia. Se tuvo en total $1 + 2 + 4 + 8 + 16 + 32 + 64 = 127$ medias y 127 desviaciones estándar.
- A partir de cada copia r_{ti} se generaron siete series normalizadas $N_i = r_{tij} - E_{ij}$ donde E_{ij} es la media de la subserie j dentro de la copia i

- Se generaron otras siete series cumulativas $C_t = \sum_{k=1}^t N_k$ donde cada valor es la suma acumulada de los N_i que le preceden.
- Se calcularon los rangos de cada una de las siete series. Estos rangos se definen como:

$$R_{ij} = \max_i\{C_1, C_2, \dots\} - \min_i\{C_1, C_2\}$$

Con i, j definidos como antes. Se contaron 127 rangos: Uno para r_{t1} , dos para r_{t2} , cuatro para r_{t3} , ..., treinta y dos para r_{t6} y sesenta y cuatro para r_{t7}

- Se reescaló cada rango dividiéndolo por la desviación estándar de su respectiva subserie: R_{11} se divide por S_{11} , R_{21} se divide por S_{21} , R_{22} se divide por S_{22} , R_{31} se divide por S_{31} y se procede hasta terminar con $(R/S)_7 = R_{ij}/S_{ij}$ donde $i = 7$ y $j = 64$.

Se obtuvieron en total 127 rangos reescalados.

- Se promediaron los rangos reescalados que son comunes, es decir, se calcula el promedio de los R_{1j}/S_{1j} con $j = 1$; se calcula el promedio de los R_{2j}/S_{2j} con $j = 1, 2$ y

se prosigue así hasta calcular el promedio de los R_{7j}/S_{7j} con $j = 1, 2, 3, \dots, 64$. Se

obtuvieron en total siete promedios de rangos escalados: Uno por cada copia r_{ti}

- Con base en la relación asintótica

$$(R/S)_i \sim cn^H$$

Reexpresada como:

$$\log(R/S)_i = H \log n + \log c \quad (24)$$

Con $n = \{n_1, n_2, n_3, \dots\}$ Se realiza un ajuste de mínimos cuadrados en una curva $\log(R/S)_i$ vs. $\log n$ como la mostrada en la Figura 3. La pendiente de la recta que ajusta a la relación en (24) será el exponente de Hurst buscado que para este caso es $H = 0,3385$

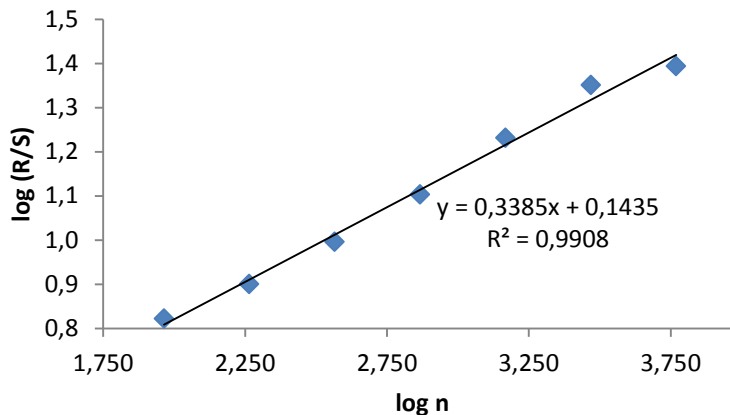


Figura 10 Resultado del análisis de rango reescalado sobre la serie K_t . En este se aprecia que para dicha serie, $H = 0,3385$

Se concluye del valor encontrado de H que la serie K_t obtenida a partir de precios constantes de la electricidad en bolsa P_t exhibe un carácter antipersistente, es decir, de reversión a la media. Se puede decir que esto es efecto del regulador del mercado así como de las estrategias de despacho de energía emprendidas para honrar el cargo por confiabilidad acordado entre los generadores y los compradores. También se hace evidente de acuerdo a [9] y [25] que la dinámica global de los precios en bolsa de la electricidad en Colombia está lejos de ser representable a través de un movimiento browniano, lo cual da cabida a pensar que en esta dinámica global hay una componente susceptible de ser modelada y otra componente aleatoria no determinada.

La serie K_t se compondrá de una parte determinística susceptible de ser modelada, de una componente de saltos y de una componente aleatoria. Dicha serie se puede expresar como:

$$dK_t = \alpha(\mu - K_t)dt + Jdq + \sigma_t dW_t \quad (25)$$

En donde el primer término es el término de reversión a la media, el segundo término es el proceso de saltos que se agregó nuevamente después de determinar los parámetros de la componente estacional de largo plazo S_t y el tercer término es aleatorio, α es la tasa de reversión a la media y μ es dicha media de largo plazo. Para la estimación de los parámetros del proceso de reversión a la media se puede proceder de dos maneras distintas, las cuáles se espera que arrojen resultados similares. Una de ellas consiste en proceder de la siguiente manera [21], [11], [26], [27]

Se lleva a cabo un ajuste por mínimos cuadrados sobre una discretización de (25) la cual se asume que es un proceso de Lévy [21] que obedece la ecuación diferencial estocástica:

$$dK_t = -\alpha(K_t - \mu)dt + Jdq + \sigma_t dW_t \quad (26)$$

Con α parámetro de reversión a la media (que se asumirá constante) y σ_t la volatilidad del proceso de difusión.

De acuerdo a [21], usando una discretización de (26) se puede tener que:

$$\begin{aligned} \Delta K_t &= -\alpha(K_t - \mu)\Delta t + J\Delta q + \sigma_t \Delta W_t \\ \Delta K_t &= -(\alpha K_t)\Delta t + \alpha\mu\Delta t + (J\Delta q) + (\sigma_t \Delta W_t) \end{aligned} \quad (27)$$

Equiparable a una ecuación lineal así:

$$\Delta K_t = \beta_1 K_t + \beta_0 + \varepsilon_t$$

En la que la variable independiente es K_t y la variable dependiente es $\Delta K_t = K_t - K_{t-1}$ Definiendo los parámetros [32]

$$\beta_1 = -\alpha\Delta t$$

$$\beta_0 = \alpha\mu\Delta t$$

$$\varepsilon_t = \sigma_t \Delta W_t$$

$$\Delta t = 1 \text{ día}$$

Acá, μ representa la media de largo plazo alrededor de la cual fluctúa y se revierte el proceso K_t y α representa la tasa de reversión a la media; el valor de Δt es consecuente con el hecho de que la serie original de precios es una serie de datos diarios; ε_t representa en este caso el error asociado a la regresión lineal, producto del proceso estocástico aleatorio $\sigma_t \Delta W_t$ y el proceso de saltos $J\Delta q$. Si se ejecuta una regresión entre ΔK_t y K_t se pueden obtener una estimación del término de reversión a la media mientras que el procesos de saltos puede modelarse separadamente y la volatilidad σ_t puede estimarse ajustando un modelo GARCH(1,1). El resultado de dicha regresión se muestra en la Figura 4. Así mismo, en la Tabla 6 se muestran los parámetros del proceso de reversión a la media $-\alpha(K_t - \mu)dt$

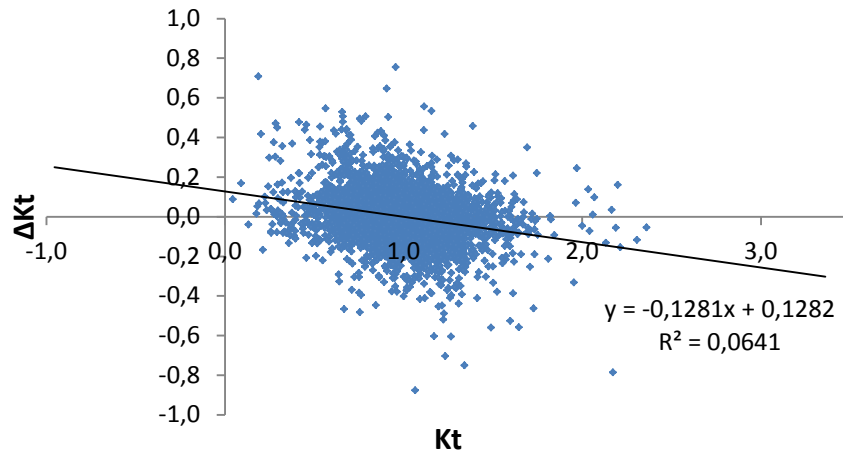


Tabla 5 Resultados del ajuste de mínimos cuadrados sobre ΔK_t para la obtención de los parámetros de reversión a la media de largo plazo μ

PARÁMETRO	β_1	β_0	α	μ
VALOR	-0,1281	0,1282	0,1281	1,0008

El segundo método a usar para obtener los parámetros del proceso de reversión a la media consiste en hacer una estimación de máxima verosimilitud. Basados en los estimadores

$$\mu = \frac{K_{ty}K_{txx} - K_{tx}K_{txy}}{n(K_{txx} - K_{txy}) - (K_{tx}^2 - K_{tx}K_{ty})}$$

$$\alpha = -\frac{1}{\Delta t} \ln \left(\frac{K_{txy} - \mu K_{tx} - \mu K_{ty} + n\mu^2}{K_{txx} - 2\mu K_{tx} + n\mu^2} \right)$$

Dónde:

$$K_{tx} = \sum_{i=1}^n K_{i-1}$$

$$K_{ty} = \sum_{i=1}^n K_i$$

$$K_{txx} = \sum_{i=1}^n K_{i-1}^2$$

$$K_{tyy} = \sum_{i=1}^n K_i^2$$

$$K_{txy} = \sum_{i=1}^n K_i K_{i-1}$$

$n = 5880$ datos

Se obtienen los resultados mostrados en la Tabla 6. Comparando con los resultados en la Tabla 5 se concluye que los valores de α y de μ no difieren significativamente de los valores obtenidos por ajuste de mínimos cuadrados

Tabla 6 Resultado del ajuste por máxima verosimilitud sobre K_t para la obtención de los parámetros de reversión a la media de largo plazo μ

PARÁMETRO	α	μ
VALOR	0,1371	1,0002

CARACTERIZACIÓN DE LA VOLATILIDAD

Con este resultado puede eliminarse la componente de reversión a la media de la serie K_t , quedando entonces solo la componente de saltos, que altera la volatilidad del proceso de precios además del término aleatorio de la ecuación (27) y así se puede generar un modelo para la volatilidad del proceso global, que puede basarse en un modelo GARCH. Con base en (27) sea la serie:

$$E_t = J\Delta q + \sigma_t \Delta W_t + 1 = \Delta K_t + (\alpha K_t) \Delta t - \alpha \mu \Delta t + 1 \quad (28)$$

Para comprobar si existe heterocedasticidad en la volatilidad de esta serie se lleva a cabo un análisis de preestimación sobre ella.

Se determinaron gráficos de la función de autocorrelación (FAC) y de la función de autocorrelación parcial (FACP) para la serie de precios, la serie de rendimientos y la serie E_t . Si se observa que las autocorrelaciones en las gráficas de FACP decrecen con más rapidez que las autocorrelaciones en las gráficas de FAC se tiene una evidencia para sospechar que la volatilidad de la serie varía con el tiempo (heterocedasticidad). Los resultados se observan en las Figuras 11 y 12. De ellas no se puede concluir que para la serie aleatoria E_t analizada haya un efecto autorregresivo AR lo que significa que posiblemente no haya presencia de dicho efecto en ella.

No obstante, se debe continuar con más chequeos para corroborar la presencia de heterocedasticidad y poder formular un modelo para la volatilidad σ_t . Se procede a realizar una prueba de Ljung-Box que, a diferencia de las FAC y FACP permite chequear autocorrelación a múltiples rezagos conjuntamente [28] Esta prueba se lleva a cabo sobre los “retornos” de la serie que se desea analizar. Como se quiere estudiar el efecto autorregresivo en la serie E_t se ejecuta la prueba sobre una serie de residuales al cuadrado generada a partir de este E_t . El resultado de dicha prueba se muestra en la Tabla 7. Se rechaza la hipótesis nula de que no todas las autocorrelaciones hasta el rezago 5 o hasta el rezago 20 son cero.

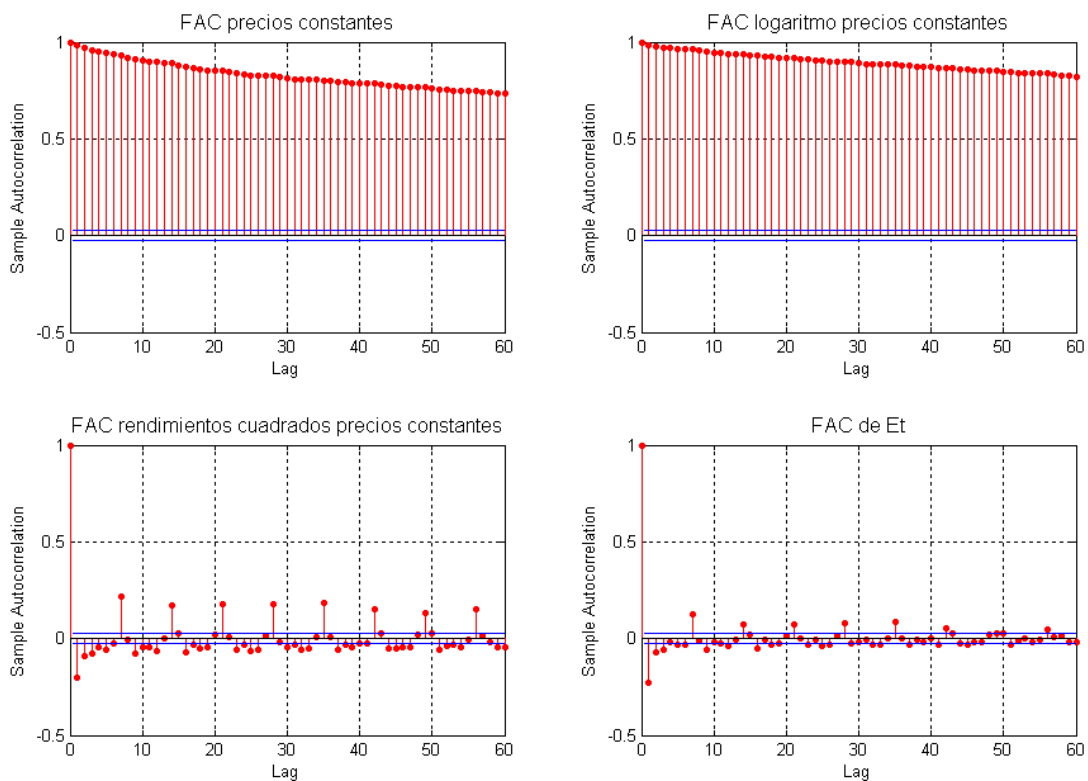


Figura 11 Gráficas de las FAC para las series temporales analizadas. Nótese el efecto estacional observado en el tercer panel y las autocorrelaciones negativas típicas de un proceso de reversión a la media [9], [21].

Tabla 7 Resultados de la prueba de Ljung-Box sobre la serie de residuales cuadrados generados a partir de la serie E_t

REZAGO	VALOR h	VALOR p	ESTADÍSTICO	VALOR CRÍTICO
5	1	0	610,7	11,07
20	1	0	1189,4	31,41

Tabla 8 Resultados de la prueba ARCH sobre la serie de residuales generados a partir de la serie E_t

REZAGO	VALOR h	VALOR p	ESTADÍSTICO F	VALOR CRÍTICO χ^2
5	1	0	1405,9	11,07
10	1	0	1430,3	18,31
20	1	0	1448,1	31,41
30	1	0	1458,9	43,77

Esto es indicativo de correlación significativa entre los valores de los retornos cuadrados calculados sobre E_t y del agrupamiento de volatilidad en la serie de retornos. Además, tal como se ve en la Figura 13, los retornos fluctúan alrededor de un nivel constante y parecen mostrar agrupamiento en la volatilidad y cambios repentinos producidos por los saltos.

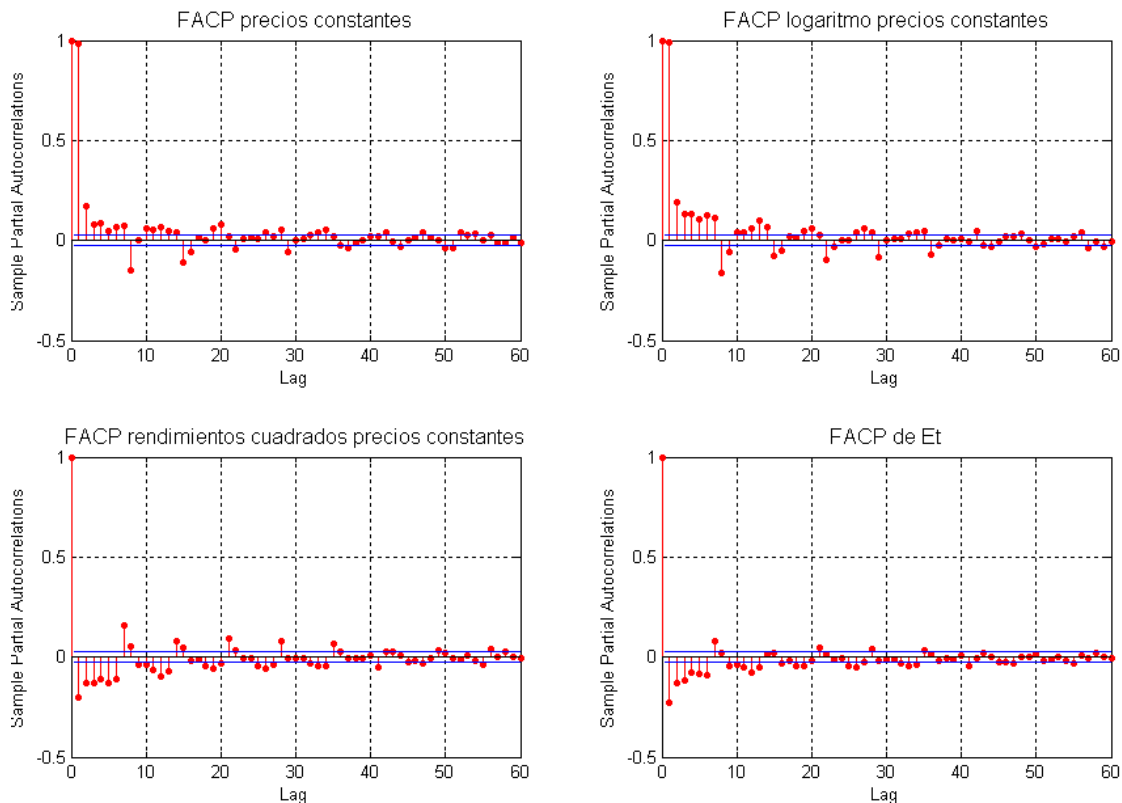


Figura 12 Gráficas de las FACP para las series temporales analizadas.

A continuación se lleva a cabo una prueba ARCH para comprobar la presencia de heterocedasticidad condicional. Los resultados de la prueba se muestran en la Tabla 8, de

estos se puede concluir que se rechaza directamente la hipótesis nula para los rezagos 5, 10, 20, 30 ya que para ellos el valor h es igual a uno y para ellos su valor p es cero. El valor del estadístico F para cada rezago chequeado es mucho mayor que el valor crítico en la distribución χ^2 . Se concluye entonces que para la serie de residuales generada a partir de E_t hay agrupamiento de volatilidad significativo y se tiene un efecto ARCH en ella. Se procede entonces a estimar un modelo GARCH para la volatilidad de E_t .

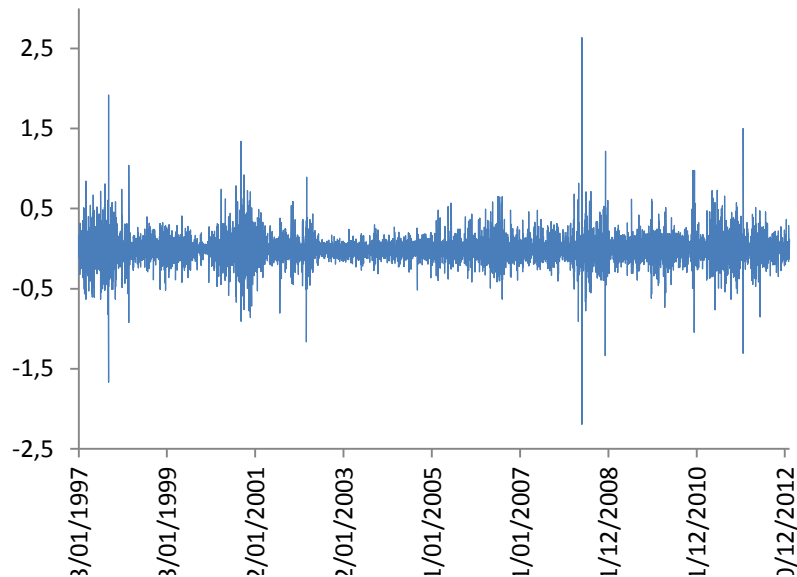


Figura 13 Serie de retornos sobre E_t

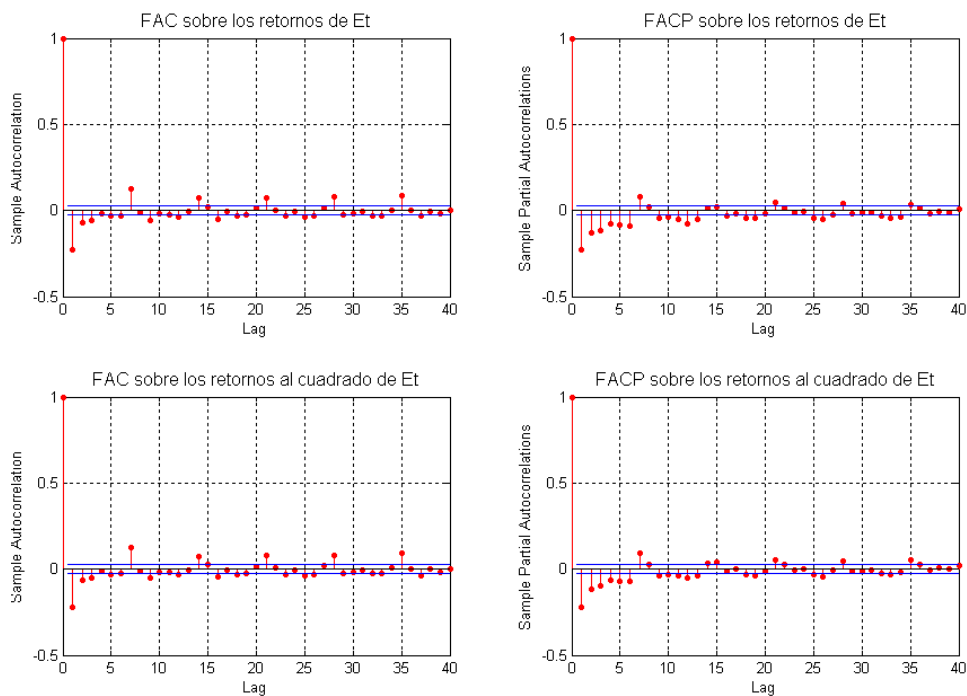


Figura 14 FAC y FACP para la serie de retornos y de retornos al cuadrado de E_t . En ella no se evidencia la presencia de autocorrelaciones.

Se ajusta para la volatilidad un modelo GARCH (1,1) debido a su naturaleza parsimoniosa que le permite requerir de menos rezagos para capturar la propiedad de varianza variable [29]. Los resultados de la estimación de parámetros para el ajuste se muestran en la Tabla 9. Y por lo tanto, una relación para estimar la volatilidad de E_t sería:

$$\sigma_t^2 = 1,1845 \times 10^{-4} + 0,9186\sigma_{t-1}^2 + 0,0748\varepsilon_{t-1}^2 \quad (29)$$

Tabla 9 Resultados de la estimación de parámetros de un modelo GARCH (1,1) a la volatilidad de E_t

PARÁMETRO	VALOR	ERROR ESTÁNDAR	ESTADÍSTICO T
C	0,9995	$1,0536 \times 10^{-3}$	948,7
K	$1,1845 \times 10^{-4}$	$1,0393 \times 10^{-5}$	11,4
GARCH (1) (G_1)	0,9186	$3,2714 \times 10^{-3}$	280,8
ARCH (1) (A_1)	0,0748	$3,6223 \times 10^{-3}$	20,7

La validez del modelo obtenido puede evaluarse a partir de un correlograma sobre la serie de residuales estandarizados al cuadrado, es decir, un correlograma sobre la serie de residuales de la serie E_t divididos por su respectiva desviación obtenida a partir de la ecuación (29) [30]. Comparando este correlograma con el correlograma de los retornos cuadrados de E_t se puede observar que el modelo obtenido explica la heterocedasticidad en los retornos. Además, la suma $G_1 + A_1$ es igual a $0,9934 < 1$, cercana a la restricción integrada de no estacionaridad, condición necesaria para un modelo GARCH [31].

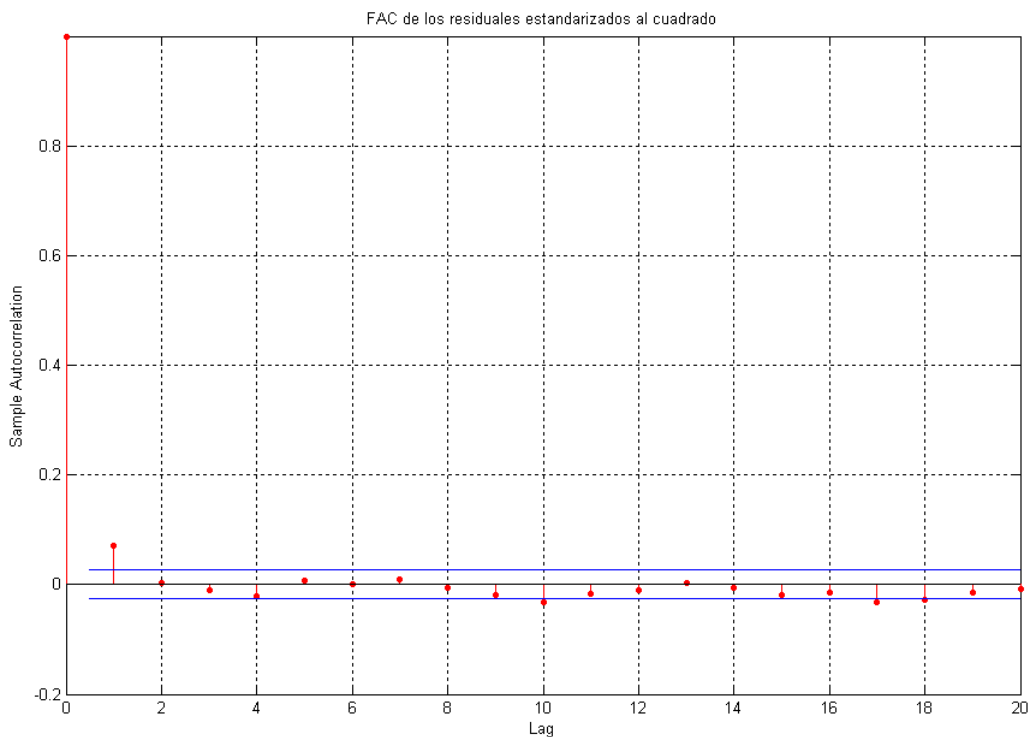


Figura 15 FAC del cuadrado de los residuales estandarizados de E_t .

Por otra parte, la serie de residuales ε_t del modelo GARCH (1,1) en (29) y exhibida en la Figura 16 no muestra un comportamiento normal como se esperaría, de acuerdo a los resultados de ajuste a distribuciones mostrados en la Figura 17. Se muestran las funciones de probabilidad acumulada y el error asociado a ellas y puede verse que el menor error está asociado con una serie t de Student. Así mismo, el ajuste a la distribución de probabilidad mostrado en la Figura 18 muestra que el mejor ajuste a los datos se logra con una distribución t de Student. Los parámetros para dicha distribución se muestran en la Tabla 10.

Tabla 10 Parámetros para la distribución t de Student ajustable a la serie de residuales ε_t

PARÁMETRO LOCACIONAL μ	PARÁMETRO DE ESCALA σ	GRADOS DE LIBERTAD ν
$-3,5828 \times 10^{-4}$	$6,9273 \times 10^{-2}$	2,8962

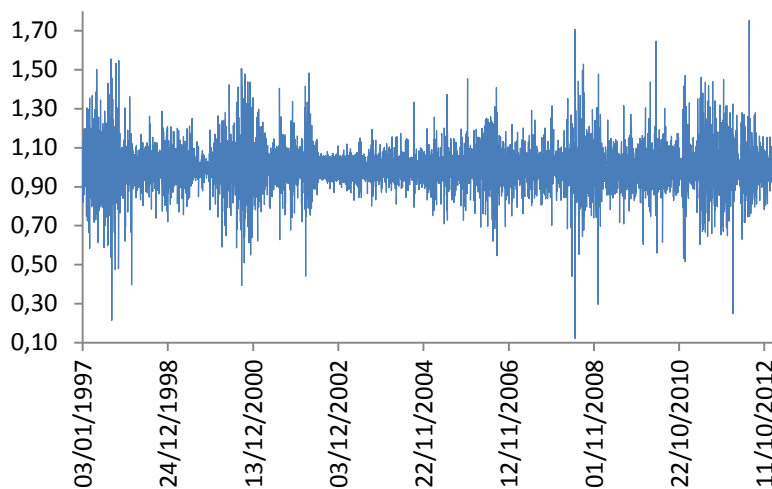


Figura 16 Serie de residuales ε_t para la serie de retornos de E_t

El resultado de este ajuste es de esperarse, porque la presencia del proceso de saltos aumenta la probabilidad de observar valores extremos, dando más peso a las colas de la distribución de probabilidad de los residuales. Se puede concluir de la alta curtosis de la distribución empírica de residuales que hay una alta probabilidad de que los residuales estén cercanos a cero.

CARACTERIZACIÓN DE LOS PARÁMETROS DEL PROCESO DE SALTOS

Los picos producidos por los saltos representan un porcentaje muy bajo del total de observaciones. En el punto 3 se identificaron los picos que genera el proceso de saltos. Estos picos pueden ser ascendentes o descendentes y en total se encontraron 442 observaciones clasificables como picos generados por el proceso de saltos, de las cuales 216 corresponden a subidas y 226 corresponden a descensos en la serie (12); por lo tanto se definirán dos subprocesos distintos tal como plantea [21]

$$J(t) = J^+(t) + J^-(t)$$

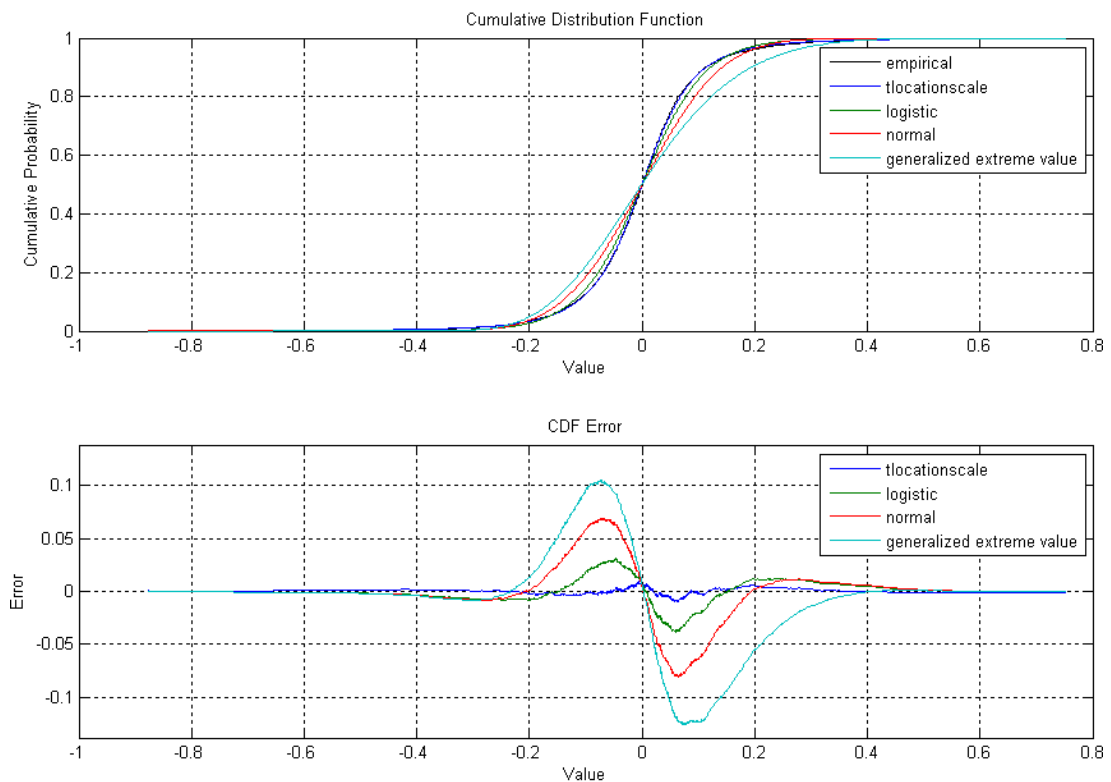


Figura 17 Funciones de distribución acumulada y de error asociado para un conjunto de distribuciones de probabilidad ajustadas a la serie de residuales ε_t

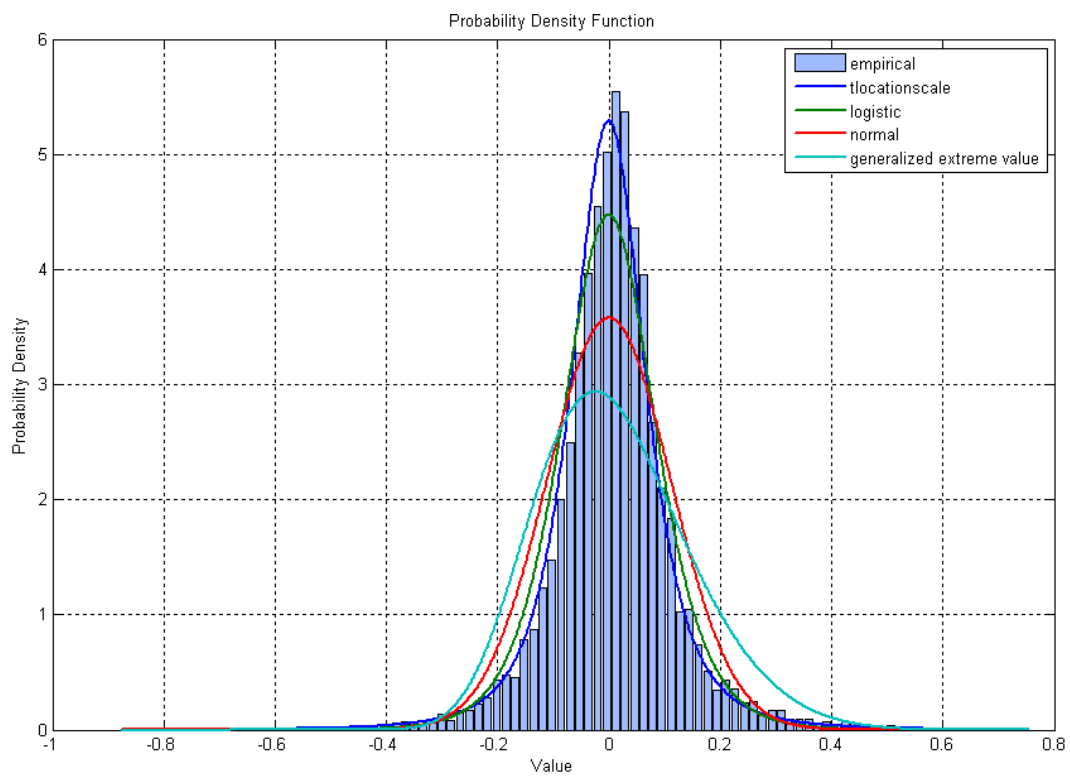


Figura 18 Ajuste de distribuciones de probabilidad a la serie de residuales ε_t

Cada uno de estos subprocessos tiene una frecuencia definida así:

$$\lambda_+ = \frac{\text{Cantidad de picos de subida}}{\text{Cantidad total de observaciones}} = \frac{216}{5881} = 0,03673$$

$$\lambda_- = \frac{\text{Cantidad de picos de bajada}}{\text{Cantidad total de observaciones}} = \frac{226}{5881} = 0,03843$$

La literatura menciona que el comportamiento de estos picos puede modelarse con un proceso compuesto de Poisson [viii - xiii] [] en el que la frecuencia de estos picos seguirá un proceso simple de Poisson dq definido como [32]:

$$dq = \begin{cases} \pm 1, & \text{si hay un salto, con probabilidad } \lambda_{\pm} dt \\ 0, & \text{si no hay salto, con probabilidad } 1 - \lambda_{\pm} dt \end{cases}$$

Y el tamaño de los picos seguirá un proceso lognormal de tal manera que:

$$\ln J_{\pm} \sim \mathcal{N}(\mu_{\pm}, \sigma_{\pm})$$

Garantiza siempre que la magnitud de los picos sea positiva.

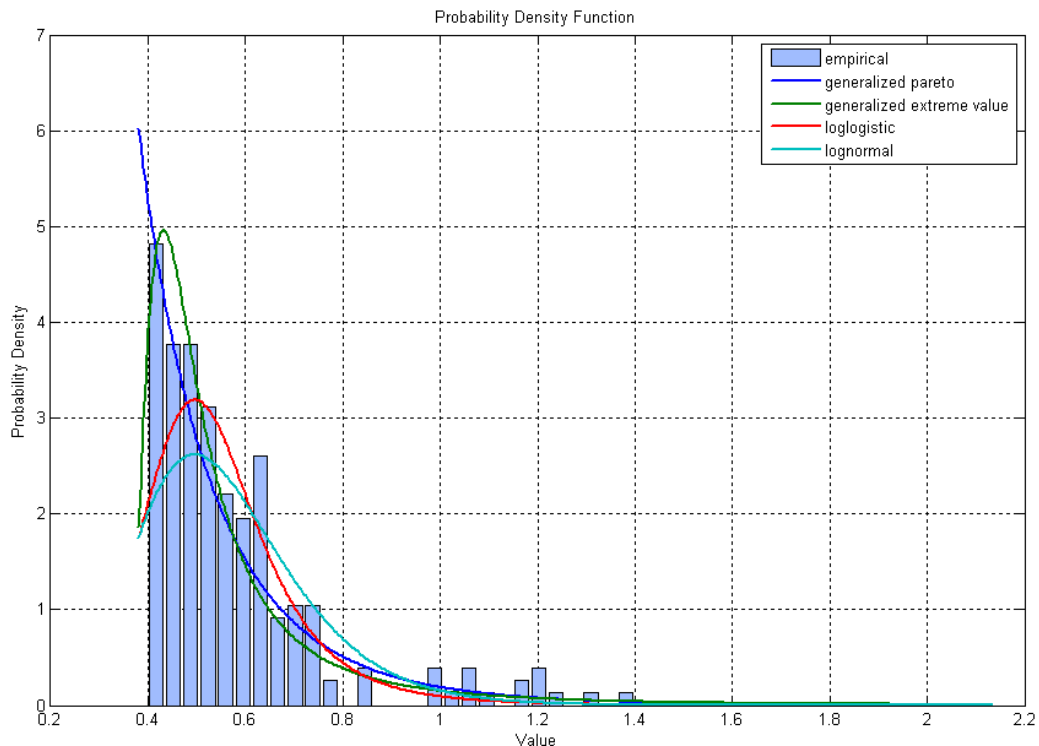


Figura 19 Ajustes de distribuciones de probabilidad a la magnitud de los picos de subida J_+

No obstante, si se hace una simulación de bondad de ajuste para los picos J_{\pm} se tienen resultados distintos: La distribución de la magnitud de los picos no se ajusta suficientemente bien a distribución alguna, teniendo como distribución más cercana la distribución de Pareto generalizada, una conclusión similar a la obtenida en un análisis realizado en [9]. Esto es un

hecho estilizado característico de la serie de precios de la electricidad en bolsa de Colombia, que se debe tener en cuenta a la hora de proponer un modelo para estos saltos.

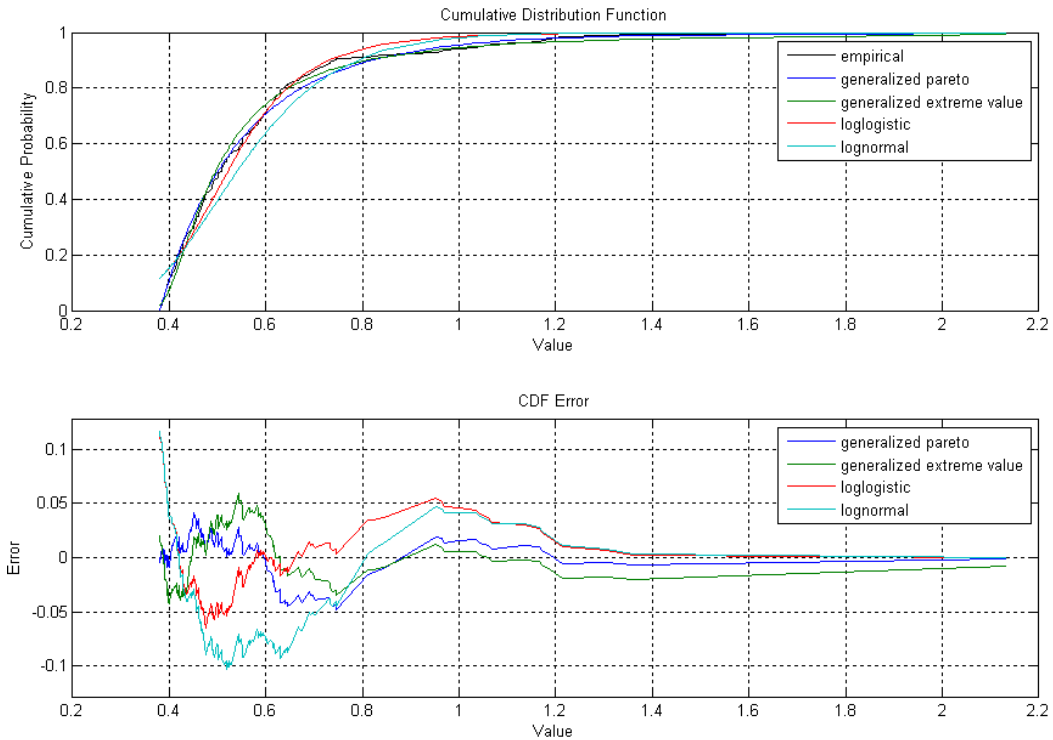


Figura 20 Funciones de distribución acumulada y de error asociado para el conjunto de magnitudes de los saltos de subida J_+

Como resultado, puede concluirse que para los precios en bolsa de la electricidad en Colombia se puede aplicar un modelo

$$P_t = e^{P'_t}$$

Dónde:

$$P'_t = \ln T_t S_t s_t K_t$$

En donde:

$$T_t = (-4 \times 10^{-8})t^2 + (6 \times 10^{-4})t + 2,5763$$

$$S_t = 0,0728 \operatorname{sen} \left[2\pi \left(\frac{t}{365} + 0,3992 \right) \right] + 0,0058 + 1,0018 \operatorname{EWMA}_t^{0,975}$$

Con:

$$\operatorname{EWMA}_t^{0,975} = (1 - 0,975) \ln P_t + 0,975 \operatorname{EWMA}_{t-1}^{0,975} \text{ una media móvil exponencial}$$

$$s_t = \begin{cases} 0,001 & \text{si es lunes} \\ 0,028 & \text{si es martes} \\ 0,031 & \text{si es miércoles} \\ 0,025 & \text{si es jueves} \\ 0,010 & \text{si es viernes} \\ -0,021 & \text{si es sábado} \\ -0,074 & \text{si es domingo} \end{cases}$$

K_t = Un proceso estocástico donde:

$$dK_t = -\alpha(K_t - \mu)dt + Jdq + \sigma_t dW_t$$

Para el cuál:

$$\alpha = 0,1371, \mu = 1,0002$$

$$\sigma_t^2 = 1,1845 \times 10^{-4} + 0,9186\sigma_{t-1}^2 + 0,0748\varepsilon_{t-1}^2 - 1$$

$\ln J_{\pm} \sim \mathcal{N}(\mu_{\pm}, \sigma_{\pm})$ y donde:

$$dq = \begin{cases} \pm 1, & \text{si hay un salto, con probabilidad } \lambda_{\pm} dt \\ 0, & \text{si no hay salto, con probabilidad } 1 - \lambda_{\pm} dt \end{cases} \text{ proceso de Poisson para el cuál:}$$

$$\lambda_+ = 0,03673$$

$$\lambda_- = 0,03843$$

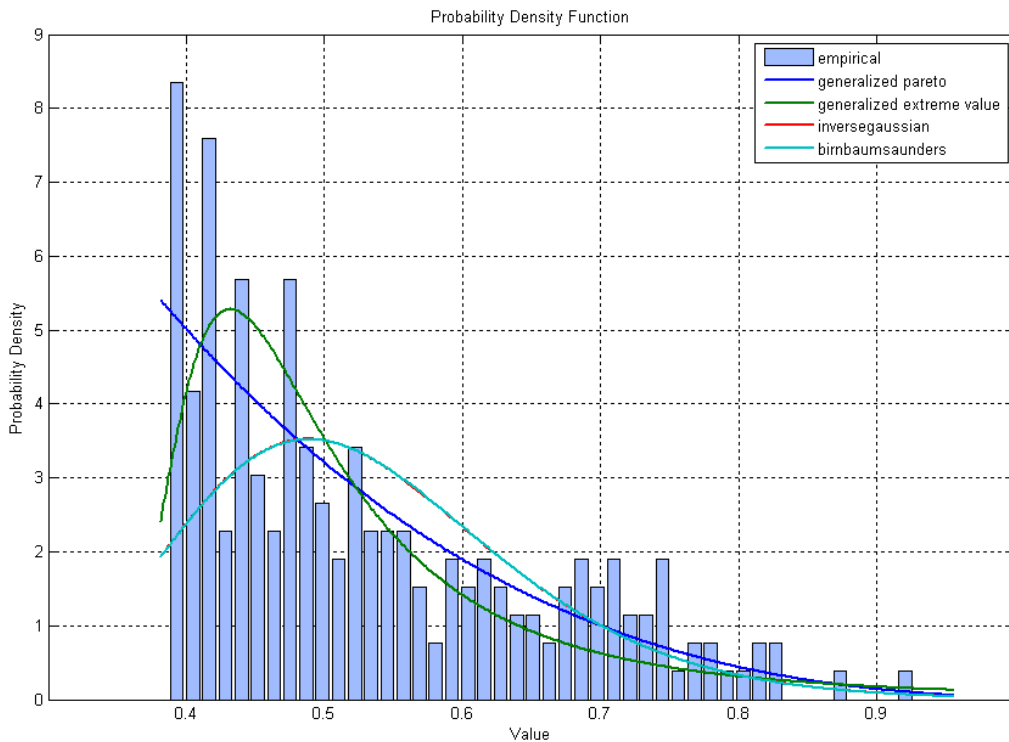


Figura 21 Ajustes de distribuciones de probabilidad a la magnitud de los picos de caída J_-

POSIBILIDAD DE APLICACIONES

Garantizando la bondad de ajuste del modelo presentado, se puede usar como base para modelar precios de derivados de electricidad tales como futuros y opciones. Para ello se debe garantizar la actualización permanente de los parámetros que conforman el modelo completo ya que, como se ha mostrado, todos ellos fueron determinados a partir de la serie histórica de precios. Si se garantiza la confiabilidad del modelo, también puede usarse para la predicción de precios futuros.

CONCLUSIONES

Se ha analizado la dinámica histórica de los precios en bolsa de la electricidad en Colombia desde el punto de vista de un modelo de difusión por saltos con reversión a la media. Se concluye que el proceso estocástico de difusión presente en los precios en bolsa no es un proceso browniano (ruido blanco normal) tal como modela la teoría básica de procesos estocásticos aplicados a la dinámica financiera [33], se concluye también que el proceso general de los precios en bolsa de la electricidad en Colombia exhibe un fuerte carácter antipersistente (de reversión en lugar de tendencia sostenida) lo que podría ser indicativo de las regulaciones del mercado en cuanto a que el regulador vela por mantener el nivel de precios en lugar de permitirle oscilar.

La cantidad de saltos que producen picos (*outliers*) es evidentemente menor en comparación de la cantidad de datos usados (menos del 8% de ellos) por lo tanto no se observa tanto riesgo como en otros mercados estudiados [ix, 5, 8, 14]. Cabe preguntarse si estos picos son producto de temporadas de clima seco o temporadas con problemas de orden público.

El trabajo presentado da espacio para mejoras y refinamiento de los planteamientos y de los modelos, toda vez que se reformulen los procesos estocásticos y se contemplen modelos más generales.

BIBLIOGRAFÍA

- [1] GAETA, Gordian. *Opportunities in Emerging Markets: Investing in the Economies of Tomorrow*. 2012. John Wiley & Sons.
- [2] STANDARD & POORS DOW INDICES. *S&P Frontier Indices Methodology*. Febrero de 2012. Tomado de la Internet en junio de 2013 del sitio www.standardandpoors.com
- [3] VELÁSQUEZ HENAO, Juan David. DYNER RESONSEW, Isaac. CASTRO SOUZA, Reinaldo. *¿Por qué es tan Difícil Obtener Buenos Pronósticos de los Mercados Competitivos?* 2007. Cuadernos de Administración, Vol. 20 (34) Universidad Javeriana.
- [4] BOTERO BOTERO, Sergio. CANO CANO, Jovan. *Análisis de Series de Tiempo para la Predicción de los Precios de la Energía en la Bolsa de Colombia*. 2008. Cuadernos de Economía vol. 48. Universidad Nacional de Colombia.
- [5] MORENO, Rodrigo. BARROSO, Luiz. RUDNICK, Hugh. BEZERRA, Bernardo. MOCARQUER, Sebastian. FLACH, Bruno. MORENO, Jorge. *Lessons from Five Years of Experience in Energy Contract Auctions in South America*. 2010. Proceedings of the 33rd International Association for Energy Economics.
- [6] COMISIÓN DE REGULACIÓN DE ENERGÍA Y GAS. *Cargo por Confiabilidad: Esquema Regulatorio Para Asegurar la Confiabilidad en el Suministro de Energía Eléctrica en Colombia, una Visión de Largo Plazo*. 2007. Documento informativo tomado de www.creg.gov.co
- [7] VILLARREAL, Julio. SOTO, Carlos. *Ensuring Generation Adequacy with Firm Energy Call Options: A Case Study for the Hydro-dependant Colombian System*. 2007. International Journal of Management Science and Engineering Management Vol. 2 No 1.
- [8] PILIPOVIC, Dragana. *Energy Risk: Valuing and Managing Energy Derivatives*. 2007. McGraw-Hill.
- [9] WERON, Rafał. *Modeling and Forecasting Electricity Loads and Prices. A Statistical Approach*. 2006. Wiley Finance Series. John Wiley & Sons.
- [10] WERON, R. BIERBRAUER, M. TRÜCK, S. *Modelling Electricity Prices: Jump Diffusion and Regime Switching*. 2004. Physica A 336.
- [11] BLANCO, Carlos. SORONOW, David. *Mean Reverting Processes: Energy Price Processes Used For Derivatives Pricing & Risk Management*. Junio 2001. Commodities Now.
- [12] GAARDER HAUG, Espen. *The Complete Guide to Option Pricing Formulas*. 2006. McGraw-Hill.
- [13] CARTEA, Álvaro. FIGUEROA, Marcelo. *Pricing in Electricity Markets: A Mean Reverting Jump Diffusion Model with Seasonality*. 2005. Applied Mathematical Finance vol. 12 No 4.

- [14] REBONATO, Riccardo. *Volatility and Correlation in the Pricing of Equity, FX and Interest-Rate Options*. 2000. John Wiley & Sons.
- [15] DELURGIO, Stephen A. *Forecasting Principles and Applications*. 1998. McGraw-Hill.
- [16] DAGUM, E.B. *X-11: The Magic Box and Four Golden Rules of Seasonal Adjustment*. 1976. Statistics Canada Research Paper 76-06-007E.
- [17] MIMMACK, Gillian. MANAS, Gary. MEYER, Denny. *Introductory Statistics for Business*. 2001. Pearson South Africa.
- [18] WILD, C.J. SEBER, G.A.F. *Chance encounters: A first course in data analysis and inference*. 1999. Wiley.
- [19] JANCZURA, Joanna. TRÜCK, Stefan. WERON, Rafał. WOLFF, Rodney C. *Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling*. 2012. Preprint submitted to Elsevier. 7
- [20] DE JONG, C. *The Nature of Power Spikes: A Regime Switch Approach*. 2006. Studies in Nonlinear Dynamics & Econometrics 10 (3)
- [21] MAYER, Klaus. SCHMID, Thomas. WEBER, Florian. *Modeling Electricity Spot Prices – Combining Mean-Reversion, Spikes and Stochastic Volatility*. 2011. CEFS Working Papers Series No 2011-02.
- [22] NOMIKOS, Nikos. ANDRIOSPOULOS, Kostas. *Modelling Energy Spot Prices: Empirical Evidence from NYMEX*. 2011. Energy Economics. Elsevier.
- [23] BIERBRAUER, M. MENN, C. RACHEV, S.T. TRÜCK, S. *Spot and Derivative Pricing in the EEX Power Market*. 2007. Journal of Banking and Finance 31.
- [24] MANDELBROT, B. *Robustness of the Rescaled Range R/S in the Measurement of Noncyclic Long-run Statistical Dependence*. 1969. Water Resources Research No 5.
- [25] SHIRYAEV, Albert. *Essentials of Stochastic Finance: Facts, Models and Theory*. 1999. World Scientific Publishing
- [26] BENTH, F. E. KIESEL, R. NAZAROVA, A. *A critical empirical study of three electricity spot price models*. 2012. Energy Economics, 34 (5)
- [27] CHANG, Yeong. *An Application of “Earning at Risk” to Assess Corporate Finance Risk*. 2006. Proceedings of the 2006 Crystal Ball Users Conference.
- [28] LJUNG, G. BOX, G. *On a Measure of Lack of Fit in Time Series Models*. 1978. Biometrika Vol. 66.
- [29] WANG, Peijie. *Financial Econometrics*. 2008. Routledge Advanced Texts in Economics and Finance. Routledge.

- [30] THE MATHWORKS INC. *GARCH Toolbox for use with MatLab®*. 2000.
- [31] BOLLERSLEV T. *Generalized Autoregressive Conditional Heteroskedasticity*. 1986. Journal of Econometrics, Vol. 31.
- [32] MCDONALD, Robert. *Derivative Markets, 2nd Edition*. 2006. Addison Wesley
- [33] HULL, John. *Options, Futures and Other Derivatives*. 2009. Pearson Prentice-Hall.