

# Modelización estocástica de los precios diarios de la electricidad del mercado mayorista español

Cortés López, Juan Carlos (jccortes@mat.upv.es)

*Instituto Universitario de Matemática Multidisciplinar*

*Universitat Politècnica de València*

Debón Aucejo, Ana María (andeau@eio.upv.es)

*Centro de Gestión de Calidad y del Cambio*

*Universitat Politècnica de València*

Rico Llopis, Manuel (marillo@ade.upv.es)

*Facultad de Administración y Dirección de Empresas*

*Universitat Politècnica de València*

## RESUMEN

El objetivo del presente trabajo es profundizar en los métodos de modelización aplicados en trabajos anteriores, construyendo un modelo para los precios de la electricidad de los días laborables en el mercado mayorista español mediante la combinación de técnicas analíticas y estadísticas. El Modelo parte de la base de las componentes clásicas, Tendencia, Ciclo, Estacionalidad y Componente Irregular, con la novedad de la inclusión de procesos estocásticos en la modelización de la última de las componentes. El Modelo propuesto se completa con un análisis estadístico de los resultados.

**Palabras clave:** Series temporales, procesos estocásticos, mercado eléctrico español.

**Área temática:** Aspectos cuantitativos del fenómeno económico (A5)

## **ABSTRACT**

The aim of this paper is to deepen in the modelization methods applied in previous papers, by building a model for electricity prices of working days in the Spanish wholesale market through the combination of analytic and statistic techniques. The Model has its grounds on the classic components, Trend, Cycle, Seasonality and Irregularity, with the novelty of the inclusion of stochastic processes for modeling the last of the components. The proposed Model is completed with a statistic analysis of the results.

**Keywords:** Time series, stochastic processes, Spanish electricity market.

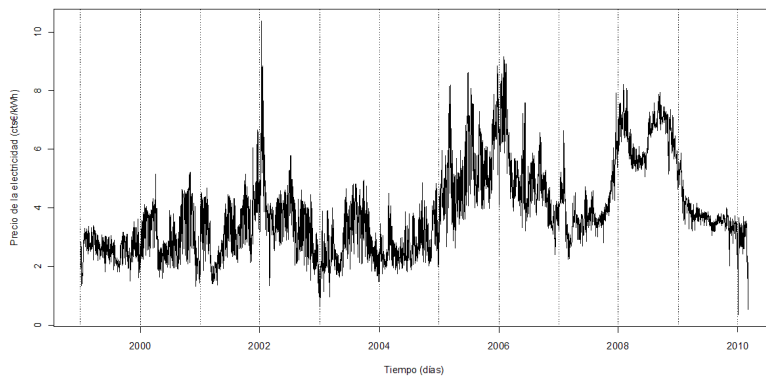
# 1 INTRODUCCIÓN

La reciente tendencia liberalizadora en los sectores eléctricos de los países desarrollados ha tenido como principal consecuencia la creación de los mercados mayoristas de la electricidad, principal manifestación del libre funcionamiento del sector. En el caso de España, la Ley 54/1997 del Sector Eléctrico es la que sienta las bases para establecer el marco regulador para la liberalización del sector, culminando en la constitución del mercado mayorista en el año 1999. Con este nuevo marco, se facilita la entrada de agentes privados a un sector vetado desde su creación a la iniciativa empresarial, previamente monopolizado en su práctica totalidad por el Sector Público (Lucía y Meneu, 2004).

Con la creación del mercado mayorista, los distintos operadores se han sometido a las leyes competitivas derivadas de la interacción de la oferta y la demanda. Este hecho es el que propicia que los precios de la electricidad sigan una evolución a priori impredecible, ya que su dinámica depende de una gran cantidad de factores, muchos de los cuales, por su naturaleza cualitativa, son de difícil medición y, por ende, de difícil inclusión en un modelo matemático. En efecto, uno de los factores cualitativos decisivos en los precios es el papel regulatorio del Estado, totalmente impredecible, y cuya cuantificación es realmente compleja y, en ocasiones, imprecisa. Es por ello que la regulación en este sector, oligopolístico, que tiende al monopolio en algunas actividades (transporte y distribución), se concreta en medidas que configuran un funcionamiento lo más competitivo posible, para lo cual se requiere un marco que garantice una sólida estabilidad.

En España, los precios de la electricidad son negociados para cada una de las horas del día, resultando ello en la existencia de 24 precios horarios, los cuales son publicados por el Operador del Mercado (OMEL). El estudio detallado del comportamiento de dichos precios ha revelado la existencia de una serie de patrones

a partir del año 2006 sobre los que basaremos nuestro modelo. Este hecho tiene su justificación en la entrada en vigor del Real Decreto-Ley 3/2006<sup>1</sup>, que modificó de manera significativa el modo de funcionamiento de los agentes del mercado. La presencia de dichos patrones es la principal motivación para elaborar un modelo de precios que los reproduzca y responda en la medida de lo posible a las necesidades de información de unos agentes que actúan en condiciones de elevada incertidumbre en el mercado mayorista.



**Figura 1:** Precios diarios desde el 01/01/1999 al 31/03/2010

El objetivo del presente trabajo es profundizar en los métodos de modelización aplicados en Cortés, Debón y Rico (2010), construyendo un modelo para los precios de la electricidad de los días laborables en el mercado mayorista español mediante la combinación de técnicas analíticas y estadísticas. El Modelo parte de la base de las componentes clásicas, Tendencia, Ciclo, Estacionalidad y Componente Irregular, con la novedad de un tratamiento previo sobre la serie y, la inclusión de procesos estocásticos en la modelización de la última de las componentes. El Modelo se completa con un análisis estadístico, que supone la validación de las hipótesis que sustentan cada parte del modelo, y la medición de la bondad del ajuste total.

---

<sup>1</sup>REAL DECRETO-LEY 3/2006, de 24 de febrero, por el que se modifica el mecanismo de casación de las ofertas de venta y adquisición de energía presentadas simultáneamente al mercado diario e intradiario de producción por sujetos del sector eléctrico pertenecientes al mismo grupo empresarial. BOE núm. 50 de 28 de febrero de 2006, página 8015, entrada 3435.

La periodicidad horaria con la que se calculan los precios de la electricidad puede introducir importantes distorsiones a la hora de elaborar un modelo, por cuanto que se captan datos extremos que afectan a la estimación de los parámetros del modelo de precios. Por esta razón, se han calculado los precios diarios realizando el promedio de los precios horarios en cada día del año, y serán sobre los que se ajuste el modelo de precios, previa realización de las correcciones que se detallan en la Sección 2. La serie total de precios tomada como base tiene como punto de partida el 1 de enero de 1999 y finaliza el 28 de febrero de 2010, representada en la Figura 1.

## **2 CORRECCIONES SOBRE LA SERIE**

Las correcciones realizadas en la elaboración del modelo de precios se han basado en las propuestas por Benth, Saltytè-Benth y Koekebakker (2008), aunque ha sido necesario realizar una selección de datos previa, en base a las particularidades del caso español (Cancelo, Espasa y Grafe, 2008). El modelo debe tener como premisa la presencia de una regularidad en el comportamiento de los precios, entendida ésta como la observación de las componentes clásicas de las series temporales. No obstante, a lo largo del tiempo de funcionamiento del mercado mayorista, ha habido distintos acontecimientos que han producido cambios en la evolución de los precios. El mercado ha pasado por una primera fase de “aprendizaje” (Franco, 2000) desde el año 1999 hasta aproximadamente 2002. A partir de entonces, presenta una evolución mucho más estable, aunque en el periodo 2002-2006 no se observan con claridad las componentes clásicas, en especial la estacionalidad, la principal determinante de los precios eléctricos según Goto y Karloyi (2004) y Platen, West y Breyman (2004).

Será a partir de 2006 cuando se pueda determinar con claridad la existencia de dichas componentes, debido fundamentalmente a la entrada en vigor del Real Decreto-Ley 3/2006 (Durban, 2009). Esta norma es la que establece el marco de

funcionamiento que determina el comportamiento de los precios en los años posteriores, en una clara diferenciación respecto a los anteriores, puesto que es la que de una forma más directa ha afectado a las decisiones que toman los agentes. Este hecho se explica porque los operadores del mercado mayorista adaptaron sus estrategias en función de esta norma con el objetivo de satisfacer las exigencias que en ella se incluían, la principal de las cuales era la obligación de adquirir energía en el mercado bilateral por parte de los distribuidores a las empresas filiales de un mismo grupo. Esto obligó a recurrir a repensar de un nuevo modo los riesgos y funciones entre las empresas de un mismo grupo energético, configurando un nuevo escenario que se mantiene hasta la fecha. Es por ello que la serie que se ha seleccionado para la modelización es la que **comienza en 1 de enero de 2006 y finaliza en 28 de febrero de 2010**. No se han tomado datos posteriores, porque si bien el modelo es aplicable en un rango de datos mayor, se adecuará en tanto no haya cambios relevantes en las fechas siguientes que perturben todo el ajuste realizado. Seguidamente se enumeran y describen las correcciones realizadas sobre la serie.

La primera corrección consiste en **eliminar los fines de semana y los festivos**. Las razones que justifican esta decisión son tanto el propio objetivo del modelo, recoger el comportamiento de los precios de los días laborables, como las importantes diferencias existentes entre los precios de los días laborables y de aquellos que se eliminan, cuyo comportamiento podría ser recogido con otro modelo, que no es objeto del presente trabajo. Con esta corrección, los años duran 250 días, porque en un año hay de media 104 días correspondientes a fines de semana y 12 u 11 días festivos dependiendo si el año es bisiesto o no, respectivamente.

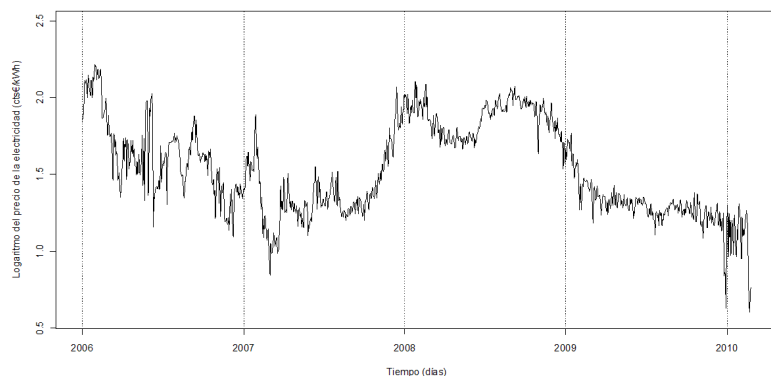
La segunda corrección que se ha realizado ha sido la de **tomar logaritmos naturales**. Esto es bastante usual al trabajar con series temporales puesto que, entre otros problemas, se elimina la heterocedasticidad y se aproxima la distribución de los datos a una Distribución Normal, aunque en esta ocasión no se consigue,

según el test Kolmogorov-Smirnov realizado con el comando *ks.test* con  $R$ , tomando el estadístico un valor de 0,002052, en términos de p-valor.

La última corrección se ocupa de **eliminar la presencia de los denominados *outliers***, datos extremos o picos. Estos picos pueden influir de manera determinante en el análisis de la tendencia y de la estacionalidad de las series temporales que definen los precios y que se desea modelizar, por lo que es conveniente eliminarlos antes de proceder a estimar los parámetros del modelo. Tras su eliminación, se debe sustituir el dato concreto por la media aritmética de los valores anterior y posterior.

El criterio que se utiliza de forma usual para determinar la existencia de *outliers* es considerar que una observación es un *outlier* si se queda fuera del intervalo determinado por  $[Q_1 - 1.5 \times IRQ, Q_3 + 1.5 \times IRQ]^2$ . En el caso concreto de la serie empleada, el rango resultante es  $[0, 573; 2, 488]$ , no habiéndose localizado ningún *outlier*.

La serie definitiva, una vez eliminados los fines de semana, festivos y *outliers* queda tal y como se representa en la Figura 2, y es la que servirá para ajustar el modelo propuesto.



**Figura 2:** Logaritmo de los precios corregidos (01/01/2006-28/02/2010)

<sup>2</sup>*IRQ* son las siglas de Rango Intercuartílico en inglés.

Del análisis de la serie corregida, se deduce la presencia de cinco componentes, cuyo comportamiento es el objeto de la modelización posterior: Tendencia, Ciclo, Estacionalidad Anual (periodo 1 año), Estacionalidad Semestral (periodo 6 meses) y la Componente Irregular. Como se ha señalado, las cuatro primeras componentes son deterministas y se modelizan combinando funciones lineales y trigonométricas, y la última se modeliza con técnicas ARIMA y procesos estocásticos.

### 3 EL MODELO DE PRECIOS. PARTE DETERMINISTA

Los pasos seguidos en la elaboración del modelo de precios diarios del mercado mayorista español de electricidad se han basado, con las adaptaciones pertinentes que recogen las características intrínsecas de dicho mercado, en la metodología desarrollada en Benth, Saltytè-Benth y Koekebakker (2008), y que ya iniciamos en un trabajo anterior (Cortés, Debón y Rico, 2010). El modelo propuesto es aditivo, y cada una de las componentes se ajusta una vez ha sido obtenida y descontada la anterior. El modelo completo se expresa en (1), donde se incluyen los elementos que se deben modelizar.

$$P(t) = P_{TC}(t) + P_{EA}(t) + P_{ES}(t) + I(t). \quad (1)$$

Se ha partido de un modelo para la tendencia  $T(t)$  y el ciclo  $C(t)$ , denotado como  $P_{TC}(t)$ , basado en una función lineal y una curva trigonométrica (vía la función coseno), respectivamente. Las razones que justifican esta decisión son la sencillez de la función coseno, su comportamiento periódico y su fácil interpretación, además de que la tendencia se modeliza de forma adecuada con una función lineal. La frecuencia seleccionada para el coseno ha sido de 750 días (3 años), ya que el ciclo observado en el período considerado tiene esta duración. La función a ajustar es



la que se muestra en (4), agregando las dos componentes detalladas en (2) y (3), correspondientes a las componente Tendencia y Ciclo respectivamente.

$$T(t) = b_0 + b_1 \cdot t, \quad (2)$$

$$C(t) = b_2 \cdot \cos \left[ \frac{2 \cdot \pi \cdot (t - b_3)}{750} \right], \quad (3)$$

$$P_{TC}(t) = b_0 + b_1 \cdot t + b_2 \cdot \cos \left[ \frac{2 \cdot \pi \cdot (t - b_3)}{750} \right]. \quad (4)$$

Para estimar el Modelo Tendencia-Ciclo, los parámetros que se deben calcular son  $b_0$ ,  $b_1$ ,  $b_2$  y  $b_3$ . Todos los parámetros se calculan mediante regresión no lineal sobre la serie de precios diarios corregida, realizando el ajuste con el programa *R* mediante la función *nls*. En primer lugar, se introducen todos los precios observados y corregidos de todos los efectos distorsionadores (fines de semana, festivos y *outliers*). En este punto se debe recordar el problema que presentan los algoritmos de ajuste de funciones no lineales, y es que los resultados dependen del punto inicial y, en ocasiones, los óptimos obtenidos no se puede garantizar que sean máximos globales.

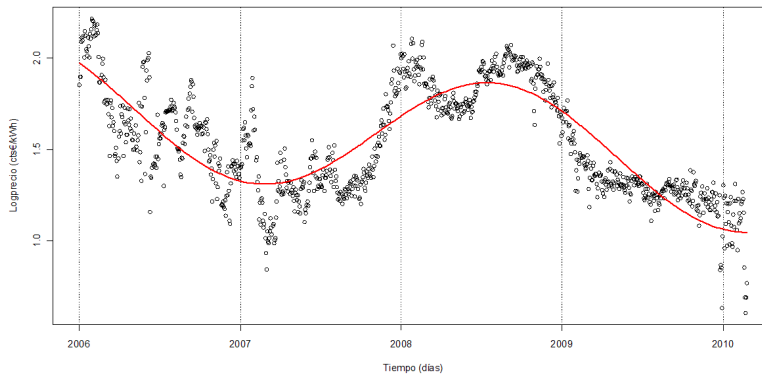
Por ello, el siguiente paso consiste en buscar un punto inicial para el ajuste del modelo dado en (4) a las observaciones de los precios corregidos. Los valores iniciales de  $b_0$  y  $b_1$  se obtienen a partir de la regresión lineal de  $P(t)$  sobre  $t$ . Posteriormente se ajusta el modelo partiendo del punto inicial con  $b_2 = 1$  para encontrar el valor inicial de  $b_2$ . Partiendo de este último punto se calcula el valor inicial para  $b_3$ . Finalmente se toma este punto como inicial del ajuste definitivo del Modelo (4) completo. El punto inicial obtenido es  $b_0 = 1,749$ ,  $b_1 = 0,000352$ ,  $b_2 = -0,3410$  y  $b_3 = -100,9$ .

Los valores de los parámetros obtenidos de este ajuste son los que se muestran en la Tabla 1. En ella se observa que todos los parámetros son significativamente

distintos de 0, a un nivel de significación del 5%. Como comentario adicional, cabe señalar la existencia de una tendencia decreciente con el valor negativo del parámetro  $b_1$ . El ajuste de esta componente se representa en la Figura 3.

Parámetro	Estimado	Error estándar	t-valor	p-valor
$b_0$	1,749	0,01039	168,44	< 2e-16
$b_1$	-0,000524	0,00001713	-20,57	< 2e-16
$b_2$	0,3410	0,0075	45,47	< 2e-16
$b_3$	-100,9	2,573	-39,23	< 2e-16

**Tabla 1:** Resultados del ajuste del modelo Tendencia-Ciclo



**Figura 3:** Ajuste del Modelo Tendencia-Ciclo a los precios corregidos

Como sucede con muchas otras variables económicas, el comportamiento de los precios de la electricidad depende del comportamiento que hayan tenido el mismo día en el año anterior, lo cual se modeliza con la introducción de un Modelo Estacional Anual, denotado por  $P_{EA}(t)$  en la Ecuación (5), para el cual se ha tomado una frecuencia de 250 días para el coseno. En (6) se expresa la operación que se realiza en este segundo paso, ajustándose el Modelo Estacional Anual sobre la serie de precios descontando la componente Tendencia-Ciclo modelizada anteriormente.

Antes de entrar en los detalles del cálculo, cabe señalar que se realizó una comprobación de la inexistencia de tendencia mediante la inclusión de un parámetro que dependía del tiempo el cual, al realizar el ajuste, resultó ser significativamente

igual a cero.

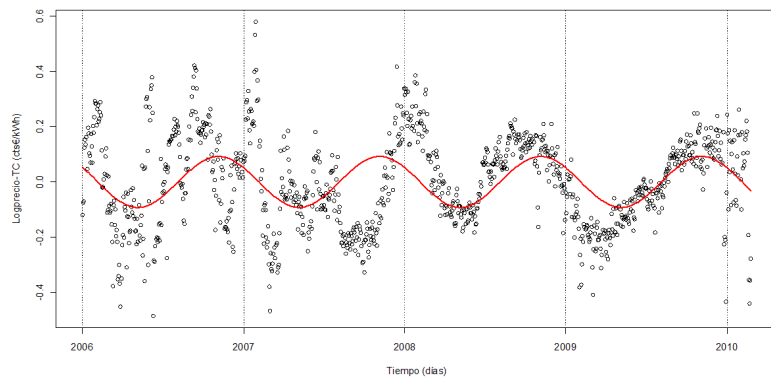
$$P_{EA}(t) = a_1 \cdot \cos \left[ \frac{2 \cdot \pi \cdot (t - a_2)}{250} \right], \quad (5)$$

$$P(t) - P_{TC}(t) = a_1 \cdot \cos \left[ \frac{2 \cdot \pi \cdot (t - a_2)}{250} \right]. \quad (6)$$

Los parámetros que se deben estimar son  $a_1$  y  $a_2$ , y al igual que en el modelo anterior, se debe realizar una primera estimación de los valores de dichos parámetros, fijando los valores iniciales  $a_1 = 1$  y  $a_2 = 0$ , siendo suficiente con una iteración para obtener los resultados de la Tabla 2, en la que se puede observar que ambos parámetros son significativamente distintos de cero a un nivel de significación del 5%. Los resultados gráficos de este ajuste se muestran en la Figura 4.

Parámetro	Estimado	Error estándar	t-valor	p-valor
$a_1$	0,091932	0,006763	13,59	< 2e-16
$a_2$	-38,115928	2,839507	-13,42	< 2e-16

**Tabla 2:** Resultados del ajuste del modelo Estacional



**Figura 4:** Ajuste del Modelo Estacional Anual

El siguiente paso, es el ajuste de la Estacionalidad Semestral. En este modelo

$P_{ES}(t)$  se ajusta nuevamente una función trigonométrica, un coseno, pero con la diferencia respecto a los anteriores en la frecuencia de la función, que en esta ocasión es de 125 días (seis meses), debido a que las oscilaciones interanuales en los precios son debidas a los patrones de consumo derivados de las estaciones climáticas, con precios altos en invierno y verano y bajos en otoño y primavera. El Modelo Estacional Semestral se expresa en (7) ajustándose sobre la serie de precios descontando las dos anteriores componentes, en (8).

$$P_{ES}(t) = c_1 \cdot \cos \left[ \frac{2 \cdot \pi \cdot (t - c_2)}{125} \right]. \quad (7)$$

$$P(t) - P_{TC}(t) - P_{EA}(t) = c_1 \cdot \cos \left[ \frac{2 \cdot \pi \cdot (t - c_2)}{125} \right] + I(t). \quad (8)$$

Los resultados obtenidos para el ajuste del Modelo Estacional Semestral son los que se muestran en la Tabla 3, habiéndose seguido los mismos pasos que en el Modelo Estacional Anual para estimar los parámetros, fijando los valores iniciales  $c_1 = 1$  y  $c_2 = 0$ . Con una iteración es suficiente para obtener una estimación de los parámetros con valores significativamente distintos de cero a un nivel de significación del 5%, representándose el ajuste en la Figura 5.

Parámetro	Estimado	Error estándar	t-Valor	p-Valor
$c_1$	0,051930	0,006431	8,075	1,86e-15
$c_2$	9,115837	2,486127	3,667	0,000258

**Tabla 3:** Resultados del ajuste del modelo Estacional Semestral

Con el ajuste de las componentes Tendencia-Ciclo, Estacionalidad Anual y Estacionalidad Semestral finaliza la parte determinista del modelo. Se probó también ajustar una estacionalidad semanal, de periodicidad 5 días, pero resultó ser no significativa, por lo que se excluyó del Modelo.

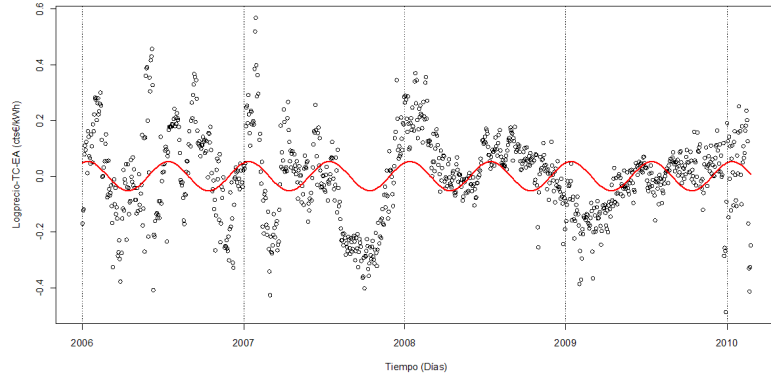


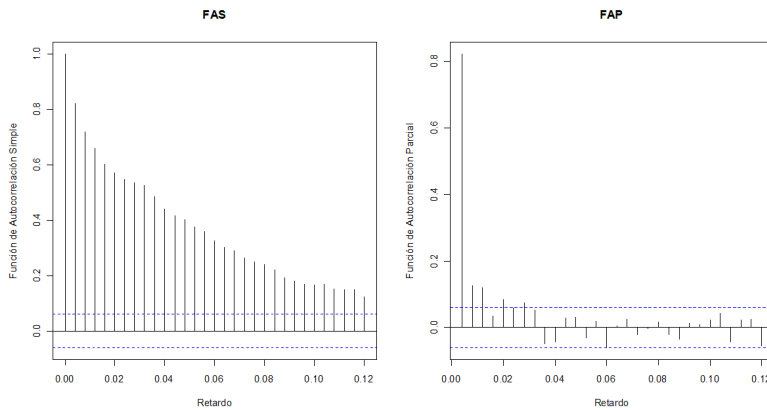
Figura 5: Ajuste del Modelo Estacional Semestral

## 4 PARTE ALEATORIA DEL MODELO

En esta sección se aborda la modelización de la Componente Irregular  $I(t)$ , que debería ser una serie estacionaria ya que, por definición, esta componente no obedece a ningún comportamiento identificable. Para comprobar la estacionariedad de la componente Irregular, se representan la FAS y la FAP de la serie descontando Tendencia, Ciclo y Estacionalidad en la Figura 6. En la FAS se produce un descenso exponencial de los coeficientes de autocorrelación, en el periodo objeto de estudio, de lo cual se puede deducir que se ha eliminado la tendencia. Por su parte, en la FAP se observa un coeficiente claramente significativo, el primero de ellos, y otros dos también significativos pero en menor medida. A medida que se alarga el retardo, hay ligeras oscilaciones de coeficientes positivos y negativos. Esto requiere proponer un nuevo submodelo para la componente Irregular, concretamente, un modelo ARIMA.

El ajuste se realiza con el paquete *forecast* del programa *R*, con la función *auto.arima*. Cabe hacer una corrección sobre el comando antes de ejecutarlo, y es que se debe definir la serie como estacionaria introduciendo el argumento *stationary=TRUE* en la función. Esta corrección se hace por dos razones: la primera, operativa, ya que cada vez que *R* debe estimar un modelo ARIMA completo, el tiempo

de cálculo es muy prolongado, puesto que considera todos los modelos posibles para la serie considerada; y la segunda, porque se ha comprobado, vía la estimación de un modelo ARIMA completo, que en la parte estacional no se ajusta ningún modelo. El modelo ARIMA ajustado con los comandos arriba descritos es el  $(1, 0, 2)$ , con los resultados de la Tabla 4, y que se procede a validar a continuación.



**Figura 6:** FAS y FAP de la componente irregular

	<b>Parámetro AR(1)</b>	<b>Parámetro MA(1)</b>	<b>Parámetro MA(2)</b>
Estimado	0,9354	-0,2508	-0,1587
Error estándar	0,0151	0,0356	0,0366
T-valor	61,9471	-7,0449	-4,3360
	ME = -0,0004067807		

**Tabla 4:** Resultados del ajuste del modelo ARIMA  $(1,0,2)$

En primer lugar, se debe comprobar que todos los parámetros son significativamente distintos de cero, confirmándose esta hipótesis con la observación de los t-valor del ajuste, cumpliéndose que todos ellos en valor absoluto son mayores a 1,658, el máximo para un intervalo de confianza del 95% en una distribución *t-Student*. Por su parte, del error se debe verificar que se cumplen las hipótesis media

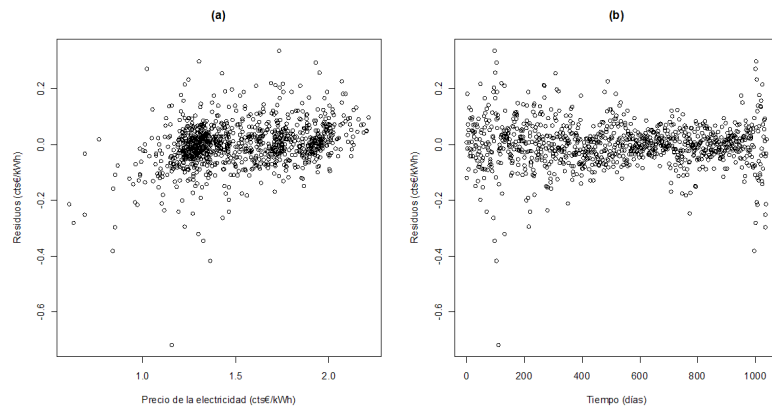
cero, varianza constante, incorrelación para cualquier retardo y distribución normal.

El contraste para aceptar que el valor de la media del error es cero se realiza mediante el intervalo (9), aceptándose la hipótesis nula si  $\varepsilon_t$  está entre el intervalo definido, rechazándose en caso contrario:

$$\left[ -z^{\alpha/2} \frac{\hat{\sigma}_\varepsilon}{\sqrt{T}}, z^{\alpha/2} \frac{\hat{\sigma}_\varepsilon}{\sqrt{T}} \right]. \quad (9)$$

El intervalo obtenido con los valores de la serie es  $[-0,00495; 0,00495]$ , entre el cual se encuentra el valor de la media del error (ME en la Tabla 4), no pudiendo rechazar la hipótesis nula, no pudiendo pensar que la media toma un valor significativamente distinto de cero.

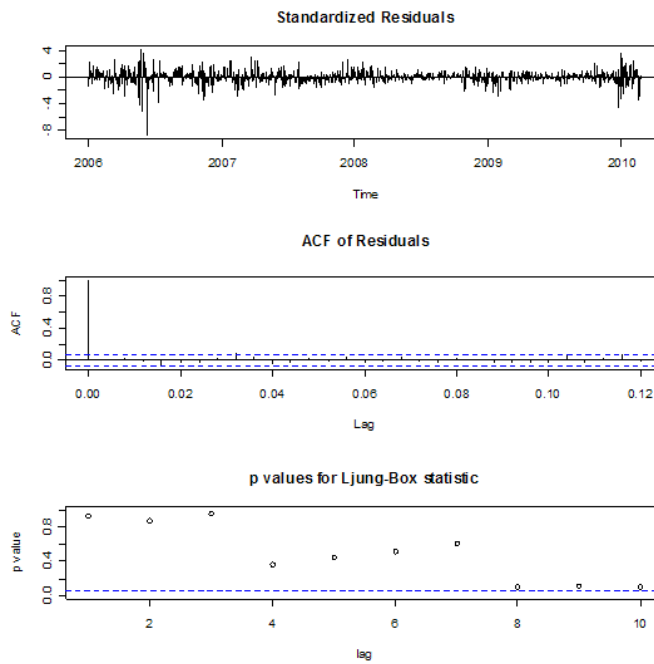
Para verificar la hipótesis de varianza constante, se representa el gráfico de los residuos frente al tiempo y frente al logaritmo del precio de la electricidad en la Figura 7, donde se aprecia a primera vista la existencia de heterocedasticidad, ya que la variabilidad de los residuos no se mantiene constante ni frente a la variable estudiada ni a lo largo del tiempo.



**Figura 7:** Residuos frente a la variable (a) y frente al tiempo (b)

La incorrelación para cualquier retardo se comprueba mediante la observación de la FAS de los residuos. Para ello, se ha representado la Figura 8, en la que

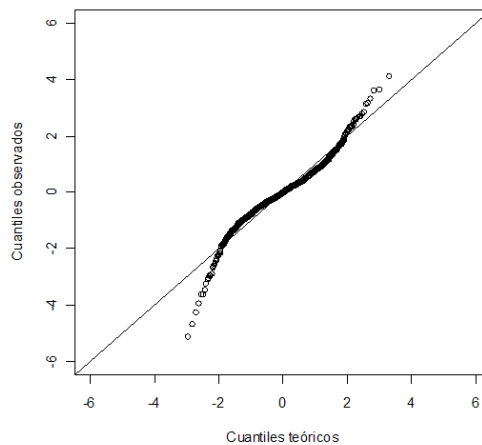
se muestra el análisis realizado con el comando *tsdiag* de *R*. En el primer gráfico se representan los residuos estandarizados de la serie ARIMA. En la FAS debe observarse que los residuos no tienen autocorrelación, como así sucede, pues solo el primero de los coeficientes de autocorrelación es 1, mostrando la relación del residuo consigo mismo. Finalmente, el test de Box-Pierce (*and Ljung-Box*) contrasta la hipótesis nula de la distribución aleatoria de los residuos. Esto proviene de la idea de que unos residuos de un modelo correctamente especificado se distribuyen independientemente. Como puede deducirse de la observación del tercer gráfico, el modelo puede admitirse como correctamente especificado pues los p-valor por encima del nivel de significación (5%) llevan a no poder rechazar esta hipótesis.



**Figura 8:** Resultados del análisis *tsdiag*



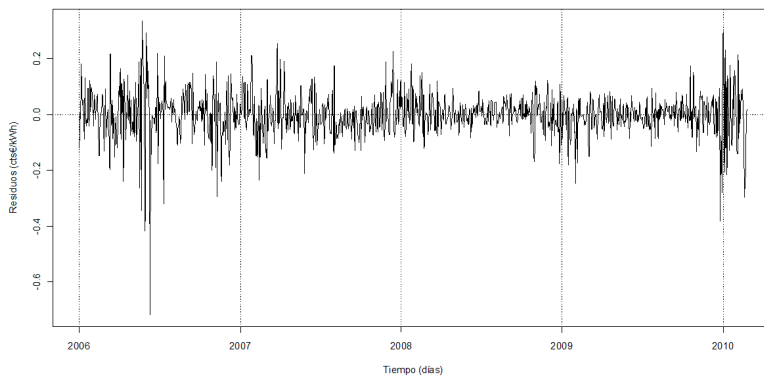
La última de las hipótesis que debe cumplir el modelo se comprueba mediante el Gráfico Q-Q (ver Figura 9), que muestra el grado en que los residuos del modelo obtenido se ajustan a una distribución normal. En la figura la mayoría de los residuos están alineados, aunque en los extremos inferior y superior hay un ligero alejamiento de la diagonal, que representa la normalidad perfecta. Se ha recurrido al test de Kolmogorov-Smirnov para comprobar la normalidad de los residuos y el p-valor obtenido en la prueba es de  $6,093e-06$ , rechazando de esta forma la hipótesis de normalidad.



**Figura 9:** Gráfico Q-Q de normalidad de los residuos estandarizados

Dado que se incumplen dos de las hipótesis que se establecen para la validación del modelo, la normalidad de los residuos y la inexistencia de heterocedasticidad, se ha realizado un análisis separado de los residuos para profundizar en la comprensión de su comportamiento y, por extensión, tratar de modelizarlos. Para ello, se ha representado la serie de los residuos, definidos como el logaritmo de los precios menos todas las componentes modelizadas en los anteriormente (Tendencia, Ciclo, Esta-

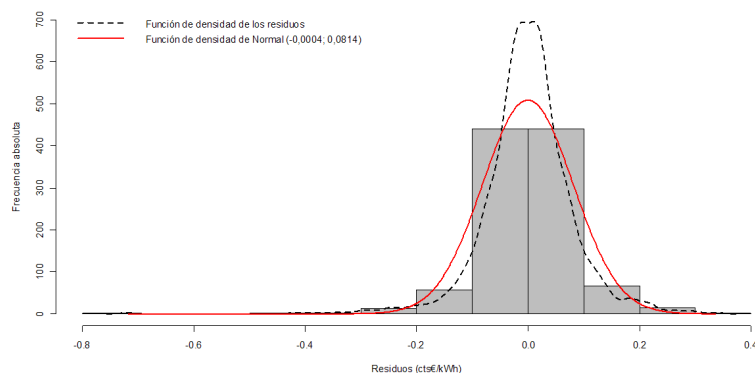
cionalidad y la parte ARIMA de la Componente Irregular), en la Figura 10, donde se observa que los mismos presentan fluctuaciones de distinta intensidad alrededor del cero, pero cada cierto tiempo aparecen saltos relativamente extremos. Este hecho motiva que se aíslen esos saltos extremos para modelizarlos por separado de los residuos.



**Figura 10:** Serie de residuos del modelo

Descriptivamente, los residuos no se aproximan a una distribución Normal, dado que, como se puede ver en la Figura 11, las observaciones de mayor frecuencia (alrededor de 500) son aquellas que están en torno al cero, pero hay un pequeño número de observaciones que toma valores extremos, principalmente valores negativos.

Para aislar los saltos extremos, se ha empleado la metodología señalada en Benth, Saltytè-Benth y Koekebakker (2008) y que se describe en detalle a continuación. El primer paso consiste en calcular la media y la desviación típica de los residuos con el objeto de establecer unos límites inferior y superior a partir de los cuales se considere que existe un salto. La media de los residuos es igual a  $\mu = -0,0004$  y su desviación típica toma un valor de  $\sigma = 0,0814$ . El límite



**Figura 11:** Histograma de los residuos

que se toma habitualmente viene determinado por la media más/menos dos veces la desviación típica. Este hecho tiene su explicación en que para una distribución Normal el 95% de los valores de la misma se sitúa entre dichos límites.

Cuando se han fijado los límites, se realiza el primer “filtrado”, considerando que existe un salto cuando un residuo particular sobrepasa los límites fijados. Si esto sucede, se sustituye el residuo por el límite correspondiente. Una vez se han localizado todos los saltos del primer filtrado y han sido sustituidos tal como se ha descrito, se cuenta el número de saltos y se vuelve a calcular la media y la desviación típica de los residuos para formar unos nuevos límites y volver a contar y sustituir los saltos. Este procedimiento se realiza cuantas veces sea necesario hasta que el número de saltos no aumente. El resumen de este proceso se muestra en la Tabla 5.

Sobre los resultados del filtrado cabe hacer una serie de comentarios. El número total de iteraciones necesarias para que no se detectaran más saltos fue de 8. Nótese cómo la desviación típica disminuye con cada una de las iteraciones, algo lógico por otra parte, dado que la mecánica de este filtrado consiste en la eliminación de los saltos extremos. En la columna “Saltos acumulados” se ha ido sumando el número de saltos que se producían en los pasos anteriores ya que, si existía un salto en un

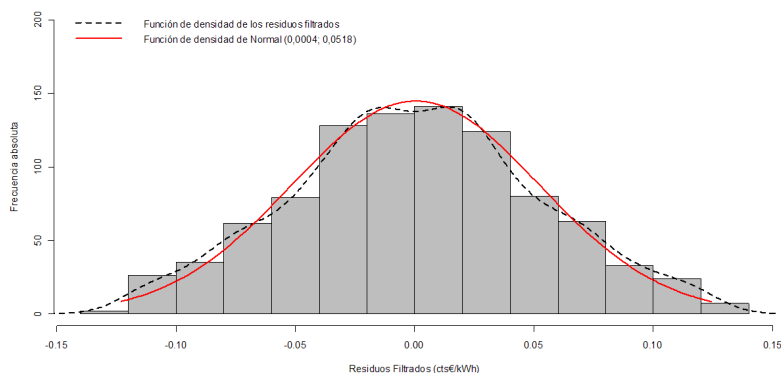
paso previo, también existirá en un paso posterior, en el que los límites son más estrechos por la disminución progresiva de la desviación típica. Por último, se ha calculado la frecuencia de saltos diaria como el cociente entre el número de saltos acumulado y el número de datos totales, que son 1.038.

<b>Iteración</b>	<b>Desviación típica</b>	<b>Salto acumulados</b>	<b>Frecuencia de saltos diaria</b>
1	0,0814	56	0,05395
2	0,0682	81	0,07803
3	0,0645	88	0,08478
4	0,0632	94	0,09056
5	0,0628	96	0,09249
6	0,0626	97	0,09345
7	0,0625	98	0,09441
8	0,0625	98	0,09441

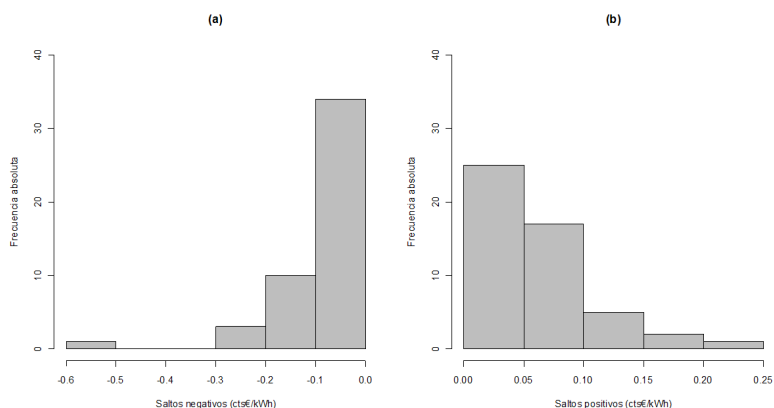
**Tabla 5:** Resumen del proceso de filtrado de los residuos

Tras realizar este procedimiento, se toman los residuos en los que se han localizado saltos y se separan de la serie de residuos. A esta nueva serie sin los saltos se le ha denominado “residuos filtrados”, cuyo histograma se representa en la Figura 12, donde se puede observar que los residuos filtrados siguen una distribución Normal  $N(0,0004; 0,0518)$ , extremo éste que se confirma tras realizar el test de Kolmogorov-Smirnov, para el cual se obtiene un p-valor de 0,8904. Asimismo, los valores de los coeficientes de curtosis y asimetría se encuentran entre los esperados para una distribución Normal, siendo -0,3619 y 0,0336, respectivamente.

Los saltos no siguen una distribución Normal, según se ha contrastado mediante el test Kolmogorov-Smirnov, obteniéndose un p-valor de 0,02213, menor que el nivel de significación del 5%. El tratamiento que se propone a continuación es el de dividir los saltos en valores positivos y valores negativos, para los cuales se ha representado su histograma en la Figura 13, y modelizarlos como se detalla a continuación.



**Figura 12:** Histograma de los residuos filtrados



**Figura 13:** Histograma de los saltos negativos (a) y positivos (b)

Para modelizar los saltos, se ha propuesto un modelo  $S(t)$  definido mediante la suma de dos procesos de Poisson compuestos, los cuales describen por separado los saltos positivos y negativos. Se define el proceso  $S(t)$  como se expresa en (10):

$$S(t) = S^+(t) + S^-(t), \tag{10}$$

donde  $S^\pm(t)$  se definen tal y como se muestra en (11):

$$S^\pm(t) = \sum_{i=1}^{N^\pm(t)} J_i^\pm, \quad (11)$$

siendo  $N^+(t)$  y  $N^-(t)$  procesos de Poisson cuyas intensidades son  $\lambda^+$  y  $\lambda^-$ , respectivamente. En (11),  $J_i^\pm$  son dos secuencias de variables aleatorias independientes e idénticamente distribuidas (i.i.d.) con las que se modeliza el tamaño de los saltos. Los pasos que se han dado para simular el proceso  $S(t)$  se detallan a continuación.

De los resultados del filtrado realizado sobre los residuos, se obtienen las intensidades  $\lambda^+$  y  $\lambda^-$ , las cuales se calculan como el número de saltos positivos (50) y negativos (48), respectivamente, divididos entre el número de datos totales. Los valores calculados son,

$$\lambda^+ = 0,04816956, \quad \lambda^- = 0,04624277,$$

respectivamente. Como se puede ver, los saltos positivos (incrementos repentinos del precio de la electricidad) son ligeramente más probables y más homogéneos que los saltos negativos, lo que justifica su elección para el estudio de la distribución de los saltos.

En primer lugar, se debe definir la función de distribución más adecuada para los saltos. La observación del histograma de los saltos positivos (véase Figura 13, gráfico (b)) sugiere el empleo, o bien de una distribución lognormal, o bien de una distribución exponencial para modelizar los tamaños de los saltos (obviamente utilizando la misma distribución para generar la longitud de saltos negativos y considerando su signo con posterioridad). Se ha probado con ambas posibilidades y por la mejora en la modelización se ha optado por el modelo exponencial, cuya función de densidad de probabilidad es de la forma dada en (12):

$$f_{Exp}(z) = \frac{1}{\mu_J} \exp(-z/\mu_J), \quad (12)$$

donde el parámetro  $\mu_J$  es el tamaño medio del salto. La estimación del parámetro  $\mu_J$  para la distribución de los saltos tanto positivos como negativos se ha hecho utilizando el enfoque de máxima verosimilitud mediante la muestra de saltos positivos, obteniéndose el valor  $\hat{\mu}_J = 0,05668014$ . A partir de la distribución exponencial cuyo parámetro ha sido estimado se genera la longitud los saltos  $J_i^+$  y  $J_i^-$  de la expresión (10), teniendo en cuenta que en el caso de los negativos han de considerarse con dicho signo.

Por lo tanto se tiene una especificación completa del proceso estocástico que gobierna la dinámica del precio diario de la electricidad del mercado mayorista español en el periodo considerado.

Finalmente, al agregar las componentes modelizadas y calcular las medidas de bondad del ajuste, se obtienen unos valores del MPE y el MAPE de 1,20% y 4,36%, respectivamente. Al deshacer la transformación logarítmica, el valor del MPE resulta en 1,57%, similar al del modelo logarítmico, y algo mayor para el MAPE, de 8,80%. Este hecho es totalmente habitual, ya que al deshacer la transformación logarítmica aumenta la variabilidad de la serie. En ambos casos, no obstante, se trata de cifras aceptables teniendo en cuenta las amplias oscilaciones que presentan los precios de la electricidad, como ocurre en Darbelay y Slama (2000) y Smith (2000).

## 5 CONCLUSIONES

En el presente trabajo se ha construido un modelo de precios para el mercado mayorista de la electricidad, con la combinación de métodos analíticos y estadísticos, recogiendo los principales patrones que determinan la dinámica de su comportamiento, partiendo de las componentes Tendencia, Ciclo, Estacionalidad y Componente Irregular. La última de las componentes, modelizada con un ARIMA, una variable aleatoria normal y un proceso estocástico de Poisson compuesto, tiene

un grado de explicación mayor sobre los precios, muestra del importante comportamiento aleatorio de los precios. Las medidas de bondad del ajuste obtenidas son aceptables dada la naturaleza de la variable económica estudiada.

En el futuro se abordará la inclusión de los precios excluidos del modelo, aquellos que corresponden a fines de semana y festivos, que presentaban un comportamiento claramente distinto a los días laborables. Para ello, sería necesaria la elaboración de un nuevo modelo que recogiera sus características intrínsecas, para agregarlos posteriormente al modelo aquí propuesto. Respecto de los procesos estocásticos, se estudiará la aplicación de la distribución Normal Inversa Gaussiana (NIG), empleada en trabajos similares.

## 6 REFERENCIAS BIBLIOGRÁFICAS

- BENTH, F.E., SALTYTÈ-BENTH, J.S. y KOEKEBAKKER, S. (2008). “Stochastic Modelling of Electricity and Related Markets”. World Scientific.
- CANCELO, J-R., ESPASA, A y GRAFE, R. (2008). “Forecasting the electricity load from one day to one week ahead for the Spanish system operator”. *International Journal of Forecasting*, 24, 588-602.
- CORTÉS LÓPEZ, J-C., DEBÓN AUCEJO, A. y RICO LLOPIS, M. (2010). “Estudio estadístico de los precios de la electricidad del mercado mayorista español”. XVIII Jornadas ASEPUMA - VI Encuentro Internacional. *Anales de ASEPUMA*, nº 18: 203.
- DARBELAY, G.A. y SLAMA, M. (2000). “Forecasting the short-term demand for electricity. Do neutral networks stand a better chance?”. *International Journal of Forecasting*, 16, 71-8722; 83.



- DURBAN, R. (2009). “El sector, la regulación y el suministro de electricidad. Descripción del sistema eléctrico español”. Comisión Nacional de la Energía.
- FRANCO UGIDOS, P. (2000). “Análisis y Desarrollo de productos financieros derivados asociados a Índices del Mercado Mayorista español de energía eléctrica”. Trabajo Final de Carrera. Escuela Técnica Superior de Ingeniería, Universidad Pontificia de Comillas.
- GOTO, M. y KAROLYI, G.A. (2004). “Understanding electricity price volatility within and accross markets”. Dice Center Working Papers, 2004-12.
- LUCÍA, J. y MENEU, V. (2004) “Los precios en los mercados reestructurados de electricidad: algunas lecciones básicas para la negociación derivada”. Departamento de Economía Financiera, Universidad de Valencia.
- PLATEN, E., WEST, J. y BREYMAN, W. (2004). “Intraday empirical analysis of electricity price behaviour”. Quantitative Finance Research Centre, University of Technology of Sydney, 140, noviembre 2004.
- SMITH, M. (2000). “Modelling and short-term forecasting of New South Wales electricity system load”. Journal of Business and Economic Statistics, 18, 465-8722; 478.