

UNIVERSIDAD PONTIFICIA COMILLAS
ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA (ICAI)
Instituto de Investigación Tecnológica

CONTRIBUCIONES A LA PREDICCIÓN A CORTO PLAZO DEL PRECIO DE LA ELECTRICIDAD MEDIANTE MODELOS DE SERIES TEMPORALES

Tesis para la obtención del grado de Doctor

Director: Dr. D. Antonio Muñoz San Roque.

Autor: Ing. D. Alberto Miguel Cruz García.



Madrid 2013

CONSTANCIA REGISTRAL DEL TRIBUNAL DEL ACTO DE LA DEFENSA DE TESIS DOCTORAL

TÍTULO: CONTRIBUCIONES A LA PREDICCIÓN A CORTO PLAZO DEL
PRECIO DE LA ELECTRICIDAD MEDIANTE MODELOS DE SERIES
TEMPORALES

AUTOR: ING. D. ALBERTO MIGUEL CRUZ GARCÍA

DIRECTOR: DR. D. ANTONIO MUÑOZ SAN ROQUE

TUTOR-PONENTE:

DEPARTAMENTO: INSTITUTO DE INVESTIGACIÓN TECNOLÓGICA

FACULTAD O ESCUELA: ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA

Miembros del Tribunal Calificador:

PRESIDENTE:

Firma:

VOCAL:

Firma:

VOCAL:

Firma:

VOCAL:

Firma:

SECRETARIO:

Firma:

Fecha de lectura:

Calificación:

*A mis padres,
vivo para que os sintáis orgullosos.*

Agradecimientos

Quiero recordar en estas líneas a las personas que han hecho posible que pueda terminar esta tesis doctoral.

A Antonio Muñoz, que confiara en mí para realizar una tesis bajo su dirección. Por su soporte económico durante la mayoría de los años que pasé en el IIT, por su acertada orientación y entera disposición.

A Miguel Ángel Sanz, mi primer supervisor en el IIT, por la oportunidad que me brindó al contratarme, hecho que ha cambiado radical y positivamente el rumbo de mi carrera.

A Juan Luis Zamora, por su ayuda prestada con los modelos de espacio de estados y el temible filtro de Kalman.

A Iskander Karibzhanov, al que no tengo el gusto de conocer en persona, por su desinteresada ayuda con las rutinas de optimización usadas para la tesis.

A Michel Rivier, antiguo director del IIT, por su apoyo en algunos momentos difíciles que no quiero olvidar.

A mis compañeros del IIT, por sus sabios consejos que me han hecho crecer como persona, por su aliento en los malos momentos y por los inolvidables momentos vividos.

A mis padres, hermanos, tíos, primos y sobrinos, que caminan a mi lado cada día. A los amigos que considero una extensión de mi familia. A ti, Mirel, por dar conmigo el último empujón a esta tesis. Vosotros habéis sido mi verdadera motivación.

Índice general

1. Introducción, motivación y objetivos	1
1.1. Introducción	2
1.2. Funcionamiento del mercado ibérico de electricidad	2
1.3. Formación del precio en el mercado diario	7
1.4. Motivación de la tesis	9
1.5. Organización del documento	10
2. Modelos de predicción de precios spot en mercados eléctricos	13
2.1. Clasificación general de los modelos de precios eléctricos	14
2.1.1. Modelos para el medio y largo plazo	14
2.1.2. Modelos para la valoración de productos derivados	16
2.2. Modelos de predicción a corto plazo	18
2.2.1. Características de las series de precios	19
2.2.2. Modelos de series temporales	28
2.2.2.1. Modelos ARIMA	28
2.2.2.2. Modelo de alisado exponencial con doble estacionalidad	31
2.2.2.3. Modelos de función de transferencia	32
2.2.2.4. Modelos vectoriales	33
2.2.2.5. Modelos de heteroscedasticidad condicional GARCH .	36
2.2.2.6. Modelos periódicos	38
2.2.2.7. Modelos por umbrales	40
2.2.2.8. Modelos de cambio de régimen	41
2.2.2.9. Modelos semiparamétricos	44
2.2.3. Modelos basados en redes neuronales	46
2.2.4. Modelos basados en técnicas de minería de datos	49
2.3. Conclusiones	50
3. Representación no paramétrica de la distribución de probabilidad de los precios de la electricidad	53
3.1. Introducción	54
3.2. Aproximación de funciones de densidad mediante series de Hermite . .	57
3.3. Modelo de función de transferencia con ruido ARIMA-ARCH estacional	64
3.3.1. Estimación de los parámetros del modelo	66
3.3.2. Diagnóstico y selección del modelo	71

3.4.	Aplicación a los precios horarios en mercados eléctricos	75
3.5.	Conclusiones	92
3.6.	Apéndice A: Polinomios de Chebyshev-Hermite	94
3.7.	Apéndice B: Propiedades de la familia de densidades Edgeworth-Sargan	97
4.	Modelo de cambios de régimen para la predicción de los precios de la electricidad	101
4.1.	Introducción	102
4.2.	Modelos en Espacio de Estados	104
4.2.1.	Representación en Espacio de Estados de un sistema dinámico	105
4.2.1.1.	Representación del modelo de función de transferencia con ruido ARIMA	106
4.2.1.2.	Resumen del filtro de Kalman	108
4.2.2.	Representación en espacio de estados del modelo de función de transferencia con ruido ARIMA-ARCH estacional	110
4.2.3.	Modelo de espacio de estados periódico	112
4.2.4.	Modelo de espacio de estados con cambio de régimen	114
4.2.4.1.	Estimación de los parámetros del modelo	118
4.2.4.2.	Diagnos y selección del modelo	122
4.3.	Aplicación a los precios medios diarios del mercado Ibérico	124
4.4.	Aplicación a los precios horarios del mercado Ibérico	136
4.5.	Conclusiones	144
5.	Aplicación a la predicción de los precios spot del día siguiente en el mercado ibérico	147
5.1.	Introducción	148
5.2.	Predicción multiperíodo	149
5.2.1.	Predicción con el modelo ARMA-ARCH	150
5.2.2.	Predicción con el modelo semiparamétrico de función de transferencia con ruido ARIMA-ARCH	152
5.2.3.	Predicción con el modelo de espacio de estados con cambio de régimen	154
5.3.	Medidas para evaluar la calidad de las predicciones	156
5.3.1.	Medidas de error para las predicciones puntuales	156
5.3.2.	Medidas para evaluar las distribuciones predictivas	157
5.4.	Análisis comparativo de modelos	158
5.4.1.	Predicción puntual	162
5.4.2.	Predicción de la función de densidad	169
5.5.	Conclusiones	189
6.	Conclusiones, aportaciones originales y futuros desarrollos	191
6.1.	Resumen y conclusiones de la tesis	192
6.1.1.	Aportaciones originales	194
6.1.2.	Futuros desarrollos	197

Capítulo 1

Introducción, motivación y objetivos

Uno de los elementos clave para el funcionamiento de la mayoría de los mercados eléctricos liberalizados es una subasta diaria donde las empresas participantes son convocadas para que envíen sus ofertas de compra y venta de energía. De la casación de estas ofertas se obtiene tanto la programación de unidades del sistema que cubrirán la demanda de energía como el precio marginal, para cada una de las 24 horas del día siguiente. Resulta crucial, por tanto, para las empresas participantes anticiparse a los movimientos del precio marginal para poder optimizar sus estrategias de oferta o protegerse frente al riesgo de grandes desviaciones. La caracterización y predicción de dicho precio marginal constituye la principal motivación de esta tesis.

En este primer capítulo se introduce el contexto en el que se desarrolla la tesis. Primero se explica el funcionamiento general de los mercados eléctricos liberalizados tomando como ejemplo el mercado ibérico de electricidad. El capítulo se centra en el proceso de formación del precio marginal y los factores por los que se ve afectado. Finalmente, se explica la motivación y los objetivos de la tesis.

1.1. Introducción

Las leyes físicas que determinan la conducción de energía a través de una red de transporte requieren mantener el equilibrio entre la inyección de energía en los puntos de generación y la extracción en los puntos de demanda (con algún margen por pérdidas en la transmisión). En todo momento la producción y el consumo deben de estar perfectamente sincronizados dado que no hay capacidad de almacenamiento. Así, la tarea del operador del sistema eléctrico es realizar el seguimiento de la demanda de electricidad y asegurar en todo momento su perfecta sincronización con las unidades de generación, requiriendo en tiempo real el funcionamiento de aquellos generadores que tengan la capacidad técnica de responder a las variaciones de la demanda y garantizando que ésta cumple con los requisitos de calidad de suministro.

El hecho de que la electricidad no sea almacenable es lo que diferencia un mercado eléctrico de otros mercados en donde se negocian por ejemplo activos financieros o materias primas. Puesto que la electricidad es un bien de consumo esencial, no existe apenas elasticidad de la demanda¹ con respecto al precio del mercado. La demanda responde fundamentalmente a factores de laboralidad y temperatura. De este modo, en un periodo de tiempo determinado (día, semana o año), se necesita que entren en funcionamiento un número variable de plantas de generación con diferentes costes de producción capaces de satisfacer los requerimientos instantáneos de electricidad. Esta variabilidad en los costes de las tecnologías de generación según los diferentes niveles de demanda, unido a otros factores como la naturaleza oligopolística de estos mercados, es lo que determina que los precios presenten una serie de características únicas y universales que serán expuestas en detalle a lo largo del documento.

Con el fin de comprender mejor la idiosincrasia de los mercados eléctricos, en la sección siguiente se describe brevemente el funcionamiento del mercado ibérico de electricidad. Esto permite asimismo establecer el marco de trabajo en el que se desarrolla la tesis, que tiene como objetivo la predicción de los precios del mercado diario de electricidad en un horizonte de corto plazo. Una vez explicado el proceso de formación de los precios eléctricos en el mercado, se expone la motivación de la tesis y la organización del resto del documento.

1.2. Funcionamiento del mercado ibérico de electricidad

El mercado ibérico de electricidad es un mercado mayorista en el que los agentes productores de electricidad venden la electricidad a distribuidores, comercializadores y grandes consumidores. Dentro de este mercado se incluye tanto la energía que se vende a través de contratos bilaterales físicos como la energía que se vende en el mercado organizado o *pool*, donde se realiza la mayor parte de las transacciones. El mercado organizado o *pool* se estructura sobre la base de una serie de mercados sucesivos en los que se va ajustando la oferta y la demanda de electricidad. El más

¹Se entiende por elasticidad de la demanda a su sensibilidad a variaciones en el precio

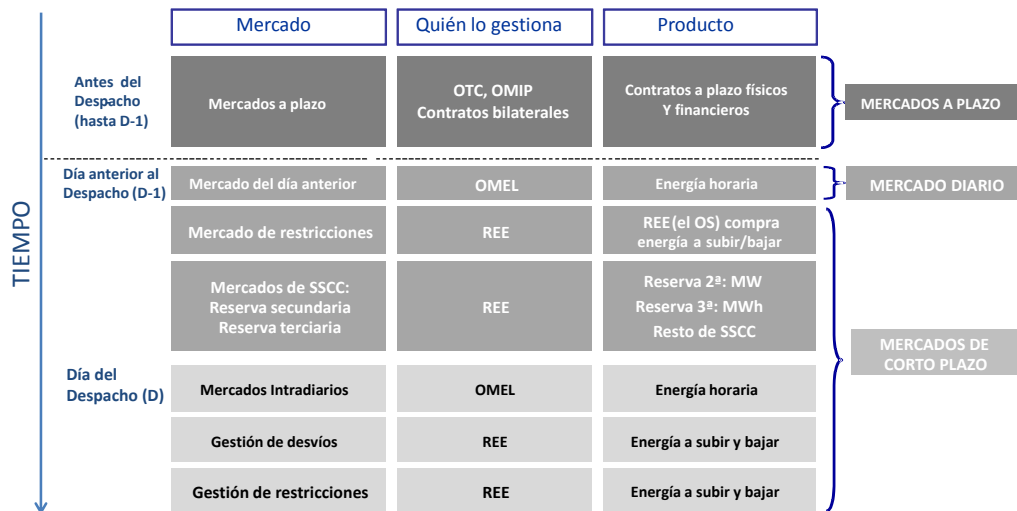


Figura 1.1: Secuencia de mercados en el mercado ibérico de electricidad.

importante de estos mercados es el mercado *diario*, en el que se negocia la energía para cada una de las 24 horas del día siguiente. Con posterioridad a la casación en el mercado *diario* las empresas oferentes y demandantes de energía pueden ajustar sus posiciones en los mercados *intradiarios*. Gracias a estos mercados, y a los procesos de resolución de restricciones técnicas, servicios complementarios y gestión de desvíos, el sistema puede acomodar las inevitables discrepancias entre la demanda real y la prevista y los generadores pueden optimizar sus programas de operación dadas las características técnicas de sus plantas.

La secuencia de mercados de generación responde al gráfico de la figura 1.1. Si atendemos al horizonte temporal en el que se negocia la energía, se distinguen tres etapas diferentes:

- En los mercados a plazo de electricidad se negocian contratos de compra-venta de electricidad con plazos de entrega superiores a 24 horas (semanas, meses, trimestres, años, antes de la entrega física de la energía).
- El día anterior a la entrega física de la energía (día D-1), los agentes del mercado negocian mediante subastas la electricidad que será consumida el día D. Estas

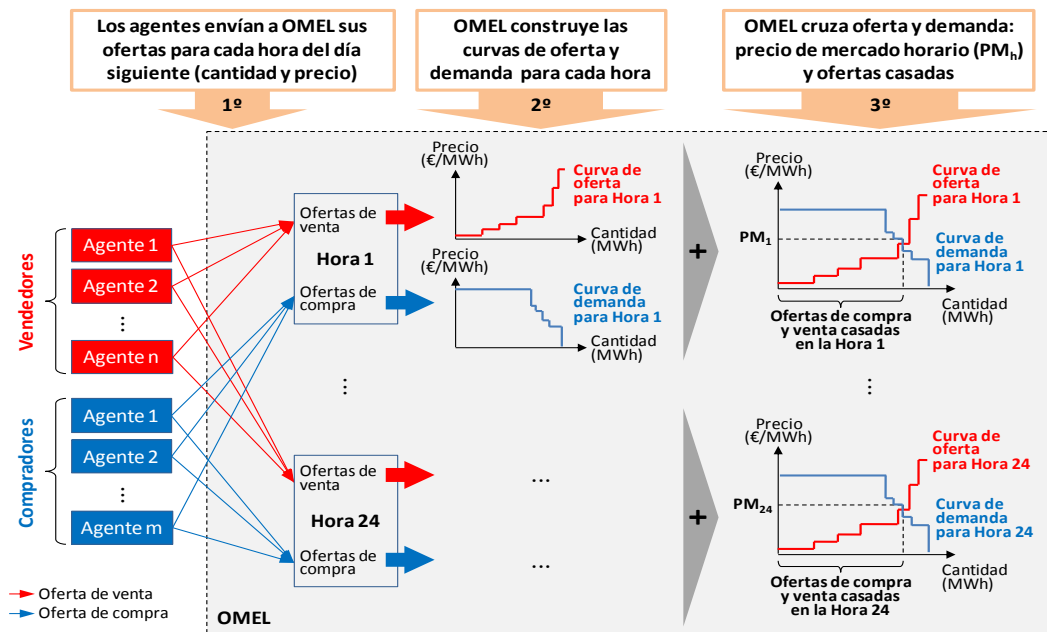


Figura 1.2: Esquema del proceso de casación de ofertas en el mercado diario. Fuente: Energía y sociedad www.energiaysociedad.es.

transacciones se llevan a cabo en el mercado diario.

- En el muy corto plazo (dentro de las 24 horas anteriores al momento de la entrega física de la energía), existen otros mercados en los que los agentes pueden ajustar sus posiciones contractuales y en los que se ofrecen una serie de servicios de gestión técnica al sistema.

Mercados a plazo. Los contratos que los agentes se intercambian con períodos de entrega superiores a un día pueden ser de dos tipos: 1) contratos bilaterales adaptados a las necesidades de los agentes que son acordados por la negociación bilateral entre las dos partes; y 2) productos estandarizados que se contratan a través de mercados organizados (en el mercado ibérico, OMIP: (bolsa de derivados del MIBEL) y CESUR (Subastas de energía para el Suministro de Último Recurso)) u OTC (acrónimo de *Over The Counter*). Este tipo de contratos cumplen un papel crucial en un mercado liberalizado desarrollado, puesto que permiten a los agentes compradores y vendedores gestionar eficazmente sus riesgos. Sin embargo, para el

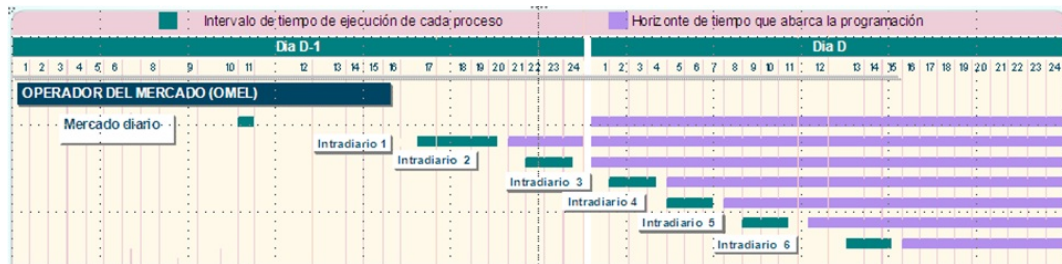


Figura 1.3: Períodos de ejecución de la secuencia de mercados intradiarios y sus horizontes de aplicación. Fuente: Red eléctrica www.ree.es.

caso del mercado ibérico estos mercados no están suficientemente desarrollados ni son lo suficientemente líquidos para permitir dicha gestión.

Para un momento dado, los precios a los que se está intercambiando electricidad a los distintos plazos futuros se conoce como *curva forward* (o *curva a plazo*). La curva forward cambia al mismo ritmo que cambian las expectativas sobre el precio del mercado diario y los factores por los que se ve afectado.

Mercado diario. El mercado del diario está organizado por el operador del mercado (OMEL), entidad privada encargada de llevar a cabo la gestión del mercado y garantizar que la contratación en el mismo se realiza en condiciones de transparencia, objetividad e independencia. Se celebra el día anterior al de la entrega de la energía y en él compradores y vendedores realizan sus ofertas de compra-venta de electricidad para cada una de las horas del día siguiente. Así, en este mercado en realidad hay 24 productos diferentes (la energía en cada una de las 24 horas del día siguiente). El operador del mercado es el encargado de realizar diariamente la casación de ofertas cruzando las curvas agregadas de compra y venta, obteniendo así la programación de unidades de generación que cubrirán la demanda prevista de electricidad y su precio, para cada una de las 24 horas del día. El proceso de casación se muestra esquematizado en la figura 1.2, y será explicado con algo más de detalle en la sección 1.3. Puesto que la predicción de dicho precio de casación constituye la principal motivación de la tesis, serán expuestos también los factores que determinan la forma de las curvas agregadas de oferta y demanda y que determinan, por tanto, el precio resultante.

Mercados intradiarios. Dentro de las 24 horas anteriores al momento del despacho de la energía, generadores y comercializadores pueden ajustar sus posiciones comprando y vendiendo energía en los seis mercados intradiarios actualmente existentes, también gestionados por OMEL, y cuyo funcionamiento es similar al del mercado diario. Los periodos de ejecución de estos mercados y su secuencia, así como los horizontes a los que aplican sus resultados se muestran la figura 1.3.

Mercados de servicios de ajuste del sistema. Bajo estos mercados se agrupan un conjunto de mecanismos de carácter competitivo gestionados por el operador del sistema. Los servicios de ajuste comprenden las restricciones técnicas, los servicios complementarios y la gestión de desvíos generación-consumo, todos ellos necesarios para mantener el sistema eléctrico en equilibrio físico y dentro de un nivel de seguridad adecuado. Sus horizontes de aplicación abarcan desde unas pocas horas hasta pocos minutos antes del despacho. Estos mercados son los siguientes:

- *Mercado de restricciones técnicas:* Una vez realizada la casación del mercado diario y teniendo en cuenta los contratos bilaterales, el OS realiza un análisis para verificar la viabilidad del programa de generación y consumo resultante. En caso de que dicho programa no sea técnicamente factible, el OS resuelve las congestiones de la red alterando el programa de generación aplicando criterios técnicos de seguridad y económicos, es decir, considerando las ofertas a subir o bajar energía enviadas por los generadores al mercado de restricciones.
- *Mercado de regulación Secundaria:* El servicio de banda de regulación permite al OS disponer de una reserva de capacidad disponible que permita resolver desequilibrios significativos entre generación y demanda. Las unidades de generación de los participantes en este mercado deben poder responder en unos pocos segundos a requerimiento automático del programa de control de la generación del OS.

El mercado se convoca después de la celebración del mercado diario y del de restricciones. Cada día, el OS estima y publica la reserva de banda de regulación (en términos de potencia) necesaria para asegurar el suministro en condiciones de fiabilidad en caso de desequilibrios producción/consumo que puedan producirse en tiempo real. A partir de dicha estimación y con carácter voluntario, las empresas generadoras con capacidad de regulación presentan sus ofertas de capacidad disponible. El coste marginal de la banda de potencia para cada hora marca el precio con el que se remunera toda la capacidad asignada en este mercado. Así mismo, el servicio remunera no sólo la banda de potencia, sino también la energía eventualmente utilizada, valorada al precio de sustitución de la energía terciaria.

- *Mercado de regulación terciaria:* El mercado de energía terciaria se celebra a última hora del día anterior al despacho, y tiene como objetivo restituir la reserva de banda en caso de que se haga uso de la banda secundaria por causa de una contingencia. En este mercado, los generadores envían ofertas por la variación máxima de su potencia a subir y a bajar. El precio de la energía terciaria utilizada a subir o a bajar es el precio marginal resultante de las ofertas realizadas por los generadores frente a una demanda establecida por el OS según sus requerimientos. Al contrario que en el caso de la reserva secundaria, los generadores sólo perciben ingresos por este servicio si es utilizado por el OS.
- *Mercados de desvíos:* El objetivo de estos mercados es resolver los desequilibrios entre la oferta y la demanda que sean identificados unas pocas horas antes

del despacho, tras la celebración de cada mercado intradiario. Una vez los agentes productores hayan comunicado sus previsiones de desvíos y junto con las previsiones de generación de energía eólica, el OS convoca el mercado de gestión de desvíos en caso de que el conjunto de desvíos previstos supere una cierta cantidad. Así, el mercado solicitará ofertas a los generadores en el sentido contrario al desvío previsto.

1.3. Formación del precio en el mercado diario

El precio del mercado diario viene dado por la intersección de la curva de oferta (construida por agregación de todas las ofertas de venta enviadas al mercado) y la de demanda (construida por agregación de todas las ofertas de compra). En un modelo de subasta de *casación simple* (*uniform-price*, como es el caso del mercado ibérico), todos los agentes de compra con ofertas por encima o igual a dicho precio pagarán ese precio por la electricidad suministrada, y los agentes de venta con ofertas por debajo o igual serán remunerados a ese mismo precio, independientemente de cuál fuese el precio establecido en sus ofertas (ver figura 1.4). Al precio así calculado se le denomina *precio marginal* o precio *spot*. Este es el mecanismo de casación adoptado por la mayoría de mercado eléctricos, aunque existen otros como el adoptado en el mercado de Reino Unido, conocido como *pay-as-bid*, donde a los agentes de venta se les paga el precio exacto que ellos ofertaron por la cantidad de electricidad indicada en su oferta (ver Vázquez (2003) para más detalles sobre modelos de casación de ofertas).

Las ofertas económicas de venta de energía eléctrica que los vendedores presenten al operador del mercado pueden ser simples o incorporar condiciones complejas en razón de su contenido. Las ofertas simples son ofertas económicas de venta de energía que los vendedores presentan para cada periodo horario y unidad de producción de la que sean titulares, expresadas como pares energía-precio. Las ofertas complejas de venta son aquellas que incorporan además todas o algunas de las condiciones técnicas o económicas siguientes: condición de indivisibilidad, gradiente de carga, ingresos mínimos y parada programada. La condición de indivisibilidad permite fijar en el primer tramo de cada hora un valor mínimo de funcionamiento. Este valor solo puede ser dividido por la aplicación de los gradientes de carga declarados por el mismo agente, o por aplicación de reglas de reparto en caso de ser el precio distinto de cero. El gradiente de carga permite establecer la diferencia máxima entre la potencia inicio de hora y la potencia final de hora de la unidad de producción, lo que limita la energía máxima a casar en función de la casación de la hora anterior y la siguiente, para evitar cambios bruscos en las unidades de producción que no pueden, técnicamente, seguir las mismas. La condición de ingresos mínimos permite la realización de ofertas en todas las horas, pero respetando que la unidad de producción no participe en el resultado de la casación del día, si no obtiene para el conjunto de su producción en el día, un ingreso superior a una cantidad fija, establecida en pesetas o euros, más una remuneración variable establecida en pesetas o céntimos de euro por cada kWh

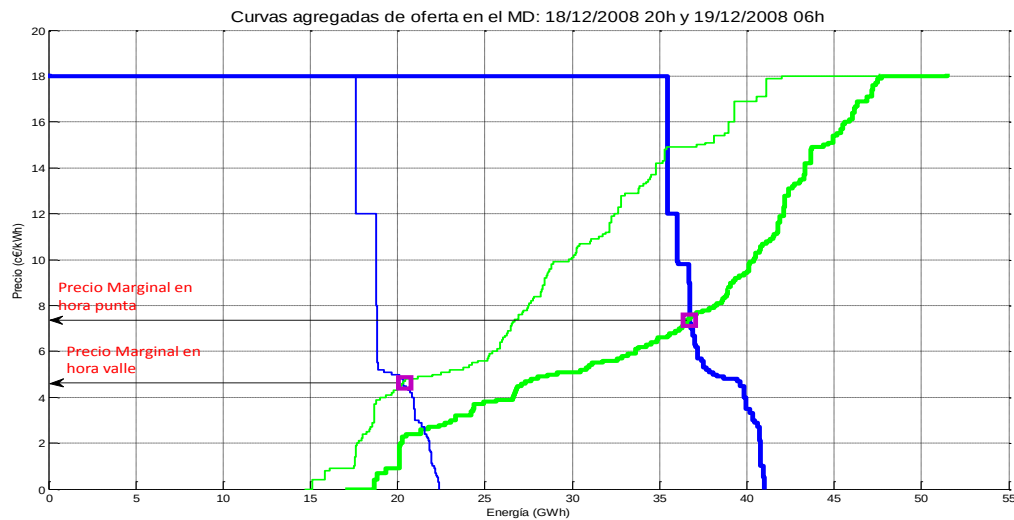


Figura 1.4: Curvas agregadas de oferta y demanda en OMEL y precios marginales resultantes en una hora punta (20h) y otra valle (06h) del 18 de diciembre de 2008.

casado. La condición de parada programada permite que si la unidad de producción ha sido retirada de la casación por no cumplir la condición de ingresos mínimos solicitada, realice una parada programada en un tiempo máximo de tres horas, evitando parar desde su programa en la última hora del día anterior a cero en la primera hora del día siguiente, mediante la aceptación del primer tramo de las tres primeras horas de su oferta como ofertas simples, con la única condición de que la energía ofertada sea decreciente en cada hora (ver www.omel.es para más detalles).

Las curvas agregadas de oferta habitualmente presentan una forma escalonada como las de la figura 1.4, en donde cada tramo o escalón refleja las ofertas correspondientes a centrales que pertenecen a una misma tecnología y el orden en el que aparecen viene determinado por sus costes de producción (o de oportunidad²). Las diferentes tecnologías empleadas en la generación de electricidad se distinguen unas de otras por su costes de operación (incluido el coste del combustible) y su flexibilidad, esto es, su capacidad para variar su potencia en un determinado periodo de tiempo. Las tecnologías más caras, como las centrales de fuel-gas, son las más flexibles. Únicamente operan unas pocas horas al año, cuando la demanda es muy alta, por lo que habitualmente se encuentra en la parte alta de la curva. En el otro

²Se entiende por coste de oportunidad el valor de la mejor opción alternativa no realizada. Se refiere a aquello de lo que un agente se priva o renuncia cuando hace una elección o toma de una decisión.

extremo, las centrales nucleares son las que presentan un menor coste variable de operación. Sin embargo, estas centrales son muy poco flexibles y tienen que operar de manera regular durante la mayor parte del año para evitar arranques y paradas. También se encuentran actualmente en la parte baja de la curva los productores de energía eólica, puesto que existe un régimen regulatorio especial que les lleva a ser tomadores de precio en el mercado³. La energía hidráulica tiene un coste de operación muy bajo y una gran flexibilidad, pero está sujeta a las disponibilidades de agua. Las centrales hidráulicas fluyentes suelen ofertarse a precios bajos puesto que sus costes de oportunidad son muy bajos. Sin embargo, las centrales hidráulicas regulables pueden aparecer en la parte alta de la curva ya que su coste de oportunidad puede ser alto (tienen la opción de reservar el agua para producir en un instante futuro en el que el precio del mercado sea alto). Finalmente, los costes de operación de las centrales de ciclo combinado de gas y las centrales de carbón se sitúan a un nivel intermedio.

Del lado de la demanda, los factores determinantes son fundamentalmente la temperatura y la laboralidad. Esto da lugar a distintos perfiles de demanda estacionales e intradiarios, que de manera general pueden clasificarse en horas de carga base u horas de punta. Una parte de la curva de demanda es a menudo ofertada a precio instrumental (el máximo posible) con el fin de asegurar el abastecimiento. Evidentemente, esto no significa que se pagará dicho precio, puesto que existe otra parte de los consumidores que sólo están dispuestos a tomar energía si su precio es menor o igual a un cierto valor. Estos consumidores tienen la posibilidad de adaptar su consumo a los precios del mercado (por ejemplo, algunos consumidores industriales o generadores hidráulicos de bombeo consumen electricidad preferentemente en los periodos de precios bajos). En la curva agregada de demanda del mercado, estos consumidores representan la parte de la curva con una cierta pendiente.

El precio resultante de la casación del mercado diario tiene una componente estratégica muy importante, puesto que depende de la manera en la que los agentes construyen sus ofertas. Puesto que el objetivo último de cada agente es maximizar el beneficio a lo largo de toda la secuencia de mercados, el precio del diario se ve también afectado por los precios de los mercados subsiguientes al diario.

1.4. Motivación de la tesis

La motivación de esta tesis surge de la necesidad de estudiar y comprender el comportamiento de los precios en mercados eléctricos, con el objetivo de predecirlos de la manera más precisa posible en un horizonte de corto plazo. Parafraseando el trabajo en Bunn and Karakatsani (2003): si la competencia en los mercados eléctricos fuera tan eficiente⁴ que los precios reflejaran los costes marginales incluso en las horas

³Esta es la situación regulatoria en los periodos de tiempo bajo estudio en este documento. Sin embargo, la entrada en vigor del real decreto 2/2013 cambia el régimen regulatorio de los productores de energía eólica, lo que puede afectar al modo en que éstos oferten en el mercado.

⁴En un mercado eficiente los precios reflejan toda la información disponible del pasado y cambian instantáneamente para reflejar nuevas informaciones.

de alta demanda, habría convergencia entre el precio de la electricidad y el coste del combustible necesario para generarla, y el modelado de los precios eléctricos sería relativamente sencillo. Si los mercados de futuros fueran líquidos⁵ y completos⁶, los precios forward tendrían una estructura similar a la de otros activos financieros y podrían aplicarse técnicas convencionales de gestión de riesgos. Sin embargo, ninguna de estas situaciones se da en la mayoría de mercados eléctricos, lo que hace necesario el análisis específico del comportamiento de los precios para su correcto modelado.

La tesis se centra en el comportamiento de los precios del mercado diario, por ser éste el elemento clave del funcionamiento de los mercados eléctricos liberalizados. Existe ya abundante literatura en la que se analizan los precios de diversos mercados y se proponen diferentes métodos de predicción. Sin embargo, a menudo el objetivo ha sido tratar de predecir únicamente el valor puntual o precio que se obtiene de la casación de las ofertas, sin prestar demasiada atención a la magnitud del posible error en la estimación o la probabilidad de obtener cualquier otro precio. Puesto que el precio resultante de la casación es el resultado del conjunto de las estrategias de los agentes participantes en el mercado, la predicción de dicho precio es un problema intrínsecamente estocástico, y es precisamente en este punto, la estimación de la función de densidad de los precios del mercado, en dónde se pone el foco de atención en la tesis. Con ello se pretende contribuir en áreas de conocimiento en donde la estimación y caracterización de los precios del mercado eléctrico resulta de gran importancia, como puede ser el diseño y optimización de estrategias de oferta o la gestión del riesgo derivado de dicho precio. El conocimiento de la distribución de probabilidad de los precios del mercado diario del día siguiente permite introducir incertidumbre en la elaboración de los escenarios utilizados para maximizar el beneficio de un agente en el mercado, y proporciona la información necesaria para una correcta gestión del riesgo que de él se deriva.

1.5. Organización del documento

Lo que resta de documento está organizado de la siguiente manera: En el capítulo 2 se hace una revisión del estado del arte de los modelos que tienen como objetivo predecir o estimar los precios de la electricidad. El capítulo se centra fundamentalmente en los utilizados para predecir los precios a corto plazo, en particular, para predecir los precios del mercado diario con un horizonte de hasta un día. Los siguientes dos capítulos describen las contribuciones más importantes de la tesis. En el capítulo 3 se presenta una manera alternativa de modelar la distribución de probabilidad de los precios eléctricos que extiende el tradicional marco de trabajo gaussiano utilizado en los modelos de series temporales. Esta metodología puede contribuir a modelar de manera más realista el comportamiento de los precios

⁵Un mercado se dice que es líquido cuando un agente puede comprar o vender cantidades significativas del producto sin alterar el precio del mismo.

⁶En un mercado completo todos los posibles pagos futuros de cada uno de los productos que se negocian se pueden replicar usando instrumentos disponibles en ese mercado.

eléctricos, sobre todo en un contexto de gestión de riesgos; En el capítulo 4 se propone un modelo de predicción de precios eléctricos capaz de mejorar las estimaciones de los modelos habitualmente utilizados en la literatura. El modelo puede enmarcarse tanto en el campo de los modelos paramétricos de series temporales como en el de las redes neuronales, puesto que mezcla la formulación utilizada en ambas disciplinas. Finalmente, en el capítulo 5 se comprueban las capacidades de predicción de los modelos propuestos en el contexto habitual al que se enfrentan los agentes del mercado eléctrico, y se comparan sus resultados con los obtenidos utilizando otros modelos más habituales en la literatura. El capítulo 6 resume las conclusiones de la tesis.

Capítulo 2

Modelos de predicción de precios spot en mercados eléctricos

En este capítulo se revisan los métodos utilizados para predecir o estimar el precio de la electricidad en mercados eléctricos. En primer lugar, se presenta una clasificación general de los distintos enfoques con los que se aborda el problema en función del horizonte de predicción o la finalidad que se persigue con el modelo. A continuación, el capítulo se centra en los modelos habitualmente utilizados para la predicción a corto plazo, especialmente para la predicción de los precios spot del día siguiente. Este grupo de modelos son de índole estadística y abarca desde los métodos paramétricos basados en el análisis de series temporales hasta los basados en técnicas de inteligencia artificial o minería de datos. El capítulo describe también las especiales características estadísticas de la series de precios, en las cuales se apoyan los modelos descritos.

2.1. Clasificación general de los modelos de precios eléctricos

La introducción de competencia tras la liberalización de los mercados eléctricos ha obligado a las empresas participantes a anticiparse a los movimientos del precio marginal, puesto que es una variable determinante en sus procesos de toma de decisiones y desarrollo estratégico. La necesidad de predicciones precisas del precio marginal abarca tanto el largo, como el medio y corto plazo. En el largo plazo, las predicciones del precio influyen en las decisiones relacionadas con el análisis de inversiones o la planificación, como por ejemplo la expansión de la generación o la planificación de la distribución. Las predicciones a medio plazo (de seis meses a un año) se utilizan para la coordinación hidrotérmica, así como a la hora de hacer una gestión eficaz del riesgo asociado al precio del mercado, de manera que los agentes puedan decidir la cantidad de energía que comercializan a través de contratos bilaterales, mercados de futuros o directamente en el *pool*. Las decisiones de corto plazo conciernen a la programación diaria de la producción de las plantas y el desarrollo de las estrategias de oferta en el mercado, con el objetivo de maximizar el beneficio.

Según el horizonte de predicción requerido o la finalidad que se persigue con el modelo, la metodología empleada para estimar los precios eléctricos resultantes de la casación del mercado del diario (precios spot) es diferente. En esta sección se introducen algunos de los modelos utilizados para la predicción de precios en el largo y medio plazo, así como para la valoración de productos derivados del precio eléctrico. El análisis detallado de este tipo de modelos está, sin embargo, fuera del alcance de la tesis. El objetivo no es otro que ilustrar las diferencias que presentan con respecto a los utilizados para la predicción a corto plazo, en particular para la predicción de los precios spot del día siguiente, los cuales se describen en profundidad en la sección 2.2.

Aunque las técnicas de predicción de precios eléctricos todavía están en una etapa temprana de madurez, existe ya mucha literatura al respecto e incluso pueden encontrarse otras taxonomías previas expuestas en Weron, 2006; Li et al., 2005; Aggarwal et al., 2009; Mateo, 2005. La clasificación presentada en este capítulo no difiere en líneas generales de las citadas.

2.1.1. Modelos para el medio y largo plazo

Los modelos que pertenecen a esta categoría no están únicamente destinados a estimar el precio del mercado, sino a representar el comportamiento del mercado completo, incluido el comportamiento estratégico de los agentes participantes, con el objetivo último de obtener información útil para la planificación a medio o largo plazo. La predicción del precio constituye por tanto una de las salidas del modelo, de la que habitualmente no se espera obtener precisión sino más bien analizar su

respuesta frente a diferentes escenarios.

Existen diferentes enfoques o metodologías que se aplican para abordar esta problemática. Como denominador común a todas ellas está la necesidad de disponer de una descripción detallada del sistema, lo que supone un significativo riesgo para la toma de decisiones.

Uno de los enfoques utilizados es el basado en costes de producción. En estos modelos se simula la producción de las unidades de generación con el objetivo de satisfacer la demanda al mínimo coste. Pueden ser usados para realizar análisis en tiempo real y obtener estimaciones del precio hora a hora. Estos modelos son apropiados para mercados muy regulados donde existe poca incertidumbre sobre el precio. Sin embargo, presentan la limitación de no considerar el comportamiento estratégico de los agentes a la hora de elaborar las ofertas del mercado, por lo que no son tan adecuados en mercados competitivos. Una extensión de un modelo de costes de producción donde se tiene en cuenta este comportamiento estratégico ha sido propuesto en Batlle and Barquín (2005). En cualquier caso, este tipo de modelos sirven como referencia para conocer cuáles son las componentes del precio.

Otro de los enfoques es el constituido por los modelos de equilibrio, que están basados en la teoría de juegos (Ventosa, 2005; Barquín et al., 2005). Aquí se tiene en cuenta que el resultado del mercado depende del conjunto de ofertas enviadas por todos los agentes, por lo que éste puede ser considerado como un juego donde los agentes participantes son los jugadores y las reglas del mismo quedan definidas por el tipo de subasta. Así, la decisión de cada agente que participa en el mercado se corresponde con un problema de optimización donde se maximizan sus beneficios sujeto a un conjunto de restricciones técnicas (arranques y paradas de sus grupos, mínimos técnicos, contratos, etc), y existe una restricción común a todos los agentes relacionada con el hecho de que siempre debe respetarse el equilibrio entre generación y demanda. Uno de los inconvenientes de estos modelos es su elevado coste computacional debido al complejo procedimiento de optimización requerido. Además, existe un gran número de componentes que tienen que ser definidos con un muy alto nivel de detalle: número de jugadores, posibles estrategias, el modo en el que interactúan los agentes, etc; lo que implica un grado de información del mercado que no está disponible para todos los agentes del mercado.

El último de los grupos de modelos utilizados son los denominados modelos de simulación (ver p.ej. Bastian et al. (1999)). Estos modelos son a menudo utilizados en sistemas eléctricos basados en precios nodales (*Locational marginal prices*), en los cuales cada unidad de energía tiene un precio diferente según el punto de la red donde sea consumida. La principal razón de esta discriminación espacial son las pérdidas en las líneas de transmisión y las congestiones. Los modelos de simulación parten de una representación de las relaciones físicas y económicas entre las variables que afectan al sistema completo, con el fin de simular su operación teniendo en cuenta sus requerimientos y restricciones. Las entradas a un modelo de este tipo son, por tanto, la red de transporte, datos de las unidades de generación, costes

de transmisión, condiciones hidrológicas, precios de los combustibles, predicciones de demanda, comportamientos estratégicos de los agentes, etc. El inconveniente de estos modelos es el volumen de datos requerido, lo que limita su aplicabilidad, y su elevado coste computacional.

2.1.2. Modelos para la valoración de productos derivados

Alrededor de la mayoría de los mercados eléctricos liberalizados se articulan también otros mercados financieros más o menos desarrollados que permiten comercializar con productos *derivados*¹ del precio spot del mercado. Estos productos financieros son una herramienta muy útil para los agentes que operan en el mercado eléctrico, tanto generadores como comercializadores, puesto que permiten transferir el riesgo asociado al precio del mercado a otros que puedan beneficiarse de tomar dicho riesgo. Como ejemplo se puede considerar la situación de una empresa comercializadora que compra electricidad en el *pool* del mercado y la vende a sus clientes finales a precios fijados. En caso de que el precio spot experimente un fuerte incremento inesperado, la comercializadora podría sufrir pérdidas muy elevadas. Del mismo modo, los beneficios de un generador que vende energía en el mercado podrían verse seriamente afectados si los costes del combustible que utiliza se incrementan más que los de otros tipos de generación, por lo que un contrato a plazo puede permitirle fijar un nivel de ingresos estable y aceptable para la cantidad objeto de contratación.

El punto clave de los modelos destinados a la valoración de productos derivados del precio eléctrico es la estimación del precio que tiene que ser pagado por la entrega física de una unidad de energía en una fecha determinada, lo que se conoce como *precio forward*. En la literatura coexisten dos enfoques diferentes que persiguen este objetivo. Uno de ellos trata de modelar las dinámicas de las curvas *forward*² directamente (ver p. ej. Borak and Weron (2008)), mientras que el otro enfoque trata de modelar los precios spot y a partir de estos deducir los precios *forward*. Partiendo de la hipótesis de ausencia de arbitraje³, la relación entre el precio spot P_t y el precio *forward* $F_{t,T}$ con fecha de vencimiento T , viene dada por (2.1) (Eydeland and Wolyniec (2002))⁴.

$$F_{t,T} = P_t e^{(r_t - d_t)(T-t)} \quad (2.1)$$

donde r_t denota la tasa de interés sin riesgo y d_t la prima asociada por

¹Derivados son contratos, instrumentos financieros, que derivan su valor de un activo subyacente, por ejemplo, el precio spot de la electricidad. Existen cuatro tipos básicos: contratos forward, futuros, swaps y opciones.

²La curva forward es el conjunto de precios a los que se podría intercambiar la electricidad en distintos plazos futuros.

³Arbitraje denota un situación en la que puede sacarse beneficio sin tomar ningún riesgo.

⁴Puesto que no entraremos en mayor detalle, se presenta la ecuación (2.1) como la habitualmente utilizada para la valoración de producto derivados. Sin embargo, para *commodities* como la electricidad los argumentos de arbitraje utilizados necesitarían ser cuidadosamente revisados (Hull (2006)).

vender el activo. En el contexto de valoración de productos derivados, los modelos habitualmente utilizados para describir la evolución de los precios spot son llamados cuantitativos (*quantitative models*). Estos modelos están inspirados en los utilizados para modelar tipos de interés y otras *comodities*, y tratan de replicar las características principales de los precios eléctricos. Los modelos conocidos como *mean-reversion jump-diffusion* son representativos de este grupo y se han utilizado para describir algunas de las características más importantes de los precios eléctricos, como son la reversión a la media y la presencia de *spikes* (Weron (2006)) (ver sección 2.2.1). La expresión general de estos modelos es a menudo descrita a través de ecuaciones diferenciales estocásticas de la forma

$$dP_t = \mu(P_t, t)dt + \sigma(P_t, t)dW_t + dq(P_t, t) \quad (2.2)$$

donde $\mu(P_t, t)$ denota la deriva (o *drift*), $\sigma(P_t, t)$ es el término que modela la volatilidad, $q(P_t, t)$ es un proceso de *Poisson* compuesto que produce infrecuentes pero grandes saltos y W_t es un movimiento browniano estándar en tiempo continuo cuya característica más relevante es que sus incrementos dW_t son variables normales i.i.d. Una vez discretizada la ecuación diferencial estocástica, estos modelos son equivalentes a algunos de los modelos de series temporales descritos en la sección 2.2.2. Así, por ejemplo, un mean-reverting diffusion (Lucia and Schwartz (2000)), que responde a (2.2) sin incluir el proceso de *Poisson* compuesto y donde

$$\mu(P_t, t) = \alpha - \beta P_t \quad (2.3)$$

es equivalente a un proceso autorregresivo de orden 1 (*AR(1)*). En Jong (2006a) se utiliza un conjunto de modelos diferentes para representar los precios spot de distintos mercados internacionales. Entre los modelos se encuentra un mean-reverting jump-diffusion que se describe en forma discretizada de la siguiente manera

$$dP_t = \alpha(\mu_1 - P_{t-1}) + \sigma_1\epsilon + \sum_{i=0}^{n_t} Z_t \quad (2.4)$$

con

$$Z_t \sim N(\mu_2, \sigma_2^2) \quad (2.5)$$

$$n_t \sim POI(\lambda) \quad (2.6)$$

dónde $POI(\cdot)$ denota un proceso de Poisson compuesto y $\lambda, \alpha, \mu_1, \sigma_1, \mu_2, \sigma_2$ son parámetros del modelo.

La analogía entre el enfoque cuantitativo y el de series temporales podría confundir al lector. La diferencia fundamental entre ambos radica en el objetivo que se persigue con el modelo. En los cuantitativos, la finalidad no es tanto la precisión sino captar adecuadamente las características esenciales del precio, con el objetivo último

de deducir el valor del derivado. Los parámetros calibrados para este tipo de modelos son consistentes con los datos del mercado y el proceso de precios que modelan sigue una evolución acorde con la distribución neutral al riesgo implícita en el mercado⁵. Esto da lugar a modelos con muy pocos parámetros, a menudo formulados en escala diaria. Los modelos de series temporales persiguen la precisión de sus estimaciones y sus parámetros se ajustan a valores históricos. Suelen involucrar un mayor número de parámetros y son más adecuados para realizar estimaciones horarias.

2.2. Modelos de predicción a corto plazo

La predicción de los precios en un horizonte de corto plazo resulta de vital importancia para los agentes del mercado en su operativa diaria. Partiendo de predicciones precisas de los precios de casación, las empresas participantes pueden ajustar sus programas de acuerdo con ellos y, por tanto, maximizar su beneficio o utilidad. En este contexto, la fuente de incertidumbre más relevante son los precios del mercado diario resultantes de la casación para el día siguiente, aunque también son necesarias estimaciones de los mismos que comprendan hasta los 7 días siguientes, así como estimaciones de los precios del resto de mercados subsiguientes al diario que comprenden horizontes que van desde unas pocas horas a una semana.

Dado que el objetivo último de estos modelos es la precisión de sus estimaciones, la metodología usada para la predicción de los precios es diferente al de los modelos anteriormente expuestos. Aquí no se trata de modelar el comportamiento estratégico de los agentes del mercado ni de describir las relaciones físicas o económicas del sistema completo, sino de modelar estadísticamente la relación existente entre el precio y su pasado cercano. Los métodos de predicción para el corto plazo se basan, por tanto, en los patrones estadísticos repetitivos que la historia de las series de precios presenta, los cuales son explotados para la predicción de los precios futuros. En sus versiones más simples, estos modelos tan solo utilizan la información de la serie histórica del precio. No obstante suele alcanzarse más precisión si se incorporan también algunas de las variables fundamentales que afectan al precio, como puede ser la demanda eléctrica o las reservas hidráulicas (Karakatsani and Bunn, 2008a; Weron, 2006). Estos métodos son adecuados para su utilización en tiempo real y están al alcance de cualquier agente del mercado.

Los métodos pertenecientes a esta categoría parten de la base de considerar el precio del mercado como la realización de un proceso estocástico que es necesario caracterizar (ver Peña (2005) para más detalles). Existen dos corrientes claramente diferenciadas. Una de ellas está formada por los métodos de predicción paramétricos, en los cuales se asume a priori un modelo estadístico del proceso estocástico, basándose en las características que presentan las series de precios. Los métodos pertenecientes a este grupo son los procedentes del análisis de series temporales

⁵La ausencia de oportunidades de arbitraje implica la existencia de una distribución de probabilidad cuya esperanza es el valor del activo valorado.

descritos en la sección 2.2.2. La otra corriente está constituida por los métodos no paramétricos procedentes del campo de la inteligencia artificial, en los cuales no se hace ninguna hipótesis previa acerca de las relaciones existentes entre el precio y su pasado reciente y/o otras variables explicativas, de manera que la estructura del modelo estadístico queda definida por el conjunto de datos utilizado para su estimación. Este último grupo de métodos son descritos en las secciones 2.2.3 y 2.2.4.

A continuación se describirán las características que habitualmente presentan las series de precios en los distintos mercados eléctricos, en las cuales se basan los métodos de predicción a corto plazo.

2.2.1. Características de las series de precios

Las series de precios spot habitualmente presentan un comportamiento considerado errático e imprevisible (Knittel and Roberts, 2001; Álvaro Escribano et al., 2002; Weron, 2006). Las causas de este comportamiento se encuentran en factores como la naturaleza no almacenable de la electricidad, la forma escalonada de las curvas de venta de energía debido a la presencia de diferentes tecnologías de generación, la estructura oligopolística de los mercados eléctricos y los frecuentes cambios de regulación y su complejo diseño. De este modo, los precios spot presentan un conjunto de características únicas y comunes a la mayoría de los mercados, que se resumen en: reversión a la media, presencia de componentes estacionales (estacionalidad diaria, semanal y anual), presencia de diferentes patrones de comportamiento según la hora del día o el día de la semana, efectos de calendario, volatilidad variable en el tiempo y distribuciones asimétricas con colas pesadas⁶ y valores extremos. En la figura 2.1 se muestran las series de precios en OMEL (mercado Ibérico, ahora OMIE), POWERNEXT (Francia), PJM (Interconexión Pensilvania - New Jersey - Maryland) y NEM (Victoria). En ella pueden apreciarse algunas de estas características, como son los patrones de comportamiento diarios y semanales, la volatilidad variable y la presencia de los valores extremos denominados *spikes*.

A la hora de explicar el comportamiento de las series de precios, la demanda eléctrica juega un papel muy relevante. Este hecho es bastante lógico, dada la instantánea naturaleza de la electricidad y puesto que el precio final de casación se obtiene tras cruzar las curvas agregadas de oferta y demanda, estas últimas habitualmente más inelásticas⁷ que las primeras. Con el resto de factores estables, un incremento en la demanda eléctrica implica un incremento en el precio spot. Habitualmente, el operador del sistema proporciona a los agentes del mercado predicciones de la demanda horaria suficientemente precisas para que estos puedan incorporarlas en sus procesos de elaboración de ofertas. Esta variable resulta, por tanto, muy útil para modelar el efecto que tienen sobre el precio diversos factores como son la actividad económica (efectos de calendario, variaciones horarias y diarias) y las

⁶En las distribuciones que presentan colas pesadas (*fat tails*), existe mayor probabilidad de obtener valores extremos que la que existiría en una distribución normal.

⁷Los bienes que presenta demanda inelástica son poco sensibles a variaciones del precio.

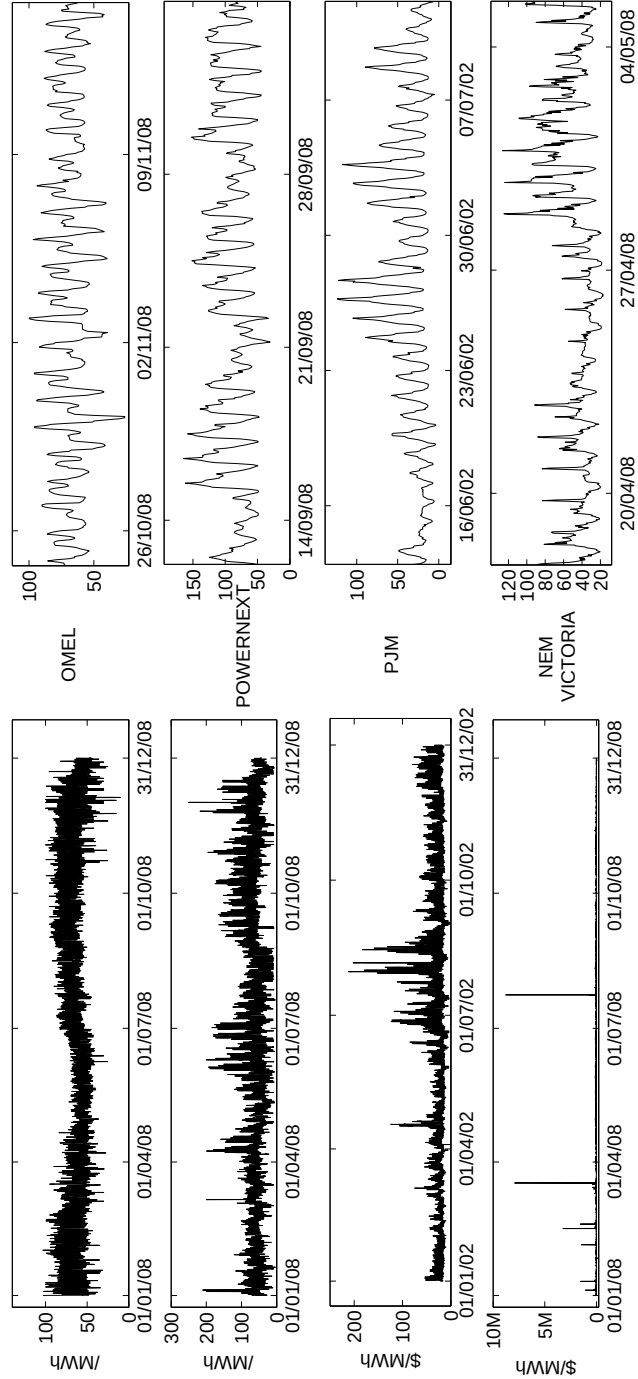


Figura 2.1: Precios en OMEL (España), POWERNEXT (Francia), PJM (Pensilvania, New Jersey, Maryland) and NEM (Victoria)

variaciones en las condiciones meteorológicas o en la temperatura. En la figura 2.2 se muestran los precios spot y la demanda eléctrica en PJM, OMEL y NEM (Victoria). En ella puede apreciarse cómo la respuesta del precio a la demanda, en general, no es lineal y varía en función de su nivel. Como se afirma en Bunn and Karakatsani (2003), los precios eléctricos presentan una estructura mucho más compleja que un simple escalado de la demanda eléctrica que refleje los costes de generación.

En el lado de la oferta, el mix de generación de cada mercado específico es el factor determinante para explicar las variaciones de los precios spot. Por ejemplo, en el mercado Ibérico (OMEL/OMIE), la previsión de generación eólica se ha convertido en una variable fundamental debido a su alto grado de penetración⁸, y un régimen regulatorio especial que lleva a los productores de energía eólica a actuar como tomadores de precio en el mercado. En el Nord Pool, una de las variables más importantes es la cantidad de recursos hidráulicos disponibles para la generación. En general, el modo en el que los generadores construyen las ofertas que son enviadas al mercado está fuertemente ligado a los costes de producción de las tecnologías de generación que cubren la demanda. Por tanto, los costes de los combustibles, costes de emisiones, reservas hidráulicas, exportaciones e importaciones, e incluso las paradas de mantenimiento de algunas plantas determinan el precio spot resultante. Sin embargo, es importante destacar que en el contexto de predicción a corto plazo, algunas de estas variables no son útiles para explicar las variaciones de alta frecuencia (variaciones horarias) de los precios. Otras variables, como la generación hidráulica, son difíciles de predecir o no están disponibles para todos los agentes del mercado.

Como ejemplo, la figura 2.3 muestra los precios medios diarios frente a la producción media diaria de las tecnologías involucradas en OMEL. Para el caso de la producción eólica, la variable mostrada en la gráfica es la previsión que el operador del sistema pone a disposición de los agentes antes de que estos envíen sus ofertas al mercado diario. Como puede apreciarse, existe una reducción significativa en el precio medio del mercado según se incrementa la previsión de generación de energía eólica. Como es lógico, el resto de tecnologías también presentan un patrón de correlación con el precio. Sin embargo la predicción de su producción resulta bastante complicada sin más información que la publicada en el mercado para todos los agentes, con la excepción de la producción nuclear, que puede ser aproximada teniendo en cuenta su disponibilidad declarada. Esto es debido a que esta tecnología también es habitualmente ofrecida al precio más bajo posible. No obstante, como se aprecia en la figura, el efecto esperado de reducción del precio según se incrementa la generación con este tecnología no parece tan claro, lo que se debe a la respuesta inelástica del precio a la generación en base⁹.

Atendiendo a los factores que pueden explicar las variaciones de los precios spot, cabe destacar el trabajo expuesto Karakatsani and Bunn (2008a), donde se plantea un

⁸El 11 % de la demanda en 2008 fue cubierta por parques eólicos. El pico de demanda máxima horaria en 2008 fue alcanzado el 15 de Diciembre (42.961MW). El 18 % de esta demanda fue cubierta por energía eólica.

⁹generación de electricidad que cubre la demanda mínima.

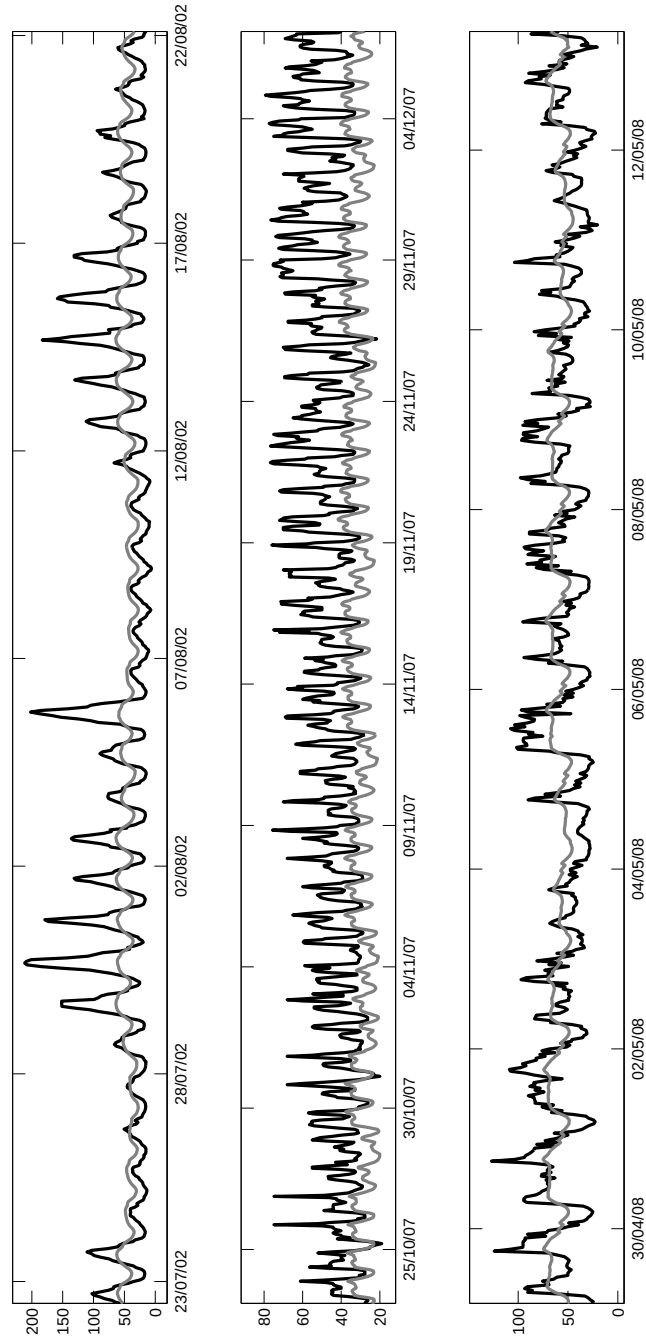


Figura 2.2: Precio(línea negra) y Demanda(línea gris) en PJM (\$/MWh - GWh), OMEL (€/MWh - GWh) y NEM (\$/MWh - GWh)

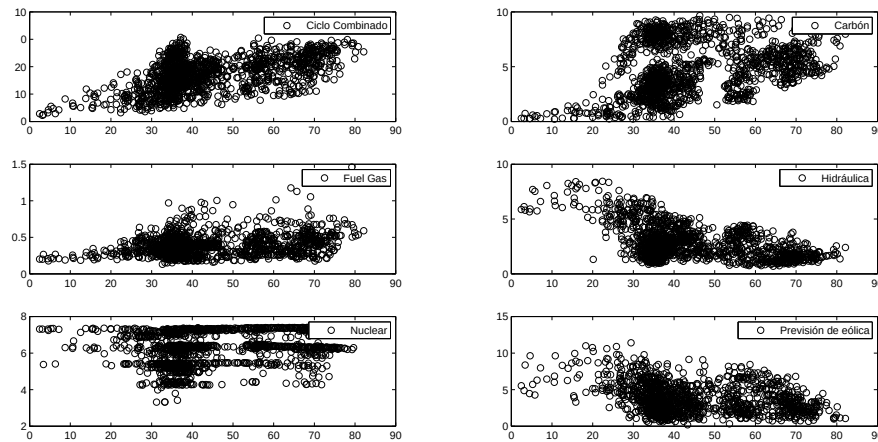


Figura 2.3: Precio medio diario (€/MWh) frente a la producción media diaria (GWh) de cada una de las tecnologías involucradas en OMEL (datos correspondientes a los balances del P48 publicados por Red Eléctrica de España, período: Enero 2007 - Agosto 2010).

modelo de regresión con coeficientes variables en el tiempo que utiliza como regresores variables fundamentales del precio. Aparte de la inclusión de una componente de tendencia y factores estacionales en el modelo, el resto de variables son:

- la demanda;
- la pendiente y curvatura de la demanda (diferencias de la demanda de primer y segundo orden, respectivamente) y el *spread* (diferencia entre el precio del diario y el precio de los mercados subsiguientes, es decir, intradiarios y de reserva) con el fin de capturar el efecto de las estrategias de los agentes en relación a los mercados subsiguientes al diario;
- los coeficientes de variación (desviación estándar dividido por la media) del precio y la demanda como índices de reacción de los agentes frente al riesgo;
- el margen de capacidad (máxima disponibilidad declarada menos la predicción de demanda) y la escasez (el margen de capacidad entre la predicción de demanda) como medidas del exceso de capacidad de generación;

No fueron considerados los costes de los combustibles debido a su lenta evolución en el tiempo, ni los precios de los contratos forward puesto que reflejan el coste del riesgo percibido por los agentes, el cual ya es capturado a través de otras variables.

En cuanto a las características intrínsecas de los precios spot, la más distintiva de ellas es la presencia de los denominados *spikes*. Estos *spikes* son valores extremos

que aparecen de manera abrupta e inesperada en un periodo muy breve de tiempo. Durante este periodo, el precio incrementa significativamente su valor, para luego volver a caer al nivel previo. Normalmente, los *spikes* se producen durante las horas de alta demanda y están presentes en la mayoría de mercados eléctricos, aunque no se dan con la misma intensidad. La figura 2.1 ilustra algunos de los spikes ocurridos en el mercado de Victoria, donde el valor del precio alcanzó varios órdenes de magnitud por encima de su valor habitual.

Los *spikes* se producen debido a la combinación de diversos factores. Por un lado, la naturaleza no almacenable de la electricidad causa que las tecnologías más caras determinen el precio durante las horas de alta demanda. Además, paradas no programadas de algunas plantas de generación o fallos en la transmisión pueden agravar esta situación. En este contexto, las estrategias de las empresas participantes en el mercado juega un papel muy importante. Por otro lado, dado que la electricidad es una *commodity* esencial, muchos compradores están dispuestos a pagar casi cualquier precio para asegurar el suministro de electricidad.

En el contexto de modelado de series temporales, es aconsejable eliminar los *spikes* del conjunto de ajuste previamente a la estimación de un modelo estadístico que capture el comportamiento normal de los precios. De otro modo estos valores extremos podrían afectar de manera dramática al valor de los parámetros y causar que el modelo final no describa adecuadamente ni los *spikes* ni el comportamiento normal de la serie. Existen algunos métodos paramétricos para detectar y corregir *outliers* de manera automática que pueden ser usados (Gómez and Taguas, 1995; Chang et al., 1988). Otras maneras de prevenir el efecto negativo de los *spikes* en el proceso de ajuste es filtrar la serie usando la transformada wavelet (Conejo et al. (2005)), o atenuar el valor de los precios cuando este sobrepase un determinado umbral (Weron (2006)). Sin embargo, es necesario tener en cuenta que los *spikes* no son valores que correspondan a errores de medida, y por consiguiente un modelo bien especificado debería ser capaz de modelar sus ocurrencias. Los modelos jump-diffusion y los de cambio de régimen han sido utilizados para tal propósito (ver por ejemplo Jong (2006b)). Estos modelos son descritos en la secciones 2.1.2 y 2.2.2.8, respectivamente.

Otra característica distintiva de los precios spot, relacionada con los spikes, es su especial distribución de probabilidad condicional. Las distribuciones empíricas obtenidas a través del tradicional ajuste gaussiano habitualmente no cumplen con la hipótesis de normalidad (o log-normalidad), incluso cuando se asume que la varianza de la serie de precios sigue un proceso GARCH (ver sección 2.2.2.5). Existe mucha literatura al respecto señalando que las distribuciones de probabilidad de los precios spot presentan colas pesadas y son asimétricas (ver p. ej. Knittel and Roberts, 2001; Weron, 2006). Este hecho no suele afectar en gran medida a la predicción puntual de los precios eléctricos, pero si a la predicción de intervalos a niveles de confianza cercanos al 100 %. ¹⁰. En la sección 2.2.2.9 se describen algunos trabajos que se han

¹⁰Se entiende por predicción puntual a la predicción del valor concreto de una serie temporal en cada instante t , es decir, la esperanza de la distribución de probabilidad predicha del proceso

preocupado de esta cuestión.

Entre las características más importantes a la hora de predecir el comportamiento de los precios eléctricos está su carácter estacional y autorregresivo. Este comportamiento suele reconocerse fácilmente tras calcular la Función de Autocorrelación Simple (FAS) de la variable diferenciada del precio eléctrico $z_t = (1 - L)(1 - L)^{S_1} \dots (1 - L)^{S_m} y_t$, siendo y_t el precio, L es el operador retardo tal que $L^n y_t = y_{t-n}$ y $S_1, \dots, S_m$ los períodos estacionales considerados¹¹ ($S_1 = 24, S_2 = 168$ para el caso de series horarias o $S_1 = 48, S_2 = 336$ en el caso de que la subasta del mercado diario comprenda 48 períodos de media hora cada uno, como sucede en NEM). En la figura 2.4 aparecen las series diferenciadas de los precios horarios en OMEL, PJM y NEM¹² así como la FAS para cada una de ellas. La inspección de la FAS muestra que existen correlaciones significativas en los retardos más bajos y en los retardos estacionales, siendo estas últimas las más importantes. Esto implica que existe información contenida en dichos retardos de la propia serie temporal que puede ser utilizada para la construcción de un modelo lineal que prediga los precios eléctricos.

Sin embargo, que exista dicha correlación lineal no significa que las verdaderas relaciones entre el precio y sus retardos sean lineales. Si consideramos ahora diferentes períodos o instantes de la serie temporal del precio de manera separada, por ejemplo, las horas llano, punta y valle¹³, o los distintos tipos de días (lunes, de martes a viernes, sábados y domingos) podemos ahora calcular la autocorrelación de la serie para cada uno de los períodos considerados a través de la Función de Autocorrelación Simple Periódica (FASP) (ver Mcleod (1993) para detalles). Las figuras 2.5 y 2.6 ilustran esta situación para los precios en OMEL. En la figura 2.5 se observa cómo el patrón de autocorrelación que afecta a los retardos múltiplos del periodo 24 es diferente y más significativo para los precios correspondientes a los martes, miércoles, jueves y viernes que para el resto de días. El hecho de que el coeficiente de autocorrelación en el retardo 24 sea, por ejemplo, diferente y mayor para un miércoles que para un sábado resulta lógico si atendemos a la diferente relación entre días consecutivos según estos sean laborables o festivos. En la figura 2.6 se observa cómo también el patrón estacional de las horas punta difiere del de las horas valle y llano. Puesto que las demandas punta diarias suelen producirse en las mismas horas del día, la figura muestra que existe mayor correlación con los precios de días anteriores cuando la hora es punta que cuando no lo es. Sin embargo, los precios en las horas punta presentan menor correlación con los de la semana anterior. Como puede verse, en

estocástico en cada instante t . Los intervalos de predicción son los valores predichos entre los cuales se encuentra el valor concreto o realizado a un determinado nivel de confianza.

¹¹Esta transformación es frecuentemente utilizada en el contexto de series temporales. Se utiliza para convertir una serie en estacionaria en media. De este modo se evita que aparezcan autocorrelaciones en la FAS causadas por la tendencia de la serie.

¹²Para el caso de NEM, los spikes han sido atenuados al valor máximo de 150\$/MWh puesto que de otro modo las autocorrelaciones quedan distorsionadas.

¹³La clasificación de las horas de un día en horas de llano, punta y valle es comúnmente utilizada en el contexto de mercados eléctricos. Las horas de punta se corresponden con las 4 horas de mayor demanda de un día, las horas de valle a las 8 horas de menor demanda y el resto se consideran horas de llano.

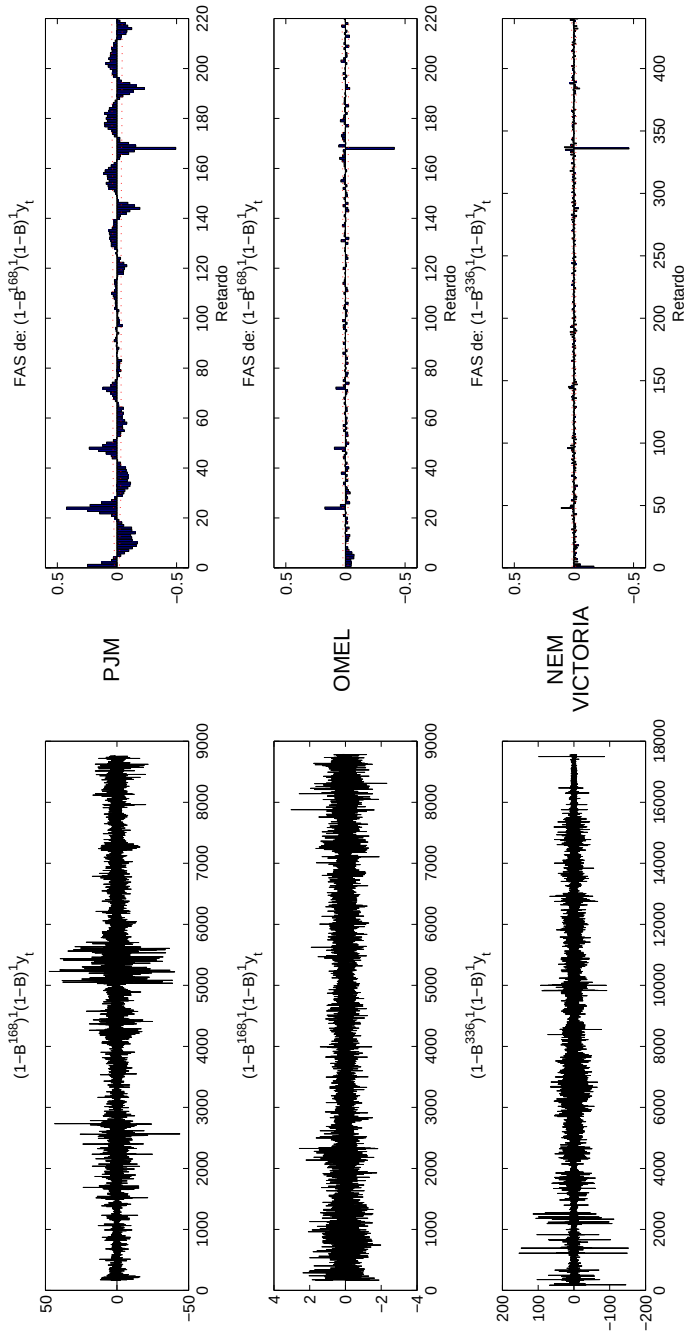


Figura 2.4: Precios diferenciados y FAS correspondiente en PJM (2002), OMEL (2008) y NEM Victoria (2008)

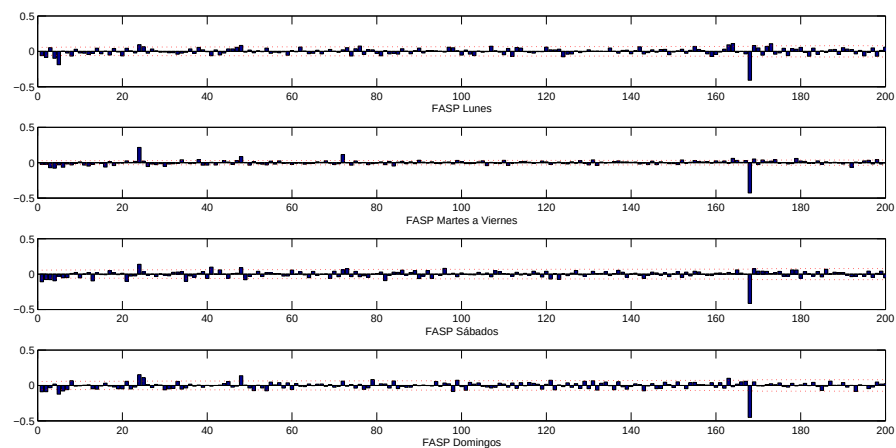


Figura 2.5: FASP de los precios diferenciados para los lunes, de martes a viernes, sábados y domingos, en OMEL (2008)

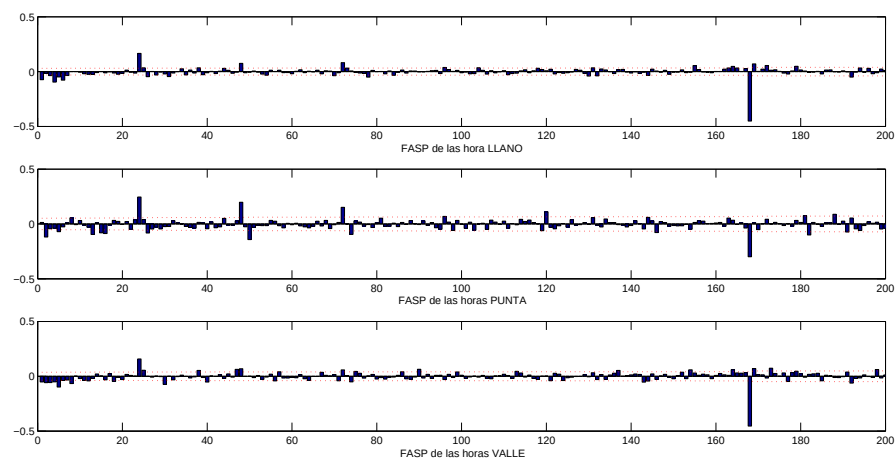


Figura 2.6: FASP de los precios diferenciados para las horas tipo Llano, Punta y Valle, en OMEL (2008)

general, las series de precios eléctricos exhiben diferentes patrones de autocorrelación intra-diarios y/o intra-semanales.

2.2.2. Modelos de series temporales

En esta sección se describen los modelos de series temporales usados para la predicción a corto plazo de los precios en mercados eléctricos. Los métodos de predicción dentro de esta categoría constituyen el enfoque paramétrico, en el cual se asume un modelo estadístico a priori en función de las características de la serie a modelar. Dado que algunas de las habituales hipótesis en este tipo de modelos no se cumplen para los precios eléctricos, han surgido también algunas técnicas de predicción híbridas que combinan métodos paramétricos y no paramétricos, los cuales son descritos al final de la sección.

2.2.2.1. Modelos ARIMA

Los modelos ARIMA (*Auto Regressive Integrated Moving Average*) han sido ampliamente utilizados como métodos eficaces de predicción, y junto con las técnicas de alisado exponencial constituyen la base de los métodos univariantes clásicos del análisis de series temporales (Peña, 2005; Hamilton, 1994). En estos modelos, el valor esperado de la variable de salida se obtiene como una combinación lineal de valores pasados reales y predichos de la propia variable. La metodología descrita en Box and Jenkins (1970) ofrece un proceso de identificación sistemático para este tipo de modelos que ha sido testada en multitud de áreas.

La expresión general del modelo ARIMA (p, d, q) responde a la ecuación (2.7)

$$\phi_p(L)\nabla^d y_t = c + \theta_q(L)\varepsilon_t \quad (2.7)$$

con

$$\phi_p(L) = 1 + \phi_1 L + \dots + \phi_p L^p \quad (2.8)$$

$$\theta_q(L) = 1 + \theta_1 L + \dots + \theta_q L^q \quad (2.9)$$

donde y_t es la variable de salida, c es el término constante y ε_t denota un

proceso de ruido blanco¹⁴ que se conoce como error de predicción o innovación¹⁵. L es el operador retardo tal que $L^i y_t = y_{t-i}$, y $\phi_p(L)$ y $\theta_q(L)$ son polinomios en el operador retardo, de órdenes p y q respectivamente, los cuales modelan los efectos autorregresivos y de media móvil, respectivamente. ∇^d es el operador diferencia tal que $\nabla^d y_t = (1 - L)^d y_t$, siendo d un entero no negativo, que describe las diferencias aplicadas para inducir estacionariedad en procesos integrados.

El predictor óptimo de la variable de salida¹⁶ con origen de predicción en T y a un horizonte de predicción k , es decir, el predictor óptimo de y_{T+k} , es su esperanza condicionada a los datos observados (ver Peña (2005) para detalles). Llamando $\varphi(L) = \phi_p(L)\nabla^d$ al operador de orden $h = p + d$ que se obtiene multiplicando el operador AR(p) por las diferencias, la expresión de la variable y_{T+k} es

$$y_{T+k} = c - \varphi_1 y_{T+k-1} - \cdots - \varphi_h y_{T+k-h} + \theta_1 \varepsilon_{T+k-1} + \cdots + \theta_q \varepsilon_{T+k-q} \quad (2.10)$$

y por tanto, tomando esperanzas condicionadas a la información disponible hasta el instante T en todos los términos de la ecuación (2.10), obtenemos la predicción de la variable y_{T+k} ,

$$E[y_{T+k}|T] = \hat{y}_{T+k|T} = c - \varphi_1 \hat{y}_{T+k-1|T} - \cdots - \varphi_h \hat{y}_{T+k-h|T} + \theta_1 \hat{\varepsilon}_{T+k-1|T} + \cdots + \theta_q \hat{\varepsilon}_{T+k-q|T} \quad (2.11)$$

En esta ecuación algunas esperanzas se aplican a variables observadas (por lo que coinciden con las observaciones reales) y otras a no observadas, cuando $T+k-j > T$ con $j = 1, \dots, h$, dependiendo del horizonte de predicción k y el orden de la parte autorregresiva h . Con relación a las innovaciones, la esperanza de las innovaciones futuras (no observadas), es decir, cuando $T+k-j > T$ con $j = 1, \dots, q$ condicionada a la historia de la serie es igual a su esperanza absoluta, es decir 0, ya que el ruido del proceso es independiente de la variable de salida. El resto de innovaciones son conocidas, por lo que sus esperanzas coinciden con los valores observados, que se calculan recursivamente mediante la siguiente ecuación

¹⁴Se entiende de aquí en adelante por ruido blanco un proceso estocástico estacionario e incorrelado de media 0. Es decir, el proceso cumple que

1. $E(\varepsilon_t) = 0$
2. $Var(\varepsilon_t) = \sigma^2$
3. $E(\varepsilon_t \varepsilon_\tau) = 0 \quad \forall t \neq \tau$

¹⁵Se denota de aquí en adelante como innovación a la componente puramente aleatoria de un modelo de serie temporal. Nótese que no siempre coincide con el error de predicción ε_t definido en (2.7), como en el caso de los modelos GARCH, donde la innovación representa el error estandarizado, es decir, el error de predicción dividido por la varianza condicional del proceso.

¹⁶óptimo en el sentido que minimiza el error cuadrático medio.

$$\varepsilon_t = y_t + \varphi_1 y_{t-1} + \dots + \varphi_h y_{t-h} - c - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (2.12)$$

desde $t = 1$ hasta T^{17} , y de esta manera, se pueden calcular recursivamente las predicciones de $y_{T+1}, \dots, y_{T+k}$ condicionales a la información disponible hasta el instante T , comenzando con el horizonte 1 hasta el horizonte k deseado.

La extensión estacional de estos modelos resulta especialmente adecuada para acomodar los patrones diarios y semanales que presentan las series de precios eléctricos. La expresión general del modelo ARIMA $(p, d, q) \times (P_1, D_1, Q_1)_{S_1} \times \dots \times (P_m, D_m, Q_m)_{S_m}$ para series con m estacionalidades responde entonces a 2.13:

$$\phi_p(L) \Phi_{P_1}(L^{S_1}) \dots \Phi_{P_m}(L^{S_m}) \nabla^d \nabla_{S_1}^{D_1} \dots \nabla_{S_m}^{D_m} y_t = c + \theta_q(L) \Theta_{Q_1}(L^{S_1}) \dots \Theta_{Q_m}(L^{S_m}) \varepsilon_t \quad (2.13)$$

con

$$\Phi_{P_s}(L^{S_s}) = 1 + \Phi_1 L^{S_s} + \dots + \Phi_{P_s} L^{P_s * S_s} \quad (2.14)$$

$$\Theta_{Q_s}(L^{S_s}) = 1 + \Theta_1 L^{S_s} + \dots + \Theta_{Q_s} L^{Q_s * S_s} \quad (2.15)$$

donde $\Phi_{P_s}(L^{S_s})$ y $\Theta_{Q_s}(L^{S_s})$ son polinomios en el operador retardo, de órdenes P_s, Q_s respectivamente, que modelan los efectos autorregresivos y de media móvil estacionales de período S_s , con $s = 1, \dots, m$. $\nabla_{S_s}^{D_s}$ son operadores diferencia $(\nabla_{S_s}^{D_s} y_t = (1 - L^{S_s})^{D_s} y_t)$ que describen las diferencias estacionales aplicadas para inducir estacionariedad, siendo D_s enteros no negativos.

El modelo para series de precios horarios descrito en (2.13), que acomoda la doble estacionalidad diaria y semanal (es decir, con $S_1 = 24, S_2 = 168$), supone utilizar un mismo juego de parámetros para todas las horas del día. Este es el enfoque adoptado en algunas publicaciones, como por ejemplo en Contreras et al., 2003; Conejo et al., 2005. Sin embargo, existe otra manera de modelar el precio de cada hora que es igualmente popular. Así, por ejemplo, Cuaresma et al., 2004; Zareipour et al., 2006b; Misiorek et al., 2006; García-Martos et al., 2007 utilizan un modelo ARIMA (o sus particularizaciones AR o ARMA) diferente para cada una de las series correspondientes a las 24 horas del día, lo que da lugar a 24 modelos de series diarias (con una sola estacionalidad de período $S_1 = 7$). Este segundo enfoque presenta la ventaja de utilizar un juego diferente de parámetros para cada hora del día, lo cual resulta beneficioso para modelar las relaciones no lineales existentes en la series de precios. Sin embargo, tiene el inconveniente de ignorar las relaciones existentes entre los precios de horas diferentes, lo que resulta perjudicial de cara a la predicción cuando éstas son significativas. Ambos enfoques se diferencian también

¹⁷es necesario sustituir las primeras q innovaciones por su esperanza, que es 0, y calcular las demás innovaciones con esta condición inicial.

en la manera en la que las predicciones del día siguiente son calculadas a partir de la información recopilada hasta la hora 24 del día origen de la predicción. Así, en el enfoque horario es necesario calcular las predicciones de las 24 horas del día siguiente $\hat{y}_{T+1|T}, \dots, \hat{y}_{T+24|T}$ como en (2.11), mientras que el cálculo es más sencillo en los modelos horarios puesto que el horizonte de predicción es siempre 1 y no es necesario tomar esperanzas puesto que todas las variables son observadas.

Actualmente los modelos ARIMA, especificados de una o otra manera, son usados como modelos de referencia o *benchmark* para testear las capacidades de predicción de otros modelos (*ver p.ej.* Swider and Weber, 2007).

2.2.2.2. Modelo de alisado exponencial con doble estacionalidad

Los métodos de suavizado exponencial constituyen una familia de modelos univariantes ampliamente utilizados para la predicción a corto plazo de series temporales (Gardner, 1985, 2006). Estos modelos suponen que la serie temporal se genera como una composición de diferentes efectos, como son la tendencia, el nivel, la componente estacional o ciclo y la componente puramente aleatoria o innovación. En el pasado, surgieron como alternativa a los métodos de descomposición que suponían tendencias y ciclos deterministas, y siguen siendo motivo de investigación hasta la fecha (Hyndman et al. (2008)).

Sin embargo, su uso no se ha extendido al campo de la predicción de precios eléctricos, probablemente debido al hecho de que no fueron originalmente diseñados para tratar series que presenten doble estacionalidad. Recientemente, en Taylor (2003) se introduce una versión del método de *Holt-Winters* capaz de acomodar este comportamiento doblemente estacional, y se aplica con éxito para la predicción a corto plazo de series de demanda eléctrica horaria en Taylor et al., 2006; Taylor, 2008. La formulación de dicha extensión del modelo de *Holt-Winters* para series horarias con doble estacionalidad diaria y semanal responde a

$$S_t = \alpha(y_t/(D_{t-24}W_{t-168})) + (1 - \alpha)(S_{t-1} + T_{t-1}) \quad (2.16)$$

$$T_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)T_{t-1} \quad (2.17)$$

$$D_t = \delta(y_t/(S_tW_{t-168})) + (1 - \delta)D_{t-24} \quad (2.18)$$

$$W_t = \omega(y_t/(S_tD_{t-24})) + (1 - \omega)W_{t-168} \quad (2.19)$$

$$\hat{y}_{t+k|t} = (S_t + kT_t)D_{t-24+k}W_{t-168+k} + \phi^k(y_t - ((S_{t-1} + T_{t-1})D_{t-24}W_{t-168})) \quad (2.20)$$

donde $\alpha, \gamma, \delta, \omega$ y ϕ son parámetros del modelo, S_t denota el nivel de la serie, T_t la tendencia, D_t la componente estacional diaria y W_t la componente estacional semanal. El término en el que está involucrado el parámetro ϕ en (2.20) se propone debido a la autocorrelación de primer orden encontrada tras la inspección de los errores de predicción a un paso de la demanda eléctrica horaria.

2.2.2.3. Modelos de función de transferencia

Existe abundante literatura mostrando la buena capacidad de predicción de los modelos de función de transferencia con ruido ARIMA o LTF (*Linear Transfer Function*) en diferentes mercados eléctricos: Nogales and Conejo (2006) y Conejo et al. (2005) en PJM (Interconexión Pennsylvania-New Jersey-Maryland), Nogales et al. (2002) en OMEL (Mercado Ibérico), Zareipour et al. (2006b) en Ontario, Misiolek et al. (2006) en CalPX (California). Estos modelos generalizan a los ARIMA incluyendo la influencia de los valores presentes y pasados de otras series temporales. Su éxito radica en una metodología de identificación y diagnosis bien establecida (ver Box and Jenkins (1970) y Pankratz (1991)), y un reducido número de parámetros que es fácilmente interpretable. Así, El modelo LTF para series horarias de precios (con doble estacionalidad, diaria y semanal) presenta la siguiente expresión general (2.21):

$$y_t = \sum_i^n G_i(L)x_{it} + \frac{\theta_q(L)\Theta_{Q_1}(L^{24})\Theta_{Q_2}(L^{168})}{\phi_p(L)\Phi_{P_1}(L^{24})\Phi_{P_2}(L^{168})\nabla^d\nabla_{24}^{D_1}\nabla_{168}^{D_2}}\varepsilon_t \quad (2.21)$$

siendo $x_{1t}, \dots, x_{nt}$ el conjunto de variables explicativas y donde $G_1, \dots, G_n$ representa una familia de funciones de transferencia lineales capaces de capturar un gran número de respuestas impulsionales con pocos parámetros (ver Pankratz, 1991). Estas funciones de transferencia responden a (2.22)¹⁸

$$G_i(L) = \frac{(w_{i0} + w_{i1}L + w_{i2}L^2 + \dots + w_{ig_i}L^{g_i})L^{b_i}}{1 + \delta_{i1}L + \delta_{i2}L^2 + \dots + \delta_{ir_i}L^{r_i}} \quad (2.22)$$

Una particularización de los modelos LTF ampliamente utilizada en la literatura son los modelos de regresión dinámica o ARIMAX (ARIMA con variables exógenas). La diferencia con el modelo LTF general es que no presenta términos en el denominador de las funciones de transferencia correspondiente a las variables explicativas, es decir $r_i = 0, \forall i$. Esta simplificación permite evaluar el modelo de una manera más sencilla, aunque a expensas de incrementar artificialmente el número de parámetros para el caso de algunos tipos de respuesta impulsional, lo que redundaría en una merma en las capacidades de generalización del modelo.

El sistema lineal de la señal de salida y_t descrito en (2.21) se escribe de manera equivalente de la siguiente manera (ver Ljung (1986))

$$y_t = \sum_i G_i(L)x_{it} + H(L)\varepsilon_t \quad (2.23)$$

¹⁸Nótese que en esta formulación puede incluirse una constante en el modelo utilizando una variable *dummy* con valor 1 para todo t y una función de transferencia asociada que incluya un único coeficiente en el numerador.

siendo

$$H(L) = \frac{\theta_q(L)\Theta_{Q_1}(L^{24})\Theta_{Q_2}(L^{168})}{\phi_p(L)\Phi_{P_1}(L^{24})\Phi_{P_2}(L^{168})\nabla^d\nabla_{24}^{D_1}\nabla_{168}^{D_2}} = \sum_{l=0}^{\infty} \psi_l L^l \quad (2.24)$$

donde los ψ'_l s se corresponden con los coeficientes de la representación del término $H(L)\varepsilon_t$ como un proceso de media móvil de orden infinito MA(∞). La predicción de la señal de salida a un paso puede calcularse entonces de la siguiente manera (2.25)

$$\hat{y}_{t|t-1} = \sum_i H^{-1}(L)G_i(L)x_{it} + [1 - H^{-1}(L)]y_t \quad (2.25)$$

donde el valor de la variable y_t en el instante t desaparece puesto que se anula y solo intervienen sus retardos (y_{t-d} , con $d > 0$). Así mismo, la predicción a horizonte k es

$$\hat{y}_{t|t-k} = \sum_i W_k(L)G_i(L)x_{it} + [1 - W_k(L)]y_t \quad (2.26)$$

siendo

$$W_k(L) = \left(\sum_{l=0}^{k-1} \psi_l L^l \right) H^{-1}(L) \quad (2.27)$$

Las variables exógenas típicamente utilizadas para la predicción del precio con modelos LTF son la previsión de demanda eléctrica (p.ej. Bunn and Karakatsani (2003)) y variables *dummies* que capturan los efectos de calendario, es decir, fines de semana y días festivos (p.ej. Weron and Misiorek (2008)). Otras variables han sido también utilizadas en función de sus efectos sobre el precio dentro de cada mercado específico, como por ejemplo, el exceso de capacidad o las predicciones de exportaciones e importaciones (Zareipour et al. (2006b)) en Ontario. En Knittel and Roberts (2001) se modela la relación no lineal entre el precio y la temperatura de California utilizando un modelo LTF que utiliza la temperatura y el cuadrado y el cubo de la misma. Un trabajo publicado del autor previo a esta tesis muestra el impacto que tiene la previsión de la generación eólica en las predicciones de los precios eléctricos de OMEL (Cruz et al. (2011)).

2.2.2.4. Modelos vectoriales

Si consideramos los valores correspondientes a las 24 horas de un día como una observación diaria compuesta de 24 componentes, es posible plantear un modelo distinto para cada hora del día que considere las relaciones existentes entre horas

diferentes. De forma análoga al caso univariable lineal, los modelos VAR (*Vector AR*) o VARX (*Vector AR* con variables exógenas) sirven a tal propósito. La expresión general de cada componente del vector de salida en un modelo VARX de orden p responde a (2.28)

$$y_{d,h} = c_h + \sum_{k=1}^p \sum_{j=1}^{24} \phi_{h,j,k} y_{d-k,j} + \sum_i \sum_{j=1}^{24} \beta_{h,j,i} x_{i,d,j} + \varepsilon_{d,h} \quad (2.28)$$

donde $y_{d,h}$ es el precio del día d y la hora h ($h = 1, \dots, 24$), $x_{i,d,h}$ son las variables exógenas y $\varepsilon_{d,h}$ cada una de las 24 componentes de un vector de ruido blanco¹⁹.

Para simplificar la notación, si se definen $y_d = (y_{d,1}, \dots, y_{d,24})'$, $C = (c_1, \dots, c_{24})'$, $x_{i,d} = (x_{i,d,1}, \dots, x_{i,d,24})'$, $\varepsilon_d = (\varepsilon_{d,1}, \dots, \varepsilon_{d,24})'$ y

$$A_k = \begin{pmatrix} \phi_{1,1,k} & \cdots & \phi_{1,24,k} \\ \vdots & \ddots & \vdots \\ \phi_{24,1,k} & \cdots & \phi_{24,24,k} \end{pmatrix}$$

$$B_i = \begin{pmatrix} \beta_{1,1,i} & \cdots & \beta_{1,24,i} \\ \vdots & \ddots & \vdots \\ \beta_{24,1,i} & \cdots & \beta_{24,24,i} \end{pmatrix}$$

entonces el modelo definido en (2.28) puede ser escrito de forma compacta como en (2.29)

$$y_d = C + \sum_{k=1}^p A_k y_{d-k} + \sum_{i=1} B_i x_{i,d} + \varepsilon_d \quad (2.29)$$

Para el caso de la aplicación de estos modelos a la predicción de los precios horarios del día siguiente, el horizonte de predicción es 1 (se predice el día d a partir de la información conocida hasta el día $d-1$ ²⁰). Es decir, el vector de precios estimado para el día siguiente se calcula de la siguiente manera (2.30)

$$\hat{y}_{d|d-1} = C + \sum_{k=1}^p A_k y_{d-k} + \sum_{i=1} B_i x_{i,d} \quad (2.30)$$

y de este modo el horizonte usado para predecir cada una de las 24 horas del día siguiente coincide con el horizonte usado al estimarlas. Esta circunstancia representa una diferencia importante con el enfoque horario (ver (2.13) o (2.21)), puesto que

¹⁹Es decir, ε_d es un vector de 24 componentes tal que $E(\varepsilon_d) = 0$, $E(\varepsilon_d \varepsilon_\tau') = Q$ para $d = \tau$, $E(\varepsilon_d \varepsilon_\tau') = 0 \ \forall d \neq \tau$.

²⁰Nótese que las variables exógenas se consideran conocidas en cualquier instante de tiempo

en este último caso los precios se estiman para la hora t a partir de la información conocida hasta la hora $t - 1$, y la predicción de los 24 precios del día siguiente se realiza de manera recursiva desde la hora 1 hasta la 24, por lo que la incertidumbre asociada a la predicción de cada hora afecta a la siguiente.

Sin embargo, el uso de este tipo de modelos conlleva la estimación de un gran número de parámetros, lo que merma considerablemente sus capacidades de generalización. Además, dado que las series de precios no son estacionarias, se hace necesario el estudio de la cointegración entre las 24 series horarias²¹. En general, el proceso de identificación de este tipo de modelos es mucho más complejo que para el caso univariable (ver Lütkepohl (2005), Part II Cointegrated Processes). Así, existen pocos ejemplos en la literatura que apliquen este tipo de modelos. Recientemente, Panagiotelis and Smith (2008) aplica un VARX de primer orden a la serie de precios horaria del mercado eléctrico de Australia donde las matrices de coeficientes son de tipo *sparse*, es decir, presentan gran cantidad de coeficientes restringidos a 0. En Koopman et al. (2010) y Carnero et al. (2010) se presenta una versión vectorial del modelo periódico descrito en la sección 2.2.2.6. El trabajo de Joutz et al. (1995), compara la capacidad predictiva de modelos VAR frente a modelos VECM (*Vector Error Correction Models*) o BVAR (*Bayesian VAR*), los cuales responden a la misma formulación del VAR descrita en (2.29) pero incorporan restricciones sobre los parámetros, para la predicción de series mensuales de consumo y precios eléctricos.

Otra posible solución al elevado número de parámetros de los modelos vectoriales son los modelos conocidos como de Factores Dinámicos (*Dynamic Factor Models*) (Peña and Poncela (2006)). El modelo propuesto en Alonso et al. (2008) para la predicción de precios eléctricos responde a la siguiente formulación (2.31)

$$y_d = \Omega f_d + \varepsilon_d \quad (2.31)$$

donde se asume que el vector de precios diarios y_d puede ser descrito por una combinación lineal de componentes no observados o factores específicos f_d mas un ruido blanco ε_d . La matriz Ω es de dimensión $24 \times r$, siendo $r < 24$ el número de factores específicos considerados. De esta manera, el conjunto reducido de factores comunes describe las dinámicas regulares y estacionales de la serie de precios, puesto que f_d sigue un proceso VARIMA $(p, d, q) \times (P, D, Q)_s$ (Vector AutoRegressive Integrated Moving Average) que responde a

$$(1 - L)^d(1 - L^s)^D\phi(L)\Phi(L^s)f_d = C + \theta(L)\Theta(L^s)w_d \quad (2.32)$$

²¹La aplicación de un modelo VAR/VARX sobre el vector de precios donde cada hora es diferenciada individualmente podría distorsionar las verdaderas relaciones existentes entre las componentes del vector de precios.

Dos procesos no estacionarios están cointegrados si existe una combinación lineal entre ellos que es estacionaria.

donde $\phi(L)$, $\Phi(L^s)$, $\Theta(L)$ y $\theta(L)$ son matrices $r \times r$ de polinomios en el operador retardo L tal que

$$\phi(L) = (I + \phi_1 L + \phi_2 L^2 + \cdots + \phi_p L^p) \quad (2.33)$$

$$\Phi(L^s) = (I + \Phi_1 L^s + \Phi_2 L^{2s} + \cdots + \Phi_p L^{Ps}) \quad (2.34)$$

$$\theta(L) = (I + \theta_1 L + \theta_2 L^2 + \cdots + \theta_p L^q) \quad (2.35)$$

$$\Theta(L) = (I + \Theta_1 L^s + \Theta_2 L^{2s} + \cdots + \Theta_p L^{Qs}) \quad (2.36)$$

y $w_d \sim N(0, Q)$ es un ruido serialmente incorrelado ($E(w_d w_\tau') = 0 \forall d \neq \tau$) que también está incorrelado con ε_d ($E(w_d \varepsilon_\tau') = 0$ para todo d y τ). Para que el modelo sea identificable es necesario imponer también $w_{ij} = 0$ $j > i$, donde $\Omega = [w_{ij}]$, y $Q = I \circ \Omega \Omega' = I$.

El modelo en (2.31)-(2.32) puede ser descrito de manera equivalente usando una formulación en Espacio de Estados como la propuesta en Kohn and Ansley (1986), que responde a (2.37)-(2.38),

$$y_d = \Omega f_d + \varepsilon_d \quad (2.37)$$

$$f_d = C + \Psi f_{d-1} + w_d \quad (2.38)$$

donde la matriz Ψ se construye en función de las matrices de polinomios $\phi(L)$, $\Phi(L^s)$, $\Theta(L)$ y $\theta(L)$. De esta manera los parámetros pueden ser estimados por máxima verosimilitud usando el filtro de Kalman²². Entre las debilidades de los modelos de Factores Dinámicos se encuentra la dificultad para identificar la estructura del modelo VARIMA dado que los factores del modelo no son observados, y la asunción de que la relación entre el vector de precios diarios y sus factores específicos se mantiene constante en el tiempo.

2.2.2.5. Modelos de heteroscedasticidad condicional GARCH

En los modelos ARIMA y LTF descritos en (2.13) y (2.21), se definió el término de error ε_t como un proceso de ruido blanco. Sin más conocimiento de la función de densidad de probabilidad de la variable de salida, es frecuente asumir que ε_t se distribuye como una normal de media cero y varianza constante, es decir, $\varepsilon_t \sim i.i.d N(0, \sigma^2)$. Para el caso de series temporales múltiples, esta hipótesis de normalidad es $\varepsilon_d \sim i.i.d N(\vec{0}, \Sigma)$ siendo Σ una matriz simétrica y semidefinida positiva. Sin embargo, como se señala en la sección 2.2.1, existe abundante literatura mostrando que las series de precios spot en mercados eléctricos se caracterizan por

²²En las secciones 4.2.1.1 y 4.2.1.2 se presenta una introducción al filtro de Kalman y la representación en Espacios de Estados propuesta en Hamilton (1994). Para otro tipo de representaciones ver, por ejemplo, de Jong and Penzer, 2004; Harvey, 1990; Hyndman et al., 2008.

presentar una estructura de la varianza compleja y no constante. Álvaro Escribano et al. (2002) y Knittel and Roberts (2001) encuentran que los precios presentan volatilidad agrupada (*volatility clustering*²³). Carnero et al. (2003) encuentra incluso diferencias en la volatilidad de la series de precios según el día de la semana. Estas circunstancias invalidan la hipótesis de homocedasticidad²⁴ asumida, lo que lleva a incorrectas especificaciones que no solo afectan a la estimación de la varianza del proceso sino que también, aunque habitualmente en menor medida, a la estimación de su media.

Algunos autores se han preocupado de modelar la varianza del los precios eléctricos asumiendo que el término de error del proceso del precio sigue un proceso GARCH (Generalized AutoRegressive Conditional Heteroskedasticity) (Bollerslev, 1986; Engle, 1982). De esta manera, se asume que existe parte de la varianza de la serie que es predecible y función lineal del cuadrado de los errores y varianzas en instantes de tiempo anteriores. Los modelos GARCH resultan especialmente válidos cuando la característica de *volatility clustering* está presente y han sido utilizados para modelar la varianza de multitud de series de índole financiera Engle (2001). Además, resultan especialmente pertinentes para modelar los precios eléctricos puesto que conducen a distribuciones con colas pesadas. La expresión general del término de error ε_t de un modelo GARCH(p,q) responde a (2.39)-(2.40)

$$\varepsilon_t = \sigma_t \epsilon_t \quad (2.39)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (2.40)$$

siendo α_i y β_j parámetros del modelo. En este caso, el término de error ε_t es el producto de un ruido blanco de varianza 1, conocido como error estandarizado o innovación, ϵ_t , y un término predecible σ_t . Si se define el proceso de ruido blanco v_t ²⁵ tal que

$$v_t = \varepsilon_t^2 - \sigma_t^2 \quad (2.41)$$

entonces la dependencia del cuadrado del término de error ε_t puede expresarse como un proceso ARMA

²³volatility clustering se refiere a la característica presente en las series temporales dónde grandes cambios tienden a ser seguidos de grandes cambios y pequeños cambios tienden a ser seguidos por pequeños cambios.

²⁴Se dice que existe homocedasticidad cuando la varianza de los errores estocásticos es la misma para cada observación en t . Decimos que existe heterocedasticidad cuando no se cumple esta situación.

²⁵Se demuestra que v_t es un proceso de ruido blanco, es decir, un ruido incorrelado de media 0 y varianza marginal constante. Sin embargo, no es un proceso de ruido blanco en sentido estricto, puesto que aunque es incorrelado no es independiente (ver Peña (2005) para detalles).

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) \varepsilon_{t-i}^2 - \sum_{j=1}^q \beta_j v_{t-j} + v_t \quad (2.42)$$

y de esta manera caben hacer las mismas consideraciones que para los modelos ARMA previamente descritos en cuanto al cálculo de las predicciones y la incorporación de términos estacionales. Se demuestra que si $\varepsilon_t \sim i.i.d N(0,1)$, el coeficiente de curtosis de la distribución de ε_t es siempre mayor que 3, por lo que su distribución marginal²⁶ será de colas pesadas (ver Peña (2005) para detalles).

El modelo GARCH y sus variantes (ver Bollerslev (2008) para una lista detallada de variantes del modelo GARCH) ha recibido mucha atención en el campo del análisis de riesgos en mercados eléctricos (ver p.ej. Chan and Gray (2006)). Con respecto a la predicción de los precios spot, ha sido utilizado en combinación con otros modelos de series temporales para modelar simultáneamente la media y varianza de la serie. Así, García et al. (2005) especifica modelos AR-GARCH para los precios del mercado Español y Californiano. Misiorek et al. (2006) aplica AR-GARCH y ARX-GARCH a los precios en CalPX. Knittel and Roberts (2001) especifica un AR-EGARCH (Exponential GARCH) para capturar la asimetría de la distribución del término del error de predicción de un modelo AR. Koopman et al. (2007) perfecciona el modelo propuesto previamente en Carnero et al. (2003) añadiéndole un término GARCH para modelar los precios medios diarios de diferentes mercados eléctricos europeos.

Los términos GARCH adicionales en cualquier modelo de serie temporal afectan a las predicciones puntuales en cuanto que influyen en la estimación de los parámetros que determinan el valor esperado de la variable de salida. A este respecto, no está claro si dichos términos GARCH adicionales contribuyen a predecir mejor el valor puntual o esperado de los precios eléctricos. García et al. (2005) señala que el modelo AR-GARCH mejora a su equivalente homocedástico AR solo en los períodos en los que los precios del mercado español y californiano presentan alta volatilidad. Misiorek et al. (2006) encuentra que el término GARCH adicional de los modelos empeora las predicciones puntuales en CalPX. En este mismo artículo se evalúa también la capacidad del modelo AR-GARCH para proveer intervalos de predicción de los precios spot a los niveles de confianza del 50 %, 90 % y 99 %, no encontrando una mejora significativa frente a otros modelos de series temporales que no consideran la varianza como una función dinámica.

2.2.2.6. Modelos periódicos

La diferente estructura de autocorrelación observada en las series de precios según el día de la semana llevan a (Koopman et al. (2007)) a especificar un modelo periódico

²⁶Se llama distribución marginal de una variable z_t a la distribución de la variable sin el conocimiento de su trayectoria hasta el instante t . Por contra, la distribución condicionada de z_t dados $z_{t-1}, \dots, z_{t-k}$ es la distribución que presenta la variable cuando se conocen los k valores anteriores del proceso.

para los precios diarios de distintos mercados europeos. Los modelos periódicos son una extensión de los modelos anteriormente expuestos permitiendo que se aplique un juego diferente de parámetros en cada predefinida estación $s \in \{1, \dots, S\}$ por las que transcurre la serie temporal (ver Franses and Paap (2004)). De esta forma, el modelo es capaz de captar comportamientos no lineales periódicos que los anteriores modelos no podían. Además, el modelo descrito en Koopman et al. (2007) añadió otras características novedosas con respecto a los ya existentes en la literatura, como son una forma alternativa para modelar el comportamiento de memoria larga que exhiben las series de precios, la inclusión de variables exógenas en el modelo para la varianza condicional y la especificación de la función de densidad del término estocástico como una distribución t-Student.

La especificación del modelo Reg-ARFIMA-GARCH (AutoRegresive Fractionally Integrated Moving Average with Exogenous Variables and GARCH disturbance) estacional periódico descrito en Koopman et al. (2007) responde a (2.43) - (2.46)

$$(1 - L^S)^{D_{s_t}}(y_t - u_t) = \varepsilon_t \quad (2.43)$$

dónde s_t toma el valor de la estación s en el instante t , los días de la semana en este caso, y el período estacional es $S = 7$ puesto que son precios diarios. Como alternativa a la inclusión de un elevado número de términos autorregresivos estacionales que capturen la autocorrelación presente en los retardos múltiplos del período estacional (lo que caracteriza a un proceso de memoria larga), se asume que $(1 - L^S)^{D_{s_t}}(y_t - u_t)$ es una serie fraccionalmente integrada, siendo $|D_{s_t}| < 0.5$. El operador de diferenciación en este caso responde a (2.44):

$$(1 - L^S)^{D_{s_t}} = \sum_{i=1}^{\infty} \frac{\Gamma(i - D_{s_t})}{\Gamma(-D_{s_t})\Gamma(i + 1)} L^{is} \quad (2.44)$$

donde $\Gamma(\cdot)$ es la función gamma. Además, la media u_t y varianza σ_t condicionales del proceso se definen como en (2.45) y (2.46), respectivamente.

$$u_t = \phi_{1s_t}y_{t-1} + \dots + \phi_{ps_t}y_{t-p} + \sum_{k=1}^K \delta_{k0s_t}x_{k,t} + \delta_{k1s_t}x_{k,t-1} + \dots + \delta_{km_s_t}x_{k,t-m} \quad (2.45)$$

$$\sigma_t^2 = \exp(\lambda_{s_t})h_t \quad h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1h_{t-1} + \sum_{k=1}^{K^*} \gamma_k z_{kt} \quad (2.46)$$

dónde ϕ_{is_t} para $i = 1, \dots, p$ y δ_{kis_t} para $k = 1, \dots, K$ y $i = 0, 1, \dots, m$ son coeficientes de regresión periódicos que se aplican a los valores pasados de y_t y valores actuales y pasados de las variables exógenas x_{kt} para modelar la media condicional. $\alpha_0, \alpha_1, \beta_1, \gamma_k$ para $k = 1, \dots, K^*$ modelan la estructura de la varianza condicional

de la serie como un modelo GARCH(1,1) que también incluye la influencia de las variables explicativas z_{kt} . Los parámetros $\lambda_1, \dots, \lambda_S$ permiten escalar la varianza según la estación s_t .

La última característica novedosa del modelo es que asume que la componente estocástica se distribuye como una t-Student de v grados de libertad, media 0 y varianza 1. La razón se encuentra en que a pesar de que asumir normalidad en las innovaciones del proceso conduce a distribuciones con colas pesadas, las distribuciones empíricas de dichas innovaciones siguen sin ajustarse adecuadamente a la distribución teórica y presentan colas más pesadas. Este hecho es una de las características distintivas de las series de precios en mercados eléctricos.

2.2.2.7. Modelos por umbrales

De manera similar a los modelos periódicos (ver sección 2.2.2.6), los modelos por umbrales constituyen una extensión de los modelos lineales anteriormente expuestos (ver 2.2.2.3) en los que se aplica un juego diferente de parámetros en cada instante t según toma valor una variable observable en relación a un valor umbral. Los modelos TAR (*Threshold-AR*) originalmente propuestos por Tong (1983) responden a la formulación descrita en (2.47)

$$\begin{cases} \phi_1(L)y_t = \varepsilon_{1,t} & \text{si } v_t \geq U \\ \phi_2(L)y_t = \varepsilon_{2,t} & \text{si } v_t < U \end{cases} \quad (2.47)$$

donde v_t es una variable exógena, existen dos regímenes o juegos diferentes de parámetros ($\phi_i(L) = 1 - \phi_{i,1}L - \dots - \phi_{i,p}L^p, i = 1, 2$) determinados por un valor umbral U ²⁷ y $\varepsilon_{i,t}$ son variables aleatorias i.i.d. . Cuando la variable exógena que determina el régimen es un retardo de la propia variable de salida, es decir $v_t = y_{t-d}$, el modelo es conocido como *SETAR* (Self-Exciting TAR).

Los modelos por umbrales constituyen una oportunidad para modelar la diferente estructura de autocorrelación que presenta los precios para, por ejemplo, valores altos o bajos, o para modelar la respuesta no lineal del precio a cambios en una determinada variable explicativa, como puede ser la demanda eléctrica. Así, en Misiorek et al. (2006) se aplican modelos TAR y TARX (TAR con variables exógenas) para los precios horarios en CalPX (California Power Exchange) y se compara su rendimiento con una batería de modelos entre los que se incluyen AR/ARX, ARX-GARCH, y modelos de Markov (ver sección 2.2.2.8), encontrando que los modelos TAR/TARX son los más precisos en cuanto a predicciones puntuales se refiere. En el citado trabajo, la variable que determina el umbral es la diferencia entre el precio medio del día anterior y el de ocho días antes, mientras que en Weron (2006) se aplica un modelo similar en el que la variable que determina el umbral es la demanda eléctrica. También en Rambharat et al. (2005) se utiliza un modelo de tipo SETAR para modelar los

²⁷La formulación en (2.47) puede generalizarse fácilmente para más de dos regímenes

precios máximos diarios en PJM en donde se tiene en cuenta también la influencia de la temperatura.

A esta categoría de modelos pertenecen también los modelos de transición suave o STAR (Smooth Transition AR) (Terasvirta, 1990; Terasvirta et al., 2005). A diferencia del caso de otros tipos de modelos no lineales, existe un conjunto de herramientas estadísticas que pueden ser empleadas para su identificación y diagnóstico (ver Eitrheim and Terasvirta, 1996; Strikholm and Terasvirta, 2005; Luukkonen et al., 1988). En estos modelos, el juego de coeficientes aplicado en cada instante t varía según una determinada función de transición suave G_t típicamente acotada entre 0 y 1, lo que permite una transición gradual entre los regímenes. Una elección frecuente de la función G_t es la función logística, dando lugar a los modelos conocidos como LSTAR (Logistic STAR) (ver 2.48-2.49)

$$y_t = \phi_p(L)y_t + \Phi_p(L)G_t y_t + \varepsilon_t \quad (2.48)$$

siendo

$$G_t = G(v_t, \gamma, U) = 1/(1 + \exp -\gamma(v_t - U)) \quad (2.49)$$

donde γ es el parámetro suavizado que determina la transición entre regímenes y p es el orden de los polinomios autoregresivos $\phi_p(L)$ y $\Phi_p(L)$.

No obstante, existen pocas referencias de la aplicación de estos modelos al precio de la electricidad, probablemente debido al hecho contradictorio que puede suponer el modelar transiciones suaves entre distintos regímenes, puesto que en muchas ocasiones los precios eléctricos experimentan cambios abruptos. Como excepción, se encuentra el trabajo de Robinson (2000), donde se señala que el grado de reversión a la media de los precios del pool de los mercados inglés y galés cambia con el nivel de la serie, y se utilizan modelos LSTAR para describir dicho comportamiento.

2.2.2.8. Modelos de cambio de régimen

En los modelos de series temporales con cambio de régimen, se asume que la serie temporal transcurre a lo largo de un pequeño número de regímenes no observables y mutuamente excluyentes R , de tal manera que $r_t \in \{1, \dots, R\}$, con $t = 1, \dots, T$, indica el régimen en el instante t , y el modelo que describe la serie temporal utiliza un juego diferente de parámetros en cada uno de los regímenes. La diferencia con los modelos periódicos o los de tipo TAR anteriormente descritos en las secciones 2.2.2.6 y 2.2.2.7 radica en el hecho de que en aquellos el régimen es una variable que está predefinida o está determinada por una variable observable y un cierto valor umbral, y por tanto los regímenes que han ocurrido en el pasado y ocurren en el presente son conocidos con certeza. Por el contrario, en los modelos de cambio de régimen, el mecanismo de *switching* o cambio entre regímenes está gobernado por

una variable aleatoria discreta y no observable r_t , y nunca se sabe con certeza qué régimen ocurre en cada instante t . Resulta necesario por tanto asumir un determinado comportamiento para la variable estocástica r_t , y así estimar o asignar probabilidades a sus ocurrencias.

La manera más simple de especificar el comportamiento del régimen r_t es asumir que evoluciona independientemente de su propio pasado y cualquier otra variable explicativa. De este modo, los parámetros adicionales del modelo son las probabilidades de ocurrencia invariantes en el tiempo de cada uno de los regímenes

$$Pr(r_t = i | \eta_{t-1}) = Pr(r_t = i) = q_i \quad (2.50)$$

con $\sum q_i = 1$ y $0 \leq q_i \leq 1$ y donde η_{t-1} representa la información conocida hasta el instante $t - 1$. Esta manera de especificar el mecanismo de cambio de régimen es una de las adoptadas en Swider and Weber (2007) para representar la distribución de probabilidad de los precios spot y de reserva en el mercado eléctrico alemán. En dicho artículo, se asumen dos regímenes de comportamiento diferentes ($i = 1, 2$) y la distribución de probabilidad del precio en cada uno de ellos responde a una normal cuya media y varianza siguen un proceso de tipo ARMA-GARCH, $y_t | r_t = i \sim i.i.d N(u_{i,t}, \sigma_{i,t}^2)$. Así, la función de densidad de probabilidad en cada régimen responde a

$$f[y_t | r_t = i; \eta_{t-1}] = \frac{1}{\sqrt{2\pi}\sigma_{i,t}} \exp \left[\frac{-(y_t - u_{i,t})^2}{2\sigma_{i,t}^2} \right] \quad (2.51)$$

donde $u_{i,t}$ y $\sigma_{i,t}^2$ son la media y varianza condicionales del modelo correspondiente al régimen i . La distribución de probabilidad del precio resultante de un modelo así especificado es una mixtura de las distribuciones asumidas en cada uno de los regímenes con pesos q_i , como se indica en (2.52)

$$f[y_t | \eta_{t-1}] = \sum_{i=1}^R f[y_t, r_t = i] = \sum_{i=1}^R q_i f[y_t | r_t = i] \quad (2.52)$$

El citado artículo señala que dicha representación resulta más adecuada que la de otros modelos ARMA o ARMA-GARCH con distribuciones gaussianas (es decir, con un único régimen), puesto que la mixtura es capaz de capturar más fielmente las colas pesadas y el comportamiento asimétrico que típicamente presentan las distribuciones empíricas de los precios.

No obstante, los modelos de cambio de régimen han sido más utilizados para modelar de manera separada el proceso estocástico de emisión de valores normales de precios del que emite altos precios o *spikes*. Para tal fin, el hecho de que las probabilidades de ocurrencia de cada uno de los regímenes sean constantes a lo largo

del tiempo supone una importante limitación a la hora de representar, por ejemplo, el mecanismo que cambia de un régimen de emisión de *spikes* a otro régimen normal y viceversa. Así, la inmensa mayoría de los modelos de cambio de régimen que modelan los precios de la electricidad asumen que el régimen sigue un proceso de Markov de primer orden dónde el valor del régimen r_t depende de r_{t-1} . En este caso, se utiliza una matriz de transición Q que contiene las probabilidades de transición de un régimen a otro

$$Q = (q_{ij}) = \begin{pmatrix} q_{11} & \cdots & q_{1R} \\ \vdots & \ddots & \vdots \\ q_{R1} & \cdots & q_{RR} \end{pmatrix}$$

donde $q_{ij} = Pr(r_t = j | r_{t-1} = i)$ denota la probabilidad de cambiar del régimen i en el instante $t - 1$ al régimen j en el instante t , con $\sum_i q_{ij} = 1 \forall j$ y $0 \leq q_{ij} \leq 1$.

Los diferentes modelos de Markov encontrados en la literatura difieren fundamentalmente en tres aspectos: el número de regímenes que presenta el modelo, las distribuciones de probabilidad asumidas dentro de cada régimen y la especificación de la matriz de probabilidades de transición Q . La mayoría de ellos presentan dos únicos regímenes, uno para modelar el comportamiento estable de los precios y otro para capturar los *spikes* (Jong, 2006b; Weron, 2006, 2009). Aunque también pueden encontrarse referencias que utilizan tres regímenes diferentes, como por ejemplo en Huisman and Mahieu, 2003; Janczura and Weron, 2010. En estos modelos existe un régimen de precios estables, otro de súbito incremento y otro de vuelta a la situación de precios estables, con una matriz de transición Q pertinentemente especificada. En cuanto a las distribuciones asumidas para cada uno de los regímenes, es común utilizar la distribución normal para el proceso de precios estables, mientras que para modelar los spikes se han utilizado diferentes distribuciones como la normal o log-normal, Pareto y la distribución de Poisson compuesta (Karakatsani and Bunn, 2008b; Janczura and Weron, 2009; Jong, 2006b). Respecto a la especificación de la matriz de probabilidades de transición Q , lo más habitual es asumir que permanece constante a lo largo del tiempo, aunque en Mount et al. (2006) y Huisman (2008) se utiliza una representación del mecanismo de cambio de régimen más realista que asume que varía con el tiempo y depende de variables explicativas como la temperatura, la demanda de energía o el margen de reserva. En este caso, cada elemento de la matriz suele especificarse a través de una función paramétrica, como por ejemplo, la función logística,

$$Pr(r_t = j | r_{t-1} = i; \eta_{t-1}; \varphi) = \exp(\beta_0 + \beta_1 z_t) / (1 + \exp(\beta_0 + \beta_1 z_t)) \quad (2.53)$$

dónde z_t es una variable explicativa y β_0, β_1 parámetros del modelo.

Para este tipo de modelos, y en general para los modelos no lineales, no existe una metodología bien establecida para identificar su estructura. Así, tanto el número de

regímenes como el número de parámetros de cada régimen han de ser determinados a priori, lo que supone una limitación a la hora de aplicar estos modelos para predecir series reales. Resulta una excepción el trabajo de Vucetic et al. (2001), donde se presenta un algoritmo capaz de determinar el número de regímenes que presenta la serie de precios a partir de sus diferentes relaciones con la demanda eléctrica. No obstante, el algoritmo presenta la fuerte restricción de no permitir frecuentes cambios de régimen y no considerar cualquier otra variable fundamental que afecte al precio.

A pesar del gran número de diferentes distribuciones que este tipo de modelos es capaz de representar y de su pertinencia para recoger algunas de las especiales características de los precios eléctricos, existe controversia sobre sus capacidades de predicción y de si realmente pueden superar el rendimiento de los modelos lineales como los ARIMA o las funciones de transferencia. El trabajo de Bessec and Bouabdallah (2005), que se basa en la descomposición del error de predicción de un modelo en sus diferentes fuentes propuesta en Krolzig (2004), pone de manifiesto que, en los modelos de Markov, la predicción del régimen futuro es el punto más crítico, y errores en dicha estimación afectan de manera dramática a su rendimiento. En Kosater and Mosler (2006), se compara el rendimiento de diferentes especificaciones de modelos de Markov con modelos autorregresivos lineales en el mercado eléctrico alemán, concluyendo que los modelos de Markov pueden mejorar las predicciones puntuales de los modelos lineales.

2.2.2.9. Modelos semiparamétricos

Se entiende por modelo semiparamétrico aquel en el que parte de la estructura del modelo estadístico (la función de densidad o alguno de los momentos condicionales de la serie, habitualmente media o varianza) es tratado de manera no paramétrica, es decir, sin asumir ninguna forma específica, mientras que el resto se especifica paramétricamente con una forma a priori como las descritas anteriormente en este capítulo. Como ya se comentó previamente en 2.2.1, la habitual hipótesis de normalidad (o log-normalidad) asumida en los modelos de series temporales no es adecuada para representar la distribución de probabilidad de los precios, puesto que las distribuciones empíricas obtenidas a menudo presentan asimetría o colas más pesadas que las representadas por una distribución gaussiana. Algunos autores han enfocado este problema asumiendo diferentes distribuciones a priori más verosímiles teniendo en cuenta las características del proceso estocástico del precio, como la *t-student* (Koopman et al. (2007)), *t-student* asimétrica (Panagiotelis and Smith (2008)), hiperbólica, inversa gaussiana o α -estable (Weron (2009)), aunque sin resultados plenamente satisfactorios, como se contrasta en el estudio comparativo en Weron and Misiorek (2007). En el trabajo expuesto en Weron and Misiorek (2008), los autores comparan el rendimiento de un conjunto de modelos gaussianos de tipo ARX y TARX con sus contrapartidas semiparamétricas, es decir, que no asumen ninguna distribución de probabilidad determinada para el ruido del proceso, para los mercados nórdico y de California. Este enfoque no paramétrico está inspirado en el

análisis previo de Cao et al. (2003). Así, la estimación de los parámetros se hace de la manera habitual, es decir, maximizando la verosimilitud de la muestra de datos disponible (2.54)

$$\hat{\varphi} = \arg \max_{\{\varphi\}} L(\varphi) = \Pi_{t=1}^T f[\epsilon_t | \eta_{t-1}; \varphi] \quad (2.54)$$

pero en lugar de utilizar la función de densidad normal estándar²⁸ como función de densidad $f[\epsilon_t | \eta_{t-1}; \varphi]$ (siendo ϵ_t las innovaciones del modelo ($\epsilon_t = (y_t - \hat{y}_{t|t-1})/\hat{\sigma}_{t|t-1}$)), los autores utilizan el estimador kernel de Parzen-Rosenblatt (Parzen, 1962; Rosenblatt, 1956).

$$f_h[\epsilon_t | \eta_{t-1}; \varphi] = \frac{1}{nh} \sum_{\tau=1}^T K\left(\frac{\epsilon_t - \epsilon_\tau}{h}\right) \quad (2.55)$$

donde el parámetro predefinido h representa el ancho de banda de la función kernel $K(\cdot)$, que puede ser cualquier función kernel como por ejemplo la función de densidad normal estándar. En el mencionado artículo, los autores llegan a la conclusión que el rendimiento de los modelos semiparamétricos generalmente es mejor que el de sus contrapartidas gaussianas en cuanto a predicciones puntuales y de intervalos se refiere. No obstante, el método presenta dos importantes desventajas: una es la sensibilidad del ajuste de los modelos al parámetro del ancho de ventana h , lo que hace necesario un método adecuado para su correcta selección (Jones et al. (1995)); el otro inconveniente es que no existe una forma analítica de la función de densidad, estando definido el valor de densidad asignado a cada punto en función de los puntos más cercanos al mismo.

Dentro del grupo de modelos semiparamétricos se encuentra también el modelo propuesto en Zhao et al. (2008), que responde a la siguiente formulación

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}, x_{it}) + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t \quad (2.56)$$

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^r \alpha_j \varepsilon_{t-j}^2 + \sum_{k=1}^m \beta_k x_{kt} \quad (2.57)$$

$$\varepsilon_t = \sigma_t \epsilon_t \quad (2.58)$$

$$\varepsilon_t \sim i.i.d N(0, 1) \quad (2.59)$$

y donde la función $f(\cdot)$ es una función no lineal cuya forma se determina a través de máquinas de vectores soporte (*Support Vector Machines*). Por tanto, el citado modelo combina una especificación no paramétrica para la media, una

²⁸Una función de densidad en su forma estándar tiene media 0 y varianza 1

especificación tipo GARCH para la varianza y utiliza la distribución normal estándar para representar la función de densidad del precio.

Una excelente revisión de métodos no paramétricos para estimar la media, varianza y la forma de la función de densidad condicionales de una serie temporal puede encontrarse en Härdle et al. (1997). En Pagan and Schwert (1990) se describen métodos no paramétricos alternativos a la familia ARCH/GARCH para modelar la volatilidad de series de índole financiera.

2.2.3. Modelos basados en redes neuronales

Dentro de las técnicas de aprendizaje automático, las redes neuronales son las que más atención han recibido para la predicción de precios spot en mercados eléctricos. En el contexto de métodos de predicción, las redes neuronales son usadas como modelos de regresión no lineal multivariable con capacidades de aproximador universal. Son especialmente adecuadas cuando las relaciones existentes entre las variables del sistema son complejas y desconocidas, y han sido utilizadas para modelar gran cantidad de procesos no lineales (Haykin, 1994; Zhang et al., 1998). No obstante, presentan algunas desventajas frente a algunos de los métodos paramétricos previamente descritos, como son el riesgo de sobreentrenamiento y la falta de interpretabilidad, lo que hace difícil encontrar soluciones parsimoniosas que generalicen adecuadamente.

La arquitectura de red feed-forward con una única capa oculta ha sido ampliamente utilizada (ver figura 2.7a). Su formulación responde a

$$y_t = \beta_0 + \sum_{j=1}^q \beta_j G(\gamma_{0j} + W_t^T \gamma_j) + \varepsilon_t \quad (2.60)$$

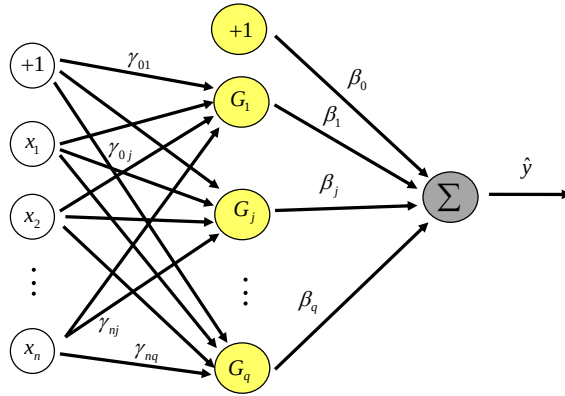
donde $W_t = (y_{t-1}, \dots, y_{t-p}, x_{1t}, \dots, x_{nt})^T$ es un vector compuesto de retardos de la variable de salida hasta el orden p y variables explicativas que pueden o no estar retardadas, $\gamma_j = (\gamma_{1j}, \dots, \gamma_{(n+p)j})^T$ es un vector de $((n+p) \times 1)$ parámetros para cada j -ésima neurona $j = 1, \dots, q$, y $\beta_0, \beta_1, \dots, \beta_q$ son parámetros escalares. La función $G : \mathbb{R} \rightarrow [0, 1]$ es una función de activación, que a menudo es la función logística

$$G(x) = 1/(1 + e^{-x}) \quad (2.61)$$

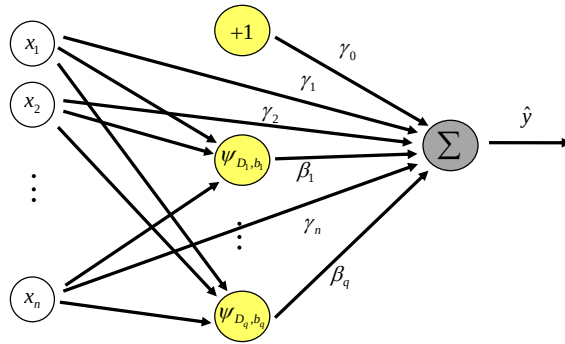
o la tangente hiperbólica

$$G(x) = (e^x - e^{-x})/(e^x + e^{-x}) \quad (2.62)$$

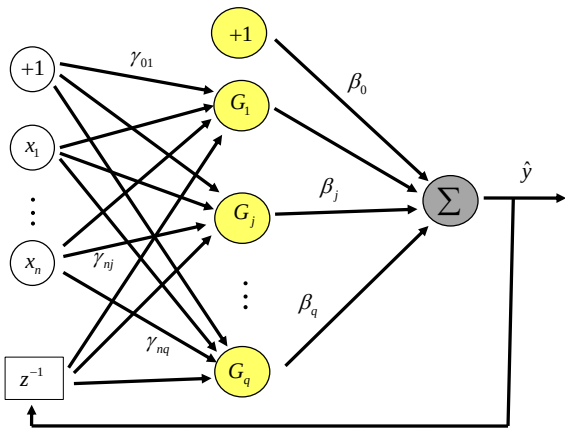
Una de las primeras publicaciones en donde se aplica la arquitectura de tipo feed-forward, en particular el Perceptrón multicapa, a la predicción de precios eléctricos



(a) Red feed-forward



(b) Red Wavelet



(c) Red recurrente

Figura 2.7: Diferentes arquitecturas de red

se encuentra en Szkuta et al. (1999). En este trabajo, la demanda y las reservas de energía, así como sus retardos y los del precio fueron las entradas seleccionadas para predecir los precios spot del mercado eléctrico de Victoria. Un perceptrón multicapa fue también usado en Gao et al. (2000) para predecir precios y demandas del mercado de California. En este caso, también se consideraron como entradas los costes de los combustibles, las exportaciones e importaciones, la temperatura, un conjunto de índices para las horas del día y los días de la semana. En Pino et al. (2008), se hace una selección de los ejemplos de entrenamiento a través de una red neuronal de tipo ART (*Adaptive Resonance Theory*), con la que posteriormente se alimenta una red feed-forward, para predecir los precios del mercado eléctrico español. En Mandal et al. (2007), se utiliza una red feed-forward para corregir la curva diaria del precio que se obtiene tras la aplicación de una técnica de predicción no paramétrica conocida como *similar day*.

Un caso particular de redes de tipo feed-forward, son las conocidas como redes *wavelet*, las cuales utilizan funciones *wavelet* como funciones de activación (ver figura 2.7b). Estas redes aprovechan las propiedades de adaptabilidad de la forma de dichas funciones para modelar señales no estacionarias y de alta frecuencia (Graps (1995)), como es el caso de las series de precios eléctricos. Su expresión general responde a (2.63)

$$y_t = \sum_{j=1}^q \beta_j \psi[D_j(W_t^T - b_j)] + W_t^T \gamma + \gamma_0 + \varepsilon_t \quad (2.63)$$

donde, el coeficiente escalar γ_0 y el vector de parámetros $\gamma = (\gamma_1, \dots, \gamma_{n+p})^T$ modelan las relaciones lineales existentes entre las entradas y la salida, y β_j son los pesos de salida de cada una de las neuronas (conocidas como *wavelons*) cuyas funciones de activación son funciones wavelet multidimensionales de parámetros la matriz diagonal D_j y el vector b_j . A menudo se toma el sombrero mexicano (*mexican hat*) como función wavelet, que responde a la siguiente expresión

$$\psi(x) = (\|x\|^2 - n)e^{-\|x\|^2/2} \quad (2.64)$$

donde x es un vector de dimensión n y $\|x\|^2$ es $x^T x$. Un ejemplo reciente de la aplicación de redes *wavelets* para los mercados Español, de Ontario y PJM se encuentra en Pindoriya et al. (2008).

Un tipo de redes más complejas que las feed-forward son las redes recurrentes Mandic and Chambers (2001) (ver figura 2.7c). En este tipo de redes, existe una retroalimentación que va de la salida a las entradas (red NOE, Nonlinear Output Error), o se presenta en alguna de las capas ocultas (red de *Elman*). Como ejemplo, en la figura 2.7c se muestra una representación gráfica de la arquitectura NOE, donde se puede apreciar que existe un elemento que retarda la señal de salida para tomarla como entrada en el instante de tiempo siguiente. Estas redes han sido

aplicadas en Andalib and Atry (2009) y Hong and Hsiao (2002), donde se compara su rendimiento con el de otros métodos de predicción como son los splines de regresión adaptativos, las redes wavelets o las tradicionales feed-forward basadas en el algoritmo de aprendizaje *backpropagation*. En ambos trabajos se pone de manifiesto las buenas capacidades de generalización de estas redes y su éxito para modelar las rápidas variaciones del precio. La desventaja de esta arquitectura es la complejidad y el coste computacional que supone estimarla debido a las retroalimentaciones que presenta.

Otro tipo de redes que han sido también utilizadas en Rodriguez and Anders (2004) para los precios spot de Ontario, y en Amjady (2006) para los de España, son las redes neuronales borrosas, donde se combinan técnicas procedentes del campo de la lógica borrosa con las redes neuronales.

Mención especial merece el trabajo desarrollado en Mateo et al. (2005), dada su conexión con el desarrollo posterior de esta tesis. En él se propone el uso de un modelo IOHMM (Input/Output Hidden Markov Model) capaz de modelar los cambios discretos en las estrategias de los agentes del mercado, los cuales se ven reflejados en las variaciones de las variables fundamentales que afectan al precio. El modelo IOHMM es una generalización de un modelo oculto de Markov basado en redes neuronales, que consta de dos módulos bien diferenciados (ver figura 2.8). Uno de los módulos es el encargado de modelar la función de distribución de probabilidad del régimen del sistema a partir de las variables fundamentales que afectan al precio y el régimen del sistema en el estado anterior. Este módulo es, por tanto, el encargado de estimar la matriz de probabilidades de transición, las cuales son función de las variables exógenas del sistema. El otro módulo consta de un conjunto de *expertos* que modelan la función de densidad del precio condicionada a cada uno de los regímenes del sistema a partir también de variables fundamentales. Tanto la red que modela la distribución de probabilidad del régimen como los expertos son implementados con perceptrones multicapa.

En general, no está claro si el rendimiento de las redes neuronales supera a los métodos paramétricos de series temporales en lo que a predicción de precios spot se refiere. Existe poca literatura comparando los métodos de ambos campos para sacar conclusiones definitivas. En Conejo et al. (2005) se compara el rendimiento de un perceptrón multicapa con el de los modelos de función de transferencia o regresión dinámica, mostrando que el primero proporciona peores resultados que los segundos. No obstante, no se ha encontrado literatura en la que se comparen los métodos de series temporales con las arquitecturas de red más sofisticadas y que mejores resultados parecen proporcionar, como son las redes recurrentes, redes wavelets, o las redes neuronales borrosas.

2.2.4. Modelos basados en técnicas de minería de datos

Como ya se ha expuesto en la sección 2.2.1, los precios spot en mercados eléctricos, en general, presentan relaciones no lineales con sus variables fundamentales,

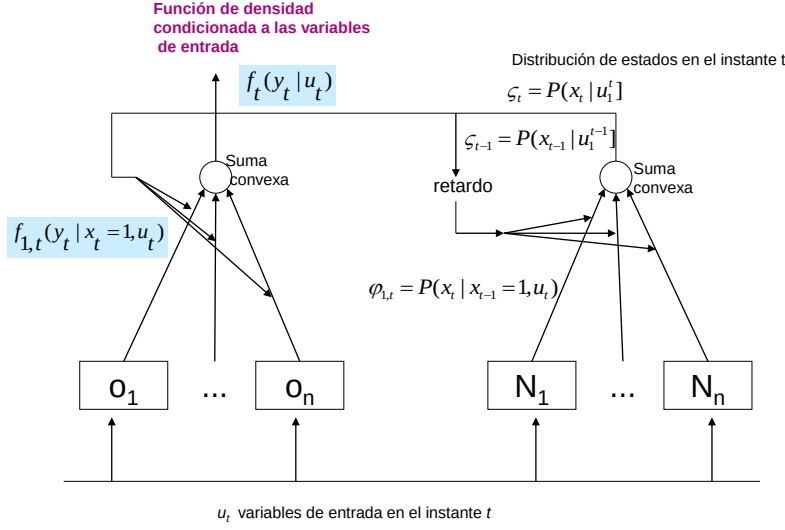


Figura 2.8: Arquitectura del modelo IOHMM. Fuente: Mateo (2005)

una volatilidad compleja y no constante y distribuciones asimétricas con valores extremos que se alejan de la habitual hipótesis de normalidad. Debido a estas especiales características y a la creciente necesidad que tienen los agentes de disponer de predicciones más precisas, algunos investigadores han propuesto métodos no paramétricos alternativos con el fin de superar los límites de los métodos paramétricos. Además de las redes neuronales que han sido expuestas en la sección anterior, existen muchas y muy diversas técnicas que se han empleado para tal propósito, habitualmente en conjunción con las redes neuronales u otros modelos de series temporales, y que no serán cubiertas en más detalle: lógica borrosa (Niimura and Nakashima (2001)), vecinos más cercanos (Lora et al. (2007)), métodos *similar-day* (Mandal et al. (2007)), splines de regresión multivariables adaptativos (Zareipour et al. (2006a)), máquinas de vectores soporte (Gao et al. (2007)), descomposición de la señal con funciones *wavelets* (Conejo et al. (2005)) y reducción de la dimensión *LLE* (*Local Linear Embedding*) (Chen et al. (2008)).

2.3. Conclusiones

Dada la naturaleza no almacenable de la electricidad y la forma escalonada de las curvas de venta debido a la presencia de diferentes tecnologías de generación, el

comportamiento de los precios resultantes de la casación de ofertas en el mercado diario resulta ser errático y complejo de predecir. Dicho comportamiento queda reflejado en un conjunto de características únicas y universales, que se resumen en: reversión a la media, presencia de componentes estacionales, diferentes patrones de comportamiento intradiarios e intrasemanales, efectos de calendario, volatilidad variable y distribuciones asimétricas con colas pesadas y *spikes*. Debido a este especial comportamiento que presentan las series de precios, la investigación en métodos de predicción paramétricos ha ido evolucionando desde los modelos univariados que tan solo utilizan la información contenida en la propia serie del precio hasta los modelos que incorporan también las variables fundamentales por las que se ve afectado; desde los modelos lineales de un solo régimen hasta los modelos que asumen diferentes modelos de comportamiento en función del estado del sistema y/o variables fundamentales; desde los métodos que tratan de predecir el valor puntual del precio hasta aquellos que tratan de caracterizar el proceso estocástico completo prediciendo su función de densidad. Así mismo, se han propuesto también gran cantidad de métodos no paramétricos con el fin de capturar las complejas relaciones no lineales que presentan los precios eléctricos, como por ejemplo las redes neuronales.

Sin embargo, debido a la complejidad de las series de precios y a la creciente necesidad de predicciones precisas, los resultados de los métodos de predicción propuestos hasta la fecha no son plenamente satisfactorios. En nuestra opinión, creemos que existen actualmente dos carencias importantes que esta tesis trata de cubrir. La primera de ellas es el planteamiento de un modelo de serie temporal que represente adecuadamente la distribución de probabilidad de los precios eléctricos, sin que esto suponga una merma en las capacidades de predicción del modelo. Como se ha comentado en la revisión de este capítulo, las distribuciones paramétricas propuestas en la literatura no se ajustan plenamente a las distribuciones empíricas obtenidas a través de los métodos descritos. La causa de esta limitación se encuentra en la forma rígida que este tipo de distribuciones impone, las cuales no permiten adaptarse de forma flexible a las series de precios de cada mercado específico. En cuanto a la solución no paramétrica basada en métodos kernel propuesta en Weron and Misiorek (2008), ésta tiene la desventaja de no proporcionar una forma analítica de la función de densidad.

La segunda limitación de los métodos de predicción en la actualidad está relacionada con la representación del comportamiento no lineal. A pesar de que los métodos lineales, como los de regresión dinámica o función de transferencia, proporcionan buenos resultados de cara a la predicción, éstos no son capaces de describir adecuadamente las relaciones no lineales presentes en la serie del precio. En cuanto a las redes neuronales, éstas en teoría son capaces de aproximar cualquier función por complejas que sean las relaciones entre las entradas y la salida. Sin embargo, en la práctica presentan un alto riesgo de sobreentrenamiento y su elevado número de parámetros puede afectar sensiblemente a sus capacidades de generalización. Los modelos de cambio de régimen constituyen una buena oportunidad para incorporar dinámicas no lineales manteniendo la estructura básica de los modelos lineales. Sin embargo, habitualmente se han utilizado modelos de

Markov, en donde se asume que la distribución de probabilidad del régimen está determinada por la distribución de probabilidad del régimen en el instante anterior. Estos modelos presentan problemas a la hora de predecir correctamente el régimen futuro, puesto que éste no puede ser determinado únicamente por el régimen en instantes anteriores. En nuestra opinión, el régimen del sistema puede ser predicho a través de las variables fundamentales que afectan al precio, las cuales influyen en los agentes a la hora de enviar sus ofertas al mercado. El modelo descrito en Mateo et al. (2005) es capaz de predecir el régimen del sistema modelando las complejas relaciones existentes entre el precio y sus variables fundamentales. Sin embargo, los modelos utilizados como expertos, redes neuronales feed-forward, adolecen de algunas de las características deseables para modelar las series de precios que sí presentan los métodos paramétricos, como son una estructura más simplificada, la retroalimentación del error y el modelado de la volatilidad como función dinámica del pasado de la serie.

Capítulo 3

Representación no paramétrica de la distribución de probabilidad de los precios de la electricidad

En este capítulo se propone un modelo semiparamétrico de serie temporal para predecir la distribución de probabilidad de los precios spot en un horizonte de corto plazo. La estructura general responde a la de un modelo de función de transferencia con ruido ARIMA-ARCH estacional. Dicha estructura permite considerar la influencia de variables fundamentales sobre el precio, así como su carácter autorregresivo, estacional y heterocedástico. La característica novedosa del modelo consiste en la especificación de la densidad de las innovaciones como una función semiparamétrica, lo que permite representar de manera flexible la asimetría y las colas pesadas que habitualmente presentan las distribuciones de los precios. Dicha función de densidad consiste en la transformación cuadrada de una serie de Hermite, la cual es capaz de aproximar densidades de formas generales con el suficiente número de parámetros. El modelo se ha aplicado a las series de precios de diferentes mercados eléctricos, comparando los resultados obtenidos con los modelos que asumen distribuciones gaussianas.

3.1. Introducción

De entre los métodos paramétricos descritos en el capítulo anterior, los de función de transferencia son uno de los que más han sido utilizados y que mejor rendimiento han mostrado de cara a la predicción a corto plazo de los precios eléctricos (Nogales and Conejo, 2006; Conejo et al., 2005; Nogales et al., 2002; Zareipour et al., 2006b; Misiolek et al., 2006). Estos modelos recogen el comportamiento autorregresivo y estacional de los precios, y además incorporan fácilmente la influencia de variables explicativas a la predicción. A diferencia de los modelos que tratan los precios como observaciones diarias de 24 componentes (como los modelos tipo VAR/VARX), el enfoque horario usado en estos modelos facilita el proceso de identificación y estimación de los mismos, lo que redundará a la postre en una mejora de las capacidades de generalización y predicción. Sin embargo, estas buenas capacidades de predicción se limitan a la predicción puntual o valor esperado del precio. En el habitual marco de trabajo que asume homocedasticidad y distribuciones gaussianas, estos modelos no representan correctamente el comportamiento estocástico de los precios eléctricos, y por tanto, no son capaces de proveer buenas estimaciones de los intervalos de predicción o, en general, de su función de densidad condicional.

La asunción de varianza condicional constante no es válida para describir el comportamiento de los precios eléctricos. Una vez eliminada la dependencia en media presente en las series de precios, es habitual que la serie del cuadrado de los errores de predicción presente una clara estructura de autocorrelación. Algunos autores han descrito este comportamiento heterocedástico usando modelos ARCH, donde parte de la varianza es predecible y función lineal de la varianza en instantes anteriores (García et al., 2005; Koopman et al., 2007; Carnero et al., 2003; Knittel and Roberts, 2001; Misiolek et al., 2006). Dado que un proceso ARCH puede escribirse como un proceso autorregresivo de los errores cuadrados del modelo, también se pueden recoger con él efectos estacionales en la varianza de la serie. Estos modelos han sido ampliamente utilizados con éxito en todo tipo de series de índole financiera. Con respecto a su aplicación a los precios eléctricos, constituyen una solución apropiada para modelar su volatilidad debido a la ya mencionada característica de *volatility clustering* presente en las series. A pesar de que existe algo de controversia sobre si su uso contribuye a predecir de manera más precisa el valor puntual o realizado de los precios eléctricos, los resultados de las investigaciones con estos modelos sugieren que incorporar un término ARCH en la especificación del modelo de función de transferencia contribuye a predecir más fielmente su densidad condicional.

Sin embargo, a pesar de que los modelos de tipo ARCH/GARCH gaussianos conducen a distribuciones marginales que presentan colas más pesadas que la distribución normal (ver Peña (2005) para más detalles), esto habitualmente no es suficiente para describir correctamente el comportamiento de los precios eléctricos, puesto que suelen presentar gran cantidad de valores inusuales (para la distribución gaussiana) y colas aún más pesadas (ver por ejemplo, (Koopman et al. (2007))). Tampoco su comportamiento asimétrico, causado por la diferente variabilidad entre

precios altos y bajos, puede ser modelado mediante una distribución simétrica. En general, las distribuciones de precios eléctricos presentan asimetría y colas pesadas y, dada la dependencia en la manera de ofertar de los agentes al mix de generación, formas diferentes según el mercado específico. Por lo tanto, en este contexto, la distribución normal resulta inadecuada e insuficiente. Algunos investigadores han tratado de modelar los precios de diferentes mercados eléctricos asumiendo otro tipo de distribuciones paramétricas con asimetría y colas pesadas ya mencionadas en el capítulo anterior, como por ejemplo la *t-student* asimétrica (Panagiotelis and Smith (2008)), la hiperbólica, la inversa gaussiana o la α -estable (Weron (2009)). Sin embargo, la forma rígida que este tipo de distribuciones impone no permite adaptarse de forma flexible a las diferentes series de precios. Además, las distribuciones paramétricas adolecen de la posibilidad de incrementar su capacidad de aproximación a la distribución empírica puesto que el número de parámetros que las definen está predeterminado.

Para caracterizar adecuadamente el proceso estocástico de los precios eléctricos resulta necesario, por tanto, una distribución capaz de modelar de manera flexible no únicamente su media y varianza condicional, sino también momentos de mayor orden. Para tal propósito, utilizar un enfoque no paramétrico es a nuestro juicio una solución apropiada, puesto que la característica principal de dicho enfoque es precisamente permitir que sea el conjunto de datos el que determine la forma que éstos presentan, sin asumir ninguna previamente. Esta idea ya ha sido aplicada a los precios eléctricos en Weron and Misiolek (2008), que está basado en el trabajo previo de Cao et al. (2003). Los modelos de series temporales presentados en estos artículos describen la densidad de las innovaciones o el comportamiento puramente estocástico de la variable de salida a través de estimadores *kernel* (Silverman (1986)). El primer artículo citado encuentra que los modelos así estimados arrojan, en general, mejores resultados que sus contrapartidas gaussianas, tanto en predicción puntual como de intervalos, para los precios eléctricos en CalPX (California) y Nord Pool. El segundo pone de manifiesto, a través de un estudio de simulación, que dicha técnica de aproximación de densidades debería ser elegida frente al habitual marco de trabajo gaussiano en situaciones en las que no existe información previa sobre la distribución de los errores de predicción del modelo. Sin embargo, el método presenta los inconvenientes de ser sensible al parámetro de ancho de banda, lo que hace necesario un método adecuado para su correcta selección, y de no presentar una forma analítica de su función de densidad, estando definido el valor de densidad asignado a cada punto en función de los puntos más cercanos al mismo.

Una manera de calcular intervalos de predicción que es aplicable cuando la hipótesis de innovaciones gaussianas no se satisface es a través del conocido método *bootstrap* (Efron and Tibshirani (1993)). Este método de remuestreo, en combinación con la estimación de *quasi máxima verosimilitud*¹ (*quasi maximum likelihood*), proporciona una manera sencilla de estimar intervalos de predicción futuros que ha

¹Se conoce como estimación de *quasi máxima verosimilitud* a aquella que maximiza una función de verosimilitud mal especificada (por ejemplo, una estimación de máxima verosimilitud bajo la hipótesis de normalidad cuando los datos reales no se distribuyen de este modo).

sido ampliamente utilizada (ver p. ej. Hyndman et al., 2008; Rodriguez and Ruiz, 2008). La idea básica consiste en tomar la distribución empírica de los errores de predicción o innovaciones del modelo obtenidos tras la estimación de sus parámetros asumiendo normalidad, para a partir de ella generar réplicas bootstrap² que puedan ser utilizadas junto con el modelo estimado para simular posibles realizaciones futuras del proceso. De este modo se obtienen distribuciones predictivas o intervalos de predicción a cualquier horizonte de tiempo basadas en los residuos obtenidos tras el ajuste del modelo en lugar de en la distribución gaussiana. El método presenta la ventaja de no asumir una forma determinada de la distribución de probabilidad del proceso estocástico para generar los intervalos de predicción, aunque como desventaja puede señalarse la “incoherencia” que supone estimar la media y varianza del proceso utilizando una distribución que es incorrecta.

Nuestra intención en este capítulo es extender la especificación del modelo de función de transferencia para que pueda proporcionar estimaciones razonables de la función de densidad de los precios eléctricos, sin que esto suponga una merma en la precisión de sus estimaciones puntuales. Para ello, el modelo adoptado incorpora un término ARCH que describe la estructura de autocorrelación que típicamente presentan los cuadrados de los errores de predicción. Además, la densidad de las innovaciones se modela como una función no paramétrica. Esto implica que la forma de dicha densidad no está fijada a priori, lo que permite que sean los datos usados en el ajuste los que determinen su forma definitiva. La característica novedosa del modelo reside en la solución adoptada para lograrlo, puesto que está basada en una transformación cuadrada de una serie de Hermite, cuyo término principal es la densidad normal estándar (Gallant and Nychka (1987)). La densidad resultante presenta, por tanto, la forma de la distribución normal modificada por la multiplicación de un polinomio al cuadrado, lo que da lugar a formas generales de densidades, entre ellas distribuciones con asimetría y colas pesadas. Esta familia de densidades es capaz de aumentar la capacidad de aproximación a la distribución empírica incrementando el orden de dicho polinomio. En comparación con el método *bootstrap*, la solución aquí propuesta guarda la coherencia entre el valor estimado de los parámetros que modelan la media y varianza del proceso y la densidad asumida para estimar los intervalos de predicción.

Lo que resta de capítulo está organizado de la siguiente manera: la sección 3.2 desarrolla el planteamiento teórico de la solución adoptada para modelar las innovaciones del modelo de precios como una función semiparamétrica (basado en Gallant and Nychka, 1987; Mauleón, 2010); en la sección 3.3 se presenta el modelo completo propuesto para la predicción de los precios eléctricos, y se detallan los procedimientos usados para su estimación e identificación; en la sección 3.4 se muestran los resultados de la aplicación del modelo a los precios horarios de distintos mercados eléctricos internacionales, mostrándose la mejora obtenida en el modelado de la función de densidad con respecto a los que asumen distribuciones gaussianas;

²Una réplica bootstrap es una permutación con reemplazamiento de los elementos del conjunto original.

finalmente, en la sección 3.5 se extraen las conclusiones de lo expuesto en el capítulo.

3.2. Aproximación de funciones de densidad mediante series de Hermite

La familia de densidades conocida como Edgeworth-Sargan (ES, ver Mauleón (2010)) responde a la siguiente formulación

$$f(\epsilon) = \rho(\epsilon) \times \left[1 + \sum_{k=1}^K d_k H_k(\epsilon) \right] \quad (3.1)$$

donde $\rho(\epsilon)$ denota la función de densidad normal estándar $N(0, 1)$, $d_1, \dots, d_K$ son un conjunto de constantes y $H_k(\epsilon)$ es el polinomio de Chebyshev-Hermite de orden k definido por la ecuación $D^r \rho(\epsilon) = (-1)^r \rho(\epsilon) H_r(\epsilon)$ (ver apéndice 3.6).

En virtud de las propiedades de estos polinomios, se comprueba que la integral de (3.1) es 1, $E(\epsilon) = d_1$ y $E(\epsilon^2) = (1 + 2d_2)$. Es decir, la densidad descrita en (3.1) está en forma estándar si $d_1 = d_2 = 0$. Además, $d_3 = E(\epsilon^3)/6$ y $d_4 = (E(\epsilon^4) - 3)/24$, por lo que recogen las medidas de asimetría y curtosis, respectivamente. Por último, la función de distribución viene dada por (3.2)

$$F(a) = \int_{-\infty}^a f(\epsilon) d\epsilon = \varrho(a) - \rho(a) \times \left[\sum_{k=1}^K d_k H_{k-1}(a) \right] \quad (3.2)$$

donde $\varrho(a)$ denota la función de distribución de la $N(0, 1)$ (ver apéndice 3.7 Kendall and Stuart, 1969; Mauleón, 2010 para estas y otras propiedades).

La familia de densidades ES describen, por tanto, formas generales de funciones de densidad según el valor que tome K y los coeficientes d_k . Además, permiten incrementar la capacidad de aproximación de una distribución dada aumentando el grado K del polinomio. Entre las ventajas de esta familia de distribuciones, aparte de su enfoque no paramétrico, está la manera de contabilizar la asimetría (a través de los polinomios impares) y el sencillo calculo que supone obtener su distribución de probabilidad. Sin embargo y a pesar de estas deseables propiedades, presentan el inconveniente de que no para todos los posibles valores de $d_1, \dots, d_K$ se cumple que $f(\epsilon) \geq 0$ para todo ϵ , lo que supone una fuerte limitación para encontrar los valores de los parámetros $d_1, \dots, d_k$ en el contexto de estimación de máxima verosimilitud.

En Gallant and Nychka (1987) y más tarde en Gallant and Tauchen (1989) se propone otra familia de densidades que responde a la formulación en (3.3)

$$f(\epsilon) = \rho(\epsilon) \times \frac{\left(\sum_{k=0}^K a_k \cdot \epsilon^k\right)^2}{\int_{-\infty}^{\infty} \rho(\epsilon) \left(\sum_{k=0}^K a_k \cdot \epsilon^k\right)^2 d(\epsilon)} \quad (3.3)$$

donde los valores a_k , $k = 0, \dots, K$ en este caso no están restringidos. Nótese que el denominador de la expresión (3.3) asegura que la integral de la función sea 1, como requiere una función de densidad. Además, dado que el numerador es un término elevado al cuadrado, el valor de $f(\epsilon)$ es siempre mayor o igual que 0 para cualesquiera valores de los a'_k s. Para conseguir una representación única, el término a_0 es fijado a 1 (ver Gallant and Tauchen (1990)). Por tanto, si $K = 0$ se tiene exactamente la densidad normal estándar. Si K es positivo, se tiene la forma de la densidad normal estándar modificada debido a la multiplicación de un polinomio al cuadrado.

Del mismo modo que las funciones ES, la familia de funciones descrita en (3.3) es capaz de aproximar formas generales de funciones de densidad entre las que se incluyen aquellas que presentan colas pesadas o asimetría. Es por ello que dichas funciones son denominadas seminoparamétricas (*Seminonparametric*, SNP), como sugieren sus autores. Dado que en este caso para cualquier valor de los a'_k s la densidad cumple que $f(\epsilon) \geq 0$ para todo ϵ , encontrar los parámetros a_k que maximicen la verosimilitud de una muestra de datos no presenta los problemas de infactibilidad de las densidades ES. Sin embargo, las densidades SNP tienen el inconveniente de no presentar una solución analítica de su función de distribución, lo que puede complicar el cálculo de los intervalos de predicción a un determinado nivel de confianza. En Gallant and Tauchen (1990) se propone un método especial que permite simular una variable aleatoria que siga la distribución definida en (3.3).

No obstante, las densidades SNP, en general, no cumplen que $E(\epsilon) = 0$ y $\sigma^2(\epsilon) = E(\epsilon^2) - E(\epsilon)^2 = 1$, lo que puede conducir a estimaciones sesgadas de la media y varianza en un modelo estocástico como el descrito en la sección (3.3). En Mauleón (2010) se establecen de forma explícita las restricciones que son necesarias imponer sobre los parámetros a_k para que las densidades SNP tengan media 0 y varianza 1. Sin embargo, su cálculo resulta computacionalmente muy costoso³. Una forma más conveniente de solventar este problema es calcular el valor de las restricciones expresando la densidad SNP en su forma equivalente ES, como se muestra en (Mauleón (2010)). Así, dado que los coeficientes de los polinomios de Hermite son constantes, el polinomio del numerador en (3.3) puede ser escrito como una suma ponderada de polinomios de Hermite, permitiendo reescribir la función de densidad de la siguiente manera

³En el citado trabajo se determina que la complejidad del cálculo de dichas restricciones es $O((2K)^2)$.

$$f(\epsilon) = v\rho(\epsilon) \times \left[1 + \sum_{k=1}^K d_k H_k(\epsilon) \right]^2 \quad (3.4)$$

donde la integral del denominador en (3.3) toma el valor $1/v$ y los coeficientes d_k definen una función de densidad para cualesquiera valores que tomen. En Mauleón (2010) se establecen también de manera explícita las restricciones que han de cumplir los coeficientes d_k en (3.4) para que la función de densidad tenga media 0 y varianza 1, cuyo cálculo resulta menos costoso que para el caso de la densidad en (3.3). Partiendo de la formulación en (3.4), la función de densidad SNP puede reescribirse de la misma forma que en (3.1) del siguiente modo

$$f(\epsilon) = \rho(\epsilon) \times \left[1 + \sum_{k=1}^{2K} \Psi_k H_k(\epsilon) \right] \quad (3.5)$$

donde los $\Psi'_k, k = 1, \dots, 2K$ son función de los parámetros no restringidos $d_k, k = 1, \dots, K$ obtenidos tras desarrollar explícitamente el término al cuadrado en (3.4) y reorganizar los resultados (ver apéndice 3.7 para más detalles). Por tanto, hay K restricciones no lineales asociadas a la última reformulación de $f(\epsilon)$, para asegurar que $f(\epsilon) \geq 0$ para todo ϵ . La densidad en (3.5) tiene media cero y varianza 1 cuando $\Psi_1 = 0$ y $\Psi_2 = 0$.

Por lo tanto y dada la equivalencia entre las distribuciones ES y SNP, es posible desarrollar un procedimiento que estime de manera flexible (no paramétrica) la densidad del término de innovación o componente puramente estocástica de una determinada variable en un modelo de predicción, asegurando las necesarias propiedades que ha de presentar, a saber: no negatividad, media cero, varianza uno e integral igual a uno. Este procedimiento es explicado en más detalle en la sección (3.3.1).

A continuación se ilustra el error de aproximación en el que puede incurrirse tras estimar los parámetros de la media y varianza de una muestra de datos asumiendo que su distribución es gaussiana cuando en realidad no lo es. Para ello se ha generado una muestra aleatoria $\{y\}$ de 1000 puntos que se distribuyen según una *t-student* con 4 grados de libertad, y se han estimado los parámetros de media y varianza de dicha población por máxima verosimilitud, asumiendo que la variable $\epsilon = \sigma^{-1}(y - \mu)$ se distribuye $N(0, 1)$ y con densidad de tipo SNP con $K=4$. Los resultados de las estimaciones de la media y varianza en ambos casos no difieren significativamente ($\hat{\mu}_{GAUSS} = 0.033, \hat{\sigma}_{GAUSS} = 1.916, \hat{\mu}_{SNP} = 0.040, \hat{\sigma}_{SNP} = 2.070$) y ambos están muy próximos a los valores reales ($\mu = 0, \sigma = 2$). Sin embargo, el problema aparece en el cálculo de los percentiles de la distribución si asumimos que la distribución asumida es la correcta para el caso gaussiano. En los gráficos de la figura 3.1 se muestran los resultados de este experimento. En ella se aprecia claramente cómo la distribución normal no estima adecuadamente la probabilidad real de los valores alejados de la media, mientras que por el contrario, la densidad SNP sí proporciona

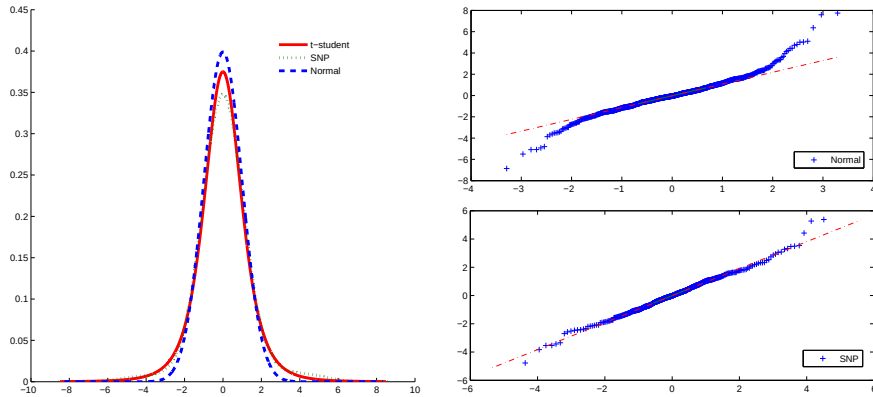
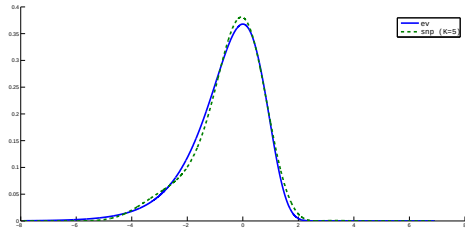


Figura 3.1: Resultado de la estimación de la función de densidad de una muestra aleatoria de 1000 puntos distribuidos según una t-student con 4 grados de libertad. El gráfico de la izquierda muestra la distribución real y las distribuciones estimadas mediante el ajuste gaussiano y SNP. La segunda columna muestra los gráficos q-q plot que enfrentan las distribuciones asumidas y empíricas.

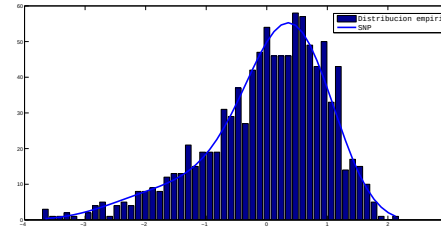
un ajuste próximo a la verdadera distribución con un valor de K pequeño.

Finalmente, para ilustrar las capacidades de aproximación de la familia de densidades SNP, en la figuras 3.2 y 3.3 se muestran los resultados de un experimento similar al anterior pero con muestras de 1000 puntos generados según distintas distribuciones, a saber: la distribución de valores extremos (EV, con parámetros de ubicación 0 y escala 1); Rayleigh (con parámetro de escala 0.5); β (con parámetros $a=2$ y $b=2$); una mixtura de gaussianas (con medias -1 y 2, varianza 0.5 y 1 y proporción 0.6 y 0.4, respectivamente); F (con 30 grados de libertad para el numerador y 10 para el denominador) y α -estable (de parámetros $\alpha=1.5$, $\beta=0$, $\gamma=1$ y $\delta=0$). Los resultados de este último experimento dejan patente que la familia de densidades SNP es capaz de ajustar un gran número de diferentes distribuciones, entre ellas distribuciones asimétricas y con colas pesadas. Dado que el término central de las densidades SNP es la densidad normal, es lógico que cuanto más alejada de ésta sea la forma de la distribución a estimar, más difícil resulte la aproximación y, en general, se requiera un valor de K mayor. Nótese que el grado del polinomio requerido está relacionado con el peso o relevancia de las colas en la distribución a estimar, puesto que los puntos cercanos a la media son modelados fundamentalmente a través de la distribución gaussiana. A este respecto, cabe mencionar la alta precisión necesaria para estimar las densidades cuando el grado del polinomio en (3.3) es alto. Por ejemplo, un valor de $K=11$ implica un orden máximo de los polinomios de Hermite en su forma ES equivalente de 22, cuyos coeficientes llegan a tener 13 dígitos (ver Kendall and Stuart (1969)). El grado de aproximación obtenido para algunas distribuciones cuya forma está muy alejada de la normal, como la distribución F o la α -estable, muestran la versatilidad y capacidad de generalización de esta familia de

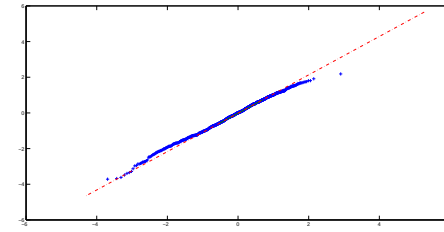
densidades. En el caso de la densidad α -estable, la aproximación en las colas no es buena debido a que la longitud de las mismas requeriría un número de datos mucho mayor para que fueran representadas adecuadamente. Es decir, existe una desviación de la distribución generada con respecto a la teórica. Esto es lo que ocurre también en cierta medida con la mixtura de gaussianas.



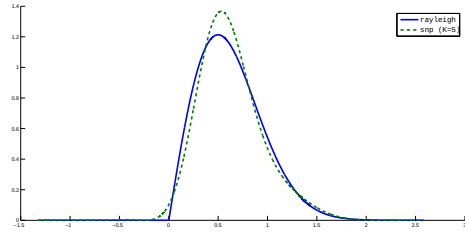
(a) EV: Densidad y estimación SNP (K=5)



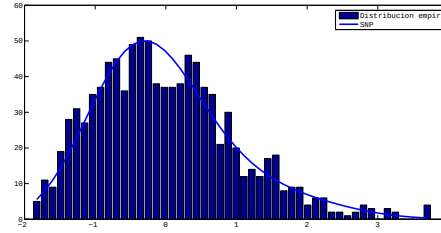
(b) EV: Histograma de residuos estandarizados y estimación SNP



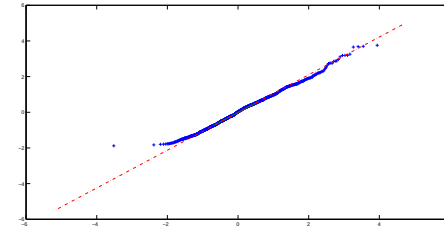
(c) EV: q-q plot SNP frente a residuos estandarizados



(d) Rayleigh: Densidad y estimación SNP (K=5)



(e) Rayleigh: Histograma de residuos estandarizados y estimación SNP



(f) Rayleigh: q-q plot SNP frente a residuos estandarizados

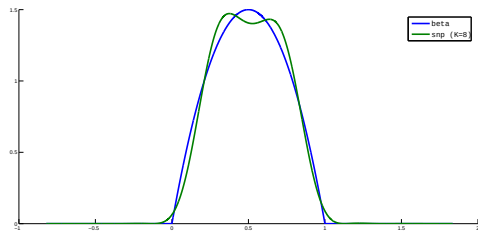
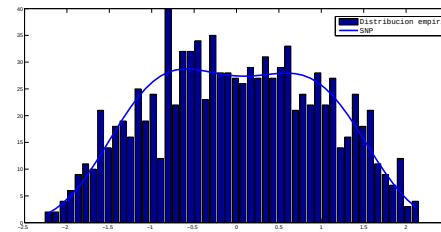
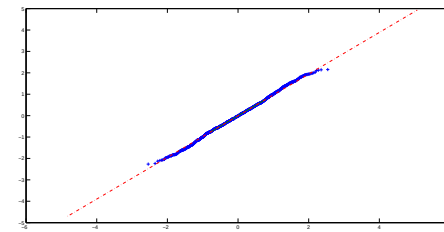
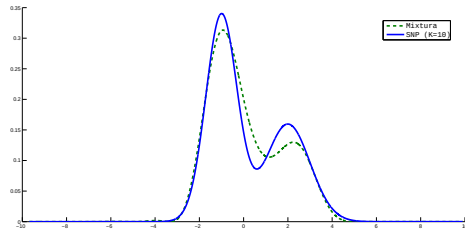
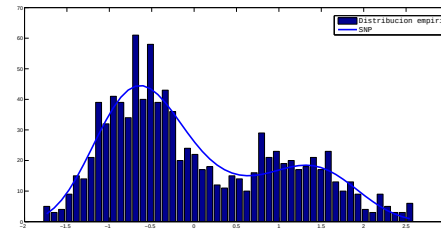
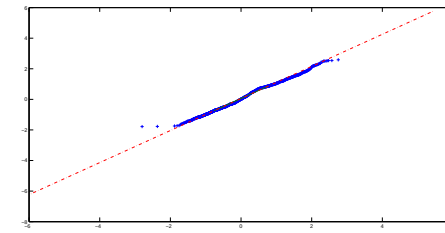
(g) β : Densidad y estimación SNP (K=8)(h) β : Histograma de residuos estandarizados y estimación SNP(i) β : q-q plot SNP frente a residuos estandarizados

Figura 3.2: Resultado de la estimación de las distribuciones: EV(0,1), Rayleigh(0.5) y Beta(2,2) a partir de muestras aleatorias de 1000 puntos.

(a) Mixtura: Densidad y estimación SNP ($K=10$)

(b) Mixtura: Histograma de residuos estandarizados y estimación SNP



(c) Mixtura: q-q plot SNP frente a residuos estandarizados

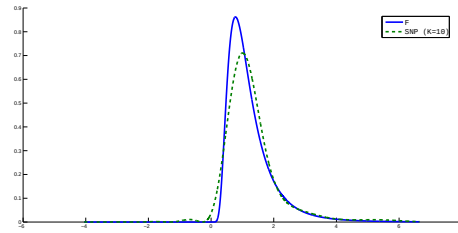
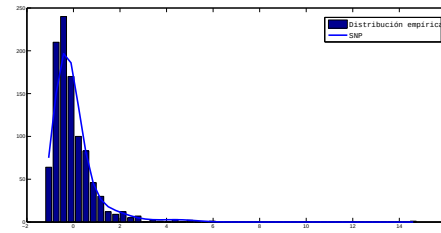
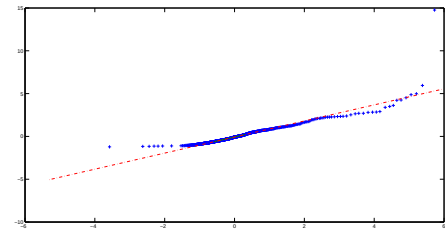
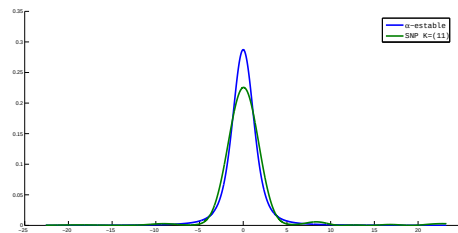
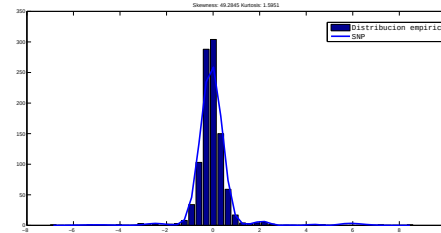
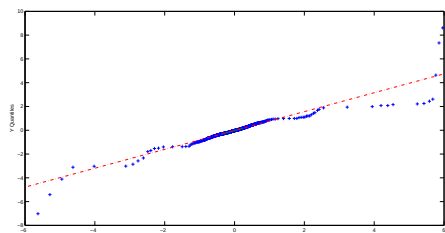
(d) F : Densidad y estimación SNP ($K=10$)(e) F : Histograma de residuos estandarizados y estimación SNP(f) F : q-q plot SNP frente a residuos estandarizados(g) α -estable: Densidad y estimación SNP ($K=11$)(h) α -estable: Histograma de residuos estandarizados y estimación SNP(i) α -estable: q-q plot SNP frente a residuos estandarizados

Figura 3.3: Resultado de la estimación de las distribuciones: Mixtura de Gaussianas $(-1,2),(0.5,1)$ y $\text{prob}=0.6$, $F(30,10)$ y α -estable $(1.5,0,1,0)$ a partir de muestras aleatorias de 1000 puntos.

3.3. Modelo de función de transferencia con ruido ARIMA-ARCH estacional

La expresión general del modelo de función de transferencia con ruido ARIMA-ARCH estacional responde a la siguiente a la siguiente formulación (3.6)-(3.15),

$$y_t = \sum_i^n G_i(L)x_{it} + \frac{\theta_q(L)\Theta_{Q_1}(L^{S_1})\dots\Theta_{Q_m}(L^{S_m})}{\nabla^d \nabla_{S_1}^{D_1} \dots \nabla_{S_m}^{D_m} \phi_p(L)\Phi_{P_1}(L^{S_1})\dots\Phi_{P_m}(L^{S_m})} \varepsilon_t \quad (3.6)$$

con

$$\varepsilon_t = \sigma_t \epsilon_t \quad (3.7)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_w \varepsilon_{t-w}^2 \quad (3.8)$$

y

$$G_i(L) = \frac{(w_{i0} + w_{i1}L + w_{i2}L^2 + \dots + w_{ig_i}L^{g_i})L^{b_i}}{1 + \delta_{i1}L + \delta_{i2}L^2 + \dots + \delta_{ir_i}L^{r_i}} \quad (3.9)$$

$$\phi_p(L) = 1 + \phi_1 L + \dots + \phi_p L^p \quad (3.10)$$

$$\Phi_{P_s}(L^{S_s}) = 1 + \Phi_1 L^{S_s} + \dots + \Phi_P L^{P_s * S_s} \quad (3.11)$$

$$\theta_q(L) = 1 + \theta_1 L + \dots + \theta_q L^q \quad (3.12)$$

$$\Theta_{Q_s}(L^{S_s}) = 1 + \Theta_1 L^{S_s} + \dots + \Theta_Q L^{Q_s * S_s} \quad (3.13)$$

$$\nabla^d = (1 - L)^d \quad (3.14)$$

$$\nabla_{S_s}^{D_s} = (1 - L^{S_s})^{D_s} \quad (3.15)$$

donde y_t es la variable de salida, x_{it} son las variables explicativas y ϵ_t es un proceso de ruido blanco de varianza 1. ε_t denota el error de predicción a un paso. $\phi_p(L)$ y $\theta_q(L)$ son polinomios en L de órdenes p y q , que describen los términos autorregresivos y de media móvil, respectivamente. $\Phi_{P_s}(L^{S_s})$ y $\Theta_{Q_s}(L^{S_s})$ describen los términos autorregresivos y de media móvil estacionales de período S_s , con $s = 1, \dots, m$ y siendo m el número de estacionalidades consideradas (por ejemplo $S_1 = 24$, $S_2 = 168$ para series horarias con estacionalidad diaria y semanal). ∇^d y $\nabla_{S_s}^{D_s}$ son los operadores que describen describen las diferencias regular y estacionales aplicadas para inducir estacionariedad en procesos integrados. $G_i(L)$ representa una familia de funciones de transferencia lineales capaces de capturar un amplio número de diferentes respuestas impulsionales con pocos parámetros. σ_t^2 modela la varianza de la serie como función lineal de los w previos errores de predicción al cuadrado.

Los parámetros han de cumplir una serie de restricciones para que el proceso sea estacionario, invertible y tenga varianza positiva (ver Chatfield, 2001; Pankratz, 1991 para detalles). Así,

$$\alpha_0 > 0 ; \alpha_j \geq 0 ; \sum_{j=1}^w \alpha_j < 1 \quad (3.16)$$

y todos los valores z que satisfacen las ecuaciones

$$1 + \phi_1 z + \dots + \phi_p z^p = 0 \quad (3.17)$$

$$1 + \theta_1 z + \dots + \theta_q z^q = 0 \quad (3.18)$$

y

$$1 + \Phi_1 z^{S_s} + \dots + \Phi_P z^{P_s * S_s} = 0 \quad \forall s = 1, \dots, m \quad (3.19)$$

$$1 + \Theta_1 z^{S_s} + \dots + \Theta_Q z^{Q_s * S_s} = 0 \quad \forall s = 1, \dots, m \quad (3.20)$$

$$1 + \delta_{i1} z + \delta_{i2} z^2 + \dots + \delta_{ir_i} z^{r_i} = 0 \quad \forall i = 1, \dots, n \quad (3.21)$$

deben quedar fuera del círculo unidad.

La formulación del modelo propuesto responde por tanto a la descrita en la sección 2.2.2.3 para los modelos de función de transferencia, los cuales han demostrado buenas capacidades para la predicción puntual de los precios eléctricos (ver p.ej. Nogales and Conejo, 2006; Misiolek et al., 2006), y añade un término ARCH que modela la varianza condicional de la serie de precios.

Es importante tener en cuenta que el modelo (3.6)-(3.15) puede escribirse también de la siguiente forma general (3.22)

$$y_t = u_t + \varepsilon_t = u_t + \sigma_t \epsilon_t \quad (3.22)$$

donde el término u_t responde a

$$u_t = \sum_i H^{-1}(L) G_i(L) x_i t + [1 - H^{-1}(L)] y_t \quad (3.23)$$

siendo

$$H(L) = \frac{\theta_q(L) \Theta_{Q_1}(L^{24}) \Theta_{Q_2}(L^{168})}{\phi_p(L) \Phi_{P_1}(L^{24}) \Phi_{P_2}(L^{168}) \nabla^d \nabla_{24}^{D_1} \nabla_{168}^{D_2}} = \sum_{l=0}^{\infty} \psi_l L^l \quad (3.24)$$

y donde los ψ'_l s se corresponden con los coeficientes de la representación del término $H(L)\varepsilon_t$ como un proceso de media móvil de orden infinito MA(∞). Nótese

que el término y_t en el instante t en (3.23) desaparece puesto que se anula, y solo intervienen sus retardos (y_{t-d} , con $d > 0$).

Dado que ϵ_t tiene media 0 y varianza 1, si denotamos como η_{t-k} a la información disponible hasta el instante $t-k$, es directo comprobar que la esperanza condicional de la variable de salida y_t es $E(u_t|\eta_{t-k})$ y que su varianza condicional es $E(\sigma_t^2|\eta_{t-k})$ sea cual sea la forma específica de la función de densidad de ϵ_t ⁴.

Para terminar la especificación del modelo propuesto, la forma de la función de densidad de ϵ_t responde a la propuesta en Gallant and Nychka (1987) descrita en la sección 3.2, es decir,

$$f(\epsilon_t) = \frac{[P_K(\epsilon_t)]^2 \rho(\epsilon_t)}{\int [P_K(s)]^2 \rho(s) ds} \quad (3.25)$$

con

$$P_K(\epsilon_t) = 1 + a_1 \epsilon_t + \dots + a_K \epsilon_t^K \quad (3.26)$$

siendo $\rho(\epsilon_t)$ la densidad normal estándar. Como se ha expuesto en la sección 3.2, esta forma no paramétrica de la función de densidad permite capturar de manera flexible la asimetría y las colas pesadas que habitualmente presentan las distribuciones de los precios eléctricos.

3.3.1. Estimación de los parámetros del modelo

La estimación de los parámetros de un modelo de serie temporal suele llevarse a cabo maximizando la verosimilitud de la muestra de datos disponible (Hamilton (1994)). Supongamos que se observa la serie temporal $Y = \{y_t\}_{t=1}^T$ y las variables explicativas $X_i = \{x_{it}\}_{t=1}^T, i = 1, \dots, n$. Si recogemos todos los parámetros que definen el modelo en el vector φ , maximizar la función de verosimilitud implica calcular el máximo de la función de densidad conjunta de las T observaciones de la muestra

$$f(y_1, y_2, \dots, y_{T-1}, y_T | X_1, \dots, X_n; \varphi) \quad (3.27)$$

con respecto al vector de parámetros φ . Es decir, la estimación de máxima verosimilitud es el valor de φ para el cual es más probable haber observado $\{y_t\}_{t=1}^T$. Teniendo en cuenta que la distribución conjunta de dos variables cualesquiera puede

⁴Si tomamos esperanzas en la ecuación (3.22)

$E(y_t|\eta_{t-k}) = E(u_t|\eta_{t-k})$ puesto que ϵ_t es incorrelado y $E(\epsilon_t) = 0$.

Elevando al cuadrado y tomando esperanzas se tiene que

$E((y_t - E(u_t|\eta_{t-k}))^2) = E(\sigma_t^2 \epsilon_t^2|\eta_{t-k}) = E(\sigma_t^2|\eta_{t-k})$, puesto que ϵ_t es incorrelado y $E(\epsilon_t^2) = 1$.

escribirse como el producto de la distribución marginal de la primera y la distribución de la segunda condicionada a los valores de la primera, podemos escribir la función de densidad conjunta en (3.27) como

$$f(y_1|X_1, \dots, X_n; \varphi) f(y_2, y_3, \dots, y_{T-1}, y_T|y_1, X_1, \dots, X_n; \varphi) \quad (3.28)$$

y repitiendo el mismo proceso para el segundo término de manera recursiva, se puede escribir la función de verosimilitud conjunta de la muestra $\{y_t\}_{t=1}^T$ de la siguiente manera

$$f(y_1, y_2, \dots, y_{T-1}, y_T|X_1, \dots, X_n; \varphi) = \prod_{t=1}^T f[y_t|\eta_{t-1}; \varphi] \quad (3.29)$$

donde η_{t-1} denota la información disponible hasta el instante $t-1$ ⁵. El problema habitualmente suele formularse en términos de la función de verosimilitud logarítmica puesto que presenta el mismo máximo y es computacionalmente más estable

$$\max_{\{\varphi\}} L(\varphi) = \sum_{t=1}^T \ln(f[y_t|\eta_{t-1}; \varphi]) \quad (3.30)$$

Por tanto, para estimar los parámetros del modelo es necesario asumir una determinada distribución de sus innovaciones, para poder entonces deducir la función de densidad de la variable de salida y_t , $f[y_t|\eta_{t-1}; \varphi]$. Sin más conocimiento de dicha distribución, es habitual asumir que $\epsilon_t \sim N(0, 1)$, lo que implica que la función de densidad de y_t es también gaussiana y responde a

$$f_{GAUSS}[y_t|\eta_{t-1}; \varphi] = n(y_t|u_{t|t-1}, \sigma_{t|t-1}) = \frac{1}{\sqrt{2\pi}\sigma_{t|t-1}} \exp\left[\frac{-(y_t - u_{t|t-1})^2}{2\sigma_{t|t-1}^2}\right] \quad (3.31)$$

donde $u_{t|t-1}$ y $\sigma_{t|t-1}^2$ son la media y varianza de y_t condicionados a la información disponible en $t-1$, respectivamente. Sin embargo, dado que las innovaciones del modelo siguen la distribución no paramétrica descrita en (3.25), la función de densidad de y_t es entonces

$$f_{SNP}[y_t|\eta_{t-1}; \varphi] = \frac{[P_K(z_t)]^2 n(y_t|u_t, \sigma_t)}{\int [P_K(s)]^2 \rho(s) ds} \quad (3.32)$$

donde los condicionantes se han omitido para simplificar la notación, $z_t = \sigma_t^{-1}(y_t - u_t)$ y $P_K(z_t)$ es un polinomio de orden K en z_t , de manera que si $K = 0$ se tiene (3.31).

⁵Nótese que las variables explicativas X_i se consideran conocidas en cualquier instante de tiempo.

Nótese que $f[y_t|\eta_{t-1}; \varphi]$ en (3.31) o (3.32) no presenta la misma expresión para cada instante $t = 1, \dots, T$ por un problema de valores iniciales. A continuación se ilustra esta situación calculando la función de verosimilitud de un simple modelo autorregresivo de orden 1 con varianza constante σ^2 (AR(1)) y ruido gaussiano, el cual responde a

$$y_t = c + \phi y_{t-1} + \varepsilon_t \quad (3.33)$$

En este caso, la función de densidad condicionada a la información precedente hasta $t - 1$ para las muestras correspondientes a los instantes de tiempo $t = 2, \dots, T$ es

$$f_{AR1}[y_t|y_{t-1}; \varphi] = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[\frac{-(y_t - c - \phi y_{t-1})^2}{2\sigma^2} \right] \quad (3.34)$$

sin embargo, para la primera muestra de la serie y_1 , la densidad marginal (no condicionada, puesto que no hay información previa al instante $t = 1$) es

$$f_{AR1}[y_1; \varphi] = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2/(1-\phi^2)}} \exp \left[\frac{-(y_1 - (c/(1-\phi)))^2}{2\sigma^2/(1-\phi^2)} \right] \quad (3.35)$$

puesto que la esperanza marginal del proceso (3.33) es $E(y_1) = \mu = c/(1-\phi)$ y su varianza $E((y_1 - \mu)^2) = \sigma^2/(1-\phi^2)$. Por tanto, para este caso la función de verosimilitud logarítmica definida en (3.30) responde a (3.36)

$$\begin{aligned} L_{AR1}(\varphi) = & \frac{-1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma/(1-\phi^2)) - \\ & - \frac{(y_1 - (c/(1-\phi)))^2}{2\sigma^2/(1-\phi^2)} - ((T-1)/2) \log(2\pi) - \\ & - ((T-1)/2) \log(\sigma^2) - \sum_{t=2}^T \left(\frac{(y_t - c - \phi y_{t-1})^2}{2\sigma^2} \right) \end{aligned} \quad (3.36)$$

la cual presenta una forma más complicada de maximizar que la que tendría si no se tuviera en cuenta la densidad marginal de la primera muestra de la serie. En general, el cálculo de la función de verosimilitud exacta para el modelo definido en (3.6)-(3.15) se convierte en un problema complejo y computacionalmente costoso a medida que el orden de los polinomios AR, MA, ARCH y los correspondiente a las funciones de transferencia aumenta debido a la diferente expresión del primer y segundo momento de la distribución para los valores iniciales de la muestra. La manera más sencilla de hacer frente a este problema es usar una representación en Espacio de Estados equivalente y evaluar la función de verosimilitud usando el

filtro de Kalman⁶. Sin embargo, una aproximación a la función de verosimilitud cuya expresión es mucho más sencilla surge si condicionamos a los valores de las primeras muestras de y_t y tomamos valores esperados en los términos que impliquen valores retardados de ε_t desconocidos. Dado que el orden $p^* = p + \sum_s^m P_s S_s$ del polinomio autorregresivo del modelo es habitualmente mayor que el de las funciones de transferencia de las variables explicativas, si denotamos por $q^* = q + \sum_s^m Q_s S_s$ y tomando $\varepsilon_0 \equiv (\varepsilon_{p^*-q^*+1}, \dots, \varepsilon_{p^*})'$ y $\varepsilon_0^2 \equiv (\varepsilon_{p^*-w+1}, \dots, \varepsilon_{p^*})'$ como dados, siendo

$$\varepsilon_{p^*-q^*+1} = \dots = \varepsilon_{p^*} = 0 \quad (3.37)$$

$$\varepsilon_{p^*-w+1}^2 = \dots = \varepsilon_{p^*}^2 = \alpha_0 / (1 - \sum_1^w \alpha_w) \quad (3.38)$$

entonces las secuencias de valores $\{u_t\}_{p^*+1}^T$, $\{\sigma_t\}_{p^*+1}^T$ y por tanto $\{z_t\}_{p^*+1}^T$ pueden ser fácilmente calculadas. La aproximación condicional a la función de verosimilitud definida en (3.30) para el caso de innovaciones gaussianas presenta entonces la forma siguiente

$$\begin{aligned} \hat{L}_{GAUSS}(\varphi) &= \log f(y_{p^*+1}, \dots, y_T | y_1, \dots, y_{p^*}, \varepsilon_0, \varepsilon_0^2; \varphi) = \\ &= \frac{T - p^*}{2} \log(2\pi) - \frac{1}{2} \sum_{p^*+1}^T \log(\sigma_{t|t-1}^2) - \\ &= -\frac{1}{2} \sum_{p^*+1}^T \frac{(y_t - u_{t|t-1})^2}{\sigma_{t|t-1}^2} \end{aligned} \quad (3.39)$$

y para el caso no paramétrico, denotando por $1/v$ la constante de proporcionalidad, la función de verosimilitud condicional responde a

⁶En las secciones 4.2.1.1 y 4.2.1.2 se presenta una introducción al filtro de Kalman y la representación en Espacios de Estados propuesta en Hamilton (1994). Nótese que para el caso de que se asuma que las innovaciones del modelo no siguen una distribución normal, el filtro de Kalman no proporciona la estimación exacta, aunque sí una buena aproximación (ver Harvey (1990) para detalles)

$$\begin{aligned}
\hat{L}_{SNP}(\varphi) &= \log f(y_{p^*+1}, \dots, y_T | y_1, \dots, y_{p^*}, \varepsilon_0, \varepsilon_0^2; \varphi) = \\
&\frac{T - p^*}{2} \log(2\pi) - \frac{1}{2} \sum_{p^*+1}^T \log(\sigma_{t|t-1}^2) - \\
&-\frac{1}{2} \sum_{p^*+1}^T \frac{(y_t - u_{t|t-1})^2}{\sigma_{t|t-1}^2} + \\
&+(T - p^*) \log(v) + 2 \sum_{p^*+1}^T \log(P_K(z_t)) \tag{3.40}
\end{aligned}$$

En comparación con la estimación exacta de los parámetros, el ajuste condicional supone un coste computacional mucho menor. Si la muestra es suficientemente larga, la contribución de las primeras muestras a la verosimilitud total es insignificante, por lo que habitualmente no existen apenas diferencias entre las estimaciones de ambos enfoques. Por este motivo, los modelos estimados en esta tesis han sido ajustados utilizando el enfoque condicional. Para no complicar la formulación, en lo que resta de documento se utilizará la expresión (3.29) o (3.30) en referencia a la función de verosimilitud condicional sin indicar explícitamente los valores iniciales condicionales o el índice de las muestras utilizadas. Para más detalles sobre la diferencia de ambos enfoques pueden consultarse Hamilton, 1994; Hyndman et al., 2008; Peña, 2005.

Una vez formulada la función de verosimilitud, es necesario maximizarla. La estimación de máxima verosimilitud es el valor de los parámetros del modelo para los cuales (3.30) se maximiza. En principio, esto requiere derivar (3.30) con respecto a cada uno de los parámetros en φ y establecer el resultado igual a cero. En la práctica, el resultado es un sistema de ecuaciones no lineales en φ para el cual no hay una solución sencilla en términos de $(y_1, \dots, y_T)$. Maximizar (3.30) requiere por tanto un procedimiento iterativo de optimización no lineal. Dependiendo de los parámetros de cada modelo particular pueden usarse diferentes procedimientos que tomen ventaja de la forma de la función de verosimilitud. Por ejemplo, para los modelos gaussianos que asumen varianza constante, es decir, $\sigma_t^2 = \alpha_0$, suele usarse la *verosimilitud concentrada*. Esto permite sacar el parámetro α_0 (es decir la varianza) fuera de la función a maximizar de manera que el problema sea equivalente a minimizar la tradicional suma de errores cuadrados (ver Gómez and Maravall, 1994; Harvey, 1990; Hyndman et al., 2008 para detalles). En tal caso resulta muy apropiado usar el algoritmo de minimización no lineal conocido como Levenberg-Marquardt (Levenberg, 1944; Marquardt, 1963; Press et al., 2002). En general, para los modelos que respondan a la formulación descrita en (3.6)-(3.15) con innovaciones gaussianas, la función de verosimilitud puede ser maximizada utilizando una rutina de optimización sin restricciones. Para ello, una opción es reparametrizar la función de verosimilitud de tal manera que los parámetros cumplan siempre las restricciones (3.16)-(3.21). No obstante, también se pueden usar los parámetros originales del modelo y comprobar que se cumplen las restricciones

una vez se han estimado los modelos, lo cual habitualmente ocurrirá si la serie a estimar es estacionaria tanto en media como en varianza (ver Hamilton (1994) p.146 para detalles). En esta tesis se ha usado un algoritmo de minimización sin restricciones especialmente apropiado para funciones de verosimilitud y que está basado en el método *quasi-Newton* y la actualización *BFGS* de la inversa del Hessiano (ver Luenberger and Ye (2008) para más detalles) cuyo código está disponible (<http://www.princeton.edu/~sims/#optimize>). Sin embargo, la estimación de los modelos con densidad SNP, requiere una rutina de optimización con restricciones debido a las restricciones adicionales sobre los coeficientes $a'_k s$ del polinomio P_K en (3.26), necesarias para que la densidad de las innovaciones tenga media 0 y varianza 1. La rutina calcula entonces la forma equivalente ES (Edgeworth-Sargan) para la densidad de las innovaciones con el fin de asegurar que se cumplen dichas restricciones (ver sección 3.2). Las dos rutinas de optimización utilizadas requieren de una estimación inicial de los parámetros. Aquellos concernientes a las funciones de transferencia de las variables explicativas pueden inicializarse a través del algoritmo de *variables instrumentales* (ver Hamilton (1994)). Tras esta primera estimación, los términos autorregresivos y de media móvil se inicializan a través del algoritmo descrito en Hannan and Rissanen (1982). Una vez eliminada la dependencia en media de la serie, una primera estimación del término ARCH se obtiene a partir del cuadrado del residuo por mínimos cuadrados lineales (*Ordinary Least Squares*). Los parámetros $a'_k s$ de la densidad de las innovaciones se inicializan a 0. Con el fin de guiar el proceso de optimización, habitualmente resulta una buena práctica estimar primero los parámetros que determinan la media y varianza del proceso asumiendo la densidad normal estándar, para luego dejar libres los parámetros $a'_k s$ relativos a la densidad no paramétrica y estimar conjuntamente todos los parámetros del modelo.

3.3.2. Diagnóstico y selección del modelo

La sección anterior describe la manera de estimar el conjunto de parámetros del modelo definido en (3.6)-(3.15), (3.25). En base a los resultados de dicha estimación, es posible comprobar la idoneidad de la estructura del modelo (significación de los coeficientes, órdenes de los polinomios, etc) y establecer unos criterios que ayuden a seleccionar la más adecuada para representar el comportamiento estocástico de la variable de salida. En esta sección se resumen estos métodos.

Uno de los primeros análisis de interés es, dada una estructura de modelo fijada, conocer cual sería el verdadero valor del vector de parámetros resultante de la estimación si observáramos todo el proceso estocástico de la variable de salida, y no sólo una particular muestra de datos. O dicho de otra manera, saber si la estimación obtenida de cada uno de los parámetros es lo suficientemente precisa o si cambiaría considerablemente si se tomara otra muestra de datos diferente. Idealmente, si se dispusiera de diversas realizaciones del proceso estocástico de la variable de salida y_t , se podría obtener la distribución de probabilidad de cada parámetro derivada de usar una muestra particular para computar la estimación. Esta es la idea básica

del método *bootstrap* (Efron and Tibshirani (1993)) utilizado en Rodriguez and Ruiz (2008). En el citado trabajo se parte de la serie de residuos estandarizados $\{\hat{z}_t\}_{t=1}^T$ obtenidos a través de la estimación del conjunto de parámetros $\hat{\varphi}$ y la serie original $\{y_t\}_{t=1}^T$, para a partir de ellos generar B réplicas bootstrap $\{\hat{z}_t^s\}_{t=1}^T$, $s = 1, \dots, B$ via permutaciones aleatorias con reemplazamiento. A partir de éstas pueden construirse las correspondientes réplicas bootstrap de la serie original $\{y_t^s\}_{t=1}^T$, $s = 1, \dots, B$ por medio del vector de parámetros estimado $\hat{\varphi}$. Finalmente, se obtiene cada una de las estimaciones del vector de parámetros correspondiente a cada una de las réplicas de la serie original, $\{\hat{\varphi}^s\}$, $s = 1, \dots, B$, con las que se puede construir la distribución de probabilidad de cada parámetro del modelo. Este enfoque presenta la ventaja de no asumir ninguna hipótesis acerca de la distribución de probabilidad de los parámetros del modelo. Sin embargo, su aplicación al modelo aquí descrito resulta muy inconveniente debido al alto coste computacional requerido, puesto que es necesario estimar tantos modelos como réplicas bootstrap se generen.

Afortunadamente, como se pone de manifiesto en Hamilton (1994), si la muestra de datos es suficientemente grande, a menudo la distribución de la estimación de máxima verosimilitud $\hat{\varphi}$ puede ser aproximada por la siguiente distribución

$$\hat{\varphi} \sim N(\varphi_0, T^{-1}\Pi^{-1}) \quad (3.41)$$

donde T es el tamaño de la muestra, φ_0 representa el verdadero vector de parámetros y la matriz Π es conocida como la *matriz de información*, la cual puede ser estimada de dos maneras diferentes. La primera de ellas pasa por calcular la matriz de segundas derivadas de la función de verosimilitud logarítmica evaluada en el punto de máxima verosimilitud, lo que puede hacerse numéricamente. Así, la *matriz de información* responde a

$$\hat{\Pi}_{2D} = -T^{-1} \frac{\partial^2 L(\varphi)}{\partial \varphi \partial \varphi'} \Big|_{\varphi=\hat{\varphi}} \quad (3.42)$$

y sustituyendo en (3.41) proporciona una forma de estimar la matriz de varianzas-covarianzas de los parámetros del modelo, siendo $L(\varphi)$ la verosimilitud logarítmica definida en (3.30). La segunda forma de estimar la *matriz de información* es conocida como *outer-product* y responde a

$$\hat{\Pi}_{OP} = T^{-1} \sum_{t=1}^T [h_t(\hat{\varphi}, \eta_{t-1})] \cdot [h_t(\hat{\varphi}, \eta_{t-1})]' \quad (3.43)$$

Donde $h_t(\hat{\varphi}, \eta_{t-1})$ denota el vector m -dimensional de derivadas del logaritmo de la densidad condicional de la t -ésima observación con respecto a los m elementos del vector de parámetros φ , con la derivada evaluada en la estimación de máxima verosimilitud, es decir,

$$h_t(\hat{\varphi}, \eta_{t-1}) = \left. \frac{\partial \log f[y_t | \eta_{t-1}; \varphi]}{\partial \varphi} \right|_{\varphi = \hat{\varphi}} \quad (3.44)$$

Como ya se puso de manifiesto en la sección (3.2), la estimación de máxima verosimilitud cuando se asume que la densidad de las innovaciones es Gaussiana es una manera razonable de estimar los parámetros de un modelo de serie temporal incluso si los datos no fueron generados por dicha distribución. Este enfoque es conocido como estimación de *quasi-máxima verosimilitud* (ver Bollerslev and Wooldridge (1992)). Las aproximaciones de la matriz de varianzas-covarianzas de los parámetros obtenidas a través de (3.42) o (3.43) son válidas y están razonablemente cerca la una de la otra si los datos fueron realmente generados por un proceso i.i.d Gaussiano. Sin embargo, no son válidas si la densidad de las innovaciones sigue otra distribución diferente. En tal caso, una aproximación adecuada de la matriz de varianzas-covarianzas del vector de parámetros es la propuesta en White (1982), que responde a (3.45)

$$E(\hat{\varphi} - \varphi_0)(\hat{\varphi} - \varphi_0)' = T^{-1} [\Pi_{2D} \Pi_{OP}^{-1} \Pi_{2D}]^{-1} \quad (3.45)$$

La estimación de quasi-máxima verosimilitud comporta por tanto la estimación de los parámetros del modelo asumiendo una distribución gaussiana y el cálculo de los errores estándar según la ecuación (3.45). A partir de esta matriz, los *errores estándar* de los parámetros se calculan tomando la raíz cuadrada de cada elemento de su diagonal. Estos errores estándar se refieren a la desviación típica de la distribución de probabilidad de la estimación de cada parámetro derivada de usar una muestra particular para computarla, y permiten realizar contrastes de hipótesis individuales sobre la significación de cada parámetro incluido en el modelo. Así, puede contrastarse si un determinado parámetro es significativamente distinto de cero a un nivel de confianza del 95 % si el valor absoluto del cociente entre el parámetro estimado y su error estándar es mayor de 1.96 (conocido como *t-test*, ver por ejemplo Pankratz (1991) para detalles).

Los errores estándar de los modelos para los precios eléctricos reportados en la sección 3.4, tanto los correspondientes a los modelos gaussianos como a los no paramétricos, son calculados de acuerdo a (3.45). Para el caso de los modelos gaussianos, este enfoque constituye la aproximación válida puesto que la distribución empírica de las innovaciones no cumple la hipótesis de normalidad. Para los modelos que asumen la distribución no paramétrica descrita en (3.32), el método proporciona una estimación robusta de dichos errores estándar (ver Hansen (1994)). Para más detalles sobre el cálculo de los errores estándar de los parámetros de un modelo pueden consultarse entre otros Bollerslev and Wooldridge, 1992; Hamilton, 1994; Newey and West, 1987.

No obstante, los tests individuales descritos no son definitivos y pueden resultar más importantes aquellos que consideran grupos de parámetros al mismo tiempo.

Así, para el caso de los modelos de precios eléctricos ensayados en la sección 3.4, resulta conveniente comprobar si la inclusión de los parámetros $a_k, k = 1, \dots, K$ que dan forma a la densidad de sus innovaciones contribuyen a obtener una mejora significativa en la estimación de su función de verosimilitud, lo que justificaría el enfoque no paramétrico adoptado. Para tal propósito puede utilizarse el test conocido como *cociente de verosimilitudes*⁷ (*likelihood ratio test*) (Hamilton (1994)). La hipótesis nula de este test establece un conjunto de restricciones sobre los parámetros del modelo, que en este caso son $a_1 = \dots = a_K = 0$. Es decir, dicha hipótesis establece un modelo para los precios eléctricos donde la distribución de probabilidad es Gaussiana. La hipótesis alternativa establece por tanto un modelo que presenta la misma estructura que el anterior pero dónde los a'_k s no están restringidos. De este modo, bajo la hipótesis nula se cumple que

$$2[L(\varphi_u) - L(\varphi_c)] \sim \chi^2(K) \quad (3.46)$$

dónde $L(\varphi_c)$ y $L(\varphi_u)$ denotan el logaritmo del valor de la función de verosimilitud en la estimación de los parámetros restringidos y no restringidos, respectivamente.

Sin embargo, el test anterior tiene una utilidad limitada puesto que únicamente permite elegir entre dos estructuras de modelos diferentes. El procedimiento adoptado en esta tesis como guía para la selección de la estructura idónea del modelo (órdenes autorregresivos y de media móvil, forma específica de las funciones de transferencia, estructura del término ARCH y orden K del polinomio SNP) está basado en las conocidas metodologías descritas en Pankratz, 1991; Box and Jenkins, 1970. Las herramientas estadísticas fundamentales en las que a su vez se basan estas metodologías son las FAS y FAP de los modelos ensayados, los patrones de respuesta impulsional de la salida frente a cada una de las variables explicativas y la significación de los coeficientes calculada según los errores estándar previamente descritos. Una vez ha sido eliminada la dependencia en media de la serie a través de las mencionadas metodologías, la estructura del término ARCH se determina tras inspeccionar la FAS del cuadrado de los residuos (ver Bollerslev (1986)). Finalmente el orden K del polinomio se determina tras la estimación de una batería de modelos con la estructura para la media y varianza ya identificada y con distintos órdenes K , y la posterior inspección visual del ajuste de la densidad asumida a la distribución empírica obtenida, mediante herramientas como q-q plots e histogramas. Por último, señalar que los criterios habituales de selección entre posibles modelos candidatos son también aplicables al caso de modelos con densidades SNP (ver Gallant and Tauchen (1990)). Estos criterios son el criterio de información de Akaike (AIC) y el criterio Bayesiano de Schwarz (SBC), los cuales responde a (3.47) y (3.48), respectivamente

⁷Este test es usualmente utilizado para comparar modelos estimados bajo la hipótesis de normalidad. No obstante, en Gallant and Tauchen (1990) se explica que la inferencia basada en este test con densidades SNP sigue siendo válida.

$$AIC = 2m - 2L(\varphi) \quad (3.47)$$

$$SBC = m \log(T) - 2L(\varphi) \quad (3.48)$$

siendo m el número de parámetros del modelo. El procedimiento usado para identificar la estructura del modelo se ilustra en la sección 3.4 con los precios horarios de diversos mercados internacionales.

3.4. Aplicación a los precios horarios en mercados eléctricos

A continuación se presentan los resultados relativos a la aplicación del modelo propuesto en la sección 3.3 a los precios horarios de varios mercados eléctricos internacionales, en particular: el mercado Ibérico (OMEL), el mercado francés (POWERNEXT), el mercado de Victoria, Australia (NEM) y la interconexión Pensilvania - New Jersey - Maryland (PJM), los cuales se muestran en las gráficas de la figura 2.1. La discusión de esta sección está fundamentalmente encaminada a justificar el modelo asumido para la media y varianza de las series de precios, y en particular, de la pertinencia de usar densidades SNP para modelar su comportamiento estocástico como alternativa a la distribución gaussiana. En el capítulo 5 se mostrará la aplicación de estos modelos a la predicción de los precios horarios del día siguiente, tras establecer la forma de predecir la densidad de dichos precios en un horizonte de predicción de hasta 24 horas.

Los modelos han sido estimados usando datos horarios correspondientes al año 2008 en el caso de OMEL, POWERNEXT y NEM, y 2002 para el caso de PJM. Esto supone 8760 muestras en el caso de los mercados de OMEL, POWERNEXT y PJM, y 17520 horas para NEM puesto que en este caso la subasta diaria comprende 48 períodos de media hora cada uno. Para los mercados de PJM y NEM, se ha utilizado la demanda eléctrica horaria real usada como proxy de la previsión de demanda puesto que no se dispone de esta última, que es a la que los agentes reaccionan. Para OMEL no solo se ha utilizado la previsión de demanda sino también la previsión de producción eólica que está disponible para los agentes en el momento de enviar sus ofertas al mercado. En el caso de POWERNEXT no se ha utilizado ninguna variable explicativa.

Para cada serie de precios se ha identificado la estructura idónea del modelo con la ayuda de la metodología y herramientas explicadas en la sección 3.3.2, asumiendo en primer lugar la hipótesis de normalidad (los modelos así estimados son referenciados como *FT-ARCH* en las tablas y gráficas). Una vez eliminada la dependencia en media y varianza de la serie, se ha estimado un modelo con la misma estructura pero asumiendo en su lugar una distribución SNP (referenciados como *FT-ARCH-SNP*). Los resultados de la estimación de ambas aproximaciones se resumen en la tabla 3.1.

En ella aparecen los valores alcanzados de verosimilitud logarítmica, así como los índices AIC y BIC que derivan de ella. En todos los mercados, el test cociente entre verosimilitudes que compara el modelo gaussiano con el no paramétrico indica que los parámetros adicionales de este último contribuyen significativamente a mejorar el valor de verosimilitud. Los índices AIC y BIC también son claramente favorables al enfoque no paramétrico.

En las tablas 3.4, 3.5, 3.6 y 3.7 aparece el detalle de los parámetros estimados y los errores estándar calculados según (3.45). Para incorporar adecuadamente la autocorrelación estacional que presentan las series de los cuadrados de los errores de predicción, el término ARCH se expresa como un modelo AR $(r) \times (R_1)_{24} \times (R_2)_{168}$ de dichas series. Es decir, en lugar de utilizar la expresión

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_w \varepsilon_{t-w}^2 \quad (3.49)$$

como en (3.8), los parámetros estimados son los que aparecen en la siguiente expresión equivalente

$$\gamma_r(L) \Gamma_{R_1}(L^{24}) \dots \Gamma_{R_2}(L^{168}) \varepsilon_t^2 = \gamma_0 + v_t \quad (3.50)$$

con

$$\gamma_r(L) = 1 - \gamma_1 L - \dots - \gamma_r L^r \quad (3.51)$$

$$\Gamma_{R_1}(L^{24}) = 1 - \Gamma_1 L^{24} - \dots - \Gamma_{R_1} L^{24 \cdot R_1} \quad (3.52)$$

$$\Gamma_{R_2}(L^{168}) = 1 - \Gamma_1 L^{168} - \dots - \Gamma_{R_2} L^{168 \cdot R_2} \quad (3.53)$$

donde $v_t = \varepsilon_t^2 - \sigma_t^2$ (ver sección 2.2.2.5).

Para comprobar el efecto que tienen los parámetros que modelan la varianza y la forma de la función de densidad en las estimación puntual del precio, se ha calculado el *MAE* (*Mean Absolute Error*) de los modelos que asumen errores gaussianos y homocedásticos, los cuales han sido estimados durante el proceso de identificación (referenciados como *FT* en las tablas y gráficas), y comparado con el de aquellos que incorporan el término ARCH adicional (FT-ARCH) y los parámetros que modelan la forma de la función de densidad (FT-ARCH-SNP). Los resultados aparecen en la tabla 3.2. En ella se aprecia cómo los errores absolutos no se ven prácticamente alterados con la incorporación de los términos ARCH, ni tampoco con la de los parámetros a'_k s que modelan la forma de la función de densidad.

Así mismo, con el ánimo de evaluar y contrastar la calidad de las densidades predictivas proporcionadas por los modelos estimados, se han utilizado algunas medidas comúnmente utilizadas para tal propósito (ver Gneiting and Raftery (2007)) y que difieren de la verosimilitud logarítmica que es la medida que se maximiza al

estimar los parámetros de los modelos. Estas medidas son el *Quadratic Score* (QS), *Spherical Score* (SS) y el *Continuous Ranked Probability Score* (CRPS), las cuales responden a

$$QS_t(\hat{f}_t, y_t) = 2\hat{f}_t(y_t) - \int_{-\infty}^{\infty} \hat{f}_t(h)^2 dh \quad (3.54)$$

$$SS_t(\hat{f}_t, y_t) = \frac{\hat{f}_t(y_t)}{\left(\int_{-\infty}^{\infty} \hat{f}_t(h)^2 dh \right)^{1/2}} \quad (3.55)$$

$$CRPS_t(\hat{F}_t, y_t) = \int_{-\infty}^{\infty} \left(\hat{F}_t(h) - 1\{h \geq y_t\} \right)^2 dh \quad (3.56)$$

donde $\hat{f}_t$ y $\hat{F}_t$ se refieren a las funciones de densidad y de distribución (acumulada) predictivas en cada instante de tiempo t , y donde la función $1\{\cdot\}$ toma valor 1 cuando la condición entre corchetes es verdad y 0 en otro caso. Estas medidas pueden ser fácilmente calculadas utilizando la formulación equivalente ES a través de la ecuación 3.2 (para el caso de las distribuciones SNP) e integración numérica. Nótese que a diferencia del QS o el SS, el CRPS tiene orientación negativa, y por tanto, cuando más bajo sea su valor mejor valoración recibe la densidad predicha. Los resultados de la media de cada una de ellas se resumen en la tabla 3.3. Como puede observarse, las medidas QS y SS son claramente favorables a los modelos FT-ARCH-SNP en todos los mercados. El CRPS de dichos modelos resulta también claramente favorable en todos los mercados en relación a los modelos FT, y en OMEL y POWERNEXT cuando se compara con los modelos FT-ARCH. En PJM y NEM esta medida resulta mínimamente superior para los modelos que no modelan la densidad de sus innovaciones a través de las densidades SNP (FT-ARCH). Este hecho viene causado por la fuerte penalización que este tipo de medida impone a la distancia entre el valor realizado de la serie y la distribución predictiva, dado que como se aprecia en la tabla 3.2 el error absoluto en estos dos mercados es ligeramente mayor para los modelos con densidades SNP (ver sección 5.3.2 para más discusión sobre medidas de predicción probabilística).

Los gráficos q-q plots de las distribuciones asumidas frente a las empíricas obtenidas (ver figuras 3.4 y 3.5) muestran de qué manera las densidades SNP se ajustan mejor a la distribución de los errores estandarizados (si procede) de cada modelo que la distribución normal, tanto cuando se asume varianza variable como cuando, y en mayor medida, se asume homocedasticidad. Así mismo, las gráficas de la figura 3.6 muestran el ajuste de la densidad no paramétrica a la distribución empírica (izquierda) y la forma de la densidad SNP aproximada frente a la de la normal estándar. En ellas puede observarse que las distribuciones obtenidas son ligeramente asimétricas a la izquierda y mucho más apuntadas que la distribución normal, hecho que es adecuadamente capturado a través de las densidades SNP. Sin embargo, la aproximación de estas densidades se aleja algo de la distribución empírica

al final de las colas. Este hecho está causado principalmente por el escaso volumen de datos atípicos, el cual hace que el ajuste en esos puntos no sea del todo preciso. Además, la magnitud de dichos valores requiere de un polinomio de orden muy alto para modelarlos, lo que resulta computacionalmente muy costoso. En las gráficas de la figura 3.7 se muestra más claramente el comportamiento en la cola derecha de las funciones de distribución SNP para cada mercado (Nótese que la escala de las abscisas en la figura es logarítmica). Aunque el ajuste obtenido no es exacto, los niveles de probabilidad proporcionados son razonables teniendo en cuenta el volumen de datos utilizado para el ajuste, y en cualquier caso, son más aproximados que los proporcionados por la distribución normal.

Finalmente, algunos de los resultados parciales obtenidos en el proceso de identificación de los modelos aparecen en la figuras 3.8, 3.9, 3.10 y 3.11. A continuación se explican los gráficos mostrados en estas figuras:

- las gráficas referenciadas por la letra a) (3.8a, 3.9a, 3.10a y 3.11a) muestran el gráfico de secuencia del residuo que se obtiene una vez se ha identificado la forma de las funciones de transferencia de las variables explicativas y las diferencias a aplicar. dicho residuo, que llamaremos *error de regresión* N_t , responde a 3.57

$$N_t = \nabla^d \nabla_{S_1}^{D_1} \dots \nabla_{S_m}^{D_m} \left(y_t - \left(C + \sum_i^n G_i(L)x_{it} \right) \right) \quad (3.57)$$

- las gráficas referenciadas por las letras b) y c) (3.8b, 3.8c, 3.9b, 3.9c, 3.10b, 3.10c, 3.11b y 3.11c) muestran las FAS's y FAP's del residuo N_t utilizadas para identificar la forma de los términos autorregresivos y de media móvil, respectivamente.
- las gráficas referenciadas por la letra d) (3.8d, 3.9d, 3.10d y 3.11d) muestran las FAS's del cuadrado de los residuos una vez se ha identificado la estructura para la media.
- las gráficas referencias por la letra e) (3.8e, 3.9e, 3.10e y 3.11e) muestran las FAS's de los residuos estandarizados $(\hat{\sigma}_t^{-1}(y_t - \hat{y}_t))$ de los modelos SNP finalmente identificados.
- las gráficas referenciadas por la letra f) (3.8f, 3.9f, 3.10f y 3.11f) muestran las FAS's del cuadrado de los residuos estandarizados de los modelos SNP finalmente identificados.
- las gráficas referenciadas por la letra g) (3.8g, 3.9g, 3.10g y 3.11g) muestran los residuos estandarizados de los modelos SNP ajustados.

En base a los gráficos de estas figuras caben hacer algunos comentarios. En general, la serie de residuo N_t , obtenida tras descontar el efecto de las variables explicativas y aplicar las diferenciaciones requeridas presenta volatilidad variable

Mercado	Modelo	Ver. log.	AIC	BIC	N. parám.
OMEL	FT-ARCH	-34.08	94.15	185.62	13
OMEL	FT-ARCH-SNP	162.74*	-285.49	-144.77	20
PJM	FT-ARCH	8146.00	-16255.99	-16129.77	18
PJM	FT-ARCH-SNP	8424.97*	-16793.94	-16597.60	28
POWERN	FT-ARCH	1771.97	-3521.94	-3444.55	11
POWERN	FT-ARCH-SNP	2216.52*	-4393.05	-4252.33	20
NEM	FT-ARCH	11130.72	-22229.43	-22105.77	16
NEM	FT-ARCH-SNP	11572.53*	-23092.90	-22891.42	26

Tabla 3.1: Resultados de la estimación de los modelos para cada mercado (* indica que el modelo rechaza la hipótesis nula del test cociente entre verosimilitudes descrito en la sección 3.3.2 al 5 % de significación).

Tipo Modelo	OMEL	PJM	POWERNEXT	NEM
FT	2.23	2.09	4.73	3.72
FT-ARCH	2.28	2.165	4.75	3.63
FT-ARCH-SNP	2.23	2.185	4.77	3.82

Tabla 3.2: *MAE's* del conjunto de entrenamiento en los distintos mercados eléctricos (OMEL y POWERNEXT (€/MWh), PJM y NEM (\$/MWh)).

y un número de valores atípicos más frecuente que en una distribución normal (aunque este último hecho ocurre en menor medida en OMEL). Las FAS's y FAP's del residuo N_t muestran una clara autocorrelación que es más relevante para los términos estacionales que para los regulares. También se observa en las FAP's un comportamiento de media móvil de período semanal que es común a las cuatro series consideradas. Una vez ha sido descontada la autocorrelación en media de la serie, el cuadrado de los residuos muestra también una clara autocorrelación regular y estacional. Las FAS's de los residuos estandarizados, y sus cuadrados, de los modelos SNP finalmente estimados no muestran autocorrelaciones significativas en los retardos de orden bajo tanto regulares como estacionales, lo que implica que se ha eliminado las dependencias lineales en media y varianza de las series.

Mercado	Tipo Modelo	QS	SS	CRPS
OMEL	FT	1.0701	1.0346	0.14644
OMEL	FT-ARCH	1.2167	1.0903	0.13928
OMEL	FT-ARCH-SNP	1.261	1.1096	0.13915
PJM	FT	2.5711	1.6038	0.059079
PJM	FT-ARCH	3.6271	1.8386	0.055787
PJM	FT-ARCH-SNP	4.0761	1.9265	0.057604
POWERN	FT	1.3926	1.1808	0.11486
POWERN	FT-ARCH	1.6202	1.2576	0.10849
POWERN	FT-ARCH-SNP	1.733	1.2999	0.10832
NEM	FT	1.5162	1.2338	0.096299
NEM	FT-ARCH	2.8404	1.629	0.081568
NEM	FT-ARCH-SNP	2.9367	1.6541	0.091476

Tabla 3.3: Valores medios del *Quadratic Score* (QS), *Spherical Score* (SS) y *Continuous Ranked Probability Score* (CRPS) evaluados en el conjunto de entrenamiento para cada uno de los mercados.

Parámetros	FT-ARCH		FT-ARCH-SNP	
	Valor	Est. Error	Valor	Est. Error
$w_{10}(dem)$	0.3133	0.0267	0.3199	0.0156
$w_{20}(eol)$	-0.1490	0.0084	-0.1493	0.0073
ϕ_1	-0.7854	0.0100	-0.8099	0.0090
θ_1	-0.9536	0.0035	-0.9554	0.0045
$\Phi_1(24)$	-0.0765	0.0163	-0.0866	0.0133
$\Phi_1(168)$	0.0573	0.0170	0.0465	0.0127
$\Theta_1(24)$	-0.7883	0.0102	-0.7937	0.0071
$\Theta_1(168)$	-0.7391	0.0123	-0.7381	0.0071
γ_0	0.0163	0.0015	0.0157	0.0016
γ_1	0.1532	0.0184	0.1422	0.0169
γ_2	0.0365	0.0101	0.0315	0.0083
$\Gamma_1(24)$	0.2481	0.0208	0.2581	0.0291
$\Gamma_1(168)$	0.2418	0.0193	0.2660	0.0257
a_1	—	—	0.0216*	0.0137
a_2	—	—	-0.1087	0.0382
a_3	—	—	-0.0173	0.0091
a_4	—	—	0.0129	0.0074
a_5	—	—	0.0029	0.0015
a_6	—	—	0.0003**	0.0005
a_7	—	—	-0.0001	0.0001

Tabla 3.4: Coeficientes estimados para los modelos FT-ARCH y FT-ARCH-SNP en OMEL (* indica que el parámetro es significativo al 10 % de significación. ** indica que el parámetro no es significativo. El resto de parámetros son significativos al 5 % de significación).

Parámetros	FT-ARCH		FT-ARCH-SNP	
	Valor	Est. Error	Valor	Est. Error
$w_{10}(dem)$	0.2865	0.0061	0.258290	0.004651
ϕ_1	-0.7306	0.0106	-0.741450	0.004976
ϕ_2	-0.0996	0.0097	-0.118970	0.005670
θ_1	-0.7938	0.0051	-0.789420	0.001926
θ_2	-0.2032	0.0047	-0.206940	0.001901
$\Phi_1(24)$	0.3010	0.0163	0.245050	0.005948
$\Phi_1(168)$	0.2014**	0.3434	-0.031438	0.009294
$\Phi_2(24)$	-0.6940	0.0071	-0.712510	0.005499
$\Phi_2(168)$	-0.0399*	0.0267	0.064616	0.005322
$\Theta_1(24)$	0.6503	0.0208	0.563630	0.008052
$\Theta_1(168)$	-0.5288*	0.3433	-0.789020	0.010507
$\Theta_2(24)$	-0.3439	0.0115	-0.379130	0.007277
$\Theta_2(168)$	-0.2313**	0.2764	0.057653	0.009152
γ_0	0.0011	0.0000	0.000512	0.000020
γ_1	0.2872	0.0081	0.285500	0.009318
γ_2	0.0477	0.0042	0.062054	0.002899
$\Gamma_1(24)$	0.2469	0.0088	0.294010	0.005687
$\Gamma_1(168)$	0.2016	0.0081	0.138070	0.003069
a_1	—	—	-0.074233	0.016604
a_2	—	—	-0.129490	0.019296
a_3	—	—	0.066386	0.016527
a_4	—	—	0.023756	0.010766
a_5	—	—	-0.019347	0.004727
a_6	—	—	-0.000852	0.002303
a_7	—	—	0.001596	0.000479
a_8	—	—	0.000390	0.000193
a_9	—	—	-0.000033	0.000016
a_{10}	—	—	-0.000024	0.000005

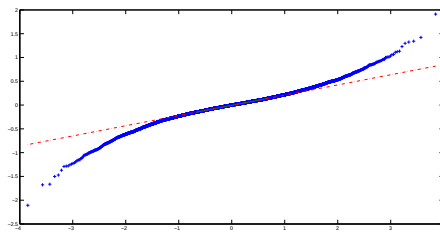
Tabla 3.5: Coeficientes estimados para los modelos FT-ARCH y FT-ARCH-SNP en PJM (* indica que el parámetro es significativo al 10 % de significación. ** indica que el parámetro no es significativo. El resto de parámetros son significativos al 5 % de significación).

Parámetros	FT-ARCH		FT-ARCH-SNP	
	Valor	Est. Error	Valor	Est. Error
ϕ_1	-0.8269	0.0074	-0.854720	0.004563
θ_1	-0.9663	0.0029	-0.973990	0.001899
$\Phi_1(24)$	-1.0361	0.0164	-1.066600	0.006946
$\Phi_1(168)$	-0.0550	0.0114	-0.100380	0.007357
$\Phi_2(24)$	0.0980	0.0121	0.124510	0.005510
$\Theta_1(24)$	-0.8324	0.0128	-0.852110	0.005933
$\Theta_1(168)$	-0.8016	0.0051	-0.824670	0.004192
γ_0	0.0161	0.0003	0.017046	0.000280
γ_1	0.2248	0.0071	0.270200	0.008407
$\Gamma_1(24)$	0.2151	0.0097	0.240880	0.008454
$\Gamma_1(168)$	0.1622	0.0075	0.173750	0.004374
a_1	—	—	0.002060**	0.013321
a_2	—	—	-0.271200	0.014203
a_3	—	—	-0.007047	0.010822
a_4	—	—	0.067935	0.005547
a_5	—	—	0.003158	0.002355
a_6	—	—	-0.005959	0.000677
a_7	—	—	-0.000443**	0.000163
a_8	—	—	0.000136	0.000024
a_9	—	—	-0.000004*	0.000003

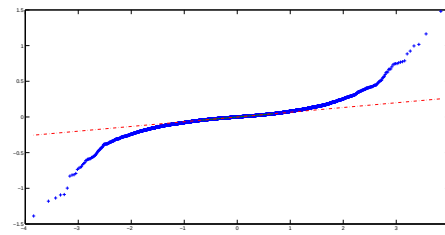
Tabla 3.6: Coeficientes estimados para los modelos FT-ARCH y FT-ARCH-SNP en POWERNEXT (* indica que el parámetro es significativo al 10 % de significación. ** indica que el parámetro no es significativo. El resto de parámetros son significativos al 5 % de significación).

Parámetros	FT-ARCH		FT-ARCH-SNP	
	Valor	Est. Error	Valor	Est. Error
$w_{10}(dem)$	0.2409	0.0036	0.283450	0.001919
ϕ_1	-0.3145	0.0130	-0.465980	0.003476
ϕ_2	-0.4853	0.0113	-0.353790	0.003155
θ_1	-0.4564	0.0129	-0.618000	0.001416
θ_2	-0.5412	0.0130	-0.379820	0.001387
$\Phi_1(48)$	-0.7356	0.0494	0.308730	0.031318
$\Phi_1(336)$	-0.0401	0.0055	-0.016274	0.005239
$\Phi_2(48)$	-0.1182	0.0396	-0.127610	0.005682
$\Theta_1(48)$	-0.5877	0.0480	-0.404890	0.032399
$\Theta_1(336)$	-0.7248	0.0020	-0.652100	0.003342
$\Theta_2(48)$	-0.1673	0.0333	-0.432210	0.028162
γ_0	0.0021	0.0000	0.002244	0.000044
γ_1	0.3648	0.0031	0.347120	0.003479
γ_2	0.1417	0.0025	0.137480	0.002505
$\Gamma_1(48)$	0.1457	0.0022	0.140040	0.002321
$\Gamma_1(336)$	0.1588	0.0032	0.181680	0.002517
a_1	—	—	-0.006701**	0.010703
a_2	—	—	-0.186740	0.011567
a_3	—	—	-0.016740	0.009546
a_4	—	—	0.033051	0.006196
a_5	—	—	0.004346	0.002427
a_6	—	—	-0.002317	0.001124
a_7	—	—	-0.000116**	0.000215
a_8	—	—	0.000237	0.000079
a_9	—	—	-0.000009*	0.000006
a_{10}	—	—	-0.000010	0.000002

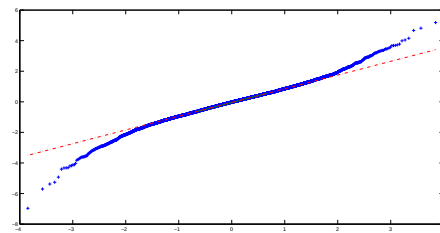
Tabla 3.7: Coeficientes estimados para los modelos FT-ARCH y FT-ARCH-SNP en NEM (* indica que el parámetro es significativo al 10 % de significación. ** indica que el parámetro no es significativo. El resto de parámetros son significativos al 5 % de significación).



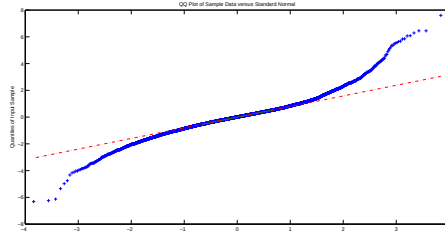
(a) Modelo FT, OMEL



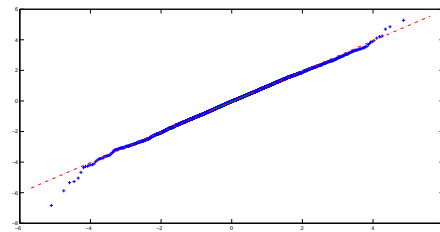
(b) Modelo FT, PJM



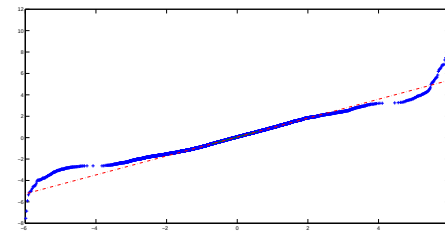
(c) Modelo FT-ARCH, OMEL



(d) Modelo FT-ARCH, PJM

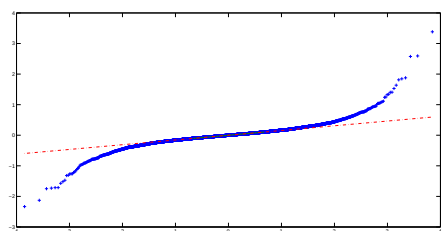


(e) Modelo FT-ARCH-SNP, OMEL

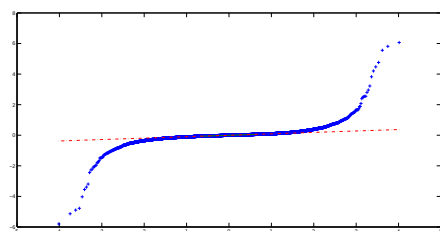


(f) Modelo FT-ARCH-SNP, PJM

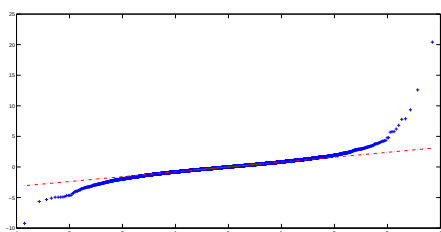
Figura 3.4: q-q plots de los modelos estimados para el mercado de OMEL (izquierda) y PJM (derecha).



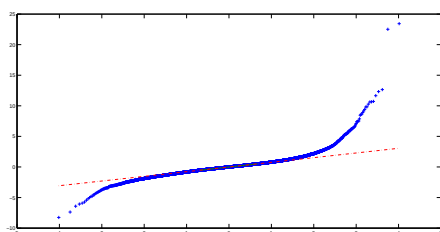
(a) Modelo FT, POWERNEXT



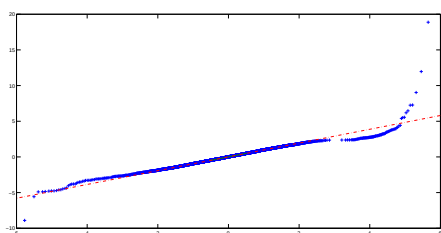
(b) Modelo FT, NEM



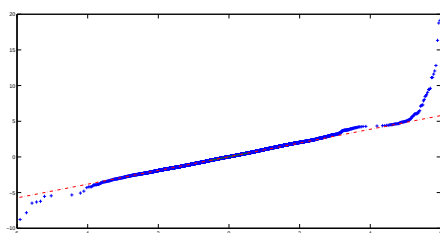
(c) Modelo FT-ARCH, POWERNEXT



(d) Modelo FT-ARCH, NEM

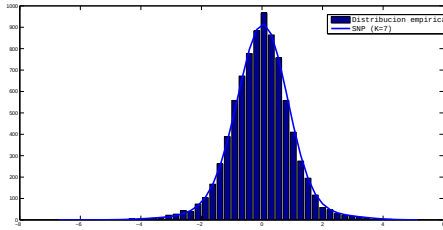


(e) Modelo FT-ARCH-SNP, POWERNEXT

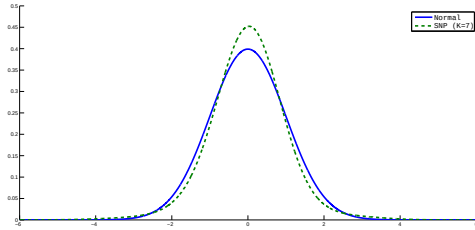


(f) Modelo FT-ARCH-SNP, NEM

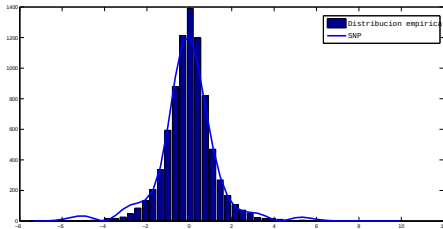
Figura 3.5: q-q plots de los modelos estimados para el mercado de POWERNEXT (izquierda) y NEM Victoria (derecha).



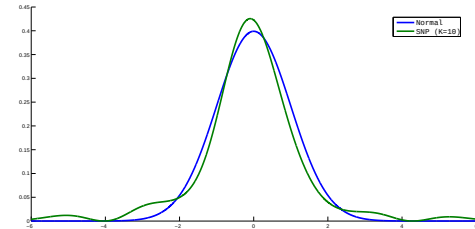
(a) Histograma, OMEL



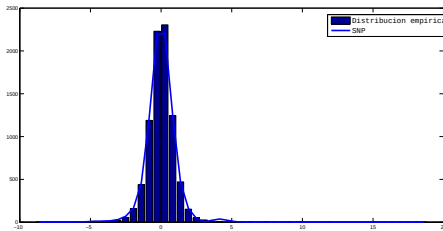
(b) Distribución teórica SNP (K=7), OMEL



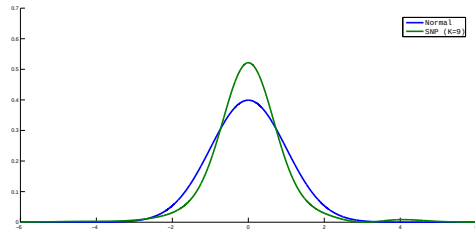
(c) Histograma, PJM



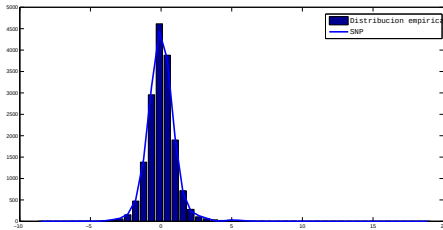
(d) Distribución teórica SNP (K=10), PJM



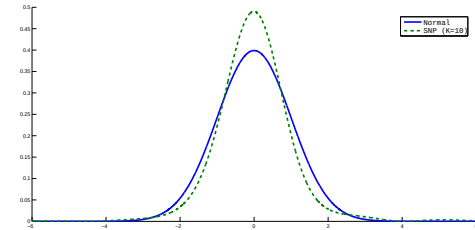
(e) Histograma, POWERNEXT



(f) Distribución teórica SNP (K=9), POWERNEXT

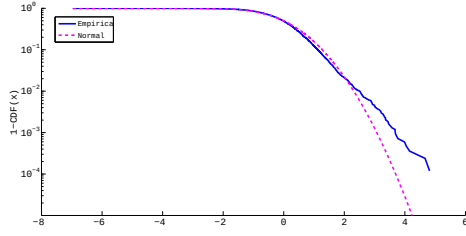


(g) Histograma, NEM

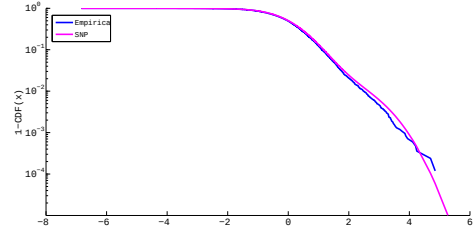


(h) Distribución teórica SNP (K=10), NEM

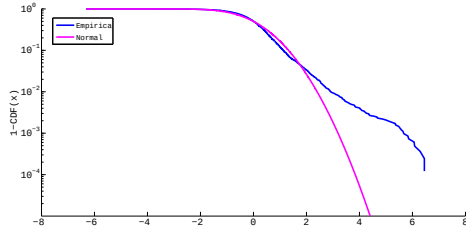
Figura 3.6: Histograma de los residuos estandarizados de los modelos FT-ARCH-SNP y Distribución teórica SNP, para cada mercado.



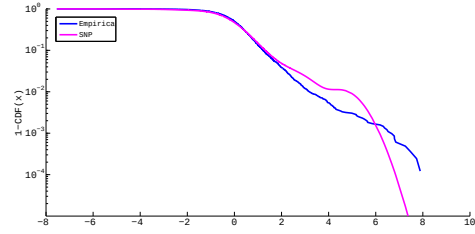
(a) Normal, OMEL



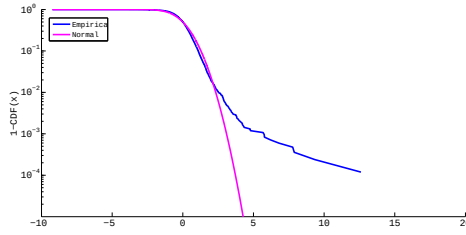
(b) SNP, OMEL



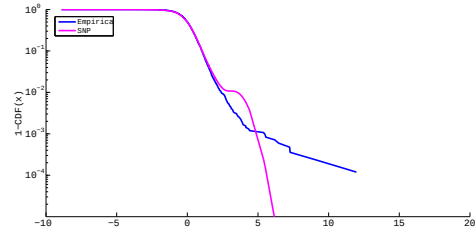
(c) Normal, PJM



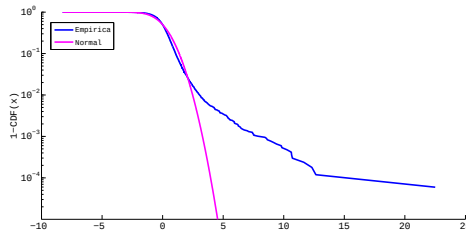
(d) SNP, PJM



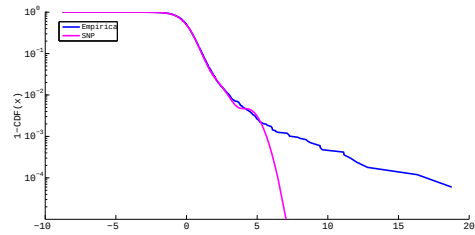
(e) Normal, POWERNEXT



(f) SNP, POWERNEXT

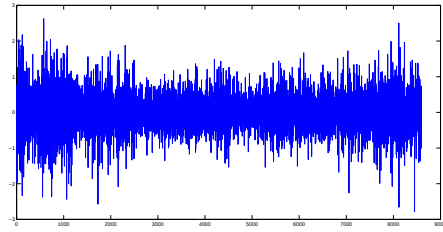
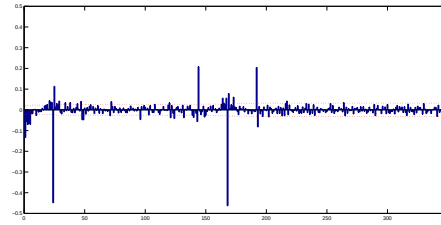
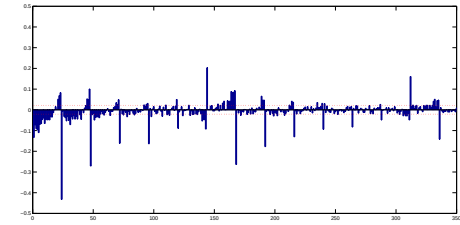
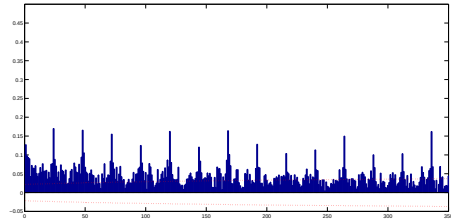


(g) Normal, NEM

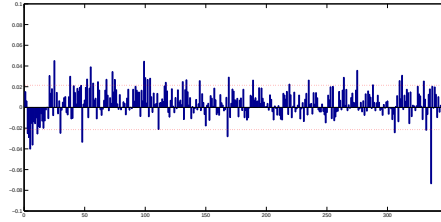


(h) SNP, NEM

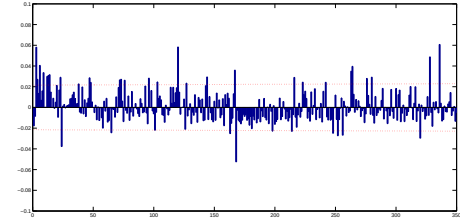
Figura 3.7: Colas derechas de las distribuciones asumidas (Normal a la izquierda y SNP a la derecha) frente a las empíricas obtenidas para cada mercado.

(a) Error de regresión N_t (b) FAS del Error de regresión N_t (c) FAP del Error de regresión N_t 

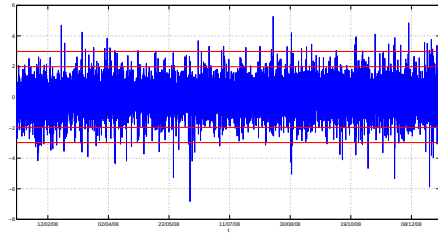
(d) FAS cuadrado residuos modelo FT



(e) FAS residuos estandarizados modelo FT-ARCH-SNP

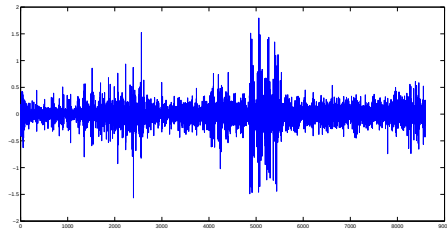
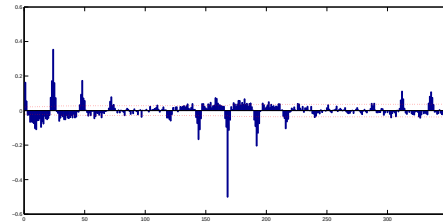
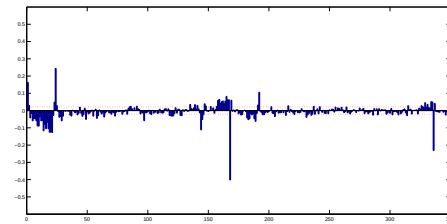
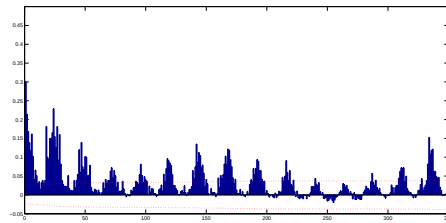


(f) FAS cuadrado residuos estandarizados modelo FT-ARCH-SNP

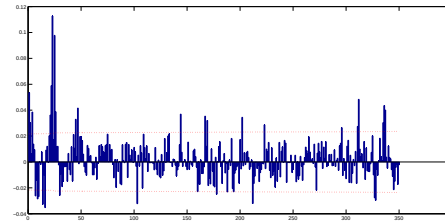


(g) Residuos estandarizados modelo FT-ARCH-SNP

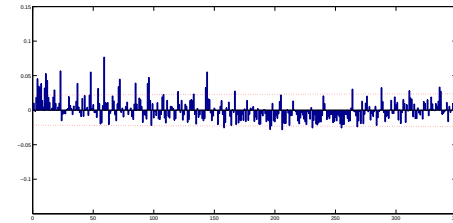
Figura 3.8: Resultados parciales del proceso de identificación de los modelos para los precios horarios en OMEL.

(a) Error de regresión N_t (b) FAS del Error de regresión N_t (c) FAP del Error de regresión N_t 

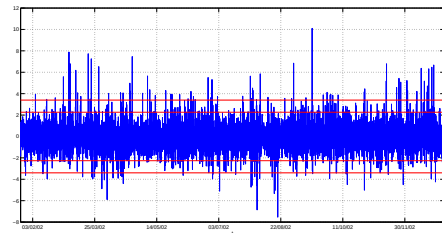
(d) FAS cuadrado residuos modelo FT



(e) FAS residuos estandarizados modelo FT-ARCH-SNP

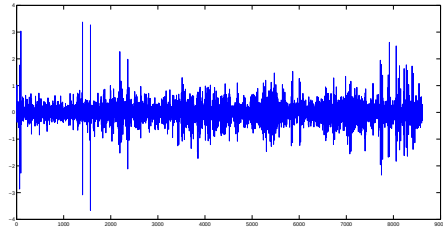
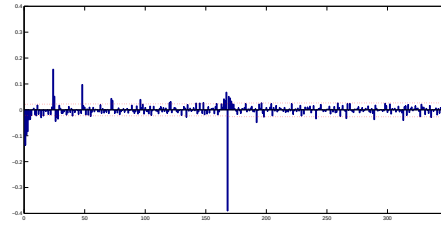
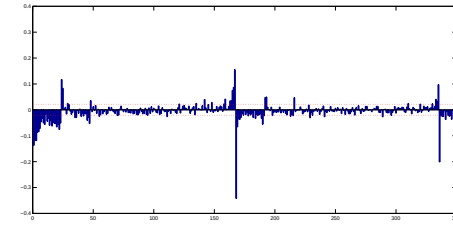
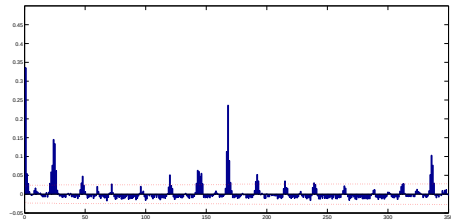


(f) FAS cuadrado residuos estandarizados modelo FT-ARCH-SNP

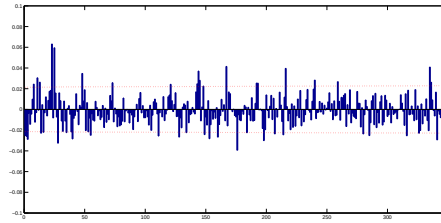


(g) Residuos estandarizados modelo FT-ARCH-SNP

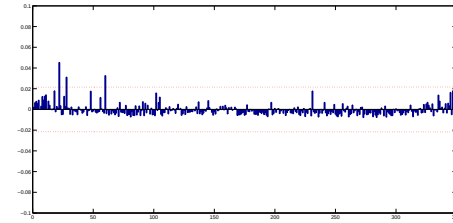
Figura 3.9: Resultados parciales del proceso de identificación de los modelos para los precios horarios en PJM.

(a) Error de regresión N_t (b) FAS del Error de regresión N_t (c) FAP del Error de regresión N_t 

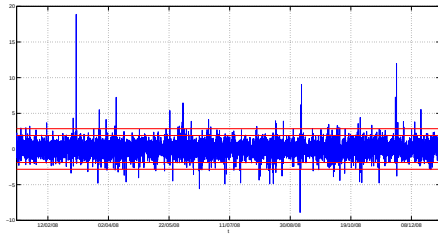
(d) FAS cuadrado residuos modelo FT



(e) FAS residuos estandarizados modelo FT-ARCH-SNP

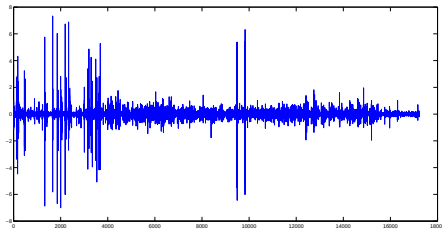
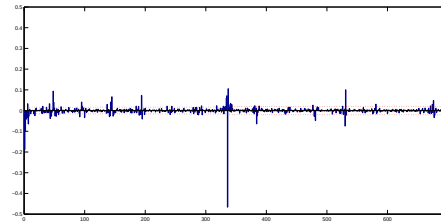
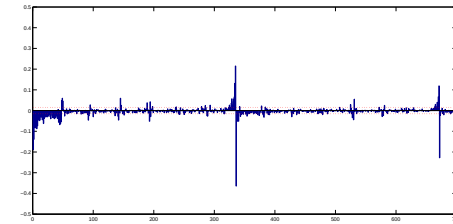
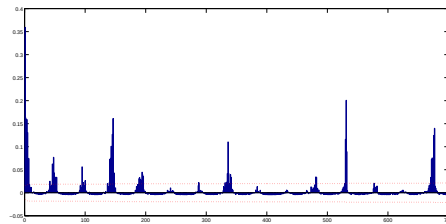


(f) FAS cuadrado residuos estandarizados modelo FT-ARCH-SNP

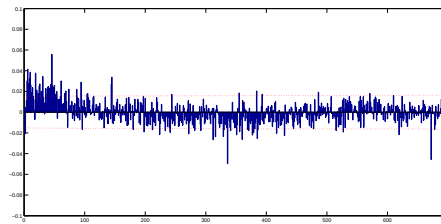


(g) Residuos estandarizados modelo FT-ARCH-SNP

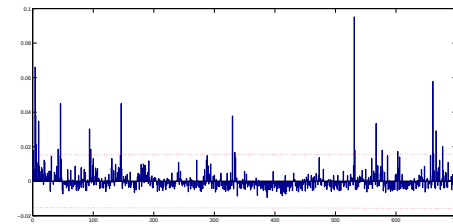
Figura 3.10: Resultados parciales del proceso de identificación de los modelos para los precios horarios en POWERNEXT.

(a) Error de regresión N_t (b) FAS del Error de regresión N_t (c) FAP del Error de regresión N_t 

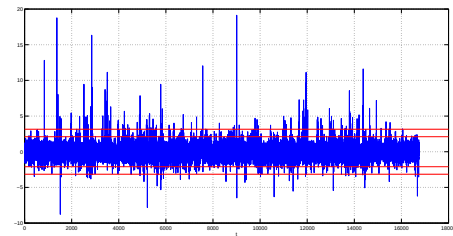
(d) FAS cuadrado residuos modelo FT



(e) FAS residuos estandarizados modelo FT-ARCH-SNP



(f) FAS cuadrado residuos estandarizados modelo FT-ARCH-SNP



(g) Residuos estandarizados modelo FT-ARCH-SNP

Figura 3.11: Resultados parciales del proceso de identificación de los modelos para los precios horarios en NEM.

3.5. Conclusiones

Las habituales hipótesis de normalidad y homocedasticidad asumidas a la hora de estimar los modelos de función de transferencia con ruido ARIMA no se cumplen para las series de precios eléctricos. Estos modelos no contemplan el carácter heterocedástico de las series, la posible asimetría de las distribuciones resultantes ni la probabilidad de que aparezcan valores inusuales o extremos. Este hecho no resulta un problema si el modelo es concebido únicamente para obtener predicciones puntuales, puesto que el marco de trabajo de estimación de quasi-máxima verosimilitud proporciona una solución suficientemente robusta. Sin embargo, resulta inadecuado si se quiere estimar de manera razonable la probabilidad de cualquier intervalo de predicción, es decir, si se quiere predecir su función de densidad.

Como algunos investigadores han señalado previamente, una conveniente solución al hecho de que las series presenten volatilidad variable es la inclusión de un término ARCH que modele la varianza del proceso. Los resultados de la sección 3.4 muestran que la inclusión de este término no sólo contribuye a estimar mejor la densidad de los precios eléctricos, sino que no supone un empobrecimiento significativo de las predicciones puntuales. Sin embargo, la forma de las distribuciones obtenidas en cada mercado particular hacen necesaria una solución lo suficientemente general y flexible para modelar los momentos mayores al segundo. En la sección 3.2 se ha presentado la familia de densidades propuesta en Gallant and Nychka (1987), a la que nos referimos como SNP. Dicho enfoque constituye una solución no paramétrica, y presenta la ventaja de ser capaz de representar formas generales de funciones de densidad e incrementar la capacidad de ajuste a una distribución dada aumentando el número de parámetros. La solución proporcionada por esta familia de densidades es, a nuestro juicio, muy adecuada para modelar las distribuciones de los precios eléctricos, puesto que son capaces de capturar fácilmente ligeras asimetrías y colas más pesadas que las de las distribuciones normales. Como desventaja se puede señalar el coste computacional que supone estimar estas densidades cuando se precisa de un polinomio de orden alto.

El modelo propuesto en este capítulo flexibiliza las hipótesis de partida habituales en el modelo de función de transferencia con ruido ARIMA, con el fin de que también pueda ser usado para proporcionar estimaciones razonables de la función de densidad condicional de los precios eléctricos. El modelo es, por tanto, una extensión semiparamétrica del modelo de función de transferencia con ruido ARIMA que incorpora un término ARCH y donde las innovaciones siguen una densidad SNP. El modelo captura las características esenciales de los precios eléctricos: influencia de variables explicativas, estacionalidad, carácter autorregresivo, y volatilidad variable en el tiempo, y trata de representar más fielmente las asimetrías y colas pesadas habitualmente presentes en las distribuciones empíricas a través de un enfoque no paramétrico. La manera de estimar los parámetros del modelo supone una generalización del habitual marco de trabajo de máxima verosimilitud asumiendo normalidad. El modelo se ha aplicado a las series de precios de OMEL,

POWERNEXT, NEM y PJM, y en todos los casos se aprecia una notoria mejoría en la representación de la función de densidad con respecto al ajuste proporcionado por distribuciones gaussianas.

3.6. Apéndice A: Polinomios de Chebyshev-Hermite

Denotando por

$$\rho(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad (3.58)$$

$$D = \frac{d}{dx} \quad (3.59)$$

y tomando sucesivas derivadas de $\rho(x)$ con respecto a x , se tiene

$$\begin{aligned} D\rho(x) &= -x\rho(x) \\ D^2\rho(x) &= (x^2 - 1)\rho(x) \\ D^3\rho(x) &= (3x - x^3)\rho(x) \\ &\dots \end{aligned}$$

El resultado es por tanto un polinomio en x multiplicado por $\rho(x)$. Se define entonces el polinomio de Chebyshev-Hermite $H_r(x)$ por la ecuación (3.60)

$$(-D)^r \rho(x) = H_r(x) \rho(x) \quad (3.60)$$

donde $H_r(x)$ es un polinomio de grado r en x y el coeficiente de x^r es la unidad. Por convención $H_0 = 1$. La fórmula cerrada viene dada por (3.61)

$$H_r(x) = \sum_{s=0}^{r/2} (-1)^s x^{r-2s} \frac{r^{[2s]}}{2^s s!} \quad (3.61)$$

donde $r^{[2s]} = (r(r-1)(r-2) \dots (r-2s+1))$, y $s = 0, \dots, ([r-1]/2)$ si r es impar.

Los diez primeros polinomios son entonces

$$\begin{aligned}
H_0 &= 1 \\
H_1 &= x \\
H_2 &= x^2 - 1 \\
H_3 &= x^3 - 3x \\
H_4 &= x^4 - 6x^2 + 3 \\
H_5 &= x^5 - 10x^3 + 15x \\
H_6 &= x^6 - 15x^4 + 45x^2 - 15 \\
H_7 &= x^7 - 21x^5 + 105x^3 - 105x \\
H_8 &= x^8 - 28x^6 + 210x^4 - 420x^2 + 105 \\
H_9 &= x^9 - 36x^7 + 378x^5 - 1260x^3 + 945x \\
H_{10} &= x^{10} - 45x^8 + 630x^6 - 3150x^4 + 4725x^2 - 945
\end{aligned} \tag{3.62}$$

Los polinomios cumplen una serie de interesantes propiedades. Así

$$\frac{d}{dx}H_r(x) = rH_{r-1}(x) \tag{3.63}$$

y de manera general para $r \geq j$

$$D^j H_r(x) = r^j H_{r-j}(x) \tag{3.64}$$

También cumplen que

$$H_r(x) - xH_{r-1}(x) + (r-1)H_{r-2}(x) = 0 \tag{3.65}$$

por lo que el polinomio de grado r puede calcularse a partir de los polinomios de grado $r-1$ y $r-2$.

La propiedad más relevante es que son ortogonales con respecto a la función de peso $\rho(x)$, de esta manera

$$\begin{aligned}
\int_{-\infty}^{\infty} H_m(x)H_n(x)\rho(x)dx &= 0, \text{ si } m \neq n \\
&= n!, \text{ si } m = n
\end{aligned} \tag{3.66}$$

Este resultado se demuestra considerando que, si $m \leq n$, por (3.60) se tiene que

$$\int_{-\infty}^{\infty} H_m(x)H_n(x)\rho(x)dx = (-1)^n \int_{-\infty}^{\infty} H_m(x)D^n \rho(x)dx \tag{3.67}$$

e integrando ahora por partes, se tiene que

$$\begin{aligned} \int_{-\infty}^{\infty} H_m(x) H_n(x) \rho(x) dx &= (-1)^n [H_m(x) D^{n-1} \rho(x)]_{-\infty}^{\infty} + \\ &+ (-1)^{n-1} \int_{-\infty}^{\infty} \frac{d}{dx} H_m(x) D^{n-1} \rho(x) dx \end{aligned} \quad (3.68)$$

puesto que $\rho(x) \rightarrow 0$ cuando $x \rightarrow \pm\infty$, el primer sumando de la expresión en (3.68) desaparece, y en virtud de (3.63) se tiene que

$$\int_{-\infty}^{\infty} H_m(x) H_n(x) \rho(x) dx = m(-1)^{n-1} \int_{-\infty}^{\infty} H_{m-1}(x) D^{n-1} \rho(x) dx \quad (3.69)$$

por lo que, continuando el proceso de manera recursiva, se obtiene $m!$ si $m = n$ y 0 en otro caso. Puesto que $H_0(x) = 1$ se deduce directamente que

$$\int_{-\infty}^{\infty} H_m(x) \rho(x) dx = 0, \text{ si } m \geq 1 \quad (3.70)$$

Para más detalles sobre las propiedades de los polinomios de Chebyshev-Hermite pueden consultarse Kendall and Stuart, 1969; Mauleón, 2010, 2003, 2006.

3.7. Apéndice B: Propiedades de la familia de densidades Edgeworth-Sargan

La expresión de la función de densidad en (3.1) integra a 1 en virtud de la propiedad en (3.70).

$$\begin{aligned}
 \int_{-\infty}^{\infty} f(\epsilon) d\epsilon &= \int_{-\infty}^{\infty} \rho(\epsilon) \times \left[1 + \sum_{k=1}^K d_k H_k(\epsilon) \right] d\epsilon \\
 &= \int_{-\infty}^{\infty} \rho(\epsilon) d\epsilon + \sum_{k=1}^K d_k \int_{-\infty}^{\infty} H_k(\epsilon) \rho(\epsilon) d\epsilon \\
 &= 1
 \end{aligned} \tag{3.71}$$

Además, el primer momento de la distribución es d_1 ,

$$\begin{aligned}
 \int_{-\infty}^{\infty} \epsilon f(\epsilon) d\epsilon &= \int_{-\infty}^{\infty} \epsilon \rho(\epsilon) d\epsilon + d_1 \int_{-\infty}^{\infty} \epsilon^2 \rho(\epsilon) d\epsilon \\
 &\quad + \sum_{k=2}^K d_k \int_{-\infty}^{\infty} \epsilon H_k(\epsilon) \rho(\epsilon) d\epsilon \\
 &= d_1
 \end{aligned} \tag{3.72}$$

por (3.66) y puesto que $H_1(\epsilon) = \epsilon$. El segundo momento es igual a $1 + 2d_2$,

$$\begin{aligned}
 \int_{-\infty}^{\infty} \epsilon^2 f(\epsilon) d\epsilon &= \int_{-\infty}^{\infty} \epsilon^2 \rho(\epsilon) d\epsilon + d_1 \int_{-\infty}^{\infty} \epsilon^3 \rho(\epsilon) d\epsilon \\
 &\quad + d_2 \left[\int_{-\infty}^{\infty} (\epsilon^4 - \epsilon^2) \rho(\epsilon) d\epsilon \right] \\
 &\quad + \sum_{k=3}^K d_k \int_{-\infty}^{\infty} \epsilon^2 H_k(\epsilon) \rho(\epsilon) d\epsilon \\
 &= 1 + (3 - 1)d_2 + \sum_{k=3}^K d_k \int_{-\infty}^{\infty} (\epsilon^2 - 1) H_k(\epsilon) \rho(\epsilon) d\epsilon \\
 &\quad + \sum_{k=3}^K d_k \int_{-\infty}^{\infty} H_k(\epsilon) \rho(\epsilon) d\epsilon \\
 &= 1 + 2d_2
 \end{aligned} \tag{3.73}$$

puesto que $H_2(\epsilon) = \epsilon^2 - 1$. Finalmente, la función de distribución responde a (3.2).

$$\begin{aligned}
\int_{-\infty}^a f(\epsilon) d\epsilon &= \int_{-\infty}^a \rho(\epsilon) d\epsilon + \sum_{k=1}^K d_k \int_{-\infty}^a H_k(\epsilon) \rho(\epsilon) d\epsilon \\
&= \varrho(a) + \sum_{k=1}^K (-1)^k d_k \int_{-\infty}^a D^k \rho(\epsilon) d\epsilon \\
&= \varrho(a) + \sum_{k=1}^K (-1)^k d_k \left[D^{k-1} \rho(\epsilon) \right]_{-\infty}^a \\
&= \varrho(a) - \rho(a) \sum_{k=1}^K d_k H_{k-1}(a)
\end{aligned} \tag{3.74}$$

siendo $\varrho(a)$ la función de distribución de la densidad normal estándar, y teniendo en cuenta que $D^{k-1} \rho(a) = -\rho(a) H_k(a)$ si k es par y su opuesto si es impar (ver Kendall and Stuart (1969) para más propiedades).

Si consideramos la expresión de la función de densidad descrita en (3.4), el valor $1/v$ puede ser determinado explícitamente. Puesto que la función de densidad integra a 1, y aplicando de nuevo la propiedad en (3.66) se tiene que

$$\begin{aligned}
1/v &= \int_{-\infty}^{\infty} \rho(\epsilon) \left[1 + \sum_{k=1}^K d_k H_k(\epsilon) \right]^2 d\epsilon \\
&= \int_{-\infty}^{\infty} \rho(\epsilon) \left[1 + \left(\sum_{k=1}^K d_k H_k(\epsilon) \right)^2 + 2 \sum_{k=1}^K d_k H_k(\epsilon) \right] d\epsilon \\
&= 1 + \int_{-\infty}^{\infty} \rho(\epsilon) \left(\sum_{k=1}^K d_k H_k(\epsilon) \right)^2 d\epsilon \\
&= 1 + \sum_{s=1}^K \sum_{r=1}^K \int_{-\infty}^{\infty} \rho(\epsilon) d_s d_r H_s(\epsilon) H_r(\epsilon) d\epsilon \\
&= 1 + \sum_{k=1}^K d_k^2 k!
\end{aligned} \tag{3.75}$$

Desarrollando explícitamente el término al cuadrado de la función de densidad expresada en (3.4) y reorganizando el resultado se tiene que

$$v \rho(\epsilon) \times \left[1 + \sum_{k=1}^K d_k H_k(\epsilon) \right]^2 = \rho(\epsilon) \times \left[\sum_{k=0}^{2K} \Psi_k H_k(\epsilon) \right] \tag{3.76}$$

donde los $\Psi'_k, k = 0, \dots, 2K$ son función de los parámetros $d_k, k = 1, \dots, K$. Puesto que la nueva expresión también integra a 1 se tiene que

$$\begin{aligned}
\int_{-\infty}^{\infty} \rho(\epsilon) \times \left[\sum_{k=0}^{2K} \Psi_k H_k(\epsilon) \right] d\epsilon &= \int_{-\infty}^{\infty} \rho(\epsilon) \times \Psi_0 H_0(\epsilon) d\epsilon \\
&+ \int_{-\infty}^{\infty} \rho(\epsilon) \times \left[\sum_{k=1}^{2K} \Psi_k H_k(\epsilon) \right] d\epsilon \\
&= 1
\end{aligned} \tag{3.77}$$

y dado que por la propiedad de ortogonalidad el segundo término de la expresión desaparece, Ψ_0 ha de ser igual a 1 (ver Kendall and Stuart, 1969; Mauleón, 2010, 2003, 2006 para más detalles sobre éstas y otras propiedades).

Capítulo 4

Modelo de cambios de régimen para la predicción de los precios de la electricidad

El handicap del modelo de función de transferencia con ruido ARIMA-ARCH descrito en el capítulo anterior reside en la imposibilidad de modelar el comportamiento no lineal que exhiben las series de precios eléctricos. En este capítulo se introduce una representación equivalente en espacio de estados, y a partir de ésta se plantea un modelo en el que las matrices que lo definen cambian con el tiempo en función de un conjunto reducido de regímenes no observados que están determinados por las variables fundamentales que afectan al precio. De este modo, es posible modelar distintas respuestas del precio a diferentes situaciones en función de las variables que condicionan el comportamiento del mercado. El modelo se ha aplicado a las series diarias y horarias de los precios eléctricos en el mercado Ibérico, comparándose la calidad de sus estimaciones probabilísticas con la proporcionada con los modelos lineales.

4.1. Introducción

El modelo de función de transferencia con ruido ARIMA-ARCH estacional propuesto en el capítulo 3 constituye una buena aproximación para representar el comportamiento de los precios en mercado eléctricos. Sin embargo, se trata de una aproximación lineal, y por tanto, con él no es posible capturar el comportamiento no lineal evidenciado tanto en la respuesta de los precios a variables explicativas como en las dinámicas autorregresivas de las propias series. Como ya se expuso en el capítulo 2, este comportamiento ha llevado a muchos investigadores a proponer métodos de predicción que sean capaces de capturarlo, como son las redes neuronales (ver por ejemplo Szkuta et al., 1999; Gao et al., 2000; Pino et al., 2008; Mandal et al., 2007; Pindoriya et al., 2008). El problema de estas técnicas es que no permiten incorporar el conocimiento previo del analista, por lo que se obtienen modelos de *caja negra*¹ en los cuales el riesgo de sobreentrenamiento, y por tanto de pérdida de generalidad, es alto (ver sección 2.2.3).

Una alternativa para modelar el comportamiento no lineal de los precios eléctricos son los modelos periódicos, por umbrales o en general de cambio de régimen. Estos modelos generalizan a los lineales permitiendo que el juego de parámetros que los definen cambie con el tiempo en función del régimen del sistema, que puede ser observado o no (ver secciones 2.2.2.6, 2.2.2.7 y 2.2.2.8 donde se describen estos modelos). Por tanto, constituyen una alternativa para modelar comportamientos no lineales sin necesidad de cambiar la estructura básica de los modelos lineales. Existen referencias en la literatura de precios eléctricos donde ya han sido aplicados (ver por ejemplo Swider and Weber, 2007; Jong, 2006b; Weron, 2006, 2009; Karakatsani and Bunn, 2008b; Janczura and Weron, 2009; Mount et al., 2006; Huisman, 2008). Sin embargo, el objetivo en estas publicaciones no es tanto la predicción precisa de los precios eléctricos sino representar fielmente sus principales características. En la mayoría de estos casos, se trata de modelos con dos regímenes ocultos (o no observados) que describen de manera separada el proceso estocástico de emisión de precios normales del que emite altos precios o *spikes*. Habitualmente se asume también que el régimen del sistema sigue un proceso de Markov de primer orden donde su valor depende del valor del régimen en el instante anterior y, en ocasiones, de variables exógenas como la temperatura, la demanda de energía o el margen de reserva. Los modelos periódicos propuestos en Koopman et al., 2007; Carnero et al., 2010 son una alternativa válida para predecir el comportamiento cambiante de los precios eléctricos según los días de la semana y/u horas del día, pero con ellos no es posible condicionar la respuesta del precio a otras variables explicativas.

En este capítulo se presenta una formulación equivalente del modelo de función de transferencia con ruido ARIMA-ARCH en la representación conocida como de espacio de estados (ver Hamilton (1994)). A partir de dicha representación, el modelo se generaliza permitiendo que las matrices que lo definen cambien en el tiempo

¹Un modelo de *caja negra* es aquel que describe la relación entre las entradas y las salidas de un sistema sin conocerse su estructura interna.

según un conjunto reducido de regímenes no observados. En línea con algunos de los trabajos citados, el modelo propuesto asume que el régimen depende de las variables fundamentales que condicionan el comportamiento del mercado. A diferencia de los anteriores, la probabilidad de que cada uno de los regímenes ocurra no se modela a través de una matriz de probabilidades de transición, sino que se estima mediante una red neuronal. Este enfoque permite que sean los propios datos de ajuste los que determinen la forma de la distribución de probabilidad del régimen sin asumir ninguna previamente.

El enfoque no paramétrico adoptado para modelar la distribución de probabilidad del régimen permite modelar las complejas relaciones existentes entre las variables fundamentales del mercado y el precio resultante. La idea básica es permitir al modelo que aprenda las diferentes situaciones del mercado que determinan distintas respuestas o dinámicas del precio. Además, permite también aumentar la capacidad de ajuste del modelo incrementando el número de regímenes. Al ser la distribución de probabilidad del régimen variable en el tiempo, la función de densidad resultante es una mixtura de las densidades asumidas para cada régimen con diferentes pesos en cada instante de tiempo, lo que proporciona una capacidad de representación mayor que la del modelo descrito en el capítulo 3, en el que la forma de la distribución permanece constante en el tiempo.

La manera de modelar la distribución de probabilidad del régimen está inspirada en el *IOHMM* (*Input Output Hidden Markov Model*) presentado en Mateo et al. (2005). El modelo allí expuesto presenta una red neuronal que modela la distribución de probabilidad del régimen oculto tomando como entradas las variables explicativas que condicionan el comportamiento del mercado y las probabilidades de cada uno de los regímenes en el instante anterior (ver sección 2.2.3). En nuestro caso, se simplifica su estructura asumiendo que sólo depende de variables fundamentales del mercado y no de su valor en instantes anteriores, lo que supone un ahorro computacional importante. Sin embargo, lo que hace completamente diferentes ambos enfoques son los modelos expertos que describen la dinámica de los precios eléctricos condicional a cada uno de los regímenes. En el caso del modelo *IOHMM*, se trata también de redes neuronales feed-forward que toman como entradas variables explicativas que podrían ser o no las mismas que determinan cada uno de los regímenes. Esta estructura permite ajustar los parámetros de cada una de estas redes expertas de forma separada en un mismo paso del algoritmo EM². En nuestro caso, los modelos expertos que describen las dinámicas del precio condicionales a cada uno de los regímenes son modelos en espacio de estados que modelan tanto la media como la varianza del proceso. Como en el caso del *IOHMM*, es posible estimar de manera aislada los parámetros que determinan la red de probabilidad del régimen en cada paso del algoritmo EM. Sin embargo en nuestro caso, puesto que los modelos expertos presentan términos que dependen del error del modelo en instantes anteriores (términos de media móvil), no es posible estimar los parámetros de cada uno de los regímenes de manera separada. El filtro utilizado para estimar la media y varianza

²El algoritmo EM se explica en la sección 4.2.4.1

condicional a partir de un conjunto de parámetros fijados se describe en 4.2.4 y el procedimiento de estimación de dichos parámetros se describe en la sección 4.2.4.1.

En la siguiente sección se presenta la formulación de los modelos de espacio de estados adoptada en Hamilton (1994), y se detalla cómo puede escribirse el modelo de función de transferencia con ruido ARIMA-ARCH en dicha formulación. A partir de la misma se presenta una extensión que permite que las matrices que lo definen cambien con el tiempo de acuerdo con un conjunto de períodos o regímenes conocidos a priori. Finalmente, se describe el modelo de cambio de régimen oculto adoptado en esta tesis. Las secciones 4.3 y 4.4 muestran los resultados obtenidos del ajuste del modelo a los precios diarios y horarios de OMEL, y se comparan con los obtenidos usando sus contrapartidas lineales. La sección 4.5 recoge las conclusiones del capítulo.

4.2. Modelos en Espacio de Estados

El objetivo de esta sección es describir el modelo de serie temporal adoptado en esta tesis para tratar de representar el comportamiento no lineal que las series de precios eléctricos presentan. Su estructura básica deriva de la formulación en espacio de estados del modelo de función de transferencia con ruido ARIMA-ARCH estacional descrito en el capítulo 3. Básicamente, el modelo extiende dicha formulación permitiendo que sus parámetros cambien en el tiempo en función de un reducido conjunto de situaciones del mercado, las cuales a su vez están determinadas por las variables fundamentales que afectan al mismo.

Para entender correctamente la estructura del modelo definitivo, se describe en primer lugar la formulación conocida como de *espacio de estados*, en la cual pueden expresarse una gran cantidad de sistemas dinámicos. Dicha formulación junto con el filtro de Kalman constituyen herramientas muy importantes dentro del campo del análisis de series temporales. Entre otras muchas aplicaciones, proporcionan la manera de estimar los parámetros de un modelo de función de transferencia con ruido ARIMA gaussiano por el método de máxima verosimilitud exacta.

Una vez descrita esta forma de representar el comportamiento de un sistema dinámico, en la sección 4.2.2 se presenta una formulación equivalente en espacio de estados del modelo lineal de función de transferencia con ruido ARIMA-ARCH estacional descrito en el capítulo 3. A continuación en la sección 4.2.3, se describe una extensión del modelo anterior que permite que las matrices cambien en el tiempo según un conjunto de regímenes no ocultos o períodos establecidos a priori. Este tipo de modelo es capaz de capturar las diferentes dinámicas que presenta una determinada variable estocástica según determinados regímenes o períodos, y por tanto, es adecuado para recoger, por ejemplo, las variaciones de comportamiento intra-diario e intra-semanal que exhiben los precios eléctricos.

El modelo adoptado en esta tesis permite descubrir de manera automática el conjunto de regímenes no observables según los cuales la serie de precios eléctricos

responde de manera diferente. El modelo es explicado en detalle en la sección 4.2.4.

4.2.1. Representación en Espacio de Estados de un sistema dinámico

A continuación se expone una formulación en espacio de estados capaz de describir el comportamiento de una gran cantidad de sistemas dinámicos diferentes. Existen en la literatura otras representaciones en espacio de estados que presentan pequeñas diferencias con respecto a la que aquí se expone tanto en la estructura como en las hipótesis del modelo (ver por ejemplo de Jong and Penzer, 2004; Harvey, 1990; Durbin and Koopman, 2001; Aoki, 1986; Hyndman et al., 2008). Sin embargo, en todas ellas se expresa el sistema en términos de un vector de valores no observados conocido como *vector de estado*, utilizado para almacenar toda la información necesaria para estimar la salida en cada instante de tiempo. El modelo en espacio de estados propuesto en Hamilton (1994) describe el comportamiento de un vector observado y_t de dimensión $(m \times 1)$ a través del siguiente sistema de ecuaciones

$$\xi_{t+1} = F\xi_t + v_{t+1} \quad (4.1)$$

$$y_t = Ax_t + H\xi_t + w_t \quad (4.2)$$

donde F, A y H son matrices de parámetros de dimensión $(r \times r)$, $(m \times n)$ y $(m \times r)$, respectivamente, y x_t un vector de dimensión $(n \times 1)$ de variables exógenas predeterminadas. La ecuación 4.1 es conocida como *ecuación de estado*, el vector ξ_t *vector de estado* y 4.2 es la *ecuación de observación*. Los vectores v_t y w_t de dimensión $(r \times 1)$ y $(m \times 1)$, respectivamente, son vectores de ruido blanco, por lo que cumplen que

$$E(v_t v'_\tau) = \begin{cases} Q & \text{para } t = \tau \\ 0 & \text{en otro caso} \end{cases} \quad (4.3)$$

$$E(w_t w'_\tau) = \begin{cases} R & \text{para } t = \tau \\ 0 & \text{en otro caso} \end{cases} \quad (4.4)$$

dónde Q y R son matrices de covarianzas de dimensión $(r \times r)$ y $(m \times m)$, respectivamente. Se asume, además, que ambos ruidos están incorrelados

$$E(v_t w'_\tau) = 0 \text{ para todo } t \text{ y } \tau \quad (4.5)$$

Si se asume también que ξ_1 es incorrelado con cualquier realización de v_t y w_t , entonces se cumple que

$$E(v_t \xi'_\tau) = 0 \text{ para todo } \tau = t-1, t-2, \dots, 1 \quad (4.6)$$

$$E(w_t \xi'_\tau) = 0 \text{ para todo } \tau = 1, 2, \dots, T \quad (4.7)$$

$$E(w_t y'_\tau) = 0 \text{ para todo } \tau = t-1, t-2, \dots, 1 \quad (4.8)$$

$$E(v_t y'_\tau) = 0 \text{ para todo } \tau = t-1, t-2, \dots, 1 \quad (4.9)$$

A continuación se muestra de qué manera puede expresarse un modelo de función de transferencia con ruido ARIMA en la representación de espacio de estados aquí descrita. Otros modelos de series temporales como los de suavizado exponencial o componentes no observados también pueden describirse a través de este tipo de representación (ver Hyndman et al., 2008; Harvey, 1990 para detalles). Una vez un modelo queda representado de esta forma, es posible aplicar el filtro de Kalman y estimar sus parámetros por máxima verosimilitud exacta. Seguidamente se resumen las ecuaciones del filtro derivado de la representación aquí descrita.

4.2.1.1. Representación del modelo de función de transferencia con ruido ARIMA

El modelo de función de transferencia con ruido ARIMA descrito en (4.10)³

$$y_t = \sum_i^n G_i(L)x_{it} + \frac{1 + \theta_1 L + \dots + \theta_q L^q}{1 + \phi_1 L + \dots + \phi_p L^p} \varepsilon_t \quad (4.10)$$

con

$$G_i(L) = \frac{w_{i0} + \dots + w_{ig} L^g}{1 + \delta_{i1} L + \dots + \delta_{ig} L^g} \quad (4.11)$$

puede ser expresado en forma de espacio de estados de la manera siguiente⁴

$$\xi_{t+1} = F\xi_t + Bx_t + v_{t+1} \quad (4.12)$$

$$y_t = H\xi_t + Ax_t \quad (4.13)$$

³En 4.10 se ha simplificado la notación del modelo presentado en la sección 2.2.2.3. Aquí no se incluyen explícitamente la constante, los términos estacionales o las diferenciaciones aplicadas a la serie. La constante puede ser considerada una variable explicativa más cuya función de transferencia presenta un único coeficiente. Los términos estacionales y las diferenciaciones están implícitos en los términos autorregresivo y de media móvil. Tampoco la expresión de las funciones de transferencia es la misma, aunque es equivalente.

⁴En la representación expuesta en Hamilton (1994) la matriz B no aparece. Esta matriz captura la influencia de los términos de las funciones de transferencia de las variables exógenas sobre el estado del sistema.

donde ξ_t es el vector de estado, $x_t = (x_{1t}, \dots, x_{nt})'$ es el vector de variables explicativas y v_{t+1} es un vector de ruido blanco con matriz de covarianzas Q .

Las matrices A, B, F, H y Q dependen de los parámetros del modelo definido en (4.10) y se construyen a partir de las siguientes

$$F_0 = \begin{pmatrix} -\phi_1 & -\phi_2 & \cdots & -\phi_r \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$H_0 = (1 \quad \theta_1 \quad \cdots \quad \theta_r)$$

siendo $r = \max(q + 1, p)$. En el caso de que el modelo no incluya variables explicativas, es decir, que responda a la descripción de un ARIMA, $F = F_0, H = H_0, A = 0$ y $B = 0$. Para el caso general, si escribimos cada función de transferencia G_i en (4.11), $i = 1, \dots, n$, de la siguiente forma

$$G_i(L) = w_{i0} + \frac{b_{i1}L + b_{i2}L^2 + \cdots + b_{ig}L^g}{1 + \delta_{i1}L + \delta_{i2}L^2 + \cdots + \delta_{ig}L^g} \quad (4.14)$$

las matrices se construyen por cajas del siguiente modo

$$A = (w_{10} \quad w_{20} \quad \cdots \quad w_{n0})$$

$$B_i = \begin{pmatrix} b_{i1} \\ b_{i2} \\ \cdots \\ b_{ig} \end{pmatrix} \quad B = \begin{pmatrix} B_1 \\ B_2 \\ \cdots \\ B_n \end{pmatrix}$$

$$F_i = \begin{pmatrix} -\delta_{i1} & 1 & \cdots & 0 \\ -\delta_{i2} & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ -\delta_{ig} & 0 & 0 & 0 \end{pmatrix} \quad F = \begin{pmatrix} F_0 & 0 & \cdots & 0 \\ 0 & F_1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & F_n \end{pmatrix}$$

$$H_i = (1 \quad 0 \quad \cdots \quad 0) \quad H = (H_0 \quad H_1 \quad \cdots \quad H_n)$$

Por último, el ruido blanco ε_t se recoge en el vector v_t ,

$$v_{t+1} = \begin{pmatrix} \varepsilon_{t+1} \\ 0 \\ \cdots \\ 0 \end{pmatrix} \quad Q = \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Nótese que tan solo aparece una fuente de ruido v_t y, por tanto, $R = 0$.

4.2.1.2. Resumen del filtro de Kalman

A partir de la representación en espacio de estados (4.1)-(4.2), el filtro de Kalman proporciona en cada instante t la estimación del vector de estado que minimiza su error cuadrático. Los pasos del filtro se resumen a continuación.

Considerando que

$$\hat{\xi}_{t|t-1} = E(\xi_t | \eta_{t-1}) \quad (4.15)$$

$$P_{t|t-1} = E[(\xi_t - \hat{\xi}_{t|t-1})(\xi_t - \hat{\xi}_{t|t-1})'] \quad (4.16)$$

son la predicción del vector de estado y su matriz de error cuadrático medio asociada, respectivamente, el algoritmo comienza con una predicción de sus valores iniciales, esto es, la media y varianza no condicionales del vector de estado, $\hat{\xi}_{1|0}$ y $P_{1|0}$. Para el caso de que el proceso sea estacionario, la solución es única (ver Hamilton (1994)) y viene dada por (4.17)-(4.18)⁵.

$$\hat{\xi}_{1|0} = 0 \quad (4.17)$$

$$vec(P_{1|0}) = [I_{s^2} - (F \otimes F)]^{-1} \cdot vec(Q) \quad (4.18)$$

donde I_{s^2} es la matriz identidad de dimensión s^2 , siendo s el tamaño del vector de estado, $\otimes$ el producto de kronecker y $vec(P_{1|0})$ los elementos de $P_{1|0}$ dispuestos como una sola columna. Dados estos valores iniciales, el primer paso es predecir el valor de y_t a través de la ecuación de observación (4.2) del siguiente modo

$$\hat{y}_{t|t-1} = H\hat{\xi}_{t|t-1} + Ax_t \quad (4.19)$$

con error cuadrático medio

$$E[(y_t - \hat{y}_{t|t-1})(y_t - \hat{y}_{t|t-1})'] = HP_{t|t-1}H' + R \quad (4.20)$$

Una vez observado y_t , se actualiza la inferencia sobre el vector de estado y su error cuadrático medio

$$\hat{\xi}_{t|t} = \hat{\xi}_{t|t-1} + P_{t|t-1}H'(HP_{t|t-1}H' + R)^{-1}(y_t - Ax_t - H\hat{\xi}_{t|t-1}) \quad (4.21)$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}H'(HP_{t|t-1}H' + R)^{-1}HP_{t|t-1} \quad (4.22)$$

⁵Existen diversas maneras de inicializar el filtro para el caso de que no se cumplan las condiciones de estacionariedad (ver Harvey (1990) para más detalles y discusión).

Finalmente, se utiliza la ecuación de estado (4.1) para predecir el estado del sistema en el siguiente instante de tiempo $\xi_{t+1|t}$, así como su correspondiente error cuadrático medio asociado

$$\hat{\xi}_{t+1|t} = F\hat{\xi}_{t|t} \quad (4.23)$$

$$P_{t+1|t} = FP_{t|t}F' + Q; \quad (4.24)$$

y con estas estimaciones se vuelve a (4.19) hasta calcular toda la muestra de datos, $t = 1, \dots, T$. A menudo los pasos (4.21)-(4.24) son calculados en uno solo del siguiente modo

$$\hat{\xi}_{t+1|t} = F\hat{\xi}_{t|t-1} + K_t(y_t - Ax_t - H\hat{\xi}_{t|t-1}) \quad (4.25)$$

$$P_{t+1|t} = (F - K_tH)P_{t|t-1}(F' - H'K_t') + Q; \quad (4.26)$$

donde la matriz K_t es conocida como la *ganancia de Kalman* y responde a

$$K_t = FP_{t|t-1}H'(HP_{t|t-1}H' + R)^{-1} \quad (4.27)$$

Si ξ_1 y $\{v_t\}_{t=1}^T$ son gaussianos, la distribución condicional de y_t es también gaussiana, con media y varianza dadas por (4.19) y (4.20), respectivamente

$$y_{t|t-1} \sim N(Ax_t + H\hat{\xi}_{t|t-1}, HP_{t|t-1}H' + R) \quad (4.28)$$

Es decir, la función de densidad de y_t responde a (4.29)

$$\begin{aligned} f[y_t|\eta_{t-1}] &= (2\pi)^{-1/2}|HP_{t|t-1}H' + R|^{-1/2} \times \\ &\times \exp\left[\frac{-1}{2}(y_t - Ax_t - H\hat{\xi}_{t|t-1})'(HP_{t|t-1}H' + R)^{-1}(y_t - Ax_t - H\hat{\xi}_{t|t-1})\right] \\ &\text{para } t = 1, \dots, T \end{aligned} \quad (4.29)$$

la cual se utiliza para la estimación de máxima verosimilitud exacta.

Un resultado importante del filtro que afecta a su eficiencia computacional son sus propiedades de convergencia. Así, si el modelo de espacio de estados (4.1)-(4.2) describe un proceso estacionario, es decir, los autovalores de la matriz F están todos dentro del círculo unidad, la matriz $P_{t+1|t} \rightarrow P$ según $t \rightarrow \infty$. Es decir, la matriz $P_{t|t+1}$ alcanza un valor constante una vez aplicado el filtro con el suficiente número de muestras. A la aplicación del filtro una vez se ha alcanzado dicho estado se le conoce como *filtro de kalman en régimen permanente (steady-state Kalman filter)*.

Para más detalles y discusión sobre esta y otras aplicaciones del *filtro de Kalman* pueden consultarse entre otros Harvey, 1990; Hamilton, 1994.

4.2.2. Representación en espacio de estados del modelo de función de transferencia con ruido ARIMA-ARCH estacional

A continuación se repite aquí por conveniencia la formulación del modelo de función de transferencia con ruido ARIMA-ARCH estacional descrito en la sección 3.3.

$$y_t = \sum_i^n G_i(L)x_{it} + \frac{\theta_q(L)\Theta_{Q_1}(L^{S_1})\dots\Theta_{Q_m}(L^{S_m})}{\nabla^d \nabla_{S_1}^{D_1} \dots \nabla_{S_m}^{D_m} \phi_p(L)\Phi_{P_1}(L^{S_1})\dots\Phi_{P_m}(L^{S_m})} \varepsilon_t \quad (4.30)$$

con

$$\varepsilon_t = \sigma_t \epsilon_t \quad (4.31)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_w \varepsilon_{t-w}^2 \quad (4.32)$$

y

$$G_i(L) = \frac{(w_{i0} + w_{i1}L + w_{i2}L^2 + \dots + w_{ig_i}L^{g_i})L^{b_i}}{1 + \delta_{i1}L + \delta_{i2}L^2 + \dots + \delta_{ir_i}L^{r_i}} \quad (4.33)$$

$$\phi_p(L) = 1 + \phi_1 L + \dots + \phi_p L^p \quad (4.34)$$

$$\Phi_{P_s}(L^{S_s}) = 1 + \Phi_1 L^{S_s} + \dots + \Phi_P L^{P_s * S_s} \quad (4.35)$$

$$\theta_q(L) = 1 + \theta_1 L + \dots + \theta_q L^q \quad (4.36)$$

$$\Theta_{Q_s}(L^{S_s}) = 1 + \Theta_1 L^{S_s} + \dots + \Theta_Q L^{Q_s * S_s} \quad (4.37)$$

$$\nabla^d = (1 - L)^d \quad (4.38)$$

$$\nabla_{S_s}^{D_s} = (1 - L^{S_s})^{D_s} \quad (4.39)$$

donde y_t es la variable de salida, x_{it} son las variables explicativas y ϵ_t es un proceso de ruido blanco de varianza 1. ε_t es conocido como error de predicción. $\phi_p(L)$ y $\theta_q(L)$ son polinomios en L de órdenes p y q , que describen los términos autorregresivos y de media móvil, respectivamente. $\Phi_{P_s}(L^{S_s})$ y $\Theta_{Q_s}(L^{S_s})$ describen los términos autorregresivos y de media móvil estacionales de período S_s , con $s = 1, \dots, m$ y siendo m el número de estacionalidades consideradas (por ejemplo $S_1 = 24$, $S_2 = 168$ para series horarias con estacionalidad diaria y semanal). ∇^d y $\nabla_{S_s}^{D_s}$ son los operadores que describen describen las diferencias regular y estacionales aplicadas para inducir estacionariedad en procesos integrados. $G_i(L)$ representa una familia de funciones de transferencia lineales capaces de capturar un amplio número de diferentes respuestas impulsionales con pocos parámetros. σ_t^2 modela la varianza de la serie como función lineal de los w previos errores de predicción al cuadrado.

En la sección 4.2.1.1 se muestra cómo puede escribirse un modelo de función de transferencia con ruido ARIMA en la representación de espacio de estados propuesta en Hamilton (1994). Puesto que el término ARCH adicional puede ser escrito como un modelo AR del cuadrado de los errores de predicción ε_t del siguiente modo,

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_w \varepsilon_{t-w}^2 + v_t \quad (4.40)$$

siendo v_t

$$v_t = \varepsilon_t^2 - \sigma_t^2 = \sigma_t^2(\varepsilon_t^2 - 1) \quad (4.41)$$

un proceso de ruido blanco, aunque no independiente⁶, entonces el modelo completo (4.30)-(4.39) puede ser expresado de manera equivalente en una representación de espacio de estados anidada del siguiente modo,

$$\xi_{t+1}^* = F\xi_t^* + Bx_t + \varpi_{t+1} \quad (4.42)$$

$$y_t = H\xi_t^* + Ax_t \quad (4.43)$$

$$\xi_{t+1}^{**} = G\xi_t^{**} + D + \nu_{t+1} \quad (4.44)$$

$$\varepsilon_t^2 = J\xi_t^{**} + \alpha_0 \quad (4.45)$$

donde ξ_t^* y ξ_t^{**} son los vectores de estado, $x_t = (x_{1t}, \dots, x_{nt})'$ es el vector de variables explicativas, $\varpi_t = (\varepsilon_t, 0, \dots, 0)'$ y $\nu_t = (v_t, 0, \dots, 0)'$. Las matrices F, B, H y A dependen de los parámetros que modelan la media condicional de y_t y se construyen de la manera descrita en la sección 4.2.1.1. De igual modo partiendo de (4.40) se construyen las matrices G, D y J , las cuales dependen de los parámetros que modelan la varianza condicional de y_t .

Partiendo de estimaciones de los vectores de estado $\xi_{t|t-1}^*$ y $\xi_{t|t-1}^{**}$, la predicción del valor de y_t y σ_t^2 en t es

$$\hat{y}_{t|t-1} = H\hat{\xi}_{t|t-1}^* + Ax_t \quad (4.46)$$

$$\hat{\varepsilon}_{t|t-1}^2 = \hat{\sigma}_{t|t-1}^2 = J\hat{\xi}_{t|t-1}^{**} + \alpha_0 \quad (4.47)$$

y una vez observado y_t , los vectores de estado pueden actualizarse según (4.48) y (4.49), respectivamente.

$$\hat{\xi}_{t|t}^* = \hat{\xi}_{t|t-1}^* + (1, 0, \dots, 0)'(y_t - Ax_t - H\hat{\xi}_{t|t-1}^*) \quad (4.48)$$

$$\hat{\xi}_{t|t}^{**} = \hat{\xi}_{t|t-1}^{**} + (1, 0, \dots, 0)' \left[(y_t - Ax_t - H\hat{\xi}_{t|t-1}^*)^2 - (J\hat{\xi}_{t|t-1}^{**} + \alpha_0) \right] \quad (4.49)$$

⁶Puesto que ε_t^2 ha de ser no negativo, la variable v_t tiene que satisfacer que $v_t \geq -(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_w \varepsilon_{t-w}^2)$ (ver Peña (2005) para detalles).

Por último, la predicción de los vectores de estado para $t+1$ se realiza del siguiente modo,

$$\hat{\xi}_{t+1|t}^* = F\hat{\xi}_{t|t}^* + Bx_t \quad (4.50)$$

$$\hat{\xi}_{t+1|t}^{**} = G\hat{\xi}_{t|t}^{**} + D \quad (4.51)$$

y repitiendo este proceso para toda la muestra de datos, $t = 1, \dots, T$, se obtienen la media y varianza condicional estimadas del proceso estocástico de la variable y_t .

La manera de actualizar los vectores de estado en el filtro desarrollado en (4.46)-(4.51) contrasta con la manera de hacerlo en el filtro de kalman explicado en la sección 4.2.1.1. La diferencia radica en la incertidumbre asociada a la estimación de los vectores de estado debido a la falta de información en los instantes iniciales. Para el caso del filtro aplicado en (4.46)-(4.51), esta incertidumbre no es considerada y por tanto se trata de una estimación condicional a las estimaciones iniciales $\hat{\xi}_{1|0}^*$ y $\hat{\xi}_{1|0}^{**}$, lo que reduce drásticamente el coste computacional.

4.2.3. Modelo de espacio de estados periódico

El enfoque periódico, en el que los parámetros que definen el conjunto de matrices del modelo en (4.42)-(4.45) cambia su valor de acuerdo con un conjunto de regímenes o períodos que gobiernan el sistema, ofrece una buena oportunidad para añadir complejidad a este tipo de modelos manteniendo su estructura básica. Así, el modelo de espacio de estados periódico utilizado en esta tesis se formula del siguiente modo

$$\xi_{t+1}^* = F_{r_{t+1}}\xi_t^* + B_{r_{t+1}}x_t + \varpi_{t+1} \quad (4.52)$$

$$y_t = H_{r_t}\xi_t^* + A_{r_t}x_t \quad (4.53)$$

$$\xi_{t+1}^{**} = G_{r_{t+1}}\xi_t^{**} + D_{r_{t+1}} + \nu_{t+1} \quad (4.54)$$

$$\varepsilon_t^2 = J_{r_t}\xi_t^{**} + \alpha_{0,r_t} \quad (4.55)$$

donde las anteriores definiciones de $\xi_t^*, \xi_t^{**}, \varpi_{t+1}$ y ν_{t+1} se mantienen y $r_t \in \{1, \dots, R\}$ con $t = 1, \dots, T$ denota el conjunto de regímenes o períodos y es conocido a priori. Como solución al elevado número de parámetros que implicaría tomar todos los elementos de las matrices en (4.52)-(4.55) como parámetros libres, cada conjunto de matrices $F_{r_t}, B_{r_t}, H_{r_t}, A_{r_t}, G_{r_t}, D_{r_t}, J_{r_t}$ y α_{0,r_t} correspondiente a cada uno de los regímenes puede especificarse a partir de los parámetros de un modelo de función de transferencia con ruido ARIMA-ARCH del mismo modo que en el caso lineal. Nótese sin embargo, que el modelo formulado en (4.52)-(4.55) con las matrices así especificadas, en general, no es equivalente a un modelo de función

de transferencia con ruido ARIMA-ARCH periódico⁷, es decir, un modelo como el descrito en (4.30)-(4.39) pero en el que el conjunto de parámetros a aplicar en cada instante t depende ahora del régimen o período r_t . La diferencia entre ambas formulaciones radica en la salida del modelo tras una transición de un determinado régimen a otro. Lógicamente, en un estado de régimen permanente (constante) ambas formulaciones tienden a reproducir el mismo sistema. La ventaja de utilizar el modelo formulado en espacio de estados frente al de función de transferencia con ruido ARIMA-ARCH periódico es de índole computacional y radica fundamentalmente en las facilidades que la representación en espacio de estados proporciona para calcular las estimaciones del modelo a cualquier horizonte de tiempo considerado. En Franses and Paap (2004) se expone una representación vectorial de algunos modelos de series temporales periódicos que puede resultar útil para su identificación. Recientemente en Hindrayanto et al. (2010) se presenta una formulación en espacio de estados (basada en la representación expuesta en De Jong (1991)) de un modelo ARIMA estacional y periódico (*Periodic SARIMA*).

Partiendo de estimaciones de los vectores de estado $\xi_{t|t-1}^*$ y $\xi_{t|t-1}^{**}$, la predicción del valor de y_t y σ_t^2 en t es

$$\hat{y}_{t|t-1} = H_{r_t} \hat{\xi}_{t|t-1}^* + A_{r_t} x_t \quad (4.56)$$

$$\hat{\varepsilon}_{t|t-1}^2 = \hat{\sigma}_{t|t-1}^2 = J_{r_t} \hat{\xi}_{t|t-1}^{**} + \alpha_{0,r_t} \quad (4.57)$$

Una vez observado y_t , los vectores de estado pueden actualizarse según (4.58) y (4.59), respectivamente.

$$\hat{\xi}_{t|t}^* = \hat{\xi}_{t|t-1}^* + (1, 0, \dots, 0)' (y_t - A_{r_t} x_t - H_{r_t} \hat{\xi}_{t|t-1}^*) \quad (4.58)$$

$$\begin{aligned} \hat{\xi}_{t|t}^{**} &= \hat{\xi}_{t|t-1}^{**} + (1, 0, \dots, 0)' \times \\ &\times \left[(y_t - A_{r_t} x_t - H_{r_t} \hat{\xi}_{t|t-1}^*)^2 - (J_{r_t} \hat{\xi}_{t|t-1}^{**} + \alpha_{0,r_t}) \right] \end{aligned} \quad (4.59)$$

Finalmente, la predicción de los vectores de estado $\hat{\xi}_{t+1|t}^*$, $\hat{\xi}_{t+1|t}^{**}$ en $t+1$ se realiza del siguiente modo,

$$\hat{\xi}_{t+1|t}^* = F_{r_{t+1}} \hat{\xi}_{t|t}^* + B_{r_{t+1}} x_t \quad (4.60)$$

$$\hat{\xi}_{t+1|t}^{**} = G_{r_{t+1}} \hat{\xi}_{t|t}^{**} + D_{r_{t+1}} \quad (4.61)$$

⁷Para algunos casos como los modelos autorregresivos o de media móvil periódicos puros la representación en espacio de estados expuesta sí es equivalente.

4.2.4. Modelo de espacio de estados con cambio de régimen

La expresión general del modelo de espacio de estados con cambio de régimen adoptado en esta tesis presenta la misma formulación que la de los modelos de espacio de estados periódicos, (4.62)-(4.65)

$$\xi_{t+1}^* = F_{r_{t+1}}\xi_t^* + B_{r_{t+1}}x_t + \varpi_{t+1} \quad (4.62)$$

$$y_t = H_{r_t}\xi_t^* + A_{r_t}x_t \quad (4.63)$$

$$\xi_{t+1}^{**} = G_{r_{t+1}}\xi_t^{**} + D_{r_{t+1}} + \nu_{t+1} \quad (4.64)$$

$$\varepsilon_t^2 = J_{r_t}\xi_t^{**} + \alpha_{0,r_t} \quad (4.65)$$

dónde las definiciones realizadas para los modelos de espacio de estados periódicos se mantienen, con la diferencia de que en este caso el régimen $r_t \in \{1, \dots, R\}$ con $t = 1, 2, \dots, T$ no es una variable observable. Este hecho supone una importante diferencia a la hora de estimar la salida del sistema en cada instante t . Así, en este caso la función de densidad de probabilidad de y_t se calcula como la suma de la función de densidad conjunta de y_t y todos los posibles valores de r_t de acuerdo con (4.66)

$$f[y_t|\eta_{t-1}; \varphi] = \sum_{j=1}^R f[y_t, r_t = j|\eta_{t-1}; \varphi] = \sum_{j=1}^R f[y_t|r_t = j; \eta_{t-1}; \varphi]Pr(r_t = j|\eta_{t-1}; \varphi) \quad (4.66)$$

donde φ denota el vector de parámetros, η_{t-1} se refiere a la información disponible hasta $t-1$ ⁸, $Pr(r_t = j|\eta_{t-1}; \varphi)$ la probabilidad de que el régimen r_t tome el valor j en el instante t y $f[y_t|r_t = j; \eta_{t-1}; \varphi]$ la función de densidad condicional al régimen j , la cual, si se asume gaussiana, responde a (4.67)

$$f[y_t|r_t = j; \eta_{t-1}; \varphi] = \frac{1}{\sqrt{2\pi\hat{\sigma}_{t,j}^2}} \exp \left[\frac{-(y_t - \hat{y}_{t|j})^2}{2\hat{\sigma}_{t,j}^2} \right] \quad (4.67)$$

siendo $\hat{y}_{t|j}$ y $\hat{\sigma}_{t,j}^2$ la media y varianza estimadas condicional a que el régimen r_t tome el valor j . Por tanto, se necesita una hipótesis adicional del comportamiento estocástico de r_t para poder derivar la función de densidad de la variable de salida y_t , y a la postre, su función de verosimilitud. Como ya se expuso en la sección 2.2.2.8, existen diversas maneras de especificar el comportamiento de r_t que han sido aplicadas en los modelos encontrados en la literatura relativa al modelado de precios eléctricos. Las diferencias entre ellas residen en la manera de representar la evolución de la probabilidad de ocurrencia de cada régimen $\{1, \dots, R\}$, especificando si su valor depende o no de su propio pasado y/o cualquier otra variable explicativa.

⁸Nótese que las variables exógenas se consideran conocidas en cualquier instante de tiempo

El modelo aquí propuesto permite descubrir de manera automática la distribución de probabilidad de r_t a partir de variables explicativas que afectan a los precios eléctricos. Asume, por tanto, que el comportamiento de las series de precios está determinado por un reducido conjunto de regímenes no observables que a su vez están determinados por variables fundamentales. El régimen del sistema, es por tanto, una variable discreta y ficticia que nos permite modelar diferentes respuestas del precio en diferentes situaciones del mercado. Dadas las complejas relaciones que existen entre los precios y las variables por las que se ve afectado, el régimen se modela a través de una red neuronal, en particular un perceptrón multicapa, que incorpora una capa adicional que hace posible que la red responda a una función de distribución de probabilidad sin imponer restricciones a los pesos de la red. Así, la probabilidad de que el régimen r_t tome el valor j responde a (4.68)-(4.70),

$$Pr(r_t = j | z_{1t}, \dots, z_{mt}) = \lambda_j^2 / (1 + \sum_{r=1}^{R-1} \lambda_r^2) \quad \forall j \in \{1, \dots, R-1\} \quad (4.68)$$

$$Pr(r_t = R | z_{1t}, \dots, z_{mt}) = 1 / (1 + \sum_{r=1}^{R-1} \lambda_r^2) \quad (4.69)$$

siendo

$$\lambda_j = \sum_{n=1}^N \frac{\omega_{nj2}}{1 - \exp(\sum_h \omega_{hn1} z_{ht} + \omega_{0n1})} \quad (4.70)$$

donde $(z_{1t}, \dots, z_{mt})'$ es un vector de variables explicativas en el instante t , N es el número de neuronas de la primera capa y ω_{hn1} , ω_{nr2} son los pesos o parámetros correspondientes a la primera y segunda capa, respectivamente. Nótese que las ecuaciones (4.68)-(4.70) permiten cumplir con las necesarias condiciones de $0 \leq Pr(r_t = j) \leq 1 \quad \forall j \in \{1, \dots, R\}$ y $\sum_{j=1}^R Pr(r_t = j) = 1$ sin imponer restricciones a los pesos de la red.

La distribución de probabilidad de la variable r_t responde, por tanto, a la forma general de un perceptrón multicapa, que tiene la propiedad de ser aproximador universal (Park and Sandberg (1991)). Esta arquitectura presenta dos ventajas frente a otros enfoques no paramétricos. La primera de ellas es su robustez frente a la inicialización de los pesos previa a su estimación, lo cual resulta importante puesto que el régimen r_t no es observable, y por tanto, es difícil conjeturar unos valores iniciales que sean apropiados. La otra ventaja es que el perceptrón multicapa es menos sensible a variables no significativas que otros aproximadores como las redes de función de base radial (*Radial Based Function Networks (RBFN)*) o, en general, estimadores kernel (Silverman (1986)).

Nótese que la asunción de la densidad condicional al régimen que aparece en (4.67) es consecuencia de que las innovaciones del modelo se distribuyen según la densidad

normal estándar, $\epsilon_t \sim N(0, 1)$, la cual a priori puede aparecer una distribución menos flexible que la familia de densidades no paramétrica asumida en el capítulo 3. Sin embargo, dado que la probabilidad de ocurrencia de cada régimen $\{1, \dots, R\}$ es variable en el tiempo, la función de densidad resultante de los precios eléctricos resulta en una mixtura de densidades normales que presenta diferentes pesos en cada instante t , como está definida en (4.66). En este aspecto, el modelo de cambio de régimen aquí propuesto proporciona una capacidad de representación mayor que la del modelo expuesto en el capítulo 3, en el que la forma de la densidad de los precios eléctricos, es decir, los momentos de la distribución mayores al segundo, permanecen constantes en el tiempo.

Los modelos de cambio de régimen que aparecen en el artículo de Hamilton (1990), germen del análisis de series temporales a través de este tipo de modelos, presentan una estructura básica que responde a la de los procesos autorregresivos, por lo que la función de densidad en (4.66) depende de un número pequeño de valores pasados de la variable de salida y_t y del régimen r_t . Como se señala en Hamilton (2008), surge un problema añadido para el caso de los modelos cuya estructura recursiva hace que la función de densidad de la variable de salida y_t (4.66) depende de la historia completa del régimen $r_t, r_{t-1}, \dots, r_1$, como es el caso de los modelos tipo ARMA, ARCH o los de espacio de estados aquí descritos. Sirva como ejemplo para ilustrar esta situación el modelo *ARMA* descrito en (4.71)

$$y_t = c_{r_t} + \phi_{r_t} y_{t-1} + \theta_{r_t} \epsilon_{t-1} + \epsilon_t \quad (4.71)$$

donde r_t , que toma los valores $\{0, 1\}$, determina el régimen del sistema y por tanto el juego de parámetros $\{c_{r_t}, \phi_{r_t}, \theta_{r_t}\}$ a aplicar en cada instante t . Nótese que en el caso de que θ_0 y θ_1 sean 0, el modelo en (4.71) presenta una estructura autorregresiva pura, y el valor de y_t depende de y_{t-1} y del valor que tome el régimen r_t en el instante t puesto que determina los parámetros c_{r_t}, ϕ_{r_t} . Sin embargo, en caso de que $\theta_{r_t}, r = 0, 1$ no sean nulos, y_t se calcula recursivamente, y por tanto depende del valor actual y todos los valores pasados del régimen $r_t, r_{t-1}, \dots, r_1$ así como de toda la historia $y_{t-1}, y_{t-2}, \dots, y_1$. Para este tipo de modelos, calcular la función de verosimilitud resulta inabordable según crece el tamaño de la muestra, por lo que la función de densidad ha de ser aproximada. A continuación se describe una manera de estimar los valores de $\hat{y}_{t|j}$ y $\hat{\sigma}_{t|j}^2$, los cuales permiten aproximar la función de densidad en (4.66). El algoritmo está basado en el filtro descrito en Kim, 1994; Kim and Nelson, 1999, donde se generalizan los modelos de Markov con cambio de régimen propuestos en Hamilton (1990) a los de espacio de estados.

Partiendo de estimaciones de los vectores de estado $\xi_{t|t-1}^{*(j)}$ y $\xi_{t|t-1}^{**(j)}$, las cuales son condicionales a que el régimen en el instante t sea el j , $r_t = j$, con $j = 1, \dots, R$, la predicción del valor de y_t y σ_t^2 condicional a $r_t = j$ se calcula del siguiente modo

$$\hat{y}_{t|t-1,j} = H_j \hat{\xi}_{t|t-1}^{*(j)} + A_j x_t \quad (4.72)$$

$$\hat{\varepsilon}_{t|t-1,j}^2 = \hat{\sigma}_{t|t-1,j}^2 = J_j \hat{\xi}_{t|t-1}^{***(j)} + \alpha_{0,j} \quad (4.73)$$

Una vez observado y_t , los vectores de estado condicionales a cada régimen j se actualizan según (4.74) y (4.75), respectivamente.

$$\hat{\xi}_{t|t}^{*(j)} = \hat{\xi}_{t|t-1}^{*(j)} + (1, 0, \dots, 0)' (y_t - A_j x_t - H_j \hat{\xi}_{t|t-1}^{*(j)}) \quad (4.74)$$

$$\begin{aligned} \hat{\xi}_{t|t}^{***(j)} &= \hat{\xi}_{t|t-1}^{***(j)} + (1, 0, \dots, 0)' \times \\ &\times \left[(y_t - A_j x_t - H_j \hat{\xi}_{t|t-1}^{*(j)})^2 - (J_j \hat{\xi}_{t|t-1}^{***(j)} + \alpha_{0,j}) \right] \end{aligned} \quad (4.75)$$

El objetivo ahora es realizar una inferencia sobre los vectores de estado en $t + 1$ basada en la información disponible en t , teniendo en cuenta los posible valores de r_t y r_{t+1} . Así, el algoritmo ahora calcula una batería de R^2 predicciones de los vectores de estado correspondientes a cada posible pareja $r_t = j, r_{t+1} = k$, del siguiente modo

$$\hat{\xi}_{t+1|t}^{*(j,k)} = F_k \hat{\xi}_{t|t}^{*(j)} + B_k x_t \quad (4.76)$$

$$\hat{\xi}_{t+1|t}^{***(j,k)} = G_k \hat{\xi}_{t|t}^{***(j)} + D_k \quad (4.77)$$

Cada iteración de este algoritmo produciría un incremento de R veces el número de casos a considerar (incluso con solo 2 regímenes habría más de 1000 casos para una muestra de tamaño 10), lo que demuestra la necesidad de utilizar una aproximación para hacer el filtro operable. La idea básica es reducir los $R \times R$ vectores de estado generados en este punto a tan solo R . La aproximación se realiza usando el término de probabilidad del régimen, el cual puede actualizarse de acuerdo con (4.78) una vez que y_t ya ha sido observado

$$\begin{aligned} Pr(r_t = j | \eta_t) &= Pr(r_t = j | \eta_{t-1}; y_t) \\ &= \frac{f[y_t, r_t = j | \eta_{t-1}]}{f[y_t | \eta_{t-1}]} \\ &= \frac{f[y_t | r_t = j; \eta_{t-1}] Pr(r_t = j | z_{1t}, \dots, z_{mt})}{\sum_{j=1}^R f[y_t | r_t = j; \eta_{t-1}] Pr(r_t = j | z_{1t}, \dots, z_{mt})} \end{aligned} \quad (4.78)$$

Para finalizar la iteración, la aproximación de los vectores de estado se realiza entonces de acuerdo con (4.79)-(4.80)

$$\widehat{\xi}_{t+1|t}^{*(k)} = \sum_{j=1}^R Pr(r_t = j|\eta_t) \times \widehat{\xi}_{t+1|t}^{*(j,k)} \quad (4.79)$$

$$\widehat{\xi}_{t+1|t}^{**(k)} = \sum_{j=1}^R Pr(r_t = j|\eta_t) \times \widehat{\xi}_{t+1|t}^{**(j,k)} \quad (4.80)$$

para todos los posibles valores del régimen, $k = 1, \dots, R$. Repitiendo este proceso desde (4.72)-(4.73) para toda la muestra de datos, $t = 1, \dots, T$, se obtienen la media y varianza estimadas condicional a cada uno de los posible regímenes, necesarias para el cálculo de la función de densidad del proceso estocástico de y_t definida en 4.66.

4.2.4.1. Estimación de los parámetros del modelo

Del mismo modo que para el caso del modelo de serie temporal descrito en el capítulo 3, la estimación de los parámetros de los modelos en espacio de estados aquí propuestos, tanto el modelo periódico como el de cambio de régimen, puede llevarse a cabo maximizando la función de verosimilitud a partir de una muestra de datos disponible, de la siguiente manera

$$\max_{\{\varphi\}} L(\varphi) = \sum_{t=1}^T \ln(f[y_t|\eta_{t-1}; \varphi]) \quad (4.81)$$

donde φ se refiere al vector de parámetros y η_{t-1} a la información disponible hasta el instante $t-1$ (η_{t-1} se omite de ahora en adelante para simplificar la notación). Los datos a partir de los cuales (4.81) se maximiza son la serie temporal de la variable de salida $Y = \{y_t\}_{t=1}^T$ y las variables explicativas $X_i = \{x_{it}\}_{t=1}^T, i = 1, \dots, n$, $Z_h = \{z_{ht}\}_{t=1}^T, h = 1, \dots, m$ (tanto las que afectan directamente a la salida del modelo como las que determinan la distribución de probabilidad del régimen para el caso de los modelos con cambio de régimen). Además, en este tipo de modelos interviene otra serie temporal que determina el régimen del sistema en cada instante de tiempo t , $S = \{r_t\}_{t=1}^T$ con $r_t \in \{1, \dots, R\}$, la cual puede ser observable o no. En el caso de que el régimen r_t sea observable, el modelo es periódico, y la estimación de los parámetros del modelo se lleva a cabo a través de los procedimientos explicados en la sección 3.3.1. En este caso, asumiendo normalidad se tiene que

$$f[y_t|r_t; \varphi] = \frac{1}{\sqrt{2\pi}\widehat{\sigma}_{t|t-1}} \exp \left[\frac{-(y_t - \widehat{y}_{t|t-1})^2}{2\widehat{\sigma}_{t|t-1}^2} \right] \quad (4.82)$$

donde $\widehat{y}_{t|t-1}$ y $\widehat{\sigma}_{t|t-1}$ se calculan según las 4.56 y 4.56, respectivamente. Sin embargo, si el régimen r_t , $t = 1, \dots, T$ no es una variable observable, la variable

de salida se distribuye entonces según (4.66), por lo que la función de verosimilitud pasa a tener la forma siguiente

$$L(\varphi) = \sum_{t=1}^T \ln \left(\sum_{j=1}^R f[y_t|r_t = j; \varphi] Pr(r_t = j; \varphi) \right) \quad (4.83)$$

la cual presenta una expresión más complicada que la de los modelos con cambio de régimen observable. Asumiendo normalidad, la función de densidad de la variable de salida condicional al régimen j , $f[y_t|r_t = j; \varphi]$ responde a (4.67), dónde los valores $\hat{y}_{t|t-1,j}$ y $\hat{\sigma}_{t|t-1,j}$ se calculan como en (4.72)-(4.73). En este caso, el vector de parámetros φ está constituido tanto por los parámetros que definen las matrices de los modelos de espacio de estados para cada uno de los regímenes como por los pesos de la red neuronal que modela la probabilidad de ocurrencia de cada régimen j , $Pr(r_t = j; \varphi)$, definida en (4.68)-(4.70). El gran tamaño de este vector de parámetros y la compleja forma de la función de verosimilitud (4.83), altamente no lineal, hacen que en la práctica resulte muy complicado que los procedimientos de estimación descritos en la sección 3.3.1 alcancen una solución aceptable.

Para resolver este problema, se ha utilizado un algoritmo alternativo para calcular estimaciones de máxima verosimilitud conocido como EM (*Expectation-Maximization*). Este algoritmo fue originalmente propuesto en Dempster et al. (1977), donde se presentó una gran cantidad de situaciones susceptibles de ser resueltas a través de su aplicación, como es el caso de ajustes en presencia de datos incompletos o variables no observadas. Fue en el trabajo de Hamilton (1990) donde se particularizó su aplicación al caso de la estimación de los parámetros de modelos de series temporales con cambio de régimen. A continuación, se resume el funcionamiento del algoritmo para la estimación de los parámetros del modelo presentado en esta tesis. Una excelente descripción del funcionamiento del mismo para el caso particular del modelo IOHMM (*Input Output Hidden Markov Models*, Bengio and Frasconi (1995)) puede encontrarse en Mateo (2005).

El Algoritmo EM es un procedimiento iterativo que consta de dos pasos en cada iteración k :

1. *Paso E*: Dado el vector de parámetros $\varphi^{(k-1)}$ obtenido en la iteración anterior ($k-1$) y los datos observados, se calcula la distribución de probabilidad de la variable r_t para toda la muestra de datos $t = 1, \dots, T$. Dicho de otra manera, dados los datos observados y el vector de parámetros en la iteración anterior, se calcula para cada uno de los posible regímenes $\{1, \dots, R\}$ la probabilidad de que fuera el responsable de la observación en cada instante de tiempo t .
2. *Paso M*: Se maximiza el valor esperado de la función de verosimilitud de y_t condicionada a la distribución de probabilidad de la variable r_t calculada en el *paso E*. De esta manera se obtiene el vector de parámetros estimados $\varphi^{(k)}$ en la iteración k .

El algoritmo asegura que tras cada iteración se mejora el valor de la función de verosimilitud (4.83), y de este modo, a partir de una estimación arbitraria de parámetros $\varphi^{(0)}$, el algoritmo converge a un máximo local (ver Dempster et al. (1977)). Puesto que la distribución de probabilidad del régimen no observable r_t es completamente desconocida, resulta muy complejo conjeturar un vector de parámetros inicial $\varphi^{(0)}$ que resulte estar próximo de un máximo aceptable, lo que sería necesario en caso de utilizar los procedimientos de estimación explicados en la sección 3.3.1. Sin embargo, como se señala en Hamilton (1990), el algoritmo EM resulta suficientemente robusto frente a una inicialización arbitraria de los parámetros.

El algoritmo se basa, por tanto, en sustituir la función de verosimilitud (4.83) por una función auxiliar que se maximiza en cada iteración, lo que permite converger a un máximo de (4.83) tras un número de iteraciones. Dicha función auxiliar responde a

$$Q(\varphi^{k-1}, \varphi^k) = E \left[L(\varphi) | S, \varphi^{k-1} \right] \quad (4.84)$$

Es decir, $Q(\varphi^{k-1}, \varphi^k)$ es la esperanza de la verosimilitud logarítmica condicional al conjunto de parámetros y la distribución de probabilidad de la variable no observada r_t obtenidos en la iteración anterior $k - 1$. Por consiguiente, en cada iteración k se maximiza $Q(\varphi^{k-1}, \varphi^k)$ con respecto al vector de parámetros φ^k . La ventaja de usar esta función en lugar de (4.83) es que permite desacoplar diferentes términos con el fin de poder estimar grupos de parámetros de manera separada. Así, se tiene que

$$\begin{aligned} Q(\varphi^{k-1}, \varphi^k) &= \sum_{t=1}^T \sum_{j=1}^R Pr(r_t = j | y_t; \varphi^{k-1}) \times [\ln(f[y_t, r_t = j; \varphi])] \\ &= \sum_{t=1}^T \sum_{j=1}^R Pr(r_t = j | y_t; \varphi^{k-1}) \times [\ln(f[y_t | r_t = j; \varphi] Pr(r_t = j; \varphi))] \\ &= \sum_{t=1}^T \sum_{j=1}^R Pr(r_t = j | y_t; \varphi^{k-1}) \times \ln(f[y_t | r_t = j; \varphi]) + \\ &\quad \sum_{t=1}^T \sum_{j=1}^R Pr(r_t = j | y_t; \varphi^{k-1}) \times \ln(Pr(r_t = j; \varphi)) \end{aligned} \quad (4.85)$$

y por tanto la función $Q(\varphi^{k-1}, \varphi^k)$ puede descomponerse en dos sumandos separados. Considerando que el vector de parámetros a estimar está constituido a su vez por el vector de los parámetros que determinan las matrices del modelo en espacio de estados, que denotaremos por φ_1 , y el vector de pesos de la red neuronal que modela la distribución de probabilidad del régimen r_t , φ_2 , es decir, $\varphi = (\varphi_1; \varphi_2)$, entonces se tiene que $Q(\varphi^{k-1}, \varphi^k)$ se escribe de la siguiente manera

$$Q(\varphi^{k-1}, \varphi^k) = Q_1(\varphi^{k-1}, \varphi_1^k) + Q_2(\varphi^{k-1}, \varphi_2^k) \quad (4.86)$$

donde

$$Q_1(\varphi^{k-1}, \varphi_1^k) = \sum_{t=1}^T \sum_{j=1}^R Pr(r_t = j | y_t; \varphi^{k-1}) \times \ln(f[y_t | r_t = j; \varphi_1^k]) \quad (4.87)$$

$$Q_2(\varphi^{k-1}, \varphi_2^k) = \sum_{t=1}^T \sum_{j=1}^R Pr(r_t = j | y_t; \varphi^{k-1}) \times \ln(Pr(r_t = j; \varphi_2^k)) \quad (4.88)$$

De este modo, el vector de parámetros φ puede ser estimado de manera separada mediante dos maximizaciones diferentes en cada iteración. En la primera maximización se estiman los parámetros en φ_1 , los cuales definen las matrices del modelo de espacio de estados, utilizando una rutina de optimización no lineal. Para asegurar la estabilidad del modelo resultante, puede ser necesario incorporar las mismas restricciones de estacionariedad e invertibilidad a las matrices que determinan cada uno de los regímenes que en el caso lineal. En la segunda maximización se estiman los pesos de la red neuronal que conforman el vector de parámetros φ_2 con respecto a $Q_2(\varphi^{k-1}, \varphi_2^k)$, lo cual puede llevarse a cabo utilizando un procedimiento de optimización no lineal sin restricciones como el descrito en la sección 3.3.1, o algún algoritmo específico apropiado para ajustar los pesos de una red neuronal (ver Haykin (1994)). Estas dos maximizaciones constituyen el *paso M* del algoritmo. Como puede apreciarse, la modularidad del algoritmo EM permitiría, por ejemplo, plantear un modelo diferente en el que la distribución de probabilidad del régimen se describiera a través de otro tipo de arquitectura diferente al perceptrón multicapa, sin más que sustituir el procedimiento de maximización correspondiente a la distribución de probabilidad del régimen.

El *paso E* no supone apenas coste computacional. Al principio de cada iteración k , a partir de una estimación de parámetros φ^{k-1} y para cada instante de tiempo t , se trata de hacer una inferencia sobre qué régimen es más probable haber sido el responsable de producir la observación y_t . De la definición de probabilidad condicional se tiene que

$$\begin{aligned} Pr(r_t = j | y_t; \varphi^{k-1}) &= \frac{f[y_t, r_t = j; \varphi^{k-1}]}{f[y_t; \varphi^{k-1}]} \\ &= \frac{f[y_t | r_t = j; \varphi^{k-1}] \times Pr(r_t = j; \varphi^{k-1})}{\sum_{j=1}^R f[y_t | r_t = j; \varphi^{k-1}] \times Pr(r_t = j; \varphi^{k-1})} \end{aligned} \quad (4.89)$$

cuyo resultado representa la probabilidad, dados los datos observados, de que el régimen no observado responsable de la observación y_t haya sido el j . Este

cálculo completa entonces todos los valores necesarios para maximizar (4.84) en cada iteración k .

4.2.4.2. Diagnósis y selección del modelo

En general, no existe una metodología consensuada que pueda ser aplicada para guiar el proceso de identificación de los modelos de cambio de régimen. Existen varias fuentes de incertidumbre a las que enfrentarse en el caso del modelo propuesto en esta tesis, como son la estructura básica del modelo (es decir, la manera en que son parametrizadas las matrices de espacio de estados), el número de regímenes ocultos existentes en la serie a modelar, las variables explicativas que afectan a la distribución de probabilidad de la variable que determina el régimen y el número de neuronas en la capa oculta de la red neuronal. Este último valor, el número de neuronas, no resulta crítico puesto que la arquitectura del perceptrón multicapa es suficientemente robusta frente a un exceso en el mismo. Sin embargo, el número de regímenes presentes en la serie sí resulta un valor relevante y difícil de decidir. Puesto que no se trata de una variable observable sino ficticia, el número de regímenes constituye a menudo una decisión del analista y su valor es previamente fijado. A este respecto, destaca el trabajo presentado en Vucetic et al. (2001), donde se propone un algoritmo capaz de descubrir de manera automática el número de regímenes ocultos en las series de precios eléctricos basándose en sus relaciones con la demanda. Sin embargo, el algoritmo no permite cambios de régimen frecuentes y tampoco tiene en cuenta las relaciones existentes entre el precio y el resto de sus variables fundamentales.

El proceso de identificación aquí propuesto comienza una vez la estructura de un modelo de función de transferencia con ruido ARIMA-ARCH estacional como el descrito en la sección 3.3 ha sido identificada. Como ya se expuso en dicha sección, algunas de las herramientas utilizadas en el proceso de identificación del modelo lineal son las funciones de autocorrelación simple y parcial (FAS y FAP) de los residuos y su cuadrado para seleccionar los órdenes de los polinomios AR, MA y ARCH, los errores estándar de los parámetros de los modelos ajustados para comprobar su significación y el criterio de información de Akaike (AIC) como índice para medir la parsimonia de los modelos estimados. A partir de dicha estructura, las matrices del modelo de cambio de régimen son especificadas del modo explicado en 4.2.2, es decir, se construyen de manera que el modelo de espacio de estados que describe cada uno de los regímenes por sí solo presenta una estructura equivalente al modelo de función de transferencia con ruido ARIMA-ARCH⁹. Nótese, sin embargo, que las matrices del modelo de cambio de régimen definitivo no tendría por qué responder exactamente a dicha especificación. Por ejemplo, un término autorregresivo podría no ser significativo en el modelo lineal, pero sí serlo una vez se han identificado distintos regímenes en la serie temporal, para todos o alguno de dichos regímenes.

⁹Nótese, como ya se expuso en 4.2.3, que la formulación del modelo de espacio de estados con cambio de régimen descrito en 4.2.4, en general, no es equivalente a la de un modelo de función de transferencia con ruido ARIMA-ARCH donde los parámetros cambian de acuerdo con el régimen del sistema.

Desafortunadamente, en este caso no se dispone de los errores estándar de los parámetros calculados a través de (3.45) puesto que el algoritmo EM estima los parámetros del modelo de forma separada en dos grupos diferentes, y calcular las matrices (3.42) y (3.43) una vez se ha alcanzado el vector de parámetros que maximiza la verosimilitud de la muestra de datos sería computacionalmente muy costoso debido a su tamaño. Resultará necesario, por tanto, una vez se hayan seleccionado el resto de valores por determinar (el número de regímenes y las variables explicativas de entrada a la red neuronal), probar algunas especificaciones de la estructura básica del modelo menos restrictivas, y decidir entre ellas usando el criterio de información de Akaike (AIC) o el criterio Bayesiano (BIC).

Dado que la variable no observable que determina el régimen en cada instante de tiempo r_t es una variable ficticia, resulta complicado conocer a priori el número de regímenes que mejor describe las dinámicas de la serie temporal a modelar. Sin embargo, una solución parsimoniosa del modelo implicará un número pequeño de regímenes. La estrategia para seleccionar dicho número y las variables de entrada a la red de probabilidad será entonces tomar todas las variables candidatas y el número máximo de regímenes a considerar, y estimar los parámetros del correspondiente modelo de cambio de régimen. Las variables que resulten realmente significativas para modelar la distribución de probabilidad del régimen pueden ser identificadas a través del análisis de la sensibilidad de la probabilidad de cada uno de los regímenes con respecto a cada una de las entradas a la red del modelo ajustado, es decir, calculando $\partial Pr(r_t = j)/\partial z_{it}$ por medio de procedimientos numéricos (ver White and Racine, 2000; Muñoz and Czernichow, 1998 para más detalles). Una vez se ha identificado un modelo con el número máximo de regímenes a considerar y donde todas las variables explicativas de la red neuronal son relevantes, se pasa a estimar un modelo que presente las entradas seleccionadas a la red y un régimen menos, para volver a identificar las variables que son relevantes en este caso. El procedimiento se repite hasta llegar a estimar los parámetros de un modelo con dos regímenes e identificar las entradas significativas a la red de probabilidad. Finalmente, de entre todos los modelos ajustados con distinto número de regímenes se elige utilizando los criterios AIC o BIC, e inspeccionando de manera visual el ajuste de la densidad asumida a la distribución empírica obtenida a través de gráficos de tipo q-q plot o histogramas. Puesto que la hipótesis del modelo es que la componente estocástica se distribuye normal estándar, $\epsilon_t \sim N(0, 1)$, estos gráficos contrastan si la serie de residuos $\{\hat{\sigma}_{t|t-1}^{-1}(y_t - \hat{y}_{t|t-1})\}_{t=1}^T$ se ajusta a dicha distribución. Dicha serie se calcula teniendo en cuenta que los momentos de primer y segundo orden de la mixtura de gaussianas en cada instante de tiempo t se calculan de la siguiente manera

$$E[y_t] = \hat{y}_{t|t-1} = \sum_{j=1}^R Pr(r_t = j | z_{1t}, \dots, z_{mt}) \times \hat{y}_{t|t-1,j} \quad (4.90)$$

$$E[(y_t - \hat{y}_{t|t-1})^2] = \sum_{j=1}^R Pr(r_t = j | z_{1t}, \dots, z_{mt}) \times \left(\hat{y}_{t|t-1,j}^2 + \hat{\sigma}_{t|t-1,j}^2 \right) - \hat{y}_{t|t-1}^2 \quad (4.91)$$

con $\hat{y}_{t|t-1,j}$ y $\hat{\sigma}_{t|t-1,j}$ calculados como en (4.72) y (4.73) (ver Schnatter (2006)). En las siguientes secciones se ilustra todo este análisis utilizando los precios medios diarios y los precios horarios del mercado de OMEL.

4.3. Aplicación a los precios medios diarios del mercado Ibérico

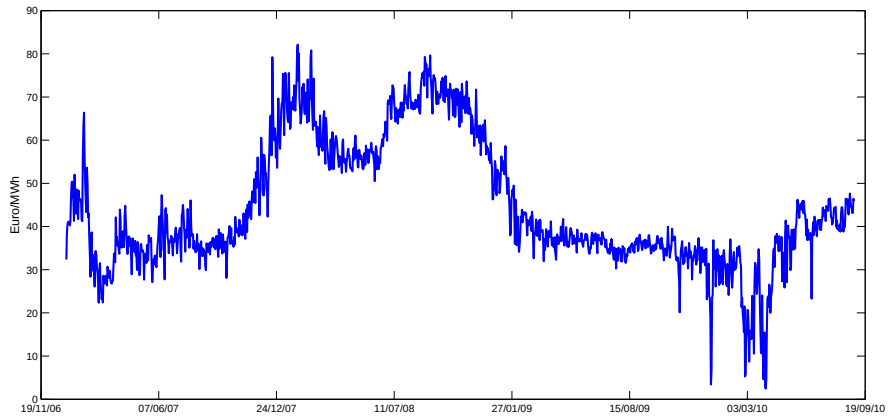
A continuación se presentan los resultados de la aplicación del modelo de cambio de régimen propuesto en este capítulo a los precios medios diarios del mercado ibérico de OMEL. Este estudio nos permitirá analizar el comportamiento de los precios sin considerar el efecto de las diferentes dinámicas intra-diarias que presentan. Más tarde, en la siguiente sección y teniendo en cuenta los resultados obtenidos tras este análisis, se aplicará el modelo a los precios horarios.

El punto de partida para este análisis es el ajuste de un modelo de función de transferencia con ruido ARIMA-ARCH (FT-ARIMA-ARCH) identificado como se explica en 3.3.2. Los datos usados son los correspondientes al período comprendido entre el 1 de enero de 2007 y el 31 de agosto de 2010, lo que suponen 1339 muestras. Las variables explicativas para este modelo son los valores medios diarios de las previsiones de demanda y producción eólica proporcionadas por el operador del sistema. En la figuras 4.1a y 4.1b se representan estas variables, así como el precio medio para todo el período, respectivamente. Como puede verse, los precios presentan un comportamiento claramente no estacionario, por lo que previamente al ajuste las variables fueron diferenciadas tanto regular como estacionalmente (diferencias diarias y semanales). El modelo FT-ARIMA-ARCH identificado responde a la siguiente formulación

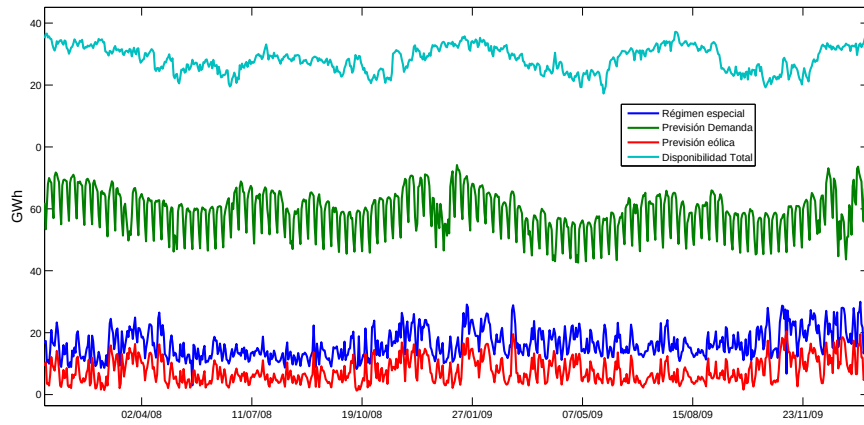
$$z_t = w_1 x_{1t} + w_2 x_{2t} + \frac{(1 + \theta_1 L)(1 + \Theta_1 L^7)}{(1 + \phi_1 L)} \varepsilon_t \quad (4.92)$$

$$v_t = -\gamma_0 + (1 - \gamma_1 L)(1 - \Gamma L^7) \varepsilon_t^2 \quad (4.93)$$

donde $z_t = (1 - L)(1 - L^7)y_t$ es el precio medio diario diferenciado tanto regular como estacionalmente, x_{1t}, x_{2t} son la previsiones medias diarias de la demanda y



(a) Precio medio diario en OMEL



(b) Disponibilidad total del sistema (incluye los grupos de tipo carbón, nuclear, hidráulicos, de fuel-gas y ciclos combinados), previsión de demanda, programa total casado de los grupos de régimen especial, y previsión de producción eólica

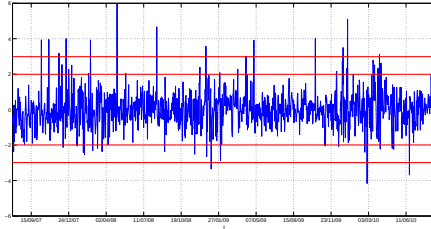
producción eólica (también diferenciadas), $w_1, w_2, \theta, \Theta, \phi, \gamma_0, \gamma_1, \Gamma$ los parámetros del modelo y v_t un proceso de ruido blanco (ver 2.2.2.5).

El modelo en (4.92) - (4.93) puede ser formulado en la representación de espacio de estados presentada en 4.2.2, y por tanto, su estructura nos sirve para especificar la de las matrices del modelo de cambio de régimen. Con el fin de identificar las variables relevantes que determinan el régimen del sistema y que constituirán por tanto las entradas a la red neuronal, se usaron aquellas empleadas en Karakatsani and Bunn (2008a) y que fueron explicadas en 2.2.1. La programación diaria resultante de la casación del mercado para cada tecnología tienen un claro efecto sobre el precio, por lo que se usaron también sus retardos de un día como proxy. Las disponibilidades de los grupos fueron también consideradas para calcular el margen de reserva diario¹⁰.

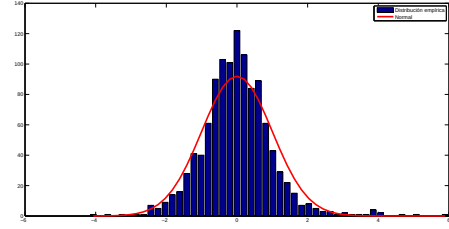
El proceso de identificación del modelo de cambio de régimen fue descrito en 4.2.4.2. Para un máximo de 4 regímenes, se estimaron diferentes modelos variando el número de variables de entrada a la red neuronal en función de la sensibilidad de la probabilidad de cada régimen frente a cada una de ellas, hasta alcanzar un conjunto de variables mínimo para un determinado número de regímenes. También se estimaron diferentes modelos variando el número de neuronas. En la tabla 4.1 se muestra un resumen con los resultados de los mejores modelos obtenidos (en términos de AIC) para cada uno de los regímenes ensayados (referenciados como RS2, RS3 y RS4), así como los resultados del modelo de función de transferencia con ruido ARIMA-ARCH en (4.92)-(4.93) (referenciado como FT-ARCH). Como puede apreciarse, el índice AIC obtenido para cada uno de los modelos indica que a pesar del incremento en el número de parámetros que supone la red neuronal, estos contribuyen de manera significativa a maximizar la verosimilitud de la muestra de datos. Como resulta lógico, a medida que se incrementa el número de regímenes (y por tanto, de parámetros) se obtiene un valor mayor para la verosimilitud logarítmica, puesto que es la función que se maximiza en el ajuste. En términos de AIC el modelo con 3 regímenes parece ser el más adecuado. Este hecho se revela más claro si atendemos a los resultados que se presentan en las figuras 4.1 y 4.2. Por ejemplo en los gráficos q-q plots, los cuales contrastan las distribuciones empíricas de los residuos estandarizados (calculados según 4.90 y 4.91) con la normal estándar asumida, muestran claramente que el modelo con 3 regímenes es el que mejor representa la distribución de probabilidad de los precios.

Para el modelo RS3, las variables explicativas más relevantes a la hora de determinar los diferentes regímenes de comportamiento fueron: las previsiones de demanda y producción eólica, el precio del día anterior, el margen de reserva y la programación total en régimen especial del día anterior (ver figuras 4.1a y 4.1b). Esta última variable, aunque bastante correlada con la previsión de producción

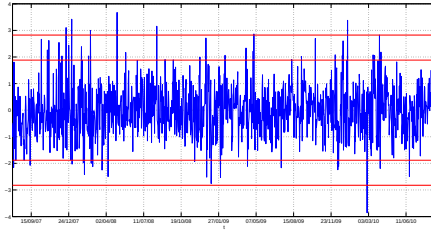
¹⁰El margen de reserva definido en Karakatsani and Bunn (2008a) es una medida del exceso de capacidad de generación, y se calcula como la máxima disponibilidad agregada de todos los grupos menos la previsión de demanda del sistema. Nótese sin embargo que para los grupos de tipo *régimen especial* (como lo son los parques eólicos) no se dispone de este dato en OMEL. Por tanto, el margen aquí calculado es la suma de las disponibilidades de los grupos con tecnología de tipo carbón, ciclo combinado, nuclear, fuel-gas e hidráulica, menos la previsión de demanda del sistema.



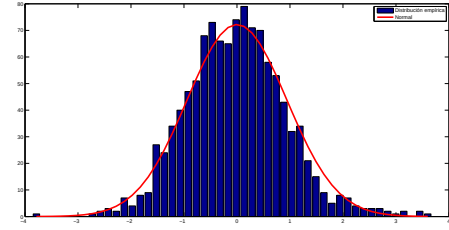
(a) Residuo estandarizado FT-ARCH



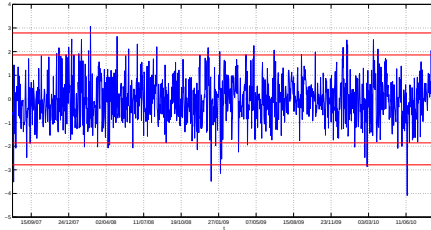
(b) Histograma FT-ARCH



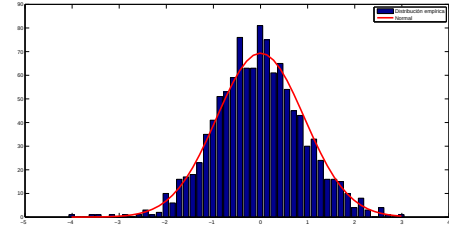
(c) Residuo estandarizado RS2



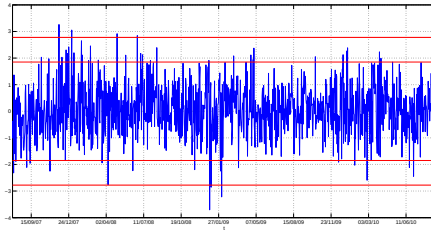
(d) Histograma RS2



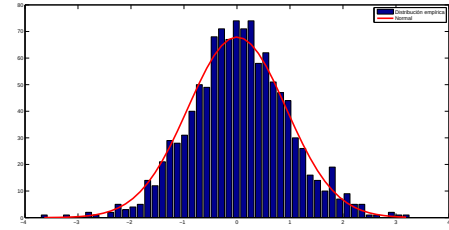
(e) Residuo estandarizado RS3



(f) Histograma RS3

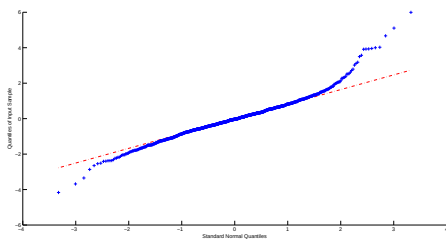


(g) Residuo estandarizado RS4

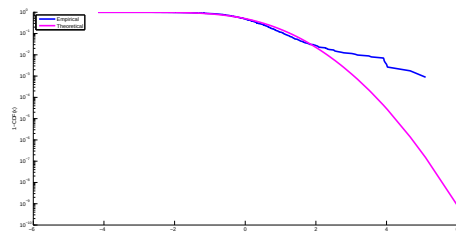


(h) Histograma RS4

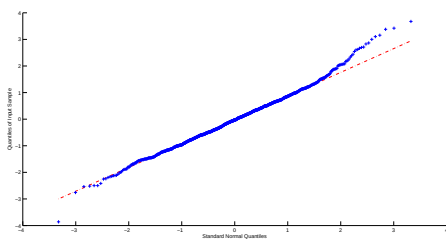
Figura 4.1: Residuos estandarizados de los modelos estimados (izquierda) y su histograma (derecha).



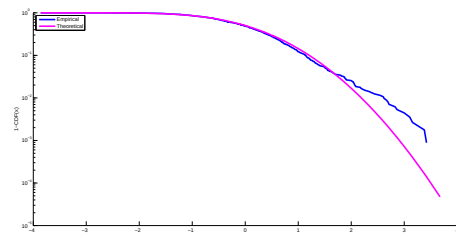
(a) q-q plot FT-ARCH



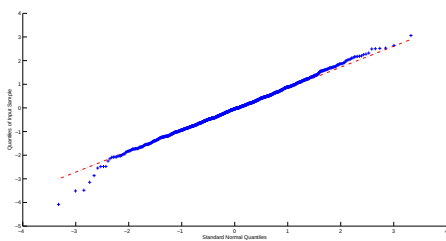
(b) cola derecha FT-ARCH



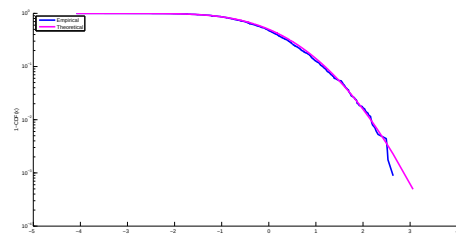
(c) q-q plot RS2



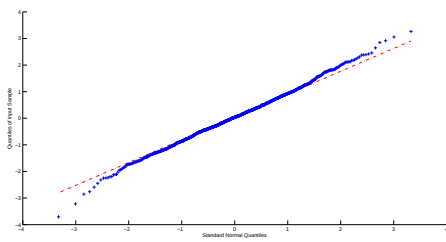
(d) cola derecha RS2



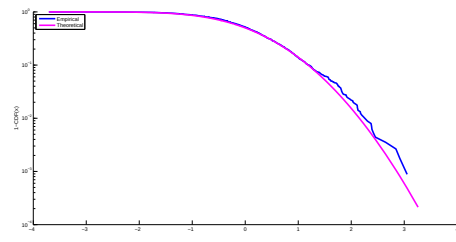
(e) q-q plot RS3



(f) cola derecha RS3

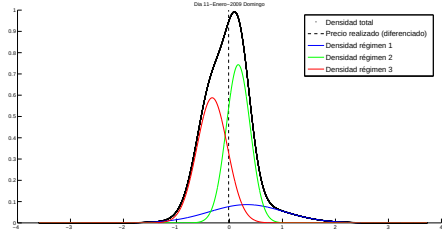


(g) q-q plot RS4

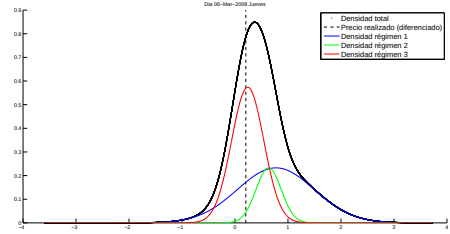


(h) cola derecha RS4

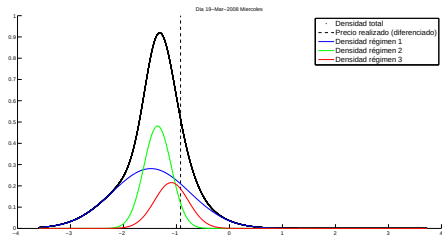
Figura 4.2: q-q plots de los modelos estimados (izquierda) y distribución de probabilidad acumulada (derecha).



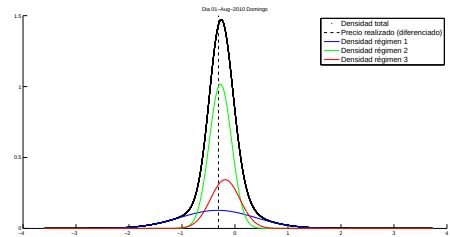
(a) 11 enero 2009, domingo



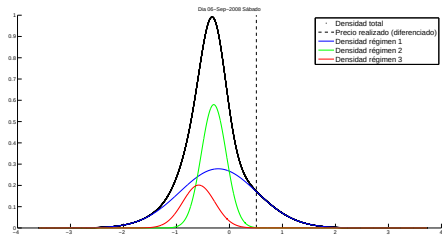
(b) 06 marzo 2008, jueves



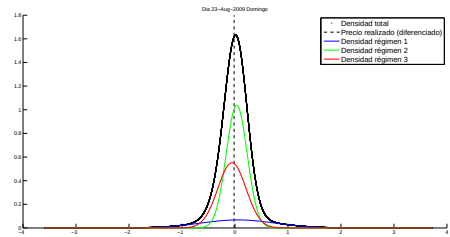
(c) 19 marzo 2008, miércoles



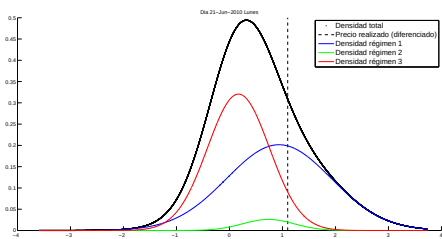
(d) 01 agosto 2010, domingo



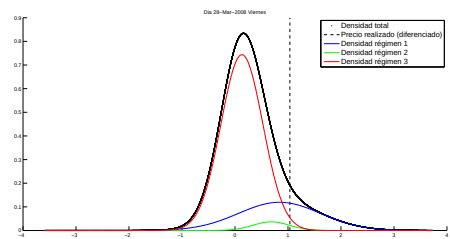
(e) 06 septiembre 2008, sábado



(f) 23 agosto 2009, domingo



(g) 21 junio 2010, lunes



(h) 28 marzo 2008, viernes

Figura 4.3: Funciones de densidad estimadas por el modelo RS3. En cada gráfico aparece la contribución de cada régimen (Régimen 1 en azul; Régimen 2 en verde; Régimen 3 en rojo) a la función de densidad total resultante (en negro). La línea punteada representa el valor observado del precio diferenciado.

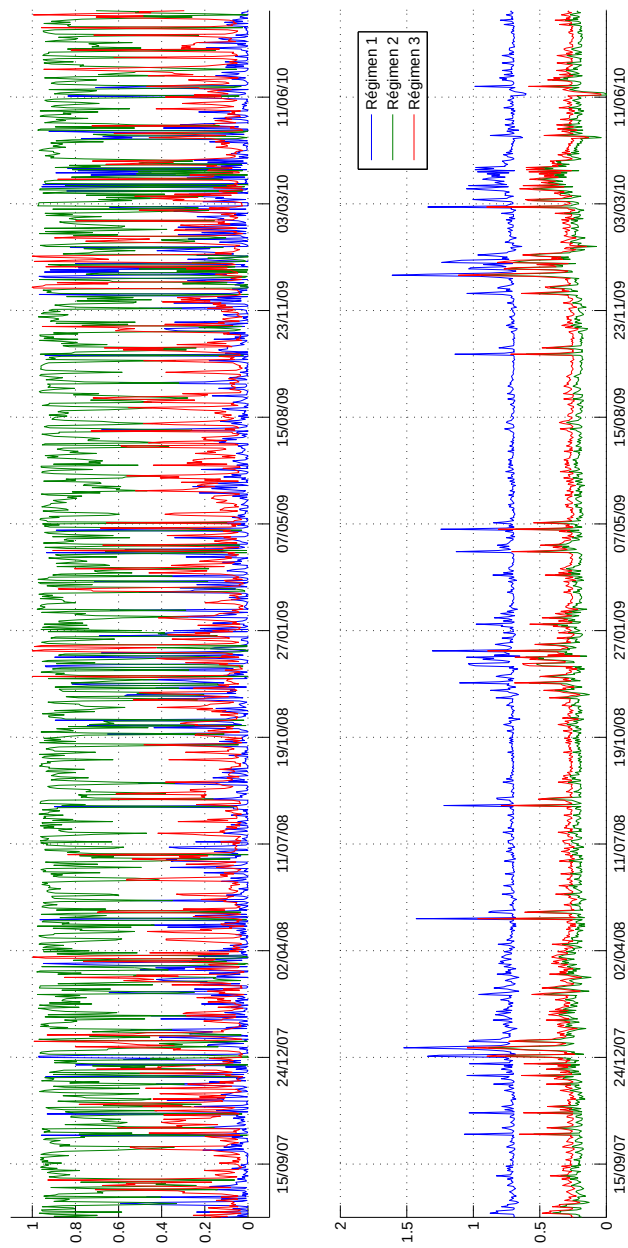


Figura 4.4: Distribución de probabilidad del régimen (arriba) y desviación estándar condicional a cada uno de los regímenes (abajo) a lo largo del período (régimen 1 en azul; 2 en verde; 3 en rojo)

Modelo	N. regímenes	Ver. Log	AIC	N. parámetros
FT-ARCH	1	-1101.4159	2218.8319	8
RS2	2	-910.2399	2020.4799	100
RS3	3	-888.8564	2017.7128	120
RS4	4	-880.5284	2041.0568	140

Tabla 4.1: Resultados de la estimación de los modelos diarios para la serie de OMEL.

eólica, se reveló como significativa probablemente debido al hecho de que proporciona información de la producción de las tecnologías no recogidas en ninguna otra variable. Respecto a las previsiones de demanda y producción eólica y el precio del día anterior, resulta lógico que se revelen como significativas, puesto que se trata de las variables incluidas en los modelos expertos encargados de las predicciones del precio condicionales a cada uno de los regímenes. El efecto del margen de reserva sobre el precio es el observado también en otra publicaciones en donde se argumenta que el precio aumenta de manera no lineal según el margen se va reduciendo y acercando a cero (ver por ejemplo Mount et al. (2006)).

La extracción del conocimiento implícito de la red neuronal ajustada, es decir, de la interacción entre las variables explicativas que determinan en qué circunstancias se activa uno u otro régimen, resulta difícil de analizar. En las figuras 4.5, 4.6 y 4.7 se muestra un gráfico similar a un *scatter plot*. En la diagonal de este gráfico se representan los valores estimados de la distribución de probabilidad del régimen para cada una de las variables explicativas de entrada a la red neuronal y por cada uno de los posibles regímenes. Es decir, cada gráfico en la diagonal representa la probabilidad estimada que tiene una muestra del conjunto de entrenamiento de pertenecer a un determinado régimen en función del valor de cada variable de entrada (probabilidad a priori). Los gráficos fuera de la diagonal son gráficos x - y entre cada par de variables, y en cada punto de la nube el valor de probabilidad estimada para cada régimen viene representado por la intensidad del color, siendo rojo probabilidad 1 y azul oscuro 0¹¹. A tenor de estas figuras puede decirse a grandes rasgos que el régimen 1 se activa en situaciones en las cuales se ha reducido la previsión de demanda con respecto a la semana anterior, y el margen de reserva es alto. El régimen 2 es el más habitual de los regímenes, por lo que puede asociarse con el régimen de comportamiento normal, y parece activarse más cuando existe una reducción en la previsión de producción eólica y de régimen especial. El régimen 3 se activa en situaciones de incremento de demanda, margen bajo e incrementos en las previsiones de producción eólica y de régimen especial.

En los gráficos de la figura 4.3 se representa las función de densidad estimada por el modelo RS3 en diferentes días del período de ajuste, así como la contribución de cada régimen a dicha distribución. Como puede observarse, las funciones de densidad

¹¹Nótese que algunas de las variables fueron diferenciadas semanalmente para evitar el posible efecto de distorsión de la tendencia en la detección de los regímenes ocultos.

resultantes de combinar las gaussianas de cada régimen con distintos pesos en cada instante de tiempo dan lugar a diferentes formas de densidades con asimetría y colas pesadas. Atendiendo a los gráficos, se observa que el régimen 1 parece ser el causante de las colas de la distribución estimada, por lo que se deduce que existe mayor incertidumbre en la predicción del precio para valores de demanda bajos que en otras circunstancias. Este hecho viene refutado también por el gráfico de la figura 4.4, en donde puede apreciarse cómo la desviación estándar del régimen 1 es superior a la de los otros dos regímenes en prácticamente todo el período de ajuste. En el gráfico superior de esta misma figura se observa también cómo el modelo capta un comportamiento diferente del precio en los períodos correspondientes a festividades (navidades, semana santa, puente de mayo, etc), puesto que se activan con mayor frecuencia los regímenes menos habituales, es decir, el 1 y el 3.

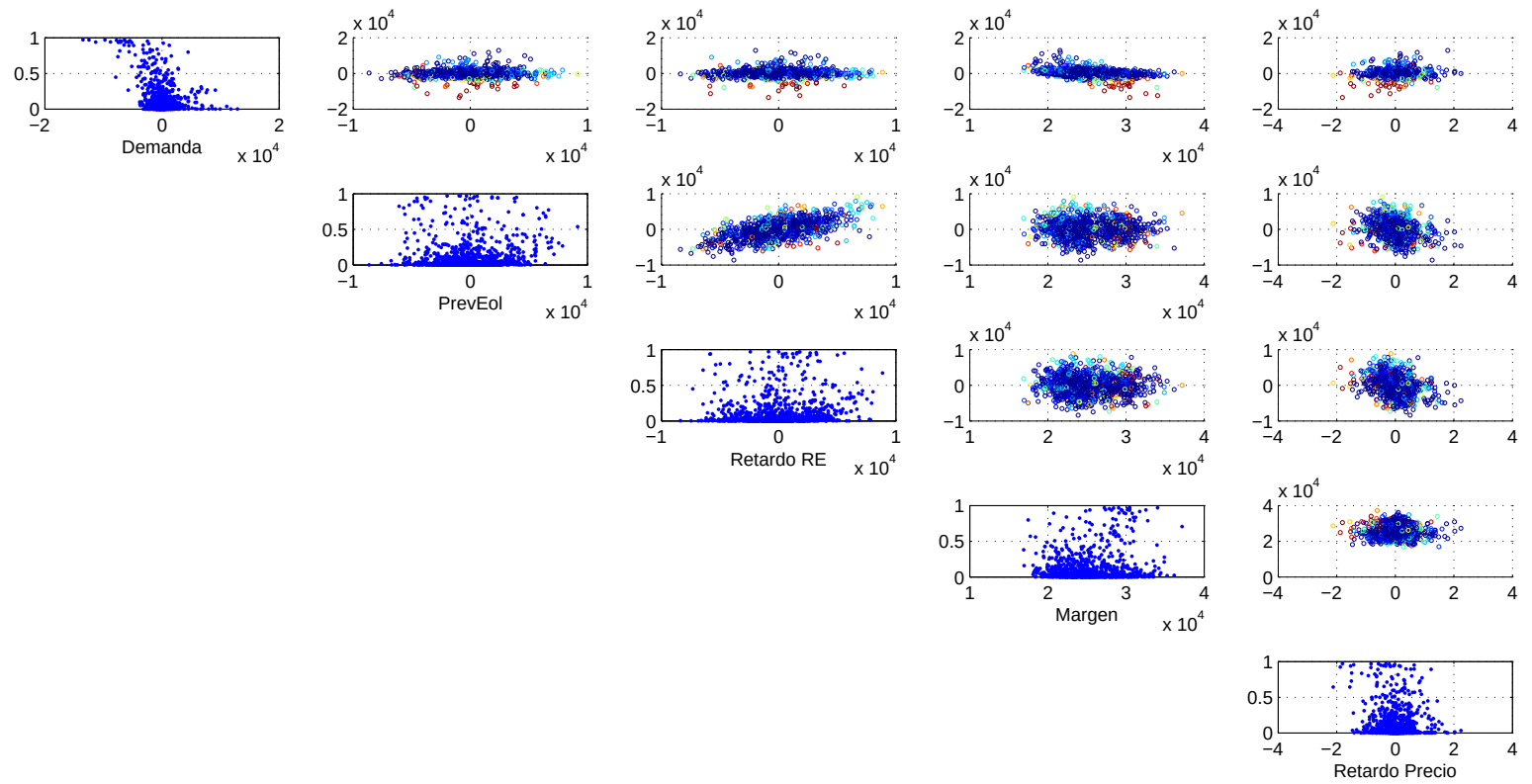


Figura 4.5: Probabilidad estimada para el primer régimen en función de las variables explicativas.

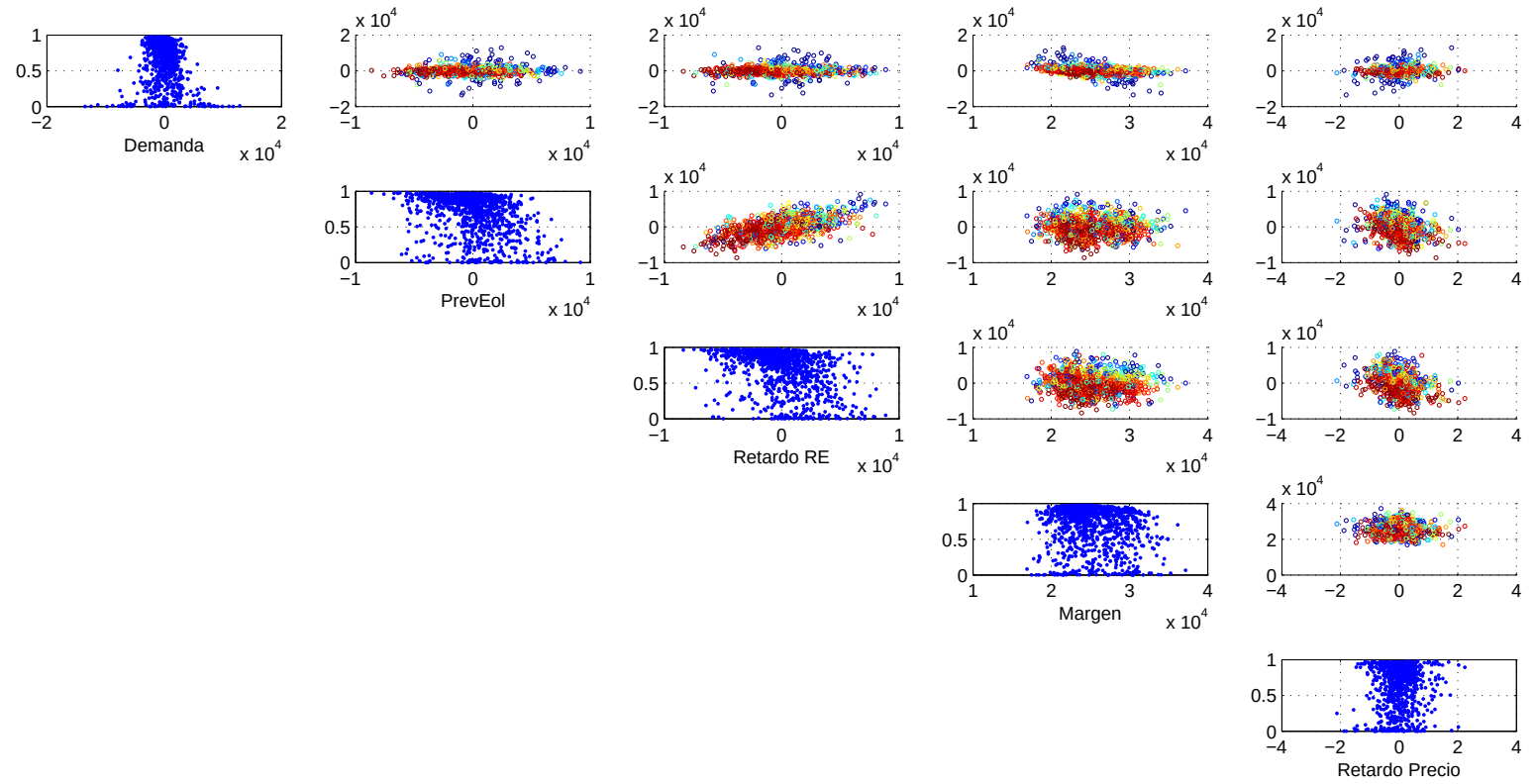


Figura 4.6: Probabilidad estimada para el segundo régimen en función de las variables explicativas.

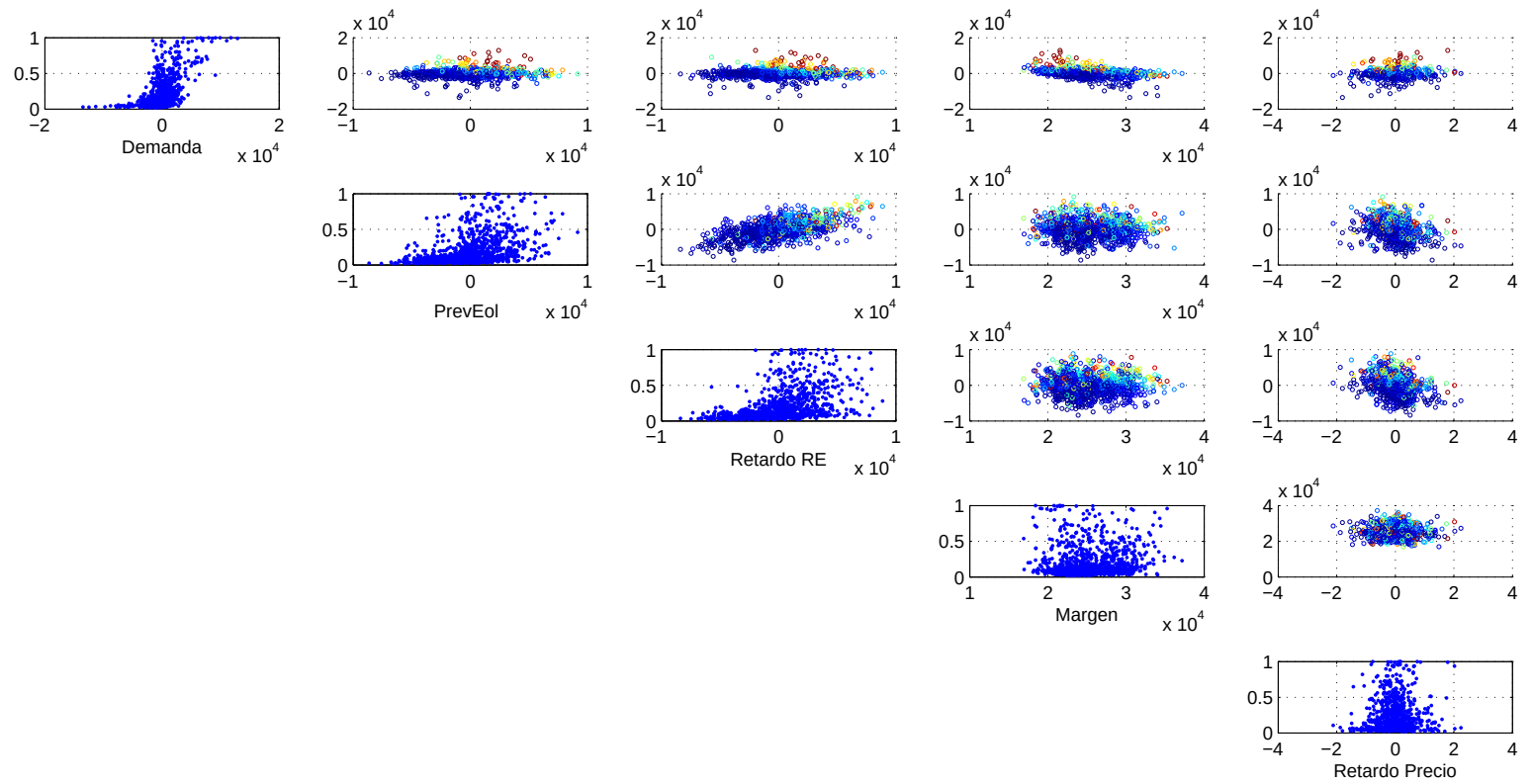


Figura 4.7: Probabilidad estimada para el tercer regimen en función de las variables explicativas.

4.4. Aplicación a los precios horarios del mercado Ibérico

En la sección anterior se ha mostrado cómo la distribución de probabilidad de los precios medios diarios puede representarse usando el modelo de cambio de régimen descrito en 4.2.4. El objetivo en esta sección es encontrar también un modelo que represente fielmente la distribución de probabilidad de los precios horarios en OMEL.

El conjunto de datos para el análisis es el mismo que en el apartado anterior, es decir, se dispone de los datos comprendidos entre el 1 de enero de 2007 y el 31 de agosto de 2010, lo que en este caso suponen 32136 muestras. El punto de partida es el modelo de función de transferencia con ruido ARIMA-ARCH previamente estimado e identificado en 3.4. Este modelo también puede ser formulado en la representación de espacio de estados expuesta en 4.2.2, y por tanto, su estructura nos sirve como punto de partida para especificar las matrices de un modelo de cambio de régimen. Su estimación e identificación presenta sin embargo algunas dificultades a nivel práctico que se pasan a exponer a continuación.

Las dificultades provienen fundamentalmente del hecho de que para capturar adecuadamente las diferentes dinámicas que presenta la serie de precios horarios en un intervalo de tiempo amplio (por ejemplo, en la sección 3.4 los modelos se estimaron tomando los datos correspondientes a un año completo, es decir, 8760 muestras) es necesario un elevado número de regímenes, lo que conlleva por tanto un aumento considerable en el número de parámetros. La consecuencia directa de esto es que, a pesar de utilizarse el algoritmo EM, resulta muy complicado que a partir de una inicialización arbitraria de parámetros se alcance un máximo aceptable debido a la compleja forma de la función de verosimilitud. Esta situación se ve agravada por el hecho de que la estimación del modelo resulta computacionalmente mucho más costosa que para el caso de usar datos en frecuencia diaria. Puesto que la serie presenta estacionalidad semanal (es decir, de período 168), los vectores de estado y las matrices del modelo equivalente son de gran tamaño (por ejemplo, los vectores de estado del modelo presentado en 3.4 para OMEL tienen 193 elementos). El efecto negativo que esto supone a la hora de implementar el filtro descrito en (4.72)-(4.80) puede ser solventado teniendo en cuenta que dichas matrices presentan gran cantidad de ceros. Sin embargo, el aumento más importante en el coste computacional no proviene del tamaño de las matrices del modelo sino del incremento en el número de regímenes necesarios, puesto que el número de operaciones a realizar para la evaluación del filtro en cada muestra se incrementa con el cuadrado de los mismos.

Una solución al problema podría ser separar la serie de precios horarios en 24 series diarias, una para cada hora, y modelar cada una de ellas por separado de manera similar a los precios medios diarios. Esta manera de tratar las series de precios horarios ha sido empleada antes por diversos autores (ver por ejemplo Cuaresma et al., 2004; Zareipour et al., 2006b; Misiorek et al., 2006; García-Martos et al., 2007). En la figura 4.8 aparecen los gráficos q-q plots de los residuos estandarizados de los modelos de cambio de régimen estimados de manera separada para cada hora del

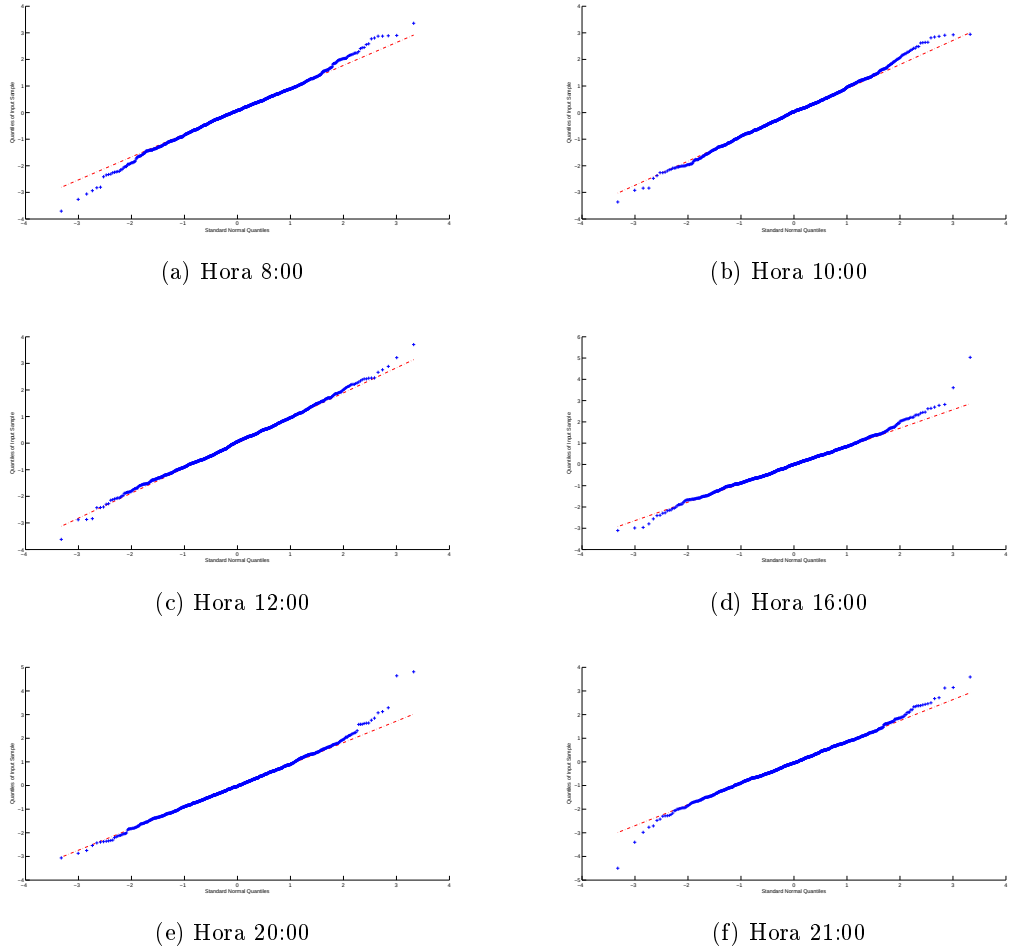
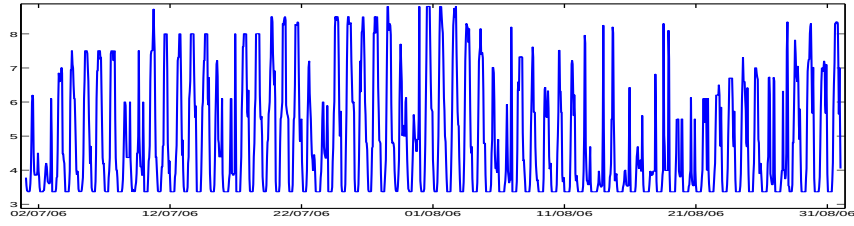


Figura 4.8: Gráficos qqplot de los residuos estandarizados de los modelos estimados de manera separada para las horas 8:00, 10:00, 12:00, 16:00, 20:00 y 21:00.

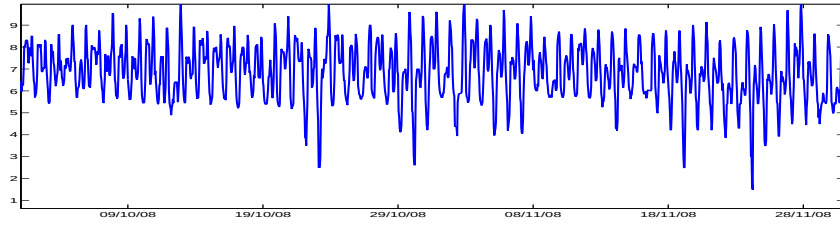
día¹². Observando dichas figuras cabrían sacar las mismas conclusiones que en el caso de los modelos aplicados a los precios medios diarios. La forma de la función de densidad asumida se ajusta bastante bien a la distribución de los residuos empíricos obtenidos. En particular, el ajuste en las colas es mucho mejor que el proporcionado por los modelos de función de transferencia con ruido ARIMA-ARCH.

Sin embargo, el inconveniente de este enfoque radica en el hecho de que no

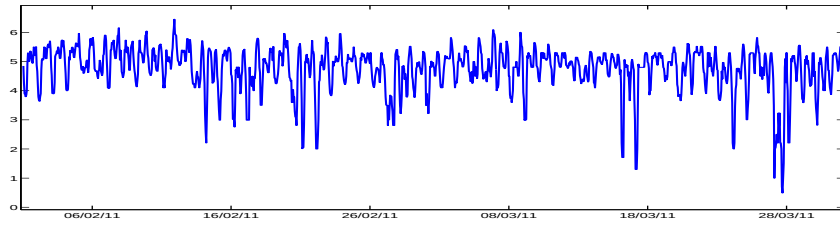
¹²El conjunto de datos para la estimación de los modelos cuyos resultados se muestran en las figura 4.8 fue separado en 24 series diferentes, una por cada hora del día (lo que da lugar a 24 conjuntos de datos de 1339 muestras cada uno). Las variables explicativas seleccionadas como entradas a las redes neuronales fueron aquellas que se revelaron como significativas en el análisis de los precios medios diarios: las previsiones de demanda y generación de energía eólica, el margen de reserva, el programa casado para los grupos de régimen especial y el precio del día anterior, todas ellas para cada hora específica. Se eligieron 3 regímenes para los 24 modelos a estimar.



(a) Período 1: 01/Jul - 31/Ago 2006



(b) Período 2: 01/Oct 30/Nov 2008



(c) Período 3: 01/Feb 31/Mar 2011

Figura 4.9: Precio del mercado de OMEL (en c€/KWh) en los distintos períodos de estimación.

considera la correlación de los precios entre horas diferentes. Para el caso de la serie de precios objeto de este estudio, esta simplificación no es despreciable en términos de predicción puntual o precisión del modelo, por lo que separar las series de precios en 24 series diarias no resulta una solución aceptable.

La solución adoptada pasa entonces por considerar un único modelo para la serie de precios horarios y utilizar conjuntos de entrenamiento de menor tamaño. Esto permite un doble ahorro computacional: por un lado, un menor número de muestras hace que la evaluación de la verosimilitud con el filtro descrito en (4.72)-(4.80) sea más rápida; por otro lado, al estimar utilizando un conjunto de datos de menor tamaño se requiere de un menor número de regímenes para representar adecuadamente la distribución de probabilidad del proceso de los precios. A la hora de utilizar un modelo así estimado para predecir *fuera de muestra*, será necesario implantar algún tipo de seguimiento de los errores de predicción para detectar posibles cambios en la dinámica de la serie que requieran reajustar los parámetros del modelo.

Modelo	N. regímenes	Ver. Log	AIC	N. parámetros	MAE	MAPE(%)
<i>Período 1</i>						
FT-ARCH	1	-207.7615	437.523	11	0.325	6.451
RS2	2	-522.5913	1213.1826	84	0.303	5.964
RS3	3	-363.6834	929.3668	101	0.292	5.711
RS4	4	-231.7987	699.5975	118	0.287	5.576
RS5	5	-82.1044	434.2089	135	0.283	5.459
<i>Período 2</i>						
FT-ARCH	1	-148.7879	325.5759	14	0.215	3.449
RS2	2	-27.6395	259.279	102	0.208	3.366
RS3	3	41.7194	162.5612	123	0.204	3.31
RS4	4	34.2455	219.5091	144	0.206	3.337
RS5	5	34.672	260.6561	165	0.206	3.314
<i>Período 3</i>						
FT-ARCH	1	-273.8978	577.7957	15	0.17	4.422
RS2	2	-933.761	647.0227	110	0.169	4.413
RS3	3	-1126.3034	583.7187	135	0.171	4.294
RS4	4	-66.2417	452.4833	160	0.165	4.252
RS5	5	-70.0336	504.0672	182	0.164	4.196

Tabla 4.2: Resultados de la estimación de los modelos horarios.

A continuación se muestran los resultados de la aplicación del modelo de cambio de régimen propuesto en este capítulo en tres períodos de tiempo diferentes: El período 1 comprendido entre el 1 de Julio y el 31 de Agosto de 2006; el período 2 entre el 1 de Octubre y el 30 de Noviembre de 2008; y el período 3 entre el 1 de Febrero y el 31 de Marzo de 2011 (ver figura 4.9). En la tabla 4.2 se resumen los resultados obtenidos de la estimación de los modelos para cada período de ajuste diferente. En ella aparecen también los resultados de aplicar sus contrapartidas lineales (que presentan un único régimen) con el fin de poder compararlos entre sí. Para el período 1 no se dispone de ninguna otra variable más que el precio, por lo que las variables de entrada a la red de probabilidad del régimen fueron la hora del día y retardos del precio (en particular, el precio retardado en 1, 2^{13} , 24 y 168 muestras). Para los períodos 2 y 3 se dispone del mismo conjunto de variables que para el caso de los modelos diarios de la sección anterior pero en frecuencia horaria. Las variables explicativas usadas para las redes de probabilidad del régimen fueron la previsión de demanda y producción eólica, la hora del día y diferentes retardos del precio.

Atendiendo al criterio de información de Akaike (AIC), los mejores modelos son el RS5 para el período 1, el RS3 para el período 2 y el RS4 para el período 3. En los tres casos, el incremento de parámetros que suponen estos modelos frente a los de función de transferencia con ruido ARIMA-ARCH (FT-ARCH) está justificado en

¹³En el caso de utilizar el modelo de cambio de régimen con el objetivo de predecir los 24 precios del día siguiente, utilizar retardos más pequeños de 24 muestras como entradas a la red de probabilidad del régimen supone una dificultad añadida puesto que no se dispone de dicha información al realizar las predicciones. Esta problemática se aborda en el siguiente capítulo.

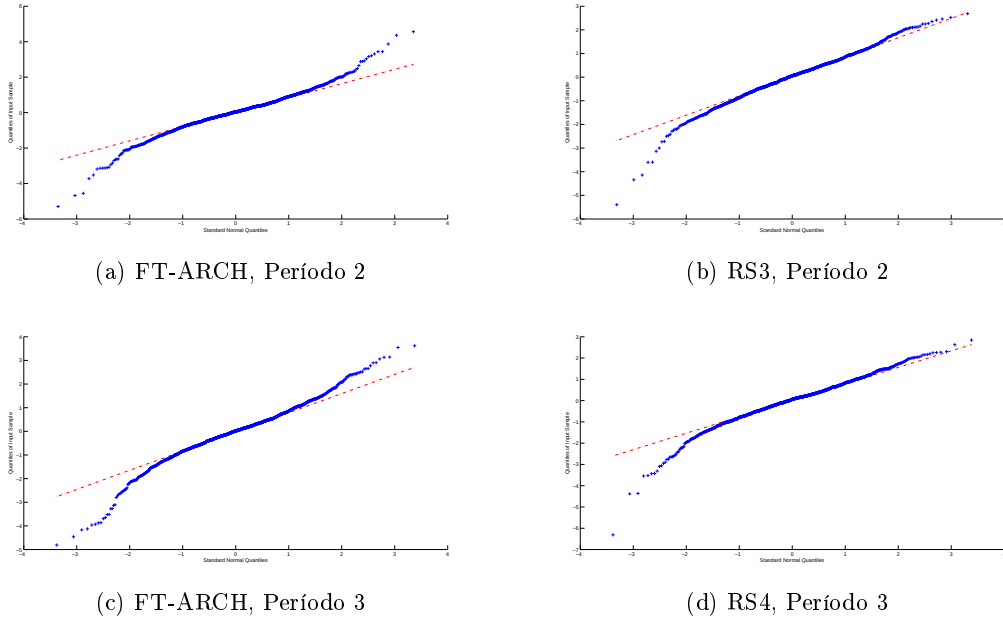
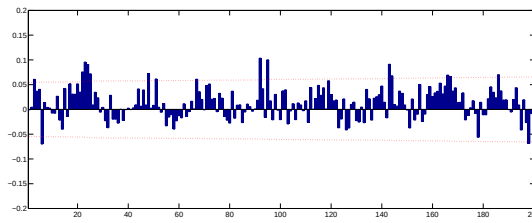


Figura 4.10: Gráficos q-q plot de los residuos estandarizados de los modelos lineales FT-ARCH (izquierda) y sus contrapartidas de cambio de régimen (derecha).

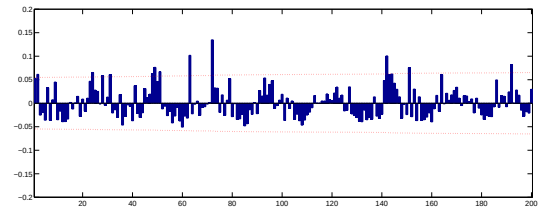
términos de verosimilitud de la muestra. También el MAE y MAPE que arrojan estos modelos supera al de sus contrapartidas lineales en los tres periodos considerados, en menor medida para los periodos 2 y 3, pero llegando a ser una disminución drástica para el periodo 1. La mejora de los modelos de cambio de régimen en los periodos 2 y 3 con respecto a los de tipo FT-ARCH queda evidenciada si atendemos a los gráficos q-q plot de los residuos estandarizados, los cuales se muestran en la figura 4.10.

Como sucedió en la sección 3.4, una vez eliminada la dependencia en media de la serie con el modelo de función de transferencia con ruido ARIMA, el cuadrado del residuo presenta un claro patrón de autocorrelación que es necesario modelar con términos de tipo ARCH. Sin embargo, dicho patrón de autocorrelación no apareció en los residuos de los modelos de cambio de régimen en ninguno de los tres periodos considerados, por lo que no fue necesario incluir ningún término para modelar la varianza condicional a excepción de la constante para cada régimen¹⁴. En la figura 4.11 se muestran las funciones de autocorrelación simple (FAS) de los residuos estandarizados de los modelos ensayados, así como también las FAS del cuadrado de dichas series. En ella se aprecia cómo las series de precios horarios no presentan ningún patrón de autocorrelación que justifique la inclusión de términos autorregresivos que modelen la varianza. Este hecho pone de manifiesto la limitación de los modelos lineales para estimar la medida del proceso de precios horarios de OMEL.

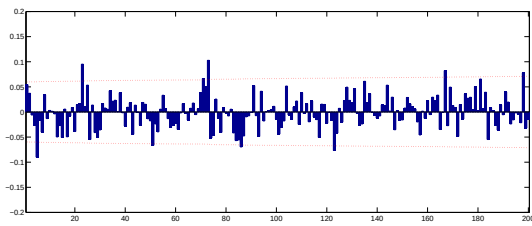
¹⁴Nótese que este hecho no ocurría en el caso de las series diarias.



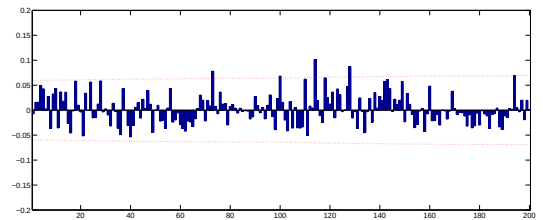
(a) FAS del residuo (RS5, Período 1)



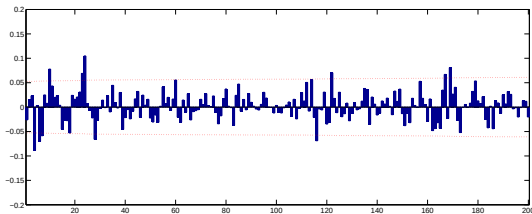
(b) FAS del cuadrado del residuo (RS5, Período 1)



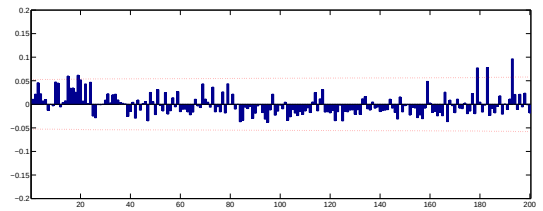
(c) FAS del residuo (RS3, Período 2)



(d) FAS del cuadrado del residuo (RS3, Período 2)



(e) FAS del residuo (RS4, Período 3)



(f) FAS del cuadrado del residuo (RS4, Período 3)

Figura 4.11: Gráficos de la Función de Autocorrelación Simple (FAS) de los residuos estandarizados (izquierda) y su cuadrado (derecha) para cada uno de los periodos ensayados.

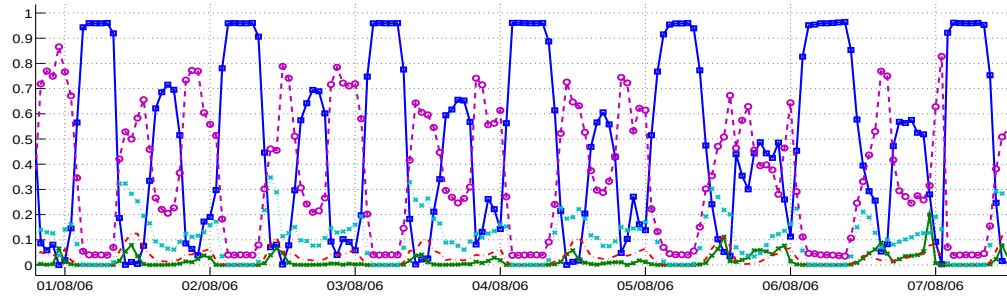
Los modelos de cambio de régimen estimados pueden escribirse por tanto mediante una única ecuación de estado y otra de salida, como en (4.94)-(4.95)

$$\xi_{t+1} = F_{r_{t+1}}\xi_t + B_{r_{t+1}}x_t + \varpi_{t+1} \quad (4.94)$$

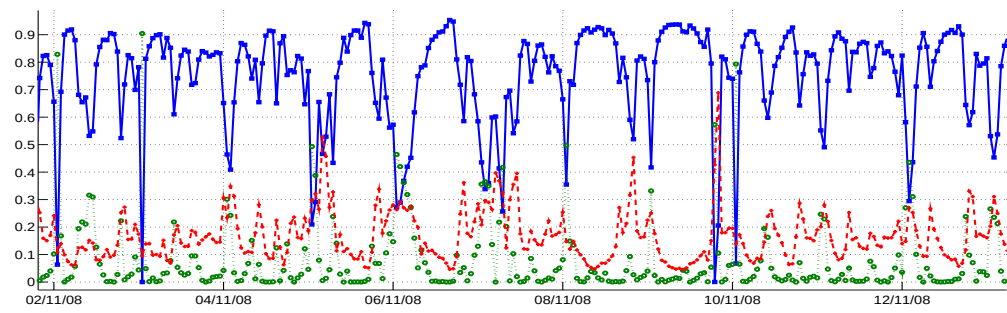
$$y_t = H_{r_t}\xi_t + A_{r_t}x_t \quad (4.95)$$

donde $r_t \in \{1, \dots, R\}$ denota el régimen en el instante t , las matrices $F_{r_t}, B_{r_t}, H_{r_t}, A_{r_t}$ quedan definidas por los parámetros estimados y $\varpi_t = (\alpha_{r_t}\varepsilon_t, 0, \dots, 0)'$, siendo α_{r_t} el parámetro de desviación típica de las innovaciones del régimen r_t .

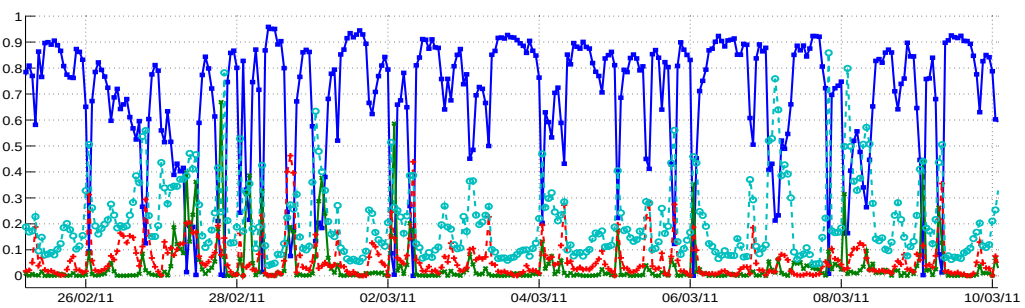
Recuérdese que la distribución de probabilidad de r_t se modela mediante una red neuronal como se describe en (4.68)-(4.70). Los gráficos de la figura 4.12 muestran la evolución de la distribución de probabilidad estimada del régimen r_t a lo largo de tiempo en los tres períodos ensayados. Como puede apreciarse, en los tres casos existe un régimen predominante que podría asociarse como el de comportamiento normal, y dentro de un mismo día se activan el resto de regímenes en mayor o menor medida según tomen valor el resto de variables explicativas. Para el caso del período 1, el régimen del sistema está determinado principalmente por la hora del día. Si atendemos a la serie de precios en este período (ver figura 4.9 y 4.13) se aprecia cómo parece estar fijado o saturado en algunas horas del día. En la figura 4.13 puede apreciarse el error del modelo FT-ARCH para capturar dicha saturación con respecto al modelo de cambio de régimen.



(a) Distribución de probabilidad estimada del régimen (período 1, modelo RS5)



(b) Distribución de probabilidad estimada del régimen (período 2, modelo RS3)



(c) Distribución de probabilidad estimada del régimen (período 3, modelo RS4)

Figura 4.12: Distribuciones de probabilidad estimadas para los regímenes de los modelos estimados durante parte del período de ajuste. La probabilidad de cada uno de los posibles valores del régimen está representada con un color diferente.

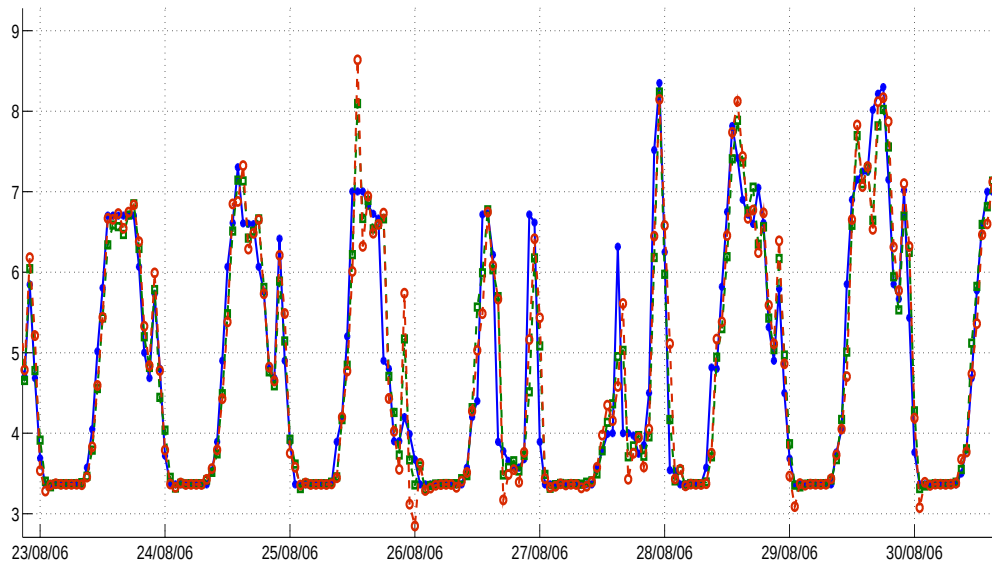


Figura 4.13: Valor del precio real (en azul) frente a la estimación del modelo FT-ARCH (en rojo, círculos) y RS5 (en verde, cuadrado) en el período 1.

4.5. Conclusiones

El comportamiento no lineal de los precios eléctricos, evidenciado tanto en la respuesta de los precios a variables explicativas como en las dinámicas autorregresivas de las propias series, no puede ser descrito adecuadamente utilizando los modelos lineales habitualmente utilizados como métodos de predicción. El modelo propuesto en este capítulo hace frente a esta debilidad generalizando el de función de transferencia con ruido ARIMA-ARCH al permitir que los parámetros del modelo varíen con el tiempo según un conjunto reducido de regímenes no observados o ficticios, los cuales a su vez están determinados por las variables explicativas que determinan el comportamiento del mercado. Respecto al modelo SNP-ARIMA-ARCH propuesto en el capítulo anterior, el aquí descrito proporciona una capacidad de representación notablemente más flexible, puesto que la función de densidad resultante es una mixtura de densidades gaussianas con diferentes pesos en cada instante de tiempo. Como contraprestación, el número de parámetros se ve incrementando considerablemente y el proceso de estimación resulta mucho más complejo.

La capacidad del modelo de aplicar un juego diferente de parámetros en cada instante de tiempo se consigue utilizando una representación en espacio de estados. A partir de dicha representación y un juego de parámetros fijados, el filtro descrito en Kim (1994) permite estimar la salida del modelo condicional a la información

conocida hasta ese momento en cada punto de la serie temporal. La distribución de probabilidad de cada uno de los regímenes se determina a través de una red neuronal que presenta como entradas las variables fundamentales que afectan al precio. Esto permite aprender de manera automática las diferentes situaciones del mercado que determinan distintas respuestas o dinámicas del precio. Los parámetros del modelo pueden estimarse usando el algoritmo EM Dempster et al., 1977; Hamilton, 1990, el cual permite maximizar de manera separada los parámetros que determinan la distribución de probabilidad de los regímenes de aquellos que modelan la media y varianza condicional del precio, siendo a la vez robusto frente a una inicialización aleatoria de los mismos.

El modelo se ha aplicado a las series de precios horarios y precios medios diarios de OMEL, poniéndose de manifiesto cómo el modelo de cambio de régimen puede modelarlas más adecuadamente que un modelo lineal como el de función de transferencia con ruido ARIMA-ARCH. Dicha mejora queda evidenciada sobre todo en las colas de la distribución condicional de los precios tanto diarios como horarios. Puesto que la distribución de probabilidad del régimen no observado está modelada por una red neuronal, resulta complejo extraer el conocimiento implícito que explique en qué circunstancias se activa uno u otro régimen. Para el modelo de precios diarios, existe un régimen claramente diferenciado de los otros dos, que se activa en situaciones de decrecimiento de previsión de demanda y que es el responsable de las colas pesadas de la distribución. En el caso de las serie horarias, los modelos de cambio de régimen estimados consiguieron capturar también diferentes dinámicas intradiarias determinadas fundamentalmente por la hora del día y su nivel de previsión de demanda.

Capítulo 5

Aplicación a la predicción de los precios spot del día siguiente en el mercado ibérico

En este capítulo se evalúan las capacidades predictivas de los dos modelos propuestos en esta tesis dentro del marco de trabajo habitual, en dónde la casación de precios se resuelve diariamente para las 24 horas del día siguiente. En primer lugar se pone en relieve la problemática de utilizar estos modelos para calcular las distribuciones predictivas en un horizonte de hasta 24 horas, puesto que en su formulación incorporan la información de la hora anterior. También se presentan las medidas utilizadas para evaluar la precisión de las predicciones, tanto las utilizadas para medir la precisión de las estimaciones puntuales como las menos habituales que consideran las distribuciones predictivas completas y no solo su valor esperado. Finalmente, se presentan los resultados de la aplicación de los dos modelos a la predicción fuera de muestra de los precios horarios de OMEL, comparando su rendimiento con los modelos de series temporales más habitualmente utilizados en la literatura de precios.

5.1. Introducción

Nuestra preocupación en este capítulo es validar las capacidades de predicción de los modelos propuestos en esta tesis en el contexto habitual en el que se desenvuelven los agentes del mercado eléctrico diario. Como se describe en el capítulo 1, la casación de los 24 precios del día $D + 1$ en el mercado diario de OMEL se resuelve en torno a las 10 de la mañana del día D , por lo que antes de ese momento los agentes necesitan disponer de las estimaciones de los mismos para poder optimizar sus estrategias de compra o venta de energía. En este marco de trabajo, los modelos de predicción son usados para que a partir de la serie histórica de precios hasta un determinado día, se predigan en bloque los 24 precios horarios del día siguiente. El precio de la primera hora del día se predecirá por tanto con un horizonte de predicción de 1, la segunda a un horizonte de 2, y así sucesivamente hasta la última hora del día que se predecirá con un horizonte de 24 horas. Una vez resuelta la casación del mercado, la información de la serie de precios es actualizada y se procede de la misma manera para el día siguiente.

La manera en la que se resuelve la casación de los precios eléctricos del mercado del diario introduce la problemática del cálculo de las predicciones a un horizonte de predicción multiperíodo. Esta problemática ha sido abordada en trabajos como Baillie and Bollerslev, 1992; Cheng et al., 2006; Lütkepohl, 2010. En general, el cálculo de las expresiones explícitas de los momentos de una distribución a un determinado horizonte de predicción se complica en función de cuál sea la distribución de las innovaciones del modelo. Una manera sencilla de abordar el problema está basada en la generación aleatoria de las innovaciones del modelo y la simulación de sus trayectorias hasta el horizonte de predicción deseado (ver Hyndman et al. (2008)). Las innovaciones pueden ser generadas a partir de la función de densidad asumida, o bien mediante réplicas bootstrap del residuo del conjunto de estimación (Efron and Tibshirani (1993)). Las distribuciones predictivas de los dos modelos propuestos en esta tesis se calculan por simulación, y el procedimiento se describe en la sección 5.2.

Para poder evaluar la capacidad predictiva de los modelos frente a un conjunto de test se utilizan diferentes medidas de precisión. Las habitualmente utilizadas en la literatura se centran en la capacidad de los modelos para realizar estimaciones puntuales, es decir, están basadas en el valor esperado de la distribución predictiva del modelo y el valor realizado de la serie temporal. Sin embargo, existe otro grupo de medidas que consideran las distribuciones predictivas completas y, por tanto, son propicias para evaluar la precisión de los intervalos de predicción proporcionados por los modelos. Tanto unas como otras son presentadas en la sección 5.3, y serán utilizadas para evaluar las predicciones de los modelos ensayados en este capítulo.

En la sección 5.4 se presentan los resultados de la aplicación de un conjunto de modelos a la serie de precios horarios de OMEL. El período de validación es lo suficientemente amplio como para extraer conclusiones generales sobre el rendimiento de los modelos. El objetivo de la comparación no es sólo validar las predicciones que proporcionan los modelos propuestos en los capítulos 3 y 4, sino también contribuir

a responder a algunas cuestiones como si incluir variables explicativas en los modelos contribuye a predecir mejor el precio del mercado diario; si los términos ARCH contribuyen a predecir mejor el valor realizado del precio; o si los modelos propuestos son adecuados para predecir los intervalos de predicción. Entre los modelos ensayados se encuentran los modelos univariantes ARIMA y el de suavizado exponencial con doble estacionalidad, los cuales son usados como *benchmarks*. El resto de modelos son el de función de transferencia con la previsión de demanda y de producción eólica como variables explicativas, el de función de transferencia con ruido ARIMA-ARCH, la extensión semiparamétrica del mismo propuesta en el capítulo 3, y el modelo de cambio de régimen propuesto en el capítulo 4. En la comparativa se incluyen también los resultados de la generación de los intervalos de predicción o distribuciones predictivas por medio del método de remuestreo o *bootstrapping*. Las conclusiones de la comparativa son discutidas en la sección 5.5.

5.2. Predicción multiperíodo

En esta sección se describe la manera de calcular las distribuciones predictivas de los modelos propuestos en esta tesis en un horizonte de predicción mayor de 1 (*multi-step prediction*). En primer lugar, en el apartado 5.2.1 se describe la manera de calcular la media y varianza de las distribuciones predictivas de los modelos de tipo ARMA-ARCH a un determinado horizonte de predicción. Las expresiones resultantes pueden ser fácilmente generalizadas para el caso de que el modelo en cuestión presente influencia de variables explicativas consideradas deterministas y no estocásticas, como es el caso del modelo de función de transferencia con ruido ARIMA-ARCH descrito en el capítulo 3. Sin embargo, en general, el cálculo de los momentos de una distribución predictiva se complica en función de cual sea la forma de las innovaciones que rigen el proceso, o bien ni siquiera están disponible en la literatura (ver Baillie and Bollerslev (1992)).

Este hecho hace que sea preferible abordar el problema de calcular las distribuciones predictivas completas por simulación. Con este enfoque, es posible obtener de forma sencilla dichas distribuciones sin más que generar aleatoriamente las innovaciones del modelo según la densidad asumida y simular sus trayectorias con origen de predicción en t hasta el horizonte $t + k$ deseado. Aunque el método es computacionalmente intensivo, resulta perfectamente viable dada la potencia de los ordenadores actuales, y puesto que se trata de procesos univariantes para los cuales se necesita un número abordable de simulaciones. El inconveniente de esta solución es que no provee de las expresiones explícitas de las distribuciones obtenidas, hecho que puede ser resuelto utilizando alguno de los métodos de aproximación de funciones de densidad no paramétricos existentes, como el método basado en expansiones de Hermite descrito en la sección 3.2 (Gallant and Nychka (1987)) o estimadores kernel (Silverman (1986)).

5.2.1. Predicción con el modelo ARMA-ARCH

Puesto que en el modelo de función de transferencia con ruido ARIMA-ARCH las variables explicativas son deterministas y no estocásticas, para simplificar la exposición sin pérdida de generalidad, a continuación se presentan las expresiones para la media y varianza de las distribuciones predictivas condicionales de un modelo ARMA-ARCH en un horizonte de predicción mayor 1 (ver Baillie and Bollerslev (1992) para detalles).

La ecuación que define el proceso para la media en un modelo ARMA o ARMA-ARCH responde a

$$y_t = c - \sum_{i=1}^p \phi_i y_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t \quad (5.1)$$

por lo que tomando esperanzas condicionadas a la información contenida hasta el instante t , la predicción de la media del proceso con origen de predicción en t y a un horizonte de predicción k responde a

$$\begin{aligned} \hat{y}_{t+k|t} &= E[y_{t+k}|t] = c - \phi_1 E[y_{t+k-1}|t] - \cdots - \phi_p E[y_{t+k-p}|t] + \\ &\quad \theta_1 E[\varepsilon_{t+k-1}|t] + \cdots + \theta_q E[\varepsilon_{t+k-q}|t] \end{aligned} \quad (5.2)$$

En esta ecuación algunas esperanzas se aplican a variables observadas, por lo que $E[y_{t+k-i}|t] = y_{t+k-i}$ cuando $t+k-i \leq t$, con $i = 1, \dots, p$, y otras a no observadas, dependiendo del horizonte de predicción k y el orden de la parte autorregresiva p . Con relación a las innovaciones, la esperanza de las innovaciones futuras (no observadas) condicionada a la historia de la serie es igual a su esperanza absoluta, es decir 0, ya que el ruido del proceso es independiente de la variable de salida. Por tanto, $E[\varepsilon_{t+k-i}|t] = 0$ cuando $t+k-i > t$, con $i = 1, \dots, q$. El resto de innovaciones son conocidas, por lo que sus esperanzas coinciden con los valores observados, que se calculan recursivamente mediante la siguiente ecuación

$$\varepsilon_t = y_t + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} - c - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \quad (5.3)$$

y de esta modo se calculan recursivamente también las predicciones $\hat{y}_{t+1|t}, \dots, \hat{y}_{t+k|t}$ condicionales a la información disponible hasta el instante t , comenzando con el horizonte 1 hasta el horizonte k deseado.

Si partimos de la representación del proceso (5.1) como un proceso de media móvil de orden infinito MA(∞) (Baillie and Bollerslev, 1992; Hamilton, 1994),

$$y_{t+k} = c + \varepsilon_{t+k} + \psi_1 \varepsilon_{t+k-1} + \cdots + \psi_{k-1} \varepsilon_{t+1} + \psi_k \varepsilon_t + \psi_{k+1} \varepsilon_{t-1} + \cdots \quad (5.4)$$

se tiene que el error de predicción a horizonte k (*k-step-ahead error prediction*) responde a

$$e_{t,k} = y_{t+k} - E[y_{t+k}|t] = \sum_{i=1}^k \psi_{k-i} \varepsilon_{t+i} \quad (5.5)$$

y el error cuadrático medio (MSE) asociado a esta predicción es entonces

$$E[e_{t,k}^2|t] = Var[y_{t+k}|t] = \sum_{i=1}^k \psi_{k-i}^2 E[\sigma_{t+i}^2|t] \quad (5.6)$$

Así, con errores homocedásticos, el MSE de dicha predicción responde a

$$Var[y_{t+k}|t] = \sigma^2 \sum_{i=1}^k \psi_{k-i}^2 \quad (5.7)$$

donde σ^2 es la varianza (constante) de los errores de predicción a un paso del modelo. En presencia de heterocedasticidad condicional, es decir, cuando el modelo incluye un término ARCH, la incertidumbre de los errores de predicción es variable en el tiempo y toma entonces la siguiente forma

$$Var[y_{t+k}|t] = \sigma^2 \sum_{i=1}^k \psi_{k-i}^2 + \sum_{i=1}^s \psi_{k-i}^2 (E[\sigma_{t+k}^2|t] - \sigma^2) \quad (5.8)$$

donde al primer término coincidente con (5.7) se le añade un segundo término que refleja las diferencias entre la varianza media o varianza no condicional de las futuras innovaciones y la predicción de la varianza de las mismas condicional a la información hasta el instante t . Si el proceso es estacionario e invertible, los ψ_i^2 tenderán a cero y $E[\sigma_{t+k}^2|t]$ a σ^2 según se incremente el valor de i .

En relación al cálculo de la predicción de la varianza de las innovaciones (es decir, $E[\sigma_{t+k}^2|t]$ en 5.6), si ésta viene descrita por un proceso ARCH, el procedimiento para calcularla es similar al de la predicción de la media para un proceso ARMA. Así, el modelo ARCH puede ser conveniente descrito de la siguiente manera

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^w \alpha_i \varepsilon_{t-i}^2 + v_t \quad (5.9)$$

siendo $v_t = \varepsilon_t^2 - \hat{\sigma}_{t|t-1}^2$ un proceso serialmente incorrelado de media 0, por lo que la expresión en (5.9) se corresponde con la de un modelo AR(w) de la serie del cuadrado de los errores de predicción a un paso. De manera análoga al caso del modelo ARMA,

la predicción de la varianza de las innovaciones del proceso a horizonte k viene dada por la recurrencia

$$E[\varepsilon_{t+k}^2|t] = E[\sigma_{t+k}^2|t] = \alpha_0 + \alpha_1 E[\varepsilon_{t+k-1}^2|t] + \cdots + \alpha_w E[\varepsilon_{t+k-w}^2|t] \quad (5.10)$$

donde $E[\varepsilon_{t+i}^2|t] = \varepsilon_{t+i}^2$ cuando $i \leq 0$. De esta manera, sustituyendo 5.10 en (5.6) tenemos la expresión del error cuadrático medio asociado a las predicciones óptimas en el modelo ARMA-ARCH general.

Resumiendo lo expuesto, los momentos de primer y segundo orden de la distribuciones predictivas del modelo ARMA-ARCH pueden ser calculados utilizando las expresiones 5.2, 5.6 y 5.10. A partir de dichas estimaciones y para el caso particular de que se asuman errores homocedásticos (sin término ARCH) y gaussianos, las distribuciones predictivas o intervalos de predicción pueden ser fácilmente obtenidos puesto que también son gaussianos¹. En presencia de un término ARCH, la forma de las distribuciones de los errores de predicción en un horizonte mayor de 1 difieren de la asumida para los errores de predicción a un paso. Para el caso de las distribuciones semiparamétricas (SNP), sus expresiones explícitas no están disponibles (ver Baillie and Bollerslev, 1992; Chatfield, 1993; Hyndman et al., 2008 para más detalles y discusión).

5.2.2. Predicción con el modelo semiparamétrico de función de transferencia con ruido ARIMA-ARCH

Otra manera de calcular la predicción de la media y de la varianza de las innovaciones de un proceso descrito por el modelo FT-ARIMA-ARCH es utilizando su representación equivalente en espacio de estados expuesta en 4.2.2. El modelo así descrito responde a

$$\xi_{t+1}^* = F\xi_t^* + Bx_t + \varpi_{t+1} \quad (5.11)$$

$$y_t = H\xi_t^* + Ax_t \quad (5.12)$$

$$\xi_{t+1}^{**} = G\xi_t^{**} + D + \nu_{t+1} \quad (5.13)$$

$$\varepsilon_t^2 = J\xi_t^{**} + \alpha_0 \quad (5.14)$$

donde ξ_t^* y ξ_t^{**} son los vectores de estado, $x_t = (x_{1t}, \dots, x_{nt})'$ es el vector de variables explicativas, $\varpi_t = (\varepsilon_t, 0, \dots, 0)'$ y $\nu_t = (v_t, 0, \dots, 0)'$. Las matrices F, B, H y A dependen de los parámetros que modelan la media condicional de y_t y se construyen de la manera descrita en la sección 4.2.1.1. De igual modo partiendo

¹En caso de errores homocedásticos todas las variables de la ecuación 5.4 son normales independientes entre sí.

de (5.9) se construyen las matrices G, D y J , las cuales dependen de los parámetros que modelan la varianza condicional de y_t .

Según esta formulación, las predicciones a horizonte k , $\hat{y}_{t+k|t}$ y $\hat{\sigma}_{t+k|t}^2$, se calculan del siguiente modo

$$\hat{y}_{t+k|t} = H\hat{\xi}_{t+k|t}^* + Ax_{t+k} \quad (5.15)$$

$$\hat{\varepsilon}_{t+k|t}^2 = \hat{\sigma}_{t+k|t}^2 = J\hat{\xi}_{t+k|t}^{**} + \alpha_0 \quad (5.16)$$

donde $\hat{\xi}_{t+k|t}^*$ y $\hat{\xi}_{t+k|t}^{**}$ son las predicciones de los vectores de estado con origen de predicción en t y a horizonte k .

Puesto que $y_{t+1}, \dots, y_{t+k}$ no son observadas, los vectores de estado no pueden ser actualizados según las ecuaciones 5.11 y 5.13. Por definición $E[\varpi_{t+1}|t] = \dots = E[\varpi_{t+k}|t] = 0$ y $E[\nu_{t+1}|t] = \dots = E[\nu_{t+k}|t] = 0$, por lo que, partiendo de las estimaciones iniciales $\xi_{t+1|t}^*$ y $\xi_{t+1|t}^{**}$, la predicción de los vectores de estado se realiza recursivamente según:

$$\hat{\xi}_{t+i+1|t}^* = F\hat{\xi}_{t+i|t}^* + Bx_{t+i} \quad (5.17)$$

$$\hat{\xi}_{t+i+1|t}^{**} = G\hat{\xi}_{t+i|t}^{**} + D \quad (5.18)$$

para $i = 1, \dots, k$.

Sin embargo, la manera más sencilla de obtener las distribuciones predictivas a un horizonte mayor de 1 es simulando la densidad SNP asumida para las innovaciones del proceso y seguir las trayectorias del modelo desde el origen de predicción t hasta el instante $t + k$ deseado. Así, para cada simulación s , $s = 1, \dots, S$, se generan las observaciones $\{y_{t+1}^{(s)}, \dots, y_{t+k}^{(s)}\}$ a partir de la información de la serie original hasta t . Utilizando la representación en espacio de estados, el procedimiento resulta sencillo. Para cada simulación s y con origen de predicción en t , los pasos a seguir se resumen en los siguientes:

1. Se simula la innovación $\epsilon_{t+1}^{[s]}$ según la distribución SNP asumida por medio de la ecuación (3.2) o a través del método descrito en Gallant and Tauchen (1990).
2. Partiendo del vector de estado $\xi_t^{**[s]}$, el valor del factor de escala para la innovación $\hat{\sigma}_{t+1|t}^{[s]}$ se calcula como la raíz cuadrada de (5.19)

$$E[\varepsilon_{t+1}^{2[s]}|t] = \hat{\sigma}_{t+1|t}^{2[s]} = J\left(G\hat{\xi}_t^{**[s]} + D\right) + \alpha_0 \quad (5.19)$$

3. A partir de los dos pasos anteriores puede calcularse el término de error $\varepsilon_{t+1}^{[s]} = \hat{\sigma}_{t+1|t}^{[s]} \epsilon_{t+1}^{[s]}$, y con ellos los vectores $\varpi_{t+1}^{[s]}$ y $\nu_{t+1}^{[s]}$, puesto que $\varpi_{t+1}^{[s]} = (\varepsilon_{t+1}^{[s]}, 0, \dots, 0)'$ y $\nu_{t+1}^{[s]} = (\varepsilon_{t+1}^{[s]} - \hat{\sigma}_{t+1|t}^{[s]}, 0, \dots, 0)'$.
4. Una vez simulados $\varpi_{t+1}^{[s]}$ y $\nu_{t+1}^{[s]}$, pueden calcularse los vectores de estado $\xi_{t+1}^{*[s]}$ y $\xi_{t+1}^{**[s]}$ a través de (5.11) y (5.13), para finalmente calcular $y_{t+1}^{[s]}$ via (5.12).
5. El proceso se repite sustituyendo t por $t+1$, hasta llegar a simular el horizonte $t+k$ deseado.

Con el suficiente número de simulaciones puede estimarse entonces cualquier intervalo de predicción tomando los quantiles pertinentes de la distribución de las observaciones simuladas para un determinado horizonte de predicción $Y_{t+k|t} = \{y_{t+k}^{(1)}, \dots, y_{t+k}^{(S)}\}$. La densidad predicha puede ser estimada usando estimadores kernel (ver Silverman (1986)) aplicados a $Y_{t+k|t}$ (ver Hyndman et al. (2008)).

5.2.3. Predicción con el modelo de espacio de estados con cambio de régimen

La formulación general del modelo de espacio de estados con cambio de régimen responde a (5.20)-(5.23)

$$\xi_{t+1}^* = F_{r_{t+1}} \xi_t^* + B_{r_{t+1}} x_t + \varpi_{t+1} \quad (5.20)$$

$$y_t = H_{r_t} \xi_t^* + A_{r_t} x_t \quad (5.21)$$

$$\xi_{t+1}^{**} = G_{r_{t+1}} \xi_t^{**} + D_{r_{t+1}} + \nu_{t+1} \quad (5.22)$$

$$\varepsilon_t^2 = J_{r_t} \xi_t^{**} + \alpha_{0,r_t} \quad (5.23)$$

dónde el régimen $r_t \in \{1, \dots, R\}$ no es una variable observable, y presenta una determinada distribución de probabilidad condicionada a un vector de variables explicativas $Pr(r_t = j|Z_t)$, con $j = 1, \dots, R$ (ver sección 4.2.4).

Del mismo modo que ocurre con la predicción a un paso, el cálculo exacto del valor esperado de la distribución predictiva a un horizonte de predicción mayor de 1 no resulta un problema computacionalmente tratable y, por tanto, requiere de una aproximación (por ejemplo, a horizonte 24 sería necesario considerar R^{24} posibles trayectorias del vector de estado que describe el proceso de la media). Así, dado el modelo en (5.20)-(5.23), la predicción de la media del proceso con origen de predicción en t y a un horizonte de predicción k se calcula ponderando las predicciones condicionales a cada uno de los regímenes, las cuales se calculan del siguiente modo

$$\hat{y}_{t+k|t,j} = H_j \hat{\xi}_{t+k|t}^{*(j)} + A_j x_{t+k} \quad (5.24)$$

con $j = 1, \dots, R$, donde j denota el régimen del sistema en $t+k$, es decir, cuando $r_{t+k} = j$. Por tanto, la predicción de la media del proceso no condicional al régimen r_{t+k} será

$$\hat{y}_{t+k|t} = \sum_{j=1}^R Pr(r_{t+k} = j|Z_{t+k}) \times \hat{y}_{t+k|t,j} \quad (5.25)$$

Para poder calcular (5.24) se necesita una estimación del vector de estado futuro correspondiente a cada régimen $\hat{\xi}_{t+k|t}^{*(j)}$, el cual ha de ser aproximado. Para ello, de manera recurrente desde $i = 1$ hasta $i = k$, se realiza una inferencia sobre los vectores de estado en $t+i+1$ basada en las predicciones del instante anterior $t+i$, teniendo en cuenta los posible valores de r_{t+i+1} y r_{t+i} . De este modo, el algoritmo calcula en cada paso una batería de R^2 predicciones del vector de estado correspondiente a cada posible pareja $r_{t+i} = j, r_{t+i+1} = h$, del siguiente modo

$$\hat{\xi}_{t+i+1|t}^{*(j,h)} = F_h \hat{\xi}_{t+i|t}^{*(j)} + B_h x_t \quad (5.26)$$

y puesto que $y_{t+1}, \dots, y_{t+k}$ no son observadas, la aproximación se realiza usando la estimación del término de probabilidad del régimen del siguiente modo

$$\hat{\xi}_{t+i+1|t}^{*(h)} = \sum_{j=1}^R Pr(r_{t+i} = j|Z_{t+i}) \times \hat{\xi}_{t+i+1|t}^{*(j,h)} \quad (5.27)$$

para todos los posibles valores del régimen, $h = 1, \dots, R$.

Del mismo modo que en la sección 5.2.2, el problema del cálculo de las distribuciones predictivas del modelo se abordará igualmente por simulación. El procedimiento es similar al utilizado para el modelo FT-ARIMA-ARCH, con la diferencia que en este caso es necesario simular también la variable que determina el régimen r_t . Para cada simulación s ($s = 1, \dots, S$) y con origen de predicción en t , los pasos a seguir se resumen en los siguientes:

1. En primer lugar se genera aleatoriamente el régimen no observado $r_{t+1}^{[s]} = j$ según su distribución de probabilidad $Pr(r_{t+1} = j|Z_{t+1})$, con $j = 1, \dots, R$. Nótese que para el caso en el que algún retardo menor que k de la variable de salida esté presente en el conjunto Z_{t+1} se tomará como valor de dicho retardo la predicción del modelo para el retardo correspondiente.
2. Se simula la innovación $\epsilon_{t+1}^{[s]}$, la cual se distribuye normal estándar, $\epsilon_t \sim N(0, 1)$.
3. Partiendo del vector de estado $\xi_t^{**[s]}$, el valor del factor de escala para la innovación $\hat{\sigma}_{t+1|t}^{[s]}$ se calcula como la raíz cuadrada de (5.28)

$$E[\varepsilon_{t+1}^{[s]}|t] = \hat{\sigma}_{t+1|t}^{2[s]} = J_j \left(G_j \hat{\xi}_t^{**[s]} + D_j \right) + \alpha_{0,j} \quad (5.28)$$

4. A partir de los dos pasos anteriores puede calcularse el término de error $\varepsilon_{t+1}^{[s]} = \hat{\sigma}_{t+1|t}^{[s]} \epsilon_{t+1}^{[s]}$, y con ellos los vectores $\varpi_{t+1}^{[s]}$ y $\nu_{t+1}^{[s]}$, puesto que $\varpi_{t+1}^{[s]} = (\varepsilon_{t+1}^{[s]}, 0, \dots, 0)'$ y $\nu_{t+1}^{[s]} = (\varepsilon_{t+1}^{[s]} - \hat{\sigma}_{t+1|t}^{2[s]}, 0, \dots, 0)'$.
5. Una vez simulados $\varpi_{t+1}^{[s]}$ y $\nu_{t+1}^{[s]}$, y puesto que $r_{t+1}^{[s]} = j$, pueden calcularse los vectores de estado $\xi_{t+1}^{*[s]}$ y $\xi_{t+1}^{**[s]}$ a través de (5.20) y (5.22), para finalmente calcular $y_{t+1}^{[s]}$ via (5.21).
6. El proceso se repite sustituyendo t por $t+1$, hasta llegar a simular el horizonte $t+k$ deseado.

Con el suficiente número de simulaciones se obtiene una aproximación a la distribución predictiva a horizonte de k , $Y_{t+k|t} = \{y_{t+k}^{(1)}, \dots, y_{t+k}^{(S)}\}$.

5.3. Medidas para evaluar la calidad de las predicciones

A continuación se presentan algunas de las medidas existentes para evaluar la calidad de las predicciones y comparar el rendimiento de diferentes modelos frente a un mismo conjunto de validación. En primer lugar se exponen brevemente las medidas más utilizadas en la literatura de precios, que son las que se centran en medir el rendimiento de un método de predicción en base a su capacidad de proveer estimaciones puntuales próximas al valor realizado de la serie temporal (ver Hyndman and Koehler (2006) para una revisión de estas medidas de error). Sin embargo, existe otro tipo de medidas quizás menos conocidas que tienen como objetivo evaluar la función de densidad predictiva completa, y no solo su valor esperado (Gneiting and Raftery, 2007; Diks et al., 2011). Indudablemente, la predicción de la función de densidad de una variable aleatoria resulta más interesante puesto que proporciona la más completa medida posible de la incertidumbre asociada a una predicción puntual, por lo que las medidas capaces de evaluar su precisión resultan también más informativas.

5.3.1. Medidas de error para las predicciones puntuales

El MAPE (Error porcentual absoluto medio) es la medida más utilizada tanto en la literatura de precios como en la industria. Sin embargo, no resulta informativa cuando los precios son iguales o próximos a cero. El MAPE* presentado a continuación es una versión del mismo que evita este efecto adverso. También el MAE (Mean Absolute Error) y RMSE (Root Mean Square Error) son medidas

frecuentemente utilizadas para medir el rendimiento de un modelo de predicción. Estas medidas responden a

$$MAPE = \frac{100}{N} \sum_{t=1}^N \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (5.29)$$

$$MAPE^* = \frac{100}{N} \frac{\sum_{t=1}^N |y_t - \hat{y}_t|}{\sum_{t=1}^N |y_t|} \quad (5.30)$$

$$MAE = \frac{1}{N} \sum_{t=1}^N |y_t - \hat{y}_t| \quad (5.31)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2} \quad (5.32)$$

siendo N el número de datos considerados y $\hat{y}_t$ la predicción del precio a un determinado horizonte de predicción (ver Hyndman and Koehler (2006) para más detalles). Es común utilizar $N = 168$ para calcular el MAPE semanal (*Weekly-MAPE*) de un modelo de predicción de precios eléctricos (Misiorek et al. (2006)).

5.3.2. Medidas para evaluar las distribuciones predictivas

El creciente interés en la predicción de distribuciones probabilísticas ha llevado al desarrollo de herramientas estadísticas para evaluar su precisión (ver Gneiting and Raftery, 2007; Diks et al., 2011). Las técnicas que se han propuesto para este propósito pueden clasificarse en dos grupos. El primero de ellos se centra en evaluar la calidad de una predicción probabilística individual con respecto al proceso de generación de datos asumido. Los tests utilizados para tal fin están basados en la función de distribución acumulada de la variable aleatoria asumida, conocida como *PIT* (*Probability Integral Transform*). La aplicación de dicha transformación *PIT* sobre la distribución predicha da como resultado la distribución uniforme en el caso en que la distribución de las predicciones y la asumida coincidan (ver por ejemplo Morales et al. (2010)). Esta técnica puede ser difícil de aplicar puesto que parte del conocimiento de la función de distribución de las densidades asumidas, cuya forma analítica a menudo resulta complicada de obtener, sobre todo en un contexto de predicción multiperíodo (ver sección 5.2). Dado que se utiliza para cotejar una distribución predicha con respecto a la distribución teórica asumida, no es aplicable para la evaluación de los modelos utilizados en el caso práctico de este capítulo, puesto que las distribuciones de probabilidad predictivas están basadas precisamente en las densidades asumidas por los modelos.

El segundo grupo de tests empleados para evaluar la calidad de las densidades predictivas utilizan una puntuación numérica (o *Score*) que permite comparar

predicciones obtenidas a través de distintos métodos. Dicha puntuación está basada en la distancia relativa de la densidad predicha con la verdadera (aunque no observada) densidad y el valor realizado de la serie temporal. La predicción con mejor puntuación, y por tanto considerada ser mejor, depende de la medida usada para cuantificar la divergencia entre dos distribuciones, existiendo de este modo diferentes medidas, como son el *Quadratic Score* (QS), *Spherical Score* (SS), *Continuous Ranked Probability Score* (CRPS) y el *Logarithmic Score* (LS), las cuales responden a

$$QS_t(\hat{f}_t, y_t) = 2\hat{f}_t(y_t) - \int_{-\infty}^{\infty} \hat{f}_t(h)^2 dh \quad (5.33)$$

$$SS_t(\hat{f}_t, y_t) = \frac{\hat{f}_t(y_t)}{\left(\int_{-\infty}^{\infty} \hat{f}_t(h)^2 dh\right)^{1/2}} \quad (5.34)$$

$$CRPS_t(\hat{F}_t, y_t) = \int_{-\infty}^{\infty} (\hat{F}_t(h) - 1\{h \geq y_t\})^2 dh \quad (5.35)$$

$$LS_t(\hat{F}_t, y_t) = \ln(\hat{f}_t(y_t)) \quad (5.36)$$

donde $\hat{f}_t$ y $\hat{F}_t$ se refieren a las funciones de densidad y de distribución (acumulada) predictivas para un instante de tiempo t a un determinado horizonte de predicción, y donde la función $1\{\cdot\}$ representa la función indicador que toma valor 1 cuando la condición entre corchetes es verdad y 0 en otro caso.

En muchas aplicaciones resulta más interesante centrarse en ciertas regiones de la densidad predicha. Por ejemplo, en la cola izquierda de la distribución para aplicaciones enfocadas a la gestión de riesgos, etc. Para tal propósito, puede utilizarse una generalización del CRPS que responde a

$$CRPS_t^w(\hat{F}_t, y_t) = \int_{-\infty}^{\infty} \varsigma(h)(\hat{F}_t(h) - 1\{h \geq y_t\})^2 dh \quad (5.37)$$

donde $\varsigma(h)$ es una función de peso no negativa definida en el intervalo unidad. Por ejemplo, una elección natural de la función de peso para evaluar la cola izquierda de una distribución es $\varsigma(h) = 1\{h \leq r\}$, donde r representa un determinado valor umbral.

5.4. Análisis comparativo de modelos

Los precios del mercado diario de OMEL fueron los utilizados para testear las capacidades predictivas de los dos modelos propuestos en la tesis. El objetivo principal del test no es otro que el de investigar los beneficios que supone de cara a la predicción el asumir una función de densidad no paramétrica en el modelo de

función de transferencia con ruido ARIMA-ARCH como se describe en el capítulo 3, o extender el modelo para incluir diferentes regímenes de comportamiento como se explica en el capítulo 4. Así mismo, el test también permitirá comprobar la idoneidad de otras características incluidas en los modelos, como la inclusión de las previsiones de demanda y producción de energía eólica como variables explicativas y los términos de tipo ARCH que modelan el ruido del modelo.

El test consiste principalmente en comparar las predicciones que arrojan los modelos que incluyen densidades tipo *SNP* y diferentes regímenes de comportamiento con su contrapartida gaussiana y de un único régimen. Además se incluyen en la comparativa modelos más simples que los anteriores, con el fin de poder analizar el impacto de las características añadidas a los modelos propuestos.

De entre los modelos ensayados, el modelo ARIMA (en lo que sigue denotado como ARIMA) y el de suavizado exponencial con doble estacionalidad (SE, ver sección 2.2.2.2) son los utilizados como *benchmark* o modelos de referencia. Su uso nos permitirá cuantificar la mejora obtenida al incluir variables explicativas o utilizar modelos más complejos. Se incluyen también dos modelos de función de transferencia con ruido ARIMA en el que uno presenta únicamente la previsión de demanda como variable explicativa (*DEM*) y el otro que presenta también la previsión de producción eólica (*EOL*). Además se incluyen un modelo que presenta estas dos variables explicativas y con una estructura de tipo ARIMA-ARCH para el ruido del proceso (*ARCH*), cuatro modelos con la misma estructura que el anterior pero con densidades seminoparamétricas que presentan un grado del polinomio diferente (de 5 a 8, denotados como *SNP5* al *SNP8*), y tres modelos de cambio de régimen (con 2, 3 y 4 regímenes, *RS2*, *RS3* y *RS4*) ((ver secciones 3.3 y 4.2.4)). Con el ánimo de validar la metodología del cálculo de los intervalos de predicción para los dos modelos propuestos, se incluyó también en la comparativa los resultados de las distribuciones predichas con el modelo *ARCH* computando los intervalos de predicción mediante *bootstrapping* (*BOOT*) (Efron and Tibshirani (1993)).

Todos los modelos ensayados fueron estimados usando el mismo conjunto de estimación, el período comprendido entre el 01 de Marzo y el 31 de Agosto de 2010. El conjunto de validación es contiguo al anterior y comprende 1 año de datos, desde el 01 de Septiembre de 2010 hasta el 31 de agosto de 2011, por tanto, es lo suficientemente extenso como para extraer conclusiones generales del comportamiento de los modelos. Las predicciones se calculan de la misma manera que tienen que hacerlo los agentes del mercado, es decir, la primera hora del día se predice con un horizonte de predicción de 1, la segunda con horizonte 2 y así sucesivamente hasta la última hora del día que se predice con horizonte 24.

Las medidas utilizadas para evaluar la precisión de los modelos son las descritas en la sección 5.3. El MAPE, MAE y RMSE fueron las utilizadas para evaluar las predicciones puntuales, mientras que para evaluar las densidades predichas se utilizó el CRPS, QS, SS y LS. Del mismo modo que en Panagiotelis and Smith (2008), para el CRPS se toma el valor mediano puesto que es una medida que se ve

Modelo	MAE	RMSE	MAPE	CRPS	QS	SS	LS
ARIMA	0.4353	0.6644	9.3117	0.2102	0.2436	0.6700	-16.0507
SE	0.4360	0.6673	9.3254	—	—	—	—
DEM	0.4369	0.6666	9.3466	0.2163	0.2283	0.6675	-18.9306
EOL	0.3921	0.6075	8.3882	0.1855	0.3429	0.7216	-16.4187
ARCH	0.3872	0.6013	8.2822	0.1812	0.5188	0.7627	-6.9103
BOOT	—	—	—	0.1867	0.3861	0.7296	-3.3258
SNP5	0.3836	0.6028	8.2055	0.1775	0.5309	0.7692	-3.3713
SNP6	0.3842	0.6044	8.2186	0.1787	0.4863	0.7569	-3.9607
SNP7	0.3890	0.6109	8.3216	0.1815	0.3418	0.7212	-5.1409
SNP8	0.3822	0.6046	8.1746	0.1774	0.4666	0.7560	-2.9207
RS2	0.3873	0.6105	8.2850	0.1795	0.3659	0.7332	-5.8328
RS3	0.3816	0.6003	8.1625	0.1755	0.4867	0.7614	-5.4669
RS4	0.3769	0.5961	8.0626	0.1755	0.5170	0.7724	-6.1575

Tabla 5.1: Resultados de la predicción del período de validación (01/Sept/2010-31/Ago/2011) con todos los modelos ensayados. En negrita aparece el mejor resultado para cada medida.

afectada por outliers, mientras que para el resto se toma su valor medio. Es necesario mencionar que puesto que las densidades predichas son calculadas por simulación, no se dispone de las expresiones analíticas de las funciones de densidad y distribución que aparecen en 5.33-5.36. Los términos integrales en estas expresiones pueden aproximarse mediante métodos de integración numérica, sin embargo la evaluación de $\hat{f}_h$ en y_t requiere de una estimación de dicha función de densidad. Así, para aproximar las densidades adecuadamente se usaron $N = 2500$ simulaciones en cada muestra, y a partir de estas se uso el estimador kernel que responde a 5.38 (Parzen, 1962; Rosenblatt, 1956)

$$\hat{f}_h(y_t) = \frac{1}{Nh} \sum_{s=1}^N K\left(\frac{y_t - y_t^{(s)}}{h}\right) \quad (5.38)$$

donde $\{y_t^{(s)}\}_{s=1}^N$ son las simulaciones del modelo a un determinado horizonte de predicción, $K(\cdot)$ es la densidad normal estándar y $h = 1.06\hat{\sigma}N^{-1/5}$.

El resumen de los resultados obtenidos se presenta en la tabla 5.1. Los resultados relativos a la evaluación de las distribuciones predictivas para el modelo SE no aparecen puesto que no se implementó ninguna rutina de simulación para dicho modelo. Tampoco se muestran los resultados relativos a las predicciones puntuales del modelo BOOT dado que son los mismos que los del modelo ARCH (presenta los mismos parámetros).

Los resultados de la tabla 5.1 reflejan que los modelos propuestos en la tesis, tanto el modelo seminoparamétrico como el de cambio de régimen, mejoran en términos globales las capacidades predictivas de su contrapartida gaussiana y de un único

Modelo	IC-50	IC-60	IC-70	IC-80	IC-90	IC-95	IC-99
ARIMA	74.28	68.21	62.29	55.05	45.92	39.46	29.41
DEM	74.74	69.20	62.80	55.45	46.97	39.63	29.58
EOL	72.37	66.59	60.01	52.76	42.41	35.50	25.74
ARCH	67.57	60.07	52.47	43.68	33.34	26.54	17.10
BOOT	72.80	67.01	59.36	49.32	35.72	25.90	10.92
SNP5	67.47	60.07	52.51	43.34	31.77	22.82	9.08
SNP6	69.83	62.88	54.79	45.50	32.74	22.53	7.10
SNP7	74.55	67.84	60.76	51.62	37.39	24.68	11.20
SNP8	70.31	63.38	55.42	46.28	32.77	22.61	8.42
RS2	79.00	73.77	67.88	60.79	40.56	32.51	19.13
RS3	74.47	68.13	60.82	52.83	35.27	27.03	15.41
RS4	69.47	62.90	55.22	46.43	34.29	26.93	15.87

Tabla 5.2: Porcentaje de muestras que quedan fuera de los intervalos de predicción predichos por cada uno de los modelos a los intervalos de confianza del 50 %, 60 %, 70 %, 80 %, 90 %, 95 % y 99 % en el periodo de validación (01/Sept/2010-31/Ago/2011, 8760 muestras). En negrita el resultado más cercano al intervalo considerado.

régimen, así como las de los modelos generalmente usados en la literatura. Los modelos de cambio de régimen parecen ser los más precisos en cuanto a predicciones puntuales se refiere. Si atendemos a este criterio, es claro que el modelo RS4 es el que muestra mejor rendimiento, puesto que es el mejor modelo para todas las medidas consideradas, seguido del modelo RS3. Nótese que no fue utilizado un mayor número de regímenes en el ensayo puesto que supondría un excesivo número de parámetros. En cuanto a la predicción de las distribuciones predictivas, los modelos de los dos tipos propuestos mejoran en general las predicciones del resto de modelos. Sin embargo, no está claro cual de los dos tipos proporciona mejores estimaciones, puesto que los índices de puntuación que evalúan las distribuciones predichas son muy parejos para ambos. Además, el modelo cuyas distribuciones predictivas se han calculado mediante bootstrapping muestra un comportamiento comparable a los de tipo SNP si atendemos a la medida de verosimilitud logarítmica, por lo que conviene analizarlos en mayor detalle.

Una manera de evaluar cómo de aproximadas son las estimaciones de los intervalos de predicción de un modelo es contabilizar cuantos puntos de una muestra de datos lo suficientemente amplia quedaron fuera de un intervalo de predicción a un determinado nivel de confianza. Así, en la tabla 5.2 se muestran los resultados de dicho test sobre las 8760 muestras horarias que supone el conjunto de validación, para los niveles de confianza del 50 %, 60 %, 70 %, 80 %, 90 %, 95 % y 99 %.

A tenor de los resultados de la tabla 5.2, los modelos ensayados parecen subestimar la incertidumbre asociada a las predicciones del modelo, puesto que para todos los intervalos de confianza considerados el número de puntos de puntos fuera de dicho intervalo excede del que teóricamente debería. Las distribuciones

reales presentan por tanto colas más pesadas que las que los modelos propuestos estiman. Este resultado, aunque algo desalentador, está en línea con otros trabajos anteriormente presentados, como es el caso de Misiorek et al. (2006). El hecho de que los intervalos predichos no sean lo suficientemente amplios está relacionado con el efecto aditivo que tiene la incertidumbre de los parámetros de los modelos en la predicción, la cual no se ha contabilizado. A este respecto, existe todavía poca literatura donde se aborde dicha problemática, y cuantificar su impacto queda fuera del alcance de la tesis.

Sin embargo, en términos comparativos, los modelos de tipo SNP parecen predecir significativamente mejor que ningún otro modelo las distribuciones predictivas de los precios. Si atendemos a los intervalos más interesantes de cara al análisis de riesgos, es decir, el intervalo del 90 % y superiores, los resultados de los modelos de cambio de régimen superan también al resto de modelos de la comparativa, a excepción del que utiliza el método de *bootstrapping*.

A continuación se presenta el análisis detallado de los resultados en cuanto a predicciones puntuales se refiere, para seguidamente analizar los resultados relativos a las predicciones de la función de densidad o intervalos de predicción.

5.4.1. Predicción puntual

Algunas conclusiones importantes pueden extraerse al analizar el MAE de los modelos ensayados para cada una de las horas del día (ver tabla 5.3). En primer lugar, el rendimiento de los dos modelos univariantes, el ARIMA y el SE, muestran un rendimiento similar. Ambos modelos presentan un rendimiento muy bueno en las primeras horas del día (hora 1 y 2) que se va deteriorando con respecto al resto de modelos a medida que el horizonte de predicción aumenta (ver figura 5.1). Esto demuestra que cuanto mayor es el horizonte de predicción más relevante son las variables explicativas. Además, el rendimiento del modelo ARIMA presenta mejor comportamiento a medida que el horizonte aumenta mientras que se ve superado por el SE a horizontes pequeños, lo que valida la utilización de ambos modelos como benchmark de la comparativa, puesto que minimiza el riesgo de estar utilizando un modelo de referencia inadecuado.

La figura 5.1 muestra también el rendimiento del modelo de función de transferencia con ruido ARIMA que incluye la previsión de demanda como variable explicativa (DEM), el que incluye también la previsión de generación eólica (EOL) y el modelo de función de transferencia con ruido ARIMA-ARCH que también incluye ambas variables (ARCH). Observando el rendimiento del modelo DEM podría desprenderse que la demanda eléctrica por si sola no mejora la capacidad predictiva de los modelos ensayados. Sin embargo, el efecto de esta variable se aprecia considerablemente en períodos de festividades, motivo por el que es incluida como variable explicativa en los modelos. En la figura 5.2 se presentan dos situaciones en las que la aportación de la previsión de demanda a la predicción del precio resulta

Hora	ARIMA	SE	DEM	EOL	ARCH	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
1	0.31	0.32	0.34	0.33	0.32	0.32	0.32	0.33	0.32	0.33	0.32	0.32
2	0.42	0.40	0.43	0.43	0.42	0.42	0.42	0.43	0.42	0.42	0.41	0.41
3	0.58	0.54	0.59	0.56	0.55	0.55	0.55	0.57	0.56	0.53	0.56	0.54
4	0.63	0.59	0.63	0.59	0.59	0.58	0.59	0.61	0.59	0.57	0.59	0.58
5	0.66	0.63	0.67	0.62	0.61	0.60	0.60	0.63	0.61	0.60	0.61	0.59
6	0.63	0.60	0.63	0.58	0.57	0.57	0.57	0.59	0.58	0.56	0.57	0.55
7	0.53	0.50	0.53	0.48	0.48	0.47	0.47	0.48	0.47	0.48	0.47	0.47
8	0.44	0.42	0.43	0.39	0.38	0.38	0.38	0.38	0.38	0.40	0.39	0.38
9	0.48	0.46	0.46	0.43	0.42	0.42	0.42	0.42	0.42	0.43	0.40	0.41
10	0.41	0.38	0.39	0.37	0.35	0.35	0.35	0.35	0.35	0.37	0.35	0.34
11	0.35	0.33	0.36	0.33	0.31	0.31	0.31	0.31	0.31	0.33	0.31	0.30
12	0.35	0.34	0.34	0.29	0.28	0.29	0.29	0.28	0.28	0.29	0.28	0.27
13	0.35	0.34	0.35	0.30	0.29	0.29	0.29	0.29	0.28	0.30	0.28	0.28
14	0.36	0.36	0.36	0.30	0.29	0.29	0.29	0.29	0.30	0.29	0.28	0.28
15	0.39	0.40	0.38	0.31	0.29	0.30	0.30	0.30	0.31	0.29	0.28	0.29
16	0.43	0.45	0.43	0.35	0.33	0.34	0.34	0.33	0.34	0.32	0.32	0.32
17	0.46	0.49	0.46	0.37	0.35	0.37	0.37	0.36	0.37	0.35	0.34	0.35
18	0.43	0.45	0.42	0.35	0.33	0.34	0.34	0.34	0.34	0.34	0.33	0.33
19	0.41	0.43	0.40	0.34	0.34	0.33	0.33	0.33	0.32	0.34	0.32	0.33
20	0.42	0.45	0.43	0.37	0.39	0.37	0.37	0.38	0.36	0.39	0.38	0.38
21	0.35	0.39	0.36	0.33	0.35	0.32	0.33	0.33	0.31	0.34	0.34	0.33
22	0.42	0.46	0.44	0.41	0.43	0.39	0.40	0.41	0.38	0.42	0.42	0.41
23	0.32	0.36	0.33	0.31	0.32	0.30	0.30	0.31	0.28	0.31	0.31	0.30
24	0.31	0.36	0.33	0.29	0.30	0.29	0.29	0.29	0.27	0.30	0.30	0.29
Total	0.44		0.44	0.39	0.39	0.38	0.38	0.39	0.38	0.39	0.38	0.38

Tabla 5.3: Error Abosoluto medio (MAE) en c€/KWh de cada modelo para cada una de las horas del día.

significativa.

Sin embargo, el impacto de la previsión de generación eólica se hace patente en todas las tablas y gráficas. En la figura 5.2 se aprecia cómo el rendimiento de los modelos que la incluyen es claramente mejor, efecto que es más notorio a medida que aumenta el horizonte de predicción. La inclusión de esta variable es sin duda la aportación más significativa en lo que al rendimiento de los modelos se refiere.

En relación a la aportación de los términos ARCH con respecto a la predicción de los modelos, dicha aportación resulta claramente positiva si atendemos a las distribuciones de probabilidad obtenidas (como veremos en la siguiente sección), sin embargo existe algo de controversia en la literatura sobre el impacto que tiene sobre las predicciones puntuales. De este ensayo se desprende que las predicciones puntuales son, en general, algo mejores que las arrojadas por los modelos con errores homocedásticos, viéndose superadas por estos tan sólo en las últimas horas del día.

Una vez analizado el impacto en la predicción de las principales características añadidas a la estructura de los modelos, pasamos a examinar en más detalle el rendimiento de los modelos propuestos en la tesis. En la figura 5.3 se muestran las diferencias del MAE de los modelos propuestos con respecto al modelo EOL². para cada una de las horas del día. Así mismo, en la figura 5.4 se muestran también las diferencias del MAPE para cada semana del conjunto de validación. De las figuras se desprende que existe una mejora significativa de los errores puntuales de predicción si se utilizan modelos de cambio de régimen con respecto al modelo de referencia

²Se toma el modelo EOL como modelo de referencia para compararlo con los modelos propuesto en la tesis puesto que es un modelo frecuentemente utilizado en la literatura de precios (ver Nogales and Conejo, 2006; Conejo et al., 2005; Nogales et al., 2002; Zareipour et al., 2006b; Misiorek et al., 2006).

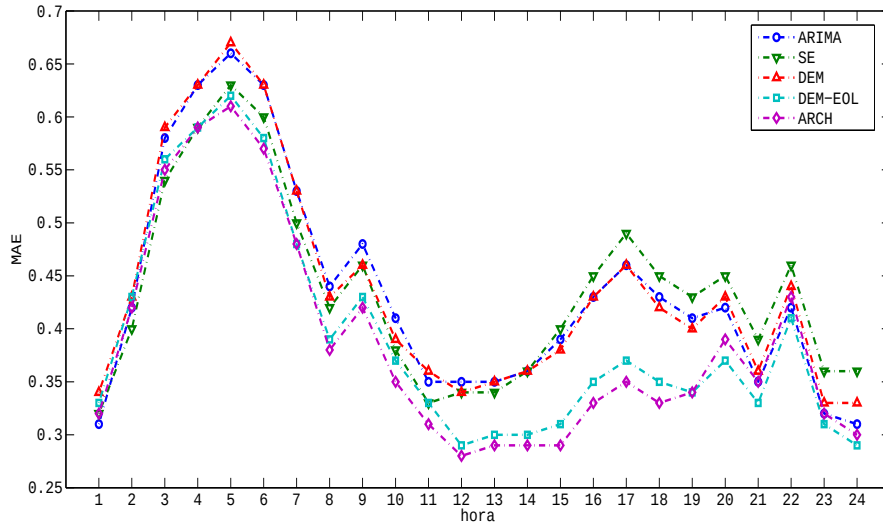
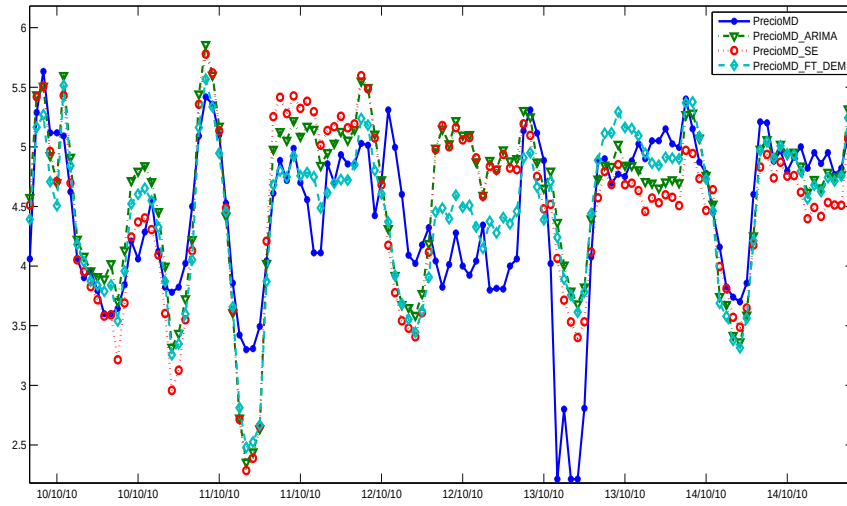


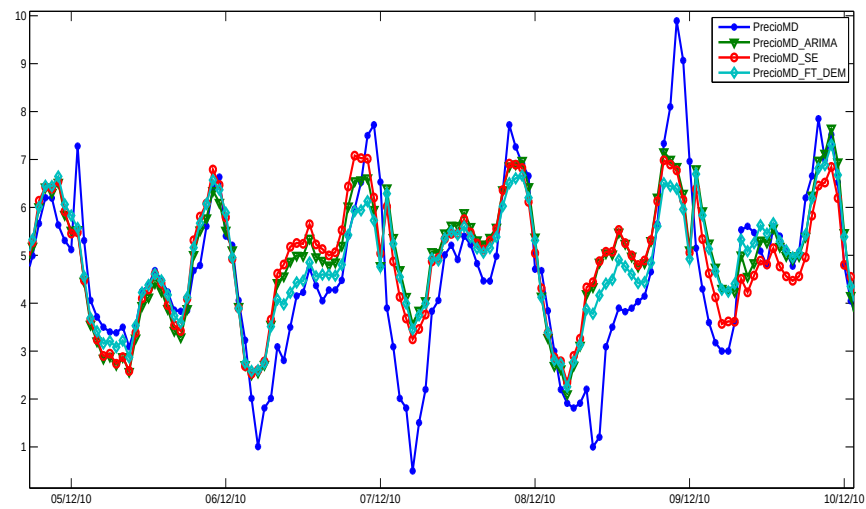
Figura 5.1: MAE (en c€/KWh) de los modelos ARIMA, SE, EOL y ARCH para cada una de las horas del día.

EOL. Los modelos RS3 y el RS4 mejoran en general las predicciones del modelo de referencia, tanto si consideramos los errores horarios como los semanales. Estos dos modelos resultan ser los mejores durante gran parte de las horas del día, y mejoran o igualan el rendimiento del modelo EOL en casi todas las semanas del período de validación. Sin embargo, de la gráfica 5.3a se desprende también que no mejoran el modelo de referencia en las cuatro últimas horas, lo que hace menos significativo su aportación global. El motivo puede encontrarse en el error a la hora de predecir el régimen a un horizonte de predicción mayor, puesto que las variables utilizadas para predecirlo, como las previsiones de demanda y producción de energía eólica, se ven afectadas también por dicho horizonte. Por último, cabe mencionar también que los resultados para el modelo RS2 ponen de manifiesto que un número insuficiente de regímenes puede tener un efecto negativo en el rendimiento del modelo.

Con respecto a los modelo de tipo SNP, la mejora obtenida en lo que a predicciones puntuales se refiere no es tan evidente. Si analizamos los errores semanales, no parece haber claras diferencias con respecto al modelo EOL. Sin embargo, si se observan si atendemos a los errores desglosados por horas. Los modelos SNP5, SNP6 y SNP8 exhiben mejor rendimiento para todas las horas del día. Además, Para las últimas horas del día el modelo SNP8 es el que presenta mejor rendimiento de todos los modelos considerados. El modelo SNP7 no muestra un rendimiento como el del resto de los modelos SNP, lo que puede ser debido a una incorrecta estimación de los parámetros, puesto que la rutina de optimización utilizada para el ajuste no asegura la obtención de un máximo de la función de verosimilitud.

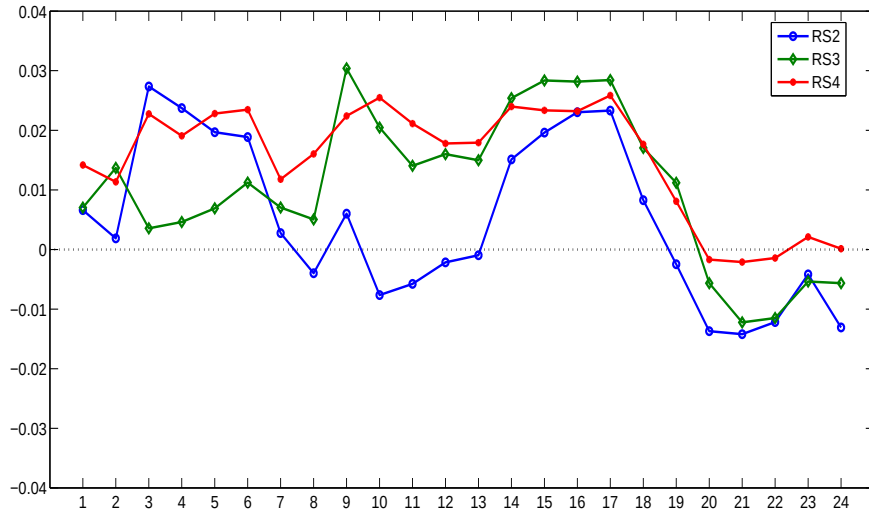


(a) 12 de Octubre, miércoles

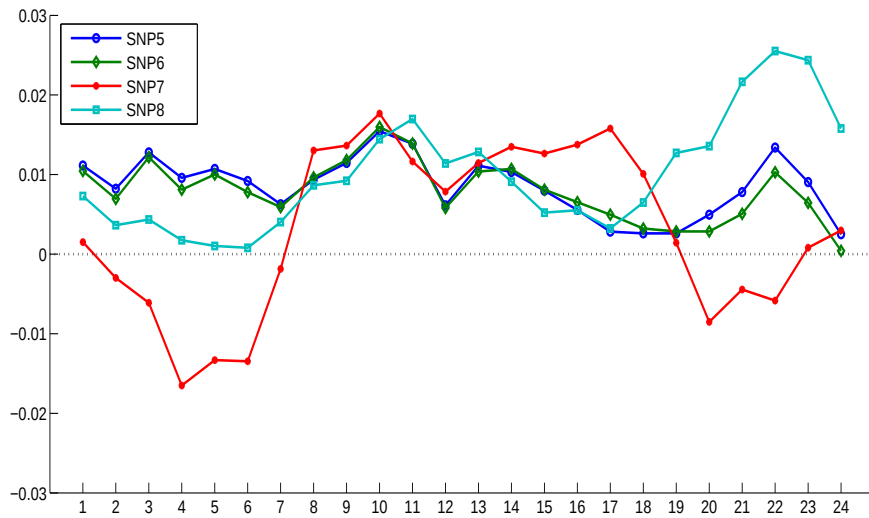


(b) 6 y 8 Diciembre, martes y jueves

Figura 5.2: Predicción de los modelos ARIMA, SE y DEM en dos períodos diferentes que incluyen festividades (Izquierda: 12 Octubre 2010, miércoles. Derecha: 6 y 8 diciembre 2010, martes y jueves).

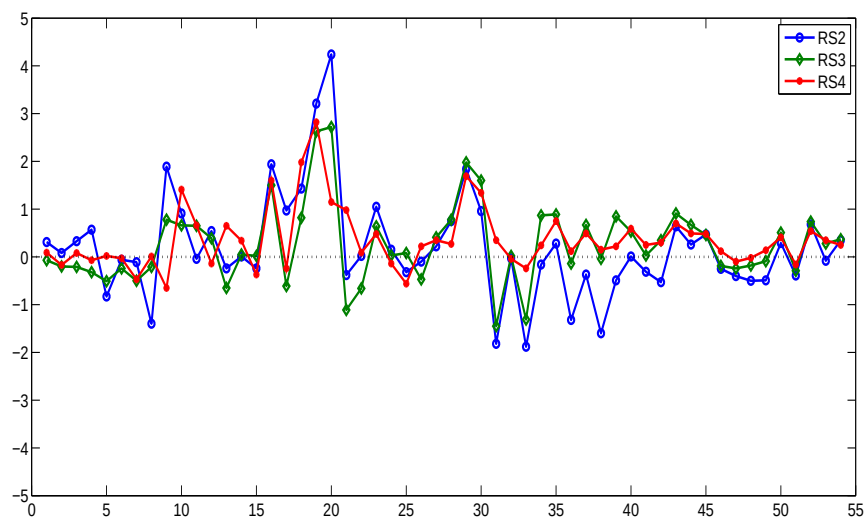


(a) Modelos de cambio de régimen

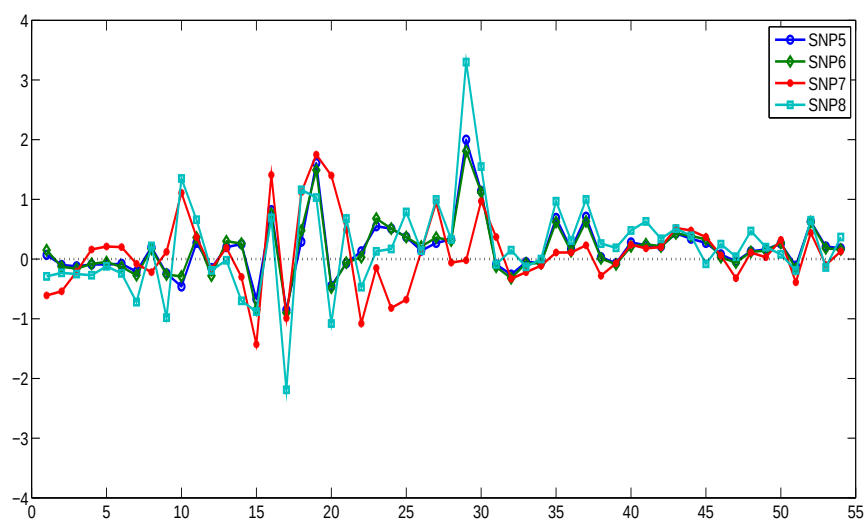


(b) Modelos con densidades SNP

Figura 5.3: Diferencias del MAE correspondiente a los modelos propuestos con respecto al modelo EOL, para cada una de las horas del día. Diferencias positivas implican que el rendimiento del modelo en cuestión supera el rendimiento del EOL.



(a) Modelos de cambio de régimen



(b) Modelos con densidades SNP

Figura 5.4: Diferencias del MAPE correspondiente a los modelos propuestos con respecto al modelo EOL, para cada una de las semanas del período de validación. Diferencias positivas implican que el rendimiento del modelo en cuestión supera el rendimiento del EOL.

Semana	ARIMA	SE	DEM	EOL	ARCH	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
2010/36	10.87	4.60	6.33	5.49	5.86	5.42	5.34	6.11	5.78	5.18	5.56	5.41
2010/37	9.14	7.23	8.82	7.24	7.48	7.34	7.36	7.78	7.47	7.16	7.44	7.40
2010/38	10.33	8.82	8.91	7.68	7.67	7.81	7.85	7.89	7.94	7.35	7.89	7.60
2010/39	7.38	5.88	6.40	5.61	5.46	5.69	5.69	5.44	5.87	5.04	5.93	5.68
2010/40	9.05	8.80	8.40	6.83	6.79	6.92	6.88	6.62	6.94	7.65	7.34	6.81
2010/41	7.22	7.07	7.14	6.26	6.14	6.35	6.39	6.07	6.50	6.33	6.50	6.30
2010/42	10.04	9.99	8.56	7.38	7.86	7.59	7.65	7.46	8.11	7.49	7.89	7.85
2010/43	5.33	5.81	5.39	4.68	5.00	4.48	4.51	4.90	4.46	6.07	4.88	4.66
2010/44	19.95	19.32	19.84	17.66	18.09	17.90	17.92	17.54	18.64	15.77	16.88	18.31
2010/45	20.95	23.15	21.22	18.45	16.91	18.91	18.74	17.34	17.10	17.54	17.80	17.05
2010/46	13.74	17.26	14.10	13.52	13.06	13.25	13.16	13.15	12.86	13.56	12.87	12.86
2010/47	13.99	14.88	13.95	12.50	12.35	12.65	12.78	12.67	12.69	11.97	12.11	12.64
2010/48	12.28	10.06	10.14	8.05	7.49	7.85	7.75	7.86	8.07	8.29	8.70	7.40
2010/49	10.06	7.81	7.72	7.82	7.28	7.56	7.55	8.11	8.52	7.81	7.77	7.47
2010/50	17.33	17.72	15.04	12.36	12.37	13.04	13.21	13.79	13.25	12.61	12.34	12.74
2010/51	11.91	13.26	11.08	10.36	9.59	9.55	9.60	8.95	9.68	8.42	8.86	8.76
2010/52	31.83	32.40	28.50	24.98	24.77	25.82	25.87	25.97	27.17	24.00	25.59	25.22
2010/53	14.70	13.37	15.61	13.20	11.64	12.91	12.72	12.08	12.04	11.77	12.38	11.21
2011/01	12.70	10.85	11.52	10.77	9.51	9.16	9.29	9.02	9.74	7.57	8.14	7.95
2011/02	35.01	34.91	34.65	29.54	27.81	29.99	30.02	28.14	30.62	25.30	26.82	28.39
2011/03	19.74	17.99	18.71	16.44	15.11	16.52	16.50	15.95	15.76	16.83	17.55	15.46
2011/04	16.23	15.64	14.98	12.47	13.54	12.34	12.45	13.55	12.94	12.46	13.14	12.38
2011/05	11.87	10.57	11.40	9.78	10.56	9.23	9.10	9.94	9.66	8.74	9.14	9.30
2011/06	6.66	5.42	6.00	5.58	7.00	5.06	5.07	6.40	5.41	5.43	5.55	5.72
2011/07	5.45	6.26	6.20	7.19	8.57	6.82	6.82	7.87	6.40	7.51	7.11	7.75
2011/08	9.19	9.29	9.46	8.80	8.77	8.66	8.59	8.66	8.66	8.90	9.26	8.58
2011/09	7.87	8.15	8.26	10.02	9.78	9.75	9.66	9.06	9.02	9.80	9.61	9.67
2011/10	6.27	6.24	6.04	5.80	5.90	5.47	5.49	5.86	5.47	5.05	5.02	5.53
2011/11	5.02	5.38	15.29	15.25	14.45	13.25	13.44	15.27	11.95	13.41	13.27	13.55
2011/12	8.07	9.74	13.09	12.94	12.29	11.80	11.82	11.97	11.39	11.98	11.34	11.60
2011/13	12.00	13.42	11.55	8.92	8.60	9.01	9.05	8.55	9.01	10.74	10.37	8.56
2011/14	8.51	10.60	9.17	7.24	7.29	7.50	7.56	7.57	7.09	7.27	7.21	7.27
2011/15	6.55	6.48	5.91	5.25	5.42	5.30	5.32	5.47	5.38	7.13	6.56	5.49
2011/16	11.88	12.63	12.43	10.45	10.26	10.50	10.52	10.56	10.45	10.61	9.58	10.21
2011/17	9.21	9.72	10.72	10.49	10.17	9.79	9.86	10.37	9.51	10.21	9.59	9.74
2011/18	6.80	6.86	5.37	5.48	5.51	5.30	5.35	5.37	5.17	6.80	5.62	5.35
2011/19	6.51	6.38	8.33	7.97	7.98	7.26	7.34	7.74	6.96	8.34	7.30	7.48
2011/20	9.05	9.59	9.10	8.34	8.42	8.31	8.33	8.61	8.08	9.94	8.38	8.19
2011/21	7.98	7.91	8.13	6.22	6.19	6.29	6.31	6.29	6.03	6.71	5.38	6.00
2011/22	5.29	5.97	5.95	5.40	5.13	5.12	5.19	5.16	4.92	5.39	4.87	4.81
2011/23	7.03	9.15	8.00	6.90	6.95	6.66	6.65	6.72	6.27	7.21	6.86	6.65
2011/24	4.67	5.36	5.16	4.76	4.60	4.55	4.55	4.57	4.42	5.29	4.41	4.46
2011/25	5.42	6.30	6.06	5.84	5.49	5.41	5.41	5.32	5.33	5.20	4.93	5.14
2011/26	4.72	5.45	5.12	5.25	5.02	4.91	4.86	4.77	4.86	4.99	4.58	4.76
2011/27	5.56	5.59	5.55	5.28	5.05	5.00	4.95	4.90	5.36	4.81	4.81	4.80
2011/28	3.10	3.22	3.15	2.90	2.85	2.82	2.87	2.84	2.65	3.15	3.09	2.78
2011/29	5.01	5.15	4.80	3.79	3.92	3.83	3.85	4.11	3.76	4.20	4.03	3.90
2011/30	4.57	5.67	4.74	5.35	5.55	5.22	5.23	5.24	4.89	5.86	5.53	5.37
2011/31	3.99	4.06	4.49	3.85	3.77	3.70	3.71	3.83	3.65	4.34	3.94	3.71
2011/32	4.83	5.08	4.83	3.69	3.41	3.43	3.42	3.37	3.61	3.40	3.18	3.28
2011/33	6.11	5.89	5.76	4.88	5.09	4.99	5.05	5.27	5.07	5.27	5.17	5.05
2011/34	3.46	3.19	5.15	5.17	4.95	4.54	4.55	4.73	4.52	4.49	4.44	4.63
2011/35	5.03	4.62	4.42	3.36	3.18	3.14	3.18	3.46	3.50	3.43	3.07	3.00
2011/36	4.06	3.40	2.93	2.89	2.73	2.70	2.73	2.76	2.52	2.54	2.52	2.64
Total	9.31	9.33	9.35	8.39	8.28	8.21	8.22	8.32	8.17	8.28	8.16	8.06

Tabla 5.4: Error Porcentual Absoluto medio semanal (MAPE semanal) de cada uno de los modelos ensayados.

5.4.2. Predicción de la función de densidad

Para evaluar el rendimiento de los modelos ensayados en lo que a predicciones probabilísticas se refiere, se utilizaron las medidas descritas previamente en la sección 5.3.2, que son: el *Logarithmic Score* (LS), *Quadratic Score* (QS), *Spherical Score* (SS) y el *Continuous Ranked Probability Score* (CRPS). Del mismo modo que para el análisis de las predicciones puntuales, los resultados fueron desglosados por hora del día y semana del período de validación, y se presentan en las tablas 5.8, 5.5, 5.6, 5.7, 5.12, 5.9, 5.10, 5.11. Algunos de los resultados de estas tablas se irán presentando de manera más clara y visual según se vayan discutiendo.

En la figura 5.5 se presentan los resultados de las medidas para los cuatro modelos utilizados como referencia (ARIMA, DEM, EOL y ARCH) para medir el rendimiento de los propuestos en la tesis, para cada una de las horas del día. También se incluye el modelo que comparte la estructura y los parámetros del modelo ARCH y cuyas distribuciones de probabilidad fueron simuladas mediante *bootstrapping* (BOOT). Así mismo, se presentan en la figura 5.6 los resultados de todas las medidas para los modelos EOL, ARCH y BOOT para cada una de las semanas del período de validación.

Con respecto a la inclusión de las variables de previsión de demanda y generación eólica, los resultados en términos comparativos son análogos a los obtenidos al validar las predicciones puntuales, lo que justifica definitivamente su inclusión. Sin embargo, el efecto de la inclusión de los términos ARCH en el modelo EOL es mucho más notorio al evaluar las distribuciones predictivas. Como puede observarse, la mejora es importante para todas las horas del día y en todas las medidas consideradas. Dicha mejora queda patente incluso en mayor medida durante las semanas de mayor volatilidad (aproximadamente de Noviembre de 2010 a Enero de 2011).

Tras un análisis preliminar, los resultados de los modelos ARCH y BOOT en términos comparativos pueden ser algo confusos. Para la medida LS el rendimiento del modelo BOOT supera al del modelo ARCH. Sin embargo, para las otras dos medidas QS y SS ocurre exactamente lo contrario. Si se analizan las expresiones de las medidas descritas 5.33-5.36, se observa que mientras que el LS tan solo tiene en cuenta el valor de verosimilitud asignado al valor realizado de la serie temporal, el QS y el SS presentan un término que penaliza la dispersión de la función de densidad. Los resultados que se observan en las gráficas significan que ateniéndose únicamente al valor de verosimilitud asignado al valor realizado de la serie de precios, es preferible el modelo BOOT, y que las densidades predichas por éste son más dispersas que las del modelo ARCH, o lo que es lo mismo, están menos concentradas. Este hecho está en consonancia con lo esperado, puesto que, como se mostró en el capítulo 3, el residuo estandarizado obtenido tras el ajuste de un modelo ARCH presenta colas más pesadas que las asumidas.

Es preciso mencionar que ninguna de estas tres medidas, LS, QS y SS, son sensibles a la distancia, es decir, no penalizan por asignar alta probabilidad a valores

Hora	ARIMA	DEM	EOL	ARCH	BOOT	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
1	-7.29	-14.90	-16.90	-4.68	-1.82	-3.25	-3.65	-3.24	-1.91	-5.09	-3.59	-3.64
2	-15.82	-20.69	-19.39	-5.33	-3.67	-3.14	-3.90	-5.33	-3.82	-6.96	-4.63	-4.76
3	-23.67	-29.97	-28.36	-9.65	-5.64	-5.64	-7.23	-7.38	-4.42	-7.21	-6.27	-8.03
4	-34.34	-38.95	-35.12	-12.23	-4.05	-4.87	-5.81	-7.74	-3.65	-5.87	-6.79	-8.79
5	-36.51	-46.09	-40.38	-15.36	-6.67	-6.42	-8.53	-9.26	-4.96	-7.28	-7.50	-9.27
6	-37.29	-43.91	-34.31	-14.41	-5.62	-7.14	-8.62	-11.08	-6.34	-8.09	-8.69	-13.94
7	-21.84	-28.82	-25.66	-8.99	-6.61	-5.69	-4.80	-8.81	-3.38	-5.05	-5.03	-9.99
8	-14.99	-16.75	-17.01	-8.08	-3.24	-2.95	-3.80	-4.49	-2.39	-3.82	-3.57	-5.42
9	-20.21	-19.67	-21.26	-11.19	-5.75	-3.51	-4.68	-8.96	-3.29	-8.28	-7.33	-8.91
10	-9.38	-11.91	-11.72	-6.11	-2.09	-4.86	-5.27	-6.13	-3.29	-5.40	-4.13	-6.09
11	-9.90	-13.32	-11.01	-4.94	-3.48	-3.22	-3.58	-3.32	-3.05	-3.95	-5.69	-5.47
12	-6.63	-9.02	-6.38	-2.16	-1.58	-1.84	-1.83	-3.47	-1.95	-2.90	-3.80	-3.18
13	-6.53	-7.77	-6.07	-4.23	-1.69	-1.64	-1.79	-3.10	-1.92	-2.87	-3.33	-3.32
14	-10.32	-9.83	-6.93	-3.57	-1.30	-1.01	-1.29	-1.49	-1.02	-2.41	-4.71	-1.96
15	-11.67	-14.25	-7.61	-4.21	-1.69	-1.08	-1.36	-2.39	-1.30	-4.84	-4.93	-2.99
16	-11.02	-15.49	-10.13	-9.54	-2.92	-3.19	-4.11	-5.34	-3.17	-8.43	-4.74	-7.03
17	-14.54	-16.43	-15.82	-9.96	-5.62	-5.19	-6.63	-6.58	-5.62	-7.53	-8.95	-9.94
18	-13.47	-14.82	-12.38	-6.22	-4.14	-3.66	-3.99	-5.99	-4.35	-6.35	-3.65	-6.96
19	-11.50	-15.15	-10.78	-3.42	-3.76	-3.30	-4.37	-3.99	-3.68	-5.07	-3.90	-5.63
20	-18.01	-14.43	-13.77	-5.82	-2.37	-2.46	-2.88	-3.21	-1.77	-5.48	-4.31	-5.54
21	-11.10	-10.38	-8.30	-3.91	-1.28	-1.62	-1.33	-2.73	-0.95	-8.53	-7.77	-5.47
22	-21.46	-22.54	-19.47	-6.55	-2.66	-3.19	-3.36	-4.61	-2.02	-9.35	-11.31	-7.02
23	-8.48	-9.57	-6.75	-3.46	-0.96	-1.34	-1.43	-3.21	-0.99	-3.82	-3.63	-2.40
24	-9.26	-9.66	-8.53	-1.82	-1.20	-0.69	-0.83	-1.54	-0.85	-5.43	-2.96	-2.04

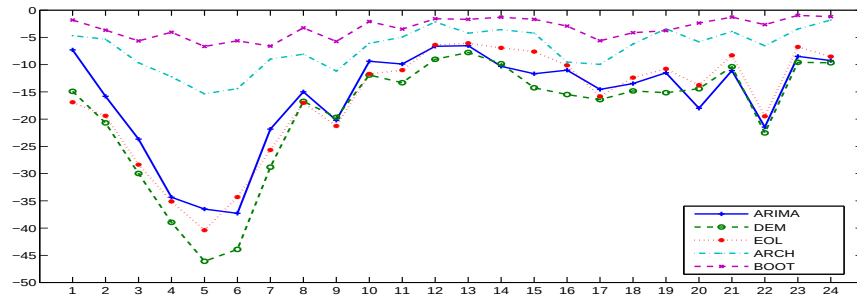
Tabla 5.5: LS medio de cada modelo ensayado para cada una de las horas del día.

Hora	ARIMA	DEM	EOL	ARCH	BOOT	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
1	0.612	0.590	0.610	0.763	0.683	0.737	0.712	0.662	0.694	0.786	0.825	0.821
2	0.203	0.223	0.193	0.396	0.265	0.358	0.300	0.160	0.282	0.219	0.484	0.412
3	-0.255	-0.223	-0.200	0.136	0.009	0.093	0.041	-0.059	0.014	0.139	0.272	0.203
4	-0.352	-0.347	-0.279	0.052	-0.116	0.050	-0.016	-0.237	-0.044	0.065	0.158	0.084
5	-0.352	-0.351	-0.262	0.054	-0.084	0.066	0.006	-0.194	-0.029	0.102	0.137	0.099
6	-0.349	-0.320	-0.254	0.050	-0.095	0.041	-0.029	-0.247	-0.050	0.086	0.113	0.080
7	-0.122	-0.112	0.012	0.264	0.133	0.263	0.218	0.053	0.172	0.283	0.309	0.289
8	0.223	0.281	0.350	0.537	0.462	0.552	0.526	0.453	0.479	0.467	0.503	0.567
9	0.237	0.271	0.292	0.522	0.401	0.535	0.496	0.348	0.487	0.465	0.611	0.576
10	0.327	0.392	0.437	0.631	0.522	0.648	0.615	0.483	0.614	0.494	0.634	0.654
11	0.507	0.479	0.595	0.699	0.562	0.723	0.695	0.539	0.671	0.528	0.657	0.717
12	0.457	0.447	0.653	0.767	0.657	0.758	0.714	0.546	0.699	0.503	0.694	0.735
13	0.454	0.467	0.654	0.761	0.649	0.766	0.741	0.600	0.702	0.477	0.655	0.729
14	0.432	0.379	0.586	0.741	0.603	0.727	0.689	0.498	0.619	0.447	0.650	0.686
15	0.318	0.308	0.535	0.678	0.518	0.666	0.613	0.433	0.533	0.485	0.597	0.634
16	0.253	0.186	0.422	0.547	0.371	0.560	0.510	0.328	0.449	0.320	0.402	0.486
17	0.161	0.087	0.370	0.474	0.312	0.447	0.396	0.237	0.354	0.151	0.297	0.410
18	0.225	0.142	0.445	0.572	0.404	0.530	0.482	0.351	0.467	0.276	0.374	0.486
19	0.251	0.259	0.452	0.539	0.366	0.576	0.520	0.389	0.543	0.344	0.437	0.526
20	0.272	0.257	0.401	0.532	0.378	0.608	0.560	0.430	0.539	0.338	0.452	0.502
21	0.581	0.483	0.557	0.650	0.531	0.757	0.715	0.565	0.709	0.451	0.579	0.645
22	0.559	0.430	0.452	0.624	0.522	0.695	0.656	0.545	0.702	0.366	0.489	0.593
23	0.641	0.629	0.607	0.725	0.594	0.822	0.792	0.661	0.837	0.540	0.646	0.729
24	0.561	0.523	0.602	0.736	0.620	0.765	0.720	0.660	0.753	0.449	0.707	0.744

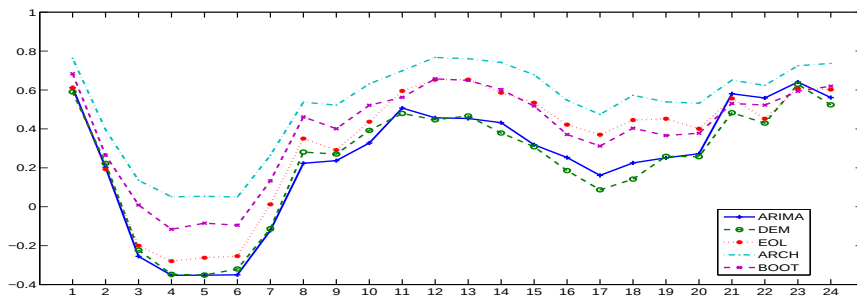
Tabla 5.6: QS medio de cada modelo ensayado para cada una de las horas del día.

Hora	ARIMA	DEM	EOL	ARCH	BOOT	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
1	0.834	0.827	0.839	0.879	0.864	0.876	0.872	0.850	0.864	0.890	0.894	0.898
2	0.652	0.665	0.656	0.703	0.672	0.697	0.682	0.640	0.674	0.647	0.722	0.704
3	0.448	0.469	0.485	0.554	0.527	0.548	0.536	0.507	0.526	0.573	0.589	0.580
4	0.404	0.413	0.449	0.508	0.467	0.514	0.497	0.439	0.487	0.526	0.532	0.523
5	0.405	0.412	0.457	0.514	0.486	0.525	0.512	0.462	0.501	0.544	0.532	0.534
6	0.406	0.426	0.461	0.522	0.493	0.519	0.502	0.456	0.500	0.549	0.534	0.537
7	0.507	0.517	0.577	0.629	0.599	0.630	0.620	0.583	0.607	0.640	0.640	0.638
8	0.661	0.691	0.724	0.770	0.756	0.774	0.769	0.766	0.756	0.727	0.739	0.778
9	0.667	0.686	0.699	0.757	0.723	0.761	0.750	0.717	0.750	0.730	0.792	0.779
10	0.708	0.740	0.763	0.819	0.789	0.824	0.814	0.784	0.818	0.757	0.816	0.829
11	0.787	0.778	0.832	0.857	0.816	0.865	0.857	0.816	0.855	0.797	0.846	0.871
12	0.765	0.764	0.857	0.891	0.861	0.883	0.869	0.825	0.870	0.810	0.879	0.888
13	0.764	0.773	0.857	0.888	0.859	0.888	0.881	0.847	0.873	0.814	0.868	0.890
14	0.754	0.734	0.828	0.883	0.845	0.873	0.861	0.810	0.842	0.812	0.875	0.879
15	0.704	0.703	0.805	0.856	0.813	0.847	0.830	0.785	0.808	0.834	0.859	0.858
16	0.674	0.649	0.756	0.801	0.755	0.801	0.788	0.743	0.775	0.772	0.786	0.801
17	0.633	0.606	0.733	0.768	0.732	0.752	0.739	0.707	0.735	0.709	0.747	0.767
18	0.662	0.629	0.766	0.809	0.766	0.787	0.775	0.749	0.783	0.748	0.768	0.798
19	0.673	0.681	0.770	0.787	0.740	0.802	0.785	0.755	0.806	0.764	0.780	0.802
20	0.683	0.680	0.748	0.768	0.727	0.805	0.790	0.755	0.792	0.748	0.768	0.776
21	0.820	0.780	0.815	0.824	0.794	0.874	0.859	0.816	0.866	0.805	0.823	0.844
22	0.810	0.756	0.769	0.797	0.773	0.835	0.825	0.792	0.850	0.763	0.776	0.811
23	0.847	0.844	0.837	0.856	0.818	0.903	0.892	0.852	0.920	0.838	0.848	0.876
24	0.811	0.798	0.835	0.865	0.832	0.878	0.862	0.854	0.886	0.800	0.861	0.875

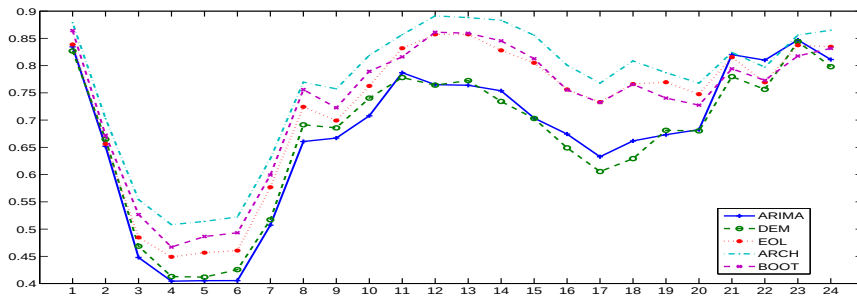
Tabla 5.7: SS medio de cada modelo ensayado para cada una de las horas del día.



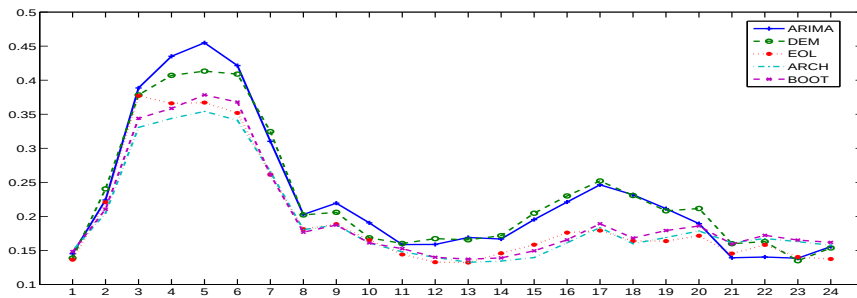
(a) LS



(b) QS

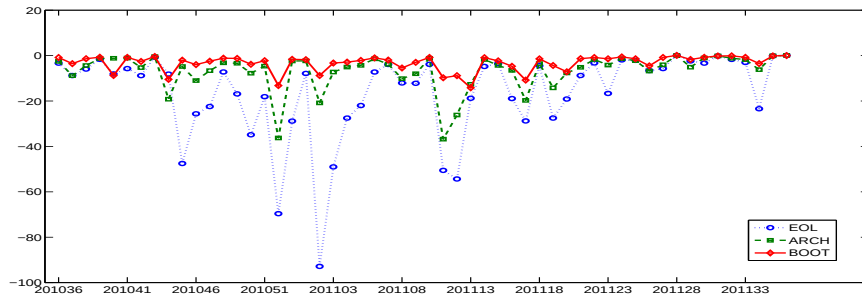


(c) SS

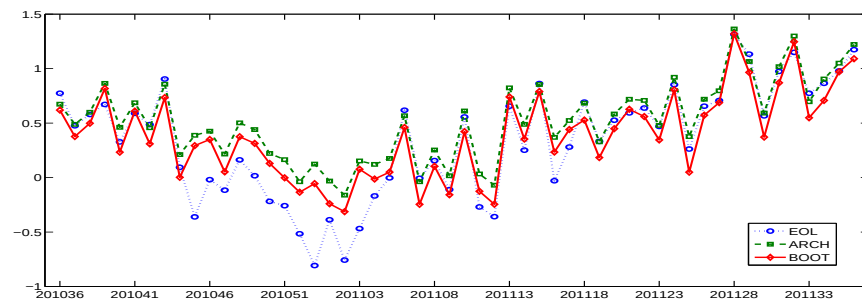


(d) CRPS

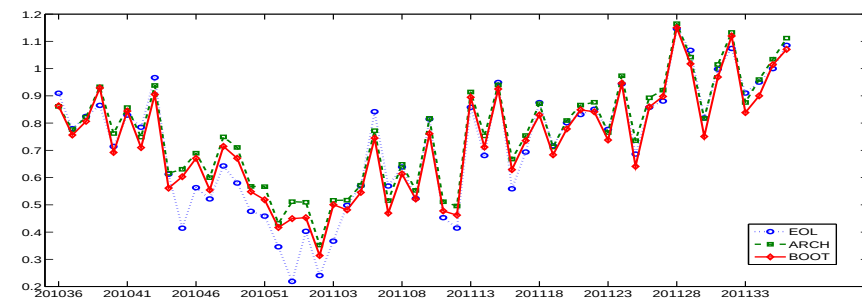
Figura 5.5: Resultados de las medidas para la evaluación de las distribuciones predictivas correspondientes a los modelos ARIMA, DEM, EOL, ARCH y BOOT para cada hora del día. Recuerdese que un modelo con mayor puntuación implica un LS, QS, SS más alto y un CRPS más bajo.



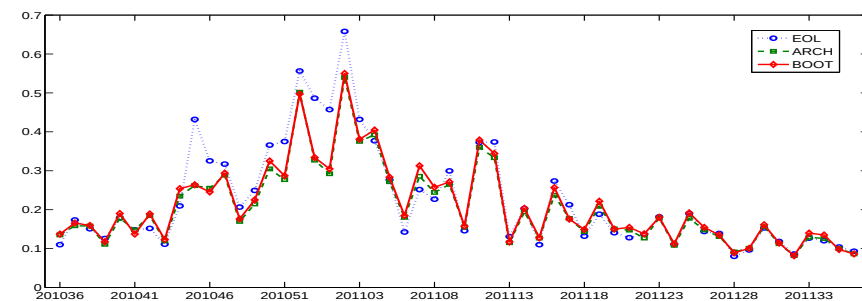
(a) LS



(b) QS



(c) SS



(d) CRPS

Figura 5.6: Resultados de las medidas para la evaluación de las distribuciones predictivas correspondientes a los modelos EOL, ARCH y BOOT para cada semana del período de validación. Recuerdese que un modelo con mayor puntuación implica un LS, QS, SS más alto y un CRPS más bajo.

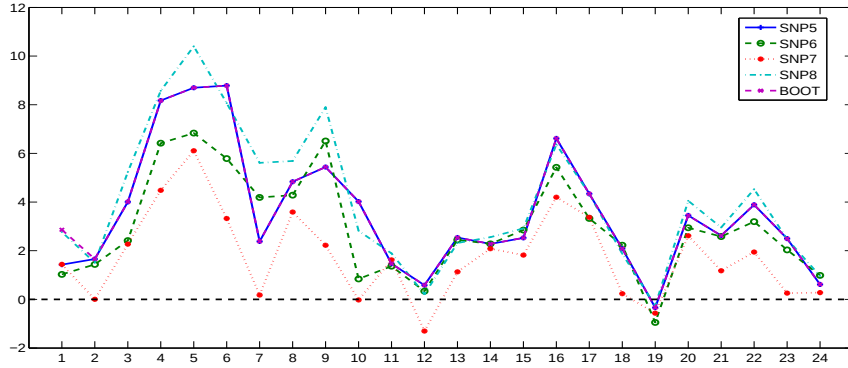
Hora	ARIMA	DEM	EOL	ARCH	BOOT	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
1	0.145	0.139	0.136	0.150	0.148	0.143	0.151	0.151	0.156	0.148	0.147	0.146
2	0.224	0.241	0.221	0.205	0.211	0.215	0.219	0.227	0.226	0.209	0.200	0.195
3	0.389	0.378	0.377	0.331	0.344	0.308	0.306	0.314	0.320	0.287	0.286	0.283
4	0.435	0.407	0.366	0.344	0.359	0.347	0.357	0.402	0.353	0.325	0.332	0.352
5	0.455	0.413	0.367	0.354	0.378	0.329	0.338	0.402	0.342	0.319	0.345	0.336
6	0.422	0.409	0.352	0.341	0.368	0.345	0.350	0.388	0.380	0.324	0.330	0.346
7	0.310	0.325	0.262	0.266	0.261	0.245	0.236	0.262	0.257	0.226	0.231	0.253
8	0.203	0.202	0.181	0.181	0.177	0.179	0.178	0.166	0.178	0.188	0.186	0.167
9	0.220	0.206	0.189	0.187	0.187	0.189	0.193	0.178	0.187	0.192	0.160	0.175
10	0.191	0.169	0.165	0.161	0.161	0.162	0.165	0.154	0.153	0.177	0.157	0.151
11	0.159	0.160	0.144	0.148	0.153	0.143	0.141	0.141	0.141	0.153	0.144	0.142
12	0.159	0.167	0.133	0.140	0.140	0.141	0.141	0.138	0.139	0.141	0.132	0.136
13	0.169	0.166	0.132	0.133	0.137	0.130	0.130	0.137	0.139	0.144	0.131	0.135
14	0.167	0.172	0.146	0.134	0.139	0.143	0.148	0.152	0.149	0.145	0.138	0.135
15	0.195	0.205	0.159	0.140	0.149	0.143	0.140	0.153	0.162	0.138	0.124	0.132
16	0.221	0.230	0.176	0.161	0.166	0.167	0.158	0.168	0.163	0.163	0.155	0.146
17	0.246	0.252	0.179	0.184	0.189	0.170	0.173	0.189	0.193	0.169	0.159	0.170
18	0.232	0.231	0.164	0.160	0.168	0.160	0.163	0.159	0.154	0.164	0.159	0.161
19	0.212	0.208	0.164	0.169	0.179	0.161	0.158	0.161	0.147	0.168	0.154	0.169
20	0.190	0.212	0.172	0.179	0.186	0.168	0.174	0.176	0.166	0.174	0.168	0.174
21	0.139	0.160	0.145	0.161	0.160	0.147	0.149	0.153	0.139	0.152	0.156	0.154
22	0.140	0.163	0.158	0.168	0.172	0.150	0.155	0.153	0.147	0.164	0.175	0.151
23	0.139	0.135	0.140	0.163	0.165	0.148	0.147	0.145	0.131	0.140	0.151	0.151
24	0.156	0.154	0.137	0.158	0.162	0.153	0.153	0.152	0.137	0.145	0.149	0.150

Tabla 5.8: Mediana del CRPS de cada modelo ensayado para cada una de las horas del día.

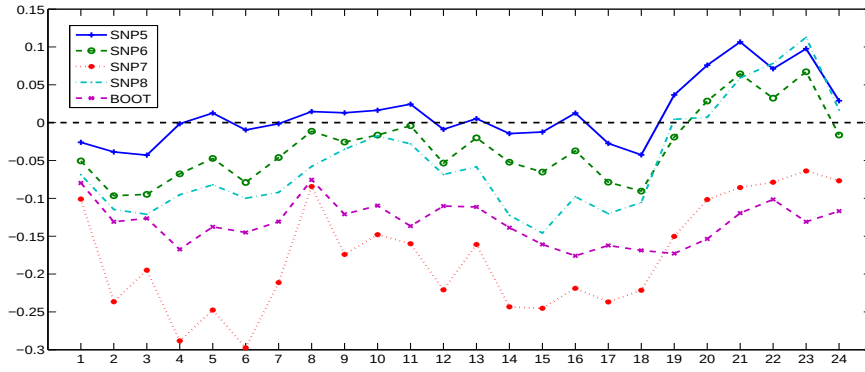
cercanos pero no iguales al valor realizado. El CRPS sí tiene en cuenta esta distancia, por lo que es quizás la medida más fiable de entre las consideradas. La gráfica 5.6d muestra que, bajo esta puntuación, existen pocas diferencias entre el modelo ARCH y BOOT (el modelo ARCH es algo mejor), y las diferencias son mucho más grandes en cambio entre estos dos y el modelo EOL, sobre todo en el período de alta volatilidad. En la gráfica 5.5d puede observarse también que el hecho de que se penalice la distancia de la distribución al valor realizado de la serie de precios hace que en las últimas horas del día el modelo EOL sea el que mejor puntuación recibe de entre los tres, puesto que como ya vimos es el que mejor predice la esperanza de la distribución. Este hecho hace pensar que los términos ARCH añadidos a los modelos resultan algo perjudiciales en los horizontes más altos.

Una vez validada la estructura del modelo ARCH, lo usaremos como modelo de referencia para analizar los resultados obtenidos con los propuestos en la tesis, e incluiremos también en la comparativa los resultados del modelo BOOT. Las figuras 5.7 y 5.8 muestran la diferencia del LS, QS y CRPS obtenido con los modelos con densidades tipo SNP y de cambio de régimen, respectivamente, con respecto al modelo de referencia ARCH y para cada una de las horas del día. Las figuras 5.9 y 5.10 muestran los mismos resultados desglosados por semana del conjunto de validación. Los resultados para el SS llevan a las mismas conclusiones que el QS por lo que no son presentados en las gráficas.

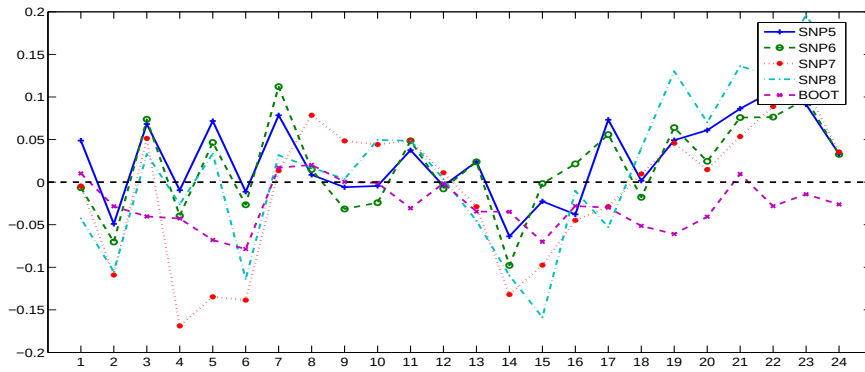
Si analizamos los resultados obtenidos para los modelos de tipo SNP, puede apreciarse cómo estos modelos asignan en general más verosimilitud que el modelo de referencia ARCH, y valores muy parecidos en comparación con el método de *bootstrapping*. Sin embargo, con respecto a la medida QS, los resultados que arrojan son peores que los del modelo de referencia, lo que implica que las distribuciones



(a) LS

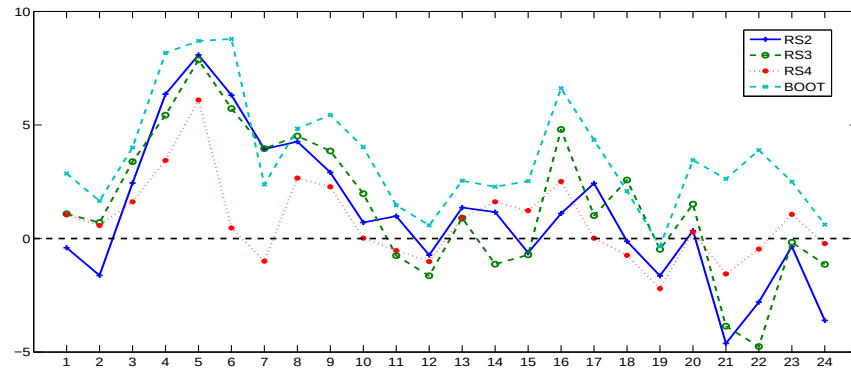


(b) QS

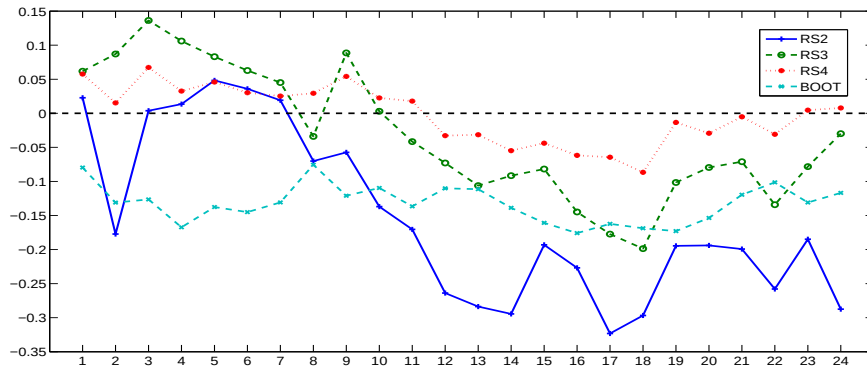


(c) CRPS

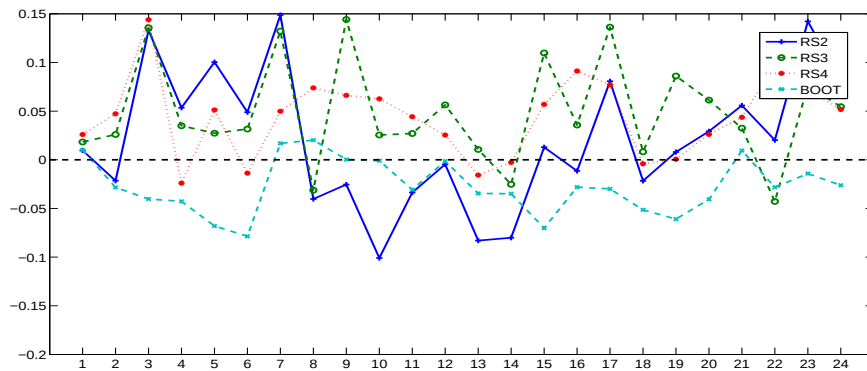
Figura 5.7: Resultados de las medidas para la evaluación de las distribuciones predictivas correspondientes a los modelos con densidades SNP y al modelo BOOT para cada hora del período de validación. Las medidas se expresan como diferencias con respecto al modelo ARCH. Diferencias positivas implican que el rendimiento del modelo en cuestión supera el rendimiento del ARCH. Las diferencias del CRPS están expresadas en porcentaje.



(a) LS

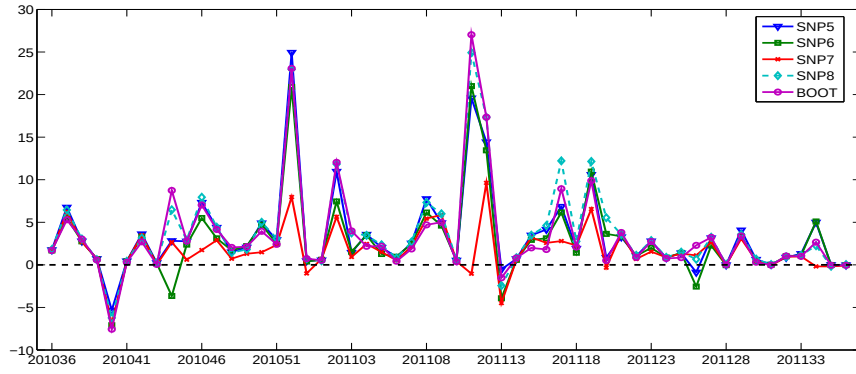


(b) QS

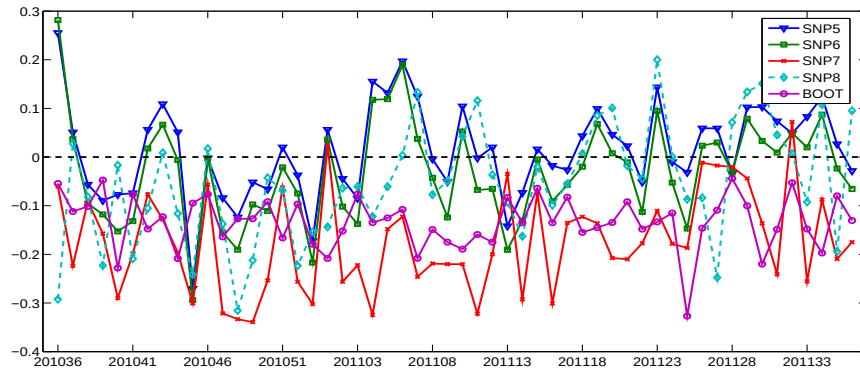


(c) CRPS

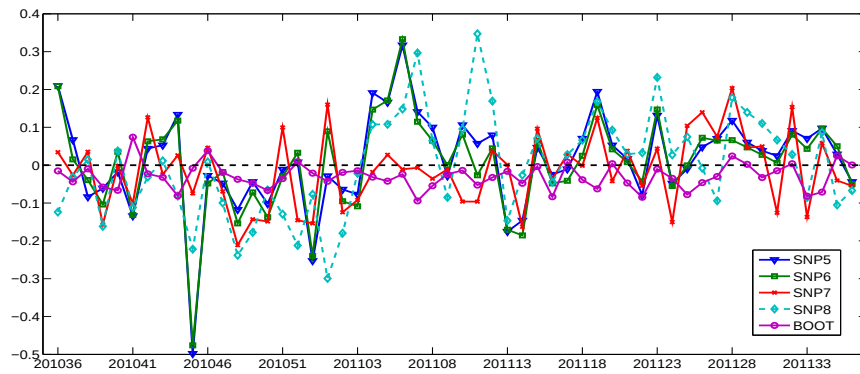
Figura 5.8: Resultados de las medidas para la evaluación de las distribuciones predictivas correspondientes a los modelos de cambio de régimen y al modelo BOOT para cada hora del período de validación. Las medidas se expresan como diferencias con respecto al modelo ARCH. Diferencias positivas implican que el rendimiento del modelo en cuestión supera el rendimiento del ARCH. Las diferencias del CRPS están expresadas en porcentaje.



(a) LS

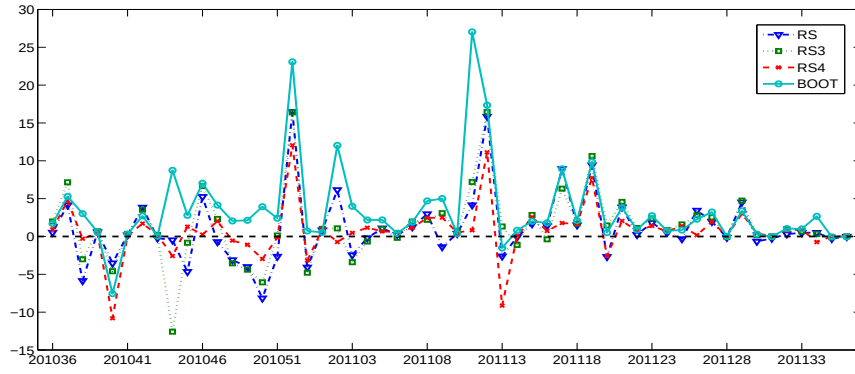


(b) QS

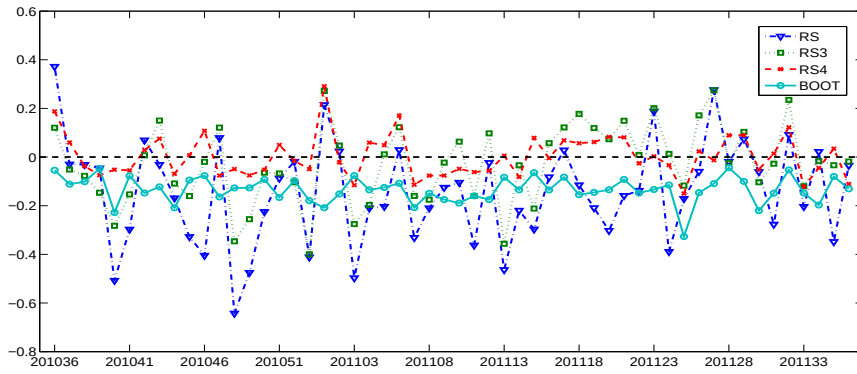


(c) CRPS

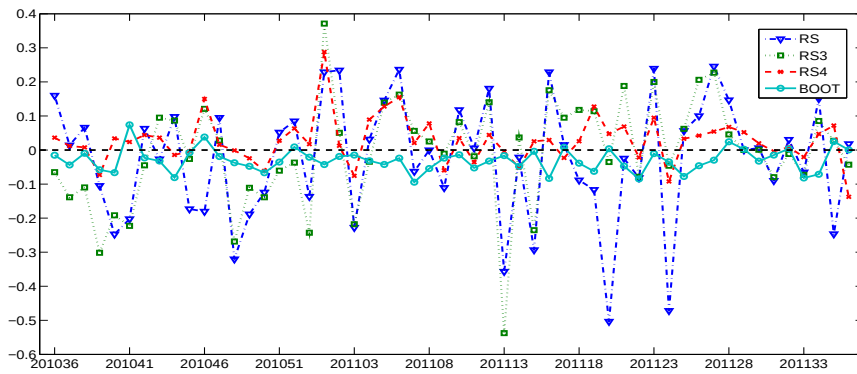
Figura 5.9: Resultados de las medidas para la evaluación de las distribuciones predictivas correspondientes a los modelos con densidades SNP y al modelo BOOT para cada semana del período de validación. Las medidas se expresan como diferencias con respecto al modelo ARCH. Diferencias positivas implican que el rendimiento del modelo en cuestión supera el rendimiento del ARCH. Las diferencias del CRPS están expresadas en porcentaje.



(a) LS



(b) QS



(c) CRPS

Figura 5.10: Resultados de las medidas para la evaluación de las distribuciones predictivas correspondientes a los modelos de cambio de régimen y al modelo BOOT para cada semana del período de validación. Las medidas se expresan como diferencias con respecto al modelo ARCH. Diferencias positivas implican que el rendimiento del modelo en cuestión supera el rendimiento del ARCH. Las diferencias del CRPS están expresadas en porcentaje.

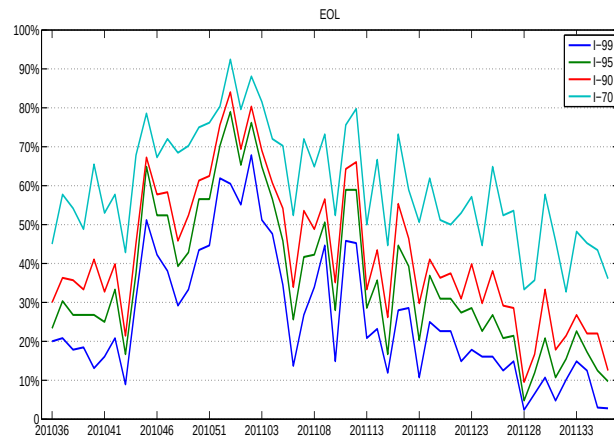
predichas son más dispersas. Esto es así para casi todas las horas del día, a excepción de las últimas en donde estos modelos no solo son más precisos en cuanto a predicción puntual se refiere (como se mostró en la sección anterior) sino que las densidades predichas están más concentradas en torno al valor realizado, lo que es corroborado también por los resultados del CRPS. Si descartamos el modelo SNP7 (por estar mal especificado, como ya se discutió en la sección anterior), y ateniéndonos al CRPS, no se aprecian grandes mejoras de los modelos SNP con respecto al modelo ARCH, pero sí frente al BOOT. Puesto que las predicciones puntuales de ambos (ARCH y BOOT) son las mismas, esto implica que las densidades predichas por los modelos SNP están más centradas en torno al valor realizado del precio que las proporcionadas por el método de *bootstrapping*.

En relación a los modelos de cambio de régimen, en general exhiben un mejor comportamiento que el modelo de referencia ARCH, y que los que presentan densidades tipo SNP. Si atendemos a los valores obtenidos de verosimilitud logarítmica, estos son peores que los del modelo BOOT aunque siempre mejores que los del modelo ARCH salvo para las últimas horas del día. Sin embargo, si atendemos al CRPS, los modelos de cambio de régimen superan claramente a los modelos ARCH y BOOT, y a los modelos con densidades tipo SNP lo largo de la mayoría de horas del día. El rendimiento del modelo RS4 es igual o mejor que el del modelo ARCH en todas las horas del día, y siempre mejor que el BOOT. El modelo RS3 es también mejor que ambos (ARCH y BOOT) en prácticamente la totalidad de las horas del día.

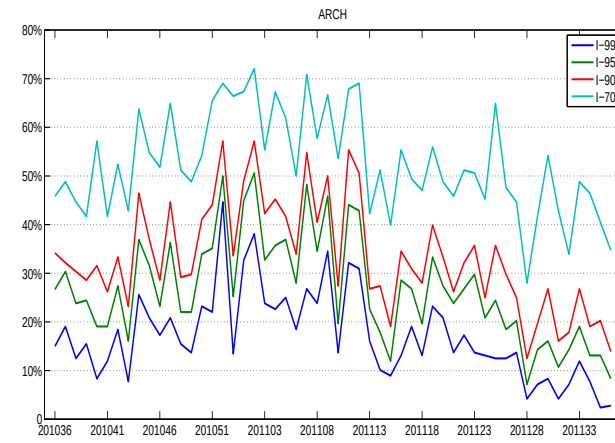
Como hemos visto, la medición de la bondad de las predicciones arrojadas por los modelos ensayados depende en parte de cuál sea la medida seleccionada para evaluarlos. Las medidas utilizadas están basadas en puntuar de diferentes formas cómo de verosímiles son las muestras realizadas para un determinado modelo y cómo de concentrada está la densidad predicha en torno a dicha muestra realizada. Puesto que sólo existe una realización de la serie de precios en cada instante de tiempo, la evaluación está sujeta al criterio subjetivo que penaliza más o menos estas características. Un criterio alternativo para evaluar la calidad de las distribuciones predictivas es contabilizar cuántos puntos de una muestra de datos lo suficientemente amplia quedaron fuera de un intervalo de predicción a un determinado nivel de confianza. El intervalo provisto será más o menos aproximado en función de lo aproximado que esté el porcentaje de puntos de todo el periodo de validación que caen dentro de dicho intervalo. Este ensayo se llevó a cabo para el período de validación bajo estudio. Así, en las tablas 5.13, 5.14, 5.15 y 5.16 se muestra el detalle de los resultados del test para los intervalos del 99 %, 95 %, 90 % y 70 %, respectivamente, y para cada una de las semanas del período de validación. En estas tablas puede apreciarse cómo los modelos de tipo SNP parecen ser los que mejor aproximan los intervalos de predicción testeados. En el caso del intervalo del 99 %, el modelo SNP6 es el mejor en 36 de las 54 semanas, el SNP5 en 14 y el SNP8 en 11. Para los intervalos del 90 % y 95 % los modelos de tipo SNP también son los dominadores.

En la figura 5.11 se representa gráficamente el resultado del test para los modelos

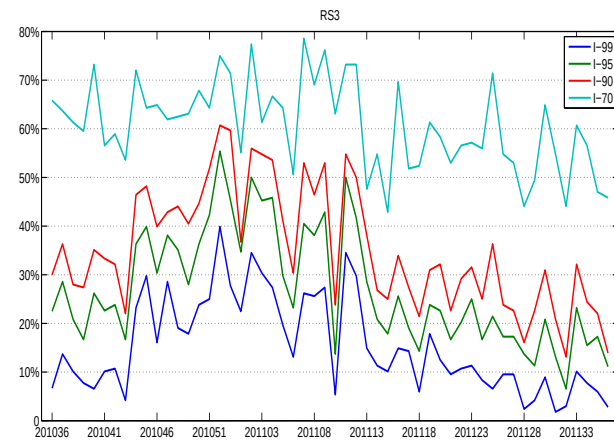
EOL, ARCH, RS3 y SNP6. En ella puede apreciarse la mejora inducida por la inclusión de los términos ARCH en el modelo de función de transferencia con ruido ARIMA. Se observa también cómo las colas de la distribución predicha del modelo SNP6 son más anchas que las del resto de modelos, y que el número de excedentes o muestras que caen fuera de los intervalos de confianza de los modelos RS3 y SNP6 presenta menos variabilidad en función de la semana analizada que el modelo ARCH o el EOL.



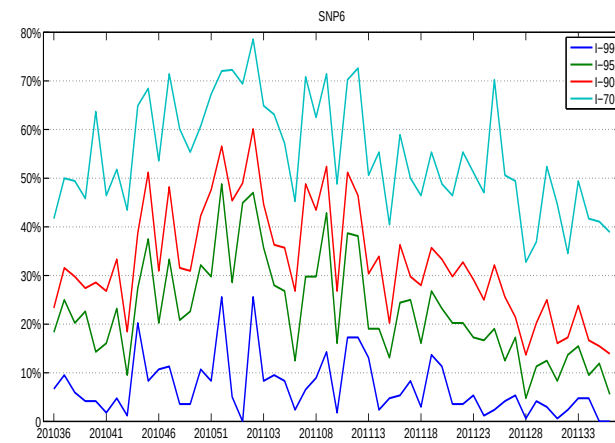
(a) EOL



(b) ARCH



(c) RS3



(d) SNP6

Figura 5.11: Porcentaje de excedentes o muestras que caen fuera de los intervalos de predicción predichos a los intervalos de confianza del 70 %, 90 %, 95 % y 99 % para los modelos EOL, ARCH, RS3 y SNP6.

Semana	ARIMA	DEM	DEM EOL	ARCH	BOOT	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
2010/36	-5.207	-2.593	-3.330	-2.547	-0.853	-0.796	-0.859	-0.951	-0.746	-2.027	-0.573	-1.608
2010/37	-18.333	-17.230	-8.808	-8.855	-3.565	-2.109	-3.347	-2.636	-2.412	-4.710	-1.706	-4.106
2010/38	-6.801	-7.992	-5.944	-4.398	-1.383	-1.628	-1.646	-1.714	-1.363	-10.261	-7.408	-4.692
2010/39	-6.673	-3.224	-1.664	-1.304	-0.730	-0.598	-0.624	-0.598	-0.691	-0.715	-0.630	-0.887
2010/40	-6.810	-5.495	-8.104	-1.213	-8.775	-6.545	-8.306	-8.160	-7.451	-4.741	-5.808	-12.025
2010/41	-5.664	-7.300	-5.741	-1.113	-0.808	-0.672	-0.692	-0.754	-0.688	-0.937	-0.776	-1.025
2010/42	-19.825	-14.302	-8.817	-5.286	-2.578	-1.670	-1.976	-1.809	-2.134	-1.456	-1.750	-3.578
2010/43	-1.027	-0.941	-0.741	-0.455	-0.411	-0.222	-0.273	-0.440	-0.233	-0.670	-0.262	-0.245
2010/44	-12.538	-11.035	-8.187	-19.215	-10.482	-16.400	-22.876	-16.551	-12.756	-19.695	-31.796	-21.797
2010/45	-28.343	-62.047	-47.514	-4.818	-2.026	-2.052	-2.442	-4.213	-1.914	-9.485	-5.657	-3.521
2010/46	-28.621	-31.687	-25.654	-11.018	-3.989	-3.701	-5.523	-9.285	-3.077	-5.767	-4.366	-10.773
2010/47	-27.981	-27.193	-22.435	-6.637	-2.488	-2.254	-3.527	-3.752	-2.218	-7.357	-4.332	-4.624
2010/48	-20.962	-13.725	-7.213	-3.190	-1.144	-1.688	-1.494	-2.474	-1.756	-6.302	-6.738	-3.743
2010/49	-26.063	-11.901	-16.888	-3.380	-1.245	-1.328	-1.181	-2.085	-1.583	-7.396	-7.765	-4.465
2010/50	-47.729	-46.803	-34.892	-7.765	-3.831	-2.868	-3.040	-6.281	-2.771	-15.927	-13.824	-10.735
2010/51	-33.428	-23.562	-18.110	-4.676	-2.268	-1.744	-2.201	-2.346	-1.561	-7.324	-4.555	-4.839
2010/52	-63.669	-58.857	-69.613	-36.295	-13.228	-11.357	-15.403	-28.332	-13.343	-19.827	-19.906	-24.241
2010/53	-36.479	-36.829	-28.865	-2.377	-1.650	-1.917	-1.987	-3.389	-1.662	-6.461	-7.170	-5.534
2011/01	-21.708	-12.440	-7.867	-2.333	-1.800	-1.783	-1.741	-1.644	-1.853	-1.496	-1.335	-1.385
2011/02	-81.514	-107.535	-92.833	-20.835	-8.798	-9.876	-13.396	-15.194	-9.009	-14.689	-19.767	-21.576
2011/03	-65.871	-55.097	-49.021	-7.257	-3.274	-5.817	-5.722	-6.336	-3.510	-9.703	-10.664	-6.787
2011/04	-32.779	-33.353	-27.532	-5.068	-2.894	-1.501	-1.580	-2.616	-1.610	-5.269	-5.741	-3.906
2011/05	-29.081	-27.229	-22.019	-4.319	-2.133	-2.293	-3.045	-2.815	-1.962	-3.240	-3.321	-3.647
2011/06	-9.184	-8.680	-7.244	-1.442	-1.023	-0.506	-0.490	-0.919	-0.653	-1.193	-1.619	-0.865
2011/07	-2.164	-2.799	-3.925	-3.837	-1.972	-1.303	-1.353	-1.683	-1.146	-2.586	-1.873	-2.709
2011/08	-11.327	-10.832	-12.056	-10.164	-5.476	-2.416	-4.025	-4.737	-2.841	-7.204	-7.972	-7.719
2011/09	-7.409	-6.102	-12.204	-7.990	-2.984	-3.074	-3.384	-2.130	-2.000	-9.377	-4.915	-5.491
2011/10	-3.481	-3.986	-1.203	-3.986	-1.203	-0.675	-0.683	-0.790	-0.622	-0.835	-0.470	-0.687
2011/11	-1.382	-60.656	-50.569	-36.810	-9.781	-17.233	-15.814	-37.853	-11.894	-32.664	-29.607	-35.994
2011/12	-16.428	-48.042	-54.357	-26.192	-8.851	-11.772	-12.740	-16.570	-8.859	-10.352	-9.764	-15.077
2011/13	-30.192	-28.353	-18.803	-12.664	-14.186	-13.190	-16.591	-17.155	-15.105	-15.278	-11.354	-21.799
2011/14	-6.083	-6.189	-4.784	-1.735	-0.932	-1.043	-1.150	-1.150	-0.903	-1.819	-2.865	-1.689
2011/15	-5.937	-5.884	-3.529	-4.334	-2.341	-0.869	-1.401	-1.110	-0.845	-2.676	-1.499	-1.780
2011/16	-21.361	-25.014	-18.930	-6.511	-4.709	-2.370	-3.404	-3.939	-1.904	-5.107	-6.885	-5.779
2011/17	-16.615	-33.236	-28.786	-19.732	-10.799	-12.886	-13.580	-16.944	-7.519	-10.801	-13.417	-17.964
2011/18	-4.755	-4.639	-4.452	-3.547	-1.463	-1.631	-2.123	-1.275	-0.793	-2.060	-1.856	-1.888
2011/19	-4.814	-31.241	-27.547	-14.225	-4.352	-3.642	-3.278	-7.674	-2.085	-4.805	-3.609	-6.488
2011/20	-21.321	-23.036	-19.145	-7.660	-7.152	-6.812	-4.025	-7.990	-2.126	-10.342	-6.213	-10.253
2011/21	-9.536	-9.939	-8.770	-5.165	-1.366	-1.894	-1.820	-1.515	-1.611	-1.314	-0.601	-3.121
2011/22	-5.286	-5.490	-3.229	-1.719	-0.847	-0.672	-0.662	-1.027	-0.601	-1.493	-0.606	-0.993
2011/23	-18.358	-19.500	-16.681	-4.178	-1.433	-1.392	-2.183	-2.626	-1.326	-2.381	-1.804	-2.791
2011/24	-1.078	-1.226	-1.914	-1.348	-0.604	-0.519	-0.496	-0.480	-0.465	-0.862	-0.505	-0.693
2011/25	-2.577	-3.449	-2.106	-2.241	-1.394	-0.867	-0.841	-0.923	-0.753	-2.565	-0.657	-0.858
2011/26	-3.674	-3.751	-6.586	-6.822	-4.538	-7.707	-9.382	-5.718	-6.164	-3.392	-4.024	-6.656
2011/27	-8.265	-9.324	-5.713	-4.065	-0.836	-0.922	-1.804	-1.414	-0.822	-2.114	-1.542	-2.118
2011/28	0.033	-0.044	0.049	0.086	0.081	0.063	0.044	0.077	0.152	-0.055	0.026	0.123
2011/29	-9.308	-7.360	-2.346	-5.109	-1.723	-1.043	-1.753	-2.038	-1.557	-0.588	-0.377	-2.098
2011/30	-7.603	-7.216	-3.343	-1.082	-0.775	-0.460	-0.588	-0.591	-0.405	-1.742	-0.922	-0.681
2011/31	-0.575	-0.958	-0.197	-0.146	-0.185	-0.106	-0.130	-0.244	-0.108	-0.403	-0.137	-0.096
2011/32	-2.302	-3.624	-1.670	-1.168	-0.132	-0.304	-0.161	-0.227	-0.229	-0.856	-0.199	-0.379
2011/33	-4.457	-4.970	-3.031	-1.772	-0.775	-0.477	-0.672	-0.788	-0.644	-1.293	-1.112	-0.845
2011/34	-0.473	-12.794	-23.436	-6.195	-3.538	-1.210	-1.105	-6.381	-3.952	-5.694	-5.863	-6.940
2011/35	-1.249	-0.658	-0.150	-0.101	-0.152	-0.103	-0.122	-0.291	-0.267	-0.384	-0.207	-0.085
2011/36	-0.396	-0.069	0.042	0.066	0.008	0.039	0.036	-0.021	0.100	0.045	0.082	0.054

Tabla 5.9: LS medio de cada modelo ensayado en cada una de las semanas del período de validación.

Semana	ARIMA	DEM	EOL	ARCH	BOOT	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
201036	-0.759	0.385	0.774	0.673	0.619	0.929	0.955	0.613	0.381	1.045	0.794	0.861
201037	0.270	0.330	0.476	0.489	0.377	0.540	0.525	0.265	0.516	0.459	0.437	0.548
201038	-0.167	0.219	0.579	0.599	0.498	0.543	0.504	0.509	0.517	0.568	0.522	0.562
201039	0.370	0.495	0.671	0.865	0.817	0.775	0.747	0.707	0.642	0.819	0.718	0.790
201040	0.085	0.155	0.327	0.461	0.233	0.384	0.308	0.171	0.444	-0.046	0.179	0.409
201041	0.221	0.323	0.589	0.687	0.610	0.613	0.555	0.488	0.478	0.389	0.533	0.632
201042	0.143	0.238	0.486	0.457	0.310	0.514	0.475	0.381	0.352	0.527	0.464	0.486
201043	0.753	0.700	0.905	0.857	0.734	0.966	0.923	0.728	0.865	0.826	1.007	0.932
201044	-0.067	-0.156	0.094	0.211	0.002	0.262	0.205	0.014	0.095	0.043	0.102	0.142
201045	-0.603	-0.698	-0.361	0.388	0.293	0.118	0.094	0.087	0.146	0.060	0.227	0.395
201046	0.162	0.130	-0.020	0.426	0.349	0.423	0.423	0.369	0.443	0.021	0.406	0.534
201047	-0.111	-0.156	-0.115	0.216	0.052	0.132	0.061	-0.105	0.079	0.295	0.337	0.140
201048	-0.394	-0.139	0.162	0.502	0.375	0.378	0.311	0.168	0.186	-0.141	0.156	0.453
201049	-0.230	0.012	0.017	0.441	0.314	0.389	0.343	0.101	0.228	-0.034	0.185	0.365
201050	-0.643	-0.449	-0.219	0.222	0.130	0.156	0.111	-0.032	0.179	-0.003	0.158	0.170
201051	-0.254	-0.373	-0.258	0.164	-0.002	0.184	0.143	0.102	0.095	0.077	0.097	0.214
201052	-0.720	-0.658	-0.516	-0.038	-0.135	-0.075	-0.113	-0.294	-0.261	-0.057	-0.141	-0.053
201053	-0.680	-0.809	-0.807	0.124	-0.055	-0.048	-0.093	-0.178	-0.030	-0.287	-0.276	0.074
201101	-0.366	-0.327	-0.387	-0.032	-0.240	0.025	0.004	-0.007	-0.175	0.183	0.240	0.259
201102	-0.789	-0.754	-0.757	-0.161	-0.313	-0.205	-0.263	-0.417	-0.225	-0.137	-0.114	-0.183
201103	-0.582	-0.624	-0.468	0.153	0.077	0.068	0.015	-0.069	0.092	-0.344	-0.123	0.037
201104	-0.640	-0.433	-0.169	0.121	-0.014	0.277	0.239	-0.203	-0.001	-0.089	-0.076	0.181
201105	-0.214	-0.201	-0.003	0.175	0.050	0.307	0.295	0.027	0.115	-0.028	0.186	0.224
201106	0.306	0.603	0.618	0.567	0.459	0.765	0.757	0.445	0.571	0.597	0.690	0.737
201107	0.638	0.266	-0.006	-0.038	-0.246	0.087	0.000	-0.284	0.095	-0.369	-0.197	-0.151
201108	0.159	0.018	0.156	0.253	0.104	0.249	0.210	0.034	0.176	0.043	0.078	0.176
201109	0.320	0.144	-0.111	0.017	-0.158	-0.034	-0.107	-0.203	-0.032	-0.108	-0.006	-0.059
201110	0.455	0.514	0.556	0.611	0.422	0.715	0.663	0.391	0.653	0.506	0.674	0.562
201111	0.732	-0.093	-0.269	0.033	-0.126	0.031	-0.034	-0.289	0.149	-0.329	-0.126	-0.028
201112	0.385	-0.372	-0.358	-0.070	-0.245	-0.050	-0.135	-0.270	-0.106	-0.093	0.027	-0.126
201113	0.100	0.199	0.656	0.823	0.740	0.681	0.633	0.788	0.730	0.360	0.467	0.829
201114	0.041	-0.104	0.251	0.488	0.354	0.415	0.358	0.196	0.326	0.269	0.455	0.407
201115	0.566	0.753	0.863	0.854	0.790	0.870	0.849	0.780	0.834	0.557	0.642	0.933
201116	-0.107	-0.137	-0.029	0.369	0.234	0.351	0.278	0.068	0.270	0.287	0.426	0.365
201117	0.174	0.249	0.280	0.524	0.441	0.497	0.467	0.388	0.469	0.552	0.645	0.592
201118	0.303	0.736	0.692	0.683	0.528	0.726	0.663	0.560	0.690	0.568	0.860	0.740
201119	0.401	0.437	0.332	0.328	0.183	0.428	0.396	0.192	0.414	0.119	0.448	0.390
201120	0.591	0.518	0.526	0.584	0.449	0.630	0.591	0.376	0.685	0.281	0.657	0.666
201121	0.080	-0.040	0.594	0.719	0.627	0.741	0.708	0.509	0.701	0.560	0.868	0.800
201122	0.807	0.498	0.638	0.708	0.560	0.656	0.595	0.531	0.663	0.570	0.716	0.682
201123	0.586	0.237	0.469	0.479	0.345	0.623	0.574	0.368	0.679	0.666	0.680	0.482
201124	0.854	0.601	0.853	0.921	0.805	0.910	0.868	0.742	0.921	0.531	0.933	0.887
201125	0.686	0.443	0.262	0.376	0.049	0.345	0.230	0.190	0.290	0.205	0.259	0.226
201126	0.992	0.794	0.655	0.719	0.573	0.778	0.742	0.707	0.635	0.659	0.890	0.743
201127	0.761	0.715	0.707	0.797	0.687	0.856	0.826	0.779	0.549	1.073	1.071	0.784
201128	1.215	1.199	1.313	1.363	1.319	1.329	1.332	1.343	1.435	1.344	1.340	1.453
201129	0.978	0.992	1.132	1.065	0.965	1.168	1.144	1.022	1.199	1.139	1.169	1.155
201130	1.063	0.973	0.566	0.592	0.371	0.694	0.625	0.455	0.743	0.531	0.488	0.541
201131	1.117	0.917	0.975	1.018	0.870	1.092	1.027	0.777	1.064	0.741	0.991	1.032
201132	0.793	0.814	1.149	1.299	1.246	1.348	1.348	1.371	1.306	1.391	1.534	1.421
201133	0.440	0.526	0.774	0.696	0.548	0.779	0.716	0.441	0.604	0.493	0.576	0.574
201134	1.121	0.824	0.865	0.905	0.708	1.028	0.992	0.817	1.012	0.927	0.888	0.860
201135	0.431	0.591	0.979	1.050	0.970	1.077	1.027	0.841	0.856	0.702	1.017	1.084
201136	0.591	1.223	1.174	1.221	1.091	1.193	1.156	1.046	1.317	1.187	1.203	1.111

Tabla 5.10: QS medio de cada modelo ensayado en cada una de las semanas del período de validación.

Semana	ARIMA	DEM	EOL	ARCH	BOOT	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
201036	0.223	0.737	0.910	0.861	0.863	0.964	0.979	0.859	0.760	1.014	0.898	0.941
201037	0.681	0.712	0.779	0.776	0.756	0.799	0.800	0.714	0.798	0.787	0.742	0.800
201038	0.487	0.663	0.825	0.822	0.807	0.799	0.791	0.814	0.800	0.825	0.789	0.818
201039	0.726	0.785	0.865	0.933	0.929	0.896	0.889	0.887	0.851	0.938	0.861	0.902
201040	0.599	0.635	0.714	0.762	0.692	0.727	0.704	0.677	0.765	0.583	0.657	0.749
201041	0.660	0.709	0.829	0.857	0.844	0.823	0.803	0.792	0.777	0.758	0.793	0.839
201042	0.625	0.672	0.785	0.749	0.710	0.777	0.768	0.753	0.721	0.777	0.737	0.761
201043	0.897	0.875	0.967	0.938	0.905	0.991	0.977	0.908	0.959	0.918	1.005	0.975
201044	0.532	0.498	0.612	0.616	0.561	0.644	0.629	0.580	0.588	0.619	0.608	0.603
201045	0.293	0.257	0.414	0.631	0.603	0.500	0.499	0.537	0.549	0.532	0.583	0.632
201046	0.633	0.624	0.563	0.690	0.671	0.690	0.694	0.690	0.709	0.545	0.682	0.736
201047	0.513	0.498	0.522	0.600	0.555	0.562	0.541	0.513	0.557	0.676	0.672	0.577
201048	0.385	0.505	0.643	0.749	0.715	0.697	0.676	0.642	0.629	0.549	0.623	0.737
201049	0.459	0.572	0.580	0.711	0.671	0.683	0.668	0.589	0.614	0.589	0.633	0.693
201050	0.275	0.370	0.476	0.568	0.548	0.532	0.517	0.483	0.547	0.523	0.555	0.550
201051	0.449	0.403	0.459	0.567	0.519	0.574	0.566	0.592	0.549	0.600	0.584	0.609
201052	0.241	0.277	0.346	0.432	0.417	0.413	0.408	0.371	0.339	0.461	0.403	0.434
201053	0.261	0.210	0.219	0.512	0.449	0.416	0.405	0.425	0.451	0.414	0.383	0.513
201101	0.398	0.424	0.403	0.510	0.453	0.533	0.535	0.579	0.470	0.658	0.664	0.650
201102	0.210	0.235	0.241	0.352	0.314	0.327	0.314	0.305	0.331	0.418	0.406	0.352
201103	0.303	0.292	0.367	0.516	0.501	0.463	0.445	0.461	0.493	0.363	0.393	0.473
201104	0.277	0.376	0.498	0.517	0.481	0.598	0.585	0.428	0.483	0.495	0.462	0.564
201105	0.467	0.479	0.570	0.572	0.546	0.654	0.655	0.567	0.589	0.580	0.609	0.626
201106	0.698	0.833	0.842	0.772	0.745	0.888	0.890	0.762	0.822	0.853	0.866	0.862
201107	0.846	0.684	0.569	0.516	0.470	0.598	0.574	0.480	0.628	0.485	0.525	0.511
201108	0.633	0.575	0.639	0.649	0.614	0.653	0.646	0.607	0.645	0.610	0.607	0.634
201109	0.703	0.630	0.523	0.553	0.521	0.536	0.520	0.521	0.569	0.551	0.563	0.535
201110	0.764	0.794	0.815	0.817	0.761	0.868	0.851	0.757	0.851	0.821	0.865	0.815
201111	0.888	0.525	0.453	0.511	0.478	0.518	0.502	0.440	0.603	0.471	0.493	0.509
201112	0.734	0.404	0.415	0.496	0.462	0.511	0.486	0.473	0.512	0.516	0.563	0.503
201113	0.606	0.654	0.858	0.914	0.895	0.852	0.838	0.920	0.878	0.718	0.746	0.916
201114	0.580	0.521	0.681	0.753	0.712	0.718	0.699	0.658	0.700	0.701	0.759	0.730
201115	0.813	0.898	0.949	0.940	0.925	0.949	0.944	0.923	0.938	0.796	0.834	0.964
201116	0.513	0.507	0.559	0.668	0.629	0.660	0.633	0.573	0.634	0.676	0.721	0.669
201117	0.638	0.677	0.694	0.754	0.736	0.744	0.739	0.728	0.747	0.763	0.801	0.783
201118	0.696	0.891	0.875	0.871	0.830	0.893	0.872	0.843	0.890	0.799	0.926	0.895
201119	0.740	0.760	0.717	0.712	0.683	0.763	0.758	0.690	0.777	0.635	0.746	0.746
201120	0.825	0.796	0.801	0.809	0.779	0.828	0.818	0.753	0.862	0.657	0.816	0.841
201121	0.597	0.550	0.831	0.866	0.849	0.873	0.862	0.802	0.871	0.811	0.930	0.900
201122	0.921	0.786	0.851	0.876	0.841	0.853	0.834	0.833	0.875	0.845	0.882	0.879
201123	0.823	0.671	0.777	0.765	0.738	0.825	0.812	0.754	0.864	0.861	0.860	0.787
201124	0.941	0.832	0.944	0.974	0.946	0.966	0.954	0.920	0.980	0.832	0.981	0.962
201125	0.866	0.762	0.686	0.736	0.641	0.724	0.689	0.699	0.723	0.754	0.737	0.711
201126	1.003	0.918	0.858	0.893	0.859	0.918	0.908	0.909	0.875	0.914	0.976	0.917
201127	0.900	0.883	0.881	0.921	0.898	0.947	0.940	0.929	0.835	1.053	1.025	0.935
201128	1.103	1.095	1.146	1.165	1.150	1.153	1.154	1.161	1.196	1.152	1.148	1.200
201129	0.997	1.004	1.067	1.042	1.018	1.085	1.079	1.037	1.103	1.057	1.069	1.076
201130	1.035	0.996	0.819	0.817	0.751	0.865	0.841	0.790	0.898	0.821	0.799	0.816
201131	1.059	0.971	0.998	1.015	0.969	1.045	1.020	0.936	1.039	0.922	1.007	1.023
201132	0.915	0.926	1.074	1.133	1.120	1.152	1.154	1.167	1.135	1.171	1.207	1.171
201133	0.758	0.799	0.910	0.876	0.839	0.908	0.886	0.793	0.853	0.834	0.844	0.844
201134	1.061	0.930	0.950	0.960	0.900	1.012	1.001	0.945	1.013	0.994	0.972	0.953
201135	0.754	0.827	1.000	1.034	1.014	1.045	1.027	0.965	0.962	0.940	1.052	1.065
201136	0.824	1.106	1.085	1.112	1.071	1.103	1.089	1.049	1.154	1.114	1.123	1.089

Tabla 5.11: SS medio de cada modelo ensayado en cada una de las semanas del período de validación.

Semana	ARIMA	DEM	DEM_EOL	ARCH	BOOT	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
201036	0.482	0.173	0.110	0.135	0.137	0.107	0.107	0.131	0.152	0.114	0.144	0.130
201037	0.191	0.198	0.174	0.159	0.166	0.148	0.156	0.163	0.164	0.156	0.181	0.157
201038	0.324	0.225	0.151	0.158	0.151	0.171	0.164	0.152	0.156	0.147	0.175	0.157
201039	0.196	0.164	0.126	0.111	0.117	0.118	0.122	0.128	0.129	0.123	0.144	0.119
201040	0.228	0.220	0.179	0.178	0.190	0.182	0.172	0.179	0.172	0.222	0.212	0.172
201041	0.221	0.178	0.148	0.148	0.137	0.168	0.168	0.163	0.165	0.178	0.181	0.145
201042	0.245	0.206	0.152	0.185	0.189	0.177	0.173	0.161	0.191	0.173	0.193	0.177
201043	0.117	0.107	0.111	0.120	0.123	0.113	0.111	0.123	0.118	0.123	0.108	0.115
201044	0.300	0.314	0.210	0.235	0.254	0.203	0.207	0.229	0.254	0.212	0.215	0.238
201045	0.571	0.657	0.432	0.262	0.264	0.392	0.387	0.281	0.320	0.307	0.269	0.263
201046	0.225	0.248	0.325	0.255	0.246	0.262	0.268	0.244	0.253	0.301	0.224	0.217
201047	0.387	0.371	0.317	0.288	0.294	0.302	0.294	0.308	0.317	0.261	0.280	0.284
201048	0.446	0.360	0.207	0.170	0.176	0.190	0.196	0.206	0.210	0.224	0.215	0.170
201049	0.366	0.253	0.249	0.215	0.225	0.224	0.230	0.246	0.253	0.255	0.239	0.220
201050	0.659	0.513	0.366	0.305	0.325	0.336	0.347	0.350	0.325	0.343	0.347	0.323
201051	0.398	0.391	0.375	0.277	0.287	0.280	0.285	0.250	0.313	0.263	0.294	0.270
201052	0.709	0.651	0.556	0.502	0.497	0.498	0.485	0.575	0.608	0.459	0.520	0.470
201053	0.562	0.665	0.487	0.327	0.334	0.410	0.406	0.378	0.352	0.372	0.407	0.322
201101	0.516	0.413	0.457	0.292	0.305	0.301	0.266	0.246	0.380	0.225	0.184	0.208
201102	0.720	0.705	0.658	0.540	0.550	0.574	0.591	0.607	0.637	0.413	0.512	0.533
201103	0.670	0.633	0.432	0.375	0.381	0.404	0.416	0.410	0.383	0.461	0.457	0.404
201104	0.597	0.456	0.377	0.392	0.405	0.317	0.335	0.399	0.350	0.380	0.406	0.357
201105	0.330	0.379	0.277	0.272	0.283	0.227	0.225	0.265	0.243	0.232	0.234	0.237
201106	0.199	0.147	0.143	0.180	0.185	0.123	0.120	0.182	0.154	0.138	0.151	0.152
201107	0.129	0.194	0.252	0.286	0.313	0.246	0.253	0.288	0.201	0.304	0.270	0.280
201108	0.242	0.270	0.227	0.244	0.258	0.220	0.228	0.253	0.228	0.244	0.238	0.225
201109	0.203	0.242	0.300	0.265	0.271	0.273	0.265	0.268	0.287	0.294	0.268	0.281
201110	0.176	0.169	0.146	0.155	0.158	0.139	0.143	0.170	0.141	0.137	0.143	0.150
201111	0.143	0.309	0.373	0.360	0.379	0.339	0.370	0.395	0.235	0.358	0.367	0.373
201112	0.179	0.364	0.374	0.334	0.344	0.307	0.319	0.321	0.277	0.273	0.287	0.319
201113	0.222	0.212	0.131	0.115	0.117	0.135	0.135	0.115	0.132	0.156	0.177	0.116
201114	0.274	0.353	0.203	0.194	0.204	0.223	0.230	0.226	0.199	0.198	0.187	0.204
201115	0.147	0.128	0.110	0.127	0.128	0.121	0.119	0.115	0.118	0.165	0.157	0.124
201116	0.309	0.306	0.274	0.237	0.257	0.243	0.248	0.246	0.248	0.183	0.195	0.230
201117	0.236	0.204	0.213	0.177	0.176	0.179	0.185	0.172	0.172	0.175	0.160	0.181
201118	0.204	0.125	0.132	0.144	0.149	0.133	0.140	0.144	0.134	0.156	0.127	0.140
201119	0.165	0.153	0.188	0.209	0.222	0.168	0.176	0.183	0.174	0.233	0.185	0.182
201120	0.136	0.152	0.141	0.151	0.150	0.143	0.144	0.157	0.137	0.226	0.156	0.143
201121	0.239	0.281	0.128	0.147	0.154	0.145	0.146	0.142	0.143	0.151	0.120	0.137
201122	0.110	0.151	0.137	0.127	0.138	0.137	0.132	0.134	0.123	0.137	0.137	0.130
201123	0.135	0.223	0.182	0.179	0.180	0.155	0.152	0.171	0.137	0.136	0.143	0.162
201124	0.118	0.149	0.111	0.108	0.112	0.114	0.114	0.125	0.105	0.159	0.113	0.118
201125	0.130	0.159	0.190	0.178	0.192	0.180	0.178	0.159	0.165	0.168	0.167	0.172
201126	0.100	0.123	0.144	0.148	0.154	0.141	0.137	0.127	0.149	0.133	0.117	0.141
201127	0.115	0.124	0.139	0.131	0.135	0.122	0.123	0.121	0.143	0.099	0.101	0.124
201128	0.091	0.098	0.080	0.091	0.089	0.081	0.085	0.073	0.075	0.078	0.087	0.085
201129	0.098	0.100	0.097	0.101	0.101	0.095	0.096	0.096	0.087	0.101	0.101	0.095
201130	0.106	0.109	0.153	0.156	0.161	0.150	0.152	0.148	0.139	0.155	0.156	0.153
201131	0.087	0.103	0.118	0.113	0.114	0.110	0.112	0.127	0.105	0.123	0.122	0.113
201132	0.117	0.125	0.086	0.082	0.082	0.075	0.075	0.069	0.080	0.079	0.083	0.081
201133	0.162	0.152	0.127	0.129	0.140	0.120	0.124	0.147	0.140	0.138	0.139	0.132
201134	0.091	0.122	0.120	0.126	0.135	0.114	0.114	0.119	0.115	0.107	0.115	0.120
201135	0.181	0.138	0.104	0.100	0.098	0.097	0.095	0.104	0.111	0.125	0.097	0.093
201136	0.170	0.088	0.093	0.086	0.086	0.090	0.091	0.091	0.092	0.085	0.090	0.098

Tabla 5.12: mediana del CRPS de cada modelo ensayado en cada una de las semanas del período de validación.

Semana	ARIMA	DEM	EOL	ARCH	BOOT	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
2010/36	58.33	18.33	20.00	15.00	10.83	5.00	6.67	10.00	6.67	15.00	6.67	13.33
2010/37	26.19	25.00	20.83	19.05	11.91	8.93	9.52	11.31	10.12	18.45	13.69	16.07
2010/38	32.14	19.05	17.86	12.50	8.93	7.14	5.95	8.93	7.74	13.10	10.12	10.71
2010/39	18.45	18.45	18.45	15.48	6.55	2.98	4.17	3.57	6.55	13.10	7.74	13.10
2010/40	28.57	23.81	13.10	8.33	6.55	4.17	4.17	4.17	4.17	14.88	6.55	7.14
201041	16.67	19.64	16.07	11.91	5.36	1.79	1.79	4.17	1.79	10.12	10.12	10.12
2010/42	27.98	23.21	20.83	18.45	10.12	7.74	4.76	8.93	9.52	14.88	10.71	15.48
2010/43	13.10	15.48	8.93	7.74	1.79	2.38	1.19	4.76	1.19	9.52	4.17	2.98
2010/44	42.86	41.67	30.95	25.60	21.43	19.64	20.24	20.83	20.24	23.81	23.21	28.57
2010/45	61.31	65.48	51.19	20.83	10.71	14.29	8.33	20.24	11.31	30.95	29.76	20.24
2010/46	38.10	35.71	42.26	17.26	13.10	11.31	10.71	14.88	10.12	23.81	16.07	14.88
2010/47	43.45	45.83	38.10	20.83	13.10	13.10	11.31	18.45	11.31	28.57	28.57	25.00
2010/48	52.98	40.48	29.17	15.48	8.93	11.91	3.57	10.71	8.93	21.43	19.05	10.12
2010/49	42.86	31.55	33.33	13.69	6.55	6.55	3.57	10.12	7.74	24.41	17.86	16.07
2010/50	69.64	55.95	43.45	23.21	15.48	10.12	10.71	17.26	11.31	28.57	23.81	25.60
2010/51	47.62	45.83	44.64	22.02	16.07	11.91	8.33	14.88	10.71	20.83	25.00	19.05
2010/52	70.24	69.05	61.91	44.64	30.95	26.19	25.60	33.93	28.57	39.29	39.88	41.67
2010/53	63.03	71.43	60.50	13.45	6.72	10.08	5.04	10.92	4.20	26.05	27.73	11.77
2011/01	53.06	48.98	55.10	32.65	12.25	10.20	0.00	18.37	4.08	30.61	22.45	18.37
2011/02	71.43	75.00	67.86	38.10	25.60	26.19	25.60	29.76	23.81	30.95	34.52	38.10
2011/03	60.71	57.74	51.19	23.81	12.50	14.29	8.33	20.83	10.71	38.69	30.36	23.21
2011/04	63.69	51.19	47.62	22.62	13.69	10.12	9.52	16.67	8.93	32.74	27.38	22.62
2011/05	44.05	46.43	34.52	25.00	17.26	14.29	8.33	20.24	15.48	28.57	19.64	23.81
2011/06	17.86	19.64	13.69	18.45	9.52	2.98	2.38	6.55	2.98	14.88	13.10	14.88
2011/07	19.64	20.83	26.79	26.79	19.05	10.12	6.55	13.69	7.14	30.95	26.19	27.38
2011/08	38.10	35.12	33.93	23.81	15.48	13.69	8.93	14.29	13.69	23.81	25.60	24.41
2011/09	27.38	33.93	44.64	34.52	23.81	17.86	14.29	17.86	14.88	30.95	27.38	35.12
2011/10	16.67	16.07	14.88	13.69	6.55	7.14	1.79	5.95	2.98	9.52	5.36	9.52
2011/11	10.12	42.86	45.83	32.14	22.02	17.26	17.26	25.00	17.26	39.29	34.52	35.12
2011/12	17.86	44.64	45.24	30.95	23.21	20.24	17.26	20.83	17.26	33.33	29.76	32.74
2011/13	27.38	26.79	20.83	16.07	13.69	13.10	13.10	13.69	13.69	19.64	14.88	15.48
2011/14	32.74	42.86	23.21	10.12	7.74	8.33	2.38	7.74	4.17	12.50	11.31	9.52
2011/15	19.05	13.69	11.91	8.93	6.55	5.36	4.76	5.95	5.95	11.91	10.12	7.74
2011/16	37.50	39.88	27.98	13.10	7.14	8.93	5.36	10.12	5.95	17.86	14.88	10.71
2011/17	29.76	32.14	28.57	19.05	11.91	10.71	8.33	16.07	11.31	16.07	14.29	15.48
2011/18	21.43	11.31	10.71	13.10	8.33	5.36	2.98	6.55	5.36	10.71	5.95	12.50
2011/19	18.45	25.00	25.00	23.21	18.45	15.48	13.69	14.29	14.29	25.00	17.86	20.24
2011/20	22.62	22.62	22.62	20.83	17.86	11.31	11.31	16.07	13.10	17.86	12.50	19.05
2011/21	29.17	30.36	22.62	13.69	8.93	7.14	3.57	7.74	4.76	20.24	9.52	11.91
2011/22	13.69	17.26	14.88	17.26	7.74	6.55	3.57	5.95	5.95	17.26	10.71	14.88
2011/23	14.88	21.43	17.86	13.69	7.74	4.76	5.36	5.95	5.36	14.29	11.31	13.69
2011/24	11.31	11.91	16.07	13.10	9.52	7.74	1.19	5.95	6.55	9.52	8.33	10.12
2011/25	20.83	22.62	16.07	12.50	6.55	5.36	2.38	8.33	3.57	19.64	6.55	10.12
2011/26	9.52	10.12	12.50	12.50	6.55	7.14	4.17	6.55	4.76	14.88	9.52	10.12
2011/27	16.07	16.07	14.88	13.69	8.33	7.14	5.36	6.55	7.14	10.71	9.52	11.91
2011/28	1.19	1.79	2.38	4.17	1.79	0.00	0.60	0.00	0.60	4.76	2.38	1.19
2011/29	12.50	14.29	6.55	7.14	4.17	3.57	4.17	4.17	8.33	4.17	5.95	5.95
2011/30	7.74	9.52	10.71	8.33	2.98	2.38	2.98	2.98	2.98	12.50	8.93	8.93
2011/31	10.12	12.50	4.76	4.17	2.38	1.19	0.60	1.19	1.19	6.55	1.79	3.57
2011/32	13.10	13.69	10.12	7.14	4.76	3.57	2.38	5.36	2.98	7.14	2.98	5.36
2011/33	20.83	19.05	14.88	11.91	7.14	3.57	4.76	7.14	4.76	13.10	10.12	10.12
2011/34	5.95	12.50	7.74	5.95	5.95	4.76	5.95	4.76	4.76	10.12	7.74	6.55
2011/35	14.29	13.69	2.98	2.38	1.19	0.60	0.00	0.60	0.60	8.33	5.95	4.76
2011/36	4.17	8.33	2.78	2.78	0.00	0.00	0.00	0.00	0.00	2.78	2.78	1.39

Tabla 5.13: Porcentaje de muestras que quedan fuera de los intervalos de predicción al 99 % de confianza en cada semana del período de validación (01/Sept/2010-31/Ago/2011, 8760 muestras). En negrita el resultado del mejor modelo en cada semana.

Semana	ARIMA	DEM	EOL	ARCH	BOOT	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
2010/36	77.50	25.83	23.33	26.67	25.83	18.33	18.33	24.17	20.83	20.83	22.50	24.17
2010/37	38.69	37.50	30.36	30.36	30.36	26.19	25.00	26.19	23.21	30.36	28.57	29.76
2010/38	52.38	34.52	26.79	23.81	23.21	19.05	20.24	20.83	17.86	25.60	20.83	24.41
2010/39	31.55	29.76	26.79	24.41	23.21	22.02	22.62	20.24	18.45	19.05	16.67	24.41
2010/40	42.26	37.50	26.79	19.05	14.88	15.48	14.29	16.07	17.86	35.71	26.19	22.02
2010/41	33.33	35.71	25.00	19.05	19.64	16.67	16.07	16.07	17.86	27.38	22.62	23.21
2010/42	41.07	37.50	33.33	27.38	25.60	23.81	23.21	22.62	23.81	25.60	23.81	27.38
2010/43	27.38	30.95	16.67	16.07	13.69	8.33	9.52	12.50	8.33	23.21	16.67	13.10
2010/44	49.41	51.19	37.50	36.91	34.52	28.57	27.38	31.55	33.33	39.88	36.31	39.29
2010/45	70.83	74.41	64.88	31.55	30.95	39.29	37.50	32.74	27.98	47.62	39.88	31.55
2010/46	44.05	45.83	52.38	23.21	22.62	21.43	20.24	20.83	20.83	43.45	30.36	22.02
2010/47	55.36	56.55	52.38	36.31	35.71	32.74	33.33	33.93	31.55	40.48	38.10	38.10
2010/48	62.50	54.17	39.29	22.02	22.62	21.43	20.83	25.00	20.83	39.88	35.12	23.81
2010/49	53.57	42.86	42.86	22.02	20.83	22.62	22.62	26.19	27.98	39.29	27.98	27.98
2010/50	77.38	62.50	56.55	33.93	31.55	31.55	32.14	36.91	31.55	40.48	36.31	35.71
2010/51	56.55	62.50	56.55	35.12	33.33	29.76	29.76	32.74	32.14	41.07	42.26	35.71
2010/52	75.60	76.19	70.24	50.00	49.41	48.21	48.81	51.19	48.81	50.00	55.36	51.19
2010/53	76.47	77.31	78.99	25.21	27.73	32.77	28.57	30.25	26.89	44.54	45.38	25.21
2011/01	63.27	57.14	65.31	44.90	44.90	42.86	44.90	44.90	48.98	38.78	34.69	32.65
2011/02	77.38	79.76	76.19	50.60	48.21	45.83	47.02	53.57	50.60	47.62	50.00	52.38
2011/03	67.86	64.88	64.88	32.74	32.74	33.33	35.71	38.10	31.55	48.81	45.24	36.31
2011/04	73.81	63.69	56.55	35.71	36.91	27.98	27.98	36.91	29.76	45.83	45.83	31.55
2011/05	51.19	57.14	45.83	36.91	36.91	28.57	26.79	32.14	30.95	42.26	29.76	32.14
2011/06	35.12	28.57	25.60	27.98	28.57	13.10	12.50	25.00	18.45	31.55	23.21	27.38
2011/07	28.57	31.55	41.67	48.21	46.43	28.57	29.76	40.48	28.57	48.21	40.48	48.21
2011/08	47.02	46.43	42.26	34.52	33.93	31.55	29.76	33.93	32.74	45.83	38.10	33.93
2011/09	39.29	44.05	50.60	45.83	43.45	43.45	42.86	42.86	44.05	48.81	42.86	47.62
2011/10	30.36	29.17	27.98	19.64	22.02	18.45	16.07	19.64	17.26	19.05	13.69	20.24
2011/11	19.05	50.60	58.93	44.05	42.86	36.31	38.69	48.21	32.74	50.00	50.00	47.02
2011/12	27.98	58.33	58.93	42.86	44.05	38.69	38.10	41.07	33.93	45.83	41.67	47.62
2011/13	41.07	36.91	28.57	22.62	20.83	17.86	19.05	19.64	18.45	30.36	28.57	20.83
2011/14	45.24	53.57	35.71	17.86	16.67	20.24	19.05	20.83	16.67	25.60	20.83	17.86
2011/15	27.98	21.43	16.67	11.91	12.50	12.50	13.10	12.50	11.91	23.21	17.86	10.71
2011/16	48.81	50.00	44.64	28.57	27.98	27.38	24.41	27.98	23.81	34.52	25.60	26.79
2011/17	42.26	36.31	39.29	26.79	25.60	26.19	25.00	26.19	23.81	23.81	19.05	26.19
2011/18	36.31	22.02	20.24	19.64	19.64	14.88	16.07	17.26	15.48	25.00	14.29	20.83
2011/19	29.76	33.93	36.91	33.33	35.12	28.57	26.79	32.14	27.38	36.31	23.81	32.14
2011/20	28.57	29.76	30.95	27.38	27.98	22.62	23.21	25.60	22.02	36.31	22.62	25.60
2011/21	42.86	47.02	30.95	23.81	23.21	20.83	20.24	22.02	20.24	30.95	16.67	23.21
2011/22	21.43	32.14	27.38	26.79	23.21	22.02	20.24	20.83	17.86	29.17	20.24	24.41
2011/23	25.60	35.71	28.57	29.76	27.98	17.26	17.26	20.24	15.48	25.60	25.00	30.95
2011/24	20.24	20.83	22.62	20.83	19.05	17.26	16.67	16.67	17.26	27.38	16.67	17.26
2011/25	26.19	33.93	26.79	24.41	24.41	20.24	19.05	17.86	23.21	35.12	21.43	25.00
2011/26	14.88	17.26	20.83	18.45	19.64	11.91	12.50	10.12	11.91	26.79	17.26	21.43
2011/27	25.00	27.38	21.43	20.24	20.24	17.26	17.26	14.88	19.05	20.83	17.26	22.02
2011/28	6.55	4.76	4.76	7.14	5.95	5.36	4.76	5.95	4.17	14.29	13.69	11.91
2011/29	18.45	19.05	11.91	14.29	13.69	11.31	11.31	13.10	11.31	17.86	11.31	12.50
2011/30	12.50	16.67	20.83	16.07	16.07	13.10	12.50	13.10	8.93	22.02	20.83	18.45
2011/31	16.67	17.86	10.71	10.71	10.71	8.93	8.33	6.55	8.93	20.83	13.10	11.31
2011/32	19.64	22.62	15.48	14.29	14.88	13.10	13.69	11.31	13.69	14.29	6.55	13.10
2011/33	26.79	26.19	22.62	19.05	18.45	16.07	15.48	16.67	18.45	27.38	23.21	20.83
2011/34	11.31	19.05	17.26	13.10	11.91	8.93	9.52	10.71	11.31	17.86	15.48	13.69
2011/35	26.19	23.81	12.50	13.10	12.50	11.31	11.91	13.69	14.29	22.62	17.26	13.10
2011/36	13.89	11.11	9.72	8.33	6.94	5.56	5.56	4.17	5.56	15.28	11.11	12.50

Tabla 5.14: Porcentaje de muestras que quedan fuera de los intervalos de predicción al 95 % de confianza en cada semana del período de validación (01/Sept/2010-31/Ago/2011, 8760 muestras). En negrita el resultado del mejor modelo en cada semana.

Semana	ARIMA	DEM	EOL	ARCH	BOOT	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
2010/36	82.50	42.50	30.00	34.17	35.00	24.17	23.33	33.33	30.00	33.33	30.00	30.83
2010/37	44.64	42.26	36.31	32.14	34.52	31.55	31.55	38.69	32.14	37.50	36.31	35.12
2010/38	60.12	45.83	35.71	30.36	30.36	29.76	29.76	30.36	27.98	31.55	27.98	31.55
2010/39	41.67	37.50	33.33	28.57	28.57	27.98	27.38	32.14	27.98	24.41	27.38	29.17
2010/40	48.21	48.81	41.07	31.55	35.71	29.17	28.57	31.55	29.76	44.05	35.12	33.93
2010/41	44.05	42.26	32.74	26.19	26.79	24.41	26.79	27.38	26.19	35.12	33.33	26.79
2010/42	50.00	46.43	39.88	33.33	35.12	32.74	33.33	36.31	33.33	33.93	32.14	31.55
2010/43	32.14	37.50	21.43	23.21	24.41	17.26	18.45	23.81	18.45	29.76	22.02	19.64
2010/44	57.14	58.33	45.24	46.43	48.21	38.10	38.69	45.83	44.05	44.64	46.43	48.21
2010/45	77.38	80.36	67.26	36.91	41.67	51.19	51.19	46.43	43.45	51.79	48.21	37.50
2010/46	48.81	50.60	57.74	28.57	33.33	29.17	30.95	36.91	30.95	51.79	39.88	26.79
2010/47	60.12	62.50	58.33	44.64	48.21	46.43	48.21	50.00	45.83	46.43	42.86	46.43
2010/48	70.24	61.31	45.83	29.17	32.14	31.55	31.55	39.88	37.50	50.00	44.05	29.76
2010/49	60.12	52.38	52.38	29.76	30.95	29.17	30.95	39.88	38.69	49.41	40.48	33.33
2010/50	78.57	68.45	61.31	41.07	41.67	42.26	42.26	48.81	40.48	47.62	44.64	41.67
2010/51	63.69	68.45	62.50	44.05	45.24	44.05	47.62	45.24	44.64	48.21	51.79	43.45
2010/52	82.14	79.76	75.60	57.14	58.33	55.95	56.55	61.91	63.69	55.95	60.71	55.95
2010/53	77.31	84.87	84.03	33.61	39.50	43.70	45.38	51.26	36.98	56.30	59.66	39.50
2011/01	69.39	67.35	69.39	48.98	51.02	53.06	48.98	51.02	57.14	40.82	36.74	40.82
2011/02	82.14	82.14	80.36	57.14	57.14	57.74	60.12	64.88	63.10	54.17	55.95	58.33
2011/03	73.81	75.60	69.05	42.26	43.45	42.26	44.64	50.00	38.10	56.55	54.76	42.26
2011/04	77.98	69.05	60.71	45.24	49.41	33.93	36.31	56.55	39.29	54.17	53.57	43.45
2011/05	59.52	61.91	54.17	41.67	43.45	35.12	35.71	42.26	38.10	51.19	41.07	38.69
2011/06	44.64	36.31	33.93	33.93	34.52	25.60	26.79	36.91	29.76	41.07	30.36	33.93
2011/07	33.33	43.45	53.57	54.76	55.95	45.83	48.81	58.33	41.67	60.71	52.98	55.36
2011/08	50.00	55.36	48.81	40.48	41.67	43.45	43.45	47.62	44.64	51.79	46.43	44.05
2011/09	44.64	49.41	56.55	50.00	52.38	50.00	52.38	53.57	51.19	53.57	52.98	52.38
2011/10	38.69	36.31	35.12	27.38	29.17	26.19	26.79	33.93	27.38	29.17	23.81	27.98
2011/11	29.76	58.33	64.29	55.36	55.95	49.41	51.19	61.91	43.45	56.55	54.76	53.57
2011/12	38.10	67.26	66.07	50.60	55.95	47.02	46.43	53.57	44.05	51.19	50.00	55.36
2011/13	47.02	47.02	33.33	26.79	28.57	29.17	30.36	29.17	26.79	36.91	38.10	27.98
2011/14	54.17	60.12	43.45	27.38	30.95	29.76	33.93	36.91	29.17	35.12	26.79	28.57
2011/15	35.71	31.55	26.19	19.05	22.62	19.64	20.24	27.38	22.02	32.14	25.00	20.83
2011/16	58.33	60.12	55.36	34.52	38.10	35.12	36.31	39.88	35.12	42.26	33.93	34.52
2011/17	48.21	45.24	46.43	30.95	32.14	29.76	29.76	34.52	31.55	33.93	27.38	31.55
2011/18	44.64	30.36	29.76	27.98	29.17	23.81	27.98	30.36	24.41	35.71	21.43	25.60
2011/19	38.10	36.31	41.07	39.88	42.26	35.12	35.71	40.48	33.33	47.02	30.95	38.10
2011/20	35.12	33.93	36.31	33.33	33.93	30.95	33.33	35.71	27.98	45.24	32.14	31.55
2011/21	48.81	54.76	37.50	26.19	29.76	27.98	29.76	33.33	26.79	33.93	22.62	29.17
2011/22	26.19	37.50	30.95	32.14	32.74	31.55	32.74	32.74	30.95	35.12	29.17	33.93
2011/23	32.14	46.43	39.88	35.71	37.50	28.57	29.17	33.93	26.79	27.98	31.55	36.31
2011/24	24.41	31.55	29.76	25.00	27.98	23.81	25.00	28.57	24.41	41.67	25.00	26.79
2011/25	33.93	37.50	38.10	35.71	39.88	30.95	32.14	32.74	30.36	44.64	36.31	37.50
2011/26	19.64	23.81	29.17	29.76	32.74	25.00	25.60	24.41	30.95	36.91	23.81	30.95
2011/27	30.95	35.71	28.57	25.00	29.76	20.83	21.43	26.79	27.38	28.57	22.62	30.95
2011/28	11.91	10.71	9.52	12.50	13.10	13.69	13.69	13.10	10.12	20.83	16.07	15.48
2011/29	22.02	21.43	16.67	19.64	22.62	20.24	20.24	20.24	20.24	24.41	22.62	22.62
2011/30	17.86	23.81	33.33	26.79	32.74	22.62	25.00	26.19	23.21	36.31	30.95	27.38
2011/31	19.64	26.79	17.86	16.07	19.64	16.67	16.07	24.41	17.26	30.95	20.83	18.45
2011/32	26.79	28.57	21.43	17.86	20.24	17.26	17.26	21.43	22.02	23.21	13.10	16.07
2011/33	32.74	30.95	26.79	26.79	33.33	25.00	23.81	32.74	27.98	36.91	32.14	32.14
2011/34	15.48	26.19	22.02	19.05	21.43	16.67	16.67	19.05	20.83	26.79	24.41	23.21
2011/35	32.74	34.52	22.02	20.24	23.21	15.48	15.48	27.98	23.81	32.74	22.02	22.62
2011/36	29.17	15.28	12.50	13.89	13.89	13.89	13.89	18.06	12.50	18.06	13.89	18.06

Tabla 5.15: Porcentaje de muestras que quedan fuera de los intervalos de predicción al 90 % de confianza en cada semana del período de validación (01/Sept/2010-31/Ago/2011, 8760 muestras). En negrita el resultado del mejor modelo en cada semana.

Semana	ARIMA	DEM	EOL	ARCH	BOOT	SNP5	SNP6	SNP7	SNP8	RS2	RS3	RS4
2010/36	90.00	58.33	45.00	45.83	48.33	40.83	41.67	51.67	57.50	60.00	65.83	46.67
2010/37	63.10	60.71	57.74	48.81	58.33	47.62	50.00	60.12	50.00	66.67	63.69	50.60
2010/38	79.17	67.86	54.17	44.64	52.38	48.21	49.41	53.57	49.41	64.29	61.31	49.41
2010/39	57.74	55.36	48.81	41.67	49.41	45.24	45.83	52.38	52.38	63.10	59.52	48.21
2010/40	69.64	63.69	65.48	57.14	66.67	63.10	63.69	69.05	56.55	78.57	73.21	60.71
2010/41	65.48	63.10	52.98	41.67	47.02	42.26	46.43	54.76	47.62	64.88	56.55	44.05
2010/42	65.48	67.26	57.74	52.38	58.93	49.41	51.79	58.33	54.17	65.48	58.93	53.57
2010/43	45.83	46.43	42.86	42.86	50.00	41.07	43.45	52.38	45.24	56.55	53.57	45.24
2010/44	72.62	75.00	67.86	63.69	74.41	63.69	64.88	73.21	67.26	74.41	72.02	66.67
2010/45	83.33	88.10	78.57	54.76	62.50	67.26	68.45	68.45	65.48	79.17	64.29	60.12
2010/46	63.69	63.10	67.26	51.79	57.14	50.60	53.57	57.14	51.79	72.62	64.88	50.60
2010/47	69.64	72.02	72.02	64.88	69.64	69.05	71.43	71.43	70.24	72.62	61.91	67.26
2010/48	80.36	72.02	68.45	51.19	60.12	55.95	60.12	67.86	64.29	76.19	62.50	56.55
2010/49	77.38	71.43	70.24	48.81	58.33	52.98	55.36	66.07	58.33	74.41	63.10	51.79
2010/50	83.33	80.36	75.00	54.17	58.33	58.33	60.71	66.67	56.55	79.17	67.86	60.71
2010/51	78.57	81.55	76.19	65.48	72.02	63.10	67.26	67.26	65.48	73.21	64.29	61.31
2010/52	88.10	83.33	80.36	69.05	72.02	72.62	72.02	77.38	76.79	79.76	75.00	68.45
2010/53	89.08	89.92	92.44	66.39	73.11	69.75	72.27	68.91	66.39	73.95	71.43	64.71
2011/01	77.55	79.59	79.59	67.35	71.43	67.35	69.39	61.22	69.39	59.18	55.10	48.98
2011/02	89.88	86.31	88.10	72.02	78.57	76.19	78.57	80.36	75.60	83.93	77.38	77.38
2011/03	86.31	88.10	81.55	55.36	61.31	63.10	64.88	67.26	58.33	79.17	61.31	63.69
2011/04	86.31	82.14	72.02	67.26	70.83	61.31	63.10	76.19	71.43	77.38	66.67	62.50
2011/05	76.79	75.00	70.24	61.91	65.48	55.95	57.14	67.26	60.71	75.60	64.29	61.31
2011/06	64.88	52.38	52.38	50.00	55.95	44.05	45.24	55.36	53.57	54.17	50.60	47.02
2011/07	48.81	66.67	72.02	70.83	77.98	68.45	70.83	76.79	69.05	82.74	78.57	71.43
2011/08	64.88	67.86	64.88	57.74	67.26	59.52	62.50	66.67	60.71	71.43	69.05	62.50
2011/09	60.12	63.10	73.21	66.67	71.43	68.45	71.43	72.02	68.45	82.74	76.19	69.05
2011/10	57.14	56.55	52.38	53.57	63.10	45.83	48.81	61.31	50.60	70.83	63.10	57.14
2011/11	51.19	70.83	75.60	67.86	73.21	67.86	70.24	75.00	64.29	77.38	73.21	68.45
2011/12	61.91	79.76	79.76	69.05	72.62	70.83	72.62	75.60	69.64	79.76	73.21	70.83
2011/13	69.05	67.86	50.00	42.26	49.41	48.21	50.60	47.62	49.41	52.98	47.62	47.02
2011/14	68.45	70.83	66.67	51.19	60.12	50.60	55.36	63.10	54.17	72.02	54.76	57.74
2011/15	54.17	47.62	44.64	39.88	42.86	36.91	40.48	45.83	40.48	54.76	42.86	39.88
2011/16	71.43	72.02	73.21	55.36	64.88	55.36	58.93	72.62	61.31	76.79	69.64	60.71
2011/17	64.29	63.69	58.93	49.41	54.76	47.62	50.00	52.98	49.41	60.71	51.79	50.00
2011/18	61.91	48.81	50.60	47.02	54.17	44.64	46.43	55.36	50.60	60.12	52.38	48.81
2011/19	60.12	56.55	61.91	55.95	59.52	54.17	55.36	63.69	54.76	67.86	61.31	55.36
2011/20	50.00	54.76	51.19	48.81	57.14	47.62	48.81	58.33	49.41	64.29	58.33	54.17
2011/21	69.64	72.62	50.00	45.83	52.98	44.64	46.43	59.52	48.21	61.91	52.98	48.81
2011/22	44.64	54.76	52.98	51.19	57.14	52.38	55.36	58.93	55.36	64.29	56.55	55.95
2011/23	49.41	66.67	57.14	50.60	58.33	48.21	51.19	58.33	47.02	64.29	57.14	50.60
2011/24	44.05	53.57	44.64	45.24	51.79	43.45	47.02	54.17	46.43	60.12	55.95	47.62
2011/25	48.81	51.19	64.88	64.88	72.02	66.07	70.24	67.26	64.88	71.43	71.43	68.45
2011/26	41.07	47.62	52.38	47.62	55.36	47.62	50.60	53.57	53.57	61.91	54.76	51.79
2011/27	46.43	48.81	53.57	44.64	54.17	45.83	49.41	53.57	55.36	60.12	52.98	48.21
2011/28	33.93	32.14	33.33	27.98	39.29	30.36	32.74	36.91	32.14	46.43	44.05	36.91
2011/29	36.91	38.69	35.71	41.67	48.81	35.12	36.91	44.64	38.10	58.33	49.41	44.64
2011/30	36.31	42.26	57.74	54.17	61.31	50.00	52.38	63.69	51.79	75.00	64.88	58.93
2011/31	31.55	40.48	45.83	42.86	52.98	41.67	44.64	51.19	44.64	60.71	54.76	48.81
2011/32	44.05	46.43	32.74	33.93	41.67	33.33	34.52	35.12	39.88	49.41	44.05	39.88
2011/33	58.93	57.14	48.21	48.81	54.76	48.21	49.41	60.12	54.76	67.26	60.71	54.17
2011/34	36.91	44.64	45.24	46.43	55.36	41.67	41.67	52.38	43.45	61.31	56.55	52.98
2011/35	60.71	53.57	43.45	40.48	45.83	39.29	41.07	47.62	47.62	57.14	47.02	40.48
2011/36	55.56	30.56	36.11	34.72	43.06	34.72	38.89	43.06	34.72	50.00	45.83	40.28

Tabla 5.16: Porcentaje de muestras que quedan fuera de los intervalos de predicción al 70 % de confianza en cada semana del período de validación (01/Sept/2010-31/Ago/2011, 8760 muestras). En negrita el resultado del mejor modelo en cada semana.

5.5. Conclusiones

En este capítulo se han aplicado los dos modelos propuestos en la tesis a la predicción de los precios horarios del mercado diario de electricidad de OMEL. La intención del estudio no es sólo predecir el precio spot del mercado diario sino estimar una distribución de probabilidad que represente adecuadamente el comportamiento estocástico de los precios eléctricos. La información de partida es pública y está disponible para todos los agentes del mercado en el momento de enviar sus ofertas de compra o venta de energía, y las predicciones se han calculado de la misma manera que tienen que hacerlo los agentes, es decir, la primera hora del día se predice con un horizonte de predicción de 1, la segunda con horizonte 2 y así sucesivamente hasta la última hora del día que se predice con horizonte 24. La problemática de utilizar los modelos propuestos en este contexto mutiperíodo se ha solventado mediante simulación. Para validar el rendimiento de los modelos se han utilizado tanto las medidas habitualmente utilizadas en la literatura para evaluar la precisión de sus predicciones puntuales como las menos conocidas que evalúan la capacidad de proveer estimaciones de su función de densidad.

Para analizar las capacidades predictivas de los modelos propuestos se ha utilizado como referencia el modelo de función de transferencia con ruido ARIMA-ARCH, que presenta la misma estructura básica y es uno de los más referenciados en la literatura de precios Nogales and Conejo, 2006; Conejo et al., 2005; Nogales et al., 2002; Zareipour et al., 2006b; Misiolek et al., 2006. Se ha validado la inclusión de la previsión de producción eólica en dicho modelo como la característica más determinante a la hora de incrementar su precisión. También se ha mostrado cómo la inclusión de los términos ARCH para modelar la varianza del residuo hace que la estimaciones de las distribuciones predictivas sean mucho más fiables. La estructura doblemente estacional que presenta el modelo de referencia queda validada por el buen rendimiento alcanzado por los modelos ARIMA y de suavizado exponencial, los cuales han sido utilizados como modelos de referencia.

Los resultados del estudio realizado demuestran que los modelos propuestos mejoran las capacidades predictivas del modelo de función de transferencia con ruido ARIMA. A pesar del incremento en el número de parámetros, los resultados obtenidos con los modelos de cambio de régimen son más precisos en cuanto a predicciones puntuales se refiere. En relación a los modelos con densidades SNP, éstos constituyen una alternativa para modelar la distribución de probabilidad de los precios eléctricos que mejora enormemente la precisión de los intervalos de predicción que proporciona, sin que la precisión de sus predicciones puntuales se vea afectada.

Capítulo 6

Conclusiones, aportaciones originales y futuros desarrollos

Este último capítulo resume los trabajos realizados en la tesis. Se subrayan las aportaciones originales de la misma y se extraen las conclusiones alcanzadas tras la aplicación de las propuestas a los casos prácticos. Además, se ponen de manifiesto los temas no abordados y se presentan las posibles líneas de investigación futuras.

6.1. Resumen y conclusiones de la tesis

El elemento clave para el funcionamiento de los mercados eléctricos liberalizados es la subasta del mercado diario, en donde las empresas participantes son convocadas para que envíen sus ofertas de compra y venta de energía. De la casación de estas ofertas se obtiene tanto la programación de unidades del sistema que cubrirán la demanda de energía como su precio marginal, para cada una de las 24 horas del día siguiente. En este contexto, los agentes participantes del mercado se ven obligados a anticiparse a los movimientos de dicho precio marginal, puesto que es una variable determinante en sus procesos de toma de decisiones.

La necesidad de predicciones precisas del precio marginal abarca tanto el largo, como el medio y corto plazo. En el largo plazo, las predicciones del precio influyen en las decisiones relacionadas con el análisis de inversiones, como por ejemplo la expansión de la generación o la planificación de la distribución. Las predicciones a medio plazo (de seis meses a un año) se utilizan para la coordinación hidrotérmica, así como a la hora de hacer una gestión eficaz del riesgo asociado al precio del mercado, de manera que los agentes puedan decidir la cantidad de energía que comercializan a través de contratos bilaterales, mercados de futuros o directamente en el *pool*. Las decisiones de corto plazo conciernen a la programación diaria de la producción de las plantas y el desarrollo de las estrategias de oferta en el mercado, con el objetivo de maximizar el beneficio.

La predicción y caracterización del precio del mercado diario en un horizonte de corto plazo ha sido la principal motivación de esta tesis. Existe ya abundante literatura en donde se proponen métodos de predicción capaces de superar el rendimiento de los modelos tradicionalmente usados como modelos de referencia. Sin embargo, a menudo el objetivo ha sido tratar de predecir únicamente el precio que se obtiene de la casación de ofertas, sin prestar demasiada atención a la magnitud del posible error en la estimación o la probabilidad de obtener cualquier otro precio. La estimación de la función de densidad de los precios del mercado, y no solo de su valor esperado, constituye un objetivo más útil y ambicioso. El conocimiento de la distribución de probabilidad de los precios del mercado diario permite introducir incertidumbre en la elaboración de los escenarios utilizados para maximizar el beneficio de un agente en el mercado, y proporciona la información necesaria para una correcta gestión del riesgo que de él se deriva.

Dada la naturaleza no almacenable de la electricidad y la forma escalonada de las curvas de venta debido a la presencia de diferentes tecnologías de generación, el comportamiento de los precios resultantes de la casación de ofertas en el mercado diario resulta ser errático y complejo de predecir. Dicho comportamiento queda reflejado en un conjunto de características únicas y universales, que se resumen en: reversión a la media, presencia de componentes estacionales, diferentes patrones de comportamiento intradiarios e intrasemanales, efectos de calendario, volatilidad variable y distribuciones asimétricas con colas pesadas y *spikes*.

El punto de partida en la investigación de métodos para la predicción de precios eléctricos a corto plazo han sido los modelos de series temporales que tanto éxito han tenido en otros campos de aplicación, como son el modelo ARIMA o el de función de transferencia. A partir de éstos, la investigación ha evolucionado hacia modelos más complejos que tratan de superar las limitaciones de los primeros y son capaces de modelar las características citadas. El amplio abanico de métodos propuestos en la literatura ha sido revisado en detalle en el capítulo 2.

Tras la revisión de los métodos existentes para la predicción a corto plazo de los precios eléctricos, se han identificado dos carencias importantes que esta tesis trata de abordar. La primera de ellas es el planteamiento de un modelo de serie temporal que represente adecuadamente la distribución de probabilidad de los precios eléctricos, sin que esto suponga una merma en las capacidades de predicción del modelo. Por un lado, las distribuciones paramétricas propuestas en la literatura no se ajustan plenamente a las distribuciones empíricas obtenidas a través de los métodos de predicción existentes. La causa de esta limitación se encuentra en la forma rígida que este tipo de distribuciones impone, las cuales no permiten adaptarse de forma flexible a las series de precios de cada mercado específico. Por otro lado, las distribuciones no paramétricas basadas en estimadores kernel presentan la desventaja de no proporcionar una forma analítica de la función de densidad.

La segunda limitación de los métodos de predicción en la actualidad está relacionada con la representación del comportamiento no lineal. A pesar de que los métodos lineales, como los de regresión dinámica o función de transferencia, proporcionan buenos resultados de cara a la predicción, éstos no son capaces de describir adecuadamente las relaciones no lineales presentes en la serie del precio. Los modelos de cambio de régimen constituyen una buena alternativa para incorporar dinámicas no lineales manteniendo la estructura básica de los modelos lineales. Sin embargo, los modelos que aparecen en la literatura han sido concebidos no tanto para la predicción precisa de los precios eléctricos sino para representar fielmente sus principales características. En cuanto a las redes neuronales, éstas en teoría son capaces de aproximar cualquier función por complejas que sean las relaciones entre las entradas y la salida. Sin embargo, en la práctica presentan un alto riesgo de sobreentrenamiento y su elevado número de parámetros puede afectar sensiblemente a sus capacidades de generalización.

Las principales aportaciones de la tesis consisten en la generalización de dos aspectos del modelo de función de transferencia con ruido ARIMA, y están en línea con las limitaciones identificadas. Atendiendo únicamente a la precisión de sus estimaciones puntuales, los modelos de función de transferencia con ruido ARIMA arrojan buenos resultados. Sin embargo, estos modelos no contemplan el carácter heterocedástico de las series, la posible asimetría de las distribuciones resultantes ni la probabilidad de que aparezcan valores inusuales o extremos. Este hecho no resulta un problema si el modelo es concebido únicamente para obtener predicciones puntuales, puesto que el marco de trabajo de estimación de quasi-máxima verosimilitud proporciona una solución suficientemente robusta. Sin embargo, resulta inadecuado

si se quiere estimar de manera razonable la probabilidad de cualquier intervalo de predicción, es decir, si se quiere predecir su función de densidad. Para superar estas limitaciones, el primero de los desarrollos realizados supone una extensión del modelo que es capaz de representar más fielmente la forma específica de la distribución del proceso de precios utilizando un tipo de distribución semiparamétrica. La aplicación del modelo a las series de precios de diferentes mercados eléctricos revela un ajuste de la distribución empírica de los residuos mucho mayor que el proporcionado por el tradicional ajuste gaussiano. Además, los resultados de la aplicación del modelo a un conjunto de precios fuera de muestra (*out-of-sample*) demuestran que los intervalos de predicción proporcionados son sensiblemente mejores que los proporcionados por el modelo gaussiano, sobre todo para los centiles superiores al 90 %.

El segundo modelo desarrollado surge de la necesidad de capturar correctamente el comportamiento no lineal evidenciado tanto en la respuesta de los precios a variables explicativas como en las dinámicas autorregresivas de las propias series de precios. La intención es mejorar las capacidades de predicción del modelo permitiendo que sus parámetros cambien con el tiempo según un conjunto de regímenes que están determinados por variables explicativas que determinan el comportamiento del mercado. El modelo permite una forma más flexible de la función de densidad de sus innovaciones que en el caso del modelo anterior, puesto que en este caso se trata de una mixtura de las densidades asumidas para cada régimen (que en este caso son gaussianas) con diferentes pesos en cada instante de tiempo. El modelo supone un fuerte incremento en el número de parámetros puesto que la distribución de probabilidad del régimen del sistema se modela a través de un perceptrón multicapa. Sin embargo, a pesar de este incremento, el rendimiento alcanzado tras la aplicación del modelo a los precios fuera de muestra del mercado de OMEL supera al del modelo de función de transferencia con ruido ARIMA, tanto si atendemos a las medidas que miden la precisión de las predicciones puntuales como a las que tienen como objetivo medir la precisión de las distribuciones predichas.

En el capítulo 5 se han comparado las capacidades predictivas de los dos modelos propuestos en la tesis con el modelo de función de transferencia con ruido ARIMA. La comparativa se ha realizado considerando el contexto habitual en el que se desenvuelven los agentes del mercado diario, en el que la predicción de los 24 precios horarios del día siguiente se realiza a partir de la información conocida antes de enviar las ofertas al operador del mercado. Los resultados del análisis realizado demuestran que los modelos propuestos mejoran las capacidades predictivas del modelo de función de transferencia con ruido ARIMA, tanto si atendemos a la precisión de sus estimaciones puntuales como a la de los intervalos de predicción proporcionados.

6.1.1. Aportaciones originales

La elaboración de la tesis doctoral ha tenido como objetivo contribuir en áreas de conocimiento en donde la estimación y caracterización de los precios del mercado

eléctrico resulta de gran importancia, como puede ser el diseño y optimización de estrategias de oferta o la gestión del riesgo derivado de dicho precio. Las principales aportaciones de la tesis son:

- El desarrollo de un modelo que supone una extensión semiparamétrica del modelo de función de transferencia con ruido ARIMA que incorpora un término ARCH y donde sus innovaciones siguen una densidad de probabilidad de tipo SNP (*Seminonparametric*, Gallant and Nychka (1987)). El modelo es capaz de capturar las características esenciales de los precios eléctricos: influencia de variables explicativas, estacionalidad, carácter autorregresivo y volatilidad variable en el tiempo, y trata de representar más fielmente las asimetrías y colas pesadas habitualmente presentes en las distribuciones empíricas a través de un enfoque semiparamétrico.
- El desarrollo de un modelo que supone una extensión del modelo de función de transferencia con ruido ARIMA-ARCH que permite que sus parámetros varíen con el tiempo según un conjunto reducido de regímenes no observados o ficticios, los cuales a su vez están determinados por las variables explicativas que determinan el comportamiento del mercado. La capacidad del modelo de aplicar un juego diferente de parámetros en cada instante de tiempo se consigue utilizando una representación en espacio de estados. A partir de dicha representación y un juego de parámetros fijados, el filtro descrito en Kim (1994) permite estimar la salida del modelo condicional a la información conocida hasta ese momento en cada punto de la serie temporal. La distribución de probabilidad de cada uno de los regímenes se determina a través de una red neuronal que presenta como entradas las variables fundamentales que afectan al precio. Esto permite aprender de manera automática las diferentes situaciones del mercado que determinan distintas respuestas o dinámicas del precio. Los parámetros del modelo pueden estimarse usando el algoritmo EM (Dempster et al., 1977; Hamilton, 1990), el cual permite maximizar de manera separada los parámetros que determinan la distribución de probabilidad de los regímenes de aquellos que modelan la media y varianza condicional del precio, siendo a la vez robusto frente a una inicialización aleatoria de los mismos.

Otras aportaciones de la tesis son también:

- El análisis cuantitativo del efecto de la inclusión de la producción eólica como variable explicativa a la hora de predecir el precio del mercado diario en un horizonte de corto plazo. Un trabajo del autor previo a la tesis discute el efecto de esta variable en la predicción de los precios eléctricos del mercado ibérico utilizando distintos modelos de series temporales, entre los que se encuentra por ejemplo una red neuronal de tipo feed-forward (Cruz et al. (2011)).
- La incorporación de términos de tipo ARCH para modelar la varianza del residuo en el modelo de función de transferencia con ruido ARIMA, para lo

que se ha utilizado una representación en espacio de estados anidada. Dichos términos no solo proporciona estimaciones de las distribuciones predictivas mucho más fiables sino que no empeoran la precisión de las estimaciones puntuales. Este resultado es importante puesto que existe algo de controversia sobre el impacto de dichos términos sobre la precisión de los modelos que tienen como objetivo la predicción de los precios eléctricos.

- El estudio de diversos casos ejemplo representativos correspondientes a diferentes mercados eléctricos, como son el mercado francés (POWERNEXT), el mercado de Victoria, Australia (NEM) y la interconexión Pensilvania - New Jersey - Maryland (PJM), y en mayor profundidad el mercado Ibérico (OMEL/OMIE).

La investigación ha dado lugar a las siguientes publicaciones y capítulos de libros:

- A. Cruz, A. Muñoz, J. Zamora, R. Espínola, "The effect of wind generation and weekday on Spanish electricity spot price forecasting", *Electric Power Systems Research*. vol. 81, no. 10, pp. 1924-1935, Octubre 2011.
- A. Muñoz, E.F. Sánchez-Úbeda, A. Cruz, J. Marín, "Short-term Forecasting in Power Systems: A guided Tour", en *Handbook of Power Systems II*. Editores Pardalos, P.M.; Rebennack, S.; Pereira, M.V.F y Iliadis, N.A.. Ed. Springer. Berlín, Alemania, 2010.

y a las siguientes ponencias en congresos:

- A. Cruz, A. Muñoz, "Predicción a corto plazo del precio de la electricidad con modelos semiparamétricos con cambio de régimen", XXXII Congreso Nacional de Estadística e Investigación Operativa. La Coruña, España, 14-17 Septiembre 2010.
- A. Muñoz, A. Cruz, "Day-ahead electricity price forecasting with seminonparametric regime switching models", 30th International Symposium on Forecasting. San Diego, EE.UU., 20-23 Junio 2010.
- A. Cruz, A. Muñoz, "Density forecasting of electricity spot prices with seminonparametric regime switching models", Workshop Industry and Price Forecasting (WIPFOR). París, Francia, 3-4 Junio 2010.
- A. Muñoz, A. Cruz, J. Zamora, R. Espínola, "Forecasting electricity prices with periodic dynamic regression models", 28th International Symposium on Forecasting - ISF2008. Niza, Francia, 22-25 Junio 2008.
- A. Cruz, A. Muñoz, J. Zamora, R. Espínola, J. Alonso, "Modelos lineales y no lineales de series temporales para la predicción a corto plazo del precio

de la energía en el mercado eléctrico español", II Simposio de Inteligencia Computacional, SICO'2007 (IEEE Computational Intelligence Society, SC). Zaragoza, España, 11-14 Septiembre 2007.

- R. Espínola, A. Cruz, A. Muñoz, J. Zamora, ".^A hybrid model to forecast electricity market prices", 27th International Symposium on Forecasting. Nueva York, EE.UU., 24-27 Junio 2007.

6.1.2. Futuros desarrollos

Existen líneas de mejora de los modelos propuestos que conviene comentar. En primer lugar, presentan la limitación de considerar como deterministas las variables explicativas. Puesto que entre éstas se encuentran por ejemplo las previsiones de demanda y de generación eólica, un tratamiento específico que considerase la incertidumbre o fiabilidad de dichas previsiones según su propio horizonte de predicción podría mejorar el rendimiento de los modelos, sobre todo para los horizontes más lejanos. La incertidumbre en la estimación de los parámetros es también un factor relevante a la hora de calcular los intervalos de predicción o distribuciones predictivas que no se ha considerado en el desarrollo de la tesis. A juicio del autor, éste es el principal motivo de que el tamaño de los intervalos predichos por los modelos ensayados esté subestimado.

También existen líneas de mejora en cuanto a la estructura de los dos modelos propuestos. Para lograr mayor flexibilidad en la forma de la densidades predichas por el modelo semiparamétrico de función de transferencia, los coeficientes del polinomio podrían ser dependientes de la historia de serie y/o variables explicativas (Härdle et al. (1997)). Esta flexibilidad permitiría relacionar la forma de la función de densidad con las variables fundamentales que afectan al precio sin aumentar demasiado el número de parámetros, independizando los parámetros que determinan la media y varianza del proceso de precios de los que determinan la forma de la función de densidad. Con respecto al modelo de cambio de régimen, la estructura de la red de probabilidad presentada se escogió por su robustez frente a una inicialización aleatoria de la misma. Sin embargo, es posible que estructuras más simplificadas o incluso una forma funcional no paramétrica permita mejorar las capacidades de generalización del modelo.

Por último, el marco de trabajo para la estimación de los parámetros de los modelos propuestos ha sido el de la maximización de la verosimilitud logarítmica de la muestra de estimación. Sin embargo, existen otras medidas alternativas a la verosimilitud logarítmica que pueden ser usadas también para la evaluación de la bondad de un determinado conjunto de parámetros (Gneiting and Raftery (2007)). Estas medidas están basadas en puntuar de diferentes formas cómo de verosímil es un determinado valor de la serie y cómo de concentrada está la densidad en torno a dicha muestra. El CRPS (*Continuous Ranked Probability Score*) es una medida adecuada para tal propósito que podría dar lugar a modelos que penalizaran en

mayor medida la desviación entre la esperanza de la distribución asumida y el valor de la serie realizado, y por tanto, dieran lugar a modelos que capturasen mejor el comportamiento en las colas.

Bibliografía

- Aggarwal, S., Saini, L., Kumar, A., January 2009. Electricity price forecasting in deregulated markets: A review and evaluation. *International Journal of Electrical Power & Energy Systems* 31 (1), 13–22.
- Alonso, A. M., García-Martos, C., Rodríguez, J., Sánchez, M. J., Mar. 2008. Seasonal dynamic factor analysis and bootstrap inference : application to electricity market forecasting. *Statistics and Econometrics Working Papers ws081406*, Universidad Carlos III, Departamento de Estadística y Econometría.
- Amjady, N., 05 2006. Day-ahead price forecasting of electricity markets by a new fuzzy neural network. *IEEE Transactions on Power Systems* 21 (2), 887–896.
- Andalib, A., Atry, F., 03 2009. Multi-step ahead forecasts for electricity prices using narx: A new approach, a critical analysis of one-step ahead forecasts. *Energy Conversion and Management* 50 (3), 739–747.
- Aoki, M., 1986. *State space modeling of time series*. Springer-Verlag New York, Inc., New York, NY, USA.
- Baillie, R. T., Bollerslev, T., 1992. Prediction in dynamic models with time-dependent conditional variances. *Journal of Econometrics* 52 (1-2), 91–113.
- Barquín, J., Centeno, E., Reneses, J., 10 2005. Stochastic market equilibrium model for generation planning. *Probability in the Engineering and Informational Sciences* 19 (4), 533–546.
- Bastian, J., Zhu, J., Banunarayanan, V., Mukerji, R., 1999. Forecasting energy prices in a competitive market. *IEEE Computer Applications in Power* 12 (3), 40–45.
- Batlle, C., Barquín, J., 02 2005. A strategic production costing model for electricity market price analysis. *IEEE Transactions on Power Systems* 20 (1), 67–74.
- Bengio, Y., Frasconi, P., 1995. An input output hmm architecture. In: *Advances in Neural Information Processing Systems*. MIT Press, pp. 427–434.
- Bessec, M., Bouabdallah, O., Mar. 2005. What causes the forecasting failure of markov-switching models? a monte carlo study. *Econometrics* 0503018, EconWPA.

- Bollerslev, T., 04 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31 (3), 307–327.
- Bollerslev, T., Sep. 2008. Glossary to arch (garch). CREATES Research Papers 2008-49, School of Economics and Management, University of Aarhus.
- Bollerslev, T., Wooldridge, J., 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews* 11 (2), 143–172.
- Borak, S., Weron, R., 2008. A semiparametric factor model for electricity forward curve dynamics. Mpra paper, University Library of Munich, Germany.
- Box, G. E. P., Jenkins, G., 01 1970. Time series analysis, forecasting and control. Holden-Day, Place of Publication: San Francisco, CA, USA. Country of Publication: USA.
- Bunn, D. W., Karakatsani, N. V., 2003. Forecasting electricity prices.
- Cao, R., Hart, J., Saavedra, n., 2003. Nonparametric maximum likelihood estimators for ar and ma time series. *Journal of Statistical Computation and Simulation* 73 (5).
- Carnero, M. A., Koopman, S. J., Ooms, M., 2003. Periodic heteroskedastic regarfima models for daily electricity spot prices.
- Carnero, M. A., Koopman, S. J., Ooms, M., 2010. Multivariate modeling of hourly electricity prices: empirical evidence from european markets. Workshop Industry and Price Forecasting (WIPFOR), Paris, France.
- Chan, K. F., Gray, P., Apr 2006. Using extreme value theory to measure value-at-risk for daily electricity spot prices. *International Journal of Forecasting* 22 (2), 283–300.
- Chang, I., Tiao, G. C., Chen, C., May 1988. Estimation of time series parameters in the presence of outliers. *Technometrics* 30, 193–204.
- Chatfield, C., 1993. Calculating interval forecasts. *Journal of Business & Economic Statistics* 11 (2), 121–35.
- Chatfield, C., 2001. Time Series Forecasting. Chapman and Hall/CRC Press.
- Chen, J., Deng, S.-J., Huo, X., 08/01 2008. Electricity price curve modeling and forecasting by manifold learning. *IEEE Transactions on Power Systems* 23 (3), 877–888.
- Cheng, H., Tan, P.-N., Gao, J., Scripps, J., 2006. Multistep-ahead time series prediction. In: Ng, W.-K., Kitsuregawa, M., Li, J., Chang, K. (Eds.), *Advances in Knowledge Discovery and Data Mining*. Vol. 3918 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 765–774.

- Conejo, A. J., Plazas, M. A., Espínola, R., Molina, A. B., 05/01 2005. Day-ahead electricity price forecasting using the wavelet transform and arima models. *IEEE Transactions on Power Systems* 20 (2), 1035–1042.
- Contreras, J., Espínola, R., Nogales, F. J., Conejo, A. J., 08/01 2003. Arima models to predict next-day electricity prices. *IEEE Transactions on Power Systems* 18 (3), 1014–1020.
- Cruz, A., Muñoz, A., Zamora, J. L., Espínola, R., 2011. The effect of wind generation and weekday on spanish electricity spot price forecasting. *Electric Power Systems Research* 81, 1924–1935.
- Cuaresma, J. C., Hlouskova, J., Kossmeier, S., Obersteiner, M., 01 2004. Forecasting electricity spot-prices using linear univariate time-series models. *Applied Energy* 77 (1), 87.
- De Jong, P., 1991. The Diffuse Kalman Filter. *The Annals of Statistics* 19 (2), 1073–1083.
- de Jong, P., Penzer, J., 10/15 2004. The arma model in state space form. *Statistics & Probability Letters* 70 (1), 119–125.
- Dempster, A. P., Laird, N. M., Rubin, D. B., 1977. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society, Series B* 39 (1), 1–38.
- Diks, C., Panchenko, V., van Dijk, D., August 2011. Likelihood-based scoring rules for comparing density forecasts in tails. *Journal of Econometrics* 163 (2), 215–230.
- Durbin, J., Koopman, S. J., 2001. *Time Series Analysis by State Space Methods*. Oxford University Press.
- Efron, B., Tibshirani, R. J., 1993. *An Introduction to the Bootstrap*. Chapman & Hall, New York.
- Eitrheim, O., Terasvirta, T., September 1996. Testing the adequacy of smooth transition autoregressive models. *Journal of Econometrics* 74 (1), 59–75.
- Engle, R., Fall2001 2001. Garch 101: The use of arch/garch models in applied econometrics. *Journal of Economic Perspectives* 15 (4), 157–168.
- Engle, R. F., July 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica* 50 (4), 987–1007.
- Eydeland, A., Wolyniec, K., December 2002. *Energy and Power Risk Management: New Developments in Modeling, Pricing and Hedging*. Wiley.
- Franses, P. H., Paap, R., 2004. *Periodic time series models*. *Advanced Texts in Econometrics*; Oxford and New York;; Oxford University Press.

- Gallant, A. R., Nychka, D. W., 1987. Semi-nonparametric maximum likelihood estimation. *Econometrica* 55 (2), 363–390.
- Gallant, A. R., Tauchen, G., 1990. A nonparametric approach to nonlinear time series analysis: estimation and simulation. In: University of Minnesota, Minneapolis.
- Gallant, R., Tauchen, G., September 1989. Seminonparametric estimation of conditionally constrained heterogeneous processes: Asset pricing applications. *Econometrica* 57 (5), 1091–1120.
- Gao, C., Bompard, E., Napoli, R., Cheng, H., 8/1 2007. Price forecast in the competitive electricity market by support vector machine. *Physica A: Statistical Mechanics and its Applications* 382 (1), 98–113.
- Gao, F., Guan, X., Cao, X.-R., Papalexopoulos, A., 01/01 2000. Forecasting power market clearing price and quantity using a neural network method. In: 2000 Power Engineering Society Summer Meeting (Cat. No.00CH37134). Vol. 4. Syst. Eng. Inst., Xi'an Jiaotong Univ. China; DOI: 10.1109/PESS.2000.866984., IEEE Power Eng. Soc. Power Eng. Soc, Place of Publication: Piscataway, NJ, USA; Seattle, WA, USA. Country of Publication: USA.
- García, R. C., Contreras, J., van Akkeren, M., García, J. B. C., 05/01 2005. A garch forecasting model to predict day-ahead electricity prices. *IEEE Transactions on Power Systems* 20 (2), 867–874.
- García-Martos, C., Rodríguez, J., Sánchez, M. J., 05/01 2007. Mixed models for short-run forecasting of electricity prices: Application for the spanish market. *IEEE Transactions on Power Systems* 22 (2), 544–552.
- Gardner, E. S., /01Jan-Mar85 1985. Exponential smoothing: The state of the art. *Journal of Forecasting* 4 (1), 1–28.
- Gardner, E. S., 00 2006. Exponential smoothing: The state of the art–part ii. *International Journal of Forecasting* 22 (4), 637–666.
- Gómez, V., Maravall, A., 1994. Estimation, Prediction, and Interpolation for Nonstationary Series with the Kalman Filter. *Journal of the American Statistical Association* 89 (426), 611–624.
- Gómez, V., Taguas, D., nov 1995. Detección y corrección automática de outliers con tramo: una aplicacion al ipc de bienes industriales no energéticos. Tech. rep.
- Gneiting, T., Raftery, A. E., 2007. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association* 102, 359–378.
- Graps, A., 1995. An introduction to wavelets. *Computing in Science and Engineering* 2, 50–61.
- Hamilton, J. D., 1990. Analysis of time series subject to changes in regime. *Journal of Econometrics* 45 (1-2), 39–70.

- Hamilton, J. D., 1994. Time series analysis. Princeton; Princeton University Press.
- Hamilton, J. D., 2008. regime switching models. In: Durlauf, S. N., Blume, L. E. (Eds.), The New Palgrave Dictionary of Economics. Palgrave Macmillan, Basingstoke.
- Hannan, E. J., Rissanen, J., 1982. Recursive estimation of mixed autoregressive-moving average order. *Biometrika* 69 (1), 81–94.
- Hansen, B. E., August 1994. Autoregressive conditional density estimation. *International Economic Review* 35 (3), 705–30.
- Härdle, W., LÄ $\frac{1}{4}$ tkepohl, H., Chen, R., 1997. A review of nonparametric time series analysis. *International Statistical Review* 65 (1), 49–72.
- Harvey, A. C., Mar. 1990. Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press.
- Haykin, S., 1994. Neural Networks: A Comprehensive Foundation, 1st Edition. Prentice Hall PTR, Upper Saddle River, NJ, USA.
- Hindrayanto, I., Koopman, S. J., Ooms, M., 2010. Exact maximum likelihood estimation for non-stationary periodic time series models. *Computational Statistics & Data Analysis* 54 (11), 2641–2654.
- Hong, Y. Y., Hsiao, C. Y., 01 2002. Locational marginal price forecasting in deregulated electricity markets using artificial intelligence. *IEE Proceedings-Generation, Transmission and Distribution* 149 (5), 621–626.
- Huisman, R., 9 2008. The influence of temperature on spike probability in day-ahead power prices. *Energy Economics* 30 (5), 2697–2704.
- Huisman, R., Mahieu, R., September 2003. Regime jumps in electricity prices. *Energy Economics* 25 (5), 425–434.
- Hull, J., 2006. Options, futures, and other derivatives. Pearson Prentice Hall.
- Hyndman, R. J., Koehler, A. B., 10 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting* 22 (4), 679–688.
- Hyndman, R. J., Koehler, A. B., Ord, J. K., Snyder, R. D., 2008. Forecasting with Exponential Smoothing. The State Space Smoothing. Springer-Verlag Berlin Heidelberg.
- Janczura, J., Weron, R., 2009. Regime-switching models for electricity spot prices: Introducing heteroskedastic base regime dynamics and shifted spike distributions. Mpra paper, University Library of Munich, Germany.
- Janczura, J., Weron, R., 2010. An empirical comparison of alternate regime-switching models for electricity spot prices. Mpra paper, University Library of Munich, Germany.

- Jones, M. C., Marron, J. S., Sheather, S. J., 1995. A brief survey of bandwidth selection for density estimation 90.
- Jong, C. D., 2006a. The nature of power spikes: A regime-switch approach. *Studies in Nonlinear Dynamics and Econometrics* 10 (3), 3.
- Jong, C. D., 2006b. The nature of power spikes: A regime-switch approach. *Studies in Nonlinear Dynamics and Econometrics* 10 (3), Article 3.
- Joutz, F. L., Maddala, G. S., Trost, R. P., 1995. An integrated bayesian vector auto regression and error correction model for forecasting electricity consumption and prices 14.
- Karakatsani, N. V., Bunn, D. W., 2008a. Forecasting electricity prices: The impact of fundamentals and time-varying coefficients. *International Journal of Forecasting* 24 (4), 764–785.
- Karakatsani, N. V., Bunn, D. W., 2008b. Intra-day and regime-switching dynamics in electricity price formation. *Energy Economics* 30 (4), 1776–1797.
- Kendall, M. G., Stuart, A., 1969. *The advanced theory of statistics*, 3rd Edition. Griffin, London.
- Kim, C.-J., 1994. Dynamic linear models with markov-switching. *Journal of Econometrics* 60 (1-2), 1–22.
- Kim, C.-J., Nelson, C. R., 1999. *State-space models with regime switching: Classical and Gibbs-sampling approaches with applications*. Cambridge and London; MIT Press.
- Knittel, C. R., Roberts, M. R., 2001. An empirical examination of deregulated electricity prices.
- Kohn, R., Ansley, C. F., 1986. Estimation, prediction and interpolation for arima models with missing data. *Journal of the American Statistical Association* (81), 751–761.
- Koopman, S. J., Dordonnat, V., Ooms, M., 06 2010. Multivariate and dynamic factor models for electricity prices and loads. *Workshop Industry and Price Forecasting (WIPFOR)*, Paris, France.
- Koopman, S. J., Ooms, M., Carnero, M. A., 03 2007. Periodic seasonal reg-arfima-garch models for daily electricity spot prices. *Journal of the American Statistical Association* 102 (477), 16–27.
- Kosater, P., Mosler, K., 9 2006. Can markov regime-switching models improve power-price forecasts? evidence from german daily power prices. *Applied Energy* 83 (9), 943–958.

- Krolzig, H.-M., 2004. Predicting markov-switching vector autoregressive processes. *Journal of forecasting*.
- Levenberg, K., 1944. A method for the solution of certain non-linear problems in least squares. *Quarterly Journal of Applied Mathematics* II (2), 164–168.
- Li, G., Liu, C.-C., Lawarree, J., Gallanti, M., Venturini, A., 01/01 2005. State-of-the-art of electricity price forecasting. In: *CIGRE/IEEE PES Symposium on Congestion Management in a Market Environment*. Dept. of Electr. Eng., Washington Univ., Seattle, WA USA, IEEE, Place of Publication: Piscataway, NJ, USA; New Orleans, LA, USA. Country of Publication: USA.
- Ljung, L., 1986. *System identification: theory for the user*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
- Lora, A. T., Santos, J. M. R., Exposito, A. G., Ramos, J. L. M., Santos, J. C. R., 08 2007. Electricity market price forecasting based on weighted nearest neighbors techniques. *IEEE Transactions on Power Systems* 22 (3), 1294–1301.
- Lucia, J. J., Schwartz, E. S., 2000. Electricity prices and power derivatives. - evidence from the nordic power exchange. *Review of Derivatives Research* 5, 5–50.
- Luenberger, D., Ye, 2008. *Linear and Nonlinear Programming*, 3rd Edition. Springer.
- Lütkepohl, H., 2005. *New Introduction to Multiple Time Series Analysis*. Berlin and New York; Springer, update Code: 200603.
- Lütkepohl, H., 2010. Forecasting aggregated time series variables: A survey. *OECD Journal: Journal of Business Cycle Measurement and Analysis* (2), 1–26.
- Luukkonen, R., Saikkonen, P., Terasvirta, T., 1988. Testing linearity against smooth transition autoregressive models. *Biometrika* 75 (3), 491–499.
- Álvaro Escribano, Peña, J. I., Villaplana, P., 2002. Modeling electricity prices: International evidence.
- Mandal, P., Senjyu, T., Urasaki, N., Funabashi, T., Srivastava, A. K., 11/01 2007. A novel approach to forecast electricity price for pjm using neural network and similar days method. *IEEE Transactions on Power Systems* 22 (4), 2058–2065.
- Mandic, D. P., Chambers, J., 2001. *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*. John Wiley & Sons, Inc., New York, NY, USA.
- Marquardt, D. W., 1963. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics* 11 (2), 431–441.
- Mateo, A., 2005. Caracterización de los precios en mercados eléctricos competitivos mediante modelos ocultos de markov de entrada salida (iohmm). aplicacion a la generación de escenarios. In: *Doctoral thesis*. Universidad Pontificia Comillas de Madrid.

- Mateo, A., Muñoz, A., García-González, J., 02/01 2005. Modeling and forecasting electricity prices with input/output hidden markov models. *IEEE Transactions on Power Systems* 20 (1), 13–24.
- Mauleón, I., 2003. Financial densities in emerging markets: an application of the multivariate es density. *Emerging Markets Review* 4 (2), 197–223.
- Mauleón, I., April 2006. Modelling multivariate moments in european stock markets. *European Journal of Finance* 12 (3), 241–263.
- Mauleón, I., 2010. Assessing the value of hermite densities for predictive distributions. *Journal of Forecasting* 29 (8), 689–714.
- McLeod, A. I., 1993. Parsimony, model adequacy and periodic correlation in time series forecasting.
- Misioerek, A., Trueck, S., Weron, R., 2006. Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models. *Studies in Nonlinear Dynamics and Econometrics* 10 (3: Nonlinear Analysis of Electricity Prices).
- Morales, J., Mínguez, R., Conejo, A., Mar. 2010. A methodology to generate statistically dependent wind speed scenarios. *Applied Energy* 87 (3), 843–855.
- Mount, T. D., Ning, Y., Cai, X., 01 2006. Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. *Energy Economics* 28 (1), 62–80.
- Muñoz, A., Czernichow, T., 1998. Variable selection using feedforward and recurrent neural networks. *Engineering Intelligent Systems for Electrical Engineering and Communications* 6, 1924–1935.
- Newey, W. K., West, K. D., May 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55 (3), 703–08.
- Niimura, T., Nakashima, T., 08 2001. Deregulated electricity market data representation by fuzzy regression models. *IEEE Transactions on Systems, Man and Cybernetics: Part C - Applications and Reviews* 31 (3), 320.
- Nogales, F. J., Conejo, A. J., 04 2006. Electricity price forecasting through transfer function models. *Journal of the Operational Research Society* 57 (4), 350–356.
- Nogales, F. J., Contreras, J., Conejo, A. J., Espínola, R., 05/01 2002. Forecasting next-day electricity prices by time series models. *IEEE Transactions on Power Systems* 17 (2), 342–348.
- Pagan, A. R., Schwert, G. W., 1990. Alternative models for conditional stock volatility. *Journal of Econometrics* 45 (1-2), 267–290.

- Panagiotelis, A., Smith, M., 0 2008. Bayesian density forecasting of intraday electricity prices using multivariate skew t distributions. *International Journal of Forecasting* 24 (4), 710–727.
- Pankratz, A., 1991. *Forecasting with Dynamic Regresion Models*. Wiley Series in Probability and Mathematical Statistics, New York.
- Park, J., Sandberg, I. W., 1991. Universal approximation using radial-basis-function networks. *Neural computation* 3 (2), 246–257.
- Parzen, E., 1962. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics* 33 (3), 1065–1076.
- Peña, D., 2005. *Análisis de Series Temporales*. Alianza.
- Peña, D., Poncela, P., Apr. 2006. Nonstationary dynamic factor analysis. *Journal of Statistical Planning and Inference* 136 (4), 1237–1257.
- Pindoriya, N. M., Singh, S. N., Singh, S. K., 08 2008. An adaptive wavelet neural network-based energy price forecasting in electricity markets. *IEEE Transactions on Power Systems* 23 (3), 1423–1432.
- Pino, R., Parreno, J., Gomez, A., Priore, P., 2 2008. Forecasting next-day price of electricity in the spanish energy market using artificial neural networks. *Engineering Applications of Artificial Intelligence* 21 (1), 53–62.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., feb 2002. *Numerical Recipes in C++: The Art of Scientific Computing*. Cambridge University Press.
- Rambharat, B. R., Brockwel, A. E., Seppi, D. J., 2005. A threshold autoregressive model for wholesale electricity prices. *Journal of the Royal Statistical Society Series C* 54 (2), 287–299.
- Robinson, T. A., 2000. Electricity pool prices: A case study in nonlinear time-series modelling. *Applied Economics* 32 (5), 527–32.
- Rodriguez, A. F., Ruiz, E., 2008. Bootstrap prediction intervals in state space models. *Statistics and econometrics working papers*, Universidad Carlos III, Departamento de Estadística y Econometría.
- Rodriguez, C. P., Anders, G. J., 02/01 2004. Energy price forecasting in the ontario competitive power system market. *IEEE Transactions on Power Systems* 19 (1), 366–374.
- Rosenblatt, M., 1956. Remarks on some nonparametric estimates of a density function. *Ann. Math. Stat.* 27, 832–837.
- Schnatter, F. S., 2006. *Finite mixture and Markov switching models*. Springer Verlag.
- Silverman, B. W., 1986. *Density estimation: for statistics and data analysis*. London.

- Strikholm, B., Terasvirta, T., Jan. 2005. Determining the number of regimes in a threshold autoregressive model using smooth transition autoregressions. Working Paper Series in Economics and Finance 578, Stockholm School of Economics.
- Swider, D. J., Weber, C., 04/01 2007. Extended arma models for estimating price developments on day-ahead electricity markets. *Electric Power Systems Research* 77 (5), 583–593.
- Szkuta, B., Sanabria, L., Dillon, T., Aug. 1999. Electricity price short-term forecasting using artificial neural networks. *Power Systems, IEEE Transactions on* 14 (3), 851–857.
- Taylor, J. W., 08 2003. Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society* 54 (8), 799–805.
- Taylor, J. W., 0 2008. An evaluation of methods for very short-term load forecasting using minute-by-minute british data. *International Journal of Forecasting* 24 (4), 645–658.
- Taylor, J. W., De Menezes, L. M., McSharry, P. E., 2006. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting* 22 (1), 1–16.
- Terasvirta, T., 1990. Specification, estimation, and evaluation of smooth transition autoregressive models. University of california at san diego, economics working paper series, Department of Economics, UC San Diego.
- Terasvirta, T., van Dijk, D., Medeiros, M. C., 10 2005. Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination. *International Journal of Forecasting* 21 (4), 755–774.
- Tong, H., 1983. *Threshold models in nonlinear time series analysis*. springer-verlag, New York.
- Ventosa, M., 05 2005. Electricity market modeling trends. *Energy Policy* 33 (7), 897–913, accession Number: 6675562; Authors: Ventosa, Mariano 0.
- Vucetic, S., Tomsovic, K., Obradovic, Z., 05 2001. Discovering price-load relationships in california’s electricity market. *IEEE Transactions on Power Systems* 16 (2), 280.
- Vázquez, C., 2003. Modelos de casación de ofertas en mercados eléctricos. In: *Doctoral thesis*. Universidad Pontificia Comillas de Madrid.
- Weron, R., 2006. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Wiley, Chichester.
- Weron, R., 07 2009. Heavy-tails and regime-switching in electricity prices. *Mathematical Methods of Operations Research (ZOR)* 69 (3), 457–473.

- Weron, R., Misiorek, A., Mar. 2007. Heavy tails and electricity prices: Do time series models with non-gaussian noise forecast better than their gaussian counterparts? MPRA Paper 2292, University Library of Munich, Germany.
- Weron, R., Misiorek, A., 0 2008. Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *International Journal of Forecasting* 24 (4), 744–763.
- White, H., January 1982. Maximum likelihood estimation of misspecified models. *Econometrica* 50 (1), 1–25.
- White, H., Racine, J., 2000. Statistical inference, the bootstrap, and neural network modeling with application to foreign exchange rates. *IEEE TRANS. ON NEURAL NETWORKS* 12, 657–671.
- Zareipour, H., Bhattacharya, K., Canizares, C. A., 2006a. Forecasting the hourly ontario energy price by multivariate adaptive regression splines. ID: 1.
- Zareipour, H., Canizares, C. A., Bhattacharya, K., Thomson, J., 11/01 2006b. Application of public-domain market information to forecast ontario’s wholesale electricity prices. *IEEE Transactions on Power Systems* 21 (4), 1707–1717.
- Zhang, G., Eddy Patuwo, B., Y. Hu, M., March 1998. Forecasting with artificial neural networks:: The state of the art. *International Journal of Forecasting* 14 (1), 35–62.
- Zhao, J. H., Dong, Z. Y., Xu, Z., Wong, K. P., May 2008. A statistical approach for interval forecasting of the electricity price. *Power Systems, IEEE Transactions on* 23 (2), 267 –276.