Universidad de Sevilla

La demanda residencial de energía eléctrica en la Comunidad Autónoma de Andalucía: un análisis cuantitativo

Ester Gutiérrez Moya

Tesis de Doctorado

Facultad de Ciencias Económicas y Empresariales

Directora: Dra. D.ª María Teresa Arévalo Quijada

2003

UNIVERSIDAD DE SEVILLA

BIBLIOTECA VIRTUAL

La Demanda Residencial de Energía Eléctrica en la Comunidad Autónoma de Andalucía: Un Análisis Cuantitativo

Tesis Doctoral presentada por Ester Gutiérrez Moya

Directora: Dra. Dª . María Teresa Arévalo Quijada

Sevilla, junio de 2003

Me gustaría expresar mi más sincero agradecimiento a la profesora Dª . María Teresa Arévalo Quijada, por su valiosa e inestimable ayuda en el desarrollo de este trabajo de tesis doctoral, así como en la evolución de mi carrera docente.

ÍNDICE

CAPÍTULO I- INTRODUCCIÓN

I.1. OBJETIVOS DEL ESTUDIO 13

I.2. ESTRUCTURA DEL TRABAJO 17

CAPÍTULO II- MODELOS PARA LA PREDICCIÓN DE LA DEMANDA DE ENERGÍA ELÉCTRICA

II.1. INTRODUCCIÓN 21

II.2. IMPORTANCIA DE LA PREVISIÓN DE LA DEMANDA DE ENERGÍA ELÉCTRICA 22

II.3. FACTORES QUE INFLUYEN EN LA PREVISIÓN DE LA DEMANDA DE ENERGÍA ELÉCTRICA 24

II.4. MODELIZACIÓN DE LA PREDICCIÓN DE ENERGÍA ELÉCTRICA 27

II.4.1. Evolución del entorno de mercado 29

II.4.2. Clasificación de los modelos según su objetivo

II.4.2.1. Modelos de energía total consumida 30

II.4.2.2. Modelos de picos de demanda 31

II.4.2.3. Modelos horarios 33

II.4.3. Métodos de previsión de la demanda

II.4.3.1. Métodos de tendencia temporal 36

II.4.3.2. Métodos de series cronológicas 37

II.4.3.3. Métodos de usuario final 38

II.4.3.4. Métodos econométricos 39

II.4.3.5. Métodos híbridos 42

II.5. FUENTES DE DATOS 43

II.6. RESUMEN DEL CAPÍTULO 45

CAPÍTULO III- ESTUDIO DE LA DEMANDA RESIDENCIAL ANUAL DE ENERGÍA ELÉCTRICA EN LA COMUNIDAD AUTÓNOMA DE ANDALUCÍA

III.1. INTRODUCCIÓN.....49
III.2. CLASIFICACIÓN DE LOS MÉTODOS DE PREDICCIÓN.....51
III.3. CLASES DE PREDICCIONES 53
III.4. EVALUACIÓN DE PREDICCIONES 56
III.5. SELECCIÓN DE VARIABLES Y MODELOS 62
III.6. MODELOS ECONOMÉTRICOS DE PREVISIÓN DE DEMANDA ANUAL DE ENERGÍA ELÉCTRICA EN LA C.A.A 71
III.7. ESTIMACIÓN BOOTSTRAP
III.7.1. Visión general del método bootstrap 151
III.7.2. Tipología de intervalos de confianza.....153
III.7.3. Predicción con bootstrap.....155
III.7.4. Aplicación práctica a través del programa *BSTP* 158
III.8. CONCLUSIONES DEL CAPÍTULO 177

CAPÍTULO IV- ESTUDIO DE LA DEMANDA RESIDENCIAL MENSUAL DE ENERGÍA ELÉCTRICA EN LA COMUNIDAD AUTÓNOMA DE ANDALUCÍA

IV.1. INTRODUCCIÓN.....185
IV.2. ESTUDIO COMPARATIVO DE MODELOS UNIVARIANTES. DATOS MENSUALES 187
IV.3. INFLUENCIA DEL FACTOR TEMPERATURA. DATOS MENSUALES
IV.3.1. Breve reseña histórica.....237
IV.3.2. Demanda base y temperatura neutra.....241
IV.3.3. Cuantificación del efecto temperatura.....242
IV.3.4. Valores de temperatura neutra más utilizados.....244
IV.3.5. El efecto temperatura en el consumo residencial de energía eléctrica en la C.A.A. 246
IV.3.6. Características de los datos mensuales.....252
IV.3.7. Análisis de los datos mensuales 261
IV.4. MODELOS DE DEMANDA EN FUNCIÓN DE LA TEMPERATURA EN LA C.A.A. DATOS MENSUALES 269
IV.5. COMBINACIÓN DE PREDICCIONES.....283

IV.6. EL FUTURO DE LOS MODELOS DE PREVISIÓN DE DEMANDA ANUAL Y MENSUAL EN LA C.A.A. 292
IV.7. CONCLUSIONES DEL CAPÍTULO 297

CAPÍTULO V- ANÁLISIS DE LA EFICIENCIA TÉCNICA DEL SECTOR ELÉCTRICO ESPAÑOL LIBERALIZADO

V.1. INTRODUCCIÓN 303
V.2. EL SECTOR ELÉCTRICO COMO OBJETO DE ESTUDIO 304
V.3. METODOLOGÍA DE ANÁLISIS DE LA EFICIENCIA TÉCNICA DEL SECTOR ELÉCTRICO ESPAÑOL
V.3.1. El concepto de eficiencia 307
V.3.2. La medición de la eficiencia 308
V.3.3. Análisis de la eficiencia técnica del sector eléctrico español 310
V.3.3.1. Aproximación paramétrica 311
V.3.3.2. Aproximación no-paramétrica a la eficiencia técnica del sector eléctrico 314
V.4. APROXIMACIÓN NO PARAMÉTRICA A LA EFICIENCIA TÉCNICA DEL SECTOR ELÉCTRICO ESPAÑOL MEDIANTE EL ANÁLISIS ENVOLVENTE DE DATOS
V.4.1. Introducción 316
V.4.2. Estudio Empírico
V.4.2.1. Las unidades objeto de estudio 318
V.4.2.2. Aplicación del análisis envolvente de datos
V.4.2.2.1. Descripción y Selección de las variables 320
V.4.2.2.2. Procedimiento para selección de modelos 324
V.4.2.3.3. Resultados empíricos 329
V.5. CONCLUSIONES DEL CAPÍTULO 341

CAPÍTULO VI-CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN

VI.1. CONCLUSIONES GENERALES 347
VI.2. FUTURAS LÍNEAS DE INVESTIGACIÓN 351

APÉNDICE 357

ÍNDICE DE TABLAS 381

ÍNDICE DE FIGURAS 389

ANEXO 395

REFERENCIAS BIBLIOGRÁFICAS 451

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

CAPÍTULO I

INTRODUCCIÓN

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

I.1. OBJETIVOS DEL ESTUDIO

El consumo de energía eléctrica es una característica relevante de las sociedades industrializadas, representando un factor básico para la producción en diversos sectores, entre los que se encuentran el industrial, el comercial y el residencial o doméstico. La demanda de electricidad ha aumentado continuamente a lo largo de las últimas décadas. El creciente nivel de vida de las sociedades industrializadas está asociado a una mayor demanda de energía eléctrica doméstica o residencial debido, entre otras causas, a la cada día mayor utilización de aparatos de calefacción y refrigeración. La proporción que la demanda doméstica o residencial, en un período determinado y en la zona geográfica servida por la compañía, representa respecto a la demanda eléctrica total depende, entre otros factores del porcentaje de la población rural frente a la urbana, de la renta de tales poblaciones, de las características climáticas de dicha zona, e incluso de la eventual celebración- en el período considerado- de acontecimientos extraordinarios de mayor o menor duración de tiempo.

El objetivo fundamental de una compañía eléctrica es satisfacer las necesidades energéticas de sus abonados (clientes) tan económicamente como sea posible, con un nivel aceptable de calidad, fiabilidad y seguridad. Estos objetivos han de lograrse en un entorno, en el que hay que contar, entre otros factores, con la competencia *directa* (entre la que se encuentran los medios sustitutivos de la electricidad y la propia competencia de otras empresas eléctricas), la competencia *indirecta* causada por la relocalización de numerosas plantas industriales en zonas de salarios más bajos o donde podían disfrutar de mayores ventajas fiscales, el cambio en el tipo de aparatos eléctricos empleados por sus clientes y la forma o frecuencia con que utilizan los mismos y la elección, por parte de algunos clientes importantes, de otras alternativas, tales como la cogeneración.

La generación de la electricidad debe satisfacer la demanda instantánea de los usuarios, ya que no es posible almacenar esta energía en cantidades considerables, lo que pone de relieve la complejidad de la predicción de la demanda eléctrica. Otros factores que dificultan el pronóstico de la demanda de la energía eléctrica son la volatilidad de los precios de los combustibles empleados en la generación de esta energía (lo que obliga a los departamentos de planificación de las empresas eléctricas a efectuar pronósticos sobre el precio futuro de tales combustibles), los incentivos sobre una utilización más efectiva de la energía eléctrica y las crecientes restricciones legales, entre las que no son las menores las medioambientales.

Los responsables de la planificación de la demanda eléctrica reconocen la existencia de ciertos límites a la exactitud de sus pronósticos, ya que siempre existe un inevitable nivel de incertidumbre y, en consecuencia, han de estar preparados para modificar sus predicciones conforme se puedan producir cambios en el mercado. Entre las medidas que tales responsables han adoptado para incrementar la flexibilidad de su planificación se encuentran la utilización de horizontes temporales más cortos para sus pronósticos, la posible incorporación de plantas generadoras de menor capacidad, la elaboración de planes de contingencia, y una información más precisa sobre el tipo de aparatos utilizados por sus clientes, así como sobre algunas características de estos últimos (número de elementos de la unidad familiar, posible utilización de termostatos, horarios de alumbrado eléctrico, etc.)

Lo anterior exige que los responsables de la planificación eléctrica tengan una actitud proactiva, es decir que se adelanten a los cambios antes de que estos tengan lugar, lo que implica la necesidad de identificar, entre otros, los siguientes elementos:

- Los objetivos estratégicos de fiabilidad en el servicio, lo cual suele hacerse mediante el establecimiento de ciertos índices siendo muy habitual la utilización del LOLP (Loss of Load Probability Load) o probabilidad de no contar con suficiente potencia del parque generador para cubrir la demanda,

la cual puede limitarse, por ejemplo, a que la mencionada incapacidad no exceda de un día en diez años.

- Las necesidades de recursos a fin de prever las posibles deficiencias de capacidad y la eventual incorporación de unidades generadoras adicionales.
- La preferencia en la utilización de los recursos.
- El valor añadido de las diferentes opciones estratégicas, contemplando posibles escenarios futuros alternativos.

Los métodos para la predicción de la demanda de energía eléctrica han sufrido una profunda modificación durante las tres últimas décadas. En el período anterior a 1970 los métodos de extrapolación de la tendencia bastaban para explicar el aumento de la energía demandada. A partir de 1970 las enormes incertidumbres que emergieron como consecuencia de la primera crisis del petróleo de 1973, hicieron necesario el desarrollo y aplicación de técnicas mucho más sofisticadas entre las que se encuentran los métodos de usuario final, los métodos econométricos y los métodos ARIMA.

Uno de los propósitos de este estudio es la formulación de modelos que contribuyan a mejorar la predicción de la demanda de energía eléctrica en el sector doméstico o residencial de la Comunidad Autónoma de Andalucía (C.A.A.). La consecución de dicho propósito se persigue mediante la construcción, estimación y validación de modelos de pronóstico a medio plazo que generalmente establecen el horizonte temporal de un año.

Los objetivos específicos de este trabajo son:

- Construir y contrastar modelos econométricos de la demanda de energía eléctrica doméstica o residencial en la Comunidad Autonómica de Andalucía (C.A.A.) basados en los datos anuales correspondientes a los años 1984-1999.
- Efectuar la predicción de dicha demanda anual basada en el modelo construido anteriormente.

- La implementación de la técnica de remuestreo denominada *bootstrap* a través de soporte informático como complemento a la estimación y previsión de la demanda.
- Construir, analizar y comparar modelos (econométricos y de series temporales) para el consumo mensual de la demanda de energía eléctrica doméstica o residencial en la Comunidad Autonómica de Andalucía (C.A.A.) basada en los datos mensuales de consumo y temperatura correspondientes a los meses del periodo enero de 1995- diciembre de 2000.
- Analizar si la predicción compuesta obtenida mediante combinación de distintas predicciones efectuadas por diferentes métodos -que utilizan datos mensuales- mejora el pronóstico de estas últimas.
- Desarrollar un estudio empírico acerca de la evolución seguida, en materia de eficiencia técnica, por los principales grupos eléctricos españoles durante el periodo 1998-2001. De esta forma, existirá una base para investigar si los modelos de previsión planteados han de estar sujetos, en el futuro, a nuevas variables de carácter desagregado o, si por el contrario, no es de esperar que varíe considerablemente la cuota de mercado de la principal empresa suministradora en la C.A.A.

I.2. ESTRUCTURA DEL TRABAJO

Para llevar a cabo estos objetivos, este trabajo está dividido en cinco capítulos, incluyendo, además, un anexo y un apéndice que constituirán un soporte para el desarrollo del mismo. A continuación, se describirá brevemente el contenido de cada uno de ellos.

Este primer capítulo tiene un carácter introductorio. El segundo capítulo se dedica a la exposición del problema de previsión de demanda de energía eléctrica, abordando su importancia, los factores que influyen en su variación y los modelos contemplados para su análisis.

En el tercer capítulo se estudian los modelos econométricos de previsión anual de demanda de energía en la Comunidad Autónoma de Andalucía. Las variables exógenas a partir de las cuales se construyen los modelos son de carácter económico, tales como los precios, el indicador de crecimiento económico y la demanda de energía retardada. Una vez estimados los modelos considerados se evalúa el carácter predictivo de los mismos, lo cual ayudará a su elección. En este tercer capítulo, se lleva a cabo la implementación de la técnica, relativamente reciente, denominada *bootstrap*, con un programa informático, *BSTP*, en entorno de ventanas creado específicamente para este trabajo. Esta técnica se utilizó inicialmente con el objetivo de conseguir medidas de la precisión de los estimadores y en este trabajo se aplica, además, para temas de inferencia.

En el cuarto capítulo se trata la demanda de energía desde un horizonte a corto-medio plazo, esto es, se desarrollan modelos de previsión de periodicidad mensual basados en la series temporales. El modelo univariante de demanda de electricidad se estima a través de distintas técnicas basadas en modelos de naturaleza no dinámicos y modelos de naturaleza dinámica. Al considerar modelos causales es necesario tener en cuenta otro tipo de variables, como, por ejemplo, la temperatura ambiental, el número de días del mes y si el año es bisiesto o no. Para ello se analiza la influencia del factor temperatura en la

demanda de energía eléctrica, y, a partir de ahí, se estudia un modelo dinámico concreto.

Todos estos modelos construidos y analizados podrían, tal vez, haber gozado de una mayor estabilidad ante condiciones del mercado constantes. Sin embargo, el 1 de Enero de 2003 tuvo lugar la liberalización completa del sector eléctrico y gasista, con la apertura al mercado doméstico o residencial. Esto puede afectar a algunas variables entre las que se encuentran los precios de la electricidad y del gas así como el producto interior bruto, ya que a partir de esta fecha los consumidores podrán contratar libremente la electricidad con cualquier suministrador. Además, a partir del año 2007, el precio máximo de la electricidad no estará regulado por el gobierno. Esto puede provocar un riesgo ya que, hasta ahora, la principal empresa suministradora en Andalucía había sido Endesa, S.A.[1]. La denominada *temperatura media* de Andalucía también podría estar afectada por los cambios demográficos experimentados por las provincias de esta Comunidad. Para conocer hasta que punto se seguirán manteniendo estas condiciones de partida se han analizado los principales grupos eléctricos de ciclo completo que operan en España.

En el quinto capítulo, se lleva a cabo el estudio de la permanencia de estas variables. Esta investigación se desarrolla a través de la técnica de Análisis por Envoltura de Datos (DEA), la cual evalúa la eficiencia técnica de los principales grupos eléctricos en España durante el periodo que lleva el sector eléctrico liberalizado.

Finalmente, se tratan las principales conclusiones obtenidas del trabajo así como las futuras líneas de investigación.

[1] Antes de 1998, denominada Compañía Sevillana de Electricidad, S.A.

CAPÍTULO II

MODELOS PARA LA PREDICCIÓN DE LA DEMANDA DE ENERGÍA ELÉCTRICA

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

II.1. INTRODUCCIÓN

Hasta la década de 1980 la predicción de la demanda de energía eléctrica había recibido escasa atención por parte de los economistas ya que el entorno económico anterior a la primera crisis del petróleo de los años 1973-74, y en particular el de las décadas de 1950 y 1960, se podía considerar relativamente estable: el bajo precio de los combustibles utilizados en la generación de energía eléctrica era bastante predecible, los tipos de interés eran moderados y no existían demasiados problemas para que los mercados de capitales proporcionasen los fondos necesarios para la expansión del sector eléctrico. Pero la mencionada crisis introdujo numerosos factores de incertidumbre que originó que los departamentos de planificación de las empresas eléctricas se viesen obligados, desde entonces, a elaborar pronósticos de demanda para los próximos cuatro o cinco años.

Los problemas de exceso de capacidad de generación de potencia, o por el contrario de capacidad insuficiente, pueden tener costes muy elevados, no sólo en términos económicos (tanto por la energía no facturada como por los perjuicios de esta índole que puedan sufrir los usuarios), sino también en la imagen que el cliente o abonado percibe en cuanto a la calidad del servicio que recibe.

La segmentación de la clientela es importante en el sector eléctrico pues sus usuarios pueden presentar características muy diferentes. No todos los factores tienen análoga relevancia en los sectores industrial, comercial y residencial, aunque el efecto *calendario* suele ser importante en casi todos los sectores. La creciente desregulación del sector eléctrico constituye un factor adicional de incertidumbre en la predicción de la demanda de energía eléctrica.

Este capítulo se encuentra organizado como sigue. En la sección II.2 se comenta la importancia de la previsión de la demanda de energía eléctrica, dado el prolongado período de tiempo necesario para la planificación y puesta en

servicio de una central eléctrica. En la sección II.3 se contemplan algunos de los principales factores que influyen en la demanda de dicha energía en los sectores industrial, comercial y residencial o doméstico. La sección II.4 se ocupa de los principales modelos y métodos que se suelen utilizar en el pronóstico de la demanda de electricidad, comentándose las principales ventajas e inconvenientes de cada uno de ellos. En la sección II.5 se detallan las fuentes datos utilizadas en este trabajo y las limitaciones que se han encontrado en la obtención directa de datos primarios, lo cual suele ser la tónica general en este tipo de estudios. Finalmente la Sección II.6 resume las principales conclusiones de este capítulo.

II. 2 IMPORTANCIA DE LA PREVISIÓN DE LA DEMANDA DE ENERGÍA ELÉCTRICA

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

La importancia de la predicción de la demanda de energía eléctrica surge, de forma obvia, de la incertidumbre asociada a una magnitud que se refiere al futuro. La mencionada predicción puede ayudar a determinar si, previsiblemente, se va a producir una carencia de capacidad generadora (y, en consecuencia, pudiera ser conveniente considerar la construcción de nuevas centrales de energía o simplemente impulsar la adopción de medidas de conservación de la energía) o, por el contrario, en el futuro existirá un exceso de capacidad que pudiera aconsejar la no utilización de parte del parque generador ya existente.

Según Kher, Sioshansi y Sorooshian(1987), *"la industria energética es un sector de capital intensivo con inversiones a muy largo plazo*". Puesto que se necesita al menos una década para planificar y construir una nueva planta generadora, una previsión correcta de la demanda de energía eléctrica es un requisito imprescindible para lograr las metas previstas de calidad y fiabilidad del servicio, ya que la creciente dependencia de la electricidad aumenta los inconvenientes causados a los consumidores si se producen deficiencias en el suministro de energía eléctrica. Baum(1998) comenta que "*los estudios de perspectiva coinciden en que la electricidad es uno de los sectores que mantendrá en el futuro una de las tasas de crecimiento más altas*".

La previsión de la demanda es una actividad esencial de los suministradores de energía eléctrica. Sin una adecuada representación de las necesidades futuras de generación eléctrica, los problemas de exceso de capacidad, o por el contrario de capacidad insuficiente, pueden tener costes sorprendentemente altos. La correcta previsión de la demanda también desempeña un importante papel en las decisiones de una compañía eléctrica respecto a qué cantidad, y en qué época, será conveniente comprar (vender) energía a otras empresas del sector.

Un argumento importante a tener en cuenta es el económico. De acuerdo con W. Labys(1999) *"Si las predicciones resultan ser demasiado bajas pueden tener lugar carencias de energía cuyos costes habitualmente son mucho mayores que el valor de la energía no suministrada. Por el contrario, si las previsiones resultan demasiado altas, los costes de oportunidad pueden ser muy elevados al tener comprometidos, de forma improductiva, cuantiosos fondos económicos durante largos períodos de tiempo*".

Es difícil que ambas desviaciones, por exceso o por defecto, no acaben repercutiendo sobre el usuario final. Si se produce una carencia de electricidad, el precio de ésta se incrementará y el abonado pagará más por la energía consumida. Si, por el contrario, las erróneas predicciones se traducen en una superabundancia de energía, los costes asociados con la clausura de algunas plantas de potencia, u otros medios de disminuir el suministro, serán trasladados al consumidor. A este respecto, cabe mencionar que, pese a la creciente preocupación medioambiental, algunos países industrializados han vuelto a considerar la opción nuclear, debate al que no escapan, por ahora de forma incipiente, algunos países de la Comunidad Económica Europea.

Según Walter Labys(1999) existen al menos tres motivos para modelar la demanda de energía. En primer lugar, el suministro razonablemente fiable de energía es vital para el funcionamiento de la economía moderna. En segundo lugar, la ampliación de los sistemas de suministro de energía requiere muchos años. En tercer lugar, las inversiones necesarias en tales sistemas son altamente

intensivas en capital, representando, en algunos países, una considerable proporción de su Producto Interior Bruto.

Por otra parte, los índices habituales de fiabilidad eléctrica miden la probabilidad de que se pueda prestar el servicio requerido y, por tanto, los programas de planificación del sector eléctrico se basan, fundamentalmente, en las predicciones que se realicen sobre las puntas de potencia y energía demandada y, de acuerdo con ellas, se pueden adoptar las eventuales decisiones de incrementar la capacidad generadora instalada. Es obvio que, cuanto más acertadas sean tales predicciones, menores serán los riesgos de incurrir en inversiones innecesarias y/o de causar insatisfacción a los usuarios.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

II.3. FACTORES QUE INFLUYEN EN LA PREVISIÓN DE LA DEMANDA DE ENERGÍA ELÉCTRICA

En la demanda de energía eléctrica conviene, frecuentemente, segmentar los kilowatios-hora consumidos en los que se observa un comportamiento similar. Aunque no existe uniformidad en su aceptación, está muy difundida la división en tres grupos de demanda: residencial, comercial e industrial aunque, a veces, resulta útil segmentar aún más las clases de clientes servidos a fin de realizar una predicción más detallada de las tendencias de los distintos tipos de usuarios. Así, por ejemplo, la tendencia del consumo de energía eléctrica de las acerías evoluciona de forma muy distinta a la de la industria alimentaria al modificarse factores externos comunes a ambas industrias tales como el precio de la electricidad, el precio del petróleo, etc.... De forma análoga en los sectores residencial y comercial, a veces, conviene hacer una clara distinción entre viviendas o locales con calefacción eléctrica y los que carecen de ella.

Como se ha comentado anteriormente, el propio precio relativo de la electricidad puede ser un factor importante en la cantidad de energía consumida por las clases de clientes anteriormente mencionadas. La influencia de otros factores puede ya no ser tan clara. Así, por ejemplo, el precio de los productos sustitutivos (petróleo, gas natural, carbón, etc..) también se suele considerar un

factor relevante en la cantidad de energía eléctrica consumida. Sin embargo, la magnitud de la renta disponible suele ser un factor más importante en el consumo del sector residencial que en el industrial.

Otros factores de considerable importancia en la demanda de energía eléctrica también pueden depender del sector considerado. Así, en el sector residencial se pueden considerar factores relevantes el número de consumidores, el precio de los electrodomésticos, el nivel de penetración en el mercado del aire acondicionado, el número de personas por vivienda y los grados diarios de calefacción o de refrigeración. También puede variar la demanda doméstica o residencial de electricidad según la hora del día, ya que se suele consumir más electricidad durante las primeras horas del día antes de producirse la incorporación a la jornada laboral, declinando posteriormente el consumo eléctrico doméstico durante el desarrollo de la misma y volviendo a aumentar dicho consumo cuando tiene lugar el regreso al hogar al finalizar la jornada laboral. Según Bartels y Fiebig(1996), el sector residencial es habitualmente uno de los que principalmente contribuye a las puntas de demanda del sistema de generación eléctrica.

En el sector comercial son factores relevantes la magnitud del espacio ocupado por las oficinas, el nivel de empleo en el sector servicios, el número de potenciales clientes en la proximidad del establecimiento comercial, el tamaño de la población escolar, los niveles de consumo de los ciudadanos, etc....

En el sector industrial se suelen considerar factores influyentes en la demanda de energía eléctrica, el nivel de producción industrial, el nivel de empleo en este sector, la productividad de los distintos establecimientos industriales, la normativa medio ambiental, etc....

No es fácil dar reglas generales de la época del año en la que las diversas compañías eléctricas alcanzan el máximo de potencia demandada o la máxima energía consumida. Entre los factores que dificultan esta generalización figuran las condiciones climáticas, socioeconómicas, política de demora de la facturación

respecto al consumo, etc.... Ni siquiera el comportamiento observado en el pasado por las mencionadas magnitudes garantiza, para una misma compañía eléctrica, la continuidad de sus características en el futuro.

La demanda de energía eléctrica está sujeta a alteraciones debido a un gran número de factores. El consumo de electricidad varía de acuerdo con la estación del año y la hora del día. Con la estación del año ya que se consume más electricidad durante el invierno y el verano que en la primavera o el otoño, estaciones estas últimas en los que el clima es más moderado. Le(1993) realizó un estudio en el que, de forma experimental, se mostraba la fuerte relación existente entre la punta de carga neta estival de varias compañías eléctricas, el nivel de renta de las zonas a las que prestaban servicio, las condiciones climáticas de dichas zonas y el nivel de penetración del aire acondicionado en los hogares. El consumo de energía eléctrica también varía según el día de la semana considerado, ya que la demanda suele ser mayor durante los días laborables que durante los fines de semana o los días festivos.

Por otra parte, la creciente desregulación del sector eléctrico en la mayoría de los países industrializados añade un elemento adicional de incertidumbre, ya que las tradicionales fronteras entre las compañías eléctricas cada día estarán menos claramente delimitadas y será el mercado el que determine el precio de la electricidad, con la consiguiente influencia de dicho precio en el consumo. Obviamente la desregulación aumenta la necesidad de una correcta predicción de la demanda de energía eléctrica para las respectivas empresas suministradoras.

II.4. MODELIZACIÓN DE LA PREDICCIÓN DE ENERGÍA ELÉCTRICA

Un sistema de predicción de la demanda eléctrica debe contar entre sus propiedades con las de optimalidad, robustez y computabilidad. Al construir un modelo para la demanda (también denominada carga) de electricidad deben considerarse las cuestiones a las que tales modelos han de responder. En este contexto, el horizonte temporal de la predicción está relacionado con los problemas a los que se enfrenta una compañía eléctrica. Básicamente se pueden considerar tres horizontes temporales para la predicción de la demanda de energía eléctrica:

- A *corto plazo*, período que a veces se suele subdividir en *a muy corto plazo* y el *corto plazo propiamente dicho.* Se suele considerar que el primero de tales períodos abarca los próximos 30 minutos en tiempo real a partir del momento en que se efectúa la predicción de la demanda. El principal objetivo a cubrir en este periodo de tiempo es la distribución, de la forma más económica posible, entre los diversos recursos de generación disponibles. Realmente se trata de realizar el seguimiento de la carga y de la predicción inmediata de ésta basándose en los datos de la demanda durante las pasadas 24 horas.

El *corto plazo propiamente dicho* se extiende desde una hora hasta una semana en el futuro y, en este período, la empresa eléctrica debe ocuparse principalmente del mantenimiento previamente establecido y de la operación diaria de su equipo generador, la coordinación hidrotérmica, el *dispatching* económico, la asignación de unidades generadoras y la gestión de la carga. Los datos en los que se basa la predicción en este período son la carga diaria y la información de la temperatura en la misma época de los años recientes del pasado. La predicción de la punta de potencia diaria suele tener lugar también para la próxima semana. En este horizonte temporal de corto plazo es importante, tanto una predicción lo más exacta posible de la demanda para determinar qué

generadores deberían ser puestos en funcionamiento y cuales deberían permanecer en reserva, como una adecuada predicción de los factores climáticos (temperatura, humedad, etc...) y de su variabilidad.

- Por *medio plazo,* se suele entender el período de predicción que se extiende en el futuro desde un mes hasta un año a partir del momento en que se efectúa el pronóstico. Esta predicción es necesaria, generalmente, para establecer el calendario de mantenimiento de las plantas generadoras y del sistema de transmisión. Los datos básicos para realizar esta predicción son los datos mensuales de puntas de carga a lo largo de varios años, así como la energía demandada y la temperatura registrada en ellos. Se tienen en cuenta los indicadores socioeconómicos y los picos de carga. La energía demandada mensualmente se predice en el horizonte temporal de un año. En este período las compañías eléctricas se enfrentan al problema de cubrir una estructura de la demanda cualitativa y cuantitativa potencialmente cambiante que han de satisfacer con una capacidad generadora que es esencialmente fija o determinada.

- Se suele denominar *largo plazo* al período de predicción que abarca desde uno a diez años en el futuro. Puesto que el tiempo necesario para planificar, construir, probar y poner en funcionamiento nueva capacidad generadora puede oscilar entre tres y diez años, un modelo econométrico puede resultar muy adecuado en este horizonte temporal de largo plazo y, cuanto más exactas sean sus predicciones, mayor es la probabilidad de satisfacer los picos de carga (puntas de demanda) y mejorar el *factor de carga*[2]. Esta predicción de largo plazo se necesita, generalmente, para la planificación del sistema generador y del sistema de transmisión o transporte, ya que una modelización y predicción adecuadas a largo plazo de la demanda puede anticipar la capacidad de generación óptima y la combinación o mezcla de potencia generadora más conveniente con que debe contar- posiblemente mediante nuevas adquisiciones- el parque generador. Los datos para la elaboración de la demanda a largo plazo suelen ser, además de los

[2] *"El factor de carga durante un período determinado se define como la razón de la potencia media demandada a la potencia máxima durante dicho período".*

mencionados anteriormente (pero con periodicidad anual), el precio de la electricidad, el de los productos sustitutivos y la evolución demográfica, así como la de los indicadores económicos más relevantes.

II.4.1. Evolución del entorno del mercado

En el entorno estable de las décadas 1950 y 1960, las técnicas de predicción simples (tales como la mera extrapolación de la tendencia) eran suficientes para proyectar los futuros picos de carga o la energía consumida. Estas técnicas tan simples proporcionaban, en las décadas mencionadas, una respuesta adecuada a las cuestiones anteriormente planteadas debido, entre otras razones, a la considerable estabilidad de los precios en general (y en particular de los combustibles empleados en la generación eléctrica), al relativamente bajo coste del capital, a la economías de escala que parecían sugerir plantas generadoras cada vez mayores en un contexto en el que parecía normal que la demanda de energía eléctrica creciera anualmente alrededor de un 7% y a una tendencia demográfica bastante predecible en muchas zonas geográficas. Entre los técnicos de planificación de las compañías eléctricas estaba muy extendida la idea, no sólo de que la demanda aproximadamente se duplicaba cada diez años, sino también la de que, hasta cierto punto, el sector era prácticamente ajeno a la ley de la oferta y la demanda. Una consecuencia de todo lo anterior es que, en las décadas mencionadas, no hubiese necesidad de dedicar gran atención a analizar los factores subyacentes en la demanda de la energía eléctrica.

A partir de 1970 se produjo un creciente interés en la modelización econométrica de la demanda eléctrica, debido entre otras razones, a la primera crisis del petróleo en 1973 y a la consiguiente severa recesión económica mundial del bienio de 1974-75, que, en algunos países, entre ellos España, por diversas razones, se prolongó bastante más tiempo. Esta disminución en las tasas de crecimiento económico indujo políticas de conservación energética y modificaciones en el uso final de la energía eléctrica. La gran volatilidad de los precios de la energía en los años posteriores a 1973 resultó tan inesperada como

la repentina elasticidad de la demanda de los usuarios a los crecientes precios de la energía eléctrica. Los técnicos de planificación eléctrica tuvieron que plantearse varias alternativas, entre las que se encontraban, continuar o no, como hasta entonces, con las plantas convencionales de fuel, incrementar la producción de las plantas nucleares o dedicar mayores cantidades de dinero a la investigación sobre las energías renovables. La mayoría de los estudios de esta época se encuentran citados en los trabajos exhaustivos de Taylor(1975), Taylor(1976) y Einhorn(1980).

La década de 1990 trajo la ola de la desregulación que, teóricamente, rompería el casi monopolio que, de hecho, había supuesto la industria de la energía eléctrica durante décadas en la mayoría de los países. La idea básica pasó a ser que la electricidad, en el futuro, sería tratada como cualquier otra mercancía, en la que el comprador se abastecería del proveedor que mejor precio ofertase. En este nuevo escenario de creciente incertidumbre, las compañías eléctricas tendrían cada vez mayor dificultad para seguir sobrecargando a sus clientes, o abonados, con los costes derivados de decisiones sobre inversiones inapropiadas como consecuencia de predicciones erróneas de la demanda.

II.4.2. Clasificación de los modelos según su objetivo

En cuanto a sus objetivos se pueden distinguir, en principio, tres categorías: modelos de energía total consumida (kWh consumidos), modelos de picos de carga y modelos horarios.

II.4.2.1. Modelos de energía total consumida

La categoría de modelos de energía total consumida incluye los modelos en los que la variable dependiente es el consumo total a lo largo de un determinado período de tiempo, el cual, frecuentemente se refiere a un año, a un trimestre o a un mes, puesto que periodos inferiores de tiempo (y menos aún los que corresponden al área del *dispatching*) no son objeto en este estudio. Casi todos estos modelos fijan su atención bien en la demanda residencial, bien en la

demanda comercial o bien en la demanda industrial, aunque algunos modelos estudian magnitudes agregadas sobre varios de estos sectores.

Es importante situar estos modelos en su perspectiva temporal adecuada ya que, al nivel de ciertas empresas eléctricas individuales, la utilidad de la información que suministran al planificador o al analista es muy limitada. Así, por ejemplo, dos compañías eléctricas con similares predicciones de demanda anual, o incluso mensual, pueden tener curvas de carga diarias muy diferentes entre sí. Sin embargo, estos modelos pueden resultar útiles al nivel agregado regional o nacional de energía total consumida.

II.4.2.2. Modelos de picos de demanda

Estos modelos tienen un considerable interés para las empresas eléctricas, ya que las puntas de carga (o picos de demanda) constituyen el fundamento de la planificación de capacidad generadora, y del calendario futuro de mantenimiento. En este aspecto se pueden considerar los siguientes enfoques:

- El primero es un modelo econométrico simple del tipo del de Spann y Beauvais (1977) en el que la punta de demanda se suele considerar como el resultado de un componente *base* y un componente debido a la temperatura, de modo que podemos escribir:

$$PD = \beta + \alpha T$$

  siendo PD la punta de demanda, β la carga *base*, T es una medida de temperatura y α es un parámetro. A veces, los modelos que responden a esta formulación genérica se suelen extender expresando β como una función de variables económicas y demográficas y T como una variable que puede incluir a otras variables climáticas tales como la temperatura de periodos precedentes (variable retardada),la velocidad del viento, la humedad, etc....

- Un segundo enfoque en la modelización de los picos de demanda se basa en la utilización del análisis de series temporales, frecuentemente a través de la familia de modelos ARIMA, los cuales suelen presentarse con diverso

grado de complejidad según se trate de modelos univariantes o multivariantes, no existiendo evidencia empírica concluyente de que la mayor complejidad del modelo utilizado implique necesariamente mayor exactitud en la capacidad predictiva de la demanda. En este sentido Abraham,B.(1983) planteó un modelo para la previsión de demanda de electricidad mensual en el estado de Iowa (E.E.U.U.):

$$(1-B^{12})z_t = \theta_0 + (1-\theta B)(1-\Theta B^{12})a_t$$

donde B es el operador retardo ($B^p z_t = z_{t-p}$)

a_t es una variable aleatoria con media cero y varianza constante.

θ es el coeficiente del polinomio de retardo regular de media móvil.

Θ es el coeficiente del polinomio de retardo regular de media móvil estacional.

Aunque los enfoques mencionados pueden proporcionar resultados adecuados para horizontes temporales de corto-medio plazo con crecimiento económico estable, suelen omitir el siguiente aspecto importante desde el punto de vista técnico, si bien no tenido suficientemente en cuenta desde una óptica estadística: no debe olvidarse que la *base* o parte sistemática da la carga, no es sino el promedio de la respuesta de un numeroso grupo de agentes consumidores a un conjunto particular de condiciones económicas, sociales y climáticas.

A pesar de que la punta de demanda no puede por menos de reflejar el comportamiento sistemático de la carga, es muy probable que la punta de carga represente, al menos parcialmente, también el valor de una perturbación extrema. Así, por ejemplo, puede ser de interés no sólo realizar la "mejor" predicción para la punta de demanda, sino determinar asimismo cual es la probabilidad de tener que cubrir eventualmente un pico de carga que supere en un cinco o un diez por ciento al valor pronosticado. Este caso, supuestos modelizados los picos de carga, pertenecería al ámbito de la Teoría de Extremos.

Aunque al haber mencionado la "mejor" predicción para la punta de demanda se supone la existencia de un criterio plausible de bondad de ajuste (por ejemplo, el mínimo-cuadrático) en este caso tendría poco sentido expresar la

punta de carga como la variable dependiente de una regresión lineal que implicase que su esperanza condicionada es una función lineal de un conjunto de variables exógenas a las que habría que sumar un término de perturbación, o error, de cuyos componentes se postulara que están idéntica e independientemente distribuidos. Como observan Engle, Granger, Mitchem, y Ramanathan(1979) y Stanton y Gupta(1980) en este caso podría resultar más adecuado, para predecir la punta de carga, el desarrollo de un modelo para la demanda L_t en un instante genérico t y posteriormente considerar el pico de demanda como el $\underset{t}{máx}(L_t)$.

II.4.2.3. Modelos horarios

BIBLIOTECA VIRTUAL

La última meta que idealmente pretende conseguir el sistema de planificación de una compañía eléctrica es la predicción adecuada de la demanda para cada una de las horas del día (objeto de los habitualmente denominados modelos *horarios*), ya que si se contase con un modelo de este tipo que fuera suficientemente fiable, resultaría muy útil en la planificación no sólo para adoptar la decisión de adicionar o sustraer capacidad generadora, sino también para considerar la adecuada mezcla de capacidad generadora, elaborar el calendario detallado de mantenimiento preventivo, etc....

Los modelos horarios que aparecen en la literatura especializada pueden clasificarse generalmente en dos categorías. La primera categoría se compone de aquellos modelos cuyo propósito es analizar varios experimentos *horarios* de precios de la electricidad, realizados al nivel de clientes individuales, incluyendo en esta última expresión a los que posiblemente forman su grupo más numeroso: el formado por abonados-clientes que constituyen un colectivo determinado de consumidores, pertenezcan o no a la misma unidad familiar.

La segunda categoría de modelos horarios es la que se desarrolla, bien al nivel regional o bien al nivel de una empresa eléctrica. Una característica notable de la literatura especializada en este campo es su falta de profundidad y su tardía aparición debido, sobre todo a la falta de interés de la industria eléctrica en estos

modelos ya que hasta mediada la década de 1970 no se disponía de un conjunto mínimo de datos *históricos*. Los investigadores pioneros en este área fueron Cargill y Meyer(1971) al realizar, a nivel de distrito, una regresión de cada una de las 24 horas diarias de consumo medio de energía eléctrica sobre los precios de la electricidad y del gas natural, la renta personal y el nivel de empleo. El modelo de estos autores no ha gozado de gran aceptación en la práctica posiblemente debido a que las condiciones climáticas quedan enmascaradas u ocultas en los consumos medios, precisamente en un modelo que pretende llegar al detalle horario en la predicción de las puntas de demanda y de los factores de carga.

Otros trabajos que merecen citarse en la ambiciosa meta de construir un modelo *horario* son los desarrollados por Data Resources Incorporated (DRI) (véase Platt(1981) y Platt(1982)) y por Quantitative Economic Research Incorporated (QUERY) (véase Engle, Granger, Mitchem, y Ramanathan(1981)) y Ramanathan y Mitchem(1981).

$$S_m = \alpha w_m + \theta x_m + \varepsilon_m$$

donde S_m son las ventas de electricidad por consumidor en el mes m.

w_m es un vector de variables temporales para el mes m.

ε_m es el término perturbación aleatoria.

α, θ son los vectores de parámetros.

En ambos casos se trata de modelos que estiman los parámetros en dos etapas a fin de intentar reducir la gran cantidad de cálculo inherente a todos los modelos *horarios*. La principal diferencia entre ambos modelos radica en que mientras en el modelo DRI la carga horaria se expresa como función de “splines” cúbicas (que simulan la trayectoria diaria de consumo) y variables climáticas, en el modelo QUERY se utilizan variables binarias para tratar el irregular consumo de los fines de semana y días festivos.

Una de las principales decisiones a adoptar en los modelos *horarios* es el inevitable compromiso entre la eficiencia computacional y la eficiencia estadística. Baste tener en cuenta que, aun tratando de caracterizar la carga horaria limitada al nivel trimestral, el modelo QUERY implica, para cada hora estudiada en 32

regiones durante 13 años (1952-1974) el cálculo y análisis de resultados de 1664 regresiones separadas, sólo para la estimación de los parámetros en su primera etapa.

Por otra parte, la frecuente aparición de multicolinealidad es difícil de evitar en los modelos DRI y QUERY ya que la existencia de correlación serial (de valor tan elevado como 0,9) de primer orden posiblemente se deba a la no inclusión explícita en ambos modelos de factores que puedan influenciar la demanda a corto plazo: programas de televisión de gran audiencia (acontecimientos deportivos, etc...), cargas inusuales de la industria, combinación de efectos climáticos adversos, etc... Procedimientos correctivos típicos (ante la evidencia de correlación entre los residuos), tal como sería utilizar el método de Cochran-Orcutt, no parecen razonables para un número tan elevado de regresiones. Por otra parte, la corrección de la multicolinealidad mediante la reducción de dimesión proporcionada por el Análisis Multivariante (componentes principales) causaría la aparición de variables ortogonales cuya naturaleza ficticia dificultaría su interpretación física en el mundo real.

Otro problema añadido en los dos modelos horarios últimamente mencionados procede de la ambigüedad de las variables a incluir en las regresiones de la segunda etapa de estimación de los parámetros. Por ejemplo, en niveles de temperatura que pueden considerarse frontera entre el calor y el frío, no siempre es fácil determinar si los aparatos electrodomésticos que estan sobrecargando el consumo (o que pueden sobrecargarlo próximamente) en una determinada región geográfica tratan de proporcionar grados de calefacción o de aire acondicionado.

Desde la detallada exposición del estado del arte en esta cuestión presentada por Einhorn(1980), la literatura posterior sobre este campo apenas presenta variantes esenciales (insistiendo sobre su muy limitada y costosa realización de los estudios *horarios*) y además, al no formar parte del objeto de esta trabajo, no se insistirá sobre este tema.

II.4.3. Métodos de previsión de la demanda

Los métodos de previsión de la demanda de energía eléctrica han evolucionado considerablemente desde la década de 1970 hasta la actualidad. En general el horizonte temporal ha sido uno de los factores primordiales que han tenido en cuenta las metodologías utilizadas, pudiendo citarse entre estas últimas las siguientes: de tendencia temporal, de series cronológicas, de utilización final, econométricas y las híbridas

II.4.3.1. Métodos de tendencia temporal

Estos métodos eran casi los únicos utilizados, con carácter general, antes de la década de 1970, debido a las tendencias estables de la demanda en dicha época, pero actualmente su utilización prácticamente ha quedado reducida al muy corto plazo. En esta técnica se obtiene el ajuste de la tendencia general histórica en kWh facturados o en los kW de las puntas de potencia.

Algunas de las técnicas utilizadas en estos métodos recurren a una transformación previa de los datos de carga a fin de que el ajuste de la tendencia explique con mayor facilidad la tendencia del consumo. Otras técnicas de esta metodología suponen una tasa de crecimiento constante de la demanda, mientras que otras se basan en una curva de saturación que tiene un período inicial de crecimiento lento, seguido de un período intermedio de crecimiento rápido, al que sigue un período final en el que declina la tasa de crecimiento alcanzando la energía demandada su nivel de saturación. Dos trabajos relativamente recientes en esta línea son los de Barakat, J., Al -Qassim y Al-Rashed, S.(1992) y Barakat, J., y Al-Rashed, S.(1993) en el que se aplica una técnica de descomposición en una compañía eléctrica en la que el pico de carga tiene una alta tasa de crecimiento.

Entre las principales *ventajas* de estos métodos se pueden citar su economía y rapidez, el no necesitar un gran apoyo técnico y su mínima exigencia en cuanto a los datos requeridos. En cuanto a sus *desventajas* puede decirse que

estos métodos no son sensibles a los cambios estructurales y que no disponen de una traza histórica para interpretar las causas de los errores pasados de predicción a fin de corregir el modelo, caso de ser necesario.

II.4.3.2. Métodos de series cronológicas

Los métodos de series cronológicas efectúan la predicción de la demanda de energía eléctrica analizando la *trayectoria* de los datos históricos y proyectándola en el futuro. En particular los métodos de Box-Jenkins suelen resultar adecuados cuando existe una fuerte autocorrelación entre los datos de la serie analizada. El modelo inicial propuesto por Box y Jenkins se basa en el estudio sistemático de los procesos ARIMA en tres fases (identificación, estimación y diagnóstico).

Si el valor de la demanda actual puede expresarse sólo como combinación lineal de un número determinado de datos pasados, estamos ante un proceso *autorregresivo* (AR). En cambio, si expresamos la demanda presente sólo como combinación lineal de los errores pasados de predicción, el proceso se denomina de *medias móviles* (MA). Cuando para expresar la demanda actual necesitamos las dos combinaciones lineales mencionadas, el proceso se conoce como ARMA y en todos los casos anteriores es necesario identificar el orden de los polinomios de retardo, estimar los parámetros del modelo identificado y proceder a la validación de este último.

Cuando los datos de la serie evidencian un comportamiento no-estacionario es necesario diferenciar los datos hasta que se logre la estacionariedad del proceso, debiendo identificar también el orden de las diferencias aplicadas. Ello da lugar a los denominados modelos ARIMA. Cuando los datos indican un comportamiento estacional de la serie el método para construir el modelo es similar al que se acaba de describir.

Se puede decir que entre las principales *ventajas* de la utilización del análisis de series temporales se encuentran su bajo coste, su moderado error en el corto plazo y que sólo utiliza datos históricos de la serie que se trata de predecir,

pudiendo citarse entre sus *inconvenientes* que, al suponer *permanente* la caracterización de la *trayectoria* de los datos históricos, se está postulando implícitamente que las condiciones socioeconómicas del pasado serán invariables en el futuro.

II.4.3.3. Métodos de usuario final

En la literatura especializada no hay unanimidad en considerar a los métodos de *usuario final* como cualitativamente distintos de los métodos de *saturación de aparatos* (generalmente electrodomésticos), ya que la metodología en ambos casos consiste en hallar la energía eléctrica consumida por un determinado tipo de aparato como el resultado de multiplicar el número de usuarios del mismo y sus tasas de utilización por el consumo energético de dicho aparato, agregando a continuación tales productos sobre los diferentes tipos de aparatos electrodomésticos que se desee incluir por considerarlos relevantes desde el punto de vista de la energía consumida. Por *saturación de un aparato* se entiende el porcentaje "estable" de hogares que lo poseen después de que hayan transcurrido varios años desde la introducción en el mercado del mismo.

Otras ligeras variaciones de estos modelos, bien sea limitarse a considerar los nuevos usuarios de determinado aparato a fin de estimar sólo el incremento de la demanda debido a este último o, bien sea considerar el nivel de saturación o penetración en una determinada población del aparato eléctrico en cuestión, no modifican básicamente el enfoque de lo que fundamentalmente se trata de estimar.

Estos modelos son aplicables, en principio, a la estimación de la demanda de energía eléctrica en los sectores residencial, comercial, industrial y servicios. No obstante, a veces, cuando se trata de estimar la demanda de energía eléctrica sólo en el sector residencial se tiende más a utilizar la expresión de *estimación por saturación de aparatos.* En cambio, cuando se trata de una estimación en los demás sectores mencionados se acostumbra a emplear preferentemente la terminología de *estimación por el método de usuario final.*

Aunque estos modelos parecen inicialmente atractivos para las empresas eléctricas porque potencialmente ofrecen información útil para políticas de planificación, sin embargo, sus limitaciones son obvias. Así, por ejemplo es difícil incorporar a estos modelos el tipo de día (laborable, festivo o fin de semana) de que se trate. Además, si se pretende incluir a todos los aparatos eléctricos, los modelos pueden resultar muy laboriosos y difíciles de analizar, sobre todo en los sectores industrial y comercial, los cuales presentan menos homogeneidad que la se observa en el sector doméstico o residencial. Por otra parte, en aquellos casos en que se han realizado experimentos de seguimiento en panel de consumidores, los costes de medición y obtención de datos han sido muy elevados.

También hay que tener en cuenta que si lo que se pretende es la predicción de la demanda a largo plazo, las expresiones de la demanda para cada aparato sólo tendrán sentido si los datos correspondientes se han obtenido sobre períodos largos de tiempo en los que han podido tener lugar considerables fluctuaciones, no sólo en el desarrollo tecnológico de los aparatos, sino también en otras variables entre las que se encuentran las socioeconómicas.

Otra objeción adicional a estos modelos consiste en cuestionar las ventajas netas de la observación de aparatos concretos frente a los de un modelo econométrico general que no requiriese mediciones. En este sentido Parti(1980) ha desarrollado un modelo sustitutivo de la medición para la estimación directa de la demanda total mensual en el sector residencial.

II.4.3.4. Métodos econométricos

Los métodos econométricos se basan en ecuaciones estimadas que relacionan la demanda de electricidad con factores externos. Generalmente se considera que algunos de los modelos de este tipo pueden ser los más complejos de todos los utilizados para la predicción de la demanda de la energía eléctrica.

En estos modelos la demanda futura se determina considerando la influencia que en esta última tienen algunas variables explicativas, cuyos valores históricos

son un requisito imprescindible para la utilización de estos modelos. Entre las mencionadas variables explicativas que suelen ser objeto de consideración en estos modelos se encuentran el número de clientes, la renta de los mismos, el precio de la electricidad y de los productos sustitutivos, el nivel de actividad económica, las variables climáticas, el nivel de equipamiento de los hogares, etc....

Pese a ser los modelos econométricos herramientas predictivas muy potentes, sin embargo algunas de sus limitaciones son obvias, no siendo la menor de ellas que se pronostica el futuro basándose en relaciones entre las variables que tuvieron lugar en el pasado. Otra limitación radica en la consistencia de la relación entre las variables relevantes, ya que dicha relación puede ser causal o simplemente casual. Tampoco se puede despreciar la influencia en el consumo de energía eléctrica de variables de nueva aparición o emergentes que, en el pasado, o no existían, o no habían sido suficientemente identificadas quedando englobadas en la perturbación aleatoria.

Con frecuencia este método se aplica en modalidad uniecuacional, y en particular en la forma de un modelo de regresión lineal. El análisis de regresión proporciona una medida de la significación de los coeficientes de las variables y de la fiabilidad de la predicción. Los límites de confianza de los valores de la predicción proporcionan una indicación del riesgo asociado a la misma. El análisis de regresión consiste en formular y contrastar un modelo que estudie el impacto de las variables explicativas (independientes) sobre la variable explicada (dependiente). Los modelos de este tipo que se consideran en este trabajo han incluido inicialmente aquellas variables que, a priori, se ha supuesto que podrían tener un impacto considerable sobre la demanda de energía eléctrica, figurando entre ellas la temperatura, el precio de la electricidad, el precio de los productos substitutivos y diversos indicadores económicos.

Algunos investigadores que se han ocupado de la demanda de energía eléctrica en el sector residencial incluyen variables demográficas en los modelos que proponen. Posiblemente el primer autor que incluyó el tamaño de la unidad familiar entre las variables explicativas fue Halvorsen(1975). Cowding y Mc

Fadden(1994) incluyen entre las variables explicativas las diferencias en el tamaño de las familias y entre la población rural y urbana. En el modelo de Crow, Robinson y Squiteri(1991), figuran como variables explicativas además de determinadas variables económicas, el número medio de miembros de la unidad familiar y el número de habitaciones del hogar.

$$KWH = a_1 - a_2 PELEC + a_3 PGAS + a_4 INC + a_5 HDD + a_6 CDD + a_7 CONS + a_8 R + a_9 S$$

donde

KWH representa la energía consumida.

PELEC representa el precio de la electricidad.

INC representa el nivel de ingresos.

HDD representa los grados equivalentes diarios de calefacción.

CDD representa los grados equivalentes diarios de aire-acondicionado.

CONS representa el número de consumidores.

R representa el número de habitaciones por hogar.

S representa el tamaño familiar medio (número de personas).

Un método moderno de predicción para la demanda de energía en Arabia Saudí que incluye como regresores variables ambientales (temperatura, radiación solar y humedad relativa) y demográficas fue propuesto por Al.Garni, Zubaid y Nizami(1998).

$$E_p = E(T, H, S, P)$$

donde

E representa la medida de consumo de energía.

E_p representa la predicción de consumo de energía.

T representa la media mensual de la temperatura ambiente (ºC).

H representa la humedad relativa.

S representa la radiación solar.

P representa la presión atmosférica.

Entre las *ventajas* de los modelos econométricos se encuentran la de que pueden medir el efecto de las causas subyacentes de la tendencia y de la incertidumbre de la predicción y que los modelos pueden ser reestimados. Entre sus *desventajas* pueden citarse la de que requieren considerable destreza en la utilización de los métodos econométricos y la de que su coste puede ser relativamente elevado.

II.4.3.5. Métodos híbridos

Cada uno de los métodos anteriores puede tener una aplicación adecuada en un determinado contexto. Así por ejemplo, cuando se dispone de datos históricos que abarcan un período de quince o más años, los modelos econométricos de *regresión* pueden ser los más convenientes. En cambio si se dispone de datos de *saturación de aparatos* o de usuario final, éstos serían los métodos apropiados. En algunos casos se recurre a modelos *híbridos* o mixtos: por ejemplo, se puede utilizar un modelo de *regresión* para predecir el porcentaje de nuevos hogares que tienen aire acondicionado central, basándose en la renta disponible familiar, utilizando posteriormente el resultado obtenido en la etapa anterior en un modelo de *saturación de aparatos*.

En la práctica, existen numerosas situaciones en las que no existe un método *mejor* de efectuar la predicción de la demanda eléctrica por lo que, a veces, se recurre a elaborar una combinación de predicciones realizadas por diferentes métodos, pretendiéndose con ello obtener una predicción más robusta que las proporcionadas por cada método individual de pronóstico.

Entre las *ventajas* de los modelos híbridos se puede considerar la de que pueden utilizar lo mejor de los de los tipos econométricos y de usuario final. Entre sus *desventajas* pueden citarse la gran cantidad de datos que necesitan, lo que redunda en su alto coste.

II.5. LAS FUENTES DE DATOS

Evidentemente una correcta previsión requiere datos exactos, ya que sin ellos sus resultados no son fiables. Sin embargo, los datos disponibles en los estudios a nivel regional de la demanda de energía eléctrica no siempre son los que se requieren para la previsión a ese nivel. Esta falta de datos adecuados implica numerosos problemas de orden práctico.

En el presente trabajo se han obtenido los datos necesarios para identificar y contrastar la importancia de aquellas variables que contribuyen a explicar la demanda de energía eléctrica en el sector doméstico de la Comunidad Autónoma de Andalucía. Lamentablemente, en numerosos casos, los datos deseados sólo están disponibles a nivel nacional, pero resultan inadecuados para la previsión de la demanda de energía eléctrica doméstica en la Comunidad Autónoma de Andalucía.

En otras ocasiones no ha sido posible disponer directamente de los datos al nivel desagregado que hubiera sido deseable optándose, en estos casos, por estimaciones por varios procedimientos que han proporcionado resultados prácticamente coincidentes. En otros casos, sólo ha sido posible obtener los valores de algunas variables con frecuencia anual, pero no con frecuencia mensual, semanal o diaria. Para aquellos datos de los que sólo se ha dispuesto de su valor anual, es obvio que el reducido número de sus valores ha desaconsejado el tratamiento mediante métodos propios de las series cronológicas.

Los datos analizados en el trabajo y las fuentes están descritos al principio de los capítulos correspondientes. Los datos del capítulo III son utilizados en forma de series temporales de periodicidad anual durante los años 1984-1999. Los datos del capítulo IV se utilizan también como series cronológicas en el periodo comprendido entre enero de 1995 y diciembre de 2001. Los datos del capítulo V han sido tratados en forma de datos de panel, o lo que se conoce como

una frontera intertemporal y los datos corresponden a cifras anuales para cada uno de los principales cuatro grupos eléctricos en España durante el periodo 1998-2001.

La serie de demanda de energía está compuesta por el consumo de energía eléctrica que tiene lugar en el mercado residencial o doméstico de la Comunidad Autónoma de Andalucía. Los datos del consumo de energía proceden de los informes anuales de la Compañía Sevillana de Electricidad, S.A., así como de las consultas realizadas a la página web del Instituto de Estadística de Andalucía en la sección de Indicadores Económicos.

El precio unitario de la electricidad ha sido recogido en los informes de Unesa, y se refieren a la tarifa de consumo 2.0 y su versión *nocturna-día*, que es la más utilizada en el mercado doméstico.

Los datos del precio del gas butano fueron facilitados por la compañía Repsol Butano (Repsol YPF). Los datos suministrados fueron los correspondientes al precio de la bombona de butano de uso doméstico (12,5 Kg). Por lo tanto, se han tenido que realizar las oportunas conversiones caloríficas, esto es, de kilogramos a kilocalorías, de kilocalorías a termias, y de termias a kilovatios por hora.

Los datos del precio de gas ciudad fueron suministrados por la empresa Gas Natural. Los datos estaban expresados en termias y tuvieron que transformarse, mediante las distintas conversiones de unidades caloríficas, a kilovatios por hora.

Los datos correspondientes al producto interior bruto corresponde a los datos visualizados en la página web del Instituto de Estadística de Andalucía.

Respecto a los datos de temperatura, habría que comentar que los datos origen fueron suministrados por la Centro Meteorológico Territorial de Andalucía Occidental y Ceuta. Los datos percibidos correspondían a las temperaturas medias mensuales para cada una de las ocho provincias de la comunidad a lo largo del periodo 1995-2001. Con el objetivo de calcular una temperatura media

mensual correspondiente a la Comunidad, se han realizado estimaciones, en función del número de habitantes y el número de recibos pagados del impuesto de bienes inmuebles en cada una de las provincias de Andalucía.

La decisión de utilizar en el análisis distintos tipos de datos, ya se esté considerando la demanda anual o mensual de energía, viene dada por la restricción que esto supone en cuanto a la limitación de datos desagregados para muchas variables.

II.6. RESUMEN DEL CAPÍTULO

BIBLIOTECA VIRTUAL

En este capítulo se ha contemplado la importancia de la predicción en la demanda de energía eléctrica, comentando las consecuencias que se pueden derivar tanto del exceso de capacidad generadora (costes de oportunidad) como de su deficiencia (energía no facturada, perjuicios a los usuarios y deterioro de la imagen del suministro). Una correcta predicción de la demanda es necesaria para que los cálculos de los parámetros que representan índices de fiabilidad y calidad del servicio prestado se encuentren dentro de límites admisibles.

Se han considerado también los factores específicamente influyentes en el consumo de electricidad en los sectores industrial, comercial y residencial o doméstico. El efecto estacional tiene una gran influencia en la demanda eléctrica, tanto por sus repercusiones sobre los períodos laborables, como por los mayores consumos debidos a la calefacción y al aire acondicionado.

Asimismo se han descrito los principales objetivos de la predicción de la demanda eléctrica según el horizonte temporal de la misma, distinguiendo entre el corto, medio y largo plazo. También ha sido objeto de comentario en este capítulo la evolución del cambiante entorno del mercado eléctrico, desde sus relativamente predecibles circunstancias de las décadas de 1950 y 1960, hasta una época de mucha mayor incertidumbre como la actual y cuyos comienzos coinciden con la primera crisis del petróleo (1973-74).

Los principales modelos de predicción de la demanda de energía eléctrica han sido clasificados atendiendo a sus objetivos y así se han dividido en modelos de energía total consumida, modelos de picos de demanda y modelos *horarios.*

En cuanto a las metodologías más frecuentemente utilizadas con este fin se han comentado las de tendencia temporal, de series cronológicas (clásicas y ARIMA), de utilización final, las econométricas y las híbridas.

Finalmente, se han detallado las fuentes de los datos utilizados en este trabajo y las dificultades inherentes a la obtención directa de algunos datos de interés y se ha comentado que cuando, en algún caso, ha habido que realizar estimaciones, éstas se han llevado acabo por procedimientos alternativos con resultados prácticamente coincidentes. Asimismo se ha comentado la bibliografía correspondiente a los puntos tratados en este capítulo.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

CAPÍTULO III
ESTUDIO DE LA DEMANDA RESIDENCIAL ANUAL DE ENERGÍA ELÉCTRICA EN LA COMUNIDAD AUTÓNOMA DE ANDALUCÍA

III.1. INTRODUCCIÓN

En este capítulo se estudia la previsión de la demanda anual de energía eléctrica residencial en la Comunidad Autónoma de Andalucía. El capítulo se estructura en siete secciones, tras la primera sección introductoria, la segunda sección clasifica los métodos de predicción prestando mayor atención a los que utilizan técnicas cuantitativas. En la sección III.3 se describen las clases de predicciones que se pueden realizar a través de un modelo cuantitativo. La sección III.4 define algunos de los estadísticos más importantes para medir el grado de bondad de la predicción realizada. En la sección III.5 se contemplan tanto distintos criterios de información para la elección de las variables explicativas que forman parte de un modelo, como la discriminación entre varios modelos candidatos.

El núcleo del capítulo es la sección, III.6, en la que se aborda empíricamente lo comentado en secciones anteriores. En el presente trabajo se realiza un análisis conjunto de los sistemas de electricidad y de gas dada la influencia cada vez mayor entre ambos sectores energéticos competitivos entre sí en diversas aplicaciones. Por ello, entre las variables explicativas a seleccionar en los modelos propuestos para la demanda de energía eléctrica en la Comunidad se ha considerado el precio de la electricidad y del gas. Posteriormente se han estimado los parámetros de los modelos propuestos, habiéndose evaluado también su poder predictivo en un horizonte *futuro* de uno y dos años, años 1998 y 1999, al haber considerado como datos históricos los años 1984 a 1997. Los resultados experimentales muestran que al efectuar la predicción *ex–post* los intervalos de confianza contienen a los valores realmente observados.

El análisis desarrollado en los modelos anteriores tiene un enfoque paramétrico, apoyándose en determinadas hipótesis de las distribuciones de las perturbaciones (normalidad, independencia, etc...). Sin embargo, tales hipótesis son difícilmente justificables debido al escaso número de datos disponibles sobre

demanda doméstica anual de electricidad en la C.A.A. por lo que se han considerado métodos basados en cálculos intensivos por medio de ordenador destinados a obtener medidas de la precisión de las estimaciones, sin realizar arriesgadas hipótesis de partida acerca de la distribución de los datos. Por ello, en la sección III.7 se aplica uno de los métodos que no se basa en las mencionadas hipótesis: es el denominado método bootstrap. La aplicación práctica en el presente trabajo de dicha metodología ha sido posible mediante la elaboración de un programa informático, *BSTP*, desarrollado específicamente para este trabajo doctoral en entorno de ventanas (windows). El bootstrap es uno de los métodos denominados de computación intensiva que no parten del supuesto de que la distribución de los datos satisface determinadas hipótesis. Las tres clases de intervalos de confianza bootstrap que se calculan en este trabajo para la predicción *ex-post* de la demanda tienen una amplitud considerablemente menor que los intervalos mínimo cuadráticos, mejorando, por tanto, la precisión del pronóstico realizado. Los resultados de los intervalos de confianza construidos contienen al valor real de la demanda.

Para finalizar el capítulo se comentan las principales conclusiones del trabajo llevado a cabo en cada una de las secciones descritas anteriormente. A lo largo de todo el capítulo se comenta la bibliografía utilizada en el mismo.

III.2. CLASIFICACIÓN DE LOS MÉTODOS DE PREDICCIÓN

Una primera clasificación de los métodos de predicción es la que los divide en *cualitativos* y *cuantitativos*. Básicamente los métodos cualitativos son aquéllos en los que el pasado no proporciona una información relevante sobre la variable que pretendemos investigar, por lo que no son los métodos econométricos el instrumento de mayor ayuda en este tipo de modelos, sino la opinión de expertos que tengan buenos conocimientos acerca de dicha variable y del contexto en que pretendemos estudiarla. Puesto que el enfoque del presente trabajo es cuantitativo nos limitamos a comentar que, entre los métodos cualitativos relevantes, figuran, de menor a mayor exigencia de formalismo matemático los denominados *braimstorming* (se intentan aflorar nuevas ideas entre un grupo de expertos), *Delphi* (utilización sistemática de juicios de opinión hasta llegar a un acuerdo) y *cross-impact* (evaluación de la función de distribución de diversos escenarios futuros). En cambio, en los métodos cuantitativos, un objetivo primordial consiste en extraer toda la información posible contenida en los datos *históricos.*

Entre los métodos cuantitativos podemos distinguir, en principio, entre el análisis *univariante* y el análisis *causal.* Los modelos de series cronológicas univariantes se basan, implícitamente al menos, en la suposición de que los valores futuros de la variable de interés son una función de sus propios valores pasados o históricos.

Los métodos causales suponen, como su nombre indica, que existe una relación *causal* entre varias variables explicativas o independientes y la variable dependiente o variable en cuya predicción estamos interesados aunque tal vez, hablar de una *relación causal* sea demasiado concreto para describir el *grado de asociación* existente entre las variables explicativas y la variable dependiente. Dentro del análisis causal son de frecuente utilización los modelos econométricos, entre los que se encuentran los modelos de *regresión*

dinámica que abordan los problemas específicos que se plantean al trabajar con series temporales.

Entre los métodos univariantes cuantitativos se pueden distinguir, en principio, los métodos de *descomposición, de alisado exponencial y modelos ARIMA*. En los métodos de descomposición o de *componentes no observables*, que aún continúan usándose en la actualidad, se suelen distinguir los siguientes componentes: la *tendencia global* o variaciones a largo plazo (mayor de un año) no estrictamente periódicas, el factor *cíclico o estacional* de carácter periódico y el movimiento *irregular* no sujeto a ninguna periodicidad y que no debe confundirse con el componente *aleatorio* de carácter no previsible. Los métodos de alisado exponencial proporcionan una tendencia de carácter *local* que se actualiza en cada punto de la serie en función de las observaciones más recientes. Las técnicas ARIMA tienen como objetivo primordial identificar el modelo generador de los datos, para posteriormente estimarlo y verificarlo. Una frontera entre los métodos ARIMA univariantes y los modelos causales es el análisis de intervención de Box y Tiao(1985). Mediante los métodos causales ARIMA, los modelos de transferencia de Box y Jenkins(1976) analizan la relación dinámica de dos series.

La diferente naturaleza de los métodos mencionados tiene su reflejo en el número de datos históricos necesarios para identificar los diferentes modelos. En opinión de Jain,C.L.(2001-2002) el número de datos a utilizar en un modelo depende de la naturaleza de éste. Así, por ejemplo, en la práctica, para calcular una media móvil se suelen utilizar, al menos las observaciones de sus seis primeros períodos, mientras que en un modelo de Box-Jenkins requiere disponer al menos de 50 datos y en un modelo de regresión es habitual sugerir que el mínimo de datos sea de cinco a diez observaciones por cada variable dependiente.

III.3. CLASES DE PREDICCIONES

Ante todo se cree conveniente hacer un breve comentario sobre la utilización del término *predicción*, o de sus sinónimos *pronóstico* y *previsión*. En sentido estricto, la verdadera predicción sería la realizada a partir de la última observación *histórica*. Sin embargo, puesto que suele ser muy útil comprobar cómo se habría comportado el modelo *dentro* del período histórico si se hubiera utilizado dicho modelo para pronosticar observaciones históricas ya conocidas con certeza (lo cual es un criterio importante para juzgar *a posteriori* la validez del modelo), por extensión, es frecuente emplear también el término *predicción* para lo que, en realidad, podría ser considerado como *bondad de ajuste* a valores históricos, siendo numerosas las publicaciones (y en especial los manuales de paquetes de *software*) que utilizan indistintamente el término *predicción* tanto para el ajuste de los períodos históricos como para el cálculo correspondientes a periodos futuros. En la práctica, el contexto de utilización suele evitar cualquier equívoco sobre esta utilización en sentido amplio del término *predicción,* y que también se utiliza en adelante en este trabajo.

Se puede distinguir entre la predicción *ex-post* y la predicción *ex-ante.* Aunque no existe uniformidad en la utilización de estos términos, se puede decir que, en sentido estricto, la primera es aquella predicción que se realiza para un período de datos reales que ya se conocen, es decir, las observaciones en las variables endógenas y en las variables exógenas del periodo pronosticado ya se conocen con certeza en el periodo en que se realiza el pronóstico. Por el contrario, una predicción *ex-ante* se refiere a un período respecto al cual no se dispone de datos reales: es decir, el período en cuyo pronóstico se está interesado pertenece al futuro.

Cuando de una serie que consta de *n* observaciones se conocen los valores reales correspondientes a sus *k* períodos finales, pero se simula que no se conocen, de forma que los mismos no intervienen en la estimación de los parámetros del modelo (el cual, por tanto, se construye utilizando sólo los primeros

n-k valores de la serie), se suele decir que se *retienen* los mencionados últimos *k* valores de la serie Y_t con objeto de compararlos con la predicción que de los mismos se hace F_t ($t=n-k+1,.....,n$) utilizando el modelo anteriormente construido. Se denomina *error de predicción* del período *t* a la diferencia $Y_t - F_t$. Esta situación, que se considera por algunos autores (Cleary, J.P. y Levenbach.(1981); Cleary, J.P. y Levenbach, H.(1982); Levenbach, H. y Cleary, J.P.(1984); Jain ,C.L. y Colosante, J.(1985); Makridakis ,S., Wheelright, S.C. y Hyndman, R.J.(1998)) como una extensión de la predicción *ex-post,* es denominada por otros autores Pindyck, R.S. y Rubinfeld, D.L. (1998) *simulación ex-post*, mientras que otros (Coccari, R.L. y Galucci, C.(1984)) se refieren a ella como una predicción *simulada ex-ante.*

BIBLIOTECA VIRTUAL

También se puede distinguir entre una predicción *condicional* y otra *incondicional*. En esta última se conocen con certeza las variables explicativas del modelo, mientras que en un pronóstico *condicional* el valor de al menos una de las variables explicativas no se conoce, por lo que hay que utilizar predicciones de las variables que sean desconocidas para realizar el pronóstico de la variable endógena o dependiente.

Otro factor a tener en cuenta en cuanto a la clasificación de los métodos de predicción es el *horizonte temporal* de la misma siendo habitual distinguir entre el *inmediato* (hasta 3 meses), el *corto plazo* (de 3 a 12 meses), el *medio plazo* (de 1 a 3 años), el *largo plazo* (de 3 a 10 años) y el *muy largo plazo* (más de 10 años).

La tipología de las predicciones puede ser considerada desde otro punto de vista. En el programa Eviews, utilizado en este trabajo, se distingue entre las predicciones *dinámicas, estáticas y estructurales.* En la primera de ellas se obtienen los pronósticos que se calculan a partir de la primera observación del período de predicción utilizando valores previamente pronosticados para la variable retardada endógena (dependiente) que pudiera contener el modelo. En la predicción *estática* las predicciones se calculan secuencialmente a partir de las observaciones anteriores, pero utilizando exclusivamente valores realmente observados, por lo que sólo si se dispone de tales observaciones se puede

efectuar la predicción estática. Es obvio que en un modelo uniecuacional en el que el único retardo existente de la variable dependiente sea el de primer orden, coinciden los valores obtenidos por las predicciones dinámica y estática al pronosticar el período siguiente al último dato histórico y éste sería el último período que se podría pronosticar por el método estático. En cambio, se pueden realizar predicciones para todos los períodos futuros que se deseen por el método dinámico, ya que los residuos desconocidos se sustituyen por su media que es nula. El método *estructural,* por omisión, ignora la correlación de los errores del modelo asignando una valor nulo a tales errores aunque presenten una estructura ARMA.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

III.4. EVALUACIÓN DE PREDICCIONES

En general, la expresión *evaluación* de una predicción o de un pronóstico en un determinado periodo se utiliza para medir el grado de bondad de la predicción realizada sobre la variable de interés en comparación con el valor realmente observado para tal variable en dicho período. Es preciso recordar que se utiliza el término predicción en el sentido amplio comentado en el apartado anterior. Es frecuente hablar de la *exactitud* de una predicción como sinónimo de *evaluación* de la misma. Cuando una predicción muestra menor error predictivo que otra se acostumbra a decir que es *más exacto*, permitiéndose esta licencia en el empleo del término *exactitud* en el contexto de la predicción.

Es conveniente tener en cuenta que, en cualquier modelo, su eventual buen *grado de ajuste* respecto a periodos históricos no garantiza su buena *capacidad predictiva* respecto al futuro. Para obtener la predicción puntual de la variable endógena o dependiente en un periodo futuro se debe conocer el valor de las variables exógenas o independientes en dicho período. Para evaluar, además, la capacidad predictiva del modelo se ha de conocer también el valor observado de la variable dependiente en tal período

Una hipótesis implícita en cualquier método de predicción es que el valor observado de la variable de interés se obtiene como suma de un *patrón o modelo* y de un *término aleatorio,* lo que equivale a admitir que, aún cuando se logre identificar el modelo adecuado o trayectoria subyacente de los datos, es habitual que exista cierta desviación entre el valor predicho y el valor observado por la variable considerada. El objetivo de un método de previsión debe ser minimizar tal desviación.

La cuestión de cual es la *mejor* forma de medir la exactitud de una predicción ha sido ampliamente debatida en la literatura especializada. Puesto que el medio más objetivo de evaluar una predicción es medir el valor realmente observado de la variable de interés en el periodo pronosticado, la

forma más simplista de expresar la exactitud de una predicción es mediante el denominado *error de predicción* correspondiente al período *t,* o diferencia entre Y_t, el valor observado por la variable considerada y F_t su valor predicho, $e_t = Y_t - F_t$. Cuando no se considera la diferencia, sino el cociente, entre las dos magnitudes del segundo miembro de esta última expresión y se multiplica tal cociente por 100, se tiene lo que habitualmente se conoce como *porcentaje de consecución* en el mundo empresarial, ámbito en el que esta última expresión se utiliza más en relación con objetivos financieros o de ventas que como medida de la exactitud de una predicción.

Para evaluar un pronóstico, además de la expresión del error de predicción, se utilizan otros parámetros a los que se suele denominar *medidas del error de predicción.* Algunos de estos parámetros son función del *error de predicción* y son objeto de comentario a continuación por aparecer en los listados de ordenador que forman parte de este trabajo.

Cuando se desea tener en cuenta el número de períodos de datos, *n,* que se han utilizado en la predicción se suele utilizar el *error medio (EM)* de la predicción definido como

$$EM = \frac{1}{n}\left\{\sum_{t=1}^{n}\left(Y_t - F_t\right)\right\}$$

El inconveniente principal del parámetro *EM* radica en que su valor puede ser pequeño a pesar de existir grandes desviaciones del valor previsto respecto al realmente observado, debido a la posible compensación de las diferencias positivas entre ambos conceptos en algunos períodos con las negativas en otros dando, de esta forma, una errónea imagen global de buen ajuste del modelo a los datos reales. Para evitar esta compensación de errores de distinto signo es frecuente utilizar el parámetro *DMA (Desviación Media Absoluta)* o *MAD (Mean Absolute Deviation)* definido como el valor medio de los valores absolutos de tales desviaciones. A este parámetro también se le suele denominar *EAM (Error Absoluto Medio*) o *MAE (Mean Absolute Error)* y se

utiliza para evaluar tanto la bondad del ajuste dentro de la muestra como la exactitud de la predicción *ex-post.*

$$EAM = DMA = \frac{1}{n}\left\{\sum_{t=1}^{n}\left|Y_t - F_t\right|\right\}$$

Otra medida muy utilizada del error de la prediccion es el denominado *Error Cuadrático Medio, ECM o MSE (Mean Squared Error)* cuya expresión involucra una función cuadrática de pérdida

$$ECM = \frac{\sum_{t=1}^{n}\left(Y_t - F_t\right)^2}{n}$$

A fin de que la medida del error quede expresada en las mismas unidades de medida que la variable objeto de pronóstico, a veces se utiliza la raíz cuadrada del parámetro anterior a la que se denomina *RECM (Raiz del Error Cuadrático Medio)* que suele aparecer en la literatura especializada como *RMS o RMSE (Root Mean Squared Error*) y que se utiliza para medir la bondad del ajuste dentro de la muestra. Según lo anterior su expresión es:

$$RECM = \sqrt{ECM}$$

Excepto el *porcentaje de consecución,* las medidas de error comentadas hasta ahora presentan el inconveniente de que cualquier cambio de escala de la variable objeto de estudio afecta a la magnitud de tales medidas. Es lógico, por consiguiente, que hayan surgido medidas del error de predicción que traten de evitar tal inconveniente, relativizando las magnitudes. Entre tales medidas adimensionales figuran el *Error Relativo o Porcentaje de Error* (*PE, Percentage Error*), el *Porcentaje Medio de Error* (*MPE, Mean Percentage Error*) y posiblemente la más utilizada de todas ellas, el *EAMP (Error Absoluto Medio del Porcentaje de Error)* o, como se le suele designar en la literatura especializada, *MAPE* (*Mean Absolute Percentage Error*) y que se utiliza para evaluar tanto la bondad del ajuste dentro de la muestra como la exactitud de la predicción en los períodos extramuestrales. Estos tres parámetros se definen respectivamente como sigue:

$$PE_t = \left(\frac{Y_t - F_t}{Y_t} \right) x\ 100$$

$$MPE = \frac{1}{n} \sum_{t=1}^{n} PE_t$$

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| PE_t \right|$$

Con excepción del *ECM* y del *RECM*, cuyas funciones cuadráticas de pérdida aumentan (penalizan) los errores de predicción de gran magnitud, las demás medidas de exactitud de la predicción anteriores ponderan con igual peso a todos los errores. El *CDT* o *Coeficiente de Desigualdad de Theil* -una de las medidas de error de la predicción más utilizadas en los estudios econométricos- representa, en opinión de Makridakis, S., Wheelright, S.C. y Hyndman, R.J.(1998), un compromiso entre las medidas absolutas y relativas del error de predicción.

En lo que sigue, designamos por $\hat{Y}_t$ y Y_t respectivamente, el valor de la variable pronosticada y el de la variable observada, y por $S_{\hat{Y}}$ y S_Y, sus correspondientes desviaciones típicas, mientras que *r* es el coeficiente de correlación simple entre la serie observada y la pronosticada. La expresión del *Coeficiente de Desigualdad de Theil* es:

$$CDT = \frac{\sqrt{\dfrac{\sum_{t=1}^{n} \left(\hat{Y}_t - Y_t \right)^2}{n}}}{\sqrt{\dfrac{\sum_{t=1}^{n} \hat{Y}_t^2}{n}} + \sqrt{\dfrac{\sum_{t=1}^{n} Y_t^2}{n}}}$$

El *CDT* es adimensional, contiene una función cuadrática de pérdida a fin de penalizar los grandes errores de predicción y su valor varía entre cero y uno, indicando la mayor proximidad a cero una mejor capacidad predictiva del modelo. El cero representa el ajuste perfecto y el valor uno indica, por el

contrario, que el modelo de predicción bajo consideración no es mejor que un simple método *ingenuo* de paseo aleatorio, es decir, el modelo considerado no tiene nada que aportar en orden a la predicción.

La descomposición del índice de Theil en los tres componentes o proporciones que siguen a continuación da lugar a sendos Coeficientes *Parciales* de Desigualdad, cada uno de los cuales varía entre cero y uno, siendo la suma de los tres igual a la unidad.

Proporción del sesgo:

$$U_1 = \frac{\left(\bar{\hat{Y}} - \bar{Y}\right)^2}{\dfrac{\sum_{t=1}^{n}\left(\hat{Y}_t - Y_t\right)^2}{n}}$$

Proporción de la varianza:

$$U_2 = \frac{\left(S_{\hat{Y}} - S_Y\right)^2}{\dfrac{\sum_{t=1}^{n}\left(\hat{Y}_t - Y_t\right)^2}{n}}$$

Proporción de la covarianza:

$$U_3 = \frac{2(1-r)S_{\hat{Y}}S_Y}{\dfrac{\sum_{t=1}^{n}\left(\hat{Y}_t - Y_t\right)^2}{n}}$$

Los dos primeros Coeficientes Parciales de Desigualdad miden, respectivamente, las discrepancias entre la media y la varianza de la serie pronosticada, $\hat{Y}$, y la realmente observada Y, es decir, los errores sistemáticos del modelo. El tercer Coeficiente Parcial mide el componente residual o no sistemático de los errores de predicción y es donde, en un buen modelo, debería encontrase la mayor proporción del error total de predicción. El resultado ideal sería $U_1 = U_2 = 0$, ya que indicaría que no existen errores sistemáticos y, en este caso, se tendría $U_3 = 1$.

A fin de comparar la magnitud de los errores de un modelo respecto a un patrón se toma como base de referencia la magnitud del error del modelo ingenuo de *paseo aleatorio.* Para un horizonte determinado, al cociente entre los errores de predicción de un modelo determinado y del modelo ingenuo (paseo aleatorio) se le denomina *RAE (Relative Absolute Error).* Más importancia tiene la media geométrica de los RAE de distintos horizontes a la que se designa por *GMRAE (Geometric Mean of Relative Absolute Errors),* Así, por ejemplo, un GMRAE=0,67 indica que el error de predicción del modelo bajo consideración es sólo el 67% del error que generaría el mencionado modelo ingenuo para los mismos datos.

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

III.5. SELECCION DE VARIABLES Y MODELOS

En esta sección se comentan algunos de los criterios que se han tenido en cuenta en la selección de un modelo de entre un conjunto de varios candidatos. En el presente trabajo no se consideran modelos que *a priori* carezcan de término constante como regresor porque de la nulidad de los regresores de los modelos considerados no se deduce que, en el mundo real, tenga lugar la anulación de la variable dependiente (demanda de energía eléctrica en el sector doméstico de la C.A.A.). Es decir, no se consideran regresiones *por el origen*, modelos estos últimos en los que la suma de los residuos no tiene por qué ser nula y en los que el coeficiente de determinación, R^2, pierde su habitual significado.

Mientras no se especifiquen algunos criterios para juzgar la *bondad* o *adecuación* de un modelo, es imposible evaluar este último. Harvey, A.C.(1981) enumera los siguientes criterios para juzgar la *calidad* de un modelo:

- Parsimonia: un modelo nunca puede ser una descripción total de la realidad, ya que para describir exactamente la realidad se tendría que desarrollar un modelo tan complejo que no sería útil en la práctica. El principio de Occam enuncia que un modelo se debe mantener tan simple como sea posible, o como dijo Milton Friedman, "*un modelo es importante si explica mucho en forma resumida...*".
- Identificalidad: Significa que para un conjunto de datos determinado, los parámetros estimados deben tener valores únicos.
- Bondad de Ajuste: El objetivo principal de los modelos de regresión es lograr explicar la mayor cantidad posible de la variación de la variable dependiente utilizando variables explicativas incluidas en el modelo. Se dice que éste es *adecuado* si el coeficiente de correlación ajustado es alto.

- Consistencia Teórica: Un modelo puede tener un coeficientes de determinación ajustados elevados pese a no tener los signos de los coeficientes de regresión adecuados
- Poder Predictivo.

En la construcción de un modelo una cuestión fundamental es determinar qué variables deben formar parte del mismo**.** Este problema de *selección de variables* en los problemas de regresión lineal surge cuando se desea modelar la relación entre la variable dependiente, Y, y el conjunto de potenciales variables explicativas $X_1, X_2, ..., X_m$, existiendo incertidumbre acerca de cual es el *mejor* subconjunto que podemos elegir de entre estas últimas. Este problema es particularmente importante cuando m es grande (mayor de 10) y/o cuando entre los regresores candidatos pueden existir varios que sean irrelevantes o redundantes entre sí.

Uno de las aspectos interesantes de este problema es la amplia variedad de aproximaciones que se han intentado para resolverlo de los que Miller(1990) hizo un estudio de las soluciones que se habían propuesto hasta 1990. Aunque este es un problema cuyas primeras soluciones datan de la década de 1960, los enfoques iniciales se basaban en criterios que valoraban sólo el buen ajuste del modelo a los datos históricos, lo cual posteriormente no siempre resultaba ser una solución satisfactoria debido a la escasa capacidad predictiva del modelo elegido al tratar de utilizarlo para pronosticar los datos de períodos futuros.

Generalmente, cuantas más variables se añadan a un modelo, mejor es el ajuste *aparente* a los datos observados. Uno de los objetivos de la técnica de selección de modelos es hallar un equilibrio entre la bondad del ajuste y la complejidad del modelo. En definitiva, la calidad de un modelo debe juzgarse por su capacidad predictiva respecto a los nuevos conjuntos de datos que se obtengan del mismo proceso.

Desde el punto de vista conceptual, se puede decir que existen dos aproximaciones para resolver el problema de la selección del modelo. La solución

estándar consiste en realizar una serie de contrastes-diagnósticos en cada etapa del análisis, y aunque esta aproximación conlleva la inevitable elección subjetiva del nivel de significación al que se van a realizar tales contrastes (subjetividad en cierto modo disminuida por las elecciones habituales del 10%, 5% o 1%), al menos este procedimiento consigue evitar los errores más importantes que se cometerían si los mencionados contrastes no se llevasen a cabo.

La aproximación alternativa a la anterior es aquélla que elige el modelo que minimiza un criterio diseñado para estimar la esperanza del error cuadrático medio de predicción. De una forma o de otra, la mayoría de los criterios presentan la forma del producto de un factor de penalización por la suma cuadrática de los residuos.

BIBLIOTECA VIRTUAL

Descartada, por su frecuente elevado número, la solución de explorar todos los posibles modelos (para m potenciales variables explicativas hay 2^m posibles modelos), una forma habitual de abordar el problema es hacerlo en dos etapas, consistiendo la primera en reducir el número de modelos a considerar (por ejemplo, atendiendo a la suma cuadrática de los residuos) a un conjunto de tamaño *razonabl*e que permita a continuación, en la segunda etapa, decidir cual es el criterio por el que va a ser elegido el *mejor* modelo.

La comparación entre los diferentes criterios es una cuestión que no suele ser fácil, por lo que la elección de un modelo atendiendo a un único criterio puede ser a veces, de dudoso valor. Es por ello por lo que un criterio pragmático de selección entre varios modelos consiste en evaluar la verdadera exactitud de las predicciones realizadas por tales modelos para períodos *fuera de la muestra*, es decir, efectuar una predicción *ex-post*, lo cual se consigue dividiendo el conjunto de datos de que se dispone en dos subconjuntos, utilizando el subconjunto correspondiente a las fechas mas recientes (*hold out set* o conjunto *retenido*) para evaluar la predicción efectuada por el modelo construido con los datos correspondientes a fechas anteriores a las del mencionado subconjunto *retenido.*

Algunas estrategias de selección de modelo muy utilizadas en el pasado, tales como la de seleccionar el modelo con mayor R^2, no siempre producen buenos pronósticos. Una de las ventajas que tiene el logaritmo de la función de verosimilitud frente al coeficiente de determinación, como criterio de comparación entre modelos alternativos, es que no se ve afectado por las transformaciones que pueda sufrir la variable dependiente y, cuando es éste el criterio de selección, se elige el modelo con mayor valor de la función de verosimilitud, evaluándose ésta en el punto definido por los parámetros estimados y siendo su expresión:

$$LnL\left(e_1,\ e_2,\ ...,\ e_n\right) = -\frac{n}{2} - \frac{n}{2}Ln\left(2\pi\right) - \frac{n}{2}Ln\left(\frac{SCE}{n}\right)$$

siendo SCE la suma cuadrática de los residuos $(e_1,\ e_2,\ ...,\ e_n)$ y n el número de observaciones.

BIBLIOTECA VIRTUAL MiGUEL DE CERVANTES

El principal objetivo de la predicción debe ser hallar el modelo que proporcione los mejores pronósticos, no el que mejor se ajuste a los datos históricos, ya que éste último puede no ser el mejor modelo predictivo por diversas razones , entre las que se encuentran las siguientes:

- El futuro puede no admitir como descripción el mismo modelo probabilístico que el pasado.
- El modelo puede involucrar demasiados parámetros, de modo que el *sobreajuste* (*overfitting*) de los datos históricos puede ser debido a haber considerado sucesos extraordinarios de improbable repetición en el futuro.

La mayoría de los criterios de selección utilizados actualmente utilizados eligen el modelo que tenga el menor error de predicción *a un paso*, es decir, de horizonte unidad, si bien, con mayor generalidad, la exactitud de los diferentes métodos puede ser evaluada comparando el error cuadrático de las predicciones *fuera de la muestra* para diversos horizontes.

Algunos criterios de selección de modelo están relacionados entre sí. El error cuadrático medio (ECM), definido con anterioridad, esta ligado con la suma cuadrática de residuos y con R^2 ya que

$$R^2 = 1 - \frac{\sum_{t=1}^{n} e_t^2}{\sum_{t=1}^{n} \left(y_t - \bar{y}_t\right)^2} = 1 - \frac{ECM}{\frac{\sum_{t=1}^{n} \left(y_t - \bar{y}\right)^2}{n}}$$

Luego elegir uno de estos tres criterios equivale a elegir los otros dos. En general no es conveniente elegir según el criterio del mínimo ECM pues éste parámetro, al añadir más regresores a un determinado modelo, no puede aumentar. Así, la inclusión de más variables en el modelo disminuye la suma cuadrática de residuos y el ECM, y aumenta el R^2, aunque no necesariamente aumenta su capacidad predictiva *fuera de la muestra*. Para corregir este defecto se define el error cuadrático medio *corregido por grados de libertad* :

$$s^2 = \frac{\sum_{t=1}^{n} e_t^2}{n-k}$$

donde k es el número de grados de libertad con que se ajustó el modelo(o sea el número de parámetros estimados incluyendo, en su caso la constante); s^2 es el estimador insesgado de la varianza de la perturbación, es decir, el cuadrado del error estándar de la regresión; n es el número de observaciones y e_t es el residuo del período t. Asimismo, como s^2 está relacionada con el coeficiente $\bar{R}^2$ mediante la expresión siguiente, el criterio que minimiza a s^2, o al error estándar de la regresión, también es el que maximiza a $\bar{R}^2$.

$$\bar{R}^2 = 1 - \frac{\sum_{t=1}^{n} e_t^2 \Big/ n-k}{\sum_{t=1}^{n} \left(y_t - \bar{y}_t\right)^2 \Big/ n-1} = 1 - \frac{s^2}{\frac{\sum_{t=1}^{n} \left(y_t - \bar{y}\right)^2}{n-1}}$$

Criterios de selección de modelos de frecuente utilización son los denominados s^2, *Akaike* y *Schwarz*. Cuando se intenta elegir entre ellos hay que tener en cuenta que estos tres criterios son funciones de la razón del número de parámetros estimados al número de observaciones, k/n, que para el valor k/n=0 el

factor de penalización de los tres criterios vale 1 y que al aumentar el valor de *k/n*, la penalización de s^2 es la que aumenta con menor rapidez de las tres, la del criterio de Akaike aumenta algo mas rápidamente y la penalización del criterio de Schwarz es la que aumenta con mayor rapidez.

Por otra parte, la cuestión de la jerarquización de los criterios s^2, Akaike y Schwarz ha de ser considerada teniendo en cuenta el concepto de *consistencia.* Se dice que un criterio de selección de modelos es consistente si:

- cuando el verdadero modelo- es decir, el proceso generador de los datos- *se encuentra* entre los modelos candidatos considerados, al aumentar el tamaño de la muestra, la probabilidad de elegir tal modelo tiende a uno.
- cuando el verdadero modelo *no está* entre los candidatos considerados (siendo, por tanto imposible seleccionarlo) la probabilidad de seleccionar la mejor aproximación al proceso generador de los datos tiende a uno al aumentar el tamaño muestral.

El *ECM* es inconsistente porque no penaliza por grados de libertad; el criterio del s^2 sí penaliza los grados de libertad pero no lo suficiente para considerarlo un criterio consistente de selección de modelos. En opinión de Diebold, F.X.(1998) el criterio de Akaike tiende a *sobreparametrizar* los modelos elegidos (suele elegir modelos demasiado grandes). En cambio, el criterio de Schwarz, que penaliza con más intensidad que el de Akaike la utilización de grados de libertad, es decir, la penalización es mayor para los modelos más complejos- sí es consistente.

Según lo anterior parece que habría que decantarse inequívocamente por el criterio de Schwarz, eligiendo aquél modelo que, dentro de la familia considerada, minimizase este último criterio. Pero la cuestión no es tan simple, ya que en la decisión anterior está implícita la hipótesis de que el proceso que realmente generó los datos, o la mejor aproximación a él, se encuentra en el conjunto *fijo* de modelos considerados. Sin embargo, esta hipótesis no es de aceptación universal, ya que un punto de vista más escéptico considera que el proceso generador de los datos, y su mejor aproximación, son mucho más complicados que cualquier modelo que se les

trate de ajustar, por lo que se debe aumentar el conjunto de modelos considerados conforme aumenta el tamaño de la muestra, lo que conduce a una propiedad de optimalidad distinta, denominada *eficiencia asintótica.*

Un criterio de selección de modelo que sea *asintóticamente eficiente* elige, a medida que aumenta el tamaño de la muestra, un conjunto de modelos cuyos errores de predicción para el horizonte unidad tienen varianzas que tienden a la que se obtendría con el modelo verdadero de parámetros conocidos con una rapì dez no menor que la de cualquier otro criterio de selección de modelo. El criterio de Akaike, aunque es inconsistente, sí es asintóticamente eficiente, mientras que el de Schwarz que sí es consistente, no es asíntóticamente eficiente. Como dicen McQuarrie, A.D.M. y Tsay, C.L.(1998), "*Existe escaso consenso sobre si la consistencia es preferible a la eficiencia asintótica o viceversa.....Ningún criterio de selección de modelos es siempre mejor que otro para cualquier tipo de modelos.....Es importante recordar que los propios criterios de selección de modelos son variables aleatorias que tienen sus correspondientes funciones de distribución*".

Posiblemente, el *AIC* (Criterio de Información de Akaike) y el *BIC* (Bayesian Information Criterion), dos criterios derivados del logaritmo de la función de verosimilitud, sean los criterios de jerarquización de modelos más utilizados actualmente. En ambos la regla de decisión es la misma**:** dentro de una determinada familia de modelos se elige aquél que minimiza el *AIC* o el *BIC* . Al minimizar el *AIC* o el *BIC* se trata de minimizar el error de predicción *fuera de la muestra* **.** Para tamaños muestrales de ocho o más observaciones el *BIC* suele elegir el mismo modelo que el *AIC*, o uno con menor número de parámetros, razón por la que muchos autores, aplicando el principio de *parquedad* o *simplicidad científica*, cuando son parecidos los modelos seleccionados por los criterios de Akaike y Schwarz, recomiendan utilizar este último criterio.

Akaike propuso su criterio en el dominio de los modelos ingenieriles donde los datos se describen frecuentemente con leyes probabilísticas de pocos parámetros. La forma general del estadístico de Akaike es:

$$AIC = \frac{\sum_{t=1}^{n} e_t^2}{n} \exp\left(\frac{2k}{n}\right)$$

Schwarz demostró que el modelo que minimiza el *BIC* es el más probable en sentido bayesiano, independientemente de la distribución *a priori*. Aunque existen varias versiones del *BIC*, en realidad, son simples transformaciones algebraicas unas de otras, siendo una de las mas utilizadas la siguiente:

$$BIC = \frac{\sum_{t=1}^{n} e_t^2}{n} n^{\left(\frac{k}{n}\right)}$$

Es frecuente que los autores y los programas informáticos (por ejemplo, Eviews) se refieran al logaritmo neperiano de los criterios de Akaike (AIC) y de Schwarz (BIC) y esta práctica es tan común que *Ln (AIC)* y *Ln (BIC),* se suelen representar, para abreviar, simplemente, por *AIC y BIC*, con las respectivas expresiones:

$$Ln(AIC) = Ln\left(\frac{\sum_{t=1}^{n} e_t^2}{n}\right) + \frac{2k}{n}$$

$$Ln(BIC) = Ln\left(\frac{\sum_{t=1}^{n} e_t^2}{n}\right) + \frac{kLn(n)}{n}$$

Otra alternativa al Criterio de Información de Akaike (AIC) es el propuesto por Hannan y Quinn(1979), cuyo estadístico se define como:

$$HQ = Ln\left(\frac{\sum_{t=1}^{n} e_t^2}{n}\right) + \frac{2LnLn(n)k}{n}$$

El estadístico de Hannan y Quinn penaliza de forma explícita el tamaño muestral y, para la mayoría de los tamaños muestrales que se presentan en la práctica, su función de penalización resulta similar a la del criterio AIC. Cuando se considera el criterio de Hannan y Quinn como el patrón de selección, se debe elegir aquel modelo que presenta el menor valor de dicho estadístico.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

III.6. MODELOS ECONOMÉTRICOS DE PREVISIÓN DE DEMANDA ANUAL DE ENERGÍA ELÉCTRICA EN LA COMUNIDAD AUTÓNOMA DE ANDALUCÍA

El propósito de este apartado es investigar las variables que influyen en la demanda anual de electricidad en la Comunidad Autónoma de Andalucía. Para ello se consideran cuatro modelos que pueden ayudar a explicar y predecir la demanda de electricidad anual. Los modelos que se estudian se expresan en la tabla III.6.1 y respecto a su forma funcional habría que decir que los cuatro modelos son del tipo log-lineal o de elasticidad constante. Esta familia de modelos log-lineal, se caracteriza porque es lineal en los coeficientes de regresión y lineal en los logaritmos de las variables. El horizonte temporal del estudio contempla los años de la serie 1984-1999 (ambos inclusive).

Tabla.III.6.1. Modelos de demanda de energía eléctrica en la C.A.A. Sector residencial.

Modelo	Expresión funcional
Modelo 1	$LnQ_i = Ln\alpha + \beta_1 LnPBKWH_i + \beta_2 LnPE_i + u_i$
Modelo 2	$LnQ_i = Ln\alpha + \beta_1 LnPGKWH_i + \beta_2 LnPE_i + u_i$
Modelo3	$LnQ_i = Ln\alpha + \beta_1 LnPE_i + \beta_2 LnPIB_i + u_i$
Modelo 4	$LnQ_i = Ln\alpha + \beta_1 LnPE_i + \beta_2 LnQ(-1)_i + u_i$

Fuente. IEA.
Elaboración. Propia.

En donde *Ln*= logaritmo natural.

Q_i [3]es la demanda de energía eléctrica del sector residencial en la Comunidad Autónoma de Andalucía en el año el *i-ésimo*.

$Q(-1)_i$ es la demanda de energía eléctrica del sector residencial retardada un año en la Comunidad Autónoma de Andalucía.

$PBKWH_i$[4] es el precio del gas butano (precio de gases licuados del petróleo) en el año *i-ésimo*, expresado en el precio unitario del kWh. Su cálculo se ha realizado a partir de la botella de 12,5 Kg de uso doméstico y aplicando las distintas conversiones entre unidades caloríficas. (Ikg butano=11800Kcal).

$PGKWKH_i$[5] es el precio del gas ciudad en el año *i-ésimo*, expresado en el precio unitario del kWh. Su cálculo se ha realizado aplicando distintas conversiones entre las unidades caloríficas. (1Th=10^6 calorías=1,162kWh).

PE_i[6] es el precio de kWh para usos residenciales (tarifa 2.0 y su versión de consumo *nocturna-día*) en el año *i-ésimo*.

PIB_i[7] es el producto interior bruto generado en la Comunidad Autónoma de Andalucía en el año i-ésimo, expresado en millones de pesetas.

α es el término constante de la regresión.

β_1, β_2 son los coeficientes de regresión del modelo.

u_i es la perturbación aleatoria del modelo, i=1, 2, 16.

Todas las cifras que están expresadas en unidades monetarias, tales como los valores de precios o el PIB han sido deflactados según el índice de precios de 1984 (referenciado a 1992).

A continuación se comenta el significado de algunos de los parámetros obtenidos en la investigación de los modelos considerados en este trabajo. La salida de ordenador comienza por recordar (en la parte superior izquierda) que se está procesando una regresión por mínimos cuadrados (LS, Least squares) y en la misma línea preliminar, pero algo más a la derecha, se indica cual es la variable dependiente y seguidamente aparece el rango de los datos históricos.

Viene a continuación una tabla en la que se puede ver cada variable explicativa, y a su derecha, cuatro estadísticos correspondientes a ella. La

[3] Informes Anuales de la Compañía Sevillana de Electricidad, S.A.
[4] Información suministrada por Repsol Butano.
[5] Información suministrada por Gas Natural.
[6] Tarifas recogidas en el Boletín Oficial del Estado y reflejadas en la memoria anual de UNESA.
[7] Banco Bilbao-Vizcaya. Renta Nacional de España y su distribución provincial. Años 1955-1993. IEA. Contabilidad Trimestral en Andalucía. Serie 1991-1996. INE. Contabilidad Regional de España. Serie 1995-1999.

"variable explicativa" ficticia C siempre vale la unidad, siendo su coeficiente estimado la ordenada en el origen de la recta de regresión. Los cuatro estadísticos mencionados que están asociados con cada variable explicativa: son, de izquierda a derecha, el coeficiente estimado ("coefficient"), su error estándar o desviación típica ("Std. Error"), el estadístico *t* y el correspondiente *p-valor* ("Prob"). Los errores estándar de los coeficientes estimados miden la variabilidad de estos y cuanto mayores sean tales errores estándar - a igualdad de los otros parámetros- mayor será la amplitud de los intervalos de confianza de los coeficientes estimados.

Los estadísticos *t* contrastan la significación de la correspondiente variable explicativa, se definen como la razón del correspondiente coeficiente de regresión estimado a su error estándar. Al 95% de grado de confianza se considera que la variable explicativa es relevante cuando *t* toma un valor absoluto igual o superior a dos.

Finalmente, en este bloque de información, figura el *p-valor,* o probabilidad de obtener un valor absoluto del estadístico *t* no inferior al que se ha obtenido, supuesta cierta la hipótesis nula que postula la irrelevancia de la correspondiente variable explicativa. Es frecuente considerar, respectivamente, los valores del *p-valor* que no superen a (0,05) 0,1 como (muy) fuerte evidencia en contra de la mencionada irrelevancia.

Los siguientes estadísticos de diagnóstico, que presenta como salida el ordenador, forman parte de los resultados obtenidos en este trabajo y tienen por objeto evaluar si el modelo de regresión es adecuado. En todos ellos *n,* y_t, $\hat{y}_t$ y e_t representan, respectivamente, al número de datos, la observación en el tiempo *t*, el valor pronosticado para y_t y el error de predicción correspondiente al tiempo o periodo *t*.

Mean dependent var (Media de la variable dependiente). El promedio muestral de la variable dependiente es un estadístico que mide la localización de los datos. Su expresión es:

$$\bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t$$

S.D. dependent var (Desviación típica de la variable dependiente). La desviación estándar de la variable dependiente mide la dispersión de las observaciones de la variable investigada, siendo su expresión:

$$\sigma = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \bar{y})^2}{n-1}}$$

Sum squared resid (Suma cuadrática de los residuos). Este parámetro, anteriormente comentado, representa como su nombre indica el valor mínimo de la suma cuadrática de los residuos, siendo útil para comparar modelos y contrastar hipótesis. Su expresión analítica es:

$$SCE = \sum_{t=1}^{n} e_t^2 \text{ , donde } e_t = y_t - \hat{y}_t$$

R-squared (Coeficiente de determinación). Si el modelo considerado incluye ordenada en el origen, este estadístico toma una valor comprendido entre 0 y 1 y expresa el porcentaje de la variable dependiente que explican los regresores, es decir la bondad del ajuste *dentro de la muestra*. El estadístico R^2 vale 1 cuando el ajuste de modelo a los datos históricos es perfecto; por el contrario, R^2 vale cero cuando el ajuste del modelo a los datos no es mejor que la simple media aritmética de éstos. Su expresión es:

$$R^2 = 1 - \frac{SCE}{SCT} = \frac{SCR}{SCT}$$

donde, $SCR = \sum_{t=1}^{n}(\hat{y}_t - \bar{y})^2$, $SCT = \sum_{t=1}^{n}(y_t - \bar{y})^2$.

Adjusted R-squared (Coeficiente de determinación ajustado). Este coeficiente, que se suele representar por $\bar{R}^2$, penaliza la inclusión de nuevas variables explicativas en el modelo, ya que, si bien al aumentar el número de regresores aumenta también la SCR, la inclusión de nuevas variables explicativas reduce los grados de libertad del modelo, por lo que no siempre compensará incorporar nuevas variables al mismo. Si existe más de un regresor el valor de este parámetro es menor que el de R^2. Una característica de $\bar{R}^2$ es que permite comparar modelos que traten la misma muestra con distinto número de variables explicativas. Su expresión analítica es:

$$\bar{R}^2 = 1 - \frac{SCE/(n-k)}{SCT/(n-1)} = 1 - \left(R^2 \frac{n-1}{n-k} \right)$$

siendo *k* es el número de coeficientes de regresión del modelo, incluida el término intersección.

S.E. of Regression (Error estándar de la regresión) Si los valores de los coeficientes de regresión β fueran conocidos, los errores de predicción serían únicamente los valores ε_t con varianza σ^2. Pero puesto que los parámetros β son desconocidos, se considera el estadístico s^2 que, bajo las hipótesis clásicas, es un estimador consistente e insesgado de la varianza de los errores σ^2. El estimador s^2 se utiliza para evaluar la bondad de ajuste del modelo de forma que cuanto mayor es s^2 peor es el ajuste del modelo y probablemente se cometerán mayores errores de predicción. En la expresión siguiente s^2 aparece corregido en el denominador por los grados de libertad a fin de obtener un buen estimador de la varianza de los errores *fuera de la muestra*, con base en los residuos *interno*s a la muestra.

$$s^2 = \frac{\sum_{t=1}^{n} e_t^2}{n-k}$$

El error estándar de la regresión, *s,* contiene la misma información que s^2, ya que es su raíz cuadrada y es más fácil de interpretar que s^2 debido a que se

expresa en unidades de la variable dependiente, siendo un estimador de σ, la desviación típica de las perturbaciones. La expresión de *s* es:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{t=1}^{n} e_t^2}{n-k}}$$

Log likelihood (Logaritmo de la razón de verosimilitud). La expresión que se indica a continuación es el valor máximo del logaritmo neperiano de la función de verosimilitud -es decir, de la función de densidad conjunta de los *n* residuos $(e_1,\ e_2,\ ...,\ e_n)$- y, al igual que la suma cuadrática de los residuos, es de escasa aplicación directa, pero es útil para comparar modelos y contrastar hipótesis. Su expresión es :

$$LnL(e_1,\ e_2,\ ...,\ e_n) = -\frac{n}{2} - \frac{n}{2}Ln(2\pi) - \frac{n}{2}Ln\left(\frac{SCE}{n}\right)$$

Durbin-Watson stat (Estadístico de Durbin-Watson). Este estadístico se utiliza para contrastar la hipótesis nula de incorrelación entre las perturbaciones aleatorias ε_t $(H_0 : \rho = 0)$ frente a la hipótesis alternativa de que tales perturbaciones siguen un proceso autorregresivo AR(1), $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$. La expresión del estadístico es :

$$d = DW = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} \approx 2(1-\hat{\rho})$$

donde e_t es el residuo correspondiente a la observación del periodo *t* y $\hat{\rho}$ el coeficiente de autocorrelación muestral. De acuerdo con la última expresión y puesto que -1< ρ<1, el estadístico *d* puede tomar valores comprendidos entre 0 y 4, de forma que en ausencia de correlación serial, es decir cuando $\hat{\rho}$=0, su valor será próximo a 2. Valores de *d* cercanos a 0 indican autocorrelación positiva de los residuos ($\hat{\rho}$=1), mientras que valores de *d* próximos a 4 son indicativos de autorrelación negativa de los residuos ($\hat{\rho}$=-1).

Este contraste presenta algunas limitaciones entre las que se encuentran las siguientes:

- El modelo debe tener término constante.
- Sólo es válido para contrastar autocorrelaciones de primer orden.
- No se puede aplicar cuando existen endógenas retardadas.
- Las zonas de indeterminación del contraste pueden ser extensas.
- Si se cuenta con menos de 15 observaciones los resultados pueden no ser fiables.

Akaike info criterion (Criterio de Información de Akaike). El criterio de información de Akaike (AIC) es un estadístico muy útil en la selección de modelos de regresión. Penaliza la inclusión de nuevos regresores en el modelo. La expresión del estadístico en forma logarítmica es:

$$AIC = nLn(2\pi) + nLn\frac{SCE}{n} + n + 2k$$

La constante $nLn(2\pi) + n$ se puede ignorar puesto que no depende de ningún parámetro que se desee optimizar por lo que, en la práctica, (después de dividir por 1/*n* para expresar el estadístico *AIC* como una tasa) la expresión utilizada es:

$$AIC = \frac{2k}{n} + Ln\left(\frac{SCE}{n}\right)$$

Schwarz criterion (Criterio de Schwarz) El estadístico de Schwarz, es más restrictivo que el de Akaike en la selección de modelos de regresión, ya que penaliza en mayor grado la inclusión de nuevos regresores en el modelo. La expresión utilizada en la práctica es:

$$SIC = \frac{k}{n} Ln\, n + Ln\left(\frac{SCE}{n}\right)$$

Tanto el estadístico de Akaike como el de Schwarz, sirven para analizar la capacidad explicativa de un modelo, permitiendo realizar comparaciones a este respecto entre modelos anidados. En realidad ambos estadísticos son estimadores de la varianza del error de predicción *fuera de la muestra* .

F-statistic. Se emplea el estadístico *F* para contrastar la hipótesis de que todas las variables explicativas (excepto la ordenada en el origen) son cero, como sostiene la hipótesis nula, frente a la hipótesis alternativa que postula que tales variables, consideradas en *conjunto*, tienen algún poder predictivo. Si el estadístico *F* excede el umbral crítico determinado por el nivel de significación del contraste, entonces al menos uno de los coeficientes de regresión es no nulo. Este estadístico se diferencia del estadístico *t*, comentado anteriormente, en que *t* contrasta la irrelevancia *individual* de cada variable explicativa. Si ninguna variable explicativa tiene valor predictivo, el estadístico *F* sigue una distribución *F* de Snedecor con (*k-1)* y (*n-k)* grados de libertad, siendo su expresión:

$$F = \frac{SCR/(k-1)}{SCE/(n-k)} = \frac{R^2}{1-R^2} \cdot \frac{n-k}{k-1}$$

BIBLIOTECA VIRTUAL

Dada la relación existente entre este estadístico y el R^2, el contraste *F* puede considerarse como un medio de evaluar si el valor de R^2 tiene un valor suficiente para considerar que el modelo tiene una capacidad explicativa adecuada.

Prob (F-statistic). Esta probabilidad indica el *p-valor* de que el conjunto de variables explicativas no tenga valor predictivo alguno, que es lo que postula la hipótesis nula. Cuanto más pequeño sea el *p-valor*, mayor evidencia habrá de que el conjunto de las variables explicativas consideradas en el modelo tiene valor predictivo.

El proceso de construcción de un modelo econométrico se inicia con la especificación de la relación a estimar y la formulación de un conjunto de hipótesis. Este procedimiento inicial, que requiere la selección entre distintas alternativas, puede incurrir, sin embargo, en errores. Por ello, es conveniente someter el modelo identificado y estimado a diversos contrastes que permitan comprobar su validez y calidad entes de utilizarlo en el trabajo empírico. Este proceso es el que se llevará a cabo para todos los modelos que a continuación se analizan para la previsión de demanda de energía eléctrica en el sector residencial en la C.A.A.

Análisis estructural del Modelo 1

Este modelo explica el consumo doméstico de electricidad en la C.A.A. en función, en primer lugar, del precio del gas butano y en segundo lugar, se expresa en función del precio del kilovatio por hora de electricidad (expresado en pesetas). La ecuación estimada del modelo se presenta en la tabla III.6.2. El coeficiente de regresión de la variable Ln PBKWH resulta significativo al 95% y expresa el *coeficiente de elasticidad cruzada* de la demanda de energía eléctrica, expresando que un incremento en el precio del butano en un 1% por peseta (1/166,386 euros) produce, por término medio, un incremento de 0,61%, siendo el precio de la electricidad una condición *ceteris paribus*. Al tomar dicha elasticidad cruzada un valor positivo revela el carácter sustitutivo del gas butano en la economía doméstica. Sin embargo, el *coeficiente de elasticidad de la demanda respecto al precio* es de 1,018, lo cual implica que para un incremento de 1% en el precio constante de la electricidad, la demanda doméstica de ésta (medida en GWh) en promedio aumenta en 1,018%, pudiéndose decir que la demanda doméstica de electricidad en la C.A.A. es elástica con respecto a su precio unitario.

Además, el modelo en su conjunto resulta ser significativo al 95% de confianza. La información suministrada por las dos variables analizadas acerca de la demanda de electricidad es del 86%.

Tabla.III.6.2. Estimación del modelo 1 de demanda de energía eléctrica en la C.A.A.

Dependent Variable: Ln Q
Method: Least Squares
Sample: 1984 1999
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PBKWH	0.619925	0.197582	3.137564	0.0079
Ln PE	1.018614	0.206389	4.935410	0.0003
C	5.468783	0.439868	12.43279	0.0000

R-squared	0.862452	Mean dependent var	8.390384
Adjusted R-squared	0.841291	S.D. dependent var	0.252566
S.E. of regression	0.100618	Akaike info criterion	-1.587608
Sum squared resid	0.131612	Schwarz criterion	-1.442748
Log likelihood	15.70087	F-statistic	40.75636
Durbin-Watson stat	0.590694	Prob(F-statistic)	0.000003

Fuente. IEA.
Elaboración. Propia.

Una vez estimada la ecuación de un modelo que se plantea como válido a la hora de predecir el consumo de energía eléctrica en la Comunidad Autónoma de Andalucía, se procede a evaluar la calidad del modelo mediante la aplicación de una serie de contrastes o tests que garanticen la viabilidad de la estructura estimada. Una guía de los contrastes aplicados se detallan en la tabla III.6.3:

Tabla.III.6.3. Contrastes de diagnóstico para la evaluación de un modelo.

Tipos de contrastes	**Aplicación**
Coeficientes	Variables irrelevantes
Residuos	Identificación de los residuos Identificación de los residuos al cuadrado Normalidad de los residuos Correlación serial de los residuos Heterocedasticidad autorregresiva condicional en los residuos Heterocedasticidad en los residuos
Estabilidad del Modelo	Errores de especificación por variables omitidas Inestabilidad de parámetros • Estimación recursiva de los residuos • Suma acumulada de los residuos al cuadrado • Estimación recursiva de los parámetros

Fuente. Pulido,A.(2001)
Elaboración. Propia.

Este proceso se realizará para los cuatro modelos de previsión anuales, comenzando por el modelo 1. De esta forma, se comienza realizando un contraste acerca de la posibilidad de omitir alguna de las variables integrantes en el modelo, por tener carácter redundante. La hipótesis nula de este constraste establece que el coeficiente de regresión de la variable a analizar es nulo. Los resultados de la aplicación del mismo a las variables Ln PBKWH y Ln PE se expresan en la tabla III.6.4 .

En este modelo, los resultados obtenidos conducen a rechazar la hipótesis nula, para las dos variables, pues los valores críticos de las tablas para el estadístico $F_{1,14}$ y χ_1^2 son, respectivamente, 4,6 y 3,84 al nivel de

significación del 5%, y los valores que obtenemos de test son superiores a éstos (9,84 para *F* y 9,02 para LR en el Ln PBKWH. 24,35 para *F* y 16,88 para LR en el Ln PE). Además la nueva estimación del modelo, pierde significación, pues el coeficiente de determinación corregido por los grados de libertad disminuye un 11,92% omitiendo la variable PBKWH y un 31,47% omitiendo la variable PE. Luego, se puede rechazar la hipótesis nula que establecía que las variables PBKWH y PE, de forma individual no eran significativas, tanto al nivel del 5% de significación como al 1%.

Tabla.III.6.4. Variables Irrelevantes. Modelo 1.

Redundant Variables: Ln PBKWH				
F-statistic	9.844309	Probability		0.007858
Log likelihood ratio	9.020042	Probability		0.002670
Test Equation:				
Dependent Variable: Ln Q				
Method: Least Squares				
Sample: 1984 1999				
Included observations: 16				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PE	1.404039	0.211856	6.627335	0.0000
C	4.948220	0.520382	9.508824	0.0000
R-squared	0.758294	Mean dependent var		8.390384
Adjusted R-squared	0.741029	S.D. dependent var		0.252566
S.E. of regression	0.128529	Akaike info criterion		-1.148856
Sum squared resid	0.231276	Schwarz criterion		-1.052282
Log likelihood	11.19085	F-statistic		43.92157
Durbin-Watson stat	0.302808	Prob(F-statistic)		0.000011

Redundant Variables:Ln PE				
F-statistic	24.35827	Probability		0.000272
Log likelihood ratio	16.88968	Probability		0.000040
Test Equation:				
Dependent Variable: Ln Q				
Method: Least Squares				
Sample: 1984 1999				
Included observations: 16				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PBKWH	1.200330	0.259361	4.628025	0.0004
C	7.568735	0.182231	41.53377	0.0000
R-squared	0.604728	Mean dependent var		8.390384
Adjusted R-squared	0.576494	S.D. dependent var		0.252566
S.E. of regression	0.164364	Akaike info criterion		-0.657003
Sum squared resid	0.378215	Schwarz criterion		-0.560430
Log likelihood	7.256028	F-statistic		21.41861
Durbin-Watson stat	0.380360	Prob(F-statistic)		0.000391

Fuente. IEA.
Elaboración. Propia.

Seguidamente se aplican un grupo de contrastes aplicados a los residuos que miden su correlación serial, su normalidad y su heterocedasticidad entre otras características.

Análisis de los residuos

En la tabla III.6.5 se representa el correlograma de los residuos para doce retardos para contrastar si se encuentran autocorrelacionados. Se observa que los coeficientes de autocorrelación (AC) y autocorrelación parcial (PAC) no alcanzan en la mayoría de los casos valores próximos a cero. El valor del estadístico *Q – de Ljung-Box*, 19,943, acompañado de un valor reducido de probabilidad (0,068) confirman la ausencia de autocorrelación entre los residuos del modelo estimado al 95% de confianza.

Tabla.III.6.5. Correlograma de los residuos. Modelo 1.

Sample: 1984 1999
Included observations: 16

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. |***. |	. |***. |	1	0.447	0.447	3.8300	0.050
. | . |	. **| . |	2	0.023	-0.221	3.8405	0.147
. *| . |	. *| . |	3	-0.179	-0.121	4.5494	0.208
.***| . |	.***| . |	4	-0.427	-0.368	8.9196	0.063
. **| . |	. | . |	5	-0.298	0.045	11.245	0.047
. **| . |	. **| . |	6	-0.243	-0.287	12.948	0.044
. *| . |	. | . |	7	-0.108	0.006	13.322	0.065
. |* . |	. *| . |	8	0.074	-0.127	13.520	0.095
. |** . |	. |* . |	9	0.198	0.141	15.141	0.087
. |** . |	. | . |	10	0.271	-0.038	18.668	0.045
. |* . |	. | . |	11	0.148	-0.008	19.926	0.046
. | . |	. *| . |	12	0.015	-0.059	19.943	0.068

Fuente. IEA.
Elaboración. Propia

Análisis de los residuos al cuadrado

Una variante del contraste anterior es el *correlograma de los residuos al cuadrado*, el cual se utiliza, principalmente, para analizar la presencia de heterocedasticidad condicional autorregresiva (ARCH) en los residuos. En la tabla III.6.6, se puede observar cómo los residuos del modelo 1 no presentan este comportamiento, ya que los coeficientes de autocorrelación y de autocorrelación parcial están cercanos a cero para los doce retardos

considerados, siendo el estadístico *Q* no significativo al nivel confianza del 95%.

Tabla.III.6.6. Correlograma de los residuos al cuadrado. Modelo 1.

Sample: 1984 1999
Included observations: 16

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. |* . |	. |* . |	1	0.183	0.183	0.6438	0.422
. | . |	. *| . |	2	-0.054	-0.091	0.7046	0.703
. | . |	. | . |	3	-0.016	0.012	0.7101	0.871
. |* . |	. |* . |	4	0.179	0.182	1.4757	0.831
. | . |	. *| . |	5	0.007	-0.070	1.4769	0.916
. | . |	. | . |	6	-0.038	-0.002	1.5174	0.958
. *| . |	. | . |	7	-0.059	-0.050	1.6306	0.977
. *| . |	. *| . |	8	-0.115	-0.144	2.1105	0.977
. *| . |	. | . |	9	-0.096	-0.045	2.4877	0.981
. | . |	. | . |	10	-0.045	-0.030	2.5830	0.990
. *| . |	. *| . |	11	-0.120	-0.120	3.4157	0.984
. *| . |	. | . |	12	-0.111	-0.034	4.3016	0.977

Fuente. IEA.
Elaboración. Propia

Normalidad de los residuos

En esta prueba se analiza la normalidad de los residuos, representando su histograma de probabilidad así como los principales estadísticos descriptivos. Se comprueba que, efectivamente, la media de los residuos es próxima a cero. El valor del coeficiente de asimetría de los mismos es 1,35 y el coeficiente de curtosis 4,80. Estos mismos coeficientes, correspondientes a una distribución normal, toman el valor 0 y 3, respectivamente, no estando próximos a estos últimos los valores obtenidos en el análisis de los residuos. De esta forma, bajo la hipótesis nula de que los errores siguen una distribución normal, el estadístico Jarque-Bera se distribuye como una χ^2_2 cuyo valor crítico en las tablas es de 5,991 al 5% de significación. Dado que la probabilidad asociada a dicho estadístico es 0,02, se puede rechazar la hipótesis nula (normalidad de los residuos) al nivel de significación del 5%.

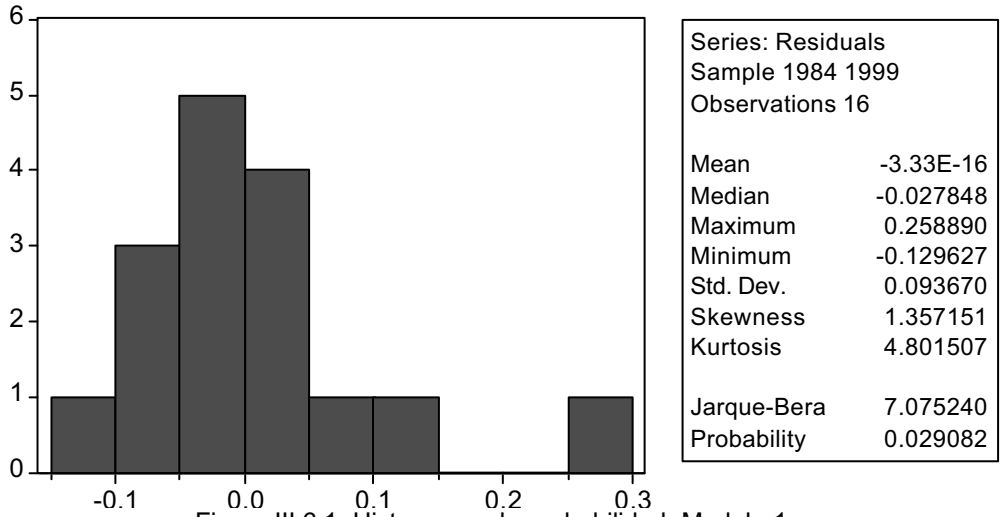


Figura.III.6.1. Histograma de probabilidad. Modelo 1.
Fuente. IEA.
Elaboración. Propia

BIBLIOTECA VIRTUAL

Correlación serial de los residuos

Una alternativa a la aplicación de estadístico *Q* para analizar la correlación serial consiste en la aplicación del contraste de Breusch-Godfrey-LM el cual es un contraste asintótico de multiplicadores de Lagrange en el que el estadístico se calcula como producto del tamaño muestral (Obs) y el coeficiente de determinación (R-squared) de una regresión auxiliar de los residuos mínimo-cuadráticos en función de *r* retardos y de las variables explicativas del modelo. Por lo tanto, este contraste se puede utilizar para analizar la presencia de autocorrelación de orden superior a 1, en general de orden *p*. La hipótesis nula que se considera, en el modelo analizado, es que no existe correlación serial de segundo orden.

Al aplicar el contraste LM a los resultados obtenidos el estadístico Obs*R-squared, que se distribuye según una χ^2_2, rechaza al nivel del 95% la hipótesis de no autocorrelación de segundo orden de los residuos.

Tabla.III.6.7. Contraste de autocorrelación serial. Modelo 1.

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	5.370741	Probability		0.023581
Obs*R-squared	7.904875	Probability		0.019208
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PBKWH	0.026294	0.166389	0.158029	0.8773
Ln PE	0.024231	0.160633	0.150849	0.8828
C	-0.069528	0.340861	-0.203976	0.8421
RESID(-1)	1.127719	0.344842	3.270250	0.0075
RESID(-2)	-0.425844	0.469200	-0.907596	0.3835
R-squared	0.494055	Mean dependent var		-3.11E-15
Adjusted R-squared	0.310075	S.D. dependent var		0.093670
S.E. of regression	0.077804	Akaike info criterion		-2.018935
Sum squared resid	0.066588	Schwarz criterion		-1.777501
Log likelihood	21.15148	F-statistic		2.685370
Durbin-Watson stat	2.226392	Prob(F-statistic)		0.087659

Fuente. IEA.
Elaboración. Propia

Heterocedasticidad autorregresiva condicional en los residuos

Seguidamente se considera el contraste denominado ARCH LM, en la tabla III.6.8, el cual es una variante de los contrastes basados en los multiplicadores de Lagrange, para analizar el comportamiento heterocedástico autorregresivo condicional (ARCH) de los residuos. De nuevo en este caso ambos estadísticos, tanto el F-statistic como el estadístico LM de Engle, no permiten aceptar la hipótesis nula de comportamiento homocedástico autorregresivo condicional en los residuos.

Tabla.III.6.8. Contraste comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 1.

ARCH Test:				
F-statistic	7.078000	Probability		0.019618
Obs*R-squared	5.287877	Probability		0.021474
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample(adjusted): 1986 1999				
Included observations: 15 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000841	0.004715	0.178344	0.8612
RESID^2(-1)	1.828564	0.687314	2.660451	0.0196
R-squared	0.352525	Mean dependent var		0.008715
Adjusted R-squared	0.302719	S.D. dependent var		0.017026
S.E. of regression	0.014217	Akaike info criterion		-5.545192
Sum squared resid	0.002628	Schwarz criterion		-5.450785
Log likelihood	43.58894	F-statistic		7.078000
Durbin-Watson stat	1.295817	Prob(F-statistic)		0.019618

Fuente. IEA.
Elaboración. Propia

Heterocedasticidad en los residuos

Para finalizar, dentro del grupo de contrastes aplicados para analizar las propiedades de los residuos, se ha contrastado la heterocedasticidad o varianza no constante del término de error, a través del contraste de White. En este test la hipótesis nula establece que los errores son homocedásticos e independientes de los regresores. A la vista de los resultados que presentan ambos estadísticos, F-statistic y Obs*R-squared, no se puede rechazar la hipótesis nula de homocedasticidad, siendo su valor muy próximo al nivel crítico.

Tabla.III.6.9. Contraste de heterocedasticidad de los residuos. Modelo 1.

White Heteroskedasticity Test:				
F-statistic	3.206892	Probability		0.056431
Obs*R-squared	8.613598	Probability		0.071518
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample: 1985 1999				
Included observations: 16				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.030498	1.472155	-2.737821	0.0193
Ln PBKWH	0.257308	0.230713	1.115274	0.2885
(Ln PBKWH)^2	-0.099408	0.162232	-0.612751	0.5525
Ln PE	3.339825	1.205746	2.769925	0.0182
(Ln PE)^2	-0.708769	0.254330	-2.786808	0.0177
R-squared	0.538350	Mean dependent var		0.008226
Adjusted R-squared	0.370477	S.D. dependent var		0.016564
S.E. of regression	0.013142	Akaike info criterion		-5.575641
Sum squared resid	0.001900	Schwarz criterion		-5.334207
Log likelihood	49.60512	F-statistic		3.206892
Durbin-Watson stat	1.445854	Prob(F-statistic)		0.056431

Fuente. IEA.
Elaboración. Propia

Errores de especificación por variables omitidas

El contraste de Ramsey estudia la idoneidad del modelo, referida a la forma funcional. A la regresión se añaden nuevas variables que son las estimaciones realizadas del modelo original elevadas al cuadrado. La hipótesis nula establece que los coeficientes de las potencias de la variable estimada son nulos, y aplicándolo a este modelo se observa que las probabilidades asociadas a los estadísticos son superiores al nivel de significación de 5%, por lo tanto se puede afirmar que la estimación lineal inicial del modelo es correcta.

Tabla.III.6.10. Contraste de Ramsey. Modelo 1.

Ramsey RESET Test:				
F-statistic	2.609718	Probability		0.132180
Log likelihood ratio	3.148485	Probability		0.075998
Test Equation:				
Dependent Variable: ln Q				
Method: Least Squares				
Sample: 1984 1999				
Included observations: 16				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PBKWH	17.48274	10.44004	1.674584	0.1199
Ln PE	28.14228	16.79116	1.676017	0.1196
C	40.37383	21.61083	1.868222	0.0863
FITTED^2	-1.603188	0.992402	-1.615462	0.1322
R-squared	0.887022	Mean dependent var		8.390384
Adjusted R-squared	0.858778	S.D. dependent var		0.252566
S.E. of regression	0.094913	Akaike info criterion		-1.659389
Sum squared resid	0.108102	Schwarz criterion		-1.466242
Log likelihood	17.27511	F-statistic		31.40524
Durbin-Watson stat	0.715323	Prob(F-statistic)		0.000006

Fuente. IEA.
Elaboración. Propia.

Estimación recursiva de los residuos

La representación de la serie de los residuos recursivos junto con las bandas de confianza, definidas con $\pm$ dos veces su desviación estándar, permite detectar gráficamente la inestabilidad en los parámetros cuando uno o varios residuos superan dichas bandas, como sucede en el modelo analizado en el año 1996, lo cual refleja una falta de estabilidad del modelo a lo largo del periodo analizado.

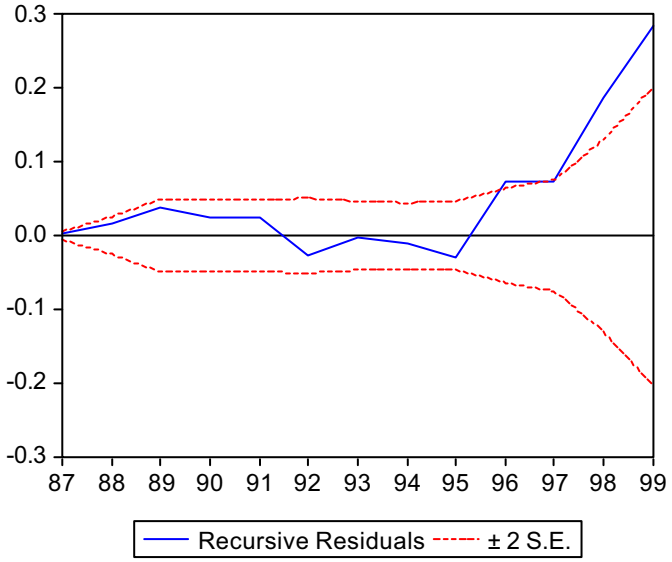


Figura.III.6.2. Residuos recursivos. Modelo 1.
Fuente. IEA.
Elaboración. Propia

Suma acumulada de los residuos recursivos

El estadístico *CUSUM* se construye a partir de la suma acumulada de los residuos recursivos. Bajo la hipótesis de estabilidad estructural, el estadístico CUSUM tiene media cero, por lo que las sumas acumuladas que se alejen de dicho valor, indican existencia de inestabilidad. En la figura III.6.3 se representan los residuos (línea continua) frente al tiempo junto con las bandas de confianza (líneas discontinuas). Se observa que el estadístico CUSUM oscila dentro de sus bandas de confianza, sin embargo, a partir del año 1996 se produce un crecimiento que le aleja del valor cero de forma continuada para el resto de la muestra, lo que constituye en su conjunto un indicio de inestabilidad del modelo.

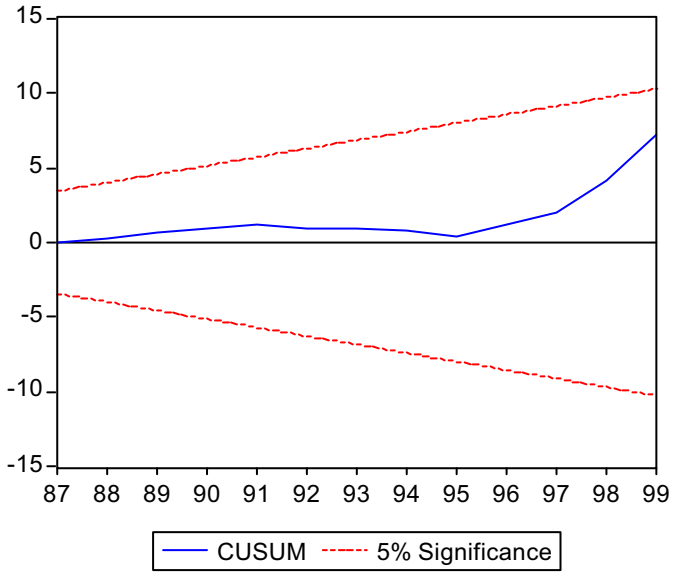


Figura.III.6.3. Suma acumulada de los residuos recursivos. Modelo 1.
Fuente. IEA.
Elaboración. Propia.

Estimación recursiva de los coeficientes

El gráfico de los coeficientes recursivos representa el comportamiento de cada uno de los estimadores al ir añadiendo observaciones a la muestra con la que se realiza la estimación. Las series de los coeficientes se presentan con sus bandas de confianza (± dos veces su desviación estándar) lo que permite detectar la presencia de inestabilidad en el modelo si los coeficientes sufren grandes cambios al ir aumentando la muestra.

La representación gráfica de los coeficientes recursivos para el modelo estimado de la demanda anual de energía eléctrica doméstica en la C.A.A. presenta los resultados reflejados en la figura III.6.4. El examen de dichos gráficos, indica un patrón en cierta medida estable en términos generales para los coeficientes, LnPE y C. Sin embargo, el coeficiente de la variable LnPBKWH tiene una evolución inestable.

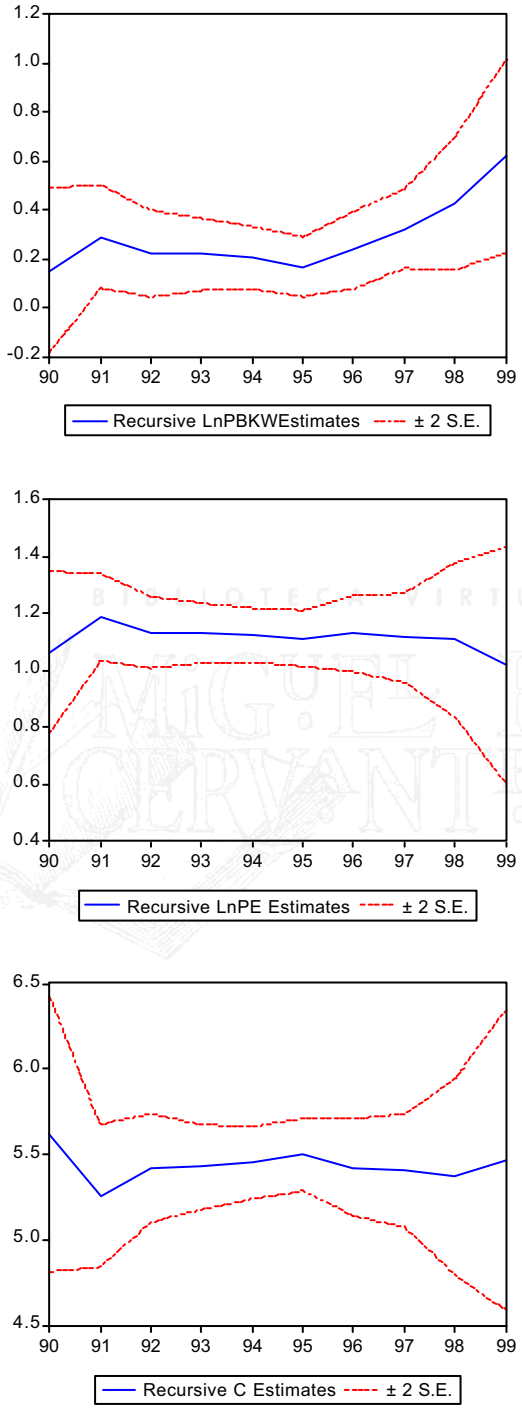


Figura.III.6.4. Gráficos de los coeficientes recursivos. Modelo 1.
Fuente. IEA.
Elaboración. Propia.

Análisis estructural del Modelo 2

El segundo modelo considerado relaciona la demanda investigada de electricidad con el precio del gas ciudad y del precio unitario del kilovatio por hora. Su estimación se presenta en la tabla III.6.11, donde se observa que el modelo es significativo globalmente, así como también de forma individual para cada una de las variables explicativas, tanto al 5% como al 1% de significación. El coeficiente de regresión de la variable Ln PGKWH expresa la elasticidad cruzada de la demanda de electricidad respecto al precio del gas ciudad estando el precio de la electricidad entre las condiciones *ceteris paribus* de la demanda. Su valor es 1,14, reflejando que un incremento en 1% por peseta (1/166,386 euros) incrementará la demanda (medida en GWh consumidos al año), en promedio, aproximadamente en un 1,1445%. Al resultar la elasticidad de Ln PGKWH positiva, indica cierto carácter de sustitución del gas-ciudad respecto a la electricidad, y mayor respecto del precio por kWh del butano, tal como se reflejó en el modelo 1, anteriormente analizado.

La elasticidad demanda-precio de electricidad resulta ser 0,78, reflejando el carácter inelástico de la demanda de electricidad respecto a su precio, ya que al incrementar en un 1% su precio la demanda se incrementa por término medio en un 0,78%.

El grado en el cual explican dichas variables la variación de la demanda de electricidad, corregidas según las variables explicativas, aumenta débilmente respecto al modelo 1 pasando de 0,8413 a 0,8479.

Tabla.III.6.11.Estimación del modelo 2 de demanda de energía eléctrica en la C.A.A.

Dependent Variable: Ln Q
Method: Least Squares
Sample: 1984 1999
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PGKWH	1.144526	0.347763	3.291109	0.0058
Ln PE	0.781117	0.249382	3.132207	0.0079
C	4.867315	0.399609	12.18021	0.0000

R-squared	0.868150	Mean dependent var	8.390384
Adjusted R-squared	0.847865	S.D. dependent var	0.252566
S.E. of regression	0.098512	Akaike info criterion	-1.629910
Sum squared resid	0.126161	Schwarz criterion	-1.485050
Log likelihood	16.03928	F-statistic	42.79829
Durbin-Watson stat	0.612162	Prob(F-statistic)	0.000002

Fuente. IEA.
Elaboración. Propia.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

Variables Irrelevantes

El análisis de la redundancia de las variables exógenas o explicativas lleva a concluir que las dos variables que especifican el modelo 2 son necesarias por aportar información relevante al mismo, tal como se refleja en la tabla III.6.12.

Tabla.III.6.12.Variables Irrelevantes. Modelo 2.

Redundant Variables: Ln PGKWH

F-statistic	10.83140	Probability	0.005847
Log likelihood ratio	9.696876	Probability	0.001846

Test Equation:
Dependent Variable: Ln Q
Method: Least Squares
Sample: 1984 1999
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PE	1.404039	0.211856	6.627335	0.0000
C	4.948220	0.520382	9.508824	0.0000

R-squared	0.758294	Mean dependent var	8.390384
Adjusted R-squared	0.741029	S.D. dependent var	0.252566
S.E. of regression	0.128529	Akaike info criterion	-1.148856
Sum squared resid	0.231276	Schwarz criterion	-1.052282
Log likelihood	11.19085	F-statistic	43.92157
Durbin-Watson stat	0.302808	Prob(F-statistic)	0.000011

Redundant Variables: Ln PE			
F-statistic	9.810722	Probability	0.007939
Log likelihood ratio	8.996501	Probability	0.002705

Test Equation:
Dependent Variable: Ln Q
Method: Least Squares
Sample: 1984 1999
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PGKWH	1.971247	0.289036	6.820066	0.0000
C	5.620764	0.407314	13.79960	0.0000

R-squared	0.768646	Mean dependent var	8.390384
Adjusted R-squared	0.752121	S.D. dependent var	0.252566
S.E. of regression	0.125746	Akaike info criterion	-1.192629
Sum squared resid	0.221371	Schwarz criterion	-1.096056
Log likelihood	11.54103	F-statistic	46.51330
Durbin-Watson stat	0.552745	Prob(F-statistic)	0.000008

Fuente. IEA.
Elaboración. Propia.

Análisis de los residuos

Se comienza el análisis de los residuos del modelo a través del estadístico de *Q* de *Ljung-Box*, de cuyo resultado se puede afirmar al 95% de confianza que los residuos no se comportan con arreglo a un ruido blanco.

Tabla.III.6.13. Correlograma de los residuos. Modelo 2.

Sample: 1984 1999
Included observations: 16

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. |***. |	. |***. |	1	0.444	0.444	3.7927	0.051
. | . |	. **| . |	2	0.022	-0.219	3.8022	0.149
. **| . |	. *| . |	3	-0.227	-0.185	4.9471	0.176
.***| . |	.***| . |	4	-0.448	-0.338	9.7672	0.045
. **| . |	. | . |	5	-0.271	0.060	11.697	0.039
. *| . |	. **| . |	6	-0.176	-0.219	12.588	0.050
. *| . |	. **| . |	7	-0.160	-0.229	13.406	0.063
. |* . |	. |* . |	8	0.103	0.097	13.790	0.087
. |** . |	. |* . |	9	0.253	0.091	16.417	0.059
. |** . |	. | . |	10	0.282	0.021	20.224	0.027
. |* . |	. *| . |	11	0.177	-0.097	22.028	0.024
. *| . |	. *| . |	12	-0.119	-0.130	23.048	0.027

Fuente. IEA.
Elaboración. Propia.

Análisis de los residuos al cuadrado

El correlograma de los residuos al cuadrado refleja que sus autocorrelaciones son cercanas a cero, no rechazándose la hipótesis de homocedasticidad autorregresiva en los residuos al nivel del 95% de confianza.

Tabla.III.6.14. Correlograma de los residuos al cuadrado. Modelo 2.

Sample: 1984 1999
Included observations: 16

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. | . |	. | . |	1	0.061	0.061	0.0715	0.789
. *| . |	. *| . |	2	-0.111	-0.116	0.3266	0.849
. | . |	. | . |	3	-0.040	-0.026	0.3618	0.948
. |** . |	. |** . |	4	0.211	0.207	1.4342	0.838
. | . |	. | . |	5	0.005	-0.032	1.4349	0.920
. *| . |	. | . |	6	-0.059	-0.017	1.5348	0.957
. | . |	. | . |	7	-0.025	-0.006	1.5550	0.980
. *| . |	. *| . |	8	-0.116	-0.177	2.0418	0.980
. *| . |	. *| . |	9	-0.078	-0.062	2.2893	0.986
. *| . |	. *| . |	10	-0.060	-0.066	2.4618	0.991
. | . |	. | . |	11	-0.028	-0.047	2.5066	0.996
. *| . |	. | . |	12	-0.086	-0.042	3.0350	0.995

Fuente. IEA.
Elaboración. Propia.

Normalidad de los residuos

La prueba de normalidad de los residuos indica que la probabilidad del estadístico de Jarque-Bera es superior levemente al 5% de significación, apoyando esta circunstancia la normalidad de los residuos.

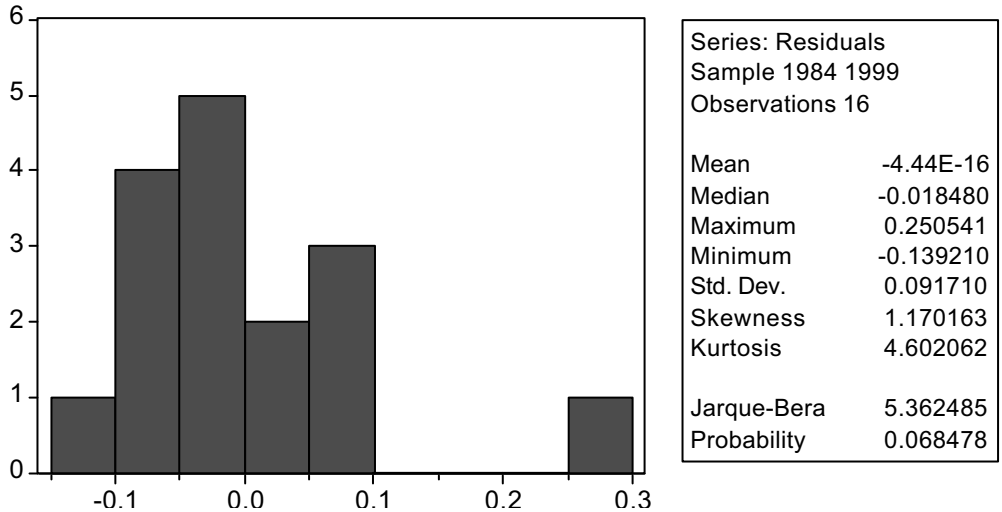


Figura.III.6.5.Histograma de probabilidad. Modelo 2.

Fuente. IEA.
Elaboración. Propia.

Correlación Serial de los residuos

El contraste de correlación serial de los residuos aplicado al modelo 2, rechaza la ausencia de autocorrelación de orden dos, al nivel de significación del 5%.

Tabla.III.6.15. Contraste de autocorrelación serial. Modelo 2

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	6.640926	Probability		0.012841
Obs*R-squared	8.751789	Probability		0.012577
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PE	-0.024883	0.182672	-0.136217	0.8941
Ln PGKWH	0.107264	0.259348	0.413592	0.6871
C	-0.083547	0.295629	-0.282608	0.7827
RESID(-1)	1.261317	0.346510	3.640063	0.0039
RESID(-2)	-0.630514	0.389731	-1.617816	0.1340
R-squared	0.546987	Mean dependent var		-6.66E-16
Adjusted R-squared	0.382255	S.D. dependent var		0.091710
S.E. of regression	0.072081	Akaike info criterion		-2.171744
Sum squared resid	0.057152	Schwarz criterion		-1.930310
Log likelihood	22.37396	F-statistic		3.320463
Durbin-Watson stat	1.998762	Prob(F-statistic)		0.051474

Fuente. IEA.
Elaboración. Propia.

Heterocedasticidad autorregresiva condicional de los residuos

En el análisis de la heterocedasticidad autorregresiva de los retardos se observa que no hay evidencia para afirmar que existe una relación ARCH entre los residuos, al 95% de confianza.

Tabla.III.6.16. Contraste de comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 2.

ARCH Test:				
F-statistic	0.637908	Probability		0.438818
Obs*R-squared	0.701620	Probability		0.402240
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample(adjusted): 1985 1999				
Included observations: 15 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.005580	0.005446	1.024509	0.3243
RESID^2(-1)	0.667071	0.835204	0.798692	0.4388
R-squared	0.046775	Mean dependent var		0.008399
Adjusted R-squared	-0.026550	S.D. dependent var		0.015856
S.E. of regression	0.016065	Akaike info criterion		-5.300731
Sum squared resid	0.003355	Schwarz criterion		-5.206325
Log likelihood	41.75549	F-statistic		0.637908
Durbin-Watson stat	1.090830	Prob(F-statistic)		0.438818

Fuente. IEA.
Elaboración. Propia.

Heterocedasticidad en los residuos

Al aplicar un contraste más completo como es el contraste de White, tabla III.6.17, se llega a concluir que en el modelo 2 no se puede rechazar la ausencia de heterocedasticidad y que los errores son independientes de las variables explicativas.

Tabla.III.6.17. Contraste de heterocedasticidad de los residuos. Modelo 2.

White Heteroskedasticity Test:				
F-statistic	2.150997	Probability		0.142238
Obs*R-squared	7.022234	Probability		0.134718
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample: 1984 1999				
Included observations: 16				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.188646	1.529801	-1.430674	0.1803
Ln PE	2.165189	1.024222	2.113984	0.0582
(Ln PE)^2	-0.467480	0.217497	-2.149364	0.0547
(Ln PGKWH)	-0.581693	1.062300	-0.547579	0.5949
(Ln PGKWH)^2	0.264924	0.379480	0.698124	0.4996
R-squared	0.438890	Mean dependent var		0.007885
Adjusted R-squared	0.234850	S.D. dependent var		0.015456
S.E. of regression	0.013520	Akaike info criterion		-5.519020
Sum squared resid	0.002011	Schwarz criterion		-5.277586
Log likelihood	49.15216	F-statistic		2.150997
Durbin-Watson stat	1.648750	Prob(F-statistic)		0.142238

Fuente. IEA.
Elaboración. Propia.

Errores de especificación en los residuos

La estimación del modelo 2 es correcta al nivel de significación del 5%, tal como se refleja en el contraste de Ramsey, tabla III.6.18.

Tabla.III.6.18. Contraste de Ramsey. Modelo 2.

Ramsey RESET Test:				
F-statistic	0.542773	Probability		0.475443
Log likelihood ratio	0.707808	Probability		0.400173
Test Equation:				
Dependent Variable: Q				
Method: Least Squares				
Sample: 1901 1916				
Included observations: 16				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PE	9.756715	12.18565	0.800673	0.4389
Ln PGKWH	14.61095	18.28204	0.799197	0.4397
C	12.92818	10.94896	1.180768	0.2606
FITTED^2	-0.695326	0.943799	-0.736731	0.4754
R-squared	0.873855	Mean dependent var		8.390384
Adjusted R-squared	0.842319	S.D. dependent var		0.252566
S.E. of regression	0.100292	Akaike info criterion		-1.549148
Sum squared resid	0.120701	Schwarz criterion		-1.356001
Log likelihood	16.39319	F-statistic		27.70960
Durbin-Watson stat	0.709306	Prob(F-statistic)		0.000011

Fuente. IEA.
Elaboración. Propia.

Una vez llegados a este punto, y habiendo comprobado que el modelo que explica la demanda doméstica de energía eléctrica en la C.A.A. frente al precio de la electricidad y al precio del gas es válido desde todos los supuestos del modelo lineal general de regresión se profundiza en la estabilidad de los residuos y de los coeficientes de regresión (a excepción de la autocorrelación de los residuos). Para completar el análisis acerca de la permanencia estructural del modelo 2 se aplican seguidamente varios contrastes basados en la estimación por mínimos cuadrados recursivos.

Estimación Recursiva de los residuos

En el gráfico, figura III.6.6, aparecen dos bandas de confianza correspondientes a dos veces la desviación típica del error típico. Los residuos se representan a través de la línea continua, observándose que a partir del año 1996 los errores se salen fuera de las bandas de confianza, lo cual es indicativo de una inestabilidad en los errores de la ecuación.

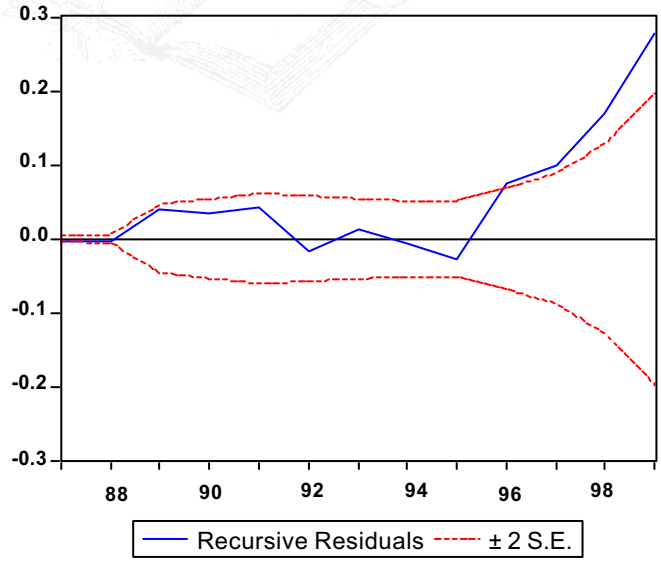


Figura.III.6.6. Residuos Recursivos. Modelo 2.
Fuente. IEA.
Elaboración. Propia.

Suma acumulada de los residuos recursivos

A través del contraste *CUSUM* se puede profundizar aún más. En este caso parece que los parámetros del modelo son estables a lo largo del periodo histórico; sin embargo al final del periodo la suma acumulada de los residuos recursivos tiende a aproximarse al límite superior de las bandas, lo cual es indicativo de falta de inestabilidad en los errores del modelo.

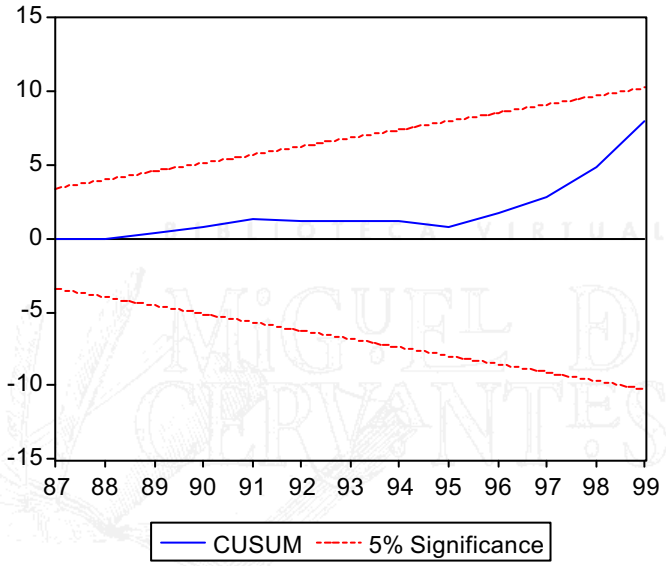


Figura.III.6.7. Suma acumulada de los residuos recursivos. Modelo 2.
Fuente. IEA.
Elaboración. Propia.

Estimación recursiva de los coeficientes

El contraste de los coeficientes recursivos, figua III.6.8, muestra la evolución de los tres coeficientes de la ecuación estimada a lo largo de todo el periodo muestral. Con respecto a la estabilidad de los coeficientes habría que decir que la evolución en los coeficientes de las variable Ln PGKWH y Ln PE son similares al modelo analizado anteriormente, modelo 1, a excepción de la evolución seguida por el término constante que fluctúa de forma más intensa y decreciente a partir del año 1995.

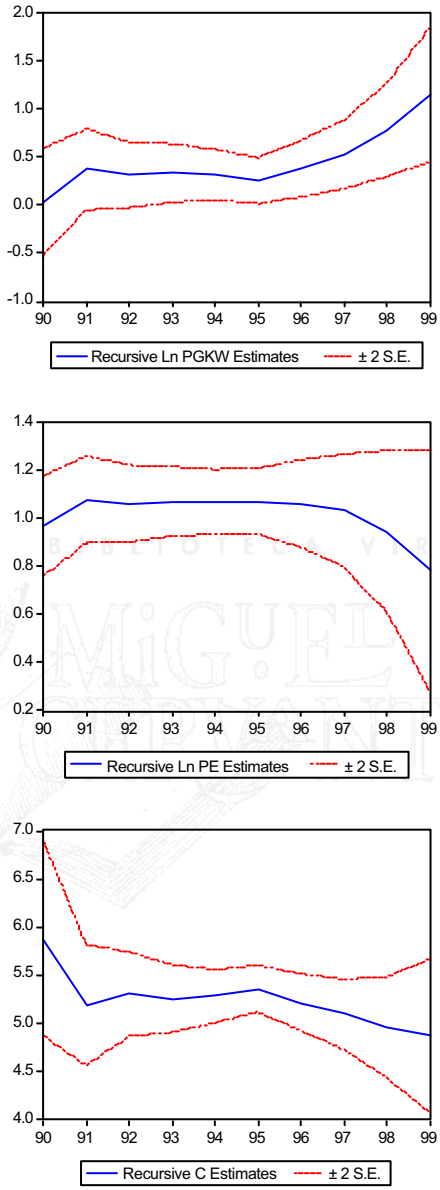


Figura.III.6.8. Gráficos de los coeficientes recursivos. Modelo 2.
Fuente. IEA.
Elaboración. Propia.

Análisis estructural del Modelo 3

En el tercer modelo considerado se relaciona la demanda anual de energía eléctrica en el sector residencial con el precio por unidad de kWh y el producto interior bruto generado en la Comunidad Autónoma de Andalucía durante el periodo 1984-1999. Para ello se efectúa una estimación mínimo cuadrática de los coeficientes del modelo, tabla III.6.19, indicando el coeficiente de regresión de la variable Ln PE, el coeficiente de elasticidad de la cantidad demandada de energía eléctrica respecto al precio (su valor es inferior a la unidad y negativo), que la curva de demanda de energía es inelástica. Además explica que al aumentar en 1% el precio del kWh la demanda sufre una disminución, por término medio, del 0,44%. Un comportamiento similar presenta la variable Ln PIB respecto a la demanda de electricidad, siendo ésta también de carácter inelástico. El coeficiente de regresión de la variable Ln PIB se interpreta como el coeficiente de la elasticidad demanda-renta, por afectar las variaciones del producto interior bruto a la riqueza personal de los ciudadanos (consumidores potenciales de electricidad) siendo el precio de la electricidad un factor *ceteris paribus* de la demanda. Su valor es 0,77, reflejando que la elasticidad de la demanda respecto a la renta es baja, sugiriendo esta circunstancia que la electricidad es un bien de primera necesidad, como ya se podría intuir.

El modelo resulta ser globalmente significativo al 95% de nivel de confianza, siendo el porcentaje de variación de demanda explicada por la variable Ln PE y Ln PIB de 0,9758, superior en un 15,08% al modelo 2 y un 15,98% al obtenido en el modelo 1.

Tabla.III.6.19.Estimación del modelo 3 de demanda de energía eléctrica en la C.A.A.

Dependent Variable: Ln Q
Method: Least Squares
Sample: 1984 1999
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PE	-0.441094	0.170500	-2.587057	0.0226
Ln PIB	0.771797	0.065975	11.69841	0.0000
C	-2.695904	0.672513	-4.008704	0.0015

R-squared	0.979032	Mean dependent var	8.390384
Adjusted R-squared	0.975806	S.D. dependent var	0.252566
S.E. of regression	0.039286	Akaike info criterion	-3.468560
Sum squared resid	0.020064	Schwarz criterion	-3.323700
Log likelihood	30.74848	F-statistic	303.4897
Durbin-Watson stat	1.458514	Prob(F-statistic)	0.000000

Fuente. IEA.
Elaboración. Propia.

Variables Irrelevantes

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

Una vez estimado el modelo, es necesario plantearse la posibilidad de poder omitir algunas de las variables explicativas consideradas. Para ello se analiza la posible ausencia de las mismas, tanto Ln PE como Ln PIB, llegando a la conclusión, tal como se puede observar en la tabla III.6.20, de que su omisión no explicaría mejor la demanda de energía eléctrica en la Comunidad Autónoma de Andalucía.

Tabla.III.6.20. Variables Irrelevantes. Modelo 3.

Redundant Variables: Ln PE

F-statistic	6.692863	Probability	0.022553
Log likelihood ratio	6.644911	Probability	0.009944

Test Equation:
Dependent Variable: Ln Q
Method: Least Squares
Sample: 1985 1999
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PIB	0.613906	0.029718	20.65803	0.0000
C	-1.288084	0.468654	-2.748478	0.0157

R-squared	0.968236	Mean dependent var	8.390384
Adjusted R-squared	0.965967	S.D. dependent var	0.252566
S.E. of regression	0.046593	Akaike info criterion	-3.178253
Sum squared resid	0.030393	Schwarz criterion	-3.081680
Log likelihood	27.42603	F-statistic	426.7542
Durbin-Watson stat	0.962208	Prob(F-statistic)	0.000000

Redundant Variables: Ln PIB			
F-statistic	136.8529	Probability	0.000000
Log likelihood ratio	39.11527	Probability	0.000000

Test Equation:
Dependent Variable: Ln Q
Method: Least Squares
Sample: 1985 1999
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PE	1.404039	0.211856	6.627335	0.0000
C	4.948220	0.520382	9.508824	0.0000

R-squared	0.758294	Mean dependent var	8.390384
Adjusted R-squared	0.741029	S.D. dependent var	0.252566
S.E. of regression	0.128529	Akaike info criterion	-1.148856
Sum squared resid	0.231276	Schwarz criterion	-1.052282
Log likelihood	11.19085	F-statistic	43.92157
Durbin-Watson stat	0.302808	Prob(F-statistic)	0.000011

Fuente. IEA.
Elaboración. Propia.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

Análisis de los residuos

El correlograma de los residuos se presenta en la tabla III.6.21, observándose que los coeficientes oscilan en torno a la media (valor cero), todos ellos dentro de las bandas de confianza, a excepción del cuarto retardo. El valor de estadístico *Q de Ljung-Box* es 12,940, acompañado de un valor de probabilidad superior al 5%, por lo que no se rechaza la hipótesis de incorrelación de los residuos.

Tabla.III.6.21. Correlograma de los residuos. Modelo 3.

Sample: 1985 1999
Included observations: 16

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. |* . |	. |* . |	1	0.179	0.179	0.6169	0.432
. |* . |	. |* . |	2	0.131	0.102	0.9685	0.616
. *| . |	. *| . |	3	-0.124	-0.171	1.3108	0.727
****| . |	****| . |	4	-0.518	-0.519	7.7555	0.101
. *| . |	. |* . |	5	-0.126	0.071	8.1698	0.147
. **| . |	. **| . |	6	-0.294	-0.190	10.652	0.100
. |* . |	. | . |	7	0.081	0.061	10.863	0.145
. |* . |	. *| . |	8	0.101	-0.178	11.231	0.189
. |* . |	. | . |	9	0.092	0.027	11.577	0.238
. | . |	.***| . |	10	0.027	-0.332	11.611	0.312
. *| . |	. *| . |	11	-0.142	-0.073	12.769	0.309
. | . |	. | . |	12	0.049	0.011	12.940	0.373

Fuente. IEA.
Elaboración. Propia.

Análisis de los residuos al cuadrado

Respecto al estudio de la heterocedasticidad condicional autorregresiva (con doce retardos) los coeficientes presentan un valor cercano a cero, siendo un poco más elevado el correspondiente a los desfases sexto, 0,322, y octavo, -0,262. Sin embargo, la probabilidad del estadístico *Q* es muy superior al nivel de significación del 5%; por lo tanto no hay evidencia muestral suficiente para rechazar la hipótesis que establece que los residuos no siguen un comportamiento ARCH.

Tabla.III.6.22.Correlograma de los residuos al cuadrado. Modelo 3.

Sample: 1901 1916
Included observations: 16

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. | . |	. | . |	1	-0.030	-0.030	0.0173	0.895
. *| . |	. *| . |	2	-0.165	-0.166	0.5768	0.749
. |* . |	. |* . |	3	0.185	0.179	1.3349	0.721
. **| . |	. **| . |	4	-0.195	-0.229	2.2499	0.690
. | . |	. |* . |	5	0.002	0.076	2.2500	0.814
. |** . |	. |** . |	6	0.322	0.236	5.2294	0.515
. *| . |	. | . |	7	-0.066	0.003	5.3700	0.615
. **| . |	. **| . |	8	-0.262	-0.261	7.8327	0.450
. | . |	. *| . |	9	-0.044	-0.148	7.9129	0.543
. *| . |	. *| . |	10	-0.139	-0.112	8.8466	0.547
. |* . |	. |* . |	11	0.088	0.157	9.2902	0.595
. | . |	. *| . |	12	0.031	-0.171	9.3600	0.672

Fuente. IEA.
Elaboración. Propia.

Normalidad de los residuos

En el análisis de los estadísticos descriptivos se observa que la media de los residuos es muy próxima a cero, y el coeficiente de asimetría es prácticamente nulo (-0,0746), no siendo el coeficiente de curtosis muy próximo a 3. Sin embargo, tras la aplicación del contraste de normalidad se puede afirmar al 95% de nivel de confianza que la distribución de los residuos se rige según una distribución normal.

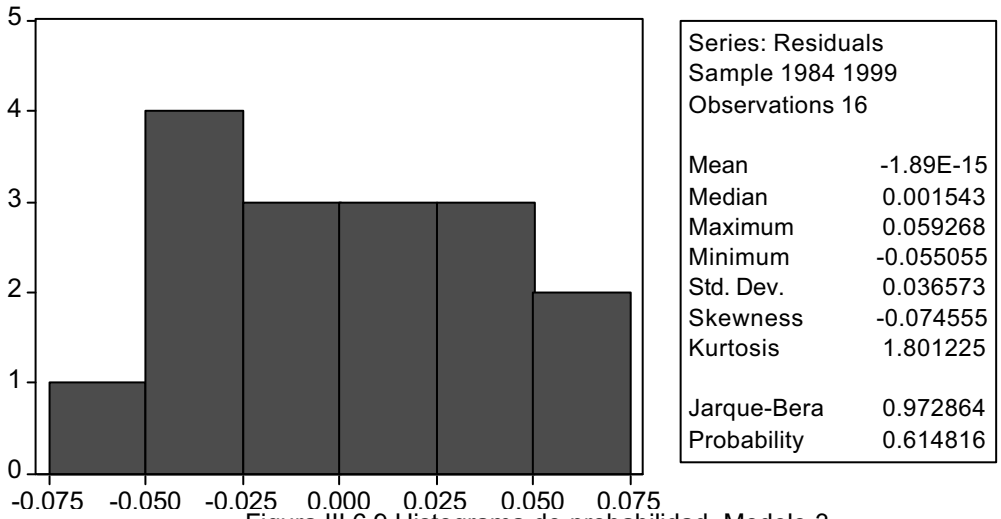


Figura.III.6.9.Histograma de probabilidad. Modelo 3.
Fuente. IEA.
Elaboración. Propia.

BIBLIOTECA VIRTUAL

Correlación Serial de los residuos

En el contraste de Breusch-Godfrey, de análisis de correlación serial de segundo orden, se observa que los estadísticos F y LM alcanzan valores muy elevados por lo que no se rechaza la hipótesis nula de ausencia de autocorrelación de orden dos en los residuos, al nivel de significación del 5%.

Tabla.III.6.23. Contraste de autocorrelación serial. Modelo 3.

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.245939	Probability		0.786157
Obs*R-squared	0.684835	Probability		0.710052
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
PE	0.002976	0.181396	0.016404	0.9872
PIB	4.97E-05	0.070171	0.000708	0.9994
C	-0.007837	0.715404	-0.010955	0.9915
RESID(-1)	0.162934	0.301167	0.541007	0.5993
RESID(-2)	0.103995	0.301683	0.344716	0.7368
R-squared	0.042802	Mean dependent var		6.11E-16
Adjusted R-squared	-0.305270	S.D. dependent var		0.036573
S.E. of regression	0.041784	Akaike info criterion		-3.262305
Sum squared resid	0.019205	Schwarz criterion		-3.020871
Log likelihood	31.09844	F-statistic		0.122969
Durbin-Watson stat	1.775278	Prob(F-statistic)		0.971228

Fuente. IEA.
Elaboración. Propia.

Heterocedasticidad autorregresiva condicional de los residuos

En el modelo 3, no se puede rechazar que los residuos no presenten un carácter heterocedástico al nivel de confianza del 95%, siendo el *p-valor* del contraste muy próximo a la unidad.

Tabla.III.6.24. Contraste de comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 3.

ARCH Test:				
F-statistic	0.009093	Probability		0.925486
Obs*R-squared	0.010484	Probability		0.918445
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample(adjusted): 1985 1999				
Included observations: 15 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.001134	0.000425	2.666844	0.0194
RESID^2(-1)	-0.023343	0.244793	-0.095356	0.9255
R-squared	0.000699	Mean dependent var		0.001103
Adjusted R-squared	-0.076170	S.D. dependent var		0.001025
S.E. of regression	0.001064	Akaike info criterion		-10.73064
Sum squared resid	1.47E-05	Schwarz criterion		-10.63623
Log likelihood	82.47980	F-statistic		0.009093
Durbin-Watson stat	1.687938	Prob(F-statistic)		0.925486

Fuente. IEA.
Elaboración. Propia.

Heterocedasticidad en los residuos

El modelo 3 presenta una estructura homocedástica en los residuos y éstos no están correlacionados con los regresores, con una probabilidad del 95%.

Tabla.III.6.25. Contraste de heterocedasticidad de los residuos. Modelo 3.

White Heteroskedasticity Test:				
F-statistic	1.106010	Probability		0.384737
Obs*R-squared	3.465752	Probability		0.325231
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample: 1984 1999				
Included observations: 16				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.063702	0.084233	0.756258	0.4641
PE	-0.021850	0.072271	-0.302338	0.7676
PE^2	0.005467	0.014882	0.367387	0.7197
PIB	-0.002656	0.001953	-1.359721	0.1989
R-squared	0.216609	Mean dependent var		0.001254
Adjusted R-squared	0.020762	S.D. dependent var		0.001159
S.E. of regression	0.001147	Akaike info criterion		-10.49073
Sum squared resid	1.58E-05	Schwarz criterion		-10.29758
Log likelihood	87.92583	F-statistic		1.106010
Durbin-Watson stat	2.211401	Prob(F-statistic)		0.384737

Fuente. IEA.
Elaboración. Propia.

Errores de especificación en los residuos

Ante la posible existencia de errores de especificación por la no inclusión en el modelo de relaciones no lineales, se realiza el contraste de Ramsey (tabla III.6.26), donde los resultados permiten rechazar la hipótesis nula de linealidad a un nivel del 5% (incluso al 1%, pues la probablidad del estadístico es 0,004102).

Tabla.III.6.26. Contraste de Ramsey. Modelo 3.

Ramsey RESET Test:				
F-statistic	12.50237	Probability		0.004102
Log likelihood ratio	11.42181	Probability		0.000726
Test Equation:				
Dependent Variable: Q				
Method: Least Squares				
Sample: 1901 1916				
Included observations: 16				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
PE	6.328546	1.918586	3.298548	0.0064
PIB	-10.06697	3.065752	-3.283687	0.0065
C	93.79049	27.29228	3.436521	0.0049
FITTED^2	0.820277	0.231987	3.535869	0.0041
R-squared	0.989731	Mean dependent var		8.390384
Adjusted R-squared	0.987163	S.D. dependent var		0.252566
S.E. of regression	0.028615	Akaike info criterion		-4.057423
Sum squared resid	0.009826	Schwarz criterion		-3.864276
Log likelihood	36.45939	F-statistic		385.5119
Durbin-Watson stat	2.210513	Prob(F-statistic)		0.000000

Fuente. IEA.
Elaboración. Propia.

Una vez realizada la prueba de Ramsey, basada en la correcta especificación lineal del modelo, se concluye que el modelo 3 no resulta válido desde el punto de vista de su estructura funcional. Detectada esta no linealidad entre los parámetros y variables, se procede a la construcción de modelos alternativos basados en productos cruzados entre las variables explicativas que constituyen el modelo 3.

De esta forma, se construyen modelos de prueba para la búsqueda de un modelo que represente más fielmente la variación de la demanda de electricidad. En primer lugar, se construye el modelo denominado, *modelo 3 cruzado 3 variables*, el cual considera las mismas variables que el modelo visto anteriormente (modelo 3), incluyendo una nueva variable que es el producto de la variable Ln PE y Ln PIB. La expresión funcional del modelo se presenta como:

$$LnQ_i = Ln\alpha + \beta_1 LnPE_i + \beta_2 LnPIB_i + \beta_3 (LnPELnPIB)_i + u_i$$

En la tabla III.6.27, se refleja la estimación del modelo donde se observa la no significatividad al 5% de la variable Ln *PIB*. Sin embargo el modelo

globalmente resulta ser significativo. Los resultados obtenidos parecen mostrar la conveniencia de no considerar la variable PIB en el modelo.

Tabla.III.6.27. Estimación del modelo 3 cruzado 3 variables de demanda de energía eléctrica en la C.A.A.

Method: Least Squares
Sample: 1984 1999
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LnPE	-7.888889	2.905835	-2.714844	0.0188
LnPIB	-0.416970	0.466527	-0.893775	0.3890
LnPELnPIB	0.482120	0.187878	2.566138	0.0247
C	15.64382	7.168921	2.182172	0.0497

R-squared	0.986461	Mean dependent var	8.390384
Adjusted R-squared	0.983076	S.D. dependent var	0.252566
S.E. of regression	0.032857	Akaike info criterion	-3.780985
Sum squared resid	0.012955	Schwarz criterion	-3.587837
Log likelihood	34.24788	F-statistic	291.4370
Durbin-Watson stat	2.081424	Prob(F-statistic)	0.000000

Fuente. IEA.
Elaboración. Propia.

En la tabla III.6.28, se aplica un contraste de redundancia de la variable Ln *PIB*, siendo no significativo y apoyando la conveniencia de excluir al *PIB* ya que así aumenta la fuerza explicativa del modelo.

Tabla.III.6.28. Variables Irrelevantes. Modelo 3 cruzado 3 variables.

Redundant Variables: PIB

F-statistic	0.798835	Probability	0.389016
Log likelihood ratio	1.031159	Probability	0.309887

Test Equation:
Dependent Variable: Q
Method: Least Squares
Sample: 1984 1999
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LnPE	-5.327008	0.473641	-11.24693	0.0000
LnPE LnPIB	0.315378	0.022049	14.30364	0.0000
C	9.243230	0.328005	28.18011	0.0000

R-squared	0.985559	Mean dependent var	8.390384
Adjusted R-squared	0.983338	S.D. dependent var	0.252566
S.E. of regression	0.032602	Akaike info criterion	-3.841537
Sum squared resid	0.013817	Schwarz criterion	-3.696677
Log likelihood	33.73230	F-statistic	443.6208
Durbin-Watson stat	1.969532	Prob(F-statistic)	0.000000

Fuente. IEA.
Elaboración. Propia.

Una vez estimado el modelo correctamente y denominando este nuevo modelo, *modelo 3 productos cruzados (modelo 3PC)*, se procede a la evaluación de los errores, desde el punto de vista de su autocorrelación, normalidad y heterocedasticidad.

Análisis de los residuos

En la tabla III.6.29 se representa el correlograma de los residuos, siendo las autocorrelaciones (AC) y las autocorrelaciones parciales (PAC) próximas a cero en la mayoría de los doce retardos efectuados a excepción del retardo 4. Sin embargo, el estadístico *Q* de *Ljung-Box* alcanza el valor 16,109 y la probabilidad 0,186 superior a 0,05, por lo que se no se puede rechazar la hipótesis de ausencia de autocorrelación al 95% de confianza.

Tabla.III.6.29. Correlograma de los residuos. Modelo 3PC.

Sample: 1984 1999
Included observations: 16

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. | . |	. | . |	1	-0.029	-0.029	0.0159	0.900
. | . |	. | . |	2	0.052	0.051	0.0721	0.965
. *| . |	. *| . |	3	-0.116	-0.113	0.3684	0.947
*****| . |	*****| . |	4	-0.645	-0.664	10.355	0.035
. | . |	. | . |	5	0.039	-0.049	10.394	0.065
. *| . |	. *| . |	6	-0.170	-0.179	11.230	0.082
. |* . |	. | . |	7	0.178	-0.027	12.249	0.093
. |** . |	. **| . |	8	0.249	-0.288	14.473	0.070
. |* . |	. | . |	9	0.072	0.048	14.684	0.100
. | . |	. **| . |	10	0.041	-0.302	14.765	0.141
. *| . |	. *| . |	11	-0.148	-0.094	16.032	0.140
. | . |	. *| . |	12	-0.033	-0.109	16.109	0.186

Fuente. IEA.
Elaboración. Propia.

Análisis de los residuos al cuadrado

La heterocedasticidad condicional autorregresiva (con doce retardos) no se presenta en el modelo, ya que la probabilidad resulta ser 0,864 superior a 0,05 con un nivel de confianza del 95%.

Tabla.III.6.30. Correlograma de los residuos al cuadrado. Modelo 3PC.

Sample: 1984 1999
Included observations: 16

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. **| . |	. **| . |	1	-0.205	-0.205	0.8077	0.369
. |** . |	. |** . |	2	0.314	0.284	2.8326	0.243
. **| . |	. *| . |	3	-0.223	-0.134	3.9302	0.269
. | . |	. *| . |	4	0.035	-0.115	3.9603	0.411
. *| . |	. | . |	5	-0.130	-0.044	4.4042	0.493
. *| . |	. *| . |	6	-0.129	-0.184	4.8831	0.559
. *| . |	. *| . |	7	-0.120	-0.157	5.3474	0.618
. *| . |	. *| . |	8	-0.068	-0.059	5.5117	0.702
. *| . |	. *| . |	9	-0.128	-0.183	6.1866	0.721
. *| . |	. **| . |	10	-0.068	-0.196	6.4090	0.780
. | . |	. | . |	11	0.055	0.020	6.5836	0.832
. |* . |	. | . |	12	0.066	0.013	6.9009	0.864

Fuente. IEA.
Elaboración. Propia.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

Normalidad de los residuos

El estadístico de Jarque-Bera resulta ser 0,423097 muy inferior al nivel crítico, además su p-valor es superior a 0,05, por lo que se evidencia al 95% de confianza la normalidad de la distribución de los residuos.

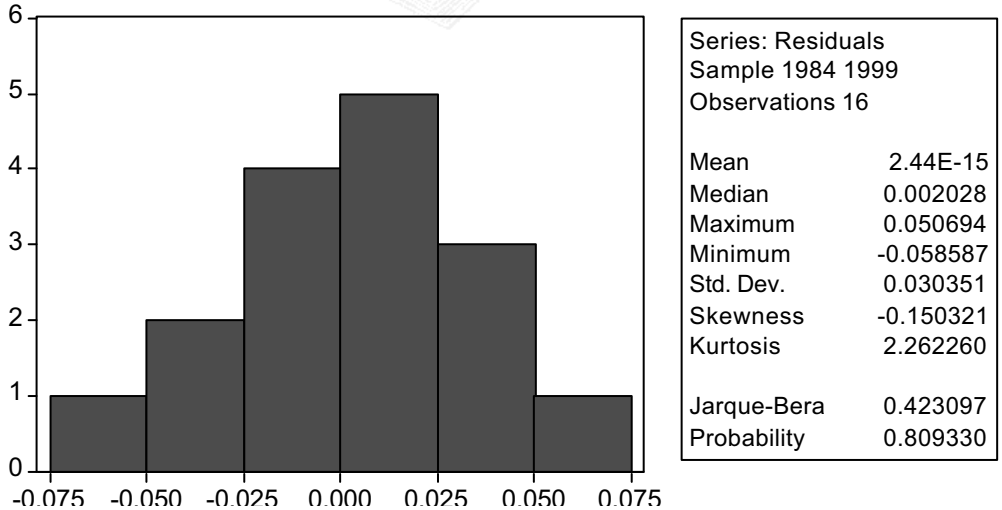


Figura.III.6.10. Histograma de probabilidad. Modelo 3PC.
Fuente. IEA.
Elaboración. Propia.

Correlación serial de los residuos

La prueba de Breusch-Godfrey de la autocorrelación de los residuos evidencia la ausencia de relación entre los mismos (orden 2) al 95% de confianza.

Tabla.III.6.31. Contraste de autocorrelacion serial. Modelo 3PC.

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.020036	Probability		0.980199
Obs*R-squared	0.058075	Probability		0.971380
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
PEPIB	0.000613	0.024720	0.024817	0.9806
PE	-0.013302	0.530825	-0.025059	0.9805
C	0.008894	0.366026	0.024300	0.9810
RESID(-1)	-0.026997	0.308624	-0.087475	0.9319
RESID(-2)	0.054605	0.312762	0.174589	0.8646
R-squared	0.003630	Mean dependent var		2.44E-15
Adjusted R-squared	-0.358687	S.D. dependent var		0.030351
S.E. of regression	0.035378	Akaike info criterion		-3.595173
Sum squared resid	0.013767	Schwarz criterion		-3.353739
Log likelihood	33.76139	F-statistic		0.010018
Durbin-Watson stat	1.900021	Prob(F-statistic)		0.999767

Fuente. IEA.
Elaboración. Propia.

Heterocedasticidad autorregresiva condicional de los residuos

La aplicación del constraste ARCH evidencia la homocedasticidad autorregresiva al 5% de nivel de significación.

Tabla.III.6.32. Contraste comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 3PC.

ARCH Test:				
F-statistic	0.581663	Probability		0.459277
Obs*R-squared	0.642406	Probability		0.422841
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample(adjusted): 1985 1999				
Included observations: 15 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.001052	0.000366	2.871099	0.0131
RESID^2(-1)	-0.208777	0.273745	-0.762668	0.4593
R-squared	0.042827	Mean dependent var		0.000865
Adjusted R-squared	-0.030802	S.D. dependent var		0.001037
S.E. of regression	0.001053	Akaike info criterion		-10.75063
Sum squared resid	1.44E-05	Schwarz criterion		-10.65622
Log likelihood	82.62970	F-statistic		0.581663
Durbin-Watson stat	1.811286	Prob(F-statistic)		0.459277

Fuente. IEA.
Elaboración. Propia.

Heterocedasticidad en los residuos

Si se realiza el contraste de White, se observa que no se puede rechazar la hipótesis nula de homocedasticidad, pues la probabilidad de rechazarla siendo cierta se eleva a 0,2672.

Tabla.III.6.33. Contraste de heterocedasticidad de los residuos. Modelo 3PC.

White Heteroskedasticity Test:				
F-statistic	1.324428	Probability		0.320796
Obs*R-squared	5.200938	Probability		0.267294
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample: 1984 1999				
Included observations: 16				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.540933	0.413326	1.308732	0.2173
Ln(PEPIB)	0.037103	0.029964	1.238246	0.2414
Ln(PEPIB)^2	-0.000472	0.000372	-1.268512	0.2308
Ln(PE)	-1.023741	0.799237	-1.280897	0.2266
Ln(PE)^2	0.206446	0.157652	1.309507	0.2171
R-squared	0.325059	Mean dependent var		0.000864
Adjusted R-squared	0.079625	S.D. dependent var		0.001002
S.E. of regression	0.000961	Akaike info criterion		-10.80618
Sum squared resid	1.02E-05	Schwarz criterion		-10.56474
Log likelihood	91.44941	F-statistic		1.324428
Durbin-Watson stat	3.360828	Prob(F-statistic)		0.320796

Fuente. IEA.
Elaboración. Propia.

Errores de especificación por variables omitidas

Los resultados del contraste de Ramsey, tabla III.6.34, permiten aceptar la hipótesis nula de linealidad a un nivel de significación del 5% pues el valor de la probabilidad del estadístico es 0,2297.

Tabla.III.6.34. Contraste de Ramsey. Modelo 3PC.

Ramsey RESET Test:				
F-statistic	1.604724	Probability		0.229270
Log likelihood ratio	2.008167	Probability		0.156454
Test Equation:				
Dependent Variable: Q				
Method: Least Squares				
Sample: 1984 1999				
Included observations: 16				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
PEPIB	-2.065234	1.879391	-1.098885	0.2934
PE	35.18634	31.98479	1.100096	0.2929
C	-28.98096	30.17606	-0.960396	0.3558
FITTED^2	0.440609	0.347819	1.266777	0.2293
R-squared	0.987263	Mean dependent var		8.390384
Adjusted R-squared	0.984078	S.D. dependent var		0.252566
S.E. of regression	0.031869	Akaike info criterion		-3.842048
Sum squared resid	0.012188	Schwarz criterion		-3.648900
Log likelihood	34.73638	F-statistic		310.0394
Durbin-Watson stat	2.019104	Prob(F-statistic)		0.000000

Fuente. IEA.
Elaboración. Propia.

Estimación recursiva de los residuos

En la figura III.6.11 se puede observar que los residuos varían en torno al valor cero. En el año 1991 sobrepasa levemente las bandas de confianza. Por lo tanto, no existe evidencia de ausencia de autocorrelación.

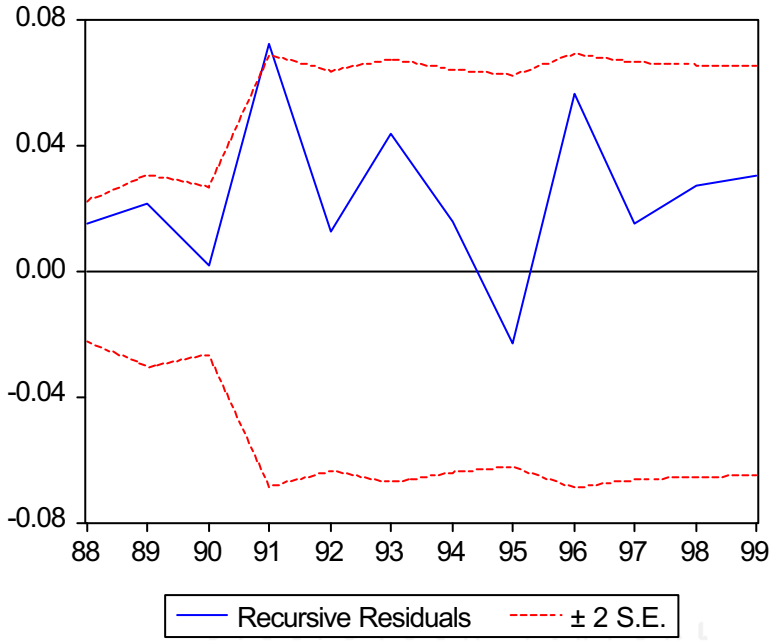


Figura.III.6.11. Residuos recursivos. Modelo 3PC.
Fuente. IEA.
Elaboración. Propia.

Suma acumulada de los residuos recursivos

El estadístico CUSUM oscila dentro de sus bandas de confianza como refleja la figura III.6.12; sin embargo, a partir de 1990 se produce un crecimiento fuerte del mismo que le aleja del valor cero de forma continuada para el resto de la muestra, lo que constituye en su conjunto un indicio de inestabilidad del modelo.

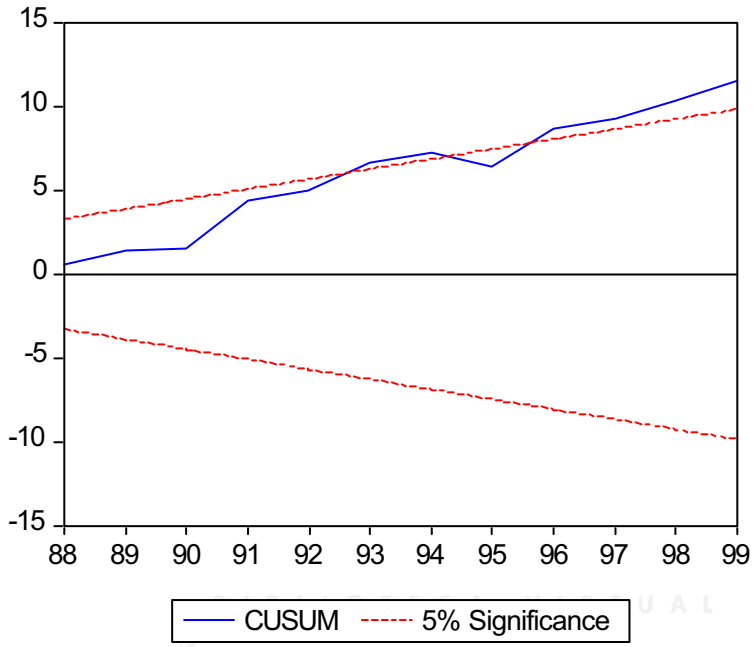


Figura.III.6.12. Suma acumulada de los residuos recursivos. Modelo 3PC.
Fuente. IEA.
Elaboración. Propia.

Estimación recursiva de los coeficientes

La estabilidad de los coeficientes, figura III.6.13, presenta elevadas fluctuaciones desde el periodo 1991-1995 para, a partir de este periodo, estabilizarse en todo su restante recorrido.

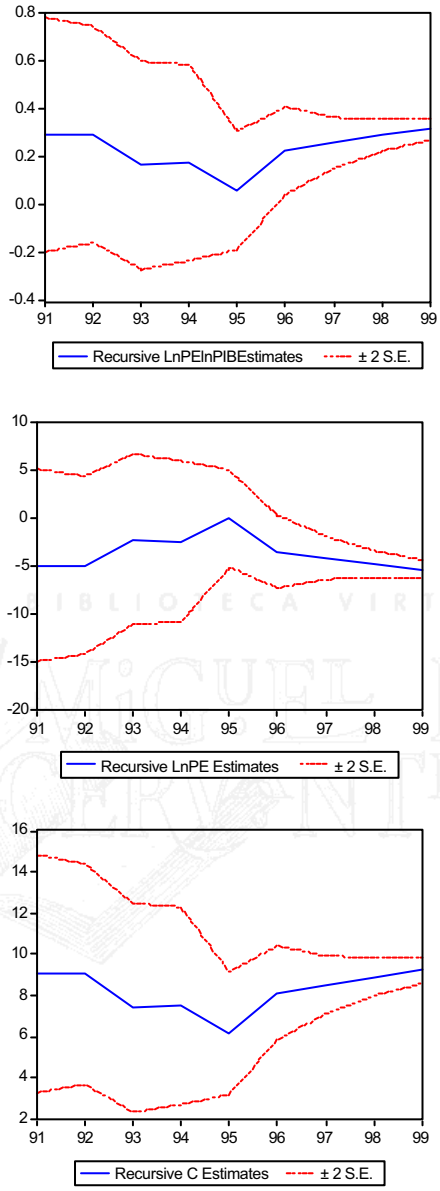


Figura.III.6.13. Gráficos de los coeficientes recursivos. Modelo 3PC.
Fuente. IEA.
Elaboración. Propia.

Análisis de estructural del Modelo 4

Finalmente, el último modelo que se ha propuesto para explicar la demanda de electricidad es el modelo 4, el cual relaciona la demanda doméstica de electricidad en la C.A.A. con el producto interior bruto y la demanda de electricidad retardada un periodo. La introducción de la demanda retardada un periodo como variable explicativa es debido a que la demanda de electricidad en un periodo puede estar relacionada con la demanda que tuvo lugar el año anterior. Por lo tanto, a los modelos econométricos que consideran como variable explicativa la variable explicada retardada un periodo se les denomina modelos autorregresivos de un único periodo.

Una vez estimado el modelo, tabla III.6.35, se observa en primer lugar que el coeficiente de elasticidad-renta es significativo al 5% (no siéndolo al 1% de significación). El valor del coeficiente es inferior a la unidad, verificando la condición de que la electricidad es un bien de primera necesidad, siendo el incremento de demanda, por término medio, de un 0,317% por cada 1% de incremento del PIB en la Comunidad Autónoma de Andalucía.

La influencia de la demanda de electricidad sufrida en el periodo anterior resulta también significativa al 5% (no siéndolo al 1% de significación). La demanda del periodo actual resulta ser inelástica frente a incrementos o decrementos en el periodo anterior; de hecho, al incrementarse en un 1% la demanda en un periodo anterior el incremento que sufre la demanda en el periodo actual es del 0,52%.

Tabla.III.6.35. Estimación del modelo 4 de demanda de energía eléctrica en la C.A.A.

Dependent Variable: Q
Method: Least Squares
Sample(adjusted): 1984 1999
Included observations: 15 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PIB	0.317360	0.139326	2.277817	0.0418
Q(-1)	0.521911	0.215503	2.421827	0.0322
C	-0.968061	0.540698	-1.790391	0.0986

R-squared	0.982366	Mean dependent var	8.415830
Adjusted R-squared	0.979427	S.D. dependent var	0.239262
S.E. of regression	0.034318	Akaike info criterion	-3.729436
Sum squared resid	0.014133	Schwarz criterion	-3.587826
Log likelihood	30.97077	F-statistic	334.2536
Durbin-Watson stat	2.467082	Prob(F-statistic)	0.000000

Fuente. IEA.
Elaboración. Propia.

Variables Irrelevantes

Respecto a la posibilidad de excluir algunas de las dos variables consideradas en el modelo 4, en la tabla III.6.36 se observa que la ausencia de alguna de ellas no favorecería una mejor explicación de la demanda de electricidad en la comunidad.

Tabla.III.6.36. Variables Irrelevantes. Modelo 4.

Redundant Variables: Ln PIB

F-statistic	5.188451	Probability	0.041843
Log likelihood ratio	5.389966	Probability	0.020253

Test Equation:
Dependent Variable:Ln Q
Method: Least Squares
Sample: 1985 1999
Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln Q(-1)	1.004679	0.044855	22.39827	0.0000
C	0.013923	0.375252	0.037102	0.9710

R-squared	0.974742	Mean dependent var	8.415830
Adjusted R-squared	0.972799	S.D. dependent var	0.239262
S.E. of regression	0.039461	Akaike info criterion	-3.503439
Sum squared resid	0.020243	Schwarz criterion	-3.409032
Log likelihood	28.27579	F-statistic	501.6824
Durbin-Watson stat	2.654396	Prob(F-statistic)	0.000000

Redundant Variables: Ln Q(-1)			
F-statistic	5.865245	Probability	0.032210
Log likelihood ratio	5.969258	Probability	0.014557

Test Equation:
Dependent Variable: Ln Q
Method: Least Squares
Sample: 1985 1999
Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PIB	0.649211	0.029565	21.95870	0.0000
C	-1.851868	0.467707	-3.959465	0.0016

R-squared	0.973747	Mean dependent var	8.415830
Adjusted R-squared	0.971728	S.D. dependent var	0.239262
S.E. of regression	0.040230	Akaike info criterion	-3.464819
Sum squared resid	0.021040	Schwarz criterion	-3.370413
Log likelihood	27.98614	F-statistic	482.1846
Durbin-Watson stat	1.307216	Prob(F-statistic)	0.000000

Fuente. IEA.
Elaboración. Propia.

Análisis de los residuos

En el estudio de la autocorrelación de los residuos, se puede observar que las autocorrelaciones de primer, cuarto y quinto orden se encuentran alejadas del valor cero, sin que por ello la hipótesis nula de no autocorrelación sea rechazada al 5% de significación, tal como se refleja en la tabla III.6.37.

Tabla.III.6.37. Correlograma de los residuos. Modelo 4.

Sample: 1985 1999
Included observations: 15

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
.***| . |	.***| . |	1	-0.326	-0.326	1.9393	0.164
. |* . |	. | . |	2	0.101	-0.006	2.1396	0.343
. |* . |	. |** . |	3	0.166	0.220	2.7228	0.436
*****| . |	****| . |	4	-0.595	-0.559	10.923	0.027
. |** . |	. *| . |	5	0.281	-0.075	12.934	0.024
. *| . |	. |* . |	6	-0.089	0.118	13.160	0.041
. | . |	. |* . |	7	-0.035	0.104	13.199	0.067
. |* . |	. **| . |	8	0.184	-0.280	14.430	0.071
. *| . |	. *| . |	9	-0.172	-0.100	15.691	0.074
. *| . |	. *| . |	10	-0.075	-0.162	15.981	0.100
. | . |	. *| . |	11	-0.012	-0.066	15.991	0.141
. *| . |	. *| . |	12	-0.071	-0.107	16.415	0.173

Fuente. IEA.
Elaboración. Propia.

Análisis de los residuos al cuadrado

Cuando se estudia la presencia de heterocedasticidad autorregresiva condicional ocurre una situación parecida a la prueba realizada anteriormente, no rechazándose la hipótesis de que los residuos no tienen un comportamiento ARCH.

Tabla.III.6.38. Correlograma de los residuos al cuadrado. Modelo 4.

Sample: 1985 1999
Included observations: 15

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. |* . |	. |* . |	1	0.068	0.068	0.0836	0.772
. **| . |	. **| . |	2	-0.295	-0.301	1.7888	0.409
. | . |	. | . |	3	-0.017	0.032	1.7950	0.616
. |**** |	. |***. |	4	0.501	0.455	7.6221	0.106
. *| . |	. **| . |	5	-0.104	-0.255	7.8972	0.162
. **| . |	. | . |	6	-0.257	0.007	9.7757	0.134
. *| . |	. *| . |	7	-0.109	-0.173	10.156	0.180
. | . |	.***| . |	8	0.032	-0.329	10.193	0.252
. *| . |	. | . |	9	-0.123	0.046	10.837	0.287
. *| . |	. *| . |	10	-0.098	-0.064	11.321	0.333
. | . |	. | . |	11	-0.022	0.038	11.353	0.414
. | . |	. |* . |	12	-0.014	0.088	11.369	0.498

Fuente. IEA.
Elaboración. Propia.

Normalidad de los residuos

La media muestral de los residuos es próxima a cero, el coeficiente de asimetría próximo a 0, ya que toma el valor –0,1847 y el coeficiente de curtosis es 2,40 próximo a 3. De esta forma se puede afirmar que los residuos del modelo 4 se distribuyen con arreglo a una ley Normal, con un nivel de confianza del 95%.

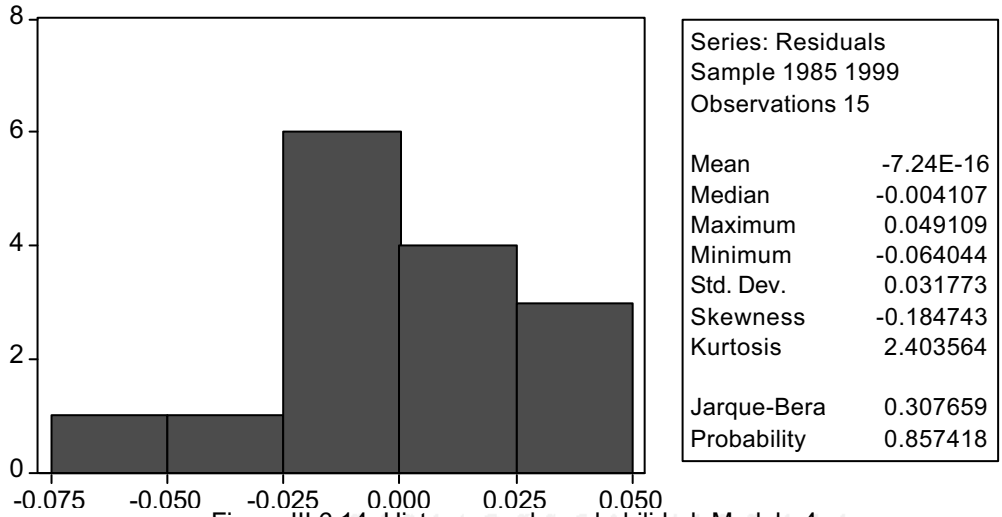


Figura.III.6.14. Histograma de probabilidad. Modelo 4.
Fuente. IEA.
Elaboración. Propia.

Correlación Serial de los residuos

Tras la aplicación del contraste de correlación serial de los residuos los estadísticos F y LM alcanzan valores elevados, por lo que se puede aceptar que exista ausencia de autocorrelación de orden dos en los residuos, al nivel de significación del 5%.

Tabla.III.6.39. Contraste de autocorrelación serial. Modelo 4.

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	1.226611	Probability		0.333881
Obs*R-squared	2.954924	Probability		0.228216
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
PIB	-0.147715	0.166642	-0.886421	0.3962
Q(-1)	0.212505	0.251390	0.845322	0.4177
C	0.556144	0.656624	0.846975	0.4168
RESID(-1)	-0.667691	0.442117	-1.510214	0.1619
RESID(-2)	-0.206545	0.400821	-0.515306	0.6175
R-squared	0.196995	Mean dependent var		-2.96E-16
Adjusted R-squared	-0.124207	S.D. dependent var		0.031772
S.E. of regression	0.033688	Akaike info criterion		-3.682164
Sum squared resid	0.011349	Schwarz criterion		-3.446147
Log likelihood	32.61623	F-statistic		0.613305
Durbin-Watson stat	1.620020	Prob(F-statistic)		0.662745

Fuente. IEA.

Elaboración. Propia.

Heterocedasticidad autorregresiva condicional de los residuos

Se puede afirmar con una probabilidad del 95% que los residuos del modelo 4 tienen una estructura homocedástica.

Tabla.III.6.40. Contraste de comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 4.

ARCH Test:				
F-statistic	0.063757	Probability		0.804927
Obs*R-squared	0.073990	Probability		0.785614
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample(adjusted): 1986 1999				
Included observations: 14 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000907	0.000419	2.165664	0.0512
RESID^2(-1)	0.075242	0.297988	0.252501	0.8049
R-squared	0.005285	Mean dependent var		0.000971
Adjusted R-squared	-0.077608	S.D. dependent var		0.001193
S.E. of regression	0.001239	Akaike info criterion		-10.41802
Sum squared resid	1.84E-05	Schwarz criterion		-10.32673
Log likelihood	74.92616	F-statistic		0.063757
Durbin-Watson stat	1.885288	Prob(F-statistic)		0.804927

Fuente. IEA.
Elaboración. Popia.

Heterocedasticidad en los residuos

Tras la aplicación del contraste de White se concluye que no existe evidencia para rechazar la hipótesis de igualdad de varianza en los residuos, al nivel del 95%. Además también se puede afirmar la independencia de éstos frente a las variables explicativas.

Tabla.III.6.41. Contraste de heterocedasticidad de los residuos. Modelo 4.

White Heteroskedasticity Test:				
F-statistic	1.026397	Probability		0.439693
Obs*R-squared	4.365915	Probability		0.358744
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample: 1985 1999				
Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.677275	0.909392	0.744756	0.4736
PIB	-0.277742	0.241025	-1.152337	0.2760
PIB^2	0.008954	0.007702	1.162559	0.2720
Q(-1)	0.358521	0.334841	1.070721	0.3095
Q(-1)^2	-0.021758	0.020034	-1.086080	0.3029
R-squared	0.291061	Mean dependent var		0.000942
Adjusted R-squared	0.007485	S.D. dependent var		0.001155
S.E. of regression	0.001151	Akaike info criterion		-10.43507
Sum squared resid	1.32E-05	Schwarz criterion		-10.19906
Log likelihood	83.26304	F-statistic		1.026397
Durbin-Watson stat	2.469459	Prob(F-statistic)		0.439693

Fuente. IEA.
Elaboración. Propia.

Errores de especificación en los residuos

La especificación y estimación inicial del modelo 4 resulta ser adecuada y, además, los errores se rigen con arreglo a una ley normal, al 5% de significación.

Tabla.III.6.42. Contraste de Ramsey. Modelo 4.

Ramsey RESET Test:				
F-statistic	0.332746	Probability		0.575661
Log likelihood ratio	0.447018	Probability		0.503755
Test Equation:				
Dependent Variable: Q				
Method: Least Squares				
Sample: 1985 1999				
Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
PE	-0.238527	0.455189	-0.524017	0.6107
Q(-1)	9.655237	15.00263	0.643570	0.5330
C	-35.92434	62.32368	-0.576416	0.5759
FITTED^2	-0.505297	0.875972	-0.576842	0.5757
R-squared	0.975487	Mean dependent var		8.415830
Adjusted R-squared	0.968802	S.D. dependent var		0.239262
S.E. of regression	0.042261	Akaike info criterion		-3.266732
Sum squared resid	0.019646	Schwarz criterion		-3.077919
Log likelihood	28.50049	F-statistic		145.9152
Durbin-Watson stat	2.858166	Prob(F-statistic)		0.000000

Fuente. IEA.
Elaboración. Propia.

Estimación Recursiva de los residuos

En la gráfica de residuos recursivos, figura III.6.15, se observa que en el año 1996 existe un problema desde el punto de vista de la estabilidad de los residuos.

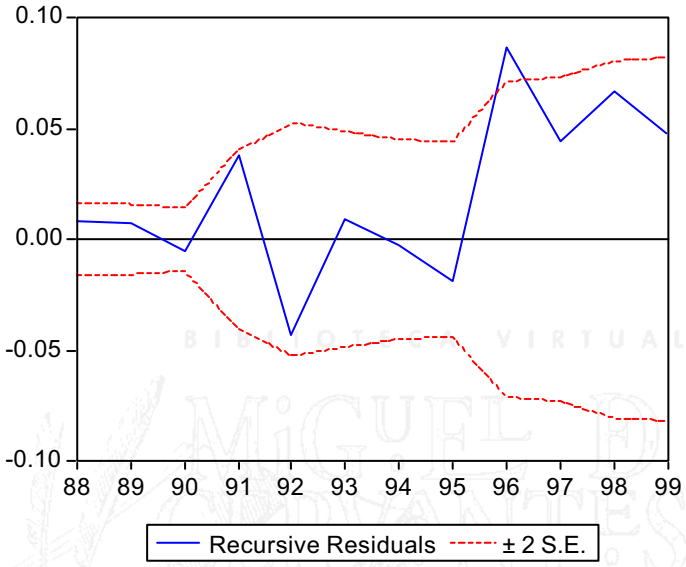


Figura.III.6.15. Residuos recursivos. Modelo 4.
Fuente. IEA.
Elaboración. Propia.

Suma acumulada de los residuos recursivos

En el gráfico de los residuos recursivos acumulados se observa que estos se encuentran en torno a cero hasta el año 1996 que tienden a salirse de la banda superior.

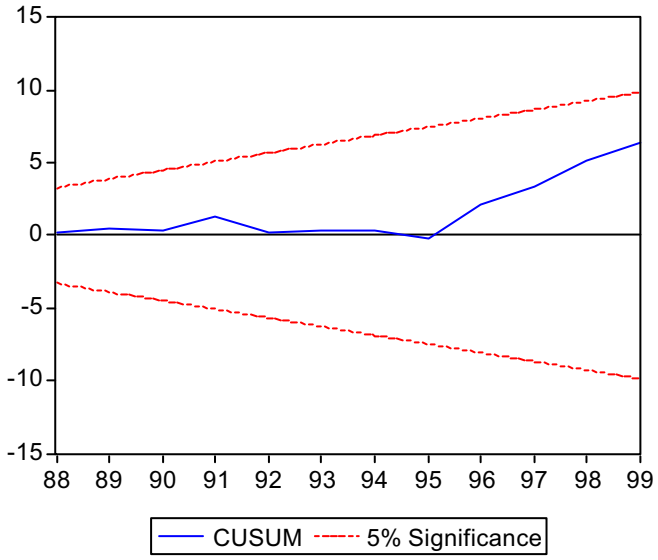
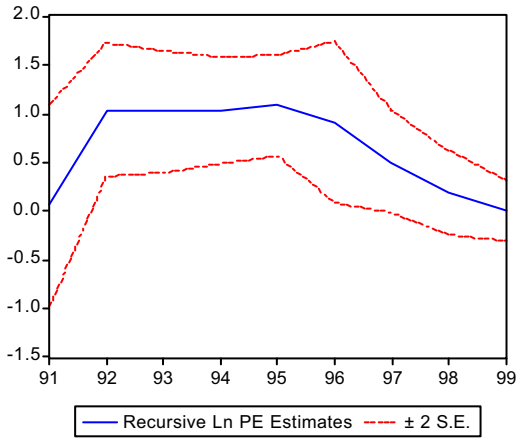
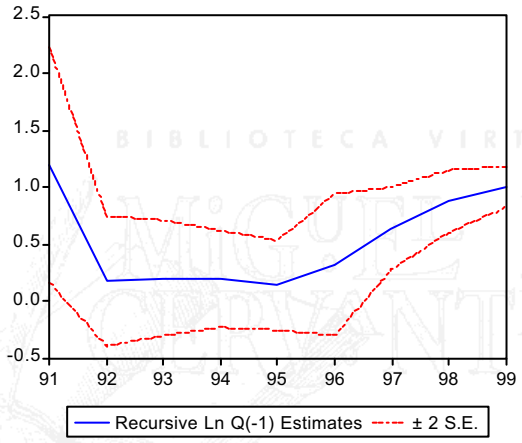


Figura.III.6.16. Suma acumulada de los residuos recursivos. Modelo 4.
Fuente. IEA.
Elaboración. Propia.

Estimación recursiva de los coeficientes

Respecto a la evolución seguida por los coeficientes del modelo, se puede observar en la figura III.6.17 un comportamiento inestable en los primeros periodos de la serie para después tender a la estabilización.





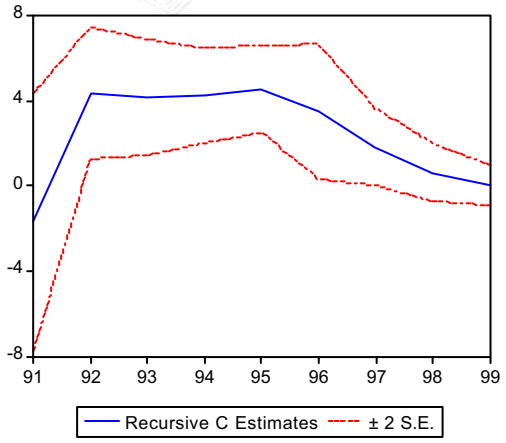


Figura.III.6.17. Gráficos de los coeficientes recursivos. Modelo 4.
Fuente. IEA.
Elaboración. Propia.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

Conclusión parcial de los modelos investigados

Una vez que se han estimado y analizado los cuatro modelos contemplados inicialmente, se llega a la conclusión de que el modelo más completo y con mejores cualidades desde el punto de vista del modelo de regresión clásico es el *modelo 3 productos cruzados*, el cual no considera de forma aislada a la transformación logarítmica del PIB sino el producto de la transformación logarítmica del PIB y la transformación logarítmica del PE.

Pues bien, llegados a este punto cabe preguntarse la cuestión de si el efecto de la variable precio de la electricidad en el año *t* puede afectar a la demanda del año *t*+1, es decir, lo que se plantea a continuación es el carácter de influencia instantánea de la variable económicas, considerada en este modelo (ya planteada por primera vez en el modelo 4). La incorporación de efectos diferidos en el tiempo en un modelo econométrico se realiza mediante la utilización de *variables retardadas* . En el caso estudiado, la expresión funcional más apropiada desde el punto de vista de la demanda residencial en la C.A.A. ha sido el modelo 3PC.

En respuesta a la pregunta planteada en párrafos anteriores, se ha construido un modelo, denominado *modelo 3 retardado* (a partir de ahora denominado Modelo 3R), que desde el punto de vista económico tiene más sentido, ya que es de esperar que un comportamiento futuro en el consumo de electricidad se base en acontecimientos pasados. Además, desde el punto de vista estadístico aporta una mayor explicación acerca de la demanda anual investigada en la C.A.A. Este modelo forma parte de los denominados de retardos distribuidos en los que las variables exógenas, PE y PIB, son las que ejercen un efecto prolongado en el tiempo sobre la variable endógena, Q. El modelo se presenta a continuación:

$$LnQ_i = \alpha + \beta_1 LnPE(-1)_i + \beta_2 LnPIB(-1)_i + u_i$$

En la tabla III.6.43, se realiza la estimación del modelo, siendo todos los coeficientes significativos al 5% y al 1%, así como el modelo considerado en su conjunto. El poder explicativo del modelo y su riesgo de error es mejorado en todos los estadísticos; por ejemplo el coeficiente de determinación ajustado se incrementa en un 0,61% respecto al modelo 3PC.

Tabla. III.6.43.Estimación del modelo 3R de demanda de energía eléctrica en la C.A.A.

Dependent Variable: Q
Method: Least Squares
Sample(adjusted): 1985 1999
Included observations: 15 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PE(-1)	-0.678152	0.160163	-4.234126	0.0012
Ln PIB(-1)	0.875518	0.066209	13.22349	0.0000
C	-3.694585	0.669560	-5.517934	0.0001

R-squared	0.990906	Mean dependent var	8.415830
Adjusted R-squared	0.989390	S.D. dependent var	0.239262
S.E. of regression	0.024645	Akaike info criterion	-4.391597
Sum squared resid	0.007289	Schwarz criterion	-4.249987
Log likelihood	35.93698	F-statistic	653.7442
Durbin-Watson stat	1.525634	Prob(F-statistic)	0.000000

Fuente. IEA.
Elaboración. Propia.

Sin embargo, tal como sucedió con el modelo 3, puede suceder que la relación entre los parámetros y variables no sea lineal, por lo que se aplica el contraste de especificación del modelo, contraste de Ramsey. En la tabla III.6.44 se presenta dicho contraste, en el cual los resultados permiten rechazar la hipótesis de linealidad a un nivel de significación del 5%, pues el valor del estadístico es de 0,016237. Una vez detectada esta no linealidad, la solución final llevaría a determinar cuáles son las potencias y productos cruzados concretos de las variables explicativas que mejor recogen esa no linealidad.

Tabla. III.6.44.Contraste de Ramsey. Modelo 3R.

Ramsey RESET Test:				
F-statistic	5.167980	Probability		0.044050
Log likelihood ratio	5.777062	Probability		0.016237
Test Equation:				
Dependent Variable:Ln Q				
Method: Least Squares				
Sample: 1985 1999				
Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LnPE(-1)	4.230222	2.163527	1.955244	0.0764
LnPIB(-1)	-5.076266	2.618724	-1.938450	0.0786
C	50.00046	23.62671	2.116268	0.0579
FITTED^2	0.393645	0.173159	2.273319	0.0441
R-squared	0.993813	Mean dependent var		8.415830
Adjusted R-squared	0.992125	S.D. dependent var		0.239262
S.E. of regression	0.021232	Akaike info criterion		-4.643402
Sum squared resid	0.004959	Schwarz criterion		-4.454588
Log likelihood	38.82551	F-statistic		588.9295
Durbin-Watson stat	2.163684	Prob(F-statistic)		0.000000

Fuente. IEA.
Elaboración. Propia.

El modelo no lineal que se construye con el objetivo de salvar esta circunstancia es el denominado *modelo 3 productos cruzados retardado* (a partir de ahora denominado Modelo 3PCR), cuya expresión funcional es la siguiente:

$$LnQ_i = \alpha + \beta_1 LnPE(-1)_i + \beta_2 [LnPELnPIB](-1)_i + u_i$$

La estimación del modelo se presenta en la tabla III.6.45, siendo los coeficientes significativos al 5% y al 1%.

El grado de explicación acerca de la demanda doméstica anual en la C.A.A. de este modelo mejora en un 0,9521% al modelo 3PC y reduce el error estándar en un 5,26%.

Tabla.III.6.45. Estimación del modelo 3PCR de demanda de energía eléctrica en la C.A.A.

Dependent Variable: Ln Q
Method: Least Squares
Sample(adjusted): 1985 1999
Included observations: 15 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PE(-1)	-5.834969	0.448684	-13.00464	0.0000
[Ln PELn PIB](-1)	0.339545	0.021084	16.10418	0.0000
C	9.606683	0.294240	32.64919	0.0000

R-squared	0.993737	Mean dependent var	8.415830
Adjusted R-squared	0.992693	S.D. dependent var	0.239262
S.E. of regression	0.020452	Akaike info criterion	-4.764623
Sum squared resid	0.005019	Schwarz criterion	-4.623013
Log likelihood	38.73467	F-statistic	952.0291
Durbin-Watson stat	2.137110	Prob(F-statistic)	0.000000

Fuente. IEA.
Elaboración. Propia.

Análisis de los residuos

En la tabla III.6.46, se representa el correlograma de los residuos con doce retardos, donde se comprueba que existe ausencia de autocorrelación, ya que se observa que los coeficientes de autocorrelación y autocorrelación parcial alcanzan, en la mayoría de los casos, valores próximos a cero. El valor del estadístico *Q*-de *Ljung-Box* es 10,510 con una probabilidad de 0,571, superior al 5% de significación.

Tabla.III.6.46. Correlograma de los residus.Modelo 3PCR.

Sample: 1985 1999
Included observations: 15

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. *| . |	. *| . |	1	-0.091	-0.091	0.1499	0.699
. | . |	. | . |	2	0.023	0.015	0.1603	0.923
. *| . |	. *| . |	3	-0.069	-0.066	0.2617	0.967
****| . |	****| . |	4	-0.461	-0.480	5.1917	0.268
. |* . |	. | . |	5	0.084	-0.015	5.3724	0.372
. **| . |	. **| . |	6	-0.256	-0.312	7.2261	0.300
. | . |	. **| . |	7	-0.018	-0.242	7.2370	0.405
. |** . |	. *| . |	8	0.208	-0.107	8.8174	0.358
. |* . |	. |* . |	9	0.141	0.131	9.6658	0.378
. | . |	. **| . |	10	0.058	-0.277	9.8371	0.455
. *| . |	. **| . |	11	-0.069	-0.229	10.137	0.518
. *| . |	. *| . |	12	-0.066	-0.125	10.510	0.571

Fuente. IEA.
Elaboración. Propia.

Identificación de los residuos al cuadrado

El modelo 3 PCR tampoco es un modelo cuyos errores tengan una estructura de heterocedasticidad autorregresiva ya que el contraste resulta no signifcativo al 95% de confianza.

Tabla.III.6.47. Correlograma de los residuos al cuadrado. Modelo 3PCR.

Sample: 1985 1999
Included observations: 15

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. **| . |	. **| . |	1	-0.295	-0.295	1.5842	0.208
. **| . |	.***| . |	2	-0.277	-0.399	3.0897	0.213
. |***. |	. |** . |	3	0.434	0.263	7.0939	0.069
. | . |	. |* . |	4	-0.056	0.104	7.1678	0.127
.***| . |	. *| . |	5	-0.332	-0.178	9.9814	0.076
. |* . |	. *| . |	6	0.138	-0.185	10.524	0.104
. | . |	. **| . |	7	-0.045	-0.256	10.590	0.158
. *| . |	. *| . |	8	-0.140	-0.073	11.300	0.185
. | . |	. | . |	9	0.041	-0.044	11.372	0.251
. | . |	. *| . |	10	-0.018	-0.085	11.388	0.328
. | . |	. | . |	11	-0.001	-0.041	11.388	0.411
. | . |	. *| . |	12	0.005	-0.125	11.390	0.496

Fuente. IEA.
Elaboración. Propia.

Normalidad de los residuos

En la figura III.6.18 se representa el histograma de probabilidad correspondiente a los residuos del modelo, tomando la media y coeficiente de asimetría un valor muy cercano a cero y el coeficiente de curtosis un valor próximo a tres. Se concluye que, al 95% de confianza, no se puede rechazar la hipótesis de normalidad de los residuos.

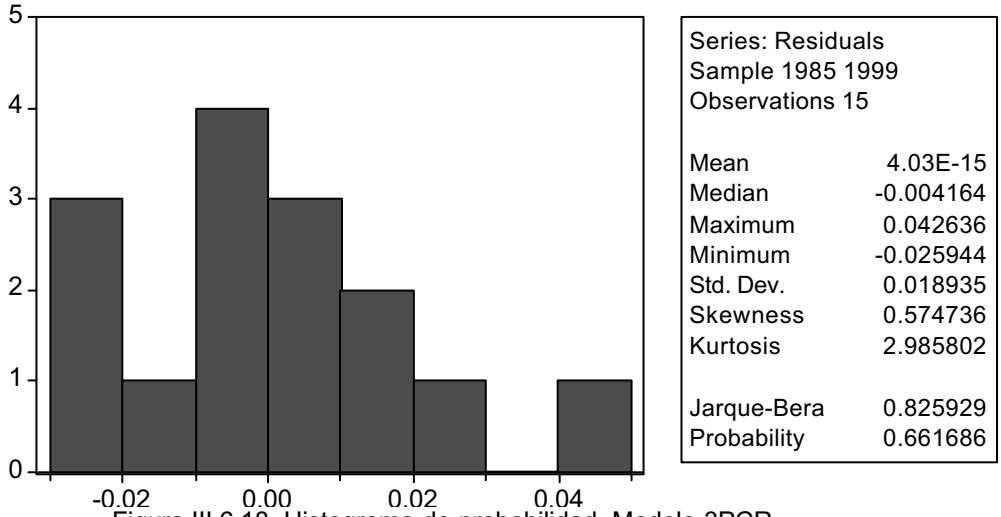


Figura.III.6.18. Histograma de probabilidad. Modelo 3PCR.
Fuente. IEA.
Elaboración. Propia.

BIBLIOTECA VIRTUAL

Correlación serial de los residuos

El contraste de Breusch-Godfrey revela que el valor muestral del estadístico no permite rechazar la hipótesis nula de incorrelación entre las perturbaciones. Además, la regresión auxiliar estimada muestra que el primer y segundo retardo no son individualmente significativos al 95% de confianza.

Tabla. III.6.48.Contraste de autocorrelación seriel. Modelo 3PCR.

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.053365	Probability		0.948302
Obs*R-squared	0.158403	Probability		0.923854
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
PERETARD	0.079990	0.612589	0.130577	0.8987
PEPIBRETARD	-0.003744	0.028694	-0.130473	0.8988
C	-0.051573	0.401142	-0.128564	0.9003
RESID(-1)	-0.117360	0.371113	-0.316238	0.7583
RESID(-2)	-0.003787	0.367783	-0.010296	0.9920
R-squared	0.010560	Mean dependent var		6.99E-15
Adjusted R-squared	-0.385216	S.D. dependent var		0.018935
S.E. of regression	0.022285	Akaike info criterion		-4.508573
Sum squared resid	0.004966	Schwarz criterion		-4.272556
Log likelihood	38.81430	F-statistic		0.026682
Durbin-Watson stat	1.940401	Prob(F-statistic)		0.998374

Fuente. IEA.
Elaboración. Propia

Heterocedasticidad autorregresiva condicional en los residuos

No existe evidencia de la existencia de heterocedastidad autorregresiva condicional al 95% de confianza, tal como se refleja en la tabla III.6.49.

Tabla.III.6.49. Contraste de comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 3PCR.

ARCH Test:				
F-statistic	1.196989	Probability		0.295399
Obs*R-squared	1.269824	Probability		0.259800
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample(adjusted): 1986 1999				
Included observations: 14 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000459	0.000163	2.819529	0.0155
RESID^2(-1)	-0.298216	0.272575	-1.094070	0.2954
R-squared	0.090702	Mean dependent var		0.000355
Adjusted R-squared	0.014927	S.D. dependent var		0.000500
S.E. of regression	0.000496	Akaike info criterion		-12.24901
Sum squared resid	2.95E-06	Schwarz criterion		-12.15772
Log likelihood	87.74308	F-statistic		1.196989
Durbin-Watson stat	2.259577	Prob(F-statistic)		0.295399

Fuente. IEA.
Elaboración. Propia.

Heterocedasticidad en los residuos

Las variables incluidas en la regresión auxiliar no tienen ningún poder explicativo sobre los residuos al cuadrado, y por lo tanto el coeficiente de determinación es pequeño como se observa en la tabla III.6.50. Por ello, el valor muestral del estadístico es suficientemente bajo como para que la probabilidad de rechazar la hipótesis nula siendo cierta sea mayor que 5% por lo que no se rechaza la hipótesis nula de homocedasticidad.

Tabla.III.6.50. Contraste de heterocedasticidad de los residuos. Modelo 3PCR.

White Heteroskedasticity Test:				
F-statistic	0.603408	Probability		0.669037
Obs*R-squared	2.916508	Probability		0.571893
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample: 1985 1999				
Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.296800	0.303851	0.976794	0.3517
LnPERETARD	-0.626555	0.601557	-1.041556	0.3221
[LnPERETARD]^2	0.122347	0.116864	1.046919	0.3198
LnPEPIBRETARD	0.024894	0.023029	1.080979	0.3051
[LnPEPIBRETARD]^2	-0.000306	0.000281	-1.090583	0.3010
R-squared	0.194434	Mean dependent var		0.000335
Adjusted R-squared	-0.127793	S.D. dependent var		0.000488
S.E. of regression	0.000518	Akaike info criterion		-12.03064
Sum squared resid	2.69E-06	Schwarz criterion		-11.79463
Log likelihood	95.22983	F-statistic		0.603408
Durbin-Watson stat	2.803611	Prob(F-statistic)		0.669037

Fuente. IEA.
Elaboración. Propia.

Errores de especificación por variables omitidas

Los resultados del contraste de Ramsey, tabla III.6.51, no permiten rechazar la hipótesis nula de linealidad a un nivel del 5%, pues la probabilidad asociada al valor del estadístico es de 0,463666.

Tabla.III.6.51. Contraste de Ramsey. Modelo 3PCR.

Ramsey RESET Test:				
F-statistic	0.576452	Probability		0.463666
Log likelihood ratio	0.766166	Probability		0.381406
Test Equation:				
Dependent Variable: LnQ				
Method: Least Squares				
Sample: 1985 1999				
Included observations: 15				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PERETARD	-24.46546	24.54247	-0.996862	0.3403
Ln PEPIBRETARD	1.415508	1.417312	0.998727	0.3394
C	26.70955	22.52817	1.185607	0.2608
FITTED^2	-0.183203	0.241296	-0.759244	0.4637
R-squared	0.994049	Mean dependent var		8.415830
Adjusted R-squared	0.992426	S.D. dependent var		0.239262
S.E. of regression	0.020823	Akaike info criterion		-4.682368
Sum squared resid	0.004769	Schwarz criterion		-4.493554
Log likelihood	39.11776	F-statistic		612.4765
Durbin-Watson stat	2.183656	Prob(F-statistic)		0.000000

Fuente. IEA.
Elaboración. Propia.

Estimación recursiva de los residuos

La representación de los residuos recursivos para el modelo 3 retardado cruzado aparece en la figura III.6.19. El resultado refleja que los residuos fluctúan en torno al valor cero en un rango $\pm 0{,}03$, no llegando a sobrepasar en ningún caso las bandas de confianza. Por lo tanto, no existe evidencia de inestabilidad en el modelo.

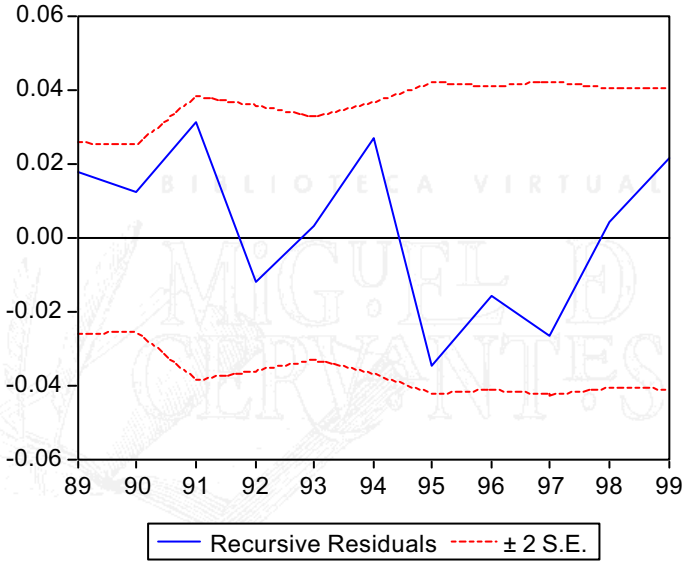


Figura.III.6.19. Residuos recursivos. Modelo 3PCR.
Fuente. IEA.
Elaboración. Propia.

Suma acumulada de los residuos recursivos

El estadístico CUSUM varía en todo su recorrido dentro de las bandas de confianza, con mayores fluctuaciones en los primeros periodos para después estabilizarse en torno a cero. Por lo tanto, se puede afirmar que no existe evidencia de inestabilidad en el modelo planteado.

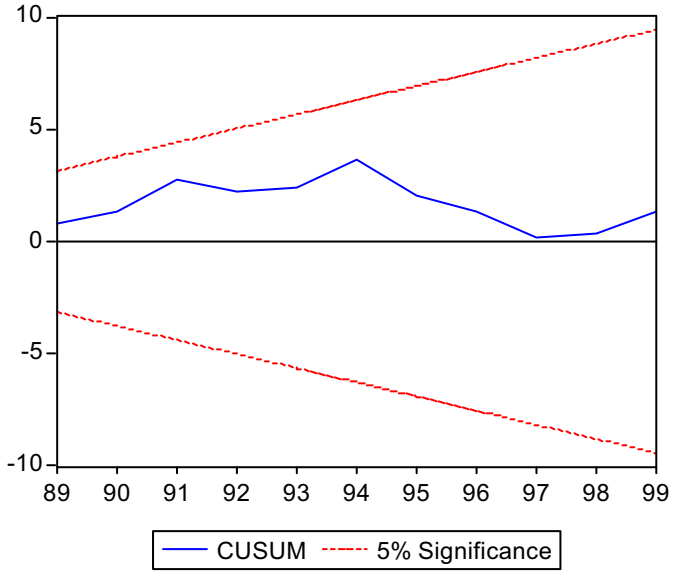


Figura.III.6.20. Suma acumulada de los residuos recursivos. Modelo 3PCR.
Fuente. IEA.
Elaboración. Propia.

Estimación recursiva de los coeficientes

Los coeficientes del modelo analizado, figura III.6.21, se caracterizan por su convergencia hacia la estabilidad a lo largo del periodo analizado.

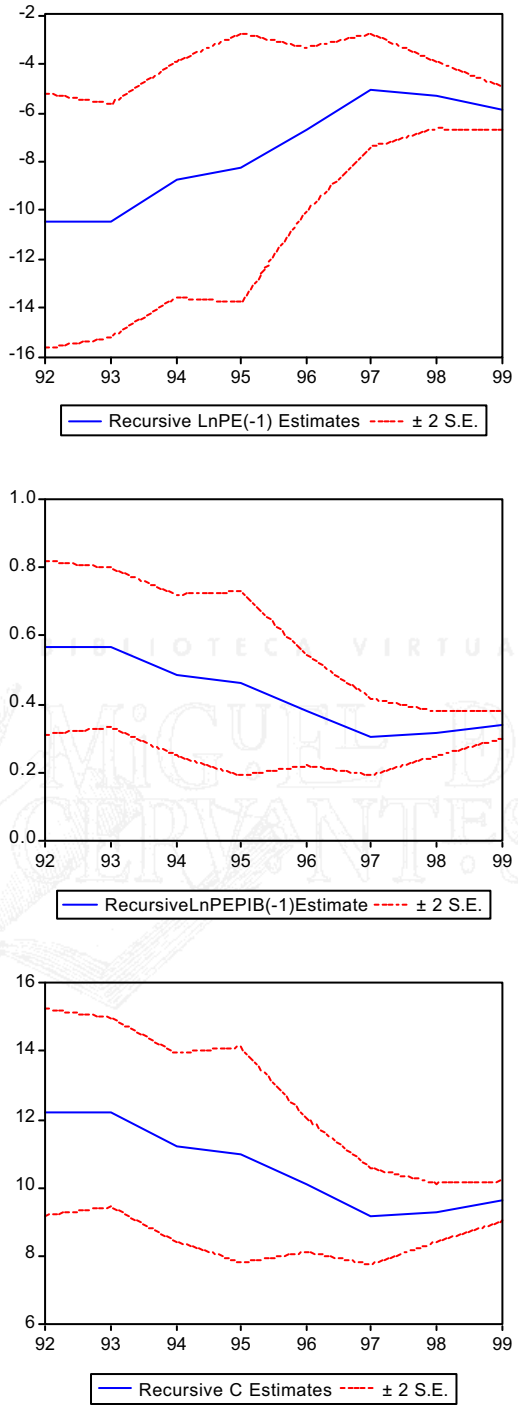


Figura.III.6.21.Gráficos de los coeficientes recursivos. Modelo 3PCR.
Fuente. IEA.
Elaboración. Propia.

Una vez analizados los seis modelos desde el punto de vista de vista de su estructura, se analizan éstos desde el punto de vista de su poder y capacidad predictiva, reflejándose en la tabla III.6.52, junto con otros estadísticos.

Conclusión de los modelos analizados

La conclusión principal que se deduce, de lo anterior es que el modelo más apropiado para estimar la demanda de electricidad en el sector residencial en la C.A.A. es el modelo 3 productos cruzados retardado. Este modelo relaciona la demanda de un periodo con la tarifa de electricidad del año anterior así como con el producto de el precio y el precio interior bruto, ambas retardadas un periodo.

De esta forma este modelo es el que mejor explica la variabilidad total de la demanda de electricidad en la C.A.A., ya que alcanza el valor 0,9927 muy próximo a su cota superior que es 1. Desde el punto de vista de los errores de ajuste al modelo, estos también se hacen mínimos en el modelo propuesto. El estadístico DW es el se encuentra más próximo a 2, por lo que en este modelo se puede afirmar con mayor evidencia, el hecho de ausencia de autocorrelación de primer orden entre los residuos. Finalmente, la capacidad predictiva de este modelo también supera a los de su misma clase, es decir, aquellos modelos que consideran las mismas variables explicativas pero su relación funcional en el modelo es distinta, así como también a los modelos que utilizan otro tipo de variables económicas.

Tabla.III.6.52. Estadísticos evaluadores del los principales modelos analizados.

Estadísticos	Modelo 1	Modelo 2	Modelo 3	Modelo 3 productos cruzados	Modelo 3 productos cruzados retardado	Modelo 4
$\bar{R}^2$	0,841291	0,847865	0,975806	0,983338	0,9927	0,979427
SE	0,100618	0,098512	0,039286	0,032602	0,020452	0,034318
SCE	0,131612	0,126161	0,020064	0,013817	0,005019	0,014133
Log likelihood	15,70087	16,03928	30,74848	33,73230	38,73467	30,97077
DW	0,590694	0,612162	1,458514	1,969532	2,137110	2,467082
Akaike	-1,587608	-1,629910	-3,468560	-3,841537	-4,764623	-3,729436
Schwarz	-1,442748	-1,485050	-3,323700	-3,696677	-4,623013	-3,587826
Forecast Error	0,1006	0,09851	0,03929	0,0326	0,02045	0,03432
MAPE	0,007963	0,007746	0,003599	0,002878	0,001722	0,003041
MAD	0,06767	0,06559	0,02996	0,02415	0,01456	0,02579
BIC	0,1176	0,1152	0,04592	0,03811	0,02398	0,04024
RMSE	0,0907	0,0888	0,03541	0,02939	0,01829	0,03069

Fuente. IEA.
Elaboración. Propia.

En la figura III.6.22 se representa gráficamente el valor que alcanzan cada uno de los estadísticos evaluadores de los principales modelos analizados. El objetivo que se persigue con esta figura es visualizar de una forma más clara estos estadísticos, para lo cual se han realizado transformaciones (cambios de escala) en algunos de ellos. De esta forma, observando la figura, serán mejores modelos los que alcancen menores valores en sus medidas (los estadísticos $\bar{R}^2$ y Log likelihood han sido normalizados). El modelo 3PCR es el que presenta mejor comportamiento explicativo y predictivo, de forma global, en todos los estadísticos.

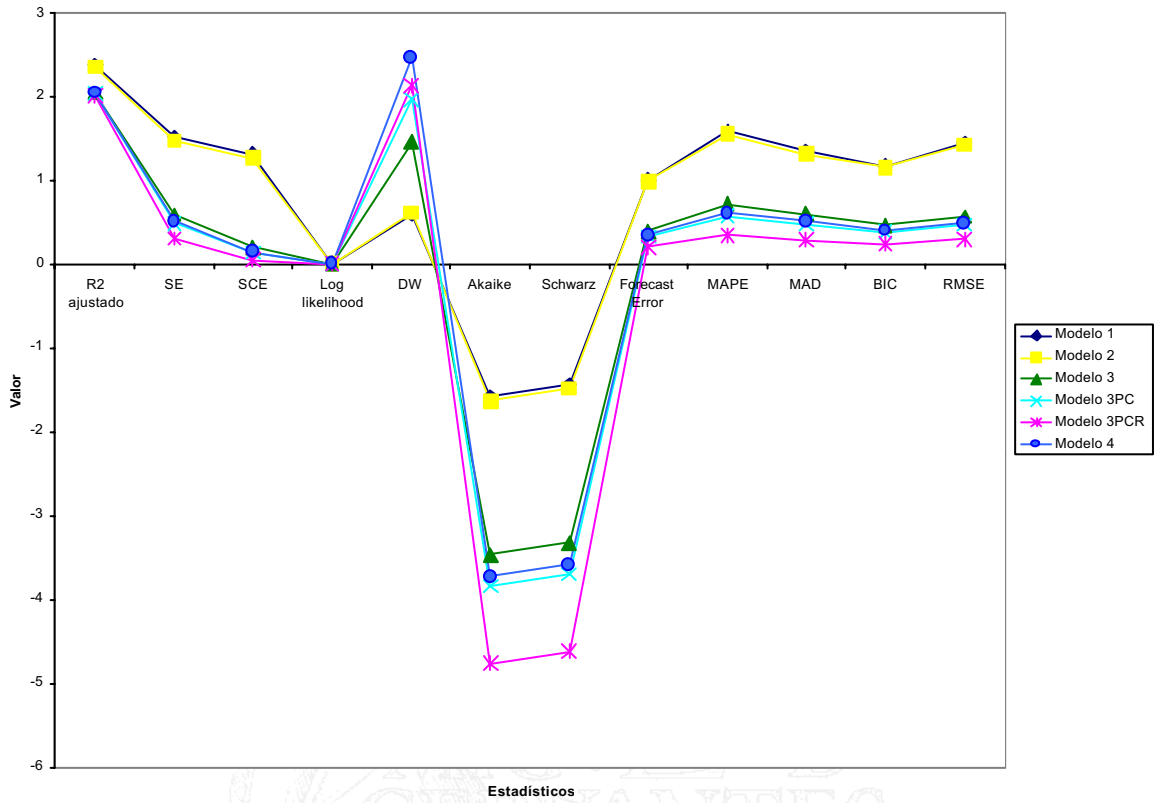


Figura.III.6.22. Comparación de los estadísticos de los modelos considerados.
Fuente IEA.
Elaboración. Propia.

Jerarquización de los modelos según la técnica de decisión Promethee

Para completar el estudio de la elección de los modelos más adecuados desde el punto de vista de la previsión de la demanda doméstica de electricidad en la C.A.A., se ha aplicado la técnica de decisión multicriterio Promethee. Esta técnica establece un orden entre varias alternativas cuando un conjunto de criterios se desean alcanzar (pudiendo encontrarse los mismos en situación de conflicto). En el análisis realizado las alternativas son cada uno de los modelos que se desean evaluar (un total de seis modelos) y los criterios son los estadísticos evaluadores de los modelos, tanto de carácter explicativo como predictivo. El peso asignado a cada criterio es idéntico y el tipo de criterio generalizado es I, el cual considera cualquier diferencia entre los modelos significativa. Los resultados del preorden completo (en función de los valores de los flujos netos) aparecen en la figura III.6.23, donde se observa que el modelo que mejores resultados presenta es el modelo 3PCR seguido del modelo 3PC. Por otra parte, entre los modelos *menos aconsejables* para

explicar y predecir la demanda residencial de energía eléctrica en la C.A.A. se encontrarían el modelo 2 y modelo 1.

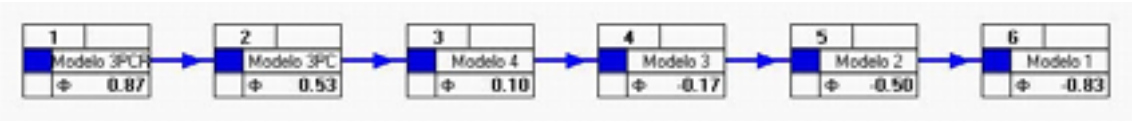


Figura.III.6.23. Preorden total. Método Promethee.
Fuente. IEA.
Elaboración. Propia.

En la tabla III.6.53 se comparan los siguientes estadísticos, el Indice de Theil (*CDT*), Proporción del Sesgo (U_1), Proporción de la Varianza (U_2) y Proporción de la Covarianza (U_3), con el objetivo de medir la capacidad predictiva de los modelos considerados. Se puede observar que todos los índices CDT son muy próximos a cero, lo que refleja la alta capacidad predictiva de todos los modelos considerados, siendo aún mayor en los modelos 3, modelo 3 PC, modelo 3 PCR y modelo 4. Y de entre estos modelos el modelo 3 PCR es el mejor capacidad predictiva alcanza. El estadístico *CDT* se descompone en los estadísticos U_1, U_2 y U_3, siendo la suma de los tres igual a la unidad. De todos los componentes del error de predicción el más difícil de disminuir es la proporción de la covarianza. Los modelos que tienen menor U_3 son los modelos 3, Modelo 3PC, modelo 3PCR y Modelo 4. La representación de las estimaciones realizadas para cada uno de los modelos, así como lo resultados los estadísticos calculados en el estudio de la capacidad predictiva se detallan en el anexo del capítulo, A.III.1-12.

Tabla.III.6.53. Índice de Theil y sus componentes.

Estadísticos	Horizonte temporal de previsión	Modelo 1	Modelo 2	Modelo 3	Modelo 3 productos cruzados	Modelo 3 productos cruzados retardado	Modelo 4
CDT	Previsión 1 periodo (1999)	0,014591	0,013913	0,001551	0,002044	0,001438	0,002858
	Previsión 2 periodos (1998, 1999)	0,016396	0,016141	0,004305	0,004069	0,001805	0,003494
U_1	Previsión 1 periodo (1999)	0,924975	0,892281	0,916447	0,813999	0,573316	0,912650
	Previsión 2 periodos (1998, 1999)	0,935817	0,916861	0,941480	0,898656	0,678730	0,932079
U_2	Previsión 1 periodo (1999)	0,003265	0,000349	0,083553	0,186001	0,426684	0,087350
	Previsión 2 periodos (1998, 1999)	0,001608	0,000092	0,058520	0,101344	0,321270	0,067921
U_3	Previsión 1 periodo (1999)	0,071760	0,107370	0,000000	0,000000	0,000000	0,000000
	Previsión 2 periodos (1998, 1999)	0,062575	0,083047	0,000000	0,000000	0,000000	0,000000

Fuente. IEA.
Elaboración. Propia.

En la tabla III.6.54, se muestran los intervalos de confianza con nivel de confianza del 95%, para cada uno de los modelos analizados y el valor real observado de la demanda de electricidad. Las estimaciones confidenciales se han realizado en primer lugar sólo para el año 1999 y posteriormente para los años 1998-1999. Como se puede observar todos los intervalos construidos para el modelo 3PCR contienen al valor realmente observado de la demanda anual de energía en el sector residencial de la C.A.A., cuyos logaritmos fueron de 8,721 y 8,804 para los años 1998 y 1999, respectivamente. Sin embargo, si se analiza el error cometido al estimar la demanda a través de la estimación puntual, las menores desviaciones son las producidas por el modelo 3PC (15 periodos) y modelo 3PCR (14 periodos). Concretamente cuando se desea realizar una estimación para el último periodo, 1999, el modelo que menor error de predicción produce en valor absoluto es el modelo 3PC. Sin embargo el modelo 3PCR pese a tener una desviación superior ésta es negativa,

significando que la demanda estimada es superior a la demanda real, no provocando dificultades en cuanto a la flexilibildad de suministro eléctrico y sí en cuanto al coste de oportunidad. Es por ello, por lo que es más aconsejable la elección, desde el punto de vista de la calidad del suministro eléctrico en la C.A.A., el modelo de productos cruzados retardado.

Cuando la previsión que se desea realizar es la de dos años, el modelo 3 productos cruzados se estima con 14 periodos históricos y el modelo productos cruzados retardados con 13 periodos. Los resultados de las estimaciones son mucho más robustas, desde el punto de vista estadístico en el modelo 3PCR que en el modelo sin variables retardadas, 3PC, a pesar de contar con un dato menos.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

Tabla.III.6.54. Estimaciones confidenciales, puntuales y errores de previsión.

Modelos	Horizonte temporal de previsión	Límite Inferior(2,5%)	Límite Superior (97,5%)	Previsión (Ln Q)	Valor Real	Error de Predicción
Modelo 1	Previsión 1 periodo (1999)	8,379	8,607	8,493	8,804	0,311
	Previsión 2 periodos (1998, 1999)	8,451 8,394	8,582 8,525	8,516 8,460	8,721 8,804	0,205 0,344
Modelo 2	Previsión 1 periodo (1999)	8,386	8,612	8,499	8,804	0,305
	Previsión 2 periodos (1998, 1999)	8,453 8,382	8,684 8,535	8,530 8,458	8,721 8,804	0,191 0,346
Modelo 3	Previsión 1 periodo (1999)	8,699	8,841	8,770	8,804	0,034
	Previsión 2 periodos (1998, 1999)	8,595 8,641	8,738 8,784	8,666 8,713	8,721 8,804	0,055 0,091
Modelo 3 productos cruzados	Previsión 1 periodo (1999)	8,6998	8,8148	8,7578	8,804	0,0462
	Previsión 2 periodos (1998, 1999)	8,6188 8,6578	8,7348 8,7728	8,6768 8,7158	8,721 8,804	0,0442 0,0882
Modelo 3 productos cruzados retardados	Previsión 1 periodo (1999)	8,829	8,901	8,865	8,804	-0,061
	Previsión 2 periodos (1998, 1999)	8,734 8,796	8,804 8,867	8,769 8,832	8,721 8,804	-0,049 -0,028
Modelo 4	Previsión 1 periodo (1999)	8,686	8,797	8,741	8,804	0,063
	Previsión 2 periodos (1998, 1999)	8,624 8,656	8,732 8,771	8,678 8,714	8,721 8,804	0,043 0,090

Fuente. IEA.
Elaboración. Propia.

Teniendo en cuenta que los valores de las variables explicativas están expresadas a través de sus logaritmos, a la hora de la interpretar la demanda real hay que aplicar la transformación inversa expresándose el resultado de ello en la tabla III.6.55. Se puede observar que el modelos 3PC y 3PCR son los que mejor estiman la demanda de electricidad en el sector residencial de la C.A.A. Sin embargo, si se observan las estimaciones para los años 1998 y 1999, éstas son muy robustas en el modelo de productos cruzados retardado.

Tabla.III.6.55.Estimaciones puntuales y errores de previsión expresado en GWh.

Modelos	Horizonte temporal de previsión	Previsión (GWh)	Valor Real (GWh)	Error de Predicción (GWh)
Modelo 1	Previsión 1 periodo (1999)	4880,485	6660,834	1780,348
	Previsión 2 periodos (1998) (1999)	4994,037 4722,058	6130,306 6660,834	1136,269 1938,776
Modelo 2	Previsión 1 periodo (1999)	4909,856	6660,834	1750,977
	Previsión 2 periodos (1998, 1999)	5064,445 4712,623	6130,306 6660,834	1065,860 1948,211
Modelo 3	Previsión 1 periodo (1999)	6438,172	6660,834	222,662
	Previsión 2 periodos (1998, 1999)	5802,243 6081,459	6130,306 6660,834	328,062 579,375
Modelo 3 productos cruzados	Previsión 1 periodo (1999)	6360,104	6660,834	300,73
	Previsión 2 periodos (1998, 1999)	5865,247 6098,511	6130,306 6660,834	265,054 562,323
Modelo 3 productos cruzados retardado	Previsión 1 periodo (1999)	7079,793	6660,834	-418,958
	Previsión 2 periodos (1998, 1999)	6431,737 6849,973	6130,306 6660,834	-301,431 -189,139
Modelo 4	Previsión 1 periodo (1999)	6254,146	6660,834	406,687
	Previsión 2 periodos (1998, 1999)	5872,290 6087,543	6130,306 6660,834	258,016 573,290

Fuente. IEA.
Elaboración. Propia.

Para finalizar, este apartado del capítulo se sintetiza el conjunto de pruebas y contrastes aplicados a los distintos modelos analizados, tal como se expresa en la tabla III.6.56.

Tabla.III.6.56. Contrastes de diagnóstico aplicados a los modelos analizados.

Tipos de contrastes	Aplicación	Modelo 1	Modelo 2	Modelo 3	Modelo 3 productos cruzados	Modelo 3 productos cruzados retardado	Modelo 4
Coeficientes	¿Variables Irrelevantes?	NO	NO	NO	NO	NO	NO
Residuos	¿Ausencia de autocorrelación?	SÍ	NO	SÍ	SÍ	SÍ	SÍ
	¿Heterocedasticidad autorregresiva primer orden?	NO	NO	NO	NO	NO	NO
	¿Normalidad?	NO	SÍ	SÍ	SÍ	SÍ	SÍ
	¿Correlación serial segundo orden?	SÍ	SÍ	NO	NO	NO	NO
	¿Heterocedasticidad autorregresiva condicional segundo orden?	SÍ	NO	NO	NO	NO	NO
	¿Heterocedasticidad de los residuos?	NO	NO	NO	NO	NO	NO
Estabilidad del modelo	¿Errores de especificación por variables omitidas?	NO	NO	SÍ	NO	NO	NO
	¿Inestabilidad de parámetros?	SÍ	SÍ	-	NO	NO	NO

Fuente. Propia.
Elaboración. Propia.

Este conjunto de pruebas realizadas a los modelos inicialmente considerados, así como a los modelos construidos a medida que los resultados se han ido obtenido, intentan contrastar las principales hipótesis del modelo lineal general. A continuación se expresa la ecuación estimada de los seis modelos analizados en profundidad.

Modelo 1:

$$LnQ_i = 5{,}468783 + 0{,}619925 LnPBKWH_i + 1{,}018614 LnPE_i$$

Modelo 2:

$$LnQ_i = 4{,}867315 + 1{,}144526 LnPGKWH_i + 0{,}781117 LnPE_i$$

Modelo 3:

$$LnQ_i = -2{,}695904 - 0{,}441094 LnPE_i + 0{,}771797 LnPIB_i$$

Modelo 3PC:

$$LnQ_i = 9{,}243230 - 5{,}327008 LnPE_i + 0{,}315378 (LnPELnPIB)_i$$

Modelo 3PCR:

$$LnQ_i = 9{,}606683 - 5{,}834969 LnPE(-1)_i + 0{,}339545 (LnPELnPIB)(-1)_i$$

Modelo 4:

$$LnQ_i = -0{,}968061 + 0{,}317360 LnPIB_i + 0{,}521911 LnQ(-1)_i$$

A lo largo de esta sección se ha comprobado que el modelo cruzado de rezagos distribuidos, modelo 3PCR, es el modelo que resulta más favorable desde el punto de vista de su capacidad explicativa y predictiva. La estimación del modelo se ha realizado a partir de la serie histórica que comprende los años 1984-1999 (15 observaciones). Una vez seleccionado el modelo, modelo 3PCR, y con los datos de las variables explicativas, PE y PIB, correspondiente a los años 2000-2002 se vuelve a realizar la estimación(tabla III.6.57) del modelo basado en 18 observaciones. Los resultados de la nueva estimación siguen favoreciendo la estructura funcional del modelo seleccionado ya que los coeficientes de regresión son significativos y el coeficiente de deteminación ajustado apenas desciende un 5%.

BIBLIOTECA VIRTUAL

Tabla. III.6.57.Estimación del modelo 3PCR. Serie histórica (1984-2002)

Dependent Variable: LnQ
Method: Least Squares
Sample(adjusted): 1985 2002
Included observations: 18 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Ln PE(-1)	-6.778312	0.620608	-10.92205	0.0000
[LnPELnPIB](-1)	0.387957	0.028447	13.63783	0.0000
C	10.05826	0.444781	22.61397	0.0000

R-squared	0.989713	Mean dependent var	8.506065
Adjusted R-squared	0.987008	S.D. dependent var	0.301454
S.E. of regression	0.045710	Akaike info criterion	-3.182007
Sum squared resid	0.031340	Schwarz criterion	-3.033611
Log likelihood	31.63806	F-statistic	362.1988
Durbin-Watson stat	0.993823	Prob(F-statistic)	0.000000

Fuente. IEA.
Elaboración. Propia

En la tabla III.6.58. se realiza una predicción *ex–ante* a medio plazo, para el año 2003, es decir, se estima que la demanda de electricidad por parte del sector doméstico en la C.A.A. para dicho año será de 8090,202 GWh. Pero se conoce que el verdadero valor de la demanda, conocido a principios del año 2004, no coincidirá exactamente con dicha cifra. Para realizar estimaciones más fiables es por lo que se realiza una estimación confidencial, pudiéndose afirmar que al 95% de nivel de confianza la demanda de electricidad en la C.A.A. para el año 2003 está contenido en el intervalo [7454,901GWh, 8779,643GWh].

Tabla. III.6.58. Estimación puntual y confidencial de la demanda de electricidad en la C.A.A. para el año 2003.

Modelos	Horizonte temporal de previsión	Límite Inferior(2,5%)	Límite Superior (97,5%)	Previsión	Valor Real
Modelo 3PCR(LnGWh)	2003	8,916	9,080	8,998	?
Modelo 3PCR(GWh)		7454,901	8779,643	8090,202	

Fuente. IEA.
Elaboración. Propia.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

III. 7. ESTIMACIÓN BOOTSTRAP

III.7.1. Visión general del método bootstrap

El bootstrap es uno de los métodos estadísticos denominados *de computación intensiva* introducido por Efron(1979). La metodología bootstrap permite efectuar inferencias estadísticas sin necesidad de postular previamente que la distribución cumpla ciertas hipótesis que, a veces, son de difícil justificación o contraste.

El exclusivo fin de este breve resumen sobre la metodología *bootstrap* es servir de introducción a las salidas de los resultados programa *BSTP*, el cual ha sido elaborado específicamente para el presente trabajo y algunas de cuyas pantallas de resultados se comentan posteriormente.

Si $X=(X_1, X_2,..., X_n)$ es una muestra aleatoria simple genérica de tamaño n, se designa con $F(x) = Pr(X \leq x)$ a la función de distribución común de las variables aleatorias X_i, lo cual en forma simbólica se escribe $X_1, X_2, ..., X_n \sim F(x)$, o simplemente, $X_i \sim F(x)$. Cuando el valor del parámetro θ de una población es desconocido y, en consecuencia, se desea utilizar un estimador $\hat{\theta} = \hat{\theta}(X_1, X_2,..., X_n)$ del mismo, es importante conocer la precisión de tal estimador, ésta es una de las primeras aplicaciones del bootstrap que más adelante se comentará.

Si en una realización muestral del vector aleatorio $X=(X_1, X_2,..., X_n)$ se observa $X_1=x_1$, $X_2=x_2$,..., $X_n=x_n$, se denominará al vector de componentes $(x_1, x_2,.....,x_n)$ la muestra *original*. Se puede decir que en el bootstrap los datos observados en la muestra *original* $x=(x_1, x_2,...,x_n)$ asumen el papel de la verdadera distribución desconocida $F(x)$, quedando ésta sustituida por su estimación $F_n(x)$, la cual suele ser frecuentemente la distribución empírica de $(x_1,x_2,..,x_n)$ que asigna peso $1/n$ a cada x_i :

$$F_n(x) = \frac{\textit{número de}\,(x_i \le x)}{n} = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le x)$$

siendo *I(A)* la función indicatriz. En adelante, a las funciones de distribución poblacional *F(x)* y $F_n(x)$ se las designará, simplemente, por *F* y F_n, respectivamente.

Si $X^*=(X^*_1, X^*_2, ..., X_n^*)$ es una muestra aleatoria simple genérica de F_n, es decir, cada X^*_i, $1 \le i \le n$, de esta muestra se obtiene independientemente (con *reemplazamiento*) de la muestra *original* $x = (x_1, x_2, ..., x_n)$, de forma que $X_i^* = x_j$, al conjunto $X^*= (X_1^*, X_2^*, .., X_n^*)$ se le denomina una *muestra bootstrap.* La notación $X^{*(b)}$ indica que nos referimos a la *b-ésima* muestra bootstrap obtenida de la muestra *original*, la cual de forma genérica, podemos designar así:

$$X^{*(b)} = (X^{*(b)}_1, X^{*(b)}_2, ..., X^{*(b)}_n)$$

Aunque para obtener el número total de tales posibles muestras bootstrap, n^n, el tiempo requerido de ordenador puede ser considerable, en la práctica no es necesario *extraer* (en realidad, efectuar una *simulación* Monte Carlo en el ordenador) tal número total de muestras ya que, a veces se logra la deseada convergencia cuando se han obtenido aproximadamente 1000 muestras, o incluso menos.

Se puede resumir lo anteriormente expuesto diciendo que el proceso bootstrap, y en particular, la construcción del estimador bootstrap de la desviación típica de un estimador, consta de las etapas siguientes:

1. *Extracción de* una muestra bootstrap $X^*=(X_1^*, X_2^*, ..., X^*_n)$, a partir de la muestra *original* $(x_1, x_2, ..., x_n)$, en la forma descrita anteriormente.
2. *Aplicación* de la *misma* función que define al estadístico de interés, $\hat{\theta}$, a la muestra construida en la etapa precedente, con lo que obtenemos $\hat{\theta}^{*(b)} = \hat{\theta}(X_1^{*(b)}, X_2^{*(b)}, ..., X_n^{*(b)})$
3. *Repetición* de las etapas 1 y 2 precedentes, *B* veces
4. *Construcción* de una distribución de probabilidad a partir de los *B*

valores $\hat{\theta}^{*(b)}$, asignando una frecuencia relativa 1/*B* a cada punto $\hat{\theta}^{*(1)}, \hat{\theta}^{*(2)},..., \hat{\theta}^{*(B)}$. Esta distribución, $G^*(\hat{\theta}^*)$, es el estimador bootstrap de la distribución muestral exacta de $\hat{\theta}$, $G(\hat{\theta})$. En la metodología bootstrap se utiliza $G^*(\hat{\theta}^*)$ para efectuar inferencias sobre $\hat{\theta}$, ya que la distribución exacta $G(\hat{\theta})$ suele ser desconocida.

5. *Construcción* del estimador bootstrap de la desviación típica del estimador $\hat{\theta}(X_1, X_2,..., X_n)$:

$$\hat{\sigma}_{\hat{\theta}}^* = \left[Var(\theta^*)\right]^{1/2} = \left[Var\hat{\theta}(X_1^*, X_2^*,..., X_n^*)\right]^{1/2} = \left\{\frac{\sum_{b=1}^{B}\left[\theta^{*(b)} - \hat{\theta}^*(\bullet)\right]^2}{B-1}\right\}^{1/2}$$

siendo $\hat{\theta}^*(\bullet) = \dfrac{\sum_{b=1}^{B}\hat{\theta}^{*(b)}}{B}$

III.7.2. Tipología de intervalos de confianza

Intervalos de Confianza No Paramétricos Vía Bootstrap

Los intervalos de confianza construidos con el método bootstrap utilizan la distribución empírica F_n y la distribución bootstrap del estadístico $G^*(t) = Pr^*\left[\hat{\theta}^* \le t\right]$. Tanto los intervalos de confianza bootstrap de tipo *standard* como los del tipo *percentil* fueron introducidos por Efron.

Intervalos de Confianza Bootstrap Standard

Un intervalo de confianza bootstrap *standard*, simétrico, de grado de confianza (1-α)100% para el parámetro θ es $\hat{\theta}^* \pm z_{1-\frac{\alpha}{2}}\hat{\sigma}_{\hat{\theta}}^*$, siendo

$\Phi(z_{1-\alpha/2}) = 1 - \dfrac{\alpha}{2}$ y Φ la función de distribución de la normal tipificada.

Intervalos de Confianza Bootstrap Tipo Percentil

El concepto de *probabilidad bootstrap*, $Pr^*(\bullet)$, está íntimamente relacionado con la función de distribución de los estimadores bootstrap de los parámetros. En la simulación Monte Carlo la probabilidad bootstrap, Pr^*, se aproxima (ya que, como se ha dicho, sólo se suelen extraer $B<n^n$ muestras) por la siguiente frecuencia relativa:

$$G_B{}^*(t) = Pr^*\left[\hat{\theta}^* \le t\right] \approx \frac{\textit{número de}\left[\hat{\theta}^{*(b)} \le t\right]}{B}$$

Si se designa por $G^{*-1}(p)$ al percentil de orden p de la función de distribución $G^*(\hat{\theta}^*)$, el *método del intervalo bootstrap tipo percentil* interpreta los percentiles $\frac{\alpha}{2}$ y 1-$\frac{\alpha}{2}$ de esta última función de distribución como los extremos inferior, $\hat{\theta}_l{}^*$ y superior, $\hat{\theta}_s{}^*$ del intervalo de confianza bootstrap simétrico para θ, de grado de confianza 1-α.

$$\hat{\theta}_l = G^{*-1}(\frac{\alpha}{2}) \; ; \qquad \hat{\theta}_S = G^{*-1}(1-\frac{\alpha}{2})$$

Tales extremos definen el intervalo que existe entre los percentiles 100$\frac{\alpha}{2}$ y 100(1-$\frac{\alpha}{2}$) de la distribución bootstrap. Si se ordenan los B valores de $\hat{\theta}_i{}^*$ en sentido ascendente, este intervalo es $(\hat{\theta}_j{}^*, \hat{\theta}_k{}^*)$, siendo $j= B\,\frac{\alpha}{2}$, $k=B\,(1-\frac{\alpha}{2})$.

Intervalo de Confianza Bootstrap Percentil Corregido de Sesgo

Si la distribución bootstrap es sesgada respecto a la mediana, es decir, si $Pr^*\left[\hat{\theta}^* \le \hat{\theta}\right] \ne 0.5$, Efron(1988) sugirió, para corregir este sesgo, un método cuyo objetivo es construir un intervalo de confianza de menor longitud que tuviera en cuenta la asimetría de la distribución de $\hat{\theta}^*$. Ordenando el conjunto de los $\hat{\theta}^*(k)$

en la forma $\hat{\theta}^*(1) \leq \hat{\theta}^*(2) \leq ... \leq \hat{\theta}^*(B)$, se determina el índice j para el que el estimador $\hat{\theta}$ convencional obtenido de la muestra *original* tiene un valor tal que $\hat{\theta}^*(j) \leq \hat{\theta} \leq \hat{\theta}^*(j+1)$.

Lo anterior sirve para calcular p_0 y z_0 en la forma siguiente:

$$\Phi(z_0) = p_0 = Pr\left[\hat{\theta}^* \leq j\right] = \frac{j}{B} \quad ; \quad z_0 = \Phi^{-1}(p_0)$$

A continuación se calculan los valores $p_l = \Phi(2\,z_0 - z_{1-\frac{\alpha}{2}})$ y $p_s = \Phi(2\,z_0 + z_{1-\frac{\alpha}{2}})$ que definen, respectivamente, los índices inferior $l = p_l\,B$ y superior $S = p_S\,B$ que permiten caracterizar, en el conjunto ordenado de los $\hat{\theta}^*(k)$ mencionado anteriormente, los dos valores de dicho conjunto, $\hat{\theta}_l^*$ y $\hat{\theta}_S^*$, que determinan los extremos del intervalo bootstrap de grado de confianza α para el parámetro θ.

III.7.3. Prediccion con bootstrap

La no verificación de las hipótesis del modelo lineal general (por ejemplo, la presencia de heterocedasticidad, no-normalidad y/o de correlación serial en los residuos) origina dificultades en las inferencias de dicho modelo al ser inestables las estimaciones mínimo-cuadráticas de los parámetros, por lo que la utilización de la metodología boostrap puede ser conveniente en estas circunstancias. La aplicación del bootstrap en la regresión también es interesante, entre otros casos, en aquellos en que la teoría matemática sólo proporciona resultados asintóticos o cuando se desea efectuar una predicción con horizonte temporal de varios períodos.

En el modelo general de regresión $Y=(y_1, y_2, ..., y_n)$ designa el vector *nx1* de la respuesta, siendo la matriz *nxk* de regresores $X=(x_1, x_2, ..., x_n)$, donde el vector

kx1, x_t representa a los regresores correspondientes a la *t-ésima* observación. El modelo lineal general se formula:

$$Y = X\beta + \varepsilon \quad \text{o equivalentemente} \quad y_t = x'_t\,\beta + \varepsilon_t \qquad (1)$$

donde ε es el vector *nx1* de errores independientes entre sí, de media nula y varianza σ^2. Se supone que *X* es una matriz de rango completo*, k,* siendo todos los elementos de su primera columna iguales a la unidad, de forma que el modelo contiene una constante en su segundo miembro. Los componentes β_i del vector de dimensión *kx1*, β, son los coeficientes de regresión o parámetros desconocidos. El estimador mínimo-cuadrático de β es:

$$\hat{\beta} = (XX')^{-1}X'Y = \beta + (XX')^{-1}X'\varepsilon \qquad (2)$$

El vector *nx1* de residuos *originales* es:

$$e = \hat{\varepsilon} = Y - X\hat{\beta} = (e_1, e_2, ..., e_n) \qquad (3)$$

El valor pronosticado puntual $\hat{y}_0$ por el modelo mínimo-cuadrático para la observación particular *futura* , $x'_0 = [1, x_{01}, x_{02}, ..., x_{0(k-1)}]$ es:

$$\hat{y}_0 = x'_0\,\hat{\beta}$$

El intervalo de confianza simétrico mínimo-cuadrático de grado de confianza $(1-\alpha)$, denominado *intervalo de predicción*, para dicho punto particular x'_0 es:

$$\hat{y}_0 \mp t_{\frac{\alpha}{2}, n-k} \sqrt{\hat{\sigma}^2\left[1 + x'_0\,(X'X)^{-1}x_0\right]} \qquad (4)$$

donde

$$\hat{\sigma}^2 = \frac{Y'Y - \hat{\beta}'X'Y}{n-k}$$

siendo $t_{\frac{\alpha}{2}, n-k}$ el valor de la abscisa que en la distribución *t* de Student de (*n-k*) grados de libertad deja a su derecha una cola de magnitud $\frac{\alpha}{2}$.

Se supone que se tienen datos históricos (X_t,,Y_t) para *t=1,2,...., n* y que los valores de las variables exógenas X_t para *t=1,2,..., n+m* son conocidas o se pronostican por algún procedimiento ajeno al modelo investigado.

Existen dos métodos para aplicar el *remuestreo* a un modelo regresión, dependiendo de la alternativa a utilizar y de la naturaleza de los regresores. Si los regresores son *fijos*, tal como es el caso que se trata, el muestreo bootstrap preserva esta característica. El algoritmo que produce este comportamiento se basa en los residuos de dicho modelo. Las sucesivas etapas del remuestreo con regresores fijos, en el ciclo *b-ésimo (=1,2 ...,B)* del proceso bootstrap, cuando el horizonte temporal del pronóstico es $t=n+m$, consisten en:

1. Cálculo para $t=1,2,...,n+m$ de las muestras bootstrap $Y_t^{*(b)}$, $t=1,2,...,n+m$, (a la que también se le suele denominar el conjunto *b-ésimo* de datos *artificiales o simulados)*, mediante la adición de $n+m$ residuos bootstrap o *remuestreados,* $e_t^{*(b)}$, al ajuste mínimo-cuadrático clásico descrito anteriormente del modelo de regresión:

$$Y_t^{*(b)} = x'_t\,\hat{\beta} + e_t^{*(b)}$$

   donde para $t=1,2,...,n+m$, cada componente $e_t^{*(b)}$ del vector $(n+m)x1$, $e^{*(b)}=(e_1^{*(b)},..., e_{n+m}^{*(b)})$ es una extracción aleatoria (con reemplazamiento) del conjunto de los n residuos originales de e_1, e_2,...., e_n de la regresión.

2. A continuación se supone que este conjunto de datos artificiales *históricos* $Y_1^{*(b)},..., Y_{n+m}^{*(b)}$ proceden de un modelo tal como el (1) en el que los coeficientes de regresión β son desconocidos y, en consecuencia, se procede a estimarlos por mínimos cuadrados con lo que se obtiene el vector $kx1$, $\beta^{*(b)}$, de coeficientes de regresión correspondientes al *b-ésimo* ciclo simulado del bootstrap.

3. Utilizando $\beta^{*(b)}$ como vector de los coeficientes de regresión se hace uso de la ecuación de regresión similar a (1) para predecir los valores futuros *simulados* $Y_t^{*(b)}$, $t=n+1,n+2,....,n+m$. Puesto que la metodología bootstrap permite calcular tanto los datos futuros *simulados* $Y_t^{*(b)}$, como los

pronósticos *simulados* $\hat{Y}_t^{*(b)}$, se puede calcular el error de predicción simulado $E_t{}^{*(b)} = Y_t^{*(b)} - \hat{Y}_t^{*(b)}$, *t=n+1,n+2,....,n+m* (lo cual no es posible en la Estadística clásica).

4. La repetición de las etapas precedentes 2) y 3) para *b=1,2,.....,B* permite obtener, entre otras distribuciones bootstrap, la del vector $\hat{\beta}^*$ y la de los errores de predicción E^*, por lo que se puede corregir el sesgo de las predicciones obtenidas (lo que tampoco es posible mediante la Estadística clásica). Mediante el programa de ordenador *BSTP*, desarrollado específicamente para el presente trabajo, se obtienen los resultados numéricos de aplicar lo anterior al caso particular de la demanda de energía eléctrica doméstica en la Comunidad Autónoma de Andalucía (datos anuales) correspondiente al período 1984-1999.

III.7.4. Aplicación práctica a través del programa BSTP

Al abrir el escritorio del ordenador aparece el icono ***BSTP,*** nombre del programa elaborado específicamente para este trabajo. Pulsando dicho icono y el botón ***Comenzar*** se da paso a una pantalla en la que en su parte superior izquierda se pueden ver los menús ***Procesos Principales, Utilidades, Formularios, Ayuda y Salir*** de los cuales sólo el primero se suele utilizar en la explotación habitual del programa. Al desplegar dicho menú aparece la opción ***Proceso Bootstrap***, pulsando la cual aparece una nueva pantalla con la cabecera ***Datos Iniciales*** y debajo de ella, en la parte izquierda de la pantalla, hay un cuadro de diálogo que solicita del usuario la información necesaria para ejecutar el programa, mientras que a la derecha de la pantalla hay una lista desplegable titulada ***Tablas*** - para que el usuario elija la base de datos con la que desea trabajar y que en el caso estudiado es ***Datos Reales***- debajo de la cual aparecen las respectivas columnas en las que el usuario debe, en primer lugar, el nombre de las variables dependientes (en nuestro caso, *Q*, que indica la demanda doméstica anual de energía eléctrica residencial en la C.A.A.) y, a continuación el nombre de las variables independientes o regresores (en nuestro caso *PE*, que

designa el precio del kWh de electricidad y *ELNPELNPIB*, que designa al número natural ***e*** elevado al producto de los logaritmos neperianos del mencionado precio del kWh y del *PIB*), figurando en la parte inferior derecha las opciones ***Aceptar*** y ***Cancelar***, a fin de que el usuario proceda en consecuencia. El modelo al que se le aplica la técnica bootstrap es al modelo 3PC, estimado sobre 14 periodos históricos (1984-1997):

$$LnQ_i = 8{,}51 - 4{,}13 LnPE_i + 0{,}258 LnPELnPIB_i$$

El análisis de la varianza del modelo anterior es el que se expresa en la tabla III.7.1(resultados del programa estadístico Minitab).

Tabla.III.7.1. Análisis de la varianza del modelo 3PC.

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	0,62468	0,31234	282,77	0,000
Residual Error	11	0,01215	0,00110		
Total	13	0,63683			

Fuente. IEA.
Elaboración. Propia.

La tabla anterior evidencia (*P-valor*= 0) la significación de las variables explicativas. Al referirse más detalladamente al mencionado cuadro de diálogo, la primera pregunta a la que se enfrenta el usuario es: ***¿El modelo es dinámico?***, a fin de que se informe si se trata de un modelo con variable endógena retardada como variable explicativa, pues en este caso, se reduce en uno el número de variables explicativas a definir en relación con los modelos no dinámicos.

A continuación el usuario debe informar sobre el valor de *N*, el número de observaciones conocidas, que puede ser igual o mayor que el número de observaciones consideradas como históricas, *n*, siendo estas últimas las únicas que se utilizan para estimar los parámetros desconocidos, siendo la finalidad de los datos correspondientes a los períodos que se encuentran entre los de orden (*n+1*) y *N* efectuar la comprobación *ex-post* del modelo construido. Seguidamente el usuario debe introducir el número de predicciones futuras posteriores a *n*, que se designa por la cantidad no negativa *F*, que no tiene más limitación que el número de observaciones de las que se conoce el valor de las variables explicativas

Seguidamente el usuario debe informar sobre ***B*** (***número de ciclos del bootstrap***)*,* que es el número de veces que se va a efectuar las distintas operaciones objeto de un ciclo bootstrap. También debe definir ***B*** (***número de ciclos del bootsrap a visualizar***)*,* pues en la fase de depuración del programa, es muy habitual que se esté interesado en comprobar que cada uno de los ciclos bootstrap opera correctamente, pero una vez que se han verificado exhaustivamente todos las rutinas y procedimientos del programa sería muy tedioso, tanto que todos los ciclos bootstrap ejecutados en los cálculos aparecieran en la pantalla como que imprimiese tales cálculos. Por ello, esta última información mencionada que se solicita del usuario tiene por finalidad que, independientemente del elevado número de ciclos que el programa ejecute**,** *B*, las salidas por impresora de tales ciclos bootstrap puedan ser muy pocas, caso de que deseemos alguna, pues en realidad se trata sólo de cálculos auxiliares previos a la presentación de los resultados finales.

A continuación se debe informar de la magnitud (***1-alfa***) (***grado de confianza para intervalos***) a fin de que quede definido el grado de precisión de tales intervalos: por defecto, los intervalos se calculan al 95%. La próxima pregunta que se hace al usuario es ***¿Logaritmizar variables?*** habiendo utilizado el término ***logaritmizar*** (se considera que no es incorrección, pues en este punto se está en el ámbito de un programa, no en el núcleo del texto de este trabajo), como sinónimo de que se desea utilizar un modelo doblemente logarítmico, cuando éste sea el caso. Por último, este cuadro de diálogo formula la pregunta ***¿Reinicializar antes 1er período bootstrap?*** que tiene por objeto saber si el usuario desea hacer uso de todos los datos históricos actualizando con datos reales las estimaciones bootstrap ya realizadas cuando esto sea posible, o por el contrario, desea que el proceso bootstrap opere libremente sin que los datos reales intervengan en la simulación un vez comenzada ésta, que es la opción más frecuente.

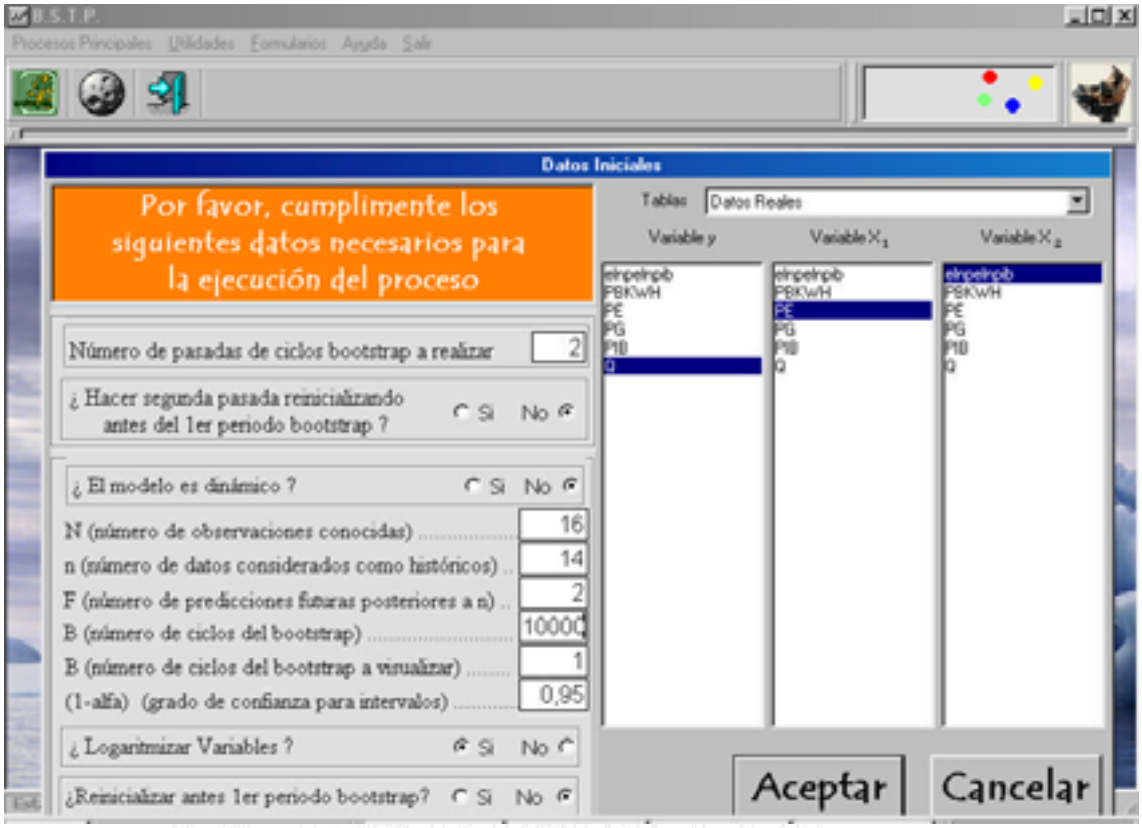


Figura.III.7.1. Cuadro de diálogo inicial. Programa BSTP.
Fuente. IEA.
Elaboración. Propia.

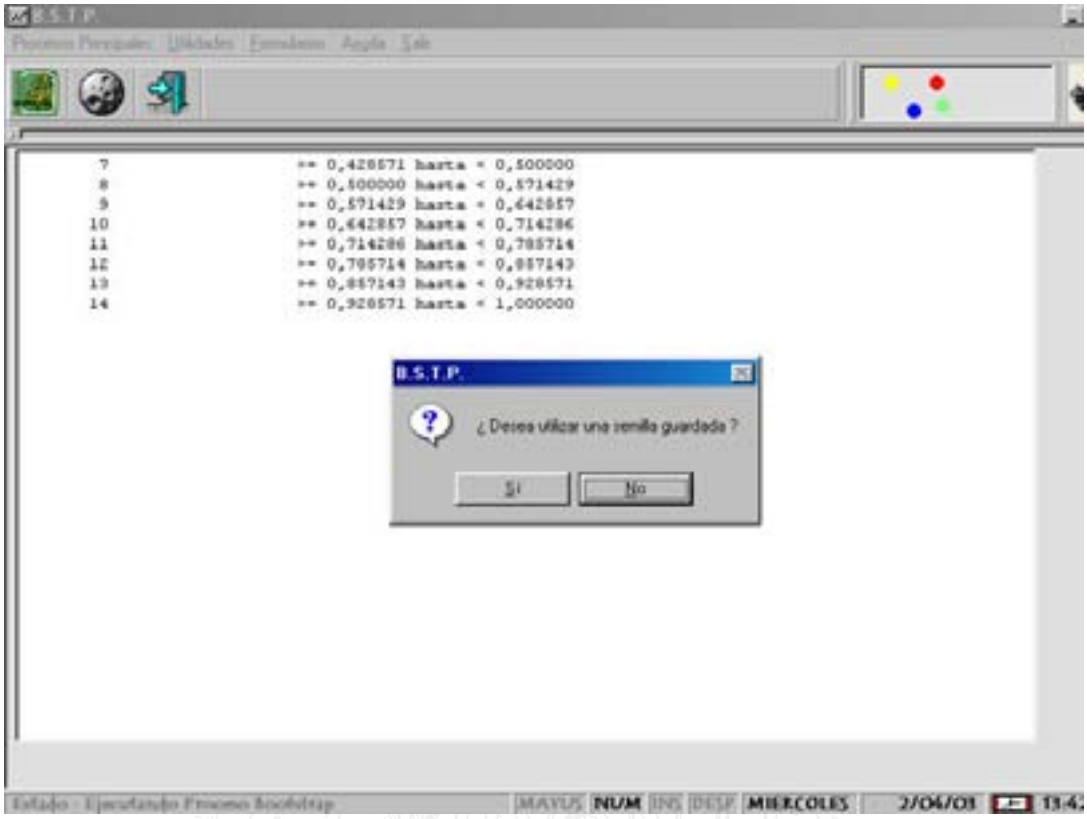


Figura.III.7.2. Opción de semilla.
Fuente. IEA.
Elaboración. Propia.

Figura.III.7.3. Selección de semilla.
Fuente. IEA.
Elaboración. Propia.

La pantalla que se representa en la figura III.7.4 es la correspondiente al *b-ésimo* ciclo del bootstrap, y en ella se presentan diversos valores de la demanda en unidades físicas (GWh) que es como también se muestran en las restantes pantallas del programa *BSTP*. Las clases de demanda que se presentan son, de izquierda a derecha, las siguientes: Q y Q^ designan, respectivamente, la demanda real y la demanda estimada por el modelo econométrico para datos anuales; Q* representa la demanda simulada; Q^*F designa la demanda anual futura simulada; E* representa la diferencia entre las dos columnas precedentes, es decir, es el error simulado de predicción expresado en unidades físicas de demanda (GWh) y, por consiguiente, sólo tiene sentido calcularla para el *futuro* (años 1998 y 1999). Finaliza esta pantalla mostrando los valores de Q^*c, es decir, la demanda futura simulada corregida por el error de predicción, algo que no es posible obtener en la *Estadística clásica*, al desconocer el verdadero valor de los parámetros poblacionales. Como en este punto se han completado los cálculos referentes al mencionado *b-ésimo* ciclo del bootstrap, si *b<B* (número de

ciclos elegido por el usuario al principio del programa BSTP) se vuelven a efectuar los cálculos correspondientes al (*b+1*)-*ésimo* ciclo del bootstrap y así se procede repetidamente hasta que se han efectuado los cálculos que se refieren al *B-ésimo* ciclo del bootstrap, momento en que comienza una nueva fase del programa *BSTP,* cuyo objetivo es sintetizar los cálculos realizados anteriores a fin de obtener las oportunas conclusiones.

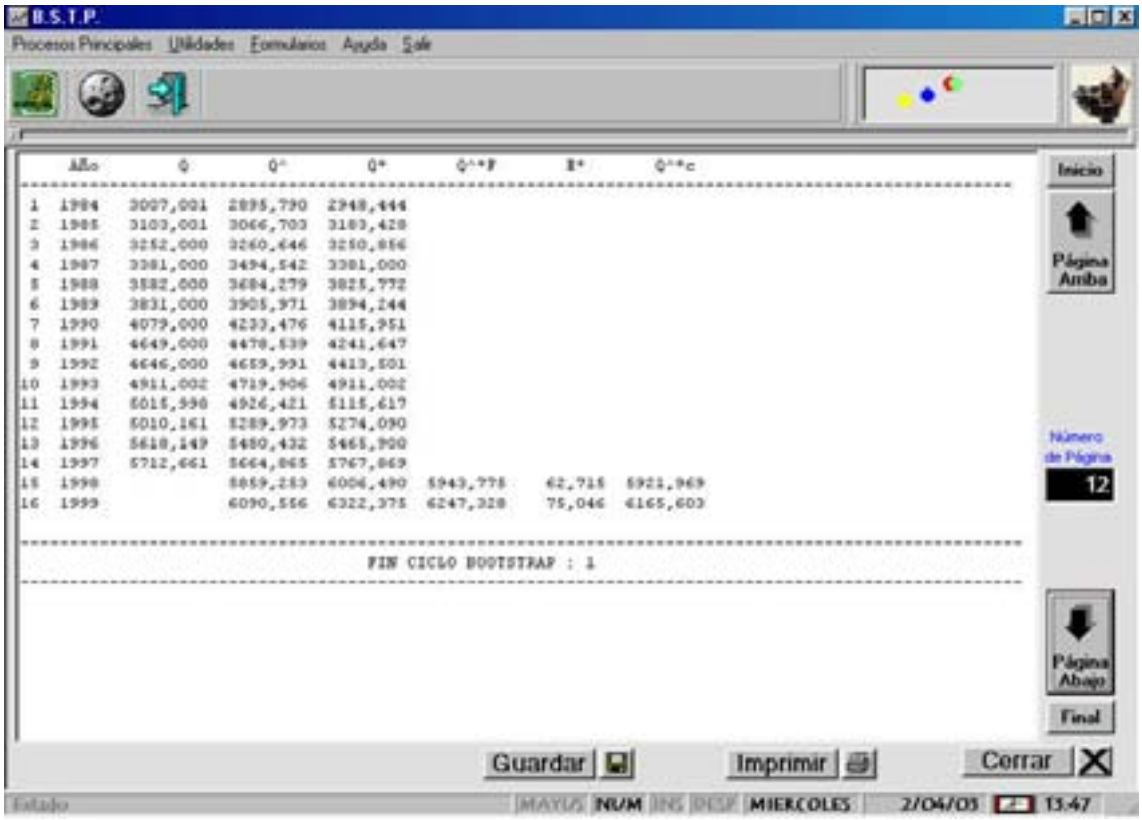


Figura.III.7.4. Cálculos de Q previos al comienzo de la Tabla X.
Fuente. IEA.
Elaboración. Propia.

El objeto de esta tabla, figura III.7.5 (tabla X), es comparar los valores de los estimadores de los parámetros calculados mediante la regresión mínimo-cuadrática y los procedentes de la simulación bootstrap. Las correspondientes cabeceras indican, en su primera línea en qué ámbito se trabaja (mínimos-cuadrados o bootstrap), siguiéndole una línea en la que se numeran las columnas de esta tabla. A continuación se pueden leer distintas cabeceras que de izquierda a derecha, se refieren al nombre del estimador del correspondiente parámetro, su estimación puntual (columna 1) y su desviación típica (columna 2), columnas estas dos últimas que pertenecen al ámbito de los mínimos cuadrados; las

restantes columnas se refieren al universo del bootstrap correspondiendo las dos primeras (columnas 3 y 4), respectivamente, al cálculo de la media y la desviación típica bootstrap calculadas mediante las correspondientes expresiones que figuran en el breve resumen introductorio sobre la metodología bootstrap con que se inicia esta sección.

Las columnas (5), (6) y (7) de esta tabla se refieren, respectivamente, a la raíz cuadrática media de las desviaciones típicas, la desviación típica de las medias bootstrap y a la significación estadística de la estimación mínimo-cuadrática respecto al bootstrap, siendo el objetivo de esto último evaluar la *robustez* del método tradicional frente a sus eventuales desviaciones respecto a las hipótesis que debería cumplir el modelo tradicional, desviación que tiene lugar en los datos que se tratan, a pesar de que el número de datos que se han podido conseguir es reducido (16, de los cuales sólo 14 se han utilizado en la construcción del modelo econométrico) el *test de Kolmogorov-Smirnov-Lilliefors* rechaza la hipótesis de normalidad de los datos de demanda anual, lo cual origina incertidumbre sobre la estabilidad de los valores de los estimadores calculados lo que, en principio, podría cuestionar la validez del modelo.

Si a_{ij} designa la celda de la fila *i-ésima* (i=0,1,2) de la columna *j-ésima* (j=1,2,3,4,5,6,7) de la figura III.7.5, el valor de las celdas de las columnas (5), (6) y (7) se calcula, respectivamente, mediante las expresiones:

$$a_{i5} = RMS_{\hat{\beta}_i^*} = \left[\frac{\sum_{b=1}^{B} \sigma_{\hat{\beta}_i^*}^{2(b)}}{B} \right]^{1/2}$$

$$a_{i6} = \frac{a_{i4}}{\sqrt{B}}$$

$$a_{i7} = \frac{a_{i3} - a_{i1}}{a_{i6}}$$

siendo *B* el número de ciclos bootstrap que el usuario eligió realizar al principio del programa *BSTP* y siendo $\sigma^{2(b)}_{\hat{\beta}_i}$ la varianza convencional (no la varianza bootstrap) del estimador $\hat{\beta}_i$ obtenido mediante el *b-ésimo* conjunto de datos artificiales, es decir, simulados por el bootstrap.

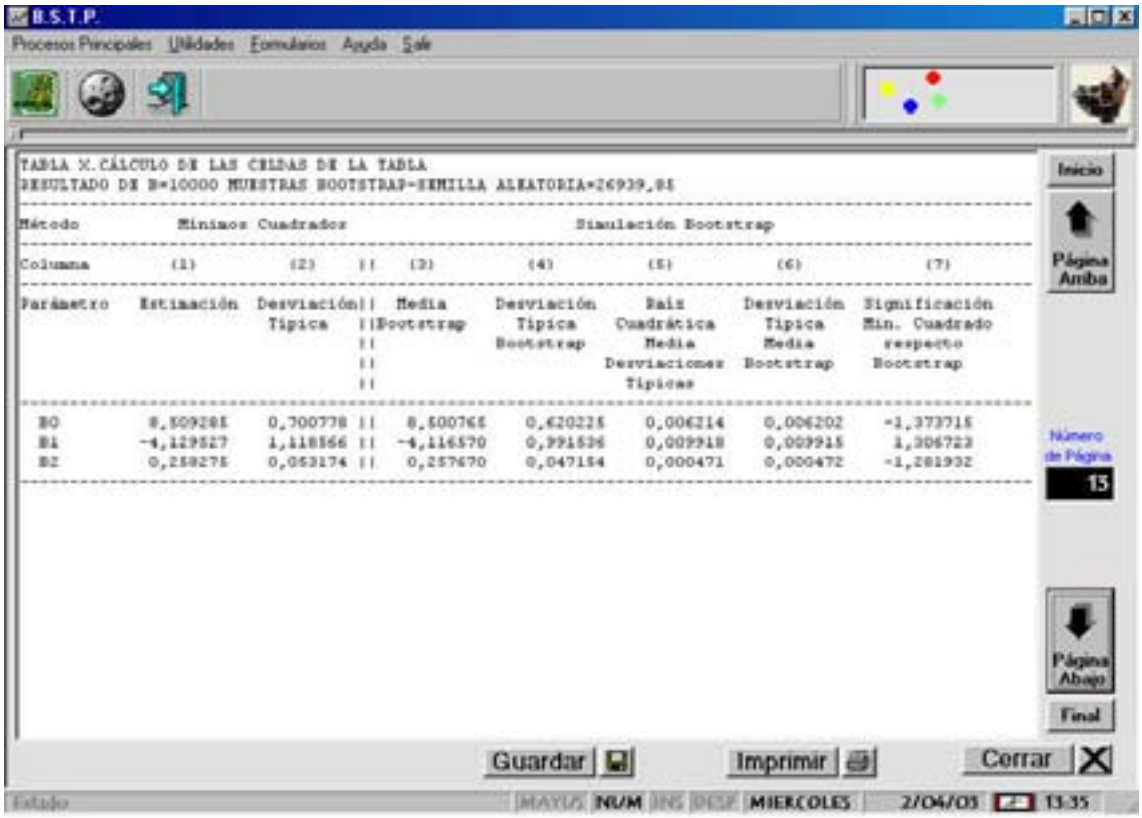


Figura.III.7.5. Cálculo de las celdas de la Tabla X.
Fuente. IEA.
Elaboración. Propia.

Las principales conclusiones que se pueden obtener de la presente tabla (realizada con dos semillas distintas de valores 26939,85 y 44237,81, cada una de las cuales se han utilizado para ejecutar el programa *BSTP* con *B*=10.000 ciclos bootstrap) son las siguientes:

1. Las estimaciones puntuales de los parámetros β_0, β_1 y β_2 calculadas mediante las metodologías tradicional (columna 1) y bootstrap (columna 3), respectivamente, son muy parecidas por lo que no existe evidencia de sesgo debido al habitualmente denominado efecto de las *muestras pequeñas* en los estimadores mínimo-cuadráticos. Esto queda corroborado

mediante los valores de las celdas de la columna (7), que en todos los casos muestran valores no significativos ($-2<t<2$) al 95%. Lo anterior indica que los estimadores puntuales mínimo-cuadráticos son satisfactorios en cuanto a su localización, a pesar del comentado reducido número de datos

2. El método mínimo-cuadrático sobreestima la variabilidad de los estimadores de los parámetros β_i ($i= 0, 1, 2$), ya que la desviación típica de este método es aproximadamente un 11% superior a la proporcionada por la metodología bootstrap, lo cual puede tener una ligera influencia en la significación de los parámetros calculados por los programas de ordenador habituales, ya que aumenta la probabilidad de que no sean declaradas significativas variables que en realidad sí lo son (*error tipo II*). Tales programas suponen la normalidad de las perturbaciones (y por tanto de las observaciones), hipótesis que, en nuestro caso y, con las reservas ya comentadas relativas al escaso número de datos que ha sido posible obtener (16, de los cuales sólo 14 se han utilizado en la construcción del modelo econométrico) resulta rechazada por el *test de Kolmogorov-Smirnov-Lilliefors*. Lo anterior sugiere que la variabilidad de los estimadores mínimo-cuadráticos no es tan satisfactoria como su localización, (comentada en la conclusión precedente) en lo cual puede influir el reducido número de datos.

3. Las estimaciones bootstrap obtenidas como valor de los parámetros β_i ($i=0, 1, 2$) son muy *precisas* (distribuciones marcadamente *leptocúrticas*, muy concentradas alrededor de la media) a juzgar por los valores tan pequeños que se observan tanto en la columna (5) como en la (6) las cuales, pese a medir la variabilidad de la distribución bootstrap con algoritmos diferentes, presentan valores muy próximos entre sí. De hecho también son muy pequeños los valores de los coeficientes de variación de las medias bootstrap.

En esta figura III.7.6 se muestra la tabla *S* cuya finalidad es comparar las predicciones mínimo-cuadráticas y las realizadas mediante la metodología

bootstrap, respecto a la cual en esta tabla se muestran los resultados de los tres tipos de intervalos de esta metodología (estandarizado, percentil y percentil *corregido por sesgo*) que ya fueron comentados en el breve resumen introductorio sobre la metodología bootstrap con que se inició la presente sección.

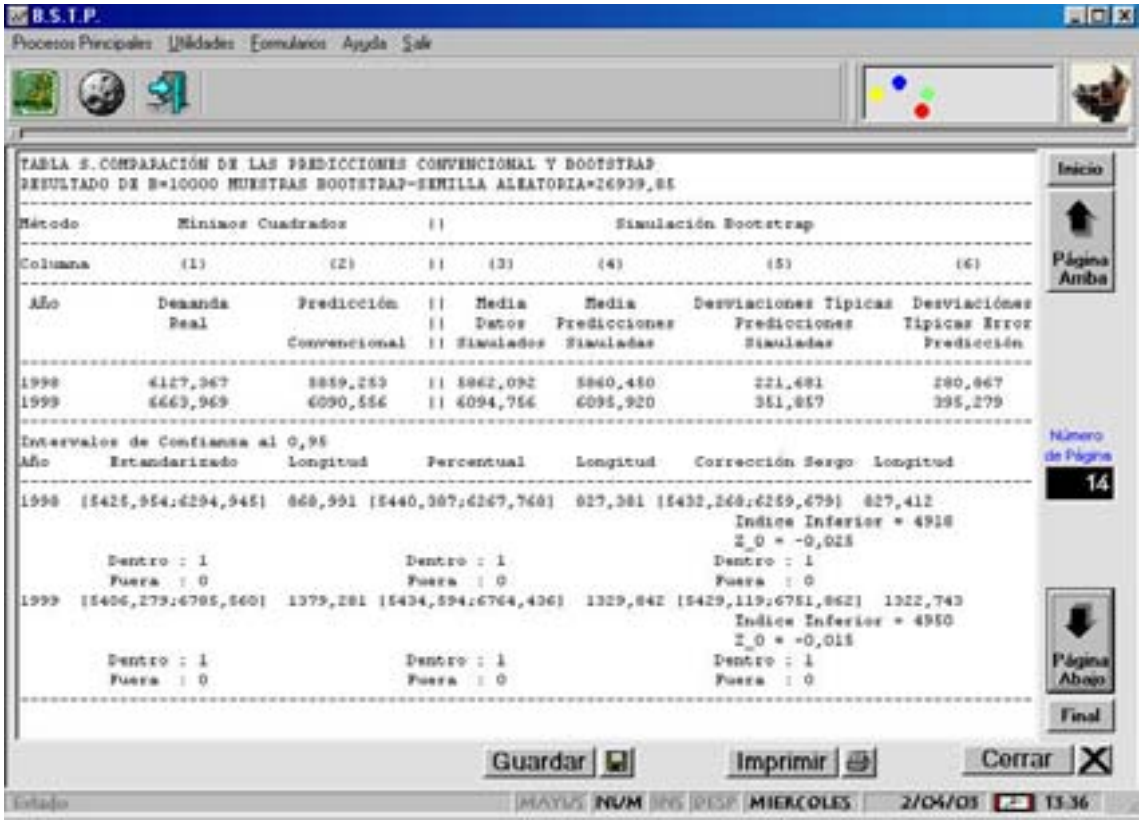


Figura.III.7.6. Comparación de las predicciones convencional y bootstrap. Tabla *S*.
Fuente. IEA.
Elaboración. Propia.

La cabecera inicial de esta tabla *S* indica a qué entorno (mínimos-cuadrados o bootstrap) se refieren los correspondientes resultados que a continuación se comentan. La columna situada más a la izquierda relaciona los años del *futuro* (1998 y 1999) de los que se va a pronosticar su demanda anual de energía eléctrica residencial en la C.A.A. Las columnas (1) y (2) se refieren respectivamente, a la demanda real conocida se evalúan las previsiones realizadas mediante predicción *ex-post* y a la pronosticada por mínimos cuadrados en cada uno de los años *futuros* 1998 y 1999.

La información restante de esta pantalla se refiere al universo simulado del bootstrap. Las columnas (3) y (4) muestran, para los años *futuros* (1998 y 1999), respectivamente, la media de la demanda anual futura simulada por el bootstrap , Q*, y la media de las predicciones futuras simuladas, Q^*F ; la diferencia entre estas dos últimas columnas es el error de predicción, E*, cuya desviación típica figura en la columna (6), mientras que la columna (5) presenta la media de las desviaciones típicas de los correspondientes valores cuyas medias figuran en la columna (4). Todas estas magnitudes están expresadas en (GWh).

Las columnas que siguen a la (6), situadas en una zona de esta pantalla que se encuentra debajo de la ya comentada, se refieren a los tres tipos de intervalos de la metodología bootstrap calculados al nivel de confianza del 95%, ya que éste fue el nivel de confianza elegido por el usuario al principio del programa *BSTP*. De cada uno de estos intervalos se presentan sus extremos inferior y superior y seguidamente la longitud o amplitud de los correspondientes intervalos a fin de obtener conclusiones sobre la precisión de los mismos. A continuación se presentan en esta pantalla dos resultados auxiliares en la construcción de intervalos de confianza *corregidos por sesgo* comentados anteriormente en el breve resumen introductorio sobre la metodología bootstrap con que se inicia esta sección: tales resultados auxiliares son interesantes porque ,en cierto modo, proporcionan una idea rápida de la asimetría de la distribución bootstrap.

Los resultados experimentalmente obtenidos con diez semillas diferentes, cada una de las cuales se utiliza con *B*=10.000 ciclos bootstrap, muestran que la mayor longitud media corresponde, en cada uno de los ciclos bootstrap, incluso individualmente considerados, al intervalo bootstrap tipo estandarizado, lo cual no resulta extraño porque es el único tipo de intervalo que, pese a pertenecer a la metodología bootstrap, todavía presenta cierta *vinculación* con la hipótesis (rechazada por el test de Kolmogorov-Smirnov-Lilliefors) de normalidad de los datos. Tal amplitud resulta exceder aproximadamente en un 4,5% y en un 3,75%, respectivamente para los años *futuros* 1998 y 1999, a las longitudes medias de los intervalos de confianza de la modalidades percentil y percentil *corregido por sesgo,* no existiendo prácticamente diferencia entre las amplitudes de estos dos últimos tipos de intervalos de confianza, debido a que el ya comentado carácter marcadamente leptocúrtico de la distribución bootstrap deja, en realidad, poco margen para la corrección de asimetría que persigue el último tipo de intervalo de confianza mencionado.

Los resultados experimentales citados arrojan para los pronósticos realizados mediante la variable Q^*F correspondientes a 1998, respecto a los intervalos tipos estandarizado, percentil y percentil *corregido por sesgo*, respectivamente, amplitudes medias de 865,40, 831,6 y 832,17 GWh. Para 1999 tales amplitudes medias son, respectivamente, 1379, 1329 y 1328 GWh, lógicamente considerablemente superiores a las de 1998 porque, al ser la *base informativa* idéntica para ambas predicciones, a mayor horizonte temporal de la predicción, mayor incertidumbre, es decir, mayor variabilidad de la predicción y, por tanto, intervalos de confianza más abiertos o de mayor amplitud.

Finalmente se presenta información en esta tabla acerca de si cada modalidad de intervalo bootstrap considerada *cubre*, en la predicción *ex-post*, al valor realmente observado de la demanda (es decir, si *incluye,* o deja *dentro* de sus extremos) al mencionado valor o si, por el contrario, *no lo cubre* (o sea, lo *excluye,* lo deja *fuera* de sus extremos). A fin de neutralizar el efecto que en esta cobertura pudiera tener la elección que el ordenador realiza de la *semilla* de inicialización del algoritmo de generación de números aleatorios, la repetida ejecución del programa *BSTP* con diez semillas independientes entre sí, cada una

de ellas utilizada con *B*=10.000 ciclos bootstrap, permite definir un *índice de cobertura* (porcentaje de intervalos que *cubren* al valor realmente observado) para cada uno de los tres tipos de intervalo de confianza bootstrap. Con ninguna de las diez semillas diferentes utilizadas en la realización experimental se detectó ningún caso de *no cobertura* del valor realmente observado a pesar de que, al nivel de confianza del 95%, cabe esperar que un 5% de los intervalos de confianza calculados dejen fuera a tal valor.

Conclusiones de la tabla S

1. Las predicciones puntuales convencionales del modelo econométrico para 1998 y 1999, que se muestran en la columna (2) de la presente tabla son, respectivamente, 5859,253 y 6090,556 (GWh) que prácticamente coinciden tanto con la medias de los datos simulados, como con la media de las predicciones puntuales simuladas por el bootstrap para dichos años las cuales figuran, respectivamente, en las columnas (3) y (4) de esta tabla. Se puede concluir que, en el aspecto predictivo *puntual,* el comportamiento de los mínimos cuadrados es muy robusto frente a eventuales desviaciones de las hipótesis básicas del modelo econométrico, teniendo en cuenta el reducido número de datos históricos del que ha sido posible disponer.

2. Los intervalos de confianza al 95% de las predicciones convencionales (mínimo-cuadráticas) para los años *futuros* 1998 y 1999 son, respectivamente, [5199,1478 ; 6595,415] y [5180,3609; 7143,0133] (GWh), con longitudes respectivas de 1396,2672 y 1962,6524 (GWh) obtenidos aplicando la fórmula (4) anterior a los valores de los regresores del modelo 3PC en dichos años. Todas estas cantidades difieren considerablemente de las correspondientes obtenidas para los intervalos de confianza de la metodología bootstrap. Así, para el primer año del *futuro*,1998, los tipos de intervalos de confianza estandarizado, porcentual y porcentual *corregido por sesgo* del bootstrap resultan para la semilla 26939,85 ser [5425,954; 6294,945], [5440,387; 6267,768] y [5432,268; 6259,679] cuyas respectivas longitudes son 868,991, 827,381 y 827,412. Para el segundo

año del *futuro*,1999, y con la misma semilla, los correspondientes intervalos de confianza bootstrap son [5406,279; 6785,560], [5434,594; 6764,436] y [5429,119; 6751,862] cuyas respectivas longitudes son 1379,281, 1329,842 y 1322,743. Por consiguiente, para 1998, las amplitudes de los intervalos bootstrap mencionadas anteriormente son aproximadamente un 37,76% inferiores a las calculadas por mínimos cuadrados en el caso del intervalo estandarizado y de un 40,74% para los tipos percentil y percentil *corregido por sesgo* (recuérdese que la amplitud de estas dos últimas modalidades es prácticamente la misma) Para el segundo año del *futuro*, 1999, la menor amplitud de los intervalos bootstrap respecto a los mínimo cuadráticos es del orden del 29,72% para los intervalos estandarizados, del 32,24% para los tipos percentil y del 32,60% para el intervalo del percentil *corregido por sesgo*. Al variar la semilla de inicialización del algoritmo de generación de números aleatorios las cantidades obtenidas son muy similares a las ya descritas.

Por tanto, experimentalmente resulta que aunque los intervalos de confianza obtenidos con la metodología convencional y boostrap satisfacen el objetivo de *cobertura* de los valores realmente observados de la demanda (la predicción *ex*-post ratifica a los modelos construidos) los intervalos confidenciales bootstrap son considerablemente más precisos que los convencionales.

La figura II.7.7 representa la tabla *W*. La finalidad de esta pantalla también es comparar las predicciones mínimo-cuadráticas y las realizadas mediante la metodología bootstrap, pero a diferencia con la pantalla anterior, la presente se refiere a la magnitud Q^*c, es decir, a la predicción de la demanda anual futura simulada *corregida por el error de predicción.*

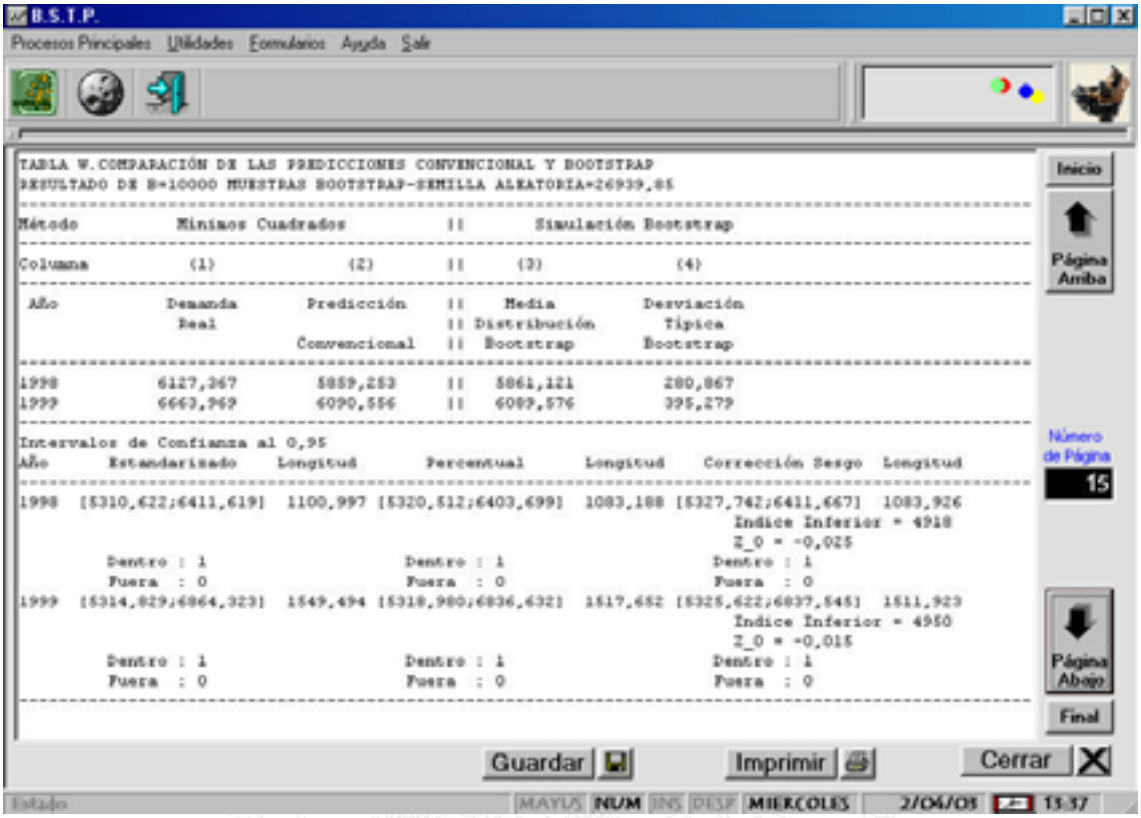


Figura.III.7.7. Comparación de las predicciones convencional y bootstrap. Tabla *W*.
Fuente. IEA.
Elaboración. Propia.

Esta pantalla referente a la tabla *W* proporciona, repitiendo, para mayor comodidad del usuario del programa *BSTP,* en sus tres primeras columnas situadas a la izquierda de la pantalla, idéntica información a la ya proporcionada en las columnas que tienen la correspondiente numeración en la *tabla precedente S* y, por tanto, dicha información se refiere al ámbito de los mínimos cuadrados. Las columnas (3) y (4) presentan, respectivamente, la media y la desviación típica de Q^*c. Las columnas que siguen a la (4), situadas en una zona de esta pantalla que se encuentra debajo de la que se acaba de comentar, se refieren a los tres tipos de intervalos bootstrap a los que se ha hecho referencia en la pantalla anterior.

Los resultados experimentales obtenidos al ejecutar el programa *BSTP* con las diez semillas independientes entre sí mencionadas anteriormente, cada una de las cuales se utiliza con *B*=10.000 ciclos bootstrap muestran que, para la variable predictiva investigada en la presente pantalla, Q^*c, la mayor amplitud

media también corresponde, en cada uno de los ciclos bootstrap, incluso individualmente considerados, al intervalo bootstrap tipo estandarizado. Tal amplitud excede aproximadamente en un 2,3 % y en un 2,1% a las longitudes medias correspondientes, respectivamente, (para los años *futuros* 1998 y 1999) de los intervalos de confianza de la modalidades percentil y percentil *corregido por sesgo*, no existiendo tampoco prácticamente diferencia entre las amplitudes de estos dos últimos tipos de intervalos de confianza bootstrap por los motivos ya apuntados.

Los resultados experimentales citados arrojan para los pronósticos realizados mediante la variable Q^*c correspondientes a 1998, respecto a los intervalos bootstrap tipos estandarizado, percentil y percentil *corregido por sesgo*, respectivamente, amplitudes medias de 1092,3, 1068,4 y 1066,9 (expresado en GWh). Para 1999 tales amplitudes medias son, respectivamente, 1514,6, 1509,1 y 1514,6 GWh, también mayores que los correspondientes a 1998 por las razones ya comentadas. El resto de la información contenida en la presente pantalla es análoga a la ya mostrada en la pantalla precedente, con la salvedad ya comentada, de que esta *pantalla W* se refiere a la magnitud Q^*c, no habiéndose detectado tampoco, con ninguna de las diez semillas diferentes utilizadas en la realización experimental ningún caso de *no cobertura* del valor realmente observados pese a la escasez del número de datos comentada anteriormente.

Se puede observar que los valores que figuran en las celdas de la columna (4) de esta tabla son iguales a los que aparecen en la columna (6) de la tabla precedente, pese a tratarse de la variabilidad de variables aleatorias cuyos órdenes de magnitud son muy diferentes. Tal igualdad se ha de cumplir necesariamente, puesto que la predicción bootstrap de la demanda corregida por el error de predicción, Q^*c, se define como la adición de la estimación mínimo cuadrática de la demanda, Q^ y el error de predicción bootstrap, E*, y como, dado un conjunto de datos, Q^ está determinada, teniendo en cuenta que la varianza (y, por tanto, la desviación típica) es invariante en las traslaciones, se sigue la igualdad entre los valores de las celdas mencionadas.

Conclusiones de la tabla W

1. Las predicciones puntuales mínimo-cuadráticas del modelo econométrico para 1998 y 1999, ya mostradas en la pantalla anterior y que, para mayor facilidad del lector se repiten en la columna (2) de la presente tabla, son prácticamente iguales a las medias de las predicciones simuladas *corregidas por el error de predicción,* Q^*c, magnitud esta última que no es posible obtener en le Estadística tradicional, por lo que se puede ratificar la conclusión obtenida en la pantalla anterior: en el aspecto predictivo *puntual* el comportamiento de los mínimos cuadrados es muy robusto frente a eventuales desviaciones de las hipótesis básicas del modelo econométrico, a pesar del reducido número de datos históricos con que se trabaja.

2. Los intervalos de confianza al 95% de las predicciones convencionales (mínimo-cuadráticas) para los años *futuros* 1998 y 1999 son, respectivamente, [5199,1478; 6595,415] y [5180,3609; 7143,0133] (GWh), con longitudes respectivas de 1396,2672 y 1962,6524 (GWh). Así, para el primer año del *futuro*, 1998, los tipos de intervalos de confianza estandarizado, porcentual y porcentual *corregido por sesgo* del bootstrap resultan para la semilla 26939,85, resultando para cada uno de los tres intervalos, estandarizado, porcentual y porcentual con corrección por sesgo, los siguientes límites para el año 1998, [5310,622; 6411,619], [5320,512; 6403,699] y [5327,742; 6411,667] cuyas respectivas amplitudes son 1100,997, 1083,188 y 1083,926. Para el segundo año del *futuro*, 1999, los correspondientes intervalos de confianza bootstrap son [5314,829; 6864,323], [5318,980; 6836,632] y [5325,622; 6837,545] cuyas respectivas amplitudes son 1549,494, 1517,652 y 1511,923. Al variar la semilla de inicialización del algoritmo de generación de números aleatorios las cantidades obtenidas son muy similares a las descritas.

Por tanto, aunque los intervalos de confianza obtenidos al nivel de confianza del 95% tanto con la metodología convencional como con la del boostrap satisfacen el objetivo de *cobertura* de los valores realmente observados en el *futuro* la precisión de los intervalos confidenciales mínimo-cuadráticos es muy inferior a la obtenida con el bootstrap para la demanda *corregida por el error de predicción*. En los intervalos de confianza construidos para el año 1998, la precisión para cada uno de los tres métodos mencionados anteriormente mejora en un 21,47%, 22,42% y un 22,36%, respectivamente, siendo su mejora de precisión para el año 1999 de un 21,05%, 22,67% y un 22,96%, respectivamente. Por tanto, experimentalmente resulta también para el pronóstico corregido por el *error de predicción,* Q^*c, que aunque los intervalos de confianza obtenidos con las metodologías convencional y boostrap *cubren* a los valores de la demanda realmente observados en el *futuro* (años 1998 y 1999), los intervalos confidenciales bootstrap son considerablemente más precisos que los convencionales.

Efecto del número de ciclos bootstrap

Manteniendo ahora fija la semilla 29939,85 de inicialización del algoritmo de generación de números aleatorios, se ha ejecutado el programa BSTP para B= 1000, 2000, 3000, 4000, 5000, 8000, 10000, 15000 y 20000 ciclos bootstrap, llegándose a la concusión de que los parámetros reflejados en las tablas comentadas anteriormente prácticamente se estabilizan cuando se han realizado 2000 ciclos observándose en sus valores oscilaciones al aumentar el número de ciclos del bootstrap en vez de tendencias ascendentes o descendentes.

III.8. CONCLUSIONES DEL CAPÍTULO

El presente capítulo ha tenido como objetivo la investigación de un modelo(s) econométrico(s) de previsión de demanda anual de energía eléctrica en la C.A.A., basándose en variables de carácter económico. El horizonte temporal de datos históricos ha abarcado desde 1984 hasta 1999 pero cuando se realiza predicción *ex-post* sólo se consideran históricos los 14 primeros años de tal intervalo. Este objetivo el capítulo se estructura en siete secciones.

La primera sección, III.1, corresponde con una introducción acerca del contenido del capítulo, en lo que comentan los motivos por los cuales se han realizado determinadas investigaciones en el mismo.

En la segunda sección, III.2, se han clasificado los métodos de predicción en modelo de predicción cualitativos y modelos cuantitativos; éstos últimos, a su vez, pueden dividirse entre los modelos que utilizan el análisis univariante y el análisis causal. En el presente capítulo los modelos que han sido propuestos para la modelización de la predicción de energía eléctrica han sido los modelos de análisis causal, dentro del cual se han estudiado los modelos econométricos.

En la sección III.3 se han comentado los dos tipos de predicción que pueden realizarse con un determinado modelo. La predicción *ex-post* es el tipo de predicción que se ha aplicado, en el presente capítulo, a los datos de demanda anual estudiada.

La sección III.4 se ocupa de la evaluación de predicciones, cuantificando el grado de bondad de la predicción realizada sobre la variable de interés al compararla con el dato realmente observado para tal variable en dicho periodo, lo cual se efectúa mediante el denominado *error de predicción*. Así, un modelo de demanda de energía tendrá mejor capacidad predictiva cuanto menor sea

su error de predicción el cual se expresa, en el presente capítulo mediante los estadísticos más habituales en la literatura especializada para este fin.

La sección III.5 se dedica a la selección de variables independientes o exógenas que forman parte del modelo, así como a los requisitos que debe cumplir un modelo candidato para resultar elegido. La enumeración de estos requisitos es un tema abierto, habiéndose comentado los más habituales dentro de la literatura econométrica entre los que se encuentran el principio de parsimonia, identificalidad, bondad de ajuste, consistencia y poder predictivo. Dentro de los mismos, algunos de ellos son medibles mediante estadísticos, los cuales han sido definidos. La sección III.6 es, en cierto modo, el núcleo del capítulo III, ya que es donde se desarrolla el estudio empírico.

BIBLIOTECA VIRTUAL

Como se ha comentado anteriormente, el capítulo ha estado destinado a la estimación de un modelo econométrico de demanda anual de energía eléctrica residencial en la C.A.A. con una fuerte bondad de ajuste a los datos históricos y con fuerte poder predictivo hacia las observaciones futuras. A partir de la demanda de electricidad histórica de las observaciones (1984-1999), se han contemplado variables explicativas altamente correlacionadas con la demanda de electricidad.

Una vez definidas estas variables independientes o exógenas (precio unitario de la electricidad, precio unitario de butano expresado en kWh, precio unitario del gas natural expresado en kWh o el producto interior bruto de la C.A.A. expresada en pesetas constantes de 1984 y la demanda de energía retardada un año) se construyen cuatro modelos econométricos los cuales son del tipo log-lineal o elasticidad constante y que constituyen los modelos a través de los cuales se ha realizado la investigación.

En cada uno de estos cuatro modelos se han realizado contrastes de diagnóstico basados en los coeficientes de regresión del modelo, estructura de los residuos y en la estabilidad del modelo. Este conjunto de contrastes ha evaluado, para cada modelo considerado, su nivel de la calidad.

El tercer modelo de previsión, modelo 3, (función del precio de la electricidad y el producto interior bruto) planteado ha sido un modelo con un fuerte poder explicativo frente al resto de modelos; sin embargo, se ha detectado en él un error de especificación debido a variables omitidas, circunstancia, que ha provocado la modificación de sus variables explicativas. La variable eliminada ha sido el producto interior bruto la cual ha sido sustituida por el producto del precio de la electricidad y el producto interior bruto. Los contrastes evaluadores del modelo han sido favorables a la elección de este modelo como explicativo de la demanda de energía eléctrica residencial en la C.A.A.

En este punto fue necesario preguntarse la cuestión acerca del efecto de la variable precio de la electricidad en un periodo y su incidencia en la demanda de un periodo posterior. En la investigación de un modelo mejor, se probaron modelos con las variables explicativas retardadas un periodo dando lugar, a un primer modelo, denominado modelo 3R, el cual considera el precio de la electricidad y el precio interior bruto, variables aleatorias retardadas un periodo. Los resultados del análisis de dicho modelo conducen, al igual que en el modelo 3, a la no linealidad de las variables en la ecuación. Todo lo anterior provocó la consideración nuevamente de la generación de una nueva variable definida como el producto cruzado de ambas variables retardadas un periodo estando también el precio de la electricidad rezagada un periodo.

De esta forma, los modelos candidatos construidos han sido seis, cuatro de los cuales son modelos ya considerados en el inicio de esta sección mientras que los dos restantes han sido construidos para salvar deficiencias en el modelo 3.

Una vez analizada su capacidad explicativa, se analizaron los modelos desde la óptica de su capacidad de predicción, teniendo en cuenta que, un modelo puede alcanzar un buen grado de ajuste sin por ello garantizar una buena capacidad predictiva respecto al futuro.

Los resultados de la capacidad predictiva de los modelos candidatos también evidencian que el modelo 3PCR es el más adecuado. Para corroborarlo se han realizado dos predicciones *ex-post* para cada uno de los seis modelos: la primera de ellas, consiste en una predicción de la demanda del año *futuro* 1999 y la segunda en una predicción para los años *futuros* 1998 y 1999. Los resultados obtenidos asignan a los modelos 3PC y 3PCR los menores errores de predicción, siendo el modelo 3PCR más robusto. Los resultados experimentales muestran que los intervalos construidos para el modelo 3PCR contienen al verdadero valor de la demanda en dichos periodos.

Modelo 3PCR:

$$LnQ_i = 9{,}606683 - 5{,}834969 LnPE(-1)_i + 0{,}339545(LnPELnPIB)(-1)_i$$

Para complementar este estudio y con el objetivo de decidir qué modelos son mejores desde el punto de vista de la bondad de ajuste y de la predicción se ha aplicado la técnica de decisión multicriterio Promethee. Esta técnica se utiliza cuando, ante posibles alternativas a las cuales se enfrenta el decisor, no existe una única que satisfaga a todos los criterios que entran en juego. Esta circunstancia es la que se observa en el estudio desarrollado. La aplicación de esta técnica conduce a proponer como mejor modelo explicativo y de predicción de la demanda anual de energía residencial eléctrica en la C.A.A. al modelo 3PCR, muy seguido en el orden de preferencia por el modelo 3PC. Por otra parte, los modelos con menor grado de ajuste y poder predictivo, por orden ascendente en orden de preferencia, han sido el modelo 1 y modelo 2.

Las estimaciones puntuales de los coeficientes de regresión calculados mediante las metodologías tradicional y bootstrap, presentan valores muy similares lo cual es indicativo de la robustez de los estimadores convencionales, en el modelo estimado en este trabajo, frente a posibles desviaciones de la hipótesis habituales del modelo lineal. El método mínimo-cuadrático sobreestima la variabilidad de los estimadores de los coeficientes de regresión resultando la desviación típica de este método aproximadamente un 11% superior a la proporcionada por la metodología bootstrap, por lo que la variabilidad de tales estimadores es menos satisfactoria que su localización. Las estimaciones

bootstrap obtenidas como valor de los coeficientes de regresión son muy *precisas* (distribuciones marcadamente *leptocúrticas*, muy concentradas alrededor de la media). Las predicciones puntuales convencionales prácticamente coinciden tanto con la medias de los datos simulados, como con la media de las predicciones puntuales simuladas por el bootstrap. Los tres tipos de intervalos de confianza bootstrap mencionados en este trabajo son considerablemente mas precisos que los intervalos de confianza convencionales.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

CAPÍTULO IV
ESTUDIO DE LA DEMANDA RESIDENCIAL MENSUAL DE ENERGÍA ELÉCTRICA EN LA COMUNIDAD AUTÓNOMA DE ANDALUCÍA

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

IV.1. INTRODUCCIÓN

El objetivo de todo sistema eléctrico es dar cobertura a la demanda de electricidad de forma eficiente. En este capítulo se estudian modelos de previsión en el horizonte temporal enero 2001-diciembre 2001. El ámbito del estudio es el consumo residencial de energía en la Comunidad Autónoma de Andalucía utilizando datos históricos mensuales desde 1995 a 2000. La demanda de electricidad ha sido objeto de diversos estudios. Sin embargo, para el caso particular de la Comunidad Autónoma de Andalucía los estudios han estado reservados básicamente a organismos públicos de gestión y control, así como a las propias compañías eléctricas.

La estructura seguida en este capítulo lo divide en siete secciones que se comentan brevemente a continuación. En la sección IV.2 se realiza un estudio detallado y comparativo de quince modelos univariantes, es decir, de aquéllos que tratan de explicar la variación de la demanda a través de sí misma. En la sección IV.3 se analiza la influencia de la temperatura ambiente en la demanda objeto del presente trabajo. En la sección IV.4 se construyen setenta y nueve modelos de regresión dinámicos para temperaturas *neutras* comprendidas en un rango de valores desde los quince hasta los veinticinco grados centígrados. Se realiza el estudio de la demanda mencionada mediante una familia de modelos de regresión dinámica, habiéndose particularizado para la temperatura ambiente de veintidós grados centígrados. La elección de esta temperatura *neutra* concreta se ha realizado acorde con la correlación existente entre las temperaturas y la demanda doméstica de electricidad en Andalucía aún cuando el presente trabajo no entra en la discusión de lo que pudiera ser una cuantificación subjetiva de tal temperatura *neutra*. El modelo elegido y detallado es, además, muy robusto frente a modificaciones de la temperatura *neutra*.

La sección IV.5 analiza los modelos que mejor explican la predicción de la demanda mensual en la Comunidad Autónoma de Andalucía a partir de datos mensuales. Con la información de las predicciones de cada uno de estos modelos

para el horizonte temporal de un año se lleva a cabo una combinación de las mismas con el objetivo de obtener una predicción más precisa para el horizonte mencionado.

La sección IV.6 esta dedicada al futuro de los modelos econométricos considerados, ante la última etapa de liberalización del sector eléctrico y gasista.

El capítulo finaliza con la sección IV.7 dedicada a exponer las principales conclusiones del mismo.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

IV.2. ESTUDIO COMPARATIVO DE MODELOS UNIVARIANTES. DATOS MENSUALES

En esta sección del capítulo se desea encontrar un modelo, a través del análisis de series temporales, que ayude a la descripción y previsión de la evolución de la demanda mensual de energía eléctrica en la C.A.A. El conocimiento de lo sucedido en el pasado a través de la observación de sus regularidades puede constituir un soporte para la elaboración de predicciones con respecto al futuro. En este apartado se analizan los datos de la demanda mensual de energía eléctrica en el sector residencial de la C.A.A., en el periodo comprendido entre enero de 1995 y diciembre de 2000. Los datos proceden de la consulta realizada a la página web del Instituto de Estadística de Andalucía en la sección de *Indicadores Económicos de Andalucía* (http://iea.junta-andalucia.es/ineco_web). Se designará por QMENSUAL a la serie cronológica definida por tales datos.

El objetivo del análisis de series temporales es estudiar el comportamiento de una (o varias) variables a lo largo del tiempo, o más formalmente, investigar la estructura de la relación entre las variables de un proceso estocástico. Un proceso estocástico es una sucesión de variables aleatorias X_t, ordenadas según el subíndice *t*, que siguen una ley de probabilidad conjunta.

Se dice que un proceso estocástico es *estacionario en sentido amplio* (débil) si su media es constante e independiente del tiempo y la covarianza entre dos variables aleatorias del proceso no depende del tiempo cronológico en que se midan las mismas, sino sólo del número de periodos (*distancia o salto*) de separación entre dichas variables aleatorias. Una consecuencia de esta última condición es que la varianza del proceso es finita y constante. En adelante, se hará referencia a los procesos estacionarios en sentido *débil,* simplemente como *procesos estacionarios*

Una realización de un proceso estocástico es una muestra de tamaño uno de cada una de las variables que componen el proceso. Se considera que una serie cronológica es parte de la realización de un proceso estocástico o que el mecanismo de generación de una serie es un proceso estocástico. Se suele decir que una serie temporal es una sucesión de valores generalmente registrados a intervalos equidistantes del tiempo.

La sistematización de las diferentes metodologías empleadas en el tratamiento de una serie temporal puede hacerse en base a criterios diversos, que dan lugar entre otros al *enfoque clásico* de descomposición. El enfoque que se trata a continuación es el enfoque *determinista* y *univariante* el cual trata de explicar la trayectoria de una variable a través de la información contenida en los datos históricos de su correspondiente serie, es decir, intenta capturar el comportamiento sistemático que muestra el pasado de la misma y en base a ello realizar predicciones respecto al futuro.

La mayoría de las series cronológicas que se utilizan en la práctica, pueden modelizarse a través de alguno de los esquemas de integración, estos son, el esquema *aditivo* y el *esquema multiplicativo*. Representando por:

T_t : valor de la tendencia.

C_t : valor del factor cíclico.

E_t : valor del componente estacional.

I_t : valor del componente irregular.

El esquema aditivo postula una relación del tipo

$$Y_t = T_t + C_t + E_t + I_t$$

es decir, el valor de la serie temporal para cualquier instante es igual a la suma de los valores correspondientes a sus cuatro componentes.

El esquema multiplicativo se expresa mediante la relación

$$Y_t = T_t \times C_t \times E_t \times I_t$$

lo que significa que el valor de la serie temporal en cualquier instante es igual al producto de los valores correspondientes a sus cuatro componentes.

La razón de utilizar una de estas dos hipótesis, en lugar de otras más complejas, radica fundamentalmente en la sencillez y operatividad de las mismas.

De esta forma, se construirán modelos de series temporales que constan de:

- *Nivel*: es un proceso no estacional, suave, de cambios de pequeña magnitud subyacente en las observaciones. No puede medirse directamente sino que debe estimarse a partir de los datos.

- *Tendencia Local*: se trata de cambios suaves y casi imperceptibles del nivel. El calificativo de *local* intenta resaltar que el cambio tiene lugar en cada punto de la serie. La predicción se basa en la tendencia local del último dato histórico, no en su tendencia global. La tendencia local no se puede medir directamente y se estima a partir de los datos.

- *Efectos Estacionales:* los índices estacionales (aditivos o multiplicativos) representan comportamientos o patrones periódicos de la serie, en cada punto de la cual se supone que tiene lugar una pequeña modificación de la estacionalidad. Al igual que el nivel y la tendencia los índices estacionales se estiman a partir de los datos.

- *Perturbaciones aleatorias:* a veces se le denominan choques aleatorios y representan las variaciones no sistemáticas. No son medibles directamente pero se pueden estimar mediante los residuos de la serie o diferencias entre los valores observados y los pronosticados.

El nivel y las perturbaciones aleatorias están presentes en toda serie. La tendencia local y la estacionalidad pueden estar presentes o ausentes. La identificación de un modelo consiste en determinar cuales de los componentes

anteriores deben estar incluidos en el mismo de forma que se obtenga una representación satisfactoria de los datos.

Una vez definidos los posibles esquemas que pueden seguir las series, se estudian los distintos modelos univariantes estudiados para la demanda de electricidad en la C.A.A., en función de su grado de complejidad.

Métodos aplicables a series sin tendencia ni estacionalidad

Modelos Ingenuos

El modelo más simple de serie temporal es el *puramente aleatorio*, $Y_t = \varepsilon_t$, habitualmente denominado *ruido blanco*. Si una serie tiene estructura de *ruido blanco* su gráfica sigue una trayectoria que no obedece a ningún patrón. Formalmente, si tal serie se designa por ε_t, decimos que ε_t es un *ruido blanco* si cumple las condiciones siguientes:

$$E[\varepsilon_t] = 0,\ \forall t;$$

$$E[\varepsilon_t^2] = \sigma^2,\ \forall t;$$

$$E[\varepsilon_t \varepsilon_{t'}] = 0,\ t \neq t';$$

Una serie *ruido blanco* es impredecible, y por tanto, no admite ningún tipo de pronóstico ya que no contiene ninguna información sobre relaciones temporales. A veces, a un ruido blanco se le denomina *innovación.*

Se dice que, para una determinada serie, un modelo es *adecuado* cuando éste ha extraído la máxima información sobre la serie y todo lo que queda de su paso por el *filtro* (los residuos) es aproximadamente *ruido blanco*. En realidad, el filtro *es* el modelo. Si los residuos aún evidenciaran cierta estructura (es decir, contuvieran alguna información útil), ésta podría utilizarse para mejorar el modelo hasta entonces contemplado.

Otro tipo de modelo ingenuo es el denominado *paseo aleatorio*, el cual es un modelo de serie temporal en el que la diferencia entre dos observaciones consecutivas en el tiempo es un *ruido blanco,* $Y_t = Y_{t-1} + \varepsilon_t$. Este tipo de modelos

se caracteriza por su simplicidad y se suelen emplear como inicio en la modelización. La predicción realizada por el modelo se hace coincidir, bajo una visión ingenua, con el último valor observado, de tal forma que se expresa:

$$\hat{Y}_{t+1} = Y_t$$

donde Y_t es el valor observado de la variable aleatoria en el instante t e $\hat{Y}_{t+1}$ la predicción que efectúa en el tiempo t sobre el valor que la serie tendrá en el periodo t+1. Con la terminología de los procesos ARIMA, la cual se abordará en el próxima sección, se dirá que un *paseo aleatorio* es un proceso no estacionario AR(1) con el coeficiente $\phi_1 = 1$.

Modelos de Medias Móviles

La media móvil es una media aritmética que se caracteriza porque toma un valor diferente en cada momento del tiempo y porque en su cálculo no entran todas las observaciones de la muestra disponible. Para la aplicación de este método hay que definir su *longitud* (p), es decir, el número de observaciones que intervienen en el cálculo de cada media móvil.

El término general de una media móvil de orden p (siendo p un número impar) centrada en el periodo t será:

$$MM(p)_{\mathrm{t}} = \frac{\left(Y_{\frac{\mathrm{t-(p-1)}}{2}} + Y_{t-1} + \ldots + Y_{\frac{t+(p-1)}{2}} \right)}{p}$$

En el caso de que p sea par, la expresión correspondiente vendrá dada por:

$$MM(p)_{\mathrm{t+0,5}} = \frac{\left(Y_{\frac{\mathrm{t-(p-1)}}{2}+0,5} + Y_{t-1} + \ldots + Y_{\frac{t+(p-1)}{2}+0,5} \right)}{12}$$

donde ahora la media móvil quedará descentrada, o centrada entre el periodo t y t+1, es decir, en t+0,5.

Los valores calculados a través de las medias móviles se utilizarán para predecir el valor de la variable en el periodo t+1.

Modelo de Alisado Exponencial Simple

Las técnicas de *alisado o suavización exponencia*l se encuentran entre las más utilizadas en la predicción, por su facilidad de utilización y de comprensión conceptual. Los estudios comparativos empíricos de Makridakis(1989) le asignan una considerable exactitud, robustez y flexibilidad de adaptación ante la posibilidad de un cambio en la estructura de los datos. Por otra parte, estos métodos suelen requerir menor número de datos que los métodos correlacionales.

Las series más propicias para la utilización de los métodos de alisado exponencial suelen ser aquéllas cuyos datos no muestran una marcada estructura correlacional y, en particular, estos métodos suelen ser convenientes en situaciones en las que se presenta al menos una de estas dos circunstancias:

- el escaso número de datos históricos no permite calcular un número adecuado de los coeficientes de correlación.
- Las correlaciones calculadas no son estables. Cuando las autocorrelaciones de los datos permanecen aproximadamente constantes, el método de Box-Jenkins o el de regresión dinámica suelen presentar mejores resultados que el alisado exponencial.

Varios de los métodos de alisado exponencial resultan ser casos particulares de la familia de modelos de Box-Jenkins, aunque la filosofía de aplicación de ambas metodologías sea completamente diferente.

Un trabajo importante de revisión de los métodos de alisado exponencial es el de Gardner, E.S.(1985) quien los clasificó en 17 métodos pudiendo considerarse, a su vez, extensiones de los mismos- y en el que sugería que los métodos de parámetro adaptativo y de Brown (considerados durante mucho tiempo como fundamentales en el alisado exponencial) podían quedar incluidos en otros más generales. Posteriormente Makridakis, S. Y Hibon, M.(1989) demostraron que el método de alisado exponencial de Holt-Winters proporciona predicciones mucho más *exactas* que el de Brown.

Los métodos de suavización exponencial se basan en un modelo estructural de datos de la serie cronológica en el que se supone que ésta presenta algunos o

todos de los componentes ya comentados anteriormente, como son el nivel, la tendencia, componentes ciclo y la perturbación aleatoria.

A continuación se comentan brevemente algunos métodos de este tipo utilizados en este trabajo. En el método de alisado exponencial *simple* los únicos componentes de la serie son el nivel y las perturbaciones aleatorias, ya que se supone que los datos consisten en fluctuaciones irregulares alrededor de un nivel aproximadamente constante, por lo que no debe emplearse este método cuando los datos muestren tendencia o estacionalidad. Si la predicción para el periodo *t* es F_t, cuando la observación Y_t esté disponible, el error de predicción correspondiente a dicho período será $e_t = Y_t - F_t$. El alisado exponencial simple trata de compensar este error de modo que el pronóstico en *t* para el periodo *t*+1 es :

$$F_{t+1} = F_t + \alpha(Y_t - F_t) \qquad (5)$$

donde α es una constante comprendida entre 0 y 1, ecuación que pone de manifiesto cómo se actualiza el nivel de la serie, $L_t = F_t$, conforme se dispone de nuevas observaciones. Por eso, cualquier predicción de horizonte temporal superior a un periodo será constante e igual al último valor observado Y_t, ya que no se dispone de los correspondientes errores, es decir, $F_{t+m} = Y_t$ (*m>1*). La ecuación (5) supone disponer del valor de α y de un valor inicial de $\hat{Y}_t = F_t$ cuestión esta última que es prácticamente irrelevante en la obtención de los sucesivos cálculos y que se suele resolver tomando como valor de $\hat{Y}_1$ el de una media aritmética de algunas observaciones iniciales de la serie. En cuanto al valor de α la mayoría de los programas de ordenador proporcionan su valor *óptimo* por el simple procedimiento de elegir aquel valor de α que minimiza la suma de los errores cuadráticos medios de predicción de las primeras 20 a 30 observaciones.

Métodos aplicables a series con tendencia y sin estacionalidad

Modelo de Holt

En el método de alisado exponencial de dos parámetros *de Holt* se supone que los datos fluctúan alrededor de un nivel variable b_t que sigue aproximadamente la tendencia de una línea recta. Las ecuaciones de predicción y actualización en este método son las siguientes:

$$L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + b_{t-1}) \quad (6)$$

$$b_t = \beta(L_t - L_{t-1}) + (1-\beta) b_{t-1} \quad (7)$$

$$F_{t+m} = L_t + b_t m \quad (8)$$

BIBLIOTECA VIRTUAL

donde α y β son parámetros con valor entre cero y uno, siendo L_t y b_t, respectivamente estimaciones del nivel y de la pendiente de la serie en el tiempo *t.* La ecuación (6) pone de manifiesto que la actualización del nivel se obtiene como una ponderación de los datos *nuevos* (primer sumando) y la mejor estimación del nuevo nivel basado en datos *anteriores* (segundo sumando) lo cual persigue que L_t tenga, aproximadamente, el nivel del último dato. La ecuación (7) actualiza la tendencia. Finalmente la ecuación (8) es la que se utiliza para predecir los datos futuros.

Métodos aplicables a series con tendencia y estacionalidad

Modelo de Holt-Winters

Los métodos anteriores son adecuados para tratar la mayoría de las series que no tengan una marcada estacionalidad ya que, caso de existir ésta, resulta más apropiado el método de Holt-Winters en el que la predicción de cada observación se obtiene multiplicando su valor desestacionalizado (el cual se supone obtenido mediante el método anterior de dos parámetros de Holt) por el índice estacional del correspondiente período. Este método incluye tres parámetros α, β y γ, cuyos valores están comprendidos entre cero y uno.

Las ecuaciones de alisado del método multiplicativo de Holt-Winters son:

Ecuación de nivel $$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + b_{t-1}) \quad (9)$$

Ecuación de pendiente $$b_t = \beta(L_t - L_{t-1}) + (1-\beta) b_{t-1} \quad (10)$$

Ecuación de factor estacional $$S_t = \gamma \frac{Y_t}{L_t} + (1-\gamma) S_{t-s} \quad (11)$$

donde s es la longitud de la estacionalidad (por ejemplo, el número de meses del año), L_t representa el nivel de la serie, b_t designa la tendencia y S_t es el índice estacional. La ecuación de actualización de la tendencia (10) es igual que la correspondiente del método de Holt. La ecuación (9) es similar a la ecuación (11) del método de Holt excepto que se desestacionaliza la última observación dividiéndola por el índice estacional que para el periodo t que se calculó hace s periodos, ya que no se puede utilizar S_t hasta que no se conozca L_t a partir de la ecuación (9). El índice estacional se calcula como la razón de la observación actual al nivel suavizado actual promediada con el último valor calculado del índice estacional correspondiente a este período. En el método de Holt-Winters no es posible obtener una expresión cerrada para los intervalos de confianza por lo que la expresión aproximada que se utiliza en este trabajo corresponde a la de Chatfield, C, y Yar, M.,(1991).

La ecuación de la predicción efectuada en el tiempo t con horizonte temporal de *un periodo*, es decir correspondiente al periodo (t+m) es:

$$F_{t+m} = (L_t + b_t m) S_{t-s+m} \quad (12)$$

Métodos Dinámicos

Modelo Box-Jenkins

En la metodología de Box-Jenkins el modelo se construye exclusivamente a partir de los datos sin necesidad de sugerir ningún modelo *a priori* que pudiera resultar arbitrario. Cuando los datos presentan una estructura correlacional considerablemente estable y no estacional el método Box-Jenkins aventaja con frecuencia a los modelos exponenciales de alisado simple y de Holt, debido a que estos dos últimos modelos son casos particulares de la familia de modelos de Box-Jenkins (que también contemplan las series estacionales), ventaja que no es empíricamente tan frecuente frente al modelo *multiplicativo* de Holt-Winters, posiblemente debido a que este método no se corresponde con ningún miembro de la familia de modelos Box-Jenkins, lo que sí ocurre, en cambio, para el modelo *aditivo* de Winters.

Entre las causas que posiblemente hayan dificultado una mayor aceptación de la metodología de Box-Jenkins en el ámbito empresarial pueden encontrarse su mayor complejidad, coste y requerimiento de número de datos históricos (al menos 50) en relación con otros métodos.

El objetivo principal de la metodología Box-Jenkins consiste en construir un modelo que reproduzca adecuadamente la estructura correlacional de los datos. Un modelo Box-Jenkins se construye a partir de las funciones muestrales de autocorrelación y de autocorrelacion parcial de los datos correspondientes a la variable investigada. La optimalidad del método depende de las propiedades de los datos y, más concretamente, de la hipótesis de estacionariedad del modelo a identificar.

Según expusieron sus autores al presentar su método en Box, G.E.P. y Jenkins. G.M.(1976), la primera etapa del mismo consiste en la *identificación* de uno o varios modelos candidatos de entre una extensa y compleja familia de modelos como una aproximación al *verdadero* proceso generador de los datos observados. Seguidamente se procede a la *estimación* de los parámetros del modelo mediante el ajuste del correspondiente modelo candidato a los datos

históricos. A continuación tiene lugar la etapa de *contrastes-diagnósticos*, generalmente mediante el análisis de la autocorrelación de los residuos (diferencias entre los datos observados y los pronósticos del modelo correspondientes al período histórico). Si el modelo candidato no supera dichos *contrastes-diagnósticos* es rechazado, procediendo de nuevo a repetir el ciclo *identificación-estimación-diagnóstico* hasta que finalmente se acepta un modelo, momento a partir del cual se trabaja bajo el supuesto de que tal modelo representa la *verdadera* distribución estadística de los datos observados y, por extensión, también de los datos futuros. Se obtiene así la distribución de la serie en el futuro *condicionada* a las observaciones del pasado y del presente. Si la hipótesis anterior es correcta, la media condicional de tal distribución futura es la predicción óptima mínimo-cuadrática, siendo la varianza de dicha distribución un estimador válido de la precisión del pronóstico. En el caso de que más de un modelo candidato supere los *contrastes-diagnósticos,* Box y Jenkins recomiendan elegir el de estructura más simple.

Una representación esquemática de la metodología Box-Jenkins es la que se representa en la figura IV.2.1:

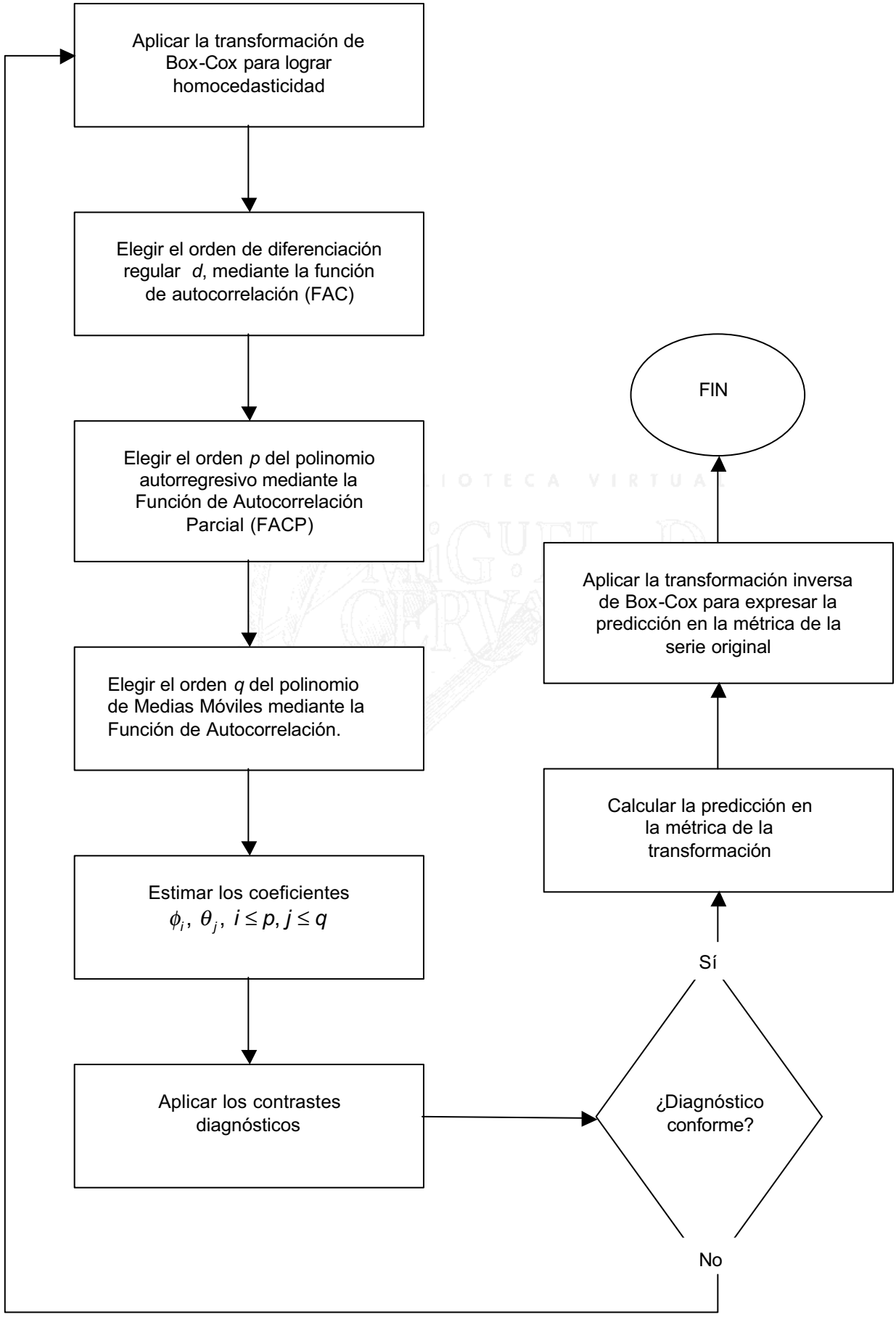


Figura.IV.2.1. Procedimiento Box-Jenkins a través de un diagrama de flujo. Fuente. Goodrich, R.L.(1990). Elaboración. Propia.

El enfoque univariante Box-Jenkins incluye una familia de modelos que utilizan una notación estándar. La representación más general de un modelo ARIMA es la siguiente:

$$\Phi(B^s)\phi(B)w_t = \Theta(B^s)\theta(B)\varepsilon_t \qquad (13)$$

$$w_t = \nabla_s^d \nabla^d Y_t^{(\lambda)} - \overline{w}$$

$$\varepsilon \sim N(0,\sigma_\varepsilon);\ E(\varepsilon_t,\varepsilon_{t'}) = 0,\ \text{para } t \neq t'$$

donde

- s es el número de periodos que define el comportamiento estacional de la serie.
- Y_t es el valor –en el tiempo t- de la serie original estudiada.
- $Y_t^{(\lambda)}$ es el resultado de la transformación de Box-Cox aplicada a Y_t para lograr la estacionariedad en varianza.
- $\overline{w}$ representa la media de w_t.
- B es el operador de retardo regular, tal que $BY_t = Y_{t-1}$
- B^s es el operador de retardo estacional, tal que $B^s Y_t = Y_{t-s}$
- ∇ es el operador *diferencia regular*; $\nabla = 1 - B$ tal que $\nabla Y_t = Y_t - Y_{t-1}$
- ∇_s es el operador diferencia estacional; $\nabla_s = 1 - B^s$, tal que $\nabla_s Y_t = Y_t - Y_{t-s}$,
- d y D son parámetros enteros no negativos que indican, respectivamente, el número de diferencias regulares y estacionales aplicadas a $Y_t^{(\lambda)}$ para lograr la estacionariedad en media de la serie.
- $\phi(B) = 1 - \phi_1 B - ... - \phi_p B^p$ es el operador autorregresivo regular (AR)
- $\theta(B) = 1 - \theta_1 B - ... - \theta_p B^q$ es el operador de media-móvil regular (MA)
- $\Phi(B^s) = 1 - \Phi B^s - ... - \Phi_P B^{P_s}$ es el operador autorregresivo estacional
- $\Theta(B^s) = 1 - \Theta_1 B^s - ... - \Theta_Q B^{Qs}$ es el operador de media móvil estacional
- p, q, P y Q son parámetros enteros no negativos que indican el orden respectivo de cada uno de los cuatro operadores anteriores, $\phi(B)$, $\theta(B)$ $\Phi(B^s)$, $\Theta(B^s)$.
- $\{\phi_i\}$, $\{\Phi_i\}$, $\{\theta_i\}$, $\{\Theta_i\}$ son parámetros continuos

- ε_t es la serie de los errores que sigue un proceso estocástico de ruido blanco $N(0,\sigma_\varepsilon)$ donde σ_ε es constante.

Por simplicidad es habitual referirse a un proceso de este tipo como un proceso ARIMA $(p, d, q)x(P, D, Q)_s$

A la operación de aplicar $\nabla^d (\nabla_s^D)$ se le suele denominar, respectivamente, *diferenciar regularmente d veces la serie* o *aplicar una diferencia regular de orden d (diferenciar estacionalmente D veces la serie* o *aplicar una diferencia estacional de orden D*) siendo habitual referirse a la serie resultante de la aplicación de tales operadores como la *serie diferenciada regularmente d veces (serie diferenciada estacionalmente D veces)*

La identificación de un modelo ARIMA supone que previamente se dispone de una serie *estacionaria* lo que exige que han de ser constantes la media y la varianza de los datos, lo cual frecuentemente requiere la transformación previa de los mismos.

La posible heterocedasticidad de los datos de la serie original se suele corregir aplicando una *transformación de Box-Cox, o de potencia,* definida como sigue:

$$Y_t^{(\lambda)} = \begin{cases} \dfrac{Y_t^{\lambda} - 1}{\lambda} & \lambda \neq 0 \\ \ln Y_t & \lambda = 0 \end{cases} \qquad (14)$$

donde λ es un número real. La tabla IV.2.1. muestra los *resultados de la variable transformada* para algunos valores frecuentes de λ, donde la última columna no debe ser interpretada en el sentido estrictamente matemático de la igualdad (14) sino como expresión de la *naturaleza de la transformación efectuada* ya que la sustracción de 1 y la posterior división por λ no altera la *naturaleza* de la serie.

Tabla.IV.2.1.Resultados de la variable transformada.

Valor de λ	Transformación	Serie Transformada	Resultado(≅)
$\lambda = 1$	Sin transformación	$Y_{(1)} = \frac{Y^1 - 1}{1}$	Y
$\lambda = 0{,}5$	Raíz cuadrada	$Y_{(1)} = \frac{Y^{0,5} - 1}{0{,}5}$	$\sqrt{Y}$
$\lambda = 0$	Logarítmica	$Y_{(1)} = \frac{Y^0 - 1}{0}$	$\ln Y$
$\lambda = -0{,}5$	Inversa de la raíz cuadrada	$Y_{(1)} = \frac{Y^{-0,5} - 1}{-0{,}5}$	$\frac{1}{\sqrt{Y}}$
$\lambda = -1$	Inversa	$Y_{(1)} = \frac{Y^{-1} - 1}{-1}$	$\frac{1}{Y}$

Fuente. Pena, J.B.(1999)
Elaboración. Propia

Para determinar experimentalmente el valor adecuado de λ, se consideran diversos valores de este parámetro y se elige aquél que consiga que el gráfico de media-desviación típica se aproxime a una línea horizontal. También se puede determinar aquel valor de λ para el que sea mínima la correlación entre la desviación típica y la media de la serie. Otros autores (Levenbach o Abraham) estiman λ por máxima verosimilitud. En los casos frecuentes en que la variabilidad de los datos aumenta con el nivel de los mismos se suele conseguir la estacionariedad de la varianza mediante la transformación logarítmica ($\lambda = 0$) o calculando la raíz cuadrada ($\lambda = 0{,}5$) de los datos originales.

En cuanto a la estacionariedad en media, el procedimiento habitual para conseguirla suele consiste en *diferenciar* regular y/o estacionalmente los datos (los cuales ya se suponen estacionarios en varianza) Y_t, construyendo una nueva serie $Z_t = Y_t - Y_{t-1}$. Caso de que después de realizar la operación mencionada todavía no se haya conseguido la estacionariedad de la media, puede ser necesario, a su vez, *diferenciar* la serie Z_t pero, en la práctica, no suele ser necesario diferenciar los datos más de una o dos veces para lograr dicha

estacionariedad, teniendo en cuenta que no se debe incurrir en una *sobrediferenciación* superflua de los datos que distorsionaría la naturaleza de la serie que se pretende analizar.

La ecuación (13) está referida a la variable diferenciada w_t correspondiendo a (13) otra ecuación en la variable original Y_t a la que se denomina la variable *integrada* (I) de w_t, lo que añadido al proceso *autorregresivo* (AR) de *medias móviles* (MA) al que se refiere la ecuación (13) explica el término ARIMA con que se suele denominar a estos procesos. Box y Jenkins no han sido los primeros autores que se han ocupado del estudio de los modelos ARIMA, aunque frecuentemente se suele identificar a estos autores con los procesos ARIMA debido a que desarrollaron un método para la construcción sistemática de tales modelos.

Una vez explicados los principales métodos y modelos de series temporales, se comienza el análisis de las series de demanda de energía eléctrica en C.A.A. durante el periodo 1995-2000. En primer lugar, se examinan tres procedimientos exploratorios para determinar el tipo de esquema al que corresponde la serie de datos de la demanda mensual de electricidad en C.A.A. durante el periodo 1995-2000. Estos procedimientos exploratorios son la *representación gráfica de la serie*, la *elaboración del gráfico desviación típica-media* y el *análisis de la variabilidad de las diferencias y cocientes estacionales.* Los criterios que se describen a continuación no revelan individualmente con exactitud el modelo que se debe seguir, ya que habrá que probar con más de uno, aunque sin embargo sirven de orientación en la búsqueda del mejor modelo.

Para ello en primer lugar se realiza una representación gráfica de la serie como el de la figura IV.2.2 con el objetivo de detectar si las componentes (tendencia, ciclo, estacionalidad y perturbación) se asocian de forma aditiva o multiplicativa. Se puede observar que la magnitud de las fluctuaciones de la demanda de electricidad son más altas cuando ésta es creciente, es decir, la

variabilidad de la serie no se mantiene constante en el tiempo, por lo que parece ser que el esquema aconsejable adoptar será el multiplicativo.

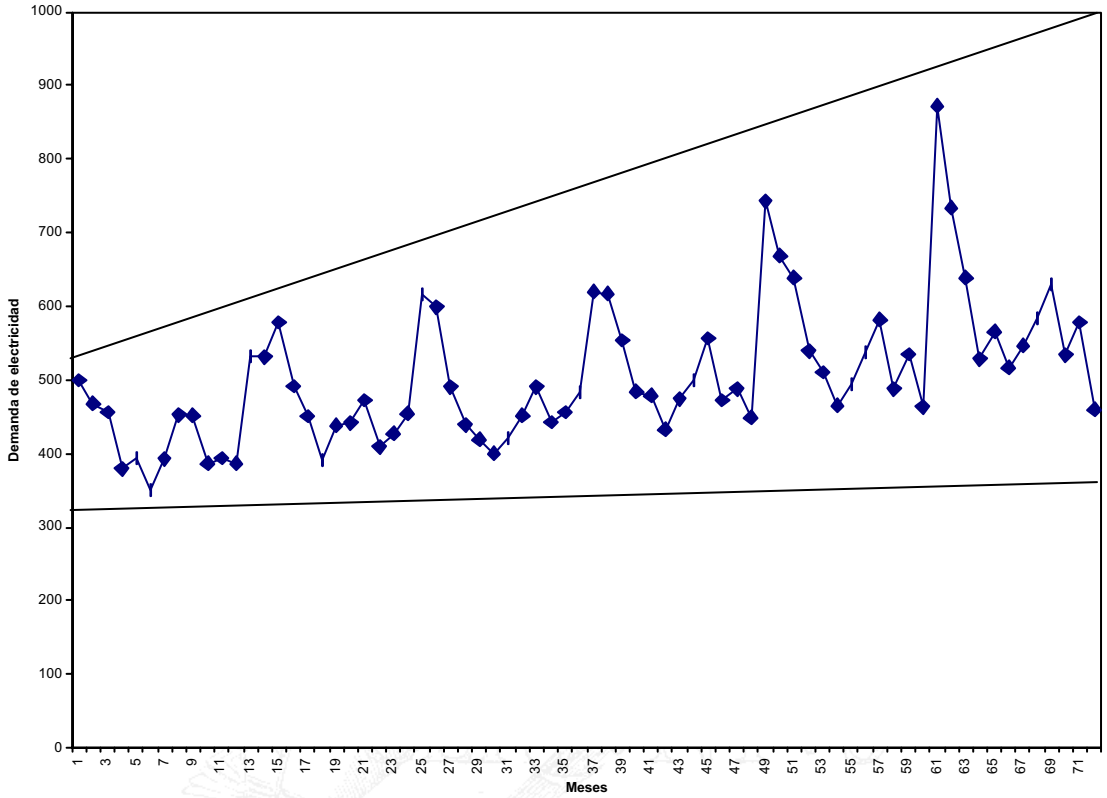


Figura.IV.2.2. Evolución de la demanda de electricidad en la C.A.A. durante el periodo 1995-2000.
Fuente. IEA.
Elaboración. Propia.

El segundo procedimiento analiza el gráfico desviación típica-media, en el cual se representa gráficamente, figura IV.2.3, a través de un diagrama cartesiano, la media y desviación típica de la demanda dentro de cada año, a fin de analizar los resultados. A primera vista, parece que la nube de puntos resultante de representar los puntos (media, desviación típica) para cada año, se distribuye más alrededor de una recta que de forma aleatoria, con lo que parece que los componentes seguirán una asociación multiplicativa.

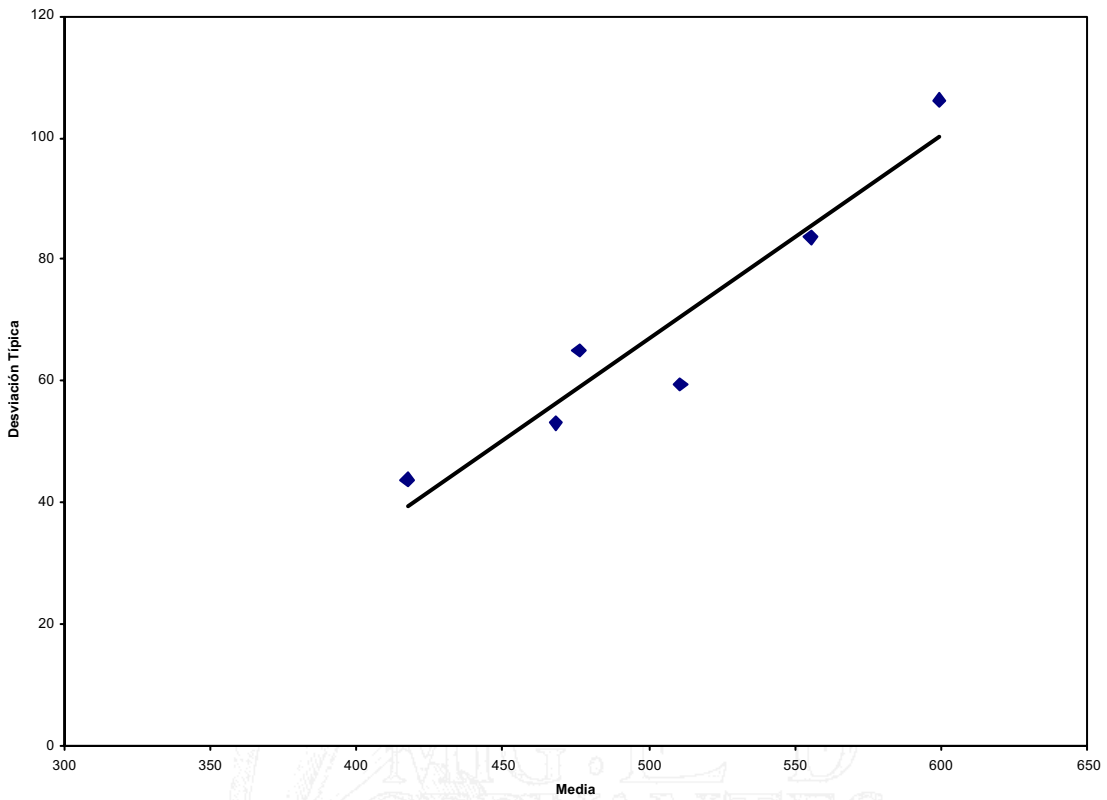


Figura.IV.2.3. Gráfico desviación típica-media. Demanda mensual de energía eléctrica durante el periodo 1995-2000.
Fuente.IEA.
Elaboración. Propia.

Los resultados obtenidos anteriormente parecen no ser muy concluyentes, por lo que se aplica el tercer procedimiento, *el análisis de la variabilidad de las diferencias y cocientes estacionales*. Para ello a continuación se definen los conceptos de diferencia estacional, $d_{t,i}$, y cociente estacional, $c_{t,i}$.

$$d_{t,i} = Y_{t,i} - Y_{t-1,i}$$

$$c_{t,i} = \frac{Y_{t,i}}{Y_{t-1,i}}$$

donde $Y_{t,i}$ es la demanda de energía eléctrica en la C.A.A. en el año *t-ésimo* del mes *i-ésimo*. *t*=1995, 1996, ..., 2000. *i*= Enero, Febrero, ..., Diciembre.

Una vez calculados las diferencias y cocientes estacionales, se calculan la media y desviación típica de ambos y seguidamente se halla el coeficiente de variación tanto para la diferencia como para el cociente, tabla IV.2.2 resultando para el caso analizado:

Tabla.IV.2.2. Media y desviación típica de las diferencias estacionales y cocientes estacionales.

	Diferencias estacionales	Cocientes estacionales
Desviación Típica	38,856	1,076
Media	36,324	0,078
Coeficiente de Variación	1,070	0,072

Fuente. IEA.
Elaboración. Propia.

$$CV(d) = \frac{Desviación\ típica(d)}{Media(d)} = 1{,}070$$

$$CV(c) = \frac{Desviación\ típica(c)}{Media(c)} = 0{,}072$$

BIBLIOTECA VIRTUAL

Por este tercer procedimiento es más evidente el hecho de que el esquema que debe ser seleccionado ha de ser el esquema multiplicativo, ya que los resultados muestran que $CV(c) < CV(d)$. De esta forma, la búsqueda de un modelo que se ajuste a los datos de demanda de electricidad se restringirá a los modelos con este tipo de asociación entre los componentes que lo integran.

Con el objeto de ilustrar estos resultados, en las figuras IV.2.4 y IV.2.5 se han representado las diferencias y cocientes de estos datos.

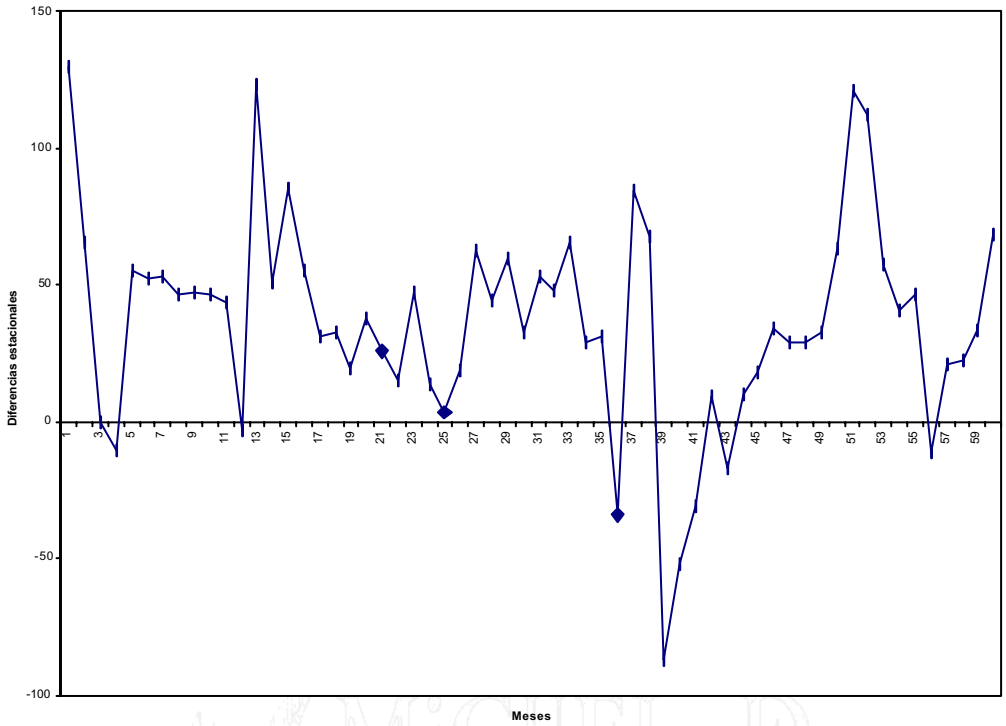


Figura.IV.2.4. Diferencias estacionales de la demanda mensual de energía eléctrica en la C.A.A. durante el periodo 1995-2000.
Fuente.IEA.
Elaboración. Propia.

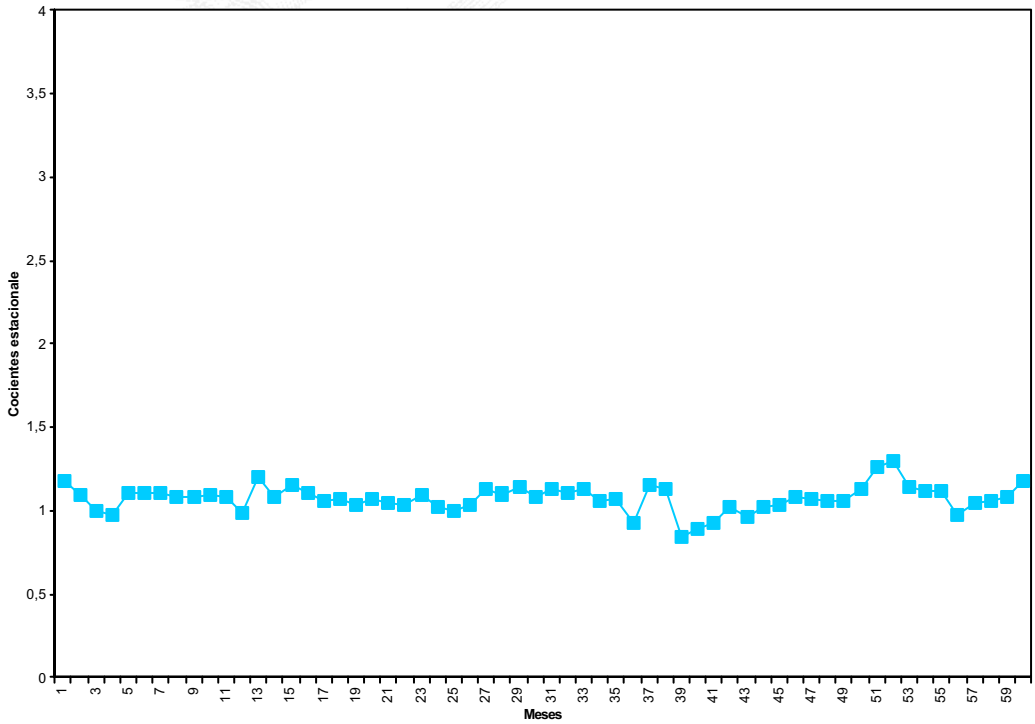


Figura.IV.2.5. Cocientes estacionales de la demanda mensual de energía eléctrica en la C.A.A. durante el periodo 1995-2000.
Fuente. IEA.
Elaboración. Propia.

El análisis de los modelos descritos anteriormente, se desarrolla a continuación expresando su forma funcional, así como los estadísticos principales que los describen. También han sido tratados algunos modelos de *ajuste de curvas*, como son el modelo lineal, cuadrático, exponencial y de crecimiento.

Modelo Ingenuo-Paseo Aleatorio

Forma funcional: $\hat{Y}_t = Y_{t-1}$

Modelo de Medias Móviles

Respecto a este modelo se han probado la media de orden 11, 12 y 13

$$MM(11)_t = \frac{(Y_{t-5} + Y_{t-1} + \ldots + Y_{t+5})}{11}$$

$$MM(12)_t = \frac{\left(Y_{\frac{t}{2}-5} + Y_{t-1} + \ldots + Y_{\frac{t}{2}+5}\right)}{12}$$

$$MM(13)_t = \frac{(Y_{t-6} + Y_{t-1} + \ldots + Y_{t+6})}{13}$$

Modelo de Alisado Exponencial, sin tendencia y estacionalidad de carácter multiplicativo

Forma Funcional: $L_t = 0{,}33067\frac{Y_t}{S_{t-12}} + (1-0{,}33067)L_{t-1}$ (ecuación de nivel)

$S_t = 0{,}99937\frac{Y_t}{L_t} + (1-0{,}99937)S_{t-12}$ (ecuación de factor estacional)

Para iniciar el proceso es preciso tener los valores iniciales de L_t y S_t. En la tabla IV.2.3 se expresa los doce índices de estacionalidad del modelo.

Tabla.IV.2.3. Índices Estacionales. Modelo de Alisado Exponencial.

Mes	S_t
Enero	1,50980
Febrero	1,25594
Marzo	1,10838
Abril	0,92913
Mayo	0,96564
Junio	0,86613
Julio	0,90828
Agosto	0,96874
Septiembre	1,04603
Octubre	0,88553
Noviembre	0,95644
Diciembre	0,78546

Fuente. IEA.
Elaboración. Propia.

Método Holt-Winters, con componentes de tendencia lineal y estacionalidad de carácter multiplicativo.

Forma funcional:

Ecuación de nivel: $L_t = 0{,}15515 \dfrac{Y_t}{S_{t-12}} + (1 - 0{,}15515)(L_{t-1} + b_{t-1})$

Ecuación de pendiente: $b_t = 0{,}04106(L_t - L_{t-1}) + (1 - 0{,}04106)\ b_{t-1}$

Ecuación del factor estacional: $S_t = 0{,}68216 \dfrac{Y_t}{L_t} + (1 - 0{,}68216)\ S_{t-12}$

Tabla.IV.2.4. Índices Estacionales. Modelo de Holt-Winters.

Mes	S_t
Enero	1,46973
Febrero	1,26168
Marzo	1,12125
Abril	0,93779
Mayo	0,95837
Junio	0,86644
Julio	0,91336
Agosto	0,97235
Septiembre	1,04697
Octubre	0,88353
Noviembre	0,94811
Diciembre	0,79295

Fuente. IEA.
Elaboración. Propia.

Modelos de Ajuste de Curva

Modelo Lineal

Forma funcional: $\hat{Y}_t = 417{,}9694 + 2{,}4367 X_t$, resultando significativos todos los coeficientes de modelo al 95% de confianza, así como el modelo globalmente.

Modelo Cuadrático Negativo

Forma funcional: $\hat{Y}_t = 417{,}1270 + 2{,}5089 X_t - 0{,}0010 X_t^2$, resultando significativos todos lo coeficientes al 95% de confianza, así como el modelo globalmente.

Modelo Exponencial

Forma funcional: $\hat{Y}_t = e^{(6{,}0477 + 0{,}0048 X_t)}$, resultando significativos todos los coeficientes al 95% de confianza, así como el modelo globalmente.

Modelo de crecimiento (curva S)

Forma funcional: $\hat{Y}_t = \dfrac{886{,}8469}{1+e^{-0{,}0113(X_t-10{,}6719)}}$, no resultando significativo ni el término constante, *a*, ni el coeficiente de tendencia, *b*, al 95% de probabilidad, pero en cambio sí lo es el coeficiente de crecimiento, *c*.

En la tabla IV.2.23(a), página 235, se expresan todos los estadísticos representativos desde el punto de vista del grado de explicación de la demanda doméstica de energía eléctrica en la C.A.A. y el error cometido en el ajuste y previsión con el objetivo de la comparar todos los modelos y seleccionar el modelo más adecuado.

A la vista de los resultados, se observa que los modelos candidatos que se disputan la representación de la demanda de energía eléctrica son, el modelo de alisado exponencial con estacionalidad multiplicativa y el modelo con tendencia y estacionalidad multiplicativa de Holt-Winters. El modelo de Holt-Winters incrementa el grado de explicación, $\bar{R}^2$, en tan sólo un 1% respecto al modelo de alisado exponencial mientras en otros estadísticos, que miden la capacidad predictiva tales como MAPE, BIC y RMSE, obtiene valores superiores respecto al modelo de alisado exponencial. A la hora de elegir un modelo también es necesario evaluar su poder predictivo a través de estimaciones, por lo que en las tablas IV.2.5 y IV.2.6, se realizan estimaciones confidenciales, al 5% de significación, para la demanda mensual de electricidad para el año 2001, para cada uno de los dos modelos, sin considerar los modelos de carácter dinámico. Los resultados de dichas estimaciones muestran que, para el modelo alisado exponencial (A.E.), tres de los doce intervalos construidos, correspondientes a la demanda de electricidad de los meses de Marzo, Octubre y Diciembre, no contendrían al valor realmente observado. Por otra parte, las estimaciones confidenciales de la demanda mensual de electricidad del modelo de Holt-Winters para el año 2001 dejan fuera al verdadero valor de la demanda en los meses de Marzo, Septiembre, Octubre y Diciembre. Por lo tanto, se recomienda el modelo de Alisado Exponencial, con estacionalidad de carácter multiplicativo, dentro de la clase de modelos no dinámicos, como modelo adecuado para explicar la variación de la

demanda residencial de electricidad en C.A.A. así como para predecir su demanda en periodos futuros. Según la descomposición clásica aplicada al modelo A.E., la variación de la demanda de electricidad en C.A.A. se debe en un 39,33% al componente tendencia ciclo, en un 52,82% al componente estacional y tan sólo un 7,85% al componente irregular.

Tabla.IV.2.5. Intervalos confidenciales (95%) del modelo de A.E.

A. E.	L. I. (2,5%)	Estimación	L.S.(97,5%)	Valor Real
Enero	773,183	883,848	994,513	910,870
Febrero	636,171	735,235	834,298	744,322
Marzo	554,694	648,853	743,012	801,157
Abril	456,594	543,917	631,24	512,637
Mayo	473,074	565,292	657,511	582,845
Junio	417,193	507,041	596,89	556,618
Julio	437,293	531,713	626,133	621,894
Agosto	467,058	567,105	667,152	634,137
Septiembre	505,594	612,352	719,111	564,881
Octubre	416,54	518,396	620,252	673,151
Noviembre	452,498	559,909	667,32	606,917
Diciembre	357,468	459,812	562,156	624,039

Fuente. IEA.
Elaboración. Propia.

Tabla. IV.2.6. Intervalos confidenciales (95%) del modelo de Holt-Winters.

Holt-Winters	L. I. (2,5%)	Estimación	L.S.(97,5%)	Valor Real
Enero	773,657	887,068	1000,479	910,870
Febrero	665,912	764,977	864,043	744,322
Marzo	593,096	682,922	772,747	801,157
Abril	496,215	573,77	651,325	512,637
Mayo	509,103	589,001	668,899	582,845
Junio	460,542	534,89	609,237	556,618
Julio	488,008	566,375	644,742	621,894
Agosto	522,368	605,635	688,901	634,137
Septiembre	565,679	654,995	744,311	564,881
Octubre	476,007	555,184	634,361	673,151
Noviembre	514,146	598,375	682,604	606,917
Diciembre	427,633	502,638	577,642	624,039

Fuente. IEA.
Elaboración. Propia

En la figura IV.2.6, se representa gráficamente el ajuste del modelo A.E. para el periodo 1995-2000, así como las estimaciones e intervalos de confianza (95% de probabilidad) para los doce meses del año 2001.

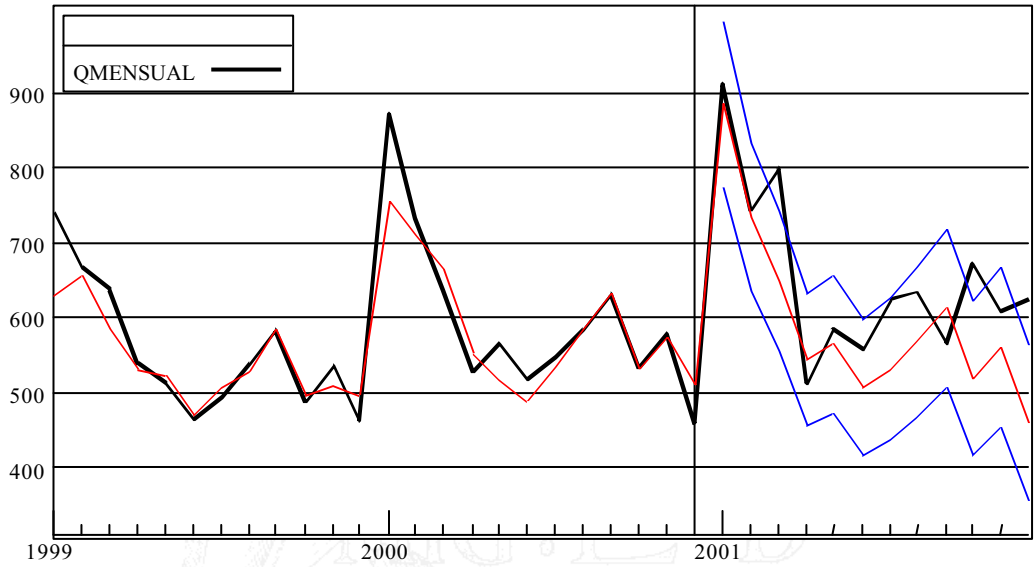


Figura.IV.2.6.Valores reales, ajustados y límites de confianza de demanda mensual de electricidad en la C.A.A. Modelo A.E. Periodo 1995-2000.
Fuente. IEA.
Elaboración.Propia.

Por último, este apartado se dedica a la modelización de la demanda de electricidad en la C.A.A. según la metodología ARIMA, ya descrita anteriormente. La serie cronológica estudiada es la ya definida por QMENSUAL que abarca los periodos de demanda mensual de electricidad desde 1985 a 2001, los cuales serán divididos en dos periodos. El primer período, al que se denominará *histórico,* comprende desde enero de 1995 hasta diciembre de 2000, utilizando las observaciones correspondientes a este intervalo temporal para modelar la serie ARIMA, es decir para efectuar las etapas de identificación de la serie, estimación de sus parámetros y realizar los contrastes-diagnóstico que permitan, en su caso, validar el modelo *dentro de la muestra*. El segundo período, que abarca desde enero de 2001 hasta diciembre de 2001, servirá para verificar el modelo *fuera de la muestra*, es decir, para realizar una evaluación *ex-post* de la predicción, basada en el modelo construido con los datos del período anterior, y compararla con los

valores realmente observados en este segundo período. A continuación se comenta el ciclo Identificación-Estimación-Diagnóstico de los modelos candidatos.

Identificación y Estimación de modelos

En la figura IV.2.2 ya se observó que, con periodicidad constante, los máximos anuales corresponden al último mes del año o al primer mes del año siguiente, épocas de mayor rigor invernal y, por tanto, de mayor consumo debido a la calefacción, existiendo anualmente otra punta de demanda, de mucha menor intensidad que el máximo anual mencionado, que corresponde a la mayor demanda estival debido al aire acondicionado. También se observa en el mismo gráfico que regularmente se producen valores muy inferiores a la media anual, los cuales corresponden a las estaciones de primavera y otoño, en los que la mayor suavidad de las temperaturas induce un menor consumo eléctrico.

El gráfico de la serie original QMENSUAL (figura.IV.2.2) sugiere la no estacionariedad en media de la serie, dado que ésta muestra una clara tendencia al alza, lo cual queda corroborado por la función de autocorrelación de la serie original, ya que los elevados valores de las autocorrelaciones que aparecen en la figura IV.2.7 no tienen tendencia a disminuir rápidamente. Las bandas de confianza de las autocorrelaciones se pueden establecer, en primera aproximación y de forma muy conservadora, para los datos del periodo *histórico* utilizado para modelizar, en $\pm\frac{2}{\sqrt{72}}$=0.2357, aunque realmente, de forma más precisa deberían calcularse mediante la expresión:

$$\pm 2\sqrt{\frac{1}{72}\left(1+2\sum_{j=1}^{k-1} r_j^2\right)}$$

siendo r_j el valor de la autocorrelación correspondiente al retardo de orden *j*.

Al mismo tiempo los *p-valores* que aparecen en esta figura muestran valores muy significativos, no apoyando la hipótesis de incorrelación serial.

Tabla.IV.2.7. Correlograma de los residuos. QMENSUAL.

Sample: 1995:01 2000:12
Included observations: 72

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. |**** |	. |**** |	1	0.579	0.579	25.128	0.000
. |*** |	. | . |	2	0.338	0.005	33.838	0.000
. | . |	**| . |	3	0.049	-0.223	34.025	0.000
. |*. |	. |*** |	4	0.153	0.330	35.860	0.000
. |*. |	. | . |	5	0.144	-0.020	37.507	0.000
. |** |	. |*. |	6	0.233	0.074	41.883	0.000
. |*. |	. | . |	7	0.132	-0.012	43.317	0.000
. |*. |	. | . |	8	0.120	0.010	44.508	0.000
. | . |	.*| . |	9	-0.007	-0.083	44.512	0.000
. |** |	. |*** |	10	0.200	0.353	47.946	0.000
. |** |	. |*. |	11	0.311	0.152	56.411	0.000
. |**** |	. |*** |	12	0.579	0.341	86.147	0.000
. |** |	***| . |	13	0.298	-0.353	94.176	0.000
. |*. |	.*| . |	14	0.115	-0.121	95.385	0.000
.*| . |	. |*. |	15	-0.066	0.138	95.794	0.000
. | . |	.*| . |	16	0.031	-0.110	95.886	0.000
. | . |	. | . |	17	0.037	-0.044	96.022	0.000
. |*. |	. | . |	18	0.087	-0.055	96.777	0.000
. | . |	. | . |	19	-0.004	-0.025	96.778	0.000
. | . |	. | . |	20	-0.022	-0.038	96.826	0.000
.*| . |	. | . |	21	-0.108	0.033	98.047	0.000
. | . |	.*| . |	22	0.020	-0.110	98.090	0.000
. |*. |	. |*. |	23	0.146	0.176	100.40	0.000
. |** |	. | . |	24	0.326	-0.015	112.17	0.000
. |*. |	. | . |	25	0.161	-0.046	115.11	0.000
. | . |	. | . |	26	0.013	0.061	115.13	0.000
.*| . |	.*| . |	27	-0.145	-0.169	117.62	0.000
.*| . |	. | . |	28	-0.095	-0.023	118.72	0.000
.*| . |	. | . |	29	-0.096	-0.024	119.87	0.000
. | . |	. | . |	30	-0.056	-0.048	120.27	0.000
.*| . |	.*| . |	31	-0.129	-0.066	122.43	0.000
.*| . |	. | . |	32	-0.151	0.005	125.48	0.000
**| . |	.*| . |	33	-0.202	-0.063	131.07	0.000
.*| . |	. | . |	34	-0.106	0.052	132.66	0.000
. | . |	.*| . |	35	-0.005	-0.063	132.66	0.000
. |*. |	. | . |	36	0.134	-0.008	135.33	0.000

Fuente. IEA.
Elaboración. Propia.

Por otra parte se observa en la figura IV.2.2 que la amplitud o dispersión de la serie aumenta a medida que transcurre el tiempo, o si se quiere, conforme aumenta el nivel de la serie, lo que también queda corroborado por el gráfico desviación-típica media (figura IV.2.3). Esta no estacionariedad en varianza de la serie original sugiere la aplicación de la transformación de Box-Cox de parámetro $\lambda = 0$ pudiendo verse en la figura IV.2.7 la transformación logarítmica de la serie original. La representación de la transformada logarítmica de la serie original, serie a la que denominaremos *lnq* (LQ), cuyo gráfico desviación típica-media (figura IV.2.8) confirma la conveniencia de la transformación realizada. La

disminución de la heterocedasticidad lograda por la transformación logarítmica aplicada es evidente, no sólo porque la desviación típica de la serie *lnq* es 0,174049 frente al valor 93,80956 que tenía en la serie original QMENSUAL tal como puede verse, respectivamente, en las figuras IV.2.9 y IV.2.10, lo cual podría atribuirse al lógico menor recorrido de la serie *lnq,* sino porque el coeficiente de variación de la serie *lnq* es aproximadamente la novena parte del de la serie original QMENSUAL (0.0280 frente a 0.1860).

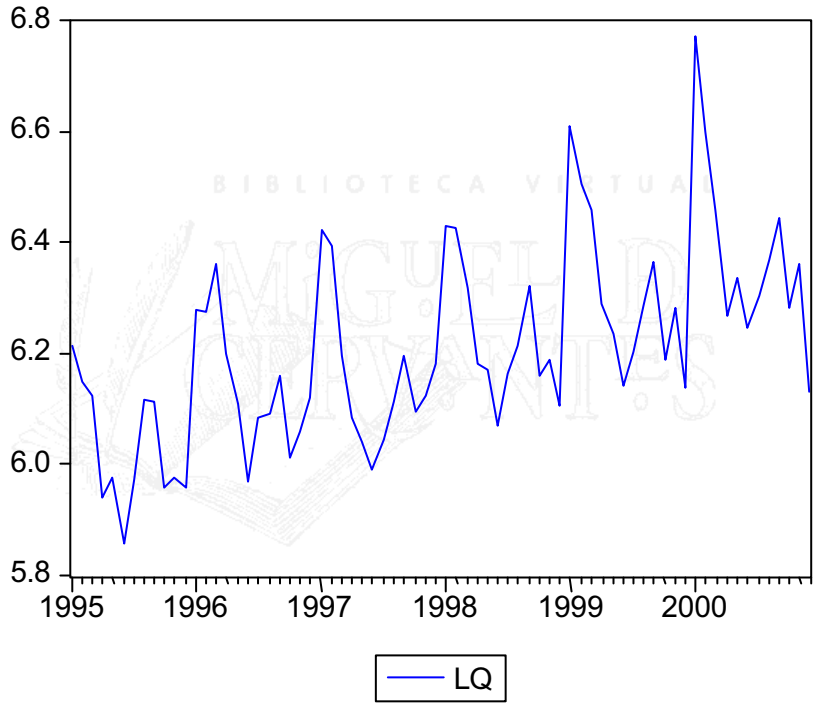


Figura.IV.2.7. Evolución demanda mensual de electricidad en la C.A.A. durante el periodo 1995-2000.
Fuente.IEA.
Elaboración. Propia.

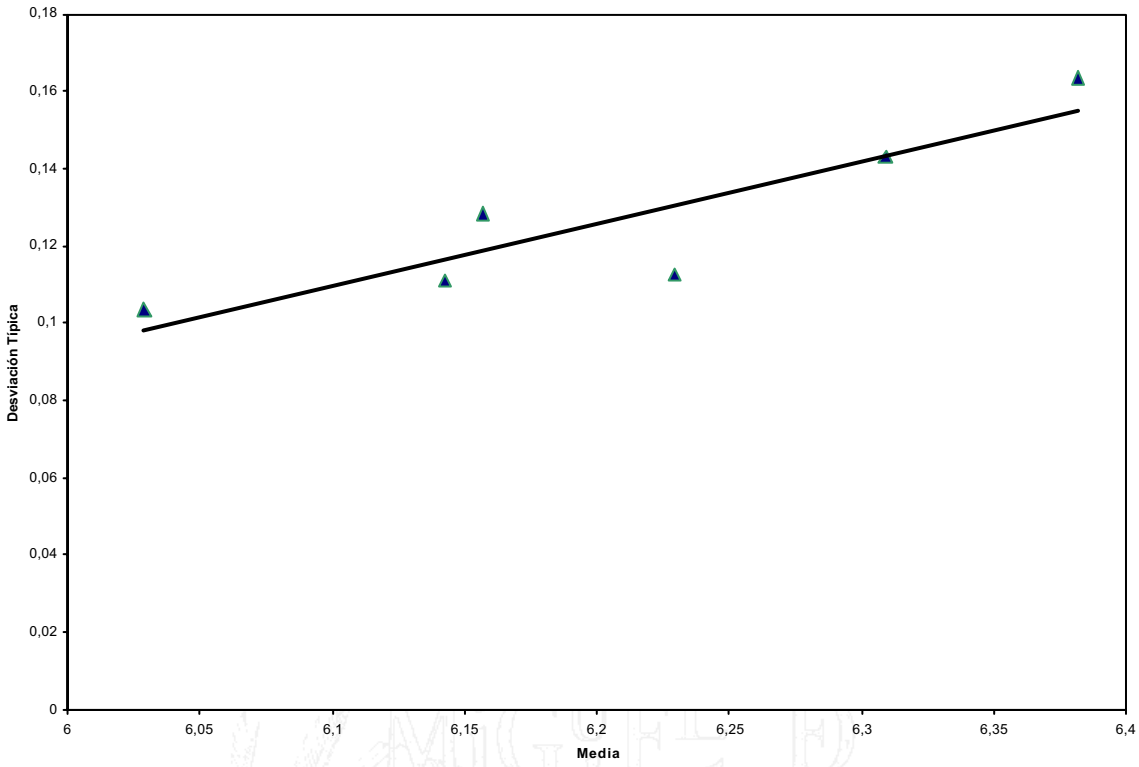


Figura.IV.2.8.Gráfico desviación típica-media. Demanda mensual de energía eléctrica durante el periodo 1995-2000.
Fuente.IEA.
Elaboración. Propia.

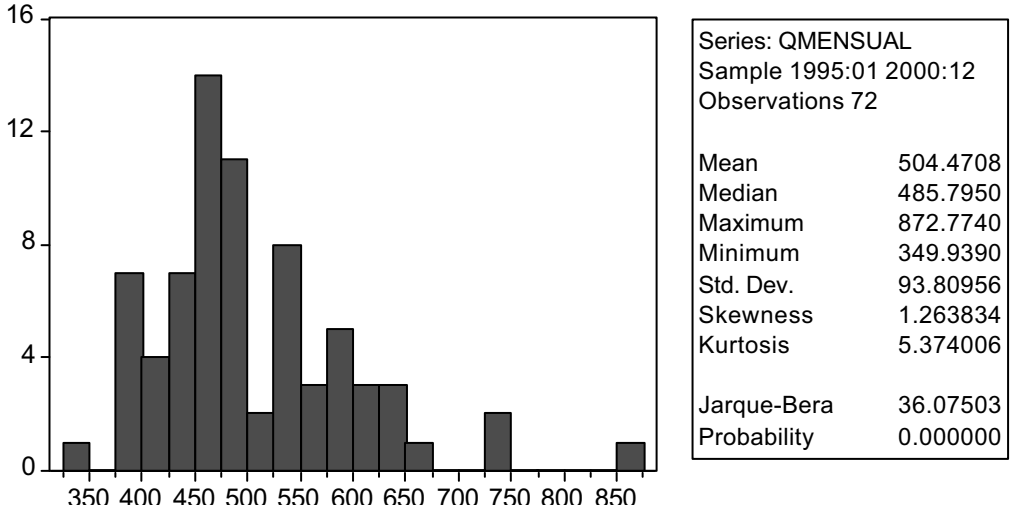


Figura.IV.2.9.Histograma de probabilidad.QMENSUAL.
Fuente.IEA.
Elaboración. Propia.

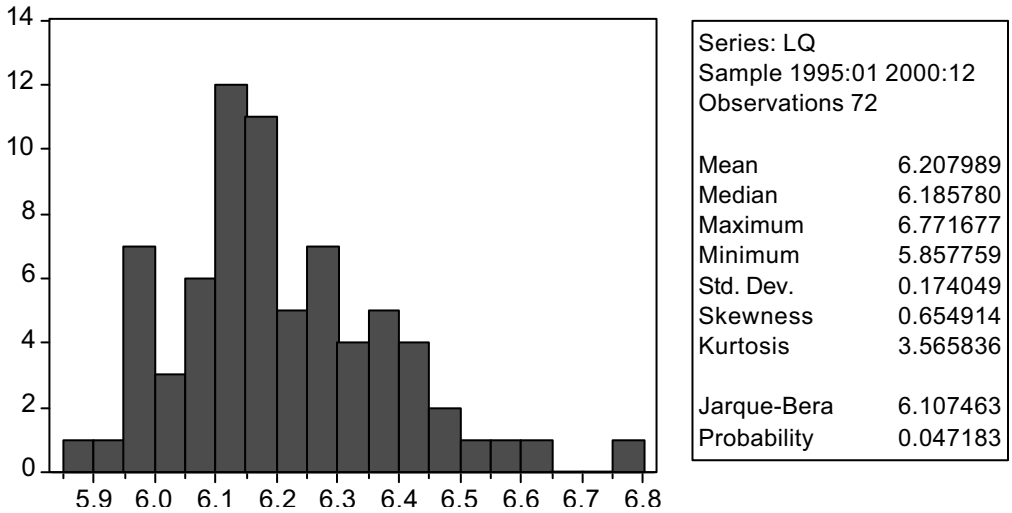


Figura.IV.2.10. Histograma de probabilidad. LQ.
Fuente. IEA.
Elaboración. Propia.

Las autocorrelaciones de la serie *lnq* aparecen en la tabla IV.2.8 en la que pueden observarse, entre otros, los elevados valores correspondientes a los retardos de orden 1, 2, 6, 12 y 24, reflejando estos últimos la estacionalidad de periodicidad 12, correspondiente a una serie de datos mensuales como la que se analiza. Lo comentado anteriormente junto con la tendencia creciente que se observa en la serie *lnq* evidencian la no estacionariedad de esta última, lo que sugiere la conveniencia de aplicar la diferencia regular de orden uno a la serie *lnq* a fin de construir una nueva serie, ∇ *lnq* de media más estable.

Tabla.IV.2.8. Correlograma de los residuos. LQ.

Sample: 1995:01 2000:12
Included observations: 72

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. |***** |	. |***** |	1	0.616	0.616	28.466	0.000
. |*** |	. | . |	2	0.377	-0.003	39.297	0.000
. |*. |	**| . |	3	0.078	-0.246	39.767	0.000
. |*. |	. |*** |	4	0.175	0.381	42.155	0.000
. |*. |	. | . |	5	0.179	-0.010	44.714	0.000
. |** |	. |*. |	6	0.282	0.073	51.153	0.000
. |*. |	.*| . |	7	0.156	-0.061	53.143	0.000
. |*. |	. | . |	8	0.127	0.024	54.493	0.000
. | . |	.*| . |	9	-0.004	-0.067	54.495	0.000
. |** |	. |*** |	10	0.200	0.354	57.935	0.000
. |** |	. |*. |	11	0.320	0.162	66.905	0.000
. |**** |	. |** |	12	0.578	0.301	96.531	0.000
. |** |	***| . |	13	0.318	-0.386	105.67	0.000
. |*. |	.*| . |	14	0.134	-0.113	107.33	0.000
.*| . |	. |*. |	15	-0.058	0.172	107.65	0.000
. | . |	.*| . |	16	0.042	-0.110	107.82	0.000
. | . |	.*| . |	17	0.059	-0.068	108.15	0.000
. |*. |	.*| . |	18	0.120	-0.059	109.58	0.000
. | . |	. | . |	19	0.010	0.022	109.59	0.000
. | . |	. | . |	20	-0.018	-0.034	109.62	0.000
.*| . |	. | . |	21	-0.108	0.034	110.83	0.000
. | . |	.*| . |	22	0.025	-0.084	110.90	0.000
. |*. |	. |*. |	23	0.151	0.165	113.39	0.000
. |*** |	. | . |	24	0.338	0.006	126.06	0.000
. |*. |	.*| . |	25	0.180	-0.075	129.72	0.000
. | . |	. | . |	26	0.026	0.036	129.80	0.000
.*| . |	.*| . |	27	-0.149	-0.165	132.42	0.000
.*| . |	. | . |	28	-0.096	-0.031	133.53	0.000
.*| . |	. | . |	29	-0.093	-0.049	134.59	0.000
. | . |	. | . |	30	-0.041	0.001	134.81	0.000
.*| . |	.*| . |	31	-0.132	-0.099	137.07	0.000
.*| . |	. | . |	32	-0.161	-0.008	140.50	0.000
**| . |	. | . |	33	-0.222	-0.022	147.24	0.000
.*| . |	. | . |	34	-0.124	0.019	149.40	0.000
. | . |	. | . |	35	-0.026	-0.054	149.49	0.000
. |*. |	. | . |	36	0.124	-0.023	151.76	0.000

Fuente. IEA.
Elaboración. Propia

La figura IV.2.11 muestra el gráfico de la serie ∇lnq, denominada DLQ, en la que si bien su media es aproximadamente cero (en la figura IV.2.12 puede verse que su valor es -0.001176) siguen apareciendo puntas de periodicidad 12 por las razones ya comentadas, lo cual aparece corroborado en el correlograma de la figura IV.2.9 en la que puede observarse que las autocorrelaciones de orden 12, 24 y 36 son estadísticamente significativas lo cual, a su vez, queda reflejado en los *p-valores* de tal figura.

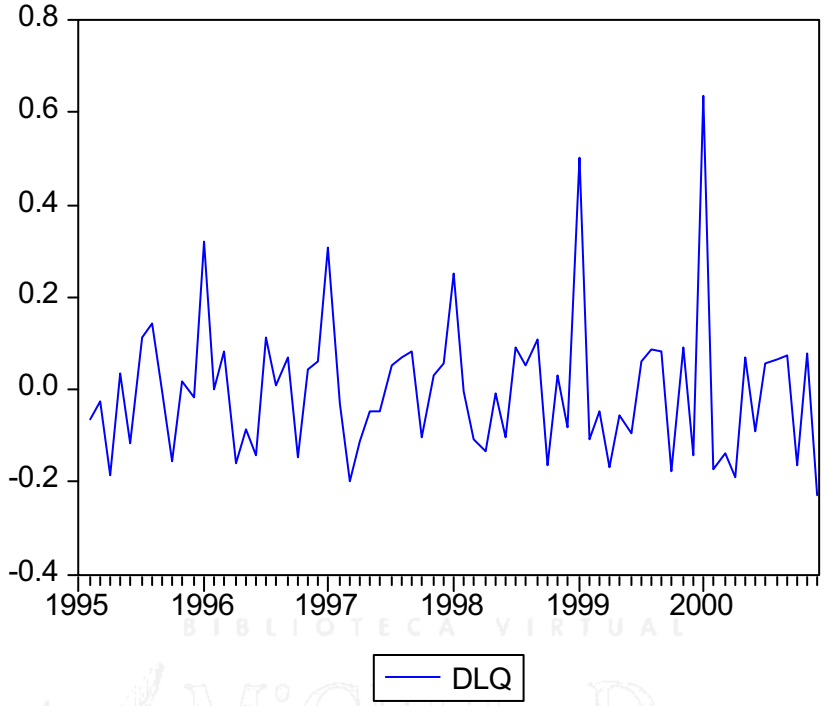


Figura.IV.2.11. Evolución de la demanda mensual de electricidad en la C.A.A. durante el periodo 1995-2000.
Fuente. IEA.
Elaboración. Propia.

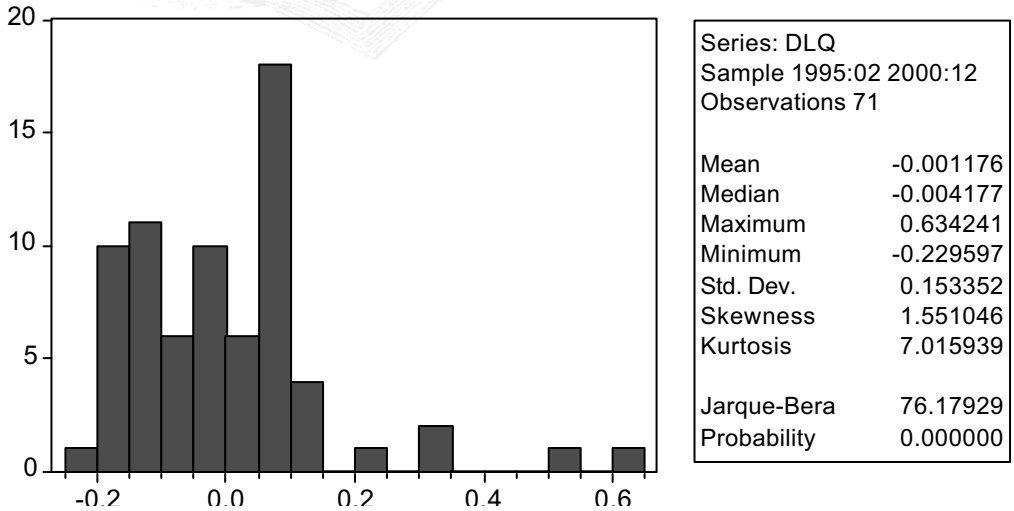


Figura.IV.2.12. Histograma de probabilidad.DLQ.
Fuente. IEA.
Elaboración. Propia.

Tabla.IV.2.9. Correlograma de los residuos.DLQ.

Sample: 1995:01 2000:12
Included observations: 71

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
.*| . |	.*| . |	1	-0.179	-0.179	2.3683	0.124
. |*. |	. | . |	2	0.075	0.045	2.7964	0.247
****| . |	****| . |	3	-0.508	-0.504	22.474	0.000
. |*. |	.*| . |	4	0.116	-0.064	23.521	0.000
.*| . |	.*| . |	5	-0.131	-0.146	24.862	0.000
. |** |	. | . |	6	0.297	0.004	31.907	0.000
.*| . |	.*| . |	7	-0.124	-0.082	33.161	0.000
. |*. |	. | . |	8	0.132	0.002	34.592	0.000
***| . |	***| . |	9	-0.429	-0.377	50.005	0.000
. |*. |	.*| . |	10	0.117	-0.139	51.162	0.000
.*| . |	**| . |	11	-0.170	-0.254	53.655	0.000
. |***** |	. |*** |	12	0.644	0.376	90.070	0.000
.*| . |	. |*. |	13	-0.092	0.095	90.827	0.000
. | . |	.*| . |	14	0.007	-0.178	90.831	0.000
***| . |	. |*. |	15	-0.373	0.081	103.73	0.000
. |*. |	. | . |	16	0.105	0.036	104.77	0.000
.*| . |	. | . |	17	-0.061	0.022	105.12	0.000
. |** |	. | . |	18	0.221	-0.051	109.88	0.000
.*| . |	. | . |	19	-0.103	0.001	110.93	0.000
. |*. |	. | . |	20	0.082	-0.056	111.62	0.000
**| . |	. | . |	21	-0.282	0.049	119.84	0.000
. | . |	.*| . |	22	0.012	-0.187	119.86	0.000
.*| . |	. | . |	23	-0.075	-0.027	120.46	0.000
. |*** |	. |*. |	24	0.426	0.068	140.46	0.000
. | . |	. | . |	25	-0.002	-0.044	140.46	0.000
. | . |	. |*. |	26	0.027	0.125	140.54	0.000
**| . |	. | . |	27	-0.290	-0.001	150.41	0.000
. | . |	. | . |	28	0.061	0.013	150.86	0.000
.*| . |	. | . |	29	-0.066	-0.027	151.40	0.000
. |*. |	. | . |	30	0.182	0.051	155.59	0.000
.*| . |	. | . |	31	-0.077	-0.039	156.35	0.000
. | . |	. | . |	32	0.043	-0.021	156.60	0.000
**| . |	. | . |	33	-0.202	-0.055	162.15	0.000
. | . |	. | . |	34	0.005	0.007	162.15	0.000
.*| . |	. | . |	35	-0.067	-0.027	162.81	0.000
. |*** |	. | . |	36	0.359	0.057	181.84	0.000

Fuente. IEA.
Elaboración. Propia.

Tales autocorrelaciones significativas sugieren la aplicación de la diferencia estacional de orden 12 construyendo así la nueva serie $\nabla\nabla_{12} lnq$, serie denominada DDLQ) cuya representación gráfica puede verse en la figura IV.2.13 de media prácticamente nula (-0.001174).

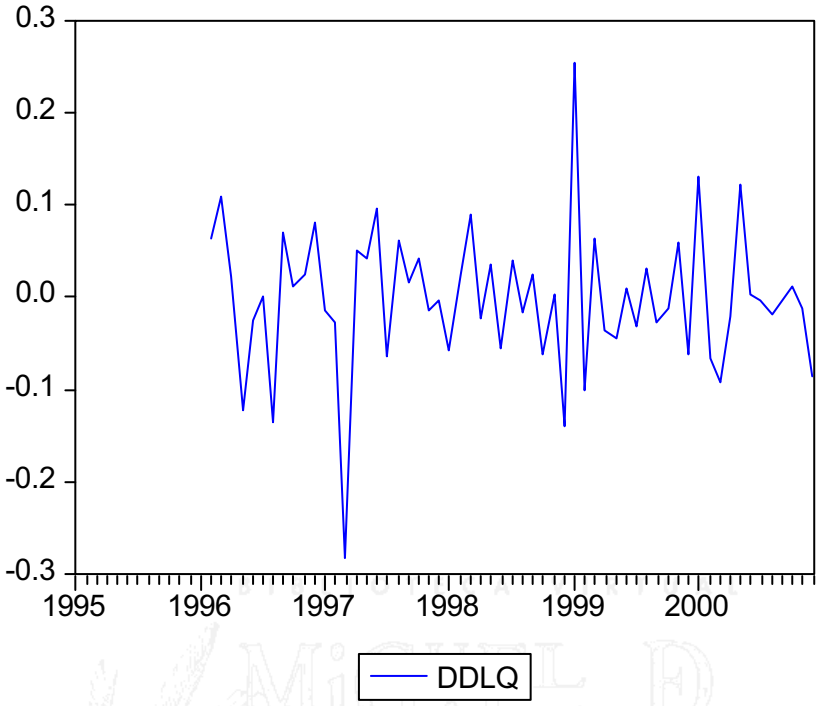


Figura.IV.2.13. Evolución de la demanda mensual de electricidad en la C.A.A. durante el periodo 1995-2000.
Fuente. IEA.
Elaboración. Propia.

En el correlograma de la tabla IV.2.11 aparecen como significativas las autocorrelaciones muestrales de orden 1,12 y 24, por lo que se identifica la serie $\nabla\nabla_{12} lnq$ como MA(1) MA(12) y en definitiva el modelo candidato de la transformada logarítmica de la serie original queda identificado como ARIMA(0,1,1) x $(0,1,1)_{12}$. Se denominará a este modelo BJ-1 del cual, una vez estimados sus coeficientes, aparecen su especificación y algunos estadísticos que permiten diagnosticar este modelo.

$$\nabla\nabla_{12} \ln q_t = (1-0{,}8103B)(1-0{,}7480B^{12})\varepsilon_t$$

$$(0{,}0701) \quad (0{,}0655)$$

Tabla.IV.2.10. Estimación de los coeficientes del modelo BJ-1.

Término	Coeficiente	Desviac.Error	Estadístico-t	p-valor
$\hat{\theta}_1$	0,8103	0,0701	11,5634	0,0000
$\hat{\Theta}_{12}$	0,7480	0,0655	11,4178	0,0000

Fuente. IEA.
Elaboración. Propia.

Tabla.IV.2.11. Correlograma de los residuos. DDLQ.

Sample: 1995:01 2000:12
Included observations: 59

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
**| . |	**| . |	1	-0.285	-0.285	5.0427	0.025
. |*. |	. | . |	2	0.066	-0.017	5.3182	0.070
**| . |	**| . |	3	-0.285	-0.295	10.541	0.014
. | . |	.*| . |	4	0.029	-0.157	10.597	0.031
.*| . |	.*| . |	5	-0.068	-0.140	10.910	0.053
. | . |	**| . |	6	-0.006	-0.192	10.912	0.091
. | . |	.*| . |	7	-0.006	-0.153	10.914	0.142
. | . |	.*| . |	8	0.026	-0.128	10.961	0.204
. | . |	.*| . |	9	0.048	-0.085	11.126	0.267
. |** |	. |** |	10	0.269	0.272	16.422	0.088
.*| . |	. |*. |	11	-0.094	0.115	17.083	0.105
.*| . |	.*| . |	12	-0.151	-0.111	18.828	0.093
.*| . |	.*| . |	13	-0.133	-0.063	20.207	0.090
.*| . |	.*| . |	14	-0.077	-0.183	20.681	0.110
. |*. |	.*| . |	15	0.104	-0.077	21.565	0.120
. | . |	. | . |	16	0.028	-0.042	21.631	0.156
. |*. |	.*| . |	17	0.067	-0.078	22.013	0.184
.*| . |	.*| . |	18	-0.074	-0.165	22.492	0.211
. | . |	.*| . |	19	0.046	-0.138	22.684	0.252
. | . |	.*| . |	20	0.052	-0.104	22.934	0.292
. |*. |	. |*. |	21	0.089	0.131	23.683	0.309
.*| . |	. | . |	22	-0.180	0.019	26.838	0.217
. | . |	. | . |	23	-0.029	0.001	26.923	0.259
.*| . |	. | . |	24	-0.060	-0.005	27.295	0.291
. | . |	.*| . |	25	0.057	-0.148	27.643	0.325
. |*. |	. | . |	26	0.099	-0.024	28.707	0.325
. | . |	. | . |	27	0.010	-0.016	28.719	0.375
. |*. |	. |*. |	28	0.079	0.115	29.448	0.390
.*| . |	. | . |	29	-0.107	-0.006	30.827	0.374
. | . |	.*| . |	30	0.029	-0.105	30.934	0.419
. | . |	. | . |	31	0.032	-0.001	31.062	0.463
.*| . |	.*| . |	32	-0.125	-0.075	33.131	0.412
. | . |	. |*. |	33	0.052	0.116	33.508	0.443
.*| . |	. | . |	34	-0.143	-0.048	36.434	0.356
. |*. |	. | . |	35	0.138	-0.017	39.294	0.283
. |*. |	. | . |	36	0.071	0.018	40.088	0.294

Fuente. IEA.
Elaboración. Propia.

Aunque en la tabla IV.2.11 pueden observarse algunas autocorrelaciones significativas distintas de las apuntadas, no debe olvidarse que trabajando al nivel

de confianza del 95 por ciento, si realmente es cierto que $\rho_k = 0$, la probabilidad de cometer el *error tipo I* es del 5 por ciento, lo que equivale a decir que aproximadamente una de cada veinte autocorrelaciones será declarada significativa cuando verdaderamente no lo es, del mismo modo que una que sea verdaderamente significativa puede ser declarada que no lo es, lo que en la práctica equivale a sospechar de las autocorrelaciones muestrales corespondientes a retardos no típicos. Además, hay que tener en cuenta que en las proximidades de los retardos estacionales pueden detectarse autocorrelaciones muestrales significativas debidas bien al efecto *contagio* de los múltiplos de la estacionalidad, bien a la interacción de dos medias móviles. No obstante, para considerar la posibilidad de que verdaderamente las autocorrelaciones de la serie $\nabla\nabla_{12} lnq$ estuviesen sugiriendo otra estructura, se considerará para la transformada logarítmica de la serie original los modelos alternativos ARIMA (1,1,0) x $(0,1,1)_{12}$ y ARIMA (0,1,2) x $(0,1,0)_{12}$ a los que se denominará BJ-2 y BJ-3, respectivamente.

$$(1+0{,}4694B)\nabla\nabla_{12}\ln q_t = \left(1-0{,}6925B^{12}\right)\varepsilon_t$$

(0,0858) (0,1216)

Tabla.IV.2.12. Estimaciones de los coeficientes del modelo BJ-2.

Término	Coeficiente	Desviac.Error	Estadístico-t	p-valor
$\hat{\phi}_1$	-0,4694	0,0858	-3,8598	0,0003
$\hat{\Theta}_{12}$	0,6925	0,1216	8,0744	0,0000

Fuente. IEA.
Elaboración. Propia.

$$\nabla\nabla_{12}\ln q_t = (1-0{,}6060B)\left(1-0{,}2935B^{12}\right)\varepsilon_t$$

(0,1260) (0,1230)

Tabla.IV.2.13. Estimaciones de los coeficientes del modelo BJ-3.

Término	Coeficiente	Desviac.Error	Estadístico-t	p-valor
b[1]	0,6060	0,1260	4,8079	0,0000
b[2]	0,2935	0,1230	2,3855	0,0204

Fuente. IEA.
Elaboración. Propia

Por otra parte, si para recoger el efecto estacional de la serie *lnq* se hubiese comenzado por diferenciar estacionalmente esta serie se habría obtenido la nueva serie ∇_{12} *lnq*, denominada D12LQ, cuya representación gráfica aparece en la figura IV.2.14. En esta figura se puede observar que la media de dicha serie es prácticamente nula (0.0705 según puede verse en la figura V.2.15). En la figura IV.2.16 se muestra el correlograma de esta serie que parece indicar un modelo AR(1) o AR(2), lo que llevaría a considerar para la transformada logarítmica de la serie original sendos modelos ARIMA(1,0,0)x(0,1,0)$_{12}$ y ARIMA(2,0,0)x(0,1,0)$_{12}$ a los que se denomina BJ-4 y BJ-5, respectivamente.

$$(1-0{,}7114B)\nabla_{12}\ln q_t = \varepsilon_t$$

(0,0899)

Tabla.IV.2.14. Estimación de los coeficientes. Modelo BJ-4.

Término	Coeficiente	Desviac.Error	Estadístico-t	P-valor
$\hat{\phi}_1$	0,7114	0,0899	7,9114	0,0000

Fuente. IEA.
Elaboración. Propia.

$$(1-0{,}5959B-0{,}1679B^2)\nabla_{12}\ln q_t = \varepsilon_t$$

(0,1279) (0,1269)

Tabla.IV.2.15. Estimación de los coeficientes. Modelo BJ-5.

Término	Coeficiente	Desviac.Error	Estadístico-t	P-valor
$\hat{\phi}_1$	0,5959	0,1279	4,6591	0,0000
$\hat{\phi}_2$	0,1679	0,1269	1,3226	0,1911

Fuente. IEA.
Elaboración. Propia.

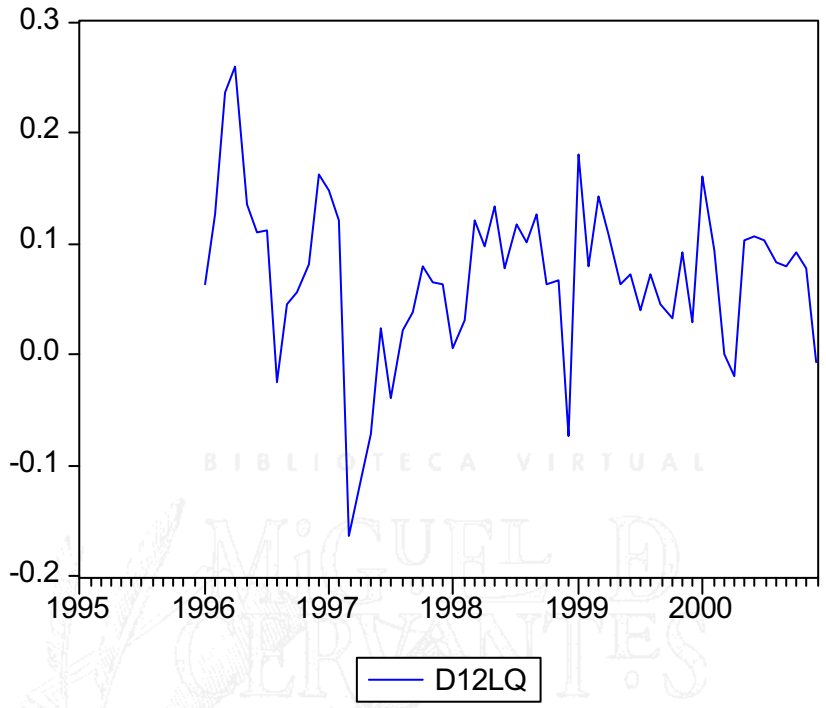


Figura.IV.2.14. Evolución de la demanda mensual de electricidad en la C.A.A. durante el periodo 1995-2000.
Fuente. IEA.
Elaboración. Propia.

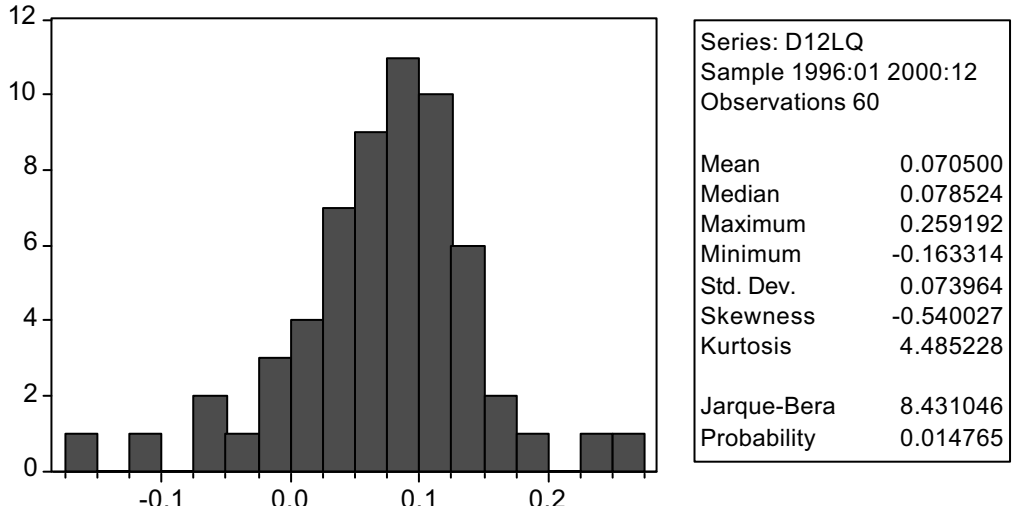


Figura.IV.2.15. Histograma de probabilidad. D12LQ.
Fuente. IEA.
Elaboración. Propia.

Tabla.IV.2.16. Correlograma de los residuos. D12LQ.

Sample: 1995:01 2000:12
Included observations: 60

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. |*** |	. |*** |	1	0.437	0.437	12.063	0.001
. |** |	. | . |	2	0.212	0.025	14.943	0.001
.*| . |	**| . |	3	-0.081	-0.226	15.376	0.002
.*| . |	. | . |	4	-0.063	0.060	15.636	0.004
.*| . |	. | . |	5	-0.078	-0.018	16.050	0.007
. | . |	. | . |	6	-0.014	0.001	16.064	0.013
. | . |	. |*. |	7	0.057	0.085	16.291	0.023
. |*. |	. |*. |	8	0.131	0.080	17.514	0.025
. |*. |	. | . |	9	0.148	0.049	19.115	0.024
. |*. |	. | . |	10	0.118	0.027	20.143	0.028
.*| . |	***| . |	11	-0.187	-0.322	22.808	0.019
***| . |	**| . |	12	-0.370	-0.243	33.435	0.001
***| . |	.*| . |	13	-0.418	-0.108	47.293	0.000
**| . |	.*| . |	14	-0.306	-0.112	54.841	0.000
.*| . |	. | . |	15	-0.128	0.013	56.199	0.000
.*| . |	.*| . |	16	-0.061	-0.080	56.513	0.000
. | . |	.*| . |	17	-0.017	-0.077	56.538	0.000
. | . |	.*| . |	18	-0.052	-0.061	56.779	0.000
. | . |	. |*. |	19	0.001	0.074	56.779	0.000
. | . |	. |*. |	20	0.002	0.095	56.779	0.000
. | . |	. | . |	21	-0.043	0.065	56.958	0.000
.*| . |	.*| . |	22	-0.177	-0.152	60.031	0.000
.*| . |	.*| . |	23	-0.128	-0.076	61.666	0.000
. | . |	. | . |	24	-0.021	-0.044	61.712	0.000
. |*. |	.*| . |	25	0.089	-0.105	62.560	0.000
. |*. |	. | . |	26	0.170	0.060	65.737	0.000
. |*. |	. | . |	27	0.143	-0.030	68.056	0.000
. |*. |	. | . |	28	0.119	-0.029	69.714	0.000
. | . |	.*| . |	29	0.003	-0.163	69.715	0.000
. | . |	. | . |	30	0.007	-0.044	69.721	0.000
. | . |	. | . |	31	-0.030	0.025	69.835	0.000
.*| . |	. | . |	32	-0.089	-0.044	70.883	0.000
. | . |	. | . |	33	-0.018	0.057	70.930	0.000
. | . |	.*| . |	34	-0.002	-0.125	70.930	0.000
. |*. |	. | . |	35	0.151	0.052	74.337	0.000
. |*. |	. | . |	36	0.142	-0.016	77.470	0.000

Fuente. IEA.
Elaboración. Propia.

Debajo de cada coeficiente estimado aparece su correspondiente desviación típica, lo que lleva a la conclusión de que todos los coeficientes estimados son significativos al 95 por ciento, excepto el término autorregresivo de segundo orden del modelo BJ-5; prescindir de este témino equivale a considerar el modelo candidato BJ-4 por lo que en adelante sólo se considerará los modelos candidatos BJ-1, BJ-2, BJ-3 y BJ-4 .

No obstante, se podría haber seguido otro curso de razonamiento alternativo y aplicar el *test de raíces unitarias de Dickey-Fuller Aumentado* a la serie ∇_{12} *lnq* para investigar como de plausible es *a priori* la identificación del modelo BJ-1 ya mencionado. Como puede deducirse de la tabla IV.2.17 siguiente, no se puede rechazar la hipótesis nula (que postula la existencia de una raíz unitaria de dicho serie) al nivel se significación del 1%, ya que el estadístico ADF (-3.775411) es, en valor absoluto, inferior al del valor crítico de MacKinnon (-4.1219). Sin embargo, este mismo test rechaza la existencia de dos raices unitarias (o equivalentemente una diferencia regular de segundo orden la serie ∇_{12} *lnq)* , ya que el estadístico ADF aplicado a la primera diferencia de la serie investigada es, en valor absoluto, superior a los valores críticos de MacKinnon a los niveles de significación del 1%, 5% y 10%, por lo que se obtiene una vía complementaria para ratificar la plausibilidad del modelo propuesto BJ.1.

Tabla.IV.2.17. Test de Raíces Unitarias de nivel de ∇_{12} *lnq.*

ADF Test Statistic	-3.775411	1%	Critical Value*	-4.1219
		5%	Critical Value	-3.4875
		10%	Critical Value	-3.1718

*MacKinnon critical values for rejection of hypothesis of a unit root.
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(D12LQ)
Method: Least Squares
Sample(adjusted): 1996:03 2000:12
Included observations: 58 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D12LQ(-1)	-0.543307	0.143907	-3.775411	0.0004
D(D12LQ(-1))	-0.019044	0.136767	-0.139247	0.8898
C	0.042740	0.027188	1.572041	0.1218
@TREND(1995:01)	-0.000140	0.000540	-0.258568	0.7970

R-squared	0.275189	Mean dependent var	-0.002303
Adjusted R-squared	0.234922	S.D. dependent var	0.078665
S.E. of regression	0.068807	Akaike info criterion	-2.448550
Sum squared resid	0.255658	Schwarz criterion	-2.306450
Log likelihood	75.00795	F-statistic	6.834057
Durbin-Watson stat	1.955984	Prob(F-statistic)	0.000548

Fuente. IEA.
Elaboración. Propia.

Tabla. IV.2.18. Test de las Raíces Unitarias de primera diferencia de ∇_{12} *lnq.*

ADF Test Statistic	-6.232305	1%	Critical Value*	-4.1249
		5%	Critical Value	-3.4889
		10%	Critical Value	-3.1727

*MacKinnon critical values for rejection of hypothesis of a unit root.
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(D12LQ,2)
Method: Least Squares
Sample(adjusted): 1996:04 2000:12
Included observations: 57 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(D12LQ(-1))	-1.351378	0.216834	-6.232305	0.0000
D(D12LQ(-1),2)	0.027190	0.134284	0.202480	0.8403
C	-0.013383	0.028130	-0.475768	0.6362
@TREND(1995:01)	0.000206	0.000611	0.337688	0.7369

R-squared	0.662542	Mean dependent var	-0.003399
Adjusted R-squared	0.643440	S.D. dependent var	0.127040
S.E. of regression	0.075859	Akaike info criterion	-2.252289
Sum squared resid	0.304993	Schwarz criterion	-2.108917
Log likelihood	68.19023	F-statistic	34.68546
Durbin-Watson stat	2.039099	Prob(F-statistic)	0.000000

Fuente. IEA.
Elaboración. Propia.

Contrastes-Diagnósticos

El análisis global de los residuos se realiza mediante el test de Ljung-Box, basado en el estadístico de este mismo nombre, una mejora estadística del test *port manteau* de Box-Pierce (ver tabla IV.2.21):

$$Q = n(n+2)\sum_{i=1}^{L}\frac{r_i^2}{n-1} \square \chi^2_{(L-k)}$$

siendo *n*= el número de puntos de la muestra, *k*= el número de parámetros y *L*= el número de coeficientes de autocorrelación considerados.

Se puede observar en la tabla IV.2.23(b) que ninguno de sus valores muestra un *p-valor* significativo al 95% que induzca a sospechar que su estructura no es la de *ruido blanco* de tales residuos, siendo el modelo BJ-1 el que muestra un valor más favorable para este estadístico. Más selectivo en este aspecto es el análisis individualizado de los residuos que puede observarse en los correlogramas que se muestran en las figuras IV.2.16 a IV.2.19 en las que el

modelo BJ-1 vuelve a ser el que más apoya la estructura de *ruido blanco* para sus residuos.

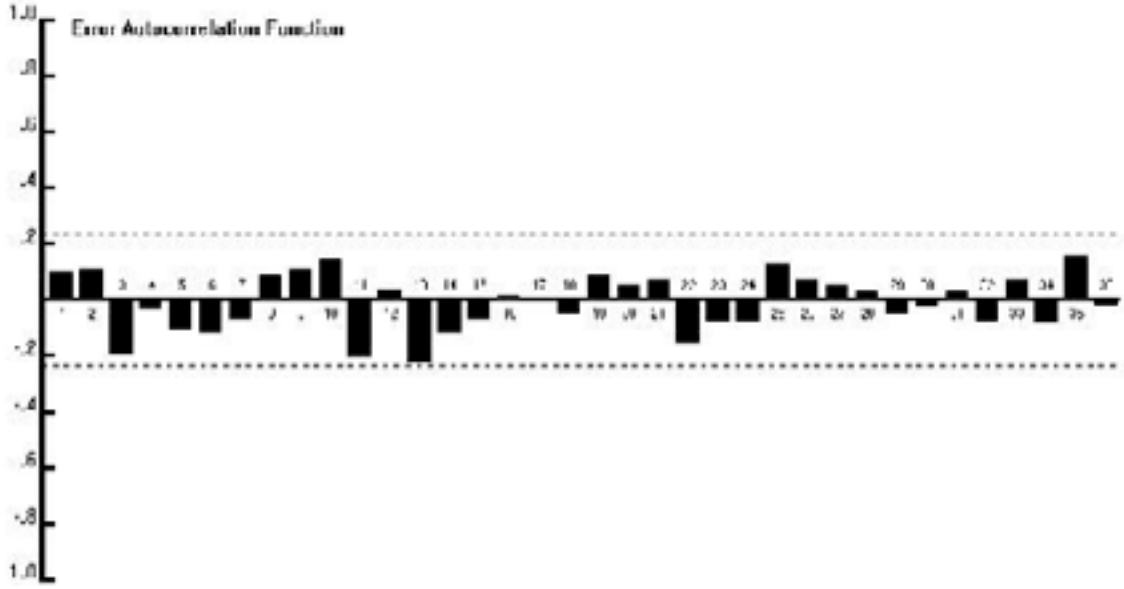


Figura.IV.2.16. Función de autocorrelación del error. Modelo BJ-1.
Fuente. IEA.
Elaboración. Propia.

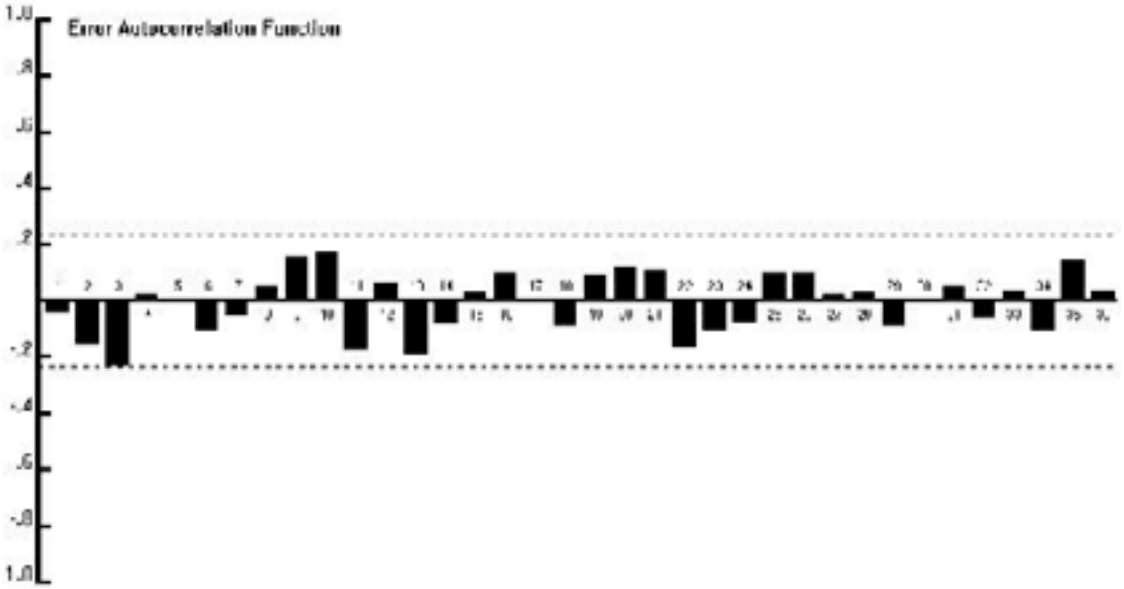


Figura. IV.2.17.Función de autocorrelación del error. Modelo BJ-2.
Fuente. IEA.
Elaboración. Propia.

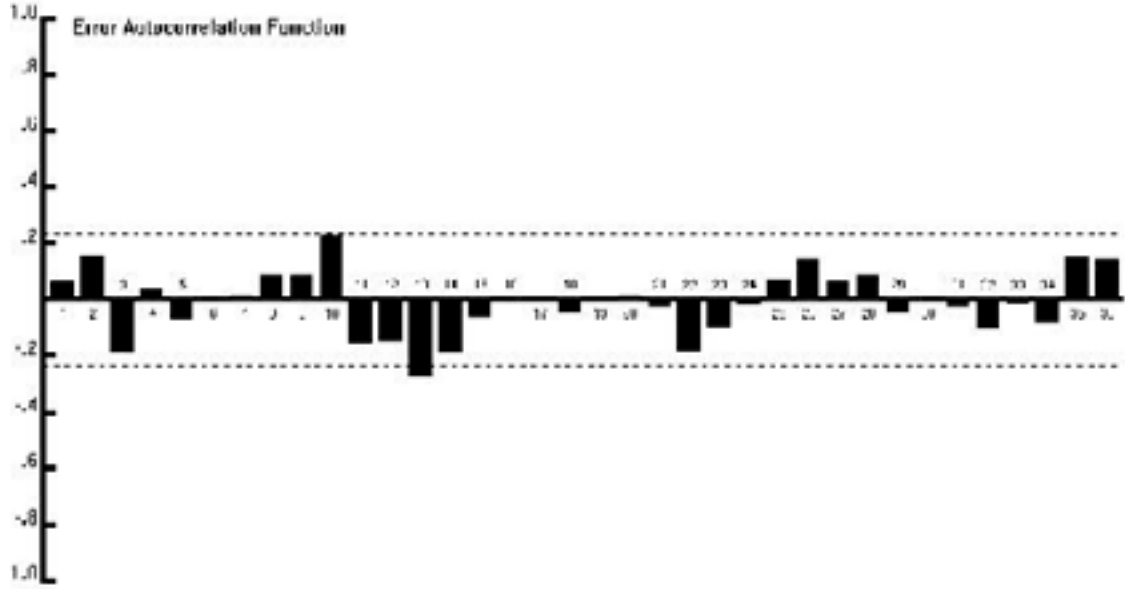


Figura.IV.2.18.Función de autocorrelación del error. Modelo BJ-3.
Fuente. IEA.
Elaboración. Propia.

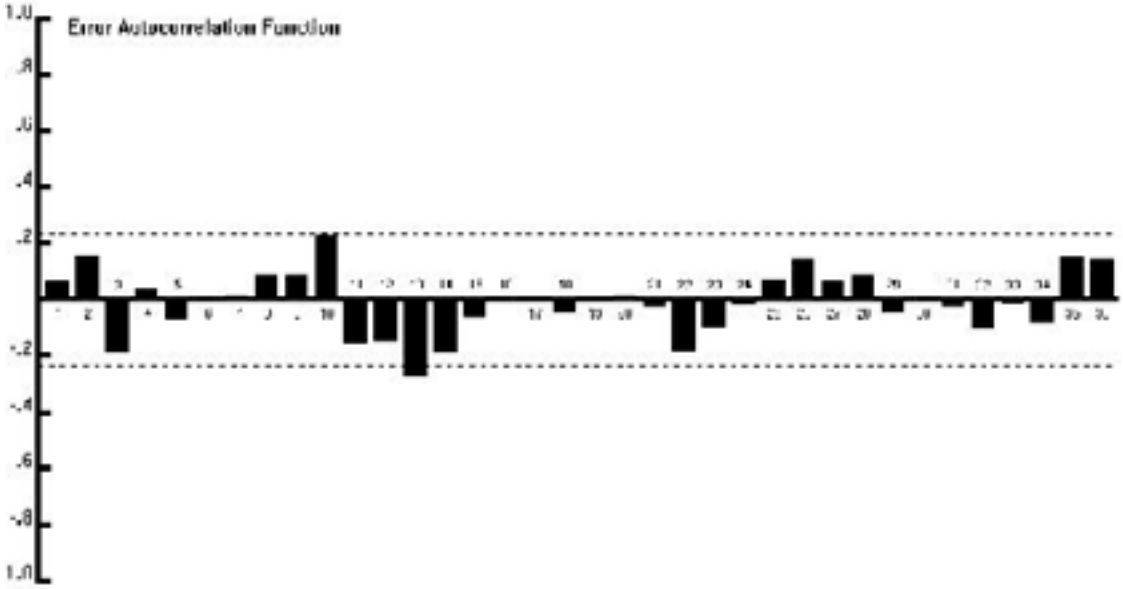


Figura.IV.2.19. Función de autocorrelación del error. Modelo BJ-4.
Fuente. IEA.
Elaboración. Propia.

Jerarquización a priori de los modelos candidatos

Por *jerarquización a priori* se quiere indicar la preferencia que racionalmente podría mostrarse por los modelos candidatos anteriores según el valor de sus parámetros *dentro de la muestra* (es decir, antes de comparar los pronósticos de los modelos con los valores realmente observados en el futuro). La expresión *dentro de la muestra* la utilizamos aquí en sentido amplio, pues ya se ha comentado anteriormente que el objetivo del parámetro BIC es intentar minimizar el error de predicción, aún cuando su valor se calcule en función de los datos históricos. Obviamente la *jerarquización* depende del criterio que para ella se elija y sobre esto no existe consenso generalizado. No obstante, es fácil observar que el modelo que se ha denominado BJ-1 es el *mejor a priori* para cualquier parámetro *dentro de la muestra* que se utilice como evaluación y, en este sentido, se podría decir que *domina a priori* a los demás modelos candidatos considerados.

Predicción

Habiendo estimado los modelos anteriores con los datos del periodo enero de 1995 a diciembre de 2000, denominado anteriormente *histórico,* se efectúa a continuación la predicción *ex-post* al 95% de grado de confianza, para el periodo enero de 2001 a diciembre de 2001 (período que se considerará *futuro*), con cada uno de los cuatro modelos mencionados, las cuales pueden verse en las tablas IV.2.19 a IV.2.22.

Tabla.IV.2.19. Intervalos confidenciales (95%) del modelo BJ-1.

BJ-1	L. I. (2,5%)	Estimación	L.S.(97,5%)	Valor Real
Enero	720,539	813,403	918,237	910,870
Febrero	661,273	748,115	846,360	744,322
Marzo	605,350	686,303	778,082	801,157
Abril	513,177	583,021	662,370	512,637
Mayo	509,720	58,284	660,6171	582,845
Junio	462,302	527,368	601,5910	556,618
Julio	497,087	568,178	649,4376	621,894
Agosto	531,783	609,031	697,5009	634,137
Septiembre	571,470	655,750	752,4614	564,881
Octubre	487,914	560,942	644,8993	673,151
Noviembre	515,920	594,256	684,4873	606,917
Diciembre	467,341	539,302	622,3438	624,039

Fuente. IEA.
Elaboración. Propia.

Tabla.IV.2.20. Intervalos confidenciales (95%) del modelo BJ-2.

BJ-2	L. I. (2,5%)	Estimación	L.S.(97,5%)	Valor Real
Enero	691,8845	788,6130	898,8645	910,870
Febrero	598,0432	693,5337	804,2714	744,322
Marzo	540,3731	645,5001	771,0790	801,157
Abril	445,9454	543,0075	661,1956	512,637
Mayo	437,9049	544,0128	675,8316	582,845
Junio	390,8199	493,9445	624,2803	556,618
Julio	413,6360	531,4897	682,9224	621,894
Agosto	436,4074	569,3817	742,8737	634,137
Septiembre	463,9102	614,1322	812,9985	564,881
Octubre	390,9069	524,6710	704,2077	673,151
Noviembre	409,5668	557,0122	757,5384	606,917
Diciembre	363,3811	500,4826	689,3118	624,039

Fuente. IEA.
Elaboración. Propia.

Tabla.IV.2.21. Intervalos confidenciales (95%) del modelo BJ-3.

BJ-3	L. I. (2,5%)	Estimación	L.S.(97,5%)	Valor Real
Enero	797,71423	911,0359	1044,3835	910,870
Febrero	677,3697	784,4914	908,5537	744,322
Marzo	589,3632	683,0043	791,5235	801,157
Abril	486,9810	564,7149	654,8568	512,637
Mayo	520,8809	604,4094	701,3325	582,845
Junio	476,0826	552,7764	641,8250	556,618
Julio	503,1430	584,5637	679,1602	621,894
Agosto	536,9132	624,1898	725,6533	634,137
Septiembre	578,9249	673,4507	783,4104	564,881
Octubre	490,6342	571,0988	664,7597	673,151
Noviembre	530,9113	618,3639	720,2219	606,917
Diciembre	421,7366	491,5087	572,8239	624,039

Fuente. IEA.
Elaboración. Propia.

Tabla.IV.2.22. Intervalos confidenciales (95%) del modelo BJ-4.

BJ-4	L. I. (2,5%)	Estimación	L.S.(97,5%)	Valor Real
Enero	755,2503	868,6467	999,0690	910,870
Febrero	616,0237	731,3966	868,3773	744,322
Marzo	529,3779	637,3983	767,4604	801,157
Abril	435,0705	527,3723	639,2563	512,637
Mayo	464,3439	564,7201	686,7944	582,845
Junio	424,1239	516,6586	629,3825	556,618
Julio	448,2524	546,5052	666,2941	621,894
Agosto	478,5226	583,6549	711,8849	634,137
Septiembre	516,2433	629,7962	768,3261	564,881
Octubre	437,7764	534,1268	651,6832	673,151
Noviembre	474,0121	578,3690	705,7008	606,917
Diciembre	376,7769	459,7395	560,9696	624,039

Fuente. IEA.
Elaboración. Propia.

Estas predicciones han sido obtenidas poniendo en forma explícita la ecuación de predicción para cada modelo, la cual se detalla a continuación para el modelo BJ-1, obteniéndose de forma análoga las de los restantes modelos candidatos:

$$z_t = \ln q_t$$

$$(1-B)\left(1-B^{12}\right)z_t = (1-0{,}81B)\left(1-0{,}75B^{12}\right)\varepsilon_t$$

$$z_t = z_{t-1} + z_{t-12} - z_{t-13} + \varepsilon_t - 0{,}81\varepsilon_{t-1} - 0{,}75\varepsilon_{t-12} + 0{,}61\varepsilon_{t-13}$$

Como puede observarse, tomando como fin de los datos *históricos* (o como origen de la predicción del *futuro*) diciembre de 2000, todos los valores realmente observados para los próximos 12 meses (excepto los correspondientes a Marzo, Octubre y Diciembre de 2001) están dentro del intervalo de confianza del 95% pronosticado por los modelos BJ-1,BJ-3 y BJ-4. El modelo BJ-2 sólo deja fuera de tal intervalo confidencial la observación correspondiente al mes de Enero y Marzo, al 95% de confianza. En cambio, al nivel de confianza del 99% todas las observaciones reales del periodo Enero 2001-Diciembre 2001 están dentro del correspondiente intervalo para los modelos BJ-1 y BJ-2, mientras que el modelo BJ-4 deja fuera de tal intervalo las observaciones de Marzo y Diciembre de 2001 y el modelo BJ-3 deja fuera del intervalo sólo la observación de Diciembre de 2001. La representación gráfica del pronóstico puntual y de los intervalos de confianza al 99% aparecen en la figura IV.2.20.

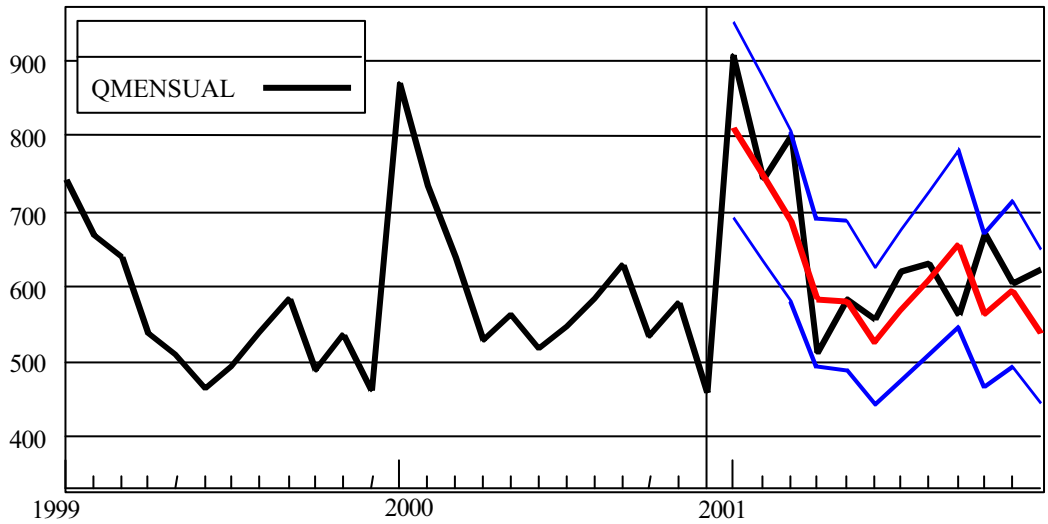


Figura.IV.2.20. Valores reales, ajustados y límites de confianza (99%) en la demanda mensual de electricidad en la C.A.A. durante el periodo 1995-2001. Modelo BJ-1.
Fuente. IEA.
Elaboración. Propia.

A continuación en la tabla IV.2.23 se presentan las estimaciones puntuales y confidenciales (al nivel de confianza del 95%) para los meses correspondientes al año 2003.

Tabla.IV.2.23. Intervalos confidenciales (95%) del modelo BJ-1 de los meses del año 2003. Serie Histórica 1995-2002.

BJ-1	L. I. (2,5%)	Estimación	L.S.(97,5%)
Enero	803,472	929,119	1074,413
Febrero	772,920	894,170	1034,440
Marzo	699,248	809,285	936,637
Abril	583,772	675,923	782,620
Mayo	602,394	697,779	808,268
Junio	530,529	614,794	712,443
Julio	587,859	681,515	790,093
Agosto	607,317	704,369	816,930
Septiembre	633,141	734,626	852,379
Octubre	589,251	683,986	793,942
Noviembre	588,757	683,696	793,945
Diciembre	561,644	655,732	765,583

Fuente. IEA.
Elaboración. Propia.

El resultado de sumar las estimaciones puntuales daría una estimación puntual de la demanda doméstica de electricidad en la C.A.A, esto es 8764,997 GWh, siendo un 7,7% superior a la predicción realizada para ese mismo año en el capítulo anterior que era de 8090,202 GWh. Además, la predicción anual realizada para el año 2003 (a través de modelos mensuales) pertenece al intervalo construido (95% de confianza) para el año 2003, ver página 150.

Tabla.IV.2.23(a). Estadísticos evaluadores de los modelos no dinámicos.

Tipología	***Modelos sin tendencia***					***M. con tendencia y estacionalidad***	***Ajuste de Curva***			
Modelo	M.Ingenuo	*Medias Móviles*			AE	M. Holt-Winters	Lineal	Cuadrático Negativo	Exponencial	Crecimiento
Estadísticos	Paseo Aleatorio	MM(11)	MM(12)	MM(13)	M. Estac. Mulp	M. Estac. Mulp				
$\bar{R}^2$	0,1836	0,176	0,2639	0,2871	0,8685	0,8773	0,2854	0,2751	0,2845	0,2755
DW	2,385	1,142	1,14	1,199	1,812	1,696	1,193	1,193	1,191	1,194
Forecast Error	84,76	85,15	80,49	79,21	34,02	33,33	79,3	79,87	79,35	79,85
MAPE	0,1084	0,1158	0,1079	0,104	0,04089	0,04104	0,1153	0,1154	0,1147	0,1154
MAD	58,23	61,59	57,46	55,48	21,44	21,24	60,01	60,06	59,77	60,08
BIC	82,85	83,23	78,67	77,42	35,6	35,67	82,97	85,47	83,03	85,45
RMSE	85,35	85,74	81,04	79,76	33,55	32,63	78,19	78,19	78,24	78,17

Fuente. IEA.
Elaboración. Propia.

Tabla.V.2.23(b). Estadísticos evaluadores de los modelos Box-Jenkins. Modelos Dinámicos.

Tipología	***Modelos Dinámicos***			
Modelo	Modelos Box-Jenkins			
Estadísticos	BJ-1	BJ-2	BJ-3	BJ-4
$\overline{R}^2$	0,8936	0,876	0,8649	0,8579
DW	1,702	1,971	1,803	2,22
Forecast Error	0,056	0,061	0,06399	0,0656
MAPE	0,0347	0,0380	0,03902	0,0422
MAD	18,5	20,46	20,69	22,87
BIC	29,51	31,86	33,25	33,34
RMSE	31,52	35,21	35,06	37,22
Ljung-Box(18) p-valor	20,3 (0,3163)	22,4 (0,2147)	31,86 (0,0228)	22,41 (0,2142)

Fuente. IEA.
Elaboración. Propia.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

IV.3. INFLUENCIA DEL FACTOR TEMPERATURA. DATOS MENSUALES

IV.3.1. Breve reseña histórica

La determinación cuantitativa del impacto de la temperatura sobre el consumo de energía eléctrica tiene particular importancia en aquellos países en los que la potencia generadora en reserva es escasa, pudiendo llegar incluso a ser deficitario el suministro de energía cuando se presentan situaciones meteorológicas muy desfavorables imprevistas (por ejemplo, la repentina llegada de un frente extremadamente frío). Más concretamente, para los distribuidores de tales países una previsión cuantitativa a muy corto plazo (por ejemplo, para el próximo día) de la carga extra, es una información fundamental, a fin de planificar una utilización racional de las fuentes de producción de las que se dispone, o incluso, en ciertos casos extremos, para evitar la incapacidad de proporcionar el servicio requerido (*blackouts*).

El estudio de la influencia de las condiciones meteorológicas en la demanda de electricidad ha sido abordado por diversos autores, pudiendo citarse entre los primeros que se ocuparon de este tema a Dryar, H.A.(1944), Williams y Leslie(1953), Davies, M.(1959), Quayle, R.G. y Diaz, H.F.(1980).

Dryar, H.A. puede ser considerado el pionero de los estudios de la influencia de la temperatura en el consumo de electricidad. Para este autor dicho consumo refleja principalmente, por una parte, el nivel de actividad económica y, por otra, las condiciones climáticas. Fue el primero que introdujo el concepto de la *demanda base,* como aquella parte de la demanda que no se debe, entre otras causas, a las variables atmosféricas. La temperatura no es la única variable considerada en este aspecto, ya que Dryar considera también la posible influencia en la demanda eléctrica de otros factores tales como el nivel

de nubosidad y la velocidad del viento. Su trabajo es, sobre todo, descriptivo y en el mismo no se propone ningún modelo analítico.

Davies, M.(1959), propuso un método para estudiar la relación estadística entre las variables meteorológicas y la demanda eléctrica dirigido especialmente a los ingenieros responsables de la planificación del servicio. Su estudio se centra principalmente en los consumos debidos a las necesidades de calefacción e iluminación. Este autor contempla intervalos de media hora en las correspondientes series temporales y propone unos comportamientos estacionales con un perfil propio para cada día de la semana. Uno de los aspectos más interesantes de este trabajo consiste en introducir un cierto retardo entre la temperatura y la demanda inducida por esta variable.

Le Comte, D.M. y Warren, H.E.(1981), formularon un modelo de regresión múltiple para establecer la relación entre la demanda de energía eléctrica semanal del verano, en nueve estados norteamericanos, y las anomalías de temperatura en ellos, ponderada por sus respectivas poblaciones.

Panuska, V.(1977), realizó un estudio probabilístico de previsiones horarias de demanda de potencia eléctrica en el que esta variable dependiente está relacionada con variables climáticas. Este trabajo establece límites probabilísticos para una demanda *nominal* (representada por una serie de Fourier) a la que se añaden, por una parte, una componente *residual* que satisface una ecuación diferencial referente a la variable dependiente temperatura y una componente *estocástica* que tiene en cuenta el factor incertidumbre de la demanda. El modelo propuesto por este autor contempla, un cierto retardo entre la variable climática y el consumo.

Taylor, L.D.(1975), presentó una panorámica general crítica de algunos estudios econométricos de la época, en los que la demanda de electricidad correspondiente a un determinado periodo de tiempo es la variable dependiente de la regresión y en los que la magnitud *grados día* (que, en definitiva, es una transformación de la variable temperatura), forma parte del conjunto de variables explicativas. La mayoría de los trabajos analizados por

Taylor se refieren al sector residencial o doméstico, con pocas referencias a los sectores comercial e industrial.

Stanton y Gupta(1969), concentraron su estudio en la relación existente entre la punta de potencia semanal demandada y las variables climáticas, postulando un modelo no-lineal en el que se sugiere una metodología para estimar la función de densidad de la punta de carga anual.

Thompson, R.P.(1976), desarrolló una metodología para predecir los picos de potencia y de energía basada en las tendencias que siguen la media y la desviación típica de sus respectivos máximos, proponiendo también directrices para el pronóstico de la carga de cada día de la semana.

Bolzern, P., Fronza, G. y Brusasca, G.(1982), estudiaron la relación de la temperatura con el consumo diario invernal de electricidad en el área de Milán, el mayor de los distritos en que ENEL (Ente Nacional para la Energía Eléctrica) considera dividida a Italia. En este trabajo la variable dependiente (carga de consumo eléctrico) es el promedio de energía consumida en el intervalo horario que comprende desde las 8:00h. hasta las 23:00h. de cada día, mientras que la variable explicativa es la temperatura media registrada en diez estaciones meteorológicas durante el mismo período horario. El análisis de componentes principales que se realiza en este trabajo arroja la conclusión de que el autovalor correspondiente al primer componente principal, por sí sólo, explica el 80% de la varianza total de la carga. Una característica interesante de este trabajo es que contempla una regresión sobre la media de dos temperaturas, la del período vigente y la del inmediatamente precedente, variable esta última sobre cuya significación estadística no existe unanimidad en estudios posteriores referentes a distintas áreas geográficas y diferentes períodos de estudio. Una de las principales conclusiones de este trabajo es que, en el ámbito de su desarrollo y respecto al período invernal, una disminución de 1ºC suele repercutir en una sobrecarga de demanda de 40 MW, por lo que una disminución de 6ºC entre dos días consecutivos (nada inusual en el área de Milán, según los autores) puede inducir magnitudes extras de potencia demandada incluso del 5%.

Ayyash, S., Salman, M. y Al-Hafi, N.(1985), utilizaron modelos lineales y cuadráticos para estudiar el impacto de la temperatura en el consumo veraniego de electricidad en Kuwait. Los datos de este trabajo son de frecuencia mensual y las temperaturas registradas, aunque sólo se obtuvieron en el Aeropuerto Internacional de Kuwait, se generalizaron a todo el país, lo cual no parece ser una fuente de considerable error, dada la reducida extensión (18000 Km^2) de dicho país. Estos autores concluyen que, en el ámbito de su estudio, un incremento del 10% en la magnitud del CDD (*cooling degree days*) mensual repercute en un consumo extra de 100-150 millones de kWh.

Ramesh, S., Natarajan, B.y Bhagat, G.(1988) desarrollaron un modelo y presentaron un análisis de la relación entre la temperatura y la humedad, por un lado, y las puntas de consumo eléctrico, por otro, referente al área de Delhi, India. Este trabajo contempla la separación absoluta entre los consumos invernal y estival, utilizando datos de frecuencia diaria, e incluso horaria, para las regresiones preliminares sobre el nivel de nubosidad e iluminación, aunque estas dos variables no resultaron ser significativas en el grado de ajuste de la ecuación. En cambio, la temperatura y la humedad resultaron ser muy significativas en el consumo eléctrico de una ciudad de características ambientales tan extremas como es Delhi.

Engle, R.F., Mustafa, C. y Rice, J.(1992) estimaron una compleja ecuación de predicción, para el día siguiente, relativa a la punta horaria de demanda eléctrica, incorporando en este modelo influencias exógenas deterministas y estocásticas, tales como el clima.

Engle, R.F., Granger, C.W., Rice, J. y Weiss, A.(1986) estudiaron un modelo no-lineal entre la demanda eléctrica del sector residencial de cuatro compañías distribuidoras norteamericanas y la temperatura, en el que la estimación de los coeficientes se realiza mediante técnicas semiparamétricas. El trabajo contempla la posibilidad de retardo temporal entre la facturación y el consumo. Aunque el estudio tiene en cuenta la posible autocorrelación de los

datos, las curvas *locales* propuestas (*splines*) carecen de intervalos de confianza, por lo que no es posible cuantificar la precisión de los resultados obtenidos

Al-Zayer, J. y Al-Ibrahim, A.(1996) desarrollaron un modelo econométrico para predecir el consumo de energía eléctrica y estudiar el impacto de la temperatura ambiente (expresada en términos de *grados día*) en Arabia Saudí. Estos autores encontraron los resultados más significativos en regresiones de horizonte anual. Asimismo abogan por considerar, en el ámbito de su estudio, formulaciones estructuralmente diferentes para la demanda eléctrica invernal y la estival.

IV.3.2. Demanda base y temperatura neutra

Dos cuestiones citadas anteriormente y, en cierto modo relacionadas entre sí, son las referentes a los conceptos de *demanda base* y *temperatura neutra* los cuales se tratan en la literatura especializada de forma algo imprecisa, dando la apariencia de que corresponden más a conceptos intuitivos que a definiciones formales, acaso porque no pueda ser de otro modo, dada la gran variedad de circunstancias que concurren en las distintas zonas geográficas demandantes de energía eléctrica. De todas formas hay que resaltar que sobre la cuantificación de estos conceptos, no suele existir gran controversia en la literatura especializada, sino una mera exposición de las opiniones de sus autores.

Ante todo hay que tener en cuenta que no toda la diferencia entre la *demanda real* y la *demanda base* debe atribuirse a factores climáticos, ya que multitud de otros factores (acontecimientos extraordinarios, traslación en el calendario de periodos no laborables entre dos años consecutivos, etc.) pueden aumentar tal diferencia y ésta puede ser mayor cuanto menor sea el período considerado o más desagregada sea la demanda. Algunos autores consideran la demanda *base* para un determinado período como aquella parte de la *demanda real* que tiene lugar en ausencia de condiciones extraordinarias (climáticas o no) y proponen diferentes modelos para estimarla, los cuales no

son susceptibles de contraste. Otros autores, al referirse a la *demanda base* la consideran, a su vez, ya desestacionalizada al considerar un ciclo semanal o anual de la carga.

Otro tanto puede decirse respecto al concepto de *temperatura neutra*, el cual puede entenderse como aquélla temperatura que no es causa de un consumo de energía eléctrica que exceda el *habitual de condiciones normales.* Lo anterior justifica que las expresiones numéricas de este concepto muestren cierta disparidad, no sólo entre zonas geográficas distantes, sino entre áreas relativamente próximas (en la Comunidad Autónoma de Andalucía no faltan ejemplos de ello), por no hablar de la distinta percepción, dentro de cierto rango, que dos individuos de una misma localidad puedan tener sobre cual es el umbral de temperatura a partir del cual se justifica un mayor incremento de energía eléctrica debido a la calefacción o al aire acondicionado. Todo ello dando por supuesto, que incluso un mismo individuo, muestre una actitud coherente consigo mismo en este aspecto a lo largo del tiempo en igualdad de circunstancias ambientales.

IV.3.3. Cuantificación del efecto temperatura

A fin de tener en cuenta la influencia de la temperatura como variable exógena en la ecuación de la demanda de energía se suele considerar una variable que expresa, las desviaciones acumulativas de la temperatura respecto a una cierta *temperatura neutra* durante un periodo determinado. A este respecto una variable habitual en la literatura especializada es la denominada *grados día (DD, degree days)*, sobre la cual pueden encontrarse definiciones no uniformes aunque, en el fondo, todas ellas expresan el concepto común de intentar cuantificar la discrepancia o *distancia* entre la temperatura ambiente representativa de un período de tiempo y una cierta temperatura *neutra* o *referencia*.

Más frecuente aún es considerar una división del concepto de los *grados día*, en *HDD* (Heating Degree Days) y *CDD* (Cooling Degree Days) correspondientes a condiciones climáticas que inducen un consumo de energía

eléctrica debido a la utilización de aparatos de calefacción y de aire acondicionado, respectivamente. Así, si denominamos T_N a la temperatura *neutra* se puede definir la variable *grados día de calefacción* (*HDD*) durante el día como la diferencia entre T_N y la temperatura media de ese día, T_{il}, cuando esta última es inferior a T_N , ya que en caso contrario el valor de *HDD* se considera nulo ese día. Así, se denomina T_{il} a la temperatura media del *l-ésimo* día del mes *i-ésimo* (l=1,2, ..., n; n=28,29,20,31; i=1,2, ..., 12). El *HDD* mensual es simplemente la suma de los *grados día de calefacción* a lo largo de todo los días del mes considerado. Formalmente:

$$HDD_{il} = \begin{cases} T_N - T_{il} & \text{si} \quad T_N > T_{il} \\ 0 & \text{si} \quad T_N \leq T_{il} \end{cases}$$

$$HDD_i = \sum_{l=1}^{n} HDD_{il} \text{ , } i=1, 2, ..., 12$$

Análogamente se puede definir la variable *grados día de refrigeración* (CDD) durante un día como la diferencia entre la temperatura media de ese día y T_N, supuesto que ésta última es superior a T_N, ya que si no se cumple esta última condición se asigna a *CDD* el valor cero para ese día. Por consiguiente se puede expresar formalmente el *CDD* mensual, CDD_i, como sigue :

$$CDD_{il} = \begin{cases} T_{il} - T_N & \text{si} \quad T_{il} > T_N \\ 0 & \text{si} \quad T_{il} \leq T_N \end{cases}$$

$$CDD_i = \sum_{l=1}^{n} CDD_{il} \text{ , } i=1, 2, ..., 12$$

En algunos estudios de la literatura especializada correspondientes a zonas geográficas de condiciones climáticas no muy extremas se considera la contribución que a la demanda de energía eléctrica realizan conjuntamente los *grados día de calefacción y refrigeración*. Sin embargo, no es infrecuente que algunos de estos estudios, correspondientes a regiones de temperaturas habitualmente extremas, consideren bien sólo los *grados día de calefacción* (Hayter(1995); Pankratz(1991)), o sólo los *grados día de refrigeración* (Ayyash, S., Salman, M. y Al-Hafi, N.(1985); Al-Zayer, J. y Al-Ibrahim, A.(1996)) por

considerar, si no inexistentes, sí prácticamente despreciables los *grados día* del tipo no considerado.

IV.3.4. Valores de temperatura neutra más utilizados

En la literatura especializada con frecuencia se encuentran no sólo cantidades puntuales, sino intervalos, para la *temperatura neutra*, entendiéndose en este último caso que, según los autores proponentes de este último concepto, cualquier valor de la temperatura perteneciente a dicho *intervalo neutro* no es causa de un mayor consumo de energía eléctrica. Es obvio, que aunque el concepto de *intervalos de temperatura neutra* tampoco es ninguna panacea (puesto que tales intervalos, a su vez, difieren entre autores cuyos trabajos se refieren a un mismo país) no obstante, intenta lograr una mayor flexibilidad que la mera formulación puntual de una única *temperatura neutra*.

Seguidamente se comentan algunos valores puntuales e intervalos de temperatura neutra, insistiendo en que ninguno de los autores que a continuación se citan se detiene lo más mínimo a defender por qué ha elegido tales valores, posiblemente porque esta eventual controversia no conduciría a resultados concluyentes.

Entre otros autores que mencionan como *temperatura neutra* la de 65º F (18,3º C) figuran Pankratz(1991) , Hayter(1995), Le Comte(1981), Engle, R.F., Scott, J.B. y Stern, G.(1988), Engle, R.F., Granger, C.W., Rice, J. y Weiss, A.(1986) y Ayyash, Salman y Al-Hafi(1985). En cambio, Tolley,G.S., Upton, C.W. y Hastings, V.S.(1977) se decantan por el valor 68º C (20º C).

Entre los autores que mencionan un *intervalo neutro* podemos citar a Mendenhall, W. y Sincich, T.(1993) quienes consideran para dicho concepto el intervalo comprendido entre 59º F y 78º F(15º C a 25,6º C) y a Chen, K.C.. y Winters, P.R.(1974) quienes contemplan el intervalo 50º F a 65º F(10º C a 18,3º C). Panuska(1977) menciona el intervalo neutro directamente en grados centígrados (16º C a 21º C). Otros autores establecen distintas pendientes para

la curva de consumo; así, por ejemplo, Engle, R.F., Mustafa, C. y Rice, J.(1992) consideran la base del HDD en 55ºF(12,7ºC) y como bases para el CDD consideran 66ºF(18,8ºC) y 76ºF(24,4ºC). Una aproximación parecida siguen Train, K e Ignelzi, P.(1984) para quienes la base del HDD está en 65ºF(18,3ºC) y los escalones del CDD los sitúan en 65ºF(18,3ºC), 75ºF(23,8ºC) y 85ºF(29,4ºC). Unas bases parecidas son contempladas por Stanton, K.N. y Gupta, P.C. (1969) para quienes el HDD se encuentra en 60ºF(15,5ºC) y el CDD en 70ºF(21,1ºC) y 75ºF(23,8ºC). Un modelo que considera como *intervalo neutro* el comprendido entre 65ºF(18,3ºC) y 75ºF(23,8ºC) es el propuesto por Shahidehpour, M. Yamin, H. y Li, Z.(2002) pudiéndose considerar como exponente de los modelos que incluyen factores atmosféricos que no sólo contemplan la temperatura ambiental sino una denominada *temperatura efectiva*, T_e, definida como sigue:

$$T_e = \begin{cases} T - \rho \dfrac{W(65-T)}{100}, & T < 65 \\ T, & 65 \le T \le 75 \\ T + \lambda \dfrac{H(T-75)}{100}, & T > 75 \end{cases}$$

donde H es la humedad relativa, W es la velocidad del viento, mientras que λ y ρ son coeficientes que tienen en cuenta los efectos de humedad y del viento.

IV.3.5. El efecto temperatura en el consumo residencial de energía eléctrica en la Comunidad Autónoma de Andalucía

En este apartado se estudia cual es la influencia que tiene la temperatura en el consumo doméstico de energía eléctrica en la Comunidad Autónoma de Andalucía. A lo largo de la investigación desarrollada no se ha tenido conocimiento de ningún estudio, anterior al presente, que haya tenido por objeto investigar el efecto de los datos mensuales de temperatura sobre el consumo de energía eléctrica del sector residencial en la Comunidad Autónoma de Andalucía.

BIBLIOTECA VIRTUAL

Ante todo se desea comentar que los valores de *temperatura neutra* (sean en forma puntual o como intervalo) que se utilizan en el presente trabajo, respecto a la Comunidad Autónoma de Andalucía, si bien intentan ser plausibles, en opinión de su autora, tienen sólo un carácter de instrumentos necesarios para el estudio metodológico, exento de cualquier interpretación dogmática, aunque sólo fuese por la ya mencionada relativa disparidad climática de nuestra Comunidad y la no menor disparidad de percepción personal respecto a este tema.

La definición anterior de los *grados día de calefacción y refrigeración* supone la disponibilidad de datos de temperatura media diaria a lo largo de varios años, periodo que suele ser el ámbito temporal de tales estudios. Lamentablemente no ha sido posible obtener tales datos para la realización de la presente memoria. En su lugar, y teniendo en cuenta que los datos que amablemente nos ha proporcionado el Banco de Datos del Instituto Nacional de Meteorología, Centro Meteorológico Territorial en Andalucía Occidental y Ceuta, se refieren a las temperaturas medias mensuales, $\overline{T}_{ij}$, de cada una de las ocho provincias andaluzas durante los años 1995 a 2001. A continuación se define brevemente, $\overline{T}_{ij}$.

En primer lugar, se registran los valores de las temperaturas máximas y mínimas diarias en cada una de las estaciones asignadas a cada provincia de Andalucía. A continuación se calcula el promedio mensual de las temperaturas máximas y mínimas para cada provincia.

$$\overline{T}_{máx_{ij}} = \frac{\sum_{l=1}^{n} T_{máx_{ijl}}}{n}; \; i=1, 2, ..., 12; \; j=1, 2, ..., 7$$

$$\overline{T}_{mín_{ij}} = \frac{\sum_{l=1}^{n} T_{mín_{ijl}}}{n}; \; i=1, 2, ..., 12; \; j=1, 2, ..., 7$$

donde:

$\overline{T}_{\text{máx}_{ij}}$: es la temperatura media de las máximas del mes *i-ésimo* y del año *j-ésimo*.

$\overline{T}_{\text{mín}_{ij}}$: es la temperatura media de las mínimas del mes *i-ésimo* y del año *j-ésimo*.

En segundo lugar, se calcula la semisuma de ambos promedios de temperaturas máximas y mínimas para cada uno de los meses de todos los años considerados, $\overline{T}_{ij}$, siendo esta magnitud la única disponible para el tratamiento de los datos:

$$\overline{T}_{ij} = \frac{\overline{T}_{máx_{ij}} + \overline{T}_{mín_{ij}}}{2}; \; i=1, 2, ..., 12; \; j=1, 2, ..., 7$$

Para salvar la dificultad comentada anteriormente, en este trabajo se ha considerado conveniente definir la variable *grados equivalentes mensuales* a fin de asignar una temperatura *representativa* de toda la Comunidad Autónoma de Andalucía para cada uno de los meses de los años estudiados.

Puesto que en el sector residencial de nuestra Comunidad se produce una contribución al consumo de electricidad tanto por calefacción como por refrigeración, en la mayoría de sus localidades, elegida una determinada temperatura neutra T_N (se elegirá aquel valor que contribuya en mayor medida al consumo de electricidad, ver tabla anexo A.IV.90) se considera el valor

absoluto de la diferencia entre la temperatura media mensual *representativa* de la Comunidad y T_N. Tal valor absoluto (o desviación respecto a la neutralidad) se multiplica por el número de días del mes de que se trate, denominando *grados equivalentes* del correspondiente mes al resultado del producto de los factores mencionados. En el cálculo anterior se ha tenido en cuenta que los años 1996 y 2000 fueron bisiestos.

La determinación de la temperatura media mensual *representativa* de la Comunidad Autónoma de Andalucía, a partir de las temperaturas medias mensuales de cada provincia, se ha realizado por dos procedimientos alternativos que arrojan prácticamente los mismos resultados, tal como puede comprobarse en tabla del anexo A.IV.1. Por una parte, se ha ponderado la temperatura media de cada provincia andaluza según la población de la misma, lo cual es frecuente en la literatura especializada cuando se desea expresar con un único número la temperatura de una determinada región, a partir de los datos suministrados por estaciones meteorológicas de la misma situadas en provincias relativamente distantes entre sí, pudiendo citarse entre las referencias de esta metodología a Lehman, R.L.y Warren, H.E.(1978), Le Comte, D.M. y Warren, H.E.(1981), Quayle,R.G. y Diaz,H.F.(1980) y el modelo del Center for Environmental Assesssment Services, NOAA, Washington.

Por otra parte, puesto que la ponderación de cada provincia andaluza en el sector doméstico o residencial se puede expresar aproximadamente a través del número de recibos (unidad de medida) del Impuesto de Bienes Inmuebles de naturaleza urbana (I.B.I.), se ha tomado esta cantidad como factor de ponderación (alternativo al correspondiente a la población, w_k) de la temperatura media provincial. Habría que precisar que se ha considerado como unidad urbana todo inmueble con relación de propiedad perfectamente delimitada a efectos fiscales. Este segundo procedimiento, ver tabla anexo A.IV.1 es menos redundante con el anterior de lo que pudiera suponerse, pues aunque los datos corroboran experimentalmente lo que ya a nivel simplemente intuitivo cabe esperar *a priori* (a mayor número de habitantes, mayor número de recibos I.B.I.) la asociación entre los dos factores de ponderación utilizados no sólo no es proporcional, sino que en algunas localidades es de sentido inverso.

Un ejemplo relevante de esto último lo constituyen las provincias de Málaga y Sevilla, pese a lo cual las diferencias entre los dos procedimientos mencionados en el cálculo de las temperaturas medias mensuales *representativas* de la Comunidad de Andalucía para los años citados son absolutamente insignificantes, por lo que, en adelante, se trabajará con las temperaturas medias resultantes de la ponderación según la población de cada provincia andaluza lo cual en cierto modo, incorpora el factor demográfico a la demanda investigada.

Los factores de ponderación según la población andaluza y según el número de recibos I.B.I. son de elaboración propia de acuerdo con los datos de las publicaciones del Instituto de Estadística de Andalucía, *Andalucía Datos Básicos 2002* y *Municipios Andaluces. Datos Básicos 2002,* respectivamente.

Los datos mensuales del consumo de electricidad en la Comunidad Autónoma de Andalucía proceden de las facturaciones realizadas por la Compañía Sevillana de Electricidad, S.A.[8] a los abonados, salvo para algunos municipios en los que los datos son estimados debido a que el suministro, y por tanto, la facturación a la que no se ha tenido acceso, es realizada por empresas intermediarias. Habría que decir que la Compañía Sevillana de Electricidad ha sido la principal y casi única empresa suministradora de electricidad en la comunidad durante el periodo 1995-2001.

Los datos medios mensuales de las temperaturas, para cada una de las ocho provincias de Andalucía, como ya se ha comentado anteriormente han sido suministrados por el Instituto Nacional de Meteorología, Centro Meteorológico Territorial de Andalucía Occidental y Ceuta. En la tabla IV.3.1 se muestra los puntos, denominación y localización geográfica, donde se han llevado a cabo la recogida de las muestras de temperaturas para cada una de las provincias de Andalucía.

[8] Denominada Compañía Sevillana de Electricidad, S.A. hasta la fusión por absorción, a finales de 1997, por parte del Grupo Endesa, S.A.

Tabla.IV.3.1. Denominación y localización de los puntos de recogida de temperaturas.

Provincia	Denominación	Coordenadas Geográficas (UTM)	Centro Meteorológico
Almería	Aeropuerto	LONG: 02-23-17W LAT: 36-50-35 ALT: 20	Andalucía Oriental
Cádiz	Cortadura	LONG: 06-15-37W LAT: 36-29-55 ALT:8	Andalucía Occidental
Córdoba	Aeropuerto	LONG: 04-51-02W LAT: 37-50-40 ALT: 91	Andalucía Occidental
Granada	Aeropuerto	LONG: 03-46-35W LAT: 37-11-24 ALT: 575	Andalucía Oriental
Huelva	Ronda Este	LONG: 06-54-35W LAT: 37-16-48 ALT: 19	Andalucía Occidental
Jaén	Cerro de los Lirios	LONG: 03-48-27W LAT: 37-46-40 ALT: 580	Andalucía Oriental
Málaga	Aeropuerto	LONG: 04-29-17W LAT: 36-40-00 ALT: 7	Andalucía Oriental
Sevilla	Aeropuerto	LONG: 05-54-13W LAT: 37-25-26 ALT: 26	Andalucía Occidental

Fuente. Centro Meteorológico Territorial en Andalucía Occidental y Ceuta. INM.
Elaboración. Propia.

Respecto a los distintos tipos datos (variables) considerados en el estudio, se tendrían que considerar por una parte, el consumo de electricidad y por otro, la temperatura ambiental. Con estos datos se han creado variables auxiliares que ayudarán a estudiar la influencia de la temperatura en el consumo de electricidad.

En primer lugar, respecto a la electricidad consumida, se han definido las siguientes expresiones:

- Consumo de electricidad (GWh): representa el total de la electricidad facturada, expresada en gigavatios por hora, en el mercado doméstico, por mes y año, en la Comunidad Autónoma de Andalucía.
- Consumo de electricidad trasladado (GWh): representa la diferencia del consumo del mes y el consumo mínimo registrado en ese año.

En segundo lugar, respecto a la temperatura ambiente, se han definido las siguientes variables:

- Temperatura ($\overline{T}_{ij}$): es la media de los temperaturas medias mensuales registradas en cada una de las ocho provincias de Andalucía ponderada según la población anual registrada durante el periodo 1995-2001.

  siendo

  $\overline{T}_{ijk}$: la temperatura media alcanzada en el mes *i-ésimo* del año *j-ésimo* en la provincia *k-ésima* de la C.A.A.

  w_k: coeficiente de ponderación según población de la provincia *k-ésima*.

  donde i= 1,2,..., 12; j=1, 2, ..., 7; k=1, 2, ..., 8.

  La temperatura calculada para cada uno de los doce meses de los años comprendidos en el periodo 1995-2001, $\overline{T}_{ij}$, es el resultado de la siguiente expresión:

$$\overline{T}_{ij} = \sum_{k=1}^{8} \overline{T}_{ijk} w_k \qquad \text{siendo} \quad \sum_{k=1}^{8} w_k = 1$$

- Geqmen22: esta variable representa los grados equivalentes mensuales considerando que la temperatura neutra es de 22ºC y es función de la temperatura, $\overline{T}_{ij}$, y del número de días naturales de los cuales consta cada mes, d_{ij}. Teniendo en cuenta que los años 1996 y 2000 son bisiestos, en el mes de febrero de estos años se consideran 29 días. Esta variable se calcula, para cada mes de los distintos años, como el producto del número de días naturales del mes en cuestión por el valor absoluto de la diferencia de la variable anterior definida, *temperatura* y la temperatura constante 22 (la temperatura considerada *neutra*). Su expresión es la siguiente:

$$\text{Geqmen22} = f(T_{ij}, d_{ij}) = d_{ij} \cdot \left|\overline{T}_{ij} - 22\right|, \quad i=1, 2, ..., 12; j=1, 2, ...7.$$

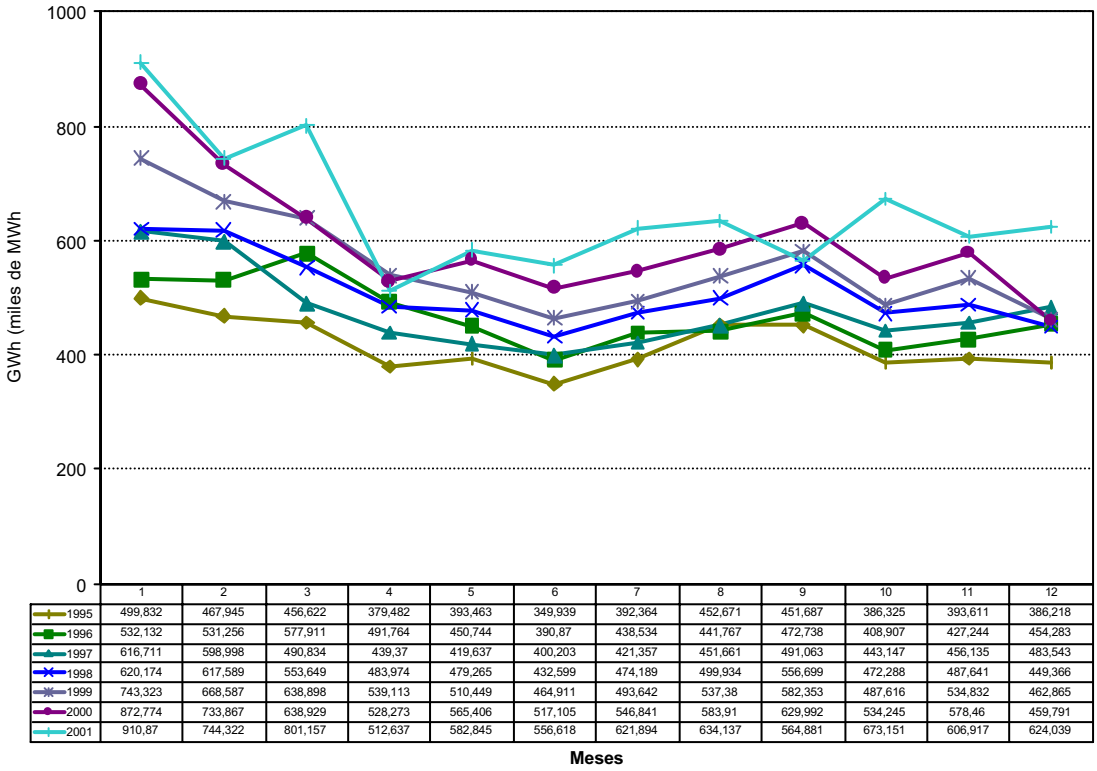
- Geqmen22(-1): representa los grados equivalentes mensuales con retardo de un mes.

IV.3.6. Características de los datos mensuales

En la figura IV.3.1 se representa un gráfico del consumo mensual de electricidad para cada uno de los años de la serie estudiada, 1995-2001. El gráfico tiene una forma definida de campana invertida, indicando el consumo mínimo (en el eje de ordenadas) de electricidad en el mes de junio para la gran mayoría de los años, a excepción de los años 1999, 2000 y 2001, cuyos consumos mínimos se registran en diciembre (años 1999 y 2000) y abril (año 2001) respectivamente. Esta circunstancia de consumos mínimos en meses inusuales puede ser debida a la distinta periodicidad con la que se realizan las facturaciones o lecturas en el periodo correspondiente. Es decir, las facturaciones se realizan con periodicidad variable, según sean suministros en alta tensión (mensual) o en baja tensión (bimensual), por lo tanto puede suceder que se correspondan con los consumos reales del periodo la facturación producida en el periodo de dos lecturas sucesivas.

En relación con esta situación puede comentarse que no se trata de un problema específico de una determinada compañía eléctrica. El trabajo anteriormente citado de Train, K. e Ignelzi, P.(1984) está primordialmente dedicado al tema del retraso que sufre la facturación respecto a la temperatura soportada por los abonados de varias distribuidoras eléctricas norteamericanas, lo cual es considerado por estos autores como una situación genérica, dado el calendario que siguen los departamentos responsables de facturación de las respectivas empresas, hasta el punto que estos autores comentan que el mencionado retraso oscurece la asociación estadística entre temperatura y facturación y proponen un sistema, que ellos mismos consideran que sólo podría aplicarse en el hipotético caso de que se tuviera acceso al número de clientes facturados cada día del periodo plurianual investigado (lo que evidentemente no es nuestro caso), de forma que se pudiera construir diariamente un valor ficticio de la temperatura como suma ponderada de las temperatura ambiental registrada por la proporción de clientes que se han facturado ese día.

En la figura IV.3.1 se observa que los picos de consumo de electricidad ocurren durante la temporada de invierno, el cual se puede considerar que aproximadamente comienza en septiembre y termina a finales de marzo. El consumo más elevado de electricidad tiene lugar en enero para todos los años estudiados desde 1995 hasta 2001, a excepción del año 1996 donde el consumo máximo de electricidad tiene lugar en el mes de marzo. En el año 2001 se registra el nivel máximo de consumo de electricidad en el mes de enero con un volumen de 910,870 GWh. El consumo menor de electricidad se produce en junio del primer año analizado, 1995, con un volumen de 349,939 GWh.



	1	2	3	4	5	6	7	8	9	10	11	12
1995	499,832	467,945	456,622	379,482	393,463	349,939	392,364	452,671	451,687	386,325	393,611	386,218
1996	532,132	531,256	577,911	491,764	450,744	390,87	438,534	441,767	472,738	408,907	427,244	454,283
1997	616,711	598,998	490,834	439,37	419,637	400,203	421,357	451,661	491,063	443,147	456,135	483,543
1998	620,174	617,589	553,649	483,974	479,265	432,599	474,189	499,934	556,699	472,288	487,641	449,366
1999	743,323	668,587	638,898	539,113	510,449	464,911	493,642	537,38	582,353	487,616	534,832	462,865
2000	872,774	733,867	638,929	528,273	565,406	517,105	546,841	583,91	629,992	534,245	578,46	459,791
2001	910,87	744,322	801,157	512,637	582,845	556,618	621,894	634,137	564,881	673,151	606,917	624,039

Figura.IV.3.1. Consumo mensual de electricidad en la C.A.A. durante el periodo 1995-2001.
Fuente. IEA.
Elaboración. Propia.

En la figura IV.3.2 se representa un gráfico donde se superpone año a año el consumo mensual de electricidad y la distribución de los grados equivalentes mensuales al nivel 22ºC (Geqmen22) para el periodo 1995-2001. El gráfico revela que el consumo de electricidad y los Geqmen22, en su conjunto, se repiten en un patrón cíclico anual. El gráfico también muestra una notable similitud al patrón cíclico considerando por si solas la variable consumo

de electricidad y la variable Geqmen22 en el periodo 1995-2001. Aunque, si bien, se observa que en este patrón que las temperaturas mínimas no se corresponden exactamente con los consumo máximos de electricidad, esto puede ser debido, como ya se ha comentado anteriormente, a la distinta periodicidad con la que se realizan las facturaciones o lecturas.

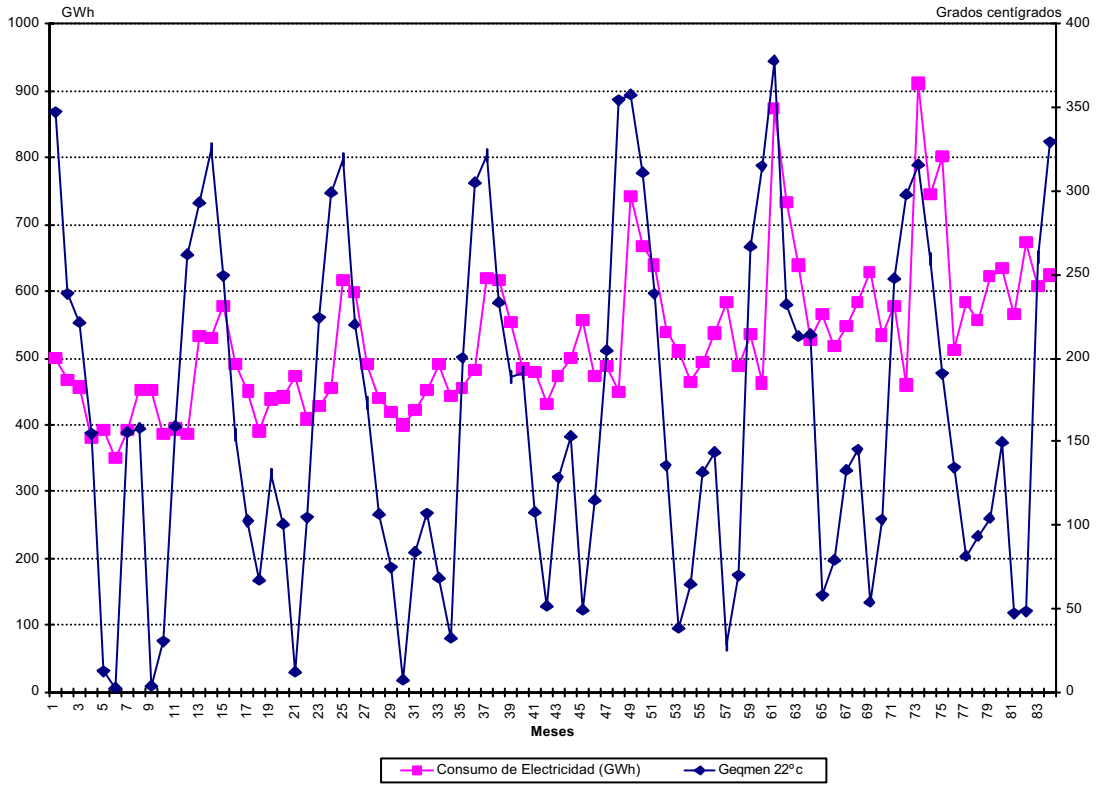


Figura IV.3.2. Consumo mensual de electricidad y grados equivalente (22º C) en la C.A.A. durante el periodo 1995-2001.
Fuente. IEA.
Elaboración. Propia.

La corrección de esta traslación sufrida por los datos se puede realizar, bien, sustrayendo al consumo mensual el consumo mínimo del año representado mediante la variable, *Consumo de Electricidad trasladado*, o bien, a través de los grados equivalentes mensuales retardados, mediante la variable *Geqmen 22(-1),* definida anteriormente. Una vez definidas estas variables se representa su evolución en el tiempo, constatándose, en la figuras IV.3.3 y IV.3.4, de forma clara el patrón cíclico que sigue el consumo de electricidad y la temperatura año a año. Así, también se observa, sin ningún tipo de traslación, que las demandas máximas de electricidad corresponden a los Geqmen22 máximos, ya sean temperaturas ambientales mínimas o máximas.

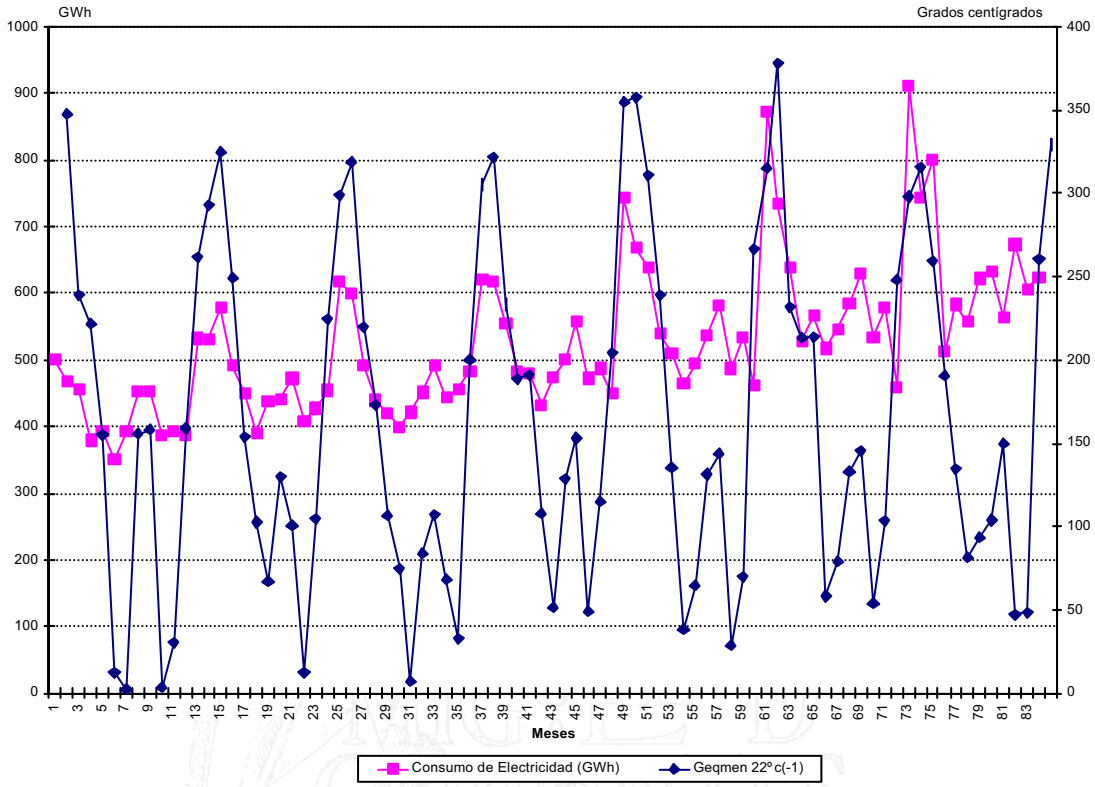


Figura.IV.3.3. Consumo mensual de electricidad y grados equivalentes retardados en la C.A.A. durante el periodo 1995-2001.
Fuente. IEA.
Elaboración. Propia.

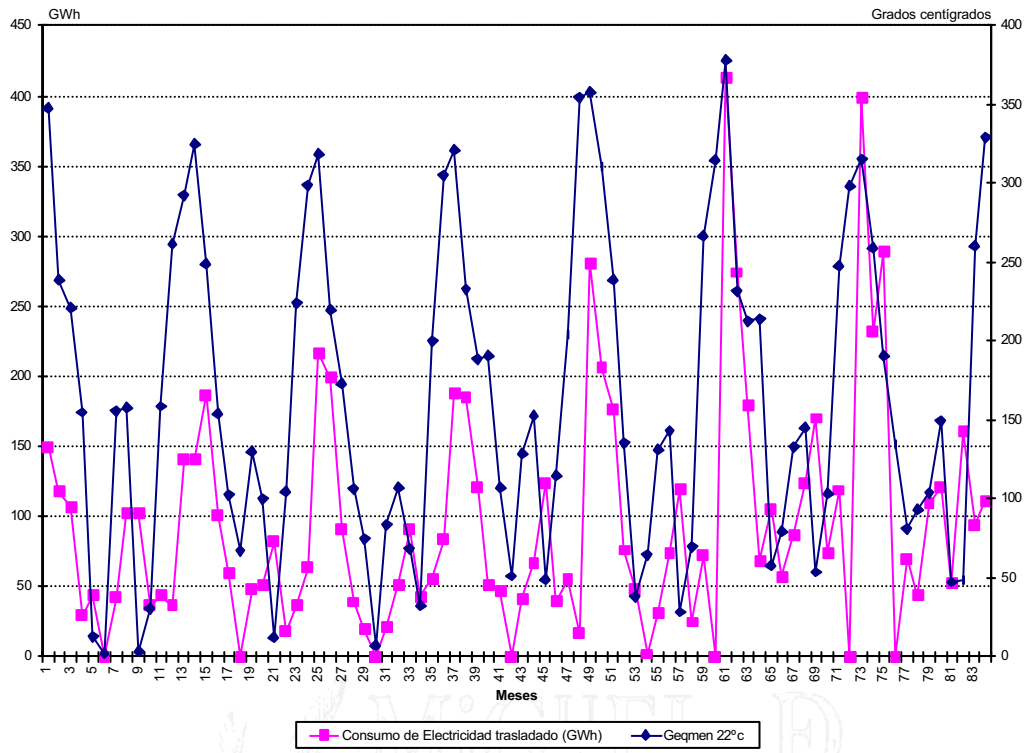


Figura.IV.3.4. Consumo mensual de electricidad trasladado y grados equivalentes (22º C) en la C A.A. durante el periodo 1995-2001. Fuente. IEA. Elaboración. Propia.

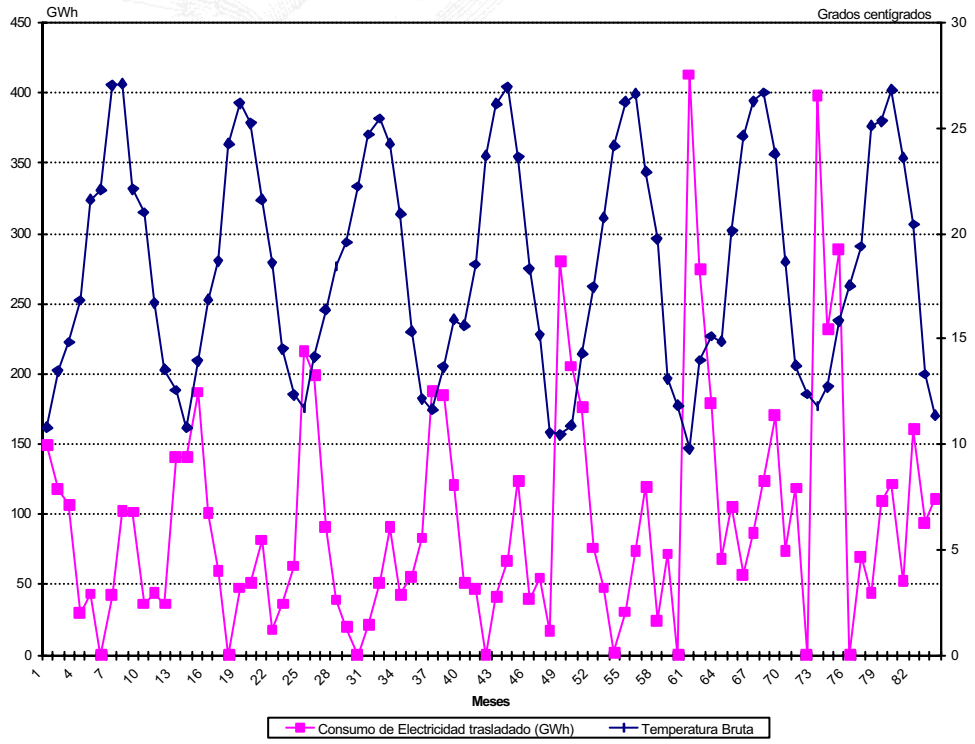


Figura.IV.3.5. Consumo mensual de electricidad trasladado y temperatura bruta en la C.A.A. durante el periodo 1995-2001.Fuente. IEA. Elaboración. Propia.

Las figuras IV.3.6 y IV.3.7 representan el patrón de consumo mensual de electricidad con respecto a Geqmen 22º y Geqmen 22º(-1), respectivamente. Se observa en ambos diagramas de dispersión una orientación de dependencia lineal positiva de los datos, apreciándose gráficamente una mayor relación entre las variables, en la figura IV.3.7, donde se representa el consumo de electricidad frente a los grados equivalentes mensuales 22ºC retardados un mes.

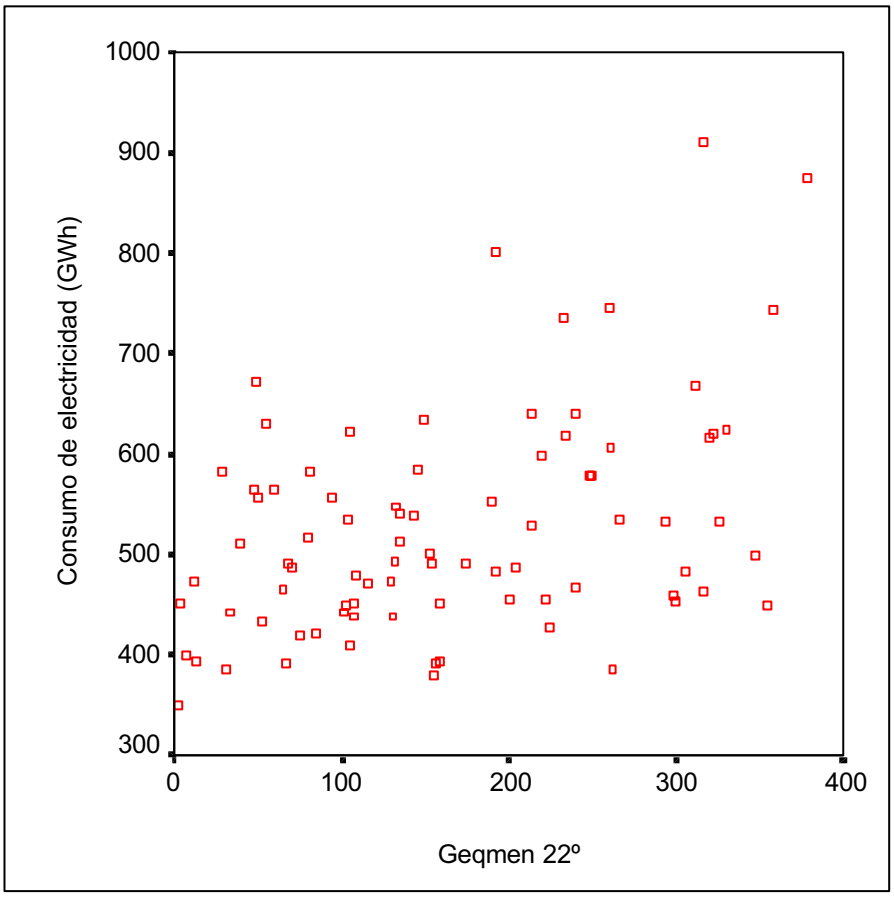


Figura.IV.3.6. Diagrama de dispersión. Consumo de electricidad versus grados equivalentes mensuales 22ºC.
Fuente. IEA.
Elaboración. Propia.

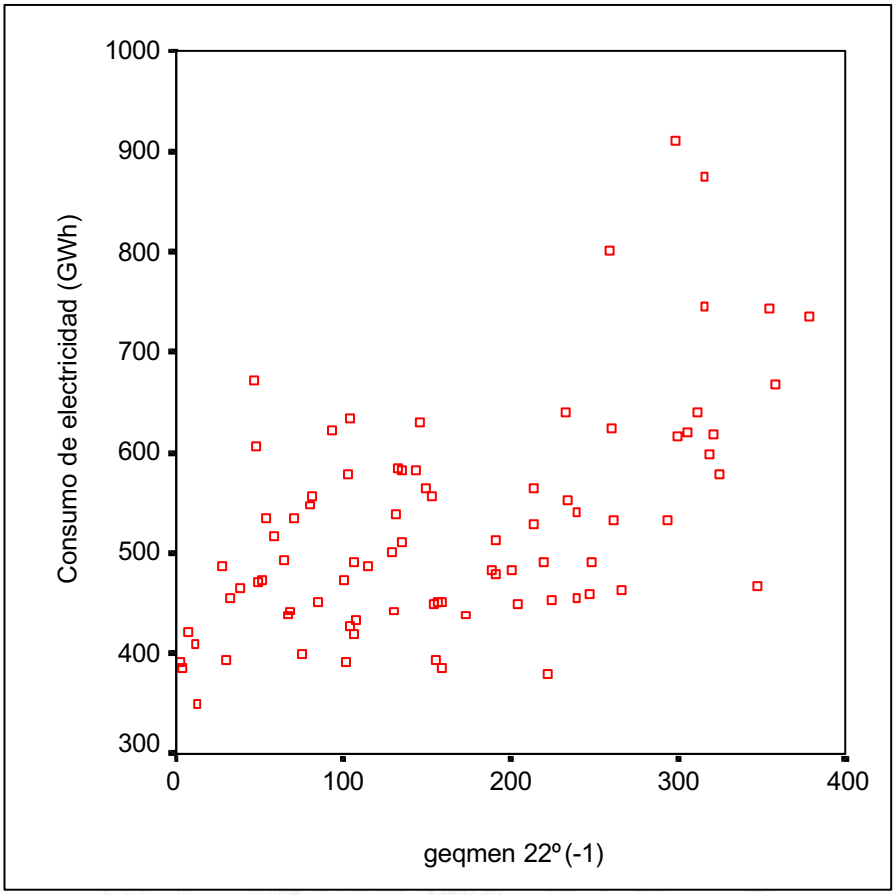


Figura.IV.3.7. Diagrama de dispersión. Consumo de electricidad versus grados equivalentes mensuales 22º C(-1).
Fuente. IEA.
Elaboración. Propia.

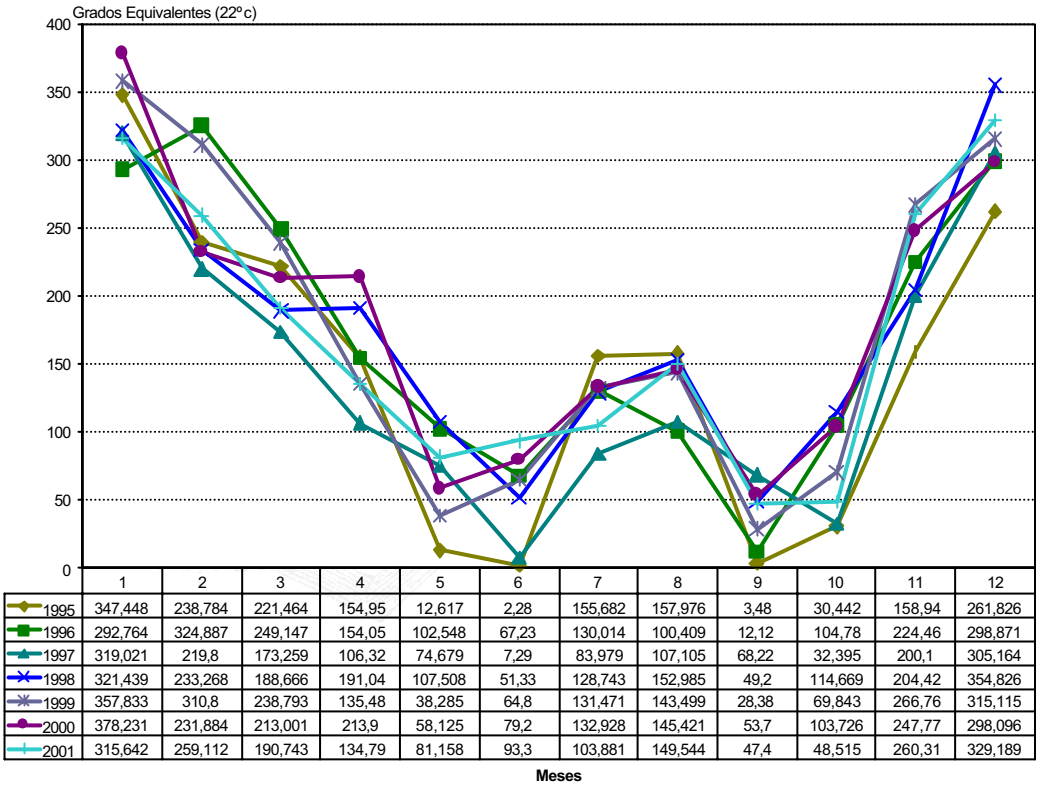
Una vez observada gráficamente la relación de dependencia positiva entre el consumo de electricidad y la temperatura, se aborda su calculo analíticamente.

El consumo mensual de electricidad se encuentra correlacionado positivamente con respecto a los Geqmen22 y Geqmen22(-1). Los coeficientes de correlación para el consumo mensual de electricidad y Geqmen22 son 0,558, 0,593* 0,746**, 0,321, 0,561, 0,475, 0,546 para el periodo 1995-2001 respectivamente. Los coeficientes de correlación para el consumo de electricidad y Geqmen 22º(-1) son 0,675*, 0,877**, 0,851**, 0,754**, 0,760**, 0,615*, 0,605*, lo que corrobora la existencia de un desfase en la facturación del consumo de electricidad. El coeficiente de correlación global para el periodo

* La correlación es significativa al nivel 0.05 (bilateral).
** La correlación es significativa al nivel 0.01 (bilateral).

1995-2001 considerando los Geqmen22 y Geqmen22(-1) son 0.422** y 0.549** respectivamente (siendo correlaciones significativas al 1%).

La figura IV.3.8 muestra el gráfico de los grados equivalentes a 22º C para cada uno de los siete años desde 1995 a 2001. Comparando entre sí las figuras IV.3.1 a IV.3.5, se observa que todos los gráficos tienen una forma parecida a una campana invertida.



	1	2	3	4	5	6	7	8	9	10	11	12
1995	347,448	238,784	221,464	154,95	12,617	2,28	155,682	157,976	3,48	30,442	158,94	261,826
1996	292,764	324,887	249,147	154,05	102,548	67,23	130,014	100,409	12,12	104,78	224,46	298,871
1997	319,021	219,8	173,259	106,32	74,679	7,29	83,979	107,105	68,22	32,395	200,1	305,164
1998	321,439	233,268	188,666	191,04	107,508	51,33	128,743	152,985	49,2	114,669	204,42	354,826
1999	357,833	310,8	238,793	135,48	38,285	64,8	131,471	143,499	28,38	69,843	266,76	315,115
2000	378,231	231,884	213,001	213,9	58,125	79,2	132,928	145,421	53,7	103,726	247,77	298,096
2001	315,642	259,112	190,743	134,79	81,158	93,3	103,881	149,544	47,4	48,515	260,31	329,189

Figura.IV.3.8. Promedio mensual de grados equivalentes (22º C) por año durante el periodo 1995-2001.
Fuente. IEA.
Elaboración. Propia.

A continuación se aplicará la *Prueba de Causalidad de Granger* para evidenciar estadísticamente la relación de causalidad entre el consumo de electricidad en el mercado residencial y los grados equivalentes de temperatura en la Comunidad Autónoma de Andalucía.

Prueba de causalidad de Granger

Para finalizar el estudio de la influencia de la temperatura en la demanda de energía eléctrica residencial en la C.A.A. se ha aplicado la prueba de causalidad de Granger. Esta prueba estudia la relación entre dichas variables. En el estudio que se presenta, tablas IV.3.2 y IV.3.3, se analiza la relación de *causalidad* que existe entre los grados equivalentes mensuales y la demanda de electricidad, es decir, se contrasta, en primer lugar si QMENSUAL *causa* GEQMEN22 y debajo se analiza lo contrario, es decir, si GEQMEN22 *causa* QMENSUAL. Como es sabido, la denominada *causalidad de Granger* no es la causalidad en el sentido ordinario de este término sino que es, en cierto modo, una relación de precedencia. Los contrastes se aplican sobre uno y dos retardos y los resultados obtenidos coinciden en rechazar la ausencia de relación causal de la temperatura en la demanda residencial de electricidad en la Comunidad Autónoma de Andalucía.

Tabla.IV.3.2. Prueba de Causalidad de Granger (un retardo).

Sample: 1995:01 2000:12 Lags: 1			
Null Hypothesis:	Obs	F-Statistic	Probability
QMENSUAL does not Granger Cause GEQMEN22	71	2.85962	0.09541
GEQMEN22 does not Granger Cause QMENSUAL		21.1112	1.9E-05

Fuente. IEA e INM.
Elaboración. Propia.

Tabla.IV.3.3. Prueba de Causalidad de Granger (dos retardos).

Sample: 1995:01 2000:12 Lags: 2			
Null Hypothesis:	Obs	F-Statistic	Probability
QMENSUAL does not Granger Cause GEQMEN22	70	1.46806	0.23791
GEQMEN22 does not Granger Cause QMENSUAL		14.4490	6.4E-06

Fuente. IEA e INM.
Elaboración. Propia.

Por lo tanto, como ya se había observado en las figuras IV.3.3, IV.3.4 y IV.3.5, se puede afirmar que la temperatura es un factor muy influyente en la demanda de energía eléctrica en la Comunidad Autónoma de Andalucía durante el periodo mensual que abarcan los años 1995-2001.

IV.3.7. Análisis de los datos mensuales

Las características generales expuestas de los datos sugieren que la temperatura ambiental debe ser un factor a tener en cuenta a la hora de estudiar el consumo de electricidad en la Comunidad Autónoma de Andalucía. Los resultados del análisis de los datos muestran una correlación considerable entre el consumo mensual de electricidad y los grados equivalentes mensuales. Esta correlación es todavía más intensa si se compara el consumo mensual con la variable que tiene en cuenta el retardo de la facturación en la temperatura, Geqmen22(-1), produciéndose en este último supuesto correlaciones mínimas de 0,605 en el año 2001 a correlaciones máximas de 0,877 en el año 1996, siendo ambas significativas, ya sea al nivel de significación del 5%, caso de la primera, o bien al 1% (aún más fuerte), caso de la segunda correlación.

Tras la realización de un análisis cuidadoso de los datos y con apoyo informático de las programas estadísticos SPSS y Eviews, se proponen dos modelos básicos lineales en los parámetros para apreciar la relación existente entre el consumo mensual de electricidad y la temperatura ambiental. Éstos son el modelo lineal y el modelo cuadrático. El modelo lineal se define como:

$$E=\alpha+\beta X$$

donde E es el consumo de electricidad y X es Geqmen22(-1). El modelo lineal ha sido aplicado de forma separada a cada uno de los años estudiados en este trabajo. Los resultados de la regresión, tabla IV.3.4, indican que el modelo lineal es estadísticamente significativo con un grado de confianza del 95%. El consumo de electricidad se encuentra positivamente correlacionado con los grados equivalentes (22ºC), teniendo los coeficientes de regresión del modelo lineal los signos esperados. Los datos del año 1996 son los que mejor se ajustan al modelo lineal obteniéndose el coeficiente de determinación ajustado más elevado siendo su valor de 0,74664.

También se ha investigado un modelo cuadrático para describir la relación entre el consumo mensual de electricidad y los grados equivalentes mensuales retardados (22º C), con el objetivo de encontrar un modelo que se ajuste mejor a los datos que el modelo lineal anteriormente considerado. En el modelo de regresión cuadrático, el cuadrado de la variable Geqmen(-1)(22º C) se incorpora como variable independiente y el modelo ha sido aplicado individualmente a cada uno de los datos anuales. Los resultados del modelo cuadrático de regresión indican que el modelo en su conjunto es más significativo al nivel del 5% para los años 1997 y 1998 no sucediendo tal circunstancia para el resto de los años. Del estudio de este modelo es preciso destacar que los coeficientes de regresión de la variable, X (Geqmen22(-1)) no resultan significativos al nivel del 5%, sucediendo lo mismo para gran parte de los años analizados respecto a la variable X^2. Por estas razones, no se propone el modelo cuadrático como ajuste adecuado al consumo de electricidad en función de los grados equivalentes retardados. El coeficiente de determinación ajustado mejora para los años 1996, 1997, 1998, 1999, 2000 y 2001, mientras que para el año 1995 decrece muy débilmente. El coeficiente de determinación ajustado, para todos los años considerados, varía desde 0,3223 a 0,8211. También se observa, al aplicar el modelo cuadrático, una disminución de la desviación estándar del modelo.

Además este análisis año a año, se ha realizado un análisis global para el conjunto de las observaciones desde el punto de vista de la bondad de ajuste a los datos de los modelos lineal y cuadrático (ver tabla IV.3.5). Los resultados de dicho estudio son similares a los obtenidos año a año, ya que el modelo lineal sigue siendo un modelo válido desde el punto de vista de su significación, tanto individual como de forma globalmente, y desde el punto de vista de los signos de los coeficientes de regresión también se obtienen los esperados. Sin embargo, el modelo cuadrático no resulta significativo respecto a sus coeficientes de regresión. Por lo tanto, se propone el modelo lineal como expresión funcional válida para el ajuste del consumo de electricidad en función de los grados equivalentes retardados.

Tabla.IV.3.4. Resultados de la regresión lineal y cuadrática para cada uno de los años comprendidos en el periodo 1995-2001.

Modelos	Lineal							Cuadrático						
Años	1995	1996	1997	1998	1999	2000	2001	1995	1996	1997	1998	1999	2000	2001
$\sigma_{regresión}$	30, 7075	27,9538	37,4031	42,7522	59,6001	91,7811	95,2957	32,5225	24,5390	28,7278	37,9382	56,4559	91,1705	78,2978
Significancia	0,0225	0,0002	0,0004	0,0046	0,0041	0,0333	0,0373	0,0864	0,0003	0,0002	0,0048	0,0079	0,0691	0,0137
α	377,9726	385,1157	395,4426	419, 5925	458,4370	479,4286	535,2826	379,3241	424,1153	444,3657	494,7979	516,5665	567,3395	738,6811
σ_{α}	14,8930	16,4800	19,0763	27,9339	31,3636	55,2968	56,1456	18,0674	24,3258	22,7087	46,2984	49,5785	99,1456	96,1498
Significancia	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0000
β_{lineal}	0,23761	0,49241	0,57204	0,53321	0,54313	0,66104	0,71126	0,2002	-0,1377	-0,3492	-0,5115	-0,4011	-0,4978	-2,4042
$\sigma_{\beta lineal}$	0,08646	0,08518	0,11159	0,14680	0,14699	0,26803	0,29626	0,2606	0,3247	0,3377	0,5586	0,6596	1,1202	1,3149
Significancia	0,0225	0,0002	0,0004	0,0046	0,0041	0,0333	0,0373	0,4645	0,6815	0,3282	0,3837	0,5582	0,6672	0,1007
$\beta_{cuadrático}$	-	-	-	-	-	-	-	0,0001	0,0018	0,0028	0,0028	0,0024	0,0028	0,0087
$\sigma_{\beta cuadrático}$	-	-	-	-	-	-	-	0,0008	0,0009	0,0010	0,0014	0,0016	0,0027	0,0036
Significancia	-	-	-	-	-	-	-	0,8819	0,0773	0,0201	0,0866	0,1771	0,3146	0,0392
$\overline{R}^2$	0,39584	0,74664	0,69676	0,52571	0,53493	0,31603	0,30220	0,3223	0,80476	0,8211	0,6265	0,5827	0,3251	0,5289
F-valor modelo	3,3780	33,4157	26,2745	13,1926	13,6525	6,0825	5,7639	3,3780	23,6700	26,2452	10,2260	8,6802	3,6493	7,1757

Fuente.IEA.
Elaboración. Propia

Tabla.IV.3.5. Resultados de la regresión para los modelos lineal y cuadrático durante el periodo de 1995 a 2001.

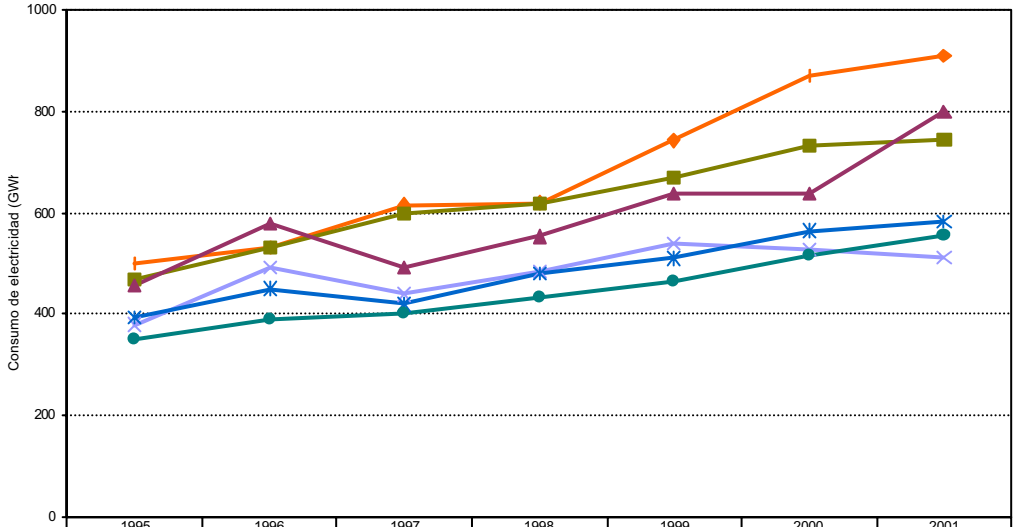
Modelos	Lineal 1995-2001	Cuadrático 1995-2001
$\sigma_{regresión}$	92,56146	91.21831
Significancia	0,0000	0,0000
α	428,432860	469,146123
σ_{α}	19,372529	29,181866
Significancia	0,0000	0,0000
β_{lineal}	0,597725	-0,075754
$\sigma_{\beta lineal}$	0,101081	0,378434
Significancia	0,0000	0,8418
$\beta_{cuadrático}$	-	0,001884
$\sigma_{\beta cuadrático}$	-	0,001021
$\overline{R}^2$	0,29290	0,31328
Significancia	-	0,0688
F-valor modelo	34,96727	19,70378

Fuente. IEA.
Elaboración. Propia.

La variación en los modelos de regresión anuales son debidas básicamente a los cambios en la climatología que tienen lugar de un año a otro y a otros factores aleatorios. Las figuras IV.3.9 y IV.3.10 representan la variación del consumo de electricidad para los primeros y últimos seis meses del año a lo largo de la serie de 1995-2001.

Cuando se compara las ecuaciones de regresión para cada uno de los meses es preciso preguntarse acerca de la igualdad de sus pendientes. Una vez calculadas las estimaciones de los coeficientes de regresión mínimo cuadráticos se ha realizado un contraste múltiple de igualdad de pendientes en las rectas[9] (se detalla en el anexo, tablas A.IV.82 a A.IV.89). La hipótesis de paralelismo de las pendientes de las rectas de regresión, tanto en el primer semestre como en el segundo semestre, no puede ser rechazada al nivel de confianza del 95%. El hecho de que las líneas resulten ser paralelas puede ser interpretado como un ratio de consumo de electricidad constante respecto a los grados equivalentes para todos los meses de cada uno de los dos semestres, es decir, si la temperatura experimenta un incremento unitario positivo (o negativo) el incremento (disminución) de la energía consumida es aproximadamente constante para todos los meses del semestre considerado.

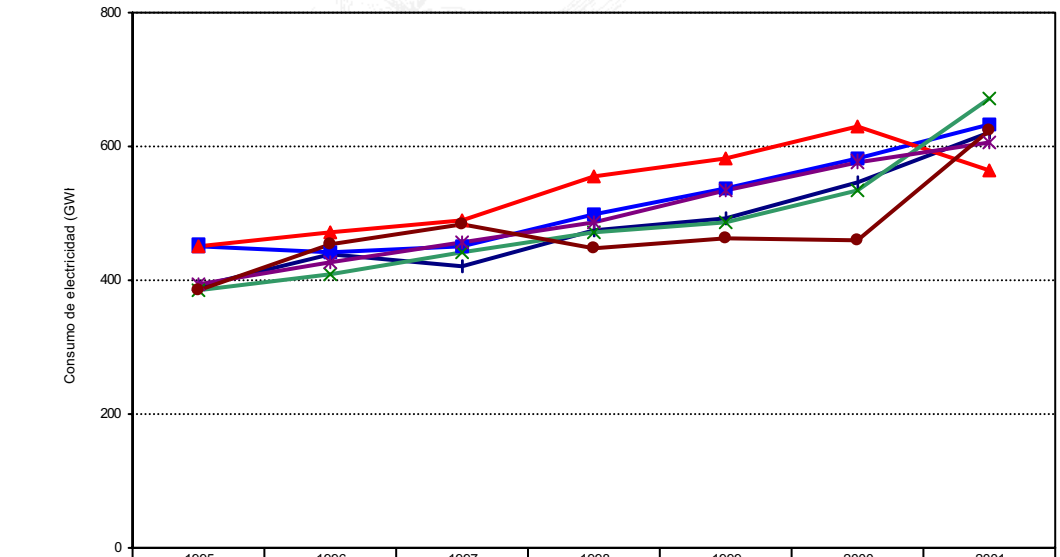
[9] Graybill, F.A.(1994).



	1995	1996	1997	1998	1999	2000	2001
ENERO	499,832	532,132	616,711	620,174	743,323	872,774	910,87
FEBRERO	467,945	531,256	598,998	617,589	668,587	733,867	744,322
MARZO	456,622	577,911	490,834	553,649	638,898	638,929	801,157
ABRIL	379,482	491,764	439,37	483,974	539,113	528,273	512,637
MAYO	393,463	450,744	419,637	479,265	510,449	565,406	582,845
JUNIO	349,939	390,87	400,203	432,599	464,911	517,105	556,618

Años

Figura.IV.3.9. Patrón de consumo de electricidad durante los seis primeros meses durante el periodo 1995-2001.
Fuente. IEA.
Elaboración. Propia

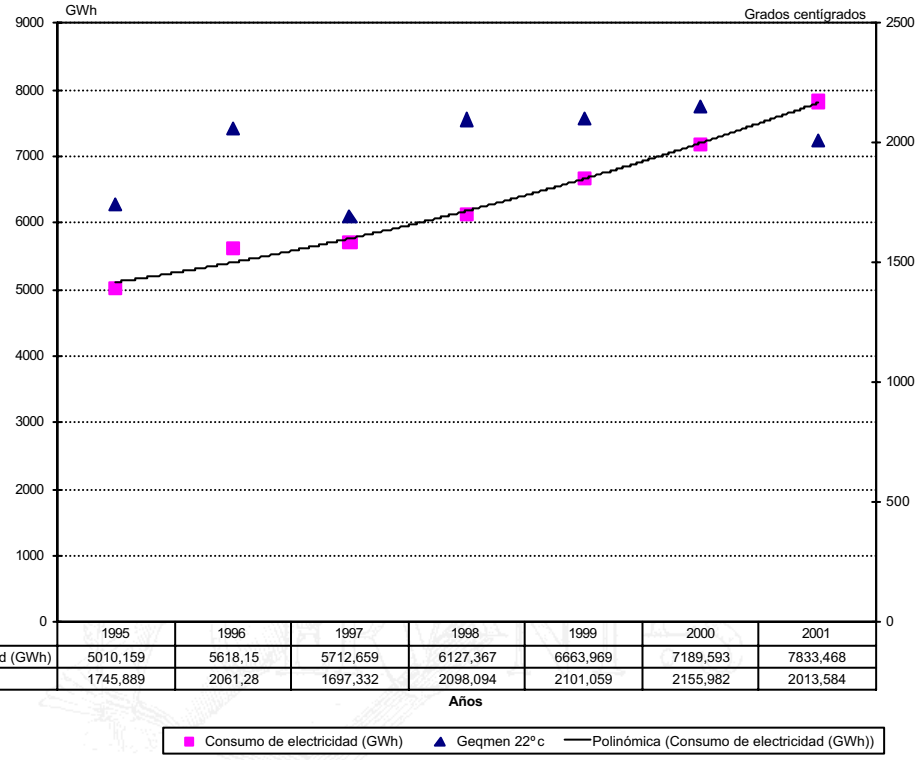


	1995	1996	1997	1998	1999	2000	2001
JULIO	392,364	438,534	421,357	474,189	493,642	546,841	621,894
AGOSTO	452,671	441,767	451,661	499,934	537,38	583,91	634,137
SEPTIEMBRE	451,687	472,738	491,063	556,699	582,353	629,992	564,881
OCTUBRE	386,325	408,907	443,147	472,288	487,616	534,245	673,151
NOVIEMBRE	393,611	427,244	456,135	487,641	534,832	578,46	606,917
DICIEMBRE	386,218	454,283	483,543	449,366	462,865	459,791	624,039

Años

Figura.IV.3.10. Patrón de consumo de electricidad durante los seis últimos meses durante el periodo 1995-2001.
Fuente. IEA.
Elaboración. Propia.

La figura IV.3.11 muestra la variación del consumo de electricidad durante el periodo investigado, 1995-2001 y los correspondientes Geqmen 22ºC. La línea continua representa el *mejor* ajuste a los datos mediante mínimos cuadrados.



	1995	1996	1997	1998	1999	2000	2001
Consumo de electricidad (GWh)	5010,159	5618,15	5712,659	6127,367	6663,969	7189,593	7833,468
Geqmen 22º c	1745,889	2061,28	1697,332	2098,094	2101,059	2155,982	2013,584

Figura.IV.3.11. Variación del consumo de electricidad y los correspondientes grados equivalentes (22ºC).
Fuente. IEA.
Elaboración. Propia.

La figura IV.3.12 muestra la media de los Geqmen22 mes a mes durante el periodo 1995-2001. Se observa que la evolución de los Geqmen22 disminuye desde los 333,19 en Enero hasta los 37,5 durante Septiembre, para ascender a 309 en Diciembre.

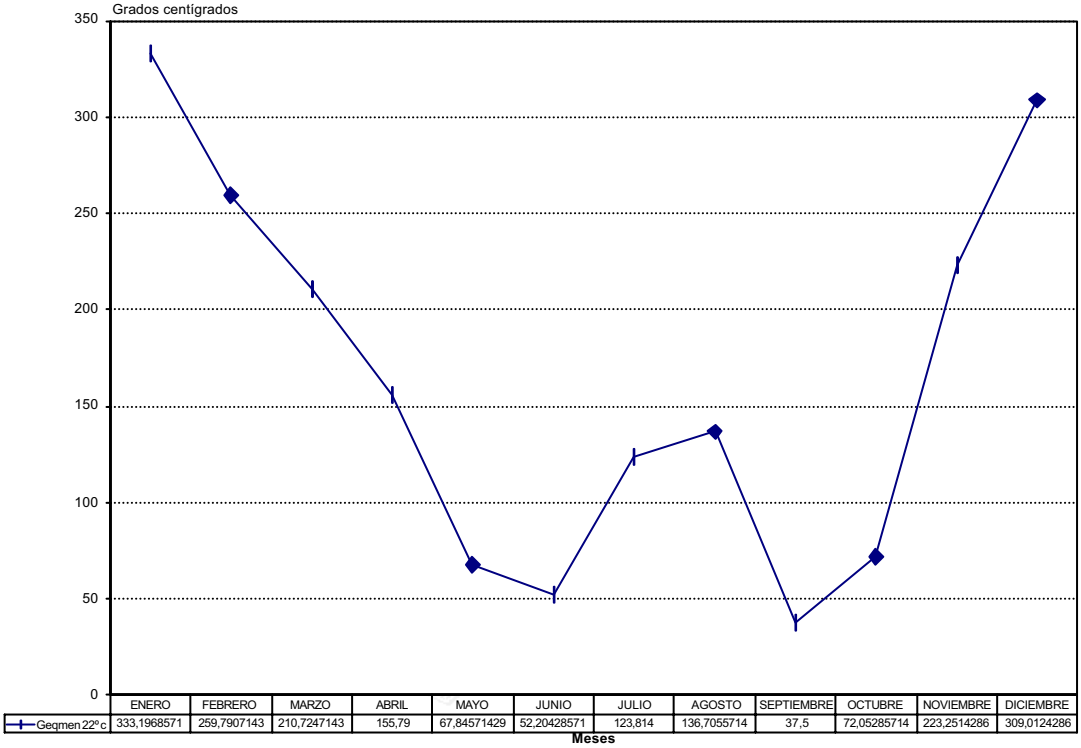


Figura.IV.3.12. Media mensual de los grados equivalentes (22º C), 1995-2001.
Fuente. IEA.
Elaboración. Propia.

IV.4. MODELOS DE DEMANDA EN FUNCIÓN DE LA TEMPERATURA EN LA C.A.A. DATOS MENSUALES

En el apartado anterior del capitulo se ha comprobado que la temperatura constituye un factor fundamental en la demanda de energía eléctrica en la C.A.A. Por ello sería conveniente estudiar dicho factor con mayor grado de detalle. Este apartado persigue estimar un modelo basado en la temperatura ambiente media de la C.A.A., para lo cual se han construido modelos cuyas variables explicativas son los grados equivalentes mensuales correspondientes a las distintas temperaturas en un rango de valores de estas últimas desde 15º C hasta 25º C (ver Anexo A.IV.2 a A.IV.80). Respecto a la forma funcional de tales modelos, para cada temperatura se han construido distintas familias de modelos pertenecientes al tipo de regresión dinámica a fin de evaluar la *robustez* de tales familias frente a eventuales modificaciones *subjetivas* de lo que *debería* considerarse temperatura *neutra* cuestión en la que se ha dicho este trabajo no profundiza sino que se ha elegido la temperatura de 22º C por motivos estrictamente objetivos de mayor correlación con la demanda, lo cual no impide que el modelo cuyo estudio se detalla resulte muy robusto frente a las modificaciones subjetivas mencionadas.

La regresión dinámica es un método de predicción que combina los datos que se presentan en forma de serie cronológica con los efectos de variables explicativas o indicadores *adelantados o precursores*. Un modelo de *regresión dinámica* consta de una variable respuesta (endógena o dependiente), Y_t, cuyo valor en el tiempo *t* está relacionado linealmente a sus propios valores retardados y de un vector de variables exógenas, X_t, que está formado por valores correspondientes al período *t* y/o precedentes de sus variables componentes X_{it}, suponiéndose habitualmente que las observaciones involucradas en la mencionada relación tienen lugar a intervalos igualmente espaciados en el tiempo. Una hipótesis básica en este modelo es que mientras la variable respuesta está influenciada por las variables independientes, éstas no lo están por la variable endógena, lo que equivale a decir que en este método sólo se

consideran modelos uniecuacionales. Una forma general de expresar un modelo regresión dinámica es:

$$P(B)Y_t = \beta X_t + \varepsilon_t \qquad (15)$$

donde *P(B)* es un polinomio en el operador de retardo *B,* similar a los que figuran en los modelos ARIMA, β es un vector cuyos componentes son los coeficientes de regresión de las X_{it} (o de sus valores retardados) y ε_t representa el valor del término aleatorio de error en el tiempo *t.* Un modelo de regresión dinámica se diferencia de los modelos ARIMA anteriormente considerados en los siguientes importantes aspectos:

- en un modelo de regresión dinámica existen una o más variables explicativas que *conducen* el proceso y que en el caso estudiado son los *grados equivalentes mensuales,* es decir, el producto de la temperatura media del correspondiente mes por el número de días del mismo.

- en un modelo de regresión dinámica no existen términos de medias móviles como los que frecuentemente forman parte de los modelos ARIMA.

- un modelo de Box-Jenkins no pretende ser más que una relación lineal que efectúe la predicción (en el sentido amplio del término que incluye el *ajuste* de los datos históricos), por lo que se suele decir que no es más que una "caja negra" (*black box*) o "filtro" que relaciona el *ruido blanco* con los datos observados, centrándose exclusivamente en el análisis de los datos de la serie, sin entrar en discusión (aunque se trate de un modelo Box-Jenkins multivariante y cuente con variables independientes relacionadas con la respuesta mediante una *función de transferencia*) de cuales son las razones por las que se han incluido en el modelo determinadas variables dependientes y no otras. Este último aspecto es muy importante en los modelos de regresión dinámica ya que se aborda frecuentemente dicha discusión recurriendo a argumentos de la Teoría Económica o, en general, razonando o motivando la inclusión de las variables explicativas en el modelo

- en un modelo Box-Jenkins *estocástico multivariante* puede existir realimentación (*feedback*) desde las variables dependientes hacia la variable dependiente lo cual no ocurre en los modelos de regresión dinámica.

Como resumen de las mencionadas diferencias podría decirse que los modelos de regresión dinámica *extienden* en un sentido (inclusión de variables exógenas) los modelos ARIMA comentados anteriormente y los *reducen* en otro (omisión de términos de medias móviles)

El modelo clásico de regresión lineal supone que los errores ε_t correspondientes a dos observaciones cualesquiera de la serie de datos no están correlacionados entre sí; sin embargo, la correlación entre los residuos de una misma serie (autocorrelación) es muy frecuente en la práctica, como puede evidenciar el correspondiente correlograma de tales residuos, lo cual significa que los datos históricos contienen información útil para predecir el futuro: así, por ejemplo, si en una serie de datos mensuales la función de autocorrelación muestral evidencia que es estadísticamente significativo el retardo de orden 12, entonces se pueden utilizar los datos del mismo mes del año anterior para disminuir el error de predicción. Además, la presencia de autocorrelación en los residuos puede indicar que se deberían añadir al modelo inicialmente considerado nuevos retardos de la variable dependiente y/o de las variables explicativas. Una solución alternativa para mejorar el modelo dinámico inicial, que generalmente requiere la estimación de menos parámetros que la anterior, la constituye el modelo generalizado de Cochrane-Orcutt en el que a la ecuación (15) anterior se añade la siguiente:

$$R(B)\varepsilon_t = \varepsilon_t - \rho_1\varepsilon_{t-1} - ... - \rho_r\varepsilon_{t-r} = a_r \qquad (16)$$

que expresa la autocorrelación de los residuos originales mediante un proceso autorregresivo especificado por el polinomio *R(B)* en el operador de retardo *B*, siendo a_t un proceso de *ruido blanco*. La ecuaciones (15) y (16) anteriores pueden integrarse como una única ecuación:

$$R(B)[P(B)Y_t - \beta X_t] = a_t \qquad (17)$$

Los procedimientos más habituales de estimación de los ρ_i en modelos lineales (Cochrane-Orcutt, Prais-Winsten o Hildreth-Lu) son iterativos y suelen estar incorporados en la mayoría de los programas de ordenador

- En la componente autorregresiva la permanencia del efecto se recoge mediante la incorporación, como variable explicativa de la variable explicada retardada algún periodo. En los modelos construidos, los distintos tipos de retardos que se aplican son los de un periodo, doce periodos (QMENSUAL(-12)) y por el efecto contagio (QMENSUAL(-11)), (QMENSUAL(-10)) y (QMENSUAL(-1)).

- En la componente de retardos distribuidos: son aquéllos en los que las variables exógenas son las que ejercen un efecto prolongado en el tiempo sobre la variable explicada y en el presente estudio siempre será la serie denominada QMENSUAL. En los modelos construidos sólo se considera un único retardo para la variables que expresa los grados equivalentes mensuales de temperatura, Geqmen^{o}C(-1).

- Ante la presencia de autocorrelación en todos los modelos, la mayoría de las estimaciones deben llevarse a cabo incorporando a los residuos del modelo original un esquema AR(12) o bien AR(24).

El número total de modelos que se han construido es de setenta y nueve para el total de temperaturas consideradas (ver anexo A.IV.2 a A.IV.80). Sin embargo, en este apartado sólo se va a analizar desde el punto de vista de su estructura y especificación el modelo 22.4, es decir, el modelo que considera como variable explicativa una temperatura neutra de 22°C según la expresión funcional que se detalla seguidamente. La razón por la que se ha elegido la temperatura neutra de 22°C es porque es la temperatura que está más correlacionada con la demanda de electricidad doméstica en la C.A.A. (ver anexo A.IV.90). Por otra parte, la expresión funcional del modelo elegido es la

más *robusta* de las aplicadas a todas las distintas temperaturas. La expresión del modelo 22.4 se expresa como:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-12} + \beta_3 x_t + \varepsilon_t$$

o bien

$$QMENSUAL=\alpha+\beta_1 QMENSUAL(-1)+\beta_2 QMENSUAL(-12)+\beta_3 GEQMEN22+\varepsilon_t$$

t=1,...,n

La estimación del modelo se refleja en la tabla IV.4.1, donde se puede observar que el modelo están bien definido, tanto global como individualmente para cada uno de los coeficientes de regresión. La capacidad explicativa de todas las variables es alta, pues explican el 95,24% de la variabilidad de la demanda doméstica de electricidad mensual de electricidad en la C.A.A.

Tabla.IV.4.1. Estimación del modelo 22.4 de demanda de energía eléctrica en la C.A.A.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	-0.103296	0.026912	-3.838294	0.0006
QMENSUAL(-12)	1.321546	0.039532	33.42978	0.0000
GEQMEN22	-0.081990	0.023397	-3.504336	0.0015
C	-54.77066	16.16541	-3.388138	0.0020
AR(12)	-0.629931	0.111498	-5.649717	0.0000
AR(24)	-0.367929	0.091287	-4.030487	0.0004

R-squared	0.952456	Mean dependent var	555.0258
Adjusted R-squared	0.944532	S.D. dependent var	93.94524
S.E. of regression	22.12571	Akaike info criterion	9.182369
Sum squared resid	14686.41	Schwarz criterion	9.446289
Log likelihood	-159.2826	F-statistic	120.1982
Durbin-Watson stat	2.106337	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.33+.90i	.33 -.90i	.17 -.94i	.17+.94i
	-.17+.94i	-.17 -.94i	-.33 -.90i	-.33+.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90 -.33i	-.90+.33i	-.94+.17i	-.94 -.17i

Fuente. IEA e INM.
Elaboración. Propia.

A continuación se analiza la estructura del modelo, cuyos residuos se contrastan a través de pruebas de autocorrelación, normalidad, heterocedasticidad, entre otras, así como la correcta especificación del mismo.

Análisis de los residuos

El correlograma de los residuos para dieciseis coeficientes se presenta en la tabla IV.4.2 Se observa que los coeficientes de autocorrelación y autocorrelación parcial de los residuos oscilan en torno a cero. Las probabilidades asociadas al estadístico *Q* de *Ljung-Box* es superior a 0,05, por lo que no se sitúan en la región de rechazo de la hipótesis nula, es decir, no se puede rechazar que exista ausencia de autocorrelación de primer orden entre los residuos al 95% de confianza.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

Tabla.IV.4.2. Correlograma de los residuos. Modelo 22.4.

Sample: 1998:01 2000:12
Included observations: 36
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. *| . |	. *| . |	1	-0.142	-0.142	0.7936	
. | . |	. *| . |	2	-0.043	-0.065	0.8694	
. | . |	. | . |	3	-0.029	-0.046	0.9038	0.342
. *| . |	. *| . |	4	-0.132	-0.150	1.6469	0.439
. |* . |	. |* . |	5	0.128	0.084	2.3748	0.498
.**| . |	.**| . |	6	-0.207	-0.205	4.3259	0.364
. |* . |	. | . |	7	0.087	0.036	4.6842	0.456
. *| . |	.**| . |	8	-0.184	-0.230	6.3425	0.386
. |* . |	. |* . |	9	0.090	0.071	6.7520	0.455
. *| . |	.**| . |	10	-0.099	-0.221	7.2674	0.508
. *| . |	. *| . |	11	-0.128	-0.114	8.1648	0.518
. |* . |	. | . |	12	0.196	0.015	10.360	0.409
. *| . |	. *| . |	13	-0.142	-0.099	11.551	0.398
. | . |	.**| . |	14	-0.021	-0.222	11.580	0.480
. | . |	. | . |	15	0.020	0.002	11.606	0.560
. |* . |	. | . |	16	0.104	0.002	12.350	0.578

Fuente. IEA e INM.
Elaboración. Propia.

Análisis de los residuos al cuadrado

Respecto a la presencia de heterocedasticidad condicional autorregresiva (ARCH) del modelo analizado se puede observar, tabla IV.4.3, que la probabilidad (p-valor) del estadístico experimental es superior al 5%, por lo tanto no se puede rechazar la ausencia de comportamiento ARCH.

Tabla.IV.4.3.Correlograma de los residuos al cuadrado. Modelo 22. 4

Sample: 1998:01 2000:12
Included observations: 36
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
.**| . |	.**| . |	1	-0.225	-0.225	1.9796	
. | . |	. *| . |	2	-0.024	-0.078	2.0019	
. |* . |	. |* . |	3	0.106	0.088	2.4713	0.116
. | . |	. | . |	4	-0.036	0.008	2.5281	0.283
. *| . |	. *| . |	5	-0.108	-0.114	3.0417	0.385
. | . |	. *| . |	6	-0.028	-0.098	3.0783	0.545
.**| . |	.**| . |	7	-0.236	-0.292	5.6997	0.337
. | . |	. *| . |	8	0.052	-0.071	5.8337	0.442
. |* . |	. |* . |	9	0.126	0.133	6.6403	0.467
. *| . |	. | . |	10	-0.065	0.045	6.8592	0.552
. | . |	. | . |	11	0.009	-0.024	6.8636	0.651
. |*** |	. |**. |	12	0.332	0.274	13.145	0.216
. *| . |	. | . |	13	-0.120	-0.003	13.994	0.233
. | . |	. *| . |	14	-0.030	-0.100	14.052	0.297
. | . |	. | . |	15	0.032	-0.038	14.117	0.366
.**| . |	. *| . |	16	-0.199	-0.171	16.834	0.265

Fuente. IEA e INM.
Elaboración. Propia.

Normalidad de los residuos

Los estadísticos descriptivos que analizan la forma de la distribución, figura IV.4.1, reflejan valores próximos a los deseables en una distribución normal. Esto es, el coeficiente de asimetría resulta ser próximo a cero y el coeficiente de curtosis próximo a tres. Bajo la hipótesis nula de una distribución normal, el estadístico de Jarque-Bera, sigue una ley χ^2_2, cuyo valor crítico en tablas es 5,991 (al 5% de significación) superior al valor experimental que es 1,67. Por lo tanto, no se puede rechazar la hipótesis de normalidad de los residuos al 95% de confianza.

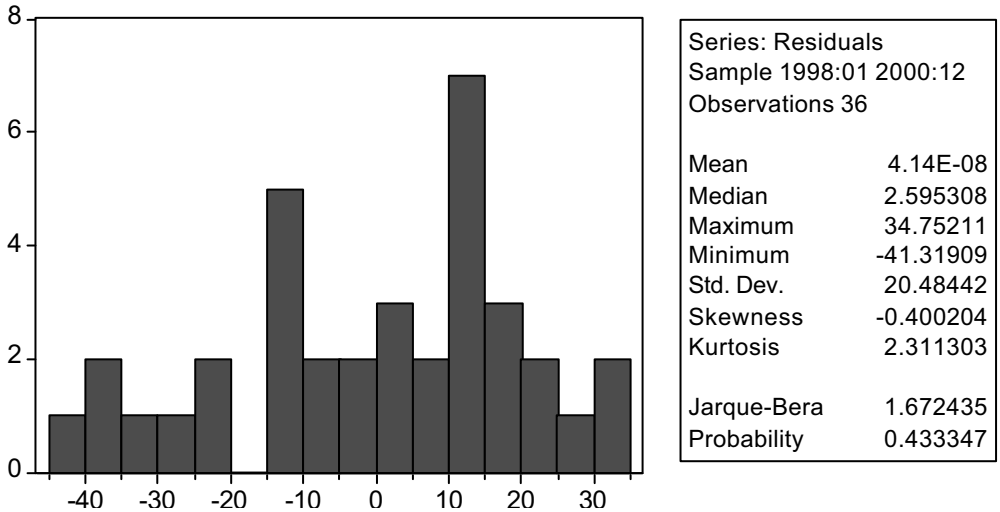


Figura.IV.4.1. Histograma de probabilidad. Modelo 22.4.
Fuente. IEA e INM.
Elaboración. Propia.

Correlación Serial de los residuos

Una alternativa a la aplicación del estadístico Q para analizar la correlación serial consiste en la aplicación del test LM. Este contraste diagnostica la presencia de autocorrelación serial. En la tabla IV.4.4 se observa que con valores tan elevados en las probabilidades asociadas (*p*-valores) a los estadísticos F y LM no hay evidencia para rechazar la hipótesis nula que establece la ausencia de autocorrelación de orden dos en los residuos, al nivel del 5%. Además, en la salida de la regresión que se adjunta, donde se incluyen como variables explicativas los dos retardos de los residuos se observa que el estadístico *t* (correspondiente al segundo retardo) no es significativo para las variables residuales.

Tabla.IV.4.4. Contraste de autocorrelación serial. Modelo 22.4.

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.421874	Probability		0.659916
Obs*R-squared	1.053085	Probability		0.590644
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	0.004028	0.028650	0.140608	0.8892
QMENSUAL(-12)	-0.009692	0.042308	-0.229072	0.8205
GEQMEN22	0.005866	0.024874	0.235832	0.8153
C	1.745157	16.60909	0.105072	0.9171
AR(12)	-0.013386	0.117735	-0.113693	0.9103
AR(24)	-0.021638	0.096035	-0.225311	0.8234
RESID(-1)	-0.184781	0.209616	-0.881521	0.3855
RESID(-2)	-0.067352	0.206931	-0.325482	0.7472
R-squared	0.029252	Mean dependent var		4.14E-08
Adjusted R-squared	-0.213435	S.D. dependent var		20.48442
S.E. of regression	22.56482	Akaike info criterion		9.263791
Sum squared resid	14256.80	Schwarz criterion		9.615685
Log likelihood	-158.7482	F-statistic		0.120535
Durbin-Watson stat	1.833839	Prob(F-statistic)		0.996239

Fuente. IEA e INM.
Elaboración. Propia.

Heterocedasticidad autorregresiva condicional en los residuos

A continuación, tabla IV.4.5, se realiza una regresión de los residuos al cuadrado en función de sus propios valores retardados dos periodos incluyendo también un término constante. De nuevo, el estadístico F es un test de la variable omitida que recoge el efecto conjunto de todos los residuos al cuadrado retardados. Los valores de las probabilidades asociadas a ambos estadísticos no permiten rechazar la hipótesis nula, es decir, los residuos no presentan comportamiento ARCH.

Tabla.IV.4.5. Contraste de comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 22.4.

ARCH Test:				
F-statistic	0.922062	Probability		0.408331
Obs*R-squared	1.909024	Probability		0.385000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Sample(adjusted): 1998:03 2000:12
Included observations: 34 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	501.2004	125.6741	3.988096	0.0004
RESID^2(-1)	-0.246801	0.183096	-1.347934	0.1874
RESID^2(-2)	-0.080745	0.160571	-0.502864	0.6186

R-squared	0.056148	Mean dependent var	381.4565
Adjusted R-squared	-0.004746	S.D. dependent var	425.8961
S.E. of regression	426.9056	Akaike info criterion	15.03510
Sum squared resid	5649699.	Schwarz criterion	15.16978
Log likelihood	-252.5967	F-statistic	0.922062
Durbin-Watson stat	1.934269	Prob(F-statistic)	0.408331

Fuente. IEA e INM.
Elaboración. Propia.

Heterocedasticidad en los residuos

El contraste de heterocedasticidad de White aplicada al modelo 22.4, tabla IV.4.6, indica la existencia de heterocedasticidad en las varianzas no especificando las variables que lo originan (al 95% de confianza).

Tabla.IV.4.6. Contraste de heterocedasticidad de los residuos. Modelo 22.4.

White Heteroskedasticity Test:				
F-statistic	3.514304	Probability		0.009786
Obs*R-squared	15.15578	Probability		0.019079
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample: 1998:01 2000:12				
Included observations: 36				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3396.462	2957.281	-1.148508	0.2601
QMENSUAL(-1)	5.622478	7.658520	0.734147	0.4688
QMENSUAL(-1)^2	-0.005674	0.006182	-0.917732	0.3663
QMENSUAL(-12)	8.763382	11.65791	0.751711	0.4583
QMENSUAL(-12)^2	-0.006058	0.010672	-0.567685	0.5746
GEQMEN22	-6.553650	3.242073	-2.021438	0.0525
GEQMEN22^2	0.018626	0.008593	2.167508	0.0385
R-squared	0.420994	Mean dependent var		407.9558
Adjusted R-squared	0.301199	S.D. dependent var		473.7855
S.E. of regression	396.0576	Akaike info criterion		14.97366
Sum squared resid	4548988.	Schwarz criterion		15.28157
Log likelihood	-262.5259	F-statistic		3.514304
Durbin-Watson stat	2.572681	Prob(F-statistic)		0.009786

Fuente. IEA e INM.
Elaboración. Propia.

Errores de especificación por variables omitidas

Hasta ahora, se ha ido comprobando la idoneidad del modelo, es decir, el cumplimiento de las hipótesis relacionadas con la parte estructural y aleatoria. Para analizar si el modelo presenta o no errores de especificación se realiza el contraste de Ramsey, tabla IV.4.7, cuya hipótesis nula establece que los coeficientes de las potencias de la variable estimada son nulos al 95%. Sin embargo, si se analiza de forma individual cada uno de los coeficientes de regresión, se observa que resultan ser no significativos al 95%, lo que expresa la necesidad de su ausencia en el nuevo modelo planteado.

Tabla.IV.4.7. Contraste de Ramsey. Modelo 22.4.

Ramsey RESET Test:				
F-statistic	4.669848	Probability		0.017775
Log likelihood ratio	10.36269	Probability		0.005620

Test Equation:
Dependent Variable: QMENSUAL
Method: Least Squares
Sample: 1998:01 2000:12
Included observations: 36
Convergence achieved after 35 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	-0.099817	0.104482	-0.955343	0.3476
QMENSUAL(-12)	-0.153375	0.140988	-1.087859	0.2859
GEQMEN22	0.050593	0.106518	0.474975	0.6385
C	610.6458	1050.043	0.581544	0.5655
FITTED^2	0.000933	0.000469	1.988193	0.0566
FITTED^3	-7.63E-08	4.25E-07	-0.179648	0.8587
AR(12)	0.359183	0.251509	1.428112	0.1643
AR(24)	0.570817	0.274408	2.080178	0.0468

R-squared	0.964348	Mean dependent var	555.0258
Adjusted R-squared	0.955435	S.D. dependent var	93.94524
S.E. of regression	19.83227	Akaike info criterion	9.005628
Sum squared resid	11012.93	Schwarz criterion	9.357521
Log likelihood	-154.1013	F-statistic	108.1953
Durbin-Watson stat	2.117661	Prob(F-statistic)	0.000000

Inverted AR Roots	1.00	.93+.25i	.93 -.25i	.86 -.50i
	.86+.50i	.68 -.68i	.68+.68i	.50 -.86i
	.50+.86i	.25 -.93i	.25+.93i	.00 -1.00i
	-.00+1.00i	-.25 -.93i	-.25+.93i	-.50 -.86i
	-.50+.86i	-.68+.68i	-.68 -.68i	-.86 -.50i
	-.86+.50i	-.93 -.25i	-.93+.25i	-1.00

Fuente. IEA e INM.
Elaboración. Propia.

Una vez analizada la naturaleza de los residuos y la estabilidad del modelo, se completa dicho análisis con la evaluación del poder predictivo del modelo 22.4 Por ello, para evaluar la demanda de energía eléctrica residencial en la C.A.A. se construyen los intervalos de confianza al 95% de probabilidad (tabla IV.4.8) en el año 2001.

Tabla.IV.4.8. Intervalos de confianza (95%) del modelo 22.4. Año 2001.

Modelo Dinámico	L. I. (2,5%)	Estimación	L.S.(97,5%)	Valor Real
Enero	950,572	990,160	1029,747	910,870
Febrero	739,300	779,098	818,895	744,322
Marzo	671,158	710,958	750,758	801,157
Abril	536,407	576,207	616,007	512,637
Mayo	586,792	626,592	666,393	582,845
Junio	500,139	539,939	579,739	556,618
Julio	562,439	602,239	642,038	621,894
Agosto	606,688	646,488	686,288	634,137
Septiembre	694,978	734,779	774.578	564,881
Octubre	525,977	565,777	605,578	673,151
Noviembre	582,491	622,291	662,091	606,917
Diciembre	428,159	467,960	507,759	624,039

Fuente. IEA e INM.
Elaboración. Propia.

La figura IV.4.2 refleja la evolución de la demanda doméstica mensual de electricidad en la C.A.A. en el periodo 1999-2000 (con valores históricos 1995-2000), así como la previsión mensual de demanda en el año 2001, indicando la línea de color rojo la estimación puntual de demanda de electricidad y las líneas de color azul los límites inferior y límites superior con una confianza del 95%.

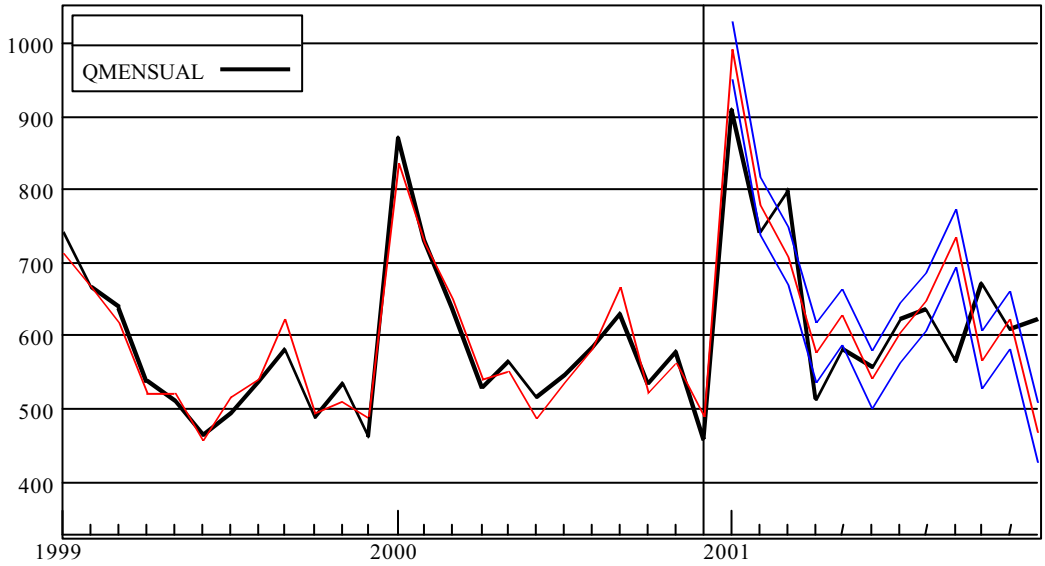


Figura.IV.4.2.Valores reales, ajustados y límites de confianza de demanda mensual de electricidad en la C.A.A. Modelo 22.4. Periodo 1995-2000.
Fuente. IEA e INM.
Elaboración. Propia.

Para finalizar el estudio de modelos de demanda mensual de electricidad habría que decir que de entre los modelos no dinámicos y dinámicos (vistos anteriormente en la sección IV.2) parecen ser candidatos aceptables los modelos de alisado exponencial y BJ-1, respectivamente. Por otra parte, si te tiene en cuenta el factor temperatura, el modelo que se ha contemplado en la presente sección considera una temperatura neutra de 22º C.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

IV.5. COMBINACIÓN DE PREVISIONES

La *combinación de predicciones* es un tema que ha sido objeto de numerosas investigaciones. Como indicativo de la importancia actual de este concepto baste comentar que la prestigiosa revista *International Journal of Forecasting,* en Diciembre de 1989, dedicó un número monográfico a este tema, denominándolo "Combining Forecasts". Con anterioridad a dicho número monográfico, y sobre todo con posterioridad al mismo, la combinación de predicciones aparece como objeto de investigación con considerable frecuencia, además de en la mencionada revista, en otras asimismo especializadas en el campo de los pronósticos tales como *The Journal of Forecasting* y *The Journal of Business Forecasting* , así como en otras publicaciones cuyo campo de atención no es exclusivamente la predicción, tales como *Operational Research Quarterly, Journal of Business Research, Journal of Marketing Research, Business Economics, Management Science*,etc. La bibliografía de Clemen(1989) sobre combinación de predicciones recogía 209 artículos en 1989, y el *Social Science Citation Index* cita 115 trabajos relevantes sobre este tema publicados entre 1988 y 2000. A una predicción que resulta como combinación de varios pronósticos individuales se le denomina también una predicción *compuesta.*

Entre los autores que proclaman las ventajas de la combinación de predicciones se encuentra Clemen(1989) para el que existen pocas dudas de que la combinación de previsiones incrementa la *exactitud* de los resultados obtenidos, incluso cuando las predicciones objeto de combinación no se limiten a las de tipo estadístico, sino que, según este autor, esta técnica admite extensiones a la combinación de predicciones estadísticas con otras de tipo subjetivo.

Asimismo**,** Collopy, F. y Armstronng, J.S.(1992), comentan que el 83% de expertos profesionales de la predicción consideran que la combinación de predicciones mejora los resultados de los pronósticos individuales.

Entre algunos de los factores que pueden influir en la proclamada mayor *exactitud* de la combinación de predicciones se encuentra la posible compensación de los errores de las predicciones individuales causados por:

- la necesidad, a veces, de pronosticar no directamente aquello en lo que se está interesado (por ejemplo, la demanda) sino un sucedáneo o sustitutivo de ello (la facturación, el número de pedidos, etc...) dada la carencia o la inaccesibilidad de datos *directos* sobre la variable objeto de interés.
- los errores de medida
- la estructura inestable de los datos
- la utilización de modelos que se limitan a minimizar errores pasados sin ulterior comprobación *ex-post*.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

Otro argumento que se aduce en favor de la combinación de predicciones es el siguiente: al despreciar un método de pronóstico individual por no ser “el mejor” de acuerdo con el criterio de selección que se haya elegido, es posible que se esté perdiendo valiosa información contenida en el procedimiento descartado debido, entre otros motivos, a que algunas variables explicativas o relaciones entre ellas contenidas en el método descartado, pueden no estar incluidas en la supuesta “mejor” predicción.

Entre las razones esgrimidas por los no partidarios de la combinación de predicciones se encuentran:

- pérdida del concepto de significación estadística.
- creencia de que un determinado método individual es superior.
- preferencia por un “modelo completo” que integre toda la información relevante.

Es posible que las posibles ventajas de la *combinación de previsiones* hubiese sido un tema menos controvertido si se hubiera establecido, desde el

principio, una clara distinción entre la *exactitud* y *el riesgo* de una predicción *compuesta*, a fin de comparar estas características con las análogas que presentan los pronósticos individuales. En cuanto a la *exactitud,* es fácil comprender que los argumentos a favor de la combinación de predicciones sean exclusivamente empíricos, ya que desde el punto de vista analítico, supuesto existente un procedimiento óptimo que proporcionase resultados con error de predicción nulo (en este punto es indiferente que tales errores fuesen *a priori* o *a posteriori,* referidos a una muestra concreta o de forma genérica), no sólo sería superfluo, sino también perjudicial, promediarlo con métodos alternativos de predicción subóptimos, ya que el resultado sería asimismo subóptimo. Por tanto las ventajas, en cuanto a *exactitud,* de la predicción compuesta proclamadas por los partidarios de la combinación de predicciones han de entenderse apoyadas exclusivamente en la posible evidencia *empírica*

Diferente es la cuestión de la ventaja de combinar predicciones individuales desde el punto de vista del *riesgo*, o medida de la dispersión. Como otros procedimientos de adecuada *diversificación del riesgo*, en campos muy diversos, dada la incertidumbre de nuestros pronósticos *ex-ante*, la combinación de predicciones es un medida de prudencia que, además cuenta con respaldo analítico. Y así como no existe evidencia analítica de que la combinación de predicciones logre siempre un resultado más *exacto* que el que proporcionan los pronósticos individuales, en cambio, Bessler, D.A. y Brandt J.A.(1981), demostraron que la varianza del error de la predicción combinada no es mayor que la menor de las varianzas de las previsiones individuales.

Condiciones idóneas de las predicciones a combinar

Los pronósticos individuales que se desea combinar deben ser insesgados, es decir, los errores de predicción de cada uno de ellos no deben ser sistemáticamente positivos o negativos, ya que la combinación de predicciones no garantiza la eliminación del sesgo.

La combinación de predicciones mejora la exactitud de los pronósticos individuales cuando estos últimos contienen información útil e independiente. Sería deseable que los errores de predicción individuales estuviesen correlacionados negativamente entre sí con objeto de que se produjese la mutua cancelación de los mismos, pero con mucha frecuencia se presenta la circunstancia de la autocorrelación positiva de los errores en los pronósticos individuales.

Asimismo es deseable que la causa de que dos modelos de pronósticos individuales proporcionen diferentes resultados de predicción sea que *extraen diferentes factores predictivos de los mismos datos o que tales modelos utilizan diferentes variables.*

En opinión de Wilson y Keating (1998), en general, una predicción combinada tendrá un error de predicción menor (medido por el RMSE*), a menos que los modelos de predicción individual ofrezcan resultados de precisión similar y que sus errores de predicción estén muy correlacionados*.

En cuanto al número óptimo de modelos a combinar Armstrong(2001) comenta que a partir del quinto método que se combina la ganancia (en términos de reducción del error de predicción) es insignificante.

Un factor importante en la combinación de predicciones es la adecuada elección del coeficiente de ponderación, o peso, que va a afectar a cada método individual de predicción. Armstrong(2001) es partidario de utilizar pesos iguales para los modelos individuales de predicción, salvo que exista una fuerte evidencia sobre las ventajas de determinados métodos en una situación concreta, de acuerdo con su teoría de que el coeficiente de ponderación que afecta a cada método de pronóstico individual debe ser elegido *ex-ante*, de forma mecánica, ya que, en su opinión, si tales coeficientes se eligen subjetivamente, una vez conocidos los resultados de cada método individual, pueden introducirse sesgos.

Por el contrario, Bates y Granger (1969) fueron los primeros en sugerir de que, a pesar de que no existe evidencia de que los métodos con mejor comportamiento en el pasado vayan a seguir siendo mejores en el futuro, si se asigna mayores pesos a los métodos que han resultado ser más exactos con los datos del pasado los resultados de la predicción global posiblemente serían mejores.

Una aproximación diferente para elegir el *mejo*r sistema de pesos es la de Wilson y Keating(1998) que permite que éstos se adapten a la estructura de los datos modificando su valor de período a período. Nelson, C.R.(1984) propuso con anterioridad que los coeficientes de ponderación óptimos entre dos predicciones individuales f_1 y f_2 se determinasen mediante regresión lineal.

Procedimiento utilizado para combinar distintos métodos de predicción

En este trabajo se seguirá el procedimiento de asignación de pesos a los pronósticos individuales propuesto por Granger(1980). Como es sabido, la media cuadrática de los residuos correspondientes a n periodos de un pronóstico calculado por el método j viene expresada por:

$$MSE_j = \frac{1}{n} SCE_j = \frac{\sum_{t=1}^{n} e_{j,t}^2}{n}$$

Si se supone que f_n y g_n son dos predicciones para el valor x_{n+1}, usando dos procedimientos diferentes. En adelante se designará a f_n y g_n simplemente por f y g respectivamente, pues ambas predicciones se supondrán calculadas utilizando los mismos n datos históricos. Una nueva predicción para dicho valor, combinando las predicciones anteriores sería:

$$c = kf + (1-k)g\,, \quad 0 < k < 1$$

cuyo error de predicción en el periodo t=n+1 vendría expresado por:

$$ec = x_{n+1} - c = kef + (1-k)eg$$

La varianza de *ec* viene expresada por:

$$var(ec) = k^2\, var(ef) + (1-k)^2\, var(eg) + 2k(1-k)\, cov(ef, eg) \qquad (18)$$

donde $cov(ef, eg)$ es la covarianza de las series *ef* y *eg*.

El valor del parámetro *k* que minimiza la varianza de la serie *ec* es:

$$k = \frac{var(eg) - cov(ef, eg)}{var(eg) + var(ef) - 2\, cov(ef, eg)} \qquad (19)$$

Cuando no exista correlación entre *ef* y *eg* esta expresión se reduce a :

$$k = \frac{var(eg)}{var(eg) + \mathrm{var}(ef)} \qquad (20)$$

Si el valor de *k* obtenido en (19), se sustituye en (18) se demuestra que la varianza del error de la serie combinada nunca será mayor que $mín\{var(ef), var(eg)\}$.

No obstante, la fórmula (19) no puede ser utilizada para calcular *k*, ya que los valores poblacionales de la varianza y la covarianza que figuran en (19) no son conocidos. Por tanto, estas varianzas se estiman mediante los errores de predicción, empleando este término en el sentido amplio que incluye el periodo de *ajuste* tal como se comentó en una sección anterior. De este modo, para calcular el valor de *k* para el pronóstico combinado del periodo *n*+1 resulta:

$$k = \frac{\sum_{t=1}^{n} (eg_t)^2}{\sum_{t=1}^{n} \left[(eg_t)^2 + (ef_t)^2 \right]} \qquad (21)$$

Como se puede observar, el valor de *k* estará siempre comprendido entre 0 y 1.

Se puede generalizar esta técnica para combinar *p* pronósticos ($p \geq 2$) correspondientes a *p* métodos diferentes de predicción que proporcionarían el nuevo pronóstico compuesto:

$$c = k_1 f_1 + k_2 f_2 + \ldots + k_p f_p$$

Si $e_{j,t}$ es el error de predicción correspondiente a la predicción *j-ésima*, f_j y al periodo t, se define:

$$A_j = \left[\sum_{t=1}^{n} \left(e_{j,t} \right)^2 \right]^{-1}$$

$$B = A_1 + A_2 + ... + A_p$$

Se demuestra que la varianza de la predicción *compuesta* es mínima cuando el coeficiente de ponderación para el *j-ésimo* método de predicción es:

$$k_j = \frac{A_j}{B}; \quad \text{siendo} \quad \sum_{j=1}^{p} k_j = 1$$

En lo que sigue se aplica la combinación de predicciones a las predicciones obtenidas anteriormente para los siguientes métodos: alisado exponencial (AE), Holt-Winters (HW), ARIMA (BJ-1) y Regresión dinámica (R-22) (todos ellos objeto de estudio en apartados anteriores de este trabajo). Según el método de Granger que se acaba de exponer se obtienen los coeficientes de ponderación k_j, j=1, 2, 3, 4 que figuran en la última fila de la tabla IV.5.1 en la que en su columna situada a la izquierda se relacionan los meses del año 2001 que son objeto de pronóstico. En la tabla IV.5.1 se pueden observar las predicciones realizadas para cada uno de los métodos anteriormente mencionados para tales meses en las columnas (1), (2), (3) y (4) mientras que en la columna (5) se muestra el valor realmente observado para el mismo periodo. Los correspondientes valores de k_j pueden verse en la última fila de dicha tabla.

Tabla.IV.5.1. Predicciones de los métodos de previsión.

Mes	EstimaciónAE (1)	Estimacion HW (2)	EstimaciónBJ-1 (3)	EstimaciónR-22 (4)	Valor Real (5)
Enero	883,848	887,068	813,403	990,16	910,87
Febrero	735,235	764,977	748,115	779,098	744,322
Marzo	648,853	682,922	686,303	710,958	801,157
Abril	543,917	573,77	583,021	576,207	512,637
Mayo	565,292	589,001	58,284	626,592	582,845
Junio	507,041	534,89	527,368	539,939	556,618
Julio	531,713	566,375	568,178	602,239	621,894
Agosto	567,105	605,635	609,031	646,488	634,137
Septiembre	612,352	654,995	655,75	734,779	564,881
Octubre	518,396	555,184	560,942	565,777	673,151
Noviembre	559,909	598,375	594,256	622,291	606,917
Diciembre	459,812	502,638	539,302	467,96	624,039
RMSE	33,55	32,63	31,52	20,2	
MSE	1125,60	1064,71	993,51	408,04	
12*MSE	13507,23	12776,60	11922,12	4896,48	
InvSCE	0,00088841	0,00093922	0,00100653	0,00245074	
k_j	0,16810399	0,17771698	0,19045424	0,4637248	

Fuente. IEA e INM.
Elaboración. Propia.

En la tabla IV.5.2 se expresan los valores que resultan para la predicción compuesta o combinada de los métodos mencionados

Tabla.IV.5.2. Predicciones del método combinado.

Metodo Combinado	
Mes	Estimación
Enero	920,30
Febrero	763,31
Marzo	690,83
Abril	571,64
Mayo	501,36
Junio	531,11
Julio	577,52
Agosto	618,74
Septiembre	684,96
Octubre	555,00
Noviembre	602,21
Diciembre	486,34

Fuente. IEA e INM.
Elaboración. Propia.

En la tabla IV.5.3 se muestran los porcentajes de los errores de predicción para cada uno de los métodos considerados. En su última columna figura el valor de dichos errores para la predicción combinada pudiendo observarse que estos últimos no alcanzan valores tan extremos como los de las predicciones individuales.

Tabla.IV.5.3. Porcentaje de error de previsión.

Mes	EstimaciónAE (1)	Estimacion HW (2)	EstimaciónBJ-1 (3)	EstimaciónR-22 (4)	Estimación Combinada
Enero	2,967	2,613	10,700	-8,705	-1,036
Febrero	1,221	-2,775	-0,510	-4,672	-2,552
Marzo	19,011	14,758	14,336	11,259	13,770
Abril	-6,102	-11,925	-13,730	-12,401	-11,510
Mayo	3,012	-1,056	0,439	-7,506	13,979
Junio	8,907	3,904	5,255	2,996	4,581
Julio	14,501	8,927	8,637	3,161	7,135
Agosto	10,571	4,495	3,959	-1,948	2,427
Septiembre	-8,404	-15,953	-16,086	-30,077	-21,259
Octubre	22,990	17,525	16,669	15,951	17,551
Noviembre	7,745	1,407	2,086	-2,533	0,775
Diciembre	26,317	19,454	13,579	25,011	22,066

Fuente. IEA e INM.
Elaboración. Propia

IV.6. EL FUTURO DE LOS MODELOS DE PREVISIÓN DE DEMANDA ANUAL Y MENSUAL EN LA COMUNIDAD AUTÓNOMA DE ANDALUCÍA

El proceso de liberalización del sector eléctrico se inicia el 20 de enero de 1995, fecha de entrada en vigor de la Ley 40/1994, de 30 de diciembre, de Ordenación del Sistema Eléctrico Nacional (LOSEN). Aunque esta ley mantiene los principios fundamentales que había venido determinando la Normativa Eléctrica (obligación de suministro, planificación conjunta de las necesidades de capacidad a largo plazo, tarifa única y explotación unificada), sin embargo, introduce elementos de concurrencia y competitividad, así como una mayor transparencia en la delimitación de las distintas actividades que forman parte del negocio eléctrico. Entre los aspectos más destacables de esta Ley se encontraban, por su trascendencia futura:

- La creación de la Comisión Nacional del Sistema Eléctrico (CNSE), con el objeto de velar por la objetividad y transparencia del Sistema.
- La especialización de los negocios que integran el suministro de energía eléctrica.
- La obligación de segregar jurídicamente las actividades que, en su caso, las compañías que forman parte del Sistema Eléctrico desarrollen en el régimen especial de producción (energías renovables, biomasa, residuos y cogeneración), así como en el exterior y en otros sectores económicos.

A través de la firma del Protocolo, en diciembre de 1996, entre el Ministerio de Industria y Energía y las empresas eléctricas, se implanta la reforma estructural del sector con el fin último de contribuir a la mejora de la competitividad de la economía española y el cumplimiento de los objetivos de convergencia de la Unión Europea. En el mismo se diseñaron las bases del nuevo sistema eléctrico.

Asimismo, la nueva regulación responde a la necesidad de transponer a nuestro ordenamiento jurídico la Directiva 96/9/CE, de 19 de diciembre de

1996, sobre normas comunes para el mercado interior de electricidad. En el caso de España, la regulación ha ido más allá, introduciendo unos niveles de liberalización y competencia que superan ampliamente los márgenes mínimos establecidos por la Unión Europea, optando además por las alternativas más aperturistas en aquellos puntos en que la Directiva ofrece varias posibilidades.
Y así, por ejemplo, en lo que se refiere a la construcción de nuevas instalaciones de generación, se ha optado por un régimen de autorización administrativa, sin condicionar el uso de energías primarias, constituyendo el sistema menos restrictivo de los previstos en la Directiva.

La segunda etapa en este proceso de liberalización se concreta en la Ley Eléctrica de 1997 (Ley 54/1997, de 27 de noviembre), cuyo fundamento se encuentra en un acuerdo entre el Gobierno y las empresas eléctricas, que se recoge en el Protocolo Eléctrico de diciembre de 1996. Mediante la Ley 54/1997 se amplía la liberalización concedida por la LOSEN y se crea el mercado mayorista de electricidad. Introduce, además, las figuras del Operador del Sistema, que se encargaría de la red física del transporte, y del Operador del Mercado, que determina los intercambios de electricidad y los precios de mercado. El propósito liberalizador de la Ley 54/1997, de 27 de noviembre, no sólo se limita a acotar la actuación del Estado en el sector, sino que además introduce cambios importantes en su regulación. Así, en la generación eléctrica reconoce el derecho a la libre instalación y organiza su funcionamiento bajo el principio de la libre competencia, por lo que la retribución de esta actividad se asienta en la organización de un mercado mayorista. Aunque el transporte y la distribución se liberalizan a través de la generalización del acceso de terceros a las redes, su retribución continuará siendo fijada administrativamente para evitar posibles posiciones de dominio determinadas por la existencia de una única red. A fin de garantizar la transparencia de esta retribución, se establece la separación jurídica entre actividades reguladas y no reguladas en cuanto a su retribución económica. Este nuevo modelo liberalizado, que sólo cuenta con las limitaciones que se derivan de las singularidades del transporte y distribución de electricidad como actividades reguladas, devuelve a las empresas el protagonismo en la toma de decisiones.

Por último, el texto legal configura la comercialización en el marco de la libertad de contratación y elección de suministro.

La decisión de liberalizar el mercado eléctrico español, iniciada en enero de 1998, replanteó las reglas de juego: las centrales ya no producirían electricidad según un orden estipulado- primero las hidroeléctricas, luego las nucleares...-, sino en función de su rentabilidad, ya que un nuevo *pool* fijaría los precios de compra-venta de la electricidad según criterios de rentabilidad.

A partir del 1 de enero de 2003, todos los consumidores de electricidad y gas de España pueden elegir suministrador. Este es uno de los hechos más importantes que ha tenido lugar en todo el proceso de liberalización del sector eléctrico ya que, por primera vez, el consumidor-cualquiera que sea su nivel de consumo-podrá recurrir a la posibilidad de cambiar de empresa suministradora para negociar las condiciones económicas del servicio.

La liberalización del sector eléctrico hará posible que el suministrador elabore una especie de *traje a medida* para el usuario. Y la estrategia de las compañías pasa, no sólo por ofrecer al consumidor suministro eléctrico, sino una amplia oferta de productos y servicios. Los grandes operadores de electricidad y gas en España consideran que la mayor amenaza para sus negocios proviene de la estrategia multiproducto (*multiutility*), que en el caso del sector eléctrico está representada por Gas Natural comercializando electricidad, y en el de gas, por las eléctricas comercializando gas.

Según un estudio realizado por la patronal eléctrica Unesa a 5000 hogares, un 3,7% de los consumidores estarían dispuestos a cambiar de compañía eléctrica, aunque afirman que esto dependerá de las ofertas que reciban de las compañías eléctricas. Un 44,2% de los encuestados reconoce no tener muy clara a esta decisión; un 14,6% dice simplemente que es probable que cambie de compañía. Un 30% de los encuestados no estaría dispuesto a cambiar de compañía, mientras que un 7,5% no sabe o no contesta. También se realizaron otro tipo de preguntas las cuales se reflejan en la figura IV.6.1.

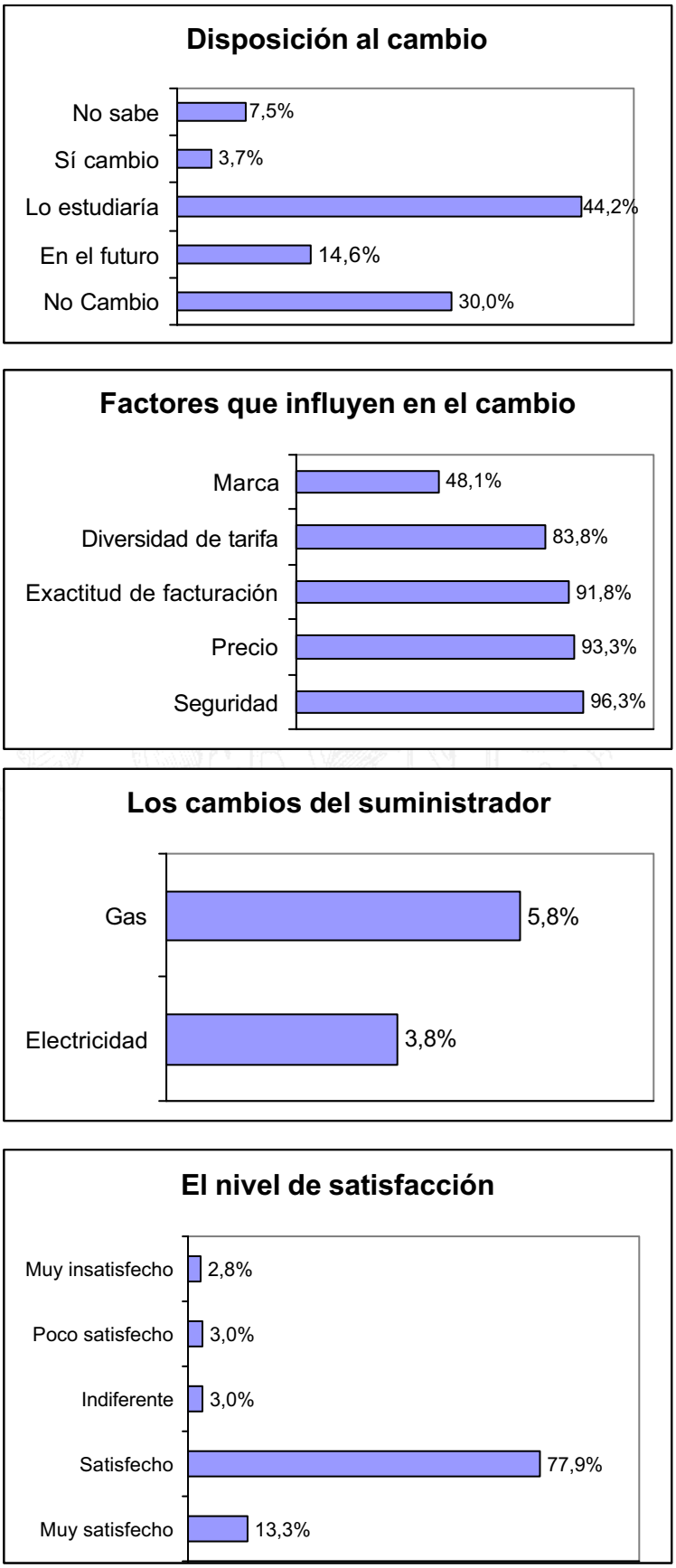


Figura.IV.6.1. El consumidor ante la apertura del sector eléctrico y gasista.
Fuente. Mercado de Dinero. Enero 2003. IBM Business Consulting Services.
Elaboración. Propia.

El sector eléctrico español mueve anualmente más de 13.500 millones de euros. Son sólo cuatro las compañías que dominan el negocio eléctrico agrupadas en la patronal Unesa: Endesa, Iberdrola, Unión Fenosa e Hidrocantábrico. El 80% del mercado está dominado por las compañías Endesa e Iberdrola.

En los capítulos III y IV de este trabajo se han construido modelos de previsión de demanda para la C.A.A. basados en datos históricos que abarcan el periodo 1984-1999 (modelos anuales) y 1995-2001 (modelos mensuales). Además se han construido intervalos de confianza que contienen a los verdaderos valores de la demanda de energía eléctrica. Desde el punto de vista de la validez de estos modelos en un mercado completamente liberalizado habría que decir que Endesa es la principal compañía suministradora, de energía durante el periodo objeto de estudio en la C.A.A. Ahora bien, con la posibilidad de que el sector doméstico pueda elegir libremente compañía suministradora puede suceder que los modelos basados en variables económicas anuales y que se refieren al conjunto de la comunidad, resulten más complejos y difíciles de construir.

El próximo capítulo tiene propósito reducir el grado de incertidumbre hacia cual será el comportamiento mayoritario del mercado residencial en la C.A.A. a partir del año 2003. En dicho capítulo se analiza y evalúa, desde el punto de vista de la eficiencia técnica, el comportamiento de los principales grupos eléctricos en España durante los años de proceso liberalizador.

IV.7. CONCLUSIONES DEL CAPÍTULO

En este capítulo se ha pretendido modelizar la previsión de demanda mensual de energía eléctrica de uso doméstico en la Comunidad Autónoma de Andalucía. Los datos históricos que se han utilizado corresponden a la mencionada demanda mensual en el periodo comprendido entre enero de 1995 y diciembre de 2000, ya que lo acontecido en el pasado, a través de sus regularidades, puede constituir un soporte adecuado para pronósticos con respecto al futuro, lo cual en este trabajo se investiga mediante las predicciones *ex -post*.

BIBLIOTECA VIRTUAL

El estudio del modelo de predicción se ha basado en el análisis de las series temporales. Para ello, en la sección IV.2 se ha realizado un estudio comparativo de modelos univariantes, es decir, modelos que explican la demanda de electricidad en función del comportamiento histórico de la propia variable aleatoria. Para llevar a cabo este estudio ha sido necesario previamente definir el esquema de integración de la serie, esto es, si se corresponde con un esquema aditivo o multiplicativo. Para modelizar la serie se han utilizado diversos procedimientos entre los que se encuentran la representación gráfica de la serie, la elaboración de un gráfico desviación típica-media, el análisis de la variabilidad de las diferencias y cocientes estacionales, estudio de correlogramas, transformación de Box-Cox, diferenciación de la serie y análisis de sus residuos. Todas estas pruebas aplicadas asignaron a la serie un esquema multiplicativo, es decir, detectaron que su variabilidad a lo largo del tiempo y aumenta con el nivel de la serie.

Estos modelos se han clasificado en modelos no dinámicos y modelos dinámicos. Por una parte, entre los modelos no dinámicos se han desarrollado diez modelos, estructurados según modelos sin tendencia, con tendencia y estacionalidad y modelos de ajuste a una curva. Por otra parte, entre los modelos dinámicos se han desarrollado cuatro modelos.

Habría que decir respecto a los resultados obtenidos de los modelos no dinámicos que, desde el punto de vista de su poder explicativo, el modelo de Holt-Winters es el que obtiene mejores valores en sus estadísticos, muy próximo al modelo de Alisado Exponencial. Por otra parte, desde el punto de vista predictivo, el modelo que mejores predicciones realiza de la demanda mensual de electricidad en la C.A.A. es el modelo de Alisado Exponencial (sin tendencia), ya que la mayoría de los estadísticos así lo expresan.

Los resultados predictivos obtenidos de los cuatro modelos dinámicos ARIMA, BJ-1, BJ-2, BJ-3 y BJ-4, desarrollados son, en términos generales, mejores que los comentados en párrafos anteriores. Por ello, en base a estas directrices se seleccionará un modelo dinámico, concretamente el modelo BJ-1, para la modelización de la demanda mensual de energía en la C.A.A. La estructura funcional del modelo BJ-1 se expresa como:

$$\nabla\nabla_{12} \ln q_t = (1 - 0{,}8103B)(1 - 0{,}7480B^{12})\varepsilon_t$$

La determinación cuantitativa del impacto de la temperatura sobre el consumo de energía eléctrica es especialmente importante en los países o regiones en los que la potencia instalada en reserva es reducida, pudiendo ser deficitario el suministro de energía cuando se presentan situaciones meteorológicas adversas. Este hecho ya se manifestó en el estado de California de los Estados Unidos en el verano de 2000 y la Comunidad Autónoma de Cataluña en las navidades de 2001. De esta forma, se ha considerado importante dedicar una sección a su estudio, en la cual se ha analizado su problemática desde el punto de vista de diversos autores. En esta sección también se han aplicado soluciones a la hora de trabajar con los datos que fueron suministrados por el Centro Meteorológico Territorial de Andalucía Occidental y Ceuta. Los resultados principales del estudio de la influencia de la temperatura en la demanda de energía eléctrica del mercado residencial en la C.A.A. llevan a confirmar la relación existente entre ambas variables y la consideración de la temperatura como variable exógena del modelo de predicción. Los resultados de la regresiones aplicadas, lineal y cuadrática,

favorecen al modelo lineal ya que sus coeficientes son significativos a nivel individual, no verificándose en el modelo cuadrático.

En la sección IV.4 se han estimado setenta y nueve modelos de regresión dinámica basados en la temperatura ambiente y la demanda. Respecto a la temperatura ambiente se han considerado modelos que abarcan temperaturas en el intervalo 15ºC-22ºC. Finalmente, se detalla el modelo correspondiente a 22ºC por ser la que mayor índice de correlación presenta con la demanda estudiada. Desde el punto de vista de su estructura y especificación el modelo elegido es robusto frente a posibles modificaciones de la temperatura neutra considerada. Las variables de este modelo denominado 22.4 son significativas tanto globalmente como de forma individual al nivel de confianza del noventa y cinco por ciento. El grado con el que la variabilidad de la variable demanda de electricidad es explicada por el resto de variables explicativas es superior al noventa y cinco por ciento. Existe ausencia de autocorrelación serial de primer y segundo orden entre los residuos del modelo. La normalidad de los residuos también se evidencia a través de la prueba de Jarque-Bera al noventa y cinco por ciento de nivel de confianza. En definitiva, se propone este modelo como expresión funcional para explicar y predecir la demanda mensual de energía eléctrica en la C.A.A.

En la sección IV.5 se estudia la combinación de predicciones de los mejores modelos estimados en secciones anteriores, obteniéndose resultados más favorables desde el punto de vista del error de estimación.

Finalmente, la sección IV.6 describe la evolución del sistema eléctrico español desde la entrada en vigor de la Ley 40/1994 hasta nuestro días. Además, en esta sección se realizan algunas reflexiones acerca de la permanencia de los modelos estimados, las cuales se esperan resolver en el siguiente capítulo.

BIBLIOTECA VIRTUAL

CAPÍTULO V
ANÁLISIS DE LA EFICIENCIA TÉCNICA DEL SECTOR ELÉCTRICO ESPAÑOL LIBERALIZADO

V.1. INTRODUCCIÓN

Actualmente los países desarrollados están llevando a cabo políticas energéticas orientadas a la consecución de tres objetivos básicos, lo cuales deben hacerse compatibles entre sí: estos son, el aumento de la competitividad de la economía con la mejora de la eficiencia energética, la integración de los objetivos medioambientales y la seguridad en el abastecimiento. Estos objetivos deben ser perseguidos en un mercado energético internacional caracterizado en los últimos años por el incremento de la oferta de energía, aunque con fluctuaciones en los precios y un crecimiento sostenido de la demanda.

En los últimos años, la eficiencia en los sectores de oferta energética, electricidad y gas, se ha fomentado a través de la liberalización creciente de los mercados, aunque experiencias como la de California (E.E.U.U.), durante el verano del año 2000, y la ocurrida en Cataluña, durante las navidades del año 2001, han mostrado que los riesgos de desequilibrio son más elevados durante la transición de la liberalización, por lo que en algunos países se está desarrollando una segunda generación de reformas normativas en este campo, siendo éste el caso de España.

En este capítulo se pretende realizar un estudio de la eficiencia técnica de las empresas eléctricas españolas de ciclo completo, las cuales tuvieron que hacer frente al nuevo entorno de carácter liberalizador que se inició en España a partir de la entrada en vigor de la nueva ley del Sector Eléctrico en 1998. El objetivo principal de este capítulo es analizar en qué medida los modelos de previsión de demanda, construidos en capítulos anteriores, van a mantener su estructura funcional en un horizonte a medio y/o largo plazo, ante los posibles desplazamientos que pudieran tener lugar en el mercado doméstico de la electricidad. Este análisis está dividido en seis secciones. La primera sección tiene un carácter introductorio. En la segunda sección se explica el objeto de la

elección del sector eléctrico para llevar a cabo el referido estudio. La tercera sección ofrece una síntesis descriptiva de algunos de los métodos más utilizados para analizar los dos tipos de eficiencia *clásicos*, esto es, la *eficiencia técnica* y la *eficiencia asignativa* o *económica*. En la cuarta sección se explica la metodología del *Análisis por Envoltura de Datos*, que ha de complementarse con el desarrollo de la técnica que puede verse en el Apéndice del trabajo. En la quinta sección tiene lugar el estudio empírico con el objetivo de conocer la evolución seguida, en materia de eficiencia técnica, por los principales grupos eléctricos españoles durante el periodo 1998-2001. Finalmente, en la sexta sección, se presentan las conclusiones más relevantes a las que se ha llegado con el citado estudio.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

V.2. EL SECTOR ELÉCTRICO COMO OBJETO DE ESTUDIO

Antes de definir las unidades objeto de estudio y evaluación, se deben realizar algunas precisiones acerca del Sistema Eléctrico Español, con objeto de encuadrar el estudio dentro de un determinado marco institucional.

El desarrollo de la Directiva Europea 96/92 y, en España, del Protocolo de diciembre de 1996 y de la Ley 54/97 supuso una transformación del sector eléctrico español desde una situación en las que las decisiones estratégicas y de inversión dependían básicamente de la Administración a otra caracterizada por la iniciativa empresarial. El objetivo de este cambio ha sido lograr gradualmente un Mercado Único Europeo de la Electricidad en un marco liberalizado, sin posiciones de dominio y, con respecto a las prácticas de la libre competencia, que permita un suministro de calidad a unos precios competitivos.

En España, el desarrollo de la Ley del Sector Eléctrico se manifestó en diversas direcciones, la más visible de las cuales fue la entrada en funcionamiento del mercado de la electricidad. La Ley 54/97 encomendaba a Red Eléctrica la constitución de la *compañía operadora del mercado español de electricidad* y su privatización. El papel de Red Eléctrica Española (R.E.E.)

como operador del sistema eléctrico fue fundamental ya que dió al funcionamiento del mercado el soporte necesario para garantizar que la seguridad del sistema no quedara comprometida en ningún momento. Además, también se creó la figura del *operador del mercado* (O.M.E.L.) que asume las funciones necesarias para realizar la gestión económica referida al eficaz desarrollo del mercado de producción de electricidad, gestionando el mercado mayorista de electricidad y ejerce las mismas respetando los principios de transparencia, objetividad e independencia.

A la entrada en vigor de la presente ley, se suprime la Comisión Nacional del Sistema Eléctrico como ente regulador del sistema eléctrico y se crea la Comisión Nacional de la Energía (C.N.E.) como ente regulador del funcionamiento de los sistemas energéticos, teniendo por objeto velar tanto por la competencia efectiva en los mismos como por la objetividad y transparencia de su funcionamiento, en beneficio de todos los sujetos que operan en dichos sistemas y de los consumidores. Algunos de los principales problemas están asociados con la implantación homogénea del modelo en todo el ámbito de la Unión Europea, como se manifestó en la cumbre de Estocolmo (tuvo lugar el 23 y 24 de marzo de 2001), donde surgieron divergencias entre los países respecto al ritmo adecuado para la consecución del Mercado Único de la Electricidad.

A pesar de las dificultades, el nuevo modelo avanza de forma irreversible en Europa y, especialmente, en España, donde los grados de desregulación y aumento de competencia del sector eléctrico son superiores a los alcanzados por la mayoría de los países de la Unión. En el año 2000, se produjeron nuevos impulsos liberalizadores en nuestro país, concretados especialmente en las medidas del Real Decreto Ley 6/2000, que, por un lado, adelantan nuevamente el calendario de apertura del mercado respecto a los plazos previamente establecidos y, por otro, establecen limitaciones para la instalación de nueva potencia en aquellas empresas que se considera poseen ya una fuerte posición en el mercado.

El tránsito gradual de uno a otro modelo está teniendo un desarrollo no exento de dificultades, por efecto de las singulares características del sector en cuanto al tipo de producto, a las especiales exigencias tecnológicas, a las inversiones requeridas y al papel que desempeña la energía eléctrica en el desarrollo de la sociedad.

Pero, además, en los últimos años se está haciendo cada vez más patente la necesidad de intensificar el esfuerzo inversor de las compañías eléctricas para afrontar la expansión de la economía española sin que se planteen problemas de estrangulamiento similares a los que ocurrieron en la experiencia californiana. Estas exigencias de nuevas inversiones vienen reforzadas, de una parte, por la práctica desaparición del tradicional exceso de capacidad que había mantenido el sector eléctrico y, de otra, por los niveles de calidad exigidos por los clientes, que requieren nuevas instalaciones capaces de garantizar elevadas cotas de servicios.

El análisis del Sector Eléctrico Español tiene especial interés por tres razones que se exponen a continuación:

- La liberalización del sector en el año 1998, comienza una nueva etapa para las empresas eléctricas españolas, ya que, hasta entonces, habían operado en un escenario de dependencia estricta hacia la Administración Pública. Esta nueva etapa abre una gran incertidumbre para las empresas, los consumidores y el ente público. Por ello se estudia el periodo que comprende los años 1998-2001 (ambos inclusive), en el que el sector ha experimentado importantes cambios liberalizadores entre otros. El año 2002 no ha sido analizado por no disponer de los datos necesarios (Hidrocantábrico celebró su junta general de accionistas el 26 de junio de 2003), además en el año 2001 es adquirida la empresa integrada dentro del grupo Endesa, Viesgo, por la empresa eléctrica italiana Enel. Esta circunstancia quebraría las hipótesis de partida para el análisis de la eficiencia del sector eléctrico en España.

- Las inversiones realizadas por las empresas eléctricas en España, en las que se pone de relieve el importante peso específico que tienen para la infraestructura del sector eléctrico español.
- La productividad y la eficiencia del sector eléctrico por ser un tema de indudable interés en España. El hecho de que las empresas eléctricas trabajen con eficiencia constituirá un factor decisivo en el posible desplazamiento de la cuota de mercado en el sector residencial de electricidad entre las distintas compañías eléctricas, pudiendo verse afectada la estructura funcional de los modelos considerados en capítulos anteriores.

V.3. METODOLOGÍA DEL ANÁLISIS DE LA EFICIENCIA TÉCNICA DEL SECTOR ELÉCTRICO ESPAÑOL

V.3.1. El concepto de eficiencia

En esta sección se comenzará por describir algunos de los métodos empleados para la valoración de la eficiencia de diferentes empresas. Así, se definirá los dos conceptos de eficiencia que, en este sentido, se suelen utilizar. Por un lado, se encuentra el concepto de eficiencia *asignativa*, que implica una correcta utilización de los recursos o inputs en cuanto a sus proporciones para producir productos y/o servicios u outputs. La eficiencia asignativa reflejará en qué medida los inputs se emplean en unas proporciones adecuadas dados sus precios y productividades en el margen. La eficiencia asignativa implica alcanzar el coste mínimo de producir un nivel dado de producto o servicio cuando se modifican las proporciones de los factores de producción utilizados, de acuerdo con sus precios y productividades marginales.

Por otro lado, el concepto de eficiencia técnica se entiende claramente al observar que la función de producción es, además de una relación entre inputs y outputs, una función frontera del conjunto de producción formado por todas las combinaciones de inputs-outputs técnicamente posibles. Por lo tanto, es posible realizar una distinción cualitativa entre unidades *eficientes*, aquéllas

situadas en la frontera de producción, e *ineficientes*, las que se encuentran por debajo de ella.

La eficiencia técnica es un concepto tecnológico que se concentra básicamente en los procesos productivos y en la organización de tareas. Centra su atención en las cantidades y no en los valores. Puede expresarse tanto en términos de productos u outputs como en términos de recursos o inputs. En el primer caso, indicaría el logro del máximo producto o servicio posible para una combinación específica de factores y, en el segundo caso, la cantidad mínima requerida de recursos, combinados en una determinada proporción, para un nivel dado de producto o servicio. Por lo tanto, la idea de la ineficiencia *técnica* está originado por problemas tecnológicos que implicarían la obtención de un número de productos u outputs excesivamente escasos, para una combinación de recursos o inputs dado.

El elemento fundamental de la eficiencia técnica, y distintivo con respecto a la eficiencia asignativa, es que parte de una proporción concreta de factores (que determina su tecnología) cuyo coste se minimiza o cuya producción se maximiza. La proporción de factores puede variar si, por ejemplo, se utiliza otra tecnología, pero no lo hace debido a los precios y las productividades marginales, como sucede en el caso de la eficiencia asignativa. Por eso, se comentaba en párrafos anteriores, que la eficiencia técnica pone el énfasis en cantidades y no en valores.

Este análisis se centrará en las aproximaciones de la ineficiencia técnica por no disponer de datos suficientes para desarrollar un estudio acerca de la ineficiencia asignativa.

V.3.2. La medición de la eficiencia

La medición de la eficiencia precisa, además de la distinción cualitativa ya señalada entre puntos eficientes e ineficientes según pertenezca o no a la frontera de producción, otra distinción de carácter cuantitativo que permita su medida a partir del concepto de distancia.

En el análisis que se realiza a continuación, de la medición de la eficiencia, se sigue el trabajo de Farrell (1957), tanto por su sencillez, como porque los estudios posteriores han tratado de mejorar la formulación teórica y los conceptos que básicamente habían sido establecidos por el mencionado autor.

Supongamos el caso más simple, tal como se refleja en la figura V.3.1, donde un conjunto de organizaciones o unidades producen un único output con un solo input. La curva OF es la función de producción que representa, como frontera, la máxima cantidad de output técnicamente posible para cada nivel de input o también, la cantidad mínima de input técnicamente necesaria para cada nivel de output. Se tendrán entonces dos posibles formas para definir esa distancia. La primera, en términos de outputs, distancia en el eje de output, es la cantidad máxima de output que se alcanzaría si la unidad fuera eficiente, dado el nivel de input consumido. Sin embargo, una mejor aproximación es utilizar el ratio $\frac{y_2 - y_1}{y_2}$ para representar la ineficiencia técnica y el ratio $\frac{y_1}{y_2}$ = 1- $\left(\frac{y_2 - y_1}{y_2}\right)$ para representar la eficiencia técnica. Es decir, $\frac{y_1}{y_2}$ para la unidad A . La segunda forma, es en términos de inputs, distancia en el eje del input, es la cantidad de input utilizado con respecto a la cantidad que sería necesario, dado un nivel de output, si la actividad fuera eficiente. Es decir, la eficiencia técnica en términos de input sería, $\frac{x_1}{x_2}$ en el caso de la unidad A.

Una de las ventajas que posee el medir la eficiencia técnica a través de este ratio es que el cociente no está expresado en ninguna unidad de medida. Este ratio de eficiencia técnica alcanzará valores comprendidos entre 0 y 1; cuanto mayor sea el valor que tome este ratio mayor será la eficiencia técnica de la unidad que se esté estudiando.

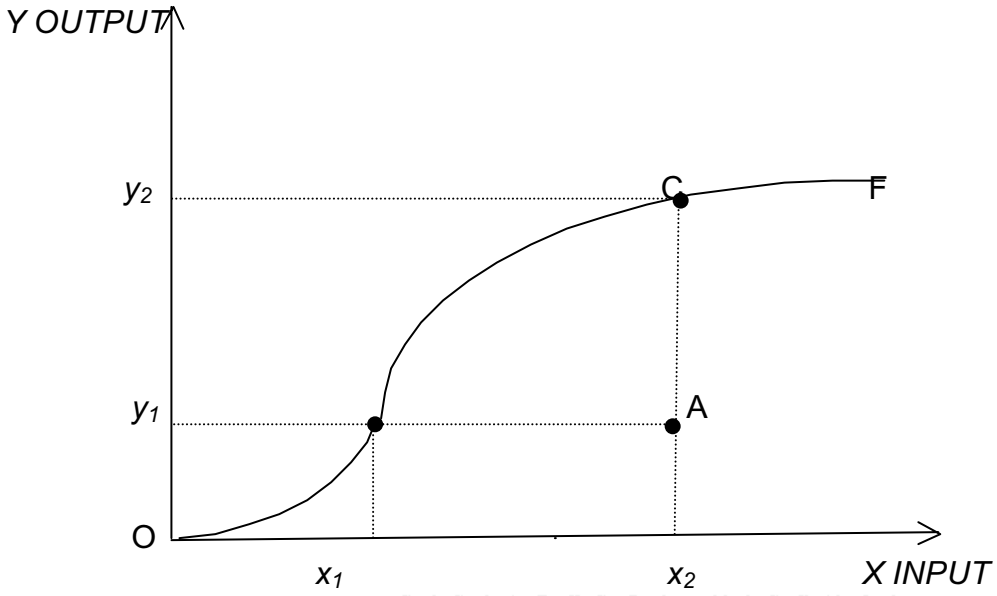


Figura.V.3.1. Formas alternativas de medición de la eficiencia.
Fuente. Pina, V. (2000)
Elaboración. Propia.

Aunque el caso explicado se limita al caso de un input-un output, los conceptos anteriores pueden generalizarse al caso de múltiples inputs y múltiples outputs.

V.3.3. Análisis de la eficiencia técnica del sector eléctrico español

Todo lo comentado anteriormente requiere un paso intermedio. Antes de calcular la eficiencia, a partir de una serie de datos referidos a un grupo de unidades, es necesario conocer el conjunto de producción y su correspondiente función frontera. Tanto uno como otro han de inferirse de los datos. Consideraciones prácticas exigen supuestos simplificadores; no se obtendrá, por tanto, *verdaderas* eficiencias sino eficiencias calculadas a partir de las mejores prácticas observadas. En general, no hay forma de saber si las medidas empíricas sobreestiman o subestiman la verdadera eficiencia técnica.

Pueden distinguirse dos metodologías principales y alternativas que difieren en la técnica utilizada para definir la frontera, éstas son, la aproximación paramétrica y la no paramétrica. A continuación se comentarán las características principales de cada una de ellas.

V.3.3.1. Aproximación paramétrica

La aproximación *paramétrica* de la frontera de producción es aquélla que precisa la especificación a priori de una forma funcional con parámetros constantes. Existen dos aproximaciones econométricas de estimar la frontera de producción. Por una parte, una de las aproximaciones consiste en calcular la frontera de producción de forma determinista, ya que considera la diferencia entre los niveles ideal y real de outputs como ineficiencia técnica, y no contempla un error aleatorio usualmente asociado a la estimación estadística. La ausencia de tal error aleatorio en las fronteras de producción deterministas implica que el ruido estadístico estaría explotado por la ineficiencia. Por otra parte, la segunda aproximación consiste en calcular las fronteras de producción estocásticas, donde tanto la ineficiencia técnica como el error aleatorio están presentes en el modelo y el problema asociado a la frontera de producción determinista es, además, evitable. Así, para medir la eficiencia técnica, la construcción de la frontera eficiente de forma estadística es un enfoque preferible al de la forma determinista.

Algunos de los tipos de enfoques paramétricos son, por ejemplo, la función de Cobb-Douglas y la función translogarítimica, estimándose sus parámetros de manera que las observaciones queden por encima o por debajo de la función. La frontera de producción de Cobb-Douglas ha sido una de las utilizadas con mayor frecuencia. Con respecto a esta función se mide la eficiencia que será distinta según la forma funcional especificada a priori[10]. La función de producción Cobb-Douglas es una función de producción homogénea de primer grado. Dicha función satisface las restricciones básicas para la frontera de producción, tales como la cuasi-concavidad y la monotonía.

[10] Muestra las diferentes combinaciones de recursos que la empresa puede adquirir, supuestos constantes los precios de los recursos o factores y dado el gasto total en el que se puede incurrir.

Además, esta función tiene la virtud de su simplicidad y validación empírica. La frontera de producción estocástica de Cobb-Douglas con dos inputs, capital (*K*) y trabajo (*L*), y un output (*Y*) puede expresarse como:

$$Y_{it} = \alpha K_{it}^{\beta_K} L_{it}^{\beta_L} e^{v_{it} - u_{it}} \qquad i = 1, ..., N; \; t = 1, ..., T$$

siendo $\beta_k + \beta_L = 1$ y donde el subíndice *i* hace referencia a la unidad evaluada y el subíndice *t* al instante de tiempo.

La utilización de la función de Cobb-Douglas impone una serie de restricciones, como es el hecho de que su aplicación presenta problemas serios, generalmente de multicolinealidad. Para resolver estos problemas se suele proponer el uso de la función *translogarítmica*:

$$Ln(Y_{it}) = \beta_0 + \beta_K \ln K_{it} + \beta_L LnL_{it} + v_{it} - u_{it}$$

donde $\beta_0 = Ln\alpha$. Los errores aleatorios v_{it} miden el error y otros factores aleatorios no imputables a la empresa, tales como las temperaturas, huelgas, suerte y accidentes en el nivel de output, junto con los efectos combinados de variables no especificadas hacia los inputs en la frontera de producción. Estos errores, v_{it}, se suponen independientes entre sí e idénticamente distribuidos según variables aleatorias con media cero, μ_v =0, y varianza unitaria, $\sigma_v^2 = 1$.

Por el contrario, los componentes residuales, u_{it}, de la *ineficiencia técnica* representan los efectos de las decisiones incurridas por la empresa, tales como una aplicación inadecuada de recursos, una comunicación ineficaz, escaso poder de decisión, excesiva información y otros. Estas *ineficiencias técnicas*, u_{it}, se suponen que son variables aleatorias independientes no negativas (no necesariamente idénticamente distribuidas) de distribuciones normales truncadas. Se supone que la distribución normal tiene media, μ_{it}, y varianza, σ_t^2. En otras palabras, $u_{it} \sim \left| N(\mu_{it}, \sigma_{it}^2) \right|$. La distribución normal truncada de u_{it}

estipula una ineficiencia técnica no negativa y que sólo dependa de algunas características específicas de la empresa.

Aún reconociendo su mayor flexibilidad, hay que decir que plantea también problemas. Éstos son de naturaleza fundamentalmente econométrica, tales como:

- El número de parámetros a estimar es elevado y, en el caso del sector eléctrico, el tamaño muestral seleccionado es reducido, lo que limita el número de grados de libertad.
- La inclusión de productos cruzados en el precio de los inputs genera una alta multicolinealidad en el modelo lo que, como es sabido, provoca que los estimadores tengan una elevada varianza.
- No facilita información sobre el nivel de eficiencia alcanzado, sino tan sólo informa de si una empresa se puede considerar eficiente o no.
- La hipótesis de que el término residual permite suministrar toda la información que concierne a la eficiencia de la actuación de la empresa en cuestión tampoco es correcta, debido a que dicho término residual puede reflejar perturbaciones aleatorias.

En definitiva, la utilización de estos métodos es ilustrativa para toda persona que se acerca al estudio de la eficiencia, pero en absoluto, es concluyente por lo que los resultados obtenidos deben considerarse con extrema cautela.

V.3.3.2. Aproximación no-paramétrica a la eficiencia técnica del sector eléctrico

En este tipo de aproximaciones no se especifica a priori una forma funcional, tal como ocurre en las aproximaciones paramétricas, sino unas propiedades formales que satisfacen los puntos del conjunto de producción: ej. libre disponibilidad de inputs y outputs, convexidad, etc. Los datos en este caso son *envueltos*, pero no por una función cuyos parámetros son estimados, sino determinando si cada punto observado puede considerarse que pertenezca o no a la frontera bajo los supuestos seleccionados. Esto se hace resolviendo un sistema de ecuaciones lineales definidas apropiadamente, una para cada observación, donde la frontera cambiará (y consecuentemente la eficiencia de cada unidad) en función de los supuestos establecidos.

Siguiendo esta última aproximación, Farell(1957), estableció las siguientes hipótesis:

- Libre disponibilidad de inputs y outputs: cualquier unidad productiva, posiblemente no observada, con el mismo nivel de output que cualquiera de las observadas, pero con más inputs, pertenece al conjunto de producción; cualquier unidad productiva, posiblemente no observada, que utiliza los mismos inputs que cualquiera otra observada, y que consigue menos outputs, pertenece al conjunto de producción.
- Convexidad: cualquier combinación convexa de puntos observados que pertenezcan al conjunto de producción pertenece también al mismo.
- Rendimientos de escala constantes.

Así, por ejemplo, en el ejemplo visto anteriormente, en el caso de un input y un output (figura V.3.1), se observa que las fronteras supuestas serían OF' (figura V.3.2).

En la frontera, OF', el supuesto de rendimientos constantes de escala convierte a la curva OF en una línea recta que pasa por el origen (figura V.3.2). La unidad más eficiente puede ser entonces fácilmente identificada: aquélla con el ratio output/input mayor. De esta forma, la pendiente de dicha frontera se determina haciendo pasar la línea recta por la organización más eficiente. La *eficiencia técnica* de la unidad A se define, como ya se ha comentado anteriormente a rendimientos de escala constantes en términos de output y en términos de input. Así, la eficiencia de los inputs se expresará a través del cociente, x_1'/x_2 y la eficiencia en términos de outputs, y_1/y_2'.

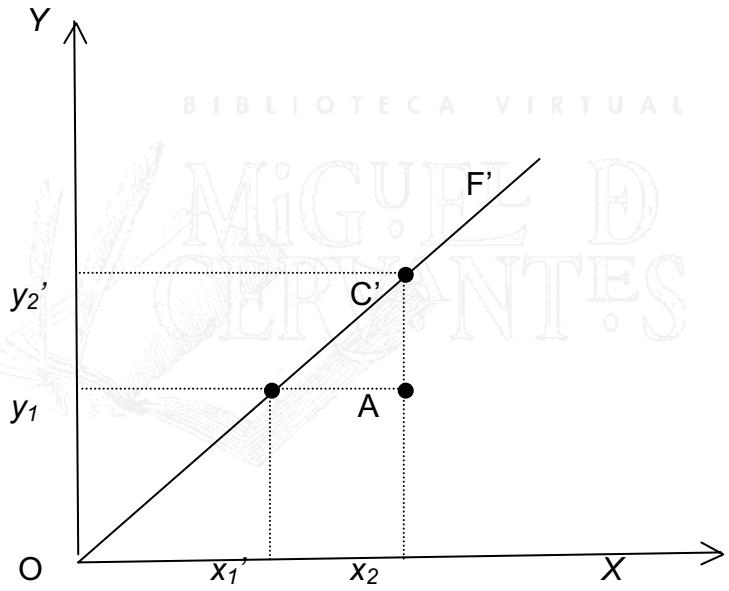


Figura.V.3.2. Función de Producción Estimada.
Fuente. Pina, V.(2000)
Elaboración.Propia.

V.4. APROXIMACIÓN NO PARAMÉTRICA A LA EFICIENCIA TÉCNICA DEL SECTOR ELÉCTRICO ESPAÑOL MEDIANTE EL ANÁLISIS ENVOLVENTE DE DATOS.

V.4.1. Introducción

En este apartado se procede a realizar un análisis de la eficiencia desde el punto de vista de la producción y distribución de las empresas eléctricas españolas que desarrollan actividades en el ciclo completo de la actividad eléctrica mediante el Análisis Envolvente de Datos. La medida de la eficiencia es una cuestión importante tanto en la generación como en el suministro de energía. Las políticas de liberalización del Sector Eléctrico en la Unión Europea y por ende, en España, han provocado la necesidad de evaluar las eficiencias entre las distintas compañías eléctricas, y de ser capaces de comunicar los resultados a los órganos de decisión. El problema de la evaluación de la eficiencia implica normalmente múltiples factores de entrada, inputs, y múltiples factores de salida, outputs.

La evaluación de la eficiencia productiva hace referencia a la productividad de una determinada unidad que genere resultados en forma de productos y/o servicios. Así, Farell(1957) expresa la productividad como el cociente de la producción fabricada o creada y los recursos consumidos, tal como se refleja en la expresión:

$$\text{Productividad} = \frac{\text{Recursos Generados}}{\text{Recursos Consumidos}} = \frac{\text{Salidas}}{\text{Entradas}}$$

La evaluación del rendimiento desde el punto de vista de la producción no debe ser realizada individualmente para cada una de las áreas de la empresa ni para cada una de las empresas, sino que es preciso contemplar la posición relativa de cada una de las empresas eléctricas con el resto de sus competidoras en el mercado de la electricidad. De manera que se pueda notar

la mejora potencial que podría desarrollar la empresa con las variables analizadas.

Los métodos paramétricos clásicos, generalmente resultan insuficientes para la resolución del problema planteado, ya que necesitan establecer una relación funcional específica que integre las variables del modelo. En el presente trabajo, la imposición de tales relaciones funcionales no parece tener lugar.

Un enfoque alternativo, que resuelve en gran medida los inconvenientes planteados, es el método de Análisis de Envoltura de Datos. Este método no paramétrico de fijación de fronteras eficientes se sustenta sobre los trabajos pioneros, de un economista, Farrell[12] acerca de la eficiencia, pero fue en el ámbito de la Investigación Operativa donde alcanzó su pleno desarrollo (Charnes, Cooper y Rhodes, (1978); Banker, Charnes y Cooper, (1984)).

En este trabajo, se aplica la técnica DEA[11] para evaluar el grado de eficiencia en el cual han trabajado las principales compañías eléctricas de ciclo completo en España, en un periodo de liberalización del mercado de la electricidad. El uso de DEA para la evaluación de las eficiencias relativas de las compañías eléctricas muestra una comprensión adecuada de la utilización de los recursos en el sistema de la generación y distribución de la electricidad, así como también suministra información acerca de futuros desplazamientos que pueden producirse en el mercado residencial hacia otras compañías del sector.

Una de las razones por las que se emplea la técnica DEA es porque se han descubierto posibilidades para su utilización, en casos en los que otro tipo de aproximaciones no han sido válidas por la complejidad de la naturaleza de numerosos inputs y numerosos outputs en muchas de estas actividades (las cuales son normalmente calificadas en unidades no-cuantificables).

[11] Acrónimo de los vocablos ingleses Data Envelopment Analysis.
[12] Este concepto es aplicado a cualquier tipo de ente, ya sea una empresa, departamento, hospital, centro de educación, etc.

La técnica DEA parte de la elección de un modelo basado en retornos de escala variables, es decir, denominado BCC. En el presente trabajo se utiliza este modelo por dos motivos:

- Expresa la reducción o aumento de los inputs u outputs[13] que debería realizar una compañía para pertenecer a la frontera eficiente.
- Permite actuar a retornos de escala variable.

El primer motivo es importante ya que permitirá reducir los inputs manteniendo los outputs, y viceversa. La importancia del segundo motivo está en que permite tener un control de la escala de producción, es decir, este hecho permitirá comparar una empresa eléctrica con pequeña capacidad de generación y distribución con otra empresa eléctrica con capacidades muy superiores.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

V.4.2. Estudio empírico

V.4.2.1. Las unidades objeto de estudio

El espacio territorial y administrativo donde las compañías eléctricas españolas se ha considerado que actúan es España, ya que la mayoría de ellas operan también en otros mercados, como son el mercado latinoamericano y resto de países europeos.

La estructura eléctrica en España, pertenece básicamente a las cuatro empresas eléctricas de ciclo completo, éstas son, Endesa, Hidrocantábrico, Iberdrola y Unión Fenosa, las cuales serán las unidades objeto de estudio, ya que suministran más del 95% de la energía vendida en el mercado español.

A continuación se expresa en la tabla V.4.1 el balance de producción peninsular, de energía eléctrica, durante los cuatro periodos estudiados, donde se observa el aumento de la demanda de electricidad en torno a un 6,71%

[13] según sea su orientación.

anual, desde el año 1998 al año 2000; sin embargo este crecimiento es menor en el año 2001, siendo de un 5%.

Tabla.V.4.1. Balance de Energía en el Sistema Eléctrico Peninsular (GWh).

	2001	**2000**	**1999**	**1998**
Hidráulica	39.374	27.422	24.171	33.989
Carbón	68.080	76.595	72.315	60.184
Fuel/Gas	12.400	10.260	9.925	5.649
Nuclear	63.705	62.094	58.852	58.996
PRODUCCIÓN BRUTA (b.a.)	183.559	176.371	165.263	158.818
-Consumos en generación	7.613	7.806	7.224	6.274
-Consumos Bombeo	4.141	4.664	3.666	2.587
PRODUCCIÓN NETA	171.805	163.901	154.373	149.957
+Intercambios internacionales	3.450	4.471	5.719	3.398
+Régimen especial	30.374	27.429	24.189	19.325
DEMANDA(b.c.)	205.629	195.800	184.281	172.680

Fuente. Informes Anuales R.E.E.
Elaboración. Propia.

V.4.2.2. Aplicación del análisis envolvente de datos

V.4.2.2.1. Descripción y selección de las variables

Para la aplicación de la técnica de Análisis Envolvente de Datos es necesario elegir una serie de inputs o entradas y outputs o salidas o logros para su desarrollo. La selección de estas variables, inputs y outputs, tratándose de un sector tan especial y complejo como es el sector eléctrico, ha estado condicionada principalmente por:

- La información disponible para este estudio se limita a la proporcionada por las distintas compañías eléctricas, Endesa, Hidroeléctrica del Cantábrico, Iberdrola y Unión Fenosa. Dicha información se materializa en los Informes Anuales de cada una de estas empresas en los años 1998,1999, 2000 y 2001. Respecto a los datos del año 2002 habría que decir que no ha sido posible conocerlos a fecha de la finalización del presente trabajo, por estar celebrándose las juntas generales de accionistas de las empresas eléctricas.
- El análisis de las correlaciones simples de los distintos factores y productos con el objetivo de conocer los inputs y outputs que han de considerarse en la formulación del modelo.

De esta forma, se han seguido una serie de fases, las cuales, son necesarias para la aplicación del análisis envolvente de datos, y que se exponen en la figura V.4.1:

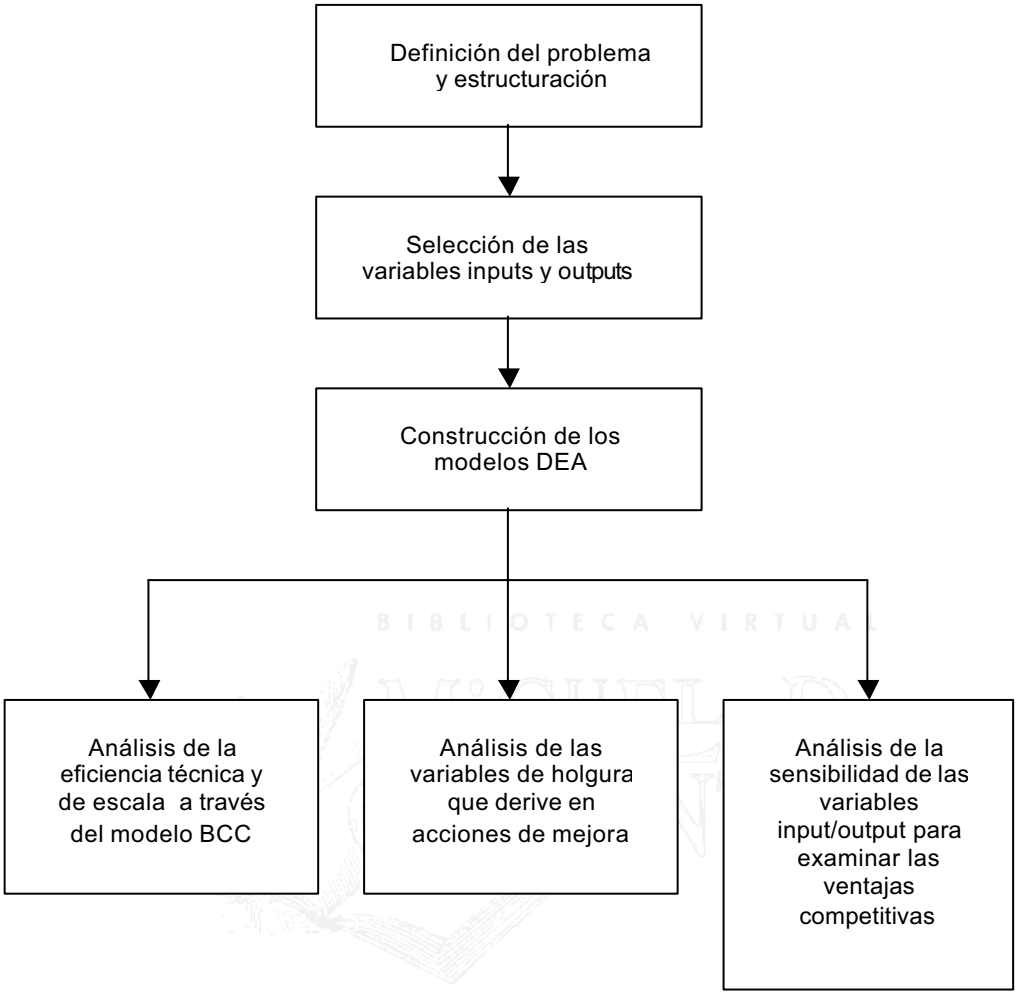


Figura.V.4.1. Estructura de la aplicación.
Fuente. Lo, Feng-Yu et al.(2001).
Elaboración. Propia.

Habría que mencionar que la aplicación del análisis envolvente de datos al Sector Eléctrico Español plantea un inconveniente: el número reducido de unidades de decisión frente a la literatura revisada, ya que son cuatro las empresas sobre las cuales versa dicho estudio. Desde el punto de vista estadístico, como ya se ha comentado anteriormente es un problema asociado a los *grados de libertad*, el cual afecta a DEA porque está orientado a la eficiencia relativa. En el desarrollo del modelo, el número de grados de libertad se aumentará con el número de DMU´s y disminuirá con el número de inputs y outputs. En el caso planteado el número de DMU's o unidades de decisión son 16, correspondientes a cada una de las empresas eléctricas de ciclo completo en los distintos espacios temporales, 1998, 1999, 2000 y 2001.

Una vez planteado el problema y las unidades objeto de estudio, se definen las distintas variables inputs y outputs relacionadas con la eficiencia de las compañías eléctricas. Se comienza definiendo los inputs:

- *Longitud de las Líneas de Energía de Alta Tensión* (LAT): es la longitud de las líneas de distribución, tanto las líneas aéreas de alta tensión como las líneas subterráneas de alta tensión, expresado en kilómetros.
- *Longitud de las Líneas de Energía de Baja Tensión* (LBT): es la longitud de las líneas de distribución, las líneas aéreas y subterráneas de media y baja tensión, expresado en kilómetros.
- Potencia (PT): integra la potencia instalada en las centrales que operan dentro del Régimen Ordinario y en Régimen Especial (Cogeneración, Eólica, Biomasa y Residuos Sólidos Urbanos y Minihidráulica Propia), expresada en megavatios.
- *Empleados* (EMP): es la plantilla total de las empresas eléctricas, la cual está integrada por titulados superiores, técnicos, administrativos y resto de categorías.

Por otra parte, las variables outputs que influyen en la eficiencia de las compañías eléctricas son las siguientes:

- *Número de Clientes* (CLI): indica la cifra de suministros con tarifas de baja (tensión <1KV) y alta (tensión>1KV) tensión.
- *Energía Facturada* (ENERFAC): son los gigawatios (10^9 vatios) suministrados por cada una de las compañías eléctricas durante un ejercicio económico.
- *TIEPI* (Tiempo de Interrupción Equivalente de la Potencia Instalada, expresado en minutos): es una medida de la calidad del suministro eléctrico hacia los clientes. Este indicador se define como:

$$TIEPI = \frac{\sum_{i=1}^{k} P_i H_i}{\sum_{i=1}^{n} P_i}$$

siendo

P_i : potencia instalada en los centros de transformación de media y baja tensión del distribuidor, más la potencia contratada en media tensión, afectada por la interrupción *i*-ésima de duración H_i, expresado en kVA.

H_i : tiempo de interrupción del suministro que afecta a la potencia P_i , expresado en horas.

k: número total de interrupciones durante el período considerado.

n: número de centros de transformación.

Tras la definición de las variables inputs y outputs, el estudio se ocupa de la relación que ejercen los inputs hacia la producción de los outputs. El objetivo fundamental, es analizar en qué medida se han seleccionado adecuadamente los inputs, y estudiar su relación con los outputs, a través un análisis de correlaciones simples entre el conjunto de inputs y outputs definidos.

De esta forma, a través de la tabla de correlaciones V.4.2 de los inputs y outputs correspondientes a los ejercicios 1998, 1999, 2000 y 2001, se trata de evidenciar el estudio de determinados factores entre otros que también se pudieran analizar.

En el análisis de la matriz de correlaciones, integrada por inputs y los outputs definidos anteriormente, se observa la existencia de una intensa relación lineal de carácter directo en todas la correlaciones bivariadas planteadas. Esto quiere decir, que los coeficientes son significativos al nivel de significación del 1% rechazándose la hipótesis de incorrelación al nivel de confianza del 99% ($H_0 : \rho = 0$). A partir de los resultados obtenidos se debe tener en cuenta, a la hora de seleccionar los posibles inputs, el output respecto al cual se va a estudiar la eficiencia técnica, ya que, como se puede observar, las relaciones de dependencia, son distintas, no siendo inferiores, en ningún caso al 0,69. Se observa que la relación de carácter lineal entre el output, el TIEPI y los inputs, LAT y PTI, es elevada. La energía facturada (ENERFAC) y el número de clientes(CLI) son variables output que tienen una relación cuasi-

perfecta con todas las variables inputs consideradas, ya que todas las correlaciones alcanzan valores muy elevados y próximos al valor máximo (la unidad). El análisis estadístico de los datos indica que, en principio, no se debe ser muy selectivo con los inputs planteados cuando se intenta explicar la variación de algunos de los outputs, ya que los coeficientes de correlación entre todos los pares de input-output alcanzan valores elevados, lo cual no provocará fuertes discriminaciones a la hora de elegir entre los distintos inputs y/o outputs que formarán parte de los modelos que se construirán en secciones posteriores.

Tabla.V.4.2. Coeficientes de Correlación de Pearson.

		TIEPI	ENERFAC	CLI
LAT	Correlación de Pearson	0,797	0,979	0,977
	Sig. (bilateral)	0,00	0,00	0,00
LBT	Correlación de Pearson	0,699	0,999	0,991
	Sig. (bilateral)	0,00	0,00	0,00
PTI	Correlación de Pearson	0,750	0,983	0,984
	Sig. (bilateral)	0,00	0,00	0,00
EMP	Correlación de Pearson	0,741	0,941	0,934
	Sig. (bilateral)	0,00	0,00	0,00

Fuente. Informes Anuales. 1998-2001.
Elaboración. Propia.

V.4.2.2.2. Procedimiento para la selección de modelos

El sector eléctrico tiene una característica importante relacionada con la dimensión de las plantas generadoras y ésta es la potencial relevancia de las economías de escala; por lo tanto, el modelo bajo la técnica del análisis envolvente de datos que se ha considerado en el estudio es el modelo BCC, de retornos a escala variables. La razón por la que se ha seleccionado un modelo DEA tipo BCC es debido a la propia naturaleza de las empresas eléctricas que no están dotadas de las mismas capacidades y dimensiones. Por lo tanto el modelo planteado tendrá que permitir la existencia de rendimientos a escala variable. Respecto a la evaluación de la eficiencia, se ha optado por elegir la orientación output ya que en el caso de considerarse con orientación input existen importantes restricciones de tiempo y coste para llevar a cabo las medidas que proponga el procedimiento por parte de las empresas, no siendo del todo objetivo. El enfoque elegido evalúa la medida en que las

unidades menos eficientes se acercan a la frontera de eficiencia, la cual está definida por las empresas más eficientes a lo largo de todo el proceso analizado. Por lo tanto, la observación de una misma compañía eléctrica en cada uno de los años (desde 1998 a 2001, ambos inclusive) es como si se tratara de empresas distintas. En otros términos, se construye lo que se denomina *una frontera intertemporal*.

A partir de ahí se lleva a cabo el análisis de la evolución relativa de los principales cuatro grupos eléctricos españoles[14], Endesa, Hidrocantábrico, Iberdrola y Unión Fenosa, que operan en el mercado eléctrico español a lo largo del periodo 1998-2001. La eficiencia es evaluada mediante programa informático EMS (Efficiency Measurement System, asociado al procedimiento DEA).

También hay que definir el número de inputs y outputs a considerar en el modelo. A este respecto, Cooper, W.W., Seiford, L.M. y Tone, K.(2000) establecen un criterio que considera que:

$$n \geq máx\{m \text{ x } s, 3(m+s)\}$$

donde n = número de DMU´s, m = número de inputs y s = número de outputs. Esta restricción aplicada al estudio, obliga a seleccionar como máximo 5 variables, como suma de inputs y outputs.

En el desarrollo del estudio ha sido necesario considerar todos los modelos posibles, centrándose el trabajo en los modelos de 1input-2outputs y 2inputs-1output, respecto a los modelos de 2inputs-3outputs y 3outputs-2inputs se comentarán a continuación las principales conclusiones extraídas de su análisis. Los modelos que consideran 1input-1output arrojan similares conclusiones a los descritos anteriormente, no haciéndose referencia en el presente trabajo.

Con respecto a los modelos formados por un total de cinco elementos (2inputs- 3outputs y 3inputs-2 outputs), entre factores y productos, se han

estudiado todas las posibles combinaciones sin repetición, resultando un total de 18 modelos planteados. En todos los casos, con características similares, los resultados son poco concluyentes, ya que por término medio aproximadamente el 45% (modelos de 3inputs-2outputs) y un 60% (modelos de 2inputs-3outpts) de las empresas analizadas en los distintos espacios de tiempo resultan ser eficientes. Siendo Iberdrola en el año 2001, la compañía a la que deben parecerse en mayor medida (el 95% de los modelos planteados) las restantes empresas eléctricas, queriendo indicar que Iberdrola es la empresa eléctrica más equilibrada desde el punto de la eficiencia tras el nuevo entorno creado.

A continuación, se comentan los principales resultados obtenidos al plantear modelos que analizan la eficiencia técnica o pura calculada bajo rendimientos de escala variables, donde se emplean un total de cuatro, entre factores y productos, (1input-3outputs y 3inputs-1output), siendo 16 los modelos analizados. En estos modelos al considerar un número menor de recursos o productos, disminuye el total de empresas eficientes. Así, en los modelos de 3inputs-1output resulta ser eficiente un promedio del 26,56% y en los modelos 1input-3outputs resultan ser eficientes por término medio el 48,43% de las unidades. Cuando se está analizando la situación de eficiencia técnica pura de las empresas respecto al output, TIEPI, la empresa *modelo* es Hicrocantábrico en 1998. Finalmente, comentar que Endesa solamente resulta ser una compañía eficiente en el año 2001, siendo en todos los ejercicios anteriores ineficiente; Hidrocantábrico es una compañía eficiente siempre en los ejercicios 1998 y 2001, siendo en algunos de los casos la única empresa que ha resultado en algún ejercicio (concretamente el año 1998) eficiente.

Una vez descritos los principales resultados obtenidos en los anteriores modelos, se comentarán más detenidamente los resultados de los modelos que están formados por 1input-2outputs y 2inputs-1output. El número de modelos resultantes de combinar los distintos inputs y outputs dan como resultado un total de 30 modelos distintos, tal como se expresa en la tabla V.4.3.

[14]Viesgo, filial del grupo Endesa e integrada posteriormente en el grupo, es comprada por Enel en el año 2001..

Desde el punto de vista económico y técnico-productivo, cada una de las dos clases de modelos tiene implicaciones distintas. Así, los modelos de 1input-2outputs, analizan con mayor detalle la utilización de un determinado recurso en la producción de los outputs, y por otra parte, los modelos de 2inputs-1output, reflejarán si el empleo de dos inputs simultáneamente mejora en la producción de un determinado output.

Respecto a la preferencia de una clase u otra de modelos habría que decir que, a las compañías eléctricas les interesará su análisis por separado, puesto que los outputs son naturaleza física bien distinta, y entre tres[15] de los cuatro inputs considerados existen relaciones intensas de carácter técnico-productivas.

Tabla.V.4.3. Modelos 1input- 2outpus y 2inputs- 1output.

1 INPUT-2 OUTPUTS			2 INPUTS-1 OUTPUT		
MODELO	**Inputs**	**Outputs**	**MODELO**	**Inputs**	**Outputs**
1	LAT	TIEPI EFACT	1	LAT LBT	TIEPI
2	LBT	TIEPI EFACT	2	LAT LBT	CLI
3	PT	TIEPI EFACT	3	LAT LBT	EFACT
4	EMP	TIEPI EFACT	4	LAT PTI	TIEPI
5	LAT	TIEPI CLI	5	LAT PTI	CLI
6	LBT	TIEPI CLI	6	LAT PTI	EFACT
7	PT	TIEPI CLI	7	LAT EMP	TIEPI
8	EMP	TIEPI CLI	8	LAT EMP	CLI
9	LAT	EFACT CLI	9	LAT EMP	EFACT
10	LBT	EFACT CLI	10	LBT PTI	TIEPI
11	PT	EFACT CLI	11	LBT PTI	CLI
12	EMP	EFACT CLI	12	LBT PTI	EFACT
			13	LBT EMP	TIEPI
			14	LBT EMP	CLI
			15	LBT EMP	EFACT
			16	EMP PTI	TIEPI
			17	EMP PTI	CLI
			18	EMP PTI	EFACT

Elaboración. Propia.

[15] LAT, LBT y PT.

Una vez planteados los modelos, se podrían agrupar dependiendo del tipo de input considerado, distinguiéndose de esta forma dos tipos de productividades, la productividad basada en el inmovilizado material y la productividad basada en los recursos humanos propios de la empresa.

Las variables inputs y outputs que constituirán la base para analizar la productividad desde el punto de vista del inmovilizado material son:

Inputs: Longitud Líneas Alta Tensión
Longitud Líneas Baja Tensión
Potencia Instalada
Outputs: Número de clientes
Energía Facturada
TIEPI

La construcción de cualquier modelo a partir de estas variables de entrada o salida proporciona información acerca de la productividad que está relacionada con los recursos básicos empleados en la actividad de generación eléctrica, así como también en las fases de transporte, distribución y comercialización.

Respecto a las variables inputs y outputs que constituirán la base para analizar la productividad centrada en los recursos humanos adquiridos por parte de la empresa, estos son:

Input: Número de empleados
Outputs: Número de clientes
Energía Facturada
TIEPI

En estos modelos se intenta obtener una visión de la productividad del factor trabajo en las empresas eléctricas. Aunque hubiera sido importante, desde otro punto de vista la desagregación de los empleados en personal administrativo, técnico y otro personal, ya que lo anterior permite separar los factores de ineficiencia provocados por la

composición de dicho personal, sin embargo, estos datos no ha sido posible conocerlos de forma completa para todos los periodos y empresas. También, habría que tener en cuenta que, hoy en día, la mayoría de las empresas eléctricas subcontratan gran parte de las tareas que realizan, con lo cual la eficiencia, desde el punto de vista del trabajo, se convertiría en una visión secundaria y además de *grados* entre empresas eléctricas, es decir, sería una cuestión acerca de en qué medida cada una de las empresas eléctricas están practicando el denominado *outsourcing*.

Los resultados de la eficiencia de las empresas eléctricas, bajo estos modelos, se describen y comentan en el apartado siguiente.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

V.4.2.3.3. Resultados empíricos

A partir de la metodología descrita en el Apéndice se obtienen los índices de eficiencia para cada uno de los 12 y 18 diferentes modelos, correspondientes a los casos en los que se considera 1input-2outputs y 2inputs-1output. A partir de los 16 datos correspondientes a los cuatro grupos eléctricos españoles se obtiene una frontera para cada modelo. De esta forma resultan un total de 30 fronteras de eficiencia (una para cada modelo planteado) y 480 (30 modelos x 16 DMU's) índices de eficiencia técnica. La ordenación de estos datos y su tratamiento estadístico permite la obtención de los resultados que se detallan a continuación.

Los índices de eficiencia toman un valor igual o superior a 1 (expresado en términos absolutos). De tal forma una empresa eléctrica será eficiente si el índice de eficiencia alcanza el valor 1, así la empresa en cuestión estará produciendo la máxima cantidad de outputs con el nivel de inputs dado.

Con la ayuda de estos modelos será posible estudiar la evolución de la eficiencia de las compañías, así como la eficiencia media de cada una de las

mismas en los distintos periodos. En primer lugar, en la tabla V.4.4 se reflejan los resultados de los modelos de 1input-2output:

Tabla. V.4.4. Indicadores de Eficiencia (expresado en valor absoluto). Modelo 1input-2outputs.

	1	2	3	4	5	6	7	8	9	10	11	12	Promedio
ENDESA(1998)	1,08	1,08	1,08	1,08	1,04	1,04	1,08	1,04	1,05	1,05	1,08	1,09	1,07
H.C.(1998)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,07	1,00	1,00	1,00	1,01
IBERDROLA(1998)	1,00	1,00	1,00	1,00	1,00	1,00	1,07	1,00	1,05	1,03	1,07	1,19	1,03
U.FENOSA(1998)	1,32	1,38	1,11	1,34	1,15	1,38	1,02	1,31	1,15	1,43	1,02	1,42	1,25
ENDESA(1999)	1,06	1,06	1,08	1,08	1,03	1,03	1,00	1,06	1,03	1,03	1,06	1,06	1,05
H.C.(1999)	1,29	1,11	1,30	1,30	1,08	1,05	1,00	1,37	1,08	1,03	1,29	1,30	1,18
IBERDROLA(1999)	1,09	1,08	1,09	1,10	1,03	1,02	1,01	1,06	1,03	1,02	1,05	1,11	1,06
U.FENOSA(1999)	1,22	1,27	1,05	1,21	1,17	1,27	1,02	1,15	1,18	1,44	1,02	1,25	1,19
ENDESA(2000)	1,00	1,00	1,00	1,00	1,00	1,00	1,02	1,00	1,01	1,01	1,02	1,02	1,01
H.C.(2000)	1,15	1,00	1,12	1,09	1,05	1,06	1,01	1,30	1,05	1,00	1,09	1,05	1,08
IBERDROLA(2000)	1,03	1,03	1,03	1,04	1,02	1,01	1,01	1,03	1,02	1,01	1,02	1,03	1,02
U.FENOSA(2000)	1,25	1,31	1,03	1,06	1,17	1,34	1,02	1,00	1,17	1,42	1,02	1,04	1,15
ENDESA(2001)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
H.C.(2001)	1,00	1,00	1,00	1,00	1,00	1,08	1,18	1,09	1,00	1,00	1,00	1,00	1,03
IBERDROLA(2001)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
U.FENOSA(2001)	1,30	1,46	1,00	1,00	1,17	1,42	1,00	1,00	1,17	1,42	1,00	1,00	1,16
Promedio	1,11	1,11	1,06	1,08	1,06	1,11	1,03	1,09	1,07	1,12	1,05	1,10	

Fuente. Informes Anuales durante el periodo 1998-2001.
Elaboración. Propia.

Los resultados obtenidos revelan que las empresas eléctricas a lo largo del periodo estudiado han actuado en intervalos próximos al nivel de eficiencia. Así como también se observa en las filas correspondientes a Endesa(2001), H.C.(2001), Iberdrola(2001) y U.Fenosa(2001), para los 12 modelos, que existe una mejora significativa en el rendimiento productivo de las empresas eléctricas a partir de la entrada del nuevo marco regulatorio en España. Según los datos del último año analizado, se podría decir que Endesa e Iberdrola están desarrollando su actividad de forma eficiente, desde el punto de vista del inmovilizado (LAT, LBT y PTI) y la calidad del servicio prestado (TIEPI), alcanzando todos los índices valor igual a 1. En una situación más desfavorable, aunque con niveles parcialmente eficientes en algunos de los modelos planteados, se encuentran Hidrocantábrico y Unión Fenosa, con eficiencias medias de 1,03 y 1,16, en el año 2001 respectivamente. Por lo tanto, estas dos empresas eléctricas podrían aumentar sus outputs por término medio en un 3% y un 16%, respectivamente, a partir de los inputs existentes. Las inversiones que debería realizar Hidrocantábrico deberían centrarse en mejorar su participación en la cuota de mercado, tal como se puede observar en las celdas correspondientes a la fila H.C.(2001) y en las columnas de los modelos 6, 7 y 8. Estos tres modelos tienen en común la longitud de las líneas

de alta tensión, como input, por lo tanto sería éste un factor que se encontraría infrautilizado.

Por otra parte, si se estudia con mayor detalle la importancia de dos inputs en la producción de un determinado output, los resultados varían sensiblemente, llegando a ser el número de escenarios eficientes menor que en el caso de 1 input- 2 outputs. Esto es debido a que se esta juzgando un único output a partir de dos inputs y da lugar a que los resultados estén condicionados por dos tipos de restricciones. Los resultados se expresan en la tabla.V.4.5.

Tabla.V.4.5. Indicadores de Eficiencia (expresado en valores absolutos). Modelo 2inputs-1output.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Promedio
ENDESA(1998)	1,89	1,05	1,13	1,89	1,05	1,14	1,89	1,05	1,14	1,89	1,05	1,14	1,89	1,05	1,14	1,89	1,08	1,17	1,36
H.C.(1998)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
IBERDROLA(1998)	1,22	1,03	1,15	1,22	1,05	1,18	1,22	1,05	1,20	1,22	1,03	1,15	1,22	1,03	1,15	1,22	1,07	1,19	1,14
U.FENOSA(1998)	1,51	1,15	1,45	1,51	1,00	1,14	1,51	1,15	1,43	1,51	1,02	1,14	1,51	1,34	1,55	1,51	1,02	1,14	1,31
ENDESA(1999)	2,27	1,03	1,06	2,27	1,03	1,06	2,27	1,03	1,06	2,27	1,03	1,06	2,27	1,03	1,06	2,27	1,06	1,09	1,46
H.C.(1999)	1,38	1,02	1,11	1,38	1,07	1,41	1,38	1,08	1,46	1,38	1,05	1,11	1,38	1,05	1,11	1,38	1,48	1,41	1,26
IBERDROLA(1999)	1,58	1,02	1,09	1,58	1,03	1,11	1,58	1,03	1,12	1,58	1,02	1,09	1,58	1,02	1,09	1,58	1,05	1,12	1,24
U.FENOSA(1999)	1,37	1,18	1,39	1,37	1,01	1,08	1,37	1,15	1,30	1,37	1,02	1,08	1,37	1,20	1,30	1,37	1,02	1,09	1,23
ENDESA(2000)	1,96	1,01	1,03	1,96	1,01	1,03	1,96	1,01	1,03	1,96	1,01	1,03	1,96	1,01	1,03	1,96	1,02	1,04	1,33
H.C.(2000)	1,36	1,02	1,00	1,36	1,05	1,12	1,36	1,04	1,09	1,36	1,06	1,00	1,36	1,06	1,00	1,36	1,30	1,09	1,16
IBERDROLA(2000)	1,52	1,01	1,03	1,52	1,02	1,04	1,52	1,02	1,04	1,52	1,01	1,03	1,52	1,01	1,03	1,52	1,02	1,04	1,19
U.FENOSA(2000)	1,46	1,17	1,35	1,46	1,01	1,05	1,46	1,02	1,07	1,46	1,01	1,05	1,46	1,03	1,07	1,46	1,02	1,05	1,20
ENDESA(2001)	2,16	1,00	1,00	2,16	1,00	1,00	2,16	1,00	1,00	2,16	1,00	1,00	2,16	1,00	1,00	2,16	1,00	1,00	1,39
H.C.(2001)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,10	1,08	1,00	1,10	1,08	1,00	1,10	1,32	1,00	1,04
IBERDROLA(2001)	1,48	1,00	1,00	1,48	1,00	1,00	1,48	1,00	1,00	1,48	1,00	1,00	1,48	1,00	1,00	1,48	1,00	1,00	1,16
U.FENOSA(2001)	1,79	1,17	1,30	1,79	1,00	1,00	1,79	1,00	1,00	1,79	1,00	1,00	1,79	1,00	1,00	1,79	1,00	1,00	1,29
Promedio	1,56	1,05	1,13	1,56	1,02	1,08	1,56	1,04	1,12	1,56	1,02	1,06	1,56	1,06	1,10	1,56	1,09	1,09	

Fuente. Informes anuales durante el periodo 1998-2001.
Elaboración. Propia.

Los indicadores de eficiencia obtenidos para los distintos modelos reflejan que Hidrocantábrico era la única empresa eficiente para todos los modelos en 1998, tras la nueva etapa de liberalización del sector eléctrico, alcanzando valores igual a 1 en todos los índices de eficiencia. Su situación de eficiencia desaparece a partir de entonces, volviendo a trabajar en eficiencia en el año 2001, en el cual las empresas eléctricas alcanzan de forma mayoritaria niveles de eficiencia.

Con el objetivo de mejorar su representación en figuras, se ha realizado una transformación a una escala 0 –1 de los valores de eficiencia alcanzados por las compañías eléctricas, significando un valor de este índice igual a 1 que la empresa es eficiente, no siéndolo en el resto de los casos.

A continuación, se analizan los modelos de 1input-2output, de forma individual, para cada una de las cuatro compañías en el periodo analizado (1998-2001). En el gráfico de la evolución de eficiencias alcanzadas por Endesa, figura V.4.2, se observa que es una compañía eficiente en el año 2001, sin embargo ha sido eficiente en años anteriores al año 2000. A excepción del modelo 7 en el año 1999, lo cual refleja que la empresa con la potencia instalada que poseía, tenía la máxima cuota de mercado que podía atender y los menores tiempos en la interrupción del suministro eléctrico.

Si se analiza más en detalle la trayectoria seguida por Endesa, habría que decir que en el año 1998 no destaca por ser eficiente en ninguno de los doce modelos, llegando, incluso a ser el grado de ineficiencia elevado por un exceso en el número de empleados de la compañía (ver modelos 4, 8 y 12, cuyos niveles de ineficiencia en la tabla V.4.4. son 1,08, 1,04 y 1,09 respectivamente). Esta circunstancia pudo ser provocada por la adquisición por parte de la compañía de un número de filiales de las cuales ya era accionista (entre las que se encontraba Sevillana de Electricidad, S.A.). En el año 1999, Endesa emplea con eficiencia la potencia instalada en relación con la cuota de mercado, como ya se ha comentado anteriormente, esto puede ser provocado como consecuencia inmediata de las nuevas incorporaciones societarias a la compañía. En el año 2000 los niveles de eficiencia son mejores aumentando sus niveles de eficiencia media en un 4% respecto al año 1999, este aumento es provocado gracias a una disminución en el TIEPI. Finalmente, en el año 2001, la eficiencia con la que trabaja la compañía se refleja en todos los modelos analizados. Esta circunstancia refleja la capacidad de reacción y flexibilidad de la compañía ante el nuevo entorno eléctrico y requerimientos del nuevo mercado.

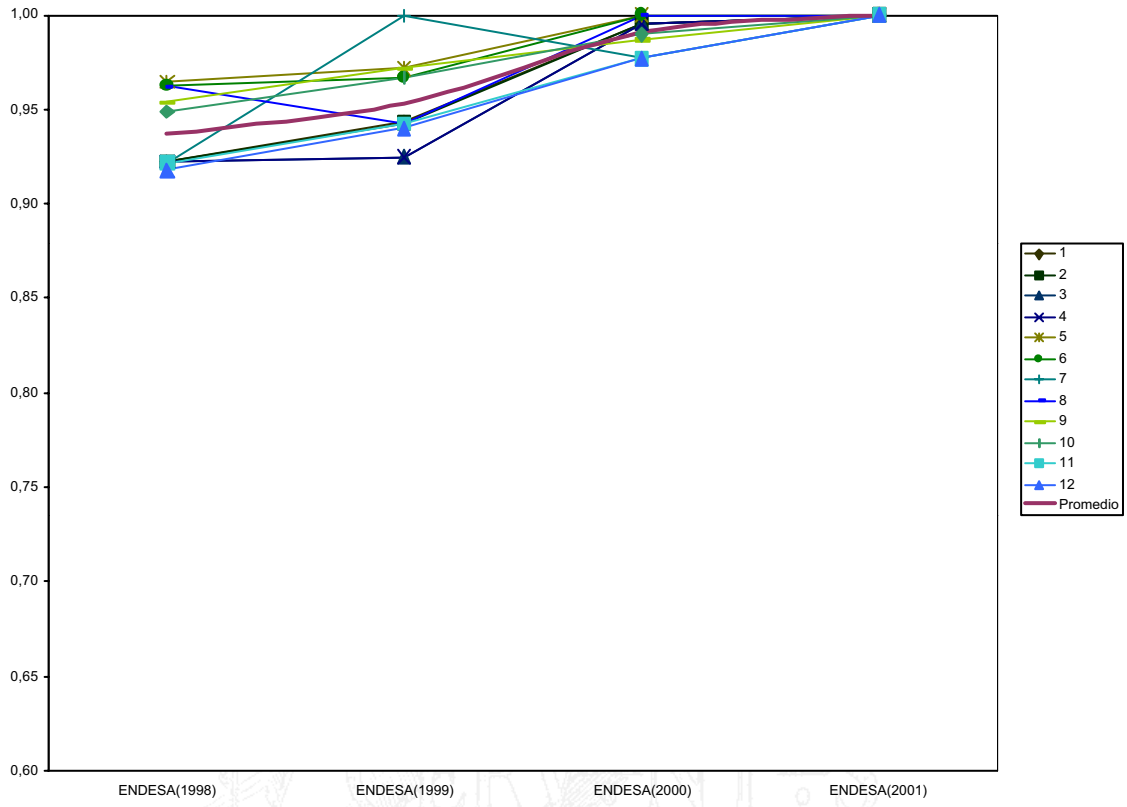


Figura.V.4.2. Evolución de los niveles de Eficiencia 1998-2001. Endesa.
Fuente. Informes anuales durante el periodo 1998-2001.
Elaboración. Propia

La tendencia seguida por la compañía Hidrocantábrico, desde el punto de vista de su eficiencia técnica, se caracteriza por su irregularidad pero con un horizonte de mejora continua en el desarrollo de sus actividades a partir del año 2000, tal como se puede observar en la tabla V.4.4 y en la figura V.4.3. Hidrocantábrico es la empresa más eficiente transcurrido el primer ejercicio de entorno liberalizador, año 1998, pero en el ejercicio 1999 descienden sus niveles en todos los modelos analizados, por una elevada plantilla para atender a las funciones desarrolladas por la empresa, lo cual redujo la productividad por empleado. Lo anterior se puede entender como una falta flexibilidad en sus operaciones eléctricas ante los nuevos dictámenes del sector. En los ejercicios posteriores esta situación mejora pero sin alcanzar los niveles de 1998, gracias al incremento de la energía facturada, tal como se observa en los modelos 2 y 10. El escaso aprovechamiento de la potencia instalada para mejorar su cuota de mercado y calidad del suministro, modelo 7 (PTI-TIEPI,CLI), se refleja en el descenso de la eficiencia a lo largo del periodo.

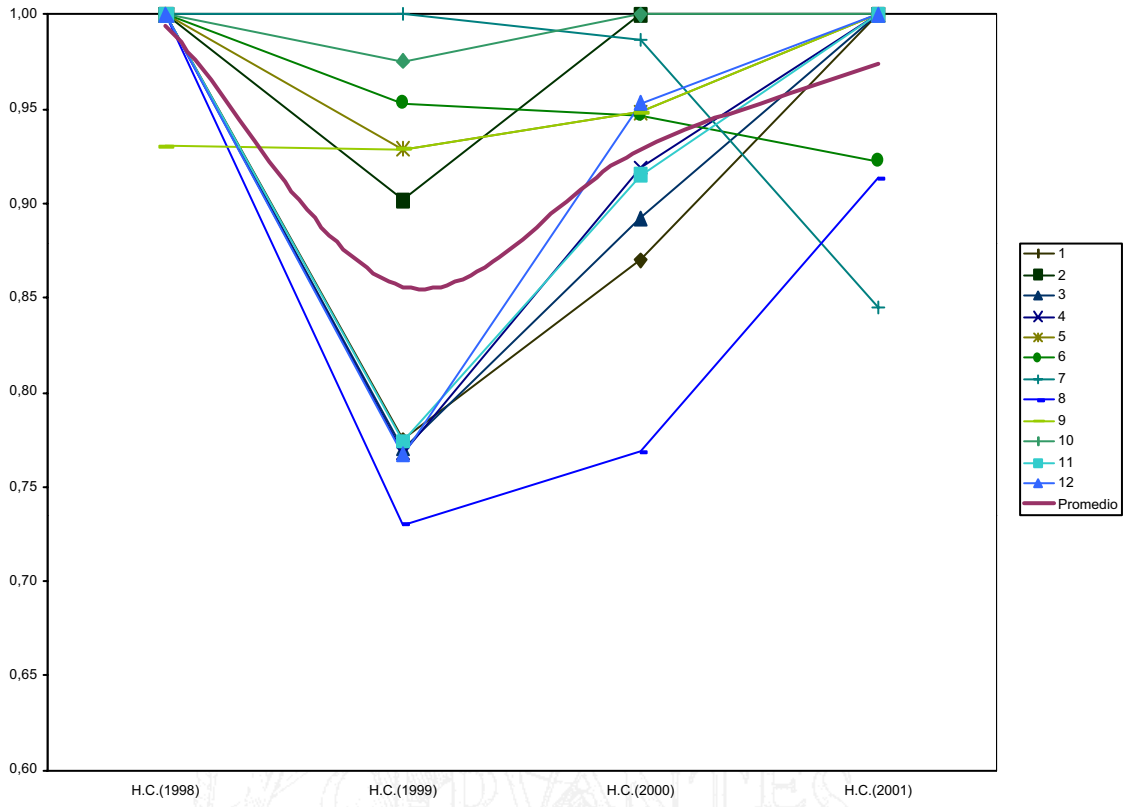


Figura.V.4.3. Evolución de los niveles de eficiencia 1998-2001. Hidrocantábrico.
Fuente. Informes anuales durante el periodo 1998-2001.
Elaboración. Propia.

En la figura V.4.4, se refleja la evolución llevada a cabo por Iberdrola, donde se pone de manifiesto su capacidad de adaptación al cambio del panorama eléctrico en España, hasta convertirse en el año 2001 en una de las dos empresas más eficientes desde el punto de vista del factor capital y del factor trabajo. Analizando la evolución seguida por Iberdrola, habría que decir que el principal punto débil con el que se encuentra en 1998 es su cuota de mercado, ya que con el inmovilizado que tenía podía atender a un mayor número de clientes. Este incremento de participación en su cuota no se alcanza hasta el año 2001. Por otra parte, también sería preciso destacar la buena conducta llevada a cabo por la compañía respecto al TIEPI en el año 1998, tal como se puede observar en la tabla V.4.4 en los modelos 1, 2, 3, 4, 5, 6 y 8 (todos ellos con valor igual a 1 en sus índices de eficiencia).

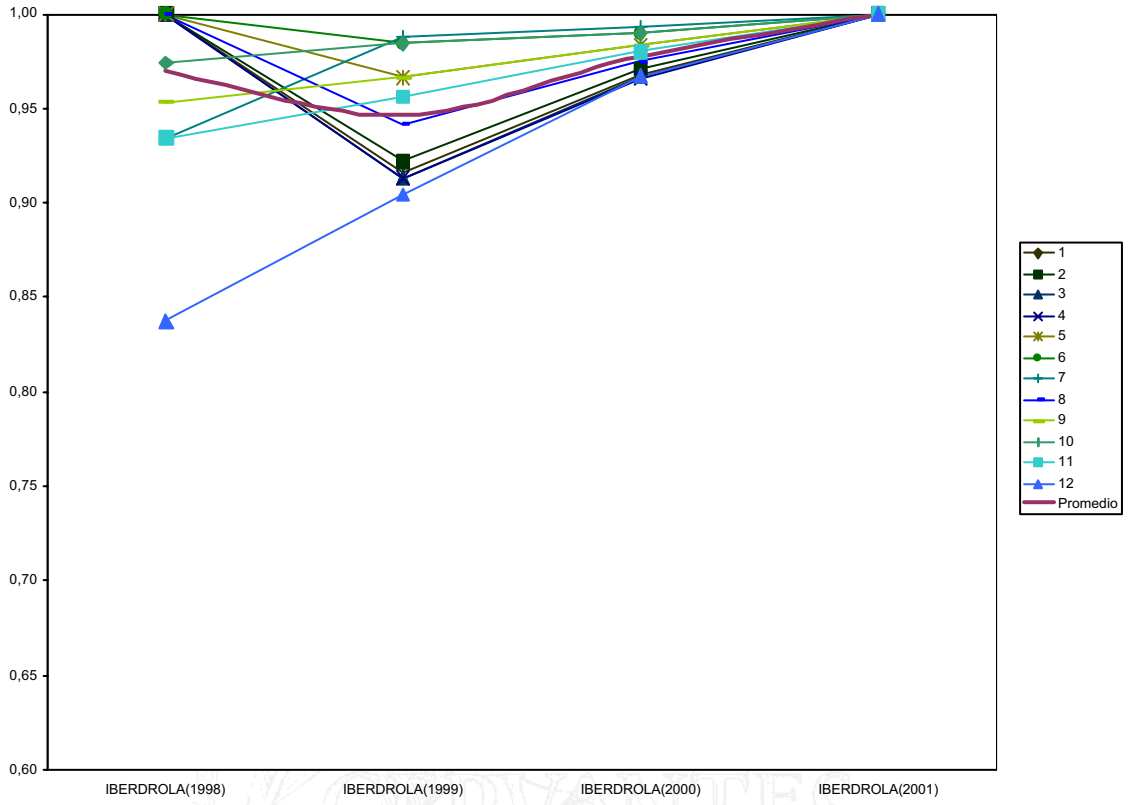


Figura.V.4.4. Evolución de los niveles de eficiencia 1998-2001. Iberdrola.
Fuente. Informes anuales durante el periodo 1998-2001.
Elaboración. Propia.

En la figura V.4.5, se puede comprobar que la empresa Unión Fenosa es la eléctrica que en peores condiciones de eficiencia ha trabajado, con niveles medios de eficiencia pertenecientes al rango [0,81, 0,88]. Las deficiencias que han llevado a Unión Fenosa a trabajar en esta situación vienen derivadas a través del factor capital, a través de la longitud de las líneas de baja tensión (modelos 2, 6 y 10) insuficiente ante la potencia instalada en sus centrales. La situación anterior ha podido provocar, en determinados instantes del tiempo, saturaciones en la distribución propia de la energía, al ser la potencia disponible muy próxima a la plena capacidad de potencia.

La evolución seguida por Unión Fenosa es irregular en todos los modelos analizados, sirva de ejemplo el modelo 2, el cual relaciona la distribución de energía con el TIEPI y el volumen de energía facturada, el periodo analizado se inicia con niveles de ineficiencia 1,38 para finalizar, en el año 2001, con niveles del 1,46.

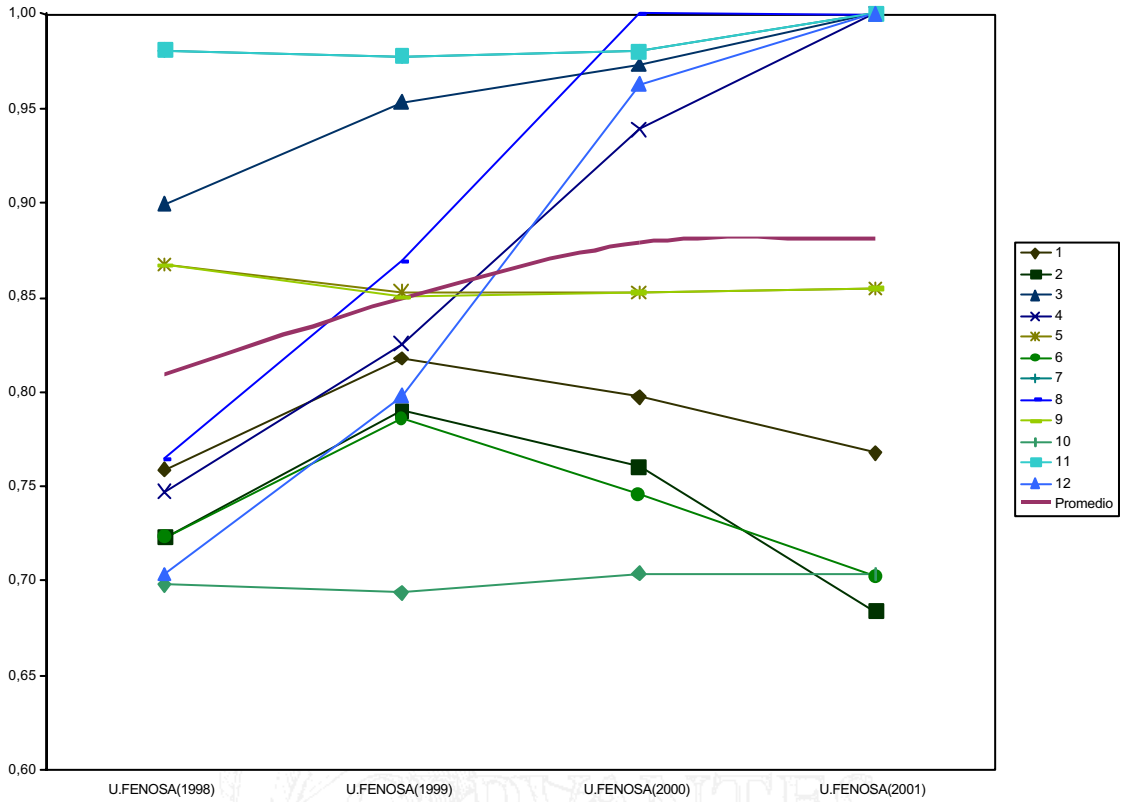


Figura.V.4.5.Evolución de los niveles de eficiencia 1998-2001. Unión Fenosa.
Fuente. Informes anuales durante el periodo 1998-2001.
Elaboración. Propia.

La figura V.4.6, representa la evolución de los índices de eficiencia de Endesa, siendo la empresa que alcanza niveles más bajos de eficiencia para los modelos que buscan la mejora de la calidad del suministro eléctrico durante el periodo 1998-2001 (modelos 1, 4, 7, 10, 13 y 16). El TIEPI se convierte cada vez más, durante el periodo analizado, en una debilidad a mejorar por parte de la compañía Endesa. Esta situación empeora en el año 2001, ya que aumentan los niveles de ineficiencia de la continuidad del servicio eléctrico en un 14% respecto del año 1998, tal como se observa en la tabla V.4.5 alcanzando un valor de 2,16.

La evolución seguida por la compañía en el resto de modelos es bien distinta mejorando el empleo del inmovilizado. El año 2001 es el único año en el que Endesa alcanza niveles de eficiencia, estos buenos resultados proceden de la maximización de la energía facturada y número de clientes a partir de los

recursos que posee. De esta forma, tanto el volumen de energía facturada como su cuota de mercado son puntos fuertes de la compañía en el año 2001.

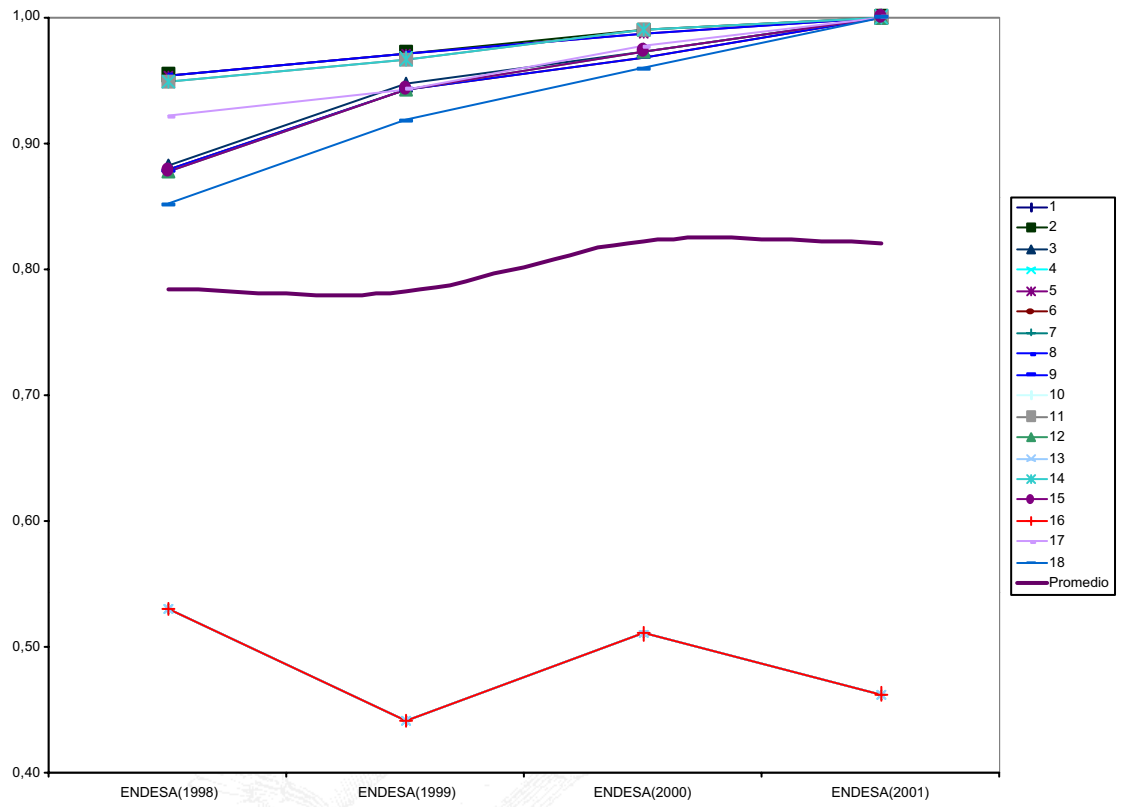


Figura.V.4.6. Evolución de los niveles de eficiencia 1998-2001. Endesa.
Fuente. Informes Anuales durante el periodo 1998-2001.
Elaboración. Propia

La evolución seguida por Hidrocantábrico se refleja en la figura V.4.7 y es muy similar a la comentada para los modelos de 1input-2outputs. En el año 1998, Hidrocantábrico alcanza niveles absolutos de eficiencia, sin embargo en el año 1999 descienden estos niveles en todos los modelos analizados para sufrir la recuperación en años posteriores.

A continuación se analizan, de forma individual, cada uno de los años objeto de estudio con la finalidad de conocer los principales puntos anómalos en la gestión de la compañía. Hidrocantábrico, en 1998, no presenta dificultades ya que emplea de forma eficiente todos los recursos considerados, alcanzando todos los índices valor igual a 1. A partir de 1999 existe una infrautilización en las líneas de alta tensión y la potencia instalada, afectando al volumen de la energía facturada y la continuidad del suministro eléctrico. En el

año 2000 la situación para la compañía es más favorable, observándose en la figura V.4.7 una tendencia creciente, llegándose a niveles de eficiencia, como es el caso de los modelos 3(LAT, LBT-EFACT), 12(LBT, PTI-EFACT) y 15(LBT, EMP-EFACT). Como se puede observar, estas mejoras afectan al volumen de ventas de la empresa. Para finalizar con el análisis de la empresa Hidrocantábrico, en el año 2001 la gestión de los recursos es más satisfactoria que en el 2000, siendo la empresa más eficiente por término medio.

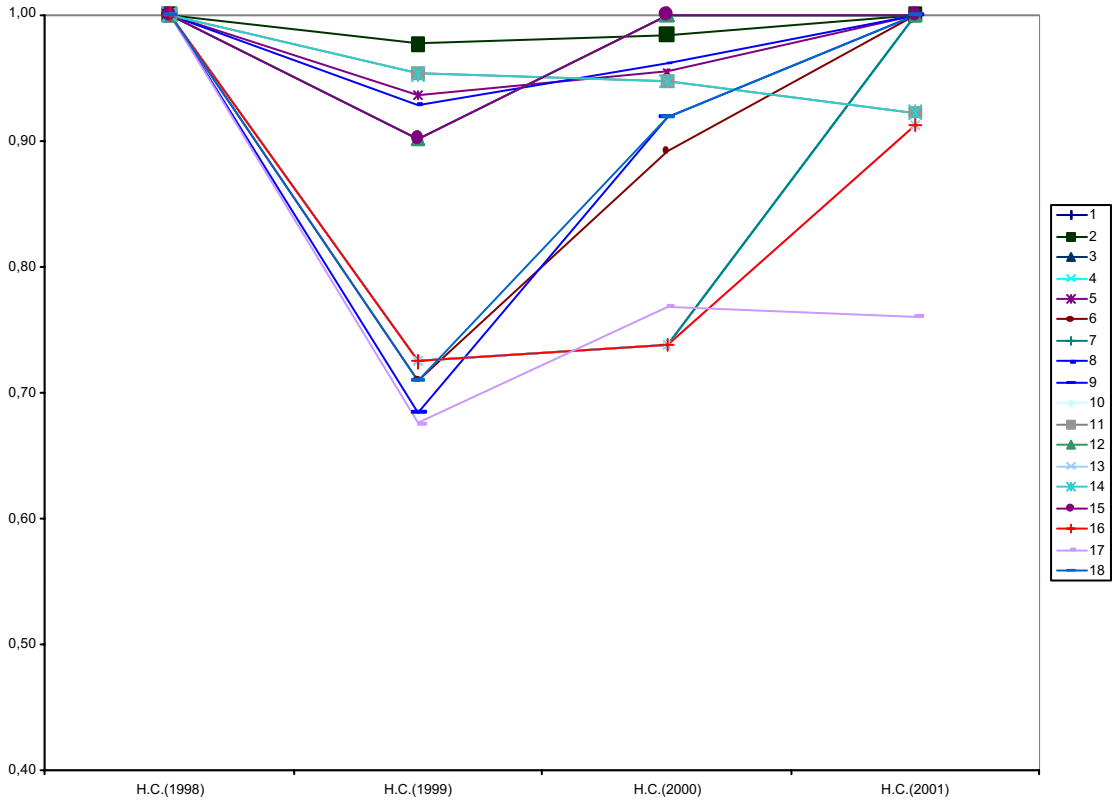


Figura.V.4.7.Evolución de los niveles de eficiencia 1998-2001. Hidrocantábrico.
Fuente. Informes anuales durante el periodo 1998-2001.
Elaboración. Propia.

El comportamiento de Iberdrola converge hacia niveles de eficiencia absoluta, tal como se refleja en la figura V.4.8, con la excepción del empleo de los recursos humanos y la capacidad de generación para mejorar la calidad del suministro, modelo 16 (EMP, PTI-TIEPI), el cual empeora a lo largo del periodo respecto al primer año de liberalización del sector eléctrico. El volumen de energía facturada es un punto débil en 1998, mejorando esta situación en años posteriores, hasta alcanzar niveles de eficiencia, por primera vez, en el año 2001.

Finalmente, habría que decir que Iberdrola tiene pendiente a lo largo de todo el periodo analizado la mejora de la calidad del suministro eléctrico.

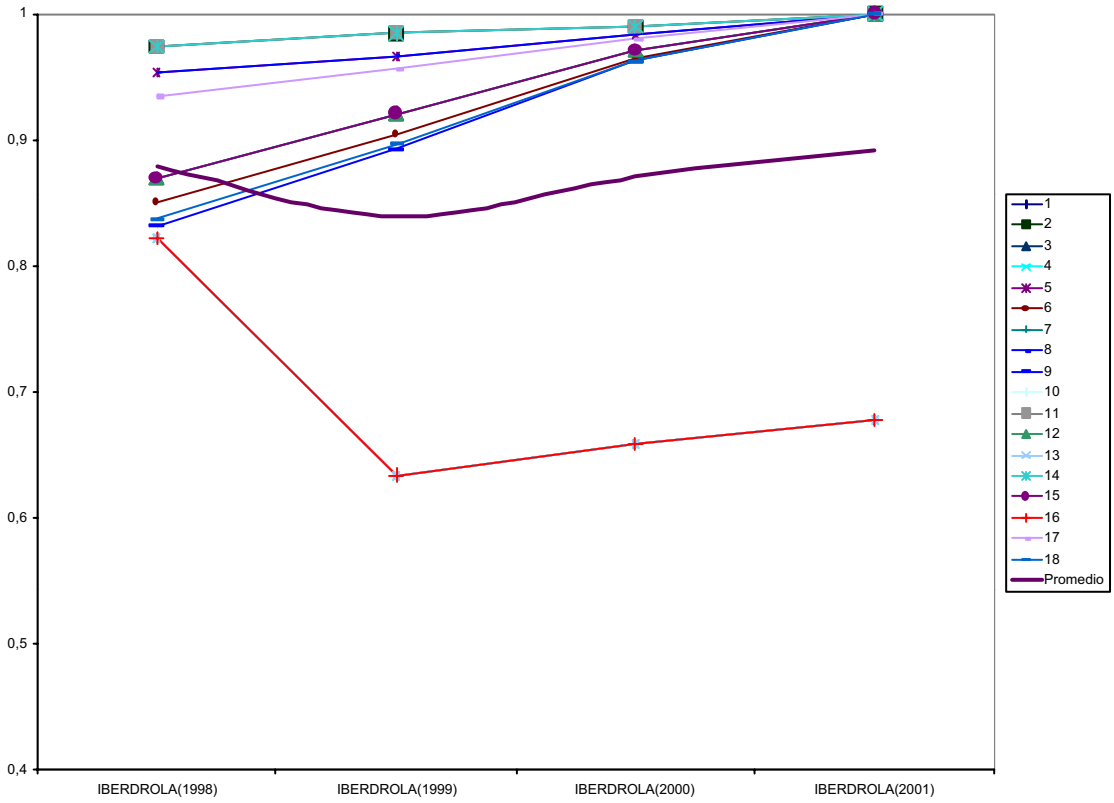


Figura.V.4.8. Evolución de los niveles de eficiencia 1998-2001. Iberdrola.
Fuente. Informes anuales durante el periodo 1998-2001.
Elaboración. Propia.

Por último, se representa la evolución de la empresa Unión Fenosa a lo largo del periodo 1998-2001. Unión Fenosa es una compañía que en el año 1998, un año con características reales semi-liberalizadoras, no actúa en eficiencia en ninguno de los outputs considerados. La energía facturada y la continuidad del suministro eléctrico son las principales debilidades de la compañía en el año 1998. A medida que se desarrolla el proceso liberalizador en España, Unión Fenosa va mejorando sus resultados. Esta mejora se realiza a partir de sus factores, a niveles de crecimiento moderado, llegando a alcanzar niveles de eficiencia, en el año 2001 en su poder de cuota de mercado y en la energía facturada, tal como reflejan los modelos 5, 6, 8, 9, 10, 11, 14,

15, 17 y 18. La disminución del TIEPI sigue siendo una mejora pendiente por Unión Fenosa en el año 2001.

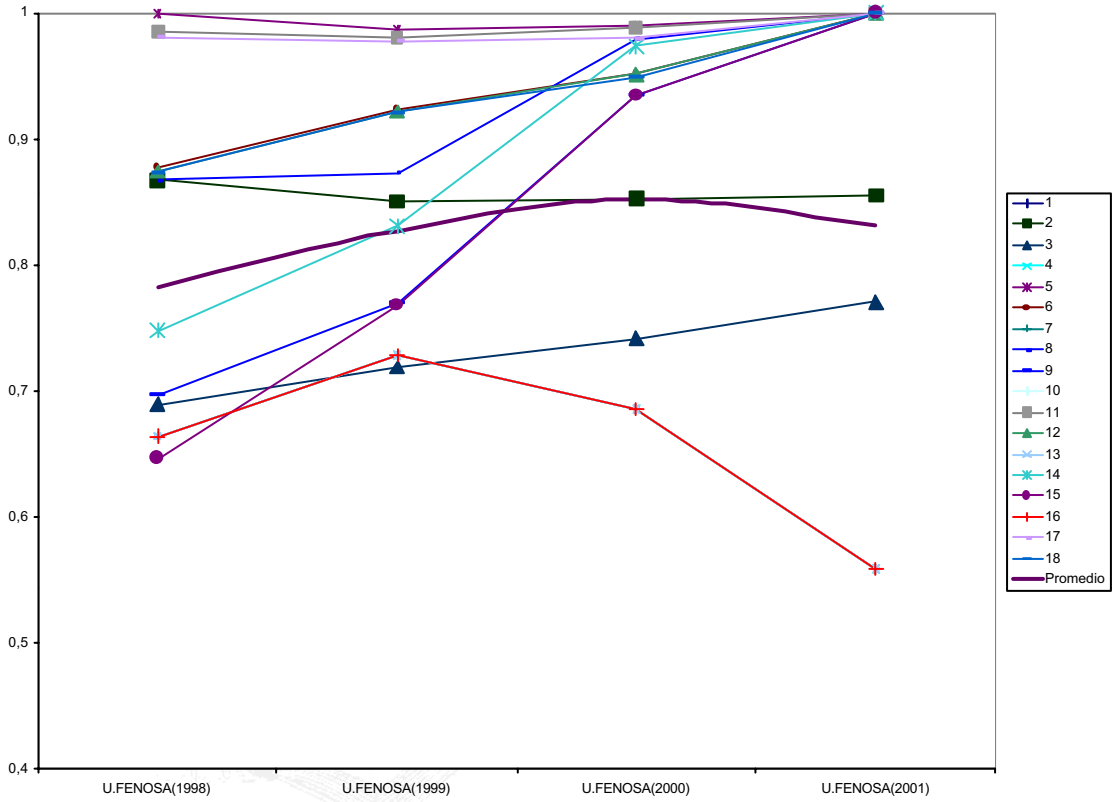


Figura.V.4.9.Evolución de los niveles de eficiencia 1998-2001. Unión Fenosa.
Fuente. Informes anuales durante el periodo 1998-2001.
Elaboración. Propia.

V.5. CONCLUSIONES DEL CAPÍTULO

Este capítulo se ha desarrollado con el objetivo de conocer las futuras tendencias que puedan tener lugar ante un mercado eléctrico completamente liberalizado, con el fin de detectar diferencias significativas en la eficiencia técnica de las empresas consideradas y su posible consecuencia hacia el desplazamiento de consumidores entre las distintas compañías del sector.

El estudio de la eficiencia a través del la técnica DEA proporciona información al sector eléctrico en España y una oportunidad para mejorar su gestión, a través de la comparación de los inputs (recursos) y los outputs (productos y logros). El análisis se ha basado en una metodología de carácter no paramétrico, DEA, que actúa como modelo de control de gestión en las organizaciones. Esta información es útil para detectar puntos anómalos en la gestión, así como también sirve de guía para detectar posibles áreas y servicios ineficientes. Los coeficientes de ineficiencia pueden ser utilizados como otro indicador para las compañías eléctricas, y para poder llevarse un seguimiento anual de la eficiencia de las compañías eléctricas en el camino hacia la completa liberalización del sector en España.

Tras el análisis practicado a las principales compañías eléctricas, se puede decir que la mayoría resultan ser eficientes en el año 2001, en pleno periodo de liberalización del sector y en un marco de significativa desaceleración de la economía internacional, es decir, las principales empresas eléctricas son ahora más eficientes en un mercado de competencia que lo eran en un mercado regulado por el Estado. En el estudio se ha podido observar que la cuota de mercado de cada empresa (medido por el número de clientes) es un factor determinante en la explicación de las eficiencias alcanzadas, lo que significa que el servicio eléctrico está organizado acorde con la demanda potencial reflejada mediante el número de clientes.

Entre las principales conclusiones que se podrían destacar de este análisis, en primer lugar, habría que decir que en su conjunto, las empresas eléctricas a lo largo del periodo 1998-2001 tienden a trabajar en eficiencia. En segundo lugar, desde el punto de vista particular de cada empresa, Endesa e Iberdrola destacan como empresas eléctricas que han sabido adaptarse a los ajustes y requerimientos del nuevo mercado de electricidad. El comportamiento seguido por las citadas compañías ha sido creciente en materia de eficiencia. Sin embargo, como punto débil principal habría que mencionar la falta de cuidado a la calidad del suministro eléctrico.

Por otra parte, Hidrocantábrico y Unión Fenosa tienen cada una de ellas una evolución peculiar. Hidrocantábrico es la única empresa eléctrica que al comienzo del proceso de liberalización está trabajando en términos de eficiencia técnica. Sin embargo, en los años sucesivos Hidrocantábrico se trata de adaptar al nuevo marco eléctrico, abandonando su posición de eficiencia. Finalmente, Unión Fenosa es la empresa que ha desarrollado una conducta más irregular en el aprovechamiento de sus recursos para la producción de energía, en los ejercicios 1998, 1999 y 2000, situación que cambia en el año 2001. Una de las principales causas que han llevado a Unión Fenosa a alcanzar niveles bajos de eficiencia es el exceso de personal.

En el estudio, el ejercicio 1998 es considerado como un año *puente* donde realmente no se detectan acciones por parte de las empresas eléctricas encaminadas a afrontar la libre competencia en las actividades de generación y comercialización. El comportamiento del ejercicio 1998 se percibe en el trabajo como un año más representativo de un mercado todavía tutelado que de un mercado regulado.

En cuanto a los dos perspectivas de eficiencia planteadas, desde el punto de vista del inmovilizado y desde el factor trabajo, no se detectan diferencias significativas entre las compañías a lo largo de los años analizados.

Con estos resultados se concluye que durante los años estudiados, en un entorno de incertidumbre, las empresas eléctricas han trabajado en eficiencia, llevando a cabo las estrategias que desde el punto de vista empresarial han considerado oportunas. A pesar de la complejidad y los efectos que han producido la separación jurídica de las actividades de generación y distribución, este hecho no ha impedido que las empresas eléctricas trabajen en eficiencia.

La principal compañía generadora, distribuidora y comercializadora en la C.A.A., Endesa, ha mejorado sus niveles de eficiencia durante el periodo de liberalización analizado. En principio se podría afirmar que este *buen* comportamiento no va a afectar a la redistribución del mercado residencial eléctrico en la C.A.A. en el corto y/o medio plazo, y, por lo tanto, a la consideración de nuevas estructuras funcionales en los modelos de previsión de demanda.

A continuación, en la tabla V.4.6 se reflejan, de manera resumida, las principales *debilidades* y *fortalezas* que se observan en los cuatro mayores grupos eléctricos en España en el año 2001. Estas *debilidades* han de entenderse como *puntos de actuación*, los cuales favorecerían la eficiencia técnica de las compañías. Por otra parte, las *fortalezas* deben ser mantenidas por las compañías, llegando incluso a mejorarlas.

Tabla.V.4.6. Principales debilidades y fortalezas de las compañías eléctricas españolas en el año 2001.

Empresas	*Debilidades*	*Fortalezas*
Endesa	TIEPI	Nº CLIENTES ENERGÍA FACTURADA
Hidrocantábrico	TIEPI Nº CLIENTES	ENERGÍA FACTURADA
Iberdrola	TIEPI	Nº CLIENTES ENERGÍA FACTURADA
Unión Fenosa	TIEPI Nº CLIENTES ENERGÍA FACTURADA	

Elaboración. Propia.

BIBLIOTECA VIRTUAL

En la tabla puede observar que las dos mayores compañías eléctricas españolas, Endesa e Iberdrola, tienen una participación en el mercado eléctrico acorde con sus dimensiones empresariales, ya que sus cuotas de mercado representan un alto porcentaje del mercado eléctrico español, así como también en el volumen de energía facturada. Esto ha podido ser provocado por la expansión que han realizado ambas compañías hacia el mercado gasista, el cual también ha sido un negocio liberalizado.

Por otra parte, las compañías Hidrocantábrico y Unión Fenosa no han ampliado su posición en el mercado respecto a las compañías anteriormente comentadas, produciéndose una infrautilización en sus recursos disponibles. Además, la fiabilidad del suministro eléctrico, medido a través del TIEPI, se debe convertir por parte de las cuatro compañías en un objetivo a superar, el cual esta pendiente a lo largo de todo el análisis desarrollado. El cumplimiento de este objetivo anterior se puede convertir en un factor fundamental para la captación de clientes en un mercado de creciente competencia.

CAPÍTULO VI
CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

VI.1. CONCLUSIONES GENERALES

El incremento del nivel de vida en las sociedades industrializadas está asociado a una mayor demanda de energía eléctrica. Una de las prioridades básicas de una compañía eléctrica ha de ser satisfacer las necesidades energéticas de sus clientes de la forma más económica posible persiguiendo, al mismo tiempo, un nivel aceptable de calidad del servicio.

En este trabajo se ha analizado la previsión de demanda doméstica de la energía eléctrica en la Comunidad Autónoma de Andalucía. La demanda de energía eléctrica se suele clasificar según el sector de actividad al que vaya dirigida, distinguiéndose habitualmente, la demanda realizada por los sectores industrial, comercial y residencial. El estudio se centra en el sector doméstico o residencial por su repercusión socioeconómica general y por su importante contribución a las puntas de demanda de electricidad.

Se ha tenido como objetivo estudiar los siguientes aspectos:

- Modelar y contrastar expresiones funcionales, basadas en modelos econométricos de demanda anual de energía eléctrica a nivel residencial que permitan predecir la demanda anual en la Comunidad Autónoma de Andalucía.
- Elaborar un programa informático en entorno de ventanas, basado en la técnica de remuestreo denominada bootstrap, que mejora la estimaciones puntuales y confidenciales de la mencionada previsión debido a la falta de software comercial adecuado para el objetivo citado.
- Modelar y contrastar expresiones funcionales, basadas en modelos econométricos y de series temporales de demanda mensual de energía eléctrica que permitan predecir de forma precisa la demanda mensual en la C.A.A. Mejorar las predicciones obtenidas mediante una combinación de algunos de los modelos.

- Analizar la eficiencia técnica de los principales grupos eléctricos de ciclo completo en España, a los cuales está afectando la última etapa liberalizadora del sector eléctrico y que supone la libre elección de suministrador por parte de los consumidores domésticos.

Para llevar a cabo todos estos objetivos, el presente trabajo se ha divido en cinco capítulos.

Una vez descritos los objetivos e intereses del trabajo de investigación, así como la clasificación y metodología de los principales modelos de previsión de demanda de electricidad, se continúa con la aplicación al caso específico de la demanda residencial de la Comunidad Autónoma de Andalucía. El capítulo III analiza, a través de los modelos econométricos, la demanda anual de energía eléctrica a nivel residencial o doméstico durante el periodo 1984 a 1999. La construcción de estos modelos se apoyan en variables económicas, tales como el precio de la electricidad y sus productos sustitutivos, así como las variables crecimiento económico, PIB y la demanda de electricidad retardada un periodo. Se consideran seis modelos que se analizan principalmente desde el punto de vista de la estructura de los residuos, estabilidad y predicción. El modelo de productos cruzados retardados, modelo 3PCR, es el que alcanza un poder explicativo mayor según el amplio conjunto de estadísticos considerados. Este modelo también ha resultado adecuado desde el punto de vista de su capacidad predictiva, ya que los intervalos construidos ,al noventa y cinco por ciento de nivel de confianza, contienen al valor real de la demanda observada de energía para los años 1998 y 1998-1999 al realizar predicción *ex-post* de estos periodos.

Modelo 3PCR:

$$LnQ_i = 9{,}606683 - 5{,}834969 LnPE(-1)_i + 0{,}339545(LnPELnPIB)(-1)_i$$

Para completar el estudio de los modelos de demanda anual se ha aplicado el método de remuestreo bootstrap al modelo 3PC. Los resultados de dicha aplicación, a través de soporte informático propio elaborado específicamente para el presente trabajo, mejoran la precisión de las

estimaciones confidenciales respecto a las obtenidas por el método convencional.

En el capítulo IV, se continúa con el análisis de la demanda residencial de energía eléctrica, pero referido a un horizonte temporal más a corto plazo que en el capítulo III. Los datos históricos de demanda que se estudian en este capítulo son mensuales, en el periodo enero de 1995 a diciembre de 2000. En este capítulo se ha realizado un estudio detallado y comparativo de quince modelos univariantes, siendo el primer modelo dinámico univariante, BJ-1, analizado el que mejor se ajusta a los datos reales y mejor previsión realiza. La estructura funcional del modelo BJ-1 se expresa como:

$$\nabla\nabla_{12} \ln q_t = (1-0{,}8103B)(1-0{,}7480B^{12})\varepsilon_t$$

También se relaciona en este capítulo la influencia que ejerce la temperatura en la demanda residencial de energía eléctrica en la Comunidad Autónoma de Andalucía. Una vez considerada la temperatura como factor a tener en cuenta en el estudio de la demanda de interés se han construido setenta y nueve modelos de regresión dinámica en los que la temperatura interviene como variable explicativa y en bs que se contemplan temperaturas *neutras* que oscilan desde los quince a los veintidós grados centígrados, tal como se puede observar en el anexo del trabajo. En la predicción *ex–post* los datos de demanda realmente observados se encuentran mayoritariamente dentro de los intervalos de confianza de los modelos construidos. Finalmente, se ha procedido a mejorar las previsiones de demanda mediante la combinación de los métodos de pronóstico no dinámicos y dinámicos analizados anteriormente. La combinación considerada de previsiones reduce los errores de predicción respecto a los de los métodos de predicción individuales.

Finalmente, en el capítulo V, se ha analizado cual será el comportamiento mayoritario del mercado doméstico en la C.A.A. a partir de 2003. Por ello se ha estudiado la eficiencia de los principales grupos eléctricos en España durante

los años que han transcurrido desde el inicio de la liberalización, es decir, el periodo 1998 a 2001. Este estudio se ha basado en la metodología de carácter no paramétrico denominada Análisis por Envoltura de Datos (DEA). Los resultados han evidenciado que la principal compañía suministradora de electricidad en Andalucía, Endesa, ha mejorado sus niveles de eficiencia hasta convertirse en una de las empresas eléctricas más eficientes.

Para finalizar nosotros pensamos que las principales contribuciones de este trabajo son:

- La construcción y contraste de modelos de demanda residencial de energía eléctrica en la C.A.A.
- La predicción de la demanda anual a partir de los modelos anteriores.
- La elaboración de un programa informático en entorno de ventanas basado en la metodología bootstrap con la que se consiguen intervalos de confianza más precisos que con la metodología convencional de los mínimos cuadrados.
- Una considerable reducción en el error de predicción respecto a la conseguida en los *mejores* modelos de predicción individuales construidos y contrastados en este trabajo mediante una combinación de tales predicciones.
- El estudio de la evolución de la eficiencia técnica de los principales grupos eléctricos de ciclo completo durante los cuatro años posteriores al inicio del proceso liberalizador en España.

V.2. FUTURAS LÍNEAS DE INVESTIGACIÓN

Actualmente una de las principales barreras para una aplicación más generalizada de la metodología bootstrap es la falta de un desarrollo adecuado de software en los programas comerciales, tales como SPSS, SAS, BMDP o MINITAB, y, aunque algunos de estos paquetes contienen ya algunas rutinas elementales de aleatorización y algún otro programa comercial como S-PLUS va contando progresivamente con una mayor librería de aplicaciones, (generalmente debida a usuarios particulares según haya sido su campo de interés) todavía como ha ocurrido en el caso de este trabajo, son mayoría las situaciones en que es necesario recurrir a una programación *ad-hoc*.

Entre las posibles extensiones del trabajo informático elaborado específicamente para este trabajo se encuentran la aplicación del bootstrap a ecuaciones de variable endógena retardada y a modelos econométricos de *m* ecuaciones simultáneas relativos a la demanda residencial de electricidad, estando, en este último caso, representados los residuos por un vector de dimensión *m* en vez de ser un escalar como en el modelo uniecuacional, generándose los datos artificiales mediante la resolución de dicho sistema de ecuaciones simultáneas.

Otras posibles extensiones de la metodología bootstrap cuando se disponga de mayor número de datos son los siguientes:

- La consideración de la aleatoriedad en las variables exógenas en lugar de considerarlas fijas.
- La aplicación de esta técnica de remuestreo a los modelos ARIMA generales, ya que actualmente la mayor parte de la investigación en este campo ha tenido como campo de aplicación modelos AR(p) con escaso número de términos autorregresivos.
- La extensión de un horizonte temporal mayor en la predicción multiperiodo con la correspondiente comparación de los modelos convencional y bootstrap.

- La evaluación de la calidad de los estimadores bootstrap lo cual al involucrar una *simulación de la simulación* puede aumentar considerablemente el tiempo del proceso informático necesario para ello.

La ampliación futura del número de datos históricos puede permitir:

- Tanto la exploración de nuevos modelos como la continuidad de la validez de algunos ya existentes que, a pesar de los escasos datos de los que hasta ahora se ha podido disponer, ya se han mostrado muy robustos en especial frente a eventuales modificaciones del valor de lo que se ha considerado como temperatura *neutra*.
- Elaborar modelos específicos para las provincias de la C.A.A., aunque no sería extraño que tales modelos difieran de una provincia a otra e incluso sean distintos de los que se consideran válidos para la C.A.A. en su conjunto, debido a las diferencias climáticas provinciales.
- La construcción de modelos *horarios* tanto a nivel de la C.A.A. como de nivel provincial posibilitando la investigación de comportamientos similares de la demanda en determinados grupos de días lo que podría ser abordado desde la perspectiva del Análisis de Conglomerados.
- La construcción de modelos *estables* de modelos de estructura ARIMA con función de transferencia correspondientes a diversos *inputs*, de los que la temperatura es un candidato obvio, pero no el único a investigar porque diversas variables socioeconómicas puedan mostrarse también influyentes en la demanda investigada.
- Entre las futuras líneas de investigación que se podrían desarrollar respecto a la eficiencia técnica a través de DEA, estarían en primer lugar, la consideración de otra compañía eléctrica, Enel (antes denominada Viesgo), a partir del año 1999. En segundo lugar, ampliar el estudio tanto en el intervalo temporal como al marco europeo.
- Uno de los requisitos fundamentales que debe cumplir todo sistema eléctrico es la fiabilidad del suministro eléctrico. El *Tiempo de Interrupción Equivalente de la Potencia Instalada* (TIEPI) es un indicador que las compañías eléctricas aplican con el objetivo de evaluar la continuidad del

suministro eléctrico, y por lo tanto la fiabilidad del mismo. En el trabajo desarrollado en el capítulo V, se detectaron las debilidades que presenta el sector eléctrico español hacia el TIEPI. A este respecto, como trabajo en curso se han planteado modelos para la explicación del TIEPI, a partir de variables explicativas basadas en el inmovilizado de las compañías. El documento "Procedimiento de Medida y Control de la Continuidad de Suministro Eléctrico" presentado por UNESA (Asociación Española de la Industria Eléctrica), ASEME (Asociación de Empresas Eléctricas) y CIDE (Centro de Investigación y Desarrollo) en cumplimiento con el artículo 108.3 del Real Decreto 1955/2000, viene a proponer los criterios y metodología a seguir para la recogida y tratamiento de los datos de la continuidad del suministro eléctrico, como son el TIEPI, NIEPI y el percentil del 80%.

BIBLIOTECA VIRTUAL

También sería interesante la comparación de la precisión de la predicción realizada con modelos econométricos y de series temporales con la efectuada mediante redes neuronales, pero es evidente que la capacidad de aprendizaje de estas últimas se verá facilitada con un mayor número de datos históricos de los que hasta ahora ha sido posible disponer.

En el presente trabajo, al tratar los datos mensuales, el horizonte temporal ha sido el año 2001. Cabe la posibilidad de que la variación del horizonte temporal modifique el *mejor* modelo a considerar. La influencia de este factor en la variable de respuesta (alguno de los parámetros considerados en la evaluación de predicciones) puede ser investigada vía Diseño de Experimentos.

Otra forma de extender la investigación realizada en este trabajo es a través del estudio de los datos anómalos o influyentes en los que distintas medidas entre las que se encuentra la *distancia de Cook* pueden contribuir a adoptar una decisión sobre los mismos.

APÉNDICE

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

El Análisis por Envoltura de Datos

El DEA se ha convertido, desde los principios de su desarrollo (1978), en una de las herramientas centrales disponibles a los *managment scientist* para el análisis de los objetivos de la organización. Las primeras aplicaciones del DEA se produjeron en las instituciones estadounidenses, hoy en día, existen centros de investigación distribuidos a lo largo de todo el mundo, los cuales han sido el origen de muchas nuevas ideas así como también nuevas aplicaciones. Esta técnica se ha utilizado con gran profusión en el sector público, como es el caso particular de los centros de educación y hospitales, también se ha utilizado en sectores privados, como el sector bancario y el sector eléctrico, Ejemplos de estos últimos son los trabajos de Miliotis(1992), Hjalmarsson (1992), Agrell (1998), Forsund(1998) y Lo Feng-Yu (2001).

La productividad de una/s organización/es es útil cuando el ente que se estudia tiene capacidad de decidir acerca de la cantidad de cada uno de los recursos que están siendo utilizados y los productos y/o servicios que se generan. Por ello, a la unidad productiva se le denomina decisoria, en la literatura anglosajona Decision Making Unit. Respecto a la consideración de las distintas DMU's estas pueden ser entidades financieras, departamentos de almacenes y supermercados, y extenderlo a fabricantes de coches, hospitales, colegios, bibliotecas públicas y más entidades. En la ingeniería, se pueden considerar DMU´s por ejemplo los aviones o sus componentes como en las máquinas de propulsión. Con el objetivo de asegurar las comparaciones relativas, cada DMU integrante de un grupo de DMU's es comparada con otra unidad de decisión con el propósito de tener un cierto grado de decisión.

Como su propio nombre indica, el Análisis Envolvente de Datos supone la identificación de una frontera eficiente que envuelve o encierra al conjunto de observaciones de la muestra objeto de estudio. El DEA es una técnica que evalúa la eficiencia de unidades de decisión, DMU's, siendo los elementos utilizados para la comparación, los recursos de entrada y los recursos de

salida, también denominados en términos anglosajones inputs y outputs respectivamente.

Si bien, se han aplicado distintos modelos de DEA para tratar problemas de gestión o producción, todos ellos coinciden en que están orientados hacia los conceptos frontera, asociados con localizaciones eficientes e ineficientes de las unidades de decisión responsables de convertir múltiples inputs en múltiples outputs. El método DEA está orientado hacia la productividad individual y la identificación de relaciones extremas entre input y output para las diferentes unidades de decisión mientras que las regresiones estadísticas, índices de productividad, están orientados hacia la producción media o output medio para un input dado.

BIBLIOTECA VIRTUAL

El análisis se realizará siempre y cuando las unidades de decisión consideradas consuman el mismo tipo de recursos para la obtención del mismo tipo de productos y/o servicios. La técnica DEA está basada en la noción económica de la optimalidad de Pareto y realiza una comparación transversal de las diferentes entradas y salidas de cada DMU con todas las demás. Entre las características de esta técnica pueden destacarse

- Flexibilidad, puesto que no requiere especificar una forma funcional de antemano.
- Capacidad para trabajar con múltiples factores, a diferencia del análisis tradicional de ratios que considera sólo dos factores de la actividad representados en el numerador y denominador.
- No se basa en la naturaleza estocástica de los datos observados.

El procedimiento DEA se puede enfocar bien hacia el aumento de los outputs o bien a la conservación de los inputs. En el contexto de los incrementos de outputs el objetivo es maximizar la cantidad de outputs obtenido, dado un nivel de inputs, denominándose *orientación de salida*. Así, una unidad es considerada técnicamente ineficiente si alguna otra unidad, o combinación convexa de unidades, no puede utilizar ningún input y producir al

menos la misma cantidad de outputs[16]. El contexto de mantener el nivel de recursos o inputs, se denomina *orientación de entrada*, y el objetivo es minimizar el consumo de recursos, dado un nivel de productos. Dada una DMU, se dice que es técnicamente ineficiente si alguna otra unidad, o alguna combinación lineal de otras unidades, produce al menos la misma cantidad de outputs y utiliza menos inputs[17]. Por otra parte, se dice que una unidad es técnicamente eficiente, con orientación de entrada, si produce los mismos outputs utilizando menos inputs.

En los últimos años se han visto una gran variedad de aplicaciones de DEA, utilizándose en la evaluación de los objetivos de muchos tipos de entidades las cuales desarrollan distintas actividades en diversos países.

Como ya se ha comentado, en DEA, los organismos objeto de estudio son denominados un DMU. La definición de DMU es bastante amplia para permitir flexibilidad en su utilización para el mayor número de aplicaciones.

Los dos modelos que frecuentemente se asocian a la metodología DEA son el modelo de CCR, desarrollado por Charnes, Cooper y Rhodes en 1978, y el modelo BCC, desarrollado por Banker, Charnes y Cooper en 1984. El modelo CCR produce una frontera eficiente a retornos de escala constantes. Las eficiencias relativas calculadas en el modelo CCR son, en términos generales, una puntuación de su eficiencia. Así, la puntuación de la eficiencia de todas las DMU´s en los modelos DEA estarán comprendidas entre los valores 0 y 1. Por otra parte, el modelo BCC produce una frontera eficiente a retornos de escala variables y evalúa tanto la eficiencia técnica como la eficiencia de escala. A continuación se tratará de exponer la metodología de ambos modelos.

A continuación se va a tratar uno de los modelos más básicos en DEA, el modelo CCR el cual para cada DMU, crea entradas ($x_0 \in \Re^m$) y salidas

[16] al menos una unidad de output.
[17] al menos una unidad de input.

($y_0 \in \Re^s$) virtuales a través de los pesos asociados ($v \in \Re^m$, $u \in \Re^s$), de momento desconocidos, a cada uno de estas entradas y salidas:

$$\text{Entradas Virtuales} = \sum_{i=1}^{m} v_i x_{i0} = v_1 x_{10} + \ldots + v_m x_{m0}$$

$$\text{Salidas Virtuales} = \sum_{r=1}^{s} u_r y_{r0} = u_1 y_{10} + \ldots + u_s y_{s0}$$

Los pesos óptimos pueden y generalmente cambiaran entre las distintas DMU´s consideradas.

Suponer que se tienen *n* DMU's: DMU_1, DMU_2...,DMU_n. Los elementos considerados entradas y salidas de las DMU´s *j=1,2,..., n* son seleccionados de la forma que se verá a continuación:

- Los valores para cada entrada y salida tienen que ser positivos para todas las DMU's.
- Los elementos (entradas, salidas y la elección de DMU´s) deberían reflejar un interés de análisis o de controlar los componentes que formaran parte de las relaciones de eficiencia de las DMU´s.
- La medida de las unidades de las distintas entradas y salidas no necesita ser congruente. Algunos pueden tener como medida el número de personas, superficie de espacio, unidades monetarias...

Consideremos un modelo simple con inputs y outputs para cada DMU *j=1, ..., n*, como (X_i, Y_j), donde $X_j=(x_{1j},...,x_{ij},...,x_{mj})$ es un vector de inputs observados e $Y_j=(y_{1j},...,y_{rj},...,y_{sj})$ un vector de outputs observados de las DMU*j*.

La relación existente entre los diferentes inputs y outputs en cada una de las unidades de decisión se representa en la siguiente figura:

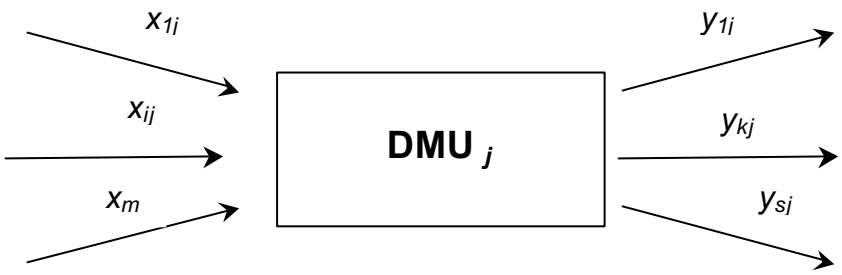


Figura.A.V.1.Flujos de Entrada y Salida de la *j*-ésima Unidad de Decisión. Elaboración. Propia.

BIBLIOTECA VIRTUAL

La técnica DEA fue desarrollada como una extensión del clásico modelo *ratio de eficiencia*, ecuación 22, el cual maximizaba para cada DMU la suma ponderada de los outputs entre la suma ponderada de los inputs, donde los pesos eran determinados por el modelo. Así, este problema de programación fraccionada es transformado en un problema de programación lineal, ver ecuación 23. Y aplicando la teoría de dualidad a dicho problema de programación lineal se obtiene su expresión dual (24).

La formulación de dicho modelo en forma de programación fraccionada, se denomina modelo ratio y su expresión es la siguiente:

$$max_{(u,v)} \quad h_0 = \frac{\sum_{r=1}^{s} u_r y_{r0}}{\sum_{i=1}^{m} v_i x_{i0}} \tag{22}$$

Sujeto a:

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1; \quad j = 1,...,n$$

$$u_r, v_i \geq 0, \quad i = 1,..., m; \quad r = 1,...,s$$

Los valores conocidos observados de las entradas y salidas de las *j* DMU´s de la muestra considerada; x_{r0} *e* y_{r0} los valores de la DMU unidad de decisión que va a ser evaluada en una prueba. De forma que las variables de ponderación o variables del modelo serían v_i^* y u_r^*, los cuales generarán el óptimo h_0^*=1, sólo si la unidad evaluada es eficiente. De esta forma, la función objetivo siempre tomará valores entre 0 y 1, para las distintas unidades estudiadas.

Por ejemplo, supóngase una función de producción que relaciona el output (*Y*) y dos inputs: el trabajo (*L*) y el capital (*K*). La función *Y* (*K,L*) muestra retornos a escala constantes (CRS). Ahora se divide los inputs entre el output, obteniendo un cociente, *L/Y*, representado a través del eje de abscisa y *K/L* representado en el eje de ordenadas. En tal espacio se dibuja una isocuanta envolvente de todas las empresas envolventes (figura A.V.2). El punto B representa una empresa ineficiente mientras que los puntos A, D, E son eficientes. Farell definió la eficiencia técnica a través del cociente: TE_{CRS}= OC/OB, que alcanzará valores entre 0 y 1. Una empresa es eficiente[18] si la eficiencia técnica es igual a la unidad.

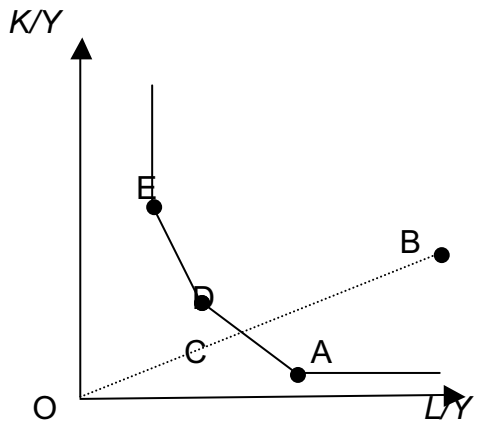


Figura A.V.2. Una medida de la eficiencia técnica: TE_{CRS}= OC/OB.
Fuente. R¹ czka, J.(2001)
Elaboración.Propia

Para poder resolver el modelo anterior, ha de expresarse en forma de problema lineal, para ello se fija el denominador a la unidad, y la expresión resulta en función del numerador, el cual tendrá que maximizarse, tal como se refleja en la expresión:

$$\max \ h_0 = \sum_{r=1}^{s} u_r y_{r0} \qquad (23)$$

Sujeto a:

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0 \quad , j= 1,...,n$$

$$\sum_{i=1}^{m} v_i x_{i0} = 1$$

$$u_r, v_i \geq \varepsilon$$

dónde $\varepsilon>0$ es una cantidad pequeña no arquimediana, y por tanto en las restricciones donde aparece los pesos nunca pueden ser nulos. Una vez resueltos los n problemas propuestos, se obtendrá un subconjunto C formado por las unidades DMU_j que han resultado ser eficientes al resolver el modelo dado, siendo h_0=1. Siendo la DMU que no cumple esta condición ineficiente respecto al subconjunto C definido, y su eficiencia será valorada como h_0<1 y su ineficiencia, (1- h_0).

Además, (v^*, u^*) son el conjunto de pesos más favorables para la DMU evaluada, DMU_0, en el sentido de maximizar h_0. Siendo:

v_i^* el peso óptimo para la *i-ésima* entrada, expresando su grado de importancia y contribución a h_0.

u_r^* el peso óptimo para la *r-ésima* salida, expresando su grado de importancia y contribución a h_0.

Sin embargo, es más frecuente utilizar las expresiones del problema dual para analizar los resultados obtenidos. La función objetivo de este modelo intenta encontrar un valor mínimo para una cantidad de factor w_0^*. Además, la función objetivo busca los mayores valores que pueden alcanzar los inputs y outputs, es decir, encuentra el punto de referencia en la función de producción

empírica que caracteriza a la peor DMU. A continuación se expresa (24) modelo dual, conocido como *forma envolvente*:

$$Min\, w_0 - \varepsilon\left[\sum_{i=1}^{m} z_i + \sum_{r=1}^{s} s_r\right] \qquad (24)$$

Sujeto a:

$$0 = w_0 x_{i0} - \sum_{j=1}^{n} x_{ij}\lambda_j - z_i \ , \ i=1,\ldots, m$$

$$y_{r0} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r, \quad r=1,\ldots,s$$

$$\lambda_j, z_i, s_r \geq 0, \quad \forall\, i, j, r$$

Las n variables λ_j son las correpondientes a las n primeras restricciones del problema primal, w_0, la variable correspondiente a la restricción restante, y z_i y s_r, que denominamos variables de holgura, son las correspodientes a las $s+m$ cotas existentes. La resolución del modelo consta de dos etapas o fases.
En la primera fase se resuelve el siguiente modelo:

$$Min\, w_0$$

Sujeto a:

$$0 = w_0 x_{i0} - \sum_{j=1}^{n} x_{ij}\lambda_j - z_i \ , \ i=1,\ldots, m$$

$$y_{r0} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r, \quad r=1,\ldots,s$$

$$\lambda_j, z_i, s_r \geq 0, \quad \forall\, i, j, r \ ,$$

$$w_0 \text{ libre}$$

Una vez resuelto la primera fase, obteniendo el valor de w_0^* se resuelve la segunda fase:

$$Min \ \ \varepsilon\left[\sum_{i=1}^{m} z_i + \sum_{r=1}^{s} s_r\right]$$

Sujeto a:

$$0 = w_0^* x_{i0} - \sum_{j=1}^{n} x_{ij}\lambda_j - z_i \ , \ i=1,\ldots, m$$

$$y_{r0} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r, \quad r=1,\ldots,s$$

$$\lambda_j, z_i, s_r \geq 0, \quad \forall\, i,\, j,\, r$$

Las funciones objetivo de ambos problemas coinciden en el óptimo y por tanto se tiene que:

$$h_0^* = w_0^* - \varepsilon\left[\sum_{i=1}^{m} z_i^* + \sum_{r=1}^{s} s_r^*\right] = \sum_{r=1}^{s} u_r y_{r0}$$

Así, pues, las soluciones obtenidas del modelo son respectivamente w_0^*, z_i^*, s_r^* y λ_j^*.

donde

w_0^* es la proporción de entradas actuales que deben utilizarse para alcanzar la eficiencia.

z_i^* es el vector columna de las variables de holgura correspondientes a la desigualdades de las entradas o inputs, sus unidades estarán expresadas según las variables de entrada.

s_r^* es el vector columna de las variables de holgura correspondientes a las desigualdades de las salidas u outputs, sus unidades estarán expresadas según sus variables de salida.

λ_j^* es un vector columna cuyas componentes son multiplicadores adimensionales, miden la proximidad de la proyección resultante de cada DMU con las unidades eficientes de las que es combinación lineal.

De forma que si en la unidad evaluada w_0^* toma el valor 1 y las variables de holgura z_i^* y s_r^* son cero entonces dicha unidad es eficiente. No resultando eficiente si al menos una variable de holgura, bien z_i^*, bien s_r^*, es distinta de cero. La variable épsilon asigna un valor por cada nivel de ineficiencia residual o estimada a través de las variables de holgura.

Así, los valores de los inputs y outputs que definirán los niveles de eficiencia serán:

$$\hat{x}_{i0} = w_0^* x_{i0} - z_i^* \qquad (25)$$

Expresándose mediante $\hat{x}_{i0}$ el nivel de inputs que deberá conseguir la unidad evaluada para trabajar en eficiencia. El ahorro equiproporcional de cada uno de los inputs empleados por la unidad evaluada sería $1 - w_0^*$, tal como se expresa a continuación:

$$\Delta x_0 = x_0 - \left(w_0^* x_0 - z_i^*\right) = (1 - w_0^*) x_0 + z_i^*$$

La ecuación 26 expresa el nivel de outputs dado para la anterior combinación de inputs de la unidad evaluada.

$$\hat{y}_{r0} = y_{r0} + s_r^* \qquad (26)$$

Una vez fijado la cantidad mínima de input/s que maximiza el output, $\hat{x}_{i0}$, puede ocurrir que el output se incremente (en este caso habría ampliaciones de carácter rectangular). Este incremento sería de s_r^*.

Hasta ahora el modelo DEA presentado tiene una *orientación bien hacia las entradas*, también denominado *modelo CCR- Input*, siendo también posible estudiar su orientación hacia las salidas, a través del *modelo CCR-Output*. Es decir, el modelo CCR puede tener como objetivo maximizar las salidas permaneciendo constantes las entradas, que va a ser el modelo que se va a expresar a continuación :

$$Min\ h_0 = \sum_{i=1}^{m} v_i x_{i0}$$

Sujeto a:

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0\ , j=1,\ 2,...,\ n$$

$$\sum_{j=1}^{n} u_r y_{r0} = 1$$

$$v_i\ ,\ u_r \geq \varepsilon$$

Se advierten las mismas consideraciones realizadas en el modelo anterior, *modelo CCR-Input*. Así, el problema dual en el *modelo CCR-Output* se obtiene de forma similar:

$$Max \quad t_0 + \varepsilon\left[\sum_{i=1}^{m} z_i + \sum_{r=1}^{s} s_r\right]$$

Sujeto a:

$$0 = x_{i0} - \sum_{j=1}^{n} x_{ij}\lambda_j - z_i\ , \qquad i=1,...,\ m$$

$$t_0 y_{r0} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r, \quad r=1,...,s,$$

$$\lambda_j, z_i, s_r \geq 0, \quad \forall\ i,\ j,\ r$$

A continuación se ilustra el uso del modelo CCR a través de un ejemplo a pequeña escala.

Ejemplo ilustrativo

La tabla A.V.1 muestra la relación entre la cantidad de input empleado y el output obtenido de 5 empresas pertenecientes a un mismo sector. Se desea conocer si se está utilizando, para cada una de las empresas, de forma eficiente el nivel de input, sabiendo que todas las empresas tienen la misma capacidad de producción.

Tabla.A.V.1. Datos problema CCR-Input.

Empresa	Input	Output
1	8	5
2	5	5
3	6	3
4	7	2
5	5	3

Elaboración. Propia.

Es posible evaluar la eficiencia de la empresa 1, resolviendo el siguiente problema de programación lineal:

$$\text{Max } h_0 = 5u$$

$$\text{s.a: } 8v = 1$$

$$5u \leq 8v \ (1) \qquad 2u \leq 7v \ (4)$$

$$5u \leq 5v \ (2) \qquad 3u \leq 5v \ (5)$$

$$3u \leq 6v \ (3)$$

$$u, v \geq \varepsilon$$

La solución óptima del problema viene dado por $v^* = 0{,}125$, $u^* = 0{,}2$, $h_0^* = 0{,}6250$. Resultando ser el nivel de eficiencia de 1 del 62,50%. El conjunto de empresas de referencia para la empresa 1 resulta ser la empresa 2 para $v^* = 0{,}125$, $u^* = 0{,}2$. Además los niveles de input y de output de la empresa 2 son empleados para la caracterización de la empresa 1 y es calificada como empresa ineficiente incluso con los mejores pesos que los datos admiten para 1.

Se procedería de forma similar con las cuatro empresas restantes y se obtendrían los resultados de la tabla A.V.2.

Tabla. A. V. 2. Resultados ejemplo CCR-Input.

Empresa	Eficiencia(%)	Conjunto de Referencia
1	62,50%	2
2	100,00%	2
3	50,00%	2
4	28,57%	2
5	60,00%	2

Elaboración. Propia.

La única empresa que trabaja en eficiencia y que pertenece al conjunto de referencia es la empresa 2, tal como se puede observar en la figura A.V.3, la cual representa la situación geométricamente. La frontera eficiente está representada por la recta que pasa solamente por el punto 2. Los índices de eficiencia, h_0^*, de la tabla A.V.2 destacan la necesidad de trasladar cada empresa a la frontera eficiente. Por ejemplo, el valor de $h_0^* = 0{,}6250$ aplicado a la empresa 1 indica que tendrá que reducir en un (1-0,6250)100% su nivel de input para trabajar dentro de la frontera eficiente. Igualmente la empresa 3 bastará con que emplee 0,5*6=3 unidades de input para situarse en la frontera eficiente, reduciendo un 50% su nivel actual de output. Igual interpretación merecen el resto de las empresas consideradas. Si se analizan el resto de empresas, se concluiría que la empresa 1 tiene que utilizar 5 (8x0,6250) unidades de input, por lo tanto tendría que reducir en tres unidades su nivel actual de input. Por último, la empresas 4 y 5 tendrían que aplicar reducciones de 5 y 2 unidades de input, respectivamente.

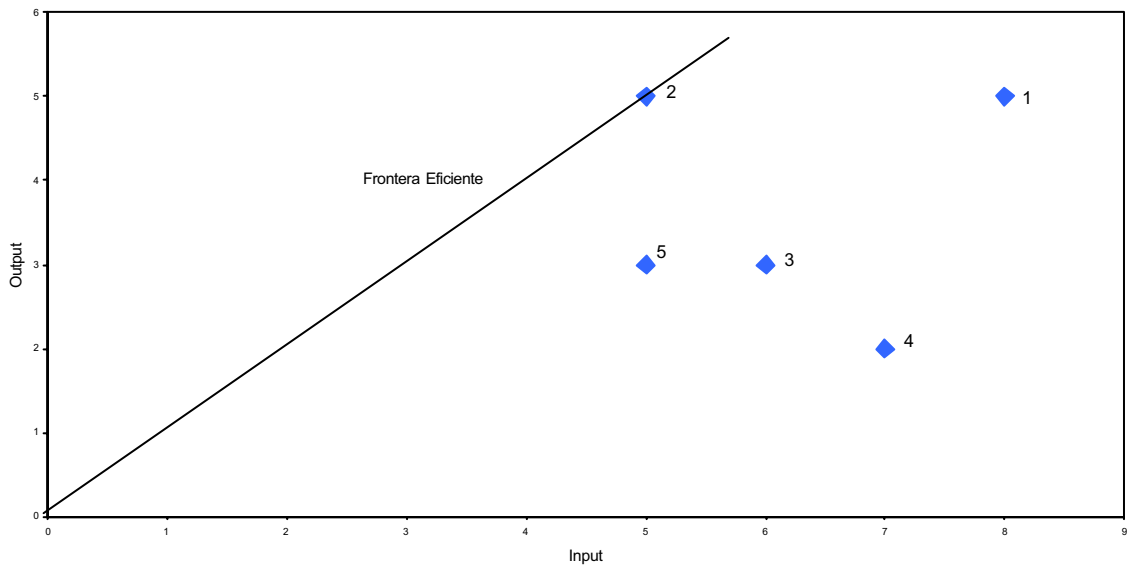


Figura A.V. 3.Situación geométrica de las empresas. CCR-Input.
Elaboración. Propia

Los resultados del problema obtenidos a través del paquete informático EMS se exponen en la figura A.V.4, donde todas las columnas, a excepción de la primera, representan variables del problema dual, cuyo significado es el siguiente:

DMU :hace referencia a la empresa que se está evaluando.

Score :es el índice de eficiencia, h_0^* o w_0^*.

Input {I}{W}: es el peso óptimo del recurso empleado, en el caso que la orientación sea de entrada.

Output {O}{W}: es el peso del producto obtenido.

Benchmarks: define el conjunto de referencia para las unidades de decisión. Entre paréntesis se expresa la intensidad (λ_j) con la que la unidad de decisión se proyecta sobre la unidad/des del conjunto de referencia.

{S}Input{I} : es la variable de holgura del recurso, z_i, en el caso considerado z_1, expresando la reducción de carácter rectangular (variación horizontal) a realizar por las unidades para proyectarse sobre la frontera eficiente.

{S}Output{O}: es la variable de holgura del producto, s_r, expresando la reducción de carácter rectangular (variación vertical) a realizar las unidades para proyectarse sobre la frontera eficiente.

EMS - [C:\Desarrollo\DEA\EjemploCCR-Anexo.xls_CRS_RAD_IN]

File Edit DEA Window Help

	DMU	Score	Input(I)(W)	Output(O)(W)	Benchmarks	(S) Input(I)	(S) Output(O)
1	1	62,50%	0,125	0,200	2 (1,00)	0,00	0,00
2	2	100,00%	0,200	0,200	4		
3	3	50,00%	0,167	0,333	2 (0,60)	0,00	0,00
4	4	28,57%	0,143	0,500	2 (0,40)	0,00	0,00
5	5	60,00%	0,200	0,333	2 (0,60)	0,00	0,00

Input Output Data C:\Desarrollo\DEA\EjemploCCR-Anexo.xls

Figura.A.V.4. Pantalla EMS. Problema CCR-Input.
Elaboración. Propia.

Una de las principales críticas al modelo de Charnes, Cooper y Rhodes (1978) fue el hecho de que el modelo que ellos plantearon calcula la eficiencia de las unidades bajo la hipótesis de retorno de escala constantes. De esta forma, la dimensión de las unidades que se comparaban debe ser parecida, es decir, las unidades no eficientes han de tener capacidad para alcanzar el nivel de eficiencia de la unidad o unidades más eficientes. A este tipo de eficiencia, planteada en el modelo anterior, se le denomina eficiencia global y tiene lugar cuando la unidad seleccionada de referencia es la de mayor productividad de todas la unidades que se están estudiando. Pero también se podrían considerar problemas donde las unidades tuvieran dimensiones diferentes al de las unidades eficientes no pudiendo ser capaces de conseguir alcanzar la productividad de estas. Así, pues el estudio se realizará a través de la eficiencia técnica, esta hipótesis se denomina *retorno de escala variable*.

Para solventar este inconveniente y poder trabajar con problemas de retorno de escala variable, Banker, Charnes y Cooper [20] (1984), añaden al modelo anterior:

$$\sum_{j=1}^{n} \lambda_j = 1$$

[20]en adelante lo denominaremos modelo BCC.

Esta restricción impone que el punto de referencia en la frontera de producción para la DMU_0 sea combinación convexa de las otras DMU´s eficientes observadas. Introduciendo la restricción en el modelo CCR-Input, resultaría el siguiente modelo:

$$Min\ w_0 - \varepsilon\left[\sum_{i=1}^{m} z_i + \sum_{r=1}^{s} s_r\right]$$

$$\text{Sujeto a:} \quad 0 = w_0 x_{i0} - \sum_{j=1}^{n} x_{ij}\lambda_j - z_i \quad ,i=1, ..., m$$

$$y_{r0} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r \quad , r=1, ..., s$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j\ , z_i, s_r \geq 0,\ \forall i, j, r$$

Siendo su expresión en forma primal la siguiente:

$$Max\ h_o = \sum_{r=1}^{s} u_r y_{rj} - u_0$$

Sujeto a:

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + u_0 \leq 0$$

$$\sum_{i=1}^{m} v_i x_{i0} = 1$$

$$u_r, v_i \geq \varepsilon;\ u_0 \text{ variable libre}$$

donde u_0: es un indicador de retornos de escala.

Al igual que el modelo CCR-Input las unidades eficientes tomarán valores $w_0^* = 1$ y $z_i^*, s_r^* = 0$.

El modelo puede estar orientado también hacia los outputs, es lo que se denomina un *modelo BCC con orientación de salida*. En este caso, el modelo para calcular los niveles de eficiencia se expresaría analíticamente de la forma siguiente:

$$Max\ t_0 + \varepsilon\left[\sum_{i=1}^{m} z_i + \sum_{r=1}^{s} s_r\right]$$

Sujeto a:

$$0 = x_{i0} - \sum_{j=1}^{n} x_{ij}\lambda_j - z_i \quad , \quad i=1, ..., m$$

$$y_{r0}t_0 = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r \ , \qquad r = 1,...,s$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j, s_i, s_r \geq 0, \ \forall i, j, r$$

El problema expresado en forma primal es el siguiente:

$$Min\ h_0 \ = \ \sum_{i=1}^{n} v_i x_{i0} - v_0$$

$$\text{Sujeto a :} \qquad \sum_{r=1}^{n} u_r y_{r0} = 1$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + v_0 \leq 0 \ , j=1,...,n$$

$$u_r, v_i \geq \varepsilon \ ; \ v_o \quad \text{variable libre}$$

donde v_0 es un indicador de retornos de escala.

Cada uno de los modelos planteados anteriormente busca determinar la eficiencia de una unidad de decisión con respecto a la frontera de producción determinada por las mejores observaciones. De esta forma, los valores de eficiencia, o valores proyectados, dependerán de la forma de la frontera y del sistema de evaluación implícito en cada modelo DEA.

Dados los valores para cada DMU, con el modelo se medirá la eficiencia para cada DMU, por lo tanto se tendrán que realizar *n* optimizaciones. La optimización produce un conjunto de valores positivos, o nulos, denominados u^* y v^*, que generarán el óptimo h^*, sólo si la DMU evaluada resulta eficiente.

Para estudiar las ineficiencias y sus causas de las distintas DMU´s, a modo de ejemplo la unidad de decisión, DMU_1, tendrá que ser comparada con una unidad virtual, o unidad de referencia $DMU_{1'}$, que es combinación lineal de DMU_2 y DMU_3.

A continuación se presenta un ejemplo del modelo de retornos de escala variable con orientación de salida, modelo BCC-Output.

Ejemplo ilustrativo

En la tabla A.V.3 se detalla la relación de input empleado y output obtenido por cinco empresas de un determinado sector, las cuales actúan en diferentes dimensiones. Suponer que se desea conocer hasta qué punto cada una de las empresas consideradas está produciendo de forma eficiente a partir de los recursos que emplea.

Tabla.A.V.3. Datos problema BCC-Output.

Empresa	Input	Output
1	30	2
2	25	6
3	30	10
4	40	8
5	40	12

Elaboración. Propia.

A partir de los datos se podría evaluar la eficiencia de la empresa 1 mediante el siguiente problema de programación lineal.

$$\text{Max } t_0 + \varepsilon\left[s_1 + z_1\right]$$

s.a:

$$t_0 + 2\lambda_1 + 6\lambda_2 + 10\lambda_3 + 8\lambda_4 + 12\lambda_5 + s_1 \geq 0$$

$$30\lambda_1 - 25\lambda_2 - 30\lambda_3 - 40\lambda_4 - 5\lambda_5 + z_1 \geq 0$$

$$\sum_{j=1}^{5} \lambda_j = 1$$

$$\lambda_j, z_1, s_1 \geq 0;\ j=1,\ 2,\ 3,\ 4,\ 5$$

$$t_0 \text{ libre}$$

donde se obtiene que la empresa 1 no es eficiente, ya que $t_0^* = 5$, es decir puede producir cinco veces más de lo que hasta ahora ha estado produciendo con un nivel de input de 2 unidades. Además todos los multiplicadores resultan nulos, $\lambda_1, \lambda_2, \lambda_4, \lambda_5 = 0$, a excepción de $\lambda_3 = 1$, indicando que la empresa 1 se debe proyectar sobre la empresa eficiente 3 para actuar con eficiencia.

En definitiva, si la empresa 1 desea llegar a alcanzar niveles de eficiencia debería utilizar y obtener los siguientes niveles de factor y producto:

$$x_1^* = \lambda_3 \cdot x_3 = 1 \cdot 30 = 30 \text{ unidades de input}$$

$$y_1^* = \lambda_3 \cdot y_3 = 1 \cdot 10 = 10 \text{ unidades de output}$$

Siendo el incremento de output:

$$\Delta y_1 = \lambda_3 y_3 - y_{10} = 10 - 2 = 8$$

y la mejora de input:

$$\Delta x_1 = z_1^* = 0$$

Así mismo, la empresa 4 tendrá que aumentar su nivel de output producido en 4 unidades, respecto al nivel actual.

En la tabla A.V.4 se detallan los índices de eficiencia obtenido para cada una de las empresas restantes. Así como el conjunto de referencia sobre el cual la

empresa en cuestión se proyecta, reflejándose entre paréntesis el valor de coeficiente multiplicador.

Tabla.A.V.4. Indices de eficiencia. Modelo BCC-Output.

Empresa	Eficiencia(%)	Conjunto de Referencia
1	500,00%	3(1,00)
2	100,00%	0
3	100,00%	1
4	150,00%	5(1,00)
5	100,00%	1

Elaboración. Propia.

En la figura A.V.5 se expresa la representación gráfica del problema.

BIBLIOTECA VIRTUAL

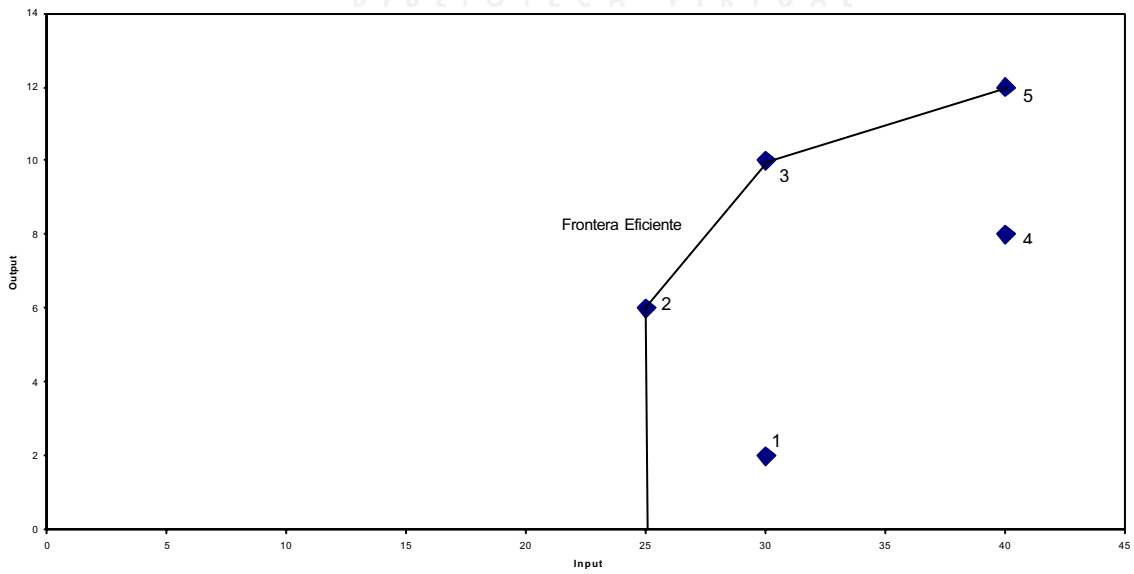


Figura.A.V.5. Situación geométrica de las empresas. BCC-Output.
Elaboración. Propia

Los resultados presentados por el programa informático EMS se exponen en la figura A.V.6.

EMS - [C:\Desarrollo\DEA\BCC-Output.xls_VRS_RAD_OUT]

File Edit DEA Window Help

	DMU	Score	Input(I)(W)	Output(O)(W)	Benchmarks	(S) Input(I)	(S) Output(O)
1	1	500,00%	0,033	0,500	3 (1,00)	0,00	0,00
2	2	100,00%	0,040	0,167	0		
3	3	100,00%	0,033	0,100	1		
4	4	150,00%	0,025	0,125	5 (1,00)	0,00	0,00
5	5	100,00%	0,025	0,083	1		

Figura.A.V.6. Pantalla EMS. Problema BCC-Output.
Elaboración. Propia.

ÍNDICE DE TABLAS

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

Capítulo III

Tabla.III.6.1. Modelos de demanda de energía eléctrica en la C.A.A. Sector residencial.....71
Tabla.III.6.2. Estimación del modelo 1 de demanda de energía eléctrica en la C.A.A.. 79
Tabla.III.6.3. Contrastes de diagnóstico para la evaluación de un modelo.....80
Tabla.III.6.4. Variables Irrelevantes. Modelo 1.....81
Tabla.III.6.5. Correlograma de los residuos. Modelo 1 82
Tabla.III.6.5. Correlograma de los residuos. Modelo 1 82
Tabla.III.6.6. Correlograma de los residuos al cuadrado. Modelo 1.....83
Tabla.III.6.7. Contraste de autocorrelación serial. Modelo 1 85
Tabla.III.6.8. Contraste comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 1.....86
Tabla.III.6.9. Contraste de heterocedasticidad de los residuos. Modelo 1.....87
Tabla.III.6.10. Contraste de Ramsey. Modelo 1 88
Tabla.III.6.11. Estimación del modelo 2 de demanda de energía eléctrica en la C.A.A 93
Tabla.III.6.12. Variables Irrelevantes. Modelo 2.....93
Tabla.III.6.13. Correlograma de los residuos. Modelo 2.....94
Tabla.III.6.14. Correlograma de los residuos al cuadrado. Modelo 2.....95
Tabla.III.6.15. Contraste de autocorrelación serial. Modelo 2.....96
Tabla.III.6.16. Contraste de comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 2.....97
Tabla.III.6.17. Contraste de heterocedasticidad de los residuos. Modelo 2.....98
Tabla.III.6.18. Contraste de Ramsey. Modelo 2 98
Tabla.III.6.19. Estimación del modelo 3 de demanda de energía eléctrica en la C.A.A 103
Tabla.III.6.20. Variables Irrelevantes. Modelo 3.....103
Tabla.III.6.21. Correlograma de los residuos. Modelo 3.....104
Tabla.III.6.22. Correlograma de los residuos al cuadrado. Modelo 3.....105
Tabla.III.6.23. Contraste de autocorrelación serial. Modelo 3.....106
Tabla.III.6.24. Contraste de comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 3.....107
Tabla.III.6.25. Contraste de heterocedasticidad de los residuos. Modelo 3.....108
Tabla.III.6.26. Contraste de Ramsey. Modelo 3 109
Tabla.III.6.27. Estimación del modelo 3 cruzado 3 variables de demanda de energía eléctrica en la C.A.A 110
Tabla.III.6.28. Variables Irrelevantes. Modelo 3 cruzado 3 variables 110
Tabla.III.6.29. Correlograma de los residuos. Modelo 3PC.....111
Tabla.III.6.30. Correlograma de los residuos al cuadrado. Modelo 3PC 112
Tabla.III.6.31. Contraste de autocorrelacion serial. Modelo 3PC.....113
Tabla.III.6.32. Contraste comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 3PC.....114

Tabla.III.6.33. Contraste de heterocedasticidad de los residuos. Modelo 3PC.....114
Tabla.III.6.34. Contraste de Ramsey. Modelo 3PC.....115
Tabla.III.6.35. Estimación del modelo 4 de demanda de energía eléctrica en la C.A.A 120
Tabla.III.6.36. Variables Irrelevantes. Modelo 4.....120
Tabla.III.6.37. Correlograma de los residuos. Modelo 4.....121
Tabla.III.6.38. Correlograma de los residuos al cuadrado. Modelo 4.....122
Tabla.III.6.39. Contraste de autocorrelación serial. Modelo 4.....123
Tabla.III.6.40. Contraste de comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 4.....124
Tabla.III.6.41. Contraste de heterocedasticidad de los residuos. Modelo 4.....125
Tabla.III.6.42. Contraste de Ramsey. Modelo 4 125
Tabla.III.6.43. Estimación del modelo 3R de demanda de energía eléctrica en la C.A.A.....130
Tabla. III.6.44. Contraste de Ramsey. Modelo 3R 131
Tabla.III.6.45. Estimación del modelo 3PCR de demanda de energía eléctrica en la C.A.A.....132
Tabla.III.6.46. Correlograma de los residuos. Modelo 3PCR.....132
Tabla.III.6.47. Correlograma de los residuos al cuadrado. Modelo 3PCR.....133
Tabla.III.6.48. Contraste de autocorrelación serial. Modelo 3PCR.....134
Tabla.III.6.49. Contraste de comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 3PCR.....135
Tabla.III.6.50. Contraste de heterocedasticidad de los residuos. Modelo 3PCR 136
Tabla.III.6.51. Contraste de Ramsey. Modelo 3PCR.....136
Tabla.III.6.52. Estadísticos evaluadores del los principales modelos analizados 141
Tabla.III.6.53. Índice de Theil y sus componentes. 144
Tabla.III.6.54. Estimaciones confidenciales, puntuales y errores de previsión 146
Tabla.III.6.55. Estimaciones puntuales y errores de previsión expresado en GWh.....147
Tabla.III.6.56. Contrastes de diagnósticos aplicados a los principales modelos analizados 148
Tabla.III.6.57. Estimación del modelo 3PCR. Serie histórica (1984-2002) 149
Tabla.III.6.58. Estimación puntual y confidencial de la demanda de electricidad en la C.A.A. para el año 2003 150
Tabla.III.7.1. Análisis de la varianza del modelo 3PC.....159

Capítulo IV

Tabla.IV.2.1.Resultados de la variable transformada 201
Tabla.IV.2.2. Media y desviación típica de las diferencias estacionales y cocientes estacionales 205
Tabla.IV.2.3. Índices Estacionales. Modelo de Alisado Exponencial 208
Tabla.IV.2.4. Índices Estacionales. Modelo de Holt-Winters 209
Tabla.IV.2.5. Intervalos confidenciales (95%) del modelo de A.E.....211

Tabla.IV.2.6. Intervalos confidenciales (95%) del modelo de Holt-Winters 211
Tabla.IV.2.7. Correlograma de los residuos. QMENSUAL.....214
Tabla.IV.2.8. Correlograma de los residuos. LQ.....218
Tabla.IV.2.9. Correlograma de los residuos. DLQ 220
Tabla.IV.2.10. Estimación de los coeficientes del modelo BJ-1.....222
Tabla.IV.2.11. Correlograma de los residuos. DDLQ 222
Tabla.IV.2.12. Estimaciones de los coeficientes del modelo BJ-2 223
Tabla.IV.2.13. Estimaciones de los coeficientes del modelo BJ-3 224
Tabla.IV.2.14. Estimación de los coeficientes. Modelo BJ-4 224
Tabla.IV.2.15. Estimación de los coeficientes. Modelo BJ-5 224
Tabla.IV.2.16. Correlograma de los residuos. D12LQ 226
Tabla.IV.2.17. Test de Raíces Unitarias de nivel de $\nabla_{12}\, lnq$ 227
Tabla.IV.2.18. Test de las Raíces Unitarias de primera diferencia de $\nabla_{12}\, lnq$.....228
Tabla.IV.2.19. Intervalos confidenciales (95%)del modelo BJ-1.....232
Tabla.IV.2.20. Intervalos confidenciales (95%)del modelo BJ-2.....232
Tabla.IV.2.21. Intervalos confidenciales (95%)del modelo BJ-3.....232
Tabla.IV.2.22. Intervalos confidenciales (95%)del modelo BJ-4.....233
Tabla.IV.2.23. Intervalos confidenciales (95%) del modelo BJ-1 de los meses del año 2003. Serie Histórica 1995-2002 234
Tabla.IV.2.23(a). Estadísticos evaluadores de los modelos no dinámicos 235
Tabla.V.2.23(b). Estadísticos evaluadores de los modelos Box-Jenkins. Modelos Dinámicos 236
Tabla.IV.3.1. Denominación y localización de los puntos de recogida de temperaturas 250
Tabla.IV.3.2. Prueba de Causalidad de Granger (un retardo) 260
Tabla.IV.3.3. Prueba de Causalidad de Granger (dos retardos) 260
Tabla.IV.3.4. Resultados de la regresión lineal y cuadrática para cada uno de los años comprendidos en el periodo 1995-2001.....263
Tabla.IV.3.5. Resultados de la regresión para los modelos lineal y cuadrático durante el periodo de 1995 a 2001 264
Tabla.IV.4.1. Estimación del modelo 22.4. de demanda de energía eléctrica en la C.A.A de 1995 a 2001.....273
Tabla.IV.4.2. Correlograma de los residuos. Modelo 22.4.....274
Tabla.IV.4.3. Correlograma de los residuos al cuadrado. Modelo 22.4 275
Tabla.IV.4.4. Contraste de autocorrelación serial. Modelo 22.4.....277
Tabla.IV.4.5. Contraste de comportamiento heterocedástico autorregresivo condicional de los residuos (ARCH). Modelo 22.4.....278
Tabla.IV.4.6. Contraste de heterocedasticidad de los residuos. Modelo 22.4.....279
Tabla.IV.4.7. Contraste de Ramsey. Modelo 22.4.....280
Tabla.IV.4.8. Intervalos de confianza (95%) del modelo 22.4.....281

Tabla.IV.5.1. Predicciones de los métodos de previsión 290
Tabla.IV.5.2. Predicciones del método combinado..... 290
Tabla.IV.5.3. Porcentaje de error de previsión..... 291

Capítulo V

Tabla.V.4.1. Balance de Energía en el Sistema Eléctrico Peninsular (GWh) 319
Tabla.V.4.2. Coeficientes de Correlación de Pearson..... 324
Tabla.V.4.3. Modelos 1 input- 2outpus y 2 inputs- 1 output 327
Tabla.V.4.4. Indicadores de Eficiencia (expresado en valor absoluto). Modelo 1input-2outputs..... 330
Tabla.V.4.5. Indicadores de Eficiencia (expresado en valores absolutos). Modelo 2inputs-1output..... 331
Tabla.V.4.6. Principales debilidades y fortalezas de las principales compañías eléctricas españolas en el año 2001 344

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

Apéndice

Tabla.A.V.1. Datos problema CCR-Input..... 368
Tabla.A.V.2. Índices de eficiencia. Modelo CCR-Input..... 369
Tabla.A.V.3. Datos problema BCC-Output 375
Tabla.A.V.4. Índices de eficiencia. Modelo BCC-Output 377

Anexo

Tabla.A.IV.1. Comparación de tres formas de calcular la temperatura media mensual en C.A.A 401
Tabla.A.IV.2. Estimación del Modelo15.1 402
Tabla.A.IV.3. Estimación del Modelo15.2 402
Tabla.A.IV.4. Estimación del Modelo16.1 403
Tabla.A.IV.5. Estimación del Modelo16.2 403
Tabla.A.IV.6. Estimación del Modelo17.1 404
Tabla.A.IV.7. Estimación del Modelo17.2 404
Tabla.A.IV.8. Estimación del Modelo18.1 405
Tabla.A.IV.9. Estimación del Modelo18.2 405
Tabla.A.IV.10. Estimación del ModeloI18,5.1 406
Tabla.A.IV.11. Estimación del ModeloI18,5.2 406
Tabla.A.IV.12. Estimación del Modelo I18,5.3 407
Tabla.A.IV.13. Estimación del Modelo19.1..... 407
Tabla.A.IV.14. Estimación del Modelo19.2..... 408
Tabla.A.IV.15. Estimación del Model19.3 408
Tabla.A.IV.16. Estimación del Modelo19.4..... 409
Tabla.A.IV.17. Estimación del Modelo 19.5..... 409

Tabla.A.IV.18. Estimación del Modelo 19.6.....410
Tabla.A.IV.19. Estimación del Modelo 19.7.....410
Tabla.A.IV.20. Estimación del Modelo 20.1.....411
Tabla.A.IV.21. Estimación del Modelo 20.2.....411
Tabla.A.IV.22. Estimación del Modelo 20.3.....412
Tabla.A.IV.23. Estimación del Modelo 20.4.....412
Tabla.A.IV.24. Estimación del Modelo 20.5.....413
Tabla.A.IV.25. Estimación del Modelo 20.6.....413
Tabla.A.IV.26. Estimación del Modelo 20,5.1.....414
Tabla.A.IV.27. Estimación del Modelo 20,5.2.....414
Tabla.A.IV.28. Estimación del Modelo 20,5.3.....415
Tabla.A.IV.29. Estimación del Modelo 20,5.4.....415
Tabla.A.IV.30. Estimación del Modelo 20,5.5.....416
Tabla.A.IV.31. Estimación del Modelo 20,5.6.....416
Tabla.A.IV.32. Estimación del Modelo 20,5.7.....417
Tabla.A.IV.33. Estimación del Modelo 21.1.....417
Tabla.A.IV.34. Estimación del Modelo 21.2.....418
Tabla.A.IV.35. Estimación del Modelo.21.3.....418
Tabla.A.IV.36. Estimación del Modelo 21.4.....419
Tabla.A.IV.37. Estimación del Modelo 21.5.....419
Tabla.A.IV.38. Estimación del Modelo 21.6.....420
Tabla.A.IV.39. Estimación del Modelo 21.7.....420
Tabla.A.IV.40. Estimación del Modelo 21,5.1.....421
Tabla.A.IV.41. Estimación del Modelo 21,5.2.....421
Tabla.A.IV.42. Estimación del Modelo 21,5.3.....422
Tabla.A.IV.43. Estimación del Modelo 21,5.4.....422
Tabla.A.IV.44. Estimación del Modelo 21,5.5.....423
Tabla.A.IV.45. Estimación del Modelo 21,5.6.....423
Tabla.A.IV.46. Estimación del Modelo 21,5.7.....424
Tabla.A.IV.47. Estimación del Modelo 21,5.8.....424
Tabla.A.IV.48. Estimación del Modelo 22.1.....425
Tabla.A.IV.49. Estimación del Modelo 22.2.....425
Tabla.A.IV.50. Estimación del Modelo 22.3.....426
Tabla.A.IV.51. Estimación del Modelo 22.4.....426
Tabla.A.IV.52. Estimación del Modelo 22.5.....427
Tabla.A.IV.53. Estimación del Modelo 22.6.....427
Tabla.A.IV.54. Estimación del Modelo 22.7.....428
Tabla.A.IV.55. Estimación del Modelo 22,5.1.....428
Tabla.A.IV.56. Estimación del Modelo 22,5.2.....429

Tabla.A.IV.57. Estimación del Modelo 22,5.3.....429
Tabla.A.IV.58. Estimación del Modelo 22,5.4.....430
Tabla.A.IV.59. Estimación del Modelo 22,5.5.....430
Tabla.A.IV.60. Estimación del Modelo 22,5.6.....431
Tabla.A.IV.61. Estimación del Modelo 23.1.....431
Tabla.A.IV.62. Estimación del Modelo 23.2.....432
Tabla.A.IV.63. Estimación del Modelo 23.3.....432
Tabla.A.IV.64. Estimación del Modelo 23.4.....433
Tabla.A.IV.65. Estimación del Modelo 23.5.....433
Tabla.A.IV.66. Estimación del Modelo 23,5.1.....434
Tabla.A.IV.67. Estimación del Modelo 23,5.2.....434
Tabla.A.IV.68. Estimación del Modelo 23,5.3.....435
Tabla.A.IV.69. Estimación del Modelo 23,5.4.....435
Tabla.A.IV.70. Estimación del Modelo 23,5.5.....436
Tabla.A.IV.71. Estimación del Modelo 23,5.....436
Tabla.A.IV.72. Estimación del Modelo 24.1.....437
Tabla.A.IV.73. Estimación del Modelo 24.2.....437
Tabla.A.IV.74. Estimación del Modelo 24.3.....438
Tabla.A.IV.75. Estimación del Modelo 24.4.....438
Tabla.A.IV.76. Estimación del Modelo 24.5.....439
Tabla.A.IV.77. Estimación del Modelo 25.1.....439
Tabla.A.IV.78. Estimación del Modelo 25.2.....440
Tabla.A.IV.79. Estimación del Modelo 25.3.....440
Tabla.A.IV.80. Estimación del Modelo 25.4.....441
Tabla.A.IV.81. Estadísticos de previsión de los modelos mensuales de temperaturas 442
Tabla.A.IV.82. Contraste de igualdad de pendientes. Primer Semestre.....444
Tabla.A.IV.83. Contraste de igualdad de pendientes. Primer Semestre.....444
Tabla.A.IV.84. Estimación mínimo cuadrática de los coeficientes de regresión
Primer Semestre.....445
Tabla.A.IV.85. Contraste de igualdad de pendientes. Primer Semestre.....445
Tabla.A.IV.86. Contraste de igualdad de pendientes. Segundo Semestre 446
Tabla.A.IV.87. Contraste de igualdad de pendientes. Segundo Semestre 446
Tabla.A.IV.88.Estimación mínimo cuadrática de los coeficientes de regresión.
Segundo Semestre.....447
Tabla.A.IV.89. Contraste de igualdad de pendientes. Segundo Semestre 447
Tabla.A.IV.90. Correlaciones bivariadas de Pearson. Demanda-Grados equivalentes 448

ÍNDICE DE FIGURAS

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

Capítulo III

Figura.III.6.1. Histograma de probabilidad. Modelo 1 84
Figura.III.6.2. Residuos recursivos. Modelo 1 89
Figura.III.6.3. Suma acumulada de los residuos recursivos. Modelo 1 90
Figura.III.6.4. Gráficos de los coeficientes recursivos. Modelo 1 91
Figura.III.6.5. Histograma de probabilidad. Modelo 2 95
Figura.III.6.6. Residuos Recursivos. Modelo 2 99
Figura.III.6.7. Suma acumulada de los residuos recursivos. Modelo 2 100
Figura.III.6.8. Gráficos de los coeficientes recursivos. Modelo 2 101
Figura.III.6.9. Histograma de probabilidad. Modelo 3 106
Figura.III.6.10. Histograma de probabilidad. Modelo 3PC 112
Figura.III.6.11. Residuos recursivos. Modelo 3PC 116
Figura.III.6.12. Suma acumulada de los residuos recursivos. Modelo 3PC 117
Figura.III.6.13. Gráficos de los coeficientes recursivos. Modelo 3PC 118
Figura.III.6.14. Histograma de probabilidad. Modelo 4 123
Figura.III.6.15. Residuos recursivos. Modelo 4 126
Figura.III.6.16. Suma acumulada de los residuos recursivos. Modelo 4 127
Figura.III.6.17. Gráficos de los coeficientes recursivos. Modelo 4 128
Figura.III.6.18. Histograma de probabilidad. Modelo 3PCR 134
Figura.III.6.19. Residuos recursivos. Modelo 3PCR 137
Figura.III.6.20. Suma acumulada de los residuos recursivos. Modelo 3PCR 138
Figura.III.6.21. Gráfico de los coeficientes recursivos. Modelo 3PCR 139
Figura.III.6.22. Comparación de los estadísticos de los modelos considerados 142
Figura.III.6.23. Preorden total. Método Promethee 143
Figura.III.7.1. Cuadro de diálogo inicial. Programa BSTP 161
Figura.III.7.2. Opción de semilla 162
Figura.III.7.3. Selección de semilla 163
Figura.III.7.4. Cálculos de Q previos al comienzo de la Tabla *X* 164
Figura.III.7.5. Cálculo de las celdas de la Tabla *X* 166
Figura.III.7.6. Comparación de las predicciones convencional y bootstrap. Tabla *S* 168
Figura.III.7.7. Comparación de las predicciones convencional y bootstrap. Tabla *W* 173

Capítulo IV

Figura.IV.2.1. Procedimiento Box-Jenkins a través de un diagrama de flujo.....198
Figura.IV.2.2. Evolución de la demanda de electricidad en la C.A. A. durante el periodo1995-2000.....203
Figura.IV.2.3. Gráfico desviación típica-media. Demanda mensual de energía eléctrica durante el periodo 1995-2000.....204
Figura.IV.2.4. Diferencias estacionales de la demanda mensual de energía eléctrica en C.A.A. durante el periodo 1995-2000.....206
Figura.IV.2.5. Cocientes estacionales de la demanda mensual de energía eléctrica en C.A.A. durante el periodo 1995-2000.....206
Figura.IV.2.6. Valores reales, ajustados y límites de confianza de demanda mensual de electricidad en la C.A.A. Modelo A. E. Periodo 1995-2000.....212
Figura.IV.2.7. Evolución demanda mensual de electricidad en CA.A. durante el periodo 1995-2000.....215
Figura.IV.2.8. Gráfico desviación típica-media. Demanda mensual de energía eléctrica durante el periodo 1995-2000.....216
Figura.IV.2.9. Histograma de probabilidad. QMENSUAL.....216
Figura.IV.2.10. Histograma de probabilidad. LQ.....217
Figura.IV.2.11. Evolución de la demanda mensual de electricidad en la C.A.A durante el periodo 1995-2000.....219
Figura.IV.2.12. Histograma de probabilidad. DLQ.....219
Figura.IV.2.13. Evolución de la demanda mensual de electricidad en la C.A.A. durante el periodo 1995-2000.....221
Figura.IV.2.14. Evolución de la demanda mensual de electricidad en C.A.A. durante el periodo 1995-2000.....225
Figura.IV.2.15. Histograma de probabilidad. D12LQ.....225
Figura.IV.2.16. Función de autocorrelación del error. Modelo BJ-1.....229
Figura.IV.2.17. Función de autocorrelación del error. Modelo BJ-2.....229
Figura.IV.2.18. Función de autocorrelación del error. Modelo BJ-3.....230
Figura.IV.2.19. Función de autocorrelación del error. Modelo BJ-4.....230
Figura.IV.2.20. Valores reales, ajustados y límites de confianza (99%) en la demanda mensual de electricidad en la C.A.A. durante el periodo 1995-2001. Modelo BJ-1.....234
Figura.IV.3.1. Consumo mensual de electricidad en C.A.A. durante el periodo 1995-2001.....253
Figura IV.3.2. Consumo mensual de electricidad y grados equivalente (22ºC) en C.A.A. durante el periodo 1995-2001.....254
Figura.IV.3.3. Consumo mensual de electricidad y grados equivalentes retardados en C.A.A. durante el periodo 1995-2001.....255

Figura.IV.3.4. Consumo mensual de electricidad trasladado y grados equivalentes (22ºC) en C.A.A. durante el periodo 1995-2001 256
Figura.IV.3.5. Consumo mensual de electricidad trasladado y temperatura bruta en C.A.A. durante el periodo 1995-2001 256
Figura.IV.3.6. Diagrama de dispersión. Consumo de electricidad versus grados equivalentes mensuales 22ºC 257
Figura. IV.3.7. Diagrama de dispersión. Consumo de electricidad versus grados equivalentes mensuales 22ºC(-1) 258
Figura.IV.3.8. Promedio mensual de grados equivalentes (22ºC) por año durante el periodo 1995-2001 259
Figura.IV.3.9. Patrón de consumo de electricidad durante los seis primeros meses durante el periodo 1995-2001 266
Figura.IV.3.10. Patrón de consumo de electricidad durante los seis últimos meses a lo durante el periodo 1995-2001 266
Figura.IV.3.11. Variación del consumo de electricidad y los correspondientes grados equivalentes (22ºC) 267
Figura.IV.3.12. Media mensual de los grados equivalentes (22ºC), 1995-2001 268
Figura.IV.4.1. Histograma de probabilidad. Modelo 22.4 276
Figura.IV.4.2. Valores reales, ajustados y límites de confianza de demanda mensual de electricidad en la C.A.A. Modelo 22.4. Periodo 1995-2000 281
Figura.IV.6.1. El consumidor ante la apertura del sector eléctrico y gasista 295

Capítulo V

Figura.V.3.1. Formas alternativas de medición de la eficiencia 310
Figura.V.3.2. Función de Producción Estimada 315
Figura.V.4.1. Estructura de la investigación 321
Figura.V.4.2. Evolución de los niveles de Eficiencia 1998-2001. Endesa 333
Figura.V.4.3. Evolución de los niveles de eficiencia 1998-2001. Hidrocantábrico 334
Figura.V.4.4. Evolución de los niveles de eficiencia 1998-2001. Iberdrola. 335
Figura.V.4.5. Evolución de los niveles de eficiencia 1998-2001. Unión Fenosa 336
Figura.V.4.6. Evolución de los niveles de eficiencia 1998-2001. Endesa 337
Figura.V.4.7. Evolución de los niveles de eficiencia 1998-2001. Hidrocantábrico 338
Figura.V.4.8. Evolución de los niveles de eficiencia 1998-2001. Iberdrola 339
Figura.V.4.9. Evolución de los niveles de eficiencia 1998-2001. Unión Fenosa 340

Apéndice

Figura.A.V.1. Flujos de Entrada y Salida de la *j*-ésima Unidad de Decisión 361
Figura.A.V.2. Una medida de la eficiencia técnica: TE_{CRS}= OC/OB 362
Figura.A.V.3. Situación geométrica de las empresas. CCR-Input 370
Figura.A.V.4. Pantalla EMS. Problema CCR-Input.....371
Figura.A.V.5. Situación geométrica de las empresas. BCC-Output 377
Figura.A.V.6. Pantalla EMS. Problema BCC-Output 378

Anexo

Figura.A.III.1. Índice de Theil. Modelo 1.(1999) 395
Figura.A.III.2. Índice de Theil. Modelo 1.(1998-1999) 395
Figura.A.III.3. Índice de Theil. Modelo 2.(1999) 396
Figura.A.III.4. Índice de Theil. Modelo 2.(1998-1999) 396
Figura.A.III.5. Índice de Theil. Modelo 3.(1999) 397
Figura.A.III.6. Índice de Theil. Modelo 3.(1998-1999) 397
Figura.A.III.7. Índice de Theil. Modelo 3PC.(1999) 398
Figura.A.III.8. Índice de Theil. Modelo 3PC.(1998-1999) 398
Figura.A.III.9. Índice de Theil. Modelo 3PCR.(1999) 399
Figura.A.III.10. Índice de Theil. Modelo 3PCR.(1998-1999) 399
Figura.A.III.11. Índice de Theil. Modelo 4.(1999) 400
Figura.A.III.12. Índice de Theil. Modelo 4.(1998-1999) 400

ANEXO

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

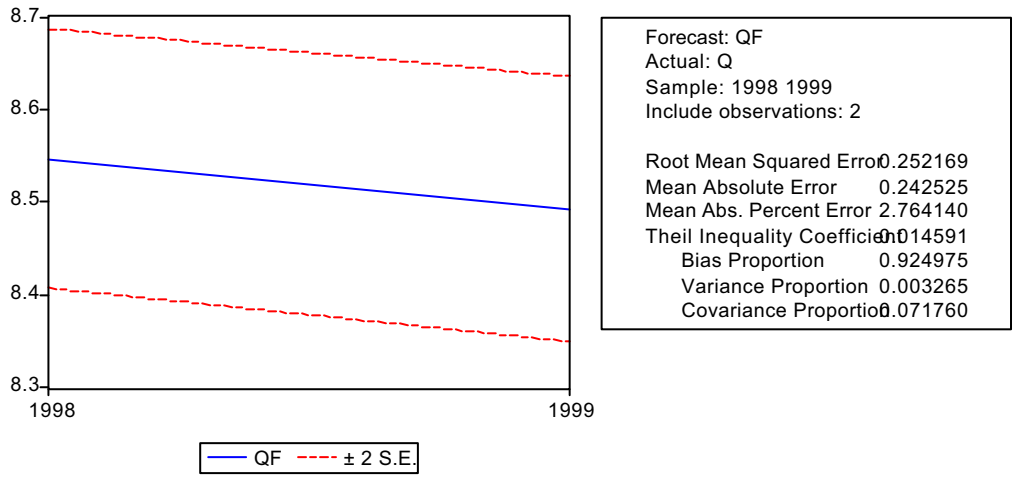


Figura.A.III.1. Índice de Theil. Modelo 1.(1999)
Elaboración. Propia.
Fuente. IEA.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

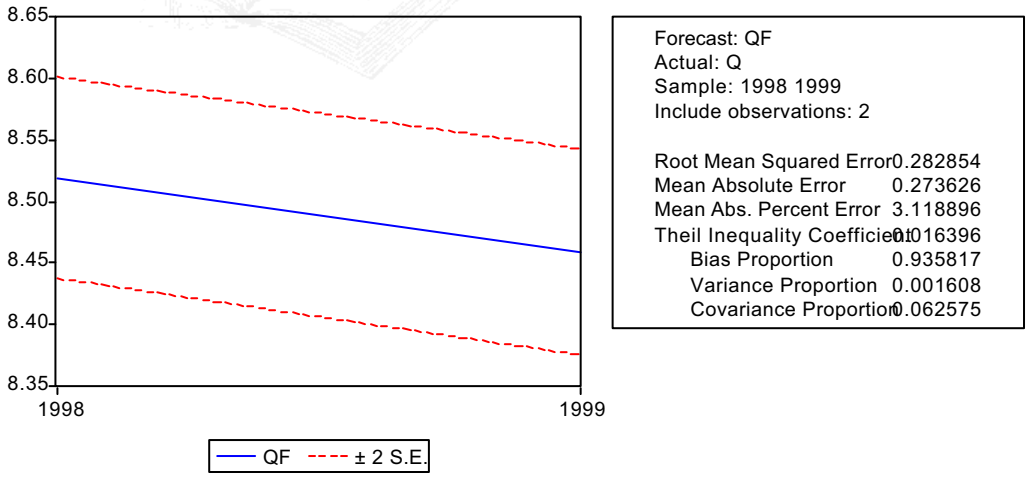


Figura.A.III.2. Índice de Theil. Modelo 1.(1998-1999)
Elaboración. Propia.
Fuente. IEA.

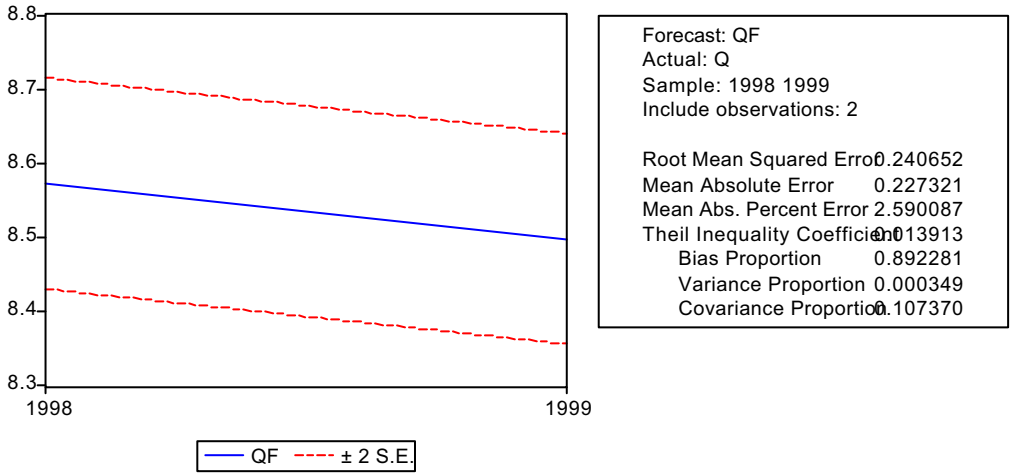


Figura.A.III.3. Índice de Theil. Modelo 2.(1999)
Elaboración. Propia.
Fuente. IEA.

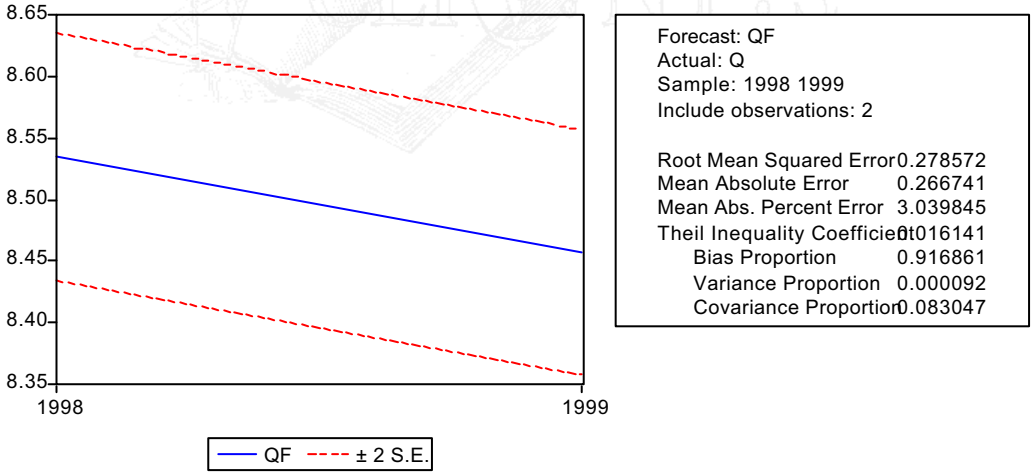


Figura.A.III.4. Índice de Theil. Modelo 2.(1998-1999)
Elaboración. Propia.
Fuente. IEA.

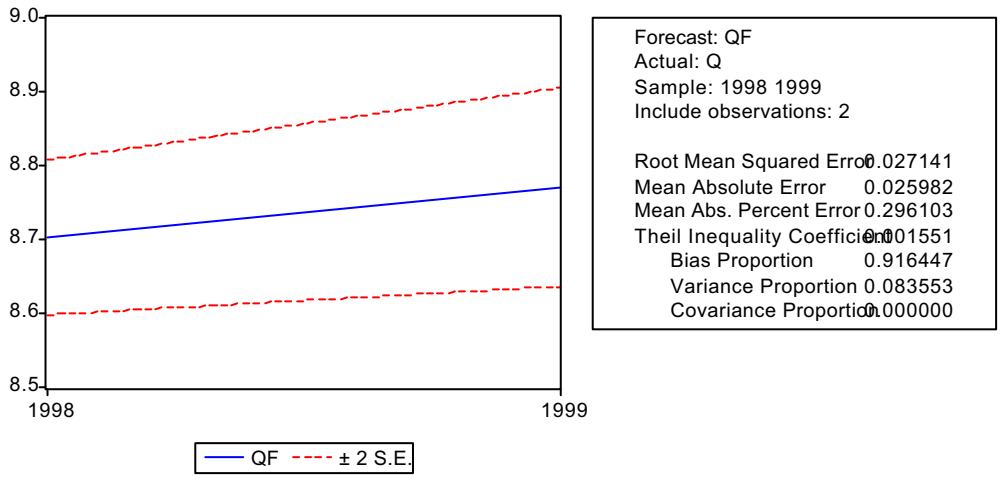


Figura.A.III.5. Índice de Theil. Modelo 3.(1999)
Elaboración. Propia.
Fuente. IEA.

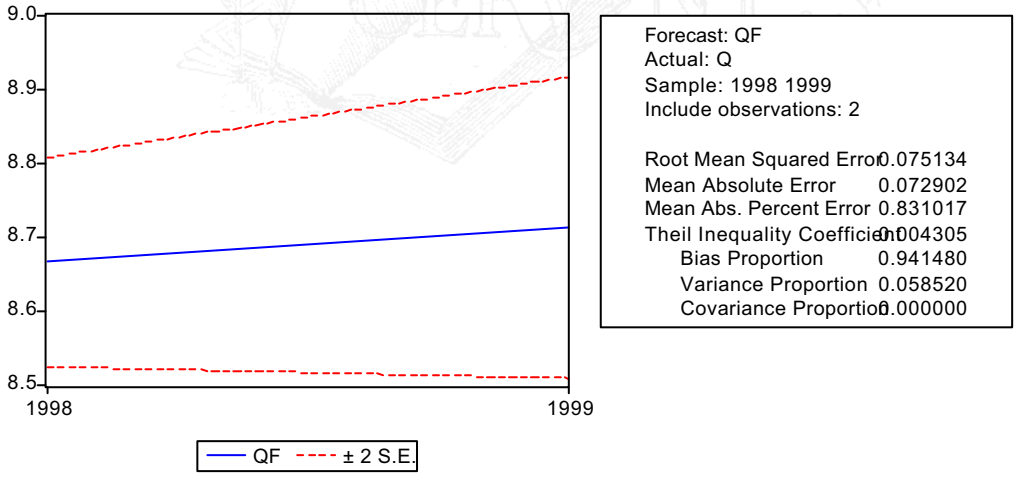


Figura.A.III.6. Índice de Theil. Modelo 3.(1998-1999)
Elaboración. Propia.
Fuente. IEA.

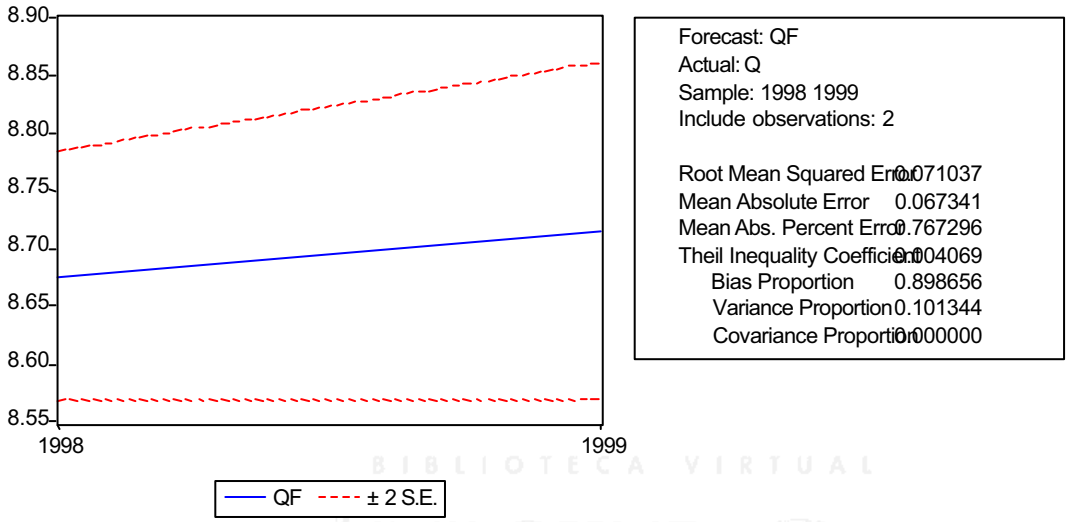


Figura.A.III.7. Índice de Theil. Modelo 3PC.(1999)
Elaboración. Propia.
Fuente. IEA.

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

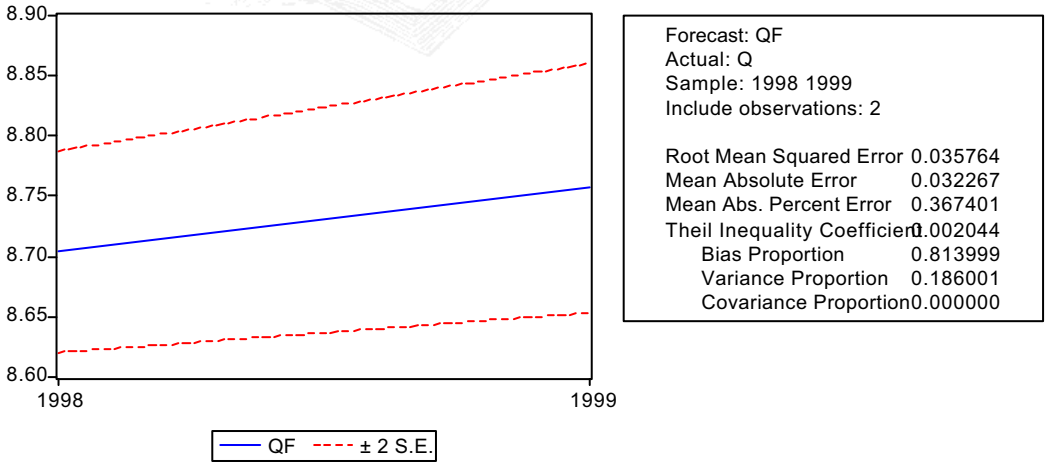


Figura.A.III.8. Índice de Theil. Modelo 3PC.(1998-1999)
Elaboración. Propia.
Fuente. IEA.

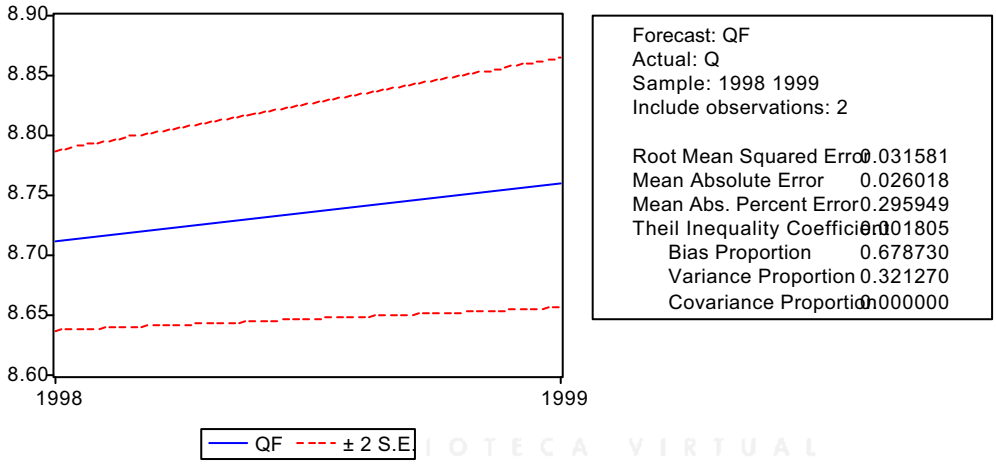


Figura.A.III.9. Índice de Theil. Modelo 3PCR.(1999)
Elaboración. Propia.
Fuente. IEA.

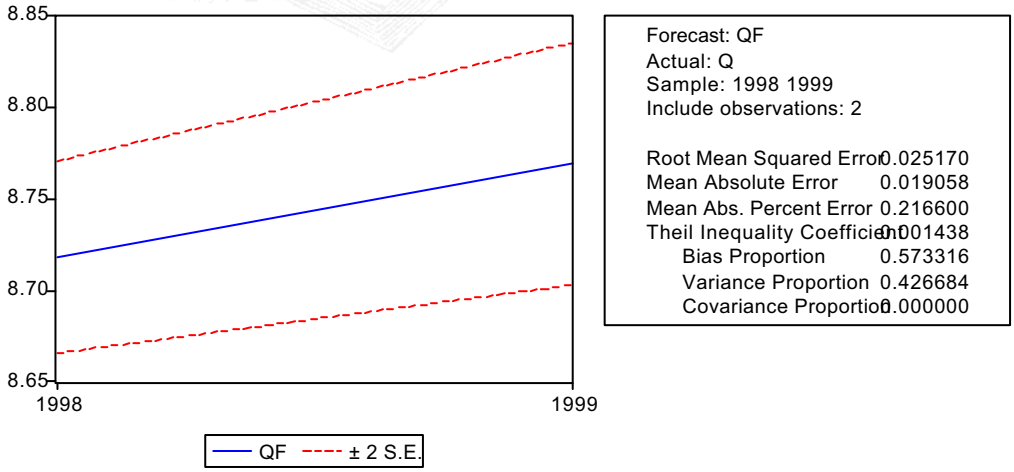
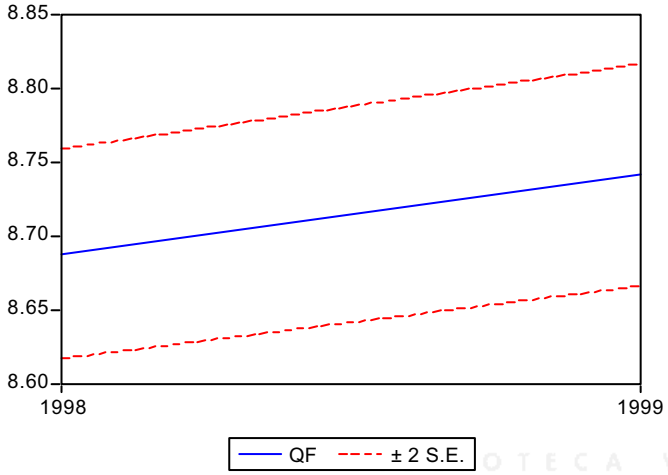


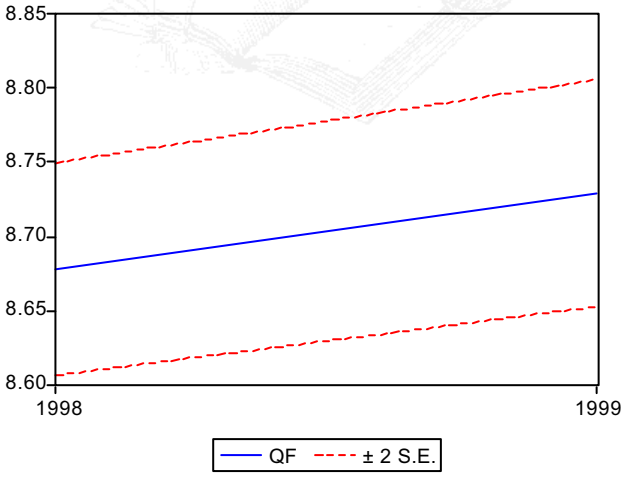
Figura.A.III.10. Índice de Theil. Modelo 3PCR.(1998-1999)
Elaboración. Propia.
Fuente. IEA.



Forecast: QF
Actual: Q
Sample: 1998 1999
Include observations: 2

Root Mean Squared Error	0.049950
Mean Absolute Error	0.047719
Mean Abs. Percent Error	0.543797
Theil Inequality Coefficient	0.002858
Bias Proportion	0.912650
Variance Proportion	0.087350
Covariance Proportion	0.000000

Figura.A.III.11. Índice de Theil. Modelo 4.(1999)
Elaboración. Propia.
Fuente. IEA.



Forecast: QF
Actual: Q
Sample: 1998 1999
Include observations: 2

Root Mean Squared Error	0.061026
Mean Absolute Error	0.058917
Mean Abs. Percent Error	0.671532
Theil Inequality Coefficient	0.003494
Bias Proportion	0.932079
Variance Proportion	0.067921
Covariance Proportion	0.000000

Figura.A.III.12. Índice de Theil. Modelo 4.(1998-1999)
Elaboración. Propia.
Fuente. IEA.

Tabla.A.IV.1. Comparación de tres formas de calcular la temperatura media mensual en C.A.A.

Mes- Año	Población	I. B. I.	(Pob+I. B. I.)/2	Mes- Año	Población	I. B. I.	(Pob+I. B. I.)/2
ene-95	10,792	10,776	10,785	ene-99	10,437	10,399	10,430
feb-95	13,472	13,415	13,443	feb-99	10,877	10,835	10,868
mar-95	14,856	14,748	14,801	mar-99	14,266	14,157	14,226
abr-95	16,835	16,683	16,661	abr-99	17,446	17,334	17,397
may-95	21,593	21,480	21,534	may-99	20,720	20,684	20,712
jun-95	22,076	23,070	22,995	jun-99	24,072	24,012	24,065
jul-95	27,022	26,945	26,980	jul-99	26,429	26,334	26,343
ago-95	27,096	27,043	27,066	ago-99	26,632	26,622	26,626
sep-95	22,116	21,742	21,792	sep-99	22,898	22,881	22,892
oct-95	21,019	20,877	20,947	oct-99	19,706	19,676	19,692
nov-95	16,702	16,530	16,636	nov-99	13,081	12,998	13,054
dic-95	13,554	13,462	13,509	dic-99	11,955	11,901	11,942
ene-96	12,556	12,476	12,517	ene-00	9,778	9,697	9,749
feb-96	10,797	10,724	10,690	feb-00	14,200	14,160	14,195
mar-96	13,963	13,838	13,900	mar-00	15,096	14,962	15,045
abr-96	16,865	16,719	16,791	abr-00	14,839	14,775	14,822
may-96	18,692	18,541	18,616	may-00	20,082	20,022	20,072
jun-96	24,241	24,082	24,159	jun-00	24,416	24,370	24,406
jul-96	26,194	26,089	26,139	jul-00	26,232	26,217	26,230
ago-96	25,239	25,183	25,210	ago-00	26,634	26,584	26,590
sep-96	21,596	21,514	21,554	sep-00	23,739	23,657	23,722
oct-96	18,620	18,488	18,554	oct-00	18,426	18,335	18,400
nov-96	14,518	14,401	14,460	nov-00	13,712	13,631	13,686
dic-96	12,359	12,246	12,303	dic-00	13,142	13,051	13,110
ene-97	11,709	11,594	11,653	ene-01	11,818	11,733	11,776
feb-97	14,150	14,015	14,083	feb-01	12,746	12,622	12,684
mar-97	16,412	16,203	16,306	mar-01	15,847	15,774	15,809
abr-97	18,456	18,280	18,367	abr-01	17,507	17,396	17,450
may-97	19,591	19,509	19,549	may-01	19,382	19,278	19,328
jun-97	22,243	22,232	22,235	jun-01	25,110	25,052	25,078
jul-97	24,709	24,635	24,364	jul-01	25,351	25,366	25,356
ago-97	25,455	25,421	25,810	ago-01	26,824	26,822	26,820
sep-97	24,274	24,136	24,204	sep-01	23,580	23,544	23,560
oct-97	20,955	20,845	20,941	oct-01	20,435	20,354	20,394
nov-97	15,330	15,216	15,274	nov-01	13,323	13,239	13,282
dic-97	12,156	12,087	12,123	dic-01	11,381	11,279	11,341
ene-98	11,631	11,559	11,596				
feb-98	13,669	13,510	13,590				
mar-98	15,914	15,731	15,822				
abr-98	15,632	15,574	15,621				
may-98	18,532	18,440	18,484				
jun-98	23,711	23,621	23,664				
jul-98	26,153	26,565	26,426				
ago-98	26,935	26,856	26,893				
sep-98	23,640	23,588	23,613				
oct-98	18,301	18,158	18,229				
nov-98	15,186	15,092	15,140				
dic-98	10,554	10,458	10,507				

Tabla.A.IV.2. Estimación del Modelo15.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1996:12 2000:12
Included observations: 49 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	0.376803	0.134566	2.800134	0.0076
QMENSUAL(-11)	-0.290186	0.122463	-2.369590	0.0223
GEQMEN15(-1)	0.237019	0.102102	2.321410	0.0250
C	206.2221	104.2351	1.978432	0.0542
AR(12)	1.147018	0.061163	18.75338	0.0000

R-squared	0.877285	Mean dependent var	533.6300
Adjusted R-squared	0.866129	S.D. dependent var	93.78636
S.E. of regression	34.31495	Akaike info criterion	10.00549
Sum squared resid	51810.69	Schwarz criterion	10.19853
Log likelihood	-240.1345	F-statistic	78.63835
Durbin-Watson stat	2.330226	Prob(F-statistic)	0.000000

Inverted AR Roots	1.01	.88+.51i	.88 -.51i	.51 -.88i
	.51+.88i	-.00 -1.01i	-.00+1.01i	-.51 -.88i
	-.51+.88i	-.88+.51i	-.88 -.51i	-1.01

Estimated AR process is nonstationary

Tabla.A.IV.3. Estimación del Modelo15.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1996:02 2000:12
Included observations: 59 after adjusting endpoints
Convergence achieved after 9 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	0.360768	0.120388	2.996706	0.0041
GEQMEN15(-1)	0.221155	0.094417	2.342322	0.0228
C	30.30701	186.6219	0.162398	0.8716
AR(12)	1.094319	0.068273	16.02848	0.0000

R-squared	0.857324	Mean dependent var	521.6882
Adjusted R-squared	0.849542	S.D. dependent var	92.17880
S.E. of regression	35.75519	Akaike info criterion	10.05666
Sum squared resid	70313.85	Schwarz criterion	10.19751
Log likelihood	-292.6714	F-statistic	110.1630
Durbin-Watson stat	1.973329	Prob(F-statistic)	0.000000

Inverted AR Roots	1.01	.87+.50i	.87 -.50i	.50+.87i
	.50 -.87i	.00+1.01i	-.00 -1.01i	-.50 -.87i
	-.50+.87i	-.87 -.50i	-.87+.50i	-1.01

Estimated AR process is nonstationary

Tabla.A.IV.4. Estimación del Modelo16.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1996:12 2000:12
Included observations: 49 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	0.338241	0.126924	2.664915	0.0107
QMENSUAL(-11)	-0.279945	0.115086	-2.432479	0.0191
GEQMEN16(-1)	0.320358	0.096875	3.306932	0.0019
C	224.8067	90.69277	2.478772	0.0171
AR(12)	1.157763	0.060084	19.26912	0.0000

R-squared	0.889645	Mean dependent var	533.6300
Adjusted R-squared	0.879613	S.D. dependent var	93.78636
S.E. of regression	32.54090	Akaike info criterion	9.899324
Sum squared resid	46592.05	Schwarz criterion	10.09237
Log likelihood	-237.5334	F-statistic	88.67848
Durbin-Watson stat	2.466328	Prob(F-statistic)	0.000000

Inverted AR Roots	1.01	.88 -.51i	.88+.51i	.51+.88i
	.51 -.88i	.00 -1.01i	-.00+1.01i	-.51 -.88i
	-.51+.88i	-.88 -.51i	-.88+.51i	-1.01

Estimated AR process is nonstationary

Tabla.A.IV.5. Estimación del Modelo16.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1996:02 2000:12
Included observations: 59 after adjusting endpoints
Convergence achieved after 9 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	0.331065	0.117398	2.820027	0.0067
GEQMEN16(-1)	0.285357	0.093184	3.062285	0.0034
C	34.23008	184.1202	0.185912	0.8532
AR(12)	1.095469	0.067793	16.15913	0.0000

R-squared	0.865989	Mean dependent var	521.6882
Adjusted R-squared	0.858679	S.D. dependent var	92.17880
S.E. of regression	34.65251	Akaike info criterion	9.994006
Sum squared resid	66043.79	Schwarz criterion	10.13486
Log likelihood	-290.8232	F-statistic	118.4709
Durbin-Watson stat	2.049615	Prob(F-statistic)	0.000000

Inverted AR Roots	1.01	.87 -.50i	.87+.50i	.50+.87i
	.50 -.87i	.00+1.01i	-.00 -1.01i	-.50+.87i
	-.50 -.87i	-.87 -.50i	-.87+.50i	-1.01

Estimated AR process is nonstationary

Tabla.A.IV.6. Estimación del Modelo17.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1996:12 2000:12
Included observations: 49 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	0.324586	0.118494	2.739262	0.0089
QMENSUAL(-11)	-0.254480	0.107045	-2.377319	0.0219
GEQMEN17(-1)	0.372168	0.086847	4.285343	0.0001
C	250.8426	71.42706	3.511871	0.0010
AR(12)	1.185965	0.062258	19.04923	0.0000

R-squared	0.902334	Mean dependent var	533.6300
Adjusted R-squared	0.893455	S.D. dependent var	93.78636
S.E. of regression	30.61304	Akaike info criterion	9.777180
Sum squared resid	41234.97	Schwarz criterion	9.970223
Log likelihood	-234.5409	F-statistic	101.6283
Durbin-Watson stat	2.625280	Prob(F-statistic)	0.000000

Inverted AR Roots	1.01	.88+.51i	.88 -.51i	.51+.88i
	.51 -.88i	.00 -1.01i	-.00+1.01i	-.51 -.88i
	-.51+.88i	-.88 -.51i	-.88+.51i	-1.01

Estimated AR process is nonstationary

Tabla.A.IV.7. Estimación del Modelo17.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN17	-0.034655	0.045226	-0.766265	0.4497
QMENSUAL(-1)	-0.104846	0.041805	-2.508017	0.0180
QMENSUAL(-12)	1.238987	0.043590	28.42389	0.0000
GEQMEN17(-1)	0.027105	0.040137	0.675325	0.5048
C	-25.66355	24.91034	-1.030237	0.3114
AR(12)	-0.520938	0.137483	-3.789105	0.0007
AR(24)	-0.317141	0.114511	-2.769512	0.0097

R-squared	0.935400	Mean dependent var	555.0258
Adjusted R-squared	0.922034	S.D. dependent var	93.94524
S.E. of regression	26.23172	Akaike info criterion	9.544481
Sum squared resid	19954.99	Schwarz criterion	9.852388
Log likelihood	-164.8007	F-statistic	69.98583
Durbin-Watson stat	1.745262	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.16i	.94 -.16i	.89 -.33i	.89+.33i
	.73+.61i	.73 -.61i	.61 -.73i	.61+.73i
	.33+.89i	.33 -.89i	.16 -.94i	.16+.94i
	-.16+.94i	-.16 -.94i	-.33 -.89i	-.33+.89i
	-.61+.73i	-.61 -.73i	-.73+.61i	-.73 -.61i
	-.89+.33i	-.89 -.33i	-.94 -.16i	-.94+.16i

Tabla.A.IV.8. Estimación del Modelo18.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1996:12 2000:12
Included observations: 49 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	0.294882	0.112130	2.629826	0.0117
QMENSUAL(-11)	-0.225446	0.101142	-2.229000	0.0310
GEQMEN18(-1)	0.397893	0.079044	5.033787	0.0000
C	274.3182	61.19649	4.482581	0.0001
AR(12)	1.220352	0.067027	18.20685	0.0000

R-squared	0.911085	Mean dependent var	533.6300
Adjusted R-squared	0.903002	S.D. dependent var	93.78636
S.E. of regression	29.20935	Akaike info criterion	9.683306
Sum squared resid	37540.20	Schwarz criterion	9.876349
Log likelihood	-232.2410	F-statistic	112.7134
Durbin-Watson stat	2.630984	Prob(F-statistic)	0.000000

Inverted AR Roots	1.02	.88+.51i	.88 -.51i	.51 -.88i
	.51+.88i	.00+1.02i	-.00 -1.02i	-.51 -.88i
	-.51+.88i	-.88+.51i	-.88 -.51i	-1.02

Estimated AR process is nonstationary

Tabla.A.IV.9. Estimación del Modelo18.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN18(-1)	-0.003258	0.041714	-0.078106	0.9382
QMENSUAL(-12)	1.188992	0.047772	24.88880	0.0000
C	-57.13858	21.52978	-2.653932	0.0124
AR(12)	-0.460426	0.151019	-3.048793	0.0047
AR(24)	-0.281483	0.127414	-2.209204	0.0347

R-squared	0.919231	Mean dependent var	555.0258
Adjusted R-squared	0.908809	S.D. dependent var	93.94524
S.E. of regression	28.36947	Akaike info criterion	9.656750
Sum squared resid	24949.64	Schwarz criterion	9.876684
Log likelihood	-168.8215	F-statistic	88.20222
Durbin-Watson stat	2.169833	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.16i	.94 -.16i	.89+.33i	.89 -.33i
	.73+.61i	.73 -.61i	.61 -.73i	.61+.73i
	.33+.89i	.33 -.89i	.16 -.94i	.16+.94i
	-.16 -.94i	-.16+.94i	-.33+.89i	-.33 -.89i
	-.61+.73i	-.61 -.73i	-.73 -.61i	-.73+.61i
	-.89+.33i	-.89 -.33i	-.94+.16i	-.94 -.16i

Tabla.A.IV.10. Estimación del ModeloI18,5.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1996:02 2000:12
Included observations: 59 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQINT185(-1)	0.334530	0.078466	4.263371	0.0001
QMENSUAL(-1)	0.272697	0.112135	2.431869	0.0183
C	90.41086	147.7516	0.611911	0.5431
AR(12)	1.114197	0.074397	14.97642	0.0000

R-squared	0.881910	Mean dependent var	521.6882
Adjusted R-squared	0.875468	S.D. dependent var	92.17880
S.E. of regression	32.52902	Akaike info criterion	9.867531
Sum squared resid	58197.55	Schwarz criterion	10.00838
Log likelihood	-287.0922	F-statistic	136.9150
Durbin-Watson stat	2.182951	Prob(F-statistic)	0.000000

Inverted AR Roots	1.01	.87 -.50i	.87+.50i	.50 -.87i
	.50+.87i	-.00 -1.01i	-.00+1.01i	-.50 -.87i
	-.50+.87i	-.87+.50i	-.87 -.50i	-1.01

Estimated AR process is nonstationary

Tabla.A.IV.11. Estimación del ModeloI18,5.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQINT185	-0.063490	0.030871	-2.056637	0.0485
QMENSUAL(-1)	-0.126376	0.035939	-3.516400	0.0014
QMENSUAL(-12)	1.285076	0.042864	29.98011	0.0000
C	-29.55874	18.34592	-1.611189	0.1176
AR(12)	-0.560730	0.122898	-4.562558	0.0001
AR(24)	-0.349302	0.102538	-3.406549	0.0019

R-squared	0.941973	Mean dependent var	555.0258
Adjusted R-squared	0.932301	S.D. dependent var	93.94524
S.E. of regression	24.44356	Akaike info criterion	9.381622
Sum squared resid	17924.63	Schwarz criterion	9.645542
Log likelihood	-162.8692	F-statistic	97.39956
Durbin-Watson stat	1.715171	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.16i	.94 -.16i	.90 -.33i	.90+.33i
	.73+.61i	.73 -.61i	.61 -.73i	.61+.73i
	.33+.90i	.33 -.90i	.16+.94i	.16 -.94i
	-.16 -.94i	-.16+.94i	-.33 -.90i	-.33+.90i
	-.61+.73i	-.61 -.73i	-.73+.61i	-.73 -.61i
	-.90+.33i	-.90 -.33i	-.94+.16i	-.94 -.16i

Tabla.A.IV.12. Estimación del Modelo I18,5.3.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	-0.115907	0.029226	-3.965826	0.0004
QMENSUAL(-12)	1.329498	0.045025	29.52818	0.0000
QMENSUAL(-10)	-0.072184	0.025845	-2.792989	0.0092
GEQINT185(-1)	-0.079968	0.031456	-2.542208	0.0166
C	-18.64241	20.61106	-0.904486	0.3732
AR(12)	-0.615939	0.112388	-5.480481	0.0000
AR(24)	-0.401263	0.094473	-4.247402	0.0002

R-squared	0.950068	Mean dependent var	555.0258
Adjusted R-squared	0.939738	S.D. dependent var	93.94524
S.E. of regression	23.06206	Akaike info criterion	9.286920
Sum squared resid	15423.90	Schwarz criterion	9.594827
Log likelihood	-160.1646	F-statistic	91.96552
Durbin-Watson stat	1.642823	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.90 -.33i	.90+.33i
	.74+.62i	.74 -.62i	.62+.74i	.62 -.74i
	.33 -.90i	.33+.90i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.33+.90i	-.33 -.90i
	-.62 -.74i	-.62+.74i	-.74 -.62i	-.74+.62i
	-.90 -.33i	-.90+.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.13. Estimación del Modelo19.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN19(-1)	-0.054139	0.048617	-1.113581	0.2740
QMENSUAL(-12)	1.230316	0.054398	22.61685	0.0000
C	-70.26234	22.74417	-3.089247	0.0042
AR(12)	-0.514800	0.146379	-3.516892	0.0014
AR(24)	-0.314060	0.121494	-2.584973	0.0147

R-squared	0.922051	Mean dependent var	555.0258
Adjusted R-squared	0.911993	S.D. dependent var	93.94524
S.E. of regression	27.86978	Akaike info criterion	9.621209
Sum squared resid	24078.47	Schwarz criterion	9.841142
Log likelihood	-168.1818	F-statistic	91.67381
Durbin-Watson stat	2.003599	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.16i	.94 -.16i	.89+.33i	.89 -.33i
	.73+.61i	.73 -.61i	.61+.73i	.61 -.73i
	.33 -.89i	.33+.89i	.16+.94i	.16 -.94i
	-.16 -.94i	-.16+.94i	-.33+.89i	-.33 -.89i
	-.61 -.73i	-.61+.73i	-.73+.61i	-.73 -.61i
	-.89+.33i	-.89 -.33i	-.94+.16i	-.94 -.16i

Tabla.A.IV.14. Estimación del Modelo19.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN19(-1)	-0.042548	0.033267	-1.278976	0.2110
QMENSUAL(-1)	-0.096140	0.024491	-3.925573	0.0005
QMENSUAL(-12)	1.327206	0.040852	32.48809	0.0000
QMENSUAL(-11)	-0.091136	0.023755	-3.836550	0.0006
C	-23.61578	18.28177	-1.291767	0.2066
AR(12)	-0.655786	0.107773	-6.084889	0.0000
AR(24)	-0.435260	0.090164	-4.827402	0.0000

R-squared	0.958305	Mean dependent var	555.0258
Adjusted R-squared	0.949678	S.D. dependent var	93.94524
S.E. of regression	21.07425	Akaike info criterion	9.106647
Sum squared resid	12879.60	Schwarz criterion	9.414553
Log likelihood	-156.9196	F-statistic	111.0876
Durbin-Watson stat	2.020090	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91 -.33i	.91+.33i
	.74+.62i	.74 -.62i	.62+.74i	.62 -.74i
	.33+.91i	.33 -.91i	.17 -.95i	.17+.95i
	-.17+.95i	-.17 -.95i	-.33+.91i	-.33 -.91i
	-.62 -.74i	-.62+.74i	-.74 -.62i	-.74+.62i
	-.91 -.33i	-.91+.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.15. Estimación del Model19.3.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN19	-0.090937	0.034420	-2.641979	0.0130
QMENSUAL(-1)	-0.134846	0.034159	-3.947606	0.0004
QMENSUAL(-12)	1.304366	0.042569	30.64124	0.0000
C	-30.08511	17.22372	-1.746726	0.0909
AR(12)	-0.582739	0.118060	-4.935953	0.0000
AR(24)	-0.366057	0.097899	-3.739121	0.0008

R-squared	0.946057	Mean dependent var	555.0258
Adjusted R-squared	0.937066	S.D. dependent var	93.94524
S.E. of regression	23.56772	Akaike info criterion	9.308644
Sum squared resid	16663.12	Schwarz criterion	9.572564
Log likelihood	-161.5556	F-statistic	105.2276
Durbin-Watson stat	1.734494	Prob(F-statistic)	0.000000

Tabla.A.IV.16. Estimación del Modelo19.4.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN19(-1)	-0.123072	0.038604	-3.188069	0.0034
QMENSUAL(-1)	-0.115933	0.027246	-4.255089	0.0002
QMENSUAL(-10)	-0.080216	0.024668	-3.251836	0.0029
QMENSUAL(-12)	1.365982	0.047350	28.84843	0.0000
C	-25.69868	19.47012	-1.319904	0.1972
AR(12)	-0.653674	0.106638	-6.129860	0.0000
AR(24)	-0.425796	0.089034	-4.782374	0.0000

R-squared	0.954074	Mean dependent var	555.0258
Adjusted R-squared	0.944573	S.D. dependent var	93.94524
S.E. of regression	22.11757	Akaike info criterion	9.203287
Sum squared resid	14186.42	Schwarz criterion	9.511194
Log likelihood	-158.6592	F-statistic	100.4093
Durbin-Watson stat	1.624356	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91 -.33i	.91+.33i
	.74+.62i	.74 -.62i	.62+.74i	.62 -.74i
	.33+.91i	.33 -.91i	.17 -.95i	.17+.95i
	-.17+.95i	-.17 -.95i	-.33+.91i	-.33 -.91i
	-.62 -.74i	-.62+.74i	-.74 -.62i	-.74+.62i
	-.91 -.33i	-.91+.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.17. Estimación del Modelo 19.5.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN19	-0.081028	0.035809	-2.262759	0.0313
GEQMEN19(-1)	-0.034453	0.040414	-0.852501	0.4009
QMENSUAL(-1)	-0.129899	0.034218	-3.796171	0.0007
QMENSUAL(-12)	1.325449	0.048725	27.20245	0.0000
C	-39.61947	20.20132	-1.961231	0.0595
AR(12)	-0.609932	0.119506	-5.103780	0.0000
AR(24)	-0.381748	0.097808	-3.903036	0.0005

R-squared	0.947321	Mean dependent var	555.0258
Adjusted R-squared	0.936422	S.D. dependent var	93.94524
S.E. of regression	23.68791	Akaike info criterion	9.340472
Sum squared resid	16272.39	Schwarz criterion	9.648379
Log likelihood	-161.1285	F-statistic	86.91810
Durbin-Watson stat	1.651927	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.90 -.33i	.90+.33i
	.74+.62i	.74 -.62i	.62 -.74i	.62+.74i
	.33 -.90i	.33+.90i	.17+.95i	.17 -.95i
	-.17 -.95i	-.17+.95i	-.33 -.90i	-.33+.90i
	-.62 -.74i	-.62+.74i	-.74+.62i	-.74 -.62i
	-.90 -.33i	-.90+.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.18. Estimación del Modelo 19.6.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN19(-1)	-0.061086	0.041754	-1.463013	0.1539
QMENSUAL(-1)	-0.087498	0.030887	-2.832816	0.0082
QMENSUAL(-12)	1.290057	0.050550	25.52021	0.0000
C	-52.52163	20.86998	-2.516612	0.0174
AR(12)	-0.590109	0.129664	-4.551046	0.0001
AR(24)	-0.355319	0.106303	-3.342511	0.0022

R-squared	0.938008	Mean dependent var	555.0258
Adjusted R-squared	0.927675	S.D. dependent var	93.94524
S.E. of regression	25.26490	Akaike info criterion	9.447721
Sum squared resid	19149.46	Schwarz criterion	9.711641
Log likelihood	-164.0590	F-statistic	90.78593
Durbin-Watson stat	1.710814	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.33+.90i	.33 -.90i	.17 -.94i	.17+.94i
	-.17+.94i	-.17 -.94i	-.33+.90i	-.33 -.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90 -.33i	-.90+.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.19. Estimación del Modelo 19.7.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1996:12 2000:12
Included observations: 49 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN19(-1)	0.469087	0.077527	6.050607	0.0000
QMENSUAL(-1)	0.232400	0.104688	2.219926	0.0316
QMENSUAL(-11)	-0.183776	0.093875	-1.957676	0.0566
C	297.0459	53.24721	5.578619	0.0000
AR(12)	1.268076	0.073424	17.27049	0.0000

R-squared	0.921485	Mean dependent var	533.6300
Adjusted R-squared	0.914347	S.D. dependent var	93.78636
S.E. of regression	27.44798	Akaike info criterion	9.558913
Sum squared resid	33149.24	Schwarz criterion	9.751956
Log likelihood	-229.1934	F-statistic	129.1005
Durbin-Watson stat	2.586659	Prob(F-statistic)	0.000000

Inverted AR Roots	1.02	.88+.51i	.88 -.51i	.51+.88i
	.51 -.88i	.00 -1.02i	-.00+1.02i	-.51 -.88i
	-.51+.88i	-.88 -.51i	-.88+.51i	-1.02
	Estimated AR process is nonstationary			

Tabla.A.IV.20. Estimación del Modelo 20.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	-0.119042	0.026156	-4.551186	0.0001
QMENSUAL(-12)	1.407235	0.045849	30.69294	0.0000
GEQMEN20	-0.089990	0.027216	-3.306497	0.0025
GEQMEN20(-1)	-0.098351	0.037646	-2.612530	0.0141
C	-74.27226	18.30264	-4.058009	0.0003
AR(12)	-0.724770	0.100730	-7.195174	0.0000
AR(24)	-0.436472	0.080234	-5.439954	0.0000

R-squared	0.961244	Mean dependent var	555.0258
Adjusted R-squared	0.953225	S.D. dependent var	93.94524
S.E. of regression	20.31795	Akaike info criterion	9.033552
Sum squared resid	11971.76	Schwarz criterion	9.341459
Log likelihood	-155.6039	F-statistic	119.8782
Durbin-Watson stat	1.895968	Prob(F-statistic)	0.000000

Inverted AR Roots	.95 -.17i	.95+.17i	.91 -.33i	.91+.33i
	.74+.62i	.74 -.62i	.62 -.74i	.62+.74i
	.33+.91i	.33 -.91i	.17+.95i	.17 -.95i
	-.17 -.95i	-.17+.95i	-.33+.91i	-.33 -.91i
	-.62+.74i	-.62 -.74i	-.74 -.62i	-.74+.62i
	-.91 -.33i	-.91+.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.21. Estimación del Modelo 20.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN20(-1)	-0.207484	0.033591	-6.176716	0.0000
QMENSUAL(-12)	1.463476	0.041420	35.33248	0.0000
QMENSUAL(-1)	-0.105234	0.019904	-5.287097	0.0000
QMENSUAL(-10)	-0.091200	0.018015	-5.062465	0.0000
C	-60.66852	15.77481	-3.845911	0.0006
AR(12)	-0.775904	0.082840	-9.366329	0.0000
AR(24)	-0.489315	0.067308	-7.269756	0.0000

R-squared	0.970970	Mean dependent var	555.0258
Adjusted R-squared	0.964964	S.D. dependent var	93.94524
S.E. of regression	17.58467	Akaike info criterion	8.744598
Sum squared resid	8967.398	Schwarz criterion	9.052504
Log likelihood	-150.4028	F-statistic	161.6604
Durbin-Watson stat	1.967548	Prob(F-statistic)	0.000000

Inverted AR Roots	.95 -.17i	.95+.17i	.91+.33i	.91 -.33i
	.74+.63i	.74 -.63i	.63 -.74i	.63+.74i
	.33 -.91i	.33+.91i	.17+.95i	.17 -.95i
	-.17+.95i	-.17 -.95i	-.33+.91i	-.33 -.91i
	-.63 -.74i	-.63+.74i	-.74 -.63i	-.74+.63i
	-.91+.33i	-.91 -.33i	-.95 -.17i	-.95+.17i

Tabla.A.IV.22. Estimación del Modelo 20.3.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN20(-1)	-0.086716	0.038105	-2.275750	0.0304
QMENSUAL(-1)	-0.088875	0.022868	-3.886437	0.0005
QMENSUAL(-12)	1.366704	0.043425	31.47270	0.0000
QMENSUAL(-11)	-0.079365	0.022913	-3.463784	0.0017
C	-45.98929	20.83221	-2.207605	0.0353
AR(12)	-0.707909	0.102635	-6.897311	0.0000
AR(24)	-0.455109	0.083424	-5.455398	0.0000

R-squared	0.962196	Mean dependent var	555.0258
Adjusted R-squared	0.954375	S.D. dependent var	93.94524
S.E. of regression	20.06673	Akaike info criterion	9.008669
Sum squared resid	11677.54	Schwarz criterion	9.316576
Log likelihood	-155.1561	F-statistic	123.0203
Durbin-Watson stat	2.184771	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91 -.33i	.91+.33i
	.74+.62i	.74 -.62i	.62 -.74i	.62+.74i
	.33 -.91i	.33+.91i	.17+.95i	.17 -.95i
	-.17 -.95i	-.17+.95i	-.33+.91i	-.33 -.91i
	-.62 -.74i	-.62+.74i	-.74 -.62i	-.74+.62i
	-.91 -.33i	-.91+.33i	-.95 -.17i	-.95+.17i

Tabla.A.IV.23. Estimación del Modelo 20.4.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN20	-0.109201	0.030379	-3.594582	0.0011
QMENSUAL(-1)	-0.133635	0.029379	-4.548610	0.0001
QMENSUAL(-12)	1.330788	0.040487	32.86917	0.0000
C	-40.63248	15.35326	-2.646505	0.0128
AR(12)	-0.626766	0.109449	-5.726580	0.0000
AR(24)	-0.384701	0.090165	-4.266627	0.0002

R-squared	0.953004	Mean dependent var	555.0258
Adjusted R-squared	0.945171	S.D. dependent var	93.94524
S.E. of regression	21.99779	Akaike info criterion	9.170773
Sum squared resid	14517.08	Schwarz criterion	9.434692
Log likelihood	-159.0739	F-statistic	121.6702
Durbin-Watson stat	1.980547	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.90 -.33i	.90+.33i
	.74+.62i	.74 -.62i	.62 -.74i	.62+.74i
	.33+.90i	.33 -.90i	.17+.95i	.17 -.95i
	-.17 -.95i	-.17+.95i	-.33+.90i	-.33 -.90i
	-.62+.74i	-.62 -.74i	-.74+.62i	-.74 -.62i
	-.90 -.33i	-.90+.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.24. Estimación del Modelo 20.5.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN20(-1)	-0.132568	0.049143	-2.697591	0.0112
QMENSUAL(-12)	1.315170	0.057531	22.86026	0.0000
C	-100.1624	22.90034	-4.373842	0.0001
AR(12)	-0.633438	0.133640	-4.739879	0.0000
AR(24)	-0.375023	0.107465	-3.489722	0.0015

R-squared	0.932204	Mean dependent var	555.0258
Adjusted R-squared	0.923456	S.D. dependent var	93.94524
S.E. of regression	25.99149	Akaike info criterion	9.481661
Sum squared resid	20942.28	Schwarz criterion	9.701594
Log likelihood	-165.6699	F-statistic	106.5629
Durbin-Watson stat	1.846842	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62+.73i	.62 -.73i
	.33 -.90i	.33+.90i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.33 -.90i	-.33+.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90+.33i	-.90 -.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.25. Estimación del Modelo 20.6.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	-0.078277	0.027151	-2.883017	0.0072
QMENSUAL(-12)	1.357546	0.051896	26.15895	0.0000
GEQMEN20(-1)	-0.125975	0.043153	-2.919276	0.0066
C	-80.72847	21.46439	-3.761042	0.0007
AR(12)	-0.683499	0.118745	-5.756036	0.0000
AR(24)	-0.402592	0.094766	-4.248292	0.0002

R-squared	0.946720	Mean dependent var	555.0258
Adjusted R-squared	0.937840	S.D. dependent var	93.94524
S.E. of regression	23.42230	Akaike info criterion	9.296266
Sum squared resid	16458.13	Schwarz criterion	9.560186
Log likelihood	-161.3328	F-statistic	106.6130
Durbin-Watson stat	1.677468	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91 -.33i	.91+.33i
	.74+.62i	.74 -.62i	.62+.74i	.62 -.74i
	.33 -.91i	.33+.91i	.17+.95i	.17 -.95i
	-.17 -.95i	-.17+.95i	-.33 -.91i	-.33+.91i
	-.62 -.74i	-.62+.74i	-.74+.62i	-.74 -.62i
	-.91 -.33i	-.91+.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.26. Estimación del Modelo 20,5.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN205	-0.108725	0.028351	-3.835005	0.0006
QMENSUAL(-1)	-0.127355	0.027758	-4.588137	0.0001
QMENSUAL(-12)	1.335807	0.039684	33.66129	0.0000
C	-46.12489	15.06692	-3.061336	0.0046
AR(12)	-0.639981	0.107731	-5.940568	0.0000
AR(24)	-0.386101	0.088451	-4.365140	0.0001

R-squared	0.954729	Mean dependent var	555.0258
Adjusted R-squared	0.947184	S.D. dependent var	93.94524
S.E. of regression	21.59027	Akaike info criterion	9.133374
Sum squared resid	13984.19	Schwarz criterion	9.397294
Log likelihood	-158.4007	F-statistic	126.5353
Durbin-Watson stat	2.079342	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.90 -.33i	.90+.33i
	.74+.62i	.74 -.62i	.62+.74i	.62 -.74i
	.33 -.90i	.33+.90i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.33 -.90i	-.33+.90i
	-.62 -.74i	-.62+.74i	-.74+.62i	-.74 -.62i
	-.90+.33i	-.90 -.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.27. Estimación del Modelo 20,5.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN205(-1)	-0.208251	0.032151	-6.477372	0.0000
QMENSUAL(-1)	-0.094179	0.019320	-4.874690	0.0000
QMENSUAL(-12)	1.471006	0.040694	36.14823	0.0000
QMENSUAL(-10)	-0.081807	0.016871	-4.849040	0.0000
C	-74.02351	16.27785	-4.547500	0.0001
AR(12)	-0.794675	0.081714	-9.725043	0.0000
AR(24)	-0.487445	0.065467	-7.445684	0.0000

R-squared	0.972231	Mean dependent var	555.0258
Adjusted R-squared	0.966486	S.D. dependent var	93.94524
S.E. of regression	17.19852	Akaike info criterion	8.700189
Sum squared resid	8577.882	Schwarz criterion	9.008096
Log likelihood	-149.6034	F-statistic	169.2208
Durbin-Watson stat	2.145465	Prob(F-statistic)	0.000000

Inverted AR Roots	.95 -.18i	.95+.18i	.91 -.33i	.91+.33i
	.74 -.63i	.74+.63i	.63+.74i	.63 -.74i
	.33+.91i	.33 -.91i	.18+.95i	.18 -.95i
	-.18+.95i	-.18 -.95i	-.33 -.91i	-.33+.91i
	-.63+.74i	-.63 -.74i	-.74 -.63i	-.74+.63i
	-.91 -.33i	-.91+.33i	-.95 -.18i	-.95+.18i

Tabla.A.IV.28. Estimación del Modelo 20,5.3.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN205	-0.083938	0.025044	-3.351662	0.0022
GEQMEN205(-1)	-0.110287	0.037041	-2.977448	0.0058
QMENSUAL(-1)	-0.106900	0.024328	-4.394068	0.0001
QMENSUAL(-12)	1.422903	0.044493	31.98003	0.0000
C	-86.81515	18.40068	-4.718040	0.0001
AR(12)	-0.754320	0.096901	-7.784424	0.0000
AR(24)	-0.442757	0.076273	-5.804859	0.0000

R-squared	0.964427	Mean dependent var	555.0258
Adjusted R-squared	0.957067	S.D. dependent var	93.94524
S.E. of regression	19.46568	Akaike info criterion	8.947849
Sum squared resid	10988.47	Schwarz criterion	9.255755
Log likelihood	-154.0613	F-statistic	131.0378
Durbin-Watson stat	2.035739	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91+.32i	.91 -.32i
	.74 -.63i	.74+.63i	.63+.74i	.63 -.74i
	.32 -.91i	.32+.91i	.17 -.95i	.17+.95i
	-.17+.95i	-.17 -.95i	-.32 -.91i	-.32+.91i
	-.63 -.74i	-.63+.74i	-.74 -.63i	-.74+.63i
	-.91 -.32i	-.91+.32i	-.95+.17i	-.95 -.17i

Tabla.A.IV.29. Estimación del Modelo 20,5.4.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN205(-1)	-0.100604	0.038722	-2.598109	0.0146
QMENSUAL(-1)	-0.083748	0.022509	-3.720573	0.0008
QMENSUAL(-12)	1.381103	0.044069	31.33948	0.0000
QMENSUAL(-11)	-0.073520	0.022816	-3.222314	0.0031
C	-56.25198	22.08984	-2.546509	0.0165
AR(12)	-0.727250	0.100715	-7.220862	0.0000
AR(24)	-0.458181	0.080781	-5.671897	0.0000

R-squared	0.963636	Mean dependent var	555.0258
Adjusted R-squared	0.956112	S.D. dependent var	93.94524
S.E. of regression	19.68101	Akaike info criterion	8.969851
Sum squared resid	11232.93	Schwarz criterion	9.277758
Log likelihood	-154.4573	F-statistic	128.0809
Durbin-Watson stat	2.242154	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91 -.33i	.91+.33i
	.74+.62i	.74 -.62i	.62+.74i	.62 -.74i
	.33+.91i	.33 -.91i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.33 -.91i	-.33+.91i
	-.62 -.74i	-.62+.74i	-.74+.62i	-.74 -.62i
	-.91+.33i	-.91 -.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.30. Estimación del Modelo 20,5.5.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN205(-1)	-0.159364	0.046182	-3.450820	0.0016
QMENSUAL(-12)	1.352508	0.056553	23.91559	0.0000
C	-113.9531	22.36198	-5.095843	0.0000
AR(12)	-0.686496	0.126659	-5.420020	0.0000
AR(24)	-0.398345	0.100442	-3.965904	0.0004

R-squared	0.937990	Mean dependent var	555.0258
Adjusted R-squared	0.929989	S.D. dependent var	93.94524
S.E. of regression	24.85761	Akaike info criterion	9.392451
Sum squared resid	19154.93	Schwarz criterion	9.612384
Log likelihood	-164.0641	F-statistic	117.2295
Durbin-Watson stat	1.856817	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91 -.33i	.91+.33i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.33 -.91i	.33+.91i	.17+.95i	.17 -.95i
	-.17 -.95i	-.17+.95i	-.33 -.91i	-.33+.91i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.91 -.33i	-.91+.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.31. Estimación del Modelo 20,5.6.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN205(-1)	-0.145831	0.041560	-3.508968	0.0014
QMENSUAL(-1)	-0.071664	0.025801	-2.777578	0.0094
QMENSUAL(-12)	1.383085	0.051179	27.02440	0.0000
C	-93.30031	21.54888	-4.329706	0.0002
AR(12)	-0.719859	0.113839	-6.323463	0.0000
AR(24)	-0.416692	0.089771	-4.641725	0.0001

R-squared	0.950620	Mean dependent var	555.0258
Adjusted R-squared	0.942390	S.D. dependent var	93.94524
S.E. of regression	22.54872	Akaike info criterion	9.220246
Sum squared resid	15253.35	Schwarz criterion	9.484165
Log likelihood	-159.9644	F-statistic	115.5077
Durbin-Watson stat	1.718932	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91+.32i	.91 -.32i
	.74+.62i	.74 -.62i	.62 -.74i	.62+.74i
	.32 -.91i	.32+.91i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.32 -.91i	-.32+.91i
	-.62+.74i	-.62 -.74i	-.74+.62i	-.74 -.62i
	-.91 -.32i	-.91+.32i	-.95+.17i	-.95 -.17i

Tabla.A.IV.32. Estimación del Modelo 20,5.7.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1996:12 2000:12
Included observations: 49 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN205(-1)	0.447332	0.081440	5.492799	0.0000
QMENSUAL(-1)	0.184811	0.112347	1.644992	0.1071
QMENSUAL(-11)	-0.121934	0.100349	-1.215105	0.2308
C	293.0224	56.59568	5.177470	0.0000
AR(12)	1.277366	0.085804	14.88700	0.0000

R-squared	0.914457	Mean dependent var	533.6300
Adjusted R-squared	0.906681	S.D. dependent var	93.78636
S.E. of regression	28.65004	Akaike info criterion	9.644638
Sum squared resid	36116.29	Schwarz criterion	9.837681
Log likelihood	-231.2936	F-statistic	117.5909
Durbin-Watson stat	2.335571	Prob(F-statistic)	0.000000

Inverted AR Roots	1.02	.88+.51i	.88 -.51i	.51+.88i
	.51 -.88i	-.00 -1.02i	-.00+1.02i	-.51 -.88i
	-.51+.88i	-.88+.51i	-.88 -.51i	-1.02

Estimated AR process is nonstationary

Tabla.A.IV.33. Estimación del Modelo 21.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN21(-1)	-0.095294	0.037941	-2.511646	0.0178
QMENSUAL(-1)	-0.097393	0.025114	-3.877977	0.0006
QMENSUAL(-12)	1.407629	0.045037	31.25506	0.0000
GEQMEN21	-0.075497	0.025466	-2.964552	0.0060
C	-87.21310	19.54058	-4.463179	0.0001
AR(12)	-0.741041	0.100596	-7.366488	0.0000
AR(24)	-0.427056	0.078877	-5.414235	0.0000

R-squared	0.962415	Mean dependent var	555.0258
Adjusted R-squared	0.954638	S.D. dependent var	93.94524
S.E. of regression	20.00872	Akaike info criterion	9.002879
Sum squared resid	11610.12	Schwarz criterion	9.310786
Log likelihood	-155.0518	F-statistic	123.7627
Durbin-Watson stat	2.013581	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91+.32i	.91 -.32i
	.74+.63i	.74 -.63i	.63+.74i	.63 -.74i
	.32 -.91i	.32+.91i	.17 -.95i	.17+.95i
	-.17+.95i	-.17 -.95i	-.32+.91i	-.32 -.91i
	-.63 -.74i	-.63+.74i	-.74 -.63i	-.74+.63i
	-.91 -.32i	-.91+.32i	-.95+.17i	-.95 -.17i

Tabla.A.IV.34. Estimación del Modelo 21.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN21(-1)	-0.174466	0.032944	-5.295809	0.0000
QMENSUAL(-1)	-0.086110	0.021818	-3.946830	0.0005
QMENSUAL(-12)	1.439719	0.043519	33.08220	0.0000
QMENSUAL(-10)	-0.067899	0.018291	-3.712068	0.0009
C	-74.31824	18.90500	-3.931143	0.0005
AR(12)	-0.764170	0.091413	-8.359525	0.0000
AR(24)	-0.457821	0.072718	-6.295866	0.0000

R-squared	0.966506	Mean dependent var	555.0258
Adjusted R-squared	0.959577	S.D. dependent var	93.94524
S.E. of regression	18.88823	Akaike info criterion	8.887621
Sum squared resid	10346.20	Schwarz criterion	9.195528
Log likelihood	-152.9772	F-statistic	139.4724
Durbin-Watson stat	1.992833	Prob(F-statistic)	0.000000

Inverted AR Roots	.95 -.17i	.95+.17i	.91+.33i	.91 -.33i
	.74 -.63i	.74+.63i	.63+.74i	.63 -.74i
	.33 -.91i	.33+.91i	.17+.95i	.17 -.95i
	-.17+.95i	-.17 -.95i	-.33 -.91i	-.33+.91i
	-.63 -.74i	-.63+.74i	-.74 -.63i	-.74+.63i
	-.91 -.33i	-.91+.33i	-.95 -.17i	-.95+.17i

Tabla.A.IV.35. Estimación del Modelo.21.3.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN21	-0.102764	0.026528	-3.873744	0.0005
QMENSUAL(-1)	-0.119379	0.026885	-4.440275	0.0001
QMENSUAL(-12)	1.335241	0.039331	33.94869	0.0000
C	-50.56420	15.18751	-3.329328	0.0023
AR(12)	-0.644013	0.107763	-5.976213	0.0000
AR(24)	-0.382567	0.088254	-4.334856	0.0002

R-squared	0.955030	Mean dependent var	555.0258
Adjusted R-squared	0.947535	S.D. dependent var	93.94524
S.E. of regression	21.51848	Akaike info criterion	9.126713
Sum squared resid	13891.35	Schwarz criterion	9.390633
Log likelihood	-158.2808	F-statistic	127.4211
Durbin-Watson stat	2.119600	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.33 -.90i	.33+.90i	.17+.95i	.17 -.95i
	-.17 -.95i	-.17+.95i	-.33 -.90i	-.33+.90i
	-.62 -.73i	-.62+.73i	-.73 -.62i	-.73+.62i
	-.90+.33i	-.90 -.33i	-.95 -.17i	-.95+.17i

Tabla.A.IV.36. Estimación del Modelo 21.4.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN21(-1)	-0.093126	0.037547	-2.480243	0.0192
QMENSUAL(-1)	-0.080948	0.023092	-3.505535	0.0015
QMENSUAL(-12)	1.376712	0.044116	31.20665	0.0000
QMENSUAL(-11)	-0.072494	0.023386	-3.099877	0.0043
C	-56.83114	23.03250	-2.467432	0.0198
AR(12)	-0.721551	0.101587	-7.102770	0.0000
AR(24)	-0.450123	0.081249	-5.540010	0.0000

R-squared	0.963118	Mean dependent var	555.0258
Adjusted R-squared	0.955487	S.D. dependent var	93.94524
S.E. of regression	19.82063	Akaike info criterion	8.983989
Sum squared resid	11392.86	Schwarz criterion	9.291896
Log likelihood	-154.7118	F-statistic	126.2150
Durbin-Watson stat	2.195977	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91 -.33i	.91+.33i
	.74+.62i	.74 -.62i	.62+.74i	.62 -.74i
	.33+.91i	.33 -.91i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.33 -.91i	-.33+.91i
	-.62 -.74i	-.62+.74i	-.74+.62i	-.74 -.62i
	-.91 -.33i	-.91+.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.37. Estimación del Modelo 21.5.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN21(-1)	-0.156447	0.042520	-3.679370	0.0009
QMENSUAL(-12)	1.359120	0.055203	24.62031	0.0000
C	-117.0179	22.00163	-5.318602	0.0000
AR(12)	-0.696630	0.124391	-5.600328	0.0000
AR(24)	-0.397565	0.098108	-4.052338	0.0003

R-squared	0.939925	Mean dependent var	555.0258
Adjusted R-squared	0.932174	S.D. dependent var	93.94524
S.E. of regression	24.46661	Akaike info criterion	9.360742
Sum squared resid	18557.07	Schwarz criterion	9.580675
Log likelihood	-163.4934	F-statistic	121.2560
Durbin-Watson stat	1.893090	Prob(F-statistic)	0.000000

Inverted AR Roots	.95 -.17i	.95+.17i	.91+.32i	.91 -.32i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.32 -.91i	.32+.91i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.32 -.91i	-.32+.91i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.91+.32i	-.91 -.32i	-.95+.17i	-.95 -.17i

Tabla.A.IV.38. Estimación del Modelo 21.6.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN21(-1)	-0.139058	0.039330	-3.535665	0.0013
QMENSUAL(-1)	-0.067273	0.026014	-2.586010	0.0148
QMENSUAL(-12)	1.381813	0.050612	27.30214	0.0000
C	-95.46052	21.90754	-4.357427	0.0001
AR(12)	-0.719168	0.113624	-6.329382	0.0000
AR(24)	-0.410276	0.089214	-4.598776	0.0001

R-squared	0.950883	Mean dependent var	555.0258
Adjusted R-squared	0.942697	S.D. dependent var	93.94524
S.E. of regression	22.48872	Akaike info criterion	9.214916
Sum squared resid	15172.27	Schwarz criterion	9.478836
Log likelihood	-159.8685	F-statistic	116.1570
Durbin-Watson stat	1.733520	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91+.32i	.91 -.32i
	.73+.62i	.73 -.62i	.62+.73i	.62 -.73i
	.32 -.91i	.32+.91i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.32 -.91i	-.32+.91i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.91 -.32i	-.91+.32i	-.95+.17i	-.95 -.17i

Tabla.A.IV.39. Estimación del Modelo 21.7.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN21	-0.065273	0.034500	-1.891977	0.0679
QMENSUAL(-12)	1.233842	0.044307	27.84784	0.0000
C	-69.47381	19.75660	-3.516485	0.0014
AR(12)	-0.517953	0.139486	-3.713304	0.0008
AR(24)	-0.313484	0.116830	-2.683249	0.0116

R-squared	0.927261	Mean dependent var	555.0258
Adjusted R-squared	0.917875	S.D. dependent var	93.94524
S.E. of regression	26.92226	Akaike info criterion	9.552030
Sum squared resid	22469.06	Schwarz criterion	9.771963
Log likelihood	-166.9365	F-statistic	98.79534
Durbin-Watson stat	2.259592	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.16i	.94 -.16i	.89 -.33i	.89+.33i
	.73+.61i	.73 -.61i	.61 -.73i	.61+.73i
	.33+.89i	.33 -.89i	.16 -.94i	.16+.94i
	-.16+.94i	-.16 -.94i	-.33 -.89i	-.33+.89i
	-.61+.73i	-.61 -.73i	-.73+.61i	-.73 -.61i
	-.89+.33i	-.89 -.33i	-.94 -.16i	-.94+.16i

Tabla.A.IV.40. Estimación del Modelo 21,5.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN215(-1)	-0.129834	0.033246	-3.905292	0.0005
QMENSUAL(-1)	-0.083639	0.025346	-3.299881	0.0026
QMENSUAL(-12)	1.393636	0.046343	30.07206	0.0000
QMENSUAL(-10)	-0.055505	0.020616	-2.692398	0.0117
C	-65.52929	21.98600	-2.980500	0.0058
AR(12)	-0.709358	0.103680	-6.841779	0.0000
AR(24)	-0.419861	0.082519	-5.088059	0.0000

R-squared	0.958554	Mean dependent var	555.0258
Adjusted R-squared	0.949979	S.D. dependent var	93.94524
S.E. of regression	21.01125	Akaike info criterion	9.100659
Sum squared resid	12802.71	Schwarz criterion	9.408565
Log likelihood	-156.8119	F-statistic	111.7838
Durbin-Watson stat	1.927694	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91 -.33i	.91+.33i
	.74+.62i	.74 -.62i	.62+.74i	.62 -.74i
	.33 -.91i	.33+.91i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.33 -.91i	-.33+.91i
	-.62 -.74i	-.62+.74i	-.74+.62i	-.74 -.62i
	-.91 -.33i	-.91+.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.41. Estimación del Modelo 21,5.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN215	-0.093365	0.024990	-3.736041	0.0008
QMENSUAL(-1)	-0.110914	0.026684	-4.156483	0.0002
QMENSUAL(-12)	1.329673	0.039343	33.79661	0.0000
C	-53.31584	15.60826	-3.415874	0.0018
AR(12)	-0.638497	0.109130	-5.850816	0.0000
AR(24)	-0.375531	0.089300	-4.205297	0.0002

R-squared	0.954109	Mean dependent var	555.0258
Adjusted R-squared	0.946461	S.D. dependent var	93.94524
S.E. of regression	21.73760	Akaike info criterion	9.146975
Sum squared resid	14175.69	Schwarz criterion	9.410895
Log likelihood	-158.6456	F-statistic	124.7448
Durbin-Watson stat	2.117344	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.33 -.90i	.33+.90i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.33 -.90i	-.33+.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90+.33i	-.90 -.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.42. Estimación del Modelo 21,5.3.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN215(-1)	-0.071866	0.034865	-2.061244	0.0483
QMENSUAL(-1)	-0.081830	0.024164	-3.386436	0.0021
QMENSUAL(-12)	1.359094	0.043489	31.25120	0.0000
QMENSUAL(-11)	-0.076144	0.024233	-3.142168	0.0038
C	-48.84561	23.29963	-2.096411	0.0449
AR(12)	-0.697504	0.104119	-6.699116	0.0000
AR(24)	-0.436932	0.083955	-5.204374	0.0000

R-squared	0.961309	Mean dependent var	555.0258
Adjusted R-squared	0.953304	S.D. dependent var	93.94524
S.E. of regression	20.30091	Akaike info criterion	9.031874
Sum squared resid	11951.68	Schwarz criterion	9.339780
Log likelihood	-155.5737	F-statistic	120.0877
Durbin-Watson stat	2.125863	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91 -.33i	.91+.33i
	.74+.62i	.74 -.62i	.62 -.74i	.62+.74i
	.33 -.91i	.33+.91i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.33 -.91i	-.33+.91i
	-.62+.74i	-.62 -.74i	-.74 -.62i	-.74+.62i
	-.91 -.33i	-.91+.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.43. Estimación del Modelo 21,5.4.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN215(-1)	-0.132915	0.039304	-3.381726	0.0020
QMENSUAL(-12)	1.339577	0.054376	24.63529	0.0000
C	-110.4689	22.01929	-5.016916	0.0000
AR(12)	-0.669217	0.126948	-5.271589	0.0000
AR(24)	-0.378780	0.100464	-3.770299	0.0007

R-squared	0.937857	Mean dependent var	555.0258
Adjusted R-squared	0.929839	S.D. dependent var	93.94524
S.E. of regression	24.88418	Akaike info criterion	9.394588
Sum squared resid	19195.90	Schwarz criterion	9.614521
Log likelihood	-164.1026	F-statistic	116.9627
Durbin-Watson stat	1.955799	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.90+.32i	.90 -.32i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.32 -.90i	.32+.90i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.32+.90i	-.32 -.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90 -.32i	-.90+.32i	-.95+.17i	-.95 -.17i

Tabla.A.IV.44. Estimación del Modelo 21,5.5.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN215	-0.067970	0.032100	-2.117434	0.0423
QMENSUAL(-12)	1.240537	0.043986	28.20298	0.0000
C	-72.03355	19.57436	-3.679996	0.0009
AR(12)	-0.528310	0.137565	-3.840436	0.0006
AR(24)	-0.317050	0.114937	-2.758462	0.0097

R-squared	0.929024	Mean dependent var	555.0258
Adjusted R-squared	0.919866	S.D. dependent var	93.94524
S.E. of regression	26.59404	Akaike info criterion	9.527497
Sum squared resid	21924.52	Schwarz criterion	9.747430
Log likelihood	-166.4949	F-statistic	101.4416
Durbin-Watson stat	2.296083	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.16i	.94 -.16i	.89 -.33i	.89+.33i
	.73 -.61i	.73+.61i	.61 -.73i	.61+.73i
	.33+.89i	.33 -.89i	.16 -.94i	.16+.94i
	-.16+.94i	-.16 -.94i	-.33 -.89i	-.33+.89i
	-.61+.73i	-.61 -.73i	-.73+.61i	-.73 -.61i
	-.89+.33i	-.89 -.33i	-.94 -.16i	-.94+.16i

Tabla.A.IV.45. Estimación del Modelo 21,5.6.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN215(-1)	-0.114852	0.036936	-3.109448	0.0041
QMENSUAL(-1)	-0.066919	0.027391	-2.443156	0.0207
QMENSUAL(-12)	1.360365	0.050367	27.00882	0.0000
C	-88.33100	22.39299	-3.944582	0.0004
AR(12)	-0.689381	0.117140	-5.885128	0.0000
AR(24)	-0.390683	0.092329	-4.231424	0.0002

R-squared	0.948184	Mean dependent var	555.0258
Adjusted R-squared	0.939548	S.D. dependent var	93.94524
S.E. of regression	23.09831	Akaike info criterion	9.268407
Sum squared resid	16005.95	Schwarz criterion	9.532327
Log likelihood	-160.8313	F-statistic	109.7943
Durbin-Watson stat	1.760069	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91 -.32i	.91+.32i
	.73 -.62i	.73+.62i	.62 -.73i	.62+.73i
	.32+.91i	.32 -.91i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.32+.91i	-.32 -.91i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.91 -.32i	-.91+.32i	-.95+.17i	-.95 -.17i

Tabla.A.IV.46. Estimación del Modelo 21,5.7.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1996:12 2000:12
Included observations: 49 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN215(-1)	0.394148	0.080699	4.884169	0.0000
QMENSUAL(-1)	0.194085	0.118667	1.635537	0.1091
QMENSUAL(-11)	-0.124327	0.106050	-1.172342	0.2474
C	282.7115	62.19530	4.545544	0.0000
AR(12)	1.251595	0.085274	14.67736	0.0000

R-squared	0.907943	Mean dependent var	533.6300
Adjusted R-squared	0.899574	S.D. dependent var	93.78636
S.E. of regression	29.72097	Akaike info criterion	9.718033
Sum squared resid	38866.78	Schwarz criterion	9.911076
Log likelihood	-233.0918	F-statistic	108.4909
Durbin-Watson stat	2.261769	Prob(F-statistic)	0.000000

Inverted AR Roots	1.02	.88+.51i	.88 -.51i	.51+.88i
	.51 -.88i	-.00 -1.02i	-.00+1.02i	-.51 -.88i
	-.51+.88i	-.88+.51i	-.88 -.51i	-1.02

Estimated AR process is nonstationary

Tabla.A.IV.47. Estimación del Modelo 21,5.8.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN215	-0.069662	0.027320	-2.549890	0.0163
GEQMEN215(-1)	-0.065204	0.039241	-1.661640	0.1074
QMENSUAL(-1)	-0.093405	0.027139	-3.441773	0.0018
QMENSUAL(-12)	1.375399	0.046108	29.83019	0.0000
C	-78.43574	20.88229	-3.756089	0.0008
AR(12)	-0.700116	0.107645	-6.503948	0.0000
AR(24)	-0.401207	0.084769	-4.732932	0.0001

R-squared	0.957847	Mean dependent var	555.0258
Adjusted R-squared	0.949126	S.D. dependent var	93.94524
S.E. of regression	21.18962	Akaike info criterion	9.117566
Sum squared resid	13021.01	Schwarz criterion	9.425473
Log likelihood	-157.1162	F-statistic	109.8287
Durbin-Watson stat	1.996483	Prob(F-statistic)	0.000000

Inverted AR Roots	.95 -.17i	.95+.17i	.91+.32i	.91 -.32i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.32 -.91i	.32+.91i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.32 -.91i	-.32+.91i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.91+.32i	-.91 -.32i	-.95+.17i	-.95 -.17i

Tabla.A.IV.48. Estimación del Modelo 22.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN22	-0.066238	0.029445	-2.249579	0.0317
QMENSUAL(-12)	1.243772	0.043611	28.51958	0.0000
C	-73.59252	19.47227	-3.779349	0.0007
AR(12)	-0.534179	0.136518	-3.912878	0.0005
AR(24)	-0.318061	0.113883	-2.792879	0.0089

R-squared	0.930112	Mean dependent var	555.0258
Adjusted R-squared	0.921095	S.D. dependent var	93.94524
S.E. of regression	26.38929	Akaike info criterion	9.512040
Sum squared resid	21588.24	Schwarz criterion	9.731973
Log likelihood	-166.2167	F-statistic	103.1425
Durbin-Watson stat	2.323598	Prob(F-statistic)	0.000000

Inverted AR Roots	.94 -.16i	.94+.16i	.90 -.33i	.90+.33i
	.73 -.61i	.73+.61i	.61 -.73i	.61+.73i
	.33 -.90i	.33+.90i	.16+.94i	.16 -.94i
	-.16 -.94i	-.16+.94i	-.33 -.90i	-.33+.90i
	-.61 -.73i	-.61+.73i	-.73+.61i	-.73 -.61i
	-.90+.33i	-.90 -.33i	-.94+.16i	-.94 -.16i

Tabla.A.IV.49. Estimación del Modelo 22.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN22(-1)	-0.104348	0.035716	-2.921551	0.0064
QMENSUAL(-12)	1.311414	0.053294	24.60724	0.0000
C	-100.7189	22.06573	-4.564494	0.0001
AR(12)	-0.629742	0.131022	-4.806398	0.0000
AR(24)	-0.356106	0.104565	-3.405593	0.0018

R-squared	0.934463	Mean dependent var	555.0258
Adjusted R-squared	0.926007	S.D. dependent var	93.94524
S.E. of regression	25.55468	Akaike info criterion	9.447764
Sum squared resid	20244.28	Schwarz criterion	9.667697
Log likelihood	-165.0597	F-statistic	110.5043
Durbin-Watson stat	2.029783	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62+.73i	.62 -.73i
	.33+.90i	.33 -.90i	.17 -.94i	.17+.94i
	-.17+.94i	-.17 -.94i	-.33+.90i	-.33 -.90i
	-.62+.73i	-.62 -.73i	-.73+.62i	-.73 -.62i
	-.90 -.33i	-.90+.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.50. Estimación del Modelo 22.3.

Dependent Variable: QMENSUAL
Method: Least Squares
Date: 02/15/03 Time: 13:48
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN19(-1)	-0.054139	0.048617	-1.113581	0.2740
QMENSUAL(-12)	1.230316	0.054398	22.61685	0.0000
C	-70.26234	22.74417	-3.089247	0.0042
AR(12)	-0.514800	0.146379	-3.516892	0.0014
AR(24)	-0.314060	0.121494	-2.584973	0.0147

R-squared	0.922051	Mean dependent var	555.0258
Adjusted R-squared	0.911993	S.D. dependent var	93.94524
S.E. of regression	27.86978	Akaike info criterion	9.621209
Sum squared resid	24078.47	Schwarz criterion	9.841142
Log likelihood	-168.1818	F-statistic	91.67381
Durbin-Watson stat	2.003599	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.16i	.94 -.16i	.89+.33i	.89 -.33i
	.73+.61i	.73 -.61i	.61+.73i	.61 -.73i
	.33 -.89i	.33+.89i	.16+.94i	.16 -.94i
	-.16 -.94i	-.16+.94i	-.33+.89i	-.33 -.89i
	-.61 -.73i	-.61+.73i	-.73+.61i	-.73 -.61i
	-.89+.33i	-.89 -.33i	-.94+.16i	-.94 -.16i

Tabla.A.IV.51. Estimación del Modelo 22.4.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	-0.103296	0.026912	-3.838294	0.0006
QMENSUAL(-12)	1.321546	0.039532	33.42978	0.0000
GEQMEN22	-0.081990	0.023397	-3.504336	0.0015
C	-54.77066	16.16541	-3.388138	0.0020
AR(12)	-0.629931	0.111498	-5.649717	0.0000
AR(24)	-0.367929	0.091287	-4.030487	0.0004

R-squared	0.952456	Mean dependent var	555.0258
Adjusted R-squared	0.944532	S.D. dependent var	93.94524
S.E. of regression	22.12571	Akaike info criterion	9.182369
Sum squared resid	14686.41	Schwarz criterion	9.446289
Log likelihood	-159.2826	F-statistic	120.1982
Durbin-Watson stat	2.106337	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.33+.90i	.33 -.90i	.17 -.94i	.17+.94i
	-.17+.94i	-.17 -.94i	-.33 -.90i	-.33+.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90 -.33i	-.90+.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.52. Estimación del Modelo 22.5.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
QMENSUAL(-1)	-0.068925	0.028993	-2.377283	0.0240
QMENSUAL(-12)	1.333802	0.049654	26.86197	0.0000
GEQMEN22(-1)	-0.087310	0.033710	-2.589999	0.0147
C	-78.20106	22.63895	-3.454271	0.0017
AR(12)	-0.652363	0.121326	-5.376937	0.0000
AR(24)	-0.369593	0.096423	-3.833056	0.0006

R-squared	0.944851	Mean dependent var	555.0258
Adjusted R-squared	0.935660	S.D. dependent var	93.94524
S.E. of regression	23.82960	Akaike info criterion	9.330746
Sum squared resid	17035.49	Schwarz criterion	9.594665
Log likelihood	-161.9534	F-statistic	102.7963
Durbin-Watson stat	1.803252	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.32i	.90+.32i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.32 -.90i	.32+.90i	.17 -.94i	.17+.94i
	-.17 -.94i	-.17+.94i	-.32+.90i	-.32 -.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90+.32i	-.90 -.32i	-.94+.17i	-.94 -.17i

Tabla.A.IV.53. Estimación del Modelo 22.6.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN22	-0.066359	0.028334	-2.342048	0.0263
QMENSUAL(-1)	-0.092619	0.028773	-3.218951	0.0032
QMENSUAL(-12)	1.344788	0.046064	29.19415	0.0000
GEQMEN22(-1)	-0.036108	0.038386	-0.940642	0.3547
C	-68.49239	21.45579	-3.192256	0.0034
AR(12)	-0.660970	0.113526	-5.822175	0.0000
AR(24)	-0.379340	0.090222	-4.204525	0.0002

R-squared	0.953809	Mean dependent var	555.0258
Adjusted R-squared	0.944252	S.D. dependent var	93.94524
S.E. of regression	22.18146	Akaike info criterion	9.209057
Sum squared resid	14268.50	Schwarz criterion	9.516963
Log likelihood	-158.7630	F-statistic	99.80382
Durbin-Watson stat	2.024326	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.90+.33i	.90 -.33i
	.73+.62i	.73 -.62i	.62+.73i	.62 -.73i
	.33 -.90i	.33+.90i	.17+.95i	.17 -.95i
	-.17 -.95i	-.17+.95i	-.33 -.90i	-.33+.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90 -.33i	-.90+.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.54. Estimación del Modelo 22.7.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN22(-1)	-0.050070	0.030831	-1.623987	0.1152
QMENSUAL(-1)	-0.084379	0.025090	-3.363134	0.0022
QMENSUAL(-12)	1.340367	0.042087	31.84763	0.0000
QMENSUAL(-11)	-0.081148	0.024830	-3.268170	0.0028
C	-39.04302	22.84091	-1.709347	0.0981
AR(12)	-0.672447	0.106388	-6.320727	0.0000
AR(24)	-0.425817	0.086857	-4.902528	0.0000

R-squared	0.959559	Mean dependent var	555.0258
Adjusted R-squared	0.951192	S.D. dependent var	93.94524
S.E. of regression	20.75483	Akaike info criterion	9.076101
Sum squared resid	12492.13	Schwarz criterion	9.384007
Log likelihood	-156.3698	F-statistic	114.6832
Durbin-Watson stat	2.078654	Prob(F-statistic)	0.000000

Inverted AR Roots	.95+.17i	.95 -.17i	.91 -.33i	.91+.33i
	.74+.62i	.74 -.62i	.62+.74i	.62 -.74i
	.33 -.91i	.33+.91i	.17 -.95i	.17+.95i
	-.17 -.95i	-.17+.95i	-.33 -.91i	-.33+.91i
	-.62 -.74i	-.62+.74i	-.74+.62i	-.74 -.62i
	-.91+.33i	-.91 -.33i	-.95+.17i	-.95 -.17i

Tabla.A.IV.55. Estimación del Modelo 22,5.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN225	-0.061865	0.026892	-2.300489	0.0283
QMENSUAL(-12)	1.244008	0.043249	28.76401	0.0000
C	-74.12815	19.43554	-3.814052	0.0006
AR(12)	-0.535577	0.136171	-3.933111	0.0004
AR(24)	-0.316906	0.113529	-2.791401	0.0089

R-squared	0.930562	Mean dependent var	555.0258
Adjusted R-squared	0.921602	S.D. dependent var	93.94524
S.E. of regression	26.30436	Akaike info criterion	9.505593
Sum squared resid	21449.50	Schwarz criterion	9.725526
Log likelihood	-166.1007	F-statistic	103.8597
Durbin-Watson stat	2.341213	Prob(F-statistic)	0.000000

Inverted AR Roots	.94 -.16i	.94+.16i	.90 -.33i	.90+.33i
	.73+.61i	.73 -.61i	.61 -.73i	.61+.73i
	.33+.90i	.33 -.90i	.16+.94i	.16 -.94i
	-.16+.94i	-.16 -.94i	-.33 -.90i	-.33+.90i
	-.61+.73i	-.61 -.73i	-.73+.61i	-.73 -.61i
	-.90+.33i	-.90 -.33i	-.94 -.16i	-.94+.16i

Tabla.A.IV.56. Estimación del Modelo 22,5.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1996:12 2000:12
Included observations: 49 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN225(-1)	0.396689	0.085651	4.631466	0.0000
QMENSUAL(-1)	0.180331	0.122920	1.467065	0.1495
QMENSUAL(-11)	-0.117581	0.109507	-1.073732	0.2888
C	262.4697	69.56767	3.772869	0.0005
AR(12)	1.221603	0.081528	14.98376	0.0000

R-squared	0.905739	Mean dependent var	533.6300
Adjusted R-squared	0.897170	S.D. dependent var	93.78636
S.E. of regression	30.07459	Akaike info criterion	9.741690
Sum squared resid	39797.18	Schwarz criterion	9.934732
Log likelihood	-233.6714	F-statistic	105.6974
Durbin-Watson stat	2.221941	Prob(F-statistic)	0.000000

Inverted AR Roots	1.02	.88+.51i	.88 -.51i	.51+.88i
	.51 -.88i	.00 -1.02i	-.00+1.02i	-.51 -.88i
	-.51+.88i	-.88 -.51i	-.88+.51i	-1.02

Estimated AR process is nonstationary

Tabla.A.IV.57. Estimación del Modelo 22,5.3.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN225(-1)	-0.081172	0.032116	-2.527434	0.0168
QMENSUAL(-12)	1.287125	0.051939	24.78124	0.0000
C	-92.19297	21.98899	-4.192688	0.0002
AR(12)	-0.596371	0.134673	-4.428297	0.0001
AR(24)	-0.338640	0.108376	-3.124666	0.0038

R-squared	0.931534	Mean dependent var	555.0258
Adjusted R-squared	0.922700	S.D. dependent var	93.94524
S.E. of regression	26.11948	Akaike info criterion	9.491486
Sum squared resid	21149.05	Schwarz criterion	9.711419
Log likelihood	-165.8467	F-statistic	105.4453
Durbin-Watson stat	2.089328	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.33+.90i	.33 -.90i	.17+.94i	.17 -.94i
	-.17 -.94i	-.17+.94i	-.33 -.90i	-.33+.90i
	-.62+.73i	-.62 -.73i	-.73+.62i	-.73 -.62i
	-.90 -.33i	-.90+.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.58. Estimación del Modelo 22,5.4.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN225(-1)	-0.065905	0.030213	-2.181336	0.0371
QMENSUAL(-1)	-0.071478	0.030193	-2.367388	0.0246
QMENSUAL(-12)	1.312529	0.048554	27.03215	0.0000
C	-69.57645	22.55400	-3.084883	0.0043
AR(12)	-0.623491	0.124615	-5.003335	0.0000
AR(24)	-0.354587	0.099816	-3.552411	0.0013

R-squared	0.942282	Mean dependent var	555.0258
Adjusted R-squared	0.932662	S.D. dependent var	93.94524
S.E. of regression	24.37834	Akaike info criterion	9.376279
Sum squared resid	17829.10	Schwarz criterion	9.640199
Log likelihood	-162.7730	F-statistic	97.95355
Durbin-Watson stat	1.838433	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.33 -.90i	.33+.90i	.17 -.94i	.17+.94i
	-.17 -.94i	-.17+.94i	-.33 -.90i	-.33+.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90+.33i	-.90 -.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.59. Estimación del Modelo 22,5.5.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN225	-0.070724	0.021850	-3.236795	0.0029
QMENSUAL(-1)	-0.097091	0.027399	-3.543568	0.0013
QMENSUAL(-12)	1.312455	0.039776	32.99591	0.0000
C	-55.07993	16.74825	-3.288698	0.0026
AR(12)	-0.619435	0.114158	-5.426115	0.0000
AR(24)	-0.360091	0.093590	-3.847530	0.0006

R-squared	0.950510	Mean dependent var	555.0258
Adjusted R-squared	0.942261	S.D. dependent var	93.94524
S.E. of regression	22.57402	Akaike info criterion	9.222488
Sum squared resid	15287.59	Schwarz criterion	9.486408
Log likelihood	-160.0048	F-statistic	115.2355
Durbin-Watson stat	2.096283	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.33 -.90i	.33+.90i	.17+.94i	.17 -.94i
	-.17 -.94i	-.17+.94i	-.33 -.90i	-.33+.90i
	-.62 -.73i	-.62+.73i	-.73 -.62i	-.73+.62i
	-.90 -.33i	-.90+.33i	-.94 -.17i	-.94+.17i

Tabla.A.IV.60. Estimación del Modelo 22,5.6.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN225(-1)	-0.065741	0.028803	-2.282412	0.0300
QMENSUAL(-1)	-0.086813	0.030122	-2.882010	0.0074
QMENSUAL(-12)	1.326269	0.046809	28.33363	0.0000
QMENSUAL(-10)	-0.044333	0.024057	-1.842785	0.0756
C	-45.91955	25.35471	-1.811086	0.0805
AR(12)	-0.625417	0.118467	-5.279266	0.0000
AR(24)	-0.371441	0.095404	-3.893361	0.0005

R-squared	0.948419	Mean dependent var	555.0258
Adjusted R-squared	0.937747	S.D. dependent var	93.94524
S.E. of regression	23.43992	Akaike info criterion	9.319424
Sum squared resid	15933.47	Schwarz criterion	9.627330
Log likelihood	-160.7496	F-statistic	88.86978
Durbin-Watson stat	1.966442	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62+.73i	.62 -.73i
	.33 -.90i	.33+.90i	.17 -.94i	.17+.94i
	-.17 -.94i	-.17+.94i	-.33 -.90i	-.33+.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90+.33i	-.90 -.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.61. Estimación del Modelo 23.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN23	-0.056688	0.024599	-2.304525	0.0281
QMENSUAL(-12)	1.242667	0.042940	28.93949	0.0000
C	-74.05677	19.43887	-3.809725	0.0006
AR(12)	-0.534215	0.136204	-3.922175	0.0005
AR(24)	-0.314559	0.113558	-2.770030	0.0094

R-squared	0.930632	Mean dependent var	555.0258
Adjusted R-squared	0.921681	S.D. dependent var	93.94524
S.E. of regression	26.29110	Akaike info criterion	9.504584
Sum squared resid	21427.89	Schwarz criterion	9.724518
Log likelihood	-166.0825	F-statistic	103.9723
Durbin-Watson stat	2.351213	Prob(F-statistic)	0.000000

Inverted AR Roots	.94 -.16i	.94+.16i	.89 -.33i	.89+.33i
	.73 -.61i	.73+.61i	.61 -.73i	.61+.73i
	.33 -.89i	.33+.89i	.16 -.94i	.16+.94i
	-.16 -.94i	-.16+.94i	-.33+.89i	-.33 -.89i
	-.61+.73i	-.61 -.73i	-.73 -.61i	-.73+.61i
	-.89 -.33i	-.89+.33i	-.94 -.16i	-.94+.16i

Tabla.A.IV.62. Estimación del Modelo 23.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Date: 02/15/03 Time: 15:07
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN23(-1)	-0.064334	0.028843	-2.230463	0.0331
QMENSUAL(-12)	1.268832	0.050576	25.08762	0.0000
C	-85.71064	21.84015	-3.924454	0.0005
AR(12)	-0.571419	0.137468	-4.156735	0.0002
AR(24)	-0.326352	0.111378	-2.930143	0.0063

R-squared	0.929368	Mean dependent var	555.0258
Adjusted R-squared	0.920254	S.D. dependent var	93.94524
S.E. of regression	26.52947	Akaike info criterion	9.522635
Sum squared resid	21818.19	Schwarz criterion	9.742568
Log likelihood	-166.4074	F-statistic	101.9737
Durbin-Watson stat	2.131398	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.61i	.73 -.61i	.61 -.73i	.61+.73i
	.33+.90i	.33 -.90i	.17+.94i	.17 -.94i
	-.17 -.94i	-.17+.94i	-.33+.90i	-.33 -.90i
	-.61+.73i	-.61 -.73i	-.73+.61i	-.73 -.61i
	-.90 -.33i	-.90+.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.63. Estimación del Modelo 23.3.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN23(-1)	-0.050838	0.026990	-1.883588	0.0693
QMENSUAL(-1)	-0.073691	0.031014	-2.376070	0.0241
QMENSUAL(-12)	1.297230	0.047428	27.35154	0.0000
C	-63.15769	22.32562	-2.828934	0.0082
AR(12)	-0.603070	0.126960	-4.750075	0.0000
AR(24)	-0.344680	0.102327	-3.368410	0.0021

R-squared	0.940494	Mean dependent var	555.0258
Adjusted R-squared	0.930576	S.D. dependent var	93.94524
S.E. of regression	24.75304	Akaike info criterion	9.406785
Sum squared resid	18381.38	Schwarz criterion	9.670705
Log likelihood	-163.3221	F-statistic	94.83019
Durbin-Watson stat	1.862621	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62+.73i	.62 -.73i
	.33 -.90i	.33+.90i	.17 -.94i	.17+.94i
	-.17 -.94i	-.17+.94i	-.33 -.90i	-.33+.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90+.33i	-.90 -.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.64. Estimación del Modelo 23.4.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN23	-0.060913	0.020389	-2.987555	0.0056
QMENSUAL(-1)	-0.092363	0.027960	-3.303405	0.0025
QMENSUAL(-12)	1.304050	0.039989	32.61025	0.0000
C	-54.78767	17.27201	-3.172049	0.0035
AR(12)	-0.608989	0.116567	-5.224378	0.0000
AR(24)	-0.352835	0.095710	-3.686516	0.0009

R-squared	0.948688	Mean dependent var	555.0258
Adjusted R-squared	0.940136	S.D. dependent var	93.94524
S.E. of regression	22.98574	Akaike info criterion	9.258637
Sum squared resid	15850.33	Schwarz criterion	9.522557
Log likelihood	-160.6555	F-statistic	110.9313
Durbin-Watson stat	2.088115	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62+.73i	.62 -.73i
	.33 -.90i	.33+.90i	.17 -.94i	.17+.94i
	-.17 -.94i	-.17+.94i	-.33 -.90i	-.33+.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90+.33i	-.90 -.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.65. Estimación del Modelo 23.5.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN23(-1)	-0.048432	0.026037	-1.860161	0.0730
QMENSUAL(-1)	-0.088957	0.031341	-2.838387	0.0082
QMENSUAL(-10)	-0.042487	0.025041	-1.696691	0.1005
QMENSUAL(-12)	1.307883	0.046034	28.41139	0.0000
C	-39.33096	25.98364	-1.513681	0.1409
AR(12)	-0.602537	0.122060	-4.936389	0.0000
AR(24)	-0.360114	0.098808	-3.644579	0.0010

R-squared	0.945958	Mean dependent var	555.0258
Adjusted R-squared	0.934777	S.D. dependent var	93.94524
S.E. of regression	23.99252	Akaike info criterion	9.366027
Sum squared resid	16693.59	Schwarz criterion	9.673934
Log likelihood	-161.5885	F-statistic	84.60310
Durbin-Watson stat	1.982175	Prob(F-statistic)	0.000000

Inverted AR Roots	.94 -.17i	.94+.17i	.90+.33i	.90 -.33i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.33+.90i	.33 -.90i	.17+.94i	.17 -.94i
	-.17 -.94i	-.17+.94i	-.33 -.90i	-.33+.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90+.33i	-.90 -.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.66. Estimación del Modelo 23,5.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN235	-0.052941	0.022728	-2.329350	0.0265
QMENSUAL(-12)	1.242469	0.042749	29.06423	0.0000
C	-74.31072	19.44827	-3.820942	0.0006
AR(12)	-0.532887	0.136012	-3.917939	0.0005
AR(24)	-0.311956	0.113325	-2.752748	0.0098

R-squared	0.930894	Mean dependent var	555.0258
Adjusted R-squared	0.921977	S.D. dependent var	93.94524
S.E. of regression	26.24135	Akaike info criterion	9.500796
Sum squared resid	21346.86	Schwarz criterion	9.720729
Log likelihood	-166.0143	F-statistic	104.3964
Durbin-Watson stat	2.356759	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.16i	.94 -.16i	.89 -.33i	.89+.33i
	.73+.61i	.73 -.61i	.61 -.73i	.61+.73i
	.33+.89i	.33 -.89i	.16+.94i	.16 -.94i
	-.16 -.94i	-.16+.94i	-.33 -.89i	-.33+.89i
	-.61+.73i	-.61 -.73i	-.73+.61i	-.73 -.61i
	-.89+.33i	-.89 -.33i	-.94 -.16i	-.94+.16i

Tabla.A.IV.67. Estimación del Modelo 23,5.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1996:12 2000:12
Included observations: 49 after adjusting endpoints
Convergence achieved after 8 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN235(-1)	0.339995	0.094185	3.609857	0.0008
QMENSUAL(-1)	0.205026	0.132452	1.547923	0.1288
QMENSUAL(-11)	-0.128681	0.118765	-1.083486	0.2845
C	234.7526	86.36053	2.718286	0.0094
AR(12)	1.187490	0.080033	14.83754	0.0000

R-squared	0.893174	Mean dependent var	533.6300
Adjusted R-squared	0.883462	S.D. dependent var	93.78636
S.E. of regression	32.01643	Akaike info criterion	9.866827
Sum squared resid	45102.29	Schwarz criterion	10.05987
Log likelihood	-236.7373	F-statistic	91.97094
Durbin-Watson stat	2.255491	Prob(F-statistic)	0.000000

Inverted AR Roots	1.01	.88+.51i	.88 -.51i	.51+.88i
	.51 -.88i	.00 -1.01i	-.00+1.01i	-.51 -.88i
	-.51+.88i	-.88 -.51i	-.88+.51i	-1.01
	Estimated AR process is nonstationary			

Tabla.A.IV.68. Estimación del Modelo 23,5.3.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN235(-1)	-0.053224	0.026138	-2.036241	0.0503
QMENSUAL(-12)	1.256519	0.049377	25.44733	0.0000
C	-81.27724	21.65774	-3.752803	0.0007
AR(12)	-0.553970	0.139257	-3.978039	0.0004
AR(24)	-0.317505	0.113363	-2.800783	0.0087

R-squared	0.928013	Mean dependent var	555.0258
Adjusted R-squared	0.918724	S.D. dependent var	93.94524
S.E. of regression	26.78279	Akaike info criterion	9.541642
Sum squared resid	22236.86	Schwarz criterion	9.761576
Log likelihood	-166.7496	F-statistic	99.90787
Durbin-Watson stat	2.160192	Prob(F-statistic)	0.000000

Inverted AR Roots	.94 -.16i	.94+.16i	.90+.33i	.90 -.33i
	.73+.61i	.73 -.61i	.61 -.73i	.61+.73i
	.33+.90i	.33 -.90i	.16+.94i	.16 -.94i
	-.16 -.94i	-.16+.94i	-.33 -.90i	-.33+.90i
	-.61+.73i	-.61 -.73i	-.73+.61i	-.73 -.61i
	-.90+.33i	-.90 -.33i	-.94 -.16i	-.94+.16i

Tabla.A.IV.69. Estimación del Modelo 23,5.4.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN235(-1)	-0.040732	0.024427	-1.667534	0.1058
QMENSUAL(-1)	-0.074967	0.031646	-2.368945	0.0245
QMENSUAL(-12)	1.286517	0.046474	27.68259	0.0000
C	-58.72758	22.14488	-2.651972	0.0127
AR(12)	-0.588299	0.128565	-4.575883	0.0001
AR(24)	-0.337526	0.104099	-3.242360	0.0029

R-squared	0.939274	Mean dependent var	555.0258
Adjusted R-squared	0.929153	S.D. dependent var	93.94524
S.E. of regression	25.00553	Akaike info criterion	9.427083
Sum squared resid	18758.29	Schwarz criterion	9.691002
Log likelihood	-163.6875	F-statistic	92.80423
Durbin-Watson stat	1.882595	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73 -.61i	.73+.61i	.61+.73i	.61 -.73i
	.33 -.90i	.33+.90i	.17 -.94i	.17+.94i
	-.17 -.94i	-.17+.94i	-.33+.90i	-.33 -.90i
	-.61 -.73i	-.61+.73i	-.73+.61i	-.73 -.61i
	-.90+.33i	-.90 -.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.70. Estimación del Modelo 23,5.5.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN235	-0.054196	0.019119	-2.834635	0.0081
QMENSUAL(-1)	-0.089023	0.028357	-3.139411	0.0038
QMENSUAL(-12)	1.298881	0.040124	32.37134	0.0000
C	-54.86582	17.65683	-3.107342	0.0041
AR(12)	-0.601086	0.118017	-5.093198	0.0000
AR(24)	-0.346775	0.096945	-3.577011	0.0012

R-squared	0.947602	Mean dependent var	555.0258
Adjusted R-squared	0.938869	S.D. dependent var	93.94524
S.E. of regression	23.22759	Akaike info criterion	9.279570
Sum squared resid	16185.62	Schwarz criterion	9.543490
Log likelihood	-161.0323	F-statistic	108.5089
Durbin-Watson stat	2.079141	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.33+.90i	.33 -.90i	.17 -.94i	.17+.94i
	-.17 -.94i	-.17+.94i	-.33+.90i	-.33 -.90i
	-.62+.73i	-.62 -.73i	-.73+.62i	-.73 -.62i
	-.90 -.33i	-.90+.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.71. Estimación del Modelo 23,5.6.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN235(-1)	-0.037172	0.023752	-1.564982	0.1284
QMENSUAL(-1)	-0.090352	0.032219	-2.804340	0.0089
QMENSUAL(-12)	1.295574	0.045268	28.61995	0.0000
QMENSUAL(-10)	-0.041628	0.025745	-1.616950	0.1167
C	-34.70083	26.38039	-1.315402	0.1987
AR(12)	-0.586550	0.124328	-4.717780	0.0001
AR(24)	-0.352405	0.101059	-3.487109	0.0016

R-squared	0.944388	Mean dependent var	555.0258
Adjusted R-squared	0.932882	S.D. dependent var	93.94524
S.E. of regression	24.33860	Akaike info criterion	9.394670
Sum squared resid	17178.65	Schwarz criterion	9.702576
Log likelihood	-162.1041	F-statistic	82.07775
Durbin-Watson stat	1.992740	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.33+.90i	.33 -.90i	.17+.94i	.17 -.94i
	-.17 -.94i	-.17+.94i	-.33+.90i	-.33 -.90i
	-.62 -.73i	-.62+.73i	-.73+.62i	-.73 -.62i
	-.90 -.33i	-.90+.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.72. Estimación del Modelo 24.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN24	-0.050019	0.021328	-2.345191	0.0256
QMENSUAL(-12)	1.242426	0.042648	29.13207	0.0000
C	-74.46546	19.45435	-3.827702	0.0006
AR(12)	-0.532036	0.135754	-3.919116	0.0005
AR(24)	-0.309586	0.113111	-2.737012	0.0102

R-squared	0.931069	Mean dependent var	555.0258
Adjusted R-squared	0.922174	S.D. dependent var	93.94524
S.E. of regression	26.20813	Akaike info criterion	9.498262
Sum squared resid	21292.84	Schwarz criterion	9.718195
Log likelihood	-165.9687	F-statistic	104.6809
Durbin-Watson stat	2.356908	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.16i	.94 -.16i	.89 -.33i	.89+.33i
	.73+.61i	.73 -.61i	.61+.73i	.61 -.73i
	.33 -.89i	.33+.89i	.16 -.94i	.16+.94i
	-.16 -.94i	-.16+.94i	-.33+.89i	-.33 -.89i
	-.61 -.73i	-.61+.73i	-.73 -.61i	-.73+.61i
	-.89 -.33i	-.89+.33i	-.94+.16i	-.94 -.16i

Tabla.A.IV.73. Estimación del Modelo 24.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN24(-1)	-0.045530	0.024065	-1.891947	0.0679
QMENSUAL(-12)	1.247411	0.048339	25.80573	0.0000
C	-77.85254	21.44425	-3.630463	0.0010
AR(12)	-0.541998	0.140528	-3.856875	0.0005
AR(24)	-0.311416	0.114858	-2.711326	0.0108

R-squared	0.927014	Mean dependent var	555.0258
Adjusted R-squared	0.917597	S.D. dependent var	93.94524
S.E. of regression	26.96788	Akaike info criterion	9.555416
Sum squared resid	22545.26	Schwarz criterion	9.775349
Log likelihood	-166.9975	F-statistic	98.43521
Durbin-Watson stat	2.178440	Prob(F-statistic)	0.000000

Inverted AR Roots	.94 -.16i	.94+.16i	.89+.33i	.89 -.33i
	.73+.61i	.73 -.61i	.61 -.73i	.61+.73i
	.33+.89i	.33 -.89i	.16+.94i	.16 -.94i
	-.16+.94i	-.16 -.94i	-.33 -.89i	-.33+.89i
	-.61+.73i	-.61 -.73i	-.73+.61i	-.73 -.61i
	-.89 -.33i	-.89+.33i	-.94 -.16i	-.94+.16i

Tabla.A.IV.74. Estimación del Modelo 24.3.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN24(-1)	-0.033618	0.022504	-1.493908	0.1456
QMENSUAL(-1)	-0.075847	0.032153	-2.358955	0.0250
QMENSUAL(-12)	1.278437	0.045667	27.99476	0.0000
C	-55.32062	21.96342	-2.518762	0.0173
AR(12)	-0.577863	0.129758	-4.453406	0.0001
AR(24)	-0.332514	0.105470	-3.152701	0.0037

R-squared	0.938337	Mean dependent var	555.0258
Adjusted R-squared	0.928060	S.D. dependent var	93.94524
S.E. of regression	25.19759	Akaike info criterion	9.442385
Sum squared resid	19047.55	Schwarz criterion	9.706305
Log likelihood	-163.9629	F-statistic	91.30377
Durbin-Watson stat	1.896506	Prob(F-statistic)	0.000000

Inverted AR Roots	.94 -.17i	.94+.17i	.90+.33i	.90 -.33i
	.73+.61i	.73 -.61i	.61 -.73i	.61+.73i
	.33+.90i	.33 -.90i	.17+.94i	.17 -.94i
	-.17 -.94i	-.17+.94i	-.33 -.90i	-.33+.90i
	-.61+.73i	-.61 -.73i	-.73+.61i	-.73 -.61i
	-.90+.33i	-.90 -.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.75. Estimación del Modelo 24.4.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN24	-0.049388	0.018155	-2.720352	0.0107
QMENSUAL(-1)	-0.086593	0.028687	-3.018541	0.0051
QMENSUAL(-12)	1.295270	0.040258	32.17459	0.0000
C	-54.88953	17.95274	-3.057446	0.0047
AR(12)	-0.595280	0.118960	-5.004034	0.0000
AR(24)	-0.341854	0.097807	-3.495200	0.0015

R-squared	0.946793	Mean dependent var	555.0258
Adjusted R-squared	0.937926	S.D. dependent var	93.94524
S.E. of regression	23.40620	Akaike info criterion	9.294891
Sum squared resid	16435.51	Schwarz criterion	9.558811
Log likelihood	-161.3080	F-statistic	106.7679
Durbin-Watson stat	2.068079	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.62i	.73 -.62i	.62 -.73i	.62+.73i
	.33+.90i	.33 -.90i	.17 -.94i	.17+.94i
	-.17 -.94i	-.17+.94i	-.33 -.90i	-.33+.90i
	-.62+.73i	-.62 -.73i	-.73+.62i	-.73 -.62i
	-.90 -.33i	-.90+.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.76. Estimación del Modelo 24. 5.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN24(-1)	-0.029411	0.022011	-1.336235	0.1919
QMENSUAL(-1)	-0.091470	0.032893	-2.780837	0.0094
QMENSUAL(-12)	1.286686	0.044574	28.86652	0.0000
QMENSUAL(-10)	-0.041309	0.026272	-1.572353	0.1267
C	-31.07391	26.60667	-1.167899	0.2524
AR(12)	-0.575469	0.125874	-4.571797	0.0001
AR(24)	-0.347242	0.102685	-3.381636	0.0021

R-squared	0.943274	Mean dependent var	555.0258
Adjusted R-squared	0.931538	S.D. dependent var	93.94524
S.E. of regression	24.58107	Akaike info criterion	9.414496
Sum squared resid	17522.65	Schwarz criterion	9.722403
Log likelihood	-162.4609	F-statistic	80.37158
Durbin-Watson stat	1.997985	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.61i	.73 -.61i	.61+.73i	.61 -.73i
	.33 -.90i	.33+.90i	.17 -.94i	.17+.94i
	-.17 -.94i	-.17+.94i	-.33 -.90i	-.33+.90i
	-.61 -.73i	-.61+.73i	-.73 -.61i	-.73+.61i
	-.90 -.33i	-.90+.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.77. Estimación del Modelo 25.1.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN25	-0.045039	0.019575	-2.300871	0.0283
QMENSUAL(-12)	1.240562	0.042682	29.06533	0.0000
C	-73.64914	19.50792	-3.775346	0.0007
AR(12)	-0.526371	0.135856	-3.874462	0.0005
AR(24)	-0.303840	0.113291	-2.681953	0.0116

R-squared	0.930784	Mean dependent var	555.0258
Adjusted R-squared	0.921852	S.D. dependent var	93.94524
S.E. of regression	26.26228	Akaike info criterion	9.502390
Sum squared resid	21380.93	Schwarz criterion	9.722324
Log likelihood	-166.0430	F-statistic	104.2177
Durbin-Watson stat	2.345230	Prob(F-statistic)	0.000000

Inverted AR Roots	.94 -.16i	.94+.16i	.89+.33i	.89 -.33i
	.73+.61i	.73 -.61i	.61+.73i	.61 -.73i
	.33 -.89i	.33+.89i	.16 -.94i	.16+.94i
	-.16 -.94i	-.16+.94i	-.33 -.89i	-.33+.89i
	-.61+.73i	-.61 -.73i	-.73+.61i	-.73 -.61i
	-.89 -.33i	-.89+.33i	-.94 -.16i	-.94+.16i

Tabla.A.IV.78. Estimación del Modelo 25.2.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN25(-1)	-0.024653	0.019926	-1.237216	0.2256
QMENSUAL(-1)	-0.077291	0.032867	-2.351655	0.0255
QMENSUAL(-12)	1.267370	0.044406	28.54073	0.0000
C	-50.31678	21.54196	-2.335757	0.0264
AR(12)	-0.563575	0.131323	-4.291501	0.0002
AR(24)	-0.326175	0.107309	-3.039590	0.0049

R-squared	0.937070	Mean dependent var	555.0258
Adjusted R-squared	0.926581	S.D. dependent var	93.94524
S.E. of regression	25.45526	Akaike info criterion	9.462734
Sum squared resid	19439.11	Schwarz criterion	9.726653
Log likelihood	-164.3292	F-statistic	89.34381
Durbin-Watson stat	1.906253	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.61i	.73 -.61i	.61 -.73i	.61+.73i
	.33+.90i	.33 -.90i	.17 -.94i	.17+.94i
	-.17 -.94i	-.17+.94i	-.33+.90i	-.33 -.90i
	-.61 -.73i	-.61+.73i	-.73+.61i	-.73 -.61i
	-.90 -.33i	-.90+.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.79. Estimación del Modelo 25.3.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN25	-0.042048	0.016984	-2.475783	0.0192
QMENSUAL(-1)	-0.083218	0.029439	-2.826748	0.0083
QMENSUAL(-12)	1.288395	0.040655	31.69058	0.0000
C	-53.87952	18.46401	-2.918083	0.0066
AR(12)	-0.583160	0.120915	-4.822915	0.0000
AR(24)	-0.333135	0.099621	-3.344036	0.0022

R-squared	0.945061	Mean dependent var	555.0258
Adjusted R-squared	0.935905	S.D. dependent var	93.94524
S.E. of regression	23.78419	Akaike info criterion	9.326931
Sum squared resid	16970.63	Schwarz criterion	9.590851
Log likelihood	-161.8848	F-statistic	103.2121
Durbin-Watson stat	2.045011	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.17i	.94 -.17i	.90 -.33i	.90+.33i
	.73+.61i	.73 -.61i	.61+.73i	.61 -.73i
	.33 -.90i	.33+.90i	.17 -.94i	.17+.94i
	-.17 -.94i	-.17+.94i	-.33 -.90i	-.33+.90i
	-.61 -.73i	-.61+.73i	-.73 -.61i	-.73+.61i
	-.90 -.33i	-.90+.33i	-.94+.17i	-.94 -.17i

Tabla.A.IV.80. Estimación del Modelo 25.4.

Dependent Variable: QMENSUAL
Method: Least Squares
Sample(adjusted): 1998:01 2000:12
Included observations: 36 after adjusting endpoints
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GEQMEN25(-1)	-0.019659	0.019659	-1.000005	0.3256
QMENSUAL(-1)	-0.093521	0.033829	-2.764553	0.0098
QMENSUAL(-10)	-0.041349	0.027018	-1.530444	0.1367
QMENSUAL(-12)	1.274973	0.043426	29.35978	0.0000
C	-25.64865	26.73339	-0.959424	0.3453
AR(12)	-0.560769	0.127758	-4.389310	0.0001
AR(24)	-0.341235	0.104770	-3.257007	0.0029

R-squared	0.941866	Mean dependent var	555.0258
Adjusted R-squared	0.929839	S.D. dependent var	93.94524
S.E. of regression	24.88420	Akaike info criterion	9.439009
Sum squared resid	17957.48	Schwarz criterion	9.746916
Log likelihood	-162.9022	F-statistic	78.30835
Durbin-Watson stat	1.995444	Prob(F-statistic)	0.000000

Inverted AR Roots	.94+.16i	.94 -.16i	.90 -.33i	.90+.33i
	.73+.61i	.73 -.61i	.61+.73i	.61 -.73i
	.33 -.90i	.33+.90i	.16 -.94i	.16+.94i
	-.16 -.94i	-.16+.94i	-.33+.90i	-.33 -.90i
	-.61 -.73i	-.61+.73i	-.73 -.61i	-.73+.61i
	-.90 -.33i	-.90+.33i	-.94+.16i	-.94 -.16i

Tabla.A.IV.81. Estadísticos de previsión de los modelos mensuales de temperaturas.

Modelos de Temperaturas	Forecast Error	MAPE	MAD	BIC	RMSE
15					
15.1.	34,37	0,04338	23,35	39,73	32,57
15.2.	35,76	0,04742	24,81	39,64	34,52
16					
16.1.	32,58	0,0417	22,42	37,65	30,87
16.2.	34,66	0,04701	24,47	38,43	33,47
17					
17.1.	30,64	0,039	20,99	35,41	29,03
17.2.	26,23	0,03385	18,16	33,36	23,54
18					
18.1.	29,24	0,037	19,99	33,79	27,7
18.2.	28,37	0,03749	20,67	33,76	26,33
Intervalo 18,5					
I18,5.1.	32,34	0,04107	21,54	35,86	31,23
I18,5.2.	24,45	0,0322	17,27	30,09	22,32
I18,5.3.	23,05	0,03036	16,45	29,31	20,68
19					
19.1.	27,87	0,03673	20,22	33,17	25,86
19.2.	21,07	0,02727	15,16	26,8	18,91
19.3.	23,57	0,03135	16,91	29	21,51
19.4.	22,12	0,02875	15,6	28,13	19,85
19.5.	23,69	0,03173	17,06	30,12	21,26
19.6.	25,26	0,03315	17,65	31,09	23,06
19.7.	27,47	0,03476	18,79	31,75	26,03
20					
20.1.	20,32	0,0292	15,98	25,84	18,24
20.2.	17,58	0,0219	12,15	22,36	15,78
20.3.	20,07	0,02716	15,14	25,52	18,01
20.4.	22	0,03027	16,49	27,07	20,08
20.5.	25,99	0,03562	19,66	30,93	24,12
20.6	23,42	0,03298	17,66	28,82	21,38
20,5					
20,5.1.	21,59	0,02996	16,42	26,57	19,71
20,5.2.	17,2	0,02333	12,93	21,87	15,44
20,5.3.	19,47	0,02751	15,2	24,75	17,47
20,5.4.	19,68	0,02698	15,05	25,03	17,66
20,5.5.	24,86	0,03552	19,67	29,58	23,07
20,5.6.	22,55	0,03203	17,23	27,75	20,58
20,5.7.	28,69	0,03871	20,6	33,16	27,18
21					
21.1.	20,01	0,0247	15,18	25,44	17,96
21.2.	18,99	0,02579	14,23	24,02	16,95
21.3.	21,52	0,02999	16,49	26,48	19,64
21.4.	20,07	0,02716	15,14	25,52	18,01
21.5.	25,99	0,03562	19,66	30,93	24,12
21.6.	22,49	0,03203	17,27	27,67	20,53
21.7.	26,92	0,03626	20,43	32,04	24,98

Modelos de Temperaturas	Forecast Error	MAPE	MAD	BIC	RMSE
21,5					
21,5.1.	21,01	0,02991	16,36	26,72	18,86
21,5.2.	21,74	0,03014	16,59	26,75	19,84
21,5.3.	20,3	0,02694	15,03	25,82	18,22
21,5.4.	24,88	0,03511	19,44	29,62	23,09
21,5.5.	26,59	0,03552	20,10	31,65	24,68
21,5.6.	23,1	0,03246	17,46	28,42	21,09
21,5.7.	29,76	0,03973	21,13	34,40	28,20
21,5.8.	21,19	0,02847	15,67	26,94	19,02
22					
22.1.	26,39	0,03492	19,81	31,41	24,49
22.2.	25,55	0,03578	19,75	30,41	23,71
22.3.	22,53	0,03211	17,48	28,65	20,22
22.4.	22,13	0,03063	16,85	27,23	20,20
22.5.	23,83	0,03311	17,78	29,32	21,75
22.6.	22,18	0.03019	16,58	28,21	19,91
22.7.	20,75	0,0272	15,17	26,39	18,63
22,5					
22,5.1.	26,3	0,0347	19,69	31,31	24,41
22,5.2.	30,13	0,03955	21,13	34,83	28,56
22,5.3.	26,12	0,03628	19,98	31,09	24,24
22,5.4.	24,38	0,03359	18,01	30,00	22,25
22,5.5.	23,44	0,03298	17,90	29,81	21,04
23					
23.1.	26,29	0,03466	19,65	31,29	24,40
23.2.	26,53	0,03659	20,13	31,57	24,62
23.3.	24,75	0,03385	18,13	30,46	22,60
23.4.	22,99	0,03179	17,43	28,29	20,98
23.5.	23,99	0,03341	18,11	30,51	21,53
23,5					
23,5.1.	26,24	0,03468	19,65	31,23	24,35
23,5.2.	32,74	0,04529	23,85	36,78	31,38
23,5.3.	26,78	0,03702	20,37	31,88	24,85
23,5.4.	25,01	0,03399	18,19	30,77	22,83
23,5.5.	23,23	0,03217	17,60	28,58	21,20
23,5.6.	24,34	0,03383	18,34	30,95	21,84
24					
24.1.	26,21	0,03494	19,77	31,19	24,32
24.2.	26,97	0,03730	20,53	32,10	25,03
24.3.	25,2	0,03425	18,33	31,01	23,00
24.4.	23,41	0,03251	17,45	28,80	21,37
24.5.	24,58	0,03408	18,48	31,26	22,06
25					
25.1.	26,26	0,03566	20,10	31,26	24,37
25.2.	25,46	0,03464	18,55	31,32	23,24
25.3.	23,78	0,03335	18,14	29,27	21,71
25.4.	24,88	0,03437	18,65	31,64	22,33

Fuente.IEA e INM.
Elaboración. Propia.

En las tablas A.IV.82. y A.IV.83. se han calculado la información muestral necesaria para efectuar el contraste.

Tabla.A.IV.82. Contraste de igualdad de pendientes. Primer Semestre.

	n	$\sum_{j=1}^{n} x_{ij}$	$\sum_{j=1}^{n} y_{ij}$	$\sum_{j=1}^{n} x_{ij}^2$	$\sum_{j=1}^{n} y_{ij}^2$	$\sum_{j=1}^{n} x_{ij} y_{ij}$
Enero	7	28	4795,816	140	3441892,436	21224,274
Febrero	7	28	4362,564	140	2781004,827	18754,198
Marzo	7	28	4158	140	2548203,466	17935,705
Abril	7	28	3374,949	140	1645989,529	14072,358
Mayo	7	28	3401,809	140	1683723,858	14495,518
Junio	7	28	3112,245	140	1415904,414	13386,195
Total	42	84	23205,383	420	8771100,729	57914,177

Tabla.A.IV.83. Contraste de igualdad de pendientes. Primer Semestre.

	n	$\bar{x}_{i*}$	$\bar{y}_{i*}$	$\sum_{j=1}^{n} x_{ij}^2 - n\bar{x}_{i*}^2$	$\sum_{j=1}^{n} y_{ij}^2 - n\bar{y}_{i*}^2$	$\sum_{j=1}^{n} x_{ij} y_{ij} - n\bar{x}_{i*}\bar{y}_{i*}$
Enero	7	4	685,1165714	28	156199,4206	2041,010
Febrero	7	4	623,2234286	28	62152,73399	1303,942
Marzo	7	4	594	28	78351,46566	1303,705
Abril	7	4	482,1355714	28	18806,56484	572,562
Mayo	7	4	485,9727143	28	30537,50515	888,282
Junio	7	4	444,6064286	28	32180,27921	937,215
Total	42	24	3315,054714	168	378227,9695	7046,716

La estimación de los coeficientes de regresión, $\hat{\beta}_{1i}$ de cada una de las seis rectas se basa en la expresión:

$$\hat{\beta}_{1i} = \frac{\sum_{j=1}^{n} x_{ij} y_{ij} - n\bar{x}_{i*}\bar{y}_{i*}}{\sum_{j=1}^{n} x_{ij}^2 - n\bar{x}_{i*}^2}$$

Los valores estimados en cada mes son:

Tabla.A.IV.84. Estimación mínimo cuadrática de los coeficientes de regresión. Primer Semestre.

Enero	$\hat{\beta}_{11}$ =72,893
Febrero	$\hat{\beta}_{12}$ =46,569
Marzo	$\hat{\beta}_{13}$ =46,561
Abril	$\hat{\beta}_{14}$ =20,449
Mayo	$\hat{\beta}_{15}$ =31,724
Junio	$\hat{\beta}_{16}$ =72,893

Y la suma de cuadrados debida a la regresión, SC_T,

$$SC_T = \sum_{i=1}^{6}\left(\sum_{j=1}^{7} y_{ij}^2 - 7\overline{y}_{i*}^2\right) - \sum_{i=1}^{3}\hat{\beta}_{1i}^2\left(\sum_{j=1}^{7} x_{ij}^2 - 7\overline{x}_{i*}^2\right)$$

Tabla.A.IV.85. Contraste de igualdad de pendientes. Primer Semestre.

	$\sum_{j=1}^{n} y_{ij}^2 - n\overline{y}_{i*}^2$	$\hat{\beta}_{1i}$	$\sum_{j=1}^{n} x_{ij}^2 - n\overline{x}_{i*}^2$	$\hat{\beta}_{1i}^2(\sum_{j=1}^{n} x_{ij}^2 - n\overline{x}_{i*}^2)$
Enero	156199,421	72,893	28	148775,779
Febrero	62152,734	46,569	28	60723,741
Marzo	78351,466	46,561	28	60701,669
Abril	18806,565	20,449	28	11708,116
Mayo	30537,505	31,724	28	28180,175
Junio	32180,279	33,472	28	31370,427
Total	378227,969	251,668	168	341459,907

resultando la suma de cuadrados

SC_T=378227,969-341459,907=36768,062

El coeficiente de regresión común β_1 bajo la hipótesis nula tiene por expresión:

$$\beta_1^* = \frac{\sum_{i=1}^{6}\left(\sum_{j=1}^{7} x_{ij} y_{ij} - 7\overline{x}_{i*}\overline{y}_{i*}\right)}{\sum_{i=1}^{6}\left(\sum_{j=1}^{7} x_{ij}^2 - 7\overline{x}_{i*}^2\right)}$$

y su estimación es $\hat{\beta}_1^*$ =41,94

La suma de cuadrados SC_{H_1} es

$$SC_{H_1} = \sum_{i=1}^{6}\left(\sum_{j=1}^{7} y_{ij}^2 - 7\overline{y}_{i*}^2\right) - \beta_1^{*2}\sum_{i=1}^{6}\left(\sum_{j=1}^{7} x_{ij}^2 - 7\overline{x}_{i*}^2\right) = 378227{,}96 - 41{,}94^2 \cdot 168 = 45887{,}25$$

y la diferencia entre las dos sumas de los cuadrados

$$SC_{H_1} - SC_T = 45887{,}25 - 36768{,}06 = 9119{,}18$$

El valor del estadístico F necesario para efectuar el contraste de hipótesis

$$F = \frac{\dfrac{SC_{H_1} - SC_T}{k-1}}{\dfrac{SC_T}{N-2k}} = \frac{\dfrac{9119{,}18}{5}}{\dfrac{36768{,}06}{30}} = 1{,}48$$

El valor crítico de $F_{5,\ 30,\ 0{,}05}$ es 2,5335 mayor que el muestral 1,48, lo que indica que no existe evidencia para rechazar la hipótesis nula de igualdad de pendientes al 95% de confianza.

Los cálculos realizados para el contraste de igualdad de pendientes en el segundo semestre del año se presentan a continuación:

Tabla.A.IV.86. Contraste de igualdad de pendientes. Segundo Semestre.

	n	$\sum_{j=1}^{n} x_{ij}$	$\sum_{j=1}^{n} y_{ij}$	$\sum_{j=1}^{n} x_{ij}^2$	$\sum_{j=1}^{n} y_{ij}^2$	$\sum_{j=1}^{n} x_{ij} y_{ij}$
Julio	7	28	3388,821	140	1678128,158	14532,773
Agosto	7	28	3601,460	140	1885858,667	15320,243
Septiembre	7	28	3749,413	140	2033674,49	15743,032
Octubre	7	28	3405,679	140	1712206,511	14778,339
Noviembre	7	28	3484,840	140	1772329,423	14960,407
Diciembre	7	28	3320,105	140	1606357,468	13984,221
Total	42	84	20950,318	420	5597661,315	45596,048

Tabla.A.IV.87. Contraste de igualdad de pendientes. Segundo Semestre.

	n	$\bar{x}_{i*}$	$\bar{y}_{i*}$	$\sum_{j=1}^{n} x_{ij}^2 - n\bar{x}_{i*}^2$	$\sum_{j=1}^{n} y_{ij}^2 - n\bar{y}_{i*}^2$	$\sum_{j=1}^{n} x_{ij} y_{ij} - n\bar{x}_{i*}\bar{y}_{i*}$
Julio	7	4	484,117	28	37541,333	977,489
Agosto	7	4	514,494	28	32928,077	914,403
Septiembre	7	4	535,630	28	25374,798	745,38
Octubre	7	4	486,526	28	55256,590	1155,623
Noviembre	7	4	497,834	28	37456,590	1021,047
Diciembre	7	4	474,301	28	31629,295	703,801
Total	42	24	2992,903	168	220186,683	5517,743

Tabla. A.IV.88.Estimación mínimo cuadrática de los coeficientes de regresión. Segundo Semestre.

Julio	$\hat{\beta}_{11}$ =34,910
Agosto	$\hat{\beta}_{12}$ =32,657
Septiembre	$\hat{\beta}_{13}$ =26,620
Octubre	$\hat{\beta}_{14}$ =41,272
Noviembre	$\hat{\beta}_{15}$ =36,465
Diciembre	$\hat{\beta}_{16}$ =25,135

Tabla.A.IV.89. Contraste de igualdad de pendientes. Segundo Semestre.

	$\sum_{j=1}^{n} y_{ij}^2 - n\bar{y}_{i*}^2$	$\hat{\beta}_{1i}$	$\sum_{j=1}^{n} x_{ij}^2 - n\bar{x}_{i*}^2$	$\hat{\beta}_{1i}^2(\sum_{j=1}^{n} x_{ij}^2 - n\bar{x}_{i*}^2)$
Julio	37541,333	34,910	28	34124,455
Agosto	32928,077	32,657	28	29861,887
Septiembre	25374,798	26,621	28	19842,548
Octubre	55256,590	41,272	28	47695,161
Noviembre	37456,590	36,466	28	37233,463
Diciembre	31629,295	25,136	28	17690,566
Total	220186,683		168	186448,081

SC_T 33738,60146

β_1^* = 32,84370833

SC_{H_1} = 38963,54106

$SC_{H_1} - SC_T$ = 5224,939602

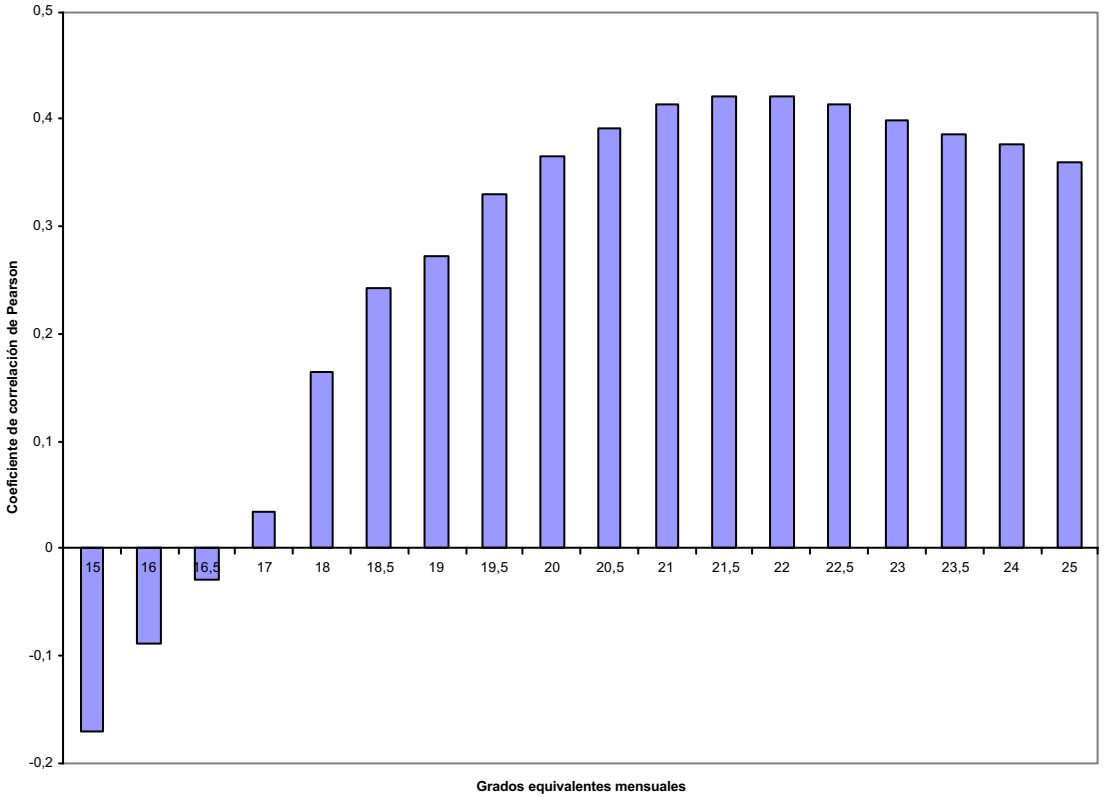
de donde el estadístico F=0,9291919<2,5335, luego se puede decir que no existe evidencia muestral, al 5% de significación, de que las rectas no sean paralelas, se puede concluir que su variación es constante a lo largo del periodo analizado.

Tabla.A,IV.90. Correlaciones bivariadas de Pearson.Demanda-Grados equivalentes.

Correlaciones		15	16	16,5	17	18	18,5	19	19,5	20	20,5	21	21,5	22	22,5	23	23,5	24	25
Demanda de electricidad	Correlación de Pearson	-0,170	-0,088	-0,029	0,034	0,165	0,242	0,273	0,330	0,365	0,391	0,413	0,422	0,422	0,413	0,399	0,386	0,377	0,360
	Sig. (bilateral)	0,123	0,427	0,797	0,761	0,134	0,027	0,012	0,002	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001
	N	84	84	84	84	84	84	84	84	84	84	84	84	84	84	84	84	84	84

Fuente. IEA e INM.
Elaboración. Propia.

BIBLIOTECA VIRTUAL



REFERENCIAS BIBLIOGRÁFICAS

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

Bibliografía General

Alvarez, E., (1997); “Economía Industrial del Sector Eléctrico: Estructura y Regulación”. Civitas.

Armstrong, J.S., (2001); "Principles of Forecasting", Kluwer.

Aznar,A., (1998); "Econometric Model Selection: A new Approach", Kluwer.

Banker, R.D., (1984); “Estimating most productive scale size using Data Envelopment Analysis”, European Journal of Operational Research, nº 17, 35-44.

Banker, R.D., Charnes, A. y Cooper, W.W., (1984); “Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis”, Management Science, Vol.30, nº 9, 1078-1092.

Bates, J.M. y Granger, C.W.J., (1969); ”The Combinations of Forecasts”, Operational Research Quarterly,20,4,451-468.

Bessler, D.A. y Brandt J.A., (1981); “Composite Forecasting: An Application with U.S. Hog Prices” , American Journal of Agricultural Economics,63,135-140,1981.

Boussofiane, A., Dyson, R.G. y Thanassoulis, E., (1991); “Applied Data Envelopment Analysis”, European Journal of Operational Research, nº 52, 1-15.

Box, G.E.P. y Jenkins, G.M., (1976); "Time Series Analysis. Forecasting and Control" Holden Day.

Box, G.E.P. y Jenkins. G.M., (1976); "Time Series Analysis: Forecasting and Control", Holden Day .

Box, G.E.P. y Tiao, G.C., (1985); "Intervention Analysis with Applications to Economic and Environmental Problems", Journal of the American Statistical Association,70, 70-79.

Cleary, J.P. y Levenbach, H., (1981); "The Beginning Forecaster" Lifetime Learning Publications.

Cleary, J.P. y Levenbach, H., (1982); "The Professional Forecaster" Lifetime Learning Publications.

Clemen,R., (1989); “Combining Forecasts: a review and annotated bibliography”, International. Journal of Forecasting, 5, 559-584.

Coccari, R. L. y Galucci, C., (1984); "Average Two best Forecasts Can Reduce Forecasting Risk", Jornal of Business Forecasting, Fall.

Collopy, F. y Armstronng, J.S., (1992); “Expert Opinions about Extrapolation and the Mystery of the Overlooked Discontinuities”, International Journal of Forecasting, 8, 4, 575-582.

Cooper, W.W., Seiford, L.M. y Tone, K., (2000); “Data envelopment analysis : a comprehensive text with models, applications, references and DEA-solver software” . Boston [etc.] Kluwer Academic Publishers.

Cooper, W.W., Thompson, R.G. y Thrall, R.M. (1996); “Extensions and New Developments in DEA”, Annals of Operations Research, nº 66, 3-45.

Charnes, A. y Zlobec, S., (1983); “Stability of Efficiency Evaluations in Data Envelopment Analysis”, Zeitschrift für Operations Research, nº 33.

Charnes, A., Cooper, W.W. y Rhodes, E. (1978); “Measuring the efficiency of decision making”, European Journal of Operational Research, nº 2, 429-444.

Charnes, A., Cooper, W.W., Golany, B. y Seiford, L., (1985); “ Foundations of Data Envelopment Analysis for Pareto-Koopmans efficient empirical production functions”, Journal of Econometrics, nº 30, 91-107.

Chatfield, C. y Yar, M., (1991); "Prediction Intervals for the Holt-Winters Forecasting Procedure", International Journal of Forecasting, 6, 1, 127-137.

Dias-Bandaranaike,R. y Munasinghe, M, (1983); "The Demand for Electricity Services and the Quality of Supply",4, 2,49-71.

Diaz,M. y Llorente, M., (2000); "Econometría", Ediciones Pirámide.

Diebold, F.X., (1998); “Elements of Forecasting", South Western College Publishing Co.

Farrell, M.J. (1957); “ The measurement of productive efficiency”, Journal of the Royal Statistical Society, Series A, nº 120, 253-282.

Fraser, H. (2001); “The Importance of an Active Demand Side in the Electricity Industry”, The Electricity Journal, november 2001.

Gardner, E.S., (1985); "Exponential Smoothing: The State of the Art". Journal of Forecasting, 4, 1-38.

George,E.L., (2000); "The Variable Selection Problem", Journal of the American Statistical Association, 95, 1304-1308.

Ghysels,E. y Osborn, D.R, (2000); "The Econometric Analysis of Seasonal Time Series", Cambridge University Press.

Golany, B. y Roll, Y., (1989); “An Aplication Procedure for DEA”, International Journal of Management Science, vol 17, nº 3, 237-250.

Goodrich, R.L., (1990); "Applied Forecasting Systems", Business Forecasting Systems.

Granger, C.W.J., (1980); "Forecasting in Business and Economics", Academic Press.

Graybill, F.A., (1994); “Regression Analysis: Concepts and Applications”, Duxbury Press.

Harvey, A.C., (1981); “The Econometric Analysis of Time Series”, John Wiley & Sons, Inc.

Homburg, C., (2000); “Using Data Envelopment Analysis to Benchmark Activities”, International Journal of Production Economics, nº 73, 51-58.

Iqbal, A., Lerme, C. y Seiford, L., (1995); “Components of efficiency evaluation in data envelopment analysis”, European Journal of Operational Research, nº 80, 462-473.

Jain,C.L y Colasante, J., (1985); "How much do we learn from *ex-post* forecasts?, Journal of Business Forecasting, Fall.

Jain,C.L., (2001-2002); "How much data should we use to prepare forecasts?, Journal of Business Forecasting, Winter.

Levenbach, H. y Cleary, J.P., (1984); "The Modern Forecaster", Lifetime Learning Publications.

Mabert, V.A. y Radcliffe, R.C., (1974); "A Systematic Modelling Methodology", Financial Management, Autumn, 59-67.

Makridakis, S. y Hibon, M., (1989); "Exponential Smoothing: the Effect of initial Values and Loss Function on Post-sample Forecasting Accuracy", Ninth International Symposium on Forecasting, Vancouver.

Makridakis, S., (1982); "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition, Journal of Forecasting, 2, 111-153.

Makridakis, S., (1984); "The Forecasting Accuracy of Major Time Series Methods", Wiley.

Makridakis,S., Wheelright,S.C. y Hyndman, R.J., (1998); Forecasting: Methods and Applications, Wiley.

McQuarrie, A.D.M. y Tsay, C.L., (1998); "Regression and Time Series Model Selection", World Scientific.

Miller, A., (1990); "Subset Selection in Regression", Chapman and Hall.

Montgomery, D.C., Peck, E. A. y Vining, G.G., (2001); "Introduction to Linear Regression Analysis", Wiley.

Nelson, C.R., (1984); " A Benchmark for the Accuracy of Econometric Forecasts of GNP", Business Economics,19,3, 52-58.

Pastor, J.T. (1996); " Translation Invariance in Data Envelopment Analysis: A Generalition" , Annals of Operations Research, nº 66, 93-102.

Pena, J.B. y Estavillo, J.A., (1999); "Cien Ejercicios de Econometría", Pirámide.

Pindyck R. S., (1980) ; "The Structure of World Energy Demand"; The MIT Press.

Pindyck, R.S. y Rubinfeld, D. L., (1998); "Econometría: modelos y pronósticos", 4ª edición, McGraw-Hill.

Pulido, A. y Pérez, J., (2001); "Modelos econométricos", Pirámide.

Stine, R. A., (1985); "Boostrap prediction intervals for regression", Journal of the American Journal Association, 82, 1072-1078.

Stoll,H.G.,1989; "Least-Cost Electric Utility Planning", Wiley Interscience.

Stulajter, F., (2002); "Predictions in Time Series Using Regression Models", Springer.

Tolley ,G.S., Upton, C.W. y Hasting, H.V., (1987); "Electric Energy Availability and Regional Growth": Ballinger Publishing Company.

Tukey, J. W., (1987); "Kinds of bootstrapping and kinds of jacknives, discussed in terms of a year of weather related data" (Tech. Rep. no. 292) Princeton, N. J. Princeton University Press, Department of Statistics.

Uriel, E., (1988); " Estadística Económica y Empresarial". Ed. AC.

Wilson,J.H. y Keating, J.H., (1998); "Business Forecasting", Mc Graw-Hill.

Yu, G., Wei, Q. y Brockett, P., 1996); "A Generalized Data Envelopment Analysis Model: A Unification and Extension of Existing Methods for Efficiency Analysis of Decision Making Units", Annals of Operations Research, nº 66, 47-89.

Bibliografía específica

Abraham, B. y Ledolter, J., (1983); “Statistical Methods for Forecasting”. Wiley-Interscience.

Adams, G. Allen, P.G. y Morzuch,.J., (1991) "Probaility Distributions of Short-Term Electricity Peak Load Forecasts", International Journal of Forecasting, 7, 3, 283-298.

Adler, N. y Berechman, J., (2001); “ Measuring Airport Quality from the Airlines’ Viewpoint: an Application of Data Envelopment Analysis”, Transport Policy, nº 8, 171-181.

Adler, N. y Golany, B., (2000); “ Evaluation of Deregulated Airline Networks using Data Envelopment Analysis combined with Principal Component Analysis with an application to Western Europe”, European Journal of Operational Research, nº 132, 260-273.

Agrell, P.J., (1998); “Efficiency Measurement and Benchmarking of Nordic Electricity Distributors”, Linköping Institute of Technology. Department of Production Economics. Working Paper WP-246.

Al-Garni, Ahmed Z., Zubaid, S.M. y Nizami, J.S., (1988); "A Regression Model for Electric-Energy-Consumption Forecasting in Eastern Saudi Arabia", Energy,19(10), 1043-1049.

Ali, A.I., (1995); “Components of Efficiency Evaluation in Data Envelopment Analysis”, European Journal of Operational Research, nº 80, 462-473.

Al-Zayer, J. y Al-Ibrahim, A., (1996); "Modelling the Impact of Temperature on Electricity Consumption in the Eastern Province of Saudi Arabia" , Journal of Forecasting, Vol 15, 97-106.

Anderson, T.R, Hollingsworth, K. y Inman, L., (2001); “ Assesing the Rate of Change in the Enterprise Database System Market Over Time Using DEA”, Management of Engineering and Technology, vol 2, 384-390.

Arocena, P., (2001); “Eficiencia Pública y Privada en la Producción de Energía Eléctrica en España”, Revista Europea de Dirección y Economía de la Empresa, vol 10, nº 1, 9-20.

Atares, M.L., (2002); “Liberalización *a medio gas*”. Periódico El Mundo. 2 de junio.

Ayyash, S., Salman, M. y Al-Hafi, N., (1985); "Modelling the impact of temperarture on summer electricity consumption in Kuwait", Energy-The International Journal, 10, no. 8, 941-949.

Barakat, E., Al-Qassim, J. y Al-Rashed, S., (1992); "Long Range Peak Demand Forecasting Under Conditions of High Growth", IEEE Transactions and Power Systems, 7, 4, 1483-1487.

Barakat, E., y Al-Rashed, S., (1993); "New Model for Peak Demand Forecasting Applied to Highly Complex Characteristics of a Fast Developing Area", IEEE, 139, 16-140.

Bartels, R. y Fiebig, D.G., (1996); "Metering and Modelling Residential End-Use Electricity Load Curves", Journal of Forecasting, 15, 415-426.

Baum, V., (1998); "Glowing in the Gloom", Petroleum Economist, 60, 6-9.

Beauvais, E.C. y Spann, R.M., (1977); "Econometric Estimation of Peak Electricity Demand", Forecasting and Modeling Time-of-Day and Seasonal Electricity Demands", EPRI, 3-23.

Bolzern, P., Fronza, G. y Brusasca, G., (1982); "Temperature effects on the winter daily electric load", Journal of Applied Meteorology, 21, 2.

Brockwell, P.J. y Davis, R.A., (2002); " Introduction to Time Series and Forecasting", Wiley.

Bunn,D.W. y Farmer, E.D., (1985); "Comparative Models for Electrical Load Forecasting", Springer-Verlag.

Byrnes, P., Färe, R. y Grosskopf, S., (1984); “Measuring Productive Efficiency: an Application to Illinois Strip Mines”, Management Science, vol 30, nº 6, 671-682.

Cao, R., Febrero-Bande, M., Gonzalez-Manteiga, W., Prada-Sánchez, J.M. y García-Jurado, I., (1997); "Saving Computer Time in Constructing Consistent Bootstrap Prediction Intervals for Autoregresive Processes", Communications in Statistics.-Simulation and Computation, 26, 991-978.

Cao,R., (1999); "An overview of bootstrap methods for estimating and predicting in time series", Test,8,1, 95-116.

Cargill, T.E. y Meyer, R.A., (1971); "Estimating the Demand for Electricity by Time of Day", Applied Economics, 3, 233-246.

Cowing, T.G. y Mc Fadden, (1994); "Microeconomic Modelling and Policy Analysis:Studies in Residential Energy Demand", Jovanovich Publishers.

Crow,R.T.,Robinson,M. y Squiteri,R.L., (1991); "Forecasting Electricity Sales and Loads: A Handbook for Small Utilities", Academic Press.

Chen, K.C. y Winters, P.R., (1974); “Forecasting Peak Demand for an Electric Utility with Hybrid exponential model”, Management Science, 20.

Davidson, A.C. y Hinkley,D.V., (1997); "Boostrap Methods and their Applications", Cambridge University Press.

Davies, M., (1959); "The Relationship between Weather and Electricity Demand", Proc.IEEE, 106C, 27-37.

Diaconis, P. y Efron, B., (1983); "Computer-intensive Methods in Statistics", Scientific American, 248,5, 116-130.

Diebold,F.X., (1999); "Elementos de Pronósticos", International Thomson Editores.

Dryar, H.A., (1944):"The effect of weather on the system load". Trans.AIEE, 63,1006-1013.

Efron, B. y Tibshirani, R.J., (1993);"An Introduction to te Bootstrap", Chapman and Hall.

Efron, B., (1979); "Bootstrap Methods : Another Look at the jacknife", Annals of Statistics, 7-126.

Efron, B., (1988);"Bootstrap confidence intervals: good or bad?", Psychological Bulletin, 104, 293-296.

Efron, B., (1989); " Better bootstrap confidence intervals", Journal of the American Statistical Association, 82, 171-200.

Einhorn, M., (1980); "Review of the Literature on Economic Analysis of the Time-of-Day Demand for Electricity", National Economic Research Associates, Inc.

Energía 2002. Foro de la Industria Nuclear Española.

Engle, R. , Granger, C.W.J. y Ramanathan, R., (1981); "Regional Load Curve Models: Long Run Forecasts and Sensitivity Analisys of the Query Model", EPRI.

Engle, R. , Granger, C.W.J. Mitchem, A. y Ramanathan, R., (1979); "Some Problems in The Estimation of Daily Load Shapes and Peaks", EPRI.

Engle, R.F., Granger, C.W., Rice, J. y Weiss, A., (1986); "Semiparametric Estimates of the Relation Between Weather and Electricity Sales", Journal of the American Statistical Association, 81, 310-320.

Engle, R.F., Mustafa, C. y Rice, J., (1992); " Modelling Peak Electricity Demand", Journal of Forecasting, 11, 241-251.

Engle, R.F., Scott, J.B. y Stern, G., (1988); “A Comparison of Adaptive Structural Forecasting Methods for Electricity Sales”, Journal of Forecasting, 7, 149-172.

Engle,R.F., Brown, S. J., y Sterm, G., (1988); "A Comparison of Adaptive Structural Forecasting Methods fo Electricity Sales", Journal of Forecasting, 7, 3, 149-172.

Faruqui, A., (1987), "On the Search for Accuracy in Electric Utility Forecasting", Journal of Forecasting, April, 93-96.

Faruqui, A., Melendy, C. y Bloom,J., (1994-95); "The Changing Role of Forecasting in Electric Utilities", The Journal of Business Forecasting, Winter .

Faruqui, A., y Seiden,K., (1990); "Estimating Efficiency Savings Enbeddeding Electric Utility Forecasts", EPRI.

Ferretti,N. y Romo, J.,(1996);"Unit root bootstrap test for AR(1) models", Biometrika, 83, 849-860.

Fischler, E.B. y Nelson ,R.F., (1986), "Integrating Time Series and End Use Methods of Forecasting Electricity Sales", Journal of Forecasting, January ,15-30.

Forsund, F.R. y Kittelsen A.C., (1998); “Productivity Development of Norwegian Electricity Distribution Utilities”, Resources and Energy Economics, nº 20, 207-224.

Fredman, D.A. y Peters,S.C., (1984); "Bootstraping a regresion equation: some empirical results", JASA, 79, 385, 97-106.

Fukuyama, H. y Weber, W.L., (2002); “ Estimating Output Allocative Efficiency and Productivity Change: Application to Japanese Banks”, European Journal of Operational Research, nº 137, 177-190.

González, G., y Soto, F., (2001); “Cobertura de la Demanda de energía eléctrica a medio y largo plazo”. Energuía.

González, J., (2002); “Liberalizar con precios regulados”, Periódico ABC. 29 de diciembre.

Halvorsen, R., (1975) ; "Residential Demand for Electric Energy", Review of Economics and Statistics, 57,12-18.

Halvorsen, R., (1975); "Residential Demand for Electric Energy", Review of Economics and Statistics, 57,12-18.

Hayter, A.J., (1995); "Probability and Statistics for Engineers and Scientists”, PWS Publishing Co.

Hjalmarsson, L. y Veiderpass, A., (1992); “Productivity in Swedish Electricity Retail Distribution”, Scandinavian Journal of Economics, nº 94 ,supplement, 193-205.

Holbert, D. Y Son, M.S. (1986); "Bootstraping a time series method: some empirical results", Communications in Statistics.-Theory and Methods, 15,12, 36669.3691.

Informe Marco sobre la Demanda de Energía Eléctrica y Gas Natural, y su Cobertura, (2001). CNE.

Jaen,M. y López,E., (2001); "Modelos Econométricos de Series Temporales", Septem Ediciones.

Kao, C. y Yang, Y.C., (1992); " Reorganization of Forest Districts Via Efficiency Measurement", European Journal of Operational Research, nº 58, 356-362.

Kher,L.K., Sioshanshi, F.P. y Soroosshian, S., (1987); "Energy Demand Modelling with Noisy Input -Output Variables", The Energy Journal,8(4), 33-49.

Kuosmanen, T. y Thierry, P., (2001); "Measuring Economic with Incomplete Information: with an application to European commercial banks", European Journal of Operational Research.

Labys, W.C., (1999); "Modelling Mineral and Energy Markets", Kluwer Academic Publishers.

Larrán, M., (1999); "Aplicación del Data Envelopment Analysis (DEA) a la Evaluación de la Eficiencia en los Hospitales Públicos Andaluces", Instituto de Estadística de Andalucía.

Le Comte, D.M. y Warren, H.E., (1981); "Modelling the Impact of Summer Temperatures on National Electricity Consumption", Center for Environmental Assessment Services, Washington.

Le, S.V., (1993); "An Econometric Approach to Peak Load Forecasting", Proceedings of Electric Load Forecasting and Generating Capacity Expansion Symposium, Columbus, Ohio.

Lehman, R.L. y Warren, H.E., (1978); "Residential natural gas consumption: evidence that conservation efforts to date have failed", Science, 199, 879-882.

Lo, Feng-Yu, Chien Chen-Fu y Lin, J.T., (2001); "A DEA study to evaluate the relative efficiency and investigate the district reorganization of the Taiwan Power Company", IEE Transactions on Power Systems, vol 16, nº 1, 170-178.

Melendy, C., (1992); "Integrated Resource Planning for the 1990s", Proceedings of the Conference of Canadian Electric Association, Quebec, 78-89.

Mendenhall, W. y Sincich, T., (1993); "A Second Course in Business Statistics:regression analysis", Macmillan Publishing Co.

Miliotis, A.P., (1992); "Data Envelopment Analysis Applied to Electricity Distribution Districs", Journal Operational Research Society, vol.43, nº 5, 549-555.

Morales, J., (2003); "La liberalización del Mercado Eléctrico sigue *a dos velas*", Revista Mercado de Dinero. Enero.

Mount, T.D., Chapman,L.D. y Tyrrell,T.J., (1975); "Electricity Demand in the United States: An Econometric Analysis· , Oak Ridge National Laoratory, Tennesee, U.S.A.

Pankratz, A., (1991); "Forecasting with Dynamic Regression Models", Wiley-Interscience Publication.

Panuska,V., (1977); "Short -Term Forecasting of Electricity Power System Load from a Weather-Dependent Model" , IFAC Symposium, Melbourne, Febrero.

Parti, R., (1980); "The Total and Applied Specific Condition Demand for Electicity in the Household Sector", The Bell Jornal of Economics, 11, 309-321.

Pérez de Lucía, A., (2002); “Liberalización y Medio Ambiente definen el cambio: Shock Energético”, Revista Dinero. Marzo.

Pina, V. y Torres, L., (2000); “Análisis of the Efficiency of the Local Government Services Delivery: An Application to Urban Public Transport”, Transportation research, nº 35, pp. 929-944.

Platt, H. D., (1981); "Regional Load Curve Models: Specification and Estimation of the DRI Model", EPRI.

Platt, H. D., (1982); "Regional Load Curve Models: Long Run Forecasts and Sensitivity Analysis of the DRI Model, EPRI.

Plourde, A. y Ryan, D., (1985); "On the Use of Double-Log Forms in Energy Demands Analysis", The Energy Journal, 6, 4, 105-113.

Quayle,R.G. y Diaz,H.F., (1980); "Heating degrees day data applied to residential heating energy consumption, J. Applied Meteor,19,241-248.

R¹czka, J., (2001); “Explaining the Perfomance of Heat Plants in Poland”, Energy Economics, nº 23, 355-370.

Ramanathan, R., y Mitchem, A., (1981);"Economic and Computational Issues in Estimating the Demand for Energy by Time-of -Day" American Statistical Association, Proceedings of the Business and Economic Statistics Section.

Ramesh, S., Natarajan, B. y Bhagat, G., (1988); "Peak Load Prediction Using Weather Variables".Energy-the International Journal,13, no.8, 671-670.

Resende, M., (2001); “Relative Efficiency Measurement and Prospects for Yardstick Competition in Brazilian Electricity Ditribution”, Energy Policy.

Rodseth, A. y Strom, S., (1986); "The Demand for Energy in Norwegian Households with Special Emphasis on the Demand for Electricity", Institute of Economics, University of Oslo.

Scully, G.W., (1998); “Reform and Efficiency Gains in the New Zealand Electrical Supply Industry”, Journal of Productivity Analysis, nº 11, 133-147.

Stanton, K.N. y Gupta, P.C., (1969); "Forecasting Annual or Seasonal Peak Demand in Electric Utility Systems" IEEE paper 69 TP 648.PWR, Summer Power Meeting, 22-27 junio.

Stanton, K.N. y Gupta, P.C., (1980); "Forecasting Annual or Seasonal Peak Demand in Electric Utility Systems", IEEE Transactions on Power Apparatus and Systems, 89, 951-959.

Stine,W., (1995); "Bootstrap of a linear Models with AR-error structure", Metrika, 42, 395-410.

Taylor, L.D., (1975); "The Demand for Electricity: A Survey", Bell Journal of Economics, 74-110, Spring.

Taylor, L.D., (1975); "The demand for electricity:a survey", Bell J. Econ. 6,74.

Taylor, L.D., (1976); "A Review of Load Forecasting Methodologies in the Electric Utility Industry", EPRI.

Thompson, R.G., (1996); “DEA Multiplier Analytic Center Sensivity with an Illustrate Application to Independent Oil Companies”, Annals of Operational Reserach, nº 66, 163-177.

Thompson, R.P., (1976); "Weather Sensitive Electric Demand and Energy Analysis on a Large Geographically Diverse Power System - Application to short term hourly electric demand forecasting", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-95, no. 1, January/February.

Tolley, G.S, Upton, C.W. y Hastings, V.S., (1977); “Electric Energy Availability and Regional Growth”, Ballinger Publishing Co.

Train, K. e Ignelzi, P., (1984); “The Billing Cicle and Weather Variables in Models of Electricity Sales”, Energy, 9, 11-12, 1041-1047.

Uri, N.D., (2000); “Changing Productive Efficiency in Telecommunications in the United States”, nº 72, 121-137.

Williams y Leslie,(1953); "Load Forecasts Reflect Weather, Electr. World,139, 86-88.

Zhang, Y. y Bartels., (1998); “The Effect of Sample Size on the Mean Efficiency in DEA with an application to Electricity Distribution in Australia, Sweden and New Zealand”, Journal of Productivity Analysis, nº 9, 187-204.

Fuentes Documentales

Boletines Estadísticos Mensuales de Energía Eléctrica. REE.

Informes Anuales de las compañías Endesa, Hidrocantábrico, Iberdrola y Unión Fenosa. Años 1998, 1999, 2000 y 2001.

Informes Anuales de la Compañía Sevillana de Electricidad, S.A., años: 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1994, 1995, 1996, 1997, 1998.

Temperaturas Mínimas, Máximas y Medias Mensuales por provincia de la Comunidad Autónoma de Andalucía, años 1995-2001. Centro Meteorológico Territorial en Andalucía Occidental y Ceuta. Instituto Nacional de Meteorología. Ministerio de Obras Públicas, Transportes y Medio Ambiente.

BIBLIOTECA VIRTUAL MIGUEL DE CERVANTES

Páginas web

Administración de la Información de la Energía en los Estados Unidos:

http://eia.doe.gov/

Agencia de la Energía de Sevilla: http://www.agencia-energía-sevilla.com/

Agencia Internacional de la Energía: http://iea.org/

Asociación Española de la Industria Eléctrica: http:\ \ www.unesa.es

Comisión Nacional de la Energía: http://www.cne.es/

Enervía: http://www.enervía.com/

Grupo Europeo de Empresas y Organizaciones Distribuidoras de Energía: http:\ \ www.geode-eu.org

Instituto de Estadística de Andalucía: http://www.iea.junta-andalucía.es/

La Guía de la Energía: http:\ \ www.energuia.com

Mercado de Electricidad: http:\ \ www.mercadoelectrico.comel.es

Ministerio de Energía: http://www.mineco.es/PlanificaciónEnergética2002-2011/

Mundoenergía: http://www.mundoenergía.com/

Programa EMS DEA: http:\ \ www.wiso.uni-dormund.de/lsfg/or/scheel/ems

Red Eléctrica Española: http:\ \ www.ree.es

Sociedad para el Desarrollo Energético de Andalucía: http:\ \ www.sodean.es

BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

Autora: **ESTER GUTIÉRREZ MOYA**

Título Tesis: **LA DEMANDA RESIDENCIAL DE ENERGÍA ELÉCTRICA EN LA COMUNIDAD AUTÓNOMA DE ANDALUCIA: UN ANÁLISIS CUANTITATIVO**

Resumen

El consumo de energía eléctrica es una característica relevante de las sociedades industrializadas representando un factor básico para la producción en diversos sectores, entre los que se encuentran el industrial, el comercial y el residencial. En este trabajo se especifican y validan modelos econométricos, univariantes y de regresión dinámica (basados en la temperatura ambiental) que contribuyen a explican la demanda de energía eléctrica doméstica en la Comunidad Autonómica de Andalucía (C.A.A.) basados en los datos anuales y mensuales respectivamente. Como complemento a la estimación y previsión de la demanda se han desarrollado y aplicado técnicas de remuestreo de carácter no paramétrico. Finalmente, se analiza la eficiencia técnica, a través del análisis por envoltura de datos, de las principales compañías eléctricas españolas de ciclo completo, las cuales tuvieron que hacer frente al nuevo entorno liberalizador que tuvo lugar a partir de la ley del Sector Eléctrico en 1998.

CDU: 620. Economía de la Energía
33. Estadística
31. Economía Ciencia Económica.

UNESCO: 1209.14
5302.02

DISSERTATION: **RESIDENTIAL ELECTRICITY DEMAND IN THE AUTONOMIC COMMUNITY OF ANDALUSIA: A QUANTITATIVE ANALYSIS**

Abstract

The consumption of electrical energy is an important characteristic of the industrialized societies representing a basic factor for the production in diverse sectors, between which they are the industrialist, commercial and the residential one. In this dissertation econometric, univariant and dynamic regression (based on the temperature) models are specified and validated that contributes to explain the demand of domestic electrical energy in the Autonomic Community of Andalusia based on the annual and monthly data respectively. As complement to the estimation and forecast of the demand has been developed and applied bootstraping process. Finally, it analyses the technical efficiency of the main electrical companies of complete cycle in Spain through the data envelopment analysis, in a period of liberation of the electricity market.