

Universidad de la República
Facultad de Ciencias Económicas y de Administración
Licenciatura en Estadística
Tutor del Trabajo: Ec. Laura Nalbarte

Estimación de la Demanda de Combustibles

MAGELA NEGRO



UNIVERSIDAD
DE LA REPUBLICA
URUGUAY

Montevideo, Octubre 2008

Agradecimientos

Quisiera agradecer al personal de ANCAP; muy especialmente a la *Gerencia de Comercialización* y al área de *Planificación Comercial*. La Gerente del área Cra. Susana Puga y la jefa de la sección: Cra. Miriam Ferreyra, han permitido que pudiera desarrollarme como profesional conjugando mis tareas diarias con mi carrera universitaria. A su vez, en todo momento recibí el apoyo necesario para llevar a cabo el presente trabajo siendo alentada en cada momento a seguir adelante. También agradezco a mis compañeras de trabajo por transitar junto a mi este camino.

Agradezco al *Instituto de Estadística de la Facultad de Ciencias Económicas y de Administración*, casa en la cual di mis primeros pasos como ayudante de investigación de la mano del profesor Ramón Alvarez. Gracias a todo el personal docente y muy especialmente a la Ec. Laura Nalbarte, tutora del presente trabajo.

A mis amigos y a Rodrigo mi compañero, por su paciencia.

A mi madre y padre, quienes me forjaron como persona y me enseñaron lo que es el esfuerzo y la constancia.

Índice de figuras

5.1. Serie de demanda de consumo de gas oil en Uruguay para el período Enero de 1999 a Junio de 2007	30
5.2. Consumo de Gas Oil por año. Período: Enero de 1999 a Diciembre de 2007	31
5.3. Salida correspondiente al Test de Dickey Fuller sobre la serie de Consumo de Gas Oil. Período Enero de 1999 a Diciembre de 2007	32
5.4. Gráfica de la serie de gas oil en primeras diferencias y su correspondiente Autocorrelograma Simple y Parcial	33
5.5. Autocorrelación Simple y Parcial junto con gráfico de los residuos del modelo ARIMA(2,1,0)(1,0,0) con valor atípico en Octubre de 2003	37
5.6. Predicción año 2007 - modelo univariado seleccionado: gas oil	41
5.7. Predicción año 2008 - modelo univariado seleccionado: gas oil	42
5.8. Consumo de Supergas en el período Enero de 1999 a Diciembre de 2007 en Uruguay. Unidad: miles de metros cúbicos	45
5.9. Consumo de Supergas por año. Período: Enero de 1999 a Diciembre de 2007	45
5.10. Salida correspondiente al Test de Dickey Fuller sobre el logaritmo del consumo de Supergas. Período Enero de 1999 a Junio de 2007	47
5.11. Autocorrelogramas y gráfica de la diferencia estacional del logaritmo de Supergas	48
5.12. Autocorrelación Simple y Parcial del modelo seleccionado	51
5.13. Predicción año 2007 - modelo univariado seleccionado para supergas	56
5.14. Predicción: año 2008	57

6.1. Gráficas de las variables del grupo: Gas Oil	60
6.2. Gas Oil, PBI de la Industria y Precio	61
6.3. Gas Oil y LLuvias	62
6.4. Residuos de las tres ecuaciones del sistema VECM	67
6.5. Normalidad de los residuos de la ecuación de Gas Oil	68
6.6. Autocorrelograma simple y parcial de le ecuación de gas oil	69
6.7. Predicción año 2007 - modelo seleccionado grupo Gas oil	72
6.8. Predicción para el año 2008 - modelo seleccionado grupo gas oil	76
6.9. Gráficas de las variables del grupo: Supergas	77
6.10. Variables: Supergas y Precio (en color negro: <i>consumo de supergas</i> , en rojo: <i>Precio</i>)	78
6.11. Variables: Supergas y Temperatura Máxima (en color negro: <i>consumo de supergas</i> , en rojo: <i>Temperatura Máxima</i>)	79
6.12. Residuos de las tres ecuaciones del sistema VECM	84
6.13. Normalidad de los residuos de la ecuación de Supergas	85
6.14. Autocorrelograma simple y parcial de le ecuación de supergas	86
6.15. Predicción para el año 2007 - modelo seleccionado grupo supergas	89
6.16. Predicción para el año 2008 - grupo supergas	91
9.1. Test de Dickey Fuller para la segunda diferencia de la serie de consumo de gas oil en el período Enero de 1999 a Diciembre del 2007	110
9.2. Autocorrelogramas Simple y Parcial de los residuos de los modelos: ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1)	113
9.3. Grafica de los Residuos de los modelos ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1)	114
9.4. Autocorrelación simple y parcial de los modelos ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1) con la inclusión de Octubre de 2003 como atípico aditivo	116
9.5. Residuos de los modelos modelos ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1) con la inclusión de Octubre de 2003 como atípico aditivo	117
9.6. Autocorrelograma simple y parcial del Modelo ARIMA(3,1,0)(1,0,0)	122

10.1. Autocorrelogramas simple y parcial de la serie de Consumo de Supergas en logaritmos y gráfica del logaritmo del Consumo de Supergas.	123
10.2. Autocorrelogramas simple y parcial de los modelos SARIMA(0,0,0)(1,1,0) y SARIMA(0,0,0)(0,1,1) para la serie en logaritmos de supergas	127
10.3. Residuos de los modelos SARIMA(0,0,0)(1,1,0) y SARIMA(0,0,0)(0,1,1) para la serie en logaritmos de supergas	128
10.4. Residuos Modelo Preliminar	130
10.5. Predicción Julio de 2007 - Junio de 2008 - Modelo univariado seleccionado para el consumo de supergas	132
11.1. Test de Dickey Fuller a log(Supergas).	133
11.2. Test de Dickey Fuller a log(Precio).	134
11.3. Test de Dickey Fuller a log(Tmáx).	135
11.4. Test de Dickey Fuller a log(Tmín).	136
11.5. Predicción para el período Julio de 2007 a Junio de 2008	139
11.6. Autocorrelograma simple y parcial del modelo 7	152
12.1. Test de Dickey Fuller a serie de Gas Oil.	158
12.2. Test de Dickey Fuller a serie de Precio de Gas Oil.	159
12.3. Test de Dickey Fuller a serie de Lluvias.	160
12.4. Test de Dickey Fuller a serie de PBIagro.	161
12.5. Test de Dickey Fuller a serie de PBI-industria.	162
12.6. Predicción período Julio de 2007 a Junio de 2008 - modelo seleccionado grupo gas oil	164

Índice de cuadros

5.1. Estimación del modelo ARIMA(2,1,0)(1,0,0) con atípico en Octubre de 2003	35
5.2. Predicción año 2007 - modelo univariado seleccionado: gas oil	40
5.4. Predicción a 6 pasos. Período: Febrero a Julio de 2008	41
5.3. Predicción año 2008 - modelo univariado seleccionado: gas oil- modelo univariado seleccionado	42
5.5. Predicción a 6 pasos. Período: Marzo a Agosto 2008	43
5.6. Predicción a 6 pasos. Período: Abril a Setiembre de 2008	43
5.7. Modelo seleccionado	50
5.8. Predicción año 2007 - modelo univariado seleccionado para supergas . .	56
5.9. Predicción año 2008 - modelo univariado seleccionado para supergas . .	57
5.10. Predicción a 6 pasos año 2008 Febrero - Julio	58
5.11. Predicción a 6 pasos año 2008 Marzo - Agosto	58
5.12. Predicción a 6 pasos año 2008 Abril - Setiembre	58
6.1. Coeficiente y significación de las variables integrantes del modelo seleccionado	65
6.2. Predicción año 2007 - modelo seleccionado grupo Gas oil	73
6.3. Predicción del año 2008 - modelo seleccionado grupo gas oil	74
6.4. Predicción para el año 2008 incorporando hasta la observación de Marzo - modelo seleccionado grupo gas oil	75
6.5. Coeficientes del modelo seleccionado para el consumo de supergas . . .	81
6.6. predicción para el año 2007 - modelo seleccionado grupo supergas . . .	90

6.7. Predicción del año 2008 - grupo supergas	92
6.8. Predicción del año 2008 incorporando la observación de Enero de 2008 - grupo supergas	93
6.9. Predicción del año 2008 incorporando hasta la observación de Mayo de 2008 - grupo supergas	93
9.1. Estimación de los modelos Rechazados para la demanda de Gas Oil: ARIMA(1,1,1)(0,0,1) y ARIMA(2,1,1)(0,0,1)	111
9.2. Estimación de los modelos Aceptados ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1)	111
9.3. Estimación de los modelos ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1) con la inclusión de indicatriz Octubre de 2003	115
9.4. Estimación de los modelos ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1) con la inclusión de Octubre de 2003, Febrero de 2002, Setiembre y Noviem- bre de 2007 como variables indicatrices	118
9.5. Predicción Jul - Dic 2007. Modelo ARIMA(2,1,0)(1,0,0)	119
9.6. Predicción: Jul - Dic de 2007. Modelo ARIMA(2,1,0)(0,0,1)	120
9.7. Predicción para el año 2008. Modelo ARIMA(2,1,0)(0,0,1)	120
9.8. Predicción: Julio 2007 - Junio 2008 - Modelo seleccionado	121
9.9. Estimación del modelo ARIMA(3,1,0)(1,0,0) con Octubre de 2003 como atípico	122
10.1. Estimación de los coeficientes de los modelos rechazados para la mo- delización del supergas	125
10.2. Estimación de los coeficientes de los modelos aceptados para la mo- delización del supergas	125
10.3. Modelos preliminares	129
10.4. Predicción Julio de 2007 - Junio de 2008 - Modelo univariado selecciona- do para el consumo de supergas	131
11.1. Autocorrelación simple y parcial del modelo seleccionado	137
11.2. Modelo seleccionado con Febrero de 2004 como atípico	138
11.3. Predicción para el período Julio de 2007 a Junio de 2008	139

11.4. Significación de las variables incluidas en el modelo 1	141
11.5. Coeficientes y significación de las variables incluidas en el modelo 2	143
11.6. Predicción del año 2008 - modelo 2	143
11.7. Coeficientes y significación de las variables del modelo 3	144
11.8. Predicción año 2008 - modelo 3	145
11.9. Coeficientes y significación de las variables del modelo 4	146
11.10 Predicción para el año 2008 - modelo 4	147
11.11 Coeficientes y significación de las variables incluidas en el modelo 5	148
11.12 Predicción para el año 2008 - modelo 5	148
11.13 Coeficientes y significación de las variables incluidas en el modelo 6	149
11.14 Predicción del año 2008 - modelo 6	150
11.15 Coeficientes y significación de las variables incluidas en el modelo 7	151
11.16 Autocorrelaciones simple y parcial- modelo 7	153
11.17 Predicción para el año 2008 - modelo 7	153
11.18 Coeficientes y significación de las variables incluidas en el modelo 8	155
11.19 Predicción para el año 2008 - modelo 8	156
11.20 Predicción del período Febrero a Diciembre de 2008 - modelo 8	156
12.1. Autocorrelograma simple y parcial del modelo seleccionado	163
12.2. Predicción para el período Julio de 2007 - Junio de 2008 - modelo seleccionado grupo gas oil	165
12.3. Modelo seleccionado con Enero de 2005 y Marzo de 2007 como atípicos	167
12.4. Coeficientes y significación de las variables integrantes del modelo seleccionado con Julio de 2007 como atípico	168
12.5. Coeficientes y significación de las variables integrantes del modelo 1	169
12.6. Predicción para el año 2008 - Modelo 1	170
12.7. Coeficientes y significación de las variables del modelo 2	171
12.8. Predicción para el año 2008 - Modelo 2	171
12.9. Predicciones para el año 2008 - Modelo 3	172
12.10 Coeficientes y significación de las variables integrantes del modelo 3	173
12.11 Predicciones para el año 2008 - Modelo 4	174

12.12 Coeficientes y significación de las variables integrantes del modelo 4 . . 175

Capítulo 1

Introducción

¿Cuál es el modelo apropiado para estimar la demanda de gas oil y supergas en el Uruguay?, ¿un modelo de series de tiempo univariante, o un modelo multivariante?, ¿cuáles serán las variables apropiadas que permitan pronosticar la demanda de esos productos?

En pos de contestar estas interrogantes, es importante visualizar que la demanda de gas oil está vinculada a un conjunto de variables relacionadas con distintas actividades económicas del país.

En la estructura de la *matriz de gas oil para el año 2002* realizada para ANCAP por EQUIPOS MORI, el total del consumo de gas oil se divide según el siguiente detalle:

- al agro: 12 %
- a la pesca: 6 %
- al gobierno: 14 %
- al transporte vehicular: 68 % (se incluyen: camiones, vehículos particulares y ómnibus)

La importancia del Agro

El sector agropecuario utiliza el gas oil para abastecer la maquinaria agrícola. Esta, es necesaria para las tareas relacionadas a la implementación de los principales cultivos. El trabajo en el campo requiere labores de cultivo constantes, desde la siembra hasta la proximidad de la cosecha de los variados productos; que se extenderán,

según el caso, por períodos mayores a los seis meses. Durante ese período de tiempo, el abastecimiento del combustible debe ser regular.

El transporte de los productos agropecuarios en nuestro país, depende de un modo casi exclusivo del transporte vehicular (camiones en su mayoría), así que ese factor también se agrega a la ecuación. Por ejemplo, el sector forestal, con perfil netamente exportador, requiere elevados consumos de combustible para la producción y para el traslado de la mercadería a zonas portuarias.

A su vez, la inestabilidad causada por los conflictos de los mercados exportadores de petróleo, la mayoría de ellos ubicados en Medio Oriente, generan impulsos al alza en el precio de venta del barril de crudo. A Noviembre del 2007 la cotización media del barril de petróleo de la OPEP¹ llegó a 92,61 dólares, mientras que la cotización media del barril para el 2006 fue de 61,00 dólares. Estas fluctuaciones, impactan gravemente en la economía nacional y desestabilizan el precio de venta público de los combustibles. De sufrir escasez del producto, el proceso productivo del agro se vería retrasado y se ocasionaría un daño a la economía nacional.

La importancia del Clima

El clima tiene un impacto directo sobre la producción agropecuaria, debido a que la siembra se vincula con las estaciones del año. Existen cultivos de invierno, de otoño, de verano y de primavera. Los últimos años se han generado fenómenos climáticos adversos para el agro y la producción agrícola. Temperaturas atípicas extremas, inundaciones, vientos fuertes y sequías prolongadas son algunas de las consecuencias que sufre nuestro país debido al *efecto invernadero*. Éstos tipos de fenómenos, retrasan la siembra disminuyendo el consumo de gas oil por parte de éste sector. El impacto del clima, en lo referente al gas oil no es sólo a través del agro, sino que también, a través del turismo. Generando un consumo más elevado en meses claves, como los relacionados a las vacaciones de verano, carnaval, turismo, vacaciones de invierno y meses que

¹Organización de Países Exportadores de Petróleo

tienen fines de semana "largos"²

De igual forma que el efecto invernadero tiene consecuencias directas sobre el agro y el consumo de gas oil, también las tiene sobre aquellos productos derivados del petróleo que tienen vinculación con el clima; como el caso del supergas. Las temperaturas mínimas, cada vez más extremas, que se han sucedido los últimos años, hacen que la previsión de éste producto resulte muchas veces insuficiente para abastecer el mercado interno. La proliferación de estufas a supergas, a partir del año 1997, ha hecho de éste producto el preferido para la calefacción del hogar (entre otras utilidades del mismo) lo que incrementa su consumo.

Como consecuencia de lo anteriormente descrito, para predecir la demanda de consumo de gas oil en Uruguay se toman como variables explicativas:

- El IVFPBI (Índice de Volumen Fijo del Producto Bruto Interno) mensual del sector agrícola, como variable indicadora de la economía.
- El IVFPBI mensual del sector industrial; debido a que incluye dentro de sus rubros la producción de papel y la industria de la madera.
- Lluvias promedio mensual en milímetros.
- Precio de venta al público del gas oil mensual, deflactado por el IPC (índice de precios al consumo.)

Para predecir la demanda de supergas, se tomarán como variables explicativas:

- Temperatura mínima mensual promedio.
- Temperatura máxima mensual promedio.

²Ley N° 16.805 Publicada D.O. 31 dic/96 N° 24696 - FERIADOS. Establece que si los feriados declarados por ley coincidieran en: sábado, domingo o lunes, se observarán en esos días. Si ocurrieren en martes o miércoles: se observarán el lunes inmediato anterior. Si ocurrieren en jueves o viernes, se observarán el lunes inmediato siguiente. Quedan exceptuados de este régimen los feriados de Carnaval y Semana de Turismo y los correspondientes al 1° y 6 de enero, 1° de mayo, 18 de julio, 25 de agosto y 25 de diciembre, los que continuarán observando en el día de la semana en que ocurrieren, cualquiera fuere el mismo.

- Precio de venta al público de supergas (por litro) mensual, deflactado por el IPC (índice de precios al consumo).

El trabajo se organizará en 5 secciones. La primera, presenta los antecedentes sobre el tema en Uruguay y el resto del mundo. La segunda sección refiere al planteamiento de los objetivos. La tercer sección, analiza la demanda de gas oil y supergas de forma univariada. La cuarta sección, analiza la demanda de gas oil y supergas de forma multivariada y por último la quinta sección presenta las conclusiones finales sobre los modelos planteados.

Capítulo 2

Objetivos

El **objetivo general** del presente trabajo consiste en la modelización de la demanda del consumo de gas oil y supergas en Uruguay, para el período comprendido entre Enero de 1999 y Diciembre de 2007.

Bajo este enfoque se plantean los siguientes **objetivos específicos**:

- Un modelo univariado de series de tiempo que se ajuste a la demanda de consumo de gas oil y supergas en el Uruguay, en el período Enero de 1997 a Diciembre de 2007. Por modelo univariado se entiende: aquel modelo que utiliza la serie de datos considerados, para explicarse asimismo.
- Incorporación de variables explicativas, de forma de encontrar aquellos modelos de series temporales multivariados que expliquen la demanda de gas oil y supergas. Por modelo de series temporales multivariado se entiende aquel modelo que utiliza la información contenida en la serie a analizar, más información proporcionada por otras variables que se suponen tienen relación con la variable endógena.
- Obtener predicciones a corto y largo plazo para la demanda de ambos productos.

Capítulo 3

Antecedentes

A continuación se hace referencia a investigaciones previas realizadas en diversos países, vinculadas con el tema a tratar en el presente trabajo.

Los antecedentes en este campo están referidos a la estimación de la demanda de gasolina, debido a que es el combustible más consumido a nivel de transporte, ya sea particular o no.

Existen dos estudios realizados a nivel nacional. El primero: *Imposición Óptima a las Naftas y el Gas Oil: Un Análisis Empírico Para Uruguay (año 2002)*, realizado por DANTE AMENGUAL y GERMÁN CUBAS. El segundo: *Elasticidad de corto y largo plazo de la demanda de gasolina en Uruguay: Una aplicación de modelos de corrección de error (Año 2002)* realizado por DIEGO ABOAL.

- AMENGUAL Y CUBAS hacen referencia principalmente, a la estimación de la demanda de las gasolinas y la demanda del gas oil, mediante modelos VAR (Vectores Autorregresivos) utilizando técnicas de cointegración. El estudio centra su interés en el carácter impositivo de los combustibles. Obtienen como resultado, que la demanda de naftas es significativamente más elástica que la demanda de gas oil. Las elasticidades precio e ingreso para el gas oil y las gasolinas resultaron ser: -0.3 y 1.13, 0.45 y 0.57, respectivamente. Atribuyen el resultado al diferente tratamiento impositivo que recibieron los productos en el período de estudio. También concluyen que existe una elevada sustitución entre los combustibles.

El modelo planteado por los autores utiliza como variables explicativas las si-

güentes series:

- Cantidad vendida en metros cúbicos de naftas (sin desagregar por subtipos) y de gas oil, per cápita.
 - Serie de precios de naftas y gas oil deflactados por el índice de precios al consumo (IPC).
 - Índice de volumen físico del Producto Bruto Interno (IVFPBI) per cápita.
- El análisis planteado por DIEGO ABOAL para estimar la demanda de naftas en el período 1986 al 2001, es realizado mediante análisis de cointegración y modelos de corrección de error (ECM). Concluye que las elasticidades precio e ingreso de corto plazo de las gasolinas fueron: -0.26 y 0.48, mientras que las de largo plazo fueron: -0.35 y 0.86, respectivamente.

El autor utiliza las siguientes variables explicativas:

- Índice de Volumen Físico (IVF) del PBI trimestral, desestacionalizado y en logaritmos
- Precio de la nafta mensual, deflactada por el IPC, en logaritmos. Se toma el promedio de los tres meses para generar la serie trimestral.
- Consumo de Nafta trimestral, desestacionalizada y en logaritmos.
- Índice de precios al consumo (IPC).

A nivel internacional los estudios revisados son:

- *Estimación de la Demanda de Combustibles en República Dominicana* realizado por MARTÍN FRANCOS RODRIGUEZ. En esta investigación se estiman las funciones de demanda de gasolinas (regular y premium), gas oil y supergas. Los métodos utilizados fueron análisis de cointegración y modelos de corrección de error. El autor llega a la conclusión de que la gasolina regular fue el producto con mayor elasticidad precio a largo plazo (-0.63); seguido por la gasolina premium (-0.41) y el supergas (-0.29).

- *PREVISÃO COM COINTEGRAÇÃO E MODELO A CORREÇÃO DE ERRO: APLICAÇÃO EMPÍRICA NO CONSUMO DE GASOLINA DO BRASIL.*, realizado por ROBERT W. SAMOBYL Y ANDERSON DE BARROS DANTAS. El estudio tiene como objetivo estimar la demanda de gasolinas en Brasil mediante técnicas de cointegración y modelos de corrección de error (ECM). Los autores concluyen que la elasticidad precio estimada de corto y largo plazo para las gasolinas fue de -0.199 y -0.2829, respectivamente. La sensibilidad de la demanda en relación al precio resultó ser inelástica en el corto y largo plazo.

Las series son expresadas en logaritmos y con periodicidad anual en el período comprendido entre los años 1955 a 1995.

Las variables que utiliza para explicar la media anual del consumo de barril diario de gasolina son:

- PBI de Brasil a precios corrientes.
- Precio al público de la gasolina.
- Precio corriente en dólares del barril.
- Media anual del consumo de barril diario de gasolina desfasado en el tiempo un período.

Capítulo 4

Aspectos Metodológicos Estadísticos

4.1. Metodología aplicada a las series de Tiempo Univariadas

4.1.1. Especificación del modelo $ARIMA(p, d, q)(P, D, Q)_s$

El objetivo del estudio de las *series temporales* es el de realizar inferencias sobre un proceso estocástico¹ desconocido a partir de los datos pasados de la variable. Una de las técnicas estadísticas que puede emplearse en este contexto es la realizada por *Box y Jenkins* (1973).

Box y Jenkins plantean la posibilidad de encontrar una familia de modelos de series de tiempo, denominados *ARIMA*, y un mecanismo recursivo para la identificación de los mismos.

La especificación del modelo univariado *ARIMA estacional multiplicativo*: $ARIMA(p, d, q)(P, D, Q)_s$ de series temporales, queda planteado según la ecuación:

$$\Phi_p(L^s)\phi_p(L)\nabla_s^D\nabla^d Y_t = \theta_q(L)\Theta_Q(L^s)\varepsilon_t \quad (4.1)$$

A partir de la fórmula anterior se logra modelizar la dependencia regular y estacional de forma conjunta. El componente regular, refiere a la dependencia entre observaciones consecutivas y el componente estacional, a la dependencia entre observaciones separadas por s períodos. La ventaja del modelo *ARIMA multiplicativo* es que

¹Ver: Anexo Metodológico: sección: 8.1

logra representar la regularidades de la serie con polinomios de orden menor.

Los componentes de la ecuación se interpretan de la siguiente manera:

- **p y P:** orden del polinomio autorregresivo de la parte regular y estacional.
- **d y D:** número de diferencias de orden regular y estacional, que la serie necesita para ser estacionaria.
- **q y Q:** orden del polinomio de medias móviles de la parte regular y estacional.

Dónde:

- $\Phi_p(L^s) = (1 - \Phi_1 L^s - \dots - \Phi_P L^{sP})$, es el operador AR estacional de orden sP .
- $\phi_p(L) = (1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)$, es el operador AR regular de orden p .
- $\nabla_s^D \nabla^d = (1 - B^s)^D (1 - B)^d$, representan las diferencias estacional y regular respectivamente.
- $\Theta_Q(L^s) = (1 - \Theta_1 L^s - \dots - \Theta_Q L^{sQ})$, es el operador media móvil estacional de orden sQ .
- $\theta_q(L) = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q)$, es el operador media móvil regular de orden q .
- ε_t es un *proceso ruido blanco*² denominado *innovación*. Este componente incluye toda la nueva información se que va sumando a la serie en cada momento t ; información que no es captada por la historia de la misma.

²Proceso ruido blanco: proceso estacionario definido por las siguientes condiciones:

1. $E(\varepsilon_t) = 0, \quad t = 1, 2, \dots$
2. $Var(\varepsilon_t) = \sigma^2 \quad t = 1, 2, \dots$
3. $Cov(\varepsilon_t, \varepsilon_{t-k}) = 0 \quad k = \pm 1, \pm 2, \dots$

4.1.2. Supuestos del modelo univariado

La metodología Box - Jenkins tiene como supuesto básico la *estacionariedad* de la serie analizada.

Una serie es *estacionaria* si la media y la varianza de la misma son constantes en el tiempo. Además, si para un intervalo de tiempo de longitud k se calcula la autocorrelación de dos variables integrantes del proceso estocástico, el resultado sólo depende del intervalo que las separa (k). Si esto sucede, la serie se denomina *estacionaria en sentido débil*.

A su vez, si se agrega el supuesto de normalidad de los residuos, la serie será *estacionaria en sentido estricto*.

Los **supuestos** que implican la *estacionariedad* del modelo univariante, son:

1. Media constante en el tiempo.

$$\mu_t = E(X_t) \quad (4.2)$$

2. Varianza constante en el tiempo (homocedasticidad).

$$\sigma_t^2 = \text{Var}(X_t) \quad (4.3)$$

3. La covarianza (ecuación 4.4) y la autocorrelación (ecuación 4.5) entre dos variables sólo depende del intervalo que las separa.

$$\text{Cov}(X_t, X_{t+k}) = \text{Cov}(X_{t+j}, X_{t+j+k}) = \gamma_k \quad (4.4)$$

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (4.5)$$

Con $\gamma_0 = \sigma^2$ (ecuación: 4.5)

4. La distribución de los residuos es: independiente e idénticamente distribuida (i.i.d) Normal(0, σ^2).

4.1.3. Metodología aplicada para la detección del modelo

Como se mencionó anteriormente la metodología *Box - Jenkins* permite la estimación del modelo a ajustar a la serie estudiada. La especificación se realiza mediante un mecanismo circular de varias etapas, cuyo objetivo es el de conocer cual es el mejor modelo que se ajusta a dicha serie. Se puede leer una profundización de cada etapa en el Anexo Metodológico, sección: 8.3.

Las etapas de dicho mecanismo son:

1. Identificación
2. Estimación
3. Diagnóstico
4. Predicción

Se debe prestar atención a los coeficientes del modelo estimado, dado que hay condiciones de estacionariedad directamente vinculados con ellos. Se amplía este tema en el Anexo Metodológico, sección: 8.4.

4.1.4. Predicción

La propiedad más interesante sobre los modelos ARIMA es la posibilidad que brindan de obtener valores futuros dado los valores pasados de la variable. Es deseable que la predicción se aproxime lo más posible al valor observado. Para ello es necesario minimizar el error de predicción; definido según:

$$e_{n+k} = z_{n+k} - \hat{z}_{n+k} \quad (4.6)$$

donde:

- z_{n+k} es el valor futuro de la variable z_n , k períodos adelante.
- \hat{z}_{n+k} es un predictor definido según valores pasados observados:

$$\hat{z}_{n+k} = \alpha_1 z_n + \alpha_2 z_{n-1} + \dots + \alpha_n z_1 \quad (4.7)$$

Como la variable z_{n+k} es desconocida, es necesario averiguar su distribución de probabilidad para poder acotar el error; es decir calcular la probabilidad de que el error este en un intervalo dado.

El procedimiento es minimizar el error cuadrático medio (ECM) dado la información pasada; es decir minimizar la siguiente ecuación (Ver desarrollo en Anexo Metodológico, sección: 8.10):

$$ECM(z_{n+k}/z_n) = E \left[\left(z_{n+k} - \hat{z}_{n+k} \right)^2 / z_n \right] = E \left[e_{n+k}^2 / z_n \right] \quad (4.8)$$

Como solución se obtiene que *el predictor que minimiza el ECM es la esperanza de la variable z_{n+k} condicionada a la información pasada.*

$$\hat{z}_{n+k} = \mu_{n+k/n} \quad (4.9)$$

4.2. Metodología aplicada a las series de Tiempo Multivariadas

Cuando interesa describir la relación existente entre un grupo de series temporales, se puede pensar en dos caminos: métodos aplicables a series estacionarias y métodos aplicables a series *no* estacionarias, entre otros.

Intentar encontrar relaciones entre variables no estacionarias mediante mediante métodos que no se aplican a las mismas se conoce como: *Relaciones Espúreas* y su aplicación lleva a resultados equívocos. Se podría pensar en analizar las series en diferencias, de esta forma estadísticamente se resuelve el problema de estacionariedad, pero se pierde información y se torna compleja la explicación de los parámetros de la regresión.

En presencia de series *no estacionarias* la técnica estadística a aplicar es la *Cointegración*. Este concepto, junto a los *Modelos de Corrección de Error* han sido los principales aportes introducidos por *Engel y Granger* (1987).

4.2.1. Cointegración

Las variables integrantes del vector $x_t = (x_{1t}, x_{2t}, \dots, x_{nt})'$ se dicen *cointegradas* si:

1. Todos los componentes del vector x_t son de orden d .
2. Existe un vector $\beta = (\beta_1, \beta_2, \dots, \beta_n)$, tal que la combinación lineal $\beta'x_t = \beta_1x_{1t} + \beta_2x_{2t} + \dots + \beta_nx_{nt}$ es de orden j , con $d > j$.

- El vector β es denominado *vector de cointegración*.

A su vez, si el vector de variables x_t tiene n componentes, entonces existen como máximo $n-1$ *vectores de cointegración linealmente independientes*.

- El número de vectores de cointegración se denomina *rango de cointegración*.

El sistema se encuentra en **equilibrio a largo plazo** si se cumple:

$$\beta_1x_{1t} + \beta_2x_{2t} + \dots + \beta_nx_{nt} = 0 \quad (4.10)$$

La desviación del equilibrio se define como:

$$e_t = \beta x_t \quad (4.11)$$

Una técnica que se utiliza para determinar la existencia o no de cointegración es el *Test de Johansen*³. A su vez, permite determinar el *rango de cointegración* o número de ecuaciones de cointegración. Básicamente el test realiza pruebas de hipótesis sobre el rango de la matriz Π en un sistema de variables no estacionarias. La definición de la matriz Π se menciona en la sección 4.2.3: Modelo Vectorial Autorregresivo con Mecanismo de Corrección de Error (VECM). El desarrollo del test está realizado en el Anexo Metodológico: Tests utilizados en el análisis de los datos, sección: 11.4.1.

4.2.2. Modelos Var

Especificación del sistema VAR para dos variables

Cuando existen dos variables y_t y z_t , en las cuales se supone no existe una causalidad unidireccional entre ellas; sino que existe retroalimentación, los modelos VAR son

³Dependiendo del rango de la matriz Π es que se opta por técnicas de cointegración propiamente dichas (VECM) o modelos VAR en niveles o en diferencias, los cuales no implican cointegración.

apropiados. La metodología VAR tiene como supuesto que las variables que se desean relacionar son *estacionarias*.

Observación: Se debe tener en cuenta que el orden que define el VAR(p) esta dado por la cantidad de retardos que resultan significativos en la especificación del VAR.

Un modelo VAR(1) en su forma *primitiva o estructural* queda definido según el sistema:

$$\begin{cases} y_t = b_{10} - b_{12}z_t + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + \varepsilon_{yt} \\ z_t = b_{20} - b_{21}y_t + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + \varepsilon_{zt} \end{cases} \quad (4.12)$$

Cuyos supuestos son:

1. Las variables y_t y z_t son estacionarias.
2. Los componentes del vector ε_t son ruido blanco. A su vez, tienen varianza finita (σ_{yt} y σ_{zt}) y no están correlacionados.

Operando y utilizando notación matricial se obtiene:

$$Bx_t = \Gamma_0 + \Gamma_1x_{t-1} + \varepsilon_t \quad (4.13)$$

Donde:

$$\begin{aligned} B &= \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}, & x_t &= \begin{bmatrix} y_t \\ z_t \end{bmatrix}, & \Gamma_0 &= \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} \\ \Gamma_1 &= \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, & \varepsilon_t &= \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \end{aligned} \quad (4.14)$$

Multiplicando por B^{-1} la ecuación anterior se obtiene el VAR en su forma *standard*:

$$x_t = A_0 + A_1x_{t-1} + e_t \quad (4.15)$$

Donde:

- $A_0 = B^{-1}\Gamma_0$
- $A_1 = B^{-1}\Gamma_1$
- $e_t = \varepsilon_t = B^{-1}\varepsilon_t$

4.2.3. Modelo Vectorial Autorregresivo con Mecanismo de Corrección de Error (VECM)

Los modelos VECM permiten modelizar las relaciones a largo plazo de las variables junto con la dinámica de corto plazo. En un sistema de variables integradas, se supone que en el largo plazo las variables tienden a moverse en conjunto, pero pueden ocasionarse desviaciones en el corto plazo. Por lo tanto, si se incluye un *término de corrección de error* se supone que el sistema se acerca al equilibrio, ya que la desviación del equilibrio en el período pasado se corrige en el próximo período y así sucesivamente.

Los *modelos de vectores autorregresivos con mecanismo de corrección de error* tienen su origen en los modelos VAR a los cuales se les aplican restricciones dadas por las relaciones de cointegración y se les agrega un término de corrección de error que se expresa como: Πy_{t-1} . El *teorema de Representación de Granger* es el que permite expresar un modelo VAR con componentes cointegrados como un modelo de corrección de error.

Por lo tanto, partiendo del modelo VAR(p) definido según:

$$y_t = \mu + \sum_{i=1}^p \Pi_i y_{t-i} + \Phi D_t + \varepsilon_t \quad (4.16)$$

Donde:

- y_t : es el vector de variables.
- Π_i : son matrices de coeficientes.
- D_t : es un vector de variables auxiliares como variables de intervención o variables que capten estacionalidad.
- ε_t : término de error. Se supone cumple con los supuestos asignados a los residuos.

Se define el modelo VECM tal que:

$$\Delta y_t = \mu + \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \Pi y_{t-1} + \Phi D_t + \varepsilon_t \quad (4.17)$$

Donde:

- Γ_i : los coeficientes de la matriz miden efectos transitorios.

- Πy_{t-1} : es el término de corrección de error.
- El lado izquierdo de la ecuación debe ser estacionario, lo mismo que el término de corrección de error.

Como se observa en la ecuación 4.17, las variables del modelo VECM son ingresadas a la modelización en diferencias. Por lo tanto, el lado izquierdo de la ecuación es estacionario y los términos del lado derecho expresados en diferencias también lo son. Es necesario analizar el término Πy_{t-1} para que se cumpla la estacionariedad, más específicamente analizar la matriz Π .

Condiciones de la matriz Π

A continuación se detallan las condiciones que debe cumplir la matriz Π (analizando su rango) para que el lado derecho de la ecuación sea estacionario.

1. **rango(Π) = n**. Aquí se tienen n ecuaciones linealmente independientes, cuyas combinaciones serán estacionarias. Corresponde aplicar un VAR en niveles.
2. **rango(Π) = 0**. Los elementos de la matriz Π son cero, por lo tanto, no existe combinación lineal que vuelva Πy_t estacionario. El VECM queda especificado como un VAR en diferencias.
3. $0 < \text{rango}(\Pi) = r < n$. Existen múltiples relaciones de cointegración. Se tienen r combinaciones linealmente independientes de las variables; lo que es igual a decir: r relaciones de cointegración; las cuales son estacionarias por definición.

Si se cumple el caso 3, se aplica la definición de cointegración. Por lo tanto, se definen dos matrices α y β de dimensión $(n \times r)$ tal que:

$$\alpha\beta' = \Pi \quad (4.18)$$

Se cumple que:

- El producto $\alpha\beta' y_{t-1}$ es estacionario, así también $\beta' y_{t-1}$.

- Las r columnas linealmente independientes son los *vectores de cointegración* y forman la matriz (β) .
- El $\text{rango}(\Pi)$ es el rango de cointegración del sistema.
- Los elementos de la matriz α representan la *velocidad de ajuste* al equilibrio de largo plazo.

Exogeneidad débil

En un sistema de variables cointegradas si una variable no responde a desvíos del equilibrio de largo plazo se dice que es *débilmente exógena*. La significación de los parámetros de velocidad de ajuste (α) es la que determina la existencia de exogeneidad débil. En particular si se restringe el valor de α a cero la variable en cuestión no se ajusta a cambios en el equilibrio de largo plazo.

Capítulo 5

Resultados: Modelos Univariados

5.1. Resultados referidos a la demanda de GAS OIL

5.1.1. Datos

La serie que se analiza corresponde al consumo de gas oil (en miles de metros cúbicos) en Uruguay en el período de Enero de 1999 a Diciembre de 2007. Los datos tienen periodicidad mensual y son proporcionados por A.N.C.A.P.¹.

5.1.2. Análisis descriptivo

Como se observa en la figura 5.1, la serie tiene pequeñas fluctuaciones dentro del período considerado. Desde el año 1999 al 2004, existe una leve tendencia decreciente en el consumo del producto, posiblemente generada por el disparo al alza del precio de dicho combustible. A partir del año 2004, el consumo de gas oil tiende a la recuperación. Suceso que puede deberse a un cambio en la conducta del consumidor, que debido al aumento en el precio de las naftas migra hacia el consumo de gas oil.

¹Administración Nacional de Combustibles, Alcohol y Portland

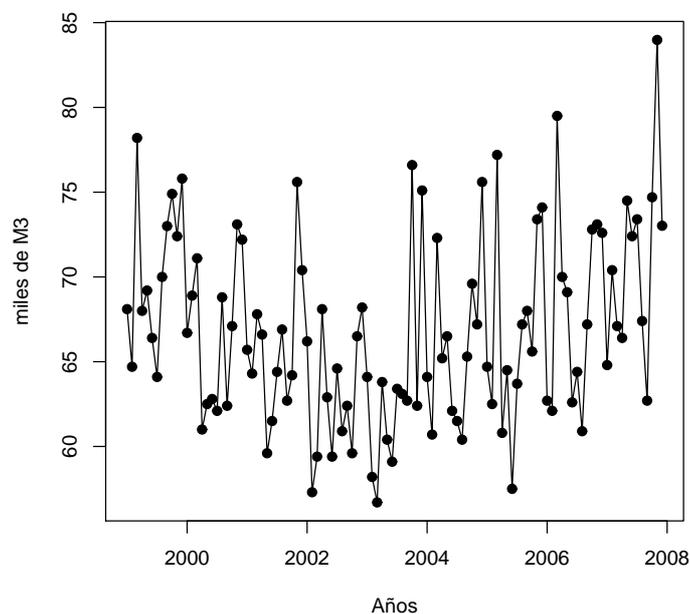


Figura 5.1: Serie de demanda de consumo de gas oil en Uruguay para el período Enero de 1999 a Junio de 2007

En el gráfico de *Gas Oil por año* Figura: 5.2, se puede apreciar que la serie (más allá de las fluctuaciones al alza o baja) conserva cierta dinámica común. Por ejemplo, si se considera el trimestre Enero - Marzo, se observa que el consumo en Enero es consumo medio en torno a los 65 mil metros cúbicos, decrece en Febrero y repunta hacia picos máximos en Marzo (cercanos a los 80 mil metros cúbicos). Este comportamiento de evolución conjunta se repite para los primeros seis meses de los años considerados, salvo algunas observaciones excepcionales. A partir de Julio, los distintos años muestran un comportamiento irregular o por lo menos no tan obvio como en el primer semestre. De todas formas, el comportamiento parejo de la primera mitad de los años evaluados puede indicar la presencia de cierta estacionalidad en el consumo. Situación que puede ser signo de estar ante una serie no estacionaria.

La no estacionariedad se analiza mediante el estudio conjunto de:

1. Test de Dicker Fuller Aumentado² a la serie de consumo de gas oil y a la serie

²El desarrollo del Test se encuentra en: Anexo Metodológico: Test de Dickey Fuller, sección: 8.6.1

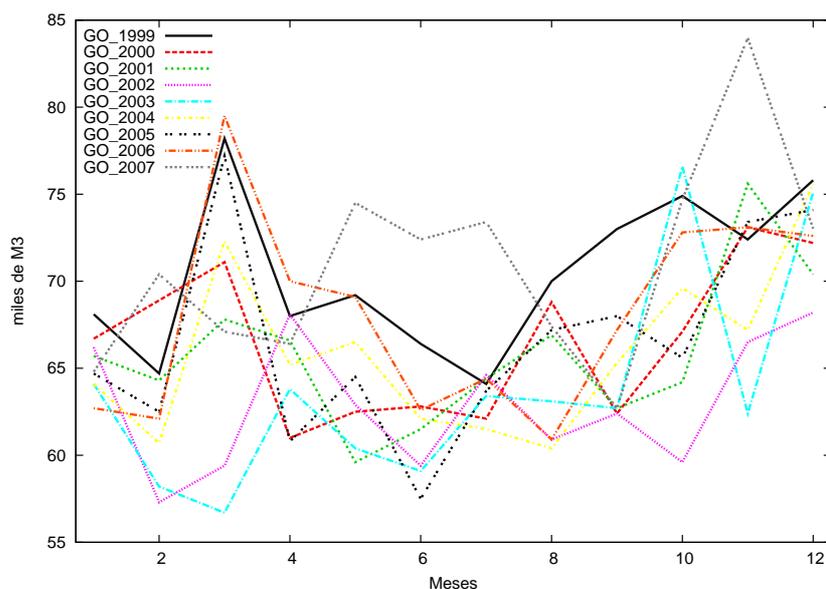


Figura 5.2: Consumo de Gas Oil por año. Período: Enero de 1999 a Diciembre de 2007

en primeras diferencias ($d(\text{gas oil})$), considerando un modelo con constante y tendencia.

2. Gráfico de la serie en primeras diferencias.
3. Autocorrelogramas simple y parcial de la serie en primeras diferencias.

Los resultados obtenidos son:

1. Referente al Test de Dickey Fuller:

Se concluye no rechazar la hipótesis nula de existencia de raíz unitaria en la serie de gas oil Ver figura: 5.3. La segunda diferencia de la serie es rechazada con un p-valor de 0.01³. Obs: El número de rezagos a utilizar en el test de Dickey Fuller se obtiene mediante el criterio de AIC (Akaike Information)⁴.

2. La Figura: 5.4, referente a la autocorrelación simple y parcial de la serie de gas oil en primeras diferencias, muestra que:

³Ver: Test para la segunda diferencia en: Anexo: Univariado Gas Oil: Test de Dickey Fuller para la segunda diferencia de la serie de consumo de gas oil en el período Enero de 1999 a Diciembre del 2007, sec: 9.1

⁴El criterio AIC se presenta en el Anexo Metodológico: sección: 8.7

```

-----
Augmented Dickey & Fuller test
-----
Deterministic regressors estimates:

      Estimate Std. Error t value Pr(>|t|)
Intercept  19.276    18.592   1.037   0.303
Trend       0.060     0.024   2.481   0.015
-----
ADF regressor estimate:

      Estimate Std. Error t value Pr(>|t|)
adf.reg  -0.344     0.295  -1.167   0.1
-----
Selected lags estimates:

      Estimate Std. Error t value Pr(>|t|)
Lag.1   -0.549     0.289  -1.900   0.062
Lag.2   -0.542     0.273  -1.983   0.051
Lag.3   -0.619     0.261  -2.370   0.021
Lag.4   -0.669     0.244  -2.742   0.008
Lag.5   -0.497     0.224  -2.219   0.030
Lag.6   -0.567     0.199  -2.856   0.006
Lag.7   -0.611     0.183  -3.331   0.001
Lag.8   -0.449     0.168  -2.671   0.009
Lag.9   -0.466     0.144  -3.244   0.002
Lag.10  -0.655     0.124  -5.290   0.000
Lag.11  -0.527     0.100  -5.254   0.000
Lag.15   0.151     0.089   1.682   0.097
Lag.20  -0.025     0.089  -0.284   0.777

Number of available observations: 87

```

Figura 5.3: Salida correspondiente al Test de Dickey Fuller sobre la serie de Consumo de Gas Oil. Período Enero de 1999 a Diciembre de 2007

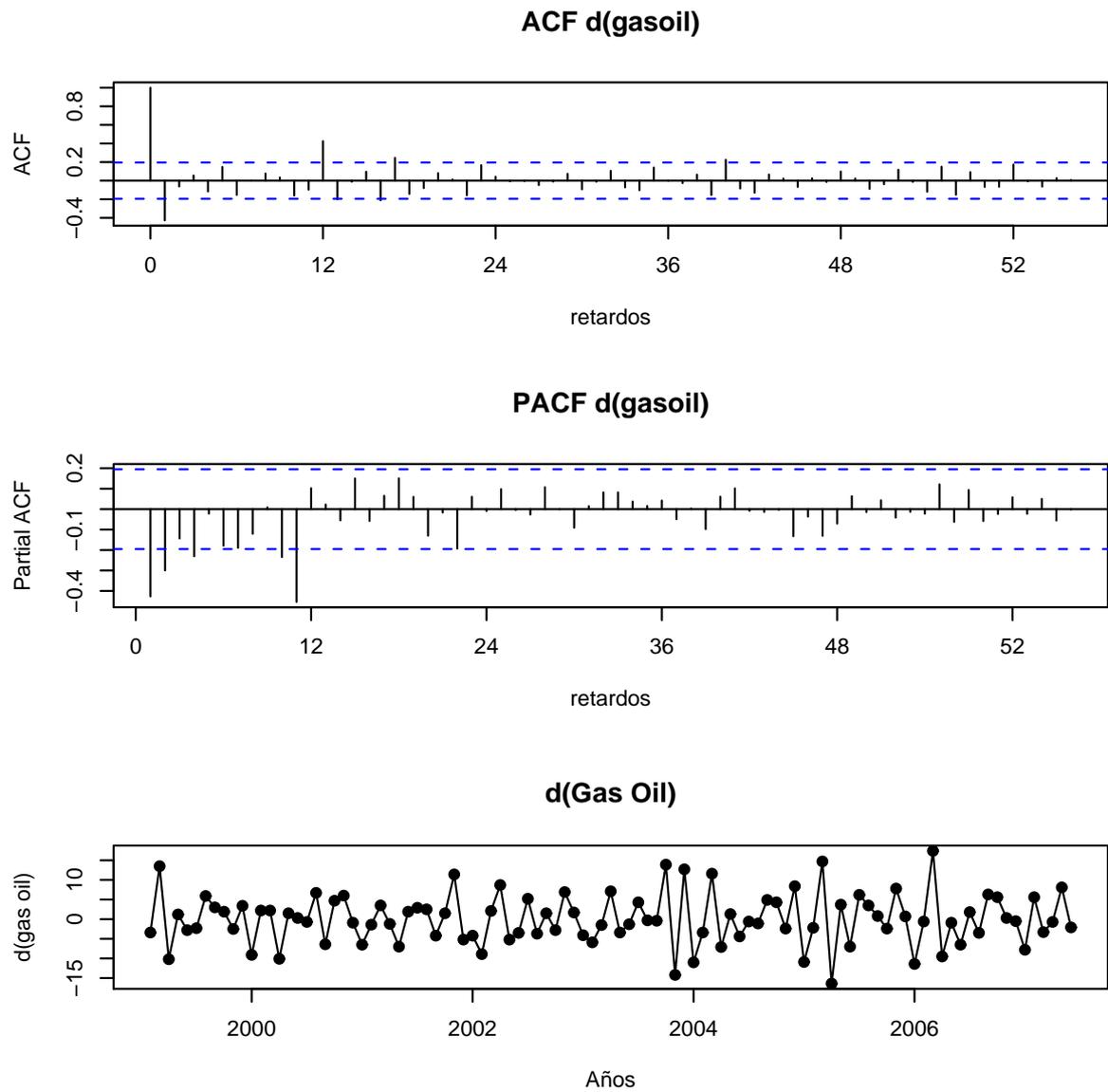


Figura 5.4: Gráfica de la serie de gas oil en primeras diferencias y su correspondiente Autocorrelograma Simple y Parcial

- La gráfica de la serie diferenciada es estacionaria en media, ya que oscila en torno a su nivel medio. No se visualiza la leve tendencia al descenso previa al 2004 ni la recuperación de la serie (posterior a dicho año).
- Los autocorrelogramas simple y parcial entran rápidamente dentro de las bandas de confianza, lo que hace suponer no necesaria la segunda diferenciación. Lo que se había verificado en la realización del test de Dickey Fuller. Ver: Anexo: Univariado Test de Dickey Fuller para la segunda diferencia de la serie de consumo de gas oil en el período Enero de 1999 a Diciembre del 2007.

Los puntos anteriormente comentados hacen concluir que la serie será analizada en primeras diferencias.

Los autocorrelogramas simple y parcial de la serie de gas oil en diferencias muestran que los retardos 1, 2 y 12 resultan significativos Ver: figura: 5.4. Este tipo de estructura hace suponer la modelización de la parte regular en conjunto con la estacional según los siguientes modelos:

1. ARIMA(2,1,0)(0,0,1)
2. ARIMA(2,1,0)(1,0,0)
3. ARIMA(1,1,1)(0,0,1)
4. ARIMA(2,1,1)(0,0,1)

5.1.3. Modelo Seleccionado

Los modelos propuestos fueron probados en las distintas etapas de la metodología empleada. Solamente los coeficientes de los modelos ARIMA(2,1,0)(0,0,1) y ARIMA(2,1,0)(1,0,0) resultan significativos.

El modelo final: modelo ARIMA(2,1,0)(1,0,0) con valor atípico aditivo en Octubre de 2003, además de cumplir con los supuestos del modelo, minimiza el error cuadrático medio en sus predicciones (en comparación con el modelo ARIMA(2,1,0)(0,0,1)). El proceso de selección del modelo junto con la justificación de porque se descartan los

modelos restantes se puede leer en: Anexo: Univariado Gas Oil, sección: 10.1.

Se elige para la modelización del consumo de Gas Oil el modelo **ARIMA(2,1,0)(1,0,0)** con valor atípico en Octubre del 2003

Estimación y Validación del Modelo Final

ARIMA(2,1,0)(1,0,0)			
	coeficiente	t-obs	t-critico
ar1	-0,48	-5,21	1,66
ar2	-0,38	-3,98	
sar1	0,53	6,34	
Oct de 2003	14,01	3,88	
AICc = 645,66			
Test Shapiro-Wilk p-value = 0,44			
TestBox-Ljung, df = 24, p-value = 0,12			

Tabla 5.1: Estimación del modelo ARIMA(2,1,0)(1,0,0) con atípico en Octubre de 2003

Coefficientes:

Referente a la significación individual de los coeficientes del modelo estimado, se observa en la tabla 5.1 que todos los coeficientes del modelo son significativamente distintos de cero al 5%. Dado que si se efectúa la prueba de hipótesis (*prueba - t*)⁵, tal que:

$$H_0) \beta_j = 0$$

$$H_a) \beta_j \neq 0$$

Cuyo estadístico es:

$$t_{(obs)}(j) = \frac{\hat{\beta}_j}{\sqrt{Var(\hat{\beta}_j)}}$$

Se cumple que $|t_{obs}| > |t_{critico}|$, entonces se rechaza la hipótesis nula. Con $t_{critico}$ igual al valor resultante de la distribución t al 95 % con 100 grados de libertad (100 igual a T-2, siendo T el total de observaciones.)

⁵Se amplía la información de esta prueba en el Anexo Metodológico, sección: 8.4

En cuanto al cumplimiento de los supuestos que tienen vinculación con los residuos se analiza: la normalidad y la autocorrelación de los mismos.

Normalidad:

El test de *Shapiro - Wilk* lleva a no rechazar la hipótesis nula de normalidad de los residuos, con un p-valor de 0,44.

Autocorrelación:

El test de *Ljung - Box* es realizado tomando k igual a 24, lo que significa que se analiza la autocorrelación de los residuos en los primeros 24 retardos. El resultado de la prueba implica no rechazar la ausencia de autocorrelación, con un p-valor de 0,12.

A continuación se muestran los autocorrelogramas simple y parcial de los residuos figura: 5.5, donde se observa que las barras de los coeficientes de autocorrelación permanecen dentro de las bandas de confianza, salvo los retardos 3 y 22 que no logran ser captados por la modelización. Sin embargo se sostiene que no hay autocorrelación en los primeros 24 retardos, según el test de Ljung Box realizado.

La descripción de los tests aplicados en los datos se realiza en el Anexo Metodológico, sección: 8.6.

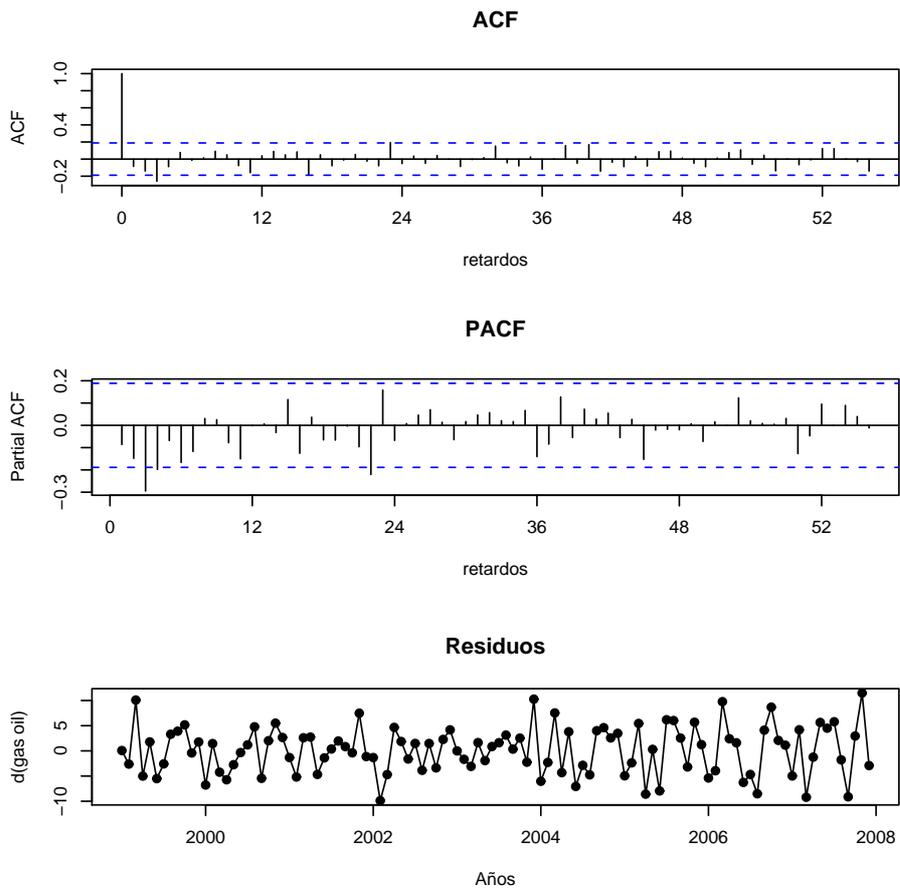


Figura 5.5: Autocorrelación Simple y Parcial junto con gráfico de los residuos del modelo ARIMA(2,1,0)(1,0,0) con valor atípico en Octubre de 2003

Interpretación del modelo seleccionado

Sustituyendo en la ecuación 4.1 (de la sección: *Aspecto Metodológico Estadístico*) los valores correspondientes a los coeficientes obtenidos en la estimación del modelo se plantea:

$$\hat{y}_t = 0,52y_{t-1} + 0,1y_{t-2} + 0,38y_{t-3} + 0,53y_{t-12} - 0,28y_{t-13} + \dots \\ \dots - 0,053y_{t-14} - 0,2y_{t-15} + 14,01I_t^{10/2003}$$

Donde se concluye que la observación en el momento t (\hat{y}_t) depende de:

- La observación del mes anterior (y_{t-1}) con peso igual a 0.52.
- La observación del 2 meses atrás (y_{t-2}) con peso igual a 0.1.
- La observación del 3 meses atrás (y_{t-3}) con peso igual a 0.38.
- La observación del año anterior (y_{t-12}). Cuyo peso en la generación de la observación al momento t es de 0,53.
- La observación de 13 meses atrás (y_{t-13}). Con peso negativo de 0,28.
- La observación de 14 meses atrás (y_{t-14}). Con peso negativo de 0,053.
- La observación de 15 meses atrás (y_{t-15}). Con peso negativo de 0,2.
- $I_t^{10/2003}$: Es la forma de especificar el efecto puntual del valor atípico: Octubre del 2003, que forma parte de la ecuación con un peso de 14,01.

Observación: La variable $I_t^{10/2003}$ es definida como variable impulso. La cual toma valor *cero* en todos los meses considerados en la muestra, salvo en Octubre del 2003 que toma valor *uno*. Lo cual implica que el efecto provocado por ese valore es puntual; es decir, el suceso ocurre en un momento del tiempo determinado $t = \text{Octubre de 2003}$ y no se traslada a lo largo de la serie, como puede ocurrir cuando el cambio en la serie es permanente o transitorio. Este tipo de valor atípico puntual se denomina **atípico aditivo**.

Por lo tanto, la serie se ajusta al modelo $ARIMA(2,1,0)(1,0,0)$ en todos los instantes donde la variable impulso es cero. Para el instante Octubre de 2003 al modelo $ARIMA(2,1,0)(1,0,0)$ se le suma el efecto puntual del atípico.

$$I_t^{10/2003} = \begin{cases} 0 & t \neq 10/2003 \\ 1 & t = 10/2003 \end{cases} \quad (5.1)$$

Performance del modelo seleccionado en Predicción

El modelo seleccionado fue probado en predicción mediante:

- **Predicción en el interior de la muestra.** Procedimiento basado en quitar intervalos de tiempo en el interior de la muestra y generar la predicción de dicho intervalo.
 1. Se realiza la predicción a 12 pasos tomando como último dato observado Diciembre de 2006 (corresponde a la predicción del año 2007). Ver tabla: 5.8 y Figura: 5.6.
 2. Se anexa la predicción a 12 pasos del período Julio 2007 a Junio 2008. Ver: Anexo Univariado Gas Oil, sec: 9.4
- **Predicción fuera de la muestra**
 1. Se realiza la predicción del año 2008 tomando como último dato observado Diciembre de 2007. Dado que se dispone de información sobre el consumo de gas oil a partir de dicha fecha, se incorporan los datos en las tablas a fin de evaluar el error porcentual de predicción.
 2. Se realizan predicciones a 6 pasos para el año 2008. En esta instancia, para generar dichas predicciones se incorporan datos nuevos en la muestra, uno a uno. Ver tablas: 5.4, 5.5 y 43.

Referido a las predicciones y el valor observado, es de esperar que la observación esté dentro de los intervalos de predicción con una confianza del 95 % ó 99 % (dependiendo de la definición del intervalo que se considere). Se desprende que para una confianza mayor (99 %) se pierde precisión ya que los intervalos generados son mas amplios.

Como resultado del análisis de predicciones se obtiene:

1. Predicción para el año 2007.

No quedan observaciones fuera de los intervalos de predicción ni al 95 % ni al 99 % Ver tabla: 5.8

2. En la predicción a 12 pasos para el período Julio de 2007 - Junio de 2008, no quedan observaciones fuera de los intervalos de predicción. Ver anexo: sec: 9.4

3. Predicciones a seis pasos para el año 2008.

Se observa que el modelo no deja observaciones fuera de los intervalos de predicción; salvo Abril, que queda fuera del intervalo al 95 % en la predicción del período marzo-agosto y abril-setiembre. Si se comparan las tablas: 5.4, 5.5 y 5.6 se verá que al ingresar información en la muestra, los errores de predicción no tienen una tendencia marcada a la disminución, sino que fluctúan.

4. En lo que respecta a la predicción a 12 pasos para el año 2008 no quedan observaciones fuera de los intervalos de predicción. tabla:5.3

			Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		
	obs	pred	Inf	Sup	Inf	Sup	ERROR
Ene-07	64,8	67,56	59,21	75,90	56,58	78,53	4 %
Feb-07	70,4	66,43	57,20	75,66	54,30	78,56	-6 %
Mar-07	67,1	76,93	67,08	86,79	63,98	89,88	15 %
Abr-07	66,4	71,45	60,31	82,59	56,81	86,09	8 %
May-07	74,5	70,77	58,78	82,76	55,01	86,53	-5 %
Jun-07	72,4	66,67	53,95	79,39	49,95	83,39	-8 %
Jul-07	73,4	67,86	54,33	81,39	50,08	85,64	-8 %
Ago-07	67,4	65,68	51,42	79,93	46,95	84,40	-3 %
Sep-07	62,7	69,56	54,64	84,49	49,95	89,18	11 %
Oct-07	74,7	73,05	57,46	88,64	52,57	93,53	-2 %
Nov-07	83,98	73,23	57,02	89,45	51,92	94,55	-13 %
Dic-07	73,02	72,92	56,10	89,74	50,82	95,02	0 %
ECM = 32,61							

Tabla 5.2: Predicción año 2007 - modelo univariado seleccionado: gas oil

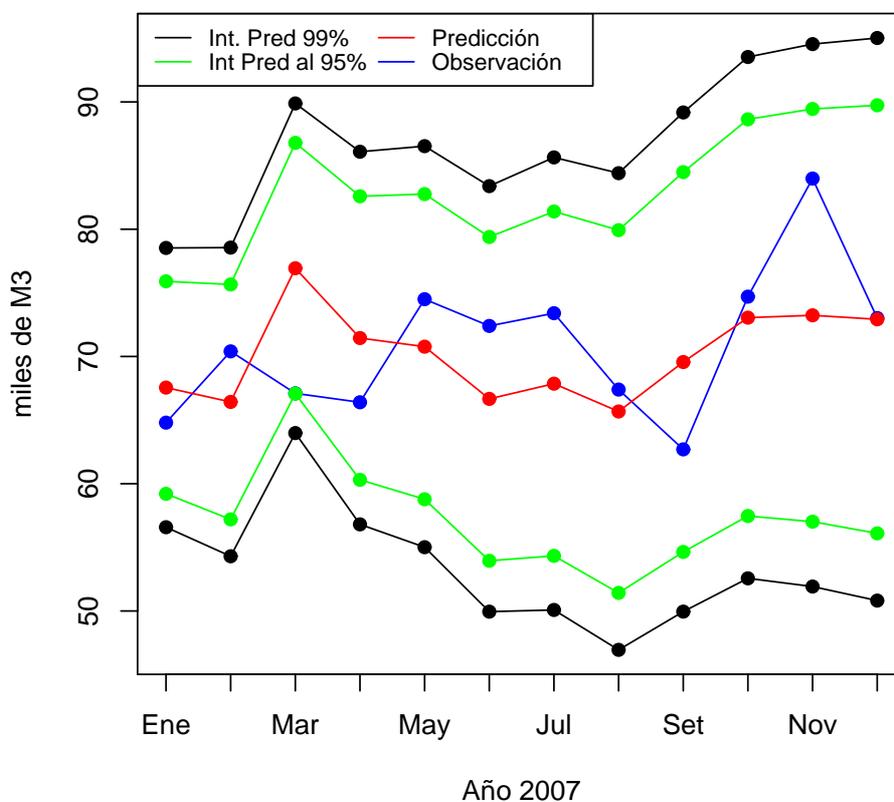


Figura 5.6: Predicción año 2007 - modelo univariado seleccionado: gas oil

		Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %			
	obs	pred	Inf	Sup	Inf	Sup	ERROR
Feb-08	67,37	76,01	67,02	85,00	64,20	87,82	13 %
Mar-08	72,51	72,23	62,10	82,35	58,92	85,53	0 %
Abr-08	81,64	71,29	60,64	81,94	57,29	85,29	-13 %
May-08	73,20	76,71	64,67	88,75	60,88	92,53	5 %
Jun-08	66,52	75,29	62,24	88,33	58,14	92,43	13 %
Jul-08	68,79	75,57	61,80	89,35	57,47	93,68	10 %

Tabla 5.4: Predicción a 6 pasos. Período: Febrero a Julio de 2008

		Intervalo de Predicción 95 %		Intervalo de Predicción 99 %			
	obs	pred	Inf	Sup	Inf	Sup	ERROR
Ene-08	68,99	70,58	61,55	79,61	58,71	82,45	2 %
Feb-08	67,37	76,78	66,63	86,93	63,44	90,12	14 %
Mar-08	72,51	72,79	62,10	83,48	58,74	86,84	0 %
Abr-08	81,64	72,29	60,20	84,38	56,40	88,18	-11 %
May-08	73,2	77,54	64,45	90,63	60,34	94,74	6 %
Jun-08	66,52	76,02	62,20	89,85	57,85	94,19	14 %
Jul-08	68,79	76,41	61,72	91,10	57,10	95,72	11 %
Ago-08	65,01	73,40	57,91	88,90	53,04	93,77	13 %
Sep-08		70,83	54,62	87,04	49,53	92,13	
Oct-08		77,22	60,30	94,14	54,98	99,46	
Nov-08		82,25	64,65	99,86	59,11	105,39	
Dic-08		76,36	58,11	94,62	52,37	100,36	

ECM = 50,01

Tabla 5.3: Predicción año 2008 - modelo univariado seleccionado: gas oil- modelo univariado seleccionado

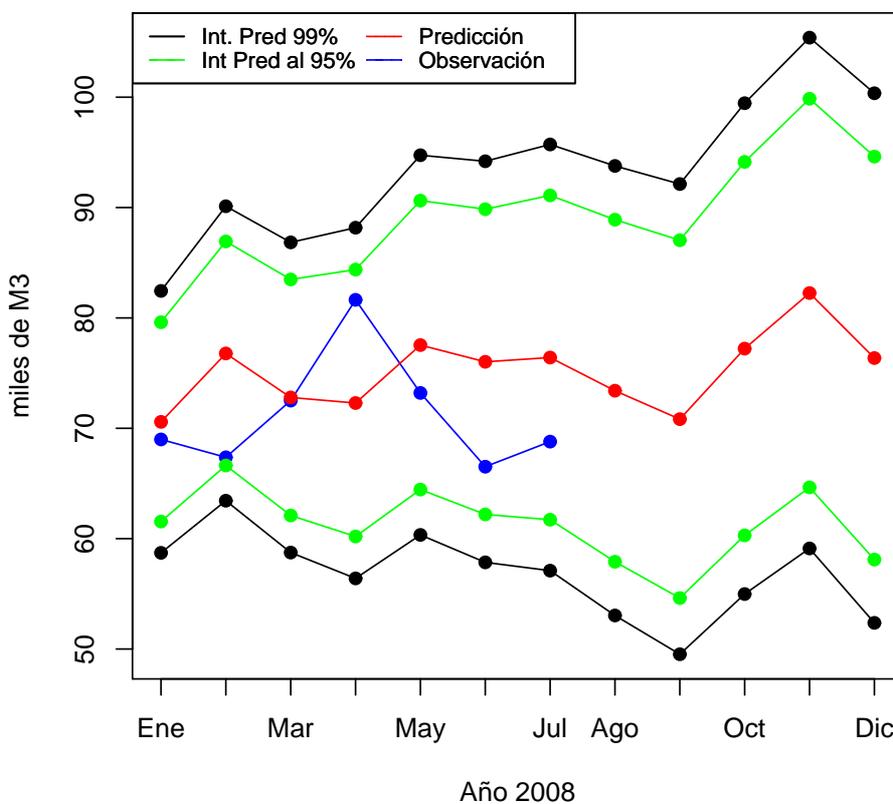


Figura 5.7: Predicción año 2008 - modelo univariado seleccionado: gas oil

			Intervalo de Predicción al 95 %		Intervalo de Predicción al 99 %		ERROR
	obs	pred	Inf	Sup	Inf	Sup	
Mar-08	72,51	67,72	58,63	76,82	55,77	79,68	-7 %
Abr-08	81,64	67,95	57,66	78,23	54,42	81,47	-17 %
May-08	73,20	71,27	60,34	82,20	56,91	85,64	-3 %
Jun-08	66,52	70,42	58,09	82,74	54,22	86,62	6 %
Jul-08	68,79	71,18	57,82	84,53	53,63	88,73	3 %
Ago-08	65,01	67,76	53,61	81,92	49,16	86,37	4 %

Tabla 5.5: Predicción a 6 pasos. Período: Marzo a Agosto 2008

			Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		ERROR
	obs	pred	Inf	Sup	Inf	Sup	
Abr-08	81,64	70,43	61,34	79,53	58,48	82,39	-14 %
May-08	73,20	73,16	62,90	83,42	59,67	86,64	0 %
Jun-08	66,52	73,39	62,50	84,29	59,07	87,71	10 %
Jul-08	68,79	73,83	61,53	86,12	57,67	89,98	7 %
Ago-08	65,01	70,22	56,91	83,52	52,73	87,70	8 %
set-08	-	67,95	53,85	82,06	49,42	86,49	-

Tabla 5.6: Predicción a 6 pasos. Período: Abril a Setiembre de 2008

5.2. Resultados referidos a la demanda de SUPERGAS

5.2.1. Datos

La serie que se presenta, corresponde al consumo de supergas (en miles de metros cúbicos) en Uruguay en el período Enero de 1999 a Diciembre 2007. Los datos tienen periodicidad mensual y son proporcionados por ANCAP.

5.2.2. Análisis Descriptivo

La figura correspondiente al consumo de supergas Figura: 5.8, sugiere una fuerte estacionalidad caracterizada por consumos elevados en los meses de invierno que se cotraponen con bajas de consumo en los meses de verano. La estructura estacional se mantiene a lo largo del período considerado, aunque las máximas de consumo, que caracterizan al invierno, han disminuido su volumen a partir del año 2004. Esto refleja la existencia de una serie con varianza no constante lo que lleva a que la misma sea no estacionaria. Dado esto, la serie debe ser transformada. La familia de transformaciones de Box - Cox⁶ tiene como finalidad corregir la falta de normalidad y la heterocedasticidad en los datos. Por lo tanto, se realiza dicha transformación tomando lambda (λ) igual a 0; lo que equivale a analizar la serie en *logaritmos*.

De forma de observar claramente lo antes comentado se presenta la gráfica de consumo de supergas por año figura: 5.9. La estacionalidad es evidente, mostrando consumos mínimos del producto (en torno a los 10 mil metros cúbicos) en los períodos comprendidos en: Enero a Marzo y Octubre a Diciembre. Los picos máximos de consumo en el entorno a los 27 mil metros cúbicos, se registran en el período de Abril a Septiembre, para los años 1999 al 2002 inclusive. En los años siguientes, el consumo máximo disminuye, estacionándose en un valor cercano a los 17 mil metros cúbicos; con la salvedad del año 2007.

⁶La transformación se desarrolla en: Anexo Metodológico: Transformación de Box - Cox, sección: 8.5

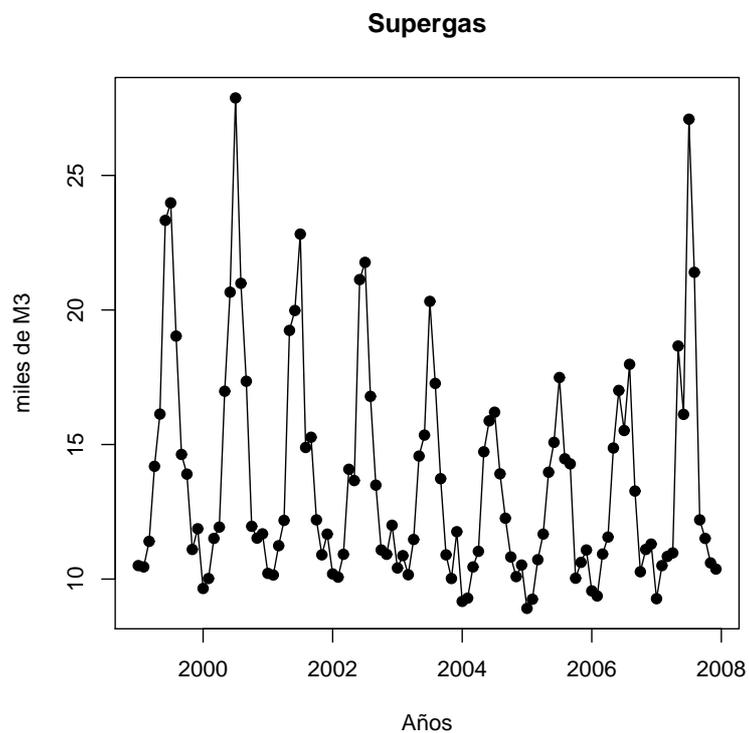


Figura 5.8: Consumo de Supergas en el período Enero de 1999 a Diciembre de 2007 en Uruguay. Unidad: miles de metros cúbicos

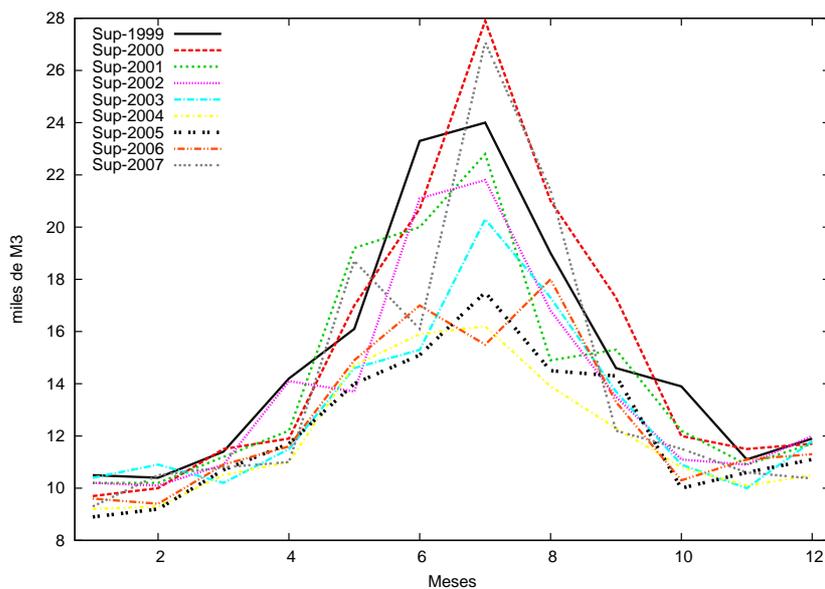


Figura 5.9: Consumo de Supergas por año. Período: Enero de 1999 a Diciembre de 2007

Para analizar la estacionariedad de la serie se realizaron: el Test de Dickey Fuller, los autocorrelogramas simple y parcial y el gráfico de la serie en primera diferencia.

Como resultado se obtiene:

1. Respecto al test de Dickey Fuller:

Al realizar el *Test de Dickey Fuller* al logaritmo del consumo de supergas se concluye no rechazar la existencia de raíz unitaria, con p-valor igual a 0,1 Ver figura: 12.5.

Si bien, en las figuras 5.8 y 5.9 no se evidencia una tendencia clara en los datos el no rechazo de la hipótesis puede deberse al efecto de una fuerte estacionalidad.

2. Respecto a los autocorrelogramas:

Los autocorrelogramas simple y parcial de la serie en logaritmos dejan en evidencia, retardos estacionales significativos prolongados en el tiempo. Por lo tanto, se decide aplicar una diferencia estacional antes de aplicar una diferencia regular. Se adjuntan los autocorrelogramas correspondientes en el Anexo: Univariado Supergas: 10.1.

A continuación se presentan los autocorrelogramas del logaritmo del consumo de Supergas con diferencia estacional.

```

-----
Augmented Dickey & Fuller test
-----
Deterministic regressors estimates:

None selected.
-----
ADF regressor estimate:
      Estimate Std. Error t value Pr(>|t|)
adf.reg  -0.005     0.004  -1.322    0.1
-----
Selected lags estimates:
      Estimate Std. Error t value Pr(>|t|)
Lag.1   -0.777     0.096  -8.062   0.000
Lag.2   -0.663     0.102  -6.482   0.000
Lag.3   -0.724     0.092  -7.843   0.000
Lag.4   -0.908     0.085 -10.635   0.000
Lag.5   -0.795     0.115  -6.904   0.000
Lag.6   -0.649     0.113  -5.739   0.000
Lag.7   -0.793     0.095  -8.382   0.000
Lag.8   -0.865     0.084 -10.318   0.000
Lag.9   -0.642     0.095  -6.789   0.000
Lag.10  -0.704     0.101  -6.952   0.000
Lag.11  -0.540     0.100  -5.393   0.000
Lag.17   0.118     0.102   1.159   0.250
Lag.18  -0.006     0.103  -0.054   0.957

Number of available observations: 83

```

Figura 5.10: Salida correspondiente al Test de Dickey Fuller sobre el logaritmo del consumo de Supergas. Período Enero de 1999 a Junio de 2007

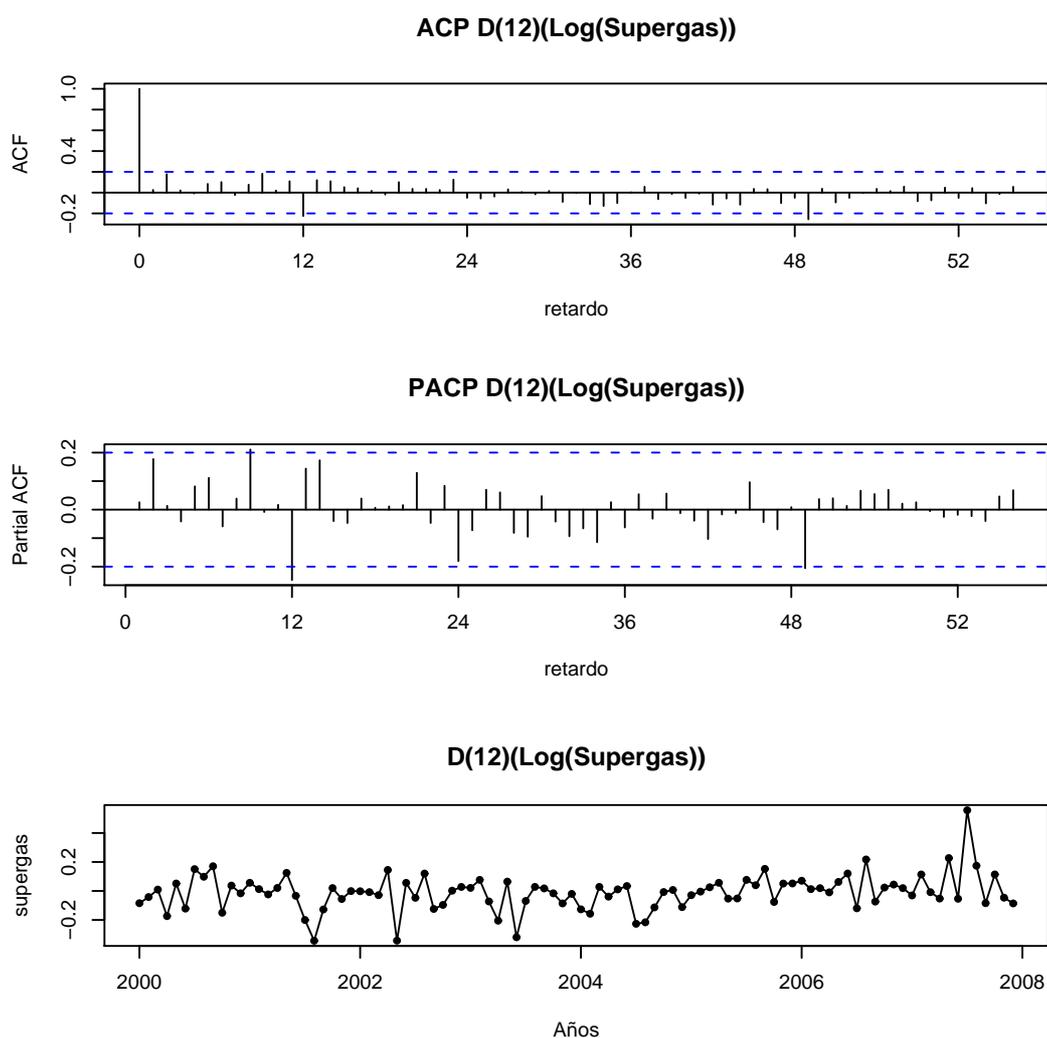


Figura 5.11: Autocorrelogramas y gráfica de la diferencia estacional del logaritmo de Supergas

La figura 5.11 muestra el retardo 12 significativo en el autocorrelograma simple y en el parcial; lo cual sugiere que la parte estacional debe ser modelizada. A su vez, en ambos gráficos queda el retardo 9 por fuera de las bandas de confianza. No se evidencian retardos significativos que supongan la modelización de la parte regular. La serie en diferencias parece ser estacionaria ya que oscila en torno a su nivel medio. Existen desviaciones importantes respecto a la media; puntos que podrían tratarse como posibles atípicos ver: figura: 5.11, gráfico: D(12)(Ln(supergas)).

Según lo anterior, se proponen los siguientes modelos para el consumo de supergas:

1. SARIMA(0,0,0)(0,1,1)

2. SARIMA(0,0,0)(0,1,2)
3. SARIMA(0,0,0)(1,1,1)
4. SARIMA(0,0,0)(1,1,0)
5. SARIMA(0,0,1)(0,1,1)
6. SARIMA(1,0,0)(0,1,1)
7. SARIMA(0,0,1)(1,1,1)
8. SARIMA(1,0,1)(0,1,1)

5.2.3. Modelo Seleccionado

Los modelos propuestos fueron estimados, resultando únicamente con coeficientes significativos: los modelos SARIMA(0,0,0)(1,1,0) y SARIMA(0,0,0)(0,1,1). A medida que se realizaban las distintas etapas de la metodología *Box - Jenkins* surgieron indicios de que la dinámica de la serie no había sido totalmente captada por la modelización propuesta, por lo que fue necesario estimar nuevos modelos. Es por ello que el modelo resultante es un SARIMA(2,0,0)(2,1,0).

El camino que lleva a la selección de este modelo se puede observar en el Anexo: Univariado Supergas: sección: 10.1.

El modelo **SARIMA(2,0,0)(2,1,0)** con atípicos aditivos en: **Agosto de 2001, Mayo de 2002, Junio de 2003, Mayo, Julio y Agosto de 2007** es elegido para la modelización de la serie de supergas.

Estimación y Validación del Modelo Final

Coefficientes:

Referente a la significación individual de los coeficientes del modelo estimado, se observa en la tabla: 5.7 que todos los coeficientes son significativos al 5 %. Dada la prueba t de significación de los coeficientes, en la cual se acepta que cada coeficiente es significativamente distinto de cero si se cumple que: $|t_{(obs)}| > t_{critico}$.

SARIMA(2,0,0)(2,1,0)	coef	t-obs	critico
ar1	0.28	2.70	1,66
ar2	0.35	3.52	
sar1	-0.54	-4.06	
sar2	-0.40	-3.38	
Ago-2001	-0.18	-2.86	
May-2002	-0.19	-2.99	
Jun-2003	-0.16	-2.64	
May-2007	0.25	3.34	
Jul-2007	0.54	7.12	
Ago-2007	0.29	3.82	
aic = -190.26			
Shapiro-Wilk p-value = 0.1326			
Box-Ljung, df=24, p-value = 0.683			

Tabla 5.7: Modelo seleccionado

Normalidad:

El test de Shapiro - Wilk lleva a no rechazar la hipótesis nula de normalidad de los residuos, con un p-valor de 0,1326.

Autocorrelación:

La prueba de Ljung - Box realizada tomando k igual a 24, resulta positiva para aceptar la hipótesis nula de ausencia de autocorrelación en los residuos, con un p-valor de 0,683.

En la figura 5.12 se observa que las barras de los coeficientes de autocorrelación permanecen dentro de las bandas de confianza, lo que verifica el resultado ya obtenido en el test de Ljung - Box y evidenciando que la dinámica de la serie ha sido captada por la modelización propuesta. Los residuos fluctúan en torno a su nivel medio.

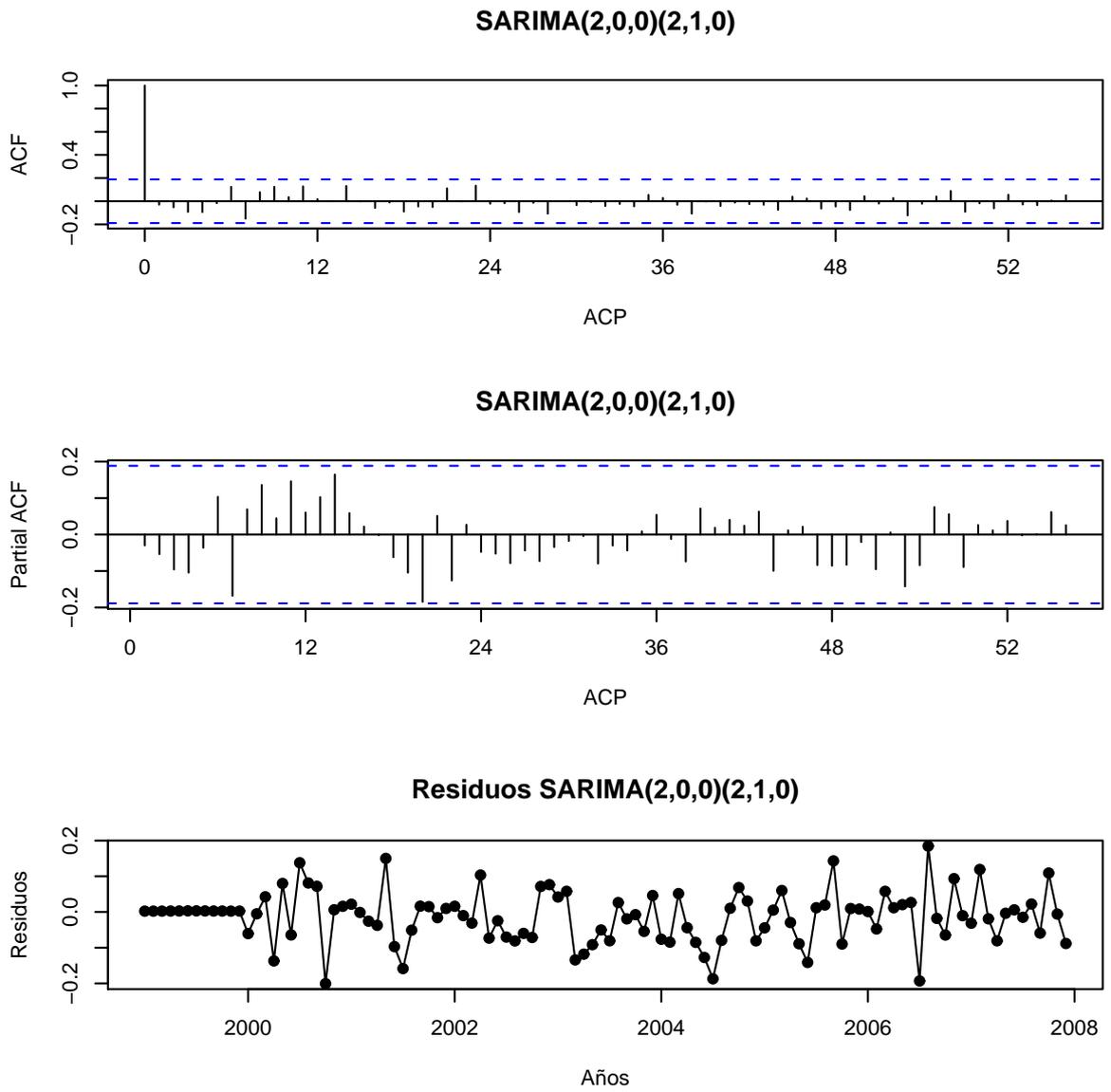


Figura 5.12: Autocorrelación Simple y Parcial del modelo seleccionado

Interpretación del modelo seleccionado

Sustituyendo en la ecuación 4.1 (de la sección: *Aspecto Metodológico Estadístico*) los valores correspondientes a los coeficientes obtenidos en la estimación del modelo se plantea:

$$\begin{aligned} \hat{\log}(y_t) = & 0,28 \log(y_{t-1}) + 0,35 \log(y_{t-2}) + 0,46 \log(y_{t-12}) - 0,13 \log(y_{t-13}) - 0,16 \log(y_{t-14}) + \dots \\ & \dots + 0,14 \log(y_{t-24}) - 0,04 \log(y_{t-25}) - 0,05 \log(y_{t-26}) - 0,35 \log(y_{t-36}) - 0,11 \log(y_{t-37}) - \dots \\ & \dots - 0,14 \log(y_{t-38}) - 0,18 I_t^{8/2001} - 0,19 I_t^{5/2002} - 0,16 I_t^{6/2003} + 0,25 I_t^{5/2007} + 0,54 I_t^{7/2007} + \dots \\ & \dots + 0,29 I_t^{8/2007} \end{aligned}$$

Por lo tanto, la observación (en logaritmos) en el momento t ($\hat{\log}(y_t)$) depende de:

- La observación del mes anterior en logaritmos ($\log(y_{t-1})$), cuyo peso en la generación de la nueva observación es de 0.28.
- La observación en logaritmos de dos meses atrás ($\log(y_{t-2})$), con peso de 0.35.
- La observación en logaritmos de 12 meses atrás ($\log(y_{t-12})$). Con peso igual a 0.46.
- La observación en logaritmos de 13 meses atrás ($\log(y_{t-13})$). Con peso igual a -0.13.
- La observación en logaritmos de 14 meses atrás ($\log(y_{t-14})$). Con peso igual a -0.16.
- La observación en logaritmos de 24 meses atrás ($\log(y_{t-24})$). Con peso igual a -0.14.
- La observación en logaritmos de 25 meses atrás ($\log(y_{t-25})$). Con peso igual a -0.04.
- La observación en logaritmos de 26 meses atrás ($\log(y_{t-26})$). Con peso igual a -0.05.

- La observación en logaritmos de 36 meses atrás ($\log(y_{t-36})$). Con peso igual a -0.35.
- La observación en logaritmos de 37 meses atrás ($\log(y_{t-37})$). Con peso igual a -0.11.
- La observación en logaritmos de 24 meses atrás ($\log(y_{t-38})$). Con peso igual a -0.14.

Es decir $\log(\hat{y}_t)$ depende de la observación del mes pasado y el anterior ($\log(y_{t-1})$ y $\log(y_{t-2})$). Más: la observación del año pasado ($\log(y_{t-12})$) y el efecto de dos meses más ($\log(y_{t-13})$, $\log(y_{t-14})$). Más la observación de dos años atrás ($\log(y_{t-24})$) y el efecto dos meses más ($\log(y_{t-25})$ y $\log(y_{t-26})$). Más: la observación de tres años atrás ($\log(y_{t-36})$) y el efecto dos meses más ($\log(y_{t-37})$ y $\log(y_{t-38})$). Cada una de las variables con sus respectivos pesos.

Se le agrega a la ecuación el efecto de los atípicos:

- $I_t^{8/2001}$: Valor atípico: Agosto de 2001, que forma parte de la ecuación con un peso de -0.18.
- $I_t^{5/2002}$: Valor atípico: Mayo de 2002, que ingresa a la ecuación con un peso de -0.22.
- $I_t^{6/2003}$: Valor atípico: Junio de 2003, que ingresa a la ecuación con un peso de -0.16.
- $I_t^{5/2007}$: Valor atípico: Mayo del 2007, que ingresa a la ecuación con un peso de 0.25.
- $I_t^{7/2007}$: Valor atípico: Mayo del 2007, que ingresa a la ecuación con un peso de 0.54.
- $I_t^{8/2007}$: Valor atípico: Mayo del 2007, que ingresa a la ecuación con un peso de 0.29.

Observación 2: Las variables $I_t^{8/2001}$, $I_t^{5/2002}$, $I_t^{6/2003}$, $I_t^{5/2007}$, $I_t^{7/2007}$ y $I_t^{8/2007}$ son definidas como variables impulso. Las cuales toman valor *cero* en todos los meses considerados en la muestra, salvo en Agosto de 2001, Mayo de 2002, Junio de 2003, Mayo, Julio y

Agosto del 2007 que toman el valor *uno*. Esto implica que el efecto de esos valores atípicos es puntual para el momento del tiempo definido, no se traslada a lo largo de la serie. Los atípicos con estas características se conocen como *atípicos aditivos*.

Por lo tanto, se interpreta que la serie se ajusta al modelo SARIMA(2,0,0)(0,1,1) en todos los instantes donde las variables impulso son cero. Para el instante Agosto de 2001 al modelo SARIMA(2,0,0)(0,1,1) se le suma el efecto puntual del atípico (lo mismo sucede para el resto de los atípicos mencionados).

Performance del modelo seleccionado en Predicción

Al igual que en el análisis de la serie de gas oil el modelo se prueba en predicción mediante:

- **Predicción en el interior de la muestra.**

Procedimiento basado en quitar intervalos de tiempo en el interior de la muestra y generar la predicción de dicho intervalo.

1. Se realiza la predicción a 12 pasos tomando como último dato observado Diciembre de 2006. Corresponde a la predicción del año 2007.
2. Se anexa la predicción del período Julio de 2007 - Junio de 2008.

Para ambos casos, se comparan las predicciones del modelo con el valor real observado calculando el error porcentual ocurrido en cada punto. Referido a las predicciones y el valor observado, es de esperar que el valor real esté dentro de los intervalos de predicción con una confianza del 95 % ó 99 % (dependiendo de la definición del intervalo que se considere). Se desprende que, para una confianza mayor (99 %) se pierde precisión ya que los intervalos generados son mas amplios.

■ Predicción fuera de la muestra.

1. Predicción del año 2008 tomando como ultimo dato observado Diciembre de 2007. Dado que se dispone de información sobre el consumo de supergas a partir de esa fecha se incorporan los datos a las tablas con el fin evaluar el error porcentual de predicción.
2. Predicciones a 6 pasos para el año 2008. En esta instancia los datos disponibles del 2008 se incorporan a la muestra uno a uno.

Como resultado del análisis de predicciones se obtiene:

1. Referente a la predicción para el año 2007 Ver tabla: 5.8 , se ve que las observaciones de Julio y Agosto quedan fuera de los intervalos de predicción. Ambas presentan los mayores errores porcentuales respecto al valor observado: 40 % y 26 % respectivamente. Estas observaciones no son rechazadas como atípicos en el modelo. Queda comprendida en el intervalo al 99 % la observación de Mayo de 2007.
2. En las predicciones para el intervalo: Julio 2007 a Junio 2008 quedan las observaciones de Julio y Agosto de 2007 fuera de los intervalos de predicción. Se observa que el modelo no fue capaz de corregir estas predicciones sino que el error porcentual al ingresar información en la muestra aumenta un 1 % (en comparación a la predicción del año 2007). Ver en Anexo:Univariado Supergas: sec: 10.3
3. Las predicciones generadas para el año 2008 son bastante cercanas a las observaciones. El error máximo incurrido es de 19 % en la observación de Junio; la cual queda por fuera del intervalo al 95 %. tabla: 5.9
4. En las tablas de predicción a seis pasos se observa que a medida que se ingresan observaciones a la muestra el modelo no genera grandes cambios en cuanto al error porcentual incurrido. En comparación con la predicción completa para el año 2008 si hay cambios: disminuye el error. La observación de junio de 2008 queda fuera del intervalo de predicción al 95 % en las tres ventanas de tiempo consideradas. Ver tablas: 5.10, 5.11 y 5.12

			Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		ERROR
	Sup	pred	Inf	Sup	Inf	Sup	
Ene-07	9,27	9,58	8,20	11,20	7,80	11,76	3 %
Feb-07	10,5	9,43	8,02	11,09	7,62	11,67	-10 %
Mar-07	10,84	10,88	9,12	12,98	8,63	13,72	0 %
Abr-07	10,97	11,50	9,60	13,76	9,07	14,57	5 %
May-07	18,66	14,75	12,28	17,73	11,59	18,78	-21 %
Jun-07	16,12	16,31	13,55	19,62	12,78	20,80	1 %
Jul-07	27,09	16,20	13,45	19,52	12,68	20,70	-40 %
Ago-07	21,4	15,82	13,12	19,07	12,38	20,22	-26 %
Sep-07	12,2	13,14	10,90	15,84	10,27	16,80	8 %
Oct-07	11,51	10,43	8,64	12,57	8,15	13,34	-9 %
Nov-07	10,6	10,66	8,84	12,86	8,34	13,64	1 %
Dic-07	10,37	11,00	9,12	13,26	8,59	14,07	6 %

ECM = 14,09

Tabla 5.8: Predicción año 2007 - modelo univariado seleccionado para supergas

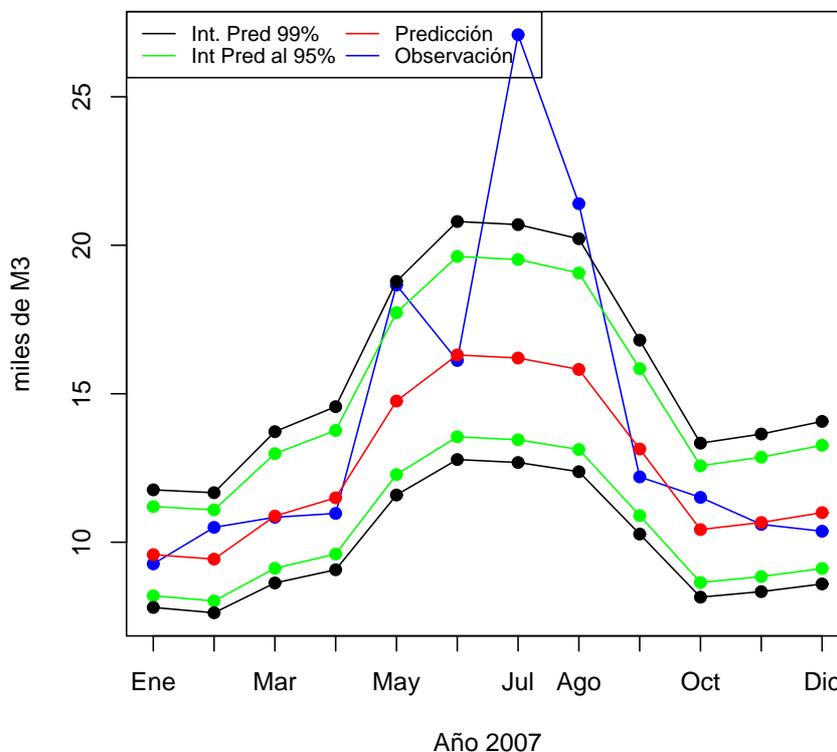


Figura 5.13: Predicción año 2007 - modelo univariado seleccionado para supergas

	Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		ERROR
	Inf	Sup	Inf	Sup	
Ene-08	7,75	10,50	7,39	11,01	-9 %
Feb-08	8,19	11,22	7,80	11,79	1 %
Mar-08	9,00	12,66	8,53	13,36	-6 %
Abr-08	9,41	13,32	8,91	14,07	-7 %
May-08	11,94	17,01	11,30	17,98	-4 %
Jun-08	13,15	18,78	12,44	19,86	-19 %
Jul-08	13,64	19,51	12,90	20,64	-1 %
Ago-08	12,99	18,59	12,28	19,67	-3 %
Sep-08	10,96	15,70	10,36	16,61	
Oct-08	8,93	12,80	8,44	13,54	
Nov-08	8,91	12,76	8,42	13,50	
Dic-08	9,00	12,89	8,50	13,64	

ECM = 1,97

Tabla 5.9: Predicción año 2008 - modelo univariado seleccionado para supergas

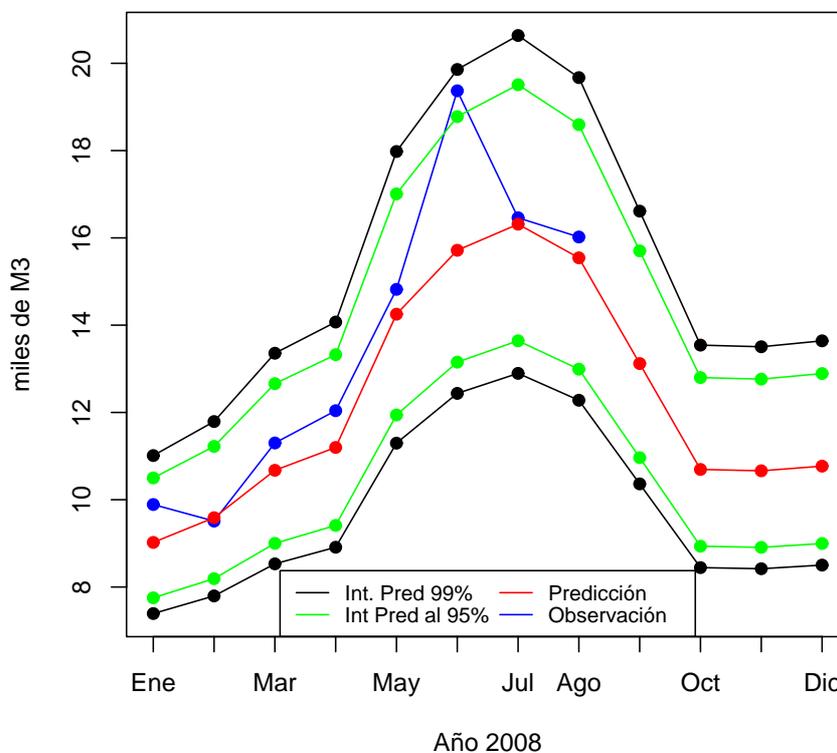


Figura 5.14: Predicción: año 2008

			Intervalo de Predicción 95 %		Intervalo de Predicción 99 %		
	Sup	pred	Inf	Sup	Inf	Sup	ERROR
Feb-08	9,51	9,82	8,43	11,43	8,04	11,99	3 %
Mar-08	11,3	11,11	9,49	13,00	9,03	13,66	-2 %
Abr-08	12,04	11,42	9,63	13,55	9,13	14,29	-5 %
May-08	14,82	14,54	12,22	17,30	11,57	18,27	-2 %
Jun-08	19,37	15,93	13,35	19,01	12,63	20,09	-18 %
Jul-08	16,46	16,48	13,79	19,69	13,04	20,82	0 %

Tabla 5.10: Predicción a 6 pasos año 2008 Febrero - Julio

			Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		
	Sup	pred	Inf	Sup	Inf	Sup	ERROR
Mar-08	11,3	11,02	9,48	12,82	9,04	13,45	-2 %
Abr-08	12,04	11,28	9,64	13,19	9,18	13,85	-6 %
May-08	14,82	14,45	12,19	17,12	11,55	18,06	-3 %
Jun-08	19,37	15,83	13,31	18,82	12,61	19,87	-18 %
Jul-08	16,46	16,42	13,77	19,58	13,03	20,70	0 %
Ago-08	16,02	15,64	13,09	18,67	12,38	19,74	-2 %

Tabla 5.11: Predicción a 6 pasos año 2008 Marzo - Agosto

			Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		
	Sup	pred	Inf	Sup	Inf	Sup	ERROR
Abr-08	12,04	11,35	9,77	13,20	9,32	13,84	-6 %
May-08	14,82	14,61	12,50	17,07	11,90	17,92	-1 %
Jun-08	19,37	15,91	13,43	18,85	12,74	19,88	-18 %
Jul-08	16,46	16,52	13,90	19,63	13,17	20,72	0 %
Ago-08	16,02	15,69	13,16	18,70	12,45	19,76	-2 %
Sep-08	-	13,23	11,08	15,79	10,48	16,69	-

Tabla 5.12: Predicción a 6 pasos año 2008 Abril - Setiembre

Capítulo 6

Resultados: Modelos Multivariados

El siguiente capítulo procura analizar si existe una relación de largo plazo entre variables que se suponen vinculadas a los productos aquí estudiados y dichos productos. Con tal propósito se estudia si las variables están cointegradas mediante el test de Johansen y posteriormente se aplican los modelos de corrección de error (VECM), que surgen como una alternativa de pronóstico a los modelos ARIMA planteados en el capítulo anterior.

6.1. Resultados referidos a la demanda de Gas Oil

6.1.1. Datos

Los datos que se disponen para el análisis de la *demanda de gas oil en Uruguay* son:

- Consumo de gas oil en el mercado interno
- Lluvias promedio del país
- IVFPBI del agro
- IVFPBI de la industria
- precio del gas oil

Las series conforman una muestra de periodicidad mensual en el período Enero de 1999 a Diciembre de 2007 (102 observaciones).

Los datos vinculados al clima, el consumo de gas oil y los precios de dicho producto son proporcionados por ANCAP, mientras que los datos del PBI de la industria y del

agro se obtienen del Banco Central del Uruguay. El precio es deflactado por el IPC y se encuentra expresado a precios de Junio de 2007.

6.1.2. Análisis descriptivo de las series

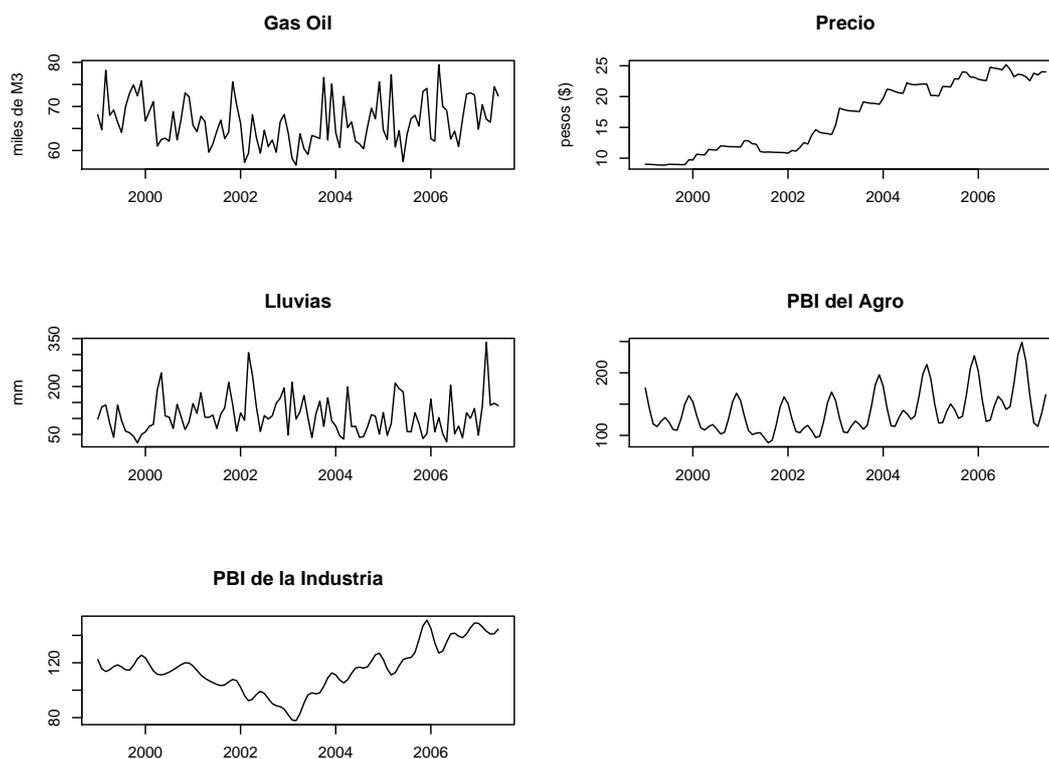


Figura 6.1: Gráficas de las variables del grupo: Gas Oil

La figura: 6.1 presenta las relaciones existentes entre las variables del grupo gas oil. Principalmente se distinguen dos períodos: antes y después del 2004.

Antes del 2004:

En el caso del gas oil y el PBI de la industria se nota que ambas series siguen la misma dinámica a lo largo del tiempo. Existe una retracción en el consumo hasta principio del 2004 lo que condice con la disminución del PBI de la industria y de forma menos notoria del PBI del agro. También se puede observar que en dicho período, el precio sufre incrementos relativos superiores. Y si se lo compara con las lluvias se puede ver que las mismas son superiores en ese período. El determinante que subyace es *el agro*. El aumento de las lluvias disminuye la necesidad de riego, como sucede con el cul-

tivo del arroz, retrasa las cosechas e incluso genera pérdidas. Esto provoca que no se destine gas oil a las bombas de riego, que la maquinaria a gas oil no este activa o si lo este (si es que se adelanta la siembra porque se prevén lluvias) y que el crecimiento del sector disminuya debido a las pérdidas que se generan.

Dado que el consumo de gasoil es inelástico al precio, la tendencia que experimenta es muy leve, en comparación a la tendencia al alza sostenida del precio (ver figura: 6.2). El agro lo necesita como insumo y lo va a seguir consumiendo, principalmente porque no puede sustituirlo a menos que ocurra un cambio tecnológico en el sector.

Después del 2004: Lo que sucede en este período es todo lo contrario a lo que sucede en el período que lo precede. Tanto el consumo, como el PBI se recuperan. Los incrementos porcentuales en los precios comienzan a ser menores y las lluvias disminuyen.

Lo antes comentado da la pauta de que las series no puedan tratarse como variables estacionarias. A su vez, el PBI del agro y de la industria muestra un marcado patrón estacional y el precio tendencia.

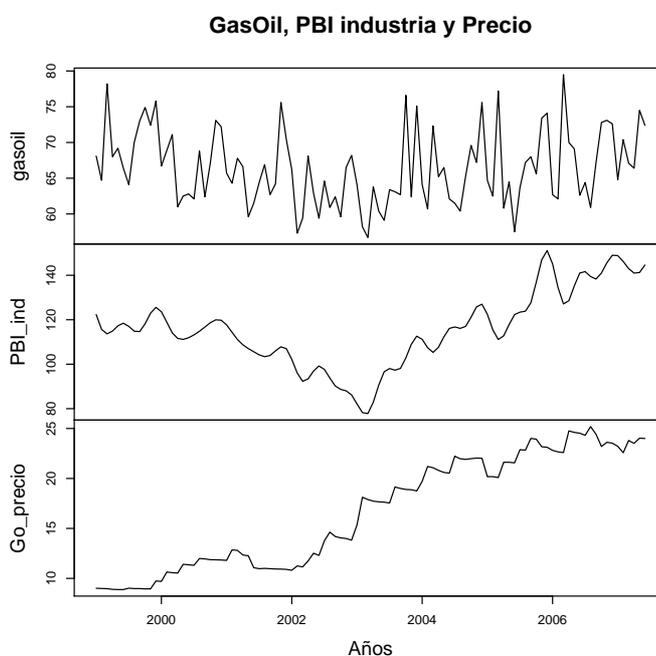


Figura 6.2: Gas Oil, PBI de la Industria y Precio

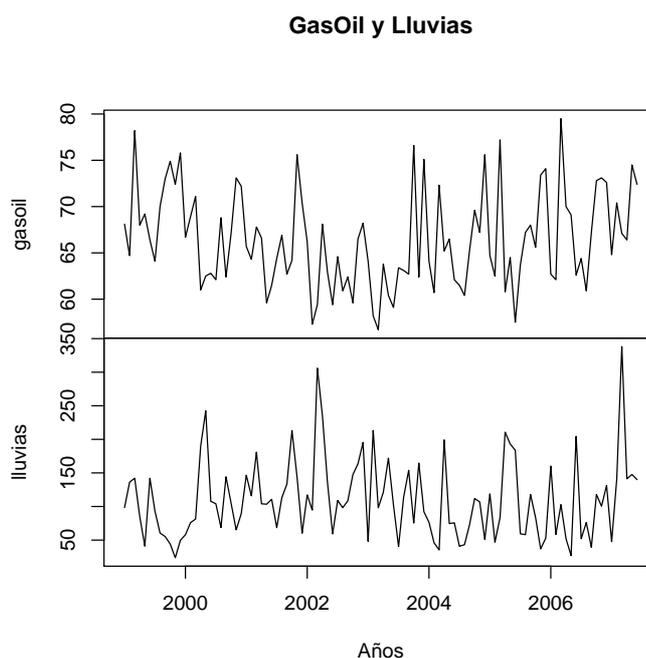


Figura 6.3: Gas Oil y Lluvias

6.1.3. Modelo seleccionado

Previo al modelo presentado en esta sección cabe acotar que se realizaron varias pruebas sobre posibles modelizaciones para el conjunto de variables del grupo gas oil. Se consideran modelos con constante en la relación de cointegración (y sin ella), con y sin variables indicatrices estacionales y con varias alternativas en la cantidad de retardos a incluir. Para aquellos modelos que pasan las pruebas de diagnóstico, se los prueba en predicción buscando obtener aquel modelo que minimice el error cuadrático medio.

Se elige para la modelización multivariada del consumo de Gas Oil el modelo VECM que incluye:

- 7 retardos
- especificado sin constante.
- se ingresan variables indicatrices estacionales.
- variables: gas oil, precio y lluvias.
- valores atípicos en: Octubre de 2003 y Marzo de 2002.

Análisis de Cointegración

Como se comenta en el marco teórico, las variables deben integrar de igual orden para que exista cointegración. Para analizar este aspecto se realiza el test de Dickey Fuller a cada una de las variables.

Los resultados obtenidos son:

- El gas oil, el PBI y el precio son variables integradas de orden 1 y las lluvias integradas de orden 0. Ver Anexo: Multivariado del grupo Gas Oil: cap: 12.

Se realiza el *test de Johansen* para evaluar la existencia o no de cointegración.

Contraste de Johansen:					
Número de ecuaciones = 3					
Orden del retardo = 8					
Periodo de estimación: 1999:09 - 2007:12 (T = 100)					
Caso 1: Sin constante					
Rango	Valor propio	Contraste de la traza	valor p	Contraste L _{máx}	valor p
0	0,20945	35,765	[0,0009]	23,502	[0,0048]
1	0,07015	12,263	[0,0504]	7,273	[0,2321]
2	0,04867	4,989	[0,0287]	4,989	[0,0303]

Se concluye que no se rechaza la hipótesis nula de existencia de una única relación

de cointegración¹. La afirmación es clara en el estadístico del *valor propio*, con p-valor de 0.23. El test fue especificado de igual forma a como se especificó el modelo final.²

El test debido a que es en si un modelo VECM, no sólo permite concluir que las variables tengan tendencia común en el largo plazo (aunque no integren de igual orden) y en ese caso estén cointegradas, sino que permite escoger la estrategia de modelización, basándose en el rango de la matriz de datos.

Estimación y validación del modelo seleccionado.

Significación del coeficiente del termino de corrección de error

El término de corrección del error (EC1) resulta significativo en el análisis. Lo que permite concluir que el modelo corrige los desequilibrios a la baja en el corto plazo. Por ejemplo, si existe un desequilibrio el mes pasado, el modelo corrige un 10 % de ese desequilibrio. Ver tabla: 6.1 coeficiente EC1.

El término de corrección de error queda determinado por la ecuación:

$$-0,099 \left(\text{gasoil}_{t-1} - 1,99\text{precio}_{t-1} - 0,42\text{Lluvias}_{t-1} \right)$$

Exogeneidad débil de las variables - Velocidad de ajuste

Se presentan las restricciones a cero de los coeficientes alfas del modelo para evaluar la existencia o no de endogeneidad (retroalimentación). La prueba se realiza mediante el test de razón de verosimilitudes, el cual distribuye $\chi^2_{(r(p-m))}$, donde r es el rango de integración; p son las filas de la matriz que contiene las restricciones en el alfa (en este caso 3) y m las columnas de dicha matriz (2).

Como resultado se obtiene:

No se rechaza que los coeficientes *alfa*³ de la variable precio y de la variable lluvia sean iguales a cero (a un nivel de significación del 5 %).

¹Observación: se realiza el test de Johansen a las variables gas oil y precio obteniendo como resultado que no están cointegradas. Se acepta la hipótesis de rango igual a cero. El test es especificado del igual forma que el modelo seleccionado.

²Observación: en la programación del modelo se deben especificar $k+1$ retardos para que resulten k en la salida del programa.

³velocidad de ajuste al equilibrio

Beta: vector cointegrante					
	gasoil	precio	Lluvias		
	1	-1,99	-0,42		
	0	0,52	0,08		
Alfa: Velocidad de ajuste					
	gasoil	precio	Lluvias		
	-0,09	-0,007	0,29		
Variable	Coefficiente	Desv. típica	Estadístico t	valor p	
d(gasoil)-1	-0,97	0,11	-8,56	<0,00001	***
d(gasoil)-2	-0,93	0,16	-5,75	<0,00001	***
d(gasoil)-3	-1,02	0,19	-5,33	<0,00001	***
d(gasoil)-4	-0,92	0,20	-4,64	0,00	***
d(gasoil)-5	-0,57	0,20	-2,88	0,01	***
d(gasoil)-6	-0,50	0,18	-2,83	0,01	***
d(gasoil)-7	-0,38	0,12	-3,21	0,00	***
d(precio)-1	-1,78	0,57	-3,11	0,00	***
d(precio)-2	-1,36	0,68	-1,99	0,05	*
d(precio)-3	0,01	0,68	0,01	0,99	
d(precio)-4	-2,18	0,68	-3,21	0,00	***
d(precio)-5	-0,53	0,68	-0,78	0,44	
d(precio)-6	0,50	0,68	0,73	0,47	
d(precio)-7	-1,55	0,68	-2,30	0,02	**
d(Lluvias)-1	-0,05	0,01	-3,96	0,00	***
d(Lluvias)-2	-0,05	0,01	-3,07	0,00	***
d(Lluvias)-3	-0,04	0,02	-2,90	0,01	***
d(Lluvias)-4	-0,03	0,01	-2,15	0,04	**
d(Lluvias)-5	-0,03	0,01	-1,93	0,06	*
d(Lluvias)-6	-0,02	0,01	-1,93	0,06	*
d(Lluvias)-7	-0,02	0,01	-1,87	0,07	*
Oct-2003	14,55	4,44	3,28	0,00	***
Mar-2002	-12,80	4,17	-3,07	0,00	***
S1	-8,88	2,17	-4,08	0,00	***
S2	-10,25	2,73	-3,75	0,00	***
S3	-2,47	3,08	-0,80	0,43	
S4	-9,75	2,81	-3,47	0,00	***
S5	-13,10	2,77	-4,73	0,00	***
S6	-12,34	2,64	-4,68	0,00	***
S7	-13,06	2,72	-4,80	<0,00001	***
S8	-14,84	2,92	-5,08	<0,00001	***
S9	-7,77	2,92	-2,66	0,01	***
S10	-7,45	2,95	-2,52	0,01	**
S11	-3,04	2,17	-1,40	0,16	
EC1	-0,10	0,02	-4,02	0,00	***

Tabla 6.1: Coeficiente y significación de las variables integrantes del modelo seleccionado

```
Restricción:  
a[2] = 0 Precio  
valor propio 1 = 0,194778  
Log-verosimilitud no restringida (lu) = -860,54218  
Log-verosimilitud restringida (lr) = -861,46139  
2 * (lu - lr) = 1,83842  
P(Chi-cuadrado(1) > 1,83842) = 0,175136
```

```
Restricción:  
a[3] = 0 Lluvias  
valor propio 1 = 0,205327  
Log-verosimilitud no restringida (lu) = -860,54218  
Log-verosimilitud restringida (lr) = -860,80202  
2 * (lu - lr) = 0,519687  
P(Chi-cuadrado(1) > 0,519687) = 0,470975
```

Lo que implica que no responden a las desviaciones del equilibrio en el período anterior. Por lo tanto, no son variables endógenas, no existe retroalimentación. Tiene sentido por dos razones lógicas: las lluvias no están determinadas por el gasoil y por otro lado el precio es fijado según políticas de precios monopolistas que tienen en cuenta una función de utilidad y no tanto el consumo del producto o como afecta en el consumo un cambio de precios.

Análisis de los residuos

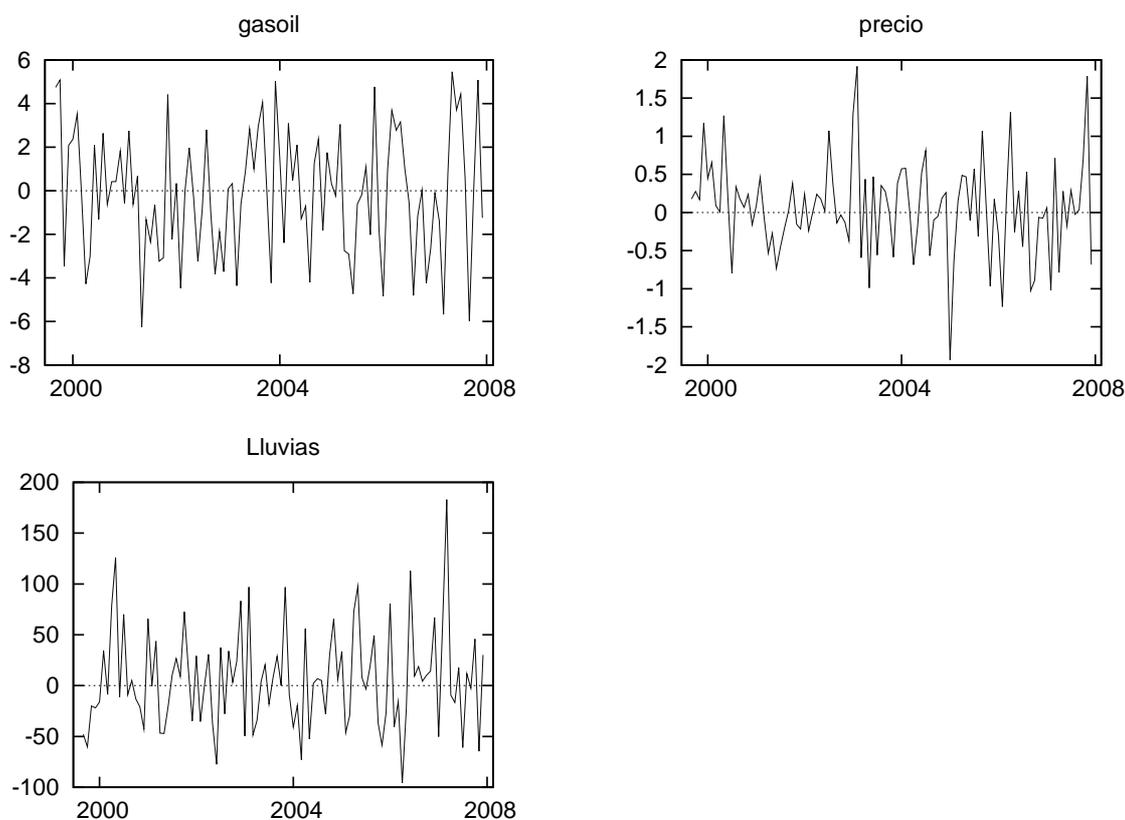


Figura 6.4: Residuos de las tres ecuaciones del sistema VECM

Dado que: la gráfica de los residuos de la ecuación precio presenta en la observación Enero de 2005 un residuo distante de la media y lo mismo para Marzo de 2007 en la gráfica de los residuos de la ecuación lluvia; se prueban los puntos: Enero de 2005 y Marzo de 2007 como posibles valores atípicos. Ambas observaciones resultaron ser rechazadas como atípicos a un nivel de significación del 5%. Dicha salida se puede ver en el Anexo: Multivariado del grupo Gas Oil, sec: 12.3

Normalidad de los residuos

El test de Normalidad de *Doornik-Hansen* realizado sobre los residuos de la ecuación de Gas Oil, lleva a no rechazar la hipótesis nula de normalidad de los mismos, con un p-valor igual a 0.4. Ver figura: 6.4. La normalidad multivariada no se rechaza a un nivel de significación del 1%:

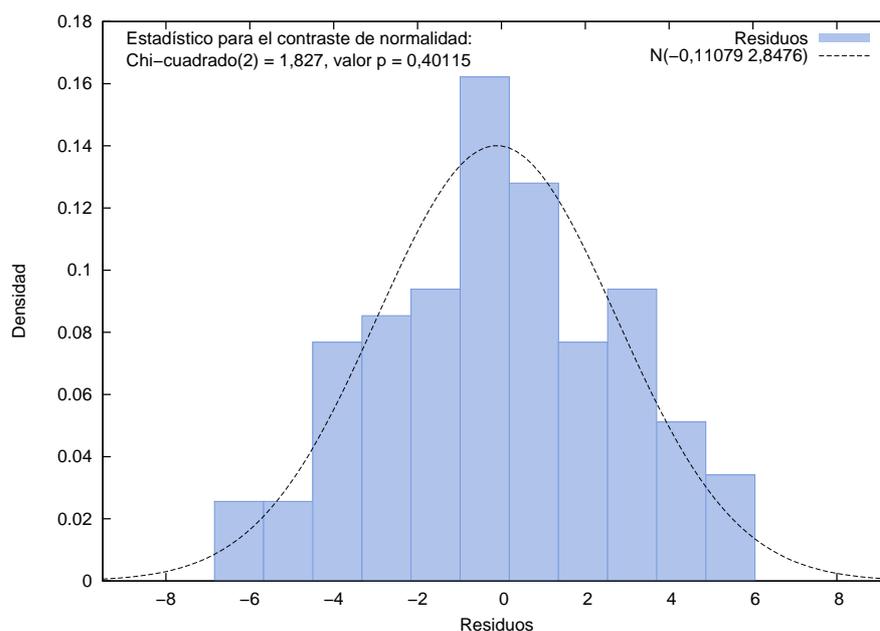


Figura 6.5: Normalidad de los residuos de la ecuación de Gas Oil

Test for multivariate normality of residuals
 Doornik-Hansen Chi-square(6) = 16,5434, with p-value = 0,011

Autocorrelación de los residuos

A continuación se presenta el test de Ljung Box, y el gráfico de los autocorrelogramas simple y parcial. Se anexa autocorrelograma simple y parcial con sus respectivos p-valores para cada retardo. Ver: Anexo: 12.1

Los resultados obtenidos son:

Se observa que el autocorrelograma simple y parcial dejan el retardo 24 por fuera de las bandas de confianza. Sin embargo, no es suficiente como para hacer rechazar la hipótesis nula de no autocorrelación en los primeros 48 retardos, con un p-valor de 0.281 (Test de Ljung-Box).

Ecuación 1: Gas Oil
 Ljung-Box $Q' = 53,2$ con valor $p = P(\text{Chi-cuadrado}(48) > 53,2) = 0,281$

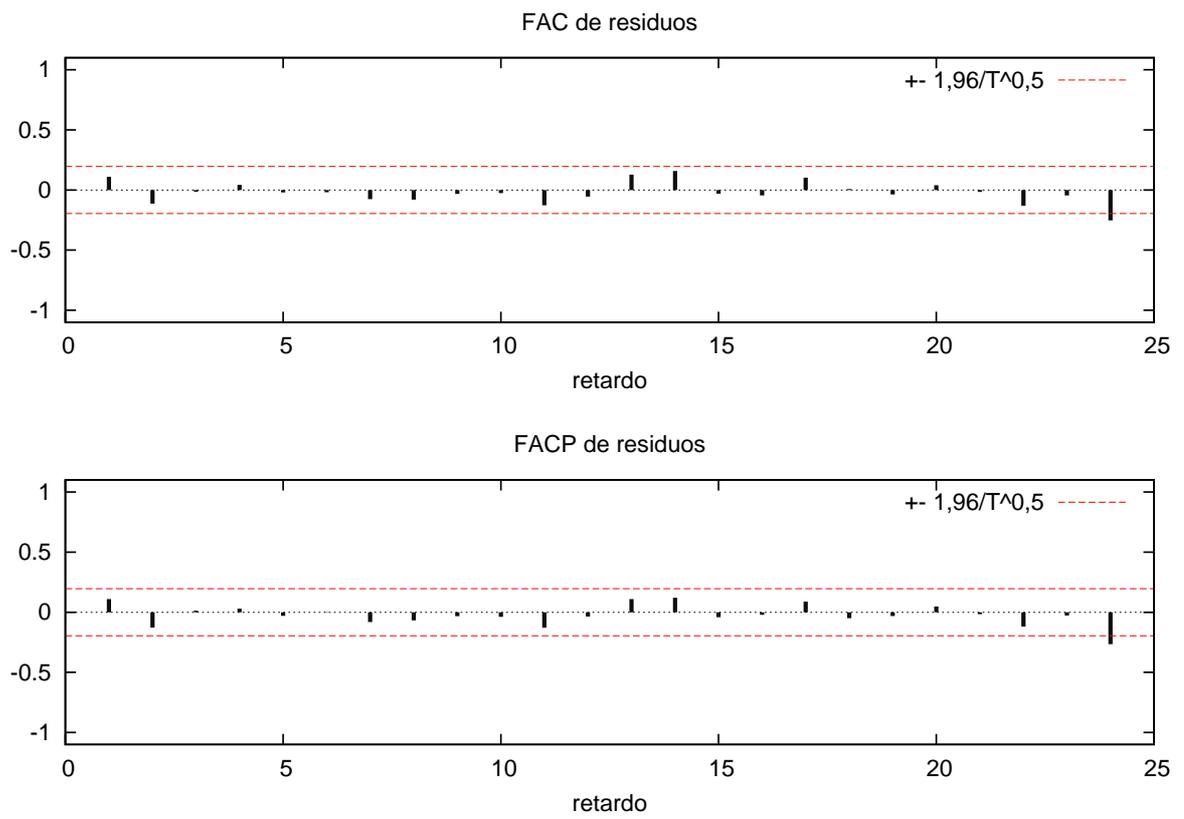


Figura 6.6: Autocorrelograma simple y parcial de la ecuación de gas oil

6.1.4. Interpretación del modelo final

La ecuación de largo plazo queda determinada por:

$$\hat{gasoil}_t = -1,99precio_t - 0,42Lluvias_t$$

La formulación del modelo es la siguiente:

$$\begin{aligned} \Delta \hat{gasoil}_t &= -0,099 (gasoil_{t-1} - 1,99precio_{t-1} - 0,42lluvias_{t-1}) + \dots \\ &\dots + \sum a_i \Delta precio_{t-i} + \sum a_i \Delta lluvias_{t-i} + \dots \\ &\dots - 12,80I_t^{3/2002} + 14,55I_t^{10/2003} + \Phi D_t \end{aligned}$$

Donde

- Los términos a_i (i) son los coeficientes de las variables diferenciadas obtenidos en la tabla: 6.1
- Las variables ingresadas de la forma: $I_t^{10/2003}$ son variables indicatrices; toman valor 1 en el momento del tiempo especificado (10/2003) y cero para el resto. Estas variables son multiplicadas por un coeficiente, el cual determina el peso con el cual ingresan en la ecuación.
- El término ΦD_t , es un término matricial compuesto por las variables indicatrices estacionales que corresponden a los meses del año e ingresan en la ecuación con su respectivo Φ_i .

Interpretación de los coeficientes

La variable *precio* presenta signo negativo de su correspondiente coeficiente *beta* (-1.99). Lo cual significa que al aumentar el precio disminuyen los volúmenes de gasoil vendidos.

Según información de ANCAP la demanda de gas oil en el largo plazo es inelástica al precio, pero en el *corto plazo* se torna elástica⁴. Por lo general los aumentos de precios se anuncian con anticipación (en el Directorio vigente). Por lo tanto, los grandes consumidores (fábricas, frigoríficos, etc.) que tengan la posibilidad de acopio del producto

⁴Elasticidad Precio de la Demanda de Gas Oil (2008) - Autores: Ec. Sabrina Rivero y Magela Negro

lo harán, incrementando los volúmenes de venta no en el momento del aumento del precio sino antes. Entonces, para el momento específico en el que aumenta el precio, aquellos grandes consumidores ya habrán adquirido el producto; lo que se traduce en menor venta para ese mes.

Con la variable Lluvias sucede lo mismo. El coeficiente *beta* tiene signo negativo (-0.415), de esta forma al aumentar las lluvias disminuyen los consumos de gas oil, situación comentada en el análisis descriptivo de la serie.

6.1.5. Performance del modelo en predicción

Para el análisis de las predicciones se evalúan 2 ventanas de predicción dentro de la muestra. La primera: predicción del año 2007 y la segunda: predicción de la ventana de tiempo: Julio de 2007 a Junio de 2008. Por último se realiza la predicción fuera de la muestra, en este caso: para el año 2008.

El mecanismo utilizado para la realización de las predicciones es cortar la muestra, realizar las predicciones en el período considerado y luego comparar el valor real con el valor predicho, de forma de corroborar como se desempeña el modelo en predicción.

Para el año 2008, se hizo un análisis pormenorizado de las predicciones. De esta forma, se evalúa la reacción del modelo sobre sus predicciones al incorporar datos a la muestra, por ello se adjuntan tablas con predicciones de enero a diciembre de 2008, de febrero a diciembre de 2008, de marzo a diciembre de 2008 y así sucesivamente.

Se comparan los errores de predicción en todos los casos evaluados esperando que disminuya a medida que se incorporan observaciones a la muestra.

Resultados obtenidos:

1. En cuanto a la predicción para el año 2007 quedan fuera del intervalo de predicción al 95 %: Mayo, Junio y Setiembre.

Por fuera de ambos intervalos (95 % y 99 %) queda la observación de Noviembre.

2. Para la predicción del período Julio de 2007 a Junio de 2008: el modelo deja las observaciones de Julio de 2007 y Abril de 2008 fuera de los intervalos de predicción

al 95 % y 99 %. La observación de Setiembre de 2007 queda fuera del intervalo de predicción al 95 %. La observación de Julio no resulta significativa como atípico en el análisis.

Véase que al ingresarse información a la muestra el modelo corrige la predicción de Noviembre de 2007 mencionada como fuera de los intervalos en el párrafo anterior. Ver: Anexo: Multivariado del grupo Gas Oil: sec: 12.3.2

- De las predicciones para el año 2008 solamente la observación del mes de Abril que queda fuera de los intervalos de predicción. Dicho mes tuvo un consumo muy elevado para la fecha: 81.64 miles de metros cúbicos. El resto de las predicciones incurren en un error no mayor al 5 %. En la tabla 6.4, se ve que al ingresar las observaciones hasta Marzo de 2008 el modelo corrige la predicción de Abril, la cual pasa de tener un error del 11 % al 9 %.

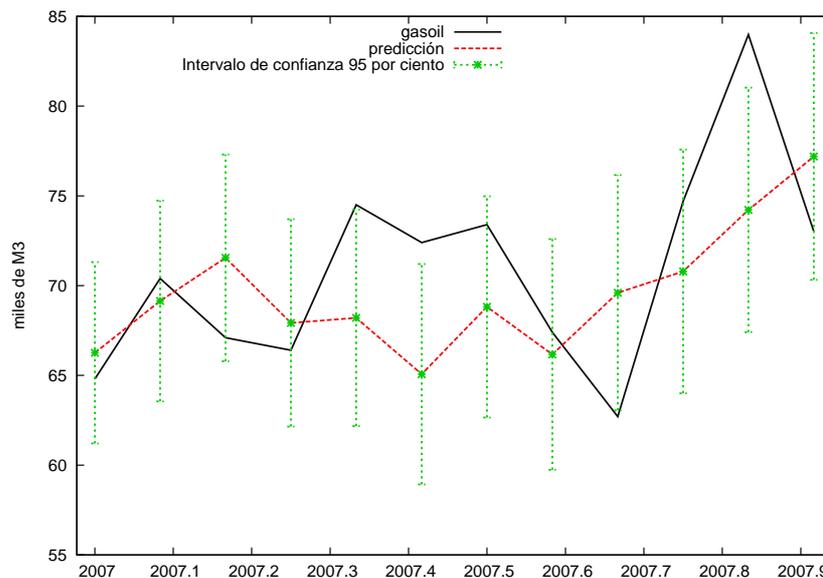


Figura 6.7: Predicción año 2007 - modelo seleccionado grupo Gas oil

			Intervalo de Predicción 95 %		Intervalo de Predicción 99 %				
Obs	gasoil	pred	Inf	Sup	Inf	Sup	ERROR	Pred. Univariada	Error univariado
Ene-07	64,8	66,26	61,21	71,31	59,62	72,90	2 %	67,56	4 %
Feb-07	70,4	69,14	63,55	74,74	61,80	76,49	-2 %	66,43	-6 %
Mar-07	67,1	71,55	65,79	77,30	63,98	79,11	7 %	76,93	15 %
Abr-07	66,4	67,92	62,14	73,69	60,33	75,51	2 %	71,45	8 %
May-07	74,5	68,20	62,17	74,23	60,28	76,13	-8 %	70,77	-5 %
Jun-07	72,4	65,07	58,93	71,20	57,00	73,13	-10 %	66,67	-8 %
Jul-07	73,4	68,82	62,65	74,98	60,71	76,92	-6 %	67,86	-8 %
Ago-07	67,4	66,16	59,73	72,59	57,71	74,61	-2 %	65,68	-3 %
Sep-07	62,7	69,60	63,05	76,16	60,99	78,22	11 %	69,56	11 %
Oct-07	74,7	70,79	64,00	77,58	61,87	79,71	-5 %	73,05	-2 %
Nov-07	83,98	74,22	67,40	81,03	65,26	83,17	-12 %	72,23	-13 %
Dic-07	73,02	77,19	70,32	84,07	68,16	86,23	6 %	72,92	0 %
ECM-multivariado = 26,46									
ECM-univariado = 32,61									

Tabla 6.2: Predicción año 2007 - modelo seleccionado grupo Gas oil

Obs	GasOil	pred	Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		ERROR	Prediccion Univariada	Error Univariado
			Inf	Sup	Inf	Sup			
Ene-08	68,99	70,59	65,03	76,15	63,29	77,90	2 %	70,58	2 %
Feb-08	67,37	70,88	64,86	76,90	62,96	78,79	5 %	76,78	14 %
Mar-08	72,51	71,36	65,18	77,54	63,24	79,48	-2 %	72,79	0 %
Abr-08	81,64	72,65	66,45	78,84	64,50	80,79	-11 %	72,29	-11 %
May-08	73,20	70,24	63,72	76,76	61,67	78,81	-4 %	77,54	6 %
Jun-08	66,52	66,77	60,09	73,46	57,98	75,56	0 %	76,02	14 %
Jul-08	68,79	71,58	64,88	78,28	62,77	80,38	4 %	76,41	11 %
Ago-08	65,01	67,59	60,64	74,54	58,46	76,72	4 %	73,40	13 %
Sep-08		72,35	65,16	79,53	62,90	81,79		70,83	
Oct-08		74,24	66,81	81,68	64,47	84,01		77,22	
Nov-08		76,94	69,47	84,40	67,12	86,75		82,25	
Dic-08		77,48	69,93	85,03	67,56	87,40		76,36	
ECM-multivariado = 15,04									
ECM-univariado = 52,01									

Tabla 6.3: Predicción del año 2008 - modelo seleccionado grupo gas oil

Obs	GasOil	pred	Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		ERROR	Prediccion Univariada	Error Univariado
			Inf	Sup	Inf	Sup			
Abr-08	81,64	74,22	68,72	79,72	67,00	81,44	-9 %	70,43	-14 %
May-08	73,20	69,21	63,26	75,17	61,38	77,04	-5 %	73,16	0 %
Jun-08	66,52	66,25	60,12	72,37	58,20	74,29	0 %	73,39	10 %
Jul-08	68,79	70,00	63,85	76,14	61,92	78,07	2 %	73,83	7 %
Ago-08	65,01	66,50	60,03	72,97	58,00	75,01	2 %	70,22	8 %
Sep-08		72,16	65,50	78,82	63,40	80,92		67,95	
Oct-08		72,76	66,08	79,43	63,99	81,52		74,36	
Nov-08		75,93	69,02	82,85	85,03	85,03		79,19	
Dic-08		76,81	69,66	83,96	86,21	86,21		73,40	
ECM-multivariado = 14,94									
ECM-univariado = 45,04									

Tabla 6.4: Predicción para el año 2008 incorporando hasta la observación de Marzo - modelo seleccionado grupo gas oil

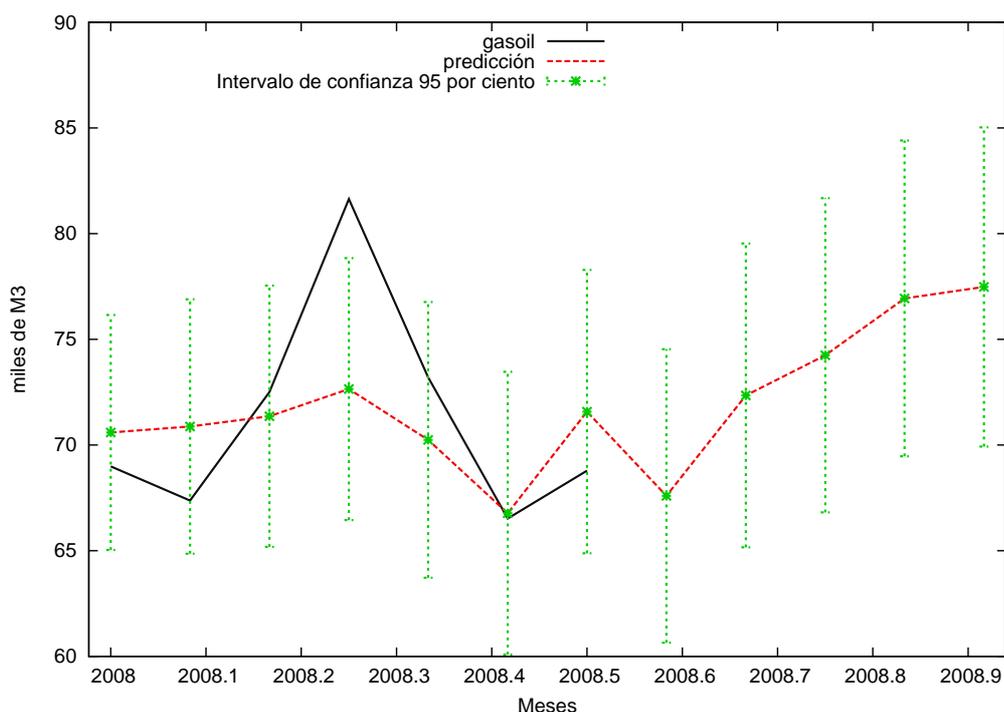


Figura 6.8: Predicción para el año 2008 - modelo seleccionado grupo gas oil

6.2. Resultados referidos a la demanda de Supergas

6.2.1. Datos

El grupo de variables relacionadas con la *demanda de Supergas en Uruguay* son:

- Consumo de Supergas en el mercado interno
- Precio por Kg de supergas
- Temperatura máxima - promedio mensual de la estación INIA: *Las Brujas* (Canelones).
- Temperatura mínima - promedio mensual de la estación INIA: *Las Brujas* (Canelones).

Las series conforman una muestra de periodicidad mensual en el período Enero de 1999 a Diciembre de 2007 (102 observaciones).

Los datos vinculados al clima son proporcionados por el INIA. El consumo de supergas y el precio de dicho producto son proporcionados por ANCAP. El precio es deflactado por el IPC y se encuentra expresado a precios de Junio de 2007.

6.2.2. Análisis descriptivo de las series

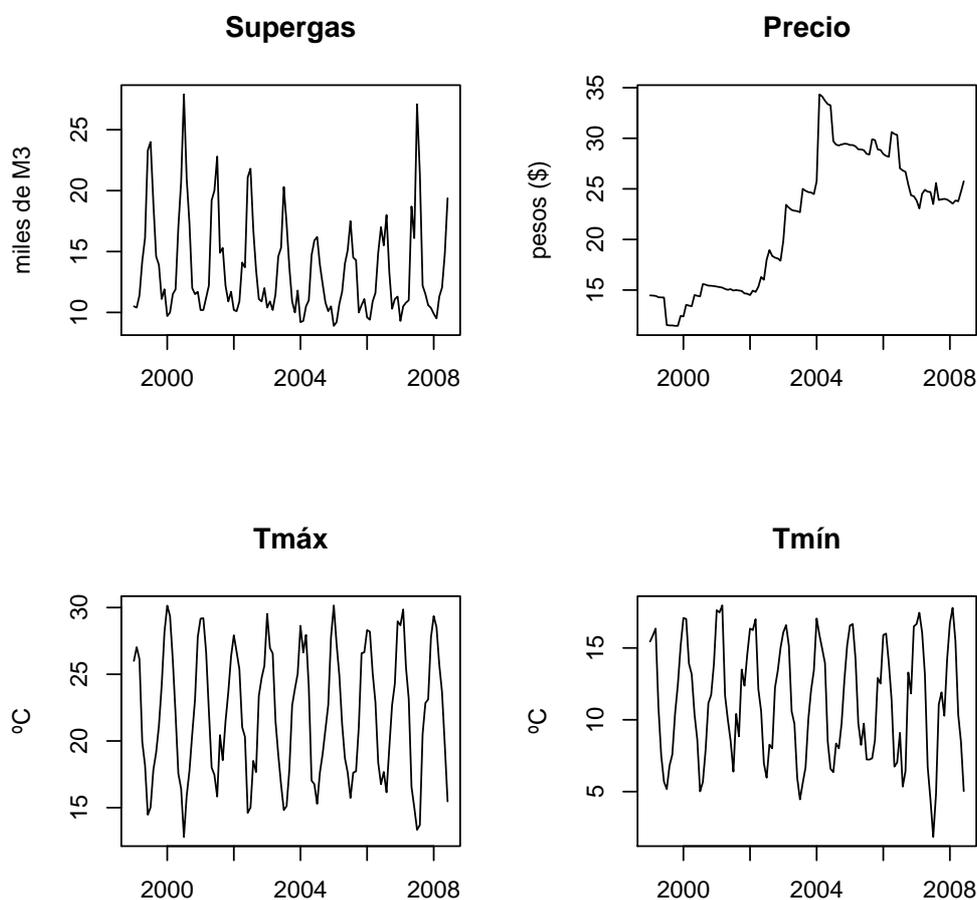


Figura 6.9: Gráficas de las variables del grupo: Supergas

Es muy clara la relación existente entre las variables climatológicas y el consumo de supergás. Básicamente son gráficas espejo. El incremento o baja en la temperatura incide de forma inversa en el consumo de supergas; por lo tanto se puede esperar que ante un incremento en la temperatura se genere una baja del consumo y ante una baja en la temperatura un incremento en el consumo. A su vez se observa que la temperatura máxima y mínima pueden albergar cierta correlación, debido que a simple vista son gráficas iguales pero con cambio de escala. Se puede esperar que teniendo conocimiento sobre una de ellas se infiera sobre la otra.⁵

⁵Este aspecto se analiza en el Anexo: Modelos Rechazados (sec: 12.3), parte: 11.4.1

A su vez en el año 2004, políticas adoptadas por parte de la administración, incrementaron el precio de forma repentina y en un corto plazo, llevando a una disminución en el consumo de supergas a partir de dicho año. Hasta ese momento la tendencia creciente en el precio generó bajas en el consumo de supergas, sobretodo en los picos máximos. Ya para el año 2005, el precio tiende a la baja y el consumo a una lenta recuperación.

Para ver más de cerca lo antes comentados se realizan dos gráficos. El primero refiere al precio contra el supergas y el segundo a la temperatura máxima contra el supergas.

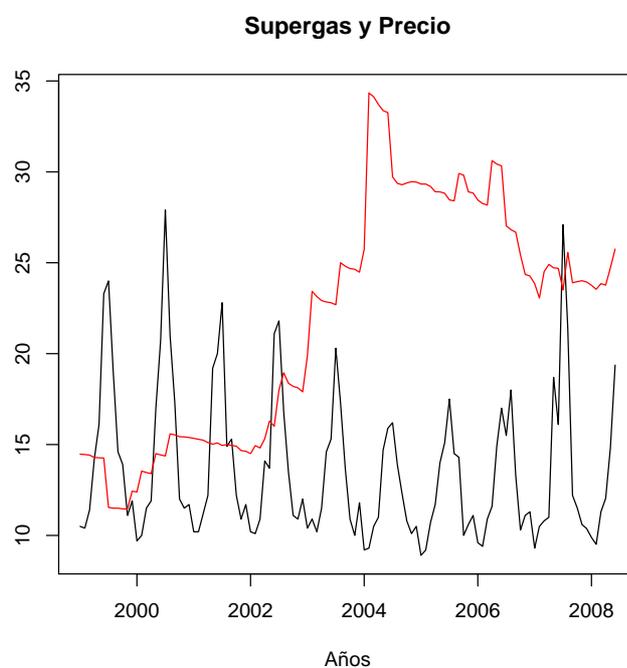


Figura 6.10: Variables: Supergas y Precio (en color negro: *consumo de supergas*, en rojo: *Precio*)

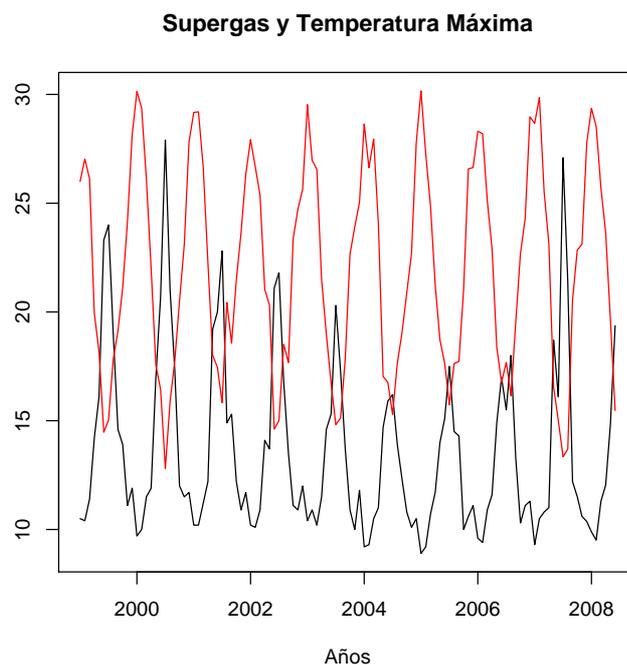


Figura 6.11: Variables: Supergas y Temperatura Máxima (en color negro: *consumo de supergas*, en rojo: *Temperatura Máxima*)

6.2.3. Modelo Seleccionado

Antes de concluir en el modelo que aquí se presenta se probaron varias modelizaciones utilizando distintas especificaciones. Por ejemplo, se consideran modelos con constante en la relación de cointegración, modelos sin constante, con y sin variables indicatrices estacionales y con varias alternativas en la cantidad de retardos a incluir. No se muestran aquí modelos que incluyan un término de tendencia ya que no resultan significativos en el análisis. Algunos de ellos no fueron estables a lo largo del tiempo, por lo que fueron desechados en el proceso de selección. Otros pasaron el diagnóstico preliminar, pero fueron desechados por su desempeño predictivo. Dichos modelos se encuentran en el anexo.

Se elige para la modelización multivariada del consumo de Supergas el modelo VECM que incluye:

- 5 retardos
- especificado sin constante y sin variables indicatrices estacionales
- las variables: supergas, precio y temperatura máxima expresadas en logaritmos.
- valores atípicos en: Julio de 2000, Agosto de 2001, Junio de 2002 y Julio de 2007

Análisis de Cointegración

Para el análisis de cointegración de las variables se realiza el test de Dickey Fuller a cada una de ellas.

Los resultados obtenidos son: el logaritmo del precio y el logaritmo del supergas son variables integradas de orden 1; el logaritmo de las temperaturas máxima y mínima son integradas de orden 0.

Test de Johansen para el modelo seleccionado

Contraste de Johansen:						
Número de ecuaciones = 3						
Orden del retardo = 6						
Periodo de estimación: 1999:07 - 2007:12 (T = 102)						
Caso 1: Sin constante						
Rango	Valor propio	Contraste de la traza		valor p	Contraste L _{máx}	valor p
0	0,54300	83,161	[0,0000]		79,873	[0,0000]
1	0,03043	3,2884	[0,7992]		3,1519	[0,7596]
2	0,001338	0,1365	[0,7808]		0,1365	[0,7715]

El *Test de cointegración de Johansen* lleva a no rechazar la hipótesis nula de existencia de una única relación de cointegración (tanto con el estadístico del valor propio como el de la traza, con p-valor de 0.80 y 0.76 respectivamente). Por lo tanto se aplican modelos VECM⁶. El test fue especificado de igual forma a como se especificó el modelo final.

⁶Observación: se realiza el test de Johansen a las variables logaritmo de supergas y logaritmo de precio obteniendo como resultado que no están cointegradas. Se acepta la hipótesis de rango igual a cero

Estimación y validación del modelo seleccionado.**Coefficientes del modelo**

Beta: vector cointegrante					
Log(Supergas)	Log(Precio)	Log(Tmax)			
1	0,2	-1,04			
0	-0,037	-0,04			
Alfa: velocidades de ajuste					
Log(Supergas)	Log(Precio)	Log(Tmax)			
-0,56	0,015	0,58			
Variable	Coefficiente	Desv. típica	Estadístico t	valor p	
d(log(sup))-1	-0,31	0,14	-2,26	0,03	**
d(log(sup))-2	-0,4122	0,13	-3,3	0	***
d(log(sup))-3	-0,32	0,12	-2,6	0,01	**
d(log(sup))-4	-0,36	0,12	-3,05	0	***
d(log(sup))-5	-0,12	0,11	-1,09	0,28	
d(log(precio))-1	-0,28	0,21	-1,36	0,18	
d(log(precio))-2	-0,08	0,21	-0,41	0,68	
d(log(precio))-3	0,01	0,2	0,05	0,96	
d(log(precio))-4	0,18	0,2	0,87	0,39	
d(log(precio))-5	-0,15	0,21	-0,73	0,47	
d(log(Tmax))-1	-1,01	0,15	-6,56	<0,00001	***
d(log(Tmax))-2	-1,16	0,16	-7,11	<0,00001	***
d(log(Tmax))-3	-1,01	0,2	-5,16	<0,00001	***
d(log(Tmax))-4	-0,64	0,21	-2,99	0	***
d(log(Tmax))-5	-0,34	0,19	-1,76	0,08	*
Jul-2000	0,39	0,11	3,48	0	***
ag-2001	-0,22	0,11	-2,03	0,05	**
jun-2002	0,27	0,11	2,5	0,01	**
jul-2007	0,38	0,11	3,38	0	***
EC1	-0,56	0,09	-6,41	<0,00001	***

Tabla 6.5: Coeficientes del modelo seleccionado para el consumo de supergas**Significación de las variables**

De las variables incluidas en el modelo, el logaritmo del precio no resulta significativo al 5 % en ninguno de los retardos considerados. Sin embargo, existe evidencia estadística que lleva a la inclusión de dicha variable en el modelo final. Las pruebas realizadas sobre la significación individual de los coeficientes del vector de cointegración hacen rechazar la hipótesis nula de que sean iguales a cero.

a un nivel de significación del 5 %. El test es especificado del igual forma que el modelo seleccionado. Observación: El modelo se especifica con $k+1$ retardos de forma que resulten k en la salida del programa.

Se anexa modelo que no incluye dicha variable. Ver: Anexo Multivariado Supergas: sec: 11.4.3. Si bien el nuevo modelo (modelo 3 en el anexo) tienen menor error cuadrático medio (ECM) que el modelo elegido, el error porcentual incurrido en cada observación es mayor (considerando un período de 3 meses a partir de Enero de 2008).

<pre> Restricción: log(Precio) b[2] = 0 valor propio 1 = 0,40906 Log-verosimilitud no restringida (lu) = 416,91037 Log-verosimilitud restringida (lr) = 403,80209 2 * (lu - lr) = 26,2166 P(Chi-cuadrado(1) > 26,2166) = 3,05192e-007 </pre>
<pre> Restricción: log(Temperatura Máxima) b[3] = 0 valor propio 1 = 0,0747365 Log-verosimilitud no restringida (lu) = 416,91037 Log-verosimilitud restringida (lr) = 380,93548 2 * (lu - lr) = 71,9498 P(Chi-cuadrado(1) > 71,9498) = 2,20745e-017 </pre>

Significación del coeficiente del termino de corrección de error

En la tabla anterior puede observarse que el término de corrección del error (EC1) es significativo en el análisis. Se concluye que el modelo corrige los desequilibrios a la baja en el corto plazo. Por ejemplo, si existe un desequilibrio el mes pasado, el modelo corrige un 56 % a la baja ese desequilibrio.

El término de corrección de error queda definido por la ecuación:

$$-0,56 (Log(supergas)_{t-1} + 0,2Log(precio)_{t-1} - 1,04Log(Tmax)_{t-1})$$

Exogeneidad débil de las variables - Velocidad de ajuste

Se restringen los coeficientes alfas del modelo a cero para evaluar la existencia o no de endogeneidad (retroalimentación). La prueba se realiza mediante el test de razón de verosimilitudes, el cual distribuye $\chi^2_{(r(p-m))}$, donde r es el rango de integración; p filas de la matriz que contiene las restricciones en el alfa (en este caso 3) y m son las columnas de la matriz que contiene las restricciones de los alphas (2).

Los resultados obtenidos son:

Restricción: $\log(\text{Precio})$ $a[2] = 0$ valor propio 1 = 0,55702 Log-verosimilitud no restringida (lu) = 421,75452 Log-verosimilitud restringida (lr) = 421,69547 $2 * (lu - lr) = 0,118093$ $P(\text{Chi-cuadrado}(1) > 0,118093) = 0,731112$
Restricción: $\log(\text{Temperatura Máxima})$ $a[3] = 0$ valor propio 1 = 0,0443451 Log-verosimilitud no restringida (lu) = 421,75452 Log-verosimilitud restringida (lr) = 382,09858 $2 * (lu - lr) = 79,3119$ $P(\text{Chi-cuadrado}(1) > 79,3119) = 5,30397e-019$

Se obtiene como resultado que el precio es una variable débilmente exógena, ya que no se rechaza la hipótesis nula de que su coeficiente α^7 sea igual a cero. Por lo tanto: no responde a las desviaciones del equilibrio en el período anterior. Implica que la variable precio no se corrige en el corto plazo con las variables restantes integrantes del sistema.

En el caso de la temperatura máxima el resultado es que se rechaza la exogeneidad débil de la variable. Puede deberse a que el comportamiento de las series de temperatura máxima y consumo de supergas hace que se puedan vincular entre ellas al punto de considerar la retroalimentación (endogeneidad) cuando la situación no es lógica!. Recuérdese que en el análisis descriptivo de las series se menciona que estas variables gráficamente se comportan como funciones espejo.

Es de esperar que el resultado en esta instancia no sea el que se obtuvo. Es decir, la *situación ideal* sería que la prueba para los alfas testeados lleve a no rechazar la hipótesis nula de exogeneidad débil. Lo que implica que las variables no son endógenas.

Análisis de los residuos

La gráfica de los residuos del precio presenta un valor muy alejado de la media: *Febrero de 2004*. Este valor no resulta significativo en el análisis, lo que implica que el aumento repentino del precio en dicho período no se traslada de forma significativa al

⁷velocidad de ajuste al equilibrio

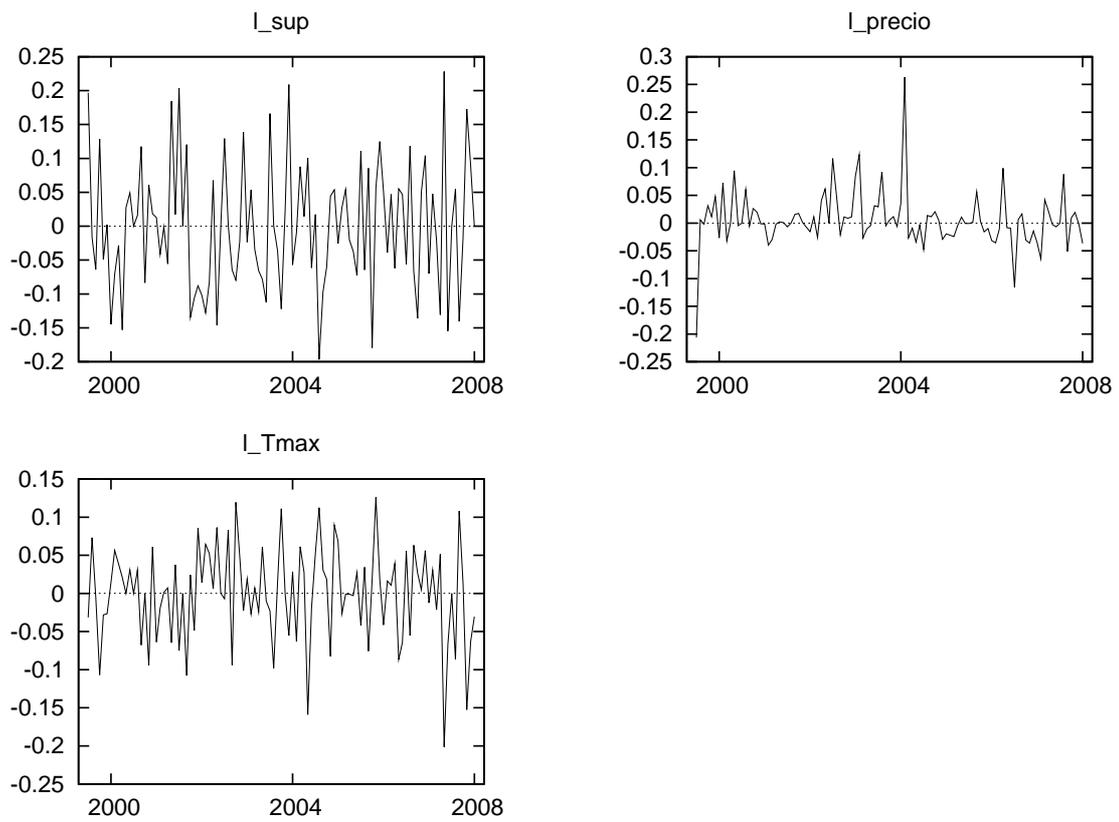


Figura 6.12: Residuos de las tres ecuaciones del sistema VECM

consumo del producto. Ver figura: 6.12 y Anexo: Multivariado Supergas, sec: 11.2.

Normalidad de los residuos

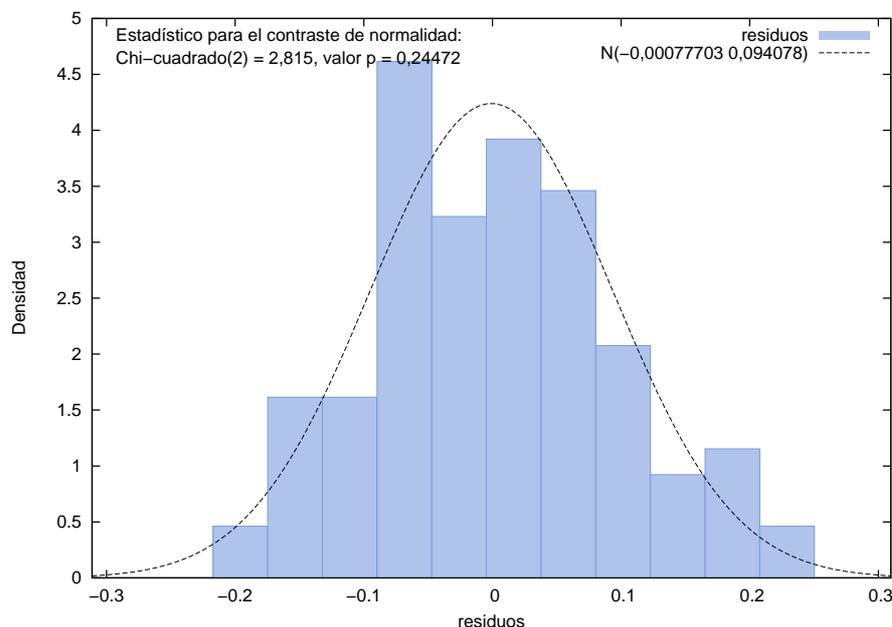


Figura 6.13: Normalidad de los residuos de la ecuación de Supergas

El test de Normalidad de *Doornik-Hansen* realizado a los residuos de la ecuación de Supergas, lleva a no rechazar la hipótesis nula de Normalidad de los mismos, con un p-valor igual a 0.26. Ver figura: 6.5. Sin embargo la normalidad multivariada no se cumple:

Test for multivariate normality of residuals
 Doornik-Hansen Chi-square(6) = 90,256, with p-value = 2,68076e-017

Autocorrelación de los residuos

A continuación se presenta el test de Ljung Box y el gráfico de los autocorrelogramas simple y parcial. Se anexa autocorrelograma simple y parcial con sus respectivos p-valores para cada retardo. Ver: Anexo: 11, sec: 12.1

Los resultados obtenidos son:

El test de Ljung-Box hace no rechazar la hipótesis nula de no autocorrelación en los primeros 48 retardos seleccionados, con un p-valor de 0.577. Se observa que el auto-

correlograma simple y parcial dejan el retardo 12 por fuera de las bandas de confianza. Esto puede ser signo de que la estacionalidad de la serie no ha sido captada. Por lo tanto, se ingresan variables indicatrices estacionales para corregir ese efecto. Los resultados obtenidos del modelo resultante no varían del aquí presentado y tampoco se mejoran las predicciones. Ver Anexo: Modelos rechazados (sec: 12.3): Modelo 6: parte: 11.4.6.

Ecuación 1:
Ljung-Box $Q' = 45,48$ con valor $p = P(\text{Chi-cuadrado}(48) > 45,48) = 0,577$

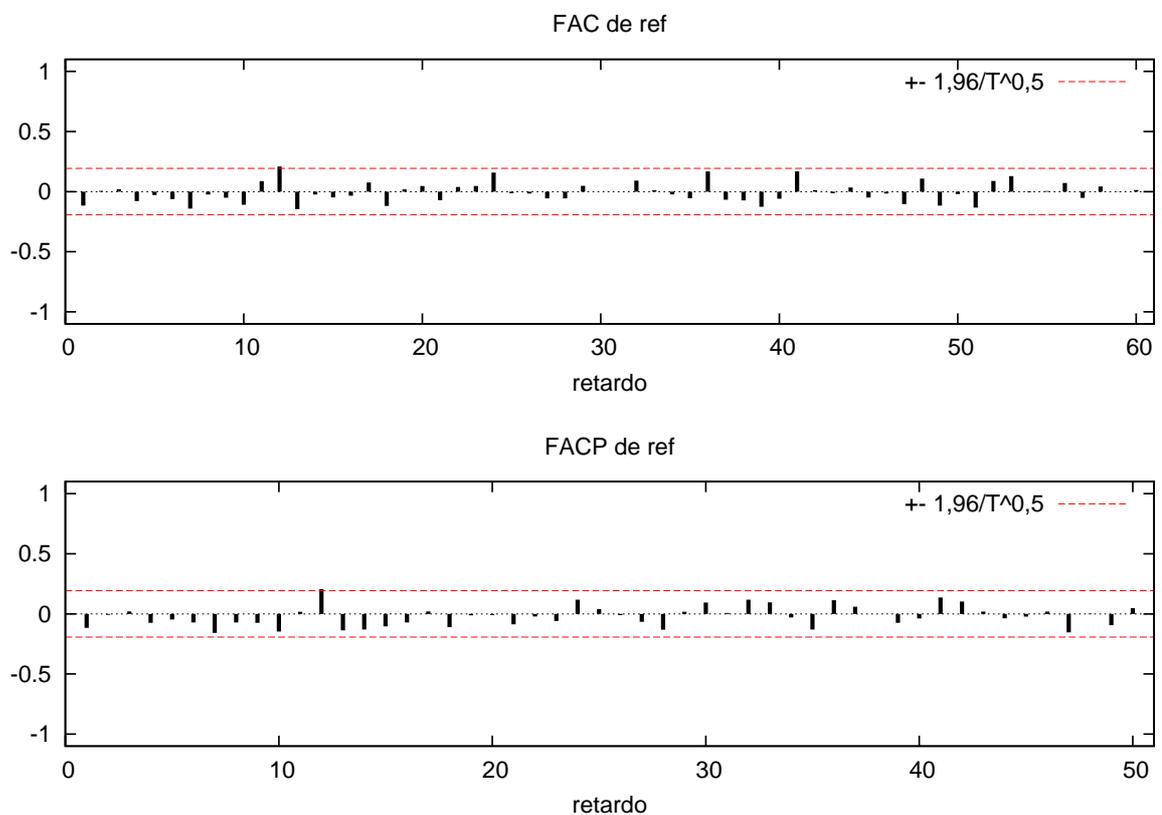


Figura 6.14: Autocorrelograma simple y parcial de la ecuación de supergas

6.2.4. Interpretación del modelo final

La ecuación de largo plazo queda determinada por:

$$\hat{Log}(supergas)_t = 0,20Log(precio)_t - 1,04Log(Tmax)_t$$

Modelo Final:

$$\begin{aligned} \Delta \hat{Log}(supergas)_t = & \underbrace{-0,56 (Log(supergas)_{t-1} + 0,2Log(precio)_{t-1} - 1,04Log(Tmax)_{t-1})}_{\text{termino de correccion de error}} \\ & + \sum a_i(i) \Delta Log(precio)_{t-i} \\ & + \sum a_i(i) \Delta Log(Tmax)_{t-i} \\ & 0,38I_t^{7/2000} - 0,22I_t^{8/2001} + 0,27I_t^{6/2002} + 0,38I_t^{7/2007} \end{aligned}$$

Donde los términos $a_i(i)$ son los coeficientes de las variables diferenciadas obtenidos en la tabla: 6.5.

Las variables ingresadas de la forma: $I_t^{7/2000}$ son variables indicatrices; toman valor 1 en el momento del tiempo especificado (7/2000) y cero para el resto. Estas variables tienen un peso con el cual ingresan en la ecuación, determinado por el real que las precede.

Interpretación de los coeficientes

La variable *logaritmo del precio* presenta signo positivo de su correspondiente coeficiente *beta* (0.2). Significa que al aumentar el precio aumentan los volúmenes de supergas vendidos.

Al igual que lo comentado en el análisis descriptivo de la serie, los aumentos en el precio influyen en el consumo de dicho producto. La particularidad de la figura: 6.10 es que el precio se mantiene a un nivel alto incluso cuando comienza a bajar en el año 2005, momento para el cual el consumo comienza a recuperarse. Pero principalmente el consumo es elevado porque la temperatura disminuye. Para dicha variable el coeficiente *beta* tiene signo negativo (-1.04). Implica que el aumento de la temperatura tiene un impacto negativo en el consumo, retrayendo la demanda. Si la temperatura aumenta, el supergas (en su mayoría consumido para calefacción del hogar) no sería tan

necesario, por lo tanto el consumo decae. También puede asociarse al comportamiento del consumo en los meses de verano.

6.2.5. Performance del modelo en predicción

La performance del modelo en predicción se evalúa de igual forma a como se realiza la evaluación del gas oil. Es decir, se evalúan 2 ventanas de predicción dentro de la muestra: año 2007 y ventana de tiempo: Julio de 2007 a Junio de 2008. Por último se realiza la predicción fuera de la muestra, en este caso: para el año 2008.

El mecanismo utilizado para la realización de las predicciones es cortar la muestra, realizar predicciones en el período considerado y luego comparar el valor real de consumo con el valor predicho, de forma de corroborar como se desempeña el modelo en predicción.

En el caso del año 2008, año corriente, se hizo un análisis pormenorizado de las predicciones. De esta forma, se evalúa la reacción del modelo sobre sus predicciones al incorporar datos a la muestra, por ello se adjuntan tablas con predicciones de enero a diciembre de 2008, de febrero a diciembre de 2008, de marzo a diciembre de 2008 y así sucesivamente.

Se comparan los errores de predicción en todos los casos evaluados esperando que disminuya a medida que se incorporan observaciones a la muestra.

Resultados obtenidos:

1. En el caso de la predicción realizada para el año 2008, se observa que el máximo error que incurre el modelo es del 11 % en las observaciones de Marzo, Junio y Julio. Sin embargo, estas observaciones no quedan fuera de los intervalos de predicción al 95 % ni al 99 %.

El modelo tiene gran capacidad de absorber la información que va sucediéndose en el tiempo. En la tabla 11.3, se ve que al ingresar la observación de Enero de 2008 el modelo corrige la predicción de Marzo de 2008, la cual pasa de tener un error del 11 % a 7 %.

2. Para la predicción de la ventana de tiempo Julio de 2007 a Junio de 2008: el modelo deja sólo la observación de Julio de 2007 fuera de los intervalos de predicción al 95 % y 99 %. Se recuerda que esta observación es un valor detectado como atípico, por lo que es razonable que no quede comprendida en dichos intervalos.
3. En cuanto a la predicción para el año 2007, sucede lo esperado. La observación Mayo y Agosto quedan fuera del intervalo de predicción al 95 % y Julio fuera de ambos intervalos 95 % y 99 %. La predicción de Agosto se ajusta a medida que va ingresando información a la muestra.

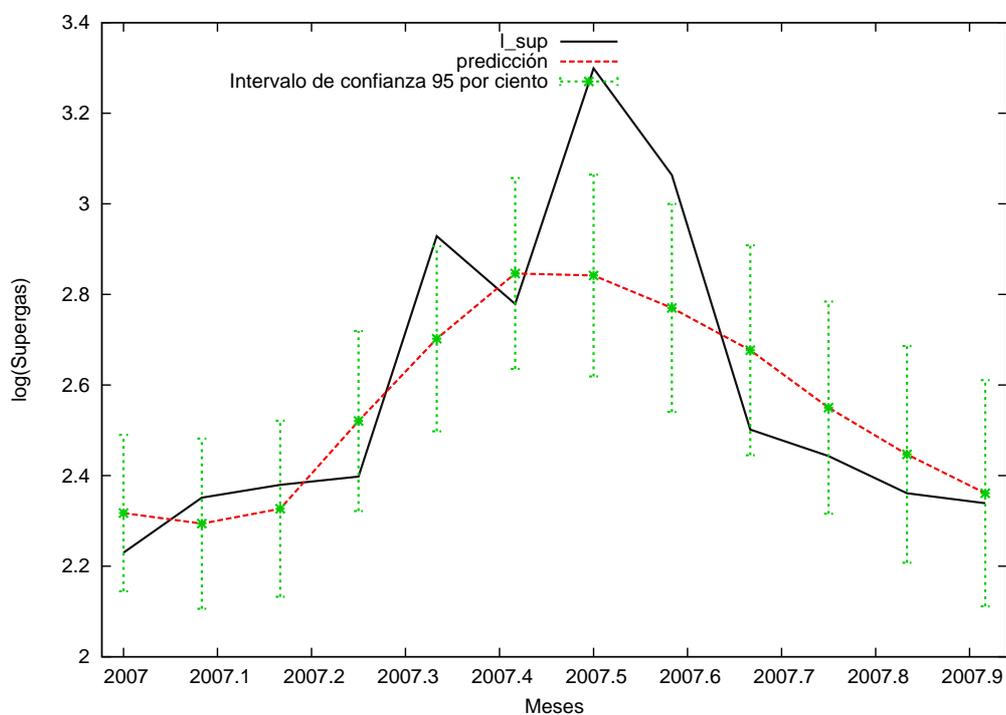


Figura 6.15: Predicción para el año 2007 - modelo seleccionado grupo supergas

			Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		ERROR	Prediccion Univariada	Error Univariado	
	Sup	pred	Inf	Sup	Inf	Sup				
Ene-07	9,30	10,15	8,54	12,06	8,09	12,73	9 %	9,58	3 %	
Feb-07	10,50	9,91	8,22	11,96	7,75	12,69	-6 %	9,43	-10 %	
Mar-07	10,80	10,25	8,44	12,44	7,94	13,23	-5 %	10,88	0 %	
Abr-07	11,00	12,43	10,19	15,17	9,58	16,15	13 %	11,50	5 %	
May-07	18,70	14,91	12,16	18,29	11,40	19,50	-20 %	14,75	-21 %	
Jun-07	16,10	17,22	13,95	21,26	13,06	22,72	7 %	16,31	1 %	
Jul-07	27,09	17,15	13,72	21,42	12,79	22,98	-37 %	16,20	-40 %	
Ago-07	21,40	15,96	12,69	20,08	11,80	21,58	-25 %	15,82	-26 %	
Sep-07	12,20	14,54	11,53	18,33	10,72	19,71	19 %	13,14	8 %	
Oct-07	11,51	12,81	10,13	16,19	9,41	17,42	11 %	10,43	-9 %	
Nov-07	10,60	11,55	9,09	14,68	8,43	15,82	9 %	10,66	1 %	
Dic-07	10,37	10,60	8,26	13,61	7,64	14,72	2 %	11,00	6 %	
ECM-multivariado = 12,96										
ECM-univariado = 14,09										

Tabla 6.6: predicción para el año 2007 - modelo seleccionado grupo supergas

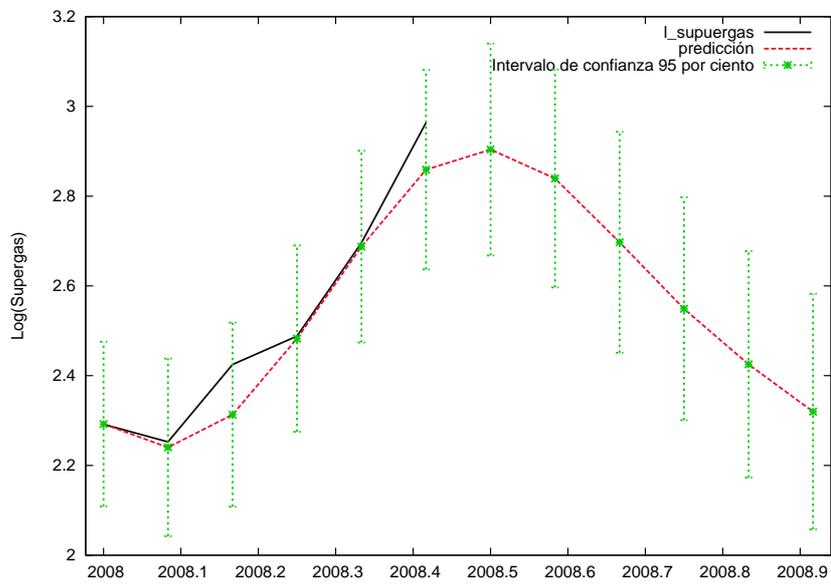


Figura 6.16: Predicción para el año 2008 - grupo supergas

			Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %				
Obs	Sup	pred	Inf	Sup	Inf	Sup	ERROR	Prediccion Univariada	Error Univariado
Ene-08	9,89	9,90	8,24	11,89	7,78	12,60	0 %	9,02	-9 %
Feb-08	9,51	9,40	7,71	11,45	7,24	12,19	-1 %	9,59	1 %
Mar-08	11,30	10,11	8,23	12,41	7,72	13,23	-11 %	10,67	-6 %
Abr-08	12,04	11,97	9,73	14,73	9,11	15,73	-1 %	11,20	-7 %
May-08	14,82	14,70	11,87	18,20	11,10	19,46	-1 %	14,25	-4 %
Jun-08	19,37	17,44	13,97	21,79	13,02	23,37	-10 %	15,72	-19 %
Jul-08	16,46	18,25	14,41	23,11	13,38	24,89	11 %	16,31	-1 %
Ago-08	16,02	17,11	13,42	21,81	12,43	23,53	7 %	15,54	-3 %
Sep-08		14,84	11,60	18,98	10,74	20,50		13,12	
Oct-08		12,80	9,98	16,40	9,23	17,74		10,69	
Nov-08		11,30	8,78	14,55	8,11	15,75		10,66	
Dic-08		10,18	7,83	13,23	7,21	14,37		10,77	
ECM-multivariado = 1,19									
ECM-univariado = 1,97									

Tabla 6.7: Predicción del año 2008 - grupo supergas

Obs	Sup	pred	Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		ERROR
			Inf	Sup	Inf	Sup	
Feb-08	9,51	9,77	8,14	11,72	7,68	12,41	3 %
Mar-08	11,30	10,53	8,65	12,82	8,13	13,64	-7 %
Abr-08	12,04	12,23	9,97	15,01	9,35	16,00	2 %
May-08	14,82	14,59	11,86	17,95	11,11	19,15	-2 %
Jun-08	19,37	17,12	13,83	21,18	12,94	22,64	-12 %
Jul-08	16,46	17,54	14,06	21,89	13,11	23,47	7 %
Ago-08	16,02	16,57	13,09	20,97	12,15	22,58	3 %
Sep-08		14,61	11,47	18,62	10,63	20,10	
Oct-08		12,76	9,97	16,31	9,23	17,62	
Nov-08		11,44	8,92	14,66	8,25	15,85	
Dic-08		10,46	8,13	13,46	7,51	14,57	
ECM-multivariado = 1,04							

Tabla 6.8: Predicción del año 2008 incorporando la observación de Enero de 2008 - grupo supergas

Obs	Sup	pred	Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		ERROR
			Inf	Sup	Inf	Sup	
Jun-08	19,37	16,46	13,75	19,70	12,99	20,85	-15 %
Jul-08	16,46	16,54	13,62	20,08	12,82	21,34	0 %
Ago-08	16,02	16,10	13,17	19,70	12,36	20,98	1 %
Sep-08		14,76	12,04	18,10	11,29	19,30	
Oct-08		13,11	10,63	16,15	9,96	17,25	
Nov-08		11,85	9,53	14,73	8,90	15,77	
Dic-08		10,85	8,60	13,68	7,99	14,71	

Tabla 6.9: Predicción del año 2008 incorporando hasta la observación de Mayo de 2008 - grupo supergas

Capítulo 7

Conclusiones y recomendaciones

El **objetivo general** del presente trabajo consiste en la modelización de las series de tiempo de gas oil y supergas en Uruguay, para el período comprendido entre Enero de 1999 y Diciembre de 2007.

Para el gas oil se encuentra un modelo univariado: $ARIMA(2,1,0)(1,0,0)$ con valor atípico en Octubre de 2003. El modelo pasa todas las etapas de diagnóstico en la metodología empleada y es el que presenta menor error cuadrático medio en sus predicciones en comparación con otros modelos evaluados. En cuanto a las predicciones, para el año 2008, el modelo genera un error porcentual promedio del 8 % sobre el valor observado.

La incorporación de información extra vinculada al producto (variables explicativas) a la muestra supone una disminución del error de predicción. En efecto, el análisis multivariado del gas oil lleva a que en promedio se incurra en un error porcentual del 4 % sobre la observación. A su vez muestra una gran ventaja: el ajuste de sus predicciones al incorporar datos a la muestra disminuyendo el error (recuérdese que en el modelo univariado al incorporar datos a la muestra el error fluctúa sin tendencia clara a la disminución).

En el caso del supergas se encuentra un modelo univariado: $SARIMA(2,0,0)(2,1,0)$ con atípicos en: Agosto de 2001, Mayo de 2002, Junio de 2003, Mayo, Julio y Agosto de 2007. Al igual que en el gas oil este modelo se obtiene como final de un proceso de selección en el cual cada etapa es cumplida según los supuestos de la metodología. A su vez es el modelo con menor error cuadrático medio en sus predicciones. En lo que respecta a las predicciones del año 2008 el modelo incurre en un error porcentual promedio de

6.25 % sobre la observación. La incorporación de observaciones a la muestra no genera grandes cambios respecto al error porcentual.

En cuanto al modelo multivariado de supergas, se destaca que presenta un menor error cuadrático medio y el error porcentual promedio en cada observación es del 5.25 %.

Los modelos multivariados para ambos productos mostraron tener un menor error en las predicciones si se lo compara con el modelo univariado.

Pese a esto, debido a que el supergas esta fuertemente ligado al clima, la elaboración de modelos de series de tiempo cuya frecuencia sea mayor (semanal o diaria) sería lo apropiado. De esa forma el impacto del cambio climático quedaría captado de forma inmediata en comparación con un modelo de frecuencia mensual como el que aquí se presenta. Por lo tanto la combinación de un modelo semanal o diario conjunto con la perspectiva anual de los modelos multivariados proporcionaría un panorama más acertado del futuro a corto y largo plazo respectivamente.

En el caso del modelo del gas oil el manejo de los modelos multivariados ha demostrado tener un mejor ajuste de las predicciones al valor observado a medida que se incorporan datos a la muestra.

Referencia Bibliográfica

- [1] Aboal, Diego (2001), '*La elasticidad de la demanda por gasolina en Uruguay: Una aplicación de Modelos de Corrección de Error*', Documento de trabajo del MEF .
- [2] Amengual, Dante y Cubas, Germán (2002), '*Imposición optima a las naftas y el gas oil: un análisis empírico para Uruguay (1988-2001)*', Banco Central del Uruguay.
- [3] Enders, Walter (2004), '*Applied Econometric Time Series*', University of Alabama, segunda edición.
- [4] Equipos Mori (2002), '*Matriz de Consumo de Gas Oil, 2001*', Equipos Mori.
- [5] Francos Rodriguez, Martín (2006), '*Estimación de la Demanda de Combustibles en la República Dominicana*', Unidad de Análisis Económico - Secretariado Técnico de la Presidencia.
- [6] Hamilton, James D. (1994), '*Time Series Analysis*', Princeton University.
- [7] Novales Cinca, Alfonso (1993), '*Econometría (Segunda Edición)*', McGraw-Hill.
- [8] Paff, Bernhard (2006), '*Analysis of Integrated and Cointegrated Time Series with R*', Springer.
- [9] Rivero, Sabrina; Negro, Magela (2008), '*Elasticidad Precio de la Demanda de Gas Oil*', Planificación Comercial - ANCAP.
- [10] Samohyl, Robert W. y De Barros Dantas, Anderson, '*PREVISÃO COM COINTEGRAÇÃO E MODELO A CORREÇÃO DE ERRO: APLICAÇÃO EMPÍRICA NO CONSUMO DE GASOLINA DO BRASIL*'.

Capítulo 8

Anexo: Metodológico

8.1. Proceso Estocástico

Se denomina *proceso estocástico*, a un conjunto de T variables aleatorias (X_1, \dots, X_T) caracterizado por la distribución conjunta de las mismas. Una *serie temporal* quedará conformada con una muestra de tamaño 1 (una realización), de cada una de las T variables aleatorias $x = (x_1, \dots, x_T)$ ordenadas en el tiempo.

La distribución de probabilidad que da lugar al proceso, hace necesario la realización de supuestos para poder hacer inferencias sobre el mismo. Los supuestos mencionados, quedan detallados en el capítulo de *Metodología* de esta tesis.

8.2. Estacionariedad y Estacionalidad

Referido a los dos primeros supuestos del modelo (Ver: página: 19) se entiende que, tanto la media como la varianza son constantes en el tiempo si para cada intervalo de longitud k que se tome, si se calcula la media y la varianza éstas no varían con el intervalo. La tercer condición implica que la covarianza entre dos variables sólo dependerá de la longitud del intervalo que las separa. Si esto sucede, la serie será *estacionaria en sentido débil*. Si a esos supuestos se le agrega el supuesto de normalidad de los residuos, la serie será *estacionaria en sentido estricto*. Una serie *no estacionaria* es aquella en la cual la media y la variabilidad de la misma cambian con el tiempo. Por ejemplo, una serie con tendencia creciente ó decreciente. Si la serie tiene pautas de evolución cíclicas, es decir, que se repiten año a año ó mes a mes (dependiendo de la frecuencia de la serie)

se dirá que son *series estacionales*.

8.3. Metodología Box-Jenkins

La metodología Box-Jenkins consta de las siguientes etapas:

1. Identificación
2. Estimación
3. Diagnóstico
4. Predicción

- En la primer etapa del método, se intenta especificar uno o varios modelos de series temporales que se ajuste a los datos. La especificación se realiza por medio de las funciones de autocorrelación y autocorrelación parcial.

La *función de autocorrelación simple* (FAS) es la representación de los coeficientes de autocorrelación del proceso en función del retardo: $\rho_k = \frac{\gamma_k}{\gamma_0}$.

La *función de autocorrelación parcial* (FAP) incluye los coeficientes de correlación entre observaciones separadas por k períodos cuando se elimina la dependencia producida por los valores intermedios.

Se define γ_h como la función de autocovarianza separada por k períodos. Por lo tanto,

$$\gamma(h) = E \{ (y_t - \mu) (y_{t-h} - \mu) \}$$

Dado que se trabaja con una muestra, es necesario una estimación de las funciones de autocorrelación; la cual se obtiene mediante:

$$r(h) = \frac{\sum_{n=h+1}^T (y_n - \bar{y}) (y_{n-h} - \bar{y})}{\sum_{n=1}^T (y_n - \bar{y})^2}$$

Si las autocorrelaciones del proceso son nulas, la *varianza* asintótica de la función de autocorrelación se estima por:

$$\text{var}(r_h) = \frac{T - k}{T(T + 2)}$$

Para T suficientemente grande, la ecuación anterior se aproxima a $1/T$.

OBSERVACIÓN: Si y_t es una secuencia independiente e idénticamente distribuida, con varianza finita, se cumple que: $r(h)$ es asintóticamente *Normal* con media cero y varianza $1/T$. De aquí se desprende que las bandas de confianza para la función de autocorrelación quedan determinadas por:

$$\pm \frac{1,96}{\sqrt{T}}$$

- En la segunda y tercer etapa, el modelo se estima utilizando software adecuado y se valida respectivamente. La validación (diagnóstico), significa someter al modelo seleccionado a pruebas de ajuste, cumplimiento de supuestos relacionados a los residuos y predicción. Si se rechaza el modelo en alguna de las pruebas, comienza la cadena nuevamente con otro modelo distinto.

Para la *significación individual de los coeficientes del modelo* se utiliza la prueba de hipótesis (prueba-t):

$$H_0) \beta_j = 0$$

$$H_a) \beta_j \neq 0$$

Cuyo estadístico es:

$$t_j = \frac{\hat{\beta}_j}{\sqrt{\text{Var}(\hat{\beta}_j)}}$$

Cuya distribución es: $t_{(T-2)}^*$ (Student, con T-2 grados de libertad). Se rechazará H_0 , si $|t_j| \geq t_{\alpha/2, (T-2)}^*$ o lo que es equivalente: p-valor $\geq \alpha$

- Una vez encontrado el modelo se procede a la predicción de los datos.

8.4. Condiciones que deben cumplir los coeficientes de los modelos SARIMA(p,d,q)(P,D,Q) para que el modelo sea estacionario

$$\Phi_p(L^s)\phi_p(L)\nabla_s^D\nabla^dY_t = \theta_q(L)\Theta_Q(L^s)\varepsilon_t \quad (8.1)$$

Los coeficientes de los polinomios involucrados en la fórmula anterior deben cumplir determinadas condiciones. En el caso del polinomio autorregresivo (AR) si $|\phi| < 1$, entonces: el proceso será *estacionario*.

En el caso del polinomio de medias móviles (MA) si, $|\theta| < 1$, el proceso será *invertible*. De esta forma, el efecto de los valores pasados decrece con el tiempo. Dado que el proceso de medias móviles se genera como combinación de procesos ruido blanco siempre será estacionario para cualquier valor que tomen sus parámetros. La condición de que $|\theta| < 1$ implica que: el peso de las observaciones pasadas aumenta con la distancia temporal, lo que no se considera lógico.

8.5. Transformación Box-Cox

La familia de transformaciones *Box - Cox* tiene como finalidad corregir la falta de normalidad y la heterocedasticidad (no homogeneidad de varianzas) en los datos. La transformación se define según:

$$z(y) = \begin{cases} \frac{y^\lambda - 1}{\lambda} & \text{si } \lambda \neq 0 \\ \ln(y) & \text{si } \lambda = 0 \end{cases}$$

8.6. Tests utilizados

8.6.1. El Test de Dickey - Fuller

Las series a estudiar pueden presentar signos de no estacionariedad, como por ejemplo, tendencia creciente o decreciente. Una forma de detectar la no estacionariedad en la media es realizar el *test de Dickey Fuller*. Mediante éste test se comprueba el orden de integración de la serie; es decir el número de diferencias que se requerirán para volver la serie estacionaria.

Si se supone que la serie esta generada por un proceso AR(1):

$$y_t = c + \phi_1 y_{t-1} + \varepsilon_t$$

la prueba de Dickey-Fuller será:

$$H_0) \phi_1 = 1$$

$$H_1) \phi_1 < 1$$

Existirá raíz unitaria si $|\phi_1| = 1$ (no rechazo H_0). Por lo tanto, el modelo será no estacionario. Reescribiendo se obtiene:

$$\Delta y_t = (\phi_1 - 1)y_{t-1} + \varepsilon_t$$

Tomando $(\phi_1 - 1) = \delta$, el planteo de este modelo equivale a testear:

$$H_0) \delta = 0$$

$$H_1) \delta < 0$$

La prueba se evalúa mediante:

$$t_\delta = \hat{\delta} / SE(\hat{\delta})$$

Donde: $SE(\hat{\delta})$ es el error estandar del parámetro estimado. Bajo H_0 , el estadístico no sigue la distribución t-Student, sino que la distribución asintótica y los valores críticos fueron generados mediante simulación y colocados en las tablas de Dickey Fuller.

Variantes a los modelos:

1. Test para raíz unitaria con constante y tendencia

$$\Delta y_t = \mu + \beta_t + \gamma y_{t-1} + \varepsilon_t$$

2. Test para raíz unitaria con constante

$$\Delta y_t = \mu + \gamma y_{t-1} + \varepsilon_t$$

3. Test para raíz unitaria

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t$$

Como el test de Dickey Fuller es planteado para series que siguen un proceso AR(1), el test se generaliza a procesos AR(p) mediante el test de DFA (Dickey Fuller Aumentado).

8.6.2. Test de Ljung - Box

Se considera el estadístico:

$$Q = n(n+2) \sum_{j=1}^k \rho^2(j) / (n-j) \sim \chi^2(k)$$

Para el cual la prueba de hipótesis planteada será:

$$H_0 : \rho_i = 0 \text{ para } i = 1, \dots, k$$

$$H_1 : \text{al menos un } \rho_i \neq 0$$

Si el valor del estadístico Q calculado excede el valor crítico de $\chi^2(k)$, por lo menos, algún valor de la función de autocorrelación para los primeros k retardos será significativamente distinto de cero. Lo que implica rechazar la hipótesis nula y concluir que existe autocorrelación en los primeros k retardos considerados.

8.6.3. Test de Kolmogorov - Smirnov

Se considera el estadístico:

$$D = \sup |F_n(x) - F(x)|$$

La prueba esta implementada para testear distribuciones continuas.

En este caso, se analiza la máxima distancia entre las curvas de distribución de un conjunto de datos ($F_n(x)$) y alguna distribución teórica ($F(x)$). Como se indica en los supuestos del modelo, los residuos del mismo deben ser normales; por lo tanto la distribución teórica elegida será *Normal*. La prueba planteada es:

$$H_0) F_n(x) = F(x)$$

$$H_1) F_n(x) \neq F(x)$$

8.6.4. Test de Shapiro - Wilk

Para corroborar si una variable se distribuye Normal se puede utilizar el estadístico W para la prueba de Shapiro-Wilk.

$$W = \frac{\sum_{i=1}^{[n/2]} (X_{(n-i+1)} - X_{(i)}) a_{(n-i+1)}}{(n-1) S^2}$$

Donde:

- X_1, X_2, \dots, X_n es una muestra aleatoria de una variable aleatoria X
- X_i como la i -ésima variable de menor valor: $X_1 \leq X_2 \leq \dots \leq X_{n-1} \leq X_n$
- $a_{(n-i+1)}$ son coeficientes de Shapiro - Wilk
- varianza muestral

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Se acepta H_0 si $W \leq W_{n,\alpha}$, donde $W_{n,\alpha}$ está tabulado en las tablas del test de Shapiro - Wilk.

8.6.5. Test Doornik - Hansen

Test basado en transformación de la simetría (b_1) y la kurtosis muestral (b_2) definidas según:

$$\sqrt{b_1} = \frac{\mu_3}{\mu_2^{3/2}}$$

$$b_2 = \frac{m_4}{m_2^2}$$

Se utilizan en el estadístico de Bowman y Shenton:

$$E = \frac{n(\sqrt{b_1})^2}{6} + \frac{n(b_2 - 3)^2}{24}$$

El cual distribuye χ_2^2

8.6.6. Test de Johansen

El test que se presenta a continuación es una generalización multivariada del *test de Dickey Fuller*. A su vez, como el test de Dickey Fuller es generalizado de un proceso AR(1) a un proceso AR(p), también el test de Johansen tiene esa propiedad.

Se define un vector de n variables x_t el cual sigue un proceso AR(p):

$$x_t = A_1 x_{t-1} + A_2 x_{t-2} + \dots + \varepsilon_t$$

Se suma y se resta x_{t-1} a ambos lados de la ecuación; se obtiene:

$$\Delta x_t = (A_1 - I) x_{t-1} + \dots + A_p x_{t-p} + \varepsilon_t$$

Se suma y se resta $(A_1 - I)x_{t-2}$, luego $(A_2 + A_1 - I)x_{t-3}$ y así sucesivamente de forma de obtener el siguiente algoritmo:

$$\Delta x_t = \sum_{i=1}^p \Delta x_{t-i} + \Pi x_{t-1} + \varepsilon_t$$

donde la matriz Φ queda definida según:

$$\Pi = - \left[I - \sum_{i=1}^p A_j \right]$$

El número de vectores de cointegración (r) depende del rango de la matriz Π , el cual se define por medio de las raíces características distintas a cero de dicha matriz.

Si se ordenan los valores propios de la matriz Π de mayor a menor, tal que: $\lambda_1 \rangle \lambda_2 \rangle \dots \rangle \lambda_n$, Johansen plantea dos posibles tests para encontrar el rango de la matriz:

Estadístico de la traza

$$\lambda_{traza} = -T \sum_{i=r+1}^n \text{Ln}(1 - \lambda_i)$$

que permite testear la hipótesis nula de que $\text{rango}(\Pi) = r$ contra la alternativa de que $\text{rango}(\Pi) = p$.

Estadístico de máximo valor propio

También es posible testear la hipótesis nula de que el $\text{rango}(\Pi) = r$ contra la hipótesis alternativa de que el $\text{rango}(\Pi) = r + 1$, mediante el estadístico del máximo valor propio:

$$\lambda_{\max} = -T \log(1 - \lambda_{r+1})$$

Se cumple que si $p = r+1$ entonces ambos estadísticos coinciden.

8.7. Akaike Information (AIC)

El criterio AIC es utilizado para la selección de los modelos que se pueden estar considerando como posibles, para modelizar la serie de datos. El modelo consta de dos componentes: el primero ($\text{Ln} \left(\frac{SCR}{T} \right)$), referido a la minimización de los errores y el segundo ($2 \left(\frac{1+P+Q+p+q}{T} \right)$), que penaliza la incorporación de parámetros adicionales.

$$AIC = \text{Ln} \left(\frac{SCR}{T} \right) + 2 \left(\frac{1 + P + Q + p + q}{T} \right)$$

Donde:

- SCR es la suma de cuadrados de los residuos definida según:

$$SCR = \sum_{i=1}^t (y_i - \hat{y}_i)^2$$

- T es el número de observaciones
- p y P son el orden del polinomio autorregresivo regular y estacional respectivamente.

- q y Q son el orden del polinomio de medias móviles regular y estacional respectivamente.

El modelos con menor valor de AIC será escogido como el que mejor se ajusta a los datos.

El AIC presenta una desventaja: sobre-estima el número de parámetros en el modelo. Una alternativa que corrige la sobre-estimación es el AICc (AIC corregido), el cual se plantea de la siguiente manera:

$$AICc = T \log(\hat{\sigma}_{MV}^2) + T \frac{(1 + k/T)}{1 - (k + 2)/T}$$

Donde:

- $(\hat{\sigma}_{MV}^2)$ es el estimador máximo verosímil de la varianza de las innovaciones.
- k es el número de parámetros estimadas para calcular las predicciones a 1 paso.

8.8. Diferenciación de la serie

La diferenciación de la serie es un método general en el cual no se realiza ninguna hipótesis sobre la tendencia a corto plazo; sólo se supondrá que ésta evoluciona lentamente en el tiempo. Se asumirá que la tendencia en el instante t es muy próxima a la tendencia en el instante $t-1$. Por lo anterior, si se genera una nueva serie, $z_t = y_t - y_{t-1}$, se quitará el efecto de la tendencia de la serie original (se denomina *serie diferenciada*):

$$z_t = \underbrace{y_t - y_{t-1}}_{\nabla y_t} = (1 - L)y_t \quad (8.2)$$

8.9. Análisis de Intervención

Las series temporales pueden ser analizadas con información del pasado o ingresando información adicional. Cuando la serie se ve afectada por sucesos puntuales conocidos, como una huelga o una inundación, la inclusión de estos efectos en el análisis puede mejorar el resultado de la estimación del modelo y de las previsiones. El

efecto producido por el fenómeno en cuestión puede ser medido mediante una variable que tome valor 1 en el momento del tiempo $t=h$ (momento del fenómeno) y cero en el resto. La variable se denomina: *variable impulso*.

Si la serie de datos está afectada por algún tipo de fenómeno que pueda representarse mediante la variable impulso; se obtendrá:

$$z_t = y_t + w_0 I_t^{(h)}$$

Dónde:

- $I_t^{(h)}$ toma valor 1 en $t=h$ y 0 en $t \neq h$.
- w_0 es la magnitud del efecto sobre la serie y_t
- y_t es una variable que sigue un modelo de series temporales SARIMA.

8.10. Predictor óptimo para modelos ARIMA

A continuación se desarrolla el predictor óptimo para modelos ARIMA tal como se mencionó en la sección 5.1.3.

Partiendo de ecuación 3.8, si se suma y resta $\mu_{n+k/n} = ECM(z_{n+k}/z_n)$ y si se desarrolla el cuadrado de la ecuación se obtiene:

$$ECM(z_{n+k}/z_n) = E \left[\underbrace{\left(z_{n+k} - \underbrace{\mu_{n+k/n}}_{ECM(z_{n+k}/z_n)} \right)^2}_{Var(z_{n+k}/z_n)} / z_n \right] + E \left[\left(\underbrace{\mu_{n+k/n} - \hat{z}_{n+k}}_{\text{termino constante}} \right)^2 / z_n \right]$$

El término: $\mu_{n+k/n} - \hat{z}_{n+k}$ es constante, ya que \hat{z}_{n+k} esta condicionado a valores pasados (según se define en la ecuación 3.7) por lo que queda fuera de la esperanza.

Dado que la varianza no depende del predictor, el mínimo de la ecuación 3.9 se obtiene igualando $\mu_{n+k/n} - \hat{z}_{n+k}$ a cero. Como resultado se obtiene:

$$\hat{z}_{n+k} = \mu_{n+k/n}$$

Por lo tanto, el predictor que minimiza el ECM es la esperanza de la variable z_{n+k} condicionada a la información pasada.

8.10.1. Intervalos de las predicciones

Si las innovaciones cumplen con los supuestos establecidos de normalidad, entonces la variable $\hat{z}_t(k)$ será una variable normal de media z_{n+k}

$$z_{t+k} \in \left(\hat{z}_t(k) \pm \lambda_{1-\alpha} \sqrt{Var(\varepsilon_t(k))} \right)$$

Donde:

- $\lambda_{1-\alpha}$ corresponde al percentil de la distribución Normal(0,1).

Capítulo 9

Anexo: Univariado Gas Oil

9.1. Test de Dickey Fuller

```

-----
Augmented Dickey & Fuller test
-----
Deterministic regressors estimates:
      Estimate Std. Error t value Pr(>|t|)
Intercept  -2.303      1.388  -1.659   0.102
Trend       0.044      0.020   2.150   0.035
-----
ADF regressor estimate:
      Estimate Std. Error t value Pr(>|t|)
adf.reg   -9.088      1.077  -8.441   0.01
-----
Selected lags estimates:
      Estimate Std. Error t value Pr(>|t|)
Lag.1      7.224      1.028   7.028   0.000
Lag.2      6.418      0.961   6.682   0.000
Lag.3      5.578      0.868   6.428   0.000
Lag.4      4.671      0.758   6.164   0.000
Lag.5      3.944      0.642   6.145   0.000
Lag.6      3.198      0.532   6.014   0.000
Lag.7      2.440      0.415   5.876   0.000
Lag.8      1.843      0.306   6.027   0.000
Lag.9      1.280      0.196   6.519   0.000
Lag.10     0.570      0.100   5.681   0.000
Lag.14    -0.050      0.053  -0.942   0.350
Lag.20     0.010      0.052   0.188   0.851

Number of available observations: 86

```

Figura 9.1: Test de Dickey Fuller para la segunda diferencia de la serie de consumo de gas oil en el período Enero de 1999 a Diciembre del 2007

9.2. Procedimiento realizado para la obtención del modelo final

Partiendo del análisis de los cuatro modelos propuestos para la modelización de la demanda de Gas Oil se obtiene:

■ **Referido a los Coeficientes de los modelos estimados:**

Resulta que únicamente los coeficientes de los modelos, $ARIMA(2,1,0)(0,0,1)$ y $ARIMA(2,1,0)(1,0,0)$, son significativos al 5%. A continuación se presentan las salidas correspondientes, tanto de los modelos aceptados como de los rechazados.

Modelo $ARIMA(1,1,1)(0,0,1)$				Modelo $(2,1,1)(0,0,1)$			
	coef	t-stat	critico		coef	t-stat	critico
ar1	0,18	1,63	1,66	ar1	0,17	1,59	1,66
ma1	-0,91	-19,46		ar2	-0,03	-0,34	
sma1	0,48	5,42		ma1	-0,92	-16,86	
				sma1	0,48	5,38	

Tabla 9.1: Estimación de los modelos Rechazados para la demanda de Gas Oil: $ARIMA(1,1,1)(0,0,1)$ y $ARIMA(2,1,1)(0,0,1)$

Modelo $ARIMA(2,1,0)(1,0,0)$				Modelo $ARIMA(2,1,0)(0,0,1)$			
	coef	t-stat	critico		coef	t-stat	critico
ar1	-0,57	-6,22	1,66	ar1	-0,52	-5,56	1,66
ar2	-0,32	-3,46		ar2	-0,31	-3,20	
sar1	0,49	5,71		sma1	0,5	5,36	
AICc = 656,78				AICc = 658,55			
Box-Ljung, df = 24, p-value = 0,05				Box-Ljung test, df = 24, p-value = 0,11			
Shapiro-Wilk, p-value = 0,67				Shapiro-Wilk, p-value = 0,53			

Tabla 9.2: Estimación de los modelos Aceptados $ARIMA(2,1,0)(1,0,0)$ y $ARIMA(2,1,0)(0,0,1)$

■ **Referido al análisis de los residuos de los modelos:**

El análisis de los residuos de los modelos aceptados permite comprobar que:

La distribución de los residuos en ambos casos resulta ser Normal. Se realiza la prueba de Shapiro - Wilk sobre los residuos de ambos modelos, obteniéndose para el modelo $ARIMA(2,1,0)(1,0,0)$ un p-valor de 0.67 y para el modelo $ARIMA(2,1,0)(0,0,1)$ p-valor igual a 0.53. Véase tabla: 9.2

La prueba de *Ljung-Box* sobre los residuos hace no rechazar la hipótesis nula de ausencia de autocorrelación en los primeros 24 retardos. Se obtiene para el modelo ARIMA(2,1,1)(1,0,0): p-valor = 0,05 y ARIMA(2,1,0)(0,0,1): p-valor = 0,11.

A continuación se presentan los autocorrelogramas simple y parcial. Se observa que los retardos 3 y 23 quedan fuera de las bandas de confianza. Sin embargo, se reitera que la prueba de *Ljung-Box* no es rechazada a una nivel de significación del 5% tomando en cuenta los primeros 24 retardos.

A su vez el gráfico de los residuos Figura: 9.3, evidencia un punto que podría ser tratado como posible *valor atípico*: **Octubre del 2003** (en ambos modelos). Se observa que el efecto provocado por este valor es puntual, es decir el suceso ocurre en el momento $t = \text{Octubre de 2003}$ y no se traslada a lo largo de la serie. Este tipo de atípico se denomina *atípico aditivo*¹

Se analizan nuevamente los modelos ingresando una variable indicatriz la cual toma valor 1 en el momento $t = \text{Octubre de 2003}$ y 0 en el resto.

¹En el Anexo Metodológico: Análisis de Intervención, sección:8.9 se puede leer más sobre el análisis de intervención en las series temporales.

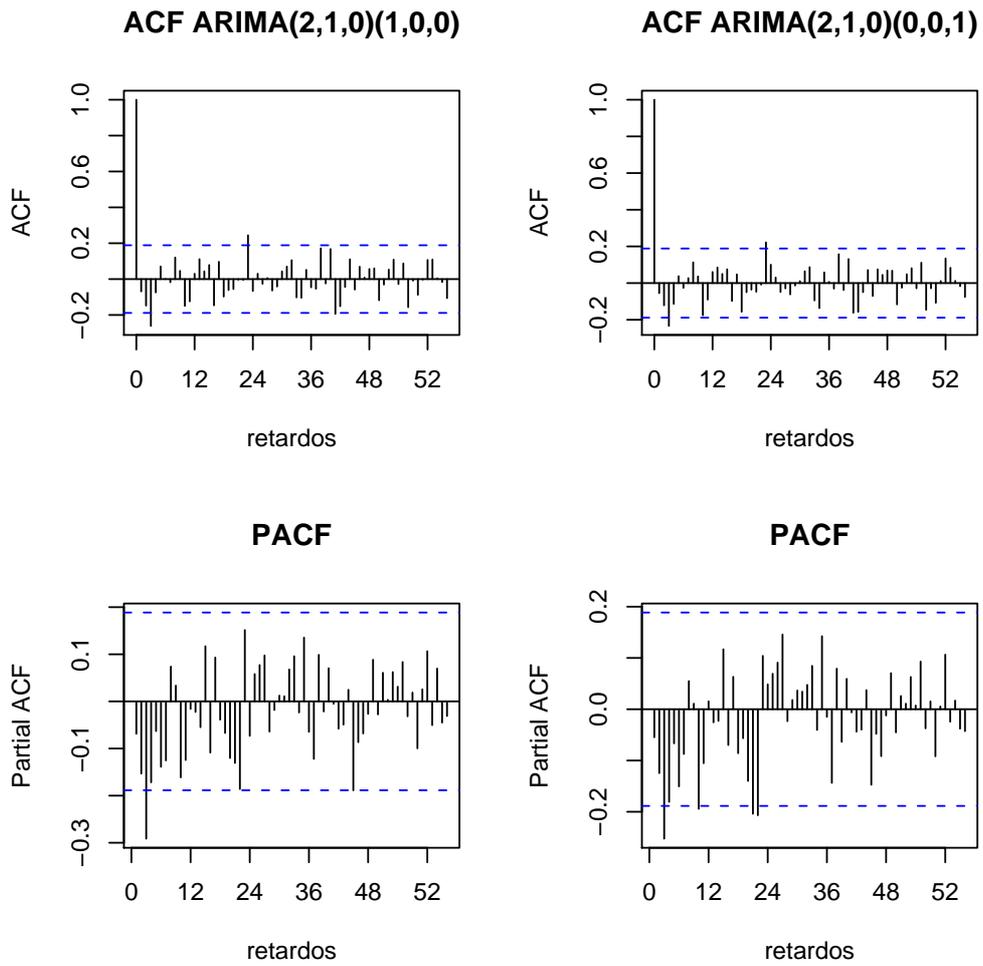


Figura 9.2: Autocorrelogramas Simple y Parcial de los residuos de los modelos: ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1)

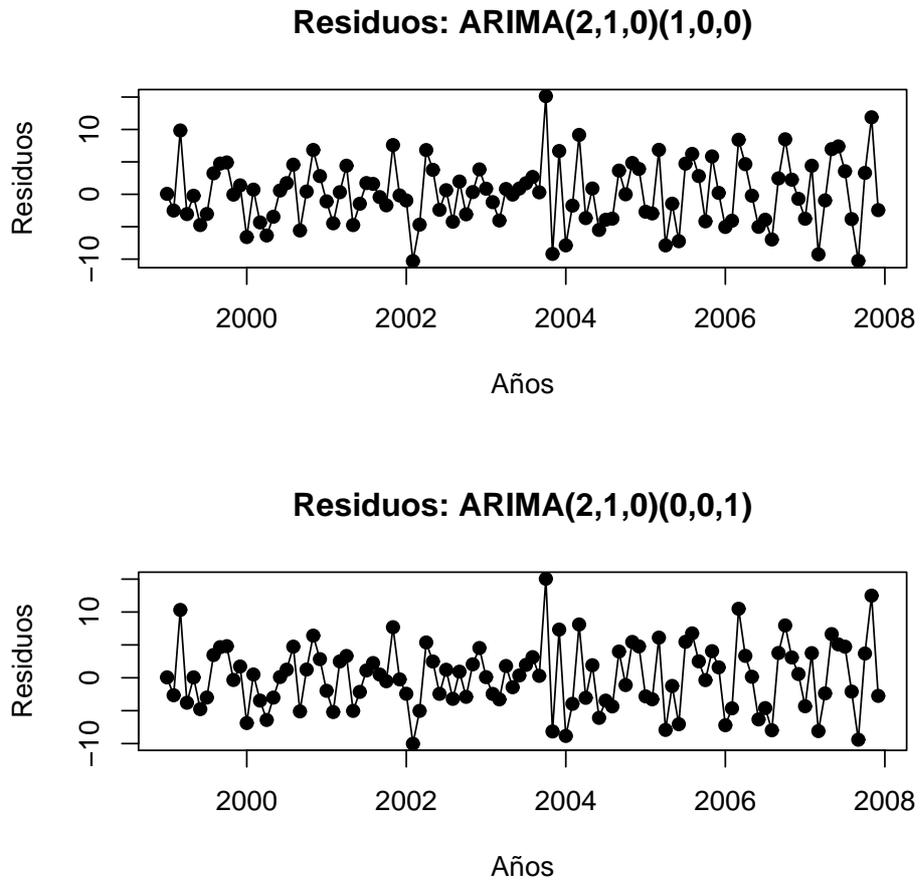


Figura 9.3: Grafica de los Residuos de los modelos ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1)

9.2.1. Resultados obtenidos del análisis de modelos

ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1) con inclusión de variable indicatriz: Octubre de 2003

A continuación se presenta la salida correspondiente al análisis planteado:

ARIMA(2,1,0)(1,0,0)				ARIMA(2,1,0)(0,0,1)			
	coef	t-stat	critico		coef	t-stat	critico
ar1	-0,48	-5,21	1,66	ar1	-0,43	-4,62	1,66
ar2	-0,38	-3,98		ar2	-0,38	-3,99	
sar1	0,54	6,34		sma1	0,54	5,66	
Oct de 2003	14,01	3,88		Oct de 2003	12,89	3,71	
AIC = 645,66				AIC = 648,59			
Box-Ljung df = 24, p-value = 0,12				Box-Ljung p-value = 0,16			
Shapiro-Wilk p-value = 0,44				Shapiro-Wilk p-value = 0,48			

Tabla 9.3: Estimación de los modelos ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1) con la inclusión de indicatriz Octubre de 2003

Según la salida anterior se puede concluir que:

1. La variable indicatriz Octubre del 2003 ha resultado significativa para ambos modelos.
2. Los coeficientes de los modelos resultaron significativos al 5 %.
3. Se observa que el AIC del modelo ARIMA(2,1,0)(1,0,0) con Octubre de 2003, resulta levemente menor que el otro modelo estudiado.
4. En cuanto al análisis de los residuos, resultaron no tener autocorrelación (según test de *Ljung - Box*) y ser normales (prueba de *Shapiro - Wilk*). Se observa que el retardo 3 sigue sin ser captado por la modelización; queda fuera de las bandas de confianza; al igual que el retardo 22. Figura: 9.4.

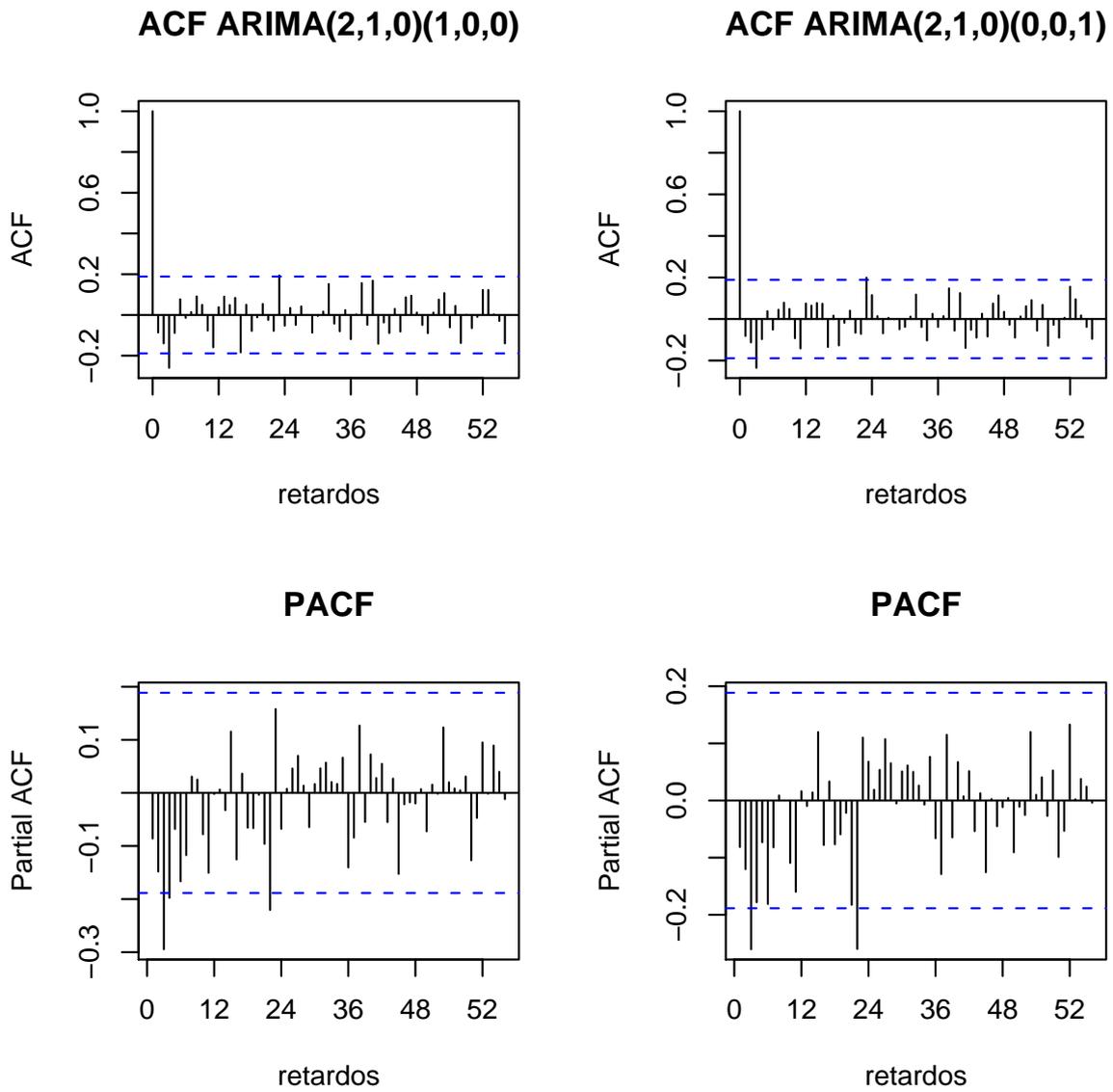


Figura 9.4: Autocorrelación simple y parcial de los modelos ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1) con la inclusión de Octubre de 2003 como atípico aditivo

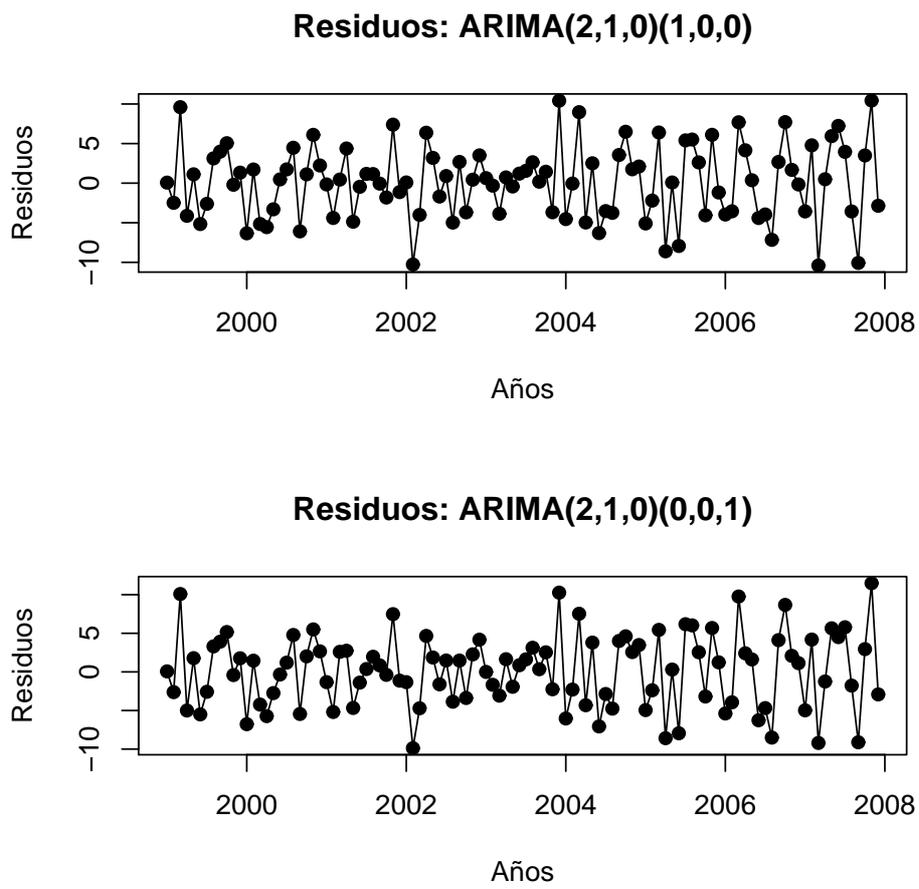


Figura 9.5: Residuos de los modelos modelos ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1) con la inclusión de Octubre de 2003 como atípico aditivo

OBSERVACIÓN: La figura 9.5, sugiere otros posibles valores atípicos: **Febrero de 2002, Setiembre y Noviembre de 2008** . Se analiza este aspecto incluyendo nuevas variables indicatrices en el modelo.

Como resultado se obtiene:

1. El AIC de los modelos al incluir las variables indicatrices no tiene mejora significativa respecto a los modelos que solo incluyen Octubre de 2003 comparar AIC de los modelos en las tablas: 9.3 y 9.4
2. La variable Febrero de 2002 resulta ser aceptada con un valor del estadístico observado muy próximo al valor crítico de la prueba t (para ambos modelos). Por tal motivo, se decide descartarla como atípico.

ARIMA(2,1,0)(1,0,0)				ARIMA(2,1,0)(0,0,1)			
	coef	t-stat	critico		coef	t-stat	critico
ar1	-0,53	-5,60	1,66	ar1	-0,46	-4,69	1,66
ar2	-0,35	-3,72		ar2	-0,33	-3,42	
sar1	0,53	6,5		sma1	0,53	5,54	
Oct-2003	13,15	3,82		Oct-2007	12,06	3,59	
Feb-02	-6,4	-1,93		Feb-2002	-5,69	-1,77	
Set-2007	-9,24	-2,51		Set-2007	-8,49	-2,30	
Nov-2007	9,99	2,56		Nov-2007	11,03	2,86	
AIC =637,14				AIC = 640,68			
Box-Ljung, df = 24, p-value = 0,29				Box-Ljung, df = 24, p-value = 0,24			
Shapiro-Wilk p-value = 0,35				Shapiro-Wilk p-value = 0,41			

Tabla 9.4: Estimación de los modelos ARIMA(2,1,0)(1,0,0) y ARIMA(2,1,0)(0,0,1) con la inclusión de Octubre de 2003, Febrero de 2002, Setiembre y Noviembre de 2007 como variables indicatrices

9.2.2. Prueba en predicción de los modelos preliminares

Se realizan predicciones a 6 pasos para ambos modelos con las indicatrices: Octubre de 2003, Setiembre y Noviembre de 2007, en el período Julio - Diciembre de 2007. Observación: como la muestra se corta en Junio de 2007 (para realizar las predicciones) únicamente se incluye como atípico la observación de Octubre de 2003.

El resultado obtenido es:

Ninguno de los dos modelos deja las observaciones de Setiembre y Noviembre fuera de los intervalos de predicción. Por lo tanto, si bien el consumo fue elevado para la fecha esta dentro de lo que se prevé. Observación: para el modelo ARIMA(2,1,0)(1,0,0) la observación de Setiembre coincide con el límite inferior del intervalo. Ver tablas: 9.5 y 9.6.

Con esta evidencia se decide no incorporar como valores atípicos las observaciones de Setiembre y Noviembre de 2007.

En cuanto a los errores porcentuales en los que incurren cada uno de los modelos: no existen grandes diferencias, por lo menos no las hay en esa ventana de predicción. Si existe diferencia en el error cuadrático medio, siendo menor para el modelo ARIMA(2,1,0)(1,0,0).

Se realiza la predicción para el año 2008 y se decide el modelo final en base a esto. Se obtiene como resultado que el modelo que minimiza el error cuadrático medio es el modelo: ARIMA(2,1,0)(1,0,0) con Octubre de 2003 como atípico. Observación la comparación del ECM para el año 2008 es entre las tablas: 9.7 y 5.3.

ARIMA(2,1,0)(1,0,0)			Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		Error Porcentual
	obs	pred	Inf	Sup	Inf	Sup	
Jul-07	73,4	69,70	60,96	78,44	58,21	81,19	-5 %
Ago-07	67,4	69,25	59,53	78,98	56,47	82,04	3 %
Sep-07	62,7	73,13	62,78	83,48	59,53	86,74	17 %
Oct-07	74,7	75,33	63,63	87,04	59,96	90,71	1 %
Nov-07	83,98	75,70	63,08	88,32	59,11	92,28	-10 %
Dic-07	73,02	75,59	62,21	88,97	58,01	93,17	4 %
ECM = 33.58							

Tabla 9.5: Predicción Jul - Dic 2007. Modelo ARIMA(2,1,0)(1,0,0)

9.3. Conclusión:

El modelo finalmente utilizado para modelizar la demanda de gas oil es el modelo: ARIMA(2,1,0)(1,0,0) con atípico aditivo Octubre de 2003 debido a que cumple con todos los supuestos de la modelización propuesta e incurre en un menor error cuadrático medio.

ARIMA(2,1,0)(0,0,1)			Intervalo de Predicción 95 %		Intervalo de Predicción 99 %		
	obs	pred	Inf	Sup	Inf	Sup	Error Porcentual
Jul-07	73,40	67,86	59,02	76,70	56,24	79,47	-8 %
Ago-07	67,40	65,85	55,71	75,99	52,52	79,18	-2 %
Sep-07	62,70	70,62	59,86	81,37	56,48	84,75	13 %
Oct-07	74,70	74,21	62,11	86,31	58,31	90,11	-1 %
Nov-07	83,98	71,99	58,83	85,15	54,70	89,28	-14 %
Dic-07	73,02	72,14	58,18	86,09	53,80	90,48	-1 %
ECM = 40,10							

Tabla 9.6: Predicción: Jul - Dic de 2007. Modelo ARIMA(2,1,0)(0,0,1)

			Intervalo de Predicción 95 %		Intervalo de Predicción 99 %		
	obs	pred	Inf	Sup	Inf	Sup	Error porcentual
Ene-08	68,99	71,43	62,28	80,58	59,41	83,46	4 %
Feb-08	67,37	78,64	68,11	89,17	64,80	92,48	17 %
Mar-08	72,51	71,17	60,11	82,22	56,64	85,70	-2 %
Abr-08	81,64	70,91	58,48	83,34	54,57	87,24	-13 %
May-08	73,20	76,98	63,43	90,53	59,17	94,79	5 %
Jun-08	66,52	76,91	62,59	91,23	58,09	95,73	16 %
Jul-08	68,79	77,71	62,53	92,90	57,75	97,67	13 %
Ago-08	65,01	76,43	60,39	92,48	55,34	97,52	13 %
Sep-08		71,73	54,94	88,53	49,66	93,80	
Oct-08		75,85	58,32	93,37	52,82	98,88	
Nov-08		82,13	63,88	100,37	58,15	106,10	
Dic-08		76,25	57,32	95,17	51,38	101,11	
ECM = 72,75							

Tabla 9.7: Predicción para el año 2008. Modelo ARIMA(2,1,0)(0,0,1)

9.4. Predicción Junio 2007 - Julio 2008 - Modelo seleccionado

			Intervalo de Predicción 95 %		Intervalo de Predicción 99 %		ERROR
	pred		Inf	Sup	Inf	Sup	
Jul-07	73,4	69,70	60,96	58,21	78,44	81,19	-5 %
Ago-07	67,4	69,25	59,53	56,47	78,98	82,04	3 %
Sep-07	62,7	73,13	62,78	59,53	83,48	86,74	17 %
Oct-07	74,7	75,33	63,63	59,96	87,04	90,71	1 %
Nov-07	83,98	75,70	63,08	59,11	88,32	92,28	-10 %
Dic-07	73,02	75,59	62,21	58,01	88,97	93,17	4 %
Ene-08	68,99	71,31	57,08	52,61	85,54	90,01	3 %
Feb-08	67,37	74,30	59,30	54,59	89,29	94,00	10 %
Mar-08	72,51	72,59	56,89	51,96	88,29	93,22	0 %
Abr-08	81,64	72,19	55,80	50,65	88,58	93,74	-12 %
May-08	73,2	76,48	59,42	54,06	93,54	98,90	4 %
Jun-08	66,52	75,38	57,69	52,13	93,07	98,63	13 %
ECM-univariado = 36,12							

Tabla 9.8: Predicción: Julio 2007 - Junio 2008 - Modelo seleccionado

9.5. Que sucede si se intenta captar el retardo 3?

Para intentar captar el retardo 3 se toma el modelo $ARIMA(3,1,0)(1,0,0)$.

Los resultados son:

Los coeficientes del modelo y la observación de Octubre de 2003 resultan significativos al 5 %.

Los residuos son normales, con un p-valor de 0.8 en la prueba de Shapiro - Wilk.

Sin embargo el test de Ljung - Box esta muy próximo a rechazarse. Como se observa en la figura: 9.6 el retardo 3 ya no queda fuera de las bandas de confianza, ahora es el retardo 4.

Esta modelización no ofrece una solución para que los retardos que se intentan captar, sean captados.

ARIMA(3,1,0)(1,0,0)			
	coef	t-stat	critico
ar1	-0,56	-5,84	1,66
ar2	-0,48	-4,74	
ar3	-0,23	-2,33	
sar1	0,54	6,52	
Oct de 2003	14,05	3,94	
AIC = 642,63			
Box-Ljung df = 24, p-value = 0,08			
Shapiro-Wilk p-value = 0,80			

Tabla 9.9: Estimación del modelo ARIMA(3,1,0)(1,0,0) con Octubre de 2003 como atípico

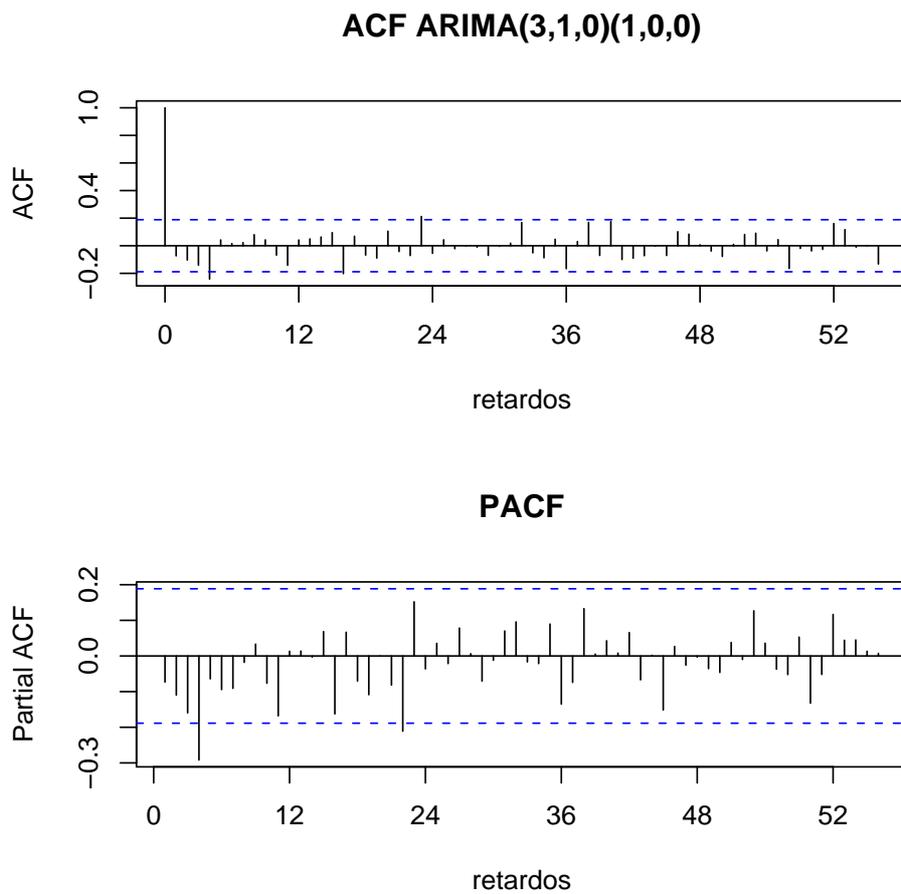


Figura 9.6: Autocorrelograma simple y parcial del Modelo ARI-MA(3,1,0)(1,0,0)

Capítulo 10

Anexo: Univariado SUPERGAS

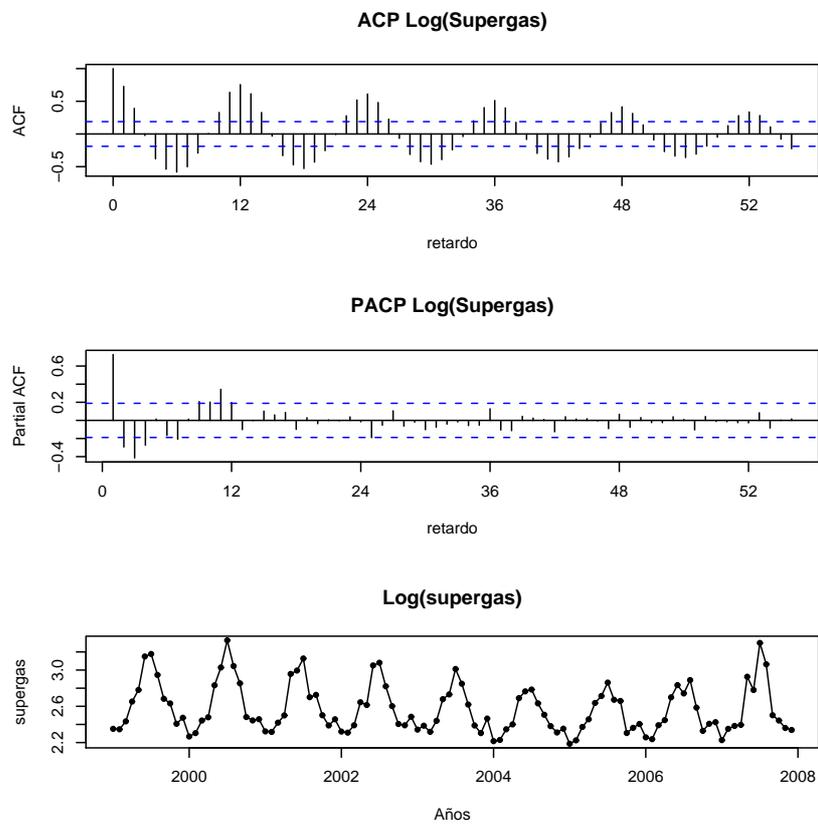


Figura 10.1: Autocorrelogramas simple y parcial de la serie de Consumo de Supergas en logaritmos y gráfica del logaritmo del Consumo de Supergas.

10.1. Procedimiento realizado para la obtención del modelo final

Partiendo de los modelos propuestos para la modelización de la demanda del logaritmo del supergas se obtiene:

1. Referido a los coeficientes de los modelos estimados

Como resultado de la estimación de los modelos sugeridos se obtiene que únicamente los coeficientes de los modelos SARIMA(0,0,0)(1,1,0) y SARIMA(0,0,0)(0,1,1) fueron significativamente distintos de cero. Ver tabla: 10.2.

La tabla: 10.1 muestra los modelos que fueron rechazados. Cabe acotar que los modelos: SARIMA(0,0,0)(1,1,1) y SARIMA(1,0,1)(0,1,1) son rechazados debido a que sus parámetros estimados: SMA(1) y AR(1) respectivamente, son cercanos a 1, lo que puede llevar a la no estacionariedad del modelo¹. Para los restantes modelos rechazados, el valor absoluto del estadístico calculado para los coeficientes estimados (t-obs) es menor al valor crítico de la prueba, por lo tanto también se rechazan.

Los modelos que pasan la significación individual de los coeficientes (SARIMA(0,0,0)(1,1,0) y SARIMA(0,0,0)(0,1,1)), presentan bajo valor-p en la prueba de Ljung - Box: 0.08 y 0.05 respectivamente. Se analiza el porque de esta situación a través de los autocorrelogramas simple y parcial.

2. Referido al análisis de los residuos

El análisis de los residuos de los modelos aceptados hace concluir que:

La distribución de los residuos en ambos casos resultó no ser Normal. Se realiza la prueba de Shapiro - Wilk sobre ambos modelos obteniéndose para el modelo ARIMA(0,0,0)(1,1,0): p-valor = 2.5e-06 y para el modelo ARIMA(0,0,0)(0,1,1): p-valor = 1.45e-06. Datos en tabla: 10.2.

¹Ver: Anexo Metodológico: sec: 8.2 y 8.4

SARIMA(0,0,0)(0,1,2)	coef	t-obs	critico	SARIMA(0,0,0)(1,1,1)	coef	t-obs	critico
sma1	-0,35	-2,88	1,66	sar1	0,56	4,49	1,66
sma2	-0,11	-0,68		sma1	-0,99	-3,24	
aic = -131,8				aic = -133,24			
Shapiro-Wilk p-value = 1,78e-06				Shapiro-Wilk p-value = 1,1e-05			
Box-Ljung, df=24, p-value = 0,12				Box-Ljung, df=24, p-value = 0,5			
SARIMA(0,0,1)(0,1,1)	coef	t-obs	critico	SARIMA(1,0,0)(0,1,1)	coef	t-obs	critico
ma1	0,13	1,43	1,66	ar1	0,23	1,98	1,66
sma1	-0,44	-3,26		sma1	-0,53	-3,66	
aic = -133,37				aic = -135,07			
Shapiro-Wilk p-value = 3,34e-06				Shapiro-Wilk p-value = 1,45e-05			
Box-Ljung, df=24, p-value = 0,14				Box-Ljung, df=24, p-value = 0,25			
SARIMA(0,0,1)(1,1,1)	coef	t-obs	critico	SARIMA(1,0,1)(0,1,1)	coef	t-obs	critico
ma1	0,09	0,83	1,66	ar1	0,97	32,32	1,66
sar1	0,48	1,02		ma1	-0,84	-12,80	
sma1	-0,91	-1,15		sma1	-0,69	-5,11	
aic = -132,2				aic = -145,75			
Shapiro-Wilk p-value = 1,76e-05				Shapiro-Wilk p-value = 0,0005331			
Box-Ljung, df=24, p-value = 0,49				Box-Ljung, df= 24, p-value = 0,68			

Tabla 10.1: Estimación de los coeficientes de los modelos rechazados para la modelización del supergas

SARIMA(0,0,0)(1,1,0)				SARIMA(0,0,0)(0,1,1)			
	coef	t-obs	critico		coef	t-obs	critico
sar1	-0,29	-2,55	1,66	sma1	-0,36	-2,85	1,66
aic = -131,93				aic = -147,12			
Shapiro-Wilk p-value = 2,5e-06				Shapiro-Wilk p-value = 1,45e-06			
Box-Ljung, df=24, p-value = 0,08				Box-Ljung, df=24, p-value = 0,05			

Tabla 10.2: Estimación de los coeficientes de los modelos aceptados para la modelización del supergas

La prueba de *Ljung-Box* sobre los residuos hace aceptar la hipótesis nula de ausencia de autocorrelación en los primeros 24 retardos. Se obtiene para el modelo ARIMA(0,0,0)(1,1,0): p-valor = 0,08 y ARIMA(0,0,0)(0,1,1): p-valor = 0.05.

Los autocorrelogramas de ambos modelos figura: 10.2 muestran que la parte regular y la parte estacional no han sido totalmente captadas por la modelización propuesta. Véase que el segundo retardo es significativo en el autocorrelograma parcial, lo que puede indicar una modelización AR(2) y para el retardo 24 se podría probar con un SMA(2) o SAR(2). Si bien la cantidad de retardos por fuera de las bandas de confianza pueden estar contemplados en el error de la prueba de Ljung - Box (5 % de los retardos seleccionados (24)), no es una situación deseable.

A su vez el gráfico de los residuos figura: 10.3, evidencia puntos comunes a ambos modelos que podrían ser tratados como posibles *valores atípicos*. Ellos son: **Agosto de 2001, Mayo de 2002, Junio de 2003 y Julio de 2007**. Se observa que el efecto provocado por estos valores es puntual. Los sucesos ocurren en el momento $t = h$, para $h = \text{Agosto de 2001, Mayo de 2002, Junio de 2003 y Julio de 2007}$, y no se trasladan a lo largo de la serie. Este tipo de atípico se denomina *atípico aditivo*.²

El incumplimiento del supuesto de normalidad, puede estar ligado la influencia de estos valores. Por lo tanto, se analizan los siguientes modelos con la inclusión de los posibles atípicos mencionados.

- SARIMA(2,0,0)(0,1,1)
- SARIMA(0,0,2)(0,1,1)
- SARIMA(2,0,0)(0,1,2)
- SARIMA(2,0,0)(2,1,0)

²En el Anexo Metodológico: Análisis de Intervención, sección:8.9 se puede leer más sobre el análisis de intervención en las series temporales.

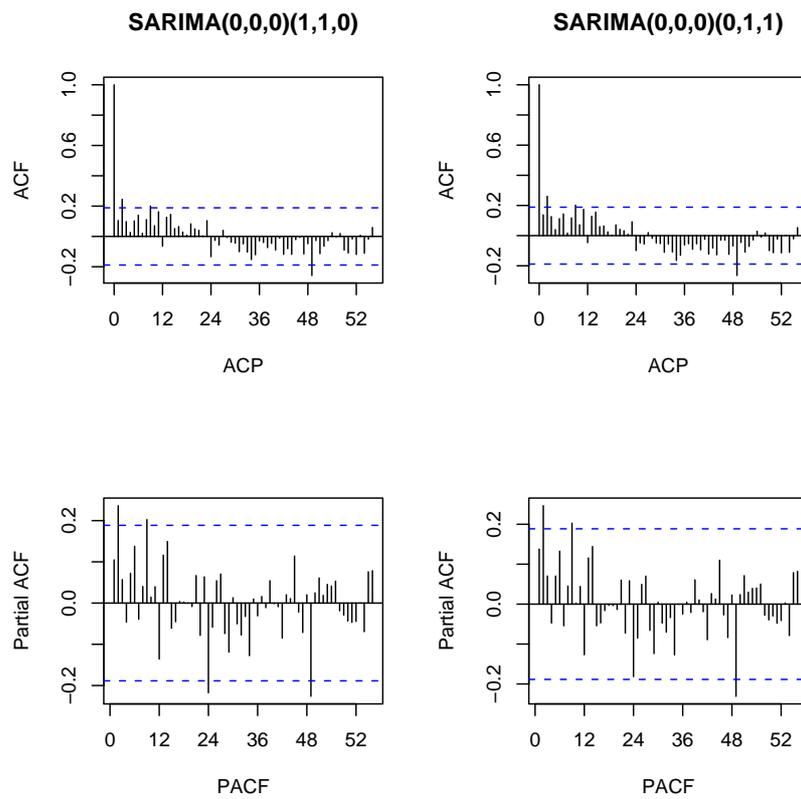


Figura 10.2: Autocorrelogramas simple y parcial de los modelos SARI-MA(0,0,0)(1,1,0) y SARIMA(0,0,0)(0,1,1) para la serie en logaritmos de supergas

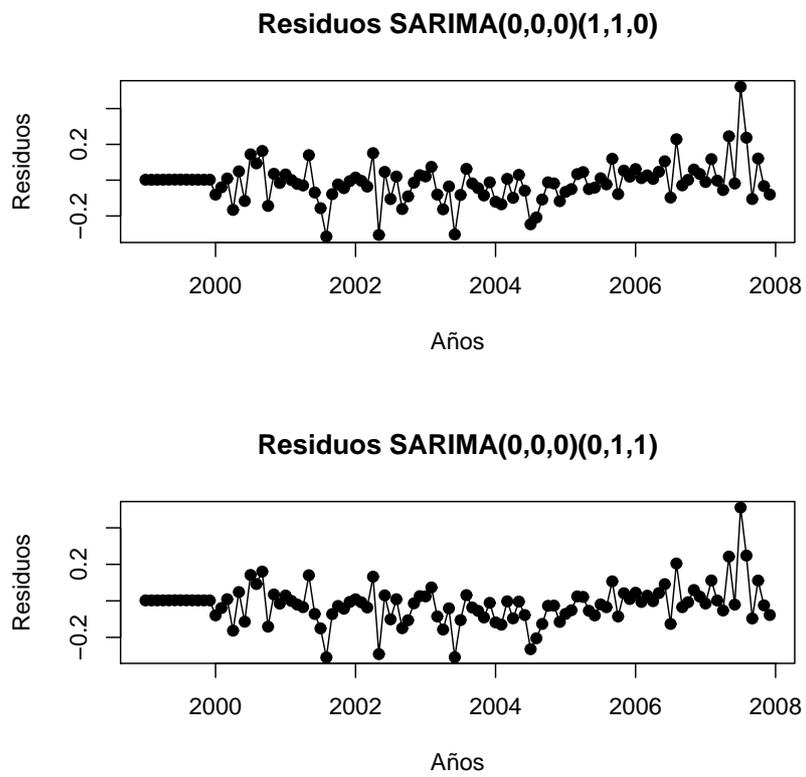


Figura 10.3: Residuos de los modelos SARIMA(0,0,0)(1,1,0) y SARI-MA(0,0,0)(0,1,1) para la serie en logaritmos de supergas

SARIMA(2,0,0)(0,1,1)	coef	t-obs	critico	SARIMA(0,0,2)(0,1,1)	coef	t-obs	critico
ar1	0.15	1.42	1,66	ma1	0.10	1.14	1.66
ar2	3.17	3.16		ma2	0.44	3.52	
sma1	-0.34	2.37		sma1	-0.27	-2.08	
Ago-2001	-0.21	-3.06		Ago-2001	-0.21	-3.29	
May-2002	-0.22	-3.02		May-2002	-0.23	-3.37	
Jun-2003	-0.17	-2.39		Jun-2003	-0.16	-2.57	
Jul-2007	0.45	4.99		Jul-2007	0.45	5.25	
aic = -172.66				aic = -173.76			
Shapiro-Wilk p-value = 0.02				Shapiro-Wilk p-value = 0.03			
Box-Ljung, df=24, p-value = 0.14				Box-Ljung, df=24, p-value = 0.21			
SARIMA(2,0,0)(0,1,2)	coef	t-obs	critico	SARIMA(2,0,0)(2,1,0)	coef	t-obs	critico
ar1	0.16	1.53	1,66	ar1	0.18	1.72	1,66
ar2	0.29	2.75		ar2	0.31	3.02	
sma1	-0.24	-1.58		sar1	-0.32	3.02	
Ago-2001	-0.22	-3.26		sar2	-0.29	-2.23	
May-2002	-0.23	-3.35		Ago-2001	-0.2	-2.87	
Jun-2003	-0.19	-2.69		May-2002	-0.21	-3.06	
Jul-2007	0.46	4.97		Jun-2003	-0.19	-2.74	
				Jul-07	0.47	5.40	
aic = -171.98				aic = -173.48			
Shapiro-Wilk p-value = 0.034				Shapiro-Wilk p-value = 1.45e-05			
Box-Ljung, df=24, p-value = 0.1635				Box-Ljung, df=24, p-value = 0.25			

Tabla 10.3: Modelos preliminares

10.1.1. Resultados de modelos preliminares

1. Referido a los coeficientes de los modelos preliminares

El único modelo con coeficientes significativos resulta ser el modelo: SARIMA(2,0,0)(2,1,0). Para el resto de los modelos presentados en la tabla: 10.3 el *t-obs* no supera en valor absoluto al *t-crítico*, por lo tanto se descartan.

2. Referido a los residuos del modelo SARIMA(2,0,0)(2,1,0):

El test de Ljung-Box no es rechazado, con un p-valor igual a 0.25. A continuación se muestra la gráfica de los residuos de dicho modelo. Allí se observa que en el año 2007 existen puntos que podrían ser tratados como atípicos: Agosto y Mayo de 2007.

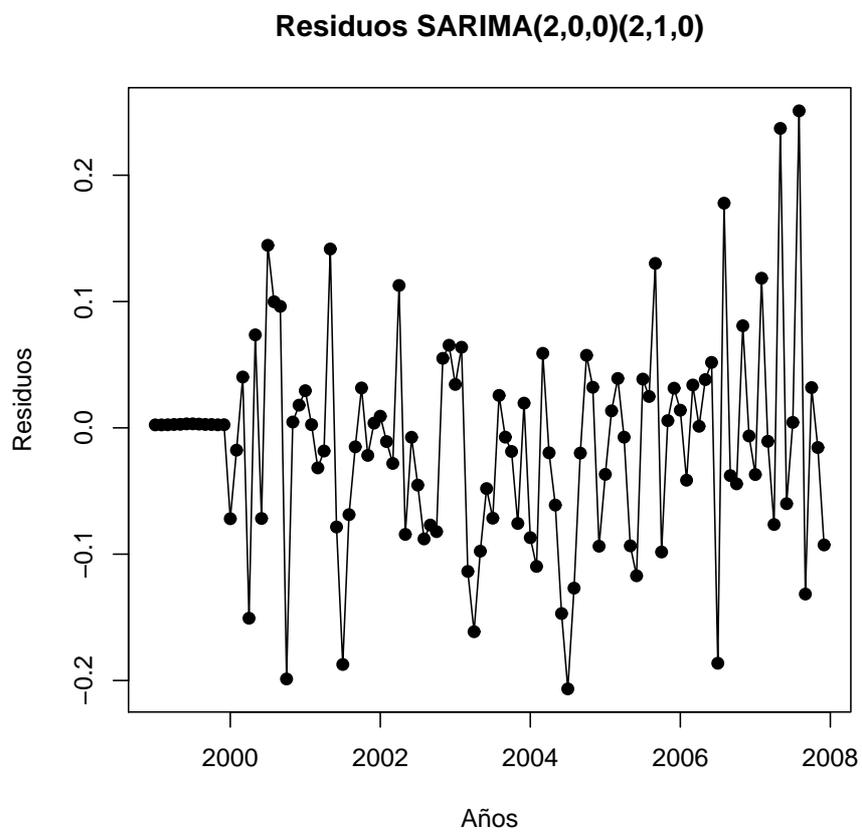


Figura 10.4: Residuos Modelo Preliminar

10.2. Conclusión:

El modelo SARIMA(2,0,0)(2,1,0) con atípicos en: Agosto de 2001, Mayo de 2002, Junio de 2003, Mayo, Julio y Agosto de 2007 es el modelo seleccionado para la modelización del supergas. El desarrollo de este modelo se encuentra en: Resultados referidos a la demanda de Supergas: sec: 6.1.3.

Se aclara que:

1. Dicho modelo mejora el test de Ljung-Box, pasando a tener un mayor p-valor que hace aceptar la prueba.
2. Los autocorrelogramas no dejan retardos fuera de las bandas de confianza.
3. Sus coeficientes son todos significativos.
4. Los residuos fluctúan alrededor de la media.

10.3. Predicción : Julio de 2007 - Junio de 2008

			Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		ERROR
	Sup	pred	Inf	Sup	Inf	Sup	
Jul-07	27,09	16,07	13,78	18,75	13,13	19,68	-41 %
Ago-07	21,4	15,72	13,40	18,44	12,74	19,39	-27 %
Sep-07	12,2	13,07	10,99	15,55	10,41	16,42	7 %
Oct-07	11,51	10,39	8,71	12,40	8,24	13,11	-10 %
Nov-07	10,6	10,63	8,88	12,73	8,40	13,47	0 %
Dic-07	10,37	10,97	9,15	13,15	8,65	13,92	6 %
Ene-08	9,89	9,19	7,66	11,03	7,24	11,68	-7 %
Feb-08	9,51	9,82	8,18	11,79	7,73	12,48	3 %
Mar-08	11,3	10,81	9,01	12,98	8,50	13,75	-4 %
Abr-08	12,04	11,32	9,43	13,59	8,90	14,39	-6 %
May-08	14,82	14,42	12,01	17,31	11,34	18,34	-3 %
Jun-08	19,37	15,92	13,26	19,11	12,51	20,25	-18 %
ECM = 14,12							

Tabla 10.4: Predicción Julio de 2007 - Junio de 2008 - Modelo univariado seleccionado para el consumo de supergas

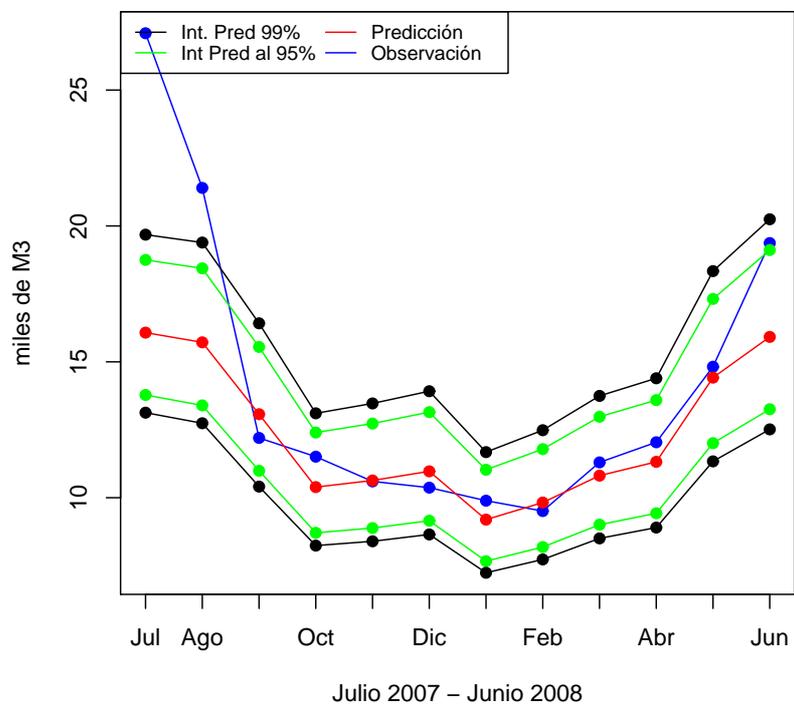


Figura 10.5: Predicción Julio de 2007 - Junio de 2008 - Modelo univariado seleccionado para el consumo de supergas

Capítulo 11

Anexo: Multivariado del Grupo Supergas

```
-----  
Augmented Dickey & Fuller test  
-----  
Deterministic regressors estimates:  
  
None selected.  
-----  
ADF regressor estimate:  
      Estimate Std. Error t value Pr(>|t|)  
adf.reg  -0.003      0.005  -0.598    0.1  
-----  
Selected lags estimates:  
      Estimate Std. Error t value Pr(>|t|)  
Lag.1    -0.776      0.100  -7.785    0.000  
Lag.2    -0.637      0.100  -6.364    0.000  
Lag.3    -0.693      0.093  -7.445    0.000  
Lag.4    -0.877      0.088 -10.007    0.000  
Lag.5    -0.674      0.105  -6.426    0.000  
Lag.6    -0.613      0.118  -5.208    0.000  
Lag.7    -0.754      0.105  -7.154    0.000  
Lag.8    -0.774      0.093  -8.315    0.000  
Lag.9    -0.684      0.099  -6.884    0.000  
Lag.10   -0.796      0.107  -7.438    0.000  
Lag.11   -0.466      0.110  -4.235    0.000  
Lag.18   -0.110      0.102  -1.076    0.285  
  
Number of available observations: 89
```

Figura 11.1: Test de Dickey Fuller a log(Supergas).

```
-----  
Augmented Dickey & Fuller test  
-----  
Deterministic regressors estimates:  
      Estimate Std. Error t value Pr(>|t|)  
Intercept    0.091      0.045   2.046   0.04  
-----  
ADF regressor estimate:  
      Estimate Std. Error t value Pr(>|t|)  
adf.reg    -0.028      0.015  -1.898   0.1  
-----  
Selected lags estimates:  
      Estimate Std. Error t value Pr(>|t|)  
Lag.5     -0.182      0.09  -2.025   0.046  
Lag.6      0.216      0.09   2.398   0.018  
  
Number of available observations: 101
```

Figura 11.2: Test de Dickey Fuller a log(Precio).

```

-----
Augmented Dickey & Fuller test
-----

Deterministic regressors estimates:
      Estimate Std. Error t value Pr(>|t|)
Intercept    7.001     1.975   3.545   0.001
Trend         0.000     0.000   0.973   0.334
-----

ADF regressor estimate:
      Estimate Std. Error t value Pr(>|t|)
adf.reg     -2.29      0.647  -3.537   0.04
-----

Selected lags estimates:
      Estimate Std. Error t value Pr(>|t|)
Lag.1       1.512     0.621   2.435   0.017
Lag.2       1.521     0.602   2.529   0.014
Lag.3       1.427     0.578   2.467   0.016
Lag.4       1.220     0.564   2.165   0.034
Lag.5       1.328     0.527   2.517   0.014
Lag.6       1.161     0.487   2.383   0.020
Lag.7       0.823     0.440   1.872   0.065
Lag.8       0.581     0.380   1.530   0.130
Lag.9       0.418     0.324   1.292   0.200
Lag.10      0.384     0.282   1.363   0.177
Lag.11      0.401     0.252   1.593   0.116
Lag.12      0.563     0.204   2.756   0.007
Lag.13      0.352     0.178   1.980   0.052
Lag.14      0.497     0.154   3.220   0.002
Lag.15      0.296     0.120   2.463   0.016
Lag.18     -0.045     0.093  -0.483   0.630

Number of available observations: 89

```

Figura 11.3: Test de Dickey Fuller a $\log(Tmáx)$.

```

-----
Augmented Dickey & Fuller test
-----

Deterministic regressors estimates:
      Estimate Std. Error t value Pr(>|t|)
Intercept    3.128     0.498   6.284   0.000
Trend        -0.002     0.001  -2.769   0.007
-----

ADF regressor estimate:
      Estimate Std. Error t value Pr(>|t|)
adf.reg     -1.27     0.202  -6.279   0.01
-----

Selected lags estimates:
      Estimate Std. Error t value Pr(>|t|)
Lag.1      0.749     0.160   4.689   0.000
Lag.2      0.551     0.156   3.541   0.001
Lag.3      0.404     0.142   2.849   0.006
Lag.4      0.352     0.136   2.587   0.011
Lag.5      0.413     0.139   2.983   0.004
Lag.6      0.209     0.124   1.686   0.096
Lag.9     -0.293     0.106  -2.758   0.007
Lag.14     0.397     0.110   3.612   0.001

Number of available observations: 93

```

Figura 11.4: Test de Dickey Fuller a $\log(Tmin)$.

11.1. Autocorrelograma simple y parcial del modelo seleccionado

RETARDO	FAC	FACP	Estad-Q.	[valor p]
1	-0,1166	-0,1166	1,429	[0,232]
2	0,0063	-0,0074	1,4333	[0,488]
3	0,0203	0,0204	1,4773	[0,688]
4	-0,0783	-0,0745	2,1402	[0,710]
5	-0,028	-0,0466	2,226	[0,817]
6	-0,0619	-0,0724	2,6499	[0,851]
7	-0,1406	-0,1583	4,8583	[0,677]
8	-0,0244	-0,071	4,9258	[0,765]
9	-0,0498	-0,0736	5,209	[0,816]
10	-0,1086	-0,1467	6,568	[0,766]
11	0,0866	0,0171	7,4415	[0,762]
12	0,2078 **	0,2058 **	12,5321	[0,404]
13	-0,145	-0,1363	15,039	[0,305]
14	-0,024	-0,13	15,1084	[0,371]
15	-0,0473	-0,1032	15,3812	[0,424]
16	-0,0335	-0,0705	15,5196	[0,487]
17	0,076	0,0214	16,2398	[0,507]
18	-0,1188	-0,1098	18,0212	[0,454]
19	0,0185	-0,0113	18,0649	[0,518]
20	0,0467	-0,0085	18,3475	[0,565]
21	-0,0715	-0,0859	19,0163	[0,584]
22	0,0381	-0,02	19,2089	[0,632]
23	0,0461	-0,0577	19,4937	[0,672]
24	0,1579	0,1179	22,8865	[0,527]

Tabla 11.1: Autocorrelación simple y parcial del modelo seleccionado

11.2. Modelo seleccionado con Febrero de 2004 como atípico

Variable	Coefficiente	Desv. típica	Estadístico t	valor p	
d(log(sup))-1	-0,32	0,15	-2,16	0,03	**
d(log(sup))-2	-0,41	0,13	-3,20	0,00	***
d(log(sup))-3	-0,32	0,12	-2,57	0,01	**
d(log(sup))-4	-0,36	0,12	-2,97	0,00	***
d(log(sup))-5	-0,12	0,12	-1,06	0,29	
d(log(precio))-1	-0,28	0,21	-1,33	0,19	
d(log(precio))-2	-0,08	0,21	-0,40	0,69	
d(log(precio))-3	0,01	0,20	0,06	0,95	
d(log(precio))-4	0,18	0,21	0,85	0,40	
d(log(precio))-5	-0,15	0,21	-0,70	0,48	
d(log(Tmax))-1	-1,01	0,16	-6,46	<0,00001	***
d(log(Tmax))-2	-1,16	0,17	-6,97	<0,00001	***
d(log(Tmax))-3	-1,01	0,20	-5,05	<0,00001	***
d(log(Tmax))-4	-0,64	0,23	-2,78	0,01	***
d(log(Tmax))-5	-0,33	0,20	-1,70	0,09	*
jul-2000	0,38	0,11	3,41	0,00	***
ago-2001A	-0,22	0,11	-2,02	0,05	**
jun-2002A	0,27	0,11	2,45	0,02	**
jul-2007A	0,38	0,12	3,34	0,00	***
Feb-2004	-0,01	0,11	-0,08	0,93	
EC1	-0,56	0,09	-5,86	<0,00001	***

Tabla 11.2: Modelo seleccionado con Febrero de 2004 como atípico

11.3. Predicción para el período Julio de 2007 - Junio de 2008

Obs	Sup	pred	Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %		ERROR
			Inf	Sup	Inf	Sup	
Jul-07	27,09	19,32	16,12	23,14	15,23	24,50	-29 %
Ago-07	21,40	17,84	14,70	21,64	13,84	22,99	-17 %
Sep-07	12,20	13,69	11,21	16,72	10,52	17,81	12 %
Oct-07	11,51	12,35	10,08	15,13	9,45	16,13	7 %
Nov-07	10,60	9,81	7,95	12,10	7,44	12,93	-7 %
Dic-07	10,37	8,74	7,01	10,89	6,55	11,67	-16 %
Ene-08	9,89	9,00	7,12	11,38	6,62	12,25	-9 %
Feb-08	9,51	9,24	7,26	11,76	6,73	12,69	-3 %
Mar-08	11,30	10,66	8,34	13,63	7,72	14,72	-6 %
Abr-08	12,04	12,76	9,96	16,35	9,22	17,67	6 %
May-08	14,82	15,03	11,68	19,33	10,79	20,93	1 %
Jun-08	19,37	17,06	13,12	22,19	12,08	24,10	-12 %
ECM-univariado = 7,21							

Tabla 11.3: Predicción para el período Julio de 2007 a Junio de 2008

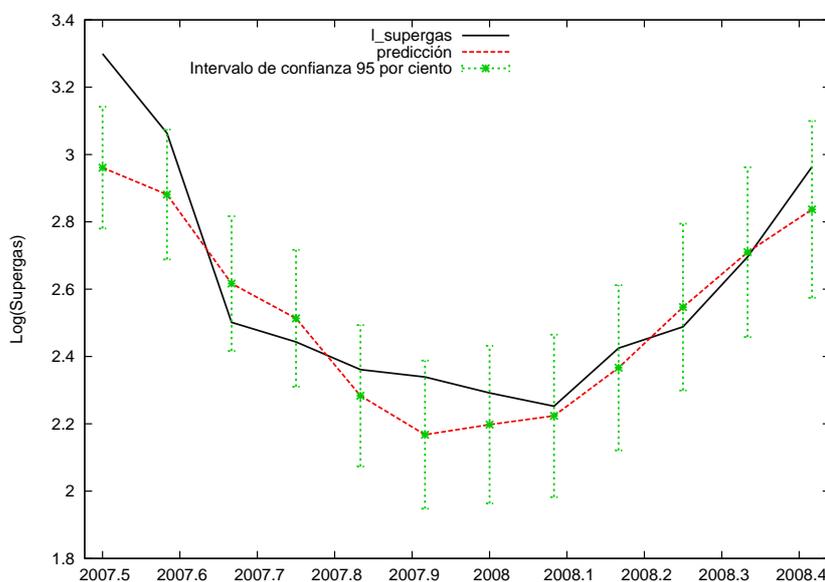


Figura 11.5: Predicción para el período Julio de 2007 a Junio de 2008

11.4. Modelos rechazados

A continuación se presenta una breve reseña de los modelos que se analizaron como posibles modelos de demanda de supergas.

11.4.1. Modelo 1 - Grupo Supergas

- 5 retardos
- variables: supergas, precio, Tmax y Tmin.
- Especificación: sin constante.
- Atípicos: Julio de 2000, Junio de 2002 y Julio de 2007

Test de Johansen

Contraste de Johansen:					
Número de ecuaciones = 4					
Orden del retardo = 6					
Periodo de estimación: 1999:07 - 2008:06 (T = 108)					
Caso 1: Sin constante					
Rango	Valor propio	Contraste de la traza	valor p	Contraste L _{máx}	valor p
0	0,64	123,81	[0,000]	109,46	[0,00]
1	0,09	14,35	[0,513]	10,01	[0,49]
2	0,04	4,35	[0,655]	4,29	[0,59]
3	0,0005	0,051	[0,878]	0,05	[0,87]

Significación de los coeficientes

Se realizan pruebas de hipótesis sobre los parámetros de la relación a largo plazo (coeficientes del vector cointegrante: *Betas*) de forma de evaluar si son significativamente distintos de cero o no.

A pesar de que el modelo incluye como variables significativas la temperatura máxima y mínima a la vez, no es razonable que se ingresen en el modelo dos variables que están altamente correlacionadas.

Beta: vector cointegrante					
supergas	precio	Tmax	Tmin		
1	0,154	-0,86	0,18		
0	(-0,03)	(-0,18)	-0,313		
Alfa: Velocidad de ajuste					
supergas	precio	Tmax	Tmin		
-0,61	0,08	0,76	0,47		
Variable	Coefficiente	Desv. típica	Estadístico t	valor p	
d(supergas)-1	0,02	0,12	0,15	0,89	
d(supergas)-2	-0,15	0,10	-1,55	0,13	
d(supergas)-3	-0,22	0,08	-2,72	0,01	***
d(supergas)-4	-0,09	0,09	-0,94	0,35	
d(supergas)-5	0,07	0,10	0,72	0,47	
d(Precio)-1	-0,07	0,13	-0,55	0,58	
d(Precio)-2	-0,07	0,13	-0,49	0,62	
d(Precio)-3	0,01	0,13	0,11	0,91	
d(Precio)-4	0,12	0,13	0,90	0,37	
d(Precio)-5	0,04	0,13	0,30	0,76	
d(Tmax)-1	-0,39	0,15	-2,72	0,01	***
d(Tmax)-2	-0,83	0,14	-5,79	<0,00001	***
d(Tmax)-3	-0,71	0,19	-3,68	0,00	***
d(Tmax)-4	-0,62	0,18	-3,48	0,00	***
d(Tmax)-5	-0,24	0,16	-1,46	0,15	
d(Tmin)-1	0,13	0,15	0,82	0,42	
d(Tmin)-2	0,34	0,19	1,81	0,07	*
d(Tmin)-3	-0,04	0,21	-0,20	0,84	
d(Tmin)-4	0,43	0,19	2,25	0,03	**
d(Tmin)-5	0,23	0,17	1,38	0,17	
jul-2007	9,28	1,67	5,55	<0,00001	***
jul-2000	7,63	1,69	4,50	0,00	***
jun-2002	4,96	1,60	3,11	0,00	***
EC1	-0,61	0,10	-6,25	<0,00001	***

Tabla 11.4: Significación de las variables incluidas en el modelo 1

Las pruebas llevan a concluir que el coeficiente de la temperatura mínima no es estadísticamente distinto de cero, por lo que la variable se quita del análisis. *Esto no sucede únicamente para este modelo sino que se repite para todos los modelos probados.*

```

Restricción: Temperatura mínima
b[4] = 0
Test of restrictions on cointegrating relations
valor propio 1 = 0,537518
Log-verosimilitud no restringida (lu) = -621,83349
Log-verosimilitud restringida (lr) = -621,97648
2 * (lu - lr) = 0,285968
P(Chi-cuadrado(1) > 0,285968) = 0,592816

```

```

Restricción:
b[2] = 0 - Precio
Test of restrictions on cointegrating relations
valor propio 1 = 0,447737
Log-verosimilitud no restringida (lu) = -621,83349
Log-verosimilitud restringida (lr) = -631,02471
2 * (lu - lr) = 18,3824
P(Chi-cuadrado(1) > 18,3824) = 1,80716e-005

```

```

Restricción:
b[3] = 0 - Temperatura máxima
Test of restrictions on cointegrating relations
valor propio 1 = 0,482579
Log-verosimilitud no restringida (lu) = -621,83349
Log-verosimilitud restringida (lr) = -627,70121
2 * (lu - lr) = 11,7354
P(Chi-cuadrado(1) > 11,7354) = 0,000613216

```

11.4.2. Modelo 2 - Grupo Supergas

El modelo 2 tiene igual especificación que el modelo anterior, sólo que se quita del análisis la temperatura mínima.

- Retardos incluidos: K=5
- Variables:supergas, precio y temperatura máxima
- Atípicos: Julio de 2000, Agosto de 2001, Junio de 2002 Mayo y Julio de 2007.
- Especificación sin constante.

Variable	Coefficiente	Desv. típica	Estadístico t	valor p	
d(Supergas)-1	-0,02	0,11	-0,22	0,83	
d(Supergas)-2	-0,14	0,09	-1,51	0,13	
d(Supergas)-3	-0,21	0,08	-2,68	0,01	***
d(Supergas)-4	-0,17	0,09	-1,86	0,07	*
d(Supergas)-5	-0,02	0,10	-0,18	0,86	
d(Precio)-1	-0,09	0,13	-0,71	0,48	
d(Precio)-2	0,02	0,13	0,13	0,90	
d(Precio)-3	0,01	0,12	0,06	0,95	
d(Precio)-4	0,13	0,13	0,97	0,34	
d(Precio)-5	-0,05	0,13	-0,39	0,70	
d(Tmax)-1	-0,46	0,10	-4,46	0,00	***
d(Tmax)-2	-0,56	0,09	-6,01	<0,00001	***
d(Tmax)-3	-0,71	0,11	-6,53	<0,00001	***
d(Tmax)-4	-0,40	0,13	-3,08	0,00	***
d(Tmax)-5	-0,20	0,12	-1,62	0,11	
jul-2007	9,55	1,62	5,90	<0,00001	***
jul-2000	7,43	1,58	4,70	0,00	***
jun-2002	4,60	1,55	2,96	0,00	***
ag-2001	-4,38	1,56	-2,80	0,01	***
may-2007	4,18	1,60	2,61	0,01	**
EC1	-0,61	0,09	-6,78	<0,00001	***

Tabla 11.5: Coeficientes y significación de las variables incluidas en el modelo

2

Obs	supergas	predicción	Error
Ene-08	9,89	9,56	-3 %
Feb-08	9,51	9,79	3 %
Mar-08	11,3	10,34	-8 %
Abr-08	12,04	11,96	-1 %
May-08	14,82	14,71	-1 %
Jun-08	19,37	17,69	-9 %
Jul-08	16,46	19,08	16 %
Ago-08	16,02	18,23	14 %
Sep-08		15,93	
Oct-08		13,30	
Nov-08		11,32	
Dic-08		10,24	
ECM-multivariado = 1,96			

Tabla 11.6: Predicción del año 2008 - modelo 2

11.4.3. Modelo 3 - Grupo Supergas

- Retardos incluidos: K=5
- Variables:supergas y temperatura máxima
- Atípicos: Julio de 2000, Agosto de 2001, Junio de 2002 y Julio de 2007.
- Especificación sin constante.

Dado que los coeficientes de la variable **precio** no resultan significativamente distintos de cero en la salida del modelo seleccionado se especifica un nuevo modelo sin incorporar dicha variable.

Si bien este modelo tienen menor error cuadrático medio (ECM) que el modelo elegido, el error porcentual en cada observación es mayor (considerando un período de 3 meses).

Variable	Coefficiente	Desv. típica	Estadístico t	valor p	
d(log(sup))-1	-0,39	0,14	-2,75	0,01	***
d(log(sup))-2	-0,46	0,13	-3,51	0,00	***
d(log(sup))-3	-0,33	0,12	-2,66	0,01	***
d(log(sup))-4	-0,38	0,12	-3,11	0,00	***
d(log(sup))-5	-0,12	0,12	-1,03	0,31	
d(log(Tmax))-1	-0,94	0,16	-6,03	<0,00001	***
d(log(Tmax))-2	-1,11	0,17	-6,68	<0,00001	***
d(log(Tmax))-3	-0,87	0,19	-4,54	0,00	***
d(log(Tmax))-4	-0,45	0,20	-2,22	0,03	**
d(log(Tmax))-5	-0,20	0,18	-1,09	0,28	
Jul-2000	0,41	0,11	3,68	0,00	***
Ago-2001	-0,19	0,11	-1,63	0,11	
Jun-2002	0,30	0,11	2,69	0,01	***
Jul-2007	0,37	0,12	3,14	0,00	***
EC1	-0,47	0,08	-5,54	<0,00001	***

Tabla 11.7: Coeficientes y significación de las variables del modelo 3

	obs	pred	Error
Ene-08	9,89	10,86	10 %
Feb-08	9,51	10,14	7 %
Mar-08	11,30	10,72	-5 %
Abr-08	12,04	12,43	3 %
May-08	14,82	14,59	-2 %
Jun-08	19,37	17,11	-12 %
Jul-08	16,46	17,56	7 %
Ago-08	16,02	16,24	1 %
Sep-08	-	14,54	-
Oct-08	-	12,91	-
Nov-08	-	11,86	-
Dic-08	-	11,17	-
ECM-multivariado = 1,03			

Tabla 11.8: Predicción año 2008 - modelo 3

11.4.4. Modelo 4 - Grupo Supergas

Se prueba incluyendo más retardos en la especificación.

- Retardos incluidos: K=12
- Variables:supergas, precio y temperatura máxima
- Atípicos: Julio de 2000, Agosto de 2001, Junio de 2002 y Julio de 2007
- Especificación: sin constante

Variable	Coficiente	Desv. típica	Estadístico t	valor p	
d(supergas)-1	-1,21	0,27	-4,44	0,00	***
d(supergas)-2	-1,29	0,29	-4,42	0,00	***
d(supergas)-3	-1,09	0,28	-3,93	0,00	***
d(supergas)-4	-1,35	0,29	-4,64	0,00	***
d(supergas)-5	-1,14	0,29	-3,93	0,00	***
d(supergas)-6	-0,92	0,28	-3,34	0,00	***
d(supergas)-7	-1,11	0,26	-4,25	0,00	***
d(supergas)-8	-1,03	0,23	-4,46	0,00	***
d(supergas)-9	-0,77	0,20	-3,90	0,00	***
d(supergas)-10	-1,07	0,18	-5,81	<0,00001	***
d(supergas)-11	-0,70	0,16	-4,39	0,00	***
d(supergas)-12	-0,34	0,15	-2,34	0,02	**
d(precio)-1	-0,01	0,14	-0,09	0,93	
d(precio)-2	-0,22	0,14	-1,57	0,12	
d(precio)-3	0,05	0,14	0,36	0,72	
d(precio)-4	-0,05	0,14	-0,34	0,73	
d(precio)-5	-0,34	0,15	-2,34	0,02	**
d(precio)-6	-0,19	0,16	-1,21	0,23	
d(precio)-7	-0,33	0,15	-2,25	0,03	**
d(precio)-8	-0,09	0,14	-0,69	0,49	
d(precio)-9	0,08	0,14	0,54	0,59	
d(precio)-10	-0,33	0,14	-2,34	0,02	**
d(precio)-11	0,04	0,15	0,26	0,80	
d(precio)-12	-0,23	0,14	-1,58	0,12	
dd(Tmax)-1	0,08	0,23	0,34	0,74	
d(Tmax)-2	0,01	0,24	0,03	0,98	
d(Tmax)-3	0,07	0,24	0,29	0,77	
d(Tmax)-4	0,08	0,24	0,32	0,75	
d(Tmax)-5	-0,03	0,21	-0,12	0,90	
d(Tmax)-6	0,15	0,18	0,81	0,42	
d(Tmax)-7	0,06	0,18	0,32	0,75	
d(Tmax)-8	0,00	0,19	0,00	1,00	
d(Tmax)-9	0,10	0,19	0,51	0,61	
d(Tmax)-0	-0,20	0,19	-1,09	0,28	
d(Tmax)-11	0,06	0,17	0,35	0,73	
d(Tmax)-12	-0,08	0,15	-0,57	0,57	
jul-2000	5,92	1,53	3,88	0,00	***
ag-2001	-8,40	1,62	-5,19	<0,00001	***
jun-2002	1,30	1,51	0,86	0,39	
jul-2007	10,04	1,57	6,40	<0,00001	***
EC1	0,30	0,24	1,27	0,21	

Tabla 11.9: Coeficientes y significación de las variables del modelo 4

Obs	supergas	predicción	Error
Ene-08	9,89	14,09	42 %
Feb-08	9,51	9,97	5 %
Mar-08	11,3	9,12	-19 %
Abr-08	12,04	15,72	31 %
May-08	14,82	14,13	-5 %
Jun-08	19,37	20,03	3 %
Jul-08	16,46	24,08	46 %
Ago-08	16,02	19,69	23 %
Sep-08		17,56	
Oct-08		12,58	
Nov-08		9,75	
Dic-08		12,62	
ECM-multivariado = 13,57			

Tabla 11.10: Predicción para el año 2008 - modelo 4

11.4.5. Modelo 5 - Grupo Supergas

A pesar de que la constante resulta significativa las predicciones para el 2008 de este modelo incurren en un error mayor al del modelo propuesto.

- Retardos incluidos: $K=6$
- Variables: supergas, precio y temperatura máxima
- Atípicos: Julio de 2000, Agosto de 2001, Junio de 2002 Mayo y Julio de 2007
- Especificación: con constante en la relación de cointegración.

Variable	Coefficiente	Desv. típica	Estadístico t	valor p	
const	-18,26	3,36	-5,43	<0,00001	***
d(supergas)-1	-0,54	0,09	-5,71	<0,00001	***
d(supergas)-2	-0,56	0,09	-6,02	<0,00001	***
d(supergas)-3	-0,50	0,09	-5,38	<0,00001	***
d(supergas)-4	-0,41	0,10	-4,18	0,00	***
d(supergas)-5	-0,13	0,10	-1,30	0,20	
d(Precio)-1	-0,13	0,14	-0,91	0,36	
d(Precio)-2	-0,01	0,14	-0,09	0,93	
d(Precio)-3	-0,06	0,14	-0,42	0,67	
d(Precio)-4	0,04	0,14	0,26	0,80	
d(Precio)-5	-0,10	0,14	-0,72	0,47	
d(Tmax)-1	-0,79	0,14	-5,84	<0,00001	***
d(Tmax)-2	-0,84	0,11	-7,41	<0,00001	***
d(Tmax)-3	-0,86	0,14	-5,91	<0,00001	***
d(Tmax)-4	-0,51	0,17	-3,09	0,00	***
d(Tmax)-5	-0,21	0,14	-1,47	0,15	
jul-2000	7,85	1,75	4,50	0,00	***
ag-2001	-4,80	1,73	-2,77	0,01	***
jun-2002	5,08	1,71	2,97	0,00	***
jul-2007	9,36	1,79	5,23	<0,00001	***
EC1	-0,01	0,00	-5,38	<0,00001	***

Tabla 11.11: Coeficientes y significación de las variables incluidas en el modelo

5

Obs	supergas	predicción	Error
Ene-08	9,89	10,88	10 %
Feb-08	9,51	10,77	13 %
Mar-08	11,30	10,99	-3 %
Abr-08	12,04	12,84	7 %
May-08	14,82	15,74	6 %
Jun-08	19,37	18,83	-3 %
Jul-08	16,46	19,73	20 %
Ago-08	16,02	18,20	14 %
Sep-08		15,28	
Oct-08		12,32	
Nov-08		10,40	
Dic-08		9,41	
ECM-multivariado = 2,49			

Tabla 11.12: Predicción para el año 2008 - modelo 5

Los modelos a continuación surgen de modificaciones realizadas al modelo seleccionado.

11.4.6. Modelo 6 - Grupo Supergas

A pesar de que la constante en la relación de cointegración resulta significativa el modelo no supera las predicciones del seleccionado.

- Retardos incluidos: K=5
- Variables en logaritmos: supergas, precio y temperatura máxima
- Atípicos: Julio de 2000, Agosto de 2001, Junio de 2002 y Julio de 2007
- Especificación: con constante en la relación de cointegración.

Variable	Coefficiente	Desv. típica	Estadístico t	valor p	
const	-4,49	0,89	-5,05	<0,00001	***
d(log(sup))-1	-0,96	0,13	-7,52	<0,00001	***
d(log(sup))-2	-0,90	0,14	-6,60	<0,00001	***
d(log(sup))-3	-0,66	0,15	-4,49	0,00	***
d(log(sup))-4	-0,61	0,14	-4,34	0,00	***
d(log(sup))-5	-0,25	0,12	-2,05	0,04	**
d(log(precio))-1	-0,36	0,22	-1,67	0,10	*
d(log(precio))-2	-0,12	0,22	-0,55	0,58	
d(log(precio))-3	-0,02	0,22	-0,10	0,92	
d(log(precio))-4	0,12	0,22	0,54	0,59	
d(log(precio))-5	-0,21	0,22	-0,95	0,35	
d(log(Tmax))-1	-1,59	0,21	-7,60	<0,00001	***
d(log(Tmax))-2	-1,59	0,20	-7,86	<0,00001	***
d(log(Tmax))-3	-1,25	0,25	-5,07	<0,00001	***
d(log(Tmax))-4	-0,73	0,27	-2,77	0,01	***
d(log(Tmax))-5	-0,37	0,21	-1,72	0,09	*
jul-2000	0,42	0,12	3,61	0,00	***
ag-2001	-0,26	0,12	-2,25	0,03	**
jun-2002	0,31	0,11	2,70	0,01	***
jul-2007	0,38	0,12	3,15	0,00	***
EC1	0,22	0,04	5,04	<0,00001	***

Tabla 11.13: Coeficientes y significación de las variables incluidas en el modelo

Obs	supergas	predicción	Error
Ene-08	9,89	10,64	8 %
Feb-08	9,51	10,08	6 %
Mar-08	11,3	10,53	-7 %
Abr-08	12,04	12,75	6 %
May-08	14,82	15,60	5 %
Jun-08	19,37	18,48	-5 %
Jul-08	16,46	19,16	16 %
Ago-08	16,02	17,34	8 %
Sep-08		14,61	
Oct-08		12,39	
Nov-08		10,89	
Dic-08		9,91	
ECM-multivariado = 1,55			

Tabla 11.14: Predicción del año 2008 - modelo 6

11.4.7. Modelo 7 - Grupo Supergas

En este modelo se ve que las variables que se intentan relacionar pierden protagonismo, pasando a captar la dinámica de la serie algunas de las variables estacionales. En la especificación del modelo surge una nueva observación como atípica: *Mayo de 2002*, el cual se ingresa en la modelización. Se adjuntan los autocorrelogramas simple y parcial de este modelo. Allí se observa que el retardo 12 deja de ser significativo, pero se resaltan los retardos 1, 4 y 23 en ambos autocorrelogramas. A pesar de esto el test global de Ljung Box lleva a no rechazar la hipótesis de no autocorrelación en los primeros 48 retardos. Dicho modelo no presenta mejoras significativas en las predicciones y el ECM resultante es mayor al del modelo seleccionado.

- Retardos incluidos: $K=6$
- Variables en logaritmos: supergas, precio y temperatura máxima
- Atípicos: Julio de 2000, Agosto de 2001, Junio de 2002, Mayo y Julio de 2007
- Especificación: sin constante
- Se incluyen variables indicatrices estacionales

Variable	Coefficiente	Desv. típica	Estadístico t	valor p	
d(log(sup))-1	-0,11	0,15	-0,74	0,46	
d(log(sup))-2	0,04	0,16	0,23	0,82	
d(log(sup))-3	0,36	0,16	2,28	0,03	**
d(log(sup))-4	0,06	0,15	0,43	0,67	
d(log(sup))-5	-0,17	0,11	-1,54	0,13	
d(log(precio))-1	0,00	0,16	0,01	0,99	
d(log(precio))-2	0,13	0,16	0,82	0,42	
d(log(precio))-3	-0,47	0,15	-3,01	0,00	***
d(log(precio))-4	0,20	0,16	1,22	0,23	
d(log(precio))-5	0,00	0,16	-0,02	0,98	
d(log(Tmax))-1	-0,42	0,19	-2,22	0,03	**
d(log(Tmax))-2	-0,09	0,21	-0,45	0,65	
d(log(Tmax))-3	0,13	0,21	0,61	0,54	
d(log(Tmax))-4	-0,09	0,19	-0,48	0,63	
d(log(Tmax))-5	-0,23	0,16	-1,48	0,14	
jul-2000	0,28	0,08	3,66	0,00	***
ag-2001	-0,36	0,08	-4,75	0,00	***
jun-2002	0,20	0,08	2,55	0,01	**
jul-2007	0,45	0,08	5,40	<0,00001	***
may-2002	-0,25	0,07	-3,29	0,00	***
S1	-0,27	0,06	-4,37	0,00	***
S2	-0,31	0,08	-3,84	0,00	***
S3	-0,26	0,10	-2,64	0,01	**
S4	-0,10	0,11	-0,92	0,36	
S5	0,24	0,11	2,13	0,04	**
S6	0,15	0,12	1,25	0,22	
S7	0,32	0,13	2,51	0,01	**
S8	0,18	0,12	1,44	0,15	
S9	0,00	0,11	-0,02	0,99	
S10	-0,16	0,09	-1,81	0,08	*
S11	-0,04	0,06	-0,77	0,44	
EC1	-0,57	0,11	-4,96	<0,00001	***

Tabla 11.15: Coeficientes y significación de las variables incluidas en el modelo

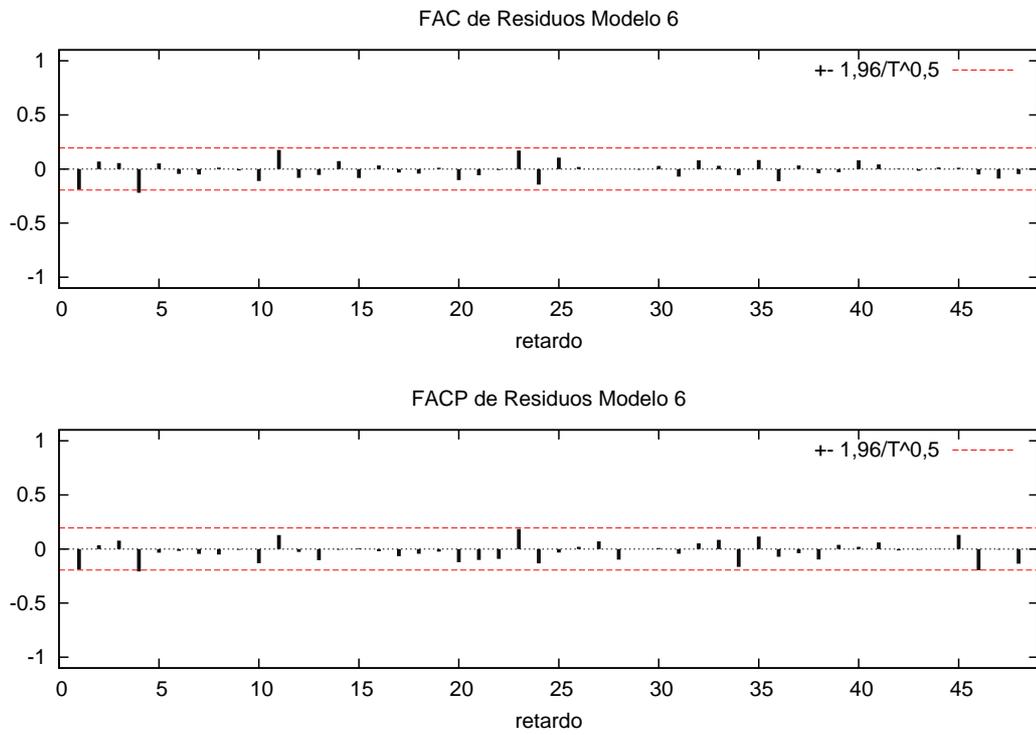


Figura 11.6: Autocorrelograma simple y parcial del modelo 7

Ecuación 1: Residuos de la ecuación de Supergas - Modelo 7
 Ljung-Box $Q' = 39,0505$ con valor $p = P(\text{Chi-cuadrado}(48) > 39,0505) = 0,818$

RETARDO	FAC	FACP	Estad-Q	[valor p]
1	-0,1894 *	-0,1894 *	3,7675	[0,052]
2	0,0694	0,0348	4,2785	[0,118]
3	0,055	0,0771	4,6025	[0,203]
4	-0,2195 **	-0,2070 **	9,8194	[0,044]
5	0,0516	-0,0336	10,1102	[0,072]
6	-0,0448	-0,0178	10,3319	[0,111]
7	-0,0502	-0,045	10,6138	[0,156]
8	0,0131	-0,0503	10,6332	[0,223]
9	-0,0128	-0,0073	10,6518	[0,300]
10	-0,111	-0,1323	12,0716	[0,280]
11	0,1762 *	0,1282	15,6887	[0,153]
12	-0,0819	-0,0265	16,48	[0,170]
13	-0,0547	-0,1045	16,8365	[0,207]
14	0,0734	-0,0066	17,4854	[0,231]
15	-0,0831	0,0058	18,3269	[0,246]
16	0,0325	-0,0208	18,4569	[0,298]
17	-0,0333	-0,0657	18,5949	[0,352]
18	-0,0418	-0,0439	18,8161	[0,403]
19	0,0115	-0,0246	18,833	[0,468]
20	-0,1024	-0,1224	20,1907	[0,446]
21	-0,0572	-0,1013	20,6184	[0,482]
22	-0,009	-0,0911	20,6293	[0,544]
23	0,1707 *	0,1849 *	24,5435	[0,374]
24	-0,1439	-0,1329	27,3606	[0,288]

Tabla 11.16: Autocorrelaciones simple y parcial- modelo 7

Obs	supergas	predicción	Error
Ene-08	9,89	9,18	-7 %
Feb-08	9,51	9,13	-4 %
Mar-08	11,3	10,00	-11 %
Abr-08	12,04	11,38	-6 %
May-08	14,82	15,57	5 %
Jun-08	19,37	16,48	-15 %
Jul-08	16,46	18,01	9 %
Ago-08	16,02	16,98	6 %
Sep-08		13,26	
Oct-08		10,64	
Nov-08		10,46	
Dic-08		10,84	
ECM-multivariado = 1,88			

Tabla 11.17: Predicción para el año 2008 - modelo 7

11.4.8. Modelo 8 - Grupo Supergas

- Retardos incluidos: $K=6$
- Variables en logaritmos: supergas, precio y temperatura máxima
- Atípicos: Julio de 2000, Agosto de 2001, Junio de 2002 y Julio de 2007
- Especificación: con constante en la relación de cointegración.
- Se incluyen variables indicatrices estacionales

Como se observa en las tablas a continuación este modelo es el que logra menor error en las predicciones para el año 2008. Por tal motivo se lo compara de forma más minuciosa con el modelo escogido. La incorporación de la observación de Enero de 2008, hace que las previsiones disminuyan su error, pero no tanto como el modelo finalmente seleccionado. Por tal motivo se descarta.

Variable	Coefficiente	Desv. típica	Estadístico t	valor p	
const	-3,15	0,84	-3,74	0,00	***
d(log(sup))-1	-0,62	0,13	-4,92	<0,00001	***
d(log(sup))-2	-0,35	0,15	-2,31	0,02	**
d(log(sup))-3	0,08	0,16	0,54	0,59	
d(log(sup))-4	-0,17	0,15	-1,16	0,25	
d(log(sup))-5	-0,31	0,12	-2,63	0,01	**
d(log(precio))-1	-0,08	0,18	-0,49	0,63	
d(log(precio))-2	0,06	0,17	0,35	0,73	
d(log(precio))-3	-0,47	0,17	-2,71	0,01	***
d(log(precio))-4	0,19	0,18	1,03	0,31	
d(log(precio))-5	-0,05	0,18	-0,30	0,76	
d(log(Tmax))-1	-0,89	0,29	-3,06	0,00	***
d(log(Tmax))-2	-0,47	0,27	-1,74	0,09	*
d(log(Tmax))-3	-0,13	0,26	-0,51	0,61	
d(log(Tmax))-4	-0,27	0,23	-1,20	0,24	
d(log(Tmax))-5	-0,34	0,18	-1,89	0,06	*
jul-2000	0,30	0,09	3,56	0,00	***
ag-2001	-0,39	0,09	-4,43	0,00	***
jun-2002	0,23	0,09	2,64	0,01	**
jul-2007	0,41	0,09	4,46	0,00	***
S1	-0,26	0,07	-3,66	0,00	***
S2	-0,30	0,09	-3,16	0,00	***
S3	-0,24	0,12	-2,07	0,04	**
S4	-0,07	0,13	-0,58	0,56	
S5	0,20	0,13	1,52	0,13	
S6	0,13	0,15	0,87	0,39	
S7	0,31	0,15	2,09	0,04	**
S8	0,16	0,14	1,10	0,28	
S9	-0,03	0,12	-0,26	0,80	
S10	-0,17	0,10	-1,80	0,08	*
S11	-0,05	0,06	-0,83	0,41	
EC1	-0,02	0,01	-3,73	0,00	***

Tabla 11.18: Coeficientes y significación de las variables incluidas en el modelo

Obs	supergas	predicción	Error
Ene-08	9,89	9,65	-2 %
Feb-08	9,51	9,41	-1 %
Mar-08	11,3	10,13	-10 %
Abr-08	12,04	11,75	-2 %
May-08	14,82	15,20	3 %
Jun-08	19,37	16,17	-17 %
Jul-08	16,46	18,26	11 %
Ago-08	16,02	16,88	5 %
Sep-08		13,08	
Oct-08		10,66	
Nov-08		10,39	
Dic-08		10,63	
ECM-multivariado = 1,99			

Tabla 11.19: Predicción para el año 2008 - modelo 8

Obs	Supergas	predicción	Error
Feb-08	9,51	9,64	1 %
Mar-08	11,3	10,45	-7 %
Abr-08	12,04	12,70	6 %
May-08	14,82	15,52	5 %
Jun-08	19,37	16,60	-14 %
Jul-08	16,46	19,54	19 %
Ago-08	16,02	17,41	9 %
Sep-08		13,62	
Oct-08		11,28	
Nov-08		10,70	
Dic-08		11,14	
ECM-multivariado = 1,88			

Tabla 11.20: Predicción del período Febrero a Diciembre de 2008 - modelo 8

Capítulo 12

Anexo: Multivariado del Grupo Gas Oil

```

-----
Augmented Dickey & Fuller test
-----

----
Deterministic regressors estimates:

      Estimate Std. Error t value Pr(>|t|)
Intercept  19.276    18.592   1.037   0.303
Trend       0.060     0.024   2.481   0.015

----
ADF regressor estimate:

      Estimate Std. Error t value Pr(>|t|)
adf.reg  -0.344     0.295  -1.167   0.1

----
Selected lags estimates:

      Estimate Std. Error t value Pr(>|t|)
Lag.1   -0.549     0.289  -1.900   0.062
Lag.2   -0.542     0.273  -1.983   0.051
Lag.3   -0.619     0.261  -2.370   0.021
Lag.4   -0.669     0.244  -2.742   0.008
Lag.5   -0.497     0.224  -2.219   0.030
Lag.6   -0.567     0.199  -2.856   0.006
Lag.7   -0.611     0.183  -3.331   0.001
Lag.8   -0.449     0.168  -2.671   0.009
Lag.9   -0.466     0.144  -3.244   0.002
Lag.10  -0.655     0.124  -5.290   0.000
Lag.11  -0.527     0.100  -5.254   0.000
Lag.15   0.151     0.089   1.682   0.097
Lag.20  -0.025     0.089  -0.284   0.777

Number of available observations: 87

```

Figura 12.1: Test de Dickey Fuller a serie de Gas Oil.

```

-----
Augmented Dickey & Fuller test
-----

Deterministic regressors estimates:
      Estimate Std. Error t value Pr(>|t|)
Intercept   0.990     0.395   2.506   0.014
Trend       0.037     0.011   3.402   0.001
-----

ADF regressor estimate:
      Estimate Std. Error t value Pr(>|t|)
adf.reg   -0.177     0.054  -3.286   0.077
-----

Selected lags estimates:
      Estimate Std. Error t value Pr(>|t|)
Lag.7     0.177     0.116   1.532   0.129
Lag.9     0.177     0.116   1.524   0.131
Lag.11    0.115     0.117   0.979   0.330
Lag.12    0.297     0.115   2.581   0.012
Lag.19    0.004     0.123   0.033   0.974

Number of available observations: 88
    
```

Figura 12.2: Test de Dickey Fuller a serie de Precio de Gas Oil.

```

-----
Augmented Dickey & Fuller test
-----

Deterministic regressors estimates:
      Estimate Std. Error t value Pr(>|t|)
Intercept  108.767    21.653   5.023   0.000
Trend      -0.281     0.245  -1.149   0.254
-----

ADF regressor estimate:
      Estimate Std. Error t value Pr(>|t|)
adf.reg    -0.82     0.103  -7.975   0.01
-----

Selected lags estimates:
      Estimate Std. Error t value Pr(>|t|)
Lag.7      0.108     0.078   1.386   0.170
Lag.9      0.094     0.083   1.124   0.264
Lag.14     0.290     0.088   3.303   0.001
Lag.20    -0.020     0.090  -0.225   0.823

Number of available observations: 87
    
```

Figura 12.3: Test de Dickey Fuller a serie de Lluvias.

```

-----
Augmented Dickey & Fuller test
-----

Deterministic regressors estimates:
      Estimate Std. Error t value Pr(>|t|)
Intercept   9.058     3.757   2.411   0.018
Trend       0.087     0.032   2.682   0.009
-----

ADF regressor estimate:
      Estimate Std. Error t value Pr(>|t|)
adf.reg    -0.1      0.04  -2.494   0.1
-----

Selected lags estimates:
      Estimate Std. Error t value Pr(>|t|)
Lag.1     1.262     0.132   9.544   0.000
Lag.2    -0.790     0.183  -4.318   0.000
Lag.3    -0.501     0.157  -3.193   0.002
Lag.4     1.151     0.153   7.532   0.000
Lag.5    -0.722     0.133  -5.445   0.000
Lag.8     0.038     0.121   0.317   0.752
Lag.9    -0.172     0.093  -1.843   0.069
Lag.12    0.267     0.113   2.367   0.020
Lag.20    0.040     0.087   0.455   0.650

Number of available observations: 87

```

Figura 12.4: Test de Dickey Fuller a serie de PBlagro.

```

-----
Augmented Dickey & Fuller test
-----
Deterministic regressors estimates:
      Estimate Std. Error t value Pr(>|t|)
Intercept    0.597      1.061   0.562   0.576
Trend         0.023      0.008   2.918   0.005
-----
ADF regressor estimate:
      Estimate Std. Error t value Pr(>|t|)
adf.reg     -0.016      0.011  -1.48   0.1
-----
Selected lags estimates:
      Estimate Std. Error t value Pr(>|t|)
Lag.1       1.391      0.103  13.448  0.000
Lag.2      -0.602      0.172  -3.508  0.001
Lag.3      -0.511      0.178  -2.873  0.005
Lag.4       0.496      0.105   4.742  0.000
Lag.7      -0.189      0.052  -3.621  0.001
Lag.10      0.109      0.052   2.071  0.042
Lag.13     -0.071      0.074  -0.956  0.342
Lag.14      0.075      0.076   0.986  0.327
Lag.17     -0.075      0.051  -1.470  0.146
Lag.19     -0.007      0.055  -0.120  0.905

Number of available observations: 88

```

Figura 12.5: Test de Dickey Fuller a serie de PBI-industria.

12.1. Autocorrelograma simple y parcial del modelo seleccionado

RETARDO	FAC	FACP	Estad-Q.	[valor p]
1	0,1095	0,1095	1,2364	[0,266]
2	-0,1134	-0,1269	2,5742	[0,276]
3	-0,0144	0,0138	2,5961	[0,458]
4	0,043	0,0296	2,7929	[0,593]
5	-0,0198	-0,0299	2,8348	[0,725]
6	-0,0182	-0,0033	2,8706	[0,825]
7	-0,0759	-0,0812	3,5016	[0,835]
8	-0,08	-0,0674	4,2118	[0,838]
9	-0,0331	-0,0328	4,3345	[0,888]
10	-0,026	-0,0383	4,4108	[0,927]
11	-0,1269	-0,1283	6,2556	[0,856]
12	-0,0555	-0,036	6,6123	[0,882]
13	0,128	0,1088	8,5346	[0,807]
14	0,1588	0,1203	11,5244	[0,644]
15	-0,0315	-0,0411	11,6436	[0,706]
16	-0,0452	-0,0212	11,8915	[0,751]
17	0,1023	0,0888	13,1781	[0,724]
18	0,0072	-0,05	13,1844	[0,781]
19	-0,0365	-0,031	13,352	[0,820]
20	0,0392	0,0487	13,5484	[0,853]
21	-0,0159	-0,0161	13,5812	[0,887]
22	-0,1314	-0,1191	15,8405	[0,824]
23	-0,0454	-0,0269	16,113	[0,850]
24	-0,2527 **	-0,2651 ***	24,6828	[0,423]

Tabla 12.1: Autocorrelograma simple y parcial del modelo seleccionado

12.2. Predicción período Junio de 2007 - Julio de 2008

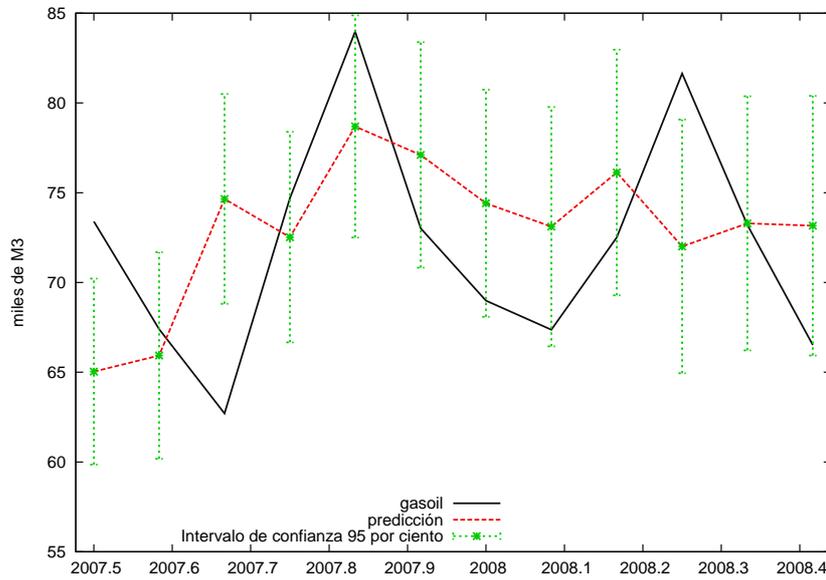


Figura 12.6: Predicción período Julio de 2007 a Junio de 2008 - modelo seleccionado grupo gas oil

			Intervalo de Prediccion 95 %		Intervalo de Prediccion 99 %				
Obs	gasoil	pred	Inf	Sup	Inf	Sup	ERROR	Prediccion Univariada	Error Univariado
Jul-07	73,40	65,04	59,85	70,22	58,23	71,85	-11 %	69,70	-5 %
Ago-07	67,40	65,93	60,18	71,68	58,37	73,49	-2 %	69,25	3 %
Sep-07	62,70	74,65	68,81	80,50	66,97	82,34	19 %	73,13	17 %
Oct-07	74,70	72,52	66,66	78,39	64,82	80,23	-3 %	74,69	1 %
Nov-07	83,98	78,69	72,50	84,88	70,55	86,83	-6 %	75,70	-10 %
Dic-07	73,02	77,10	70,82	83,38	68,85	85,36	6 %	75,59	4 %
Ene-08	68,99	74,41	68,08	80,74	66,10	82,73	8 %	71,31	3 %
Feb-08	67,37	73,11	66,44	79,78	64,35	81,87	9 %	74,30	10 %
Mar-08	72,51	76,13	69,28	82,97	67,13	85,12	5 %	72,59	0 %
Abr-08	81,64	72,01	64,94	79,07	62,72	81,29	-12 %	72,19	-12 %
May-08	73,20	73,29	66,21	80,38	63,99	82,60	0 %	76,48	4 %
Jun-08	66,52	73,16	65,93	80,39	63,65	82,67	10 %	75,38	13 %
ECM-multivariado = 39,72									
ECM-univariado = 36,12									

Tabla 12.2: Predicción para el período Julio de 2007 - Junio de 2008 - modelo seleccionado grupo gas oil

12.3. Modelos rechazados

Los modelos que se presentan a continuación son algunos de los modelos probados, la mayoría de ellos descartados por su desempeño en predicción.

12.3.1. Modelo seleccionado con Enero de 2005 y Marzo de 2007 como atípicos

Variable	Coefficiente	Desv. típica	Estadístico t	valor p	
d(gasoil)-1	-0,92	0,11	-8,02	<0,00001	***
d(gasoil)-2	-0,88	0,16	-5,45	<0,00001	***
d(gasoil)-3	-0,98	0,19	-5,10	<0,00001	***
d(gasoil)-4	-0,87	0,20	-4,41	0,00	***
d(gasoil)-5	-0,50	0,20	-2,53	0,01	**
d(gasoil)-6	-0,43	0,18	-2,43	0,02	**
d(gasoil)-7	-0,36	0,12	-3,02	0,00	***
d(precio)-1	-2,03	0,58	-3,50	0,00	***
d(precio)-2	-1,30	0,67	-1,93	0,06	*
d(precio)-3	-0,06	0,67	-0,09	0,93	
d(precio)-4	-2,06	0,67	-3,05	0,00	***
d(precio)-5	-0,68	0,68	-1,00	0,32	
d(precio)-6	0,53	0,68	0,78	0,44	
d(precio)-7	-1,46	0,66	-2,19	0,03	**
d(Lluvias)-1	-0,05	0,01	-4,03	0,00	***
d(Lluvias)-2	-0,05	0,01	-3,21	0,00	***
d(Lluvias)-3	-0,04	0,02	-2,88	0,01	***
d(Lluvias)-4	-0,03	0,01	-2,09	0,04	**
d(Lluvias)-5	-0,02	0,01	-1,75	0,08	*
d(Lluvias)-6	-0,02	0,01	-1,85	0,07	*
d(Lluvias)-7	-0,01	0,01	-1,83	0,07	*
Oct-2003	13,74	4,39	3,13	0,00	***
Mar-2002	-13,32	4,12	-3,23	0,00	***
Mar-2007	-8,19	4,19	-1,95	0,06	*
Ene-05	0,50	3,95	0,13	0,90	
S1	-9,35	2,18	-4,28	0,00	***
S2	-10,55	2,69	-3,92	0,00	***
S3	-1,38	3,09	-0,45	0,66	
S4	-10,27	2,79	-3,68	0,00	***
S5	-13,09	2,74	-4,78	0,00	***
S6	-12,29	2,60	-4,72	0,00	***
S7	-12,66	2,69	-4,70	0,00	***
S8	-14,72	2,88	-5,12	<0,00001	***
S9	-7,92	2,89	-2,74	0,01	***
S10	-7,13	2,93	-2,43	0,02	**
S11	-3,12	2,13	-1,46	0,15	
EC1	-0,11	0,03	-4,32	0,00	***

Tabla 12.3: Modelo seleccionado con Enero de 2005 y Marzo de 2007 como atípicos

12.3.2. Modelo seleccionado con Julio como atípico

Variable	Coficiente	Desv. típica	Estadístico t	valor p	
d(gasoil)-1	-1,06	0,12	-9,12	<0,00001	***
d(gasoil)-2	-1,02	0,17	-5,97	<0,00001	***
d(gasoil)-3	-1,07	0,20	-5,31	<0,00001	***
d(gasoil)-4	-0,94	0,20	-4,61	0,00	***
d(gasoil)-5	-0,69	0,20	-3,40	0,00	***
d(gasoil)-6	-0,65	0,18	-3,59	0,00	***
d(gasoil)-7	-0,40	0,12	-3,41	0,00	***
d(precio)-1	-1,75	0,60	-2,90	0,01	***
d(precio)-2	-1,38	0,66	-2,10	0,04	**
d(precio)-3	-0,04	0,65	-0,06	0,95	
d(precio)-4	-2,21	0,65	-3,37	0,00	***
d(precio)-5	-0,14	0,67	-0,21	0,84	
d(precio)-6	0,46	0,68	0,67	0,50	
d(precio)-7	-1,66	0,67	-2,46	0,02	**
d(Lluvias)-1	-0,05	0,01	-3,98	0,00	***
d(Lluvias)-2	-0,04	0,01	-3,08	0,00	***
d(Lluvias)-3	-0,04	0,01	-2,65	0,01	**
d(Lluvias)-4	-0,03	0,01	-1,91	0,06	*
d(Lluvias)-5	-0,03	0,01	-2,13	0,04	**
d(Lluvias)-6	-0,02	0,01	-1,54	0,13	
d(Lluvias)-7	-0,01	0,01	-0,65	0,52	
Oct-2003	18,42	4,41	4,18	0,00	***
Mar-2002	-13,95	4,00	-3,49	0,00	***
Jul-2007	8,36	4,46	1,87	0,07	*
S1	-8,34	2,10	-3,96	0,00	***
S2	-10,03	2,62	-3,83	0,00	***
S3	-2,81	3,01	-0,94	0,35	
S4	-9,30	2,80	-3,33	0,00	***
S5	-12,53	2,75	-4,56	0,00	***
S6	-12,99	2,59	-5,01	<0,00001	***
S7	-16,20	2,88	-5,62	<0,00001	***
S8	-15,02	2,99	-5,02	<0,00001	***
S9	-6,61	2,95	-2,24	0,03	**
S10	-8,89	2,93	-3,03	0,00	***
S11	-3,78	2,13	-1,78	0,08	*
EC1	-0,09	0,02	-3,67	0,00	***

Tabla 12.4: Coeficientes y significación de las variables integrantes del modelo seleccionado con Julio de 2007 como atípico

12.3.3. Modelo 1 - Grupo Gas Oil

- 7 retardos
- variables: gas oil, precio y lluvia.
- Especificación: con constante en la relación de cointegración y en la relación de equilibrio.
- Con variables indicatrices estacionales

■ Atípicos: Octubre de 2003 y Marzo de 2002

Se destaca que la constante es significativa en el análisis, sin embargo el modelo incurre en mayor error cuadrático medio (a comparación del modelo que se finalmente se selecciona.)

Beta: vector cointegrante					
gasoil	precio	Lluvias			
1	5,81	1,74			
0	(2,25)	(0,51)			
Alfa: Velocidad de ajuste					
gasoil	precio	Lluvias			
0,03	0,002	-0,24			
Variable	Coefficiente	Desv. típica	Estadístico t	valor p	
const	-10,22	2,91	-3,52	0,00	***
d(gasoil)-1	-1,08	0,12	-9,10	<0,00001	***
d(gasoil)-2	-1,01	0,17	-6,01	<0,00001	***
d(gasoil)-3	-1,08	0,20	-5,42	<0,00001	***
d(gasoil)-4	-0,97	0,20	-4,73	0,00	***
d(gasoil)-5	-0,61	0,20	-2,99	0,00	***
d(gasoil)-6	-0,53	0,18	-2,91	0,00	***
d(gasoil)-7	-0,39	0,12	-3,25	0,00	***
d(precio)-1	-1,84	0,58	-3,15	0,00	***
d(precio)-2	-1,35	0,69	-1,94	0,06	*
d(precio)-3	0,14	0,69	0,20	0,84	
d(precio)-4	-2,08	0,70	-2,98	0,00	***
d(precio)-5	-0,40	0,69	-0,59	0,56	
d(precio)-6	0,69	0,69	1,00	0,32	
d(precio)-7	-1,36	0,69	-1,97	0,05	*
d(lluvias)-1	-0,06	0,02	-4,03	0,00	***
d(lluvias)-2	-0,05	0,02	-3,29	0,00	***
d(lluvias)-3	-0,05	0,02	-3,10	0,00	***
d(lluvias)-4	-0,04	0,02	-2,36	0,02	**
d(lluvias)-5	-0,03	0,01	-2,10	0,04	**
d(lluvias)-6	-0,02	0,01	-2,09	0,04	**
d(lluvias)-7	-0,02	0,01	-1,96	0,05	*
Oct-2003	15,28	4,47	3,42	0,00	***
Mar-2002	-12,94	4,24	-3,05	0,00	***
S1	-8,91	2,18	-4,09	0,00	***
S2	-10,40	2,74	-3,80	0,00	***
S3	-2,32	3,09	-0,75	0,46	
S4	-9,63	2,83	-3,41	0,00	***
S5	-13,30	2,79	-4,77	0,00	***
S6	-12,70	2,68	-4,75	0,00	***
S7	-13,38	2,75	-4,86	<0,00001	***
S8	-14,84	2,93	-5,06	<0,00001	***
S9	-7,59	2,92	-2,60	0,01	**
S10	-7,44	2,97	-2,51	0,01	**
S11	-3,10	2,18	-1,42	0,16	
EC1	0,03	0,01	3,80	0,00	***

Tabla 12.5: Coeficientes y significación de las variables integrantes del modelo

			Intervalo de Prediccion 95 %		
Obs	gasoil	predicción	Inf	Sup	Error
Ene-08	68,99	71,43	65,89	76,97	4 %
Feb-08	67,37	71,97	66,00	77,94	7 %
Mar-08	72,51	72,09	65,95	78,24	-1 %
Abr-08	81,64	72,97	66,81	79,14	-11 %
May-08	73,20	70,68	64,19	77,18	-3 %
Jun-08	66,52	67,51	60,85	74,16	1 %
Jul-08	68,79	71,57	64,88	78,25	4 %
Ago-08	65,01	67,50	60,59	74,40	4 %
Sep-08		72,36	65,12	79,61	
Oct-08		74,28	66,73	81,83	
Nov-08		77,12	69,54	84,71	
Dic-08		77,41	69,71	85,12	
ECM-multivariado = 15,46					

Tabla 12.6: Predicción para el año 2008 - Modelo 1

12.3.4. Modelo 2 - Grupo Gas Oil

- 3 retardos
- variables: gas oil, lluvia y PBI del agro .
- Especificación: sin constante en la relación de cointegración.
- Atípicos: Octubre de 2003
- Con variables indicatrices estacionales

Variable	Coficiente	Desv. típica	Estadístico t	valor p	
d(gasoil)-1	-0,77	0,11	-7,06	<0,00001	***
d(gasoil)-2	-0,56	0,13	-4,36	0,00	***
d(gasoil)-3	-0,22	0,11	-1,97	0,05	*
d(lluvias)-1	-0,04	0,01	-3,19	0,00	***
d(lluvias)-2	-0,02	0,01	-1,97	0,05	*
d(lluvias)-3	-0,01	0,01	-0,93	0,35	
d(PBI-agro)-1	-0,20	0,18	-1,10	0,28	
d(PBI-agro)-2	0,05	0,23	0,21	0,83	
d(PBI-agro)-3	-0,24	0,19	-1,26	0,21	
Oct-2003	10,95	4,41	2,48	0,02	**
S1	-8,17	2,95	-2,77	0,01	***
S2	-14,52	4,52	-3,22	0,00	***
S3	-12,54	5,28	-2,37	0,02	**
S4	-18,94	5,01	-3,78	0,00	***
S5	-18,03	4,37	-4,13	0,00	***
S6	-19,36	4,19	-4,62	0,00	***
S7	-12,84	3,45	-3,72	0,00	***
S8	-9,68	3,58	-2,70	0,01	***
S9	-8,36	3,83	-2,18	0,03	**
S10	-5,96	3,60	-1,66	0,10	
S11	-1,38	2,98	-0,46	0,64	
EC1	-0,14	0,04	-3,38	0,00	***

Tabla 12.7: Coeficientes y significación de las variables del modelo 2

Obs	gasoil	pred	Intervalo de Predicción 95 %		ERROR
			Inf	Sup	
Ene-08	68,99	65,87	58,71	73,02	-5 %
Feb-08	67,37	58,89	51,65	66,13	-13 %
Mar-08	72,51	56,55	48,87	64,23	-22 %
Abr-08	81,64	52,99	44,48	61,51	-35 %
May-08	73,20	60,57	51,27	69,87	-17 %
Jun-08	66,52	64,72	54,96	74,49	-3 %
Jul-08	68,79	68,38	58,26	78,50	-1 %
Ago-08	65,01	63,32	52,91	73,72	-3 %
Sep-08		53,05	42,36	63,73	
Oct-08		49,43	38,25	60,61	
Nov-08		56,72	44,73	68,71	
Dic-08		69,69	56,91	82,47	
ECM-multivariado = 165,37					

Tabla 12.8: Predicción para el año 2008 - Modelo 2

12.3.5. Modelo 3 - Grupo Gas Oil

- 12 retardos
- variables: gas oil, lluvia y PBI del agro.
- Especificación: con constante.
- Atípicos: ninguno

			Intervalo de Prediccion 95 %		ERROR
	gasoil	pred	Inf	Sup	
Ene-08	68,99	83,59	78,57	88,61	21 %
Feb-08	67,37	82,24	76,99	87,48	22 %
Mar-08	72,51	61,48	56,19	66,77	-15 %
Abr-08	81,64	40,73	35,16	46,29	-50 %
May-08	73,20	23,34	17,58	29,11	-68 %
Jun-08	66,52	85,16	79,07	91,25	28 %
Jul-08	68,79	160,56	154,23	166,89	133 %
Ago-08	65,01	175,77	169,02	182,52	170 %
Sep-08		42,82	34,35	51,30	
Oct-08		-130,79	-140,82	-120,75	
Nov-08		-183,76	-194,91	-172,61	
Dic-08		68,29	50,60	85,97	
ECM-multivariado = 3219					

Tabla 12.9: Predicciones para el año 2008 - Modelo 3

Variable	Coefficiente	Desv. típica	Estadístico t	valor p	
const	45,55	16,18	2,82	0,01	***
d(gasoi)-1	-0,20	0,31	-0,65	0,52	
d(gasoi)-2	-0,27	0,30	-0,90	0,37	
d(gasoi)-3	-0,02	0,29	-0,08	0,94	
d(gasoi)-4	0,01	0,30	0,05	0,96	
d(gasoi)-5	0,44	0,31	1,43	0,16	
d(gasoi)-6	0,47	0,31	1,50	0,14	
d(gasoi)-7	0,40	0,32	1,27	0,21	
d(gasoi)-8	0,53	0,31	1,74	0,09	*
d(gasoi)-9	0,60	0,29	2,10	0,04	**
d(gasoi)-10	0,44	0,27	1,64	0,11	
d(gasoi)-11	0,12	0,22	0,56	0,58	
d(gasoi)-12	0,09	0,14	0,66	0,52	
d(Lluvias)-1	-0,02	0,01	-1,86	0,07	*
d(Lluvias)-2	-0,01	0,01	-0,84	0,40	
d(Lluvias)-3	0,00	0,01	-0,36	0,72	
d(Lluvias)-4	0,01	0,01	0,84	0,40	
d(Lluvias)-5	0,01	0,01	0,70	0,49	
d(Lluvias)-6	0,01	0,01	0,88	0,38	
d(Lluvias)-7	-0,01	0,01	-0,47	0,64	
d(Lluvias)-8	0,02	0,01	1,16	0,25	
d(Lluvias)-9	0,00	0,01	0,14	0,89	
d(Lluvias)-10	0,03	0,01	1,84	0,07	*
d(Lluvias)-11	0,02	0,01	1,92	0,06	*
d(Lluvias)-12	0,02	0,01	1,64	0,11	
d(PBI-agro)-1	0,12	0,51	0,23	0,82	
d(PBI-agro)-2	-1,09	1,52	-0,72	0,47	
d(PBI-agro)-3	0,77	2,69	0,29	0,78	
d(PBI-agro)-4	0,39	3,04	0,13	0,90	
d(PBI-agro)-5	-3,29	4,00	-0,82	0,41	
d(PBI-agro)-6	2,10	5,73	0,37	0,72	
d(PBI-agro)-7	2,89	5,95	0,49	0,63	
d(PBI-agro)-8	-6,61	4,60	-1,44	0,16	
d(PBI-agro)-9	3,82	4,19	0,91	0,37	
d(PBI-agro)-10	0,60	4,59	0,13	0,90	
d(PBI-agro)-11	-2,21	3,39	-0,65	0,52	
d(PBI-agro)-12	0,35	1,31	0,27	0,79	
S1	-25,22	8,93	-2,82	0,01	***
S2	-28,83	10,30	-2,80	0,01	***
S3	-28,17	11,95	-2,36	0,02	**
S4	-24,30	10,73	-2,27	0,03	**
S5	-34,93	10,47	-3,34	0,00	***
S6	-28,27	10,87	-2,60	0,01	**
S7	-25,83	10,50	-2,46	0,02	**
S8	-17,66	10,17	-1,74	0,09	*
S9	-25,06	12,12	-2,07	0,04	**
S10	-4,43	10,37	-0,43	0,67	
S11	-3,43	8,84	-0,39	0,70	
EC1	-0,92	0,33	-2,81	0,01	***

Tabla 12.10: Coeficientes y significación de las variables integrantes del modelo 3

12.3.6. Modelo 4 - Grupo Gas Oil

- 16 retardos
- variables: gas oil, lluvia y PBI de la industria .
- Especificación: con constante en la relación de cointegración.
- Atípicos: Julio de 2004
- Sin variables indicatrices estacionales

			Intervalo de Prediccion 95 %		ERROR
	gasoil	pred	Inf	Sup	
Ene-08	68,99	83,38	79,32	87,44	21 %
Feb-08	67,37	76,83	71,27	82,40	14 %
Mar-08	72,51	72,44	66,54	78,34	0 %
Abr-08	81,64	74,26	68,07	80,46	-9 %
May-08	73,20	75,09	68,76	81,43	3 %
Jun-08	66,52	75,76	69,22	82,30	14 %
Jul-08	68,79	81,50	74,65	88,34	18 %
Ago-08	65,01	79,68	72,58	86,79	23 %
Sep-08		73,81	65,90	81,72	
Oct-08		77,28	69,15	85,40	
Nov-08		83,64	75,09	92,19	
Dic-08		71,30	62,56	80,04	
ECM-multivariado = 102,1					

Tabla 12.11: Predicciones para el año 2008 - Modelo 4

Variable	Coficiente	Desv. típica	Estadístico t	valor p	
const	-43,12	17,79	-2,42	0,02	**
d(gasoi)-1	-1,50	0,23	-6,45	<0,00001	***
d(gasoi)-2	-1,28	0,28	-4,57	0,00	***
d(gasoi)-3	-1,08	0,30	-3,55	0,00	***
d(gasoi)-4	-0,93	0,31	-3,03	0,00	***
d(gasoi)-5	-0,96	0,32	-3,01	0,00	***
d(gasoi)-6	-1,13	0,32	-3,48	0,00	***
d(gasoi)-7	-1,42	0,33	-4,30	0,00	***
d(gasoi)-8	-1,12	0,36	-3,12	0,00	***
d(gasoi)-9	-0,79	0,37	-2,15	0,04	**
d(gasoi)-10	-0,70	0,34	-2,08	0,04	**
d(gasoi)-11	-0,56	0,31	-1,78	0,08	*
d(gasoi)-12	-0,15	0,30	-0,51	0,62	
d(gasoi)-13	-0,11	0,27	-0,41	0,68	
d(gasoi)-14	-0,31	0,22	-1,39	0,17	
d(gasoi)-15	0,15	0,18	0,88	0,38	
d(gasoi)-16	-0,19	0,13	-1,47	0,15	
d(Lluvias)-1	-0,14	0,06	-2,38	0,02	**
d(Lluvias)-2	-0,13	0,05	-2,39	0,02	**
d(Lluvias)-3	-0,09	0,05	-1,82	0,08	*
d(Lluvias)-4	-0,09	0,04	-2,04	0,05	**
d(Lluvias)-5	-0,08	0,04	-2,23	0,03	**
d(Lluvias)-6	-0,09	0,03	-2,68	0,01	**
d(Lluvias)-7	-0,10	0,03	-2,99	0,00	***
d(Lluvias)-8	-0,06	0,03	-2,03	0,05	**
d(Lluvias)-9	-0,05	0,03	-1,81	0,08	*
d(Lluvias)-10	-0,03	0,03	-1,14	0,26	
d(Lluvias)-11	-0,03	0,03	-1,32	0,19	
d(Lluvias)-12	-0,04	0,02	-1,74	0,09	*
d(Lluvias)-13	-0,03	0,02	-1,53	0,13	
d(Lluvias)-14	-0,03	0,02	-2,02	0,05	*
d(Lluvias)-15	0,00	0,01	-0,26	0,79	
d(Lluvias)-16	0,00	0,01	0,29	0,78	
d(PBI-ind)-1	2,74	0,49	5,59	<0,00001	***
d(PBI-ind)-2	-4,88	0,91	-5,40	<0,00001	***
d(PBI-ind)-3	3,25	0,95	3,43	0,00	***
d(PBI-ind)-4	1,18	1,00	1,18	0,24	
d(PBI-ind)-5	-1,71	0,95	-1,80	0,08	*
d(PBI-ind)-6	1,02	0,93	1,09	0,28	
d(PBI-ind)-7	0,39	0,90	0,44	0,66	
d(PBI-ind)-8	1,13	0,88	1,29	0,20	
d(PBI-ind)-9	-1,79	0,90	-1,98	0,05	*
d(PBI-ind)-10	1,49	0,90	1,66	0,11	
d(PBI-ind)-11	0,37	0,91	0,40	0,69	
d(PBI-ind)-12	-1,28	0,87	-1,47	0,15	
d(PBI-ind)-13	1,97	0,83	2,37	0,02	**
d(PBI-ind)-14	-0,07	0,82	-0,08	0,93	
d(PBI-ind)-15	-1,08	0,75	-1,45	0,16	
d(PBI-ind)-16	0,94	0,43	2,20	0,03	**
Jul-2004	-17,82	4,56	-3,91	0,00	***
EC1	0,42	0,17	2,43	0,02	**

Tabla 12.12: Coeficientes y significación de las variables integrantes del modelo 4