



FACULTAD DE
CIENCIAS ECONÓMICAS
Y DE ADMINISTRACIÓN



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

**UNIVERSIDAD DE LA REPÚBLICA
FACULTAD DE CIENCIAS ECONÓMICAS Y DE ADMINISTRACIÓN**

**TRABAJO FINAL DE GRADO PARA OBTENER EL TÍTULO DE
LICENCIADO EN ECONOMÍA**

**Modelos SARFIMA diarios no estacionarios: aplicación al consumo de energía
eléctrica.**

por

Mauricio Giacometti Ferrer

TUTOR: Adrián Fernández Poncet

**Montevideo
URUGUAY
2007**

A las personas que hacen que cada día intente superarme

A mis abuelos

Germán Ferrer

María Julia Melo

Abstract

El presente trabajo tiene por objetivo investigar el poder predictivo de los modelos SARFIMA para el análisis de series temporales diarias, con una aplicación a la serie diaria de consumo eléctrico del Uruguay.

Este tipo de modelos permite el estudio de características de largo plazo de la serie en cuestión y disminuye la posibilidad de sobrediferenciación de la serie ya que los parámetros de los polinomios de diferenciación son estimados y pueden ser puestos a prueba mediante estadísticos con distribución conocida.

La estimación es llevada a cabo con el método de mínimos cuadrados, mediante la utilización del algoritmo de Gauss-Newton. A diferencia de otros métodos de estimación, éste es llevado a cabo en una única etapa, lo que da lugar a estimadores más eficientes.

Para poner a prueba la calidad de las predicciones se contrasta el poder predictivo de estos modelos contra modelos SARIMA mediante indicadores como el MAE y MAPE. Los resultados indican que un modelo que permita la diferenciación fraccional es una herramienta poderosa para la predicción de este tipo de series.

Palabras clave: predicción del consumo de energía eléctrica, diferenciación fraccional estacional, series diarias.

1) Introducción

A partir de la década de los 70 se popularizan los modelos de series de tiempo ARIMA, gracias a la obra de Box y Jenkins; estos modelos simples pretenden describir el comportamiento de una variable en función de sus valores pasados sin tener que recurrir para ello a una teoría económica bien desarrollada.

Esta modelización se generaliza para poder modelizar datos que presentan estacionalidad, es decir, que presentan oscilaciones periódicas, con período menor al año. Estos modelos son denominados SARIMA(p,d,q)(P,D,Q)_s, en donde d y D son seleccionados de manera de que la serie resultante de la diferenciación sea estacionaria.

A diferencia de los modelos SARIMA(p, d, q)(P, D, Q)_s, los modelos SARFIMA(p, d, q)(P, D, Q)_s presentan una ventaja en su formulación y es que permiten que los parámetros “d” y “D” sean estimados y tomen valores racionales en vez de enteros, es decir, permiten la existencia de raíces fraccionales, por lo que se deberá diferenciar fraccionalmente a la variable de estudio para que resulte integrada de orden 0.

La idea de diferenciación fraccional surge ya en 1695 cuando L'Hopital pregunta a Leibniz (inventor de la notación $\partial^n y / \partial x^n$) ¿que pasaría si “n” fuese ½? Desde ese momento diferentes matemáticos han desarrollado la idea de que “n” fuese real o complejo, hasta llegar al caso de diferenciación discreta racional, que es la utilizada en este caso.

En palabras de Granger y Joyeux[13] los econometristas diferencian sus series “para lograr estacionariedad”, o sea que suponen que la serie original tiene varianza infinita. A su vez, aquellos que estudian las series de tiempo en el dominio de la frecuencia se muestran aversos a la diferenciación de las series porque aseguran que este filtro quita información en las frecuencias bajas, o sea aquellas asociadas al largo plazo.

Los procesos SARFIMA (Seasonally Autorregresive Fractionally Integrated Moving Average) vienen, de esta manera, a llenar un vacío en la modelización de series de tiempo estacionales, permitiendo reconciliar estas dos posturas al dejar que los parámetros de diferenciación tomen valores reales, donde el 0 y el 1 (o cualquier valor entero) son casos particulares.

Desde un punto de vista empírico, la lenta convergencia de las autocorrelaciones de algunas series hacia cero no concuerda con la idea de los modelos SARMA, pero tampoco con la de los modelos SARIMA, ya que sus autocorrelaciones no deberían decaer sino mantenerse estables. Es por esto que se hace necesario generalizar la modelización para permitir mayor flexibilidad.

La posibilidad de que el parámetro de diferenciación “d” tome valores no enteros fue introducido en econometría por Granger y Joyeux y Hoskin, pero este trabajo toma la idea de otras ciencias como la hidrometría con trabajos como el de Hurst, en donde se refieren a este tipo de modelos como capaces de capturar las correlaciones existentes entre observaciones distantes en el tiempo, o fenómeno de “long memory” o “long range dependence”.

El hecho de que los parámetros “d” y “D” tomen valores diferentes a los naturales, (usualmente 0, 1, 2), lo que fue llamado por Maynard y Phillips (2001) como “alternativas extremas”, permite distinguir a los procesos entre aquellos de “memoria larga” y aquellos antipersistentes o sobrediferenciados. Entre los procesos de “memoria larga” se encuentran

aquellos estacionarios, los no estacionarios con varianza finita pero que retornan a la media y aquellos que no retornan a la media ni tienen varianza finita.

Como forma de visualizar los beneficios de esta manera de estudiar las series temporales es que en el presente trabajo se aplica esta modelización al consumo diario de energía eléctrica de Uruguay.

Para estimar los parámetros se propone la utilización del método de minimización de la suma de los errores al cuadrado, y para lograr este objetivo se utiliza el algoritmo de Gauss-Newton. Estos estimadores son iguales a aquellos encontrados por el método de cuasi máxima verosimilitud y por ello “heredan” sus propiedades, lo que permite la utilización de estimadores como Wald, razón de verosimilitud o multiplicadores de Lagrange para probar hipótesis lineales sobre los parámetros.

Con el objetivo de mostrar la relevancia de este tipo de modelización se estiman a su vez modelos SARIMA similares, para poder hacer una comparación en términos de poder predictivo a distintos horizontes temporales.

La presente monografía se estructura de la siguiente manera: la próxima sección comenta las características del proceso generador de datos. En la tercera sección se analiza el procedimiento de estimación. A continuación, en el cuarto apartado se analiza la relevancia que tiene este tipo de análisis al estudio del consumo eléctrico, así como las características de los datos a estudiar. Luego se presentan los resultados en cuanto a los valores estimados y su poder predictivo. Las conclusiones se encuentran en el sexto capítulo y en el último capítulo son expuestas algunas consideraciones sobre el método y lineamientos para futuras investigaciones.

2) Propiedades básicas de los modelos SARFIMA

El proceso generador de datos SARFIMA(p,d,q)(P,D,Q)₇ es definido como:

$$\Phi_p(L)\Phi_P(L)(1-L)^d(1-L^7)^D Y_t = \Theta_q(L)\Theta_Q(L)\varepsilon_t \quad (2.1)$$

Y_t es un proceso centrado, ε_t es un ruido blanco y los polinomios $\Phi_p(L) = 1 - \sum_{i=1}^p \phi_i L^i$,

$\Phi_P(L) = 1 - \sum_{i=1}^P \phi_{is} L^{is}$, $\Theta_q(L) = 1 + \sum_{i=1}^q \theta_i L^i$ y $\Theta_Q(L) = 1 + \sum_{i=1}^Q \theta_{is} L^{is}$, tienen todas sus raíces fuera del círculo unidad y no poseen ninguna raíz en común entre ellos.

Además, los polinomios $(1-L)^d$ y $(1-L^7)^D$ son los polinomios de diferenciación regular y estacional respectivamente y los parámetros d y D, los órdenes de diferenciación, que a diferencia de la modelización SARIMA, pueden tomar valores reales.

Para d, D > -1, podemos expandir cualquiera de estos binomios de diferenciación, gracias a un desarrollo de Taylor de orden infinito, como:

$$(1-L)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-1)^k L^k \quad (2.2)$$

$$(1-L^7)^D = \sum_{k=0}^{\infty} \binom{D}{k} (-1)^k L^{7k} \quad (2.3)$$

con coeficientes

$$\binom{a}{k} = \frac{a!}{k!(a-k)!} = \frac{\Gamma(a+1)}{\Gamma(k+1)\Gamma(a-k+1)} \quad a = d, D \quad (2.4)$$

en donde $\Gamma(\cdot)$ es la función Gamma definida como: $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$.

Los modelos que no admiten la presencia de raíces fraccionales modelan variables estacionarias que tienen autocorrelaciones que decaen a una tasa exponencial, mientras que los modelos que si permiten la existencia de raíces fraccionales aceptan que las autocorrelaciones de la variable de estudio presenten caídas a tasas hiperbólicas, o sea, más lentas.

Según Katayama[20], las autocovarianzas $\gamma(j)$ del proceso dependen de $(d+D)$ y de d a

$$\text{medida que } j \rightarrow \infty. \gamma(j) \sim \alpha_1 j^{2(d+D)-1} + j^{2D-1} \sum_{i=2}^{1+\left(\frac{s-1}{2}\right)} \{\alpha_i \cos(j\lambda_i) + \beta_i \sin(j\lambda_i)\} \quad (2.5)$$

donde β_i, α_i son constantes apropiadas que no dependen de j y λ_j está definida como la

$$\text{frecuencia j-ésima, } \lambda_j = \frac{2(j-1)\pi}{s-1} \forall j=1, \dots, 1+\left(\frac{s-1}{2}\right). \quad (2.6)$$

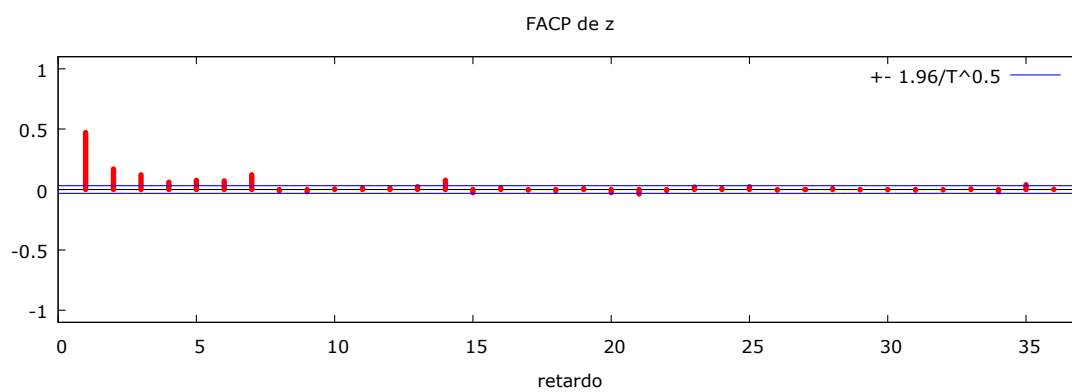
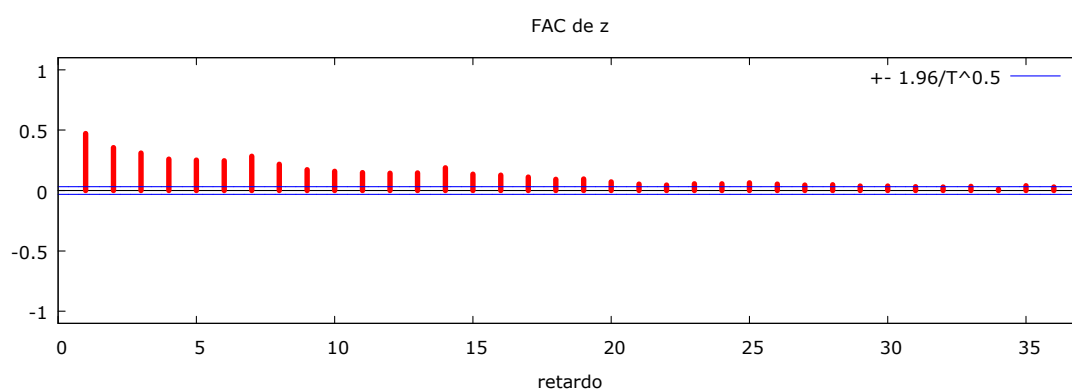
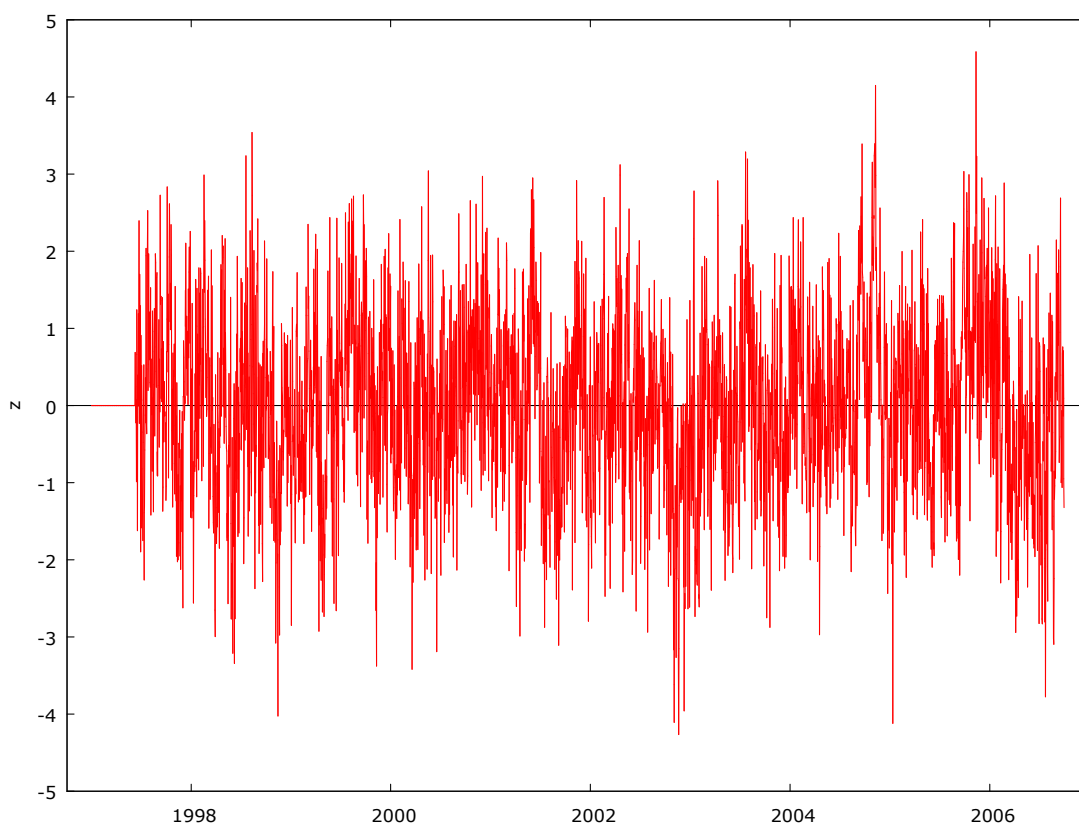
El proceso será estacionario si $(d+D) < \frac{1}{2}$ y $D < \frac{1}{2}$, mientras que será invertible si $(d+D) > -\frac{1}{2}$ y $D > -\frac{1}{2}$.

Entonces, si $\max (d+D,D) \in \left(-\frac{1}{2}, 0\right)$ entonces el proceso Y_t tiene memoria intermedia $\left(\sum_{j=0}^{\infty} |\gamma(j)| < \infty\right)$ y si $\max (d+D,D) \in \left(0, \frac{1}{2}\right)$ se dice que el proceso Y_t tiene memoria larga $\left(\sum_{j=0}^{\infty} |\gamma(j)| = \infty\right)$.

Para ilustrar el comportamiento de procesos fraccionalmente integrados se presenta a continuación 3 series simuladas. La primera es estacionaria e invertible, la segunda es no estacionaria, pero retorna a la media y la tercera serie no es estacionaria ni tiene varianza finita.

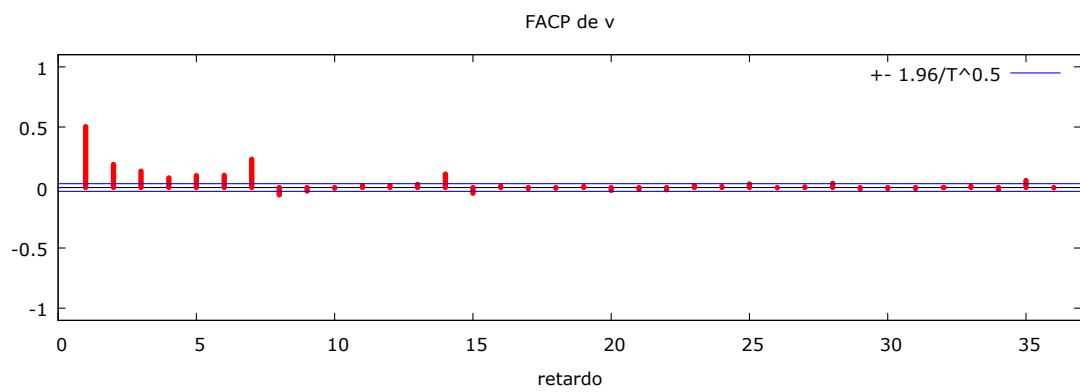
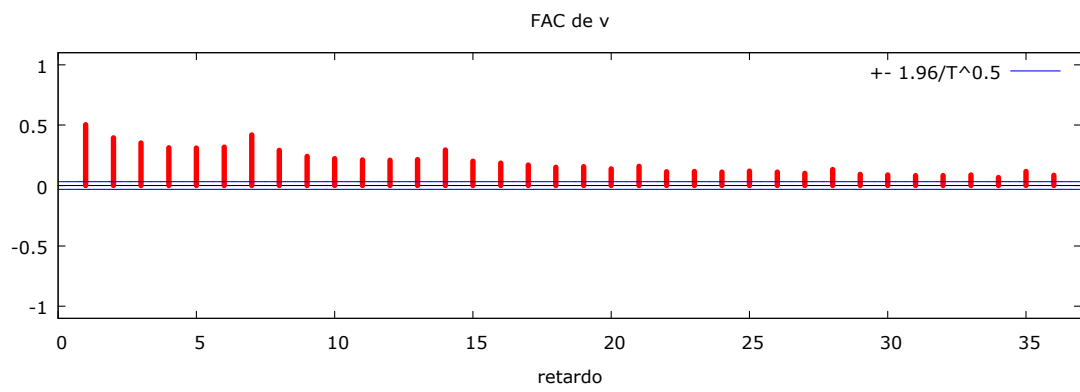
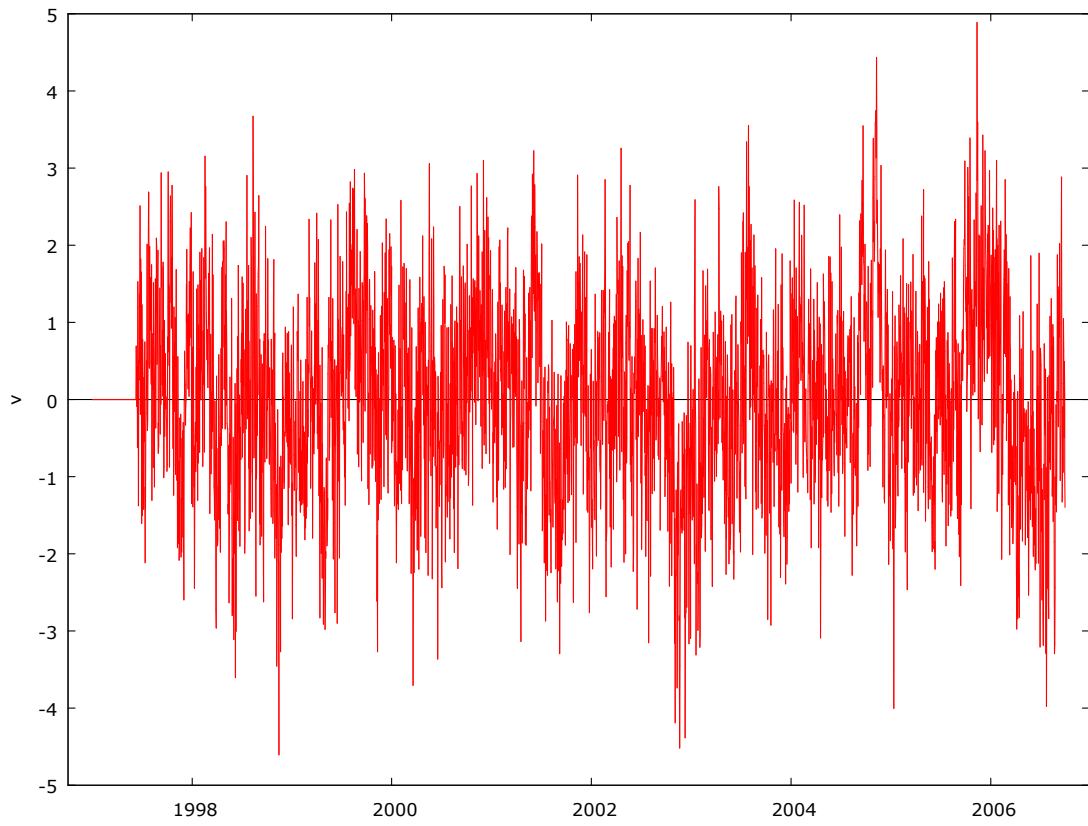
Ruido blanco gaussiano Z con $d=0.35$ y $D=0.1$

$$(1-L)^{0.35} (1-L^7)^{0.1} Z_t = \varepsilon_t$$



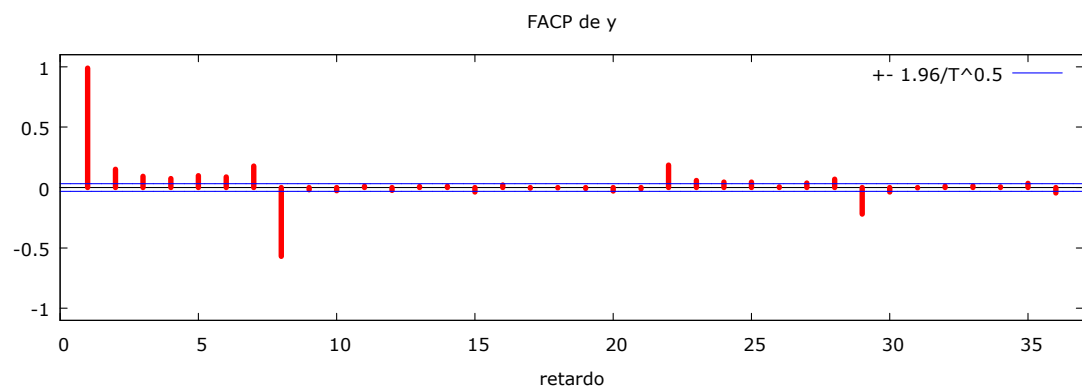
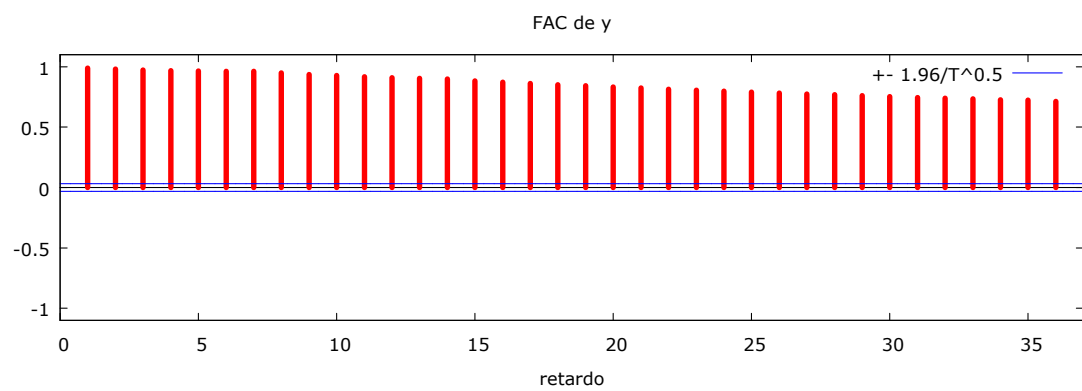
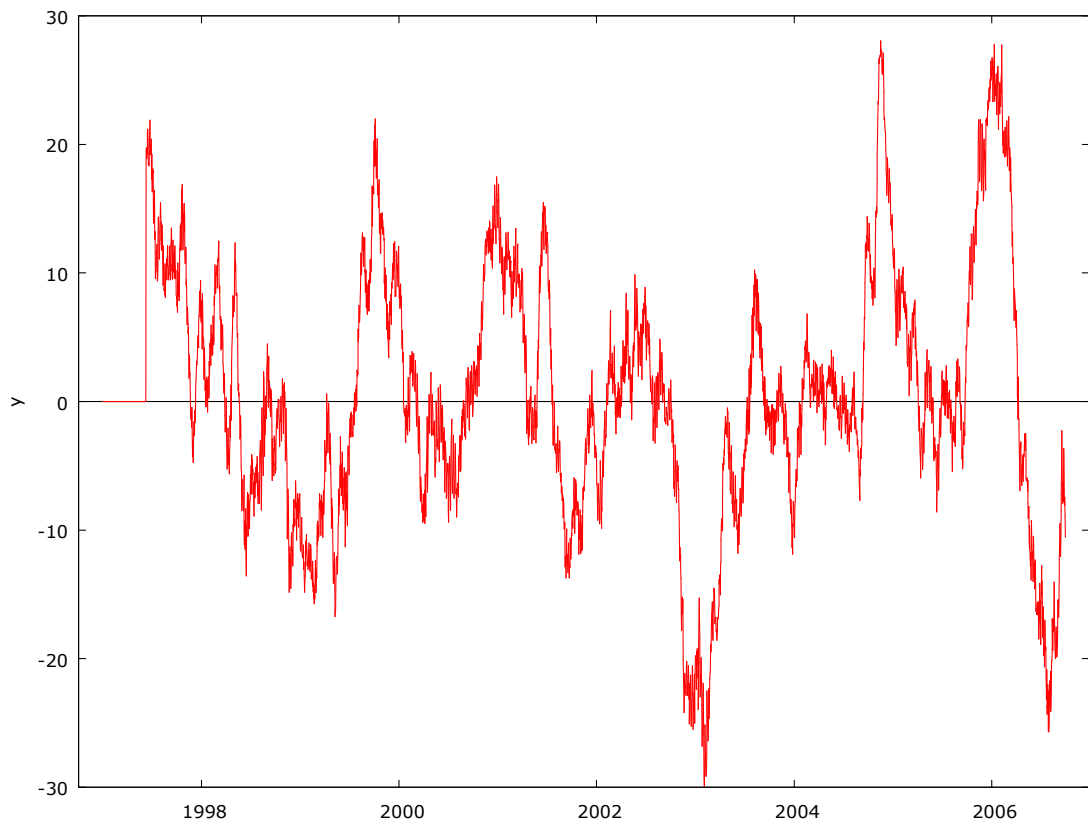
Variable fraccionalmente integrada V con $d=0.35$ y $D=0.2$

$$(1-L)^{0.35} (1-L^7)^{0.2} v_t = \varepsilon_t$$



Variable fraccionalmente integrada Y, con $d=0.85$ y $D=0.7$

$$(1-L)^{0.85} (1-L^7)^{0.7} y_t = \varepsilon_t$$



3) Estimación

Estimación por mínimos cuadrados no lineales

Siguiendo a Chan[10], el método seguido a lo largo de este capítulo podría enmarcarse como una “aproximación AR” al fenómeno de la “memoria larga”.

Se considera un proceso de gran persistencia y_t , definido por su expansión $AR(\infty)$

$$y_t + \pi_1(\theta)y_{t-1} + \pi_2(\theta)y_{t-2} + \dots = \varepsilon_t$$

donde π_j son los coeficientes de $\Phi(L)\Theta(L)^{-1}(1-L)^d$ y θ conforma el vector de los parámetros de estos polinomios y el parámetro de diferenciación, d . En la práctica, al contar con un número finito de observaciones, se considera el modelo truncado en algún valor m .

El estimador de los pesos π_j puede encontrarse maximizando la función de cuasi máxima verosimilitud:

$$\ell(\theta) = - \sum_{t=m+1}^T \left[y_t + \pi_1(\theta)y_{t-1} + \pi_2(\theta)y_{t-2} + \dots + \pi_m(\theta)y_{t-m} \right]^2 \quad (3.1)$$

que en caso de que las perturbaciones sean normales se puede estimar un procedimiento equivalente que es el de Mínimos cuadrados no lineales (MCNL).¹

Este método consiste en minimizar la suma de cuadrados residuales.

$$\text{Min} \sum_{t=1}^T e_t^2 \quad (3.2)$$

En donde los errores están definidos por:

$$e_t = Y_t - f(X_t, \hat{\beta}) \quad (3.3)$$

f es una función continua y diferenciable en los parámetros $\hat{\beta}$.

Esta minimización se logra igualando el vector de pendientes al vector nulo.

$$\frac{\partial \sum_{t=1}^T e_t^2}{\partial \hat{\beta}} = 0 \quad (3.4)$$

En el caso de que la función f no sea una función lineal en $\hat{\beta}$, entonces el sistema de ecuaciones definido en (3.4) no tiene una solución analítica y la solución a este sistema podría encontrarse mediante métodos numéricos como el algoritmo de Gauss-Newton.

Este algoritmo está definido por:

$$\hat{\beta}^{k+1} = \hat{\beta}^k - \left(J_e(\hat{\beta}^k)^T J_e(\hat{\beta}^k) \right)^{-1} J_e(\hat{\beta}^k)^T e(\hat{\beta}^k) \quad (3.5)$$

Donde k indica la cantidad de iteraciones, $J_e(\hat{\beta}^k)$ es la matriz jacobiana de la función errores evaluada en el vector $\hat{\beta}^k$.

¹ Ver anexo [C]

Propiedades de los estimadores MCNL

Dado que los métodos de cuasi máxima verosimilitud (cuando suponemos distribución normal de las perturbaciones) y de mínimos cuadrados no lineales arrojan el mismo vector de estimadores, estos últimos “heredan” las características de los primeros.

Los estimadores de cuasi máxima verosimilitud comparten las mismas propiedades asintóticas que los estimadores máximo verosímiles. Por lo tanto son:

1. Consistentes
2. Asintóticamente normales
3. Asintóticamente insesgados
4. Asintóticamente eficientes

$$\hat{\beta} \xrightarrow{d} N\left(\beta, \frac{\sigma^2}{n} Q_0^{-1}\right)$$

$$Q_0 = p \lim \frac{1}{n} X^{0'} X^0$$

$$Asy\hat{V}(\hat{\beta}) = \hat{\sigma}^2 (X^{0'} X^0)^{-1}$$

$$X^0 = -J_e(\hat{\beta}^K) = J_f(\hat{\beta}^K)$$

$$\hat{\sigma}^2 = \frac{e'e}{T}$$

Procedimiento de estimación

El proceso SARFIMA fue definido como:

$$\Phi_p(L) \Phi_p(L) (1-L)^d (1-L^s)^D Y_t = \Theta_q(L) \Theta_q(L) \varepsilon_t$$

Conviene hacer la siguiente descomposición:

$$\Phi_p(L) \Phi_p(L) (1-L)^{m_1} (1-L)^{m_2} (1-L^s)^{g_1} (1-L^s)^{g_2} Y_t = \Theta_q(L) \Theta_q(L) \varepsilon_t \quad (3.6)$$

donde: $d = m_1 + m_2$ y $D = g_1 + g_2$ y m_1 y g_1 son la parte entera de los parámetros de los polinomios de diferenciación regular y estacional respectivamente.

En este caso particular, dado que la serie es de frecuencia diaria se sustituye s por 7.

Para que sea posible la estimación, el proceso debe ser estacionario, por lo tanto se seleccionará m_1 y g_1 de manera de lograr la estacionariedad. Luego de seleccionar estos órdenes, se procede a la diferenciación correspondiente. Esta variable ya filtrada es estacionaria y le llamaremos X_t .

$$(1-L)^{m_1} (1-L^7)^{g_1} Y_t = X_t \quad (3.7)$$

Es por lo tanto a esta serie filtrada a la que se la someterá al proceso de estimación.

El error en t (e_t) queda definido como aquella parte de la variable original que no pudo ser capturada por la modelización evaluada en los parámetros estimados.

Vale destacar que la naturaleza de e_t y de ε_t son diferentes, ya que ε_t es la perturbación del modelo y recoge aquellos efectos que influyen sobre Y_t pero que no están presentes en el pasado de la variable, mientras que e_t recoge además de estos efectos, los desvíos que existen entre los verdaderos parámetros y los parámetros estimados.

Por lo que e_t queda definido como:

$$\frac{\widehat{\Phi}_p(L)\widehat{\Phi}_p(L)(1-L)^{\widehat{m}_2}(1-L^7)^{\widehat{g}_2}}{\widehat{\Theta}_q(L)\widehat{\Theta}_q(L)}X_t = e_t \quad (3.8)$$

El citado método consiste en encontrar el vector

$$\widehat{\beta} = \begin{pmatrix} \widehat{m}_2 \\ \widehat{g}_2 \\ \widehat{\phi}_{i1} \\ \widehat{\phi}_{i2*7} \\ \widehat{\theta}_{j1} \\ \widehat{\theta}_{j2*7} \end{pmatrix} / \min \sum e_t^2 \left(\widehat{m}_2, \widehat{g}_2, \widehat{\phi}_{i1}, \widehat{\phi}_{i2*7}, \widehat{\theta}_{j1}, \widehat{\theta}_{j2*7} \right) \quad (3.9)$$

Es decir, utilizar el método de mínimos cuadrados no lineales para estimar los parámetros $\widehat{m}_2, \widehat{g}_2, \widehat{\phi}_{i1}, \widehat{\phi}_{i2*7}, \widehat{\theta}_{j1}, \widehat{\theta}_{j2*7} \quad \forall i_1 = 1, 2, \dots, p. i_2 = 1, 2, \dots, P. j_1 = 1, 2, \dots, q. j_2 = 1, 2, \dots, Q.$

Para ello se utilizará el algoritmo de Gauss-Newton, tal como fuera mencionado en el marco teórico.

Tomando la expresión 2.4 y 2.2, la diferenciación fraccional puede definirse para $m_2 > -1$:

$$\begin{aligned} (1-L)^{m_2} &= 1 + \frac{m_2}{1!}(-L)^1 + \frac{m_2(m_2-1)}{2!}(-L)^2 + \dots \\ &= \sum_{j=0}^{\infty} \frac{m_2(m_2-1)(m_2-2)\dots(m_2-(j-1))}{j!}(-1)^j L^j \end{aligned} \quad (3.10)$$

Luego de sucesivas transformaciones podemos expresar esta expansión como²:

$$(1-L)^{m_2} = \sum_{j=0}^{\infty} \frac{\Gamma(j-m_2)}{\Gamma(j-1)\Gamma(-m_2)} L^j \quad (3.11)$$

$$(1-L)^{m_2} = \sum_{j=0}^{\infty} C_j L^j \quad (3.12)$$

donde los términos de la sumatoria, gracias a la propiedad recursiva de la función gamma ($\Gamma(x) = (x-1)\Gamma(x-1)$), son:

$$C_j = \frac{(j-m_2-1)}{j} C_{j-1} \quad (3.13)$$

Excepto $C_0=1$

De esta manera, el binomio de diferenciación fraccional queda expresado como:

$$(1-L)^{m_2} \approx \left(\sum_{j=0}^J C_j L^j \right) \quad (3.14)$$

De manera análoga podemos tomar el binomio de diferenciación estacional y aplicarle la misma expansión, tomando L^7 en vez de L y g_2 en vez de m_2 . A los pesos resultantes de esta transformación se les llamará F_s .

² Ver demostración en Anexo F

$$(1-L^7)^{\delta_2} \approx \left(\sum_{s=0}^S F_s L^{7s} \right) \quad (3.15)$$

$$F_s = \frac{(j - g_2 - 1)}{j} F_{s-1} \quad (3.16)$$

Con $F_0 = 1$

No confundir el subíndice s con el orden estacional s que en este caso fue fijado en 7 dado que la serie es de frecuencia diaria. De ahora en adelante s será el subíndice de F .

Luego de estas aclaraciones podemos describir el proceso como:

$$\Phi_p(L) \Phi_p(L) (1-L)^{m_1} (1-L^7)^{g_1} \left(\sum_{j=0}^J C_j L^j \right) \left(\sum_{s=0}^S F_s L^{7s} \right) Y_t = \Theta_q(L) \Theta_q(L) \varepsilon_t \quad (3.17)$$

Si sustituimos a los parámetros por sus estimadores y despejamos coherentemente obtendremos los e_t .

Por lo tanto vemos que e_t depende de los parámetros estimados $\hat{\phi}_{i1}, \hat{\phi}_{i2*7}, \hat{\theta}_{j1}, \hat{\theta}_{j2*7}, \hat{C}_j(\hat{m}_2)$ y de $\hat{F}_s(\hat{g}_2)$, en definitiva tenemos a $e_t(\hat{m}_2, \hat{g}_2, \hat{\phi}_{i1}, \hat{\phi}_{i2*7}, \hat{\theta}_{j1}, \hat{\theta}_{j2*7})$.

En cuanto a los límites de las sumatorias (J y S), la teoría indica que se logra la igualdad cuando tienden a ∞ , pero se tomaron los valores $S=20$ y $J=20$ entendiéndose que era ya una muy buena aproximación si tomamos en consideración la forma de la función a minimizar.³ De esta forma e_t queda expresado en función de $Y_{t-1}, Y_{t-2}, \dots, Y_{t-(7P+p+J+7S)}$, o sea Y_{t-185} aproximadamente por lo que se podría empezar a calcular la suma de cuadrados residuales desde el valor 186, por lo que se pierden 185 datos, lo que arroja una pérdida de aproximadamente medio año. Dada la cantidad de datos con los que se cuenta, esta pérdida no es tan significativa (en torno al 5% de los datos originales).

Para la obtención de las estimaciones de los parámetros, de la matriz de varianzas y covarianzas y de la suma de cuadrados residuales, se confeccionó un programa que actualmente está en estado de prototipo.

Dicho software fue desarrollado en lenguaje Java en ambiente de desarrollo Eclipse utilizando además de las librerías nativas, la librería Jama para realizar las operaciones con matrices.

Procedimiento de estimación en la práctica

El modelo propuesto para ajustar los datos es relativamente nuevo y el método de estimación solo ha sido implementado en pocos paquetes econométricos, pero éstos no incluyen la estimación de parámetros relacionados con el comportamiento estacional de la variable, por eso la necesidad de crear un programa que se adecue a las características de esta serie.

Tomando en cuenta lo anterior, se desprende que este procedimiento es altamente experimental, por lo que se procedió a definir los parámetros iniciales, la cuantía de la actualización y el stopping rule del algoritmo de Gauss-Newton de manera ad-hoc, ya que los procedimientos

³ ver gráfico de la función SCR en función de d en anexo G

implementados para la estimación de los parámetros de modelos SARIMA no pueden ser seguidos en este caso.

El algoritmo de Gauss-Newton puede ser re-expresado como:

$$\hat{\beta}^k = \hat{\beta}^{k-1} + \text{actualización}$$

$$\text{actualización} = -\text{porcentajeactualización} \left(J_e \left(\hat{\beta}^k \right)^T J_e \left(\hat{\beta}^k \right) \right)^{-1} J_e \left(\hat{\beta}^k \right)^T e \left(\hat{\beta}^k \right)$$

Tanto la elección del valor inicial de los parámetros, como la mejor elección de la actualización, como el stopping rule tuvo varias etapas. En las próximas líneas se intentará describirlas.

La elección del valor inicial de los parámetros: 1) Parámetros elegidos arbitrariamente. Este procedimiento es muy trabajoso, porque hay que ingresar individualmente los $(p + P + q + Q)$ datos y no siempre se llega a la convergencia, ya que, dado que la forma de la función es altamente no lineal, ésta presenta muchos mínimos locales, que “atraen” a los valores de los parámetros y los desvían de la senda de convergencia. Estos mínimos no permiten que se llegue a una región estable, por lo que se debía comenzar nuevamente. 2) Tomando en cuenta esta característica, se hacen 2 cambios.

Primero, la elección de los parámetros es pseudoaleatoria, tomando una variable pseudoaleatoria $X \sim U[0,1)$, en algunos casos modificada. Para la elección de los parámetros ϕ_{i1} , ϕ_{i2*7} iniciales, se generan 2 vectores (de dimensión p y P respectivamente) y luego se los ordena de mayor a menor. Para los parámetros θ_{i3} , θ_{i4*7} se toma la variable $Y=2X-1$ de manera que los valores iniciales de estos parámetros oscilen entre 1 y -1. Segundo, para intentar eliminar la posibilidad de que se comience en parámetros que determinen la no estacionariedad, o no invertibilidad del proceso, se sigue el procedimiento anterior, pero se elimina el vector generado en caso de que la suma de los elementos del mismo sea mayor o igual a 1.

La elección de la cuantía de la actualización: El segundo componente del algoritmo también pasó por varias etapas. 1) el porcentaje de actualización es 50%, valor fijo independientemente de la iteración y valor de la suma de cuadrados residuales, luego este valor se fijó en 10% porque se observaba que la actualización en algunas circunstancias era demasiado elevada y se sobrepasaba el mínimo y hasta inclusive se llegaba a pasar al entorno de otro mínimo o inclusive se llegaba a casos donde el algoritmo divergía. En última instancia se lo deja a elección del investigador en cada caso, pero es este último porcentaje el que dio mejores resultados. 2) luego de observar la cantidad de veces que el algoritmo no llegaba a un punto válido (valores codificados por Java como NaN-not a number-) se observa que esta divergencia está asociada al hecho de que el valor de m_2 y/o g_2 sobresale demasiado de la región que hace que la variable sea estacionaria e invertible ($\max(m_2+g_2, g_2)$ no pertenece al intervalo $(-\frac{1}{2}, \frac{1}{2})$). Es por esta razón que cuando en una iteración dada, cualquiera de los mencionados parámetros toma valores menores a -1, se les asigna el valor -1 y se les asigna el valor 1 si toman valores mayores que 1, de esta manera se minimizan las probabilidades de que el algoritmo se desvíe de la senda de convergencia. En muchos casos este “callejón” hace que el algoritmo arroje valores “NaN” y se empiece el proceso iterativo nuevamente ahorrando decenas de iteraciones que en la amplia mayoría de los casos daría valores no aceptables.

Stopping rule: tal como las dos otras partes ésta pasó por diferentes estadios. 1) el algoritmo deja de iterar cuando la diferencia entre la scr^k y la scr^{k-1} es menor a 0.001 en valor nominal. 2) se cambia esta regla por la de que la diferencia **porcentual** entre las dos scr explicadas anteriormente sea de 1%. 3) la elección de este porcentaje queda a criterio del investigador, pero 1% es el valor que mejores resultados arrojó.

Obtención de los resultados

El procedimiento antes descrito genera un archivo log de tipo txt que deja un registro de los valores de convergencia de las distintas inicializaciones del algoritmo. De esta manera se deja que el programa inicialice por lo menos 100 estimaciones y luego, mediante la inspección del archivo, se elige para cada set de valores (p, q, P, Q) la scr mas baja y se analiza el comportamiento de sus residuos en cuanto a sus autocorrelaciones. Luego de haber obtenido la mejor estimación de cada modelo (p,q)(P,Q) se elige aquel con menor valor del criterio de Akaike.

En lo que refiere a la elección del mejor modelo SARIMA, el procedimiento es un tanto mas sencillo ya que gretl retorna una única estimación de los parámetros para cada modelización.

El modelo que posea menor valor del criterio de Akaike será el modelo elegido. Posteriormente se le realizan las pruebas de validación correspondientes para asegurarse de que cumple con todas las condiciones (hipótesis) requeridas.

4) Consideraciones acerca del consumo de electricidad en Uruguay

En este apartado se mencionarán algunos aspectos relevantes para la modelización de la serie de consumo de energía eléctrica en Uruguay, como ser las cualidades de esta energía, los factores que influyen su consumo, la importancia de una correcta previsión del consumo y cierra el capítulo una breve descripción del mercado eléctrico uruguayo.

Cualidades de la energía eléctrica

- a) **Transporte instantáneo para consumo inmediato:** cuando se prende un electrodoméstico o una lámpara de luz, inmediatamente se está demandando energía, que es transportada a la velocidad de la luz hasta el cliente.
- b) **Imposibilidad de almacenamiento:** la corriente alterna no puede ser almacenada en baterías como la corriente continua, lo que hace que los desfases entre oferta y demanda no se traduzcan en variación de existencias, sino que se resuelven “por el lado corto del mercado”.
- c) **Transporte por redes fijas:** Solo puede transportarse por redes (conductores) dimensionadas de acuerdo a previsiones de largo plazo.
- d) **Ausencia de sustitutos:** en muchos casos directamente no existen sustitutos (PC, TV, etc.) y en otros los sustitutos son muy imperfectos, iluminación y refrigeración por ejemplo.

Factores que influyen en la demanda de energía eléctrica

Podemos desagregar los factores como lo hace Gutiérrez Moya[16], por sector o lo podemos hacer por horizonte temporal.

Por sector, tenemos el sector residencial, el comercial y el sector industrial.

En el sector residencial, los factores que más influyen en la demanda son número de clientes, precio de los electrodomésticos, grado de penetración del aire acondicionado, número de personas por vivienda, y variables climáticas, entre las que se menciona la temperatura como la más importante.

Para el sector comercial, mencionaré el espacio ocupado por oficinas, el número de personas empleadas en el sector servicios, el nivel de consumo de los ciudadanos, el día de la semana, etc. Por último, en el sector industrial, los factores más influyentes son el nivel de producción industrial, su empleo de mano de obra, la normativa medio ambiental, el día de la semana, entre otros.

En tanto, por horizonte temporal, podemos mencionar el precio relativo de la electricidad con respecto a otros recursos energéticos que tiene relevancia a mediano y largo plazo, porque la capacidad instalada de equipos eléctricos está dada en el corto plazo y el ingreso de la población por las mismas razones expuestas anteriormente, los consumidores solo pueden ajustar la intensidad con la que utilizan los equipos eléctricos, pero no pueden cambiar de equipamiento en el corto plazo.

O lo podemos hacer como Arrillaga, Rosas y Sibille[2], quines lo hacen por tipo de variable.

Variables climáticas: el uso que más energía eléctrica utiliza es el de acondicionamiento del ambiente, por esto me centro en las variables que influyen sobre la sensación térmica.

*“Como es sabido, la **temperatura del aire exterior** no siempre es un indicador seguro y digno de confianza para determinar el frío que una persona puede sentir, si está expuesta al aire libre. Existen otros parámetros meteorológicos que influyen como la **velocidad del viento**, la **radiación** y la **humedad relativa**. El término sensación térmica es usado para describir el grado de incomodidad que un ser humano siente, como resultado de la combinación de la temperatura y el viento en invierno y de la temperatura, la humedad y el viento en verano”.*⁴

A las variables expresadas anteriormente, se debería agregar la **precipitación**, ya que es importante para la determinación del “nivel de incomodidad” que una persona puede sentir sobre todo en invierno y para quitar esta sensación utiliza electrodomésticos (lava-secarropas, termofón, estufas, bombas de achique, etc)

Debería agregarse también la **insolación** como factor importante en la decisión de utilizar energía eléctrica para la iluminación de ambientes y exterior.

Efecto calendario: en el consumo de energía eléctrica, la utilización se da con diferente intensidad dependiendo del día de la semana, días feriados (laborables y no laborables) y próximos a feriados (fines de semana largos por ejemplo). Este efecto se debe principalmente a la disminución de la actividad que se da en los sectores comercial e industrial.

Grado de urbanización: el acervo de electrodomésticos se da principalmente en los centros más poblados, por esto se encuentra una correlación positiva entre el grado de urbanización y el consumo de electricidad.

Variables monetarias y socioculturales: la teoría clásica del consumidor relaciona positivamente el consumo personal con su ingreso, con el precio de los bienes sustitutos y negativamente con el precio del producto en cuestión y con el precio de los productos complementarios.

Cuando se mencionan variables socioculturales se refiere a las costumbres y los usos, por ejemplo culinarios o de elección de fuente de energía para la calefacción.

Importancia de la estimación y previsión del consumo de energía eléctrica

Muchas son las razones esgrimidas para fundamentar la importancia de buenas estimaciones y predicciones del consumo de energía eléctrica a diferentes horizontes, entre ellas, para la optimización de las alternativas de *generación*, una eficiente *política tarifaria* y la correcta *planificación de la red*.

Este trabajo utiliza series diarias y por lo tanto las predicciones tendrán más validez en el corto plazo (la política tarifaria y la planificación de la red son tareas que requieren predicciones de largo plazo), por lo que el tipo de modelos se podrá utilizar de la mejor manera en la óptima elección de recursos para la generación y/o compra de energía.

En Uruguay el despacho nacional de cargas (DNC) es el encargado de entregar la energía con una tensión y frecuencia adecuada al menor costo posible. Para cumplir con este objetivo estima la curva de carga⁵ diaria para optimizar los recursos en la generación de energía. Según W. Labys (1999) “Si las predicciones resultan ser demasiado bajas pueden tener lugar carencias de energía cuyos costes habitualmente son mucho mayores que el valor de la energía no

⁴Información correspondiente a la página: <http://teleformacion.edu.aytolacoruna.es/FISICA/document/fisicaInteractiva/Calor/Temperatura/SensacionTermica.htm>

⁵ función que relaciona el consumo de energía con el tiempo (en minutos, horas, días, etc.).

suministrada. Por el contrario, si las previsiones resultan demasiado altas, los costes de oportunidad pueden ser muy elevados al tener comprometidos, de forma improductiva, cuantiosos fondos económicos durante largos períodos de tiempo”.

Las fuentes de generación de energía de las que se nutre Uruguay son en orden ascendente en relación a su costo: generación hidráulica propia, compra a Comisión técnico mixta de Salto Grande (CTM), importaciones desde Argentina y Brasil y generación térmica propia (a fuel oil y gasoil).

Es por estos motivos que el DNC corrobora que la generación hidráulica sea suficiente para abastecer la demanda y si considera (a cierto nivel de confianza) que no va a ser suficiente, entonces comienza a abastecerse de las otras fuentes hasta el caso extremo de prender las centrales térmicas.

Dada la diferencia de costos en la generación (sobre todo de la generación térmica), se hace importante mejorar los intervalos de predicción, para una correcta elección de las fuentes de aprovisionamiento de la energía eléctrica.

Para mostrar la importancia de la precisión en la predicción de la demanda diaria de energía eléctrica, Bunn y Farmer demuestran que para 1984, un incremento de un 1% en el error de predicción, causa un aumento de costos promedio de 10 millones de libras anuales. Este trabajo fue realizado para un solo operador eléctrico en Inglaterra.

Descripción del sistema eléctrico uruguayo

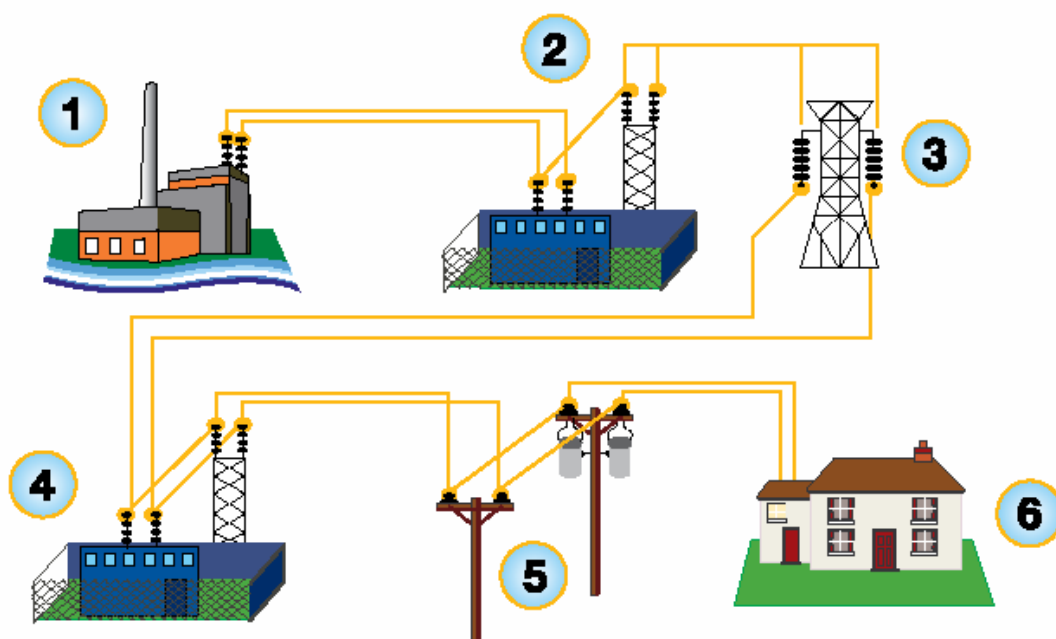
Un sistema eléctrico se compone de tres etapas:

Generación: se encarga de convertir otros tipos de energía en energía eléctrica y de transformarla a una tensión adecuada para su transmisión.

Transmisión: se lleva la energía a alta tensión hasta las zonas de consumo, donde se entrega la energía a distribución

Distribución: Es la encargada de alimentar el consumo final.

En el siguiente gráfico se observa la dinámica del sistema eléctrico



Fuente: Edison Electric Institute

La electricidad es producida por la planta generadora (1), que la entrega a una subestación elevadora de voltaje (2). Luego es transmitida hasta la zona de consumo (3), donde entra en una subestación reductora (4) que la entrega a baja tensión a la distribución por las líneas urbanas (5) hasta que llega al consumo final (6).

Marco regulatorio

Antes de la aprobación de la ley 16.832 de junio de 1997, las tres etapas del sistema eléctrico uruguayo eran llevadas a cabo por la empresa estatal UTE de manera monopólica. Pero con esta ley, se materializa el nuevo marco regulatorio del sistema eléctrico nacional, que excluye del concepto de servicio público la generación de energía eléctrica con destino a su comercialización, de esta manera se termina el monopolio que UTE ostentaba en esta parte del sistema eléctrico, pero no se termina el monopolio en las fases de transmisión y distribución.

Además, se crea el mercado mayorista, donde generadores y grandes consumidores realizan transacciones. Este mercado lejos está de ser de competencia perfecta, por lo que se crea la Administración del mercado eléctrico (ADME) que es la encargada de regular el mercado, realizar el predespacho del sistema, priorizar las operaciones y determinar que centrales deben entrar a operar para cubrir el consumo.

En este marco regulatorio los actores del mercado mayorista pueden comprar y vender energía en el mercado spot o en el mercado de contratos (cantidad y precio preacordado). Pero la decisión de qué generador produce la energía la hace el DNC de acuerdo a criterios económicos. Las diferencias entre el contrato y el despacho se compensan en el mercado spot.

Como fue expresado anteriormente, UTE compete por la generación de energía, pero ésta debe ser llevada hasta el consumidor final, que puede ser un “consumidor regulado” que no accede a la compra directa de un generador o un “gran consumidor”. Como las tareas de transmisión y distribución son mantenidas bajo la administración de UTE, éste cobra un peaje si otros agentes desean transportar energía por sus redes hasta un “gran consumidor” o exportar.

5) Resultados

Estimación

Los resultados que se presentarán a continuación fueron obtenidos con los siguientes programas:
 Estimación del mejor modelo SARIMA: gretl mediante el filtro de Kalman
 Estimación del mejor modelo SARFIMA: software hecho a medida para esta monografía
 Predicción: en ambos casos se llevó a cabo la predicción con el programa antes mencionado, utilizando los parámetros del mejor modelo estimado.

Las estimaciones y predicciones fueron hechas en un PC con procesador Pentium IV de 1.8 GHZ y 512 MB de memoria.

Las estimaciones de los modelos SARIMA tomaron en promedio 5:30 minutos y las estimaciones de los modelos SARFIMA 10 minutos.

Estimación SARMA(3,3)(1,1)₇ utilizando las 3552 observaciones 97/01/09-06/09/30
 Estimado usando el filtro de Kalman (MV exacta)
 Variable dependiente: $(1-L)(1-L^7)$ Consumo

Variable	Coeficiente	Desv. típica	Estadístico t	valor p	
phi_1	0.860556	0.129349	6.6530	<0.00001	***
phi_2	0.543448	0.203338	2.6726	0.00753	***
phi_3	-0.452207	0.085094	-5.3142	<0.00001	***
Phi_7	0.125637	0.019457	6.4572	<0.00001	***
theta_1	-1.11345	0.124414	-8.9495	<0.00001	***
theta_2	-0.553621	0.225003	-2.4605	0.01387	**
theta_3	0.681887	0.108049	6.3109	<0.00001	***
Theta_7	-0.991019	0.00282987	-350.1994	<0.00001	***

Media de la var. dependiente = -8.1684e-005

Desviación típica de la var. dependiente. = 0.0508172

Media de las innovaciones = -0.00073337

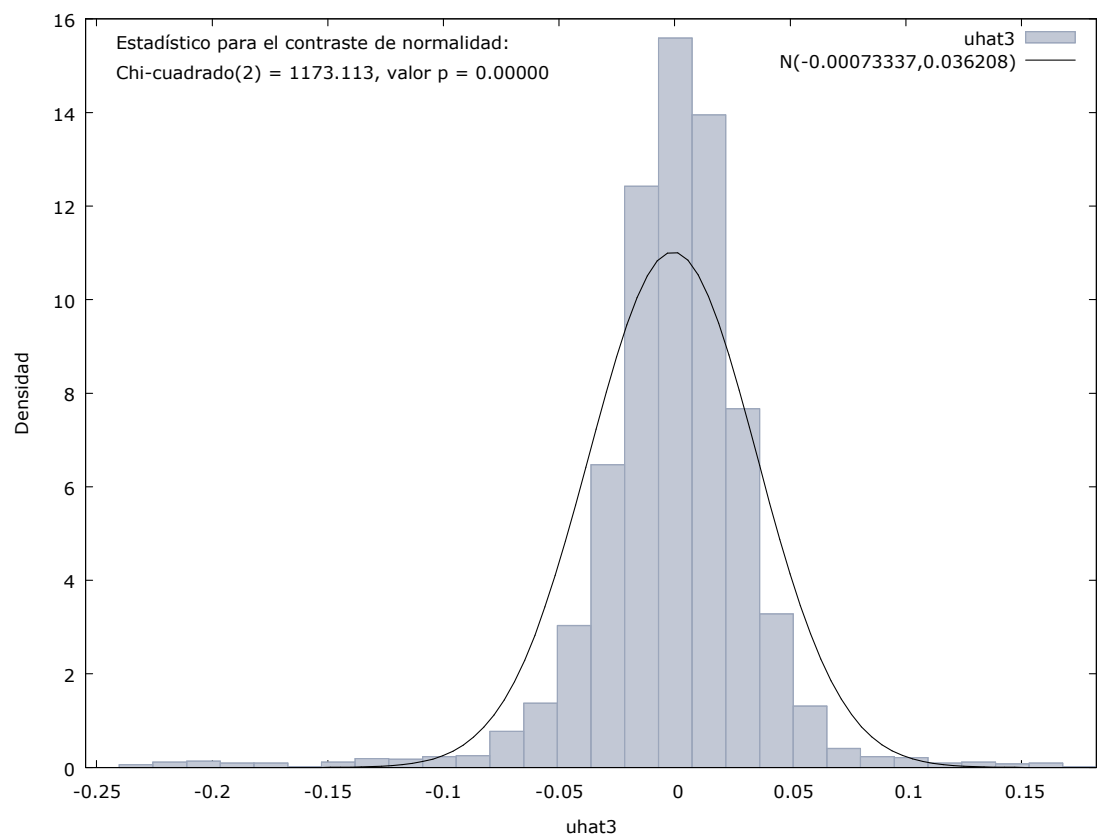
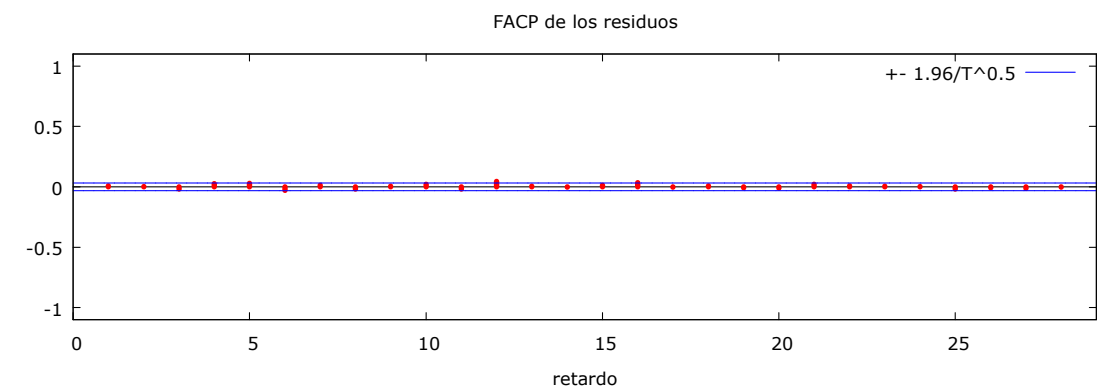
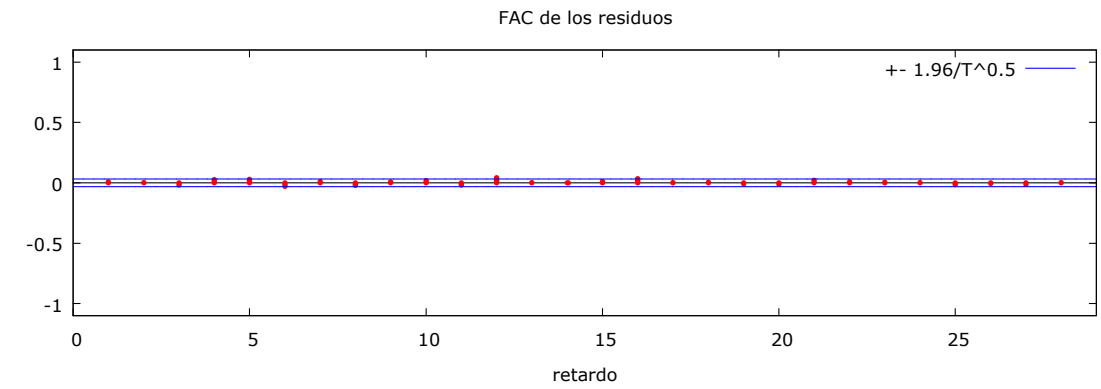
Scr = 4.57536672

Varianza de las innovaciones = 0.00128811

Log-verosimilitud = 6778.46308

Criterio de información de Akaike = -6.65486327

		Real	Imaginaria	Módulo	Frecuencia
AR					
	Raíz 1	1.4760	0.0000	1.4760	0.0000
	Raíz 2	1.0945	0.0000	1.0945	0.0000
	Raíz 3	-1.3688	0.0000	1.3688	0.5000
AR (estacional)					
	Raíz 1	7.9594	0.0000	7.9594	0.0000
MA					
	Raíz 1	1.0578	0.0781	1.0606	0.0117
	Raíz 2	1.0578	-0.0781	1.0606	-0.0117
	Raíz 3	-1.3036	0.0000	1.3036	0.5000
MA (estacional)					
	Raíz 1	1.0091	0.0000	1.0091	0.0000



Pruebas de hipótesis

Sobre estacionariedad e invertibilidad del proceso

En este caso el orden de algunos polinomios es superior a 2, lo que hace difícil la obtención de la condición de estacionariedad y/o invertibilidad. Lo que puede observarse en estos casos es que las raíces estimadas de los polinomios AR y MA son superiores a la unidad lo que indica que el proceso sería estacionario e invertible.

Las pruebas de hipótesis que pueden llevarse a cabo en este caso son acerca de las raíces AR y MA estacionales.

$$\begin{array}{ll} H_0) |\phi_7| = 1 & H_0) |\theta_7| = 1 \\ H_1) |\phi_7| < 1 & H_1) |\theta_7| < 1 \end{array}$$

Estas hipótesis se llevan a cabo mediante la construcción de intervalos de confianza al 95% de confianza. Ninguna de las dos hipótesis puede ser rechazada, aunque la segunda está muy cerca al valor que haría que el proceso fuera no invertible.

Sobre las perturbaciones

- *Prueba de normalidad*

Para la realización de esta prueba se utiliza el estadístico de Jarque-Bera

$$H_0) \varepsilon_t \sim Normal \quad JB = \frac{n}{6} \left(S^2 + \frac{(K-3)^2}{4} \right) \sim \chi^2_2$$

$$H_1) \varepsilon_t \not\sim Normal$$

El valor del estadístico se presenta en el histograma de los errores y se desprende de allí que se rechaza la hipótesis nula.

Este resultado sin embargo, no debe ser tomado como concluyente, ya que se observa que los errores mantienen la forma de una densidad normal y otra sería la conclusión si se hubiera hecho análisis de intervención.

- *Prueba de incorrelación serial*

El estadístico que calcula el programa gretl es el de Box-Pierce

$$H_0) \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_M \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad Q = T \sum_{i=1}^M \widehat{\rho_i^2} \sim \chi^2_{M-k}$$
$$H_1) \text{algún } \rho_i \neq 0$$

donde k es la cantidad de parámetros del modelo.

El valor crítico para esta prueba es $\chi^2_{28-8}^{2(0.95)} = 31.41$ y el valor del estadístico en la muestra es 30.77, por lo que no se rechaza la hipótesis nula con M=28 y k=8.

Modelo SARFIMA estimado					
N=3384		Variable dependiente: (1-L ⁷)LnConsumo			
Estimaciones SARFIMA(1,2)(1,1) utilizando las 3552 observaciones 97/01/09-06/09/30					
Variable	Coeficiente	desv est	estadístico t	P-valor	Vza
m ₂	0.48337272	0.07192483	6.720526639	1.82081E-11	0.00517318
g ₂	-0.33233353	0.06815976	-4.875802841	1.08518E-06	0.00464575
phi_1	0.71261128	0.28114415	2.534682912	0.011254958	0.07904204
Phi_7	0.38114386	0.08399658	4.537611719	5.69497E-06	0.00705543
theta_1	-0.42660776	0.3276879	-1.301872147	0.192960225	0.10737936
theta_2	-0.20305019	0.08467067	-2.39811724	0.016479575	0.00716912
Theta_7	-0.82624674	0.04096246	-20.1708257	0.000000000	0.00167792

Media de la var. dependiente = $-8.1684e-005$

Desviación típica de la var. dependiente. = 0.0508172

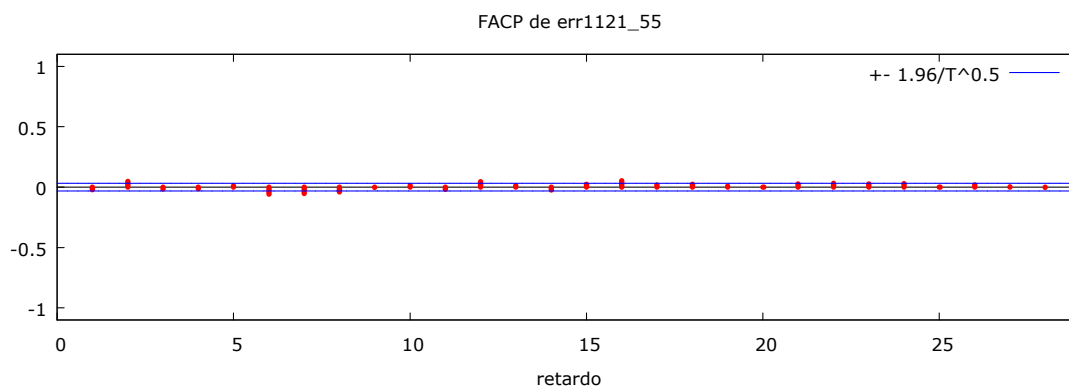
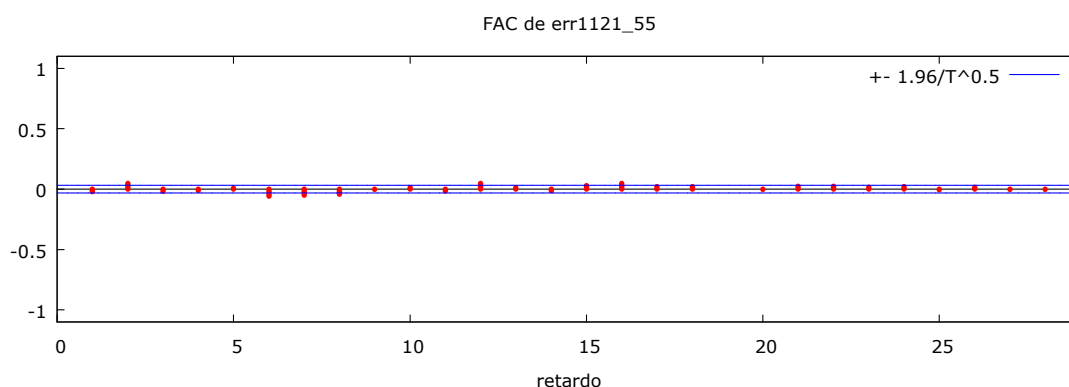
Media de las innovaciones = 0.000406

Scr = 4.50153000

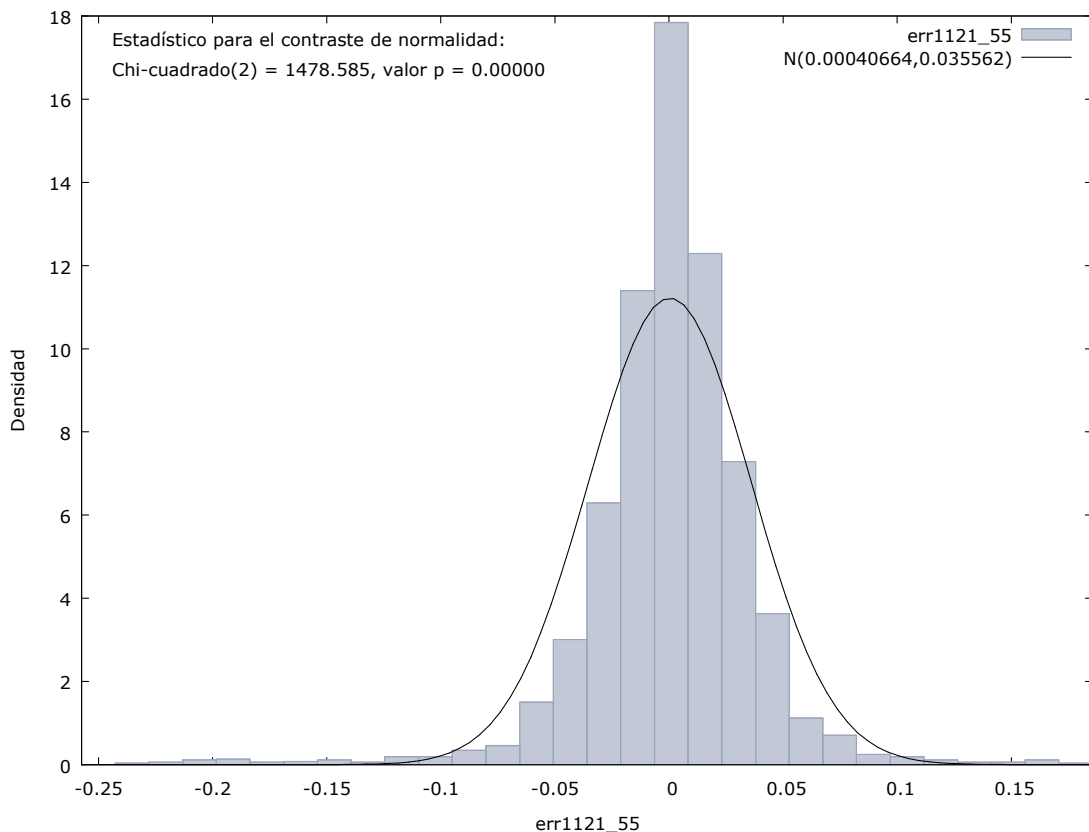
Varianza de las innovaciones = 0.00133024

Log-verosimilitud = 6317.38142

Criterio de información de Akaike⁶ = -6.6714136



⁶ el cálculo del AIC en ambos casos fue calculado a partir de la fórmula $AIC = \ln\left(\frac{e'e}{T}\right) + \frac{K}{T}$



Pruebas de hipótesis

Sobre estacionariedad e invertibilidad del proceso

- Estacionariedad e invertibilidad asociada a los parámetros m_2 y g_2 :

$$H_0) |m_2 + g_2| < \frac{1}{2}$$

$$H_1) |m_2 + g_2| \geq \frac{1}{2}$$

$$H_0) |g_2| < \frac{1}{2}$$

$$H_1) |g_2| \geq \frac{1}{2}$$

Para ello se realizó un intervalo de confianza para $m_2 + g_2$ y para g_2 al 95% de confianza. Del análisis de estos intervalos se concluye que no se pueden rechazar ninguna de las 2 hipótesis, por lo que el proceso en cuestión (Δ_7 Ln Consumo) sería estacionario e invertible.

ic para m_2+g_2			ic para g_2		
m_2+g_2 est	límite inf	límite sup	g_2 est	límite inf	límite sup
0.151039186	-4.20E-02	3.44E-01	-0.332333533	-0.46592665	-0.198740416

- Estacionariedad e invertibilidad asociada a las raíces de los polinomios AR y MA.

Es contrastada mediante pruebas de hipótesis que verifican que las raíces de los polinomios MA y SMA son mayores que la unidad en módulo.

$$H_0) |\phi_1| < 1 \quad H_0) |\phi_7| < 1 \quad H_0) |\theta_7| < 1 \quad H_0) |\theta_2| < 1$$

$$H_1) |\phi_1| \geq 1 \quad H_1) |\phi_7| \geq 1 \quad H_1) |\theta_7| \geq 1 \quad H_1) |\theta_2| \geq 1$$

$$\begin{array}{ll} H_0) \theta_2 + \theta_1 < 1 & H_0) \theta_2 - \theta_1 < 1 \\ H_1) \theta_2 + \theta_1 \geq 1 & H_1) \theta_2 - \theta_1 \geq 1 \end{array}$$

Todas estas hipótesis se probaron con los siguientes intervalos de confianza

ic para phi_1			ic para Theta_7		
phi_1 est	límite inf	límite sup	Theta_7 est	límite inf	límite sup
0.712611284	0.16156874	1.263653827	-0.826246735	-0.90653317	-0.745960305

ic para theta_1			ic para phi_7		
theta_1 est	límite inf	límite sup	Phi_7 est	límite inf	límite sup
-0.426607756	-1.06887605	0.215660537	0.381143864	0.21651057	0.54577716

ic para theta_2		
theta_2 est	límite inf	límite sup
-0.203050191	-0.3690047	-0.03709568

ic para thita_2 -thita_1			ic para thita_2+thita_1		
thita_2 -thita_1 est	límite inf	límite sup	thita_2+thita_1 est	límite inf	límite sup
0.223557565	-4.41E-01	8.88E-01	-0.629657947	-1.3E+00	3.28E-02

La única hipótesis que no se puede rechazar es la que involucra a phi_1, por lo que no se puede afirmar, con un 95% de confianza que el proceso sea estacionario porque podría contener una raíz unitaria en el polinomio AR regular.

En lo que respecta a las raíces de los restantes polinomios, puede afirmarse con el mismo nivel de confianza que no contienen raíces unitarias.

Sobre las perturbaciones

- *Prueba de normalidad*

Las conclusiones sobre la normalidad de las perturbaciones son idénticas a las obtenidas en el caso de los errores obtenidos del modelo SARIMA.

- *Prueba de incorrelación serial*

$$H_0) \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_M \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad Q = T \sum_{i=1}^M \widehat{\rho}_i^2 \sim X_{M-k}^2$$

$$H_1) \text{algún } \rho_i \neq 0$$

El valor crítico para esta prueba es $X_{28-7}^{2(0.95)} = 32.67$ y el valor del estadístico en la muestra es 72.065 por lo que se rechaza la hipótesis nula con $M=28$ y $k=7$.

Si bien se rechaza la hipótesis de incorrelación serial de las perturbaciones, puede apreciarse en el correlograma que el valor de las correlaciones no es muy elevado (raras veces supera el 4%).

Debe acotarse que el modelo anterior cuenta con la ventaja de que las correlaciones de los primeros lags fueron “espuriamente” disminuidas dado que el modelo está sobrediferenciado y la aplicación de la doble diferenciación (regular y estacional) elimina ciertos componentes del proceso.

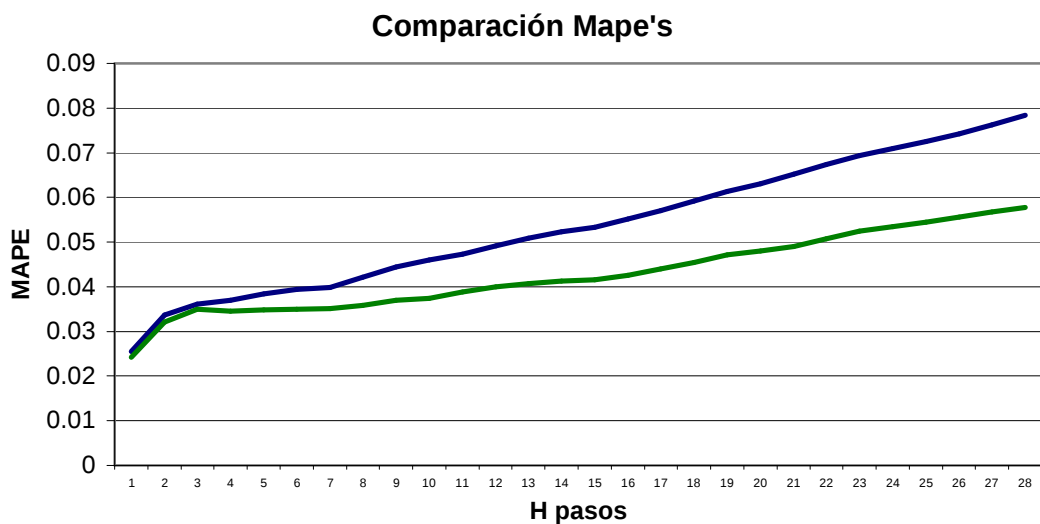
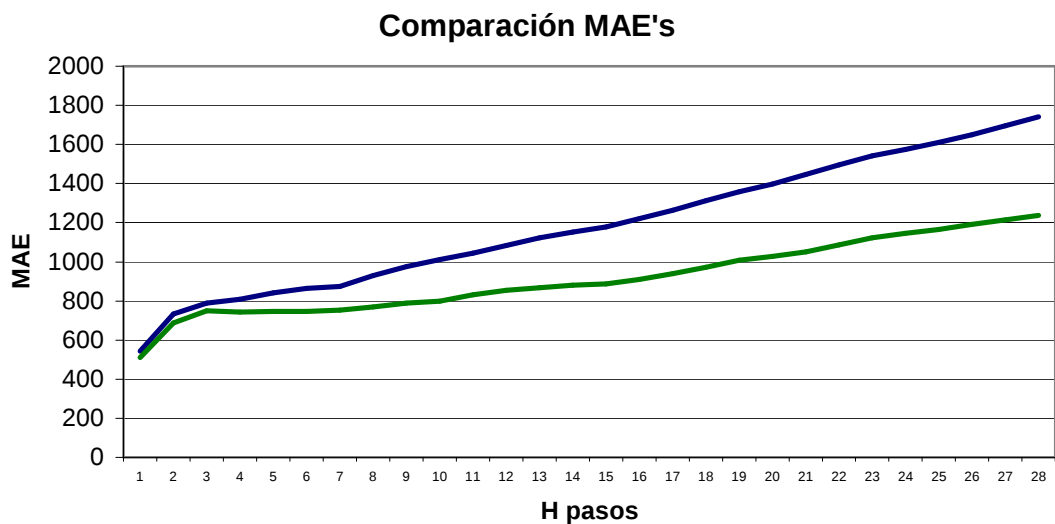
Predicción

Para comparar el poder predictivo de ambas modelizaciones se calcularon los MAE's y los MAPE's de ambos modelos para el período 1/10/2006 a 31/12/2006 y las predicciones fueron hechas hasta h=28.

$$MAE_h = \frac{\sum_{i=1}^n \left(|Y_t - \widehat{Y}_{t/h}| \right)}{n}$$
$$MAPE_h = \frac{\sum_{i=1}^n \left(\frac{|Y_t - \widehat{Y}_{t/h}|}{|Y_t|} \right)}{n}$$

$\widehat{Y}_{t/h}$ significa la predicción para el período t hecha con información disponible hasta h días antes.

Los resultados siguientes pertenecen en verde al indicador del mejor modelo SARFIMA y en azul al mejor modelo SARIMA:



A continuación es interesante visualizar las diferencias entre las estimaciones de las 2 modelizaciones.

	m1	m2	g1	g2
SARIMA	1	0	1	0
SARFIMA	0	0.48	1	-0.33

	phi_1	phi_2	phi_3	theta_1	theta_2	theta_3	Phi_7	Theta_7
SARIMA	0.86	0.54	-0.45	-1.11	-0.55	0.68	0.12	-0.99
SARFIMA	0.71			-0.42	-0.20		0.38	-0.82

	$\bar{e}$	$\hat{\sigma}_{\varepsilon}^2$	log verosimil	AIC
SARIMA	-0.000733	0.00128	6778.46308	-6.65486327
SARFIMA	0.000406	0.00133	6317.38142	-6.6714136

6) Conclusiones

Luego de analizar los resultados se desprenden dos tipos de conclusiones:

Desde el punto de vista de la correcta diferenciación de la serie y desde el punto de vista de la bondad de predicción.

A) Desde el punto de vista de la correcta especificación

En el capítulo concerniente a las características de los modelos SARFIMA, se hace referencia a las características que un proceso debe poseer para que sea estacionario e invertible. Luego la sección referida a la estimación, menciona que debemos suponer que la perturbación distribuye normal para poder estimar los parámetros por el método de mínimos cuadrados y que las propiedades de los mismos coincidan con las de los estimadores máximo verosímiles. Para suponer este tipo de distribuciones es necesario que estemos trabajando con series estacionarias, es por este motivo que se hace tan importante la identificación de los parámetros m_1 y g_1 en una primera etapa, para que después no surjan problemas en el momento de la validación del modelo. Si bien es importante no trabajar con series poco diferenciadas que tienen la consecuencia mencionada anteriormente, también es muy importante no trabajar con series sobrediferenciadas, ya que la sobrediferenciación tiene consecuencias sobre la varianza del proceso y en última instancia sobre los intervalos de confianza de los estimadores como de los valores predichos de la variable dependiente.

En definitiva debemos analizar cuales son los valores de los parámetros d y D para clasificar correctamente el proceso que generó la serie $\ln\text{Consumo}$.

Del análisis precedente surge que $\hat{d} = m_1 + \hat{m}_2 = 0 + 0.48337272 = 0.48337272$
 $\hat{D} = g_1 + \hat{g}_2 = 1 - 0.33233353 = 0.66766647$. Tanto el valor de m_1 como el de g_1 no llevan “gorro” porque no fueron estimados, sino seleccionados de manera que la serie cumpla con las características requeridas.

De lo anterior se desprende que el proceso en cuestión puede ser catalogado como de memoria larga que no retorna a la media y tiene memoria infinita.

A su vez fue demostrado que los parámetros m_2 y g_2 son diferentes a cero, lo que prueba que el correcto tratamiento de este proceso debe hacerse tomando en cuenta la existencia de raíces fraccionales. También se observa que el modelo que permite la diferenciación fraccional es más parsimonioso que el mejor modelo SARIMA, lo que es un buen indicador de la relevancia de este tratamiento de las series.

B) Desde el punto de vista de la bondad de predicción

Tal como se observa en las gráficas de MAE y de MAPE expuestas en el capítulo resultados, no solamente las predicciones hechas con modelos SARFIMA presentan menores errores absolutos y relativos, sino que estos errores son menores en términos relativos (SARFIMA con respecto a SARIMA), a medida que aumenta el paso (h); lo que debe verse como una consecuencia de no eliminar la información contenida en las frecuencias bajas, asociadas al largo plazo que generalmente son eliminadas para que el correlograma represente lo que usualmente identificamos como perteneciente a una serie estacionaria.

Es útil en este punto citar a los precursores de este tipo de modelización. Dicen Granger y Joyeux[13]...”modelos que permiten d fraccionales, no necesariamente presentan mejores predicciones de corto plazo, pero pueden presentar mejores predicciones a largo plazo, donde la modelización de la información contenida en las frecuencias bajas es vital”.

Es por esto que las predicciones de corto plazo son similares en los dos modelos, pero conforme nos alejamos del último valor conocido, la calidad de las predicciones que surgen del modelo SARFIMA mejora en relación a las del modelo SARIMA.

Estos resultados pueden apreciarse en las siguientes tablas:

MAPE	h=1	h=2	h=3	h=4	h=5	h=6	h=7
SARIMA 3 3 1 1	0.0255	0.0337	0.0362	0.0369	0.0384	0.0395	0.0398
SARFIMA 1 2 1 1	0.0243	0.0320	0.0349	0.0346	0.0348	0.0349	0.0351

MAPE	h=8	h=9	h=10	h=11	h=12	h=13	h=14
SARIMA 3 3 1 1	0.0422	0.0444	0.0460	0.0473	0.0490	0.0509	0.0523
SARFIMA 1 2 1 1	0.0359	0.0370	0.0373	0.0388	0.0400	0.0407	0.0412

MAPE	h=15	h=16	h=17	h=18	h=19	h=20	h=21
SARIMA 3 3 1 1	0.0534	0.0552	0.0571	0.05921	0.0613	0.0630	0.0652
SARFIMA 1 2 1 1	0.0415	0.0425	0.0439	0.0454	0.0471	0.0479	0.0491

MAPE	h=22	h=23	h=24	h=25	h=26	h=27	h=28
SARIMA 3 3 1 1	0.0674	0.0694	0.0709	0.0726	0.0743	0.0763	0.0785
SARFIMA 1 2 1 1	0.0508	0.0524	0.0535	0.0544	0.0556	0.0567	0.0578

7) Consideraciones

Acerca del orden de diferenciación:

En la modelización SARIMA: La elección de d y D es hecha en base a los correlogramas y al criterio de tomar aquella diferenciación de la serie que presente menor desvío estándar. Tanto el test ADF como HEGY⁷ concluyen que $d=D=0$, lo que puede significar 2 cosas: 1º los referidos test tienen poca potencia si los procesos que estudian presentan raíces fraccionales (particularmente cerca de los límites que marcan la región estacionaria), o la serie presenta una tendencia determinística, característica que no se puso a prueba.

En la modelización SARFIMA: la elección de m_1 y g_1 fue realizada en base a la validación. Es decir se probó con diferentes valores de m_1 y g_1 hasta que se llegó a validar el modelo elegido, (con $m_1=0$ y $g_1=1$) mediante el modelo que obtuviera el vector de error con las menores autocorrelaciones posibles y los valores de m_2 y g_2 que hicieran que el proceso fuera estacionario e invertible.

Acerca del algoritmo utilizado:

Nada asegura de que se haya llegado al mínimo global de la función scr , este algoritmo solamente asegura haber llegado a un mínimo. Como forma de disminuir la probabilidad de elegir un valor de $\hat{\beta}$ que diera un mínimo local de la función objetivo se dejó que el algoritmo se inicializara un mínimo de 100 veces y luego se inspeccionó el archivo de resultados en busca de la menor scr .

El algoritmo puede ser realizado de manera más eficiente, como por ejemplo mediante una mejor elección de β inicial y mejor elección de la magnitud del salto, que se adecue a la diferencia de scr entre una iteración y otra.

Acerca de la modelización elegida:

No elimina totalmente las autocorrelaciones de los primeros 28 lags. Se puede argumentar en este punto que el método de estimación no presenta recursos como el backcast de los términos MA. Este recurso podría mejorar sensiblemente los valores de las autocorrelaciones.

Se rechaza la hipótesis de normalidad de las perturbaciones. Este hecho puede deberse a que en este trabajo no se consideró el análisis de intervención de la serie, lo que hace que existan varios errores negativos que influyen el estadístico de simetría de la serie. En un posterior trabajo se podrían considerar además del análisis de intervención, la incorporación de variables explicativas fuertemente exógenas como las climáticas (temperatura principalmente).

En este trabajo no se permite la posibilidad de ocurrencia de un cambio estructural, como el que puede haber ocurrido en junio de 2002.

⁷ Ver anexo E

Anexo A) Nociones del concepto de diferenciación fraccional

Los conceptos fundamentales de diferenciación fraccional fueron desarrollados por Liouville en un artículo de 1832.

La derivada fraccional de una función de orden a es comúnmente definida por medio de integrales transformadas de Fourier o Mellin. Un punto importante es que la derivada de $f(x)$ es una propiedad local de la función solamente si es de orden entero, pero la derivada no entera de la misma función es una característica más global de la misma. Por lo que se necesitará mayor información acerca del comportamiento de la función, inclusive para valores alejados del valor de x en el cual se evalúa la derivada.

Una característica de esta rama de la matemática es que al día de hoy no se conoce la interpretación geométrica de la operación diferenciación cuando el orden es no entero.

Derivada $\frac{1}{2}$ de una función sencilla.

Definamos $f(x)$ como un monomio a la k -ésima potencia

$$f(x) = x^k .$$

La derivada primera tiene la fórmula usual

$$f'(x) = \frac{d}{dx} f(x) = kx^{k-1} .$$

Si repetimos el procedimiento obtendremos el resultado general

$$\frac{d^a}{dx^a} x^k = \frac{k!}{(k-a)!} x^{k-a} ,$$

En el cual, si sustituimos el operador factorial por la función gamma

$$\frac{d^a}{dx^a} x^k = \frac{\Gamma(k+1)}{\Gamma(k-a+1)} x^{k-a} .$$

Entonces, por ejemplo, la derivada $\frac{1}{2}$ de x queda:

$$\frac{d^{\frac{1}{2}}}{dx^{\frac{1}{2}}} x = \frac{\Gamma(1+1)}{\Gamma(1-\frac{1}{2}+1)} x^{1-\frac{1}{2}} = \frac{\Gamma(2)}{\Gamma(\frac{3}{2})} x^{\frac{1}{2}} = 2\pi^{-\frac{1}{2}} x^{\frac{1}{2}} = \frac{2x^{\frac{1}{2}}}{\sqrt{\pi}} .$$

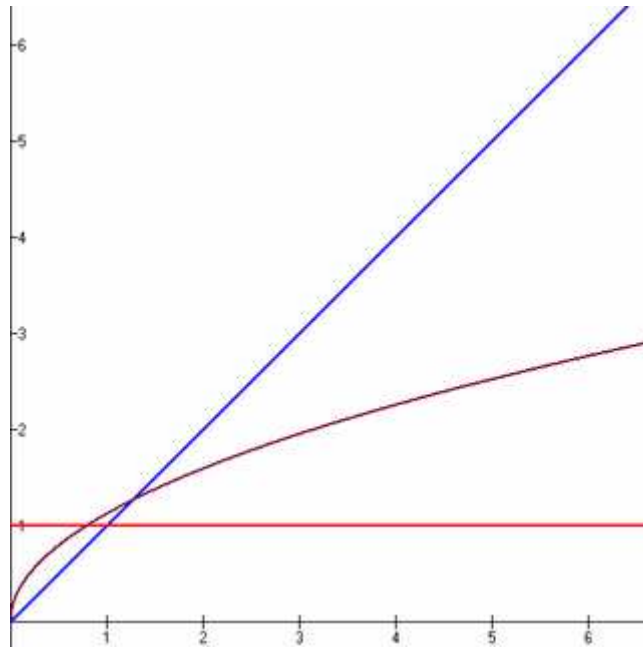
derivando nuevamente

$$\frac{d^{\frac{1}{2}}}{dx^{\frac{1}{2}}} 2\pi^{-\frac{1}{2}} x^{\frac{1}{2}} = 2\pi^{-\frac{1}{2}} \frac{\Gamma(1+\frac{1}{2})}{\Gamma(\frac{1}{2}-\frac{1}{2}+1)} x^{\frac{1}{2}-\frac{1}{2}} = 2\pi^{-\frac{1}{2}} \frac{\Gamma(\frac{3}{2})}{\Gamma(1)} x^0 = \frac{1}{\Gamma(1)} = 1$$

Obteniéndose el resultado esperado, ya que:

$$\left(\frac{d^{1/2}}{dx^{1/2}} \frac{d^{1/2}}{dx^{1/2}} \right) x = \frac{d}{dx} x = 1$$

Gráficamente:



La línea azul representa $f(x)=x$; la línea roja, su derivada primera y la marrón, la derivada $1/2$.

Aquellos interesados en profundizar en estos conceptos se les recomienda seguir con la lectura del trabajo de Miller y Ross[22].

Anexo B) Procedimiento para la construcción del programa

B.1) Módulo estimación

El procedimiento para la construcción del software que realiza la estimación fue el siguiente: Teniendo en cuenta que la estructura del proceso fue modelada de la forma:

$$\Phi_p(L)\Phi_p(L)(1-L)^d(1-L^7)^D Y_t = \Theta_q(L)\Theta_q(L)\varepsilon_t \quad (C.1)$$

Donde:

$$(1-L)^{m_2} = \sum_{j=0}^J C_j L^j \quad (C.2)$$

$$(1-L^7)^{g_2} = \sum_{s=0}^S F_s L^{7s} \quad (C.3)$$

Se tomaron los valores S=J=20.

Los polinomios AR, SAR, MA y SMA se re-definieron de la siguiente manera:

$$\left(-\sum_{i1=0}^p \phi_{i1} L^{i1}\right), \left(-\sum_{i2=0}^P \phi_{i2*7} L^{i2*7}\right), \left(\sum_{i3=0}^q \theta_{i3} L^{i3}\right), \left(\sum_{i4=0}^Q \theta_{i4*7} L^{i4*7}\right) \quad (C.4)$$

respectivamente con $\phi_0 = -1$ y $\theta_0 = 1$.

Por lo que la estructura queda expresada:

$$\left(\sum_{i1=0}^p \phi_{i1} L^{i1}\right) \left(\sum_{i2=0}^P \phi_{i2*7} L^{i2*7}\right) \left(\sum_{j=0}^J C_j L^j\right) \left(\sum_{s=0}^S F_s L^{7s}\right) Y_t = \left(\sum_{i3=0}^q \theta_{i3} L^{i3}\right) \left(\sum_{i4=0}^Q \theta_{i4*7} L^{i4*7}\right) \varepsilon_t \quad (C.5)$$

Tomando los valores estimados de los parámetros y despejando se define el error en t.

$$\left(\sum_{i1=0}^p \hat{\phi}_{i1} L^{i1}\right) \left(\sum_{i2=0}^P \hat{\phi}_{i2*7} L^{i2*7}\right) \left(\sum_{j=0}^J \hat{C}_j L^j\right) \left(\sum_{s=0}^S \hat{F}_s L^{7s}\right) Y_t - \left(\sum_{i3=0}^q \hat{\theta}_{i3} L^{i3}\right) \left(\sum_{i4=0}^Q \hat{\theta}_{i4*7} L^{i4*7}\right) e_t = e_t \quad (C.6)$$

o análogamente:

$$\sum_{i1=0}^p \sum_{i2=0}^P \sum_{j=0}^J \sum_{s=0}^S \hat{\phi}_{i1} \hat{\phi}_{i2*7} \hat{C}_j \hat{F}_s Y_{[t-(i1+7*i2+j+7*s)]} - \sum_{i3=0}^q \sum_{i4=0}^Q \theta_{i3} \theta_{i4*7} e_{t-(i3+i4*7)} = e_t \quad (C.6')$$

Exceptuando el caso de que $i_3=i_4=0$

La función que calcula los errores suma en i_1, i_2, j y s a $Y_{[t-(i1+7*i2+j+7*s)]}$, y en i_3 e i_4 a los errores pasados, por lo que no toma en cuenta aquellos errores anteriores al momento $p+7*P+j+7*s$.

Una función calcula los pesos $\hat{C}_j$ y $\hat{F}_s$ en función de los valores de $\hat{m}_2$ y $\hat{g}_2$.

Dados los valores de todos los parámetros, otra función calcula los e_t .

Luego se calcula la matriz jacobiana.

$$J_e(\hat{\beta}) = \begin{pmatrix} \frac{\partial e_1}{\partial \hat{m}_2} & \frac{\partial e_1}{\partial \hat{g}_2} & \frac{\partial e_1}{\partial \hat{\phi}_{i1}} & \frac{\partial e_1}{\partial \hat{\phi}_{7*i2}} & \frac{\partial e_1}{\partial \hat{\theta}_{i3}} & \frac{\partial e_1}{\partial \hat{\theta}_{7*i4}} \\ \frac{\partial e_2}{\partial \hat{m}_2} & \frac{\partial e_2}{\partial \hat{g}_2} & \frac{\partial e_2}{\partial \hat{\phi}_{i1}} & \frac{\partial e_2}{\partial \hat{\phi}_{7*i2}} & \frac{\partial e_2}{\partial \hat{\theta}_{i3}} & \frac{\partial e_2}{\partial \hat{\theta}_{7*i4}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial e_T}{\partial \hat{m}_2} & \frac{\partial e_T}{\partial \hat{g}_2} & \frac{\partial e_T}{\partial \hat{\phi}_{i1}} & \frac{\partial e_T}{\partial \hat{\phi}_{7*i2}} & \frac{\partial e_T}{\partial \hat{\theta}_{i3}} & \frac{\partial e_T}{\partial \hat{\theta}_{7*i4}} \end{pmatrix} \quad (C.7)$$

La matriz jacobiana tiene $p+P+q+Q+2$ columnas y $T-p+7*P+J+7*S+m_1+g_1$ filas. Se toman todas las derivadas excepto para θ_0 y ϕ_0 dado que no son parámetros.

En esta matriz de derivadas vale aclarar algunos puntos:

- 1) el subíndice 1 responde a la observación $p+7*P+J+7*S+m_1+g_1+1$ y así sucesivamente.
- 2) no se utiliza la función e_t sino la aproximación dada en (C.6').

El punto 2 permite que se simplifiquen las derivadas expresadas en función de $\hat{C}_j$ y $\hat{F}_j$.

Para el cálculo de esta matriz se diferenciaron 4 “familias” de derivadas:

(C.6') puede verse como: $A_t - B_t = e_t$.

$$\text{Derivadas en } \hat{\phi}: \frac{\partial e_t}{\partial \hat{\phi}_{i1,i2*7}} = \frac{\partial A_t}{\partial \hat{\phi}_{i1,i2*7}} = A_t \left(\hat{\phi}_{i1,i2*7} = 1 \right) \quad (C.8)$$

$$\text{Derivadas en } \hat{\theta}: \frac{\partial e_t}{\partial \hat{\theta}_{i3,i4*7}} = \frac{\partial B_t}{\partial \hat{\theta}_{i3,i4*7}} = B_t \left(\hat{\theta}_{i3,i4*7} = 1 \right) \quad (C.9)$$

$$\text{Derivada en } \hat{d}: \frac{\partial e_t}{\partial \hat{d}} = \frac{\partial A_t}{\partial \hat{d}} = A_t \left(\frac{\partial \hat{C}_j(\hat{d})}{\partial \hat{d}} \right) \quad (C.10)$$

$$\text{Derivadas en } \hat{D}: \frac{\partial e_t}{\partial \hat{D}} = \frac{\partial A_t}{\partial \hat{D}} = A_t \left(\frac{\partial \hat{F}_s(\hat{D})}{\partial \hat{D}} \right) \quad (C.11)$$

$$\text{siendo } \frac{\partial \hat{C}_0}{\partial \hat{m}_2} = \frac{\partial \hat{F}_0}{\partial \hat{g}_2} = 0 \quad (C.12)$$

De esa manera queda conformada la matriz jacobiana, su transpuesta y el vector e_t .

El primer error que se puede calcular es el correspondiente al momento $p+7*P+J+7*S+m_1+g_1+1$, aquellos errores previos a este momento se supuso que eran igual a su valor promedio, o sea 0.

El algoritmo se inicializó de la siguiente manera:

Los valores de los parámetros $\hat{\phi}'_t$ y $\hat{\theta}'_t$ se inicializaron con valores aleatorios de manera tal que la suma de los parámetros (en valor absoluto) sea menor que 1, de esta manera se minimiza la probabilidad de que se parta de puntos donde el proceso es no estacionario o no invertible. Y los valores de d y D se inicializaron con el vector nulo dado que la teoría anterior así lo imponía., o sea, propone que con un orden de diferenciación entero la serie queda $I(0)$.

B.2) Módulo predicción

Nuevamente se toma la ecuación C.6, pero en este caso se despeja Y_t .

$$Y_t = \left(\sum_{i3=0}^q \hat{\theta}_{i3} L^{i3} \right) \left(\sum_{i4=0}^Q \hat{\theta}_{i4*7} L^{i4*7} \right) e_t - \left(\sum_{i1=0}^p \hat{\phi}_{i1} L^{i1} \right) \left(\sum_{i2=0}^P \hat{\phi}_{i2*7} L^{i2*7} \right) \left(\sum_{j=0}^J \hat{C}_j L^j \right) \left(\sum_{s=0}^S \hat{F}_s L^{7s} \right) Y_t$$

(C.13)

Exceptuando el caso $i_1+i_2+s+j=0$

Esta estructura se cumple para todo t y si agregamos el supuesto de que no se producen cambios estructurales en los próximos n períodos, podremos tomar esta estructura como válida para predecir los siguientes h días, teniendo como datos desde el h -ésimo día anterior.

En este sentido se define la predicción a h días de la siguiente manera:

$$\begin{aligned} \widehat{E(Y_{T+h/T})} = & \left(\sum_{i3=0}^q \hat{\theta}_{i3} L^{i3} \right) \left(\sum_{i4=0}^Q \hat{\theta}_{i4*7} L^{i4*7} \right) e_{(T+h)/T} - \\ & - \left(\sum_{i1=0}^p \hat{\phi}_{i1} L^{i1} \right) \left(\sum_{i2=0}^P \hat{\phi}_{i2*7} L^{i2*7} \right) \left(\sum_{j=0}^J \hat{C}_j L^j \right) \left(\sum_{s=0}^S \hat{F}_s L^{7s} \right) (Y_{(T+h)/T}) \end{aligned}$$

Esta ecuación debe leerse el valor estimado esperado para el día $T+h$ dada la información hasta el día T .

Para ello utiliza los errores conocidos hasta el día T y a los errores futuros se les asigna su valor esperado (es decir 0). Un tratamiento diferente reciben las Y de momentos futuros, ya que se

sustituye $Y_{(T+h-i)/T}$ por $\widehat{E(Y_{(T+h-i)/T})} \forall h > i, i > 0$ que fue encontrado i pasos antes.

Anexo C) Pruebas de validación del programa a medida

Varias pruebas fueron corridas para asegurarse de que el programa estuviera haciendo lo que la teoría indicaba.

Para ello se estimó un modelo SARIMA (1,1,1)(1,1,1)₇ mediante el método de máxima verosimilitud con el método de Kalman. Su salida es presentada a continuación:

Modelo 1: estimaciones ARMA utilizando las 3552 observaciones 97/01/09-06/09/30
Variable dependiente: $(1-L)(1-L^7)\text{LnConsumo}$

Variable	Coeficiente	Desv. típica	Estadístico t	valor p	
phi_1	0.630234	0.0185542	33.9672	<0.00001	***
Phi_1	0.122582	0.0179344	6.8351	<0.00001	***
theta_1	-0.912588	0.00933058	-97.8061	<0.00001	***
Theta_1	-0.990864	0.00285611	-346.9283	<0.00001	***

Media de la var. dependiente = -8.1684e-005

Desviación típica de la var. dependiente. = 0.0508172

media de las innovaciones = -0.000749549

Varianza de las innovaciones = 0.00129656

Log-verosimilitud = 6752.76

Criterio de información de Akaike = -13495.5

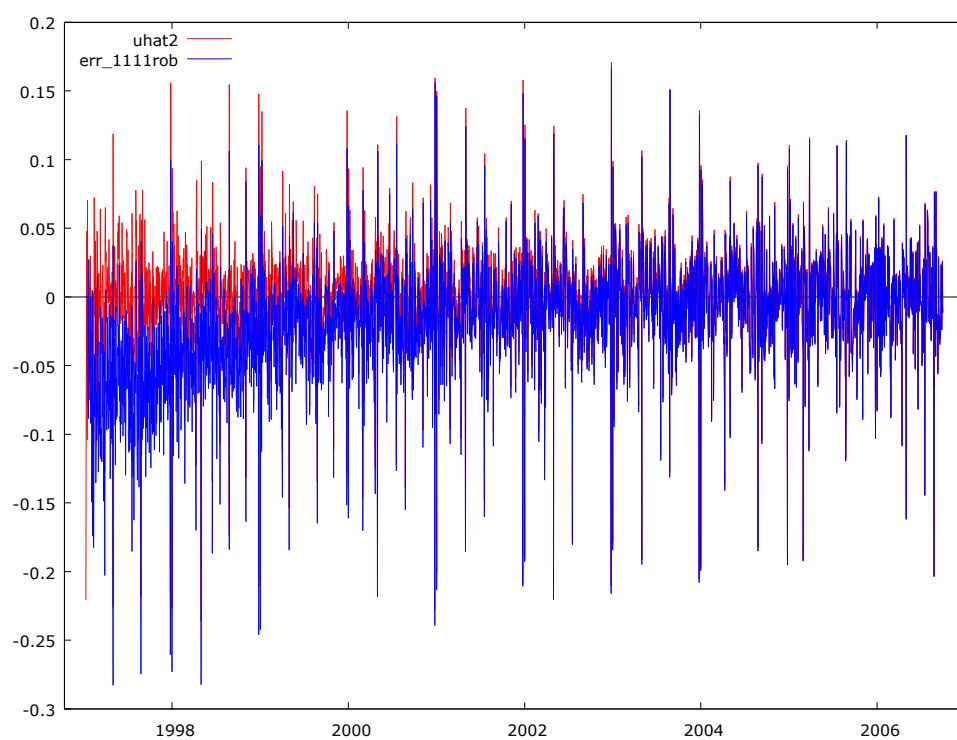
Criterio de información Bayesiano de Schwarz = -13464.6

Criterio de Hannan-Quinn = -13484.5

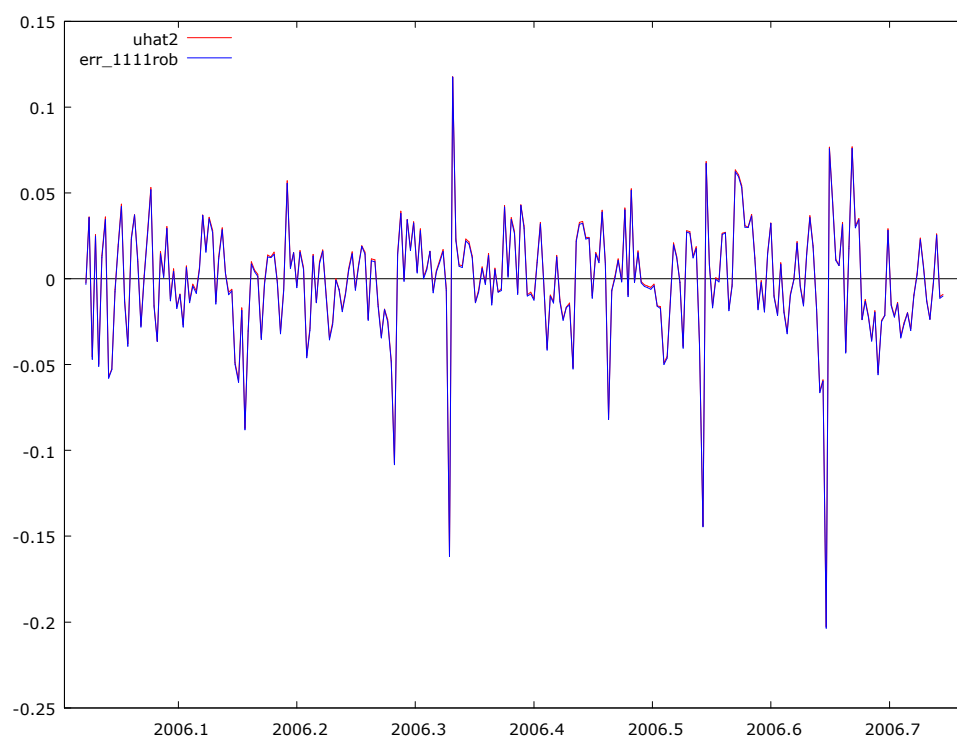
		Real	Imaginaria	Módulo	Frecuencia
AR					
	Raíz 1	1.5867	0.0000	1.5867	0.0000
AR (estacional)					
	Raíz 1	8.1578	0.0000	8.1578	0.0000
MA					
	Raíz 1	1.0958	0.0000	1.0958	0.0000
MA (estacional)					
	Raíz 1	1.0092	0.0000	1.0092	0.0000

Acto seguido, el vector errores fue evaluado en el programa de estimación utilizando los parámetros estimados mediante el procedimiento anterior.

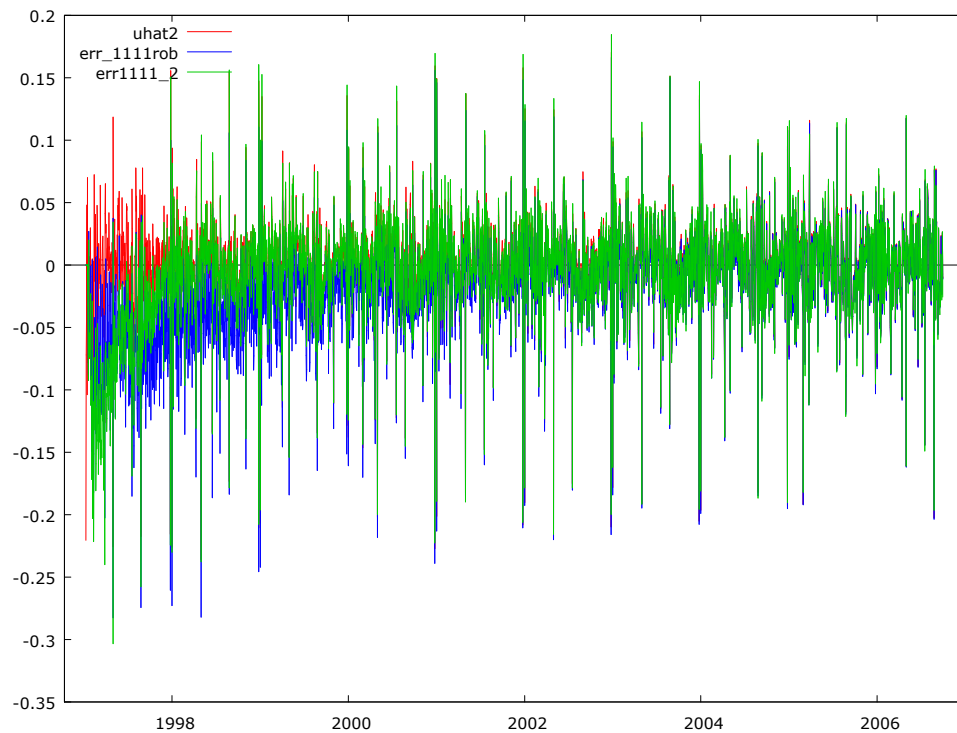
La gráfica siguiente muestra al vector errores calculado mediante estos dos programas.



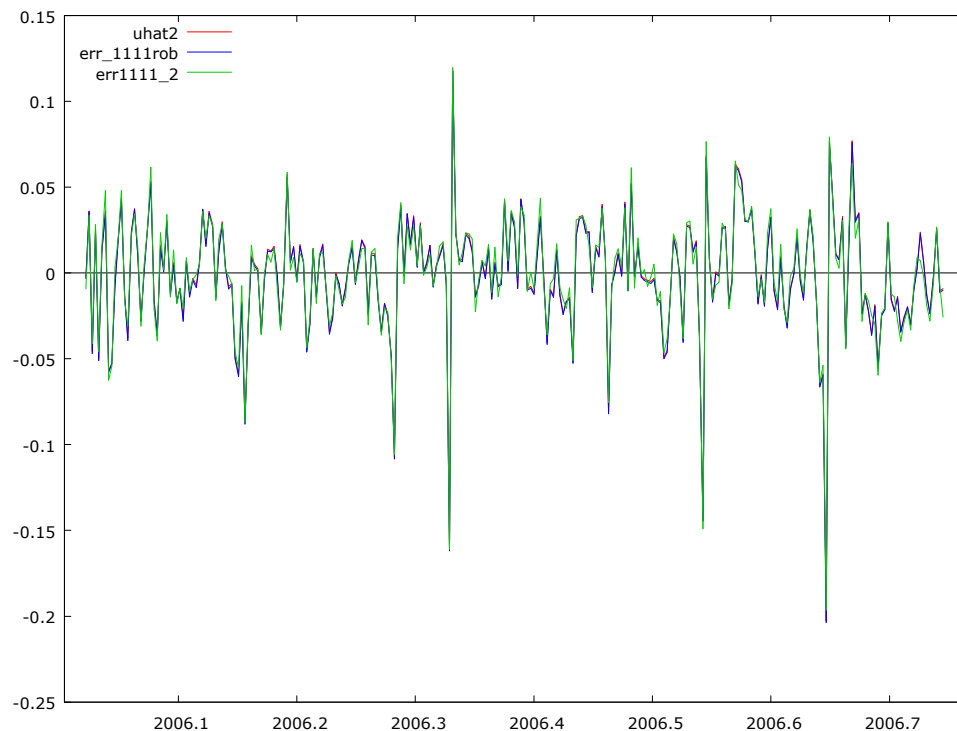
Ahora, si nos enfocamos en el año 2006, obtenemos la siguiente gráfica



Cuando estimamos el modelo con el programa hecho a medida (estos errores están en verde) los errores comparados quedan de la siguiente manera:



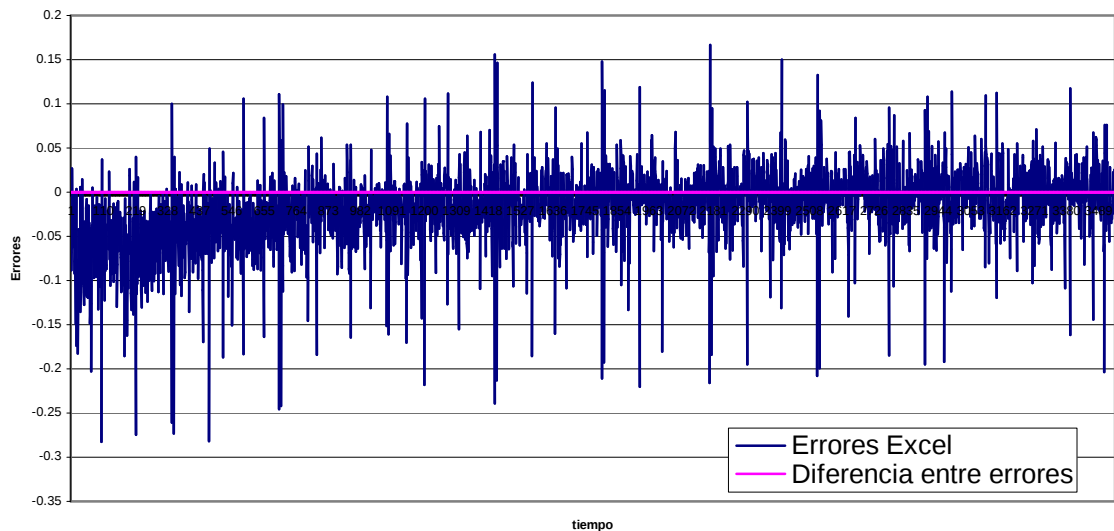
Si nos enfocamos nuevamente en el año 2006, obtenemos el siguiente gráfico:



De esta manera se puede apreciar que el método implementado en este trabajo es asintóticamente equivalente al método de Kalman. Por lo que los errores finales son idénticos y las conclusiones acerca de la bondad de predicción no se verán afectadas, dado que en predicción, solo los últimos errores son tomados en cuenta.

Si bien los errores que surgen del programa son negativos al principio de la muestra y esto podría parecer como un signo de estar mal calculados, un cálculo hecho en Excel demuestra que están bien calculados y que los errores que surgen del programa gretl están calculados al principio de la muestra presentan una corrección (backcast) que hace que queden centrados. La diferencia radica en que parten de errores iniciales diferentes.

Si bien la diferencia entre los errores calculados por Excel y el programa no es nula, es ciertamente despreciable como se observa en la siguiente gráfica y se debe al diferente nivel de precisión utilizado por los citados programas.



Anexo D) Resultados de las estimaciones con datos simulados

Las siguientes simulaciones se hicieron con un tamaño muestral de 3551.

Verdadero proceso: $(1-L)(1-L^7)Y_t = \varepsilon_t$ donde $\varepsilon \sim N(0,1)$.

Estimación:

SCR: 3466.533987366699

(Jacobian * JacobianT)^-1

6.881499417926226E-7 -3.160131933070397E-7

-3.160131933070396E-7 2.064440613183322E-6

d = 0.9839298458081122

D = 0.9653051920994286

Verdadero proceso: $(1-L)X_t = \varepsilon_t$ donde $\varepsilon \sim N(0,1)$.

Estimación:

SCR: 3395.1204569724773

(Jacobian * JacobianT)^-1

5.179018265853838E-7 -1.2204790119495254E-6

-1.2204790119495254E-6 1.8073197699857086E-4

d = 0.9839017958544769

D = 0.025920981086486824

Verdadero proceso: $Z_t = \varepsilon_t$ donde $\varepsilon \sim N(0,1)$.

Estimación:

SCR: 3466.412940017291

(Jacobian * JacobianT)^-1

1.8021440814214944E-4 -2.389991917020891E-5

-2.3899919170208913E-5 1.8661771500941244E-4

d = 0.0039981326643316354

D = -0.018200010299706085

	Y	Aii	lim superior	lim inferior
d	0.9839298	6.88E-07	9.86E-01	9.82E-01
D	0.9653052	2.06E-06	9.97E-01	9.63E-01
SCR	3466.534			
desvest	0.97414341			

	X	Aii	lim superior	lim inferior
d	0.9839018	5.18E-07	9.85E-01	9.83E-01
D	0.025921	1.81E-04	5.18E-02	7.06E-08
SCR	3395.1205	0.9640571		
desvest	0.9640571			

	Ruido bco	Aii	lim superior	lim inferior
d	0.0039981	1.80E-04	3.02E-02	-2.22E-02
D	-0.0182	1.87E-04	8.41E-03	-4.48E-02
SCR	3466.4129	0.9741264		
desvest	0.9741264			

Necesariamente la varianza del proceso está mal calculada, dado que los intervalos de confianza no contienen en muchos casos los valores reales de los parámetros.

Debemos notar que en el único caso en donde el verdadero valor del parámetro cae dentro del intervalo de confianza es el caso del ruido blanco, único proceso estacionario estimado en este set de pruebas.

Anexo E) Resultados del test de HEGY

Modelo 1: estimaciones MCO utilizando las 3635 observaciones 97/01/18-06/12/31
Variable dependiente: sd_d_lnC

Variable	Coeficiente	Desv. típica	Estadístico t	valor p	
const	0.0163782	0.00479518	3.4156	0.00064	***
dummy_2	-0.0102035	0.00716807	-1.4235	0.15469	
dummy_3	-0.01055	0.00717871	-1.4696	0.14175	
dummy_4	-0.0632814	0.00679939	-9.3069	<0.00001	***
dummy_5	-0.0930317	0.00687769	-13.5266	<0.00001	***
dummy_6	0.0663329	0.00751863	8.8225	<0.00001	***
dummy_7	-0.0024965	0.00717363	-0.3480	0.72785	
d_Z1	-0.487812	0.0312441	-15.6129	<0.00001	***
d_Z2	-0.200971	0.0131641	-15.2666	<0.00001	***
d_Z3	0.0600616	0.0135796	4.4229	0.00001	***
d_Z4	-0.18061	0.011764	-15.3528	<0.00001	***
d_Z5	0.010946	0.0119712	0.9144	0.36059	
d_Z6	-0.192751	0.0124946	-15.4268	<0.00001	***
d_Z7	-0.00905837	0.0126709	-0.7149	0.47472	
sd_d_lnC_1	0.191774	0.0359822	5.3297	<0.00001	***
sd_d_lnC_2	0.137183	0.0350745	3.9112	0.00009	***
sd_d_lnC_3	0.114821	0.034398	3.3380	0.00085	***
sd_d_lnC_4	0.119699	0.0297009	4.0302	0.00006	***
sd_d_lnC_5	0.114861	0.0293364	3.9153	0.00009	***
sd_d_lnC_6	0.0991679	0.0289497	3.4255	0.00062	***
sd_d_lnC_7	0.0244379	0.0283311	0.8626	0.38842	
sd_d_lnC_8	0.0256494	0.0276367	0.9281	0.35342	
sd_d_lnC_9	-0.00494652	0.0263822	-0.1875	0.85128	
sd_d_lnC_10	0.0120814	0.0255534	0.4728	0.63639	
sd_d_lnC_11	2.45697e-05	0.019402	0.0013	0.99899	
sd_d_lnC_12	0.0283456	0.0190403	1.4887	0.13665	
sd_d_lnC_13	0.0181512	0.0187247	0.9694	0.33243	
sd_d_lnC_14	-0.0281459	0.0183082	-1.5373	0.12430	
sd_d_lnC_15	-0.0104133	0.0178096	-0.5847	0.55878	
sd_d_lnC_16	-0.0129125	0.0170058	-0.7593	0.44772	
sd_d_lnC_17	-0.012906	0.0165412	-0.7802	0.43531	

Media de la var. dependiente = -4.25684e-006

Desviación típica de la var. dependiente. = 0.0518665

Suma de cuadrados de los residuos = 4.94234

Desviación típica de los residuos = 0.0370317

$R^2 = 0.49444$

R^2 corregido = 0.490231

Estadístico F (30, 3604) = 117.491 (valor p < 0.00001)

Estadístico de Durbin-Watson = 1.99978

Coef. de autocorr. de primer orden. = 3.30022e-005

Log-verosimilitud = 6838.61

Criterio de información de Akaike = -13615.2

Modelo 6: estimaciones MCO utilizando las 3635 observaciones 97/01/18-06/12/31
Variable dependiente: sd_lnC

<i>Variable</i>	<i>Coeficiente</i>	<i>Desv. típica</i>	<i>Estadístico t</i>	<i>valor p</i>	
dummy_1	0.491086	0.0945422	5.1944	<0.00001	***
dummy_2	0.485421	0.0946916	5.1263	<0.00001	***
dummy_3	0.486451	0.0947258	5.1354	<0.00001	***
dummy_4	0.425603	0.0947122	4.4936	<0.00001	***
dummy_5	0.388114	0.0946346	4.1012	0.00004	***
dummy_6	0.57381	0.0945091	6.0715	<0.00001	***
dummy_7	0.492487	0.0945956	5.2062	<0.00001	***
d1t	2.45418e-06	1.62789e-06	1.5076	0.13175	
d2t	2.19225e-06	1.62297e-06	1.3508	0.17686	
d3t	1.32672e-06	1.62192e-06	0.8180	0.41342	
d4t	3.98448e-06	1.61849e-06	2.4619	0.01387	**
d5t	7.82815e-06	1.61921e-06	4.8346	<0.00001	***
d6t	-3.65402e-06	1.62758e-06	-2.2451	0.02482	**
d7t	1.10591e-06	1.62867e-06	0.6790	0.49716	
Z1	-0.00690849	0.00136734	-5.0525	<0.00001	***
Z2	-0.124176	0.0116803	-10.6313	<0.00001	***
Z3	-0.133617	0.0112681	-11.8580	<0.00001	***
Z4	-0.243277	0.0186961	-13.0122	<0.00001	***
Z5	-0.176716	0.0185283	-9.5376	<0.00001	***
Z6	-0.371438	0.0240288	-15.4580	<0.00001	***
Z7	-0.0981329	0.0243406	-4.0317	0.00006	***
sd_lnC_1	0.110155	0.0316744	3.4777	0.00051	***
sd_lnC_2	-0.0490333	0.0317062	-1.5465	0.12207	
sd_lnC_3	-0.00192179	0.0297583	-0.0646	0.94851	
sd_lnC_4	-0.000880379	0.0267993	-0.0329	0.97380	
sd_lnC_5	-0.00784954	0.0267577	-0.2934	0.76927	
sd_lnC_6	-0.00785651	0.026721	-0.2940	0.76876	
sd_lnC_7	-0.0567096	0.0267291	-2.1216	0.03394	**
sd_lnC_8	-0.0106648	0.02674	-0.3988	0.69004	
sd_lnC_9	-0.0267841	0.0267036	-1.0030	0.31592	
sd_lnC_10	0.0337	0.0244549	1.3780	0.16827	
sd_lnC_11	-0.0188799	0.0206277	-0.9153	0.36011	
sd_lnC_12	0.0259277	0.0206064	1.2582	0.20839	
sd_lnC_13	-0.00541235	0.0205977	-0.2628	0.79275	
sd_lnC_14	-0.0360225	0.0205967	-1.7489	0.08039	*
sd_lnC_15	0.0108377	0.0205934	0.5263	0.59873	
sd_lnC_16	-0.000796844	0.0205674	-0.0387	0.96910	
sd_lnC_17	0.017128	0.0165782	1.0332	0.30160	

Media de la var. dependiente = 0.00035203

Desviación típica de la var. dependiente. = 0.0616593

Suma de cuadrados de los residuos = 4.86923

Desviación típica de los residuos = 0.0367926

$R^2 = 0.647577$

R^2 corregido = 0.643952

Estadístico F (38, 3597) = 173.934 (valor p < 0.00001)

Estadístico de Durbin-Watson = 2.00035
 Coef. de autocorr. de primer orden. = -0.000232337
 Log-verosimilitud = 6865.7
 Criterio de información de Akaike = -13655.4

La conclusión del test en primera instancia, rechaza la presencia de una raíz unitaria en las frecuencias bajas, indicando que no existe raíz unitaria regular en las primeras diferencias.
 El análisis de la segunda salida indica que no existen raíces unitarias asociadas a ninguna frecuencia, por lo que no debería hacerse ninguna diferencia entera, ni regular ni estacional a la serie del Ln del Consumo

Las tablas del estadístico utilizado en el test son las siguientes:

Auxiliary Regression	T	$t': \pi_{add}$							
		0.01	0.025	0.05	0.10	0.90	0.95	0.97.5	0.99
No drift	240	-2.29	-1.93	-1.62	-1.25	1.26	1.62	1.94	2.31
No dummies	480	-2.31	-1.94	-1.62	-1.26	1.27	1.64	1.95	2.30
No trend	1000	-2.30	-1.95	-1.63	-1.29	1.28	1.63	1.95	2.30
Drift	240	-2.26	-1.90	-1.60	-1.24	1.27	1.63	1.96	2.33
No dummies	480	-2.32	-1.94	-1.63	-1.26	1.27	1.63	1.94	2.30
No trend	1000	-2.30	-1.95	-1.63	-1.27	1.27	1.63	1.94	2.30
Drift	240	-2.62	-2.21	-1.86	-1.45	1.48	1.89	2.24	2.63
Dummies	480	-2.67	-2.25	-1.90	-1.48	1.52	1.92	2.28	2.69
No trend	1000	-2.70	-2.29	-1.93	-1.52	1.54	1.95	2.29	2.70
Drift	240	-2.25	-1.89	-1.58	-1.22	1.26	1.62	1.93	2.32
No dummies	480	-2.31	-1.93	-1.62	-1.25	1.27	1.63	1.95	2.31
Trend	1000	-2.33	-1.96	-1.63	-1.26	1.27	1.63	1.95	2.34
Drift	240	-2.60	-2.21	-1.85	-1.45	1.45	1.85	2.21	2.63
Dummies	480	-2.66	-2.26	-1.91	-1.50	1.49	1.89	2.24	2.65
Trend	1000	-2.70	-2.30	-1.93	-1.51	1.51	1.94	2.31	2.71
No drift	240	-2.65	-2.27	-1.90	-1.49	1.54	1.97	2.33	2.76
Dummies	480	-2.76	-2.34	-1.98	-1.55	1.58	2.01	2.37	2.81
Seas. dummies.	1000	-2.87	-2.45	-2.07	-1.62	1.63	2.08	2.44	2.85

Auxiliary Regression	T	$t': \pi_1$				$t': \pi_{env}$				$F_{evenadd}$			
		0.01	0.025	0.05	0.10	0.01	0.025	0.05	0.10	0.90	0.95	0.97.5	0.99
No drift	240	-2.52	-2.16	-1.89	-1.57	-2.50	-2.17	-1.87	-1.54	2.32	3.01	3.66	4.63
No dummies	480	-2.58	-2.19	-1.91	-1.58	-2.54	-2.18	-1.89	-1.57	2.37	3.05	3.77	4.67
No trend	1000	-2.57	-2.21	-1.91	-1.60	-2.54	-2.20	-1.91	-1.58	2.39	3.08	3.78	4.70
Drift	240	-3.34	-3.04	-2.78	-2.49	-2.48	-2.13	-1.86	-1.53	2.30	2.98	3.65	4.54
Dummies	480	-3.43	-3.11	2.84	-2.56	-2.50	-2.18	-1.89	-1.55	2.35	3.05	3.73	4.65
Trend	1000	-3.44	-3.12	-2.86	-2.57	-2.58	-2.23	-1.93	-1.58	2.38	3.08	3.79	4.73
Drift	240	-3.36	-3.04	-2.78	-2.47	-3.83	-3.52	-3.25	-2.95	5.25	6.25	7.18	8.33
No dummies	480	-3.39	-3.10	-2.81	-2.51	-3.86	-3.54	-3.30	-2.99	5.43	6.40	7.35	8.50
No trend	1000	-3.41	-3.10	-2.85	-2.56	-3.88	-3.58	-3.33	-3.03	5.54	6.53	7.48	8.63
Drift	240	-3.89	-3.60	-3.34	-3.06	-2.48	-2.16	-1.80	-1.55	2.30	2.97	3.67	4.52
No dummies	480	-3.98	-3.67	-3.41	-3.12	-2.51	-2.18	-1.89	-1.56	2.34	3.03	3.71	4.60
Trend	1000	-3.98	-3.67	-3.41	-3.12	-2.53	-2.20	-1.92	-1.58	2.40	3.10	3.81	4.70
Drift	240	-3.91	-3.57	-3.31	-3.03	-3.84	-3.52	-3.25	-2.95	5.24	6.23	7.16	8.32
Dummies	480	-3.93	-3.61	-3.35	-3.06	-3.86	3.55	-3.29	-2.98	5.40	6.37	7.30	8.60
Trend	1000	-3.98	-3.67	-3.41	-3.12	-3.89	-3.58	-3.33	-3.03	5.56	6.59	7.54	8.69
No drift	240	-4.07	-3.76	-3.49	-3.19	-4.92	-4.59	-4.33	-4.03	8.88	10.12	11.29	12.81
Dummies	480	-4.03	-3.71	-3.46	-3.17	-4.81	-4.51	-4.25	-3.96	8.71	9.90	11.03	12.42
Seas. Trends.	1000	-3.98	-3.69	-3.42	-3.14	-4.72	-4.44	-4.19	-3.91	8.60	9.82	10.88	12.22

Fuente: Antonio Rubia[29].

Anexo F) Demostración del cálculo recursivo de los Cj y los Fs

Los multiplicadores C y F se obtienen expandiendo el binomio mediante el desarrollo de Taylor, lo que se conoce como binomio de Newton.

$$\begin{aligned}
 (1-L)^d &= 1 + \frac{d}{1!}(-L)^1 + \frac{d(d-1)}{2!}(-L)^2 \\
 &\quad + \frac{d(d-1)(d-2)}{3!}(-L)^3 + \dots \\
 &= \sum_{j=0}^{\infty} \frac{d(d-1)(d-2)\dots(d-(j-1))}{j!} (-1)^j L^j
 \end{aligned} \tag{E.1}$$

Ahora si multiplicamos los factores por -1:

$$(1-L)^d = \sum_{j=0}^{\infty} \frac{(-d)(1-d)(2-d)\dots((j-1)-d)}{j!} L^j \tag{E.2}$$

Luego multiplicamos por $1 = \frac{\Gamma(j-j-d)}{\Gamma(-d)}$ y revertimos el orden de los factores para obtener:

$$(1-L)^d = \sum_{j=0}^{\infty} \frac{(j-1-d)(j-2-d)\dots(j-j-d)\Gamma(j-j-d)}{j!\Gamma(-d)} L^j \tag{E.3}$$

Ahora que tenemos esta expresión, podemos repetirla usando la propiedad recursiva de la función

$$\Gamma(x) = (x-1)\Gamma(x-1) \tag{E.4}$$

Hasta llegar a la fórmula que vemos en la mayoría de los artículos referidos a este tema.

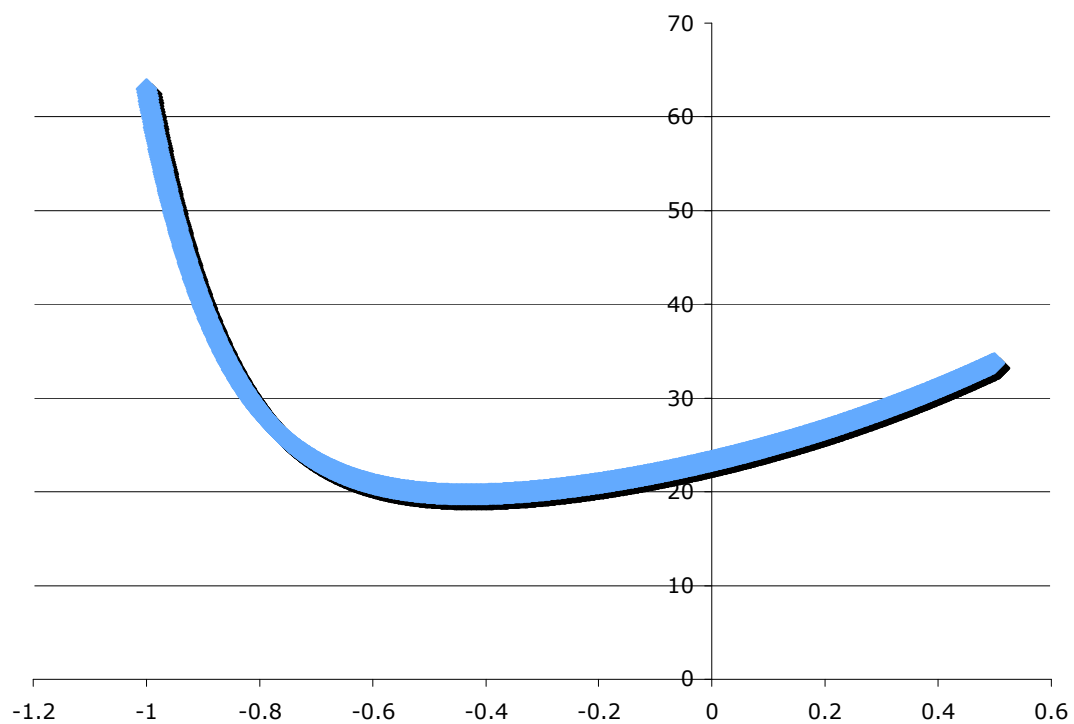
$$(1-L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} L^j \tag{E.5}$$

de esta manera los pesos Cj son difíciles de calcular para valores “grandes” del argumento de la función gamma. Por ejemplo, la función gamma evaluada en 171 es aproximadamente 7.257E306. Esta aproximación usa también la propiedad recursiva de la función gamma (E.4) y llega a la fórmula recursiva de los Cj.

$$\begin{aligned}
 C_0 &= \frac{\Gamma(0-d)}{\Gamma(1)\Gamma(-d)} = 1 \\
 C_j &= \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} = \frac{(j-d-1)\Gamma(j-d-1)}{j\Gamma(j)\Gamma(-d)} = \\
 &= \frac{(j-d-1)}{j} C_{j-1}
 \end{aligned} \tag{E.6}$$

La demostración arriba explicada fue obtenida del trabajo de Mc Carthy, DiSario y Saraoglu [23] y puede ser extendida fácilmente al caso de $(1-L^7)^D$.

Anexo G) Gráfica de SCR en función de d



Bibliografía

1. AL-SHAKARCHI, M. R. G.; GHULAIM, M. M. 2000. Short-Term Load Forecasting for Baghdad Electricity Region. *Electric Machines and Power Systems*. 28, 355–371.
2. ARRILLAGA, P, ROSAS F. SIBILLE, M. 2002. Estimación de curvas de carga para subestaciones de UTE por área de distribución tipo. Monografía para la obtención del título de Lic. En Economía. Univ. de la República. Uru. 126 p.
3. BAILLIE, RICHARD.1996. Long memory processes and fractional integration in econometrics”. *Journal of Econometrics* .73. 5-59.
4. BISOGNIN, CLEBER. 2003. Estimación e previsão em processos com longa dependência sazonais. Tesis de Maest. en Matemática. Universidad Federal de Río Grande del Sur, Bra. 132 p.
5. BISOGNIN, CLEBER. 2003. Estimación e previsão em processos com longa dependência sazonais. Tesis Maest. en Mat. Universidad Federal de Río Grande del Sur, Brasil. 132 p.
6. BISOGNIN, CLEBER; LOPES, REGINA. 2007. Estimating and forecasting the long-memory parameter in the presence of periodicity. *Journal of Forecasting*. 26 no publicado.
7. BOND, D; HARRISON, M.J; O'BRIEN, E.J. 2005. Testing for Long Memory and Nonlinear Time Series: A Demand for Money Study. *Economic Papers tep20021*, Trinity College Dublin, Department of Economics.
8. BRIETZKE; LOPES; BISOGNIN. 2005. A closed formula for the Durbin-Levinson's algorithm in seasonal fractionally integrated processes. *Mathematical and Computer Modelling*. 42. 1191-1206.
9. CANCELO, JOSE R.; ESPASA, ANTONI. 1993. Modelling and forecasting daily series of electricity demand. *Investigaciones económicas*. Vol. XX (3) 359-376.
10. CHAN, N. H., PALMA, W. 2006. Estimation of long-memory time series: A survey of likelihood-based methods.www.bus.lsu.edu/economics/faculty/chill/personal/aie/chan.pdf
11. CHUNG, CHING-FAN. 1996. Estimating a generalized long memory process. *Journal of Econometrics*. 73 237-259.
12. Edison Electric Institute. 2006. Key facts about the electric power industry. www.eei.org/industry_issues/industry_overview_and_statistics/nonav_key_facts/index.htm
13. GRANGER, CLIVE; JOYEUX, ROSELYNE. 1980. An introduction to long memory time series models and fractional differencing. *Journal of time series analysis*. 1 1980,15-29.
14. GUÉGAN, DOMINIQUE. 2004. The use of special day information in a demand forecasting model for Nordic power market. Independent research projects in applied mathematics, Systems Analysis Laboratory, Helsinki University of Technology. www.sal.tkk.fi/Opinnot/Mat-2.108/pdf-files/elau04a.pdf

15. GREENE, WILLIAM H. "ANÁLISIS ECONOMETRICO". Tercera edición. Madrid: Prentice Hall, 1999. 913 p.
16. GUTIÉRREZ MOYA, ESTER. 2003. La demanda residencial de energía eléctrica en la Comunidad Autónoma de Andalucía: un análisis cuantitativo. Tesis Dr. en Economía. Universidad de Sevilla, Es. 448 p.
17. HENLEY, ANDREW; PEIRSON, JOHN. 1997. Non-Linearities in Electricity Demand and Temperature: Parametric Versus Non-Parametric Methods. *Oxford Bulletin of Economics and Statistics*. 59 (1):149-162
18. International Conference.2004. Modelling and forecasting electricity loads: a comparison. Lodz, Pol.WERON, RAFAL; MISIOREK ADAM. 2004. Modelling and forecasting electricity loads: a comparison. International Conference, The European Electricity Market 2004". Lodz, Polonia. p.135-142.
19. Jornadas Anuales de Economía del BCU .2006.Montevideo, Uru. La memoria larga y la cointegración fraccional en las series temporales: una aplicación a la paridad de los poderes de compra en los países del cono sur.
20. KATAYAMA, NAOYA.2004. Seasonally and Fractionally Differenced Time Series. Discussion paper series, no 11. Hitotsubashi Research unit. Hitotsubashi University. 30 p.
21. LE COMTE, D. M., AND WARREN H., 1981: Modelling the impact of summer temperatures on national electricity consumption. *Journal of Applied Meteorology*, 20, 1415-1419
22. MILLER K.S., ROSS B. An introduction to the fractional calculus and fractional differential equations Publisher: John Wiley & Sons; 1 edition, 1993. 384 p.
23. MCCARTHY, JOSEPH, DISARIO, ROBERT, SARA OGLU, HAKAN. 2003.A recursive algorithm for fractionally differencing long data series. *Journal of Modern Applied Statistical Methods* 2 (1)272-278.
24. OCKER, DIRK. 1999. Stationary and non stationary FARIMA Models- model choice, forecasting, aggregation and intervention. Tesis Doctoral Universidad de Konstanz, Alemania. 166 p.
25. PARKE, WILLIAM. 1999. What is fractional integration?". Department of economics. University of North Carolina, Chapel Hill. Working paper no. 99-01.
26. PÉREZ ESPARTERO, ANA. 1999. Modelos con memoria larga para la media : procesos ARFIMA. Serie documentos de trabajo. no 1 Dpto. Economía Aplicada (Estadística y Econometría) Universidad de Valladolid.
27. REISEN, VALDERIO. ARFIMA: O modelo ARIMA para d fracionário. CNPq, ABE, 1995. 80 p.

28. REISEN, VALDERIO; RODRÍGUEZ, ALEXANDRE; PALMA, WILFREDO. 2006. Estimation of seasonal fractionally integrated processes. *Computational Statistics & Data Analysis* 50(2):568-582
29. RUBIA, ANTONIO. 2001. Testing for seasonal unit roots in daily electricity demand: evidence from deregulated markets. Instituto Valenciano de Investigaciones Económicas, Working papers serie EC no. WP-2001-21. 28p.
30. SOARES, LACIR JORGE; ROCHA SOUZA, LEONARDO. 2003. Forecasting Electricity Load Demand: Analysis of the 2001 Rationing Period in Brazil. *Economics Working Papers* (Ensaio Económico, 486)
31. SOARES, LACIR JORGE; ROCHA SOUZA, LEONARDO. 2006. Forecasting Electricity Demand Using Generalized Long Memory. *International Journal of Forecasting* 22 (1):17-28.
32. SPREMOLLA, ALESSANDRA. 2001. PERSISTENCIA EN EL DESEMPLEO DE URUGUAY. *Cuad. econ.*, vol.38, no.113, p.73-89. ISSN 0717-6821
33. TAYLOR, JAMES W., DE MENEZES, LILIAN M.; MCSHARRY, PATRICK E. 2006. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting*. 22 p. 1-16.
34. TAYLOR, LESTER D. 1979. Time-of-Day and Seasonal Demand for Electric Power. *Growth and Change*. 10 (1): 6 p.
35. URIEL, EZEQUIEL. 1985. Análisis de series temporales, Modelos Arima. Madrid. Paraninfo. 270 p.

Agradecimientos

Deseo agradecer de manera especial y en primer lugar el apoyo emocional en todo momento recibido de parte de mi familia.

Mención especial merece el aporte de mi novia Daniela Jaureguiberry en esta monografía, ya que ella creyó y cree en mí más de lo que yo mismo podría.

En cuanto al consejo profesional agradezco especialmente a mi tutor, Adrián Fernández, que a pesar de su agenda se hizo un tiempo para asesorarme cada vez que lo requería para marcarme el camino mas coherente.

En cuanto a dudas que parecían ser existenciales apareció en muchas ocasiones la Silvia Rodríguez para calmar mi ansiedad y siempre poner al problema en su lugar.

Varias fueron las personas que ayudaron desde el exterior en este trabajo, entre ellos destaco al profesor Reisen desde Vitoria con quien mantuve contacto vía mail y fue él quien me envió uno de los primeros materiales de integración fraccional a los que tuve acceso. Gracias también al profesor Cleber Bisognin quien respondió a mis preguntas por mail e inclusive me recibió en su despacho de Porto Alegre de la Universidad Federal de Rio Grande do Sul. Mucho ayudaron desde España los comentarios de Gerardo Prieto Blanco quien además fomentó mi interés en los procesos tratados en esta monografía. Por último debo agradecer los aportes realizados a la programación que hizo desde Bryant College en Rhode Island el profesor McCarhty.

Dentro de UTE deseo mostrar mi agradecimiento a Luis Rodríguez por su experiencia dentro de la empresa y a Gustavo Pérez por sus sugerencias en cuanto a la modelización del consumo de electricidad y por último a Germán Di Lorenzo por su gestión para la obtención de los datos.

En lo referente a la búsqueda bibliográfica destaco la labor del personal de biblioteca, tanto de Facultad de Ciencias Económicas como la del Banco Central del Uruguay.

Por último un reconocimiento para quienes hicieron posible que PROLEGO pudiera existir: Ing. Juan J. Silveira, Ing. Andrés Meyer e Ismael Garrido. Detrás de la PC pasaron horas escribiendo un código que apenas entendían para que era útil, para ellos: muchas gracias.

Mis disculpas a todas aquellas personas cuya mención pude olvidar.

Abstract.....	1
1) Introducción	2
2) Propiedades básicas de los modelos SARFIMA	4
3) Estimación.....	9
Estimación por mínimos cuadrados no lineales	9
Propiedades de los estimadores MCNL	10
Procedimiento de estimación.....	10
Procedimiento de estimación en la práctica	12
4) Consideraciones acerca del consumo de electricidad en Uruguay.....	15
Cualidades de la energía eléctrica	15
Factores que influyen en la demanda de energía eléctrica	15
Importancia de la estimación y previsión del consumo de energía eléctrica.....	16
Descripción del sistema eléctrico uruguayo	17
Marco regulatorio	18
5) Resultados	19
Estimación	19
Pruebas de hipótesis	21
Pruebas de hipótesis	23
Predicción	25
6) Conclusiones	27
A) Desde el punto de vista de la correcta especificación	27
B) Desde el punto de vista de la bondad de predicción.....	27
7) Consideraciones	29
Acerca del orden de diferenciación:	29
Acerca del algoritmo utilizado:	29
Acerca de la modelización elegida:	29
Anexo A) Nociones del concepto de diferenciación fraccional	30
Derivada $\frac{1}{2}$ de una función sencilla.....	30
Anexo B) Procedimiento para la construcción del programa.....	32
B.1) Módulo estimación	32
B.2) Módulo predicción.....	34
Anexo C) Pruebas de validación del programa a medida.....	35
Anexo D) Resultados de las estimaciones con datos simulados	39
Anexo E) Resultados del test de HEGY	41
Anexo F) Demostración del cálculo recursivo de los C_j y los F_s	44
Anexo G) Gráfica de SCR en función de d	45
Bibliografía	46
Agradecimientos	49