

Predicción mediante modelos ARFIMA y FOU de energía afluente

Prediction using ARFIMA and FOU models of affluent energy

Juan Kalemkerian¹

Recibido: Agosto 2017

Aprobado: Octubre 2017

Resumen.- En este trabajo se estudian predicciones a partir de modelos ARFIMA y FOU para la serie de datos semanales de energía afluente generada por las represas hidroeléctricas de Uruguay entre 1909 y 2012. Se describe la serie de datos, y mediante la estimación del exponente de Hurst se muestra la conveniencia de modelar a través de procesos de memoria larga. Se presentan dos familias de modelos de series de tiempo de este tipo, los ARFIMA y los FOU (Ornstein-Uhlenbeck fraccionarios). Se estiman sus parámetros y se compara el rendimiento de los mismos teniendo en cuenta su poder predictivo.

Palabras clave: modelos ARFIMA; Ornstein-Uhlenbeck fraccionarios; memoria larga.

Summary.- In this work we study predictions from ARFIMA and FOU models for the weekly data series of affluent energy generated by hydroelectric dams in Uruguay between 1909 and 2012. The estimation of Hurst coefficient suggests modeling through long memory time series. We present two families of time series models of this type, ARFIMA and FOU (fractional Ornstein-Uhlenbeck) models. Their parameters are estimated and taking into account their predictive power, their performance is compared.

Keywords: ARFIMA model; fractional Ornstein-Uhlenbeck; long range dependence.

1. Introducción.- En estadística cuando estamos frente a un conjunto de datos $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ que responden a mediciones de cierta variable a través del tiempo, se dice que estamos frente a una serie de tiempo. Los modelos que existen para las series de tiempo, tienen en cuenta que la observación de un dato en el instante t depende de las observaciones anteriores y es de uso común el trabajar con modelos llamados estacionarios (donde la dependencia entre una variable medida en el instante t y la medición de la variable en el instante $t+k$ no depende de t). Si las variables consideradas además están centradas (es decir tienen esperanza nula), se utiliza como medida de dependencia el coeficiente de correlación entre X_t y X_{t+k} que llamaremos

$\rho(k) := \rho(X_t, X_{t+k}) = \frac{E(X_{t+k} X_t)}{V(X_t)}$ y recibe el nombre de función de autocorrelación de la

serie. En la mayoría de las aplicaciones se cumple que $\rho(k) \rightarrow 0$ cuando $k \rightarrow +\infty$ ya que habitualmente cuando dos observaciones están lejanas en el tiempo tienden a ser independientes

¹ Docente Facultad de Ciencias UdelaR, Docente Facultad de Ingeniería, Universidad de Montevideo, Uruguay.
jkalem@emat.edu.uy

entre sí. Cuando $\sum_{k=1}^{+\infty} |\rho(k)| < +\infty$ se dice que estamos en presencia de un proceso de memoria corta, lo cual se puede leer como que la tendencia a cero de $\rho(k)$ es rápida, mientras que cuando $\sum_{k=1}^{+\infty} |\rho(k)| = +\infty$ el proceso se llama de memoria larga ya que ahora la tendencia a cero de $\rho(k)$ es más lenta, lo cual se interpreta de modo en que el valor de una observación en un instante de tiempo, puede estar siendo influido por algunas observaciones pasadas bastante alejadas en el tiempo. A partir del trabajo pionero de Hurst [1] se ha detectado el fenómeno de la memoria larga en varias series de tiempo que tienen que ver con datos hidrológicos y climáticos, entre otros. En muchos modelos de serie de tiempo estacionarios, se cumple que $\rho(k) \cong ck^{2H-2}$ cuando $k \rightarrow +\infty$ para cierta constante c positiva. De aquí se deduce que si $H > 1/2$ estamos ante un proceso de memoria larga. El valor de H recibe el nombre de exponente de Hurst. El presente trabajo es el producto de la investigación realizada en el proyecto ANII² “Seguro climático en contexto no estacionario” en lo que tiene que ver con la modelación estadística de la serie de energía afluente generada por las represas hidroeléctricas de Uruguay mediante series de tiempo de memoria larga. En la sección 2 se describe la serie de datos y se trazan los objetivos a perseguir en el trabajo. En la sección 3, se muestra mediante la estimación del exponente de Hurst, la conveniencia de modelar la serie de tiempo mediante modelos de memoria larga. También se presentan dos familias de modelos de memoria larga abreviados como ARFIMA y FOU, se definen los mismos, así como se enuncian algunas de sus propiedades fundamentales. En la sección 4 se hace una breve explicación de cómo se pueden realizar predicciones mediante estos modelos. A su vez se muestran los principales resultados obtenidos para esta serie de datos. En la sección 5 se plantean las conclusiones a las que se arriba. En la sección 6 se plantea el trabajo a futuro para seguir profundizando en el poder predictivo de estos modelos.

2. Descripción de los datos y objetivos.- Se trabajó con los datos semanales de la energía afluente generada por las represas hidroeléctricas desde la primer semana de 1909 a la semana 26 de 2013, totalizando 5435 datos. Cada dato corresponde a la energía afluente semanal generada por las represas de Salto Grande, Rincón del Bonete y Palmar en MWh. A partir de los aportes (en m³/segundo) semanales de estas tres represas se calcula la energía afluente que generan. El objetivo principal del trabajo es el de ajustar diversos modelos a esta serie de datos de energía afluente y principalmente analizar su comportamiento en cuanto a la predicción de la energía acumulada en cada semestre. Lograr un modelo que reproduzca bien el comportamiento de la serie observada y que permita realizar buenas predicciones, es de vital importancia para la mejora en la planificación energética de un país. Los cuartiles de las distribuciones relativas a la energía acumulada en cada semestre de cada año se ven la Tabla I y en el diagrama de caja que se ve en la Figura III.

Acumulados	Mínimo	Primer cuartil	Mediana	Segundo cuartil	Máximo
Semestre 1	222.737,3	1.925.545,0	2.696.106,0	3.727.722,8	7.159.239,7
Semestre 2	616.352,3	3.338.084,6	4.107.523,2	4.855.440,5	7.578.012,2

Tabla I.- Cuartiles de la variable acumulados semestrales por semestre en MWh.

² Agencia Nacional de Investigación e Innovación

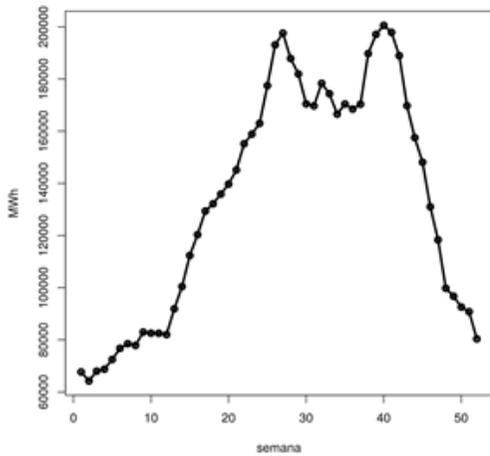


Figura I.- Medias semanales de la energía afluente MWh.

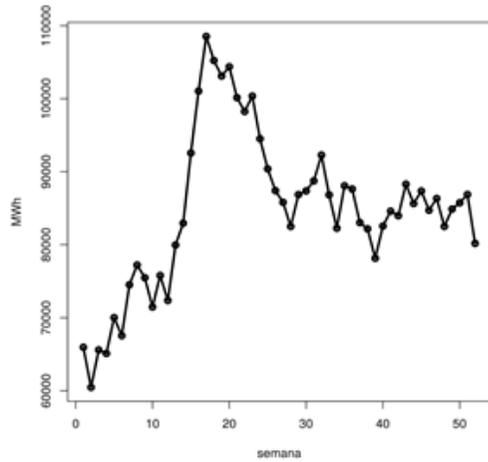


Figura II.- Desviaciones típicas de la energía afluente en MWh.

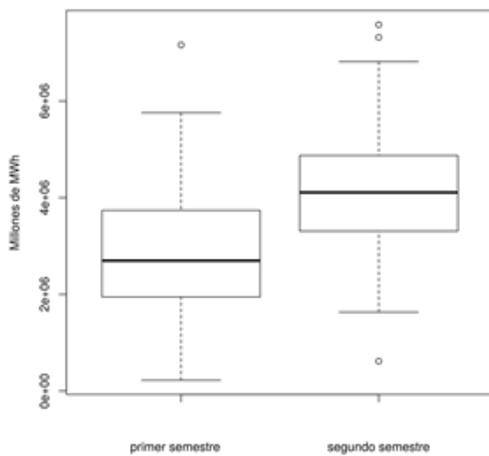


Figura III.- Diagrama de caja con la energía afluente acumulada del primer y segundo semestre de los 104 años.

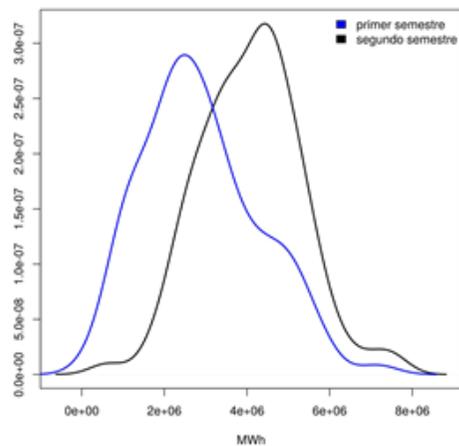


Figura IV.- Estimación por núcleos de la densidad de las variables energía afluente acumulada de cada semestre en los 104 años.

La serie de tiempo tiene una clara componente estacional, como se puede ver en las figuras I y II donde se grafican los promedios semanales así como las desviaciones típicas correspondientes a las 52 semanas del año. También se ve claramente en la figura V, donde se grafica la función de autocorrelación, con rezagos de hasta 300 semanas. Para modelar la serie a través de un modelo estacionario, se le quitó la componente estacional, lo que se llevó a cabo restándole a cada dato correspondiente a determinada semana i (para $i=1,2,3,\dots,52$) el valor de la media semanal de la semana i y dividiéndolo por la desviación estándar semanal de la semana i . En la figura VI, se grafica la función de autocorrelación de la serie desestacionalizada, donde se observa cómo la componente estacional ha sido removida.

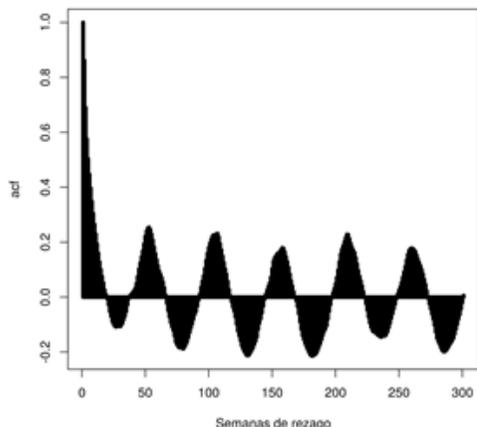


Figura V.- Función de autocorrelación de la serie original.

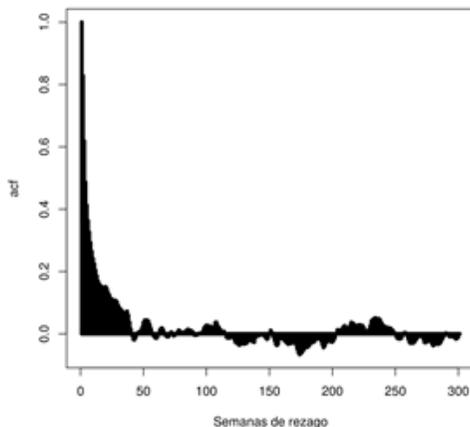


Figura VI.- Función de autocorrelación de la serie desestacionalizada.

3. Modelos considerados. - Debido a que la serie de datos depende fuertemente de datos hidrológicos (aportes a las represas de Salto Grande, Rincón del Bonete y Palmar), se consideró razonable modelar a través de modelos de los llamados de “memoria larga”. Para verificarlo, se hizo un análisis sobre la conveniencia de modelar a través de procesos de memoria larga, lo cual se llevó a cabo estimando el exponente de Hurst por varios métodos distintos. En [2], se puede ver una descripción de diversos métodos para estimar el parámetro de Hurst. En las figuras VII y VIII se ven las estimaciones de H por dos de ellos: el método R/S y el método de la varianza agregada.

Si bien se observan diferencias grandes entre la estimación de H por un método u otro, se destaca que es algo que ocurre frecuentemente, pero en ambos casos se ve que la estimación de H supera claramente al valor $\frac{1}{2}$. Esto mismo ha ocurrido para la estimación de H utilizando otros métodos.

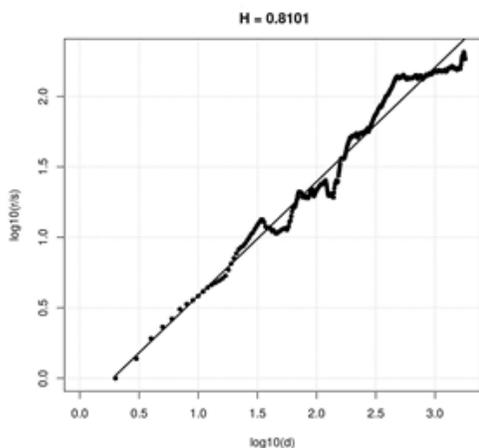


Figura VII.- Estimación de H por el método R/S .

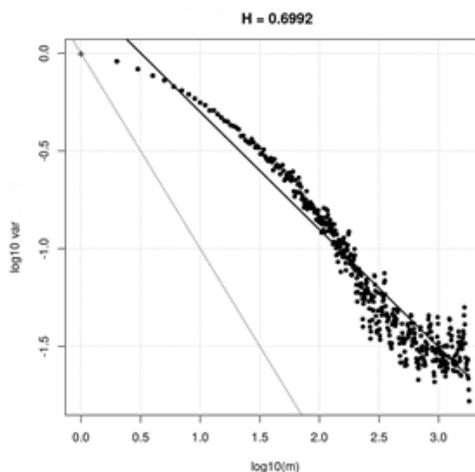


Figura VIII.- Estimación de H por el método de la varianza agregada.

Una vez que se observó la razonabilidad de ajustar la serie por una serie de tiempo de memoria larga, se recurrió a los modelos autorregresivos de medias móviles fraccionalmente integrados (ARFIMA o FARIMA), así como a los modelos de Ornstein-Uhlenbeck fraccionarios (FOU).

3.1 Modelos ARFIMA.

Definición. Dado un proceso estocástico $\{X_t\}$ estacionario y centrado, se dice que es un ARMA(p, q) si y sólo si existen un proceso $\{\varepsilon_t\}$, p, q números naturales y $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$ números reales tales que $\Phi(B)X_t = \Theta(B)\varepsilon_t$ donde B es el operador shift, $BX_t = X_{t-1}$, Φ y Θ son los polinomios $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$, y $\{\varepsilon_t\}$ es ruido blanco, es decir, que son variables no correlacionadas, idénticamente distribuidas con media cero.

Definición. Dado un proceso estocástico $\{X_t\}$ estacionario y centrado, se dice que es un ARFIMA(p, d, q) si y sólo si, existe $d \in (-1/2, 1/2)$, un proceso $\{\varepsilon_t\}$, p, q números naturales, y $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$ números reales tales que $\Phi(B)(1-B)^d X_t = \Theta(B)\varepsilon_t$ donde B es el operador shift, Φ y Θ son los polinomios $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$, el proceso $\{\varepsilon_t\}$ es ruido blanco, y el operador $(1-B)^d$, se define como sigue: $(1-B)^d = \sum_{s=0}^{+\infty} \frac{\Gamma(1+d)}{\Gamma(d-s+1)\Gamma(s+1)} (-1)^s B^s$ donde Γ es la función gamma.

Se observa que $\{X_t\}$ es un ARFIMA(p, d, q) si y sólo si $Y_t^{(d)} = (1-B)^d X_t$ es un ARMA(p, q).

Más explícitamente, se tiene que $Y_t^{(d)} = \sum_{s=0}^{+\infty} \frac{\Gamma(1+d)}{\Gamma(d-s+1)\Gamma(s+1)} (-1)^s X_{t-s}$. Se observa que

cuando le aplicamos la transformación $(1-B)^d$ a la serie $\{X_t\}$, se obtiene una nueva serie $Y_t^{(d)}$ que es un ARMA(p, q), que es un proceso de memoria corta. Se prueba que bajo ciertas hipótesis (ver [3] o [4]), un proceso ARFIMA(p, d, q) tiene una estructura de correlación con la siguiente expresión asintótica, $\rho(k) \cong ck^{2d-1}$ cuando $k \rightarrow +\infty$, de aquí se deduce que el proceso es de memoria larga si y sólo si $0 < d < 1/2$, lo que equivale a decir que si definimos $H = d + 1/2$, entonces la condición se traduce a $1/2 < H < 1$. Dado que las autocorrelaciones asintóticas son como las encontradas por Hurst (1951), dicha igualdad vincula el parámetro de memoria larga de un modelo ARFIMA, con el parámetro de Hurst.

Los modelos ARFIMA han sido propuestos por Granger & Joyeux (ver[5]) y tienen varias ventajas. Por un lado se tiene libertad de poder elegir la cantidad de parámetros que uno desee, lo cual le da una gran flexibilidad. Por otro lado, son modelos que contemplan tanto la memoria larga (a través de su parámetro d), como la memoria corta, a través de los p coeficientes AR y q coeficientes MA. A su vez es una virtud de estos modelos, el hecho de que el ruido blanco puede tener cualquier distribución no necesariamente normal. Los modelos ARFIMA han sido utilizados con éxito para modelar varios casos de series de tiempo de memoria larga pertenecientes a diversas áreas del conocimiento como por ejemplo la hidrología, climatología, estudio de volatilidad en algunas series económicas.

3.1.1 Selección del modelo.- En un modelo ARFIMA(p,d,q), los valores p y q son fijados de antemano por lo que a la hora de ajustar un modelo de éstos, el primer paso es elegir mediante algún tipo de criterio los valores de p y q para luego estimar los parámetros del modelos que son $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q, d, \sigma$. Se tomaron valores de p y q variando entre 0 y 4 y en cada uno de estos casos se estimaron los parámetros por máxima verosimilitud, se utilizó para este objetivo el paquete arfima de R. Para la elección de los valores de p y q se tomaron en cuenta cuatro criterios: *AIC*, *BIC*, (definidos líneas más abajo) ajuste de correlaciones y el trabajo realizado sobre los aportes a las represas hidroeléctricas uruguayas (en metros cúbicos sobre segundo) realizado por Kalemkerian ([6]).

Los criterios de información de Akaike (*AIC*) propuesto por Akaike ([7]), y el de información bayesiano (*BIC*) propuesto por Schwartz ([8]), implican elegir aquel modelo minimice las cantidades definidas como sigue:

$$AIC: = 2k - 2 \ln(L(\hat{\beta}))$$

$$BIC: = k \ln(n) - 2 \ln(L(\hat{\beta}))$$

siendo $\hat{\beta}$ el estimador máximo verosímil de

$\beta = (\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q, d, \sigma)$, L la función de verosimilitud del modelo y k la cantidad de parámetros que incluye el modelo ($k = p + q + 2$).

La interpretación es la siguiente: minimizar *AIC*, equivale a maximizar $-AIC = 2 \ln(L(\hat{\beta})) - 2k$ que se puede interpretar como elegir el modelo que tenga mayor verosimilitud, pero penalizando de acuerdo a la cantidad de parámetros que pueda tener, ya que es bien sabido que cuantos más parámetros tenga el modelo, se sobreajusta el mismo y la verosimilitud aumenta. Análogamente se interpreta el criterio *BIC*.

Los resultados obtenidos se adjuntan en las tablas II y III.

Los mismos se obtuvieron a partir de la serie observada hasta la última semana de 2012 (5408 datos).

q\p	0	1	2	3	4
0	-5.248,84	-6.472,769	-6.670,57	-6.682,234	-6.683,039
1	-6.571,997	-6.678,425	-6.678,396	-6.684,986	-6.684,106
2	-6.673,838	-6.679,542	-6.682,659	-6.684,732	-6.684,303
3	-6.677,147	-6.690,134	-6.688,812	-6.688,824	-6.689,302
4	-6.677,833	-6.688,757	-6.689,253	-6.689,102	-6.227,836

Tabla II.- Valores del AIC para distintos valores de p y q .

q\p	0	1	2	3	4
0	-5.229,053	-6.446,386	-6.637,592	-6.642,660	-6.636,87
1	-6.545,614	-6.645,446	-6.638,823	-6.638,816	-6.631,341
2	-6.640,86	-6.639,968	-6.636,490	-6.631,968	-6.624,943
3	-6.637,573	-6.643,964	-6.636,047	-6.629,463	-6.623,346
4	-6.631,704	-6.635,992	-6.629,892	-6.623,146	-6.155,284

Tabla III.- Valores del BIC para distintos valores de p y q .

Vale la pena comentar que en el ajuste de los distintos modelos ARFIMA, se observó una gran variabilidad de los parámetros estimados en el sentido de que modelos de muy alta y similar verosimilitud entre sí, se dan para valores muy disímiles de los parámetros. Por ejemplo, para la estimación del parámetro de memoria larga d , en algunos casos se dio que el modelo de máxima verosimilitud a veces se obtenía para parámetros d negativos, o incluso cercanos a 0.5. Los valores del AIC y BIC que se ven en la tabla, no siempre se corresponde con el modelo de mayor verosimilitud, si no que se tomó el de mayor verosimilitud entre los que dieron una estimación de d positiva. Cabe destacar asimismo que cuando el modelo más verosímil resultó con $d < 0$, muchas veces el que le seguía era el más verosímil entre los de memoria larga, además de tener una verosimilitud muy cercana a la óptima. La variabilidad se hace más notoria incluso cuando aumentamos los valores de p y q .

El criterio de minimizar el AIC da como resultado $p = 1$ y $q = 3$. Si se sigue el criterio de minimizar el BIC la sugerencia sería modelar con $p = 1$ y $q = 1$. Se observa también que los valores para p y q variando entre 1 y 4 de ambos índices, son bastante similares. El tercer criterio a considerar fue el de minimizar las diferencias en valor absoluto entre las autocorrelaciones observadas y las del modelo ajustado punto a punto en las primeras 300 diferencias. El valor 300 fue elegido arbitrariamente, representando aproximadamente un 5% de rezago de la longitud total de la serie. Para calcular las autocorrelaciones del modelo estimado en cada caso (según el valor de p y de q), se simularon 100 series de longitud igual a la observada (con los parámetros estimados por máxima verosimilitud) y se promediaron las autotcorrelaciones punto a punto.

q\p	0	1	2	3	4
0	69.07	30.33	29.42	12.82	10.25
1	75.68	13.71	18.25	7.72	11.94
2	30.79	19.08	14.27	6.23	9.52
3	26.23	6.208	6.28	6.24	5.94
4	20.57	6.104	6.33	6.37	60.65

Tabla IV.- Valores de las suma de diferencias en valor absoluto entre las autocorrelaciones observadas y las simuladas según los distintos valores de p y de q .

Ajustar las autocorrelaciones tiene la importancia de que si deseamos modelar a través de un proceso que sea gaussiano (a los datos desestacionalizados y centrados), la función de autocorrelación determina unívocamente al proceso. Los resultados se muestran en la Tabla IV. Más explícitamente, los resultados de la Tabla IV son los distintos valores de la función

$$f(p, q) = \sum_{i=0}^{300} \left| \rho(i) - \bar{\rho}(i) \right| \text{ siendo } \rho(i) \text{ la autocorrelación de rezago } i \text{ de la serie observada}$$

desestacionalizada, mientras que $\bar{\rho}(i)$ es el promedio de las autocorrelaciones de las 100 series

$$\text{simuladas, es decir } \bar{\rho}(i) = \frac{1}{100} \sum_{j=1}^{100} \rho_j(i) \text{ donde } \rho_j(\cdot) \text{ es la función de autocorrelación de la } j\text{-ésima simulación.}$$

Como se ve, el criterio de seleccionar el modelo minimizando la Tabla IV, sugiere utilizar $p = 4$, $q = 3$. Como se ve en la Tabla IV, en este caso el comportamiento es más errático que en las tablas II y III. Como cuarto criterio se tuvo en cuenta el caso en que más se ajustaran las curvas de intensidad para la serie de datos de aportes complexivos a las represas hidroeléctricas trabajo que puede verse en [6], que son curvas que permiten medir de alguna forma la persistencia de sequías, este criterio dio como resultado el caso $p = 3$, $q = 1$. Si bien éste caso no optimizó ninguno de los otros tres criterios, estuvo muy cerca de los valores óptimos de los mismos y será el que utilizamos como representante de los modelos ARFIMA que se

mostrarán en el resto del trabajo. Cabe destacar que se trabajó con los otros tres modelos en donde se obtuvieron los óptimos según cada criterio y no se observaron diferencias significativas en cuanto a su performance. Para $p = 3$, $q = 1$ el modelo estimado quedó con los siguientes parámetros:

$$\hat{\phi} = (\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3) = (1.28876, -0.632609, 0.172675), \quad \hat{\theta} = 0.452115, \quad \hat{d} = 0.18569, \\ \hat{\sigma}^2 = 0.28994 \text{ (varianza estimada del ruido blanco).}$$

Es decir que si le llamamos $\{X_t\}$ a la serie original y $\{D_t\}$ a la serie desestacionalizada y centrada, el modelo ajustado quedaría en la forma:

$$(1 - 1.28876B + 0.632609B^2 - 0.172675B^3)(1 - B)^{0.18569} D_t = \\ (1 - 0.452115B) \varepsilon_t.$$

3.2 Modelos de Ornstein-Uhlenbeck fraccionarios.

Definición. Dado $H \in (0, 1)$ se dice que $\{B_H(s)\}_{s \in \mathbb{R}}$ es un puente browniano fraccional si es un proceso que cumple las siguientes condiciones:

- es gaussiano, centrado,
- tiene trayectorias continuas (casi seguramente),
- tiene covarianzas dadas por $E(B_H(s)B_H(t)) = \frac{1}{2}(|s|^{2H} + |t|^{2H} - |t-s|^{2H})$.

El movimiento browniano fraccional fue definido por Mandelbrot & Van Ness ([9]) como ejemplo de una familia de procesos autosimilares.

Todo puente browniano fraccional, tiene incrementos estacionarios, ergódico, y es de memoria larga para $H > 1/2$. Se observa además que coincide con el puente browniano estándar para el caso en que $H = 1/2$.

Definición. Se dice que $\{X_t\}_{t \in \mathbb{R}}$ es un proceso de Ornstein-Uhlenbeck fraccionario si y sólo si se puede representar como $X_t = \sigma \int_{-\infty}^t e^{-\lambda(t-s)} dB_H(s)$ siendo $\{B_H(s)\}_{s \in \mathbb{R}}$ un puente browniano fraccional. Notación: $\{X_t\}_{t \in \mathbb{R}} \sim FOU(\lambda, \sigma, H)$.

Todo FOU, es estacionario, gaussiano, centrado, ergódico y de memoria larga para $H > 1/2$ (Cheridito et al, 2003) y puede ser visto como una interpolación a tiempo continuo de un proceso ARFIMA($1, d, 0$) (ver [10]).

También se prueba en [11] que $\{X_t\}_{t \in \mathbb{R}} \sim FOU(\lambda, \sigma, H)$ entonces, es la única solución estacionaria de la ecuación diferencial estocástica

$$dX_t = -\lambda X_t dt + \sigma dB_H(t).$$

Los parámetros del $FOU(\lambda, \sigma, H)$ fueron estimados mediante las fórmulas propuestas en [12].

Para obtener las fórmulas que nos permiten estimar los parámetros, se considera la serie observada (que debe ser desestacionalizada) como X_1, X_2, \dots, X_n que se toman como las observaciones a tiempo equiespaciado de un proceso continuo $\{Y_t\}_{0 \leq t \leq T}$, indizado en un intervalo de tiempo $[0, T]$ para algún T más explícitamente se considera $X_i = Y_{i\Delta}$ para cada i . Se necesita en primer lugar elegir el valor de T , con el mismo, se define $\Delta_n := T/n$. Luego se debe elegir $a = (a_0, a_1, \dots, a_k)$ que puede ser cualquier filtro de longitud $k+1$ y orden $L \geq 2$, esto es $\sum_{i=0}^k a_i i^p = 0$ para todo $p = 0, 1, 2, \dots, L-1$, $\sum_{i=0}^k a_i i^L \neq 0$. Dado un filtro $a = (a_0, a_1, \dots, a_k)$ de orden L y longitud $k+1$, se define el filtro $a^2 := (a_0, 0, a_1, 0, \dots, 0, a_k)$ que queda de orden L y longitud $2k+1$. Luego se define la variación cuadrática de la serie asociada a un filtro a mediante $V_{n,a} := \frac{1}{n} \sum_{i=0}^{n-k} \left(\sum_{j=0}^k a_j X_{i+j} \right)^2$. Por último, se estiman los parámetros mediante las siguientes fórmulas:

$$\hat{H} = \frac{1}{2} \log_2 \left(\frac{V_{n,a^2}}{V_{n,a}} \right),$$

$$\hat{\sigma} = \left(\frac{-2V_{n,a}}{\Delta_n^{2\hat{H}} \sum_{i=0}^k \sum_{j=0}^k a_i a_j |i-j|^{2\hat{H}}} \right)^{1/2},$$

$$\hat{\lambda} = \left(\frac{2 \sum_{i=1}^n X_i}{n \hat{\sigma}^2 \Gamma(2\hat{H} + 1)} \right)^{-\frac{1}{2\hat{H}}}.$$

Considerando la serie desestacionalizada hasta la última semana de 2012 (son 104 años que totalizan 5408 datos), se tomó como unidad de tiempo un año, por lo que se tomó $T = 104$, y se utilizó el siguiente filtro de Daubechies de orden 2 (ver [13]):

$$a = (0.482962913144543, -0.8365163037378, 0.224143868242013, 0.1294395225512).$$

Los valores estimados con sus desviaciones estimadas se presentan en el siguiente cuadro:

	λ	σ	H
Valores estimados	0.43056	0.69239	0.70737
Desviación estándar	0.01594	0.06045	0.01436

Se probaron las estimaciones utilizando otros posibles filtros sin que existieran diferencias significativas en las estimaciones de los modelos. Sí se observaron diferencias importantes si uno cambia los posibles valores de T . La estimación de H no depende de T pero sí las de σ y λ por

lo que al variar T es lógico que varíen las estimaciones. La elección de $T = 104$ se realizó teniendo en cuenta la naturaleza de los datos, para poder interpretar el intervalo de tiempo en años.

4. Predicciones semestrales de los distintos modelos.- Antes de comenzar a describir los resultados obtenidos con los dos modelos, es de destacar que los datos deben carecer de componente estacional, por lo que a la serie original es necesario desestacionalizarla, por ello, tanto para la utilización de los modelos ARFIMA como el de FOU, se estudiaron dos alternativas. En una de ellas, se desestacionalizó la serie de la manera más frecuentemente utilizada, que fue la comentada en la sección 1 (restando a cada dato semanal la media semanal y dividiendo entre su desvío), pero también se probó antes de desestacionalizar, aplicar a los datos la siguiente transformación $Y_t = \phi^{-1}(F(X_t))$ siendo los datos observados $\{X_t\}$, ϕ , la función de distribución de la normal típica y F la distribución empírica de los datos, para luego desestacionalizar la nueva serie por el procedimiento anteriormente descrito. Esta transformación en algún sentido regulariza la serie y fue aplicada con éxito primeramente en [14] aplicada semana a semana y en [6] al conjunto de todos los datos, en la modelación de la serie de aportes hidrológicos a las represas de Uruguay. En dichos casos se obtuvieron mejoras en la modelación de las series estudiadas, sin embargo, en esta oportunidad la situación fue a la inversa, tanto las verosimilitudes obtenidas, como los valores de los índices AIC y BIC , fueron mejores las ajustadas a los modelos desestacionalizados sin utilizar esta transformación, que la desestacionalizada luego de aplicar la transformación. También el valor predictivo fue mejor el obtenido por los modelos no transformados. En los resultados que se mostrarán en el resto del trabajo, se omiten los obtenidos mediante dicha transformación.

4.1 Predicción a partir del modelo ARFIMA(p, d, q).- Si $\{D_t\}$ es una serie desestacionalizada

ARFIMA(p, d, q), y le llamamos $c_{s,d} = \frac{(-1)^s \Gamma(1+d)}{\Gamma(d-s+1)\Gamma(s+1)}$ entonces $Y_t^{(d)} = \sum_{s=0}^{+\infty} c_{s,d} D_{t-s}$ es un

ARMA(p, q) lo cual implica que

$$Y_t^{(d)} = \phi_1 Y_{t-1}^{(d)} + \dots + \phi_p Y_{t-p}^{(d)} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \text{ utilizando ahora que } c_{0,d} = 1 \text{ obtenemos}$$

$$D_t = -\sum_{s=1}^{+\infty} c_{s,d} D_{t-s} + \phi_1 \sum_{s=0}^{+\infty} c_{s,d} D_{t-1-s} + \dots + \phi_p \sum_{s=0}^{+\infty} c_{s,d} D_{t-p-s} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \text{ lo que}$$

permite calcular el valor de D_t si uno tiene los valores observados hasta el tiempo $t-1$ de la serie así como los valores $\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q}$. Estos últimos valores no son observables, pero son ruido blanco.

Esta última fórmula nos permite obtener predicciones, ya que si tenemos la serie observada hasta el instante $t-1$ con los que ajustamos el modelo, sustituyendo los valores de los parámetros por sus estimaciones, y los $\varepsilon_{t-1}, \dots, \varepsilon_{t-q}$ por los residuos del modelo, mientras que el valor de ε_t lo sustituimos por su valor esperado que es 0, estamos así obteniendo una fórmula para utilizar en la predicción. Utilizando como valor al instante t el valor predictivo calculado \hat{D}_t y repitiendo el procedimiento se pueden obtener predicciones de tantos valores futuros como se desee.

4.2 Predicciones a partir del modelo FOU.- Debido a que todo modelo FOU (λ, σ, H) es un proceso gaussiano, centrado, se tiene que la distribución conjunta de las observaciones D_1, D_2, \dots, D_t es normal multivariada con vector de medias nulo y matriz de covarianzas Σ que depende de los parámetros (λ, σ, H) . Si queremos obtener predicciones k pasos para adelante, lo hacemos calculando la esperanza de $(D_{t+1}, D_{t+2}, \dots, D_{t+k})$ condicionado al vector (D_1, D_2, \dots, D_t) .

Si particionamos la matriz de covarianzas del vector $(D_1, D_2, \dots, D_t, D_{t+1}, D_{t+2}, \dots, D_{t+k})$ en $\Sigma = \begin{pmatrix} \Sigma_{t,t} & \Sigma_{t,k} \\ \Sigma_{k,t} & \Sigma_{k,k} \end{pmatrix}$ siendo $\Sigma_{t,t}$ la matriz de covarianzas de (D_1, D_2, \dots, D_t) y $\Sigma_{k,k}$ la matriz de covarianzas de $(D_{t+1}, D_{t+2}, \dots, D_{t+k})$. Es conocido que en el caso de la normal multivariada, la estimación de la esperanza de $(D_{t+1}, D_{t+2}, \dots, D_{t+k})$ condicionado al vector (D_1, D_2, \dots, D_t) es igual a $(\hat{D}_{t+1}, \hat{D}_{t+2}, \dots, \hat{D}_{t+k})^T = \hat{\Sigma}_{k,t} (\hat{\Sigma}_{t,t})^{-1} (D_1, D_2, \dots, D_t)^T$ donde $\hat{\Sigma}_{k,t}$ y $\hat{\Sigma}_{t,t}$ se obtienen sustituyendo en las matrices $\Sigma_{k,t}$ y $\Sigma_{t,t}$ los valores de λ, σ y H por sus estimaciones $\hat{\lambda}, \hat{\sigma}$ y \hat{H} . Todas las predicciones que aparecen en la Figura IX y siguientes así como los valores de la Tabla V, y debido a las dificultades para invertir $\hat{\Sigma}_{t,t}$ para valores grandes de t en las predicciones que se hicieron en base a esta metodología, se tomaron en cuenta únicamente los 520 datos anteriores al instante t , de este modo se invirtieron matrices de 520×520 .

4.3 Performance de las predicciones de ambos modelos.- Se realizaron predicciones a 5 años con 11 inicializaciones distintas. Más explícitamente, se consideraron los datos observados desestacionalizados y centrados hasta la última semana de 1990, con dichos datos se estimó el modelo FOU y el ARFIMA con $p=3$ y $q=1$, con dichos modelos ajustados se procedió a predecir los valores con un horizonte de 5 años. Luego se repitió el procedimiento con los datos hasta la última semana de 1992 (se estimaron los modelos y se realizaron predicciones a 5 años), luego con los datos hasta la última semana de 1994 y así sucesivamente, inicializando cada dos años hasta utilizar los datos hasta la última semana de 2010. Lo que se observó en general es que la mayor parte de las veces el ARFIMA realizó mejores predicciones aunque no siempre resultó así. En la Figura IX, vemos un ejemplo de un año (1991) que fue predicho mejor por ARFIMA que por FOU y en la Figura X un ejemplo donde el resultado fue al revés. Por otro lado, se observó también que el comportamiento del ARFIMA resultó más estable, en el sentido de que en todos los casos se apartó muy poco de los promedios semanales. En la Figura XIII se ve un ejemplo donde la forma de la curva de predicción 5 años hacia adelante es similar en ambos casos y similar a la de promedios semanales, mientras que en la Figura XII se observa cómo el cambio de inicialización en la curva de predicción afecta mucho más al FOU que al ARFIMA. Se observa también que en la medida que hacemos predicciones a más largo plazo, las predicciones semanales de ambos modelos tienden a los promedios semanales. Esto se ve en las figuras XI y XII, mientras que en la Figura XIII se observa además de las curvas con las predicciones de ambos modelos, la curva con los promedios semanales en color verde, donde se observa claramente cómo en la medida que pasan los años las dos curvas de predicciones tienden a la de promedios semanales. Dejando de lado la comparación entre los modelos ARFIMA y FOU, se observa que semana a semana, las predicciones no son precisas. Este hecho es lógico puesto que por ejemplo en el caso de la figura IX, se utiliza la información de la serie hasta la última semana de 1990 y con ella se predicen las 52 semanas siguientes. En la medida en que van transcurriendo

las semanas de 1991, la información nueva que generan estas nuevas semanas no está siendo usada para predecir las siguientes semanas del año, y por ello, las predicciones de cada semana, tienden al promedio de esa semana. Esto no implica, como podría parecer a primera vista, que los modelos considerados no sean buenos. Si se incorporara la información de cada semana que se va observando al ajuste del modelo y luego se predice la semana siguiente, se logran predicciones mucho más cercanas a las realmente observadas.

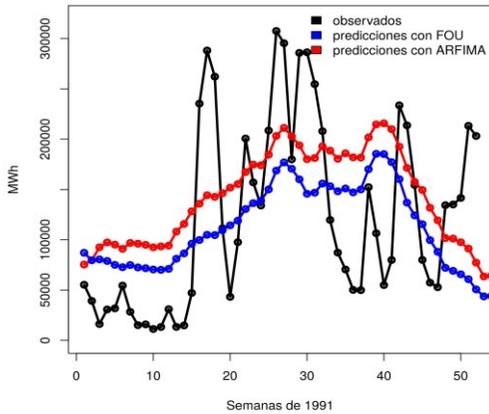


Figura IX.- Predicciones semanales para 1991.

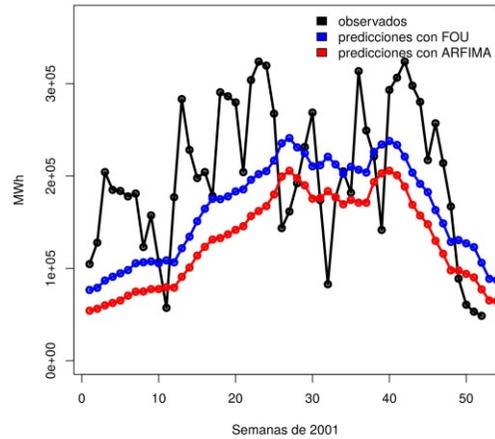


Figura X.- Predicciones semanales para 2001.

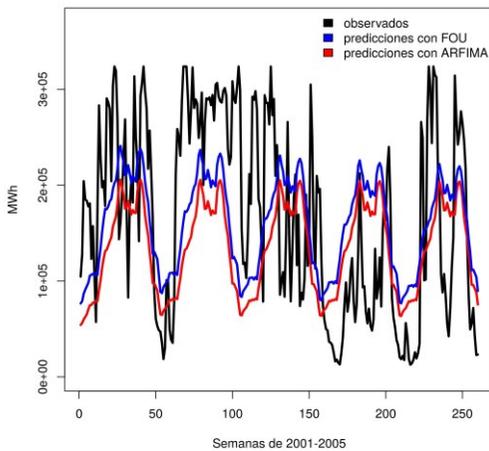


Figura XI.- Predicciones a 5 años, modelos estimados hasta la última semana de 2000.

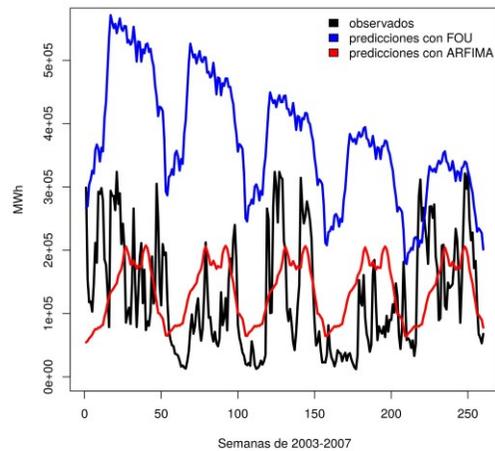


Figura XII.- Predicciones a 5 años, modelos estimados hasta la última semana de 2002.

Debido al interés por predecir la energía afluente semestral, para cada una de las inicializaciones definidas líneas más arriba, se procedió a acumular las predicciones por semestre y a compararlas con las realmente observadas. Como era de esperarse, dado que ARFIMA en general tuvo mejores predicciones semana a semana que FOU, en los acumulados semestrales también tuvieron un comportamiento similar, es decir, mejores predicciones del ARFIMA en general, aunque en algunos casos FOU logró mejores predicciones. Para cada una de las 11 inicializaciones también se procedió a simular 100 curvas de predicción, para obtener intervalos

de predicción al 95% para cada uno de los semestres. A modo de ejemplo, en la Tabla V se muestran los resultados para el modelo estimado hasta la última semana de 2006, los valores reales para cada semestre para los años 2007-2011, los predichos por cada uno de los modelos y los intervalos de predicción al 95% así como los valores mínimo y máximo simulados por cada semestre.

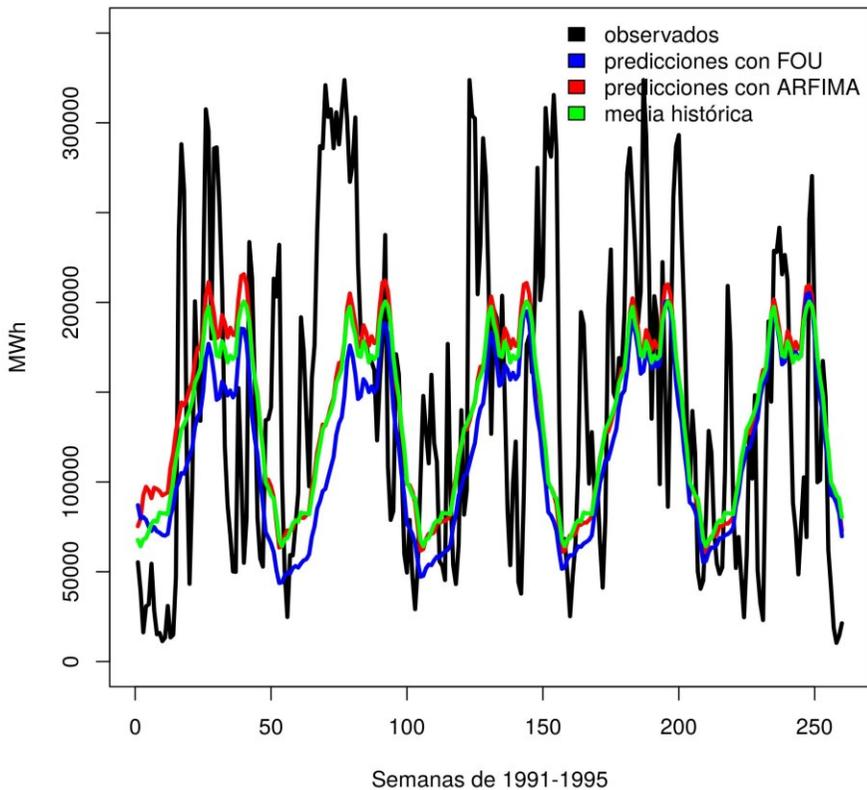


Figura XIII.- Predicciones semanales a 5 años, modelos estimados hasta la última semana de 1990, superpuestos con la curva de promedios semanales.

En la fila donde dice ARFIMA100, se hace el promedio de las predicciones de las 100 curvas. Estos valores se muestran en las figuras XIV y XV. En muchas ocasiones, se observa lo que ocurre en la Figura XIV y es que los valores observados se escapan durante los 5 años fuera de los intervalos de predicción al 95%, inclusive varias veces se escapan aún de las franjas del mínimo y del máximo obtenidos de las 100 simulaciones. Esto nos muestra el modesto desempeño que tienen las predicciones cuando uno se aleja de la última observación, como se ve en la Figura XIII, las predicciones además de parecerse mucho a la curva de los promedios semanales, tienen poca varianza como se observa de las curvas roja (intervalos de predicción al 95%) y azul (mínimos y máximos de las 100 simulaciones) de las figuras XIV y XV.

	Primer semestre	Segundo semestre	Primer semestre	Segundo semestre	Primer semestre
--	------------------------	-------------------------	------------------------	-------------------------	------------------------

	2007	2007	2008	2008	2009
REAL	4.412240	4.446235	1.404098	3.377962	0.869556
ARFIMA100	2.796792	3.918708	2.809075	4.021593	2.812773
MINIMO	1.464456	2.629202	1.613902	2.704164	1.208540
MAXIMO	4.089541	5.844321	4.124988	5.636088	4.073193
INF 95%	1.784214	2.946789	1.658774	2.960446	1.822494
SUP 95%	3.984718	5.088234	3.893858	4.959410	3.883271
	Segundo semestre 2009	Primer semestre 2010	Segundo semestre 2010	Primer semestre 2011	Segundo semestre 2011
REAL	5.972594	4.537426	3.995628	2.712170	4.282879
ARFIMA100	4.051263	2.794881	4.098451	2.872259	4.045553
MINIMO	2.704164	1.611089	2.688968	1.552078	2.415487
MAXIMO	5.636088	3.874258	5.288510	4.382919	5.510244
INF 95%	2.960446	1.715710	3.029256	1.871064	3.055835
SUP 95%	4.956410	3.787667	5.134616	4.088332	5.222814

Tabla V.- Predicciones semestrales (en millones de MWh) a partir de 100 simulaciones ARFIMA durante 5 años con el modelo estimado hasta la última semana de 2006.

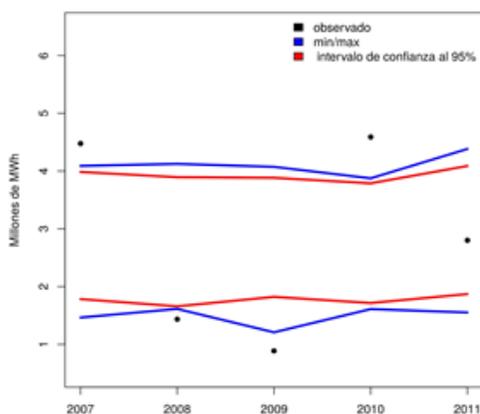


Figura XIV.- Predicciones de los primeros semestres para 2008-2011.

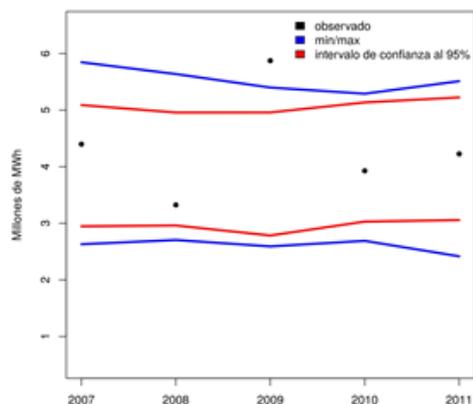


Figura XV.- Predicciones de los segundos semestres para los años 2008-2011.

El modelo ARFIMA además de haberse comportado en general mejor que el modelo FOU para predecir, también fue el que tuvo mejor ajuste. No se incluyen en este trabajo las medidas que lo muestran (dado que el trabajo está centrado en la comparación predictiva), pero las mismas

incluyen la estimación de la densidad de los datos, la densidad de los aportes de cada uno de los semestres así como el ajuste a la función de autocorrelación.

En [15] se ve que bajo ciertas hipótesis (las que se verifican en este caso), un ARFIMA puede ser representado como un $AR(\infty)$. A continuación mostraremos los 10 primeros sumandos del mismo, para poder interpretar la contribución de los datos anteriores aun instante dado t .

En este caso, utilizando el modelo estimado como en todo este trabajo con $p=3$, $q=1$ se utilizaron todos los datos de la serie observada hasta la última semana de 2012, resulta que la descomposición de la serie desestacionalizada queda de la siguiente manera:

$$D_t = -1.0222D_{t-1} + 0.33414D_{t-2} - 0.0873D_{t-3} - 0.0285D_{t-4} - 0.0118D_{t-5} \\ - 0.0058D_{t-6} - 0.0036D_{t-7} - 0.0027D_{t-8} - 0.0023D_{t-9} - 0.0021D_{t-10} + \dots + \varepsilon_t.$$

Se observa que los coeficientes tienden a cero de manera muy lenta, lo cual está lógicamente vinculado al hecho de que estamos utilizando un modelo de memoria larga.

5. Conclusiones.- Este trabajo de investigación estuvo centrado en ajustar modelos ARFIMA y FOU, para luego ver la capacidad predictiva a largo plazo, con inicializaciones en distintos años, prediciendo 5 años hacia adelante, sin actualizar la información que se va obteniendo al pasar de las semanas donde se van obteniendo nuevos datos. Se muestra cómo ambos modelos en su predicción van tendiendo a la curva de las medias semanales. Si bien en algunos casos FOU predijo mejor que ARFIMA en los acumulados de algunos semestres, en términos generales, ARFIMA predijo mejor que FOU, además de haber obtenido un mejor ajuste a la serie observada. Por otro lado se observó un carácter mucho más volátil de los modelos FOU, mientras que los ARFIMA tuvieron un comportamiento mucho más estable. Como conclusión final, se puede decir que de este trabajo se infiere que la serie de datos estudiada en este trabajo presenta un exponente de Hurst mayor que $\frac{1}{2}$ por lo que es razonable ajustarle un modelo de memoria larga. De las dos familias de modelos consideradas, los ARFIMA tuvieron un mejor desempeño.

6. Trabajo a futuro.- Varios aspectos se pueden abordar a partir de lo realizado hasta aquí. En primer lugar, se pueden mejorar notoriamente las predicciones semanales si para estimar el valor de una semana, se ajusta el modelo hasta la semana inmediata anterior, para luego con toda esta información realizar la predicción de la semana en cuestión. Luego de ello, es posible también mejorar las predicciones eligiendo convenientemente una combinación lineal de las predicciones realizadas por cada modelo. En segundo lugar, se pueden explorar otros tipos de modelos tanto para el ajuste como para la predicción que surjan de la agregación de dos o más FOU por ejemplo, para tener mayor cantidad de parámetros y poder comparar su rendimiento frente a los ARFIMA. Si nos centramos únicamente en las predicciones, se podría profundizar el estudio incorporando modelos de memoria corta (ARMA por ejemplo) y métodos modernos de aprendizaje automático como las redes neuronales. Finalmente, los ARFIMA ajustados en todos los casos (variando p y q), fueron estimados por máxima verosimilitud. Este presupone modelos ARFIMA con ruidos blancos gaussianos, pero no se han explorado los resultados estimando en primer lugar d por algún método que no presuponga la distribución del ruido blanco para luego ajustar por mínimos cuadrados los parámetros correspondientes a los parámetros AR y MA y utilizar de esta forma ruidos blancos no gaussianos.

7. Referencias

- [1] HURST, H. E., *Long term storage in reservoirs*, Trans. Am. Soc. Civil Eng., 1951, 116, 770-799.

- [2] TAQQU, M. S., TEVEROVSKY, M. and WILLINGER, W., *Estimators for long range dependence: an empirical study*, Fractals, 1995, 3, 785-788.
- [3] BERAN J., *Statistics for Long Memory Processes*, 1994, Chapman and Hall, New York.
- [4] PALMA, W., *Long Memory Time Series-Theory and Methods*, 2007, John Wiley, Hoboken, NJ.
- [5] GRANGER, C. W. J. and JOYEUX, R., *An introduction to long memory time series models and fractional differencing*, J. Time Ser. Anal., 1980, 1(1), 15-29.
- [6] KALEMKERIAN, J., *Performance of the Different FARIMA Models Adjusted to the Hydrological Contributions to the Dams*, Journal of Basic and Applied Scientific Research, 2014, 4(10) 27-39.
- [7] AKAIKE, H., *A new look to at the statistical model identification*, IEEE Transactions on Automatic Control 1974, (19), 6 716-723.
- [8] SCWARTZ, G. E., *Estimating the dimension of a model*. Annals of Statistics 1978, 6(2): 461-464.
- [9] MANDELBROT, B. And VAN NESS, J., *Fractional Brownian motions, fractional noises and applications* SIAM Review, 1968, 10: 422-437.
- [10] COMTE and RENAULT, *Long memory continuous time models*. Journal of Econometrics, 1996, 73, 101-149.
- [11] CHERIDITO, P., KAWAGUCHI, H and MAEJIMA, M., *Fractional Ornstein-Uhlenbeck processes*. Electronic Journal of Probability, 2003, 8(3):1-14.
- [12] BROUSTE, A., and IACUS, S., *Parameter estimation for the discretely observed fractional Ornstein-Uhlenbeck process and the Yuima package*, Computational Statistics, 2013, 28(4): 1529-1547.
- [13] DAUBECHIES, I., *Ten Lectures on Wavelets*, 1992, SIAM.
- [14] CHAER, R. and ZEBALLOS, R., *Modelo Simplificado de Central con Embalse con fines didácticos*, IeeLatinamericaTransactions, 2006, vol.4, no.3, mayo 2006.
- [15] BROCKWELL, P. J. and DAVIES R. A., *Time Series: Theory and Method* 2nd ed., 1991, Springer, New York.