Conclusiones Fisiológicas y Estimadores Fractales*

María Eugenia Torres**, Gastón Schlotthauer

En este trabajo comparamos tres métodos diferentes utilizados para estimar el exponente de Hurst, y analizamos su eficiencia cuando son aplicados a series de datos de diferentes longitudes. Se analizan series temporales de fBm sintetizada pura y con tendencias sinusoidales superpuestas. Mostraremos que los tres métodos aquí discutidos, DFA, basado en wavelets y de variaciones discretas, no sólo son altamente dependientes de la longitud de la señal, sino también del orden o número de los momentos (polinómico, regularidad wavelet o variaciones discretas). Para longitudes de datos suficientemente grandes (superiores a 2^{12}), los métodos basados en wavelets y de variaciones discretas mostraron ser menos sesgados y más estables para señales fBm simuladas. Mostraremos que el método de DFA, más utilizado en el ambiente biomédico, es el que proporciona peores estimaciones, arrojando resultados ambiguos cuando son aplicados a señales biológicas de diferentes longitudes o con diferentes parámetros de estimación, sin que pueda considerarse a ninguno de los otros dos como métodos confiables en el momento de desear obtener resultados de relevancia física o fisiológica. Los resultados obtenidos indican que debería procederse con más cautela cuando se trata de obtener conclusiones fisiológicas a partir de estimaciones realizadas a partir de señales reales.

Palabras clave: Dependencias de largo alcance - autosimilaridad - exponente de Hurst - fisiología

*) Artículo que expone resultados del proyecto de investigación PID-UNER 6060/2000-2005 y PICT-2002 11-12700, Agencia Nacional de Promoción Científica y Tecnológica, desarrollados en la Facultad de Ingeniería, Universidad Nacional de Entre Ríos en 2000-2005. Informe Final PID-UNER aprobado por Res. CS 141/06; presentado para publicación en agosto 2006 y aceptado en diciembre 2006.
**) Doctora en Matemáticas, Profesora Titular Ordinaria de *Matemática III*, Fac. Ingeniería, UNER, Directora del proyecto. E-mail: metorres@ceride.gov.ar

Physiological conclusions and fractal estimators*

María Eugenia Torres** ; Gastón Schlotthauer

In this paper we compare three different methods used to estimate Hurst exponent and we analyze their efficiency while they are applied in data series of different lengths. We analyze synthesized fBm time series, pure and with superimposed sinusoidal trends. We show that the three methods here discussed, DFA, wavelets based and discrete variations, no only are highly dependent on the signal length, but also on the order or number of moments (polynomic, wavelet regularity or discrete variations). For large enough data length (higher than 2^{12}), the methods based on wavelets and discrete variations have shown to be the less biased and the more stable ones for simulated fBm signals. We show that DFA method, the most widely used among the biomedical community, is the one that provides worst estimators, displaying ambiguous results while applied to real biological signals of different lengths and for different parameters in the estimator. More over, nor DFA neither the other two methods provide reliable estimators if physical or physiological conclusions are wanted. The obtained results indicate that more caution should have to be taken while trying to derive physiological conclusions from the Hurst exponent estimations obtained from real data.

Key words: Long range dependency - Self similarity - Hurst exponent - Physiology

*) This paper presents the results of the research projects PID-UNER 6060/2000-2005 and PICT-2002 11-12700, National Agency for the Promotion of Science and Technology, developed at the Faculty of Engineering, National University of Entre Rios -UNER- 200-2005, Final Report approved in 2006, submitted in August 2006 and accepted in Dicember 2006.
**) Maths Doctor, Regular Professor in *Mathematics III*, Faculty of Engineering, UNER, Director of the aboved mentioned research projects. E-mail: metorres@ceride.gov.ar

I. Introducción

El diagnóstico clínico y las investigaciones básicas dependen de manera crítica de la habilidad para registrar y analizar señales fisiológicas. Ejemplos de esto son los registros de tasa cardiaca de pacientes con alto riesgo de muerte súbita, los registros electroencefalográficos (EEG) en epilepsia y otros desórdenes, así como fluctuaciones hormonales u otras señales de mensajeros moleculares en dinámicas neuroendocrinas.

Sin embargo, ni el análisis clínico del enfermo ni el análisis de laboratorio de dichas señales han ido de la mano de la mayoría de los avances tecnológicos que permiten el registro y almacenamiento de conjuntos de datos masivos correspondientes a señales continuamente fluctuantes. Sorprendentemente, aún cuando recientemente se ha demostrado que esas señales, típicamente complejas, tienen una naturaleza no lineal, no estacionaria y de naturaleza de no-equilibrio, las herramientas para analizar dichos datos con frecuencia aún suponen linealidad, estacionariedad y condiciones de casi-equilibrio. Tales técnicas convencionales incluyen métodos estadísticos (media, desvió Standard y otras características como histogramas), junto con el análisis espectral.

Un excitante descubrimiento de mediados de la década de 1990 es que dichos conjuntos de datos complejos pueden contener *información oculta*, definida aquí como información no extraíble con métodos convencionales de análisis. Tal información promete ser de valor clínico (predicción de muerte cardiaca súbita en pacientes ambulatorios, o catástrofes cardio-pulmonares durante procedimientos quirúrgicos), así como relacionarla con mecanismos de funciones saludables o patológicas. El análisis fractal es uno de los enfoques nuevos más prometedores para extraer dicha información oculta en las series temporales fisiológicas. Esto en parte se debe al hecho de que la ausencia de características de escalas temporales (o espaciales) - el sello del comportamiento fractal - puede conferir importantes ventajas biológicas, relacionadas con la adaptabilidad de respuesta (**Goldberger** y col., 1990; **Bunde y Havlin**, 1994; **Khokha** y col., 1996; **Goldberger**, 1996; **Goldberger**, 1997).

Sorprendente es que condiciones de autosimilaridad, correlación de largo alcance (LRD) o mas genéricamente de *scaling* no sólo han sido reportadas en el análisis de señales biológicas tales como electroencefalogramas, variabilidad cardiaca, DNA o recuento de glóbulos

rojos, sino que además aparecen ligados al análisis de series temporales de la economía, climatología en incluso sismología. En bioingeniería, esto ha sido utilizado como base para proponer técnicas de diagnóstico y realizar importantes afirmaciones de índole fisiológica y biofísica, que en ciertos casos ponen en duda hechos biológicos aceptados por décadas. Es entonces imperativo realizar una análisis certero de este tipo de fenómenos en datos experimentales. La presencia de tendencias suaves de índole determinística (de origen externo -aparatos- o interno, pero relacionada con mecanismos biológicos tales como al respiración), superpuestas a los datos, con frecuencia pueden impedir una correcta estimación de los parámetros de un fenómeno de scaling realmente existente.

Es el exponente de Hurst el parámetro que permite determinar la existencia del fenómeno de scaling. En este trabajo compararemos tres métodos usados para su estimación. Utilizando datos simulados y datos biomédicos reales mostramos que estos estimados son sesgados y altamente sensibles a la longitud de la serie de datos. Los resultados permiten afirmar que debería tenerse mas cuidado al momento de realizar afirmaciones de relevancia fisiológica derivadas de conclusiones de fenómenos de scaling obtenidas con series temporales cortas.

I.1. Objetos Fractales y Procesos Autosimilares

A los efectos de favorecer la comprensión, presentaremos aquí algunos conceptos básicos de objetos fractales, procesos autosimilares y procesos estacionarios con dependencias de largo alcance. Discutiremos algunos de los métodos mas conocidos en la actualidad los que serán de utilidad para analizar señales fisiológicas con dinámicas fractales, tales como la regulación del latido cardíaco humano, y la regulación del paso de la marcha humana, que está bajo el control voluntario del sistema nervioso central. Distintos autores han utilizado estas técnicas para analizar y comparar las salidas de dichos sistemas en salud y en enfermedad.

I.2. Breve reseña histórica

Una forma geométrica se dice autosimilar de manera determinística

si la misma estructura geométrica es observada, independientemente de la distancia desde donde se la mire y la forma. Para una definición matemáticamente exacta ver **Mandelbrot** (1983) y **Mandelbrot** (1977). Ejemplos clásicos de la geometría fractal son el triángulo de Sierpinsky (ver figura 1) y la curva de Von Koch (figura 2). En ambas observamos que una parte es geométricamente similar al todo. Los objetos fractales geométricos mas simples pueden ser construidos por medio se sistemas iterados de funciones (IFS) (**Barnsley**,1988; **Falconer**, 1990).

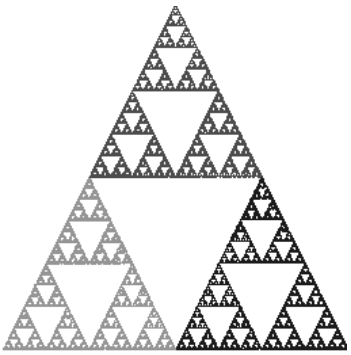


Figura 1: Triángulo de Sierpinsky

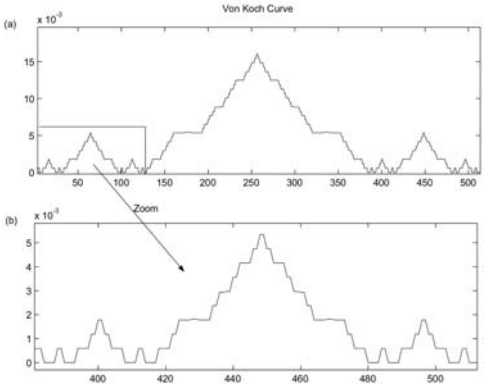


Figura 2: Curva de Von Koch

Los procesos autosimilares fueron introducidos por Kolmogorov en 1941 (**Kolmogorov**, 1941). En particular, sus investigación sobre turbulencias lo motivaron a definir el movimiento Browniano (Bm) y otros procesos relacionados. Sin embargo, no fue sino hasta 1968 que se comprendió la relevancia estadística de tales procesos, cuando Mandelbrot y coautores (**Mandelbrot** y **Van Ness**, 1968; **Mandelbrot** y **Wallis**, 1968; **Mandelbrot** y **Wallis**, 1969 (a, b y c)) los introdujeron en la estadística. La idea de autosimilaridad es aún mas antigua. Así por ejemplo, Mandelbrot (**Mandelbrot**, 1977; **Mandelbrot**, 1983) refiere a los dibujos de Leonardo da Vinci de campos turbulentos que exhiben coexistencia de remolinos ("eddies") de todos los tamaños y en consecuencia los identifica como autosimilares.

Desde el punto de vista histórico son de particular interés las series de datos correspondientes a las altitudes mínimas de agua en el río Nilo

entre los años 622 y 1281, medidos en Roda Gauge, cerca de El Cairo (**Tousson**, 1925, pp 366-385). El análisis de éstas series temporales y otras similares, dieron lugar al descubrimiento de lo también denominado *efecto Hurst* (**Hurst**, 1951). Motivado por los resultados experimentales de Hurst, Mandelbrot y coautores introdujeron posteriormente el *ruido Gaussiano fraccionario* (fGn) como un modelo con *memoria de largo alcance* (long memory) (**Mandelbrot**, 1968; **Mandelbrot** y **van Ness**, 1968; **Mandelbrot** y **Wallis**, 1969).

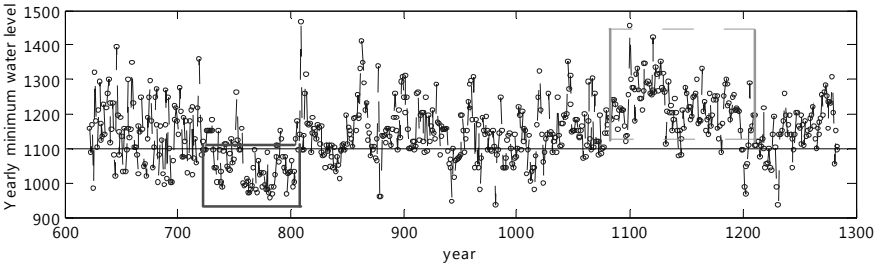


Figura 3: Niveles del Río Nilo entre los años 622 y 1281

I.3. El efecto de José y el efecto de Hurst

Desde tiempos remotos, el río Nilo ha sido conocido por su característico comportamiento de largo alcance: largos períodos de sequía eran seguidos por nuevos períodos de inundaciones. Estas últimas tenían el efecto de fertilizar el suelo, de modo que se correspondían con períodos de cosechas abundantes. Sobre bases especulativas, se podría pensar que este hecho ya fue descrito en la Biblia (Génesis 41, 29 - 30): "Vendrán siete años de gran abundancia en toda la tierra de Egipto, y detrás de ellos vendrán siete años de escasez, que harán se olvide toda la abundancia en la tierra de Egipto, y el hambre consumirá la tierra" (Sagrada Biblia, 1965). No existen registros de aquellos lejanos tiempos, sin embargo, los correspondientes a los años 622 - 1281 (ver Fig. 3) exhiben un comportamiento de *largo alcance* que podrían dar una explicación a los siete buenos años seguidos de siete malos a los que se refiere el Génesis. Coexistían largos períodos en que el nivel máximo tendía a mantenerse elevado con otros períodos, también prolongados,

con niveles bajos. En su conjunto, la serie temporal aparece como estacionaria. Cuando sólo se miran períodos cortos, aparecen ciclos o tendencias locales. De todos modos, en el conjunto global no hay ciclos persistentes aparentes. Por el contrario, parecería que ocurriesen ciclos en casi todas las frecuencias, superpuestos y en sucesión aleatoria. En consecuencia, no existe una tendencia global. En referencia a la expresión bíblica, Mandelbrot la denominó *efecto de José* (Joseph effect) (**Mandelbrot**, 1977 a, b; **Mandelbrot** y **Van Ness**, 1968, **Mandelbrot** y **Wallis**, 1968)

El famoso hidrólogo Hurst observó estas características en 1951 cuando investigaba como regularizar el flujo del río Nilo. Mas específicamente, su descubrimiento puede describirse del siguiente modo: Supongamos que deseamos calcular la capacidad de un reservorio tal que sea ideal en un período de tiempo dado, entre *t* y *t*+*k*. Supongamos, por simplicidad, que el tiempo es discreto y que no hay pérdidas de acumulación debidas a evaporación o pérdidas del reservorio. Entendemos por capacidad ideal que la salida es uniforme, que en tiempo *t*+*k* el reservorio está tan lleno como en el instante *t* y que el mismo nunca se rebalsa. Indiquemos con X_i el ingreso en el tiempo *i* y sea

$Y_j = \sum_{i=1}^{j} X_i$ la acumulación de flujo ingresante hasta el instante *j*.

Entonces la capacidad ideal será igual a:

$$R(t,k) = \max_{0 \le i \le k}\left[\left(Y_{t+i} - Y_t\right) - \frac{i}{k}\left(Y_{t+i} - Y_t\right)\right] \qquad (1)$$

$$= -\min_{0 \le i \le k}\left[\left(Y_{t+i} - Y_t\right) - \frac{i}{k}\left(Y_{t+i} - Y_t\right)\right]'$$

R(t,k) se denomina el rango ajustado (rescaled range). Para estudiar las propiedades que son independientes de la escala, se estandariza *R(t,k)*:

$S(t,k) = \sqrt{k^{-1}\sum_{i=t+1}^{t+k}\left(X_i - \overline{X}_{t,k}\right)^2}$, donde $\overline{X}_{t,k} = k^{-1}\sum_{i=t+1}^{t+k} X_i$. Observar que $S^2(t,k)$ es igual a $\frac{k-1}{k}$ veces la varianza usual de $X_{t+1},\ldots,X_{t+k}$. El cociente

$$R/S = \frac{R(t,k)}{S(t,k)}$$

se denomina rango ajustado re-escalado o estadístico R/S. Graficando

$\log(R/S)$ vs. $\log(k)$ para varios valores de k, Hurst observó que para valores grandes de k, $\log(R/S)$ se movía entorno a una línea recta cuya pendiente era mayor a ½. Probabilísticamente, esto significa que para valores grandes de k,

$$\log\{E[R/S]\} \cong a + H\log k \text{, con } H > 1/2 \cdot \quad (2)$$

Este descubrimiento empírico estaba en contradicción con los resultados para procesos de Markov, procesos de mezclas y otros procesos estocásticos usualmente considerados por ese tiempo. Para cualquier proceso estacionario, con dependencia de corto plazo, R/S debería comportarse asintóticamente como una constante por $\sqrt{k}$. Entonces, para valores grandes de k, $\log(R/S)$ debería dispersarse aleatoriamente entorno a un línea recta con pendiente 1/2. Los descubrimientos de Hurst sobre los datos del río Nilo y sobre otros registros hidrológicos, geológicos y climatológicos, en relación a que R/S se comporta como una constante por k^H para algún $H>$ 1/2 son conocidos con el nombre de *efectos de Hurst*. Mandelbrot y coautores demostraron que dicho efecto puede modelizarse mediante el denominado ruido Gaussiano fraccionario con **parámetro de autosimilaridad** $H \in (1/2, 1)$ (H por Hurst). Veremos mas adelante como este proceso se define de manera exacta y porqué H es denominado parámetro de autosimilaridad. Dicho modelo corresponde a un proceso estacionario cuya correlación $\rho(k)$ decae a cero con una ley exponencial de k del tipo

, (3)

para k tendiendo a infinito, donde c_p es una constante positiva.

Un proceso estacionario con correlaciones de decaimiento lento ec. (3) se denomina *proceso estacionario con memoria larga* o "long-range dependence" (LRD) o con dependencia fuerte (en contraste con aquellos procesos que tienen correlaciones sumables, también denominados procesos con *memoria corta* o "short-range correlations"' o dependencia débil). La ecuación (3) es esencialmente equivalente a una densidad espectral con un polo en cero. Recordar que la densidad espectral f está definida por (**Priestley**, 1981):

$$f(\lambda) = \frac{\sigma^2}{2\pi} \sum_{k=-\infty}^{\infty} \rho(k) e^{ik\lambda}$$

Entonces la ec. 3 implica que $f(\lambda) \cong c_f |\lambda|$, para $\lambda \to 0$ donde c_f es una constante positiva. Entonces, para $\lambda < 1$, f tiende a infinito en el origen. En contraste, para depende $\lim_{n\to\infty} n^{-1} \sum_{i \neq j} \rho(i,j) < \infty$ ance

para procesos con co $\sum_{k=-\infty}^{\infty} \rho(k) =$ ies que decaen tan rápido a cero que constante < ∞.

En la figura 4 observamos que para los datos del río Nilo, α = 1-0.80249 = 0.1975, si se toma la regresión lineal desde $\lambda = 2^{-4}$, siendo c_f=11.6825.

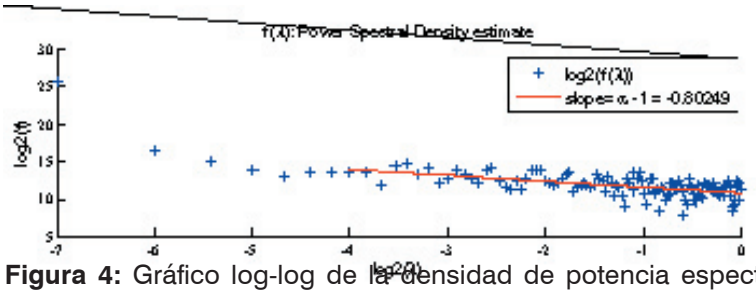


Figura 4: Gráfico log-log de la densidad de potencia espectral de los datos del Río Nilo entre los años 622 y 1281. (Reproducido de **Beran**, 1994).

II. Definiciones

AUTOSIMILARIDAD. Un proceso estocástico Y_t se dice *autosimilar* con parámetro de autosimilaridad H si y sólo si $Y_t = c^{-H} Y_{ct}$ en el sentido de las distribuciones, para cualquier factor de estiramiento C.

INCREMENTOS ESTACIONARIOS. Se dice que un proceso estocástico Y_t tiene incrementos estacionarios si y sólo si las distribuciones del proceso $X_t = Y_{t+1}-Y_t$ no dependen de *t*.

FBm. El movimiento fraccionario Browniano B_H es el único proceso Gaussiano autosimilar con incrementos estacionarios. El proceso de sus

Gaussiano autosimilar con incrementos estacionarios. El proceso de sus incrementos G_H (k) = B_H (k+1)-B_H (k) es conocido como *ruido Gaussiano fraccionario* (fGn).

SI *0<H<0.5* se dice que le proceso es *autosimilar (SS)*, si *0.5<H<1* el proceso tiene *dependencias de largo alcance (LRD)* y si *H=0.5* el proceso es el clásico *movimiento Browniano (BM)*.

El concepto de autosimilaridad quiere decir que las propiedades estadísticas en ambos miembros de la ecuación son idénticas. En otras palabras, un proceso autosimilar, *y(t)*, con parámetro *H* tiene la misma distribución de probabilidad que un proceso adecuadamente re-escalado $a^H y(t / a)$. En consecuencia, los pasos muestrales típicos de un proceso autosimilar se ven cualitativamente iguales, independientemente de la distancia desde la cual se los mire. En contraste con la autosimilaridad determinística (por ejemplo el triángulo de Sierpinsky), esto no significa que la misma figura se repita a medida que nos acercamos. Es mas bien la impresión general la que se mantiene igual.

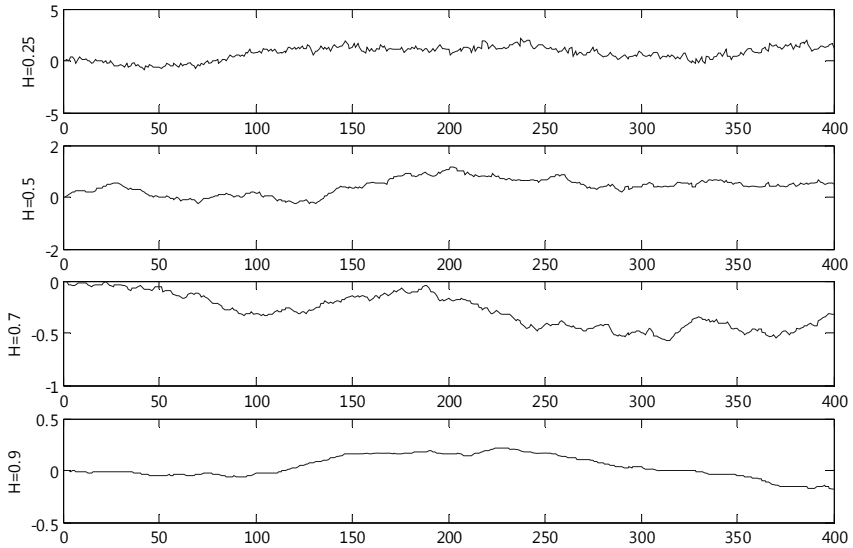


Figura 5: Movimiento Browniano fraccionario para H = 0.25, 0.5, 0.7 y 0.9.

Aun cuando no es estacionario, el fBm tiene incrementos estacionarios, lo que significa que las propiedades estadísticas de los procesos $B_H(t+s) - B_H(t)$ sólo dependen de la variable de retardo *s*. Más aun, este proceso de incrementos es autosimilar. Esta autosimilaridad inherente a la estructura del fBm, tiene como consecuencia que una realización individual de tal proceso es una curva fractal. Es por este motivo que el parámetro *H* del fBm está relacionado con la tosquedad (*rougthness*) de las muestras de fBm: cuanto mas grande sea *H*, mas suave es la serie temporal. Como ejemplo, mostramos la Figura 5, donde se han generado muestras de fBm para $H = 0.25, 0.5, 0.7$ y 0.9.

Las series temporales acotadas pueden ser transformadas en procesos autosimilares por medio de su integración. De todos modos, otro desafío al que deben hacer frente los investigadores que aplican este tipo de análisis a señales fisiológicas es que las mismas con frecuencia son altamente no-estacionarias. Una definición general y simplificada caracteriza a una serie estacionaria cuando su media, desviación Standard y los momentos de orden superior, así como sus funciones de correlación son invariantes ante traslaciones temporales. Las señales que no obedecen estas condiciones son no estacionarias. El procedimiento de integración realizado en el método de Hurst puede aun aumentar la no-estacionariedad de los datos originales.

Se han introducido diferentes métodos para superar este problema. En **Peng** y col. (1994) y **Peng** y col. (1995) se introdujo uno basado en modificaciones al análisis del error cuadrático medio del random walk, denominado análisis de remoción de fluctuaciones (Detrended fluctuations analysis - DFA) ampliamente aplicado al análisis de datos biológicos. Otros enfoques incluyen el análisis multifractal, basado en análisis wavelets (**Abry** y col., 1993; **Abry** y col., 1995) y el método de variaciones discretas (**Coeurjolly**, 2001).

III. Métodos de estimación de parámetros de auto similaridad

El problema cuando analizamos señales reales radica en estimar el valor de *H*. En **Kantelhardt** y col. (2001) se menciona que, mientras que el análisis espectral (la transformada de Fourier) y el método de la transformada wavelet modulo máximo (WTMM) (**Muzy** y col., 1991; **Muzy** y col., 1993; **Bacry** y col., 1993) analizan directamente la serie temporal,

el método de remoción de fluctuaciones (Detrending Fluctuation Analysis, DFA) está basado en la teoría de random walk (**Shlesinger** y col., 1987; **Avraham** y **Havlin**, 2000), de modo similar al análisis re-escalado de Hurst (**Hurst** y col., 1965; **Feder**, 1988).

DFA cuantifica la presencia o ausencia de correlaciones de largo alcance (**Ho**, 1997). Este método ha sido ampliamente utilizado en los últimos años (**Buldyrev** y col., 1993; **Ossadnik**, 1994; **Peng** y col., 1994; **Peng** y col., 1995; **Haussdorf** y col. 1995; **Haussdorf** y col., 1996). Se afirma que, dado que en los métodos basados en la teoría de random walk (incluido el DFA), se trabaja sobre las sumas acumuladas de las series temporales, se reduce el nivel de ruido producido durante la adquisición de los registros. El factor de Fano (**Fano**, 1947) y el de Allan (**Barnes** y **Allan**, 1966) (ver también **Viswanathan** y col., 1997), emplean un contexto similar, pero dichos métodos no reducen las tendencias en los datos.

Para realizar una detección confiable de correlaciones de largo alcance es necesario distinguir las tendencias de las fluctuaciones a largo alcance intrínsecas en los datos. Las tendencias pueden ser causadas por efectos externos (Por ejemplo los efectos invernadero, las variaciones estacionales de los registros de temperatura) y habitualmente se supone que poseen comportamientos oscilatorios suaves y monótonos. La presencia de tendencias fuertes en los datos pueden dar lugar a falsas detecciones de correlaciones de largo alcance si sólo un método es aplicado o si los resultados no son cuidadosamente interpretados y según **Ossadnik** y col. (1994) y **Peng** y col. (1995), ésta es justamente una de las ventajas del método DFA dado que sistemáticamente elimina tendencias de diferentes órdenes, tal como lo hace el método basado en wavelets. Mostraremos que esta afirmación no es correcta, y que por el contrario, el método de DFA es altamente sensible al tipo y cantidad de tendencia global presente en la señal.

A tal fin, en este trabajo comparamos tres métodos de estimación del exponente de Hurst, reconocidos en la literatura por presentar interesantes características tanto en términos de su desempeño como de su robustez ante la superposición de tendencias de tipo aditivo (sinusoidales, exponenciales, polinómicas, etc.)

III.1. Análisis de Fluctuaciones con Tendencias Removidas

Este método, conocido como DFA, consiste básicamente en analizar las variaciones de las sumas acumuladas normalizadas en media, a las cuales se les remueven las tendencias polinomiales. Dada una serie temporal *(Y)*, el perfil del registro de longitud n se determina: $X(i) = \sum_{k=1}^{i}(Y_k - \overline{Y})$, donde $\overline{Y}$ indica el valor medio de *(Y)*. El perfil *X(i)* se corta en $n_s \cong [n/s]$ segmentos no solapados de igual longitud s para diferentes valores de *s*. Luego, se calcula la tendencia local para cada segmento mediante un ajuste lineal no pesado de los datos correspondientes. La serie temporal sin tendencia de duración de segmento s está definida como: $X_s(i) = X(i) - p_N(i),$ donde $p_N(i)$ es el polinomio de ajuste del η-ésimo segmento. Para cada uno de los n_S segmentos se calcula la variancia de la función de fluctuación de DFA según $F(s) = 1/n_s \sum_{\eta=1}^{n_s} F_s^2(\eta),$ donde $F_s^2(\eta) = 1/s \sum_{i=1}^{s} X_s^2[(\eta-1)s+i]$ es la función de fluctuación. Puede mostrarse que *F(s)* s^{α} para valores grandes de *s*, donde el exponente de fluctuación α está relacionado con el exponente de correlación (**Kantelhardt** y col., 2001) Ha sido demostrado que $\hat{H}_{DFA}$ =1- α/2 cuando *Y* es fBm con parámetro *H* (**Beran**, 1994). El estimador de *H* consiste en un ajuste lineal no pesado en coordenadas $log_2(s)$ versu$(F(s))$ *(s)*:

. (4)

Parámetros prácticos: Como estamos interesados en comparar el método con aquellos basados en wavelet, cortamos el perfil en segmentos equi-espaciados loga [illegible] las décadas $s=2.s_{min}$ y $s=2.s_{max}$. Los valores de s_{min} y $\hat{H}_{DFA} = 1 - \frac{1}{2}\log$ nado de manera tal que $2.s_{min} \leq 2+N$, donde *N* es el orden del polinomio, y s_{max} de forma tal que tenga al menos 10 ventanas de longitud máxima. Para señales cortas hemos seleccionado s_{min} = 3 y $s_{max}=\log_2(n)-5$ ($\log_2(n)-4$).

III.2. Análisis Ondita

Sea $\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t-k)$ la plantilla de la ondita madre ψ. Sea $\{d_Y(j,k) = \langle Y, \psi_{j,k} \rangle\}$ el conjunto de los llamados coeficientes ondita discretos del proceso Y. Sea

donde n_j es el número disponible de coeficientes ondita en la octava *j*.

discretos del proceso Y. Sea

$$S_n(j) = \frac{1}{n_j} \sum_k \left| d_Y(j,k) \right|^2,$$

donde n_j es el número disponible de coeficientes ondita en la octava j. Observar que $S_n(j)$ es, desde un punto de vista estadístico la varianza de los coeficientes ondita en la escala *j-ésima*. Esencialmente $n_j = 2^{-j}\, n$ donde n es la longitud de los datos. El estimador de H consiste en un ajuste lineal no pesado en las coordenadas $log_2(2^j) = j$ versus $log_2\, S_n(j)$:

$$\hat{H}_W = 1/2 \sum_{j=j_1}^{j_2} w_j \log_2 S_n(j) - 1/2, \quad (5)$$

donde $w_j = (S_0\, j - S_1)/(S_0\, S_2 - S_{12})$, con $S_m = \sum_{j_1}^{j_2} j^m, m = 0,1,2.$

Cuando Y es fBm con parámetro H, se ha demostrado que es un estimador para H eficiente y débilmente sesgado (**Abry** y col., 1993; **Abry** y col., 2000; **Coeurjolly**, 2001).

Parámetros prácticos: En este trabajo, utilizamos onditas Daubechies con N momentos nulos y hemos seleccionado $j_1 = 3$ y $j_2 = log_2(n) - 5$.

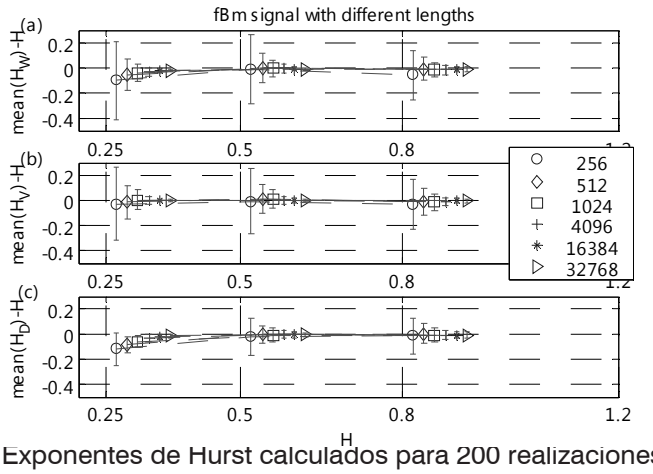


Figura 6: Exponentes de Hurst calculados para 200 realizaciones de fBm, simulados para cada longitud diferente. (a) Sesgo y desviación estándar correspondientes a $\hat{H}_D$ (b) Sesgo y desviación estándar correspondientes a $\hat{H}_W$. (c) Sesgo y desviación estándar correspondientes a $\hat{H}_V$.

III.3. Variaciones Discretas

Sea $f(t)=\sum_l a_l \delta(t-l)$, donde $\delta(t)$ indica la masa de Dirac. Se pide que a_l satisfaga: $\sum_{j=0}^{1} j^r a_j = 0$, para $r = 0,...,N-1$; y $\sum_{j=0}^{1} j^N a_j \neq 0$.

Note que el entero *N* juega un rol totalmente equivalente al del número de momentos nulos en el caso del método ondita. Con $f_{j,k}(t)=2_{-j/2}\, f(2^{-j}\, t-k)$ denotamos las plantillas de *f*.

Sea $\{V_Y(j,k)=\langle Y, f_{j,k}\rangle\}$ el conjunto de coeficientes de variación discreta del proceso *Y*. Sea $T_n(j)=1/n_j \sum_k |V_Y(j,k)|^2$, donde n_j es el número disponible de coeficientes en la octava *j*. Esencialmente $n_j=2^{-j}\, n$ donde n es la longitud de los datos. El estimador de H consiste en un ajuste lineal no pesado el las coordenadas $log_2(2^j) = j$ versus $log2\ T_n(j)$:

$$\hat{H}_V = 1/2 \sum_{j=j_1}^{j_2} w_j \log_2 S_n(j) - 1/2 \ , \quad (6)$$

donde los w_j se definen como arriba.

Cuando Y es fBm con parámetro H, se ha demostrado que $\hat{H}_W$ es un estimador de H eficiente, no sesgado y robusto (**Istas** y **Lang**, 1997; **Coeurjolly**, 2001).

Parámetros prácticos: en este trabajo, utilizamos sucesiones a_l tales que $\{s\, a_1 = (-1)^l\, 1!/(N!\,(N-1)!\,), \quad l = 0,\ldots,N\}$, donde s es una constante tal que $\sum_l a^2 l = 1$ y j_1 y j_2 han sido elegidos como antes.

IV. Resultados y discusión

IV.1. Metodología

Con el objeto de estudiar la performance de estimación de los tres estima-dores descriptos arriba, hemos sintetizado 200 realizaciones de fBm, $B_H(k)$, para diferentes valores de *H* y diferentes longitudes *n*, utilizando el método de matriz circulante (**Wood** y **Chan**, 1994). Hemos superpuesto tendencias aditivas T(k) a $B_H(k)$. La serie temporal *Y* a ser analizada se obtiene según: $Y_k = B_H(k) + \sigma_T\, T(k)$ donde σ_T = (max(T)-min(T))/std(B_H) para los casos en los que discutamos aspectos relativos a

la longitud de la señal. Los valores medios y las desviaciones estándar para los tres estimadores $\hat{H}_W, \hat{H}_V$ y $\hat{H}_D$ se obtienen promediando sobre todas las realizaciones.

Los valores de los parámetros prácticos *elegidos son: nbreal = 200 (para fBm) y 100 (para fBm+trend),* 2^8 n 2^{15}, H=[0.1, 0.25, 0.4, 0.5, 0.65, 0.7, 0.8, 0.9, 0.95], N=1,2,3,4; σ_T fue calculada en cada realización.

En la Sección IV.2, T(t) $\approx$0 y en la Sección IV.3 $T(t) = A\ \sin(2\pi\ f\ t)$ con $A = 1,\ f = 2^{-11}$.

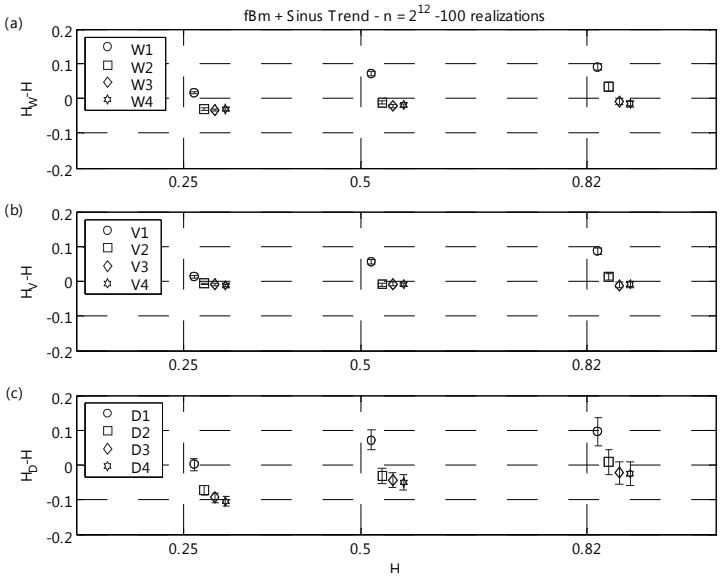


Figura 7: Exponente de Hurst estimado para 100 realizaciones de fBm (longitud $n=2^{12}$) con tendencia sinusoidal superpuesta. Los sesgos y desviación estándar corresponden a: (a) $\hat{H}_D$. (b) $\hat{H}_W$ .y (c) $\hat{H}_V$.

IV.2. Simulaciones numéricas I: Dependencia de la longitud de la señal

En esta sección, no se superpusieron tendencias a los incrementos de fBm (T(t) 0) y comparamos la performance del estimador con respecto a la duración n de la observación. Los sesgos y las desviaciones estándar

se comparan en la Fig. 6 para los tres estimadores.

En la Fig. 6 presentamos los resultados obtenidos para H = 0.25, 0.5 y 0.8. Podemos apreciar que el sesgo está relacionado con la longitud de la señal. Notamos que para valores grandes de n, los tres estimadores presentan sesgos casi despreciables. Sin embargo, cuando el número de muestras n desciende, todos los estimadores muestran un sesgo que aumenta mientras desciende H. Además se nota que para los valores más pequeños de H y n, el sesgo de $\hat{H}_D$ es significativamente grande comparado con los de $\hat{H}_V$ y $\hat{H}_W$, que permanecen comparables con una pequeña ventaja respecto al primero.

IV.3. Simulaciones numéricas II: Tendencia sinusoidal superpuesta a fBm

En esta sección, estudiamos los efectos de la presencia de una tendencia superpuesta T(t) = A sen(2 πf t), con A=1, y f $=2^{-11}$ sobre la performance de los estimadores.

En la Fig. 7 se muestran los resultados obtenidos para 100 realizaciones de longitud 2^{12}, comparando la performance de nuestros estimadores como función del orden. Nuevamente notamos que el sesgo de $\hat{H}_V$ y $\hat{H}_W$ es despreciable para un orden $\geq$ 2, y que sus desviaciones estándar son más pequeñas que las correspondientes a $\hat{H}_D$. Todos los estimadores mostraron un sesgo que se incrementa en tanto H aumenta cuando el orden (momento) es 1, pero que desciende mientras H crece para órdenes (momentos) mayores que 1. El sesgo en este caso es despreciable para los estimadores basados en ondita y en variaciones discretas, con una ventaja en el este último, dado que $\hat{H}_V$ presenta un sesgo de casi la misma magnitud para todos los valores de H considerados.

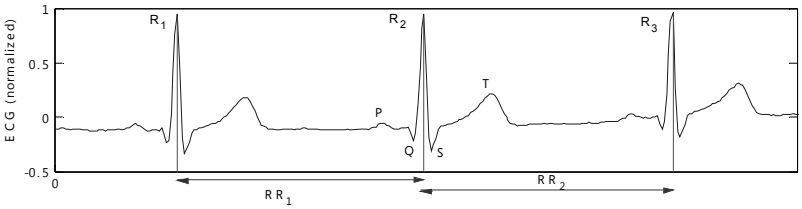


Figura 8: Electrocardiograma, donde se indican las ondas P, Q, R, S y T. Se señalan dos intervalos entre ondas RR: RR1 y RR2

IV.4. Datos Fisiológicos Reales: señales RR

Consideremos ahora señales temporales RR obtenidas de la base de datos Physionet European ECG (European ST-T database, en http://www.physionet.org/physiobank/database/edb/) correspondiente a fluctuacio-nes del ritmo cardíaco.

En la Figura 8 mostramos un electrocardiograma (ECG) típico y las ondas P, Q, R, S y T características y se indican además dos segmentos RR cuyas longitudes darán lugar a lo que se denominan los intervalos RR. Como puede observarse en la Figura 4, la serie puntual de intervalos RR normal presenta importantes irregularidades. En (**Goldberge**r y col., 2000; **Goldberger** y col., 2002; **Peng** y col., 2000) los autores aseguran que las series temporales del ritmo cardíaco humano muestran fluctuaciones "ruidosas". De acuerdo a los paradigmas fisiológicos clásicos basados en la homeostasis, tales sistemas deberían estar diseñados para disminuir el ruido y dirigirse a un estado similar a un equilibrio constante. Sin embargo, el análisis de las fluctuaciones de los latidos cardíacos bajo condiciones de estado de reposo aparente revela, de acuerdo a los mismos autores, la presencia de correlaciones de largo alcance. El descubrimiento de esta organización de largo alcance impone un desafío destacable a los esfuerzos contemporáneos para entender y, eventualmente, simular sistemas de control fisiológico. Modelos plausibles deberían considerar tales memorias de largo alcance.

Hemos calculado el exponente de Hurst para estos datos considerando diferentes longitudes de señal: $n = 2^8, 2^{10}, 2^{12}, 2^{13}$ y 2^{14}. En cada caso comparamos los tres métodos para diferentes polinomios y órdenes de onditas N = 1, 2, 3, 4, 5.

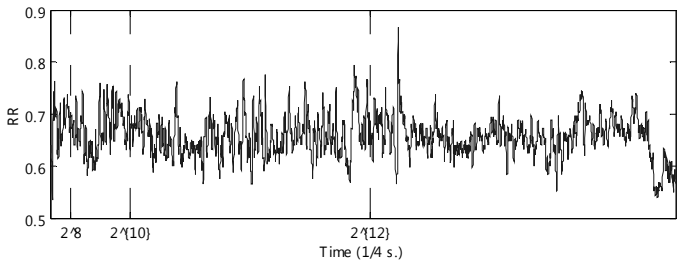


Figura 9: Segmento de señal RR.

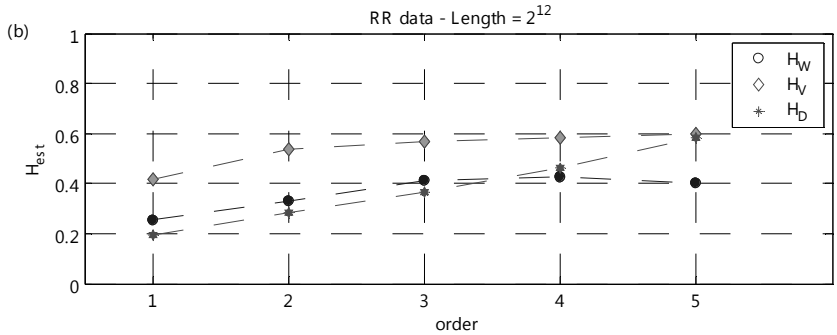


Figura 10: Valor estimado de H para un segmento de señal RR mostrada en (a) de longitud 212.

Valores considerados de los parámetros: s_{min}= 3; s_{max}=12-4 = 8; y órdenes 1, 2, 3, 4 y 5.

Como puede observarse en la Fig 10, los valores estimados de H para los tres métodos aquí discutidos son altamente sensibles al orden N utilizado en la estimación, cuando son aplicados al segmento RR de longitud 2^{12} que se muestra en la figura anterior. En efecto, observando la Fig. 10, no estamos en condiciones de asegurar si el estimado de H, $\hat{H}$, es o no mayor a 0.5. El método de variaciones discretas ofrece $\hat{H}_V > 0.5$ para momentos N=2-5, mientras que el método basado en onditas da $\hat{H}_W < 0.5$ para todos los valores de momento aquí considerados. En cuanto

a lo concerniente a DFA, $\hat{H}_D$ se incrementa constantemente en tanto el orden N va de 1 a 5.

Observemos ahora en la Fig. 11 el comportamiento de los tres métodos para diferentes longitudes y órdenes (momentos). Podemos apreciar que para N=1, tanto los métodos de ondita y DFA ofrecen estimaciones de H <0.5, mientras que el método de los incrementos hace lo mismo sólo para señales de longitud n>210. En tanto más aumentemos el orden (momento) observamos que, independiente-mente de la longitud de la señal, todos los estimados de H también aumentan. En particular, $\hat{H}_D$ y $\hat{H}_V$ llegan a valores levemente mayores que 0.5 en el caso N=5.

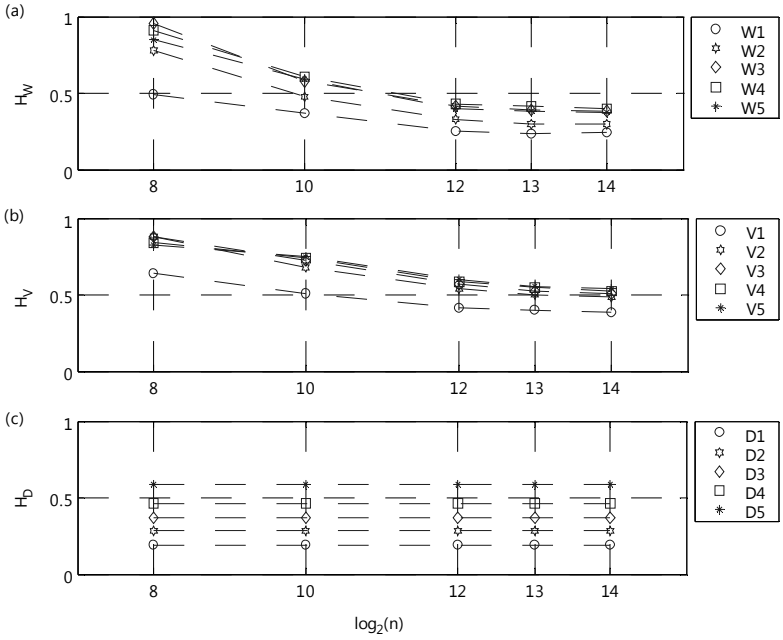


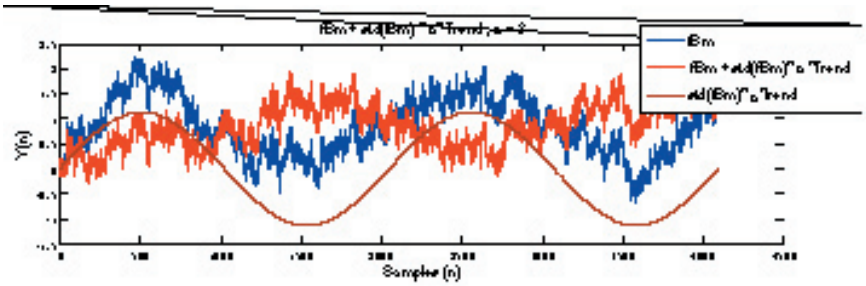
Figura 11: Valores estimados de H para segmentos de la señal RR mostrada en (a) de longitud 2^8, 2^{10}, 2^{12}, 2^{13} y 2^{14}. Valores de parámetros considerados:s_{min}= 3; s_{ma}x= longitud- 4; y órdenes 1,2,3,4y 5.

Si hubiésemos considerado solamente las estimaciones basadas en los métodos de variaciones discretas, hubiésemos dicho que esta serie temporal RR corres-ponde a un proceso LRD probablemente moviéndose hacia un estado de antipersistencia. Sin embargo, los resultados del DFA nos dicen que un LRD no es una opción tan clara.

Resultados similarmente contradictorios han sido encontrados en otras señales reportadas en la bibliografía. Destacamos que hemos usado en todos los casos las mismas bases de datos que los autores han puesto a disposición en Internet para reproducir sus resultados. En efecto, los resultados se reproducen para ciertos valores de momento y long adecuada de las series temporales, sin embargo, nuestro análisis muestra que los mismos no son suficientes para poder realizar conclusiones de índole biofísica o fisiológica.

IV.5. Simulaciones numéricas III: Sensibilidad a la "cantidad" de tendencia sinusoidal superpuesta a fBm.

Por último, mostramos algunos de los resultados que hemos obtenido al analizar la sensibilidad de estos estimadores con relación a la "cantidad" de tendencia contenida en la señal analizada (**Torres**, 2005). A modo de ejemplo mostramos en la Fig. 12 una señal de fBm con H=0.9, a la cual se le ha adicionado una tendencia sinusoidal haciendo $Y(t) = fBm(t) + \Delta.std(fBm).T(t)$, con *T(t)* como antes y Δ=3. Se puede apreciar que la presencia de la tendencia no aporta información visual de la periodicidad, excepto la provista por fBm, la cual aparenta contener cierta periodicidad. Se realizaron las estimaciones sobre 200 muestras para valores de H= 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 y 0.9, con valores de Δ=0.5, 1, 2 y 3 y 2^{12} muestras. Los valores medios de los estimados de H y su desviación Standard (std) se muestran en la tabla 1. De su observación surge que los tres métodos son sensibles a la cantidad de tendencia presente en la señal, si bien queda claro que los basados en wavelets y en variaciones discretas, presentan mejor performance, en particular para H<0.5, aun para valores grandes de Δ.



Esto queda puesto en evidencia en la Figura 13 donde mostramos los errores promedios obtenidos para $\Delta=0.5$, correspondientes a los estimadores de momentos (órdenes) igual a 1 y a 4. Observar que el error con DFA está entre 1 y 2, para todos los valores de H, siendo el valor teórico $0<H<1$. En particular, el error relativo para DFA, para valores de H pequeños es del orden del 800% y disminuye a cerca del 100% para H próximo a 0.9. Para los otros dos métodos el error relativo es de aproximadamente el 100% y asciende a un valor máximo del 150%. En la Tabla 2 mostramos el promedio de los valores medios y de los desvíos Standard de los estimados de H, según los tres métodos, para diferentes Δ y para los mo-mentos de orden 1,2, 3 y 4.

Si se observan los errores de la estimación realizada con el método DFA queda claro que cualquier inferencia que se realice a partir del mismo adolecerá de serias imprecisiones al momento de su interpretación en el mundo real de la física o de la biología. Similares resultados se obtuvieron para otro tipo de tendencias.

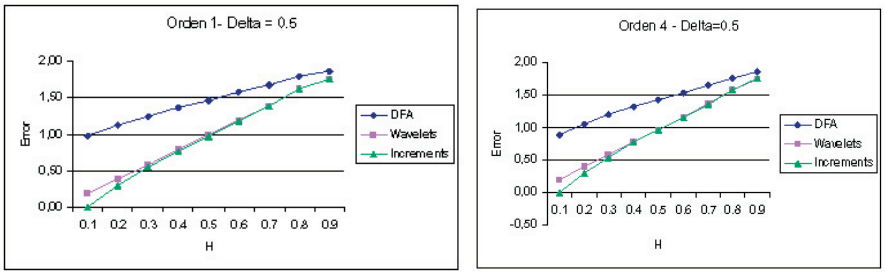


Figura 13: Errores promedio obtenidos en la estimación del exponente de Hurst para fBm con tendencia sinusoidal, para $\Delta=0.5$. Estimados con momentos (grados) de orden 1 (izquierda) y 4 (derecha).

V. Conclusiones

El objetivo de este trabajo fue comparar diferentes métodos para estimar el exponente de Hurst, y analizar su eficiencia cuando son aplicados a series de datos de diferentes longitudes. Hemos discutido el comportamiento de tres estimadores en el caso de 200 realizaciones de fBm sintetizada y 100 fBm con tendencia sinusoidal superpuesta. Hemos visto que los tres métodos aquí discutidos, DFA, basado en wavelets y de variaciones discretas, no sólo son altamente dependientes de la longitud de la señal, sino también del orden o número de los momentos (polinómico, regularidad wavelet o variaciones discretas). Para longitudes de datos suficientemente grandes (superiores a 2^{12}), los métodos basados en wavelets y de variaciones discretas mostraron ser menos sesgados y más estables para señales fBm simuladas.

Cuando analizamos señales de intervalo RR, que han sido reportadas como poseedoras de evidencias de dependencia de largo alcance (LRD), observamos que para longitudes de señal donde los tres métodos proporcionan estimados poco sesgados para series simuladas, en estas son altamente sensibles al orden o momento del método utilizado. De igual modo se observó alta sensibilidad a la longitud de la señal, con comportamientos similares para cada orden y valor de momento en los métodos basados en wavelets y de variaciones discretas. En conclusión, en el caso de datos reales, dependiendo de los métodos y parámetros

utilizados, se pueden obtener resultados indicando tanto autosimilaridad o una dependencia a corto o largo alcance. Los autores han observado comportamiento similar en otros datos provistos por otros colegas que sospechaban LRD o autosimilaridad en ellos.

Esta última observación y el hecho de que en las señales biológicas con frecuencia hay periodicidades inmersas, nos motivó a explorar la sensibilidad de los tres métodos ante tendencias sinusoidales superpuestas. Observamos que el método de DFA es el que proporciona peores estimaciones, sin que pueda considerarse a ninguno de los otros dos como métodos confiables al momento de desear obtener resultados de relevancia física o fisiológica.

Los resultados obtenidos indican que debería procederse con más cautela cuando se trata de obtener conclusiones fisiológicas a partir de estimaciones realizadas a partir de señales reales. Debe evitarse trabajar con insuficiente núme-ro de muestras. Los investigadores deberían utilizar varios estimadores para identificar dependencia de largo alcance o autosimilaridad. Sería también necesa-rio un análisis de descomposición de la señal, dado que este provee información útil sobre posibles tendencias ocultas, que distorsionan los valores estimados del exponente de Hurst.

las leyes fisiológicas bien establecidas. Una adecuada comprensión de los métodos utilizados tanto desde el punto de vista teórico como de sus limitaciones experimentales es fundamental en todo trabajo científico serio.

	Δ = 0.5	Δ = 1	Δ = 2	Δ = 3
Orden 1				
DFA	1,456 ± 0,047	1,482 ± 0,049	1,537 ± 0,053	1,587 ± 0,056
Wavelets	0,990 ± 0,031	1,027 ± 0,034	1,104 ± 0,040	1,173 ± 0,044
Increments	0,949 ± 0,031	0,996 ± 0,035	1,096 ± 0,042	1,184 ± 0,047
Orden 2				
DFA	1,433 ± 0,044	1,437 ± 0,045	1,446 ± 0,046	1,458 ± 0,049
Wavelets	0,982 ± 0,028	0,989 ± 0,029	1,008 ± 0,033	1,030 ± 0,037
Increments	0,941 ± 0,029	0,955 ± 0,031	0,990 ± 0,038	1,028 ± 0,045
Orden 3				
DFA	1,420 ± 0,043	1,420 ± 0,043	1,420 ± 0,044	1,420 ± 0,044
Wavelets	0,976 ± 0,029	0,977 ± 0,029	0,978 ± 0,029	0,980 ± 0,029
Increments	0,939 ± 0,028	0,941 ± 0,029	0,947 ± 0,029	0,954 ± 0,030
Orden 4				
DFA	1,414 ± 0,043	1,414 ± 0,043	1,414 ± 0,043	1,414 ± 0,043
Wavelets	0,974 ± 0,029	0,975 ± 0,029	0,975 ± 0,029	0,975 ± 0,029
Increments	0,936 ± 0,029	0,937 ± 0,029	0,938 ± 0,029	0,939 ± 0,029

Tabla 2: promedio de los valores medios y de los desvíos Standard de los estimados de H, según los tres métodos, para diferentes Δ y para los momentos de orden 1,2, 3 y 4. Tendencia sinusoidal

Referencias bibliográficas

ABRY P., FLANDRIN P., TAQQU M. S., VEITCH D. (2000). *Wavelets for the analysis, estimation and synthesis of scaling data.* New York: Wiley Interscience.

ABRY P., GONÇALVES P., FLANDRIN P. (1993.) Wavelet-based spectral analysis of 1/f processes. En: *Proceedings of IEEE International Conference on Acoustic, Speech*

	= 0.5	= 1	= 2	= 3
		Orden 1		
DFA	1,456 ± 0,047	1,482 ± 0,049	1,537 ± 0,053	1,587 ± 0,056
Wavelets	0,990 ± 0,031	1,027 ± 0,034	1,104 ± 0,040	1,173 ± 0,044
Increments	0,949 ± 0,031	0,996 ± 0,035	1,096 ± 0,042	1,184 ± 0,047
		Orden 2		
DFA	1,433 ± 0,044	1,437 ± 0,045	1,446 ± 0,046	1,458 ± 0,049
Wavelets	0,982 ± 0,028	0,989 ± 0,029	1,008 ± 0,033	1,030 ± 0,037
Increments	0,941 ± 0,029	0,955 ± 0,031	0,990 ± 0,038	1,028 ± 0,045
		Orden 3		
DFA	1,420 ± 0,043	1,420 ± 0,043	1,420 ± 0,044	1,420 ± 0,044
Wavelets	0,976 ± 0,029	0,977 ± 0,029	0,978 ± 0,029	0,980 ± 0,029
Increments	0,939 ± 0,028	0,941 ± 0,029	0,947 ± 0,029	0,954 ± 0,030
		Orden 4		
DFA	1,414 ± 0,043	1,414 ± 0,043	1,414 ± 0,043	1,414 ± 0,043
Wavelets	0,974 ± 0,029	0,975 ± 0,029	0,975 ± 0,029	0,975 ± 0,029
Increments	0,936 ± 0,029	0,937 ± 0,029	0,938 ± 0,029	0,939 ± 0,029

Referencias bibliográficas

ABRY P., FLANDRIN P., TAQQU M. S., VEITCH D. (2000). *Wavelets for the analysis, estimation and synthesis of scaling data.* New York: Wiley Interscience.

ABRY P., GONÇALVES P., FLANDRIN P. (1993.) Wavelet-based spectral analysis of 1/f processes. En: *Proceedings of IEEE International Conference on Acoustic, Speech and Signal Processing, ICASSP-93*, 27 al 30 de abril 1993, vol. III, pág. 237-240.

---------- (1995). Wavelets, spectrum estimation and 1/f processes. En: A. Antoniadis; G. Oppenheim (Eds.) (1995). *Wavelets and Statistics, Lectures Note in Statistics*. New York: Springer-Verlag, 103, pág. 15 -30.

AVRAHAM D. B, HAVLIN S. (2000). *Diffusion and Reactions in Fractals and Disordered System*. Cambridge: Cambridge University Press.

BACRY E., MUZY J. F., ARNEODO A. (1993). Singularity spectrum of fractal signals from wavelet analysis: exact results. En: *J. Stat. Phys.*, 70, pág. 635- 674.

and Signal Processing, ICASSP-93, 27 al 30 de abril 1993, vol. III, pág. 237-240.

---------- (1995). Wavelets, spectrum estimation and 1/f processes. En: A. Antoniadis; G. Oppenheim (Eds.) (1995). *Wavelets and Statistics, Lectures Note in Statistics*. New York: Springer-Verlag, 103, pág. 15 -30.

AVRAHAM D. B, HAVLIN S. (2000). *Diffusion and Reactions in Fractals and Disordered System*. Cambridge: Cambridge University Press.

BACRY E., MUZY J. F., ARNEODO A. (1993). Singularity spectrum of fractal signals from wavelet analysis: exact results. En: *J. Stat. Phys.*, 70, pág. 635- 674.

BARNES J. A., ALLAN D. W. (1966). A statistical model of flicker noise. En: *Proc. IEEE*, 54, pág. 176-178.

BARNSLEY M.F. (1988). *Fractals Everywhere*. Boston: Academic Press, Inc.

BERAN J. (1994). *Statistics for Long-Memory Processes*. New York, Chapman & Hall.

BULDYREV S. V., GOLDBERGER A. L., HAVLIN S., PENG C-K., STANLEY H. E., SIMONS M. (1993). Fractal landscapes and molecular evolution: modeling the myosin heavy chain gene family. En: *Biophys J*, 65, pp. 2673-2679.

BUNDE A., HAVLIN S., Eds. (1994). *Fractals in Science*. Berlin: Springer-Verlag.

COEURJOLLY J. F. (2001). Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. En: *Statistical Inference for Stochastic Processes*, 4, pp.199-227.

FALCONER K. J. (1985). *The Geometry of Fractal Sets* Cambridge: Cambridge University Press.

FANO U. (1947). "Ionization yield of rations. II. the fluctuations of the number of ions". En: *Phys. Rev.*, 72, pp. 26-29.

FEDER J. (1988). *Fractals*. New York: Plenum Press.

GOLDBERGER A. L. (1997). Fractal variability versus pathologic periodicity: complexity loss and stereotypy in disease. En: *Perspect Biol Med*, 40, pp. 543 -561.

GOLDBERGER A. L. (1996). Non-linear dynamics for clinicians: Chaos theory, fractals, and complexity at the bedside. En: *Lancet*, 347, pp.1312-1314.

GOLDBERGER A. L., AMARAL L. A. N., GLASS L., HAUSDORFF J. M., IVANOV P. C., MARK R. G., MIETUS J. E., MOODY G. B., PENG C.-K., STANLEY H. E. (2000). Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. En: *Circulation*, 101(23), pp. e215- e220.

GOLDBERGER A. L., AMARAL L. A. N., HAUSDORFF J. M., IVANOV P. C., PENG C.-K., STANLEY H. E. (2002). Fractal dynamics in physiology: Alterations with disease and aging. En: *Proceedings of the National Academy of Science of the United States of America*, 99(1), pp. 2466-2472.

GOLDBERGER A. L., RIGNEY D. R., WEST W. J. (1990). Chaos and fractals in human physiology. En: *Sci. Am.*, 262, pp. 42- 49.

HAUSDORFF J. M., PENG C-K., LADIN Z., WEI J. Y., GOLDBERGER A. L. (1995). Is walking a random walk?. Evidence for long-range correlations in the stride interval of human gait. En: *J. Appl. Physiol.*, 78, pp. 349-358.

HAUSDORFF J. M., PURDON P., PENG C-K., LADIN Z., WEI J. Y., GOLDBERGER A. L. (1996). Fractal dynamics of human gait: stability of long-range correlations in stride interval fluctuations. En: *J. Appl. Physiol.*, 80, pp. 1448-1457.

HO K. K. L., MOODY G. B., PENG C-K., MIETUS J. E., LARSON M. G., LEVY D., GOLDBERGER A. L. (1997). Predicting survival in heart failure cases and controls using

fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics. En: *Circulation*, 96, pp. 842-848.

HURST H. E. (1951). Long-term storage capacity of reservoir". En: *Trans Am Soc Civil Engineers*, 116, pp. 770-799.

HURST H. E., BLACK R. P., SIMAIKA Y. M. (1965). *Long-Term Storage. An Experimental Study*. London: Constable.

ISTAS J., LANG G. (1997). Quadratic variations and estimation of the local Hölder index of a Gaussian process. En: *Ann. Inst. Henri Poincaré*, 33(4), pp. 407 -436.

KANTELHARDT J. W., KOSCIELNY-BUNDE E., REGO H. H. A., HAVLIN S., BUNDE A. (2001). Detecting long-range correlations with detrended fluctuation analysis. En: *Physica A*, 295, pp. 441-454.

KHOKHA M.K., IANNACONNE P., Eds. (1996). *Fractal Geometry in Biological Systems: An Analytical Approach*. Boca Raton: CRC Press,

KOLMOGOROV A. N. (1941). The local structure of turbulence in incompressible viscous fluid for very large Reynolds number. En: *Dokl. Akad. Nauk USSR*, 30, pp. 299-303.

MANDELBROT B. B. (1977a)."Limit theorems of the self--normalized range for weakly and strongly dependent processes". En: *Z. Wahr. verw. Geb.*, 31, pp. 271-285.

---------- (1977b). *Fractals: Form, chance and dimension*. San Francisco: W. H. Freeman.

---------- (1983). *The Fractal Geometry of Nature*. San Francisco: W. H. Freeman.

MANDELBROT B. B., VAN NESS J. M. (1968). Fractional Brownian motions, fractional noises and applications. En: *SIAM Rev.*, 10(4), pp. 422-437.

MANDELBROT B. B., WALLIS J. R. (1969 a). Computer experiments with fractional Gaussian noises. En: *Water Resources Res.*, 5(1), pp. 228-267.

---------- (1968). Noah, Joseph and operational hydrology. En: Water Resources Res., 4(5), pp. 909-918.

---------- (1969 b). Robustness of the rescaled range r/s and the measurement of non-cyclic long-run statistical dependence. En: *Water Resources Res.*, 5, pp. 967-988.

---------- (1969 c). Some long-run properties of geophysical records. En: *Water Resources Res.*, 5, pp. 321-340.

MUZY J. F., BACRY E., ARNEODO A. (1993). Multifractal formalism for fractal signals: the structure function approach versus the wavelet transform modulus maxima method. En: *Phys. Rev. E*, 47 (2), pp. 875-884.

MUZY J. F., BACRY E., ARNEODO A. (1991). Wavelets and multifractal formalism for singular signals: application to turbulence data. En: *Phys. Rev. Lett.*, 67(25), pp. 3515-3518.

OSSADNIK S. M., BULDYREV S. V., GOLDBERGER A. L., HAVLIN S., MANTEGNA R. N., PENG C-K, SIMONS M., STANLEY H. E. (1994) Correlation approach to identify coding regions in DNA sequences. En: *Biophys. J.*, 67, pp. 64-70.

PENG C-K., BULDYREV S. V., HAVLIN S., SIMONS M., STANLEY H., GOLDBERGER A. L. (1994). On the mosaic organization of DNA nucleotides. En: *Physical Rev. E*, 49(2), pp. 1685-1689.

PENG C-K., HAUSDORFF J. M., GOLDBERGER A. L. (2000). Fractal mechanisms in neural control: Human heartbeat and gait dynamics in health and disease. En:

WALLECZEK J., Ed., *Self-Organized Biological Dynamics and Nonlinear Control*. Cambridge University Press: Cambridge.

PENG C-K., HAVLIN S., STANLEY H. E., GOLBERGER A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. En: *Chaos*, 1, pp. 82-87.

PRIESTLEY M. B. (1981). *Spectral analysis of time series*. London: Academic Press.

SHLESINGER M. F., WEST B. J., KLAFTER J. (1987). Lévy dynamics of enhanced diffusion: Application to turbulence. En: *Phys. Rev. Lett.*, 58 (11) pp. 1100-1103.

TOUSSON O. (1925). *Mémoires sur l'Historie du Nile*. Mémoires de l'Institut d'Egypte.

TORRES, M.E. (2005). Scaling exponent estimators. Are physiological conclusions always reliable?. (En preparación).

VISWANATHAN G. M., PENG C.-K., STANLEY H. E., GOLDBERGER A. L. (1997). Deviations from uniform power law scaling in nonstationary time series. En: *Phys. Rev. E*, 55, pp. 845-849.

WOOD A. T. A., CHAN G. (1994). Simulation of stationary Gaussian processes in [0;1]d. En: *J. of Computational and Graphical Statistics*, 3, pp. 409-432.