

ANÁLISIS DE SERIES DE TIEMPO SOBRE LAS BASES DE LA TEORÍA FRACTAL

Jorge O. Pierini^{1,2,3} y *Gerardo M. E. Perillo*^{1,4}

¹ Instituto Argentino de Oceanografía. jpierini@criba.edu.ar

² Departamento de Física de la Universidad Nacional del Sur

³ Comisión de Investigaciones de la Provincia de Buenos Aires

⁴ Departamento de Geología de la Universidad Nacional del Sur

RESUMEN

El análisis espectral ha sido convencionalmente utilizado como un método útil y eficiente para evaluar series de tiempo. Especialmente cuando el espectro sigue la ley de la potencia, su exponente es considerado como el índice que representa la irregularidad de la serie de tiempo. En tal sentido se evaluaron datos de corrientes y olas obtenidos en una marisma en Villa del Mar y en el Canal La Lista, ambos sitios en el estuario de Bahía Blanca. En el primer caso, los datos fueron adquiridos con un correntómetro acústico que mide las tres componentes del flujo a 25 Hz y, en el segundo, con otro correntómetro (ADP) que mide perfiles verticales de velocidad con una frecuencia de 9 Hz. El análisis se efectuó sobre extensas series de datos asumiendo que las fluctuaciones son estadísticamente estacionarias; sin embargo, se sabe que en ambientes dominados por marea, estas características varían en un intervalo corto de tiempo. Ello condiciona la longitud de la serie y la interpretación de los resultados. La investigación cuantitativa de series de tiempo no lineales basadas en teorías de caos determinísticas o estocásticas son una herramienta poderosa en el conocimiento de sistemas turbulentos y no periódicos, principalmente, en datos observacionales. En particular pueden proveer descripciones e interpretaciones en series de tiempo irregulares, las cuales no podrían ser gobernadas por un proceso físico estocástico ni tampoco empleando un método lineal. Por lo tanto, el trabajo presenta un análisis de un conjunto de datos oceanográficos, explica el método empleado, discute puntos críticos en la determinación de la dimensión fractal como índice para describir la irregularidad de las series temporales oceanográficas e incorpora una herramienta diferente para el análisis de este tipo de series.

Palabras claves: dimensión fractal-análisis espectral-series de tiempo-espectro de potencia.

ABSTRACT

The spectral analysis has been conventionally used as a useful and efficient method for analyzing a time series. Especially when the spectrum follows the power law, your exponent is considered to be the index for representing the irregularity of the time series. The quantitative investigation of a nonlinear time series based in deterministic or stochastic chaos theories are a powerful tool to understand turbulent and no periodic systems, mainly in observational data. In particular it can provide descriptions and interpretations for irregular time series, which nevertheless might not be governed by a stochastic physical process and which are only poorly understood by linear methods. The present study considers oceanographic data measured in a experimental flume using a ADV Field Sontek and La Lista channel with a Acoustic Doppler Profiler Sontek equipments, which were analyzed using spectral analysis who show noise fluctuations over the power spectrum. The spectrum was calculated over a large series data assuming fluctuations statistically stationarity; however, since the statistical characteristics of fluctuations often vary for a short time interval, which conditions the number of the data used and the interpretation of the results. Therefore, this paper present a oceanographic data analysis, explain the method employed and discuss critical points in the determination of fractal dimension, like the index to describe the oceanographic time series irregularity and also incorporates a different tool to analyze this type of series.

Keywords: fractal dimension-spectral analysis-time series-power spectrum-oceanographic data.

INTRODUCCION

El análisis de series de tiempo es una herramienta muy valiosa en el estudio de datos oceanográficos. Sin embargo, muchos de los métodos de análisis son más precisos cuando la longitud de la serie es larga, ya que pueden conducir a errores significativos si se aplican a datos experimentales de escasa o irregular cantidad. La turbulencia es un fenómeno natural que ocurre en todos los cursos de agua, usualmente se encuentra caracterizada por un movimiento desordenado y caótico sobre un amplio rango de escalas y frecuencias. Un gran número de fenómenos naturales exhiben una representación de la ley de la potencia y se emplea el análisis espectral para evaluar este tipo de series de tiempo. En particular, cuando la densidad espectral $P(f)$ muestra un ruido blanco sin un marcado pico, este se aproxima a un espectro de energía de la forma $P(f) \propto f^{-\alpha}$, donde el proceso autoregresivo de orden uno muestra apreciablemente más energía en bajas frecuencias que en las altas ($\alpha > 0$). Esta aproximación para este tipo de procesos es la descripción más simple, donde la variable α indica la irregularidad de la serie temporal (Yamamoto y Hughson, 2006).

Mandelbrot (1977) incorporó un nuevo concepto para describir las características de las series de tiempo irregulares. Sería deseable conocer si el comportamiento de un sistema durante un cierto número de años puede ser extrapolado a intervalos de tiempo sustancialmente distintos. Si ello fuera cierto, el sistema en estudio podría ser considerado como un fractal en el tiempo y la metodología utilizada para el tratamiento de los objetos fractales sería de aplicación para su estudio.

La idea básica que subyace en las técnicas de análisis fractal es el concepto de auto-similitud y autoafinidad. Dicho concepto supone que un objeto fractal presenta un aspecto semejante independientemente del grado de detalle con que sea examinado. En

el caso de series temporales, el mayor o menor grado de detalle vendrá dado por la menor o mayor longitud de los intervalos temporales sobre los que se promedian las variables.

La persistencia a largo plazo es uno de los factores más importantes para caracterizar el patrón de una serie temporal. Ello se debe a que ésta es una característica con propiedad de larga memoria (explicar), donde la dependencia temporal persiste aún entre observaciones distantes. Estas series pueden ser caracterizadas por patrones cíclicos diferentes pero no periódicos (aclarar). Mandelbrot (1982), Turcote (1992) y Hall y Wolff (1995) señalan que estos procesos de larga memoria tienen dimensión fractal.

Este trabajo presenta la aplicación del análisis fractal para evaluar las propiedades de series de corrientes y olas obtenidas en dos sitios diferentes del estuario de Bahía Blanca, a través de la comprensión de su estructura de correlación. Lógicamente, en un sistema hidrodinámico complejo, como es el que se evalúa, puede ser un parámetro que proporcione una buena información sobre su comportamiento.

DIMENSION FRACTAL: PARAMETRO DE HURST

Aunque el concepto de fractal fue introducido formalmente a mediados de los años 70 (Mandelbrot, 1975, 1977), Hurst (1965) describió un procedimiento de análisis estadístico asimilable a las técnicas fractales, desarrollado y dirigido expresamente al estudio de series temporales discretas sobre las crecidas del río Nilo. Analiza la autosemejanza de una serie temporal consigo misma en el transcurso del tiempo y proporciona una información equivalente a la dimensión de correlación. Este procedimiento, denominado análisis del rango reescalado (rescaled range analysis), calcula el estadístico R/S para diferentes intervalos temporales y permite establecer una relación

entre tres variables: la amplitud del intervalo de desviaciones acumuladas de una magnitud medible (rango estadístico, R), la desviación estándar de la misma (S) y el período de estudio (τ).

Actualmente se aplica en diversos campos, por ejemplo, en el análisis de las series temporales de variación de cotización en las bolsas (Willmann *et al.*, 2002), señales periódicas (Markovic y Koch, 2005), y siempre que interese comprobar la persistencia de una determinada información cuando el sistema al que se refiere evoluciona con el tiempo. Su desarrollo matemático puede resumirse en los siguientes pasos (ver Hurst, 1965 para la descripción completa del método):

1. Descomponer la serie en un conjunto de intervalos de una amplitud temporal determinada, τ , calcular la media para cada intervalo y las desviaciones acumuladas de esta media. Para un vector $x(t)$ que contiene T observaciones espaciadas uniformemente en el tiempo desde $t=1$ hasta $t=T$, se define \hat{u}_0 como

$$\hat{u}_0 = \tau^{-1} \sum_{t=1}^{\tau} x(t)$$

y

$$X(t, \tau) = \sum_{t=1}^T (x_t - \hat{u}_0)$$

con varianza muestral dada por

$$S^2(\tau) = (\tau^{-1} \sum_{t=1}^{\tau} (x_t - \hat{u}_t)^2)^{1/2}$$

Así $X(t, \tau)$ mide la diferencia entre la suma de las observaciones desde 1 a T comparada al promedio de las primeras τ observaciones.

2. Hallar el intervalo de variabilidad de esta desviación como máximo menos mínimo y normalizarlo dividiendo por la desviación estándar. A esta amplitud normalizada se la suele llamar R/S. A cada valor de τ se asigna el valor promedio de R/S sobre todos los

subintervalos que determina. Se elige entonces otro τ y se procede de modo semejante, por lo que se obtiene un conjunto de valores τ y sus R/S asociados. Se define el parámetro H, exponente de Hurst, como $H = \log(R/S) / \log(\tau)$. Su cálculo práctico puede hacerse representando $\log(R/S)$ frente a $\log(\tau)$, de este modo la pendiente de la recta teórica obtenida es H. A esta representación gráfica se la llama “diagrama de Pox”.

$$R(\tau) = \max_{1 < t < \tau} \{X(t, \tau)\} - \min_{1 < t < \tau} \{X(t, \tau)\}$$

y el estadístico R/S se obtiene de $R(\tau) / S(\tau)$

Este análisis es robusto en el sentido de que su comportamiento está influenciado solamente por la persistencia a largo plazo. Detecta ciclos no periódicos aún cuando éstos tengan longitud mayor o igual al período muestral en estudio. Además, es un método adecuado para detectar correlaciones a largo plazo en procesos aleatorios. Diferentes autores han utilizado el análisis R/S para detectar persistencia a largo plazo en variables geofísicas (Mandelbrot y Wallis, 1969; Lovejoy y Schertzer, 1983; Lovejoy y Mandelbrot, 1985; Fluegeman y Snow, 1989; Bodri, 1994; Oñate, 1997). Mandelbrot y Wallis (1969) y Mandelbrot (1975) indican que el análisis R/S también permite deducir que muchos fenómenos naturales no son procesos aleatorios independientes, pero tienen correlación a largo plazo significativa y el $R/S \gg (\tau)^H$.

En general, el exponente de Hurst toma valores en el rango $0 < H < 1$. El valor de $H = 0.5$ indica aleatoriedad de la señal: es el valor típico de un ruido browniano. Mientras que valores de $H > 0.5$ son indicativos de persistencia es decir, la tendencia de la serie temporal es continuar creciendo o decreciendo, y será mayor con H tendiendo a 1. Un exponente menor que 0.5 indica antipersistencia, o sea, la tendencia es, en cierto sentido, contraria a si misma. En este caso las curvas tien-

den a oscilar de modo más errático, se trata de un sistema ergódico. Entonces se puede inferir que el exponente de Hurst es una medida de la información que la señal conserva al cabo de un cierto tiempo.

METODOLOGIA

En este estudio los datos turbulentos se obtuvieron sobre una marisma en la localidad de Villa del Mar. Se empleó un correntómetro ADV3D (Sontek) montado sobre un tetrápodo y ubicado a 5 cm sobre la superficie de la marisma. Este equipo mide las tres componentes del flujo en punto. La marisma está dominada por *Spartina alterniflora* pero con baja densidad y con tallos de hasta 25 cm. Los datos fueron adquiridos a una frecuencia de 25 Hz durante más de 75 min mientras la marea cubría el sensor con una profundidad máxima de 1 m. Durante el experimento se registraron vientos intensos del Oeste que implicaron el desarrollo de olas de 5 a 15 cm de altura y 2 s de período (Perillo, 2006), por lo que la serie de tiempo presenta una combinación de corrientes de marea y olas.

Por otro lado, se obtuvieron perfiles de velocidad y dirección de la corriente con un Acoustic Doppler Profiler (Sontek) en el Canal La Lista, en la parte interna del estuario, durante un ciclo completo de marea. Si bien el equipo mide perfiles verticales, en este caso se empleó la serie de tiempo correspondiente

a la celda ubicada a 4 metros de profundidad donde se obtuvieron datos con un intervalo de muestreo de 60 s.

Para ambas series de tiempo se evaluó la dimensión fractal de las propiedades turbulentas de u' , w' y $u'w'$, se descompuso en intervalos de amplitud temporal determinada τ y se calculó el cociente normalizado, R/S , entre el rango muestral $R(\tau)$ y la desviación estándar muestral $S(\tau)$. Este procedimiento se repitió para valores de $\tau = 2^n$, con n variando desde uno hasta el logaritmo de la longitud máxima de la serie. El exponente de Hurst será la pendiente de la línea recta obtenida al graficar $\log(R/S)$ versus $\log(\tau)$. Mientras que la dimensión fractal se obtiene de $D = 2 - H$.

RESULTADOS Y DISCUSION

La Tabla 1 muestra los estadísticos descriptos por cada serie de datos turbulentos (u' , w' y $u'w'$). De acuerdo con nuestras observaciones, se necesitan al menos 20000 datos para evaluar la dimensión fractal. En las Figuras 1 y 2 se observa la dimensión fractal evaluada por el método R/S en función del número de datos graficados. El parámetro fractal alcanza un valor constante por encima de los 20000 datos. Además se calcularon para cada serie u' , w' y $u'w'$ la dimensión fractal media, parámetro que puede ser empleado para clasificar el comportamiento temporal de la serie (Fig. 1 y 2).

Tabla 1. Estadísticos de las Series

Serie \ Estadístico	ADV3D			ADP		
	u'	w'	$u'w'$	u'	w'	$u'w'$
Cantidad de datos	102975	102975	102975	655	655	655
Int.de muestreo (seg)	0.04	0.04	0.04	60	60	60
Media (cm.s ⁻¹)	-	-	89.35	-	-	20.65
Mediana (cm.s ⁻¹)	-0.90	4.09	-26.24	6.93	-0.21	12.73
Mínimo (cm.s ⁻¹)	-55.38	-293.01	-3452.2	-60.47	-3.81	-148.69
Máximo (cm.s ⁻¹)	56.86	216.75	12888	60.47	4.59	234.37
D. Standard (cm.s ⁻¹)	13.26	34.84	465.19	38.75	1.43	54.24

Mediante la variación de la longitud de la serie de tiempo de las fluctuaciones de velocidad, se evaluaron la dimensión fractal observando la existencia de una cantidad mínima de datos donde la dimensión fractal se vuelve prácticamente constante (Figs. 3, a 5). Valor que establece que las propiedades turbulentas de las series oceanográficas pueden ser aproximadas por la teoría fractal. Por otro lado, se observan en la evolución temporal de la dimensión fractal (Figs. 3 a 5) oscilaciones no periódicas en función de la cantidad de datos de la serie u' , w' y $u'w'$. A los efectos de determinar la componente que más aporta a este comportamiento, se procedió a calcular la correlación cruzada entre las series para ambos conjuntos de datos (Tabla 2).

Los valores de la Tabla 2 indican que las oscilaciones de la dimensión fractal se ven más influenciadas por la componente w' , el parámetro fractal pone de manifiesto las diferentes alteraciones de las fluctuaciones de la componente w' (ADV3D), originadas principalmente por otros procesos, de hecho recordemos que durante la adquisición de los datos había viento y olas. Perillo (2006) estimó

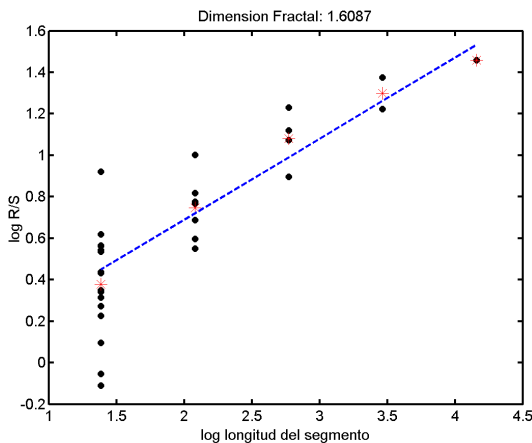


Figura 1. Gráfico de Pox para la serie temporal de $u'w'$ obtenida con ADP en Canal La Lista. Los * establecen el valor medio, la pendiente de la recta es el parámetro de Hurst ($r^2= 0.97$, $F=105.99$, $p=0.002$) y la dimensión fractal evaluada como $D = 2 - H = 1.6087$.

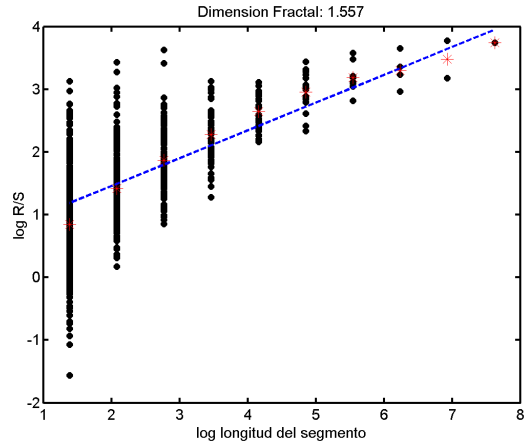


Figura 2. Gráfico de Pox para la serie temporal de $u'w'$ obtenida con ADV3D en Villa del Mar. Los * establecen el valor medio, la pendiente de la recta es el parámetro de Hurst ($r^2= 0.96$, $F=173.63$, $p=0.001$) y la dimensión fractal evaluada como $D = 2 - H = 1.557$.

que el esfuerzo de corte originado por las olas es 100 veces mayor al esfuerzo de las corrientes. Esta evolución de la dimensión fractal se encuentra muy vinculada al comportamiento hidrodinámico, corroborado por los altos valores significativos de correlación. Por otro lado, la menor influencia de correlación se evidencia para los datos de ADP, debido a que estos datos fueron obtenidos a 4 m de profun-

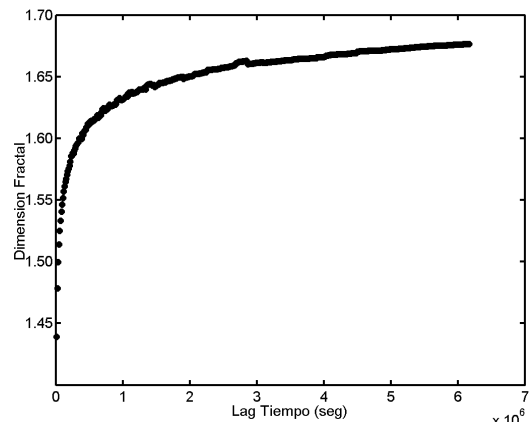


Figura 3. Gráfico de la Dimensión Fractal en función de la cantidad de datos para la serie temporal de u' obtenida con ADV3D en Villa del Mar.

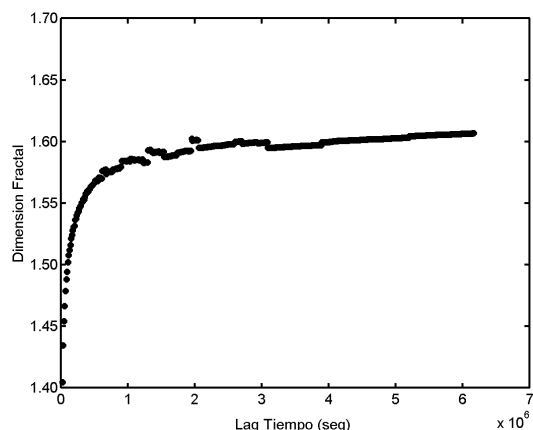


Figura 4. Gráfico de la Dimensión Fractal en función de la cantidad de datos para la serie temporal de w' obtenida con ADV3D en Villa del Mar.

idad donde la influencia de la componente w' se encuentra afectada por procesos de menor escala.

Los resultados de esta investigación muestran que la dimensión fractal es potencialmente una herramienta para describir en forma explícita, objetiva y automática series de datos oceanográficos. Se estudiaron las variaciones de las componentes turbulentas utilizando metodologías de análisis fractal, ya que estas suministran la información necesaria

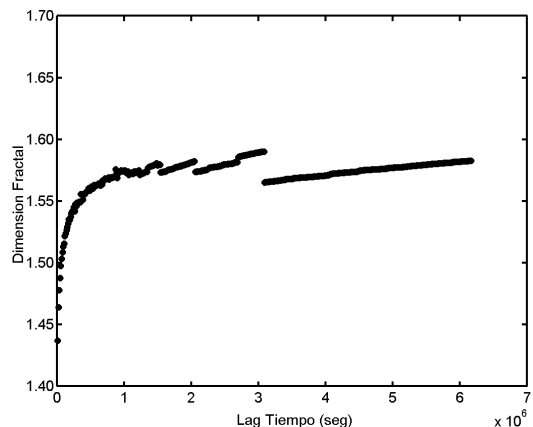


Figura 5. Gráfico de la Dimensión Fractal en función de la cantidad de datos para la serie temporal de $u'w'$ obtenida con ADV3D en Villa del Mar.

Tabla 2. Correlación de $u'w'$ vs u', w' para ambas series de datos, correlaciones significativas con $p < 0.05$

para describir su comportamiento. En las series analizadas de ADV3D se determinó que la dimensión fractal se vuelve constante a partir de una cantidad de 20000 datos, equivalentes a 800 segundos de registro. De igual forma para el conjunto de datos del ADP se vislumbra el mismo comportamiento, pero en este caso a una escala menor. Además se evaluó estadísticamente la importancia de cada una de las componentes turbulentas, observándose que las fluctuaciones verticales son las que generan oscilaciones en la dimensión fractal, manifestando la evolución hidrodinámica de la serie. Al igual que sucede en otras ramas de la ciencia, estos nuevos métodos de estudio no van a sustituir a las tradicionales técnicas de análisis aunque se convertirán probablemente en una herramienta de investigación complementaria.

REFERENCIAS

- Bodri, L. 1994. Fractal analysis of climatic data: Mean annual temperature records in Hungary. *Theoretical and Applied Climatology*, 49, 53–57.
- Fluegeman, R.H. and R.S. Snow. 1989. Fractal analysis of long-range paleoclimatic data: oxygen isotope record of Pacific core V28-239. *PAGEOPH* 131, 307–313.
- Hall, P. and R.C. Wolff. (1995). On the strength of dependence of a time series generated by a chaotic map. *Journal of time series analysis*. 16(6), 571-583.
- Hurst, H.E. 1956. Methods of using long-term storage in reservoirs. *Proc. Inst. Civil Engrs., Part 1*, 519.

Análisis de series de tiempo sobre las bases de la teoría fractal.

- Lovejoy, S. and B.B. Mandelbrot. 1985. Fractal properties of rain, and a fractal model, *Tellus*, 37A, 209-232.
- Mandelbrot, B.B. and J.R. Wallis. 1969. Some Long-Run Properties of Geophysical Records. *Water Resources Research*, 5, 2, 321-340.
- Mandelbrot, B.B. 1975. *Les objets fractals: forme, hasard et dimension*. Paris: Flammarion, pp.187.
- Mandelbrot, B. 1982. *The fractal geometry of nature*. New York, W.H. Freeman y Co, pp.461.
- Markovic, D. and M. Koch. 2005. Sensitivity of Hurst parameter estimation to periodic signals in time series and filtering approaches. *Geophysical Research Letters*, 32(17), L17401.
- Oñate, J.J. 1997. Fractal analysis of climatic data: annual precipitation records in Spain. *Theoretical and Applied Climatology*, 56, 83-87.
- Perillo, M.M. 2006. *Análisis de las corrientes y olas de la marisma de Villa del Mar, Argentina*. Informe Técnico, Departamento de Física, Universidad Nacional del Sur, pp.20.
- Schertzer, D. and S. Lovejoy. 1983. IUTAM symposium on turbulence and chaotic phenomena in fluids, *Turbulence and chaotic phenomena in fluids, proceedings*, 5-10 Sept., Kyoto, Japan, 141-144.
- Turcotte, D.L. 1992. *Fractals and Chaos in Geology and Geophysics*. Cambridge: Cambridge University Press, pp.221.
- Willmann, R.D., G.M, Schutz and D. Challet. 2002. Exact Hurst exponent and crossover behavior in a limit order market model. *Physica A* 316, 430-443.
- Yamamoto Y. and R.L. Hughson. 2006. Extracting fractal components from time series. *Physica* 68D, 250-264.