

Análisis de Fourier y Wavelet de Las Series de Tiempo de la Estación Meteorológica ESPOL-FIMCBOR.

Isaac Mancero Mosquera^a, Xavier Ochoa Chehab^b

^a Facultad de Ciencias Naturales y Matemáticas, Escuela Superior Politécnica del Litoral,
Km. 30.5 Vía Perimetral, Guayaquil, Ecuador
imancero@espol.edu.ec

^b Facultad de Ingeniería en Electricidad y Computación, Escuela Superior Politécnica del
Litoral, Km. 30.5 Vía Perimetral, Guayaquil, Ecuador
xavier@cti.espol.edu.ec

Resumen. Mediciones históricas de temperatura del aire, presión atmosférica y velocidad del viento de la estación meteorológica de la ESPOL correspondiente al año 2000-2001 se han analizado mediante la transformación wavelet, la cual permite establecer si las oscilaciones previamente detectadas por medio de la transformación de Fourier son de tipo estacionario o transitorio, y estudiar con mejor detalle la naturaleza de los fenómenos medidos.

Palabras Clave: Transformación wavelet, Fourier, señal transitoria, estacionaridad.

1 Introducción

Los fenómenos cíclicos son inductivamente conocidos por la humanidad desde tiempos inmemoriales. La denominada variabilidad diurna corresponde a una periodicidad de 24 horas y está relacionada al ciclo astronómico del día y la noche. El ciclo estacional (o anual) es también parte de la experiencia de los seres humanos y evidente a los sentidos. Sin embargo, ciertos fenómenos no son predecibles con exactitud precisamente porque hay variaciones alejadas del comportamiento cíclico o estacionario. El estudio de este tipo de variabilidad revela las limitaciones de la transformación de Fourier, que se basa en funciones sinusoidales de carácter perenne, mientras provee un amplio campo de experimentación con la transformación wavelet [1]. En el presente artículo se documentan los resultados de la aplicación de métodos de análisis basados en ambas transformaciones, a series de tiempo de datos históricos de temperatura, humedad relativa, presión atmosférica, precipitación y velocidad del viento; recolectados en la estación meteorológica de ESPOL-FIMCBOR.

2 Materiales y datos

2.1 Instalación experimental

La estación meteorológica en referencia se localizó en el Campus Gustavo Galindo de

la Escuela Superior Politécnica del Litoral, en las coordenadas 79°57'46''O y 2°08'48''S; la cual recae en la actual Facultad de Ingeniería Marítima, Ciencias Biológicas, Oceánicas y Recursos Naturales; con una elevación 90 metros sobre el nivel del mar. Los instrumentos, de marca Vaisala®, incluían sensores de temperatura, presión atmosférica, humedad relativa, tres sensores de velocidad de viento orientados ortogonalmente para medir las componentes Norte-Sur, Este-Oeste y Vertical; y un sensor de intensidad de lluvia. las mediciones eran tomadas periódicamente por un micro controlador e integradas en una media aritmética cada hora, y almacenadas para su posterior descarga a un computador. De este modo, se generaba un reporte horario de todos los parámetros excepto de la intensidad de lluvia (que mide milímetros de lluvia por unidad de tiempo) cuya sección fue configurada para obtener mediciones cada diez minutos de la precipitación.

Las unidades de medida utilizadas en el presente trabajo son: Presión atmosférica en hectopascales (equivalente a milibares), temperatura del aire en grados Celsius, humedad relativa en valor porcentual, velocidad del viento en metros por segundo y la intensidad de precipitación integrada para producir unidades en milímetros por hora. Las mediciones de viento fueron multiplicadas por un factor de menos uno, pues la convención meteorológica establece que el vector de viento sea medido según su procedencia, mientras la convención matemática establece la orientación según hacia dónde se dirigen (el sentido del vector). Puesto que el presente trabajo se enfoca en la aplicación de una herramienta del procesamiento digital de señales, incidentalmente a series meteorológicas, se ha decidido hacer el cambio del vector de viento a la convención matemática. Luego de un periodo inicial de pruebas, la estación comenzó a funcionar a partir del año 2000, alimentada con baterías y paneles solares. El período bajo estudio en el presente artículo se extiende desde las 09h00 del 30 de Mayo del 2000, hasta el 4 de Junio de 2001 a las 14h00 de la hora local (GMT-5).

2.2 Descripción estadística de los datos

Las series de tiempo, excepto el viento, se presentan en los paneles de la Figura 1. La variación estacional es especialmente evidente en la serie de intensidad de lluvia, que comienza a aumentar en Enero del 2001, alcanzando su máxima intensidad en la segunda mitad de Marzo, y decayendo en el mes de Mayo. La estación lluviosa es localmente conocida como “invierno”. Por otra parte, el periodo entre Junio y Diciembre del 2000 corresponde a la estación seca. La estación lluviosa está caracterizada por un aumento de la humedad, temperaturas más altas, y decrecimiento de la energía (o varianza) del viento. La máxima intensidad de precipitación (I.P.) es 38.1 mm/h durante la estación lluviosa, pero es apenas 1.8 mm/h durante la estación seca. La humedad relativa (H.R.) oscila en torno a 82.5% durante el periodo de lluvias, mientras que decae a 75.9% en la estación seca. La estación lluviosa es más caliente (25.3°C) que la estación seca (23.1°C) mientras que la presión es más alta (1012.3 hPa) en el periodo seco que en el periodo de lluvias (1011.5 hPa). El viento tiene una magnitud promedio de 2.52 m/s en la estación seca, mientras es 1.01 m/s en el periodo de lluvias. Estos resultados se resumen en la Tabla 1.

Los mayores cambios en la variabilidad (desviación estándar) son observados en el viento, que desciende de 5.5 m/s (estación seca) a 0.9 m/s (estación lluviosa); y en la

precipitación que aumenta de 0.04 mm/h a 1.82 mm/h respectivamente. La variabilidad es similar en ambas estaciones en las series de presión, temperatura y humedad relativa, presentando solo diferencias menores.

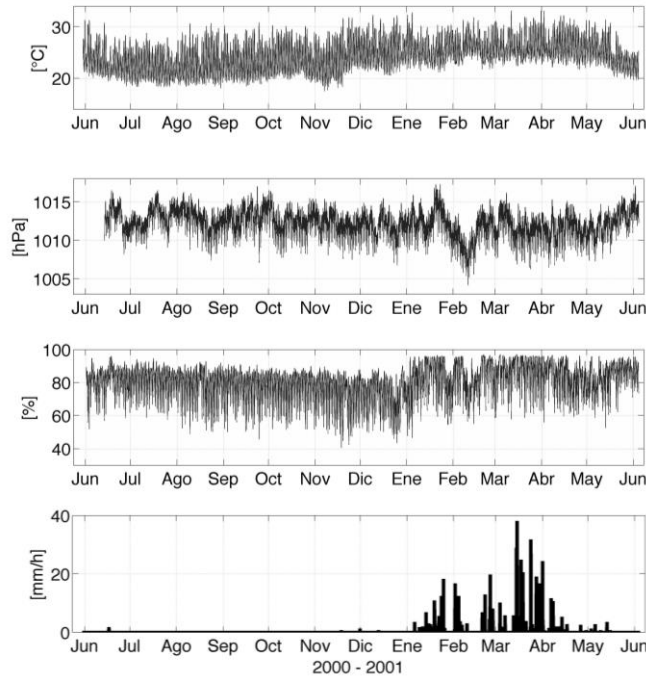


Fig. 1. Del panel superior al inferior, en orden, las series de temperatura del aire, presión atmosférica, humedad relativa e intensidad de precipitación

Tabla 1. Descripción de las series mediante el tamaño de la muestra (N), valor promedio, desviación estándar (S) y valores extremos (mínimo y máximo).

	Estación seca					Estación lluviosa					
	T [°C]	P [hPa]	H.R. [%]	I.P. [mm/h]	Vm [m/s]	T [°C]	P [hPa]	H.R. [%]	I.P. [mm/h]	Vm [m/s]	
N	5175	4832	5121	5175	5170	N	3711	3711	3711	3711	
\bar{x}	23.1	1012.3	75.89	0.002	2.52	\bar{x}	25.3	1011.5	82.5	0.3	1.01
S	3.1	1.68	10.8	0.04	5.5	S	2.5	2.0	10.2	1.82	0.9
Mín.	17.54	1006.8	40.5	0	0	Mín.	19.6	1004.2	46.6	0	0
Máx.	33.22	1017	96.1	1.8	7.9	Máx.	33.2	1017.3	97.1	38.1	6.5

El viento se representa mediante sus componentes en las direcciones geográficas Este-Oeste y Norte-Sur; y la componente vertical (Figura 2). Por medio del Análisis de Componentes Principales [2] se ha determinado que la varianza de la serie está mayormente distribuida a lo largo de un eje en la dirección 148.74°, que se denomina Componente Principal (CP1); la segunda componente (CP2) es perpendicular (238.74°) y no tiene correlación estadística con CP1; el método encuentra una tercera

dirección CP3 ortogonal a las dos primeras. Los vectores de viento pueden representarse en el nuevo sistema de componentes principales, con coordenadas CP1, CP2, CP3. Las series correspondientes se muestran en el panel inferior de la Figura 2.

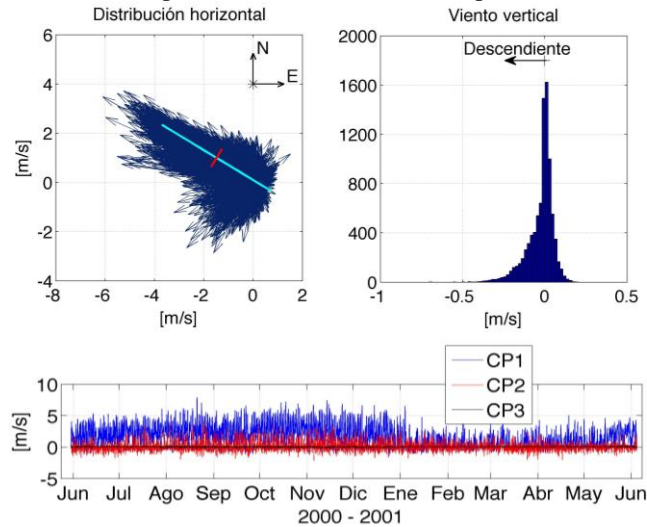


Fig. 2. Los ejes de CP1 y CP2 se muestran en celeste y rojo (*panel superior izquierdo*). Las series CP1 a CP3 (*panel inferior*) son las coordenadas de la velocidad del viento en un sistema ortogonal orientado según los ejes principales.

Tabla 2. Descripción del viento en sus componentes geográficas y principales

Componentes	Geográficas			Principales		
	E-O	N-S	V	CP1	CP2	CP3
\bar{x} [m/s]	1.2829	0.7491	-0.0241	1.4852	-0.0245	0.0293
s [m/s]	1.4158	1.0028	0.0790	1.6098	0.6498	0.0509
s^2 [%]	-	-	-	85.92	14.00	0.08
Orientación	0°	90°	-	31.26°	121.26°	-10.13°
Elevación	0°	0°	90°	-2.03°	1.78°	87.29°
Componentes	Estación seca			Estación lluviosa		
	CP1	CP2	CP3	CP1	CP2	CP3
\bar{x} [m/s]	2.0805	-0.0707	0.0261	0.6559	0.0399	0.0339
s [m/s]	1.6577	0.7334	0.0572	1.0920	0.5039	0.04

Las estadísticas descriptivas se presentan en la Tabla 2. Cada descripción hecha sobre una componente en la nueva representación es independiente de las demás, lo que facilita su estudio. CP1 se orienta en la dirección 148.74° pero con una elevación de -2.03° , lo que indica que su valor promedio de 1.48 m/s es un vector descendente. Esto se explica porque los instrumentos se localizaron en una zona irregular, no una planicie. Por el contrario, valores negativos de CP1 son ascendentes, pero estos son menos frecuentes. CP1 contribuye con más del 85.92% de la varianza total de los

datos; CP2 contribuye con el 14% mientras CP3 con apenas el 0.08%. Tanto CP1 como CP2 muestran mayor magnitud durante la estación seca. El valor medio de CP2 cambia de positivo a negativo, mientras el cambio en CP3 es pequeño. La desviación estándar es mayor durante la estación seca, principalmente en CP1. En total, el 99.92% de la variabilidad total se explica solo con las contribuciones de CP1 y CP2.

3 Métodos

3.1 La transformación de Fourier

Aunque la transformación de Fourier está bien documentada en la literatura, una breve descripción es hecha para tener un tratamiento cabal. Esta transformación es un operador integral de la forma:

$$F_f = \left\langle f \mid e^{i\omega t} \right\rangle = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt, \quad (1)$$

donde t representa el tiempo, ω la frecuencia, $\langle \mid \rangle$ indica el producto interno en el espacio de funciones con energía finita $L^2(\mathbb{R})$. Este último implica, en el caso de datos discretos, la proyección escalar de f sobre un conjunto ortogonal de funciones $\exp(j\omega t)$ tal que los módulos al cuadrado de los coeficientes resultantes indican cuán fuerte es la energía asociada a oscilaciones de frecuencia ω [3]. Los algoritmos prácticos requieren que no haya datos faltantes, pero en el presente estudio los datos faltantes se presentan solo al inicio de las series de presión (343 horas sin registro), humedad (24) y velocidad de viento (5 datos faltantes); por lo cual la mejor solución es el truncamiento. Otros cursos de acción contemplan la interpolación o el relleno con ceros. Las series fueron estandarizadas previo a la transformación, para facilitar la comparación de sus espectros, según se recomienda cuando se procesan parámetros con unidades de medida diferentes (como grados Celsius y m/s) que no admiten una comparación directa.

Para disminuir el error de estimación, se realiza la partición en *ventanas* de 512 datos de longitud, ponderando cada segmento con una función de Hamming, o sea 16 subintervalos sin datos faltantes. Este método prevé el cálculo del espectro para cada “ventana” y su promediado posterior mediante la expresión:

$$Psd_T = \frac{1}{k} \overset{\circ}{\underset{\text{over } i}{\hat{a}}} \frac{Psd_i}{U_w W_L}, \quad (2)$$

donde Psd_T es el espectro de energía total estimado de la suma de los espectros parciales de cada ventana de datos Psd_i . Luego, Psd_i es el módulo al cuadrado de la transformada de Fourier del i -ésimo segmento de datos, U_w es la energía de la

ventana de Hamming, mientras k , el número de ventanas utilizadas, depende de la longitud de la misma, la longitud del record siendo particionado y el porcentaje de traslapamiento entre segmentos consecutivos, si así se decide hacerlo [3]. El espectro de Fourier existe entre la frecuencia cero y el límite superior de Nyquist con una resolución de frecuencia de $1/W_L$.

3.2 La transformación wavelet

El operador wavelet tiene también una forma integral:

$$\mathcal{W}f = \left\langle f \mid \psi \left(\frac{t-t_0}{s} \right) \right\rangle = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} f(t) \psi^* \left(\frac{t-t_0}{s} \right) dt, \quad (3)$$

donde $*$ indica conjugación compleja. A diferencia de la transformación de Fourier, ψ es una función oscilatoria de duración finita, con promedio cero, denominada *wavelet madre*. Dado su carácter transitorio, la transformación de f a lo largo de todo su dominio requiere el auxilio del parámetro de traslación t_0 . Esto convierte la expresión (3) en una convolución o filtrado lineal de f con un filtro equivalente de tipo pasa-banda, en virtud de la media cero de la función wavelet. El parámetro “ s ” expande o contrae la función cambiando su escala y por consiguiente las frecuencias de oscilación y la respuesta de frecuencia del filtro asociado. Sin embargo, “ s ” mantiene unidades de tiempo correspondiente a periodicidad.

Los coeficientes de transformación se distribuyen en el plano t - s permitiendo analizar su evolución en el dominio tiempo. Un catálogo de funciones wavelet puede ser encontrado en [1], con información sobre sus características y utilización. En la forma discreta, t_0 varía según el intervalo de muestreo, $t_0 = nDt$, $n \in \mathbb{Z}$; mientras s varía en modo *diádico* (según potencias de 2), $s = 2^j$, $j \in \mathbb{Z}$ donde j se denomina *nivel de descomposición*. En consecuencia, cada cambio de escala permite capturar porciones del espectro en diferentes intervalos de frecuencia. Estos intervalos se indican como periodos, se denominan *escalas de tiempo* y sus límites se expresan como $[2^j \Delta t, 2^{j+1} \Delta t]$. Aunque en teoría es posible que $j \rightarrow \infty$, se elige un máximo nivel de descomposición finito J .

En el presente trabajo se calculan dos tipos de transformación wavelet, la *transformada continua* y la *transformación ortogonal*. El primer tipo se ajusta a la definición (3), pues se elige una función wavelet, se muestrea hasta obtener la misma cantidad de datos que la serie a transformar y se efectúa la el filtrado, generalmente en el dominio de la frecuencia, haciendo uso del teorema de convolución. Esto es, se calculan las transformadas de Fourier de los datos y de la función wavelet, se multiplican en el dominio de la frecuencia, y se calcula la transformación de Fourier inversa del producto. Este proceso se repite para cada nivel $j \leq J$. La función elegida para este caso es la denominada wavelet de Morlet, pues se trata de una función analítica y tiene la propiedad de menor dispersión de energía en el plano tiempo frecuencia, lo que la hace adecuada para los análisis exploratorios de la energía de una serie de datos.

La transformación ortogonal, por otra parte, hace uso del hecho que algunas funciones wavelet producen una base ortogonal de $L^2(\mathbb{R})$ y los coeficientes de transformación son por lo tanto estadísticamente independiente entre ellos. No siempre se conoce las expresiones matemáticas de este tipo de wavelets, pero sí los coeficientes del filtro asociado, por lo cual el cálculo de la transformación se realiza mediante esquemas de filtrado iterativo descritos con detalle en el trabajo de Mallat [4]. El esquema se basa en un *filtro wavelet* discreto $g[n]$, asociado a la función ψ , y en un *filtro de escalado* $h[n]$, que forma un par conjugado o en cuadratura con g :

$$g[n] = (-1)^{1-n}h[1-n]. \quad (4)$$

El filtro h corresponde a una función de escalado ϕ , que genera una familia de funciones $\phi_j((t-t_0)/s)$ que son una base ortogonal del denominado subespacio de *aproximación* H , que es el complemento ortogonal del subespacio *wavelet* W , cuya base ortogonal la conforma la familia de funciones $\psi_j((t-t_0)/s)$.

El esquema iterativo comienza con el filtrado de la serie de datos $x[n]$ con los filtros h y g de la función wavelet elegida. Las series resultantes se denominan *aproximaciones* y *detalles* de la serie original, o $a_1[n]$ y $d_1[n]$. La serie $d_1[n]$ son los coeficientes wavelet del primer nivel de descomposición ($j=1$) y contiene información entre la frecuencia de Nyquist f_N y $f_N/2$. La serie $a_1[n]$ contiene información complementaria, entre la frecuencia cero y $f_N/2$.

Para obtener los coeficientes del segundo nivel, la serie $a_1[n]$ es muestreada, tomando un dato sí, un dato no, lo cual duplica el intervalo de tiempo entre ellos a $2\Delta t$ y reduce su cantidad total a la mitad. Esta serie muestreada se filtra nuevamente con h y g . Las secuencias resultantes son $a_2[n]$ y $d_2[n]$, donde $d_2[n]$ son los coeficientes wavelet de segundo nivel ($j=2$). Los filtros no han cambiado, pero con el intervalo entre muestras duplicado, la serie $d_2[n]$ contiene información entre $f_N/4$ y $f_N/2$ mientras $a_2[n]$ contiene información entre la frecuencia cero y $f_N/4$.

Nuevamente se muestrea $a_2[n]$ y se efectúa el filtrado sucesivo; ahora el intervalo entre muestras es $4\Delta t$. El proceso se detiene en la práctica cuando el muestreo de $a_j[n]$ produce una cantidad de datos similar al número de coeficientes del filtro que se está utilizando. En este último nivel J , tanto $a_J[n]$ y $d_J[n]$ tienen la misma cantidad de coeficientes, y la suma total de los coeficientes de transformación es igual a la cantidad de datos de la serie original. La secuencia remanente $a_J[n]$ contiene la información correspondiente a la frecuencia cero y por lo tanto sobre la media del proceso.

La transformación ortogonal permite una reconstrucción exacta de la serie original pues contiene toda la información necesaria para ello con una cantidad mínima de coeficientes, por ello es muy utilizada en tareas de compresión de datos. Además, dado que los coeficientes provienen de vectores linealmente independientes, una amplia variedad de aplicaciones y pruebas estadísticas es permitida. Los cálculos de la transformación ortogonal en el presente trabajo se realizaron utilizando el filtro Symmlet, asociado a la homónima función wavelet. Este filtro, diseñado por Daubechies [5], tiene soporte compacto y una respuesta de fase casi lineal, lo que permite relacionar un coeficiente específico a un tiempo t en la serie original. Por

ejemplo, si un test estadístico establece la validez de alguna información con respecto a un coeficiente, es posible saber también el tiempo asociado a este coeficiente. En la Tabla 3 se presentan los intervalos de periodicidad para 10 niveles de descomposición realizados con las series de presión, temperatura, y componentes de viento CP1 y CP2.

4 Resultados y discusión

4.1 Espectros de energía de Fourier

Las diferencias estacionales en los datos pueden explicarse por los ciclos naturales, en particular, los denominados forzantes astronómicos: la rotación, traslación terrestre, los ciclos lunar y solar con sus respectivos periodos, etc. La aplicación de la transformación de Fourier permite determinar si tales forzantes son importantes mediante el espectro de energía. El método descrito en la sección 3.1 utiliza ventanas de datos lo que permite establecer límites de error mediante la expresión general:

$$\left[\frac{\chi^2_{n-1}}{c_n^2(1-a/2)}, \frac{\chi^2_{n+1}}{c_n^2(a/2)} \right] P_{sd_T}, \quad (5)$$

donde χ^2 son los grados de libertad de la variable Chi cuadrado, toma el valor de 2 para todas las frecuencias diferentes de cero, o $2M$ si se usan M ventanas de datos. En el presente cálculo, los valores son los factores 1.7495 y 0.64672, que multiplicados por el espectro P_{sd_T} produce un intervalo de confianza de 95%. Los resultados se muestran en la Figura 3. La precipitación fue omitida en este análisis.

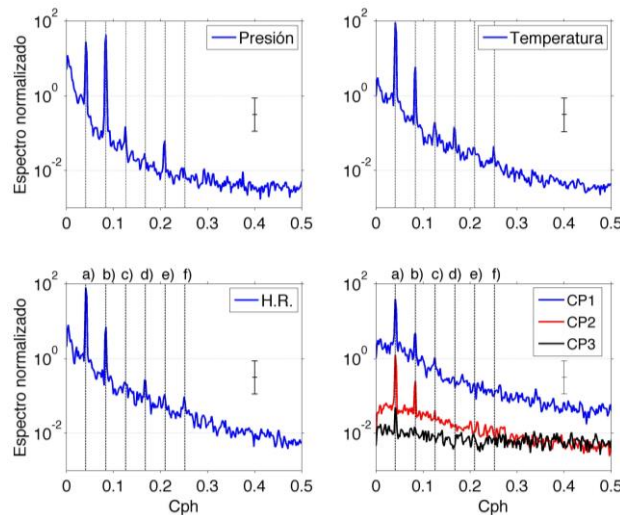


Fig. 3. Estimación de los espectros de energía y sus intervalos de significatividad al 95%.

Se pueden observar dos máximos locales, estadísticamente significantes y bien diferenciados. El primero, indicado con “a” en la Figura 3, de 0.041 ciclos por hora (Cph), o un periodo de 24.38 horas. El segundo, “b”, de 0.084 cph (11.90 horas). Éstos se denominan “grupo diurno” y “grupo semidiurno” porque están relacionados a la rotación terrestre [6]. En la serie de tiempo de presión atmosférica, ambos máximos tienen similar orden de magnitud, mientras en las series restantes el grupo diurno es dominante. Otros máximos locales de menor magnitud aparecen en frecuencias armónicas superiores, indicadas como “c”, “d”, “e” y “f” en la Figura 3.

La influencia astronómica en fenómenos climáticos no siempre es igual. Se espera que la temperatura del aire esté correlacionada con la rotación y traslación terrestre (periodos de 24 horas y un año); mientras la presión atmosférica, al ser un fenómeno gravitacional, recibe influencia lunar. En ambos casos tal periodicidad es permanente, pero el fenómeno de brisa que se manifiesta en el viento con un periodo de 24 horas, suele ser estacional. La transformación de Fourier nos presenta información sobre las frecuencias presentes en los registros, pero no indica si tales oscilaciones son perennes o no.

4.2 Análisis wavelet

Con la transformación wavelet ortogonal es posible estimar la varianza de los coeficientes wavelet, escala por escala. Se utiliza el estimador propuesto por Percival [7], modificado para incluir una corrección de sesgo, no eliminar coeficientes al inicio y fin de las series y, adicionalmente, aplicar el cálculo a las series de aproximación remanente [1], lo que no era originalmente previsto en el trabajo de Percival. Los resultados se muestran en la Tabla 3.

Tabla 3. Descomposición de la varianza por escalas, de las series de temperatura (T), presión (P), CP1 y CP2. El último nivel ($j=10$) incluye también las aproximaciones remanentes. La suma total estima la varianza de las series originales, difiriendo en un pequeño porcentaje.

nivel j	Escala:	T [°C ²]	P [hPa ²]	CP1 [(m/s) ²]	CP2 [(m/s) ²]
1	2 - 4h	0.035711	0.009819	0.097363	0.052327
2	4 - 8h	0.117569	0.084128	0.176930	0.062322
3	8 - 16h	0.884068	0.905415	0.387911	0.091361
4	16 - 32h	5.035192	0.672871	0.850444	0.147757
5	32 - 64h	0.615497	0.144306	0.216348	0.032939
6	64 - 128h	0.196893	0.188070	0.102038	0.013179
7	128 - 256h	0.126544	0.191242	0.062648	0.006609
8	256 - 512h	0.093387	0.271176	0.022544	0.003251
9	512 - 1024h	0.117066	0.307483	0.018272	0.002187
10	1024 - 2048h	0.170825	0.305876	0.025482	0.001050
10	Aprox. [$\geq 2048h$]	1.953028	0.372759	0.625702	0.008477
	Total	9.345783	3.453146	2.585682	0.421459
	Error (%)	-0.57	-1.3	-0.22	-0.17

Se espera que la suma de las varianzas por escalas sea igual a la varianza de la serie analizada. El error de estimación se mantiene por debajo del 0.6% en magnitud en todos los casos excepto la presión atmosférica, cuyo error es de magnitud 1.3%. Los niveles de detalles 3 y 4, que contienen la señal semidiurna (12h) y diurna (24h) respectivamente, son los de mayor varianza en concordancia con los resultados del análisis de Fourier. Los siguientes en importancia son los niveles 5 y 2, mientras que a partir de la serie de detalles de $j=6$ la energía decae abruptamente. Las aproximaciones remanentes del último nivel ($j=10$) tiene mayor varianza que su correspondiente serie wavelet. Los valores detallados en la Tabla 3 conforman el denominado espectro wavelet global de energía, donde las escalas corresponden al inverso de la frecuencia de Fourier, extendiéndose desde el límite de Nyquist (2 horas de periodo) hasta la frecuencia cero (periodo infinito, incluido en la aproximación).

Las variaciones de la energía con el tiempo se analizan con la transformación continua utilizando el método propuesto por Torrence y Compo [8]. Este consiste en no promediar los coeficientes de transformación, sino en mapear sus módulos al cuadrado con referencia a sus respectivas escalas y tiempos. Según la praxis, las series fueron normalizadas para facilitar la comparación entre ellas, pues tienen diferentes unidades de medida. Una codificación por colores permite observar las regiones del plano t-s con relativamente alta energía. Los resultados se presentan en la Figura 4.

Para determinar la significatividad, los cuadrados de los coeficientes son comparados con un modelo de ruido que puede ser blanco o rojo. En el presente caso, el ruido rojo es el adecuado a datos meteorológicos por ser decreciente con la frecuencia, conforme al comportamiento del espectro de fondo observado en la Figura 3. Tal como por encima de este espectro sobresalen los picos espectrales de Fourier, se busca determinar los coeficientes wavelet que producen energía que sobrepasa un cierto umbral. El modelo de ruido tiene la forma:

$$P(w) = (1 - a^2) \left(1 - a^2 - 2a \cos(2\pi k / N) \right)^{-1}, \quad (6)$$

donde “a” es la autocorrelación de lag 1 obtenida de cada serie de tiempo. El límite del 95% es obtenido mediante una distribución Chi cuadrado similar a la utilizada en los espectros de Fourier [8]. Los coeficientes cuyos módulos al cuadrado superan este umbral son señalados como significativos y delimitados como se muestra en la Figura 4, mediante líneas de contorno blancas.

Se puede observar que la energía de la temperatura es predominante en la escala que contiene el grupo diurno, y persistente durante todo el periodo de mediciones. Por otra parte, los máximos locales de energía que aparecen entre 8 y 16 horas (conteniendo el grupo semidiurno) son recurrentes pero relativamente más débiles. En el caso de la presión atmosférica, la energía predominante está en la banda semidiurna y persistente a lo largo del periodo de estudio, mientras la escala del grupo diurno es relativamente débil. Los grupos diurno y semidiurno pueden considerarse persistentes en ambas series. Los datos de presión presentan un importante máximo local alrededor de la escala de 1024 horas, que abarca los meses Enero a Marzo del 2001, es decir, una clara señal transitoria durante la estación lluviosa.

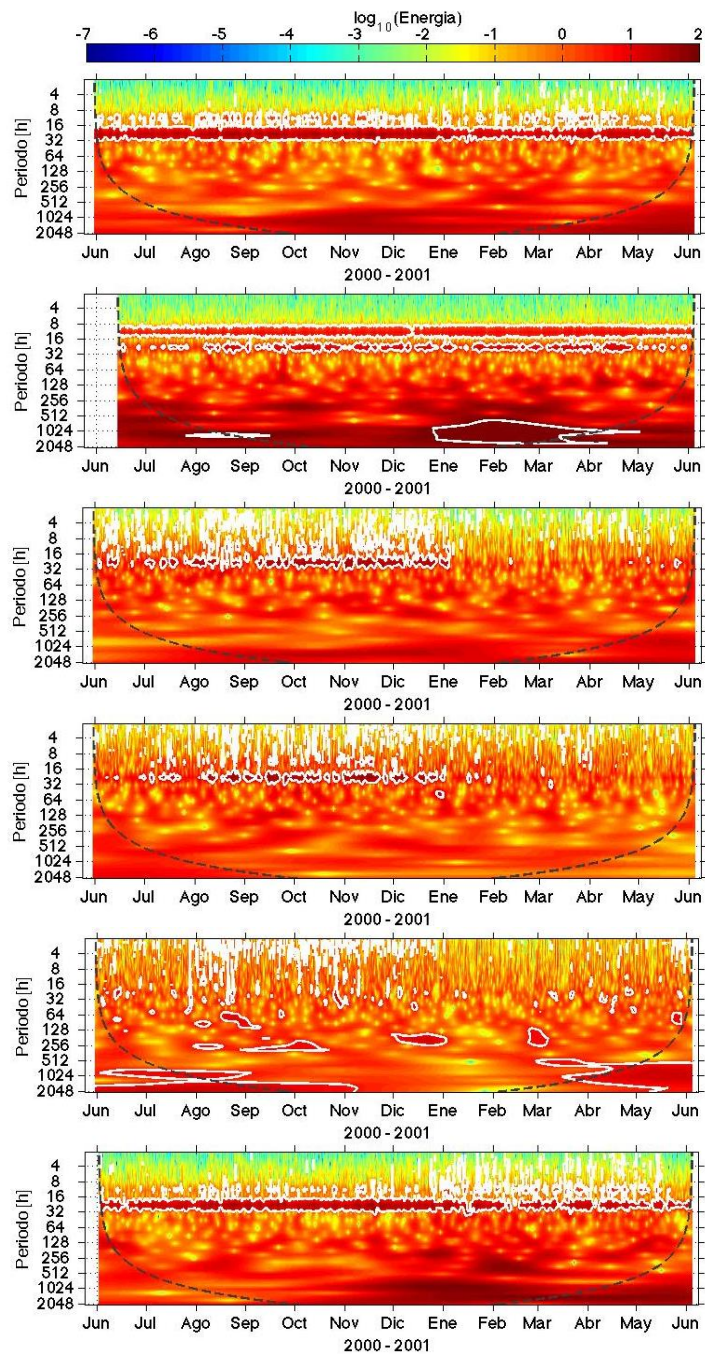


Fig. 4. Espectros normalizados wavelet de (*en orden descendiente*) la temperatura, presión, componentes del viento CPI, CP2 y CP3; y humedad relativa, respectivamente. Las *líneas de contorno blancas* indican regiones de energía significativa.

Considerando las componentes del viento, la energía predominante de CP1 y CP2 aparece en el grupo diurno, y está confinada a la primera mitad del registro, entre Junio 2000 e inicios de Enero del 2001, es decir, la estación seca. Esporádicas manifestaciones de energía aparecen a escalas de tiempo menores (altas frecuencias), siempre más frecuentes durante la estación seca. CP3 presenta un espectro de Fourier casi plano, con excepción de la región del grupo diurno, indicando que la mayor parte de esta serie es ruido blanco. El espectro wavelet es consistente con esta observación pues contiene numerosos máximos locales dispersos, que aparecen principalmente en la banda diurna y en los periodos cortos de la estación seca. El espectro wavelet de la serie de humedad se presenta también, mostrando similares características que la temperatura. Esto se debe a que ambos parámetros están físicamente relacionados.

4.3 Separación de componentes perennes y transitorias

La energía contenida en las bandas diurna y semidiurna presenta las características de una señal estacionaria en el caso de la temperatura y la presión, pero es definitivamente transitoria en las componentes del viento, con respecto al periodo de estudio. Esto es consistente con la idea de una influencia astronómica en la presión y temperatura, mientras que la variabilidad del viento puede ser explicada por la brisa de montaña. La brisa se debe al gradiente térmico que aparece en laderas montañosas, el mismo que es destruido durante el periodo de lluvias, de aquí su carácter estacional. La contribución de los forzantes astronómicos puede ser estimada con el Análisis Armónico. Este método, estrechamente relacionado al análisis de Fourier, consiste en ajustar por mínimos cuadrados a los datos una serie finita de sinusoides de la forma:

$$z = \sum_i A_i \cos(x_i t - j_i), \quad (7)$$

donde las frecuencias x corresponden a ciclos astronómicos conocidos [9]. Las amplitudes (A) se estiman para cada término o *constituyente*, y las fases j se obtienen con respecto a la señal astronómica en Greenwich. Una lista completa de tales constituyentes puede encontrarse en [6] y detalles sobre el método en el trabajo de Pawlowicz [10]. El Análisis Armónico no es aplicable a las series de viento, pues la premisa que las oscilaciones son perennes en el periodo de estudio no se cumple.

Como resultado de (7), se obtiene una serie sintética z cuya varianza permite cuantificar la influencia astronómica en los datos con relación a la varianza de la serie original. Se encontró que el 82.2% de la varianza de la temperatura es atribuible al forzante astronómico; así también el 31.2% de la varianza de la presión, y el 67.4% en el caso de la humedad. En total, se hallaron 67 constituyentes astronómicas de las cuales las 6 más importantes se presentan en la Tabla 4, mostrando sus amplitudes (A_m), fase (Ph) y el porcentaje de contribución a la varianza astronómica total (C.V.)

Las constituyentes más fuertes de la temperatura del aire son iguales a aquellas de la humedad relativa por la relación física que existe entre ambos parámetros. En ellos predominan las frecuencias relacionadas al sol, con más del 96% de contribución de solo 4 constituyentes solares. Por otra parte, la presión contiene importante influencia

lunar, incluyendo el ciclo del mes lunar de casi 28 días, y dos oscilaciones producto de interacciones entre sol y luna, con respecto a la tierra. La remoción de estas constituyentes de los datos $x(t)$ produce series residuales $r(t) = x(t) - z(t)$, cuya variabilidad ya no es atribuible a oscilaciones de origen astronómico.

Tabla 4. Las 6 constituyentes astronómicas más importantes encontradas en las series de temperatura del aire, humedad relativa (H.R.) y presión atmosférica.

Temperatura		Frecuencia	Periodo	Am.	Ph.	C.V.
<i>Constituyente:</i>		[cph]	[h, d]	[°C]	[°]	[%]
S1	Radiacional	0.0416667	24 h	4.53	244.1	80.05
S _A	Solar anual	0.0001141	365.18 d	1.921	47.25	14.39
S2	Solar semidiurna	0.0833333	12 h	0.894	47.7	2.65
S _{SA}	Solar semianual	0.0002282	182.58 d	0.442	30.38	0.76
K1	Lunar diurna	0.0417807	23.9345 h	0.391	123.87	0.60
P1	Solar diurna	0.0415526	24.0659 h	0.359	299.14	0.50
total						98.96%

H.R.		Frecuencia	Periodo	Am.	Ph.	C.V.
<i>Constituyente:</i>		[cph]	[h, d]	[pc.]	[°]	[%]
S1	Radiacional	0.0416667	24 h	14.96	156.24	81.38
S _A	Solar anual	0.0001141	365.18 d	5.175	113.72	9.73
S2	Solar semidiurna	0.0833333	12 h	3.184	40.2	3.69
S _{SA}	Solar semianual	0.0002282	182.58 d	2.195	326.42	1.75
K1	Lunar diurna	0.0417807	23.9345 h	1.347	50.83	0.66
P1	Solar diurna	0.0415526	24.0659 h	1.516	187.46	0.84
total:						98.04%

Presión		Frecuencia	Periodo	Am.	Ph.	C.V.
<i>Constituyente:</i>		[cph]	[h, d]	[hPa]	[°]	[%]
S2	Solar semidiurna	0.0833333	12 h	1.367	306.48	85.55
S _{SA}	Solar semianual	0.0002282	182.59 d	0.267	193.53	3.28
M _M	Lunar mensual	0.0015122	27.55 d	0.263	180.77	3.16
MS _M	Sinódica Lunisolar	0.0013098	31.81d	0.259	184.74	3.07
MS _F	Sinódica Lunisolar	0.0028219	14.77 d	0.162	17.74	1.21
P1	Solar diurna	0.0415526	24.0659 h	0.143	93.47	0.93
total						97.19%

Espectros wavelet fueron calculados para tales series residuales. La presión atmosférica residual (Figuras 5) presenta una importante variabilidad en las escalas de tiempo en torno a 1024 horas, presente en los meses de invierno de 2001(estación lluviosa). Máximos locales aislados pueden observarse entre 64 y 128 horas de periodicidad, mientras energía remanente en la banda diurna se presenta también aislada. La serie residual de temperatura (Figura 6) presenta un máximo significativo en las escalas de tiempo mayores de 512 horas entre septiembre de 2000 y febrero de 2001; que no era relativamente importante cuando la señal astronómica estaba

presente. También se muestra energía remanente en la banda diurna, y un máximo local aislado entre 256 y 512 horas, durante la primera quincena de febrero de 2001.

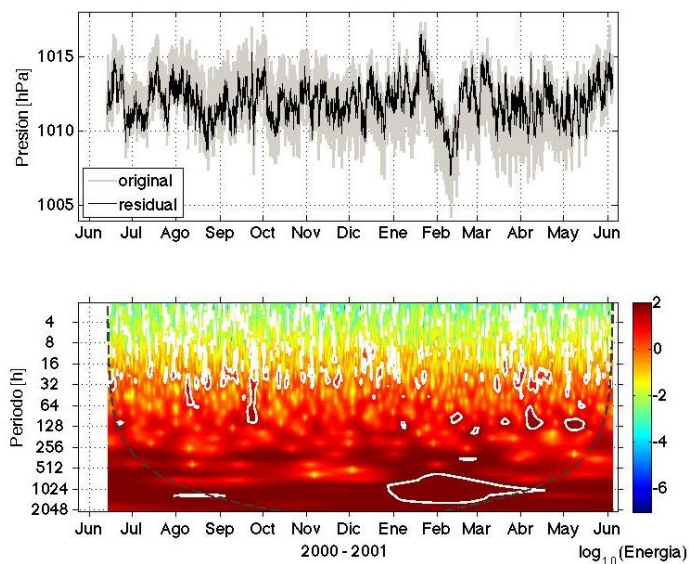


Fig. 5 La presión residual sobrepuesta a la serie original (panel superior), junto con su correspondiente espectro wavelet (panel inferior).

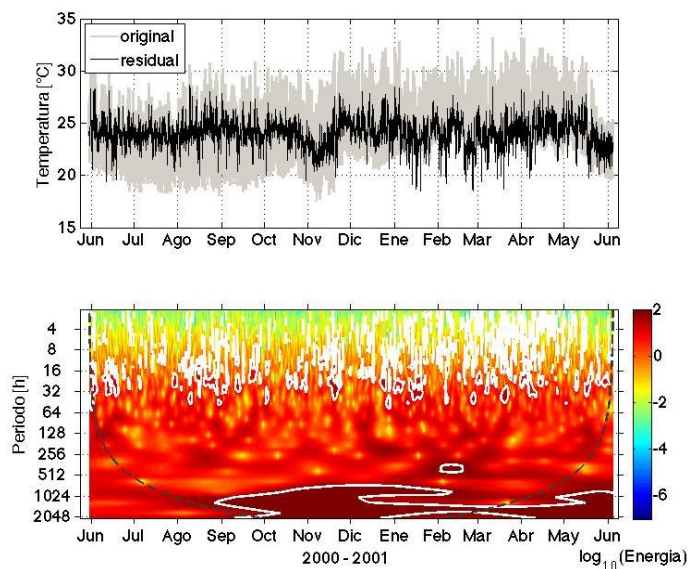


Fig. 6 La temperatura residual sobrepuesta a la serie original (panel superior), junto con su correspondiente espectro wavelet (panel inferior).

Las causas de la variabilidad infra-anual aun tienen que ser exploradas. El régimen biestacional es acentuado en la región ecuatorial, en contraposición a las cuatro estaciones presentes a mayores latitudes. Ambas estaciones están presentes en los registros que abarcan poco más de un año. Un test de punto de cambio de la media es efectuado utilizando el estimador basado en sumas acumulativas normalizadas P_k , de Tanizaki [11], modificado para evitar tratar con media y varianza simultáneamente [1]:

$$P_k = \frac{\hat{a}_{i=1}^k X_i}{s}, \quad (8)$$

Tales sumas tienen un valor máximo y mínimo E^+ y E^- , respectivamente; pero su extremo absoluto E , estima el momento de cambio en la media cuando supera los niveles críticos mostrados en Tabla 5.

Tabla 5. Niveles críticos para el estadístico E halladas por Monte Carlo [1], donde α es el nivel de significatividad y N el tamaño de la muestra. La columna “As.” contiene los niveles críticos para muestras grandes (asintóticos).

$\alpha = .05$	N= 8	16	32	64	128	256	512	1024	As.
L_{sup}	3.086	4.949	7.517	11.051	15.971	22.864	32.847	46.509	268.99
L_{inf}	1.061	1.553	2.287	3.391	4.994	7.179	10.288	14.830	86.701

Dado que la energía de las escalas empieza a decaer a partir del nivel $j=5$ (Tabla 3), la prueba se aplica a las aproximaciones de una descomposición ortogonal hasta $j=5$; en este nivel, el número de coeficientes es 256. La Tabla 6 resume los resultados de la prueba.

Tabla 6. Resultados del test de cambio en la media de las series. E es el estadístico de prueba, que cae fuera del intervalo crítico, luego la hipótesis nula de ausencia de cambio es rechazada (R). El momento estimado del cambio es presentado también.

	E	Hipótesis nula	Fecha	hora GMT+5	valor (antes)	promedio (después)
Temperatura	109.71	R	21 Noviembre 2000	01h00	22.65 °C	25.2 °C
Presión	57.09	R	29 Octubre 2000	09h00	1012.55 hPa	1011.58 hPa
Humedad	64.81	R	1 Enero 2001	09h00	75.88%	82.49%
Precipitación	47.46	R	14 Enero 2001	17h00	0.007 mm/h	0.316 mm/h
CP1	104.67	R	21 Diciembre 2000	17h00	2.12 m/s	0.7 m/s
CP2	37.99	R	20 Enero 2001	01h00	0.065 m/s	-0.047 m/s
CP3	53.8	R	6 Agosto 2000	09h00	0.044 m/s	0.026 m/s

Los valores E superan ampliamente los niveles críticos, indicando que efectivamente un cambio en la media tuvo lugar. El índice donde E fue encontrado es

indicativo del momento en que el cambio tuvo lugar, expresado como una fecha en Tabla 6.

En el caso de la temperatura, hay un cambio en el 21 de Noviembre de 2000 a las 01h00. La temperatura promedio desde el inicio del periodo de registro hasta la fecha donde E alcanza su máximo es 22.65°C, mientras en la segunda parte del registro es 25.19°C. Los resultados muestran un cambio significativo de la media en la serie presión el 29 de Octubre de 2000; y un cambio significativo en la serie de humedad relativa el 1 de Enero de 2001. El cambio en la humedad relativa es por lo tanto en correspondencia con el cambio estacional. El punto de cambio estimado en la serie de precipitación es el 14 de Enero de 2001, por lo cual, según los resultados, el cambio en la media de la humedad relativa precede el inicio de las lluvias. La intensidad de precipitación antes del punto de cambio es 0.007 mm/h; luego aumenta a 0.316 mm/h, pero la información que interesa es la cantidad total de lluvia (mm). Antes del punto de cambio, la estación registró un total de 37.72 mm de lluvia; luego del punto de cambio, la precipitación totalizó 1067.78 mm de lluvia en total.

5 Conclusiones

Las transformaciones de Fourier y wavelet son lineales, lo que permite la separación de componentes estacionarios y transitorios bajo la premisa de superposición lineal, como se muestra en el presente trabajo. La transformación de Fourier es adecuada para el estudio de señales armónicas perennes, pues se basa en funciones sinusoidales. Por su parte, la transformación wavelet permite establecer si el contenido de frecuencia es permanente o no, siendo útil en el estudio de señales transitorias. Componentes estacionarios en las series de presión y temperatura, atribuibles a forzantes astronómicos, se removieron mediante Análisis Armónico, y las series residuales se analizaron nuevamente con la transformación wavelet poniendo, de este modo, las señales transitorias en perspectiva. En este sentido, la transformación wavelet facilita los análisis de varianza (ANOVA) con descomposiciones de datos en diferentes escalas de temporales. Los cambios en la varianza pueden ser observados en específicas escalas, y se puede estimar el momento de los cambios. Los cambios en la media se pueden observar también, luego de eliminar la influencia de fluctuaciones con media cero, lo que es posible con el operador wavelet debido al característico promedio cero de las funciones wavelet.

Un año de datos no es suficiente para establecer una climatología de una región específica, pero permite tener importante información sobre los procesos a escala infra-anual. Las series temporales que mayores cambios estacionales presentan, sin contar la obvia presencia de lluvias en la serie de precipitación, son las componentes de viento horizontal, y la serie de presión atmosférica. La energía del viento a escala diurna decae bruscamente en Enero 2001, mientras la serie de presión presenta un cambio en escalas de tiempo en torno 1024 horas (1.4 meses). La variabilidad de la temperatura presenta un cambio estacional marcado en escalas grandes (mayores de 512 horas), pero no en correspondencia con el inicio de la estación lluviosa. Luego, la actividad del viento, la presión atmosférica, y la precipitación son las series que mejor pueden dar información con respecto al cambio estacional, en la localidad indicada y

bajo los resultados de los datos presentados.

Estudios adicionales con series de más de un año son necesarios para comprender la denominada variabilidad interanual. Es de esperar que fenómenos tales como El Niño Oscilación Sur, o su contraparte “La Niña”, aparezcan en series más largas. Sin embargo, el conocimiento de la variabilidad infra anual es importante para determinar y cuantificar lo que serían “años anómalos” desde el punto de vista climático en un futuro. Las metodologías de análisis wavelet tienen por tanto una buena perspectiva de utilización.

Agradecimientos

Este artículo presenta parte de los trabajos desarrollados para la tesis Del Análisis de Fourier al Análisis Wavelet, presentada en la Facultad de Ingeniería en Electricidad y Computación de la Escuela Superior Politécnica del Litoral en 2014.

Las series de tiempo fueron recolectadas en la Estación Meteorológica de la Facultad de Ingeniería Marítima y Ciencias Biológicas, Oceánicas y Recursos Naturales (FIMCBOR), de ESPOL; como parte del proyecto ECLIMA, y provistas por el supervisor jefe del proyecto, el Dr. José Luis Santos Dávila, a quién expresamos nuestro agradecimiento.

Referencias

1. Mancero Mosquera, I.: Del Análisis de Fourier al Análisis Wavelet. Tesis, Escuela Superior Politécnica del Litoral, Guayaquil (2014).
2. Preisendorfer, R. W.: Principal Component Analysis in Meteorology and Oceanography. Elsevier Science Ltd., Amsterdam (1988).
3. Jenkins, G. M., & Watts, D. G.: Spectral Analysis and Its Applications, Holden Day, San Francisco (1968).
4. Mallat, S.: A Wavelet Tour of Signal Processing. Wiley, Amsterdam (1987)
5. Daubechies, I.: Ten Lectures On Wavelets. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (1992).
6. Pugh, D. T.: Tides, Surges and Mean Sea-Level. Wiley, (1987)
7. Percival, D. B.: On Estimation of the Wavelet Variance, Biometrika, Volumen 82-3, pp. 619-631 (1995).
8. Torrence, C. & Compo, G.: A practical guide to wavelet analysis. Bulletin of American Meteorological Society, Volumen 79-1, pp. 61-78 (1998).
9. Foreman, M. G. G.: Manual for Tidal Currents Analysis and Prediction, Pacific Marine Science Report 78-6, Institute of Ocean Sciences, Patricia Bay, Sidney, British Columbia, (1996).
10. Pawlowicz, R., Beardsley, B. & Lentz, S.: Classical Tidal Harmonic Analysis Including Error Estimates in Matlab using T_TIDE, Computers & Geosciences, Volume 28, pp. 929-937 (2002).
11. Tanizaki, H.: Asymptotically Exact Confidence Intervals of CuSum and CuSumSq Tests: A Numerical Derivation Using Simulation Technique, Communications in Statistics, Simulation and Computation, Volumen 24-4, pp. 1019-1036 (1995).