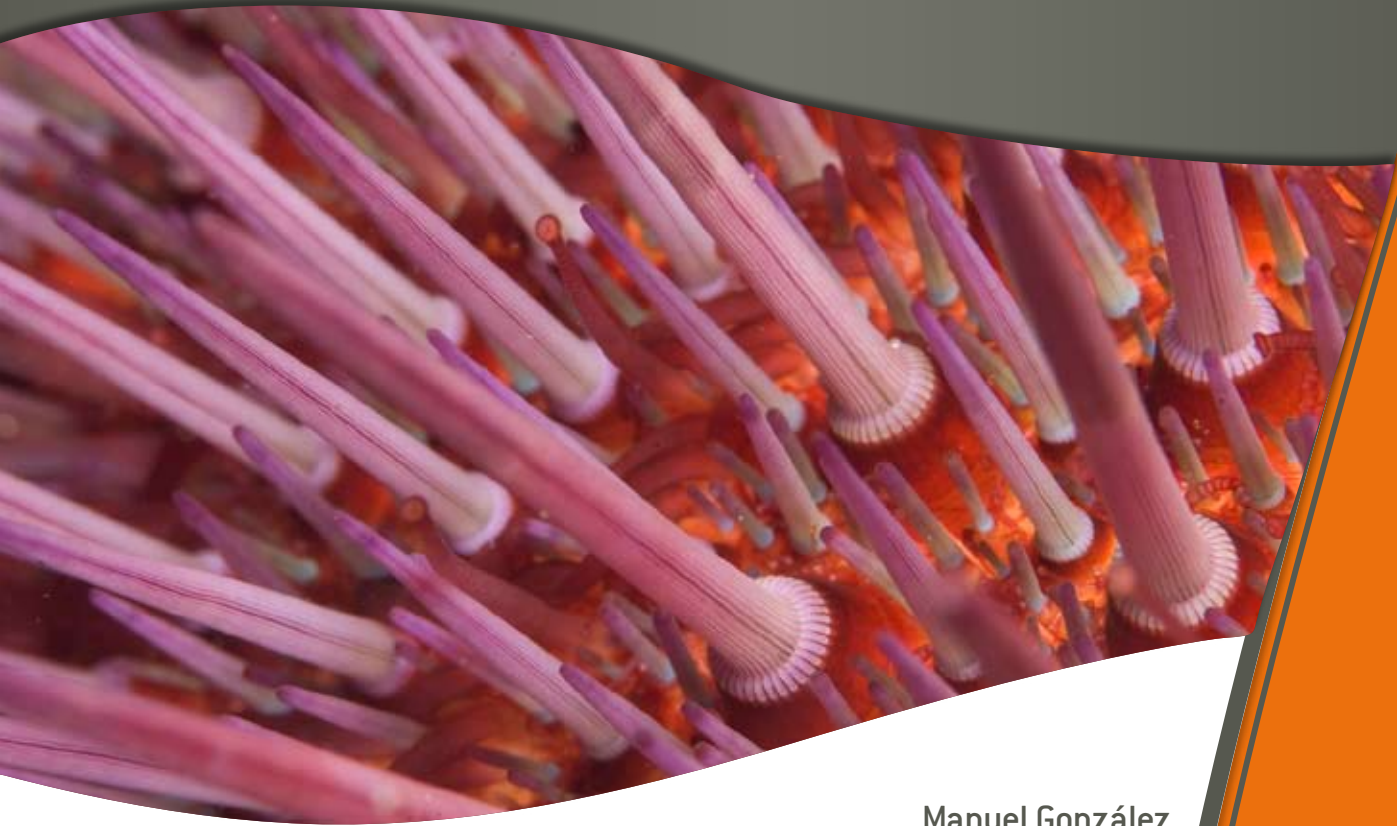


20(5) Filtros digitales para suprimir la variabilidad estacional y plurianual en series climáticas de datos



Manuel González
Almudena Fontán

González, M. y Fontán, A., 2013. Filtros digitales para suprimir la variabilidad estacional y plurianual en series climáticas de datos. *Revista de Investigación Marina, AZTI-Tecnalia*, 20(5): 71-76

La serie '*Revista de Investigación Marina*', editada por la Unidad de Investigación Marina de Tecnalia, cuenta con el siguiente Comité Editorial:

Editor: Dr. Ángel Borja

Adjunta al Editor: Dña. Mercedes Fernández Monge e Irantzu Zubiaur
(coordinación de las publicaciones)

Comité Editorial: Dr. Lorenzo Motos
Dr. Adolfo Uriarte
Dr. Michael Collins
Dr. Javier Franco
D. Julien Mader
Dña. Marina Santurtun
D. Victoriano Valencia
Dr. Xabier Irigoien
Dra. Arantza Murillas
Dr. Josu Santiago

La '*Revista de Investigación Marina*' de Tecnalia edita y publica investigaciones y datos originales resultado de la Unidad de Investigación Marina de Tecnalia. Las propuestas de publicación deben ser enviadas al siguiente correo electrónico aborja@azti.es. Un comité de selección revisará las propuestas y sugerirá los cambios pertinentes antes de su aceptación definitiva.



Edición: 1.ª Mayo 2013

© AZTI-Tecnalia

ISSN: 1988-818X

Unidad de Investigación Marina

Internet: www.azti.es

Edita: Unidad de Investigación Marina de Tecnalia

Herrera Kaia, Portualdea

20010 Pasaia

Foto portada: © Álex Iturrate

© AZTI-Tecnalia 2013. Distribución gratuita en formato PDF a través de la web: www.azti.es/RIM

Filtros digitales para suprimir la variabilidad estacional y plurianual en series climáticas de datos

Manuel González^{1*} y Almudena Fontán^{1**}

Resumen

En las series climáticas de datos (temperatura atmosférica, temperatura superficial del mar, nivel del mar, presión atmosférica, etc.) se detecta variabilidad en el rango de horas, días, meses, estaciones a un año, y de escala multianual. Este trabajo discute como eliminar la variabilidad de período inferior a uno dado, por ejemplo la estacional, mediante filtros digitales simétricos, específicos para series de datos diarios, como la temperatura atmosférica del Observatorio de Monte Igueldo (1928-2012, AEMET). Los resultados del análisis espectral muestran que el comportamiento de este tipo de filtros sobre una serie de datos con una baja proporción de huecos (menos del 1%) es muy bueno, al permitir eliminar de forma casi exacta la variabilidad de período inferior al periodo de corte con un coste computacional muy bajo. No obstante, la ventana de promediado necesaria es relativamente alta (entre 3 y 5 veces el número de datos del período de corte); lo cual, en series largas, supone una pequeña pérdida de información al inicio y final de la serie temporal.

Palabras clave: series climáticas, variabilidad estacional, filtro digital, temperatura atmosférica.

Abstract

In climatic time-series (atmospheric and sea surface temperature, sea level, atmospheric pressure, etc.), variability ranging from hours, days, months, seasons to several years can be detected. The present study discusses the application of symmetric digital filters, especially designed for daily time-series, such as the atmospheric temperature of the Observatory of Igueldo (1928-2012, AEMET), to suppress variability with periods of less than a given period. The results of the spectral analysis show that the application of the filter, in time-series with a low proportion of gaps (less than 1%), permits to eliminate the variability below the cut-off period, with a low computational cost. However, the length of the window required for the digital filtering is relatively long (between 3 and 5 times the cut-off period). This leads to loss of information at the beginning and at the end of the time-series.

Key words: climatic series, seasonal variability, digital filter, atmospheric temperature.

Introducción

La marea astronómica y oscilaciones inerciales constituyen la variabilidad de alta frecuencia (en el rango semidiurno y diurno) presente en las corrientes marinas o en el nivel del mar. La variabilidad de baja frecuencia, tanto en la temperatura atmosférica como la temperatura del mar es debida, en su mayor parte, al ciclo estacional y, en menor medida a la escala plurianual. Tanto para series de nivel del mar como corrientes o temperaturas, es habitual disponer de series de datos con frecuencia de muestreo diaria con las que caracterizar la variabilidad estacional o las tendencias a escala decadal.

En el campo de los Métodos Numéricos, los algoritmos de filtrado de series se encuentran entre las técnicas a emplear para la descripción estadística de datos (Press *et al.*, 2007). Un

filtro digital consiste en una media móvil ponderada y su diseño requiere un tamaño de ventana de promediado y el cálculo de los pesos. Esta técnica es frecuente en el análisis de señales en muy diferentes ámbitos: química, comunicación, electrónica, economía, etc. (Meyer-Baese, 2007). Por ejemplo, una de las variantes más difundidas de esta metodología es el filtro Savitzky-Golay (SG). El filtro SG es ampliamente utilizado en aplicaciones espectrométricas y sus fundamentos matemáticos fueron descritos por primera vez en el trabajo de Savitzky y Golay (1964), que constituye el quinto artículo más citado de la revista *Analytical Chemistry* (Schafer, 2011).

En Oceanografía esta técnica es usual para el tratamiento de series de datos de marea. Las oscilaciones de la superficie del mar (semidiurnas, diurnas, estacionales y plurianuales como la nodal de 18,61 años de período (Pugh, 1987; Woodworth, 2012)) son la causa de que no exista un “nivel medio del mar instantáneo” (Lie, 1978), sino que éste esté asociado a un período de promediado. Para eliminar parte de la variabilidad, y obtener una estima del nivel medio instantáneo a partir de datos horarios, se emplean los filtros de Doodson de 19, 72 y

¹ AZTI-Tecnalia, Unidad de Investigación Marina, Muelle de Herrera s/n, Zona Portuaria, 20110 Pasaia, Gipuzkoa, España.

* Email: mgonzalez@azti.es,

** Email: afontan@azti.es

168 datos (Pugh, 1987) que retiran de la señal una gran parte de la variabilidad de alta frecuencia (componentes diurnas y semidiurnas), producida por la marea astronómica.

La presencia de huecos en la serie altera la aplicabilidad del filtro digital ya que algunos de los valores necesarios para el cálculo de la serie filtrada no están disponibles. Lo mismo ocurre en el inicio y el final de la serie, aun cuando la serie esté completa, como sucede en cualquier otra media móvil. No obstante, según Pugh (1987), la variabilidad natural de los procesos oceanográficos es suficientemente importante como para que la presencia de dichos huecos no altere de forma sustancial la aplicabilidad de esta técnica de filtrado de datos.

Frente a las medias móviles, el filtro digital no supone una complicación de cálculo apreciable. La forma típica de los pesos de los filtros digitales es la de “sombrero mexicano”. Así, dentro la ventana de promediado $[t-n\Delta t, t+n\Delta t]$, el peso de las medidas más próximas al punto de promediado (t) es mayor que el de las más alejadas.

En este trabajo se muestra una metodología para el cálculo de filtros digitales adecuados al tratamiento de series de datos diarios para eliminar la frecuencia de período de meses y años. Este filtro permite retirar exactamente la variabilidad asociada a ciertos períodos prefijados (el ciclo anual, Sa, y semianual, Ssa, por ejemplo) y de forma aproximada la variabilidad de período inferior a uno dado, sin modificar la contribución del resto de componentes.

Metodología y Datos

Disponiendo de una serie de datos equiespaciados (Δt) de la variable $z(t)$ para obtener una estima de $z_0(t)$ en la que no se encuentren señales de período inferior a un cierto umbral puede utilizarse un filtro digital simétrico, denotado como el operador Ψ , consistente en una media móvil ponderada centrada:

$$\tilde{z}_o(t) = \Psi[z(t)] = \frac{w^0 z(t) + \sum_{k=1}^n w^k [z(t+k\Delta t) + z(t-k\Delta t)]}{w^0 + 2 \sum_{k=1}^n w^k} \quad (1)$$

En la que el $\vec{w} = (w^0, w^1, \dots, w^n)$ es el vector de pesos.

La función respuesta, R , en la frecuencia ω , de este filtro es (Thompson, 1983):

$$R(\omega) = w^0 + 2 \sum_{k=1}^n w^k \cos(\omega k) \quad (2)$$

Generalmente, en el diseño del filtro se impone que el denominador de (1) sea igual a 1, por lo cual, en el caso que la serie no tenga huecos no es necesario calcularlo.

Según Thompson (1983), los pesos $\vec{w}$ del filtro digital pueden calcularse como:

$$\begin{aligned} [A \cdot D] \cdot \vec{u} &= \vec{s} \\ \vec{w} &= D \cdot \vec{s} + \left[D \cdot A^T (A \cdot D \cdot A^T)^{-1} \right] \vec{u} \end{aligned} \quad (3)$$

Siendo D una matriz cuadrada de tamaño el número de pesos $(n+1)$ y A una matriz rectangular de tantas filas como frecuencias que se retiran de forma exacta ($\omega_{j=1, \dots, n_f}$) más uno (correspondiente al término estacionario, $\omega=0$, en el cual la respuesta del filtro es igual a 1):

$$D_{(n+1, n+1)} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1/2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1/2 \end{bmatrix} \quad (4)$$

$$A_{(n_f+1, n+1)} = \begin{bmatrix} 1 & 2 & \dots & \dots & 2 \\ 1 & 2\cos(\omega_1) & 2\cos(2\omega_1) & \dots & 2\cos(n\omega_1) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 2\cos(\omega_{n_f}) & 2\cos(2\omega_{n_f}) & \dots & 2\cos(n\omega_{n_f}) \end{bmatrix} \quad (5)$$

$$s_o = \frac{1}{\pi} \int_0^\pi L(\omega) d\omega; \quad s_{k=1, \dots, n} = \frac{2}{\pi} \int_0^\pi L(\omega) \cos(\omega k) d\omega \quad (6)$$

$L(\omega)$ es una función *a priori* que representa la forma de la respuesta $R(\omega)$ (2) objetivo que se desea obtener. Obviamente, resulta ventajoso que $L(\omega)$ sea tal que las integrales en las cuales aparezca (6) tengan solución analítica y que puedan ser calculadas exactamente.

Existen diferentes alternativas para $L(\omega)$: función de Hanning, Hamming, Blackman, Kaiser, Barlett, etc. (Meyer-Baese, 2007). La elección de una u otra depende esencialmente del resultado obtenido sobre una serie de datos concreta (Press *et al.*, 1992). En este trabajo se ha usado la función de Hanning que ha sido ampliamente empleada para el diseño de filtros digitales óptimos para el análisis de datos de nivel de marea y de corrientes (Thompson, 1983):

$$L(\omega) = \begin{cases} 1, & \omega < \Omega_1 \\ \frac{1}{2} + \frac{1}{2} \cos\left(\frac{\pi(\omega - \Omega_1)}{\Omega_2 - \Omega_1}\right), & \Omega_1 < \omega < \Omega_2 \\ 0, & \Omega_2 < \omega \end{cases} \quad (7)$$

Donde Ω_1 y Ω_2 son dos constantes.

De este modo, los parámetros que definen el filtro son Ω_1 y Ω_2 , el número de pesos $(n+1)$ y las frecuencias ω_j que el filtro elimina exactamente. Para el cálculo de los pesos y de la función respuesta (3), (4), (5) y (6), se ha desarrollado un código en el lenguaje de programación Fortran-90.

Para evaluar la bondad del filtro obtenido se ha empleado el análisis espectral, mediante el método de computación rápida del periodograma de Lomb (Press *et al.*, 1992). Este método permite obtener una estima de la parte de variabilidad que no ha sido eliminada por el filtro digital.

La temperatura atmosférica diaria (máxima y mínima, Figura 1) utilizada en este trabajo proviene del Observatorio de Monte Igueldo de Donostia-San Sebastián (González *et al.*, 2010) para el período 1928-2012 (AEMET). En 85 años de duración, esta serie diaria tiene muy pocos huecos: no hay datos de temperatura mínima en 58 días y de máxima en 162 días.

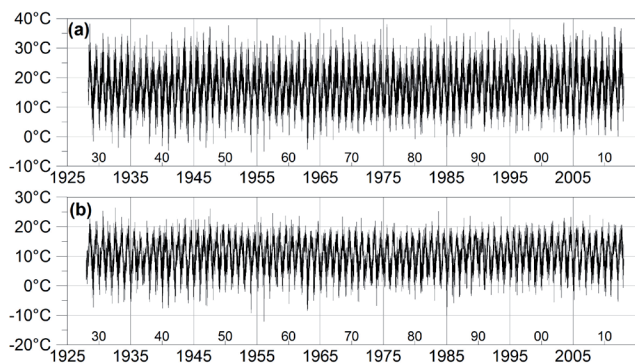


Figura 1. Datos diarios de temperatura atmosférica del Observatorio de Monte Igueldo de Donostia-San Sebastián (AEMET). (a) Temperatura máxima. (b) Temperatura mínima.

Resultados

En la Figura 2a puede verse la respuesta del filtro calculada mediante (2) y en la Figura 2b los pesos del mismo $\vec{w}$. El filtro ha sido obtenido mediante la metodología anteriormente expuesta en el caso de $\Omega_1=0,785^\circ\text{día}^{-1}$, $\Omega_2=0,997^\circ\text{día}^{-1}$, $(n+1)=1886$ e imponiendo que la respuesta para la frecuencia Ssa (semianual) y la Sa (anual) sean nulas. Puede observarse como la respuesta del filtro en el rango inferior a 1 año es próxima a cero y que, a partir de 16 meses la respuesta es prácticamente igual a 1.

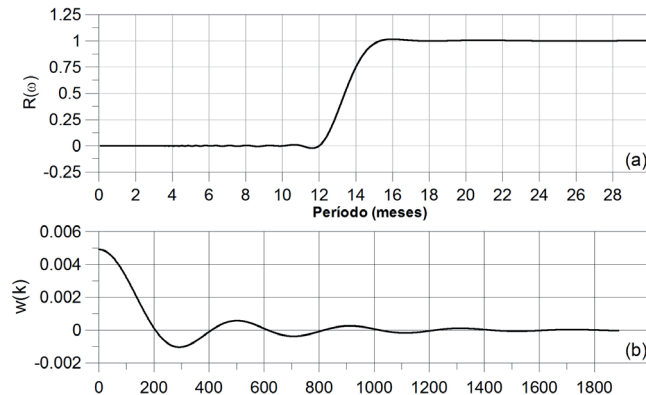


Figura 2. Filtro obtenido con $\Omega_1=0,785^\circ\text{día}^{-1}$, $\Omega_2=0,997^\circ\text{día}^{-1}$, $(n+1)=1886$ y $R(\omega_{\text{Ssa}})=R(\omega_{\text{Sa}})=0$. (a) Respuesta. (b) Pesos.

En la Figura 3 se ha representado el resultado del filtro aplicado a la serie de datos de temperatura máxima y mínima diaria de Igueldo. Se muestra también la media móvil centrada de semiancho de ventana 182 días, que es el tamaño mínimo de ventana necesario para retirar de la serie la componente semianual (Ssa) y anual (Sa). En ambos casos se ha impuesto que el denominador de (1) sea superior a 0,75, en caso contrario no se representa el valor de la serie filtrada. Se observa que el filtrado digital supone una pérdida de datos en el inicio y final de la serie.

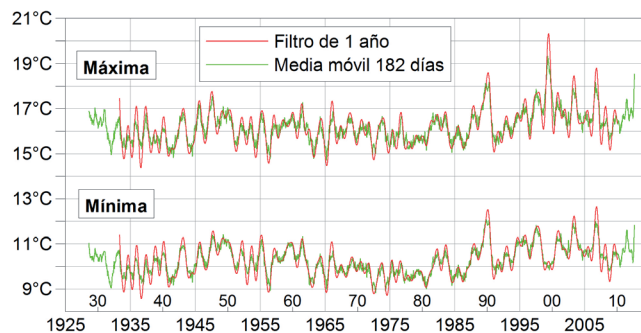


Figura 3. Filtrado de la serie de temperatura de Igueldo mediante el filtro digital obtenido con $\Omega_1=0,785^\circ\text{día}^{-1}$, $\Omega_2=0,997^\circ\text{día}^{-1}$ y $(n+1)=1886$ y con una media móvil centrada de semiancho de ventana 182 días.

En la Figura 4 se encuentra el espectro de la serie de datos original, el de la serie filtrada con el filtro digital y el de la serie filtrada mediante la media móvil centrada de 182 días de semiancho de ventana. La disminución de la energía presente en el resultado del filtro digital en la banda inferior a 1 año es varios órdenes de magnitud inferior a la de la serie original y también menor que la del espectro de la serie filtrada con la media móvil. En el rango superior a 1 año la distribución de energía en el dominio de la frecuencia no cambia prácticamente y se sigue detectando, por ejemplo, el pico energético de la componente QBO (oscilación cuasi-bienal, Baldwin *et al.*, 2001).

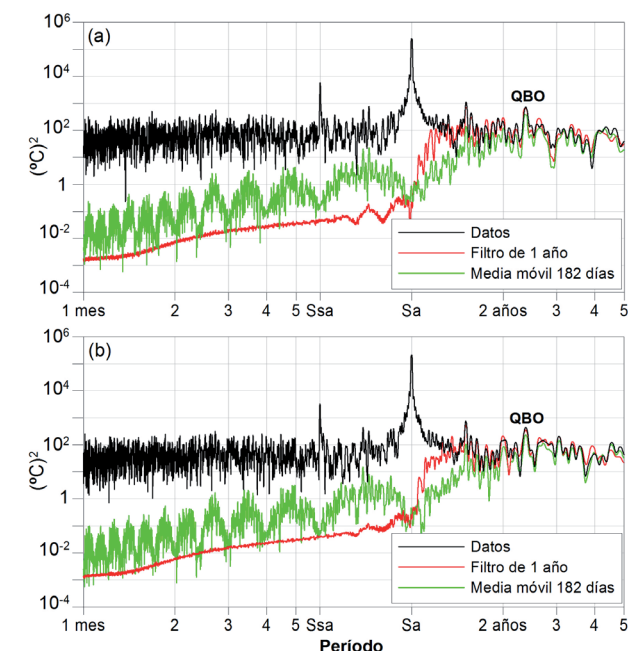


Figura 4. Espectros de la temperatura diaria de Igueldo, de la serie filtrada con el filtro digital de 1 año y de la media móvil centrada de semiancho de ventana 182 días. (a) temperatura máxima. (b) temperatura mínima.

Para retirar la variabilidad de período inferior a 4 años se ha diseñado un filtro de este tipo con $\Omega_1=0,1565^\circ\text{día}^{-1}$, $\Omega_2=0,2539^\circ\text{día}^{-1}$, $(n+1)=4286$ cuya respuesta para las componentes S_a , S_{sa} y de período igual a 4 años es igual a cero. En la Figura 5a puede verse la función respuesta (2) y en la Figura 5b los pesos del filtro.

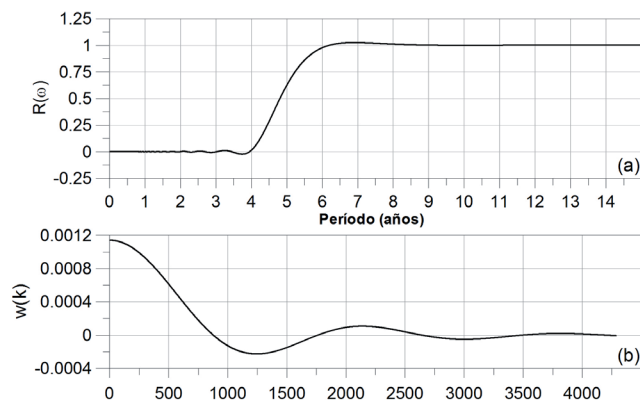


Figura 5. Filtro obtenido con $\Omega_1=0,1565^\circ\text{día}^{-1}$, $\Omega_2=0,2539^\circ\text{día}^{-1}$, $(n+1)=4286$, y $R(\omega_{Ssa})=R(\omega_{Sa})=R(\omega_{4\text{ años}})=0$. (a) Respuesta. (b) Pesos.

En la Figura 6 se muestra el resultado del filtro, de período de corte 4 años, aplicado sobre la serie de datos de temperatura máxima y mínima de Igueldo. También se ha representado el cálculo de la media móvil centrada de semiancho de ventana de 2 años, que es el tamaño mínimo de ventana necesario para retirar la contribución de las componentes estacionales y la del período de 4 años. Se ha empleado también el mínimo valor de 0,75 en el denominador de (1).

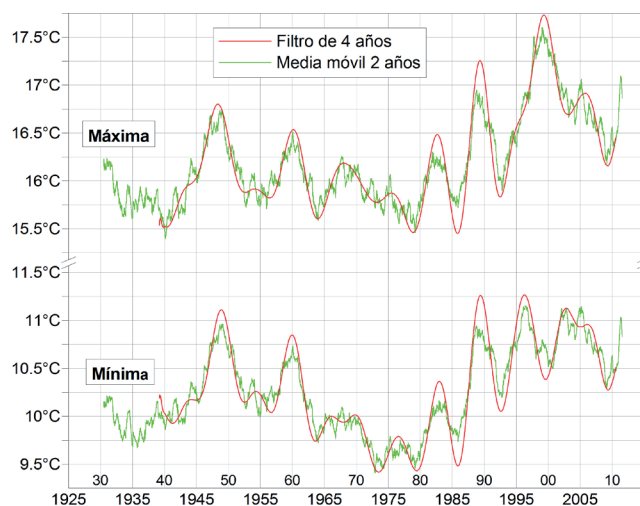


Figura 6. Filtrado de la serie de temperatura de Igueldo mediante el filtro obtenido con $\Omega_1=0,1565^\circ\text{día}^{-1}$, $\Omega_2=0,2539^\circ\text{día}^{-1}$ y $(n+1)=4286$ y con una media móvil centrada de semiancho de ventana 2 años.

El resultado del espectro de energía de la serie tratada con el filtro digital de 4 años y mediante la media móvil de 2 años de

semiancho de ventana se encuentra en la Figura 7. Como puede observarse en el espectro, en general el filtro digital es mucho más eficaz que la media móvil, retirando mayor cantidad de energía en la banda objetivo (período inferior a 4 años) de forma uniforme. Asimismo, el filtro digital elimina en mayor medida la contribución de la componente QBO que el filtrado mediante la media móvil.

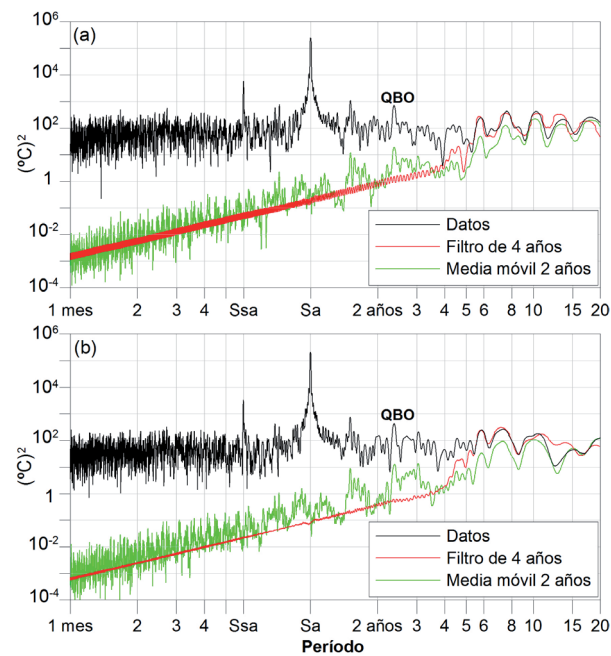


Figura 7. Espectros de la temperatura diaria de Igueldo, de la serie filtrada con el filtro digital de 4 años y de la media móvil centrada de semiancho de ventana de 2 años. (a) temperatura máxima. (b) temperatura mínima.

Discusión

El filtro digital, consistente en una media móvil ponderada, no supone mayor complejidad matemática que una media móvil. Esta técnica, al contrario que por ejemplo la de estima de componentes mediante análisis armónico, tiene las ventajas de las aproximaciones no paramétricas (Yang y Zurbenko, 2010) ya que no asume una forma de variabilidad prefijada en la serie de datos por lo que es de aplicación a temperatura atmosférica, temperatura del mar, nivel del mar, corrientes, etc.

Frente a otras técnicas, como el filtro Kolmogorov-Zurbenko (KZ) y su versión adaptativa (KZA) (Zurbenko *et al.*, 1996), que han sido empleadas en otros trabajos para el estudio de series de este tipo (González *et al.*, 2011), presenta la ventaja de retirar de forma específica la energía de la banda de frecuencias objetivo, manteniendo inalterada la contribución de las componentes de período superior.

En cuanto al tamaño de ventana necesario, los resultados obtenidos indican que para retirar la variabilidad estacional (de período inferior a 1 año, 365 datos de frecuencia de muestreo

diaria) es necesario promediar en, aproximadamente, 5 veces ese número de datos (1886 datos). La relación entre período de corte y semiancho de ventana de trabajo es muy parecida a la del filtro con datos horarios “120i913” (Thompson, 1983) que tiene un semiancho de ventana de 120 (con $\Omega_1=9^\circ.\text{hora}^{-1}$ y $\Omega_2=13^\circ.\text{hora}^{-1}$) para retirar la variabilidad de período inferior a 1 día (24 datos). Para eliminar la contribución de período inferior a 4 años (1460 datos diarios) se necesitan unos 4300 datos; por tanto, para la relación entre semiancho de ventana y período de corte es cercana a 3. En el caso de series de datos horarios (característicos de las series de datos de marea), para retirar de forma exacta la variabilidad estacional (período inferior a 1 año), se necesita algo menos de 19000 datos (González *et al.*, 2013), lo cual supone una relación de poco más de 2.

Conclusiones

La aplicación de la técnica propuesta por Thompson (1983) permite obtener, de forma relativamente sencilla, filtros digitales adecuados para series de datos equiespaciados.

En este trabajo se presentan dos ejemplos de aplicación sobre series de datos diarios: uno para retirar la variabilidad estacional y otro para retirar la de período inferior a 4 años.

El uso del filtro a las series de temperatura de Monte Igueldo (Donostia-San Sebastián, AEMET), que apenas tienen huecos, pone de manifiesto la potencia del método ya que elimina la variabilidad estacional y la de período inferior a 4 años, sin modificar la distribución de energía en los períodos superiores a los de corte.

La principal desventaja del método es el tamaño de ventana de promediado necesario, lo cual supone una pérdida de información al inicio y al final de la serie filtrada. Con datos diarios, para retirar los períodos inferiores a 1 año (365 datos) se requieren unos 1900 datos; en el caso de la contribución de período inferior a 4 años (1460 datos diarios) se requiere un filtro de casi 4300 datos.

Esta proporción entre semiancho de ventana de promediado y período de corte pone de manifiesto la importancia del espaciado temporal de los datos; ya que, cuanto menor sea éste, menor es el número de datos necesario y menor es la pérdida de datos en la fases inicial y final de la serie.

Agradecimientos

Los autores desean agradecer a la Dirección de Ciencia y Tecnología del Gobierno Vasco (proyecto Saiotek-DECLIM) la financiación de una parte de este trabajo, a AEMET la puesta a disposición de la información de temperatura atmosférica del Observatorio de Monte Igueldo de Donostia-San Sebastián y a la Dra. Ainhoa Caballero sus valiosas aportaciones en la revisión de este trabajo.

Esta es la contribución número 626 de AZTI-Tecnalia (Unidad de Investigación Marina).

Bibliografía

- Baldwin, M.P., Gray, L.J., Dunkerton, T.J., Hamilton, K., Haynes, P.H., Randel, W.J., Holton, J.R., Alexander, M.J., Hirota, I., Horinouchi, T., Jones, D.B.A., Kinnersley, J.S., Marquardt, C., Sato, K., Takahashi, M., 2001. The quasi-biennial oscillation. *Reviews of Geophysics*, 39: 170–229.
- González, M., Fontán, A., Borja, A., Del Campo, A., Esnaola, G., Ferrer, L., Goikoetxea, N., Mader, J., Uriarte, Ad., Valencia, V., 2010. Trend analysis of multidecadal datasets of air and sea surface temperatures within the southeastern Bay of Biscay. *Thalassas*, 26 (2): 23-31.
- González, M., Fontán, A., Esnaola, G., Valencia, V., 2011. Variaciones multidecadales notables en la temperatura atmosférica, temperatura superficial y nivel del mar en el sudeste del golfo de Vizcaya, detectadas mediante filtros “Kolmogorov-Zurbenko”. *Revista de Investigación Marina*, 17(8): 1-15.
- González, M., Fontán, A., Esnaola, G., 2013. Digital filtering of seasonal and long-term sea level components. Application to sea level data in Brest (France), Newlyn (UK) and Santander, Coruña and Vigo (Spain). *Proceedings 7ª Asamblea Hispano-Portuguesa de Geodesia y Geofísica*, Donostia-San Sebastián Junio de 2012.
- Lie H-J., 1978. A review of some filters for the calculation of daily mean sea level. *Journal of the Oceanological Society of Korea*, 30 (2): 1-4.
- Meyer-Baese, U., 2007. Digital Signal Processing with Field Programmable Gate Arrays. *Springer Series on Signals and Communication Technology*. 3ª Ed., 774 pp
- Press, W.H., S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, 1992. Numerical Recipes in Fortran 77: The Art of Scientific Computing. *Cambridge University Press*, 936 pp.
- Press, W.H., B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, 2007. Numerical Recipes: The Art of Scientific Computing. *Cambridge University Press*, 1235 pp.
- Pugh, D.T., 1987. Tides, surges and mean sea-level. Chichester, UK, John Wiley & Sons Ltd, 472 pp.
- Savitzky, A. y Golay, M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36(8): 1627-1639.
- Schafer, R.W., 2011. What Is a Savitzky-Golay Filter?. *IEEE Signal Processing Magazine*, 28(4): 111–117.
- Thompson, R.O.R.Y., 1983. Low-pass filters to suppress inertial and tidal frequencies. *Journal of Physical Oceanography*, 13: 1077–1083.
- Woodworth, P.L., 2012. A Note on the Nodal Tide in Sea Level Records. *Journal of Coastal Research*, 28: 316–323.
- Yang, W. y Zurbenko, I., 2010. Nonstationarity. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(1): 107–115.
- Zurbenko, I.G., Porter, P.S., Rao, S.T., Ku, J.Y., Gui, R., Eskridge, R.E., 1996. Detecting discontinuities in time series of upper air data: development and demonstration of an adaptive filter technique. *Journal of Climate*, 9: 3548–3560.

