



## REVISIÓN DEL MÉTODO DE TRATAMIENTO DE DATOS EN LA APLICACIÓN DE LA DIRECTIVA COMUNITARIA 2002/49/CE DE GESTIÓN Y CONTROL DEL RUIDO AMBIENTAL

PACS: 43.50.Rq

Cueto Ancela, José L.; Rivas Calvete, Silvia; Hernández Molina, Ricardo  
Laboratorio de Acústica y Vibraciones. Universidad de Cádiz.  
C.A.S.E.M .Campus Río San Pedro  
Puerto Real 11510.Cádiz. España  
Tel: 956 016 051/956 016 140  
Fax:956 016 051  
E-Mail: lav@uca.es

### ABSTRACT

Nowadays European Directive develops useful procedures in order to evaluate by means of index, the acoustic pollution affecting urban area. To apply properly those methods, it is necessary to use long data sequence. Is usual to find in this kind of series some peculiar measures or events witch could change indicatively our acoustic pollution index data.

So, we suggest a simple data treatment method as a first step in the way of develop better process of acoustic index construction.This methodology we permit us to introduce the "acoustic climate" concept.

### RESÚMEN

La nueva Directiva Comunitaria define los protocolos a aplicar a la hora de elaborar índices que definan el grado de afección acústica de una determinada zona. Para el cálculo de dichos índices, lo mas apropiado es partir de una serie de datos lo suficientemente larga. Se comprueba fácilmente que en dichas series pueden ocurrir sucesos y anomalías que pueden alterar el valor del índice de afección de forma significativa.

Sugerimos a continuación un sencillo método de tratamiento de este tipo de series, como paso previo para un cálculo más aproximado de dichos índices, y a partir del que podemos introducir el concepto de "clima acústico".

### INTRODUCCIÓN

El objetivo principal del presente estudio, es el de comprobar la utilidad de ciertas herramientas de tratamiento de series temporales, en la estimación del «indicador de ruido» «*Lden*» (Indicador de ruido día-tarde-noche). Este indicador queda detallado en la Directiva 2002/49/CE del Parlamento Europeo y del Consejo de 25 de Junio de 2002 sobre Evaluación y Gestión del Ruido Ambiental.

En principio, cabe suponer que al disponer de una serie lo suficientemente larga de datos sobre el ruido de una zona, la estimación del *Lden* se hará cada vez más fiable. Para obtener una

serie de esta naturaleza es indispensable disponer de un equipo de intemperie que sea capaz de almacenar todos los datos requeridos, durante periodos de semanas, meses e incluso años. En muchos casos, esto no es posible, y por lo tanto, aparecen errores indetectables a priori en el cálculo del índice comentado. Así sería útil razonable desarrollar técnicas que permitan reducir el error en la estimación del *Lden*.

## CONCEPTO DE “CLIMA DE RUIDO AMBIENTAL”

Se puede definir el “clima de ruido ambiental” de un área, como aquel ruido que permanece constante a lo largo del tiempo. La diferencia entre lo que llamaremos clima y las medidas puntuales que se puedan realizar, corresponde a la variabilidad del sistema, o dicho de otra manera, a la componente aleatoria de ese sistema analizado. El clima, por tanto, será aquella componente o componentes deterministas, normalmente una componente continua, una serie de componentes periódicas y/o pseudoperiódicas y, posiblemente una tendencia.

El uso del índice *Lden* debe tender hacia un valor que indique el clima de ruido. *Lden* es por tanto un promedio y además tiende a despreciar la importancia de ciertos sucesos y distorsiones. Sin embargo para que dicho índice sea fiable debería ser obtenido a partir de una serie de datos correspondiente a un año (o un mes al menos), ya que de lo contrario dichos sucesos y distorsiones puede que tengan la suficiente entidad para desvirtuar el índice.

Como esto no será posible siempre, presentamos una herramienta que permita paliar los posibles errores. Así, en el caso de poseer series cortas el índice *Lden* ha de calcularse sobre esas componentes deterministas y no sobre la serie cruda (original) de datos. En el desarrollo de este trabajo, dispondremos de series de LAeq calculados cada hora.

Así, cuando se pretende definir el clima de un área afectada por ruido de tráfico, la introducción de factores (focos de ruido) fuera de lo normal, por ejemplo el paso desacostumbrado ambulancias durante esa medida, introducirá un error en la estimación. Es esencial en estos casos tener en cuenta la relación existente entre el tiempo que afecta el factor distorsionante y el tiempo utilizado para el cálculo del *Lden*. Cuanto menor sea esta relación, menos afectará ese factor al cálculo.

## FILTRADO DE LA SERIE TEMPORAL BASADA EN LA DETERMINACIÓN DE SUBESPACIOS

Supongamos que la serie de datos medidos es  $\mathbf{x}[n]$ ,  $n=0,1,\dots,N-1$ , está compuesta por la suma de  $P$  componentes pertenecientes al clima de ruido inmersas en variabilidad aleatoria a la que hemos hecho alusión. La matriz de autocovarianza  $\mathbf{R}_{xx}$  calculada a partir de  $\mathbf{x}[n]$  es una matriz cuadrada y definida positiva  $M \times M$ .  $M$  es una fracción de  $N$  que permite que la matriz de autocovarianza rinda valores estadísticamente fiables,  $P < M < N$ . Si clima y variabilidad aleatoria son dos procesos no correlacionados, podemos expresar  $\mathbf{R}_{xx} = \mathbf{R}_{ss} + \mathbf{R}_{nn}$ , como:

$$\mathbf{R}_{xx} = \mathbf{S} \mathbf{P}_{ss} \mathbf{S}^H + \mathbf{s}_n^2 \mathbf{I} = \mathbf{E} \mathbf{\Lambda} \mathbf{E}^H \quad (1)$$

$\mathbf{P}$  es la matriz diagonal  $M \times M$  de autovalores reales y no negativos  $\lambda_{k,k} = \lambda_k$ , para  $k=1,2,\dots,M$ , y están ordenados en orden decreciente  $\lambda_1 > \lambda_2 > \dots > \lambda_M$ . La matriz  $\mathbf{E}$  es la  $M \times M$  formada por “ $k$ ” autovectores  $\mathbf{E}_k[m]$ , ordenados en columnas de izquierda a derecha,  $k=1,2,\dots,M$ , cuyos lags son  $m=0,1,2,\dots,M-1$ .  $\mathbf{S}$  es la matriz de Vandermonde  $N \times P$ , compuesta por las componentes (en realidad son exponenciales complejas, cada par de ellas definen una oscilación) del clima  $\mathbf{S} = [\mathbf{s}_1[n], \mathbf{s}_2[n], \dots, \mathbf{s}_P[n]]$  y  $\mathbf{P}_{ss}$  expresa la energía de cada componente. El superíndice  $H$  es por hermitiana traspuesta y  $\mathbf{s}_n^2$  es la varianza de la variabilidad aleatoria, siendo la matriz  $\mathbf{I}$ , que es la matriz de identidad  $M \times M$ . Por tanto, cada autovalor expresa la varianza de la serie temporal en la dirección diseñada por su respectivo autovector.

En este escenario, los primeros (más energéticos)  $P$  autovectores de  $\mathbf{R}_{xx}$  (de la serie cruda) son los autovectores de  $\mathbf{R}_{ss}$  (del clima). Así podemos separar los autovalores (y sus respectivos

autovectores) en variabilidad aleatoria ( $\lambda_k = s_n^2$ ) y clima de ruido ( $\lambda_k > s_n^2$ ) eigenfunctions. Por tanto podemos escribir

$$R_{xx} = R_{ss} + R_{nn} = E_s \Lambda_s E_s^H + E_n \Lambda_n E_n^H \quad (2)$$

donde  $\lambda_n = s_n^2 I$  (solo cuando la variabilidad se comporta como ruido blanco y  $R_{xx}$  es una matriz de autocovarianza real). La herramienta SVD (Descomposición en Valores Singulares) se considera un método fiable y robusto para separar los Subespacios de Clima y Variabilidad de una serie de datos perturbada por ruido blanco [1] [2].

Las columnas de  $E_s$  (autovectores que forman el Subespacio Clima) expanden el mismo espacio que  $S$  (así que  $E_n$  expande el espacio ortogonal complementario). Bajo las condiciones en las que  $S$  contiene oscilaciones reales los pares de autovectores aparecen como pares seno/coseno, con autovalores de igual valor [3]. Nosotros usaremos esos autovectores como base para capturar las oscilaciones que expresan el clima de ruido ambiente. Este proceso de filtrado seleccionando filtros "a posteriori" es preferible y se muestra muy superior que la elección de una base ortogonal de funciones fijas (por ejemplo compuesta de senos y cosenos).

Basado en técnica de análisis de autovectores utilizamos la herramienta SSA (Singular Spectral Analysis) aplicada al dominio temporal [3] [4] [5] [6]. Esta herramienta rinde la reconstrucción (RCs) de la evolución en el tiempo (amplitud y fase) de las componentes del clima de una oscilación pseudoperiódica con mayor detalle que los métodos clásicos [4].

La serie descrita por el vector  $x[n]$  se "embebe" en un espacio M-dimensional, de tal manera que se forma una matriz de estado, esta matriz de datos (N-M+1)xM es llamada matriz de trayectorias [7]. Así se ha forzado a una serie a convertirse en una serie multivariante (M variables). La matriz de autovectores debe ser interpretada como un banco de filtros ("autofiltros") de Media Movil. Proyectando la Matriz de datos sobre los filtros obtenemos una nueva matriz multivariante filtrada [3] [4] de la siguiente manera

$$a_k[n] = \sum_{m=0}^{M-1} x[n+m] E_k[m] \quad (3)$$

donde tenemos cada k-ésima Componente Principal (PC), y donde  $n=0,1,2,\dots,N-M$  y los lags temporales son  $m=0,1,2,\dots,M-1$ . Desafortunadamente los PCs no tienen la longitud de la serie original  $x[n]$  (N) y carecen de información sobre la fase. Convolucionando los M autovectores, con sus respectivas PCs y ponderando los resultados convenientemente, obtenemos las llamadas Componentes Reconstruidas (RC) de longitud N. Por lo tanto cada RC representa la versión filtrada de la serie original  $x[n]$ , capturada por el k-ésimo autofiltro  $E_k[m]$ . La expresión que describe el proceso de reconstrucción es la siguiente:

$$\hat{s}_k[n] = \beta \sum_{m=0}^{M-1} a_k[n-m] E_k[m] \quad (4)$$

El factor, " $\beta$ ", para cada vector RC es  $\beta=1/M$  para la parte central de la serie ( $M/2 \leq n < N-M/2$ ). En las colas, cuando  $0 \leq n < M/2$  y  $N-M/2 < n < N$ , el factor de ponderación depende del número de elementos de cada PC que intervienen en la reconstrucción:  $1 < \beta < 1/(M-1)$  [4]. Por supuesto, la suma de todas las RCs rinde  $x[n]$ , y la suma de las P principales RCs da el clima de ruido ambiente.

## TRATAMIENTO DE UNA SERIE REAL

La serie de prueba (fig.1) fue tomada en octubre del 2002 en la estación situada en la calle de tipo residencial, General Ricardos de Madrid. La serie tiene una duración de 744 horas, y está registrada desde las 00:00 del martes 1 hasta las 24:00 del jueves 31 ambos de octubre (a esta serie la llamaremos GR-31). Hay que decir que el LAeq de la hora 00:00 corresponde al recogido en el intervalo que va desde las 23:01 del día anterior hasta las 24:00.

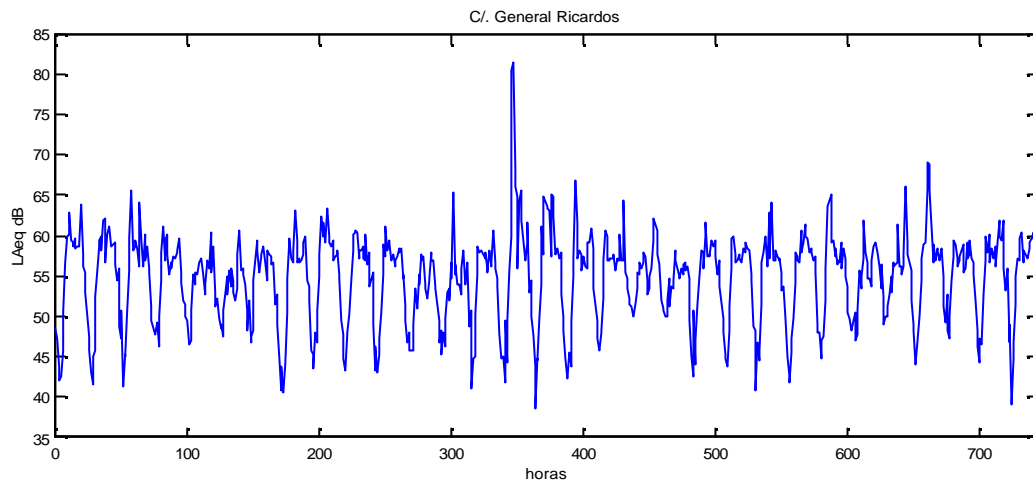


Figura 1

Calculando los índices correspondientes, se obtiene:

GR-31	<i>Ldía</i>	<i>Ltarde</i>	<i>Lnoche</i>	<i>Lden</i>
	61,7	58,3	50,5	61,7

Se ha extraído de esta serie una parte en la que se manifiesta una gran variabilidad que no se da en el resto, desde el lunes 14 hasta el miércoles 16 de octubre, en total 72 horas de análisis (a esta serie la llamaremos GR-03). Si se extrae desde la hora 312 hasta la hora 383 (ambas inclusive, lo que significa que introducimos las 23:00 del domingo 13 y excluimos las 23:00 del miércoles 16) se obtiene :

GR-03	<i>Ldía</i>	<i>Ltarde</i>	<i>Lnoche</i>	<i>Lden</i>
	69,4	58,1	49,4	67

Estas diferencias son lo suficientemente importantes como para tomarlas en consideración. Existen dos opciones a la hora de elaborar herramientas de filtrado convenientes:

1. Calcularlo a partir de la serie corta, pese a los inconvenientes que ello conlleva sobre la calidad de la estimación
2. O bien a partir de otras estaciones de control con niveles de correlación cruzada altos y que posean series de datos largas, con las que confeccionar unas herramientas de filtrado más precisas.

Siguiendo esta segunda opción, se tomó la estación de la Calle Fernández Ladreda (a esta serie la llamaremos FL-31) como estación de control (el coeficiente de correlación entre GR-31 y FL-31 es de 0,83), utilizando la serie temporal de medidas de ruido tomadas durante el mismo mes de octubre (fig.2).

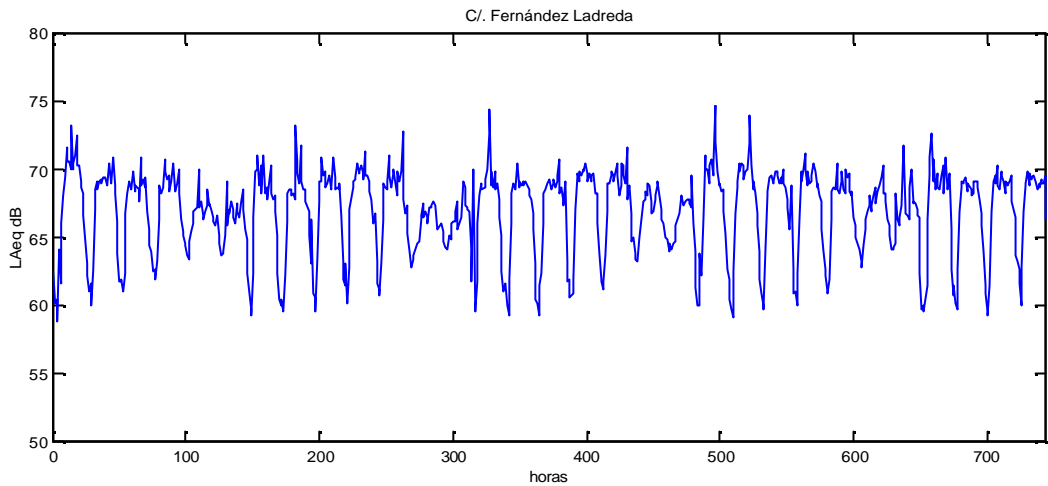


Figura 2

Así se han sometido ambas series (GR-31 y FL-31) al análisis propuesto. Se deduce que las componentes significativas para  $M=24$ , en ambos casos son aquellas que corresponden a los autovalores más energéticos  $k=1,2,3,4$  ( $P=4$ ).

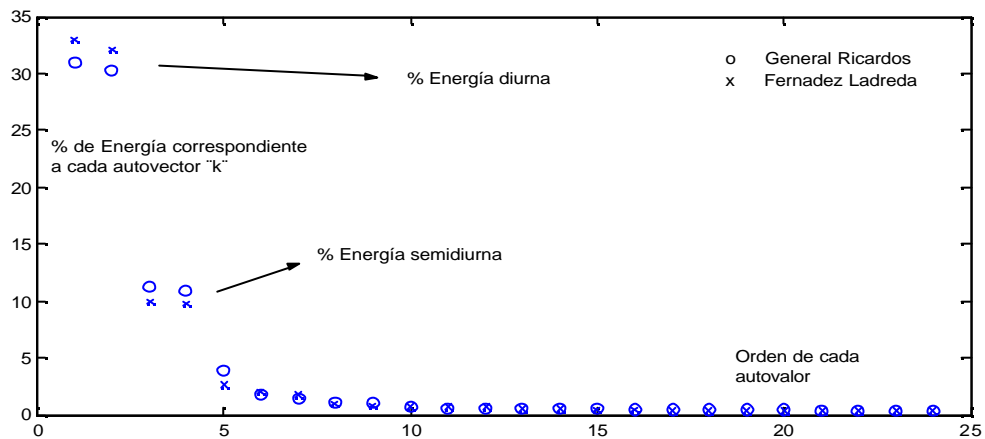


Figura 3

Encontramos que los autovectores  $k = 1, 2, 3$  y  $4$ , correspondientes a la serie “de prueba” GR-31 y a la serie “de control” FL-31, son prácticamente iguales en frecuencia, amplitud y fase. Utilizando una herramienta de análisis frecuencial “MUSIC” [8] (MULTiple Signal Classification), encontramos que estos autovectores definen unas componentes principales de frecuencia  $0,0415$  ciclos/hora y  $0,0833$  ciclos/hora, lo que se corresponde con periodos de aproximadamente  $24$  (diurno) y  $12$  horas (semidiurno), respectivamente, no olvidemos que la mayor parte de la energía corresponde al tráfico.

Podemos echar de menos una componente semanal del análisis frecuencial. Esta no aparece porque la “ $M$ ” la dimensión embebida con la que hemos realizado el análisis es  $M=24$  horas, demasiado pequeña para reconocer componentes con periodos más largos. Hemos escogido esta “ $M$ ” porque la serie corta que nos preocupa (GR-03) tiene  $72$  horas. De todas maneras la componente semanal aparece modulando la señal diurna. El resto, como se ve en la gráfica de autovalores (fig.3) corresponden a “ruido blanco”, sin información sobre el clima de ruido.

Aplicamos el análisis propuesto a la serie GR-03. Hayamos las Componentes Significativas  $1$  a  $4$  de esta serie corta filtrándolas con los autovectores extraídos de la serie de control FL-31. Sumadas estas componentes (RCs) principales forman la serie que aparece en la figura 4 (línea sólida) y que llamaremos GR-03-CLI, como se ve a continuación.

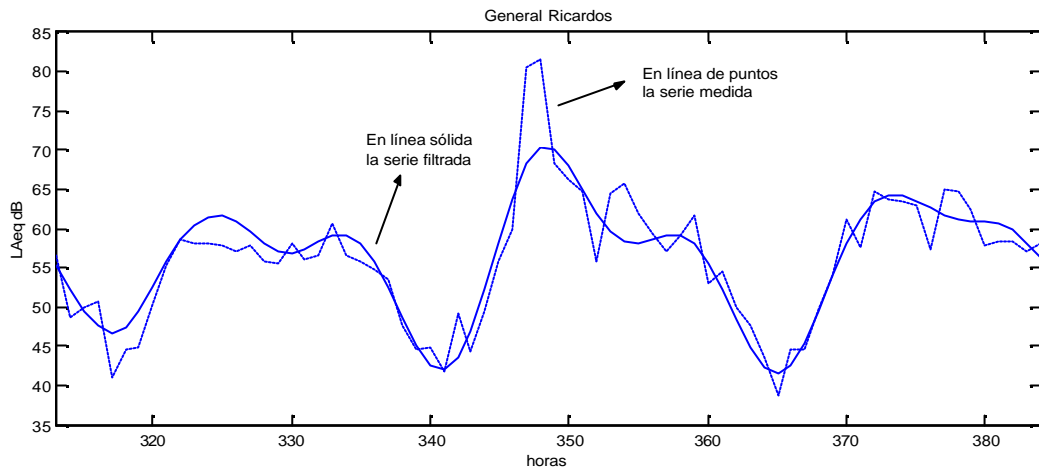


Figura 4

Es decir, hemos reconstruido las componentes de las oscilaciones diurna y semidiurna (lo que hemos venido llamando clima de ruido), y encontramos que aun a pesar de seguir contando solo con 72 horas de análisis, tenemos lo siguiente:

GR-03-CLI	<i>L</i> día	<i>L</i> tarde	<i>L</i> noche	<i>L</i> den
	63,1	58,5	49,3	62,2

### CONCLUSIONES

Filtrar la serie larga de ruido significa que vamos a caracterizar mediante el índice *Lden* solo el clima de ruido y no aquellas variaciones carentes de información al respecto de la afección acústica del lugar a largo plazo. En la gráfica de la figura 5 se presenta la serie filtrada (GR-31-CLI) en relación a la original (GR-31). Aquí, GR-31-CLI ha sido calculada mediante los autovectores generados por la propia GR-31 (también hemos calculado FL-31-CLI a partir de los autovectores generados por FL-31, y encontramos que el coeficiente de correlación entre GR-31-CLI y FL-31-CLI es de 0,95).

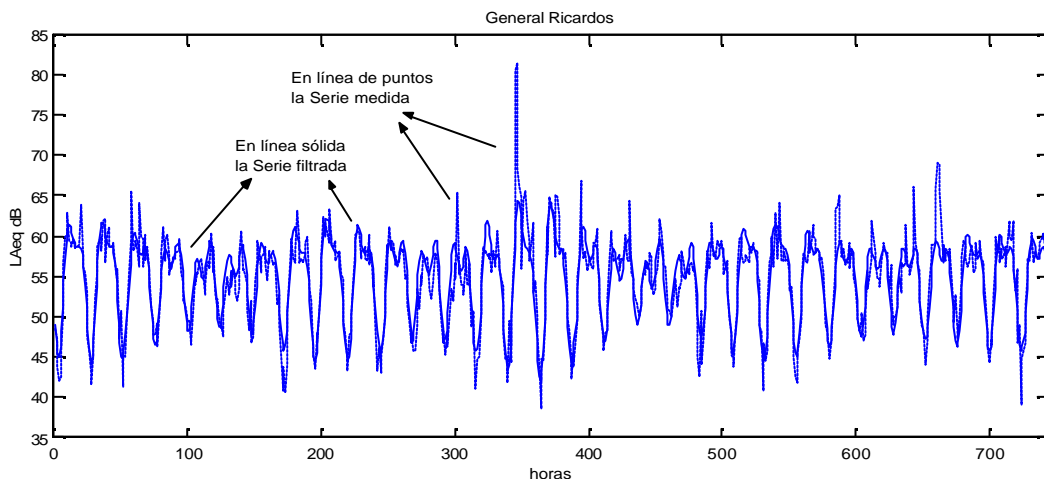


Figura 5

Sabemos que al calcular el índice *Lden* sobre la serie cruda larga y la serie larga del clima de ruido, las diferencias que se dan son marginales, y no suelen sobrepasar uno o dos decibelios. Esto se debe, claro está, a la promediado energético que se da en todo el proceso de cálculo. En nuestro caso, el clima de ruido en esta zona calculado sobre las 744 horas es de:

GR-31-CLI	<i>Ldía</i>	<i>Ltarde</i>	<i>Lnoche</i>	<i>Lden</i>
	58,3	57,8	50	60

No ocurre lo mismo con series cortas, donde algunos factores irregulares pueden afectar a la estimación del índice con la calidad necesaria. En ese caso, averiguar cuales son las componentes del clima acústico de la zona, evitara errores, que como hemos visto en nuestro caso llegan a los 5,3 dBs, pero se pueden producir desviaciones mayores si hubiéramos cogido un periodo mas pequeño (por ejemplo con 48 horas hubiéramos tenido desviaciones de hasta de 8 dBs).

## AGRADECIMIENTOS

Gracias al Departamento de ruido del Ayuntamiento de Madrid por los datos facilitados con los que se ha realizado este estudio, y en especial a Placido Pereda y Beatriz López por su interés y su amabilidad.

## BIBLIOGRAFÍA

- [1] Kumaresan, R. and Tufts, D.W. (1980). **Data adaptive principal components signal processing**. IEEE Proc. Conf. on Decision and Control. pp. 949-954
- [2] Tufts, D.W.; Kumaresan, R. and Kirsteins, I. (1982). **Data adaptive signal estimation by singular value decomposition of a data matrix**. Proceedings of the IEEE, vol.70, pp.684-685
- [3] Vautard, R. and Ghil, M. (1989). **Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series**. Physica D, 35. pp.395-424
- [4] Vautard, R., Yiou, P. and Ghil, M. (1992). **Singular spectrum analysis: A toolkit for short, noisy chaotic signals**. Physica D, 58. pp.95-126
- [5] Broomhead, D.S. and King, G. (1986). **Extracting qualitative dynamics from experimental data**. Physica D, 20. pp.217-236
- [6] Cueto, J.L. and Hernandez, R. (1999). **Estimation of Signals with Slowly Changing Frequencies from Noisy Time Series**. Sixth International Congress on Sound and Vibration. Copenhagen
- [7] Takens, F. (1981). **Detecting strange attractors in turbulence. Dynamical Systems and Turbulence**. Edited by Rand, D.A. & Young, L.S. Springer-Verlag, Berlin. pp. 366-381
- [8] Schmidt R. (1986). **Multiple Emitter Location and Signal Parameter Estimation**. IEEE transactions on Antennas and Propagation. Vol. AP34; pp 276-290.