

# Métodos estadísticos para la identificación automática de fondos marinos

Javier Tarrío-Saavedra<sup>1,2,3</sup>, Noela Sánchez Carnero<sup>4,5</sup>, Andrés Prieto<sup>2,3,6</sup>

<sup>1</sup>Departamento de Matemáticas, Escola Politécnica Superior, Universidade da Coruña, Spain

<sup>2</sup> ITMATI, Santiago de Compostela, Spain

<sup>3</sup> CITIC, A Coruña, Spain

<sup>4</sup>Centro Para el Estudio de Sistemas Marinos (CESIMAR). Centro Nacional Patagónico (CENPAT-CONICET). Argentina.

<sup>5</sup>Grupo de Oceanografía Física. Universidad de Vigo. Spain

<sup>6</sup>Departamento de Matemáticas, Facultade de Informática, Universidade da Coruña, Spain

`javier.tarrío@udc.es`

**Resumen.** En este trabajo se propone el uso de metodologías estadísticas basadas en la clasificación no supervisada para la identificación automática de tipos de fondos marinos en zonas costeras. Para ello se parte de datos acústicos obtenidos mediante una ecosonda de haz simple (a 200kHz) acoplada a un barco de pequeñas dimensiones. Cada una de las curvas acústicas resultantes ha sido preprocesada mediante la aplicación de correcciones de tiempo (alargamiento del eco con la profundidad), potencia (atenuación de la onda con la distancia) y longitud de ping (deformación del eco debido a ésta), con el objeto de eliminar su dependencia con respecto a la profundidad. Los datos experimentales han sido obtenidos en un ambiente controlado en la región de Cabo de Palos (Murcia, España), estudiándose tres tipos diferentes de fondo: arenoso, arenoso con vegetación y roca. Las técnicas estadísticas adaptadas y aplicadas a este caso particular pertenecen a la clasificación clúster a partir de series de tiempo. De hecho, teniendo en cuenta que en los problemas de identificación reales no se conocen de antemano las clases de fondo existentes, se ha abordado el problema de identificación mediante una perspectiva de clasificación no supervisada basado en el cálculo previo de matrices de disimilitudes y la aplicación de métodos de clasificación clúster jerárquicos. Los resultados obtenidos, identificándose correctamente el 93% del total de fondos, sin apenas confusión entre sus clases, apoyan el uso de técnicas de clasificación automática en este ámbito para la correcta caracterización de los fondos marinos y su aprovechamiento.

**Palabras claves:** Fondo marino, Sónar, Clasificación clúster, Series de tiempo.

## 1 Introducción

Cuando se requiere estudiar el fondo marino, la columna de agua representa la primera y difícilmente franqueable barrera. El simple hecho de su presencia limita drásti-

amente el uso de las dos principales metodologías utilizadas en el medio terrestre: i) métodos directos de toma de datos, en este caso a través de buceo, restringidos por razones de seguridad, o del uso de dragas, limitadas por la necesidad de infraestructura e inversión de tiempo para la toma de datos, e ii) imágenes de satélite, limitadas a áreas muy someras donde la atenuación de la señal electromagnética en el agua todavía permite la obtención de datos [1]. Teniendo en cuenta las restricciones de las técnicas anteriormente mencionadas, la acústica submarina representa la mejor alternativa metodológica para el estudio del fondo marino [2]. La señal acústica es una onda longitudinal que, en un medio denso como el agua, se transmite a elevadas velocidades (alrededor de los 1500 m/s) y grandes distancias antes de ser atenuada por debajo del ruido ambiente. En consecuencia, su uso y estudio permite trabajar a grandes profundidades. En las ecosondas científicas, esta señal, generada mediante un transductor, se propaga hacia el fondo en forma de onda esférica. Al llegar al fondo, la señal acústica lo insonifica con su frente de onda, primero verticalmente e, inmediatamente después, oblicuamente, debido a la esfericidad del frente de onda. Parte de esta energía acústica es reflejada por la capa más superficial y otra parte penetra levemente en el fondo (en función de la frecuencia e intensidad de la onda), para ser reflejada a su vez por las capas inmediatamente más profundas [3].

Dado que se conocen o se pueden modelar las características de la onda inicial, la geometría de la propagación de la onda (tanto en su avance por la columna de agua como en su insonificación de la superficie del fondo) y la atenuación a lo largo de la columna de agua, la adquisición de los datos acústicos reflejados en el fondo nos permite inferir características del mismo. En esta línea, se han desarrollado trabajos de clasificación de hábitats [4], cartografía de vegetación submarina [5], estudios de geomorfología submarina [6], cartografía de comunidades bentónicas [7], entre otros.

Esquemáticamente la onda acústica es generada en el transductor, viaja hacia el fondo donde es reflejada y asciende hacia la superficie, una vez allí sufre una nueva reflexión en la interfaz agua-aire que le hace volver a propagarse hacia el fondo, donde de nuevo sufre reflexión y asciende. Desde el momento en el que el transductor emite la onda, este comienza a adquirir información de la onda reflejada. Durante el avance de la onda a través de la columna de agua, la energía acústica reflejada es muy baja, debido a cambios en las características del medio, o a algún elemento (como peces) encontrado en la trayectoria. Cuando llega al fondo se produce una fuerte reflexión debida a la gran diferencia de impedancias acústicas entre el agua y el sedimento o la roca del fondo. Este proceso se repite con el segundo rebote. De este modo, el transductor recoge la intensidad acústica que recibe a lo largo del tiempo en forma de un ruido de fondo, de un nivel de intensidad muy bajo, hasta que recibe el primer eco, procedente del primer rebote del ping. Este eco comienza con un máximo correspondiente a la reflexión de la parte de la onda que se propaga verticalmente seguido de una "cola" que se debe a la retrodispersión oblicua de la onda esférica; la forma de esta cola depende del tipo de fondo en el que se refleje, y puede presentar diferentes máximos debidos a la interferencia de la onda, a la penetración en el fondo, etc. A ese primer eco le sigue otro segundo eco procedente de la segunda reflexión, con una forma similar aunque atenuada y suavizada tras los dos rebotes previos en el fondo y la superficie del agua.

Debido a la elevada velocidad de avance de la onda acústica en el agua la tasa de adquisición de datos con ecosondas es muy elevada. Por ejemplo, trabajando en aguas someras, a 50 m de profundidad, con pulsos de longitud media (1,024 $\mu$ s) y adquiriendo tanto el primer como el segundo rebote de la sonda en el fondo, las ecosondas científicas permiten la adquisición de unos 450 pings por segundo (aunque lo habitual son unos 10 pings por segundo). Esta elevada tasa de muestreo tiene como resultado bases de datos ingentes que requieren de la aplicación de análisis estadísticos para la extracción de información relevante.

Al comienzo de la aplicación de métodos acústicos para el estudio de los fondos marinos en la década de 1980 [8], se utilizaba el total de energía integrada dentro de cada una de estas curvas (primer y segundo eco) para realizar una clasificación no supervisada de cada uno de los puntos muestreados [9]. Con esta aproximación metodológica se obtuvieron buenos resultados [10], si bien al tener tanto nivel de simplificación (toda la curva caracterizada por dos únicas variables) limita la clasificación obtenida. Con el objetivo de capturar más información presente en las curvas acústicas surgieron posteriormente aproximaciones multivariantes que extraían variables (features) en base a diferentes tipos de análisis de la señal: análisis de Fourier, análisis fractal, integración energética, distribución de potencia a lo largo del eco, etc. Esta metodología sigue siendo a día de hoy la aproximación principal para el estudio acústico de fondo, bien a través de software comercial [11], bien con aplicaciones libres [12]. A pesar de que estas aproximaciones amplían drásticamente el número de variables desde el enfoque inicial (de 2 en RoxAnn, hasta a 166 QTC), al someter la señal a un proceso de caracterización mediante variables, no siempre totalmente independientes (funcional o estadísticamente), siempre se está desechando una parte de la información relativa a la forma de la curva del eco que podría responder a características relevantes del fondo. Por último, muy recientemente se han propuesto alternativas de clasificación, tanto supervisada como no supervisada, usando técnicas de Análisis de Datos Funcionales (FDA) y clasificación clúster de series de tiempo [13], línea de investigación sigue el presente trabajo. Este tipo de alternativas de clasificación estadística tienen la ventaja de automatizar el proceso de identificación de fondos y de minimizar la necesidad de conocimiento previo acerca del problema físico, incrementando la rapidez de la respuesta sin necesidad de personal entrenado.

## 2 Obtención de datos

Mediante una ecosonda de haz simple (EA400P Simrad) acoplada a un bote de 5m, se han obtenido curvas acústicas correspondientes a una frecuencia de 200kHz en un área controlada de Cabo de Palos (Murcia, España), a fecha del 19 de julio de 2014. Estas curvas se corresponden con tres clases diferentes de fondos atendiendo al sustrato: fondo arenoso (SP1), fondo arenoso con escasa vegetación (SP2) y fondo rocoso (SP3). Para eliminar el efecto de la profundidad, se ha procedido a aplicar correcciones de tiempo (alargamiento del eco con la profundidad), potencia (atenuación de la onda con la distancia) y longitud de ping (deformación del eco debido a ésta), esta última a través de la convolución mediante una función tipo núcleo que compensa la

relación entre el pulso-longitud y la profundidad [7]. Todas las transformaciones anteriormente descritas han sido aplicadas mediante el software ECOSONS [12]. Como resultado, por cada ping se obtiene una curva de intensidad acústica en función del tiempo, compuesta por dos ecos o picos, el primero más relacionado con la rugosidad del fondo, mientras que el segundo con su dureza [9]. Finalmente se han obtenido 1383 curvas acústicas, una por cada ping, con valores de intensidad acústica integrados en 710 tiempos diferentes. Por un lado, se recogen los valores de la variable  $Z$  o verdaderas clases de fondo marino, mientras que por otro se obtienen los valores de intensidad sonora medidos en decibelios (dB). Estos valores en función del tiempo forman las series de tiempo que se utilizan para estimar los distintos tipos de fondo marino.

Es importante destacar que, de las 1383 curvas acústicas obtenidas, se han tenido en cuenta para el análisis 678, 236 de la clase SP1 (25% de las totales de dicha clase) 160 del tipo SP2 y 282 de la clase SP3. Se procede de esta manera para conseguir un diseño balanceado y, por tanto, que los índices que miden el desempeño de la clasificación no se vean sesgados de forma espuria por el tamaño muestral de cada clase.

### 3 Clasificación no supervisada a partir de series de tiempo

En este estudio se proporciona una solución alternativa a la identificación del tipo de fondo marino en relación al substrato que los define. Por tanto, es un problema que se puede abordar desde la perspectiva de la clasificación estadística. Las técnicas de clasificación pueden dividirse en dos grupos: (a) Técnicas de clasificación supervisada, mediante las cuales se estima la clase de fondo  $Z$  (variable aleatoria cualitativa) a partir de un modelo estadístico, función de un vector de  $m$  características cuantitativas  $X = \{X_1, X_2, \dots, X_T\}$ , cuyos parámetros han sido estimados mediante una muestra de entrenamiento, conociéndose previamente el número y tipo de clases existentes. (b) Técnicas de clasificación no supervisada o clúster, que permiten agrupar individuos (en este caso fondos marinos) en un número variable de clases o grupos según diferentes criterios de similitud aplicados a los vectores de características  $X$  que definen a cada individuo. Las técnicas clúster se pueden aplicar sin conocer previamente ni el número ni las clases existentes ni su número. Por ello, teniendo en cuenta que, en los estudios de este ámbito, no se suele tener información previa acerca de qué clases de fondo existen en una región de terminada, de entre todas las técnicas de clasificación, en este trabajo se proponen herramientas de clasificación no supervisada.

En este caso particular, se propone la aplicación de metodologías de clasificación clúster de series de tiempo, dado que el vector de características  $X$  está compuesto por  $T$  observaciones de intensidad sonora en función del tiempo. Es importante destacar que, en este contexto, las diferencias entre individuos, representados por series de tiempo, pueden depender tanto de la proximidad de cada una de las observaciones en una rejilla de tiempos como del carácter dinámico de la serie (estructura de dependencia, correlaciones, etc.) [14]. Por ello, el primer paso del procedimiento es la definición de una medida adecuada de disimilitud que permita definir la distancia entre dos curvas o series de tiempo. Una vez elegida, se aplica a las series de tiempo obteniendo

como resultado la matriz de disimilitudes, de dimensión  $n \times T$ , donde  $n$  es el número de observaciones de fondo marino que se están estudiando y  $T$  el número de tiempos diferentes en los que se evalúa la intensidad sonora (puntos por los que está constituida cada serie de tiempo). Finalmente, se aplica un método de clasificación clúster [14], preferiblemente jerárquico [13], a la matriz de disimilitudes, obteniendo como resultado la asignación de un grupo a cada una de las observaciones de fondos.

En este trabajo se ha estudiado el desempeño de 6 disimilitudes diferentes pertenecientes a la tipología “model free” [14], dada su sencillez y rápido tiempo de ejecución. Para ilustrarlas, se han definido  $\mathbf{X}_T = (X_1, X_2, \dots, X_T)$  y  $\mathbf{Y}_T = (Y_1, Y_2, \dots, Y_T)$  como las realizaciones particulares de los procesos de valor real  $X = \{X_t, t \in \mathbb{Z}\}$  e  $Y = \{Y_t, t \in \mathbb{Z}\}$ , respectivamente, siguiendo las indicaciones de Montero y Vilar [14]. La más sencilla es la disimilitud euclídea,  $d_{L_2}$ , definida por

$$d_{L_2}(\mathbf{X}_T, \mathbf{Y}_T) = \left( \sum_{t=1}^T (X_t - Y_t)^2 \right)^{\frac{1}{2}} \quad (1)$$

que únicamente tiene en cuenta las distancias punto a punto entre las series de tiempo. Una alternativa a  $d_{L_2}$ , es la distancia denominada Dynamic Time Warping,  $d_{DTW}$  que en su definición tiene en cuenta tanto la distancia punto a punto como las diferencias de forma de cada par de series. De hecho, una de sus principales ventajas es que es capaz de reconocer formas similares en curvas incluso después de haber sido transformadas (desplazadas y escaladas) [14] como es el caso de la presente base de datos. Para su cálculo se define  $r = ((X_{a_1}, Y_{b_1}), \dots, (X_{a_m}, Y_{b_m}))$ , donde  $a_i, b_j \in \{1, \dots, T\}$  de forma que  $a_1 = b_1 = 1$ ,  $a_m = b_m = T$ , mientras que  $a_{i+1} = a_i$  o  $a_i + 1$  y  $b_{i+1} = b_i$  o  $b_i + 1$ , con  $i \in \{1, \dots, m-1\}$ , siendo  $m$  todos las posibles parejas de observaciones.

$$d_{DTW}(\mathbf{X}_T, \mathbf{Y}_T) = \min_r \left( \sum_{i=1,2,\dots,m} |X_{a_i} - Y_{b_i}| \right) \quad (2)$$

Estas dos distancias,  $d_{DTW}$  y  $d_{L_2}$  no tienen en cuenta la estructura de dependencia dentro de cada serie de tiempo, asumen que cada observación dentro de la serie es independiente de las demás. Una forma de medir esta dependencia es mediante el cálculo de distancias basadas en las funciones de autocorrelación (ACF) [14],

$$d_{ACF}(\mathbf{X}_T, \mathbf{Y}_T) = \sqrt{(\hat{\rho}_{X_T} - \hat{\rho}_{Y_T})^T \Omega (\hat{\rho}_{X_T} - \hat{\rho}_{Y_T})} \quad (3)$$

y en las funciones de autocorrelaciones parciales (PACF),  $d_{PACF}$ , cuya distancia se define de forma similar, siendo  $\Omega = I$  la matriz de pesos, y  $\hat{\rho}_{X_T} = (\hat{\rho}_{1,X_T}, \dots, \hat{\rho}_{L,X_T})$  y  $\hat{\rho}_{Y_T} = (\hat{\rho}_{1,Y_T}, \dots, \hat{\rho}_{L,Y_T})$  los coeficientes de autocorrelación de  $X_T$  e  $Y_T$ , con un retardo  $L$  para el que se obtiene  $\hat{\rho}_{j,X_T} = 0$  o  $\hat{\rho}_{j,Y_T} = 0$  con  $j > L$  [14].

Otra alternativa de modelo libre la proporciona el uso de la disimilitud basada en la diferencia entre los periodogramas (estimaciones de la densidad espectral de las series de tiempo) de cada par de series de tiempo [14]. Definiendo los periodogramas de  $\mathbf{X}_T$  e  $\mathbf{Y}_T$  como  $I_{X_T}(\lambda_k) = T^{-1} |X_t e^{-i\lambda_k t}|^2$  e  $I_{Y_T}(\lambda_k) = T^{-1} |Y_t e^{-i\lambda_k t}|^2$ , respectivamente, con  $\lambda_k = \frac{2\pi k}{T}$ ,  $k = 1, \dots, \frac{T-1}{2}$ , se puede definir a su vez la distancia euclídea entre dos periodogramas como

$$d_P(\mathbf{X}_T, \mathbf{Y}_T) = \frac{1}{T-1} \sqrt{\sum_{k=1}^{\frac{T-1}{2}} (I_{X_T}(\lambda_k) - I_{Y_T}(\lambda_k))^2} \quad (4)$$

mientras que su versión normalizada con transformación logarítmica tiene la expresión

$$d_{LNP}(\mathbf{X}_T, \mathbf{Y}_T) = \frac{1}{T-1} \sqrt{\sum_{k=1}^{\frac{T-1}{2}} (\log NI_{X_T}(\lambda_k) - \log NI_{Y_T}(\lambda_k))^2} \quad (5)$$

donde  $NI_{X_T}(\lambda_k) = \frac{I_{X_T}(\lambda_k)}{\hat{\gamma}_{0,X_T}}$ , siendo  $\hat{\gamma}_{0,X_T}$  la varianza muestral de  $\mathbf{X}_T$  (expresiones análogas para  $\mathbf{Y}_T$ ). La distancia  $d_{LNP}(\mathbf{X}_T, \mathbf{Y}_T)$  es útil cuando sólo interesa medir diferencias en la estructura de correlación, no en la escala.

Una vez calculadas las matrices de disimilitudes, se propone aplicar el algoritmo de clasificación clúster jerárquico de distancia máxima [13-15].

## 4 Resultados y discusión

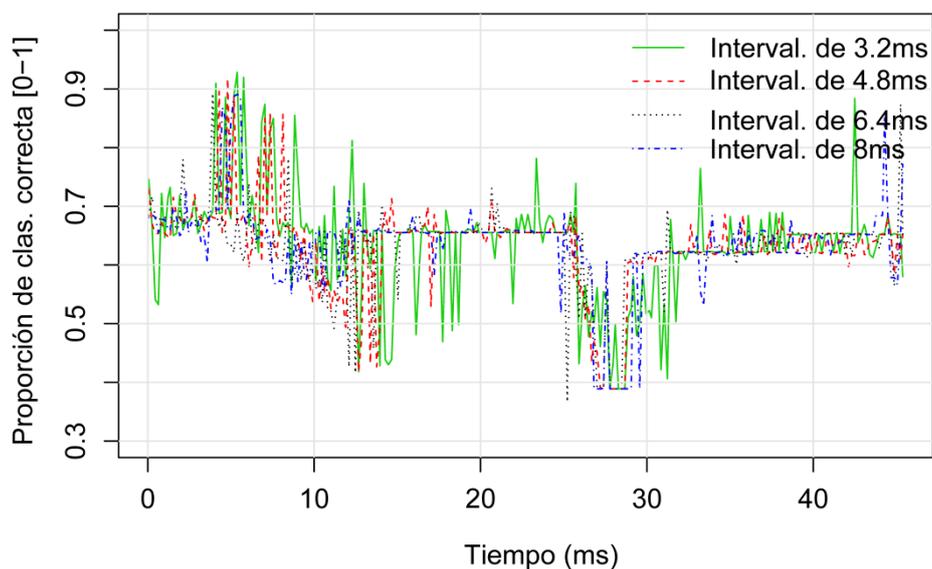
En la presente sección, con el fin de identificar el tipo de fondo marino según su sustrato, se aplica el algoritmo de clasificación jerárquico de distancia máxima a las matrices de disimilitudes obtenidas a partir del cálculo de las distancias descritas en la Sección 3. Para evaluar el desempeño del método de clasificación según la matriz de disimilitudes utilizada, se propone el cálculo de tres índices de uso común en este ámbito: proporción de clasificación correcta [16], Cluster Evaluation Index (CEI) [14] e índice kappa [17]. Una de las principales aportaciones de este estudio es la identificación de los intervalos de tiempo correspondiente a los datos acústicos que mejor definen a cada clase de fondo, es decir, en aquellos donde las diferencias entre curvas de la misma clase es mínima y la diferencia entre curvas acústicas de fondos marinos distintos es máxima. Por ello se evalúa la metodología de clasificación propuesta en segmentos de curva comprendidos en intervalos de tiempo de longitud (3.2, 4.8, 6.4 y 8.0ms) y tiempo inicial (de 0 a 40ms, con un paso de 1ms) variables.

La Figura 1 muestra la proporción de clasificación correcta obtenida, usando la distancia DTW, para cada intervalo de las curvas acústicas en función del tiempo de comienzo del intervalo. Las mejores clasificaciones se obtienen a partir de los intervalos situados en la cola del primer eco, particularmente en aquellos de longitud menor (3.2 o 4.8ms). En la Figura 2, se observa el intervalo de tiempo al que se obtiene la mayor proporción de clasificación correcta, el más informativo, representado sobre las curvas de intensidad acústica reales correspondientes a las tres clases de fondo marino. Las diferencias en el primer eco están relacionadas con la rugosidad del fondo marino.

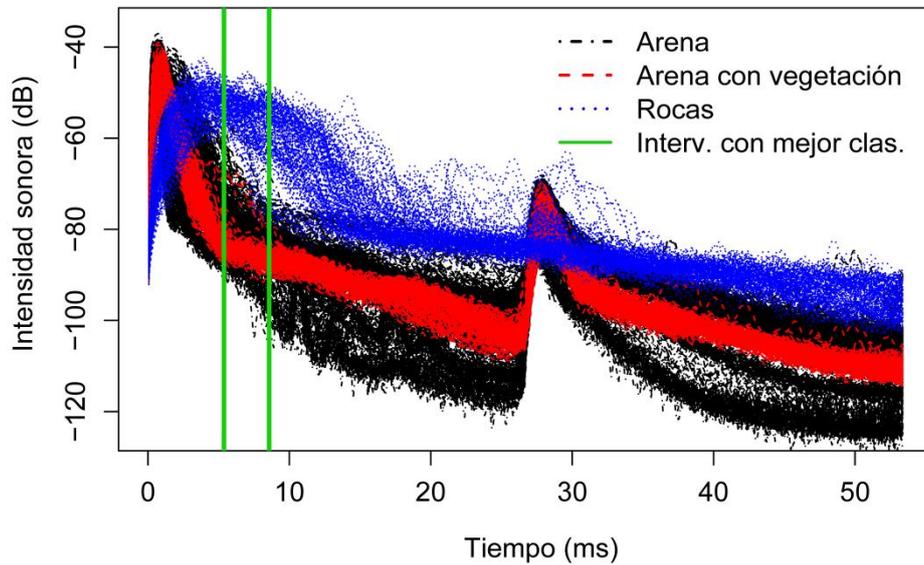
En la Tabla 1 se muestran clasificaciones casi perfectas aplicando el algoritmo clúster jerárquico de distancia máxima a las matrices de disimilitudes construidas a partir de la distancia DTW y euclídea. En efecto, mediante el uso de la distancia DTW, se obtiene una proporción de clasificación correcta del 93%, resultado coheren-

te con el correspondiente intervalo de confianza al 95% para el índice kappa de (0.86,0.92), dentro de los límites de la clasificación casi perfecta (definida para los valores comprendidos entre 0.81 y 1). El hecho de que sean precisamente estas disimilitudes las que proporcionen mejores resultados de clasificación, está relacionado con que las diferencias entre las curvas acústicas correspondientes a las distintas clases de fondo obedecen más a diferencias punto a punto y de forma. Esto demuestra que las correcciones aplicadas a los ecos han sido eficientes, ya que el efecto de la profundidad desplaza y dilata al eco dentro de la curva. Por otro lado, si se observan los resultados obtenidos usando las disimilitudes  $d_{LNP}$  (81% de clasificación correcta) y  $d_{PAFC}$  (84%), se verifica que los fondos marinos también se diferencian por la estructura de autocorrelación de sus correspondientes series de tiempo acústicas. Teniendo en cuenta estos resultados, el uso de otras métricas que combinen distancia punto a punto y estructura de dependencia, como son las disimilitudes COR y CORT, está justificado [13].

Por último, en la Tabla 2 se indican las matrices de confusión correspondientes al uso de cada tipo de disimilitud, en el intervalo para el que se obtiene la mayor proporción de clasificación correcta, en cada caso. En la diagonal de cada matriz de confusión se muestra la proporción de curvas a las que se le ha asignado la clase correcta, mientras que fuera de la diagonal se sitúan las proporciones de clasificación errónea. Mediante el uso de la disimilitud DTW o la distancia euclídea se clasifican correctamente todos los tipos de fondo (valores de la diagonal próximos a 1), incluso aquellos muy similares como son los etiquetados como SP1 y SP2. En cambio, mediante la disimilitud PAFC existe una pequeña confusión entre dichas clases, mientras que el uso de la disimilitud LNP lleva a confundir SP3 con SP1.



**Figura 1.** Proporciones de clasificación correcta para cada intervalo de tiempo estudiado, en función del tiempo de inicio de cada uno. Es el resultado de aplicar el algoritmo clúster jerárquico a la matriz de disimilitudes DTW correspondiente a 4 longitudes de intervalo diferentes.



**Figura 2.** Curvas de intensidad sonora en función del tiempo (1<sup>er</sup> y 2<sup>o</sup> eco) correspondientes a cada una de las 3 clases de fondo. Se destaca el intervalo de tiempo para el que se obtiene una mejor identificación aplicando el algoritmo jerárquico a la matriz de disimilitudes DWT.

**Tabla 1.** Índices que miden el desempeño de las distintas alternativas de clasificación clúster ensayadas: proporción de clasificación correcta, CEI e intervalo de confianza (95%) de kappa.

	Proporción de clasificación correcta	CEI	Kappa, intervalo de confianza (95%)	Intervalo de tiempo (ms)
$L_2$	0.92	0.92	(0.86,0.92)	(4.99,8.19)
DTW	0.93	0.93	(0.86,0.92)	(5.38,8.58)
ACF	0.65	0.66	(0.42,0.54)	(0.06,3.26)
PACF	0.85	0.84	(0.72,0.81)	(0.256, 8.26)
P	0.77	0.72	(0.59,0.69)	(25.22,33.22)
LNP	0.81	0.81	(0.65,0.75)	(4.67,12.67)

## 5 Conclusiones

Se ha propuesto la utilización de técnicas clúster para la clasificación de series de tiempo con el fin de identificar el tipo de fondo marino a partir de datos acústicos proporcionados por ecosondas. La metodología propuesta proporciona una alternativa para la identificación automática de tipos de fondos marinos, con la ventaja de no precisar de personal experto o un conocimiento profundo del contexto del problema.

Atendiendo a diversos índices que evalúan el desempeño de los métodos de clasificación utilizados (proporción de clasificación correcta, CEI y kappa), se han obtenido clasificaciones casi perfectas aplicando el algoritmo clúster jerárquico de distancia máxima a las matrices de disimilitudes construidas a partir de la distancia euclídea o

la DTW. De hecho, mediante el uso de la distancia DTW, se obtiene una proporción de clasificación correcta del 93% y un intervalo para el kappa de (0.86,0.92), es decir, se consigue una clasificación casi perfecta (correspondiente a un kappa entre 0.81 y 1). Estos resultados indican que las diferencias entre las series de tiempo correspondientes a las distintas clases de fondo se deben principalmente a las distancia punto a punto y a la forma que presentan. Aunque también hay que destacar que la estructura de dependencia de cada serie también permite distinguir los distintos tipos de fondos (en menor medida), como muestran los resultados obtenidos a partir de las disimilitudes basadas en la función de autocorrelaciones parciales (85% de clasificación correcta) y en el periodograma normalizado (81%).

Las mayores diferencias entre clases y, por tanto, las mejores clasificaciones, se han obtenido en intervalos de tiempo cortos correspondientes a la cola (distancia DWT y euclídea) del primer eco, relacionado con la rugosidad del fondo.

**Tabla 2.** Matrices de confusión resultado de aplicar el método clúster jerárquico a cada matriz de disimilitud (en los intervalos de la Tabla 1). Proporciones en tanto por uno.

	<b><math>L_2</math></b>			<b>DTW</b>		
	Estimadas			Estimadas		
Reales	SP1	SP2	SP3	SP1	SP2	SP3
SP1	0.84	0.14	0.02	0.84	0.15	0.01
SP2	0.04	0.96	0.00	0.03	0.97	0.00
SP3	0.00	0.00	1.00	0.00	0.00	1.00
	<b>ACF</b>			<b>PACF</b>		
	Estimadas			Estimadas		
Reales	SP1	SP2	SP3	SP1	SP2	SP3
SP1	0.61	0.39	0.00	0.79	0.18	0.03
SP2	0.11	0.89	0.00	0.24	0.76	0.00
SP3	0.44	0.05	0.51	0.01	0.00	0.99
	<b>P</b>			<b>LNP</b>		
	Estimadas			Estimadas		
Reales	SP1	SP2	SP3	SP1	SP2	SP3
SP1	0.91	0.03	0.06	0.89	0.10	0.01
SP2	0.68	0.30	0.02	0.16	0.84	0.00
SP3	0.01	0.00	0.99	0.30	0.03	0.67

## Agradecimientos

Este trabajo ha sido subvencionado la Xunta de Galicia a través del proyecto Simulación numérica de problemas hidroacústicos de alta frecuencia en medios marinos - SIMNUMAR (EM2013/052), a través de Centros Singulares de Investigación de Galicia (ED431G/01 2016-2019) y Grupos de Referencia Competitiva (ED431C2016-015), siendo cofinanciado por los European Regional Development Funds (ERDF).

## Referencias

1. Kenny, A., Cato, I., Desprez, M., Fader, G., Schüttenhelm, R., Side, J.: An overview of seabed-mapping technologies in the context of marine habitat classification. *ICES Journal of Marine Science: Journal du Conseil* 60, 411–418 (2003).
2. Lurton, X.: *An Introduction to Underwater Acoustics: Principles and Applications*. 2nd edn. Geophysical Sciences Subseries, XXXVI, Springer Praxis Books (2010) (724 pp.).
3. Sánchez-Carnero, N., Aceña, S., Rodríguez-Pérez, D., Couñago, E., Fraile, P., Freire, J.: Fast and low-cost method for VBES bathymetry generation in coastal areas. *Estuarine, Coastal and Shelf Science* 114, 175–182 (2012).
4. Snellen, M., Siemes, K., Simons, D.G.: Model-based sediment classification using single beam echosounder signals. *The Journal of the Acoustical Society of America* 129, 2878–2888 (2011).
5. Sánchez-Carnero, N., Rodríguez-Pérez, D., Couñago, E., Aceña, S., Freire, J.: Using vertical sidescan sonar as a tool for seagrass cartography. *Estuarine, Coastal and Shelf Science* 115, 334–344 (2012).
6. Rovere, A., Vacchi, M., Firpo, M., Carobene, L.: Underwater geomorphology of the rocky coastal tracts between Finale Ligure and Vado Ligure (western Liguria, NW Mediterranean Sea). *Quaternary International*, 232(1), 187–200 (2011).
7. Rodríguez-Pérez, D., Sánchez-Carnero, N., Freire, J.: A pulse-length correction to improve energy-based seabed classification in coastal areas. *Continental Shelf Research* 77, 1–13 (2014).
8. Orłowski, A.: Application of multiple echoes energy measurements for evaluation of sea-bottom type. *Oceanologia* 19, 61–78 (1984).
9. Voulgaris, G., Collins, M.B.: *USP RoxAnn Ground Discrimination System: A Preliminary Evaluation*. Technical Report, University of Southampton (1999).
10. Henriques, V., Guerra, M.T., Mendes, B., Gaudêncio, M.J., Fonseca, P.: Benthic habitat mapping in a Portuguese marine protected area using EUNIS: an integrated approach. *Journal of Sea Research* 100, 77–90 (2015).
11. Somerton, D.A., McConnaughey, R.A., Intelmann, S.S.: Evaluating the use of acoustic bottom typing to inform models of bottom trawl sampling efficiency. *Fisheries Research* 185, 14–16 (2017).
12. Rodríguez-Pérez, D., Sánchez-Carnero, N., Freire, J.: ECOSONS software. Available at <http://www.kartenn.es>, last accessed 2017/07/27 (2014).
13. Tarrío-Saavedra, J., Sánchez-Carnero, N., Prieto, A.: Supervised and unsupervised seabed classification using Functional Data Analysis and Time Series clustering from acoustic curves. Submitted.
14. Montero, P., Vilar, J.A.: Tslust: An R package for time series clustering. *Journal of Statistical Software* 62, 1–43 (2014).
15. Everitt, B. S., Landau, S., Leese, M., Stahl, D.: *Hierarchical clustering*. In: *Cluster Analysis*, 5th edn., pp. 71–110. Wiley (2011).
16. Francisco-Fernández, M., Tarrío-Saavedra, J., Naya, S., López-Beceiro, J., Artiaga, R.: Statistical classification of early and late wood through the growth rings using thermogravimetric analysis. *Journal of Thermal Analysis and Calorimetry*, 127, 499–506 (2017).
17. Cook, R.J.: 2005. Kappa. In: *Encyclopedia of Biostatistics*, 4. John Wiley and Sons (2005).