

***ESTADO DEL CONOCIMIENTO ECOSISTÉMICO,  
SANITARIO, AMBIENTAL Y OCEANOGRÁFICO  
ASOCIADO A LAS POBLACIONES DE CETÁCEOS DE  
LA MACARONESIA***



**Interreg**

Fondo Europeo de Desarrollo Regional



# ***CAPÍTULO IV***

## ***MONITORIZACIÓN ACÚSTICA AUTOMÁTICA***

***Informe realizado por***

***Dr. Fernando L. Rosa González, Dr. Jonas Phillip Lüke, Ing. Angélica Ravelo,***

***Ing. Iván Rodríguez Méndez***



**Departamento de  
Ingeniería Industrial**  
Universidad de La Laguna

## Tabla de contenidos:

|  |    |
|--|----|
| 1. Acústica submarina .....                        | 2  |
| 1.1. Bases de la acústica submarina .....          | 2  |
| 1.2. Bioacústica marina .....                      | 3  |
| 1.3. Registro acústico pasivo .....                | 5  |
| 1.4. Bioacústica Física .....                      | 6  |
| 1.5. Adquisición y detección .....                 | 9  |
| 1.5.1. Nodos inteligentes .....                    | 9  |
| 1.5.2. Almacenamiento .....                        | 11 |
| 1.5.3. Detección .....                             | 12 |
| 1.6. Localización y clasificación .....            | 17 |
| 1.6.1. Localización con TDOA por correlación ..... | 17 |
| 1.6.2. Clasificación manual de eventos .....       | 18 |
| 2. Monitorización acústica automática .....        | 19 |
| 2.1. Detección automática de especies .....        | 21 |
| 3. Bibliografía .....                              | 24 |

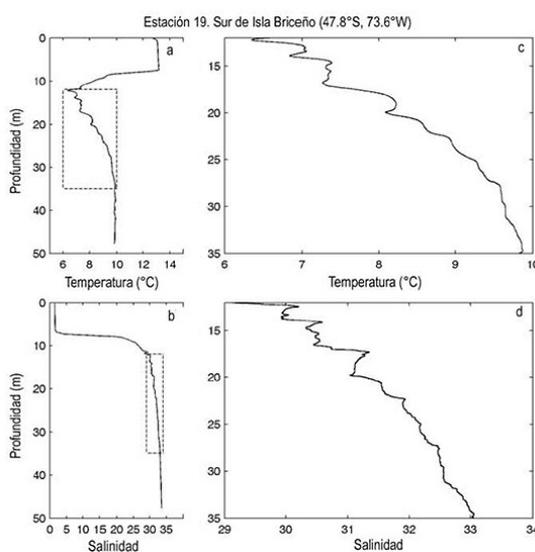
## 1. Acústica submarina

### 1.1 Bases de la acústica submarina

Es bien conocido que el sonido es una onda de presión que se propaga a una velocidad que depende de las propiedades del medio. La ecuación que gobierna esta propagación se puede expresar como en la ecuación 1.1, en la que el Laplaciano de la presión es proporcional a la aceleración en el cambio de la presión en el tiempo. La constante de proporcionalidad está precisamente relacionada con la velocidad de propagación al cuadrado.

$$\nabla^2 p = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} \quad (1.1)$$

Como esta relación debe cumplirse en cada punto del volumen de agua donde se encuentra la onda y en cada uno de esos puntos las condiciones materiales pueden cambiar, también lo puede hacer la velocidad de propagación  $c$  lo que hace que la ecuación deje de ser lineal en las variables espaciales. La dependencia de la velocidad de propagación con la temperatura, salinidad y presión fue establecida por Mackenzie en 1981 [1] y como se puede ver en la ecuación 1.2 depende no linealmente de la temperatura expresada en grados centígrados, la salinidad en partes por mil y la profundidad en metros. También es importante saber que esta expresión tiene un límite de aplicación para la temperatura entre 0 y 30 grados, para la profundidad entre 0 y 8000 metros y para la salinidad entre 30 y 40 partes por mil. La velocidad, o más correctamente su módulo, la celeridad, la obtendremos en metros por segundo.



**Figura 1.1:** Perfiles de la estación 19 al Sur de la Isla Briceño 47.8S, 73.6W.

$$c(d, s, T) = A_0 + A_1T + A_2T^2 + A_3T^3 + A_4(s-35) + A_5d + A_6d^2 + A_7T(s-35) - A_8Td^3 \quad (1.2)$$

Los perfiles de temperatura y salinidad son por lo tanto elementos que condicionan como se propaga el sonido debajo del mar. Si observamos la figura 1.1 se puede ver que tanto la temperatura como la salinidad tienen variaciones importantes lo que hace que no sea fácil predecir como se va a propagar el sonido.

Uno de los efectos del aumento gradual de la celeridad sonora es que provoca un cambio de dirección en la propagación generando el efecto de un canal de propagación. En la figura 1.2 se puede ver un esquema de la idea de canal sonoro que se genera cuando tenemos un valor mínimo de velocidad a cierta profundidad, un kilómetro en la figura, y si ascendemos en la columna de agua la velocidad aumentará, también lo hará si descendemos haciendo que la propagación vaya girando gradualmente para volver hacia el eje donde la velocidad es mínima. De este modo la atenuación del sonido se disminuye y se aumenta el rango donde se alcanza a recibir el sonido.

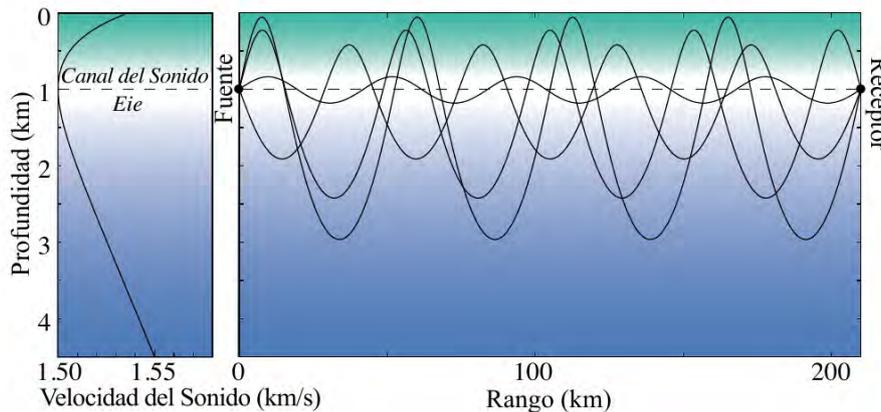
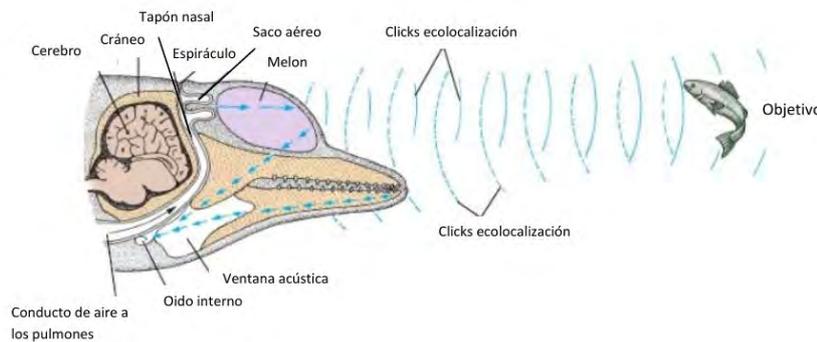


Figura 1.2: Esquema de propagación en un canal sonoro con un perfil de velocidad que genera un eje.

## 1.2 Bioacústica marina

La bioacústica puede ser definida de modo sencillo como el uso de la acústica para estudiar cualquier asunto relativo a los animales marinos, tales como la capacidad auditiva, la producción de sonidos, la comunicación, el comportamiento alimentario o cualquier otro de su historia natural [2]. Los animales producen sonido por medio de distintos mecanismos. En el caso de los cetáceos su aparato fonador está compuesto por sacos aéreos que pasan a través de válvulas como los labios de

mono que producen la onda que se propagará a través de su melón para ser proyectado hacia el exterior. En la figura 1.3 se puede ver un esquema de como un delfínido produce sonido en forma de clicks de ecolocalización. Estos sonidos impulsivos se propagan a 1500 metros por segundo reflejándose en los objetivos y volviendo hacia el animal que escuchará estos ecos de modo espacial reconociendo donde se encuentra su objetivo. Para más información sobre la morfología nasal y su papel en la generación de sonido consultar [3].

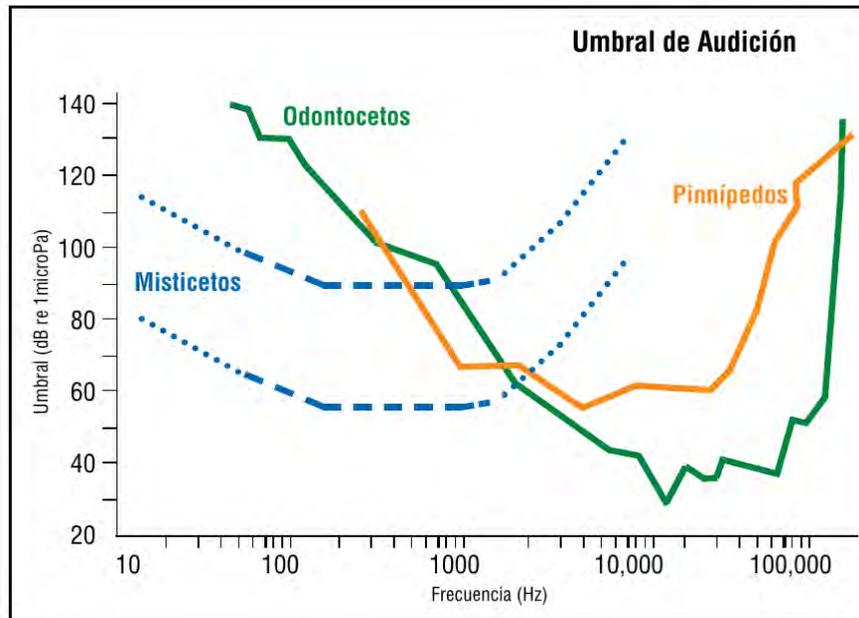


**Figura 1.3:** Esquema de propagación en un canal sonoro con un perfil de velocidad que genera un eje.

Uno de los principales problemas que se está detectando es el aumento de embarcaciones deportivas y en general tráfico marítimo que introduce energía sonora en el medio marino. Se sospecha que esto puede ser la causa del descenso de algunas poblaciones pero más directamente en la pérdida de sensibilidad para la audición. Comúnmente hablamos de sordera en el caso humano, sin embargo la sensibilidad para escuchar puede reducirse en diferentes bandas del espectro acústico. En la figura 1.4 se muestran las curvas de los umbrales de audición en condiciones normales de diferentes especies de animales marinos.

En el caso de los Mysticetos, al ser un gran número de especies se tiene una gran varianza de modo que se muestran dos perfiles entre los que encontraremos el umbral de audición de este parvorden del infraorden de los cetáceos. Es de destacar en los odontocetos que prácticamente tienen sensibilidad para escuchar desde 300 Hertzios hasta 100 Kilohertzios. También los pinnípedos son sensibles a las llamados ultrasonidos por exceder la capacidad auditiva del ser humano que está cifrada en 20KHz. Los estudios llevados a cabo sobre delfines no permiten saber con seguridad las causas de los varamientos pero ya que los odontocetos confían en su oído para ecolocalizar,

orientarse y alimentarse cualquier defecto de esta escucha podría ser la causa de este tipo de sucesos. A través de la medición de los potenciales evocados se puede saber si los animales tienen alguna deficiencia en oír. Los rangos de reducción severa están entre 70 y 90 decibelios y por encima de 90 decibelios para la pérdida profunda del oído en humanos [4].



**Figura 1.4:** Umbrales de audición sanos de mistocetos, odontocetos y pinnípedos.

### 1.3 Registro acústico pasivo

Son muchos los sistemas que se han desplegado en el mar para medir sonido. Específicamente para medir sonido de cetáceos en 2007 Philpott investigó la posibilidad de utilizar un detector pasivo de delfines para monitorizar delfines Tursiops en el estuario Shannon en Irlanda. Las mediciones se realizaron en el año 2005 y les permitieron correlacionar los registros con observaciones visuales realizadas hasta a 8 kilómetros del observador [5]. En la figura 1.5 se ve el aspecto del contenedor de la electrónica que será sumergido en el fondo marino con un fondeo simple y que podrá ser recuperado tirando de la cuerda que llega a la boya de superficie.

Más adelante apareció una tecnología más avanzada denominada C-Pod que venía acompañada de un software propietario denominado cpod.exe. La presencia de filtros para detectar en las señales adquiridas así como otros procesadores no estaban disponibles para su análisis de modo que se realizaron paralelos con adquirentes denominados trampas sonoras que se

analizaban posteriormente por medio de un software abierto denominado PAMGUARD. La comparación se realizó en 2016 para comprobar que ambos sistemas tenían muy baja correlación, seguramente por tener dos diferentes aproximaciones al problema de detectar especies [6].



**Figura 1.5:** Sistemas de boyas para monitoreo con acústica pasiva. a) T-Pod b) PAMBuoy.

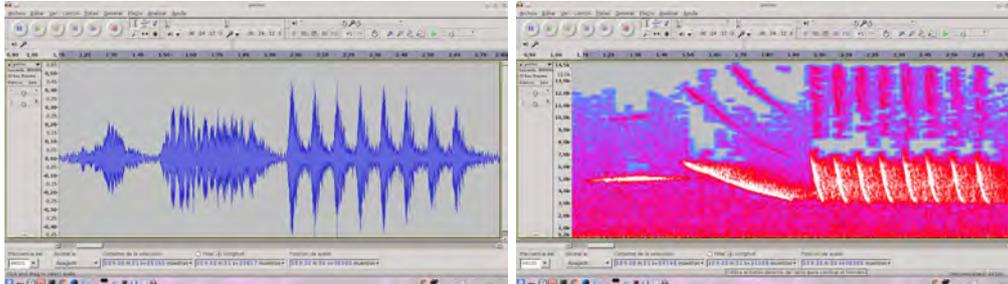
Un sistema de detección basado en PAMBuoy se desplegó en agosto de 2012 para el seguimiento pasivo de Belugas [7]. Este hardware llevaba el software PAMGUARD pero necesitaba todavía la automatización de muchos de los procesos para la detección de los distintos eventos. Para ello se utilizaron revisiones manuales realizadas por humanos.

Recientemente se han realizado pruebas con instrumentos no guiados por humanos como se puede leer en la revisión realizada por Ursula Verfuss y sus colaboradores [8].

#### 1.4 Bioacústica Física

Las características físicas del sonido tanto para su producción como para su propagación y medida, son ambitos que están profundamente relacionados con las ciencias físicas. El sonido que se adquiere en un ordenador con un micrófono cuando canta un pájaro transduce la presión sonora en voltaje que es digitalizado y almacenado en un registro binario de la computadora. Esta conversión se debe hacer a un ritmo lo suficientemente alto como para poder observar los cambios que se producen en la presión por el paso de la onda sonora. Si las frecuencias que se quieren medir son del orden del habla humana, limitada en 20 Kilohertzios el ritmo de muestreo regular deberá ser mayor que el doble 40KHz, así se suele tomar como frecuencia de alta fidelidad para las señales que usamos los humanos una frecuencia de muestreo de 44100 muestras por segundos.

El análisis de Fourier de una señal permite definir un cierto espacio tiempo-frecuencia que se denomina espectrograma y que de algún modo nos permite ver como evoluciona en el tiempo el contenido de frecuencias de la fuente sonora.



**Figura 1.6:** a) Onda de presión de un gorrión cantando. b) Espectrograma de la presión.

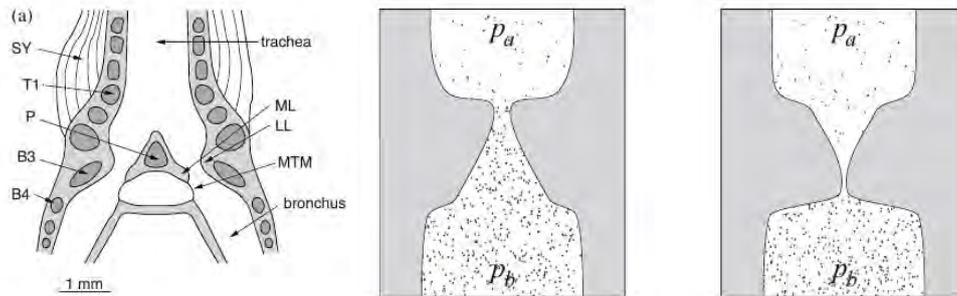
Supongamos la señal del canto de un gorrión como la que se muestra en la figura 1.6.a. En la misma figura 1.6.b se muestra el espectrograma y se puede ver como la señal comienza con un tono de 5000 Hz durante 40 centésimas de segundo para luego hacer un descenso desde 6KHz a 4KHz en 0.4 segundos comenzando una serie rápida de ocho descensos desde 7KHz a 3.5KHz durando cada uno apenas 0.05 segundos. Esta rápida cadencia de sonidos no pueden ser gestionadas por el sistema nervioso para controlar la vibración en si misma contrayendo y relajando un músculo para guiar esas rápidas frecuencias. El único modo de que esto sea posible es que haya un mecanismo físico que oscile a esas frecuencias y que la musculatura pueda cambiar ciertos parámetros que permitan controlar la producción sonora.

Para tratar de explicar este mecanismo Gabriel Mindlin propone como sistema de producción sonora la ecuación de Van der Pool [9]. Esta ecuación se muestra en la expresión 1.3 en la que se ve una relación entre como se mueve una membrana ante la presión subglotal representada por  $p$ . Deben tenerse en cuenta parámetros como la amortiguación del movimiento de la membrana por rozamiento que se representa por medio de la letra  $\beta$  y que se opone a la presión subglotal que actúa como escitadora del movimiento. El término  $k$  es la constante de recuperación lineal del resorte, en nuestro caso de la membrana y tiene que ver con su longitud y su elasticidad y finalmente será la responsable de la frecuencia a la que oscile el sistema. Por último, se ha añadido un término no lineal que añade limitación a la oscilación  $c$ . El término  $f_0$  se denomina de apertura y se utiliza para modelar la apertura del tracto vocal para respirar.

$$\dot{x} = y$$

(1.3)

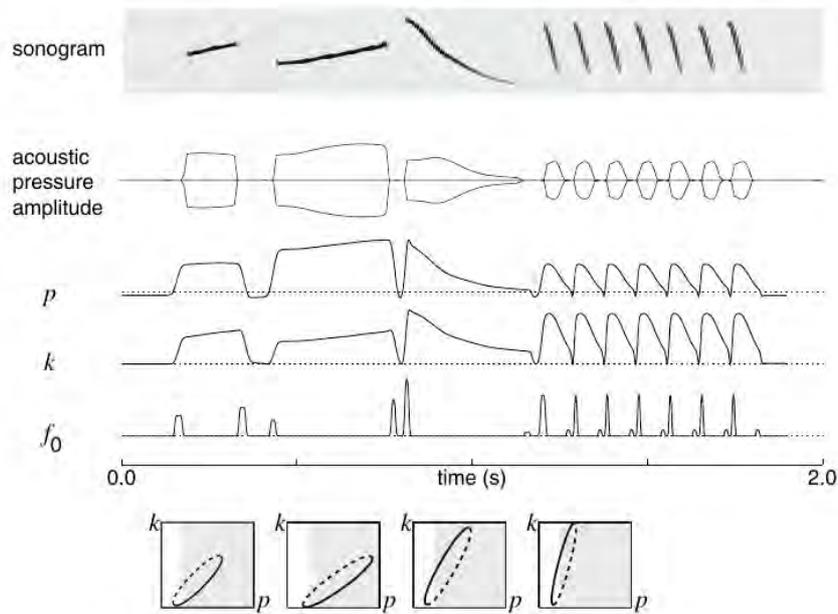
$$\dot{y} = -kx + (\beta - p)y - cx^2y - f_0$$



**Figura 1.7:** a) Siringe de un pájaro. b) Esquema de generación de la onda sonora.

En la figura 1.7 se muestra el esquema de una siringe donde se observa un doble aparato fonador y como la oscilación produce la onda sonora que seguirá hacia el pico del pájaro. El siguiente paso en este planteamiento es obtener los parámetros que permitan obtener la onda que se ha medido.

La inversión de los parámetros la realiza Mindlin y en una figura de su libro que se muestra en la figura 1.8. Como se puede ver el espectrograma, también llamado sonograma se encuentra en la parte alta como resultado de analizar la señal etiquetada como “acoustic pressure amplitude” que es el resultado de integrar la ecuación diferencial con la secuencia de parámetros que se encuentra en las gráficas de  $p$ ,  $k$  y  $f_0$ . Esto permite entender que la tensión muscular se mantiene apretando los pulmones para generar una presión cuando se necesita pero a un ritmo mucho más lento que el de la oscilación que se produce en la membrana productora. Esto también ocurre tensionando los tejidos cambiando  $k$ , y por último otros músculos abren y cierran el conducto para romper o permitir la oscilación, además de permitir respirar al animal. En la parte baja de la figura se observan los gráficos de evolución de los parámetros conductores del canto en gráficas  $k-p$  que permiten ver la sincronía desfasada entre los distintos parámetros para conseguir ese canto tan estereotipado del gorrión.



**Figura 1.8:** de arriba a abajo a) Espectrograma de la onda del canto de gorrión. Siringe de un pájaro. b) Amplitud de la onda de presión. c) Valor del parámetro presión en función del tiempo. d) Valor del parámetro de recuperación elástica en función del tiempo. e) Valor del parámetro de apertura en función del tiempo. f) Ciclos  $k$ - $p$  de las distintas partes del canto.

## 1.5 Adquisición y detección

### 1.5.1 Nodos inteligentes

La necesidad en el año 2005 de realizar instrumentación específica para adecuarse a la adquisición de datos a partir de hidrófonos instalados en las piscinas de OrcaOcean en Loro Parque en el Puerto de la Cruz en Tenerife. Implicó la instalación de un sistema de adquisición multicanal que pudiese obtener datos a frecuencias de muestreo de 200000 muestras por segundo y que nos permitiera recoger sin distorsión las señales en el rango auditivo de las orcas. Otra de las razones de este muestreo tan alto para los estándares de audio humano, es que la resolución temporal tiene mucha influencia en la capacidad de extraer los tiempos de diferencia de llegada a los distintos hidrófonos que finalmente influye en la posibilidad de localizar la fuente de sonido. Inicialmente la instalación se realizó de modo analógico y se digitalizaba en un ordenador central dotado de una tarjeta multicanal síncrona. Debido al tamaño de la instalación, que implica cables de mas de 100 metros de longitud para la transmisión de las señales se optó por el sistema actual de nodos inteligentes lo más próximos a los hidrófonos.

En la figura 1.9 se puede ver la presencia de un nodo Maestro que recibe la información de los ocho nodos esclavos. El nodo seleccionado que se puede observar en color verde está adquiriendo a 200000 muestras por segundo y tiene activado el filtro pasabajos a 97KHz. Cada nodo tiene un sistema operativo que permite comandar instrucciones por un operario que accediese al sistema. Así por ejemplo, se puede establecer un ritmo de medidas, tomar datos un minuto cada cinco minutos o se puede cambiar las frecuencias de muestreo. También se puede observar que el sistema se puede acceder desde Internet de modo que la monitorización del sistema se puede realizar en remoto.

La electrónica que se utilizó originalmente para la digitalización de las señales es la que se muestra en la figura 1.10. En ella se muestra los diseños originales de los sistemas de control de carga de baterías aislantes, del sistema de filtros y la placa Olimex que se utilizó con un procesador ARM de las primeras generaciones. Así mismo se muestra la caja prototipo inicial que se usó para la comprobación de todo el sistema y que resultó exitosa.

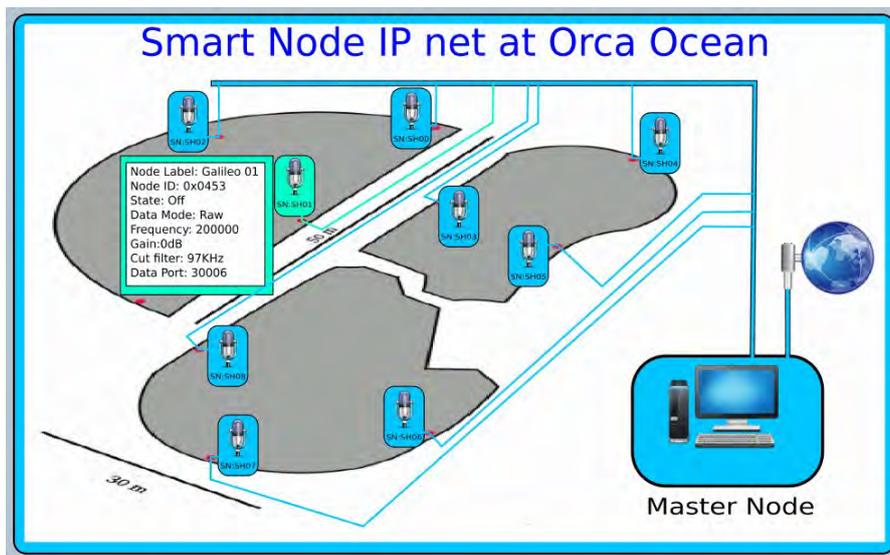
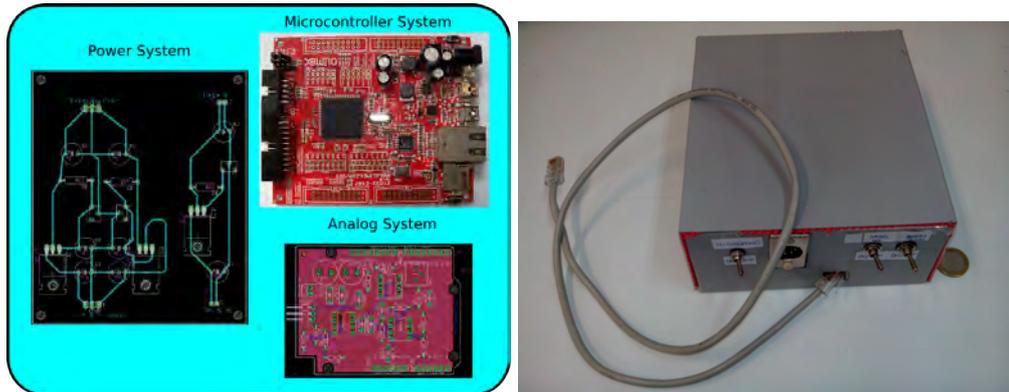


Figura 1.9: Esquema de conexionado IP de OrcaNET.



**Figura 1.10:** a) Caja de comprobación inicial del nodo inteligente de OrcaNET. b) Placa de desarrollo con microcontrolador de Olimex y diseños propios de control de carga y de filtrado analógico

### 1.5.2 Almacenamiento

A partir de ahora denominaremos orcaNode a los esclavos de la red y orcaMaster al maestro de la red. En esta red inteligente de nodos esclavos el Maestro sirve de repositorio de datos. Los nodos son configurados a través del orcaMaster y le envían los datos adquiridos en modo Raw. En la figura 1.11 se puede ver como los datos son almacenados en el sistema utilizando el sistema de ficheros para guardar los datos con su fecha sin tener que alargar los nombres de los ficheros en exceso. Como la adquisición se realiza de modo continuo los ficheros toman nombres de los fragmentos de cinco minutos en los que se puede dividir el día. Así, por ejemplo, en la figura se ve bajo el icono de Raw Data un árbol de ficheros que se encuentran en un directorio llamado “/media/Elements/.data” donde destacan dos elementos. Los datos están en un disco externo al ordenador que se puede cambiar cuando se llene, y se utiliza el formato “.data” que es el nombre de un directorio oculto.

Dentro del directorio de datos Raw se tiene un directorio para cada año que comienza con la letra “A”, A2015. Dentro de este directorio estarán los meses precedidos con la letra “M”, M08. Dentro de cada mes estarán los directorios de los días con su número precedido por la letra “D”, D03. Para la automatización se han normalizado el número de caracteres que se utilizan en cada uno de estos números, cuatro para el año, dos para los meses y dos para los días. Dentro encontraremos para cada nodo, dos en la figura, un directorio con el nombre del nodo, y por último los ficheros que tienen

la siguiente estructura "F131S000E299.raw", donde las letras F, S y E, significan, fragmento de cinco minutos del día, comienzo en segundos dentro del fragmento y segundo final dentro del fragmento respectivamente.

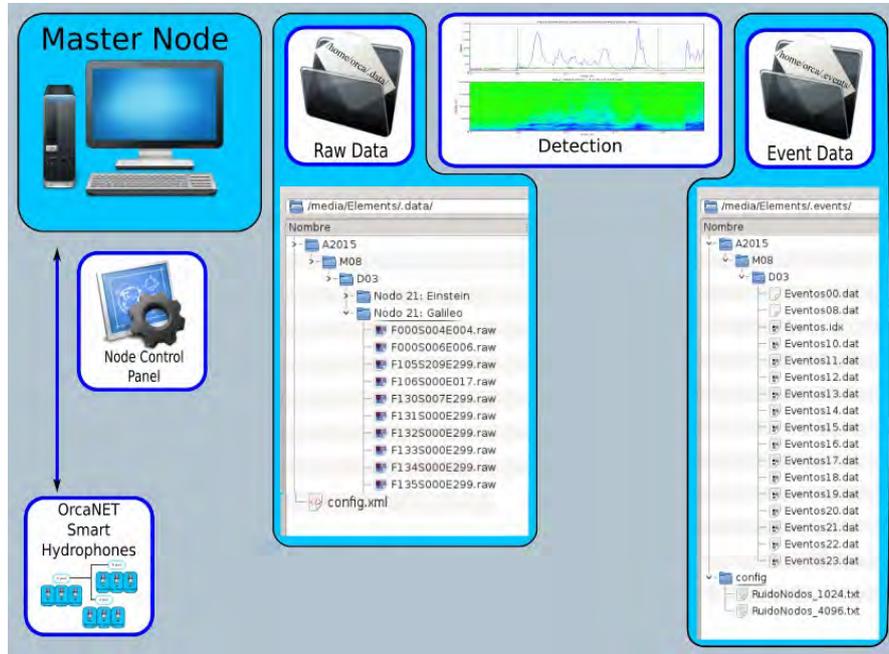


Figura 1.11: Esquema de almacenamiento y detección de eventos.

### 1.5.3 Detección

Una vez que los datos están almacenados en un fragmento de cinco minutos se realiza el procesamiento para detectar los eventos sonoros que se hayan producido. Gran parte del trabajo en este aspecto se centra en encontrar aquellos aspectos de la señal que son extraordinarios en el comportamiento estadístico de la señal. Las señales eléctricas que se han incorporado a la señal que provienen de ruido eléctrico, de acoplamientos de los transformadores de potencia u otras fuentes similares deben poder eliminarse para detectar adecuadamente los eventos sonoros de nuestro interés.

El proceso utiliza el procedimiento esquematizado en la figura 1.12. La adquisición de datos de un nodo termina en el fichero de datos Raw de cinco minutos. Sobre los datos muestreados a 200000 muestras por segundo se desplaza una ventana de análisis de 4098 muestras a la que se le calcula el espectro de Fourier discreto mediante una transformada rápida de Fourier utilizando una ventana de apodizado. Durante el proceso de calibración de ruido del nodo en el que se estima la

media y la desviación en cada una de las bandas espectrales promediando en el tiempo que se elija para la estimación del ruido. Este tiempo puede variar pero al tratarse de estimaciones estadísticas se recomienda utilizar un periodo de cinco minutos en el que el ambiente sonoro se mantenga sin eventos como puertas, golpes o llamadas de animales. De este modo se obtiene una máscara de variabilidad del ruido en cada banda de frecuencia de modo que sólo se utilizarán aquellas medias de ruido cuya variabilidad sea menor que un umbral. Con este sistema se puede evitar que el ruido eléctrico de la adquisición que se amplifica por el proceso de análisis espectral no tome parte en la evaluación de los niveles de disparo de presencia de eventos. En la esquema de la figura se utiliza la expresión  $NSTD$  para la sigma del ruido en cada banda de frecuencia, y  $tsig$  para el umbral de rechazo de las varianzas altas. Una vez que sabemos que bandas de frecuencia usar realizamos una distancia cuadrática media entre la potencia de ruido y la potencia en la señal medida en cada banda, integrando en la variable frecuencia. A este indicador lo denominamos **distancia media** ya que se ha calculado la distancia media de la potencia de la señal a la potencia del ruido en las bandas de frecuencia estables.

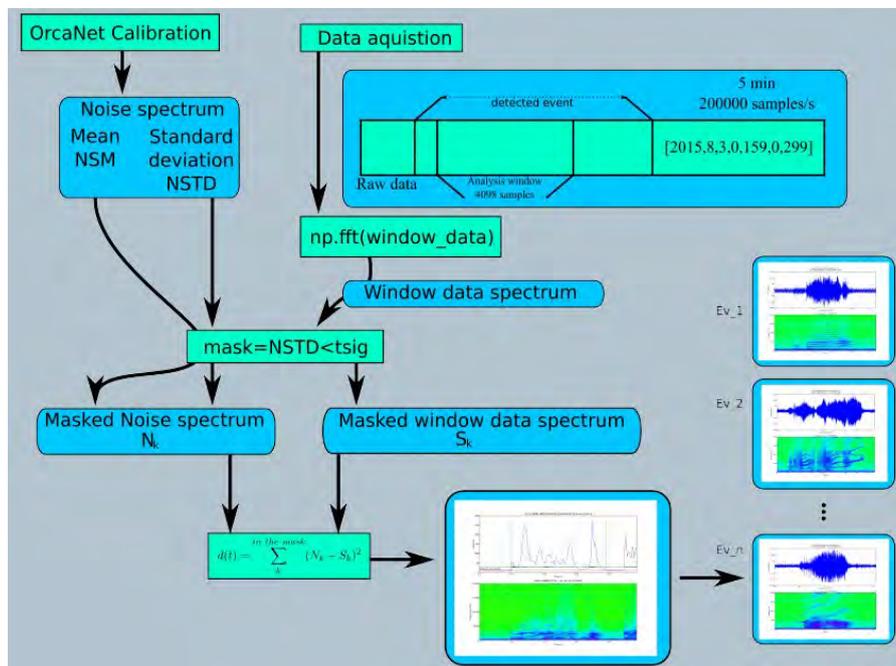
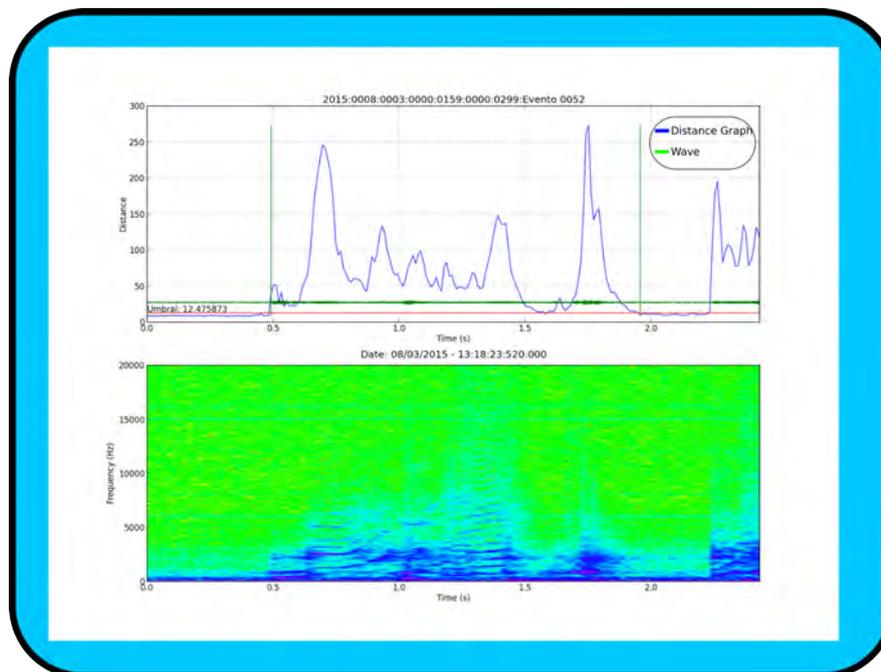


Figura 1.12: Esquema de almacenamiento y detección de eventos.

En la figura 1.13 se muestra en la gráfica inferior el espectrograma correspondiente al evento sonoro del canto de una orca el 3 de agosto de 2015 a las 13:15 horas, etiquetado como el evento número 52. En la parte superior en azul se muestra la evolución temporal de la distancia media en verde se muestra la onda sonora, en malva se ha pintado una línea horizontal que se utiliza como umbral de disparo en el valor umbral: 12.455873, y dos líneas verticales que son las líneas de corte del evento detectado. Se puede ver que la distancia media no se comporta de un modo estable durante una llamada de modo que se tienen que abordar algoritmos que prevengan la separación de un evento en varios o la unión de varios eventos en uno.



**Figura 1.13:** Detalle de un ejemplo de detección.

El resultado de detectar los eventos sonoros es una lista de eventos que deben ser almacenados en una base de datos de fácil acceso. El sistema elegido es el mismo que para los datos en bruto y se muestra en el lado derecho de la figura 1.12 donde en el directorio oculto ".events" se establece el árbol de fecha y en el interior del día se guarda un fichero índice con los eventos detectados en la red de hidrófonos en un fichero denominado "Eventos.idx". Los datos de cada evento son almacenados en ficheros llamados "EventosXX.dat" donde XX se sustituye por la hora del evento con dos dígitos. También se puede ver en la figura los ficheros donde se encuentra los datos del ruido resultante del proceso de calibración en ficheros específicos para el tamaño de la ventana

utilizado, en la figura aparecen los ficheros para los tamaños de ventana de 1024 y 4096, respectivamente. Unos ejemplos de eventos que corresponden a cantos de Orca de las piscinas en OrcaOCEAN se muestran en la figura 1.14.

El carácter estereotipado de las llamadas de Orca que se observa en la figura 1.14 fue documentada en las costas de la Colombia Británica por Ford y Fisher en el año 1986 [10]. Un extracto de la tabla de llamadas de varios de los *Pods* (grupos sociales estables), se muestra en la figura 1.15.

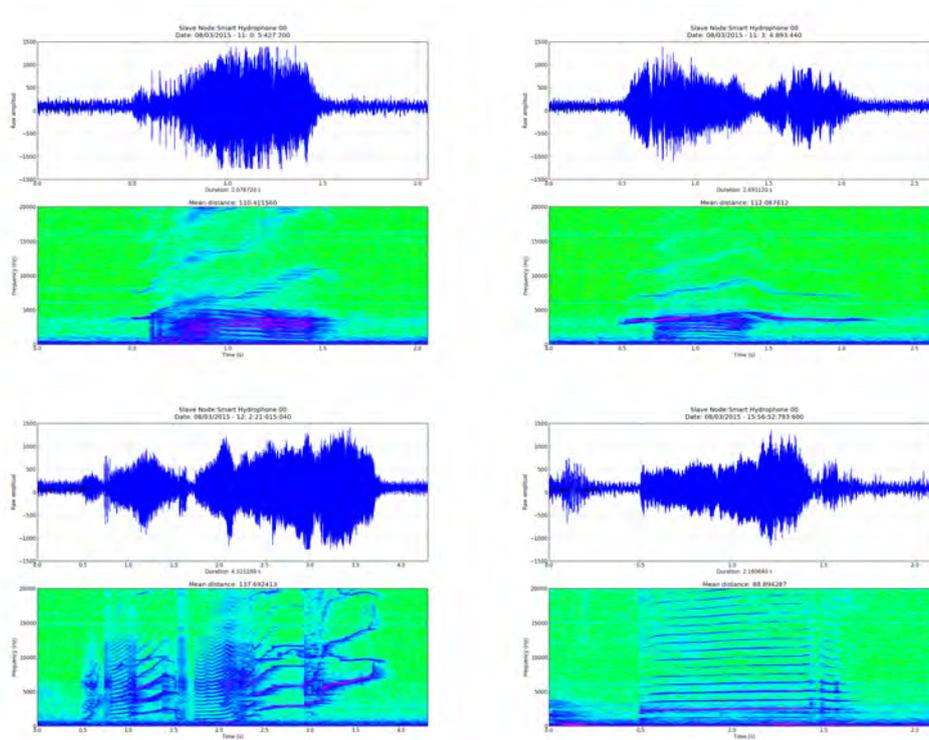


Figura 1.14: Cuatro eventos del día 3 de agosto de 2015.

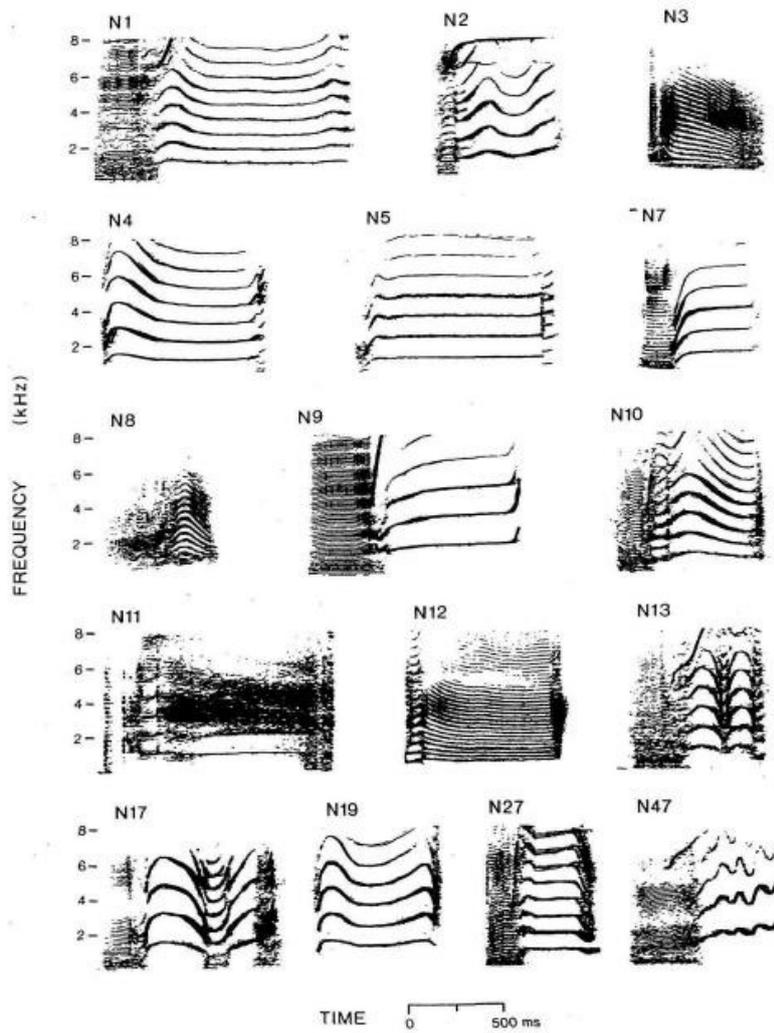


Figura 1.15: Llamadas estereotipadas de orcas de Vancouver.

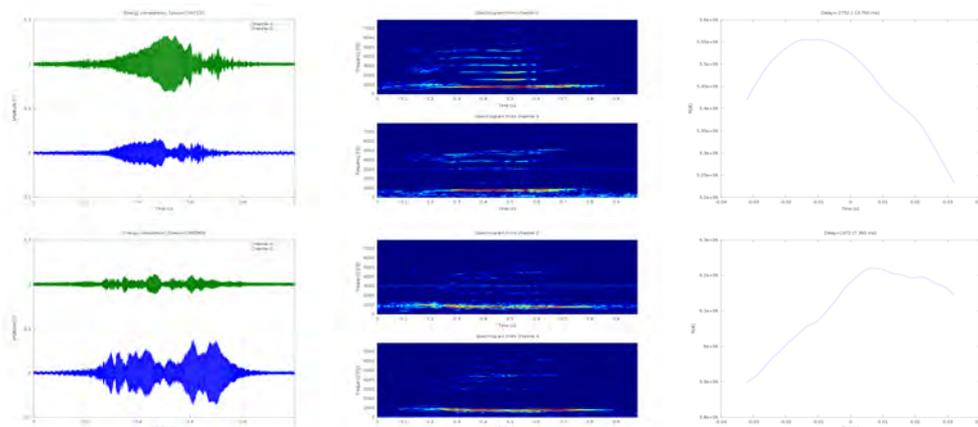


Figura 1.16: Dos ondas de dos sensores, su espectrograma y su correlación.

Monitorización Acústica Automática

## 1.6 Localización y clasificación

### 1.6.1 Localización con TDOA por correlación

Utilizando más de un receptor de sonido se puede registrar el mismo evento sonoro desde sitios diferentes, la diferencia de tiempo de llegada entre ellos puede ser utilizada para obtener la localización de la fuente del evento. En la figura 1.16 se muestran las ondas del mismo evento registrado con dos hidrófonos diferentes. Aunque las ondas son bastante diferentes la potencia de su contenido espectral no lo son de modo que si se realiza una correlación en tiempo de los dos espectrogramas de potencia se puede obtener la diferencia de tiempo de llegada entre los dos, que se denomina TDOA, de sus siglas en inglés.

Como se puede ver en la figura 1.16 la correlación tiene su máximo desplazado a la izquierda correspondiendo a un retraso de  $tdoa = -13,760$  ms en un caso y en el otro a un adelanto de  $tdoa = 7,360$  ms. Lo que permite localizar al animal que produce la vocalización situandolo sobre una isocrona como se muestra en la figura 1.17 donde se ha situado el icono de una orca sobre la isocrona correspondiente al retraso del primer caso.

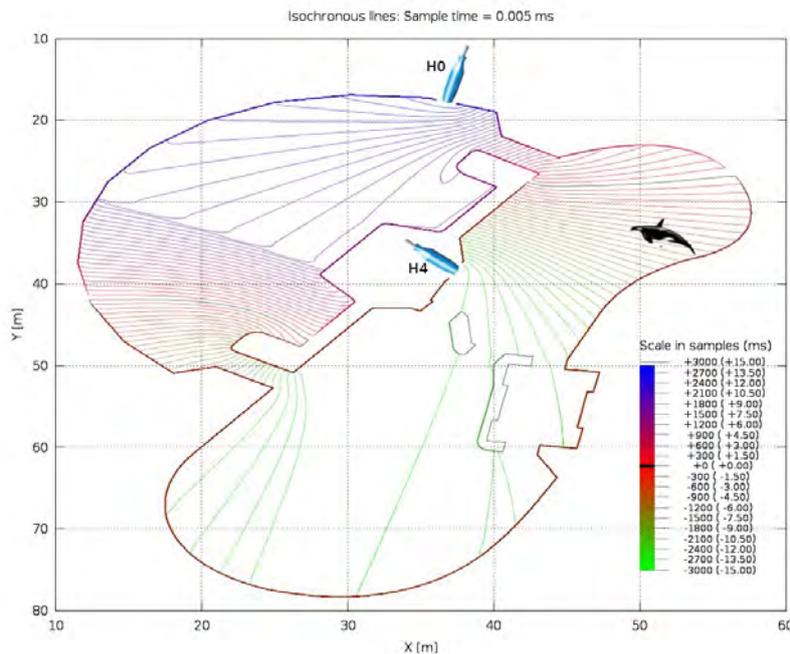


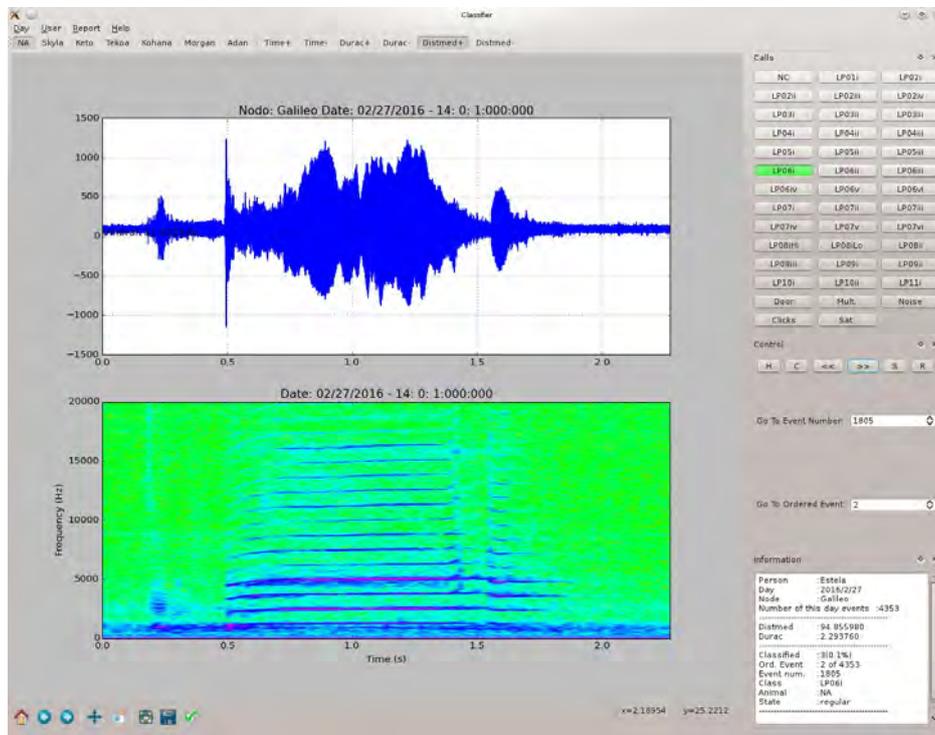
Figura 1.17: Isocronas para dos hidrófonos y situación en la isocrona  $tdoa = 0,005$  ms.

### 1.6.2 Clasificación manual de eventos

La identificación del tipo de llamada estereotipada que es un evento, o si el evento pertenece a un golpe del cierre de una puerta o el chapoteo del agua cuando un animal salta y vuelve a entrar en la piscina lleva mucho tiempo para un humano ya que el número de eventos detectados durante un día puede llegar a ser del orden de las decenas de mil. Es por lo tanto necesario el desarrollo e implementación de un clasificador automático de los eventos sonoros que permitan establecer niveles sintácticos en la producción de llamadas entre animales junto con su localización que permita identificar que animal es el que está vocalizando.

La utilización de inteligencia artificial para la clasificación automática de eventos necesita una gran cantidad de eventos previamente identificados. Es por esto que algunos clasificadores manuales han realizado el procesamiento de algunos miles de buenos eventos clasificándolos en categorías del dialecto. Estas clasificaciones se utilizan para el entrenamiento o aprendizaje de inteligencias artificiales.

La herramienta para realizar este tipo de clasificaciones manuales se muestra en la figura 1.18. En la figura se aprecia la gran cantidad de tipos de llamadas presentes en el dialecto.



**Figura 1.18:** Evento sonoro en la aplicación de clasificación manual.

## 2. Monitorización acústica automática

En la Ley 42/2007, de 13 de diciembre, del Patrimonio Natural y de la Biodiversidad en su Anexo V expone las especies animales y vegetales de interés comunitario que requieren una protección estricta. Bajo esta protección estricta se encuentran todos los cetáceos. En la tabla 2.1 se encuentra una relación de los cetáceos resumida que no pretende ser completa sino alumbrar la diversidad de especies que están bajo este paraguas de protección. Muchas de estas especies transitan las aguas de la Macaronesia. Cada una de estas especies tiene un modo diferente de producir sonidos y las estrategias para la identificación automática de una especie pueden no funcionar con otra.

**Tabla 2.1:** Clasificación común de cetáceos protegidos.

| Parvorden                   | Familia                                | Nombre común               |
|-----------------------------|--|----------------------------|
| Mysticeti                   | Balaenidae                             | Ballena franca             |
|                             |  | Ballena de Groenlandia     |
|                             | Balaenopteridae                        | Rorcuales                  |
|                             |  | Yubarta                    |
| Eschrichtiidae              | Ballena gris                           |                            |
| Cetotheriidae Neobalaenidae | Ballena pigmea                         |                            |
| Odontoceti                  | Delphinidae                            | Delfines oceánicos y orcas |
|                             | Iniidae, Platanistidae y Pontoporiidae | Delfines de río            |
|                             | Phocoenidae                            | Marsopas                   |
|                             | Physeteridae                           | Cachalote común            |
|                             | Kogiidae                               | Cachalotes enano y pigmeo  |
|                             | Ziphiidae                              | Zifios                     |
|                             | Monodontidae                           | Beluga y narval            |

La situación de boyas con hidrófonos o el transecto con gliders que tengan instalados hidrófonos se enfrentan con dos problemas diferentes y que deben ser abordados de modo distinto. El primero es la estimación del ambiente sonoro *Ambient Noise measurements* y el segundo es la identificación de especies *Marine mammals detection and monitoring*.

Para grabar sonidos en el mar es conveniente consultar el Manual Técnico de Acústica Pasiva de Castellote, revisar el documento técnico sobre impactos y mitigación de la contaminación acústica

marina y el magnífico libro de Au y Hastings sobre los principios de la bioacústica marina [2], [11], [12].

En el capítulo de Introducción ya se ha hablado suficiente sobre como se adquieren los datos pero es de especial atención para la automatización de la detección de una especie determinada utilizar la tasa de muestreo que mejor se acomoda a su sensibilidad acústica, ya que normalmente una especie produce sonidos útiles en el rango que puede oírlos y por lo tanto usarlos para comunicarse. En el caso de los cetáceos gran parte de los sonidos que realizan pueden grabarse con tasas de muestreo de 44100 muestras por segundo como en el caso del habla humana, sin embargo, determinados clicks de ecolocalización tienen sus picos de potencia en frecuencias mayores de 22050 hertzios y que por ello requieren tasas de muestreo más elevadas.

En la figura 2.1 obtenida de la documentación que acompaña al software PAMGuard se muestran los audiogramas de algunas especies marinas, y como aproximadamente en 100KHz está el final de la banda sensible. Por lo tanto, una frecuencia segura para adquirir estas señales sería 200000 muestras por segundo. Si cada una de estas muestras adquiridas ocupa en nuestro almacenamiento digital dos, tres o cuatro bytes tendremos cada segundo 400000, 600000 o 800000 bytes por segundo. Si nos ponemos en el caso más favorable para el almacenamiento estaremos en el caso menos favorable para el error de cuantificación de la amplitud sonora. Si un día con un hidrófono grabamos 32GB en un mes tendremos 1TB lo que implica que si queremos una monitorización permanente no podemos almacenar tal cantidad de datos de modo continuo.

De esta situación surgen las estrategias de reducción del tiempo de grabación a intervalos de tiempo, lo que también reduce los recursos de alimentación, ya que mientras no se está grabando no se tiene porque gastar energía. En todo caso esto hace que necesitemos reducir los datos. En el caso de querer la estimación de ambiente sonoro lo que haremos es realizar promedios de cantidades que son significativas en periodos de tiempo preestablecidos. En este caso podemos estimar indicadores de biodiversidad que corresponden a un número de cuatro bytes cada minuto. Esto corresponde a 169KB lo cual implica una reducción del 99,9 % en el almacenamiento. En todo caso para la estimación del ambiente sonoro utilizaremos más de un indicador de modo que la reducción será algo menor pero no significativa. En el caso de la determinación de especies para su monitorización la reducción se realiza manteniendo el evento sonoro como elemento significativo de la detección. En este caso la reducción se reduce en alrededor de un 80 % ya que solo parte del

tiempo la especie objetivo vocaliza y por lo tanto genera eventos sonoros. En la siguiente sección haremos la detección del canto de un delfín y su caracterización como ejemplo.

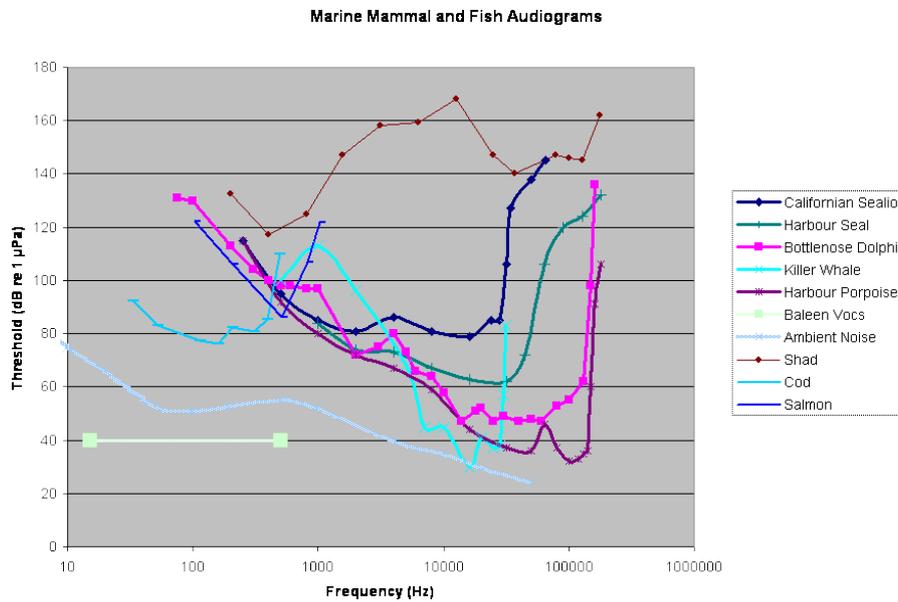
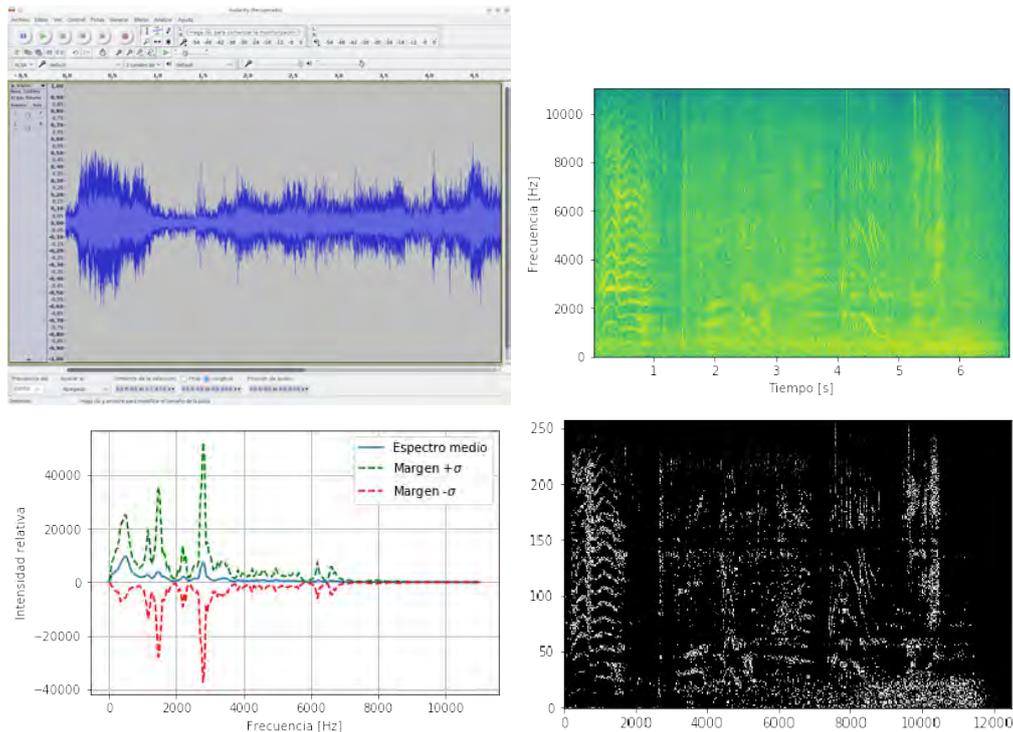


Figura 2.1: Esquema de almacenamiento y detección de eventos.

## 2.1 Detección automática de especies

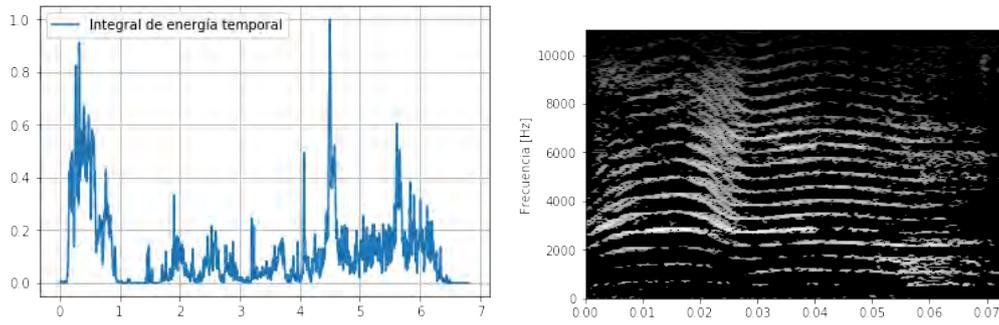
Imaginemos que obtenemos la grabación de un delfín cuya onda se muestra en la figura 2.2.a muestreado a 22050 muestras por segundo y realizamos un espectrograma que tendrá como frecuencia máxima 11025 hertzios como se muestra en la figura 2.2.b. Podemos observar seis segundos en los que el sonido evoluciona con distintas formas, procedemos a establecer la media y la variancia del ruido como se muestra en la figura 2.2.c y utilizando un umbral de energía en banda de frecuencia podemos seleccionar los canales de frecuencia-tiempo más intensos como se muestra en la figura 2.2.d.



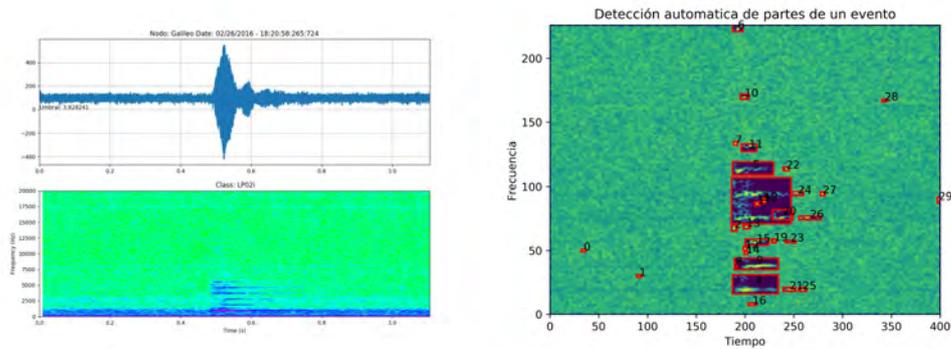
**Figura 2.2:** a) Onda de un delfín. b) Espectrograma del delfín. c) Media y Sigma del ruido. d) Espectrograma del delfín sin ruido.

Para ver más claramente el impacto de este proceso se muestra en la figura 2.3.a la energía en cada instante una vez que hemos retirado el ruido, donde se puede observar como al principio hay un evento muy relevante que se puede separar como el primer evento de modo automático quedando como aparece en la figura 2.3.b donde se puede ver claramente la evolución de esta parte tan intensa del canto. Esto permite utilizar algoritmos que identifiquen las características de este evento y luego tomar acciones asociadas como la identificación de la especie que es capaz de producir este sonido.

Ahora tomaremos un ejemplo real de un evento detectado automáticamente el día 26 de febrero de 2016 a las 18 horas, 20 minutos, 50 segundos, 265 milisegundos y 724 microsegundos, que sabemos por la clasificación manual que se trata de una clase



**Figura 2.3:** a) La integral en frecuencia del espectrograma de energías permite ver los momentos más intensos. b) Espectrograma del primer Evento más intenso sin ruido.



**Figura 2.4:** a) Evento automáticamente adquirido. b) Análisis de las partes del evento en 29 partes que permiten estimar que tipo de canto es.

LP02i que se muestra en la figura 2.4.a. Al realizar la aproximación anterior y proseguir en la identificación de las partes que lo componen podemos obtener 29 partes de modo automático, como se muestra en la figura 2.4.b, en el que ha simple vista podemos ver las componentes de frecuencia de un canto que está en nuestro dialecto conocido y por lo tanto se puede etiquetar de modo automático.

### 3. Bibliografía

- [1] K. V. Mackenzie, «Nine-term equation for sound speed in the oceans», *The Journal of the Acoustical Society of America*, vol. 70, n.o 3, ago. de 1981.
- [2] W. Au y M. Hastings, *Principles of Marine Bioacoustics*, ép. Modern Acoustics and Signal Processing. Springer New York, 2009, ISBN: 9780387783659.
- [3] T. W. Cranford, M. Amundin y K. S. Norris, «Functional Morphology and Homology in the Odontocete Nasal Complex: Implications for Sound Generation», *Journal of Morphology*, vol. 228, págs. 223-285, 1996.
- [4] D. Mann, M. Hill-Cook, C. Manire y col., «Hearing loss in stranded odontocete dolphins and whales», *PLoS One*, vol. 5, n.o 11, e13824, 2010.
- [5] E. Philpott, A. Englund, S. Ingram y col., «Using T-PODs to investigate the echolocation of coastal bottlenose dolphins», *Journal of the Marine Biological Association of the United Kingdom*, vol. 87, n.o 1, págs. 11-17, 2007.
- [6] J. Sarnocinska, J. Tougaard, M. Johnson y col., «Comparing the performance of C-PODs and SoundTrap/PAMGUARD in detecting the acoustic activity of harbor porpoises (*Phocoena phocoena*)», en *Proceedings of Meetings on Acoustics 4ENAL, ASA*, vol. 27, 2016, pág. 070 013.
- [7] A. Maginnis y D. Gillespie, «PAMBuoy® Beluga Monitoring», 2013.
- [8] U. K. Verfuss, A. S. Aniceto, D. V. Harris y col., «A review of unmanned vehicles for the detection and monitoring of marine fauna», *Marine pollution bulletin*, vol. 140, págs. 17-29, 2019.
- [9] G. B. Mindlin y R. Laje, *The physics of birdsong*. Springer Science & Business Media, 2006.
- [10] J. Ford y H. Fisher, «Group-specific dialects of killer whales (*Orcinus orca*) in British Columbia», *Behavioral Biology of Killer Whales*, págs. 129-161, ene. de 1986.
- [11] C. A. Chicote y M. Castellote, «Manual del Técnico de ACústica Pasiva para operaciones Off-shore generadoras de ruido en aguas españolas», inf. téc., 2013.
- [12] N. Aguilar, A. Tejedor y col., «Documento Técnico sobre Impactos y Mitigación de la Contaminación Acústica Marina», inf. téc., 2012.