Underwater noise recognition of marine vessels passages: two case studies using Hidden Markov Models

Manuel Vieira

Departamento de Biologia Animal and cE3c - Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, Bloco C2. Campo Grande, 1749-016 Lisboa, Portugal

M. Clara P. Amorim

MARE – Marine and Environmental Sciences Centre, ISPA – Instituto Universitário, Rua Jardim do Tabaco 34, 1149-041 Lisboa, Portugal

Present address: Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, Bloco C2. Campo Grande, 1749-016 Lisboa, Portugal

Andreas Sundelöf

Department of Aquatic Resources, Swedish University of Agricultural Sciences, Havsfiskelaboratoriet, Turistgatan 5, 453 30 Lysekil, Sweden.

Nuno Prista

Department of Aquatic Resources, Swedish University of Agricultural Sciences, Havsfiskelaboratoriet, Turistgatan 5, 453 30 Lysekil, Sweden.

Paulo J. Fonseca

Departamento de Biologia Animal and cE3c - Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, Bloco C2. Campo Grande, 1749-016 Lisboa, Portugal

Passive acoustic monitoring (PAM) is emerging as a cost-effective non-intrusive method to monitor the health and biodiversity of marine habitats, including the impacts of anthropogenic noise on marine organisms. When long PAM recordings are to be analysed, automatic recognition and identification processes are invaluable tools to extract the relevant information. We propose a pattern recognition methodology based on hidden Markov models for the detection and recognition of acoustic signals from marine vessels passages and test it in two different regions, the Tagus estuary in Portugal and the Öresund strait in the Baltic Sea. Results show that the combination of hidden Markov models with PAM provides a powerful tool to monitor the presence of marine vessels and discriminate different vessels like small boats, ferries and large ships. Improvements to enhance the capability to discriminate different types of small recreational boats are discussed.

I. Introduction

Underwater noise has been increasing during the last decades (Markus & Sánchez, 2018), altering soundscapes throughout most aquatic environments (Watts *et al.*, 2007; Normandeau Associates Inc., 2012). Consequently, anthropogenic noise is now recognised as a pollutant under the international legislation (e.g. descriptor 11 on the European Commission Marine Strategy Framework Directive, MSFD, 2008/56/EC, inclusion in the US National Environment Policy Act, and as a permanent item on the International Maritime Organization Marine Environmental Protection Committee agenda). Although recent studies have demonstrated that boat noise can affect the behaviour and physiology of various aquatic species (e.g. Graham & Cooke, 2008; Picciulin *et al.*, 2012; Bruintjes & Radford, 2013; Castellote *et al.*, 2012; Rolland *et al.*, 2012; Holles *et al.*, 2013; Voellmy *et al.*, 2014; Nedelec et al., 2015; Edmonds *et al.*, 2016; Marley *et al.*, 2017; Putland *et al.* 2018) present knowledge on the prevalence of man-made noise is still limited.

Single hydrophone passive acoustic monitoring (PAM) coupled with automatic recognition methods is a promising tool for continuous assessment of anthropogenic noise in the marine environment. This is particularly important in the case of marine vessel noise, the main source of continuous man-made ocean noise (McDonald *et al.* 2006). The main sources of vessel noise are machinery, cavitation by the propeller and other structures, and hydrodynamic processes. The recorded noise can vary depending on vessel conditions such as speed, orientation, manoeuvring, and distance to the hydrophone, especially at low depths (Trevorrow *et al.*, 2008; Zak 2008; Averbuch et al., 2011; Traverso *et al.* 2015). PAM has been recently used for the determination of boat visits to artificial and natural reefs off Florida (Simard *et al.*, 2016) and boat passages in a river (Averbuch et al., 2011). Capacity to discriminate noise from vessels of different size, hull-material and engine type has been documented (table 1), as well as the use of Coherent Hydrophone Arrays to detect and track ships (Huang *et al.* 2017, Zhu *et al.*, 2018; table 1). However, widespread usage of PAM for monitoring boat traffic has remained limited in part due to difficulties in analysing the large acoustic datasets generated by long term acoustic monitoring.

Several approaches have been attempted to study extensive acoustic recordings. The simpler and more commonly employed methods involve automatic detection that make use of e.g. energy thresholds or a matched filter to locate the chosen acoustic pattern in the recordings (table 1). Such methods are sometimes followed by common procedures of multivariate statistical analysis to categorize sound types (e.g. discriminant function analysis; Averbuch et al., 2011). With the improvement of models and techniques for automatic speech recognition in the past few decades, the recognition of acoustic patterns has become increasingly faster, more accurate, and robust. Robust methods using machine learning, such as Gaussian mixture models (GMMs; Reynolds and Rose, 1995), artificial neural networks (ANN; Lippmann, 1988; Yu et Oh, 1997), and hidden Markov models (HMMs; Baker, 1975; Jelinek, 1976; Jelinek *et al.*, 1975; Rabiner, 1989; Young and Bloothooft, 1997) have been successfully used to recognize and classify human speech, other animals' vocalizations (Somervuo *et al*., 2006; Scheifele *et al*., 2015; Vieira *et al*., 2015; Putland *et al*., 2017; Ranjard *et al*., 2017, Vieira *et al*., 2019) and anthropogenic noise (Feroze *et al*., 2018). Methods used in speech and scene recognition (e.g. HMMs, ANN) are capable of dealing with extensive recordings permitting recognition and classification of each sound. In particular, HMMs can be used to statistically model both temporal and spectral variations of acoustic patterns through robust algorithms allowing optimization of relevant mathematical criteria. Furthermore, due to the extensive research on speech recognition, this method is currently available in several freeware applications (Young et al., 2006).

In aquatic environments, HMMs have been mainly adapted and successfully applied to the recognition of vocalizations of marine mammals and fish (marine mammals: Pace *et al.*, 2012; Putland *et al.*, 2017; fish: Vieira *et al.*, 2015; Vieira *et al.*, 2019). Given that HMM methods are based on temporal and spectral variations, and since these disparities are also known to occur among marine vessel noise, it is plausible to adapt HMMs to recognize the passages of marine vessels. To date, however, HMMs have not been applied for detecting and classifying marine vessels possibly because this method was not initially developed for classification of stationary signals. Temporal variations of sounds from marine vessels occur but are mainly related to sound propagation. Table 1 shows some of the few studies on marine vessels sound detection and classification (table 1).

In this paper we developed a HMMs-based automatic recognition method to detect and recognize different vessel types and test it in two case studies: (1) recognition of small boats recorded as acoustic snapshots at several marinas across the Öresund strait (Sweden); and (2) recognition of different types of marine vessels recorded with PAM in a channel of the Tagus estuary (Portugal) with boat passages to a nearby ferryboat terminal.

Our specific goal with the Öresund strait case-study was to test the association of PAM and HMM for the recognition and quantification of boats circulating at the entrance of several marinas. The counting of boat passages can be particularly useful in, e.g., recreational fisheries surveys where direct estimates of fishing effort are frequently needed (Hyder et al. 2018) but very difficult to obtain (e.g. number of fishing trips, Pollock et al. 1994). We tested discrimination of boat types to separate the number of trips per boat type. This is especially relevant since the relative importance of each boat type to the recreational fishing differs (e.g. open deck private boats are used much more often for recreational fishing than sail boats).

In the Tagus estuary case-study our aim was to create an automatic recognition system capable of identifying the presence of noise of some marine vessels. This system could be useful to evaluate the impacts of the marine vessels passages on the vocal activity of soniferous fish, such as the Lusitanian toadfish and the meagre (Amorim *et al*., 2006 Prista, 2014) and other aquatic organisms, or to monitor its impact on aquatic soundscapes.

II. METHODS

A. Data collection

A.1 Öresund strait (Sweden)

Acoustic recordings were made from 4 to 7 of July 2017 in thirteen marinas along the Öresund strait (Sweden, figure 1): Domsten, Vikingstrand, Helsingborg, Knähaken, Råå, Borstahusen, Landskrona, Lindeshamn, Lomma, Malmö Västra Hamnen, Limhamn, Klagshamn and Strandhem. Sounds were registered with a High Tech 94 SSQ hydrophone (sensitivity of –165 dB re 1 V/µPa, flat frequency response up to 6 kHz ± 1 dB) and a Tascam DR-40 Portable Digital Recorder (48 kHz, 16 bit resolution). The hydrophone was deployed at a water depth of 0.6 - 1.2 m, depending on the marina. Each recording was accompanied by photos of the boat involved so that sounds and boat type could later be matched. Overall, the acoustic recordings lasted 1 to 6 hours depending on the boat traffic intensity and contained sounds from boats with different characteristics (table 2; figure 2 and photos in figure S-2). The soundscapes of these ports and marinas were dominated by boat noise, with almost no other sound either from biological or non-biological origin.

A.2 Tagus estuary (Portugal)

The data set consisted of ca. 6 days round-the-clock recordings of sounds obtained from 15 to 20 of May 2017, in the Tagus estuary (Air Force Base 6, Montijo, Portugal; 38º42'N, 8º58'W). Water depth varied approximately between 3 - 6 m, depending on tide. The signal from a High Tech 94 SSQ hydrophone was recorded (4 kHz, 16 bit resolution) by a 16 channel stand-alone data logger (Measurement Computing Corporation LGR-5325, Norton, Virginia, USA). The hydrophone was anchored at about 20 cm from the bottom to a stainless steel holder projecting from a concrete base where the cable was attached to minimise current-induced hydrodynamic noise,

Recordings contained sounds from different types of vessels passing the recording site. Each vessel was manually classified into 3 broader categories according to their acoustic properties (duration and Lloyd's mirror effect; Carey, 2009; figure 3, and photos in figure S-3) previously subjected to visual identification. Ferries passages had a bigger duration than the smaller private boats and were also confirmed using their departure schedule. The soundscape of this estuary channel was dominated by vessel noise and sounds from biological origin (e.g. fish choruses).

B. Pattern recognition

The proposed noise recognition systems were adapted from those described in Vieira *et al.* (2015) and Young *et al.* (2001) using HMMs. The overall flowchart of the method is shown in figure 4.

B.1. Signal processing

The first stage in the signal processing splits the waveform signal into a sequence of elementary segments according to a predefined window duration (see figure 4). This window should be longer than a cycle of the lower relevant frequency but short enough to provide temporal resolution while also assuring stable properties. After some preliminary tests, we chose a window of 200 ms with a 50% overlap to avoid losing information on the transition between two consecutive elementary segments (O'Shaughnessy, 1987). To try to extract the most relevant information from the signal, we selected the following features: cepstrum, Mel-frequency cepstral (MFC), delta, and acceleration coefficients (more information about these features in Table S-3).

B.2. The HMM time alignment structure

Each sound type has an average expected duration that is directly related to the number of states. For example, a human phoneme is usually modelled by three states (McDermott *et al.*, 1990). However, because there are no phonemes in marine vessels noises, we assumed that the number of states should be equal to or higher than the number of different consecutive stable parts of the sound, taking into account the stochastic variability and the median duration of these sounds. Note that we used models with a linear topology in which all the states could transit to the same state, to the next or to the following one (except the initial and final states where self-transitions are meaningless as they only serve as signal boundary markers; figure 4). This type of transitions between states should give enough flexibility to each model to reflect the vessels noise variations (e.g. different durations of stable noise caused by different speed). After some preliminary tests (figure S-1), we considered 224 states for marine vessels sounds and 5 for background noise (silence) models. To analyse the Tagus estuary dataset we added extra models with 224 states for modelling non-biological patterns with high energy and duration (e.g. consecutive non-biological pulses with high energy), and biological patterns (e.g. some fish choruses; see figure 3).

For each sound type, a representative subset of samples (e.g. passages of a particular class of boats) was used to train the HMMs. The transition probabilities and the elementary segment probability densities of each state were estimated with the Baum–Welch algorithm (Baum *et al.*, 1970; figure 4).

In the recognition phase, each vessel noise was matched against the estimated HMM for each sound type. This was achieved by using a Viterbi algorithm (Forney, 1973) that produced a likelihood measure for each HMM. The vessel noise was assigned to the sound type corresponding to the HMM with the highest likelihood.

For computations we used the HMM Toolkit (HTK, University of Cambridge, UK), a group of modules written in C to create automatic recognition systems for human speech (Young *et al.*, 2006).

B.3. Automatic recognition systems

B.3.1 Öresund strait (Sweden)

Automatic HMM-based systems were prepared to (1) recognise boat noise (without discrimination of boat type), (2) recognise each boat type and additionally a system to (3) discriminate boats arriving and boats leaving the port. To take full advantage from the available data and overcome the variability caused by bias in training data selection, a resampling method was used based on a random subsampling validation (Efron, 1981). Details the resampling procedure are described below. All trials were repeated 100 times.

1 - The boat noise recognition system (without discrimination of boat type) was based on one HMM that considered all registered boat types (table 2). Each training set used to

produce a recognition system included 20 boat sounds randomly selected from the overall dataset. This procedure was repeated 100 times. Note that some boat types had small sample size with less than 15 recorded sounds (table 2). The system was tested with the field recordings (each with a different duration between 5 and 75 min) and optimised by testing different frequency bandwidths adjusted to the spectrum of the boat noises recorded in the field. The preliminary tests considered different frequency cut-offs; low (0, 20, 200, 500, and 1000 Hz) and high (1000, 5000, 10000 and 20000 Hz). Here we show the results using different low (20, 500, and 1000 Hz) and high (2000, 10000 and 20000 Hz) frequency cut-offs.

2 - The boat type recognition system was created using a different HMM for each of 12 boat types (commercial fishing boat, recreational fishing tour boat, open deck private boats with outboard engine, open deck private boats with inboard engine, rigid inflatable boats (RIB), sail boat with inboard engine, sail boat with outboard engine, jetski, small yacht with inboard engine, small yacht with outboard engine, double ender boat, medium to large yacht; figure S-2) and using a total of 208 boat sounds. These categories were selected to monitor how many boats of each type transited in this area as a proxy to the recreational fishing effort. From these, four sounds were randomly sampled and included in the training set for each boat type. Sounds used in the training set were included in the testing set. A full system, involving all boat types showed low identification rate possibly because of the low number of samples. Consequently, we developed a system using only the most common boat types (open deck boat with outboard engine and sail boat with inboard engine) using the same protocol except that sounds used in the training set were not included in the testing set. Training sets using 4 and 8 sounds were tested. We present the results of the best classification system we obtained after a range of other alternatives were tested. This system involves using 1 second segments of the recordings centred in the maximum sound pressure level of each boat sound.

3 - The automatic recognition system to discriminate sound of boats arriving and leaving the ports was trained for each boat noise type using sounds from the most common boat (open deck private boats with outboard engine). A total of 49 boat noise samples were used. From these, four sounds were randomly resampled and included in the training set for both HMMs (boats arriving or leaving). Sounds used in the training set were not included in the testing set.

B.3.2 Tagus estuary (Portugal)

An automatic HMM-based system was prepared to recognise marine vessel types. This procedure included the noise produced by small private boats without AIS (mostly open deck private boats with outboard engine), ferries, and other anthropogenic unknown source (possibly large ships at distances higher than 1 km). We considered "small boats" as vessels with less than 12 m (mostly open deck private boats with one outboard engine) and ferries as the ca. 50 m long passenger vessels that connect the localities of Lisbon and Montijo (figure S-3).

The marine vessels' type recognition system was trained for each sound type using sounds from the two first recording days (sounds from 142 passages were used). The ferries and other type of anthropogenic noise of unknown origin classes were subdivided into two models each, to reduce the diversity between each model and increase the overall identification rate. The small boats class was represented only by one HMM. Additionally, we used 13 sounds (with low energy noise with no obvious abiotic or biotic sources) for the background noise model, 13 sounds for modelling non-biological patterns with high energy, and 77 sounds for the biological pattern models, namely the fish choruses (figure 3).

The system was tested with the recordings of the subsequent four days (a total of 96 hours with 286 vessels sounds). Several frequency bandwidths were tested (0 to 2000 Hz, 1000 to 2000 Hz, 1200 to 2000 Hz). We only present results using 1200 to 2000 Hz since this bandwidth showed the best results as it avoided the interference of fish choruses (see examples of choruses in figure 5).

B.4. Evaluation of the recognition system

For each optimal alignment, the number of substitution errors (i.e., when one signal type is recognised as another signal type, S), deletion errors (i.e., when a sound type occurs but is not detected by the system – a false negative, D), insertion errors (i.e., when a signal is detected by the system but it did not occur - a false positive, I) the total number of labels in the reference transcriptions (N) were determined (Young *et al.* 2000). The performance of the recognition systems was then evaluated by computing the percentage of correctly recognized sounds (identification rate) using:

$$\text{Identification rate (\%)} = \frac{\mathrm{N} - \mathrm{D} - \mathrm{S}}{\mathrm{N}} \times 100,$$

or by computing the recognition accuracy using:

$$\text{Accuracy (\%)} = \frac{\mathrm{N} - \mathrm{D} - \mathrm{S} - \mathrm{I}}{\mathrm{N}} \times 100.$$

Additionally, we calculated the ratio between vessel hits (number of sounds events identified by the system) presented by the recognition system and the total number of vessels passages in each file. This can be relevant to verify if the number of hits can be used as a proxy of the number of vessels that passed by.

III. RESULTS

A. Sound Properties

A.1 Öresund strait (Sweden)

Over 10 vessel types were recorded in the Swedish ports and marinas during the field work. Most sounds came from boats with less than 10 m long (table 2). Power spectral density (PSD) plots of the noise produced by each boat type are represented in figure 2. Overall, dominant frequencies of noises from several boats were within the range 200-2000 Hz. Although the PSD mean values varied among boat types (figure 2), the large overlap difficulted the distinction of boat types. There was some variation among the background noise recorded at each port, but it was on average 20.7 ± 4.6 dB below boat noise. The duration of the vessel sounds presented a high variability that can be related to different underwater seascapes (topography, presence of sound propagation barriers, water depth, etc), boat velocity, engine sound intensity, distance to the hydrophone, and some vessel manoeuvres (table 2). None of the recorded boats showed a noticeable Doppler effect, but almost all showed a Lloyd's mirror effect. Doppler effect causes a frequency shift on the sound wave emitted as a result of the motion of the emitter, shifting from higher to lower frequencies with the approach and then departure of the boat from the recording hydrophone (Urick, 1983). The Lloyd's mirror effect is the result of out-of-phase reflections of the sound. This effect also shows a shift on the frequencies observed according to the distance of the source, but is usually symmetrical between approach and departure (Carey, 2009). Only some boats parking or starting the engine near the entrance of the port (where the hydrophone was deployed) showed acoustic signature that could be related to the manoeuvres (figure 5).

A.2 Tagus estuary (Portugal)

There were three types of anthropogenic noises detected during the recordings: small private boats without AIS, ferries, and anthropogenic sounds of unknown source. Most traffic was from ferries. Power spectral density (PSD) plots of each sound type are represented in figure 3. The duration of vessel passage sounds varied from ca. 20 s for small boats, to ca. 50s for ferries, while the noise from an anthropogenic unknown source presented a high variation (from 20 s to several min). The latter include engine-type noises apparently stationary, most probably large transport ships located very distant from the recorder device. Lloyd's mirror effect was evident on most ferries' recordings (see figure 3), while only some small boats showed clearly this effect. None of the recorded noises from an anthropogenic unknown source exhibited a noticeable Doppler and Lloyd's mirror effect. We detected choruses produced by fish species (see figure 3 and figure 5), namely Lusitanian toadfish (Amorim *et al.*, 2008), Meagre's long grunts (Lagardère et Mariani, 2006), and series of isolated pulses (Pereira, 2019). The sounds produced by these species were only detected between ca. 50 -1200 Hz.

B. Vessels recognition

B.1 Öresund strait (Sweden)

Automatic HMM-based systems were prepared to (1) recognise boat noise (without discrimination of boat type), (2) recognise each boat type and (3) discriminate boats arriving and boats leaving the port.

1 - The recognition systems considering all boats as one class (without discrimination of boat type), presented correct identification rates ranging from 75 to 100% (table S-1). Accuracy ranged from 25 to 86%, being highly affected by the randomly selected training data (table S-1). Each recognition system segmented the boat sounds differently, sometimes one boat was segmented in several hits, leading to lower accuracy value calculated using HTK algorithm (Young et al. 2000; see figure 4). Different frequency bandwidths (figure 6 and figure S-4) were tested. Increasing the lower frequency of the filter bandwidth led to an increase in the number of segments generated by the recognition system, which proved useful in cases where the sound from different boats was partially overlapped. On the other hand, decreasing the bandwidth's lower frequency had an opposite effect that could be useful to count boats in case of repeated variations of boat velocity (including repeated turning off and on of the engine; figure 5B). As expected, a reduced number of hits, was found when boat noises overlapped. Figure 5A shows an example of the output of the boat noise recognition system applied to a 15 min long recording using a 20-10000 Hz frequency bandwidth. The number of hits varied from an underestimation of the real boat passages of 83% to an overestimation of 110% (figure S-5).

2 - Several frequency bandwidth combinations were tested to create identification systems for each boat type. The 20-5000 Hz bandwidth produced the best output, resulting

in an overall mean identification rate of 15.9 ± 3.4 % (mean ± standard deviation; accuracy with the same value). Notice that the overall mean identification rate is obtained by averaging 100 outputs simulated with the identification system. Each boat type was thus poorly recognized by the system.

Because the low identification rate could be due to the small number of samples available for some boat types, we tested a simplified system considering only the two most common boats: open deck with outboard engine and sail boat with inboard engine. Using the same 20–5000 Hz bandwidth the overall mean identification rate of these two boat types improved to 62.6 ± 5.8% using 4 sounds in the training set (accuracy with the same value, table S-2), and 63.0 ± 7.4% using 8 sounds in the training set. This identification rate was above the value expected by chance alone (50%), despite the overlapping characteristics of the sounds produced by these two boat types (figure 2, figure S-6).

3 - The classification according to the direction of the boat (arriving or leaving the port) achieved an identification rate of circa 50% (51.0 ± 7.7%), a value that could be expected by chance alone.

B.2 Tagus estuary (Portugal)

The 1200–2000 Hz bandwidth allowed the best results by the marine vessel noise type automatic recognition system. A mean identification rate of 90.9 ± 8.2 % (and an accuracy with the same value) was obtained for all vessels using recordings from four days. This system achieved a higher identification rate when considering only the ferryboats (95%), while small boats and anthropogenic unknown sources were recognized with mean identification rates of 67 % and 86 %, respectively. Some mistakes in the classification of small boats were due to misidentifications with a ferry. Note that the small boats were less common, with only 24 detectable passages during the four days in contrast to 169 ferries passages. Table 3 represents the mean confusion matrix. The total number of hits on the four days tested varied from an underestimation of vessel passages of 71 % to a small underestimation of 95 % (due to some substitution errors). Although the anthropogenic unknown source had a high correct classification of sound events, the number of hits should not be interpreted has number of passages or number of sound sources, because it appears to be a unique stationary source.

Figure 7 illustrates the presence of marine vessels at the passive acoustic monitoring station in the Tagus estuary (Portugal), estimated using the automatic recognition system. Figure 7A shows the quantification of vessels by the number of hits, while figure 7B represents the total time per 2 hours where a marine vessel sound was detected. As expected, ferries start passing by at 6 am on working days, and the peak traffic periods are 6-10 am and 6-10 pm. On a Saturday (20 may 2017) the number of ferries reduces. Comparing Fig. 7A and 7B, we can observe that small boats had a smaller duration due to higher velocity and/or less source noise intensity than ferries. Note that if a vessel stays stationary during a long period of time and/or changes engine power significantly it could cause an overestimation of the number of vessels.

IV. Discussion

We show that automatic recognition methods based on hidden Markov models coupled with PAM is a valid and easible option for monitoring the presence of different types of marine vessels in a variety of aquatic systems (e.g., port channels, Marine Protected Areas). These tools rendered good vessel identification rates being both cost- and time-effective while free of privacy-related issues associated with other alternatives (e.g., video surveillance). Furthermore, this kind of automatic recognition systems can have other applications, from monitoring of biological activity to characterization of background noise levels and disturbances due to human activities. Although this method can be effective for detection and classification of vessels in specific estuaries and marinas, it would probably not provide a universal recognition system. Each system should be trained using a library of sounds collected in the locations under study and conditions.

A. Öresund strait (Sweden)

Our specific goal with the Öresund strait case-study was to test PAM and HMM in the recognition, classification and quantification of boat passing the entrance of several marinas (map in figure 1 and boat types in figure 2). The counting of boat passages can be particularly useful in e.g. recreational fisheries surveys, where it is frequently necessary to sample and estimate (or validate) total effort (number of fishing trips) carried out by private boats (Pollock *et al.* 1994).

The automatic recognition system developed in the present study was able to detect the presence of boats on recordings of underwater sounds. The system featured a combination of cepstrum, Mel-frequency cepstral (MFC), delta, and acceleration coefficients and reached an identification rate above 95%, being little influenced by the different frequency bandwidths tested (20-2000 Hz, 500-2000 Hz, 1000-2000 Hz, 20-10000 Hz, 500-10000 Hz, 1000-10000 Hz, 20-20000 Hz, 500-20000 Hz and 1000-20000 Hz). The use of different bandwidths caused only a small variation in the detection rate generated by the boat recognition system (cf. figure 5 and figure S-4). Nevertheless, some inaccuracies do exist such as multiple recognitions of the same boat mostly due to variations on velocity (including turning the engine off and on) common at the entrance of ports and marinas, which may cause an overestimation of boat passages. Future work should consider a step to join sequential hits which would minimize this type of overestimation. Another issue was the overlapping of noise from two different boats that could sometimes be identified as a single boat thus causing an underestimation of vessel counts. The improvement of the algorithm accuracy warrants longer term recordings (to obtain a more complete set of reference boat types) and testing.

The development of an automatic recognition system capable to differentiate boat types (table 2) could be a considerable advantage in the context of recreational fishing effort estimation because some boat types are more likely to be used for recreational fishing (e.g., recreational fishing tour boats, open deck vessels) than others (e.g., commercial fishing vessels, sail boats). Testing such ability was the focus of the second system developed in the Öresund case-study. In the trials where we discriminated all 12 visually identified boat types, the recognition system reached a low identification rate, only barely surpassing the value expected by chance alone (for 12 possible choices it is expected a probability of approximately 8 %, or 1/12). This result was likely due to the small sample size for most boat types. The current categorization based mostly on the size and use of the vessels could also be responsible for the poor performance, although the mean confusion matrix did not reveal clear patterns of recurrent misclassification. When the HMM was developed with the two most common boats (open deck boat with outboard engine and sail boat with inboard engine), sample sizes were larger and so was the discrimination capability of the automatic recognition system (a mean identification rate of 62.6 ± 5.8 % was obtained, despite the spectral similarities of the noise produced by those boats). This suggests that it should be possible to develop a system with a reasonable number of boats, provided that an initial large dataset is used, offering exciting opportunities to monitor the activity of different boats. Considering the present difficulties of quantifying recreational fishing effort in many regions of the world, even a very simple and autonomous system with only two boats types (such as the one developed here) would bring significant improvements to the understanding of the impacts and dynamics of those fisheries. In this experiment we used 1 second recordings that also limit the Lloyd’s mirror effect on the HMM’s recognition abilities. The temporal characteristics of the Lloyd’s mirror effect depends on several factors (e.g. boat velocity and source level), an additional information that, if available, could help better distinguish vessels passages. In the Tagus estuary the Lloyd's mirror effect was a key information to distinguish marine vessels classes.

An additional perspective on the capabilities of the PAM-HMM system is given by the third system developed in the Öresund strait. Here our goal was to test the capabilities of the method to distinguish between outgoing and incoming vessels. Such distinction could be useful to assess circadian rhythms of fishing effort in particular and marina usage in general. The majority of the boat sounds recorded did not exhibit a detectable difference regarding the direction of the movement at the entrance of the marinas, where the speeds are very low and therefore no clear Doppler effect is expected. Only boats parking or starting the engine near the entrance of the port (where the hydrophone was deployed) showed a signature as reported by Averbuch *et al.* (2010). Averbuch *et al.* (2010), presented an algorithm based on the combination of the Linear Discriminant Analysis (LDA) and the Classification and Regression Trees (CART) to detect the arrival and mooring, and departure of passengers' vessels, in cases where the sound shows a clear sequence of expected manoeuvres. This restricts the use of such a system to specific conditions where it is possible to record the mooring and the engine start of all the vessels thus calling for a more comprehensive recognition system.

B. Tagus estuary (Portugal)

Here the usefulness of HMM-based automatic recognition systems to extensively recognise marine vessels in relatively noisy estuary conditions is demonstrated. In fact, the sounds used in the present study were registered in a complex natural estuarine environment not only presenting fluctuations of environmental parameters affecting sound (e.g. current speed, wind, temperature, turbidity, salinity) but also of biological sounds such as fish choruses.

The results of the HMM-based recognition system using as features a combination of cepstrum, Mel-frequency cepstral (MFC), delta, and acceleration coefficients and a frequency bandwidth of 1200-2000 Hz, showed a good performance. In this case we restricted the sound frequency bandwidth to 1200-2000 Hz to avoid overlapping with fish choruses (see figure 4 for an example of overlap between the frequency range encompassing marine vessels noise and fish vocalizations). This system achieved a high identification rate when considering only the ferryboats (ca. 95 %). As shown by Vieira *et al.* (2015), a larger number of sounds used in the training phase usually improves the model's recognition ability, an advantage of the large data set available. Extending the bandwidth to lower frequencies in locations without the presence of such biological sounds may further improve vessels passages detection.

In the case of the anthropogenic noise of unknown origin, which may include distant stationary or passing vessels, the system showed a good performance in recognizing the sounds. However, the number of hits must be considered with care since it might not be a good proxy to the number of sources, that can be under- or overestimated. Nevertheless, the high precision of the automatic system in detecting this noise allowed measuring its total duration. Assessing the presence/duration of unidentified anthropogenic noise may be useful to characterise soundscapes and human impact.

Future work using this system may allow evaluating the effects of the presence of vessels in fish behaviour, especially relevant in fish breeding and nursery areas such as estuaries. This is especially important since marine vessel noise components under 1 kHz overlap with the fish hearing range, affect fish larval stages, induce stress-responses, interfere with communication and with the detection of predators and prey (Vasconcelos *et al.* 2007, 2011; Picciulin *et al.* 2012; Voellmy *et al.* 2014; Nedelec *et al*. 2015) In fact, marine vessels noise components within 20 - 1000 Hz, overlap with the hearing range of both the Lusitanian toadfish (Vasconcelos et al., 2007, 2011) and the meagre (M. Beauchaud and P. J. Fonseca, unpublished results), and may interfere with fish communication.

C. Comparison between Öresund strait (Sweden) and Tagus estuary (Portugal)

Monitoring the general increase of boating activity can take advantage from PAM allied to automatic recognition methods, especially if focussed on private boats not required to use AIS (Automated Identification System). In fact, in contrast with large scale fishing vessels that are monitored though the Vessel Monitoring System (VMS), the presence of small boats may be difficult to monitor (Pollara *et al.*, 2017) since they are usually not equipped with AIS and, due to their size, they are not generally well detected by radar. Therefore, the development of small boats' detection systems is a most needed but relatively unexplored research field (table 1). In fact, although some work exists on characterization of sounds produced by boats and on the categorization of anthropogenic noise (table 1), only limited attempts have been made to automatically recognize private boats, and to separate boats and what appears to be noise from large ships .

HMM-based boat recognition methods together with PAM could be an important tool to monitor the presence of small scale and recreational fishing activity on marine parks with restriction areas. The automatic recognition systems in this study were not entirely successful in discriminating amongst boats recorded in the Öresund strait. Several boat types produced rather similar waterborne noise. Nevertheless, the recognition system proved reliable to discriminate between groups of less similar vessels (small boats, ferries and anthropogenic noise of unknown source) in the Tagus estuary. An important difference between the two studied areas relate to the place where boats were recorded. While at the Öresund strait the recordings were made at the entrance of marinas, where it was common to observe boats manoeuvring and many recordings overlapped two or more boat noises, at the Tagus estuary almost all small boats and ferries passed by at a constant velocity and there were almost no overlaps of vessel noises. In order to use PAM as a proxy for estimating number of boat passages one should avoid sites where manoeuvring boats are expected to occur.

V. Conclusion

The increase in the use of small recreational boats together with the need to monitor and manage protected areas and fisheries call for an operationally reliable and cost-efficient tool to be used on a continuous basis to monitor and recognize passing boats. In addition, our knowledge regarding the impact of boat noise on aquatic organisms is still limited and could greatly benefit from such a tool. Automatic recognition methods of AIS non-trackable boats coupled with PAM can offer such a tool but is a relatively unexplored research field (table 1). Here we present an automatic recognition system able to pinpoint the passage of marine vessels in one environment with a soundscape characterized by the presence of biological sounds (Tagus estuary) and in environments with almost no biological sounds (several marinas at Öresund strait). Despite the difficulties in differentiating boat types, it demonstrates the capability to recognise boats from ferries and stationary anthropogenic of unknown source with high accuracy. Therefore, this recognition system, which adapts a free and established system for human speech recognition (HTK, Young *et al.*, 2000), can be an accessible and important tool in studies where long-term monitoring of boating and shipping is required. The performance and efficacy of this recognition method would be better exploited on local dimensions, by training the system with typical signal types (and propagation characteristics) of each specific location, including common sounds of geophony and biophony.

Supplementary material

The supplementary tables (S-1 to S-3) and figures (S-1 to S-6) are available at ICESJMS online.

Acknowledgments

This study was funded by the Science and Technology Foundation, Portugal [grant SFRH/BD/115562/2016 to M.V.; strategic projects UID/MAR/04292/2013 to M.C.P.A. by MARE and UID/BIA/00329/2013 to P.J.F. by cE3c; and project PTDC/BIA-BMA/29662/2017]. Part of this work was performed through, and financed by, the Swedish National Work Plan for data collection in the fisheries and aquaculture sectors 2017-2019, in relation to EU regulations (EU) 2017/1004 and (EU) 2016/1251. The Swedish National Work Plan is co-funded by the EMFF (Regulation (EU) No 508/2014) and by Swedish national funds through Swedish Agency for Marine and Water Management. We would like to thank Maria Hansson and Filip Bohlin for their crucial assistance allowing this investigation. We also thank Marta Bolgan and three anonymous reviewers for providing helpful comments that led to the improvent of the manuscript.

References

Amorim, M. C. P., Simo~es, J. M., and Fonseca, P. J. 2008. Acoustic communication in the Lusitanian toadfish, Halobatrachus didacty- lus: evidence for an unusual large vocal repertoire. Journal of the Marine Biological Association of the United Kingdom, 88: 1069–1073.

Amorim, M. C. P., Vasconcelos, R. O., Marques, J. F., and Almada, F. 2006. Seasonal variation of sound production in the Lusitanian toadfish Halobatrachus didactylus. Journal of Fish Biology, 69: 1892–1899.

Averbuch, A., Zheludev, V., Neittaanma¨ki, P., Wartiainen, P., Huoman, K., and Janson, K. 2011. Acoustic detection and classifi- cation of river boats. Applied Acoustics, 72: 22–34.

Baker, J. 1975. The DRAGON system—an overview. IEEE Transactions on Acoustics, Speech, and Signal Processing, 23: 24–29.

Baum, L., Petrie, T., Soules, G., and Weiss, N. 1970. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. Annals of Mathematical Statistics, 41: 164–171.

Bruintjes, R., and Radford, A. N. 2013. Context-dependent impacts of anthropogenic noise on individual and social be- haviour in a cooperatively breeding fish. Animal Behaviour, 85: 1343–1349.

Carey, W. M. 2009. Lloyd's mirror-image interference effects. Acoustics Today, 5: 14–20.

Castellote, M., Clark, C. W., and Lammers, M. O. 2012. Acoustic and behavioural changes by fin whales (Balaenoptera physalus) in re- sponse to shipping and airgun noise. Biological Conservation, 147: 115–122.

Edmonds, N. J., Firmin, C. J., Goldsmith, D., Faulkner, R. C., and Wood, D. T. 2016. A review of crustacean sensitivity to high am- plitude underwater noise: data needs for effective risk assessment in relation to UK commercial species. Marine Pollution Bulletin, 108: 5–11.

Efron, B. 1981. Nonparametric estimates of standard error: the jack- knife, the bootstrap and other methods. Biometrika, 68: 589–599.

Feroze, K., Sultan, S., Shahid, S., and Mahmood, F. 2018. Classification of underwater acoustic signals using multi-classifiers. In Applied Sciences and Technology (IBCAST), January 2018. 15th International Bhurban Conference on IEEE, pp. 723–728.

Forney, G. 1973. The Viterbi algorithm. Proceedings of the IEEE, 61: 268–278.

Graham, A. L., and Cooke, S. J. 2008. The effects of noise disturbance from various recreational boating activities common to inland waters on the cardiac physiology of a freshwater fish, the large- mouth bass (Micropterus salmoides). Aquatic Conservation: Marine and Freshwater Ecosystems, 18: 1315–1324.

Holles, S., Simpson, S. D., Radford, A. N., Berten, L., and Lecchini, D. 2013. Boat noise disrupts orientation behaviour in a coral reef fish. Marine Ecology Progress Series, 485: 295–300.

Huang, W., Wang, D., Garcia, H., Godø, O., and Ratilal, P. 2017. Continental shelf-scale passive acoustic detection and characteri- zation of diesel–electric ships using a coherent hydrophone array. Remote Sensing, 9: 772.

Hyder, K., Weltersbach, M. S., Armstrong, M., Ferter, K., Townhill, B., Ahvonen, A., Arlinghaus, R. et al. 2018. Recreational sea fish- ing in Europe in a global context—participation rates, fishing ef- fort, expenditure, and implications for monitoring and assessment. Fish and Fisheries, 19: 225–243.

Jelinek, F. 1976. Continuous speech recognition by statistical meth- ods. Proceedings of the IEEE, 64: 532–556.

Jelinek, F., Bahl, L., and Mercer, R. 1975. Design of a linguistic statis- tical decoder for the recognition of continuous speech. IEEE Transactions on Information Theory, 21: 250–256.

Lagardere, J. P., and Mariani, A. 2006. Spawning sounds in meagre Argyrosomus regius recorded in the Gironde estuary, France. Journal of Fish Biology, 69: 1697–1708.

Lippmann, R. 1987. An introduction to computing with neural nets. IEEE ASSP Magazine, 4: 4–22.

Markus, T., and Sanchez, P. P. S. 2018. Managing and regulating un- derwater noise pollution. In Handbook on Marine Environment Protection, pp. 971–995. Ed. by M. Salomon and T. Markus. Springer, Cham.

Marley, S. A., Kent, C. P. S., Erbe, C., and Parnum, I. M. 2017. Effects of vessel traffic and underwater noise on the movement, behaviour and vocalisations of bottlenose dolphins in an urbanised estuary. Scientific Reports, 7: 13437.

McDermott, E., Iwamida, H., Katagiri, S., and Tohkura, Y. 1990. Shifttolerant LVQ and hybrid LVQ–HMM for phoneme recogni- tion. In Readings in Speech Recognition, pp. 425–438. Ed. by A. Waibel and K.-F. Lee. Morgan Kaufmann, San Francisco, CA.

McDonald, M. A., Hildebrand, J. A., and Wiggins, S. M. 2006. Increases in deep ocean ambient noise in the Northeast Pacific west of San Nicolas Island, California. Journal of the Acoustical Society of America, 120: 711–718.

Nedelec, S. L., Simpson, S. D., Morley, E. L., Nedelec, B., and Radford, A. N. 2015. Impacts of regular and random noise on the behaviour, growth and development of larval Atlantic cod (Gadus morhua). Proceedings of the Royal Society B, 282: 20151943.

Normandeau Associates, Inc. 2012. Effects of noise on fish, fisheries, and invertebrates in the U.S. Atlantic and Arctic from energy in- dustry sound-generating activities. A workshop report prepared under Contract No. M11PC00031 for the Bureau of Ocean Energy Management, US Department of the Interior.

Ogden, G. L., Zurk, L. M., Jones, M. E., and Peterson, M. E. 2011. Extraction of small boat harmonic signatures from passive sonar. The Journal of the Acoustical Society of America, 129: 3768–3776. O’Shaughnessy, D. 1987. Speech Communication: Human and Machine Addison-Wesley Series in Electrical Engineering, pp.204–211. IEEE Press, Reading, MA.

Pace, F., White, P., and Adam, O. 2012. Hidden Markov modeling for humpback whale (Megaptera Novaeanglie) call classification. In Proceedings of Meetings on Acoustics ECUA2012, 17. ASA. 070046 pp.

Pereira, B. 2019. Acoustic repertoire of the meagre Argyrosomus re- gius: the influence of development, gender and context in tempo- ral and spectral variation of calls. MSc thesis, University of Algarve.

Picciulin, M., Sebastianutto, L., Codarin, A., Calcagno, G., and Ferrero, E. A. 2012. Brown meagre vocalization rate increases dur- ing repetitive boat noise exposures: a possible case of vocal com- pensation. The Journal of the Acoustical Society of America, 132: 3118–3124.

Pollara, A., Sutin, A., and Salloum, H. 2017. Passive acoustic methods of small boat detection, tracking and classification. In Technologies for Homeland Security (HST), April 2017. IEEE International Symposium on IEEE, pp. 1–6.

Pollock, K. H., Jones, C. M., and Brown, T. L. 1994. Angler Survey Methods and Their Applications in Fisheries Management. American Fisheries Society Special Publication 25, Bethesda, MD.

Prista, N. 2014. Argyrosomus regius (Asso, 1801) fishery and ecology in Portuguese waters, with reference to its relationships to other European and African populations. Doctoral dissertation, University of Lisbon.

Putland, R. L., Montgomery, J. C., and Radford, C. A. 2018a. Ecology of fish hearing. Journal of Fish Biology, 95: 39–52.

Putland, R. L., Ranjard, L., Constantine, R., and Radford, C. A. 2018b. A hidden Markov model approach to indicate Bryde’s whale acoustics. Ecological Indicators, 84: 479–487.

Rabiner, L. R. 1989. A tutorial on hidden Markov models and se- lected applications in speech recognition. Proceedings of the IEEE, 77: 257–286.

Ranjard, L., Reed, B. S., Landers, T. J., Rayner, M. J., Friesen, M. R., Sagar, R. L., and Dunphy, B. J. 2017. MatlabHTK: a simple inter- face for bioacoustic analyses using hidden Markov models. Methods in Ecology and Evolution, 8: 615–621.

Reynolds, D. A., and Rose, R. C. 1995. Robust text-independent speaker identification using Gaussian mixture speaker models. IEEE Transactions on Speech and Audio Processing, 3: 72–83.

Rolland, R. M., Parks, S. E., Hunt, K. E., Castellote, M., Corkeron, P. J., Nowacek, D. P., Wasser, S. K. et al. 2012. Evidence that ship noise increases stress in right whales. Proceedings of the Royal Society B: Biological Sciences, 279: 2363–2368.

Scheifele, P. M., Johnson, M. T., Fry, M., Hamel, B., and Laclede, K. 2015. Vocal classification of vocalizations of a pair of Asian Small-Clawed otters to determine stress. The Journal of the Acoustical Society of America, 138: EL105–109.

Simard, P., Wall, K. R., Mann, D. A., Wall, C. C., and Stallings, C. D. 2016. Quantification of boat visitation rates at artificial and natu- ral reefs in the eastern Gulf of Mexico using acoustic recorders. PLoS One, 11: e0160695.

Slamnoiu, G., Radu, O., Rosca, V., Pascu, C., Damian, R., Surdu, G., and Radulescu, A. 2016. DEMON-type algorithms for determina- tion of hydro-acoustic signatures of surface ships and of divers. In IOP Conference Series: Materials Science and Engineering, 145. IOP Publishing, 082013 pp.

Somervuo, P., Harma, A., and Fagerlund, S. 2006. Parametric repre- sentations of bird sounds for automatic species recognition. IEEE Transactions on Audio, Speech, Language Process, 14: 2252–2263.

Traverso, F., Gaggero, T., Rizzuto, E., and Trucco, A. 2015. Spectral analysis of the underwater acoustic noise radiated by ships with controllable pitch propellers. In OCEANS 2015-Genova, pp. 1–6. IEEE.

Trevorrow, M. V., Vasiliev, B., and Vagle, S. 2008. Directionality and maneuvering effects on a surface ship underwater acoustic signa- ture. The Journal of the Acoustical Society of America, 124: 767–778.

Urick, R. J. 1983. Principles of Underwater Sound, 3rd edn. McGraw-Hill, New York, NY.

Vahidpour, V., Rastegarnia, A., and Khalili, A. 2015. An automated approach to passive sonar classification using binary image features. Journal of Marine Science and Application, 14: 327–333.

Vasconcelos, R. O., Amorim, M. C. P., and Ladich, F. 2007. Effects of ship noise on the detectability of communication signals in the Lusitanian toadfish. Journal of Experimental Biology, 210: 2104–2112.

Vasconcelos, R. O., Sisneros, J., Amorim, M. C. P., and Fonseca, P. J. 2011. Auditory saccular sensitivity of the vocal Lusitanian toad- fish: low frequency tuning allows acoustic communication throughout the year. Journal of Comparative Physiology A, 197: 903–913.

Veirs, S., Veirs, V., and Wood, J. D. 2016. Ship noise extends to fre- quencies used for echolocation by endangered killer whales. PeerJ, 4: e1657.

Vieira, M., Fonseca, P. J., Amorim, M. C. P., and Teixeira, C. J. C. 2015. Call recognition and individual identification of fish vocal- izations based on automatic speech recognition: an example with the Lusitanian toadfish. Journal of the Acoustical Society of America, 138: 3941–3950.

Vieira, M., Pereira, B. P., Pous~ao-Ferreira, P., Fonseca, P. J., and Amorim, M. 2019. Seasonal variation of captive meagre acoustic signalling: a manual and automatic recognition approach. Fishes, 4: 28.

Voellmy, I. K., Purser, J., Flynn, D., Kennedy, P., Simpson, S. D., and Radford, A. N. 2014. Acoustic noise reduces foraging success in two sympatric fish species via different mechanisms. Animal Behaviour, 89: 191–198.

Watts, R. D., Compton, R. W., McCammon, J. H., Rich, C. L., Wright, S. M., Owens, T., and Ouren, D. S. 2007. Roadless space of the conterminous United States. Science, 316: 736–738.

Yang, S., Li, Z., and Wang, X. 2000. Vessel radiated noise recognition with fractal features. Electronics Letters, 36: 923–925.

Young, S., and Bloothooft, G. (Ed.) 1997. Corpus-based Methods in Language and Speech Processing, 2. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X., Moore, G. et al. 2006. The HTK Book (for HTK Version 3.4), pp. 1–359. Cambridge University Press, Cambridge, UK.

Yu, H. J., and Oh, Y. H. 1997. A neural network for 500 vocabulary word spotting using acoustic sub-word units. In Acoustics, Speech, and Signal Processing, 1997. ICASSP-97, 1997 IEEE International Conference on IEEE, 4, pp. 3277–3280.

Zak, A. 2008. Ships classification basing on acoustic signatures. WSEAS Transactions on Signal Processing, 4: 137–149.

Zhu, C., Garcia, H., Kaplan, A., Schinault, M., Handegard, N., Godø, O., Huang, W. et al. 2018. Detection, localization and classifica- tion of multiple mechanized ocean vessels over continental-shelf scale regions with passive ocean acoustic waveguide remote sensing. Remote Sensing, 10: 1699.

Table 1 Examples of relevant articles on recognition and detection of marine vessels through the underwater noise produced.

	Objective	System	Feature	Reference
Extraction of small boat harmonic signatures from passive sonar	c		HEAT	Ogden *et al.*, 2011
DEMON-type algorithms for determination of hydro-acoustic signatures of surface ships and of divers	c	-	DEMON	Slamnoiu *et al.*, 2016
Ship noise extends to frequencies used for echolocation by endangered killer whales	c	-	-	Veirs *et al.* 2016
Passive acoustic methods of small boat detection, tracking and classification	c	-	DEMON	Pollara *et al.* 2017
Continental shelf-scale passive acoustic detection and characterization of diesel-electric ships using a coherent hydrophone array.	c, D	-	POAWRS	Huang *et al.* 2017
Detection, Localization and Classification of Multiple Mechanized Ocean Vessels over Continental-Shelf Scale Regions with Passive Ocean Acoustic Waveguide Remote Sensing	c, D	-	POAWRS	Zhu *et al.* 2018
Quantification of Boat Visitation Rates at Artificial and Natural Reefs in the Eastern Gulf of Mexico Using Acoustic Recorders	D	**	[b]	Simard *et al.*, 2016
Ships classification basing on acoustic signatures	I(5)	ANN		Zak 2008
Acoustic detection and classification of river boats	T(2)	LDA CART		Averbuch *et al.*, 2010
An Automated Approach to Passive Sonar Classification Using Binary Image Features	T(4)*	ANN		Vahidpour *et al.*, 2015
Vessel radiated noise recognition with fractal features	T(6)	***	[a]	Yang 2000

ANN, artificial neural network; c, ship noise characterization; CART, Classification and Regression Trees; D, boat detection with no categorization; DEMON, Detection of Envelope Modulation on Noise algorithm; HEAT, Harmonic Extraction and Analysis Tool; LDA, Linear Discriminant Analysis; I(n), individual ship recognition system with n different ships; POAWRS, Passive ocean acoustic waveguide remote sensing technique using an array of hydrophones; T(n), marine vessel type recognition system with n categories; [a], Fractional Brownian motion feature and Fractal dimension feature; [b], to each sound was calculated the FFT average (fast Fourier transform, to produce an averaged power spectrum of file), the peak identification (to identify harmonics typical of boat noise within averaged power spectrum), and the amplitude threshold; *distinction of boat and ships (with weight of 1 248, 2 592, 3 660 t and 35 573 tons); ** The algorithm operated using five steps: median filter, band-pass filter, FFT average, peak identification, and amplitude threshold to determine if the overall root mean-square amplitude of the 10-second acoustic file was a threshold level above that of surrounding files.*** Fractal dimension features.

Table 2 Different types of boat recorded at the port and marinas of the Öresund strait (Sweden); according to shape of the boat, hull material, type of engine, and number of engines. We defined boats as all small vessel for travelling on water, propelled by an engine. The term "vessels" was used to include boats, ferries and ships.

Type of boat	**Hull material**	**Size**	**Type of Engines**	**Number engines**	**Recreational Fishing**	**Number of boats**		**Duration**
E.g., Recreational: sail, yacht, open deck; Commercial: fishing, cruise; ferry, other	E.g., wood, metal, other?	In meters	Inboard (i), outboard (o)	e.g., 1 or 2, unknown (?)	Can be used on recreational fishing?	Number of boats passages sounds recorded without overlap	Number of boats recorded	Approximate range of sound durations recorded.
	metal		i	1	No			
commercial fishing boat	plastic	10-15	i	1	No	8	8	40 s – 3 min
	wood		i	1	No			
recreational fishing tourboat	wood	15-20	i	?(1)	Yes	3	3	35 s – 3 min
	plastic		i	1	Yes			
	plastic		o	1	Yes			
open deck private boats	plastic	7-12	o	2	Yes	55	49	40 s – 2 min
	aluminium		o	1	Yes			
open deck private boats	plastic	7-12	i	1	Yes	4	4	40 – 60 s
RIBs	plastic	5-10	o	1	No	13	11	25 – 90 s
	plastic		o	2	No			
sail boat	plastic	10-20	i	1	No	55	53	1 – 2 min
sail boat	plastic	10-20	o	1	No	11	11	1 – 3 min
jetski	plastic	3-4	i	1	No	5	3	30 s – 1 min
small yacht	plastic	7-12	i	1	Yes	25	23	40 s – 2 min
small yacht	plastic	7-12	o	1	Yes	9	9	40 s – 2 min
double ender boat	plastic	7-12	i	1	Yes	9	9	30 s – 2 min
medium to large yacht	plastic	12-30	i	?(1)	Yes	11	11	30 s – 2 min

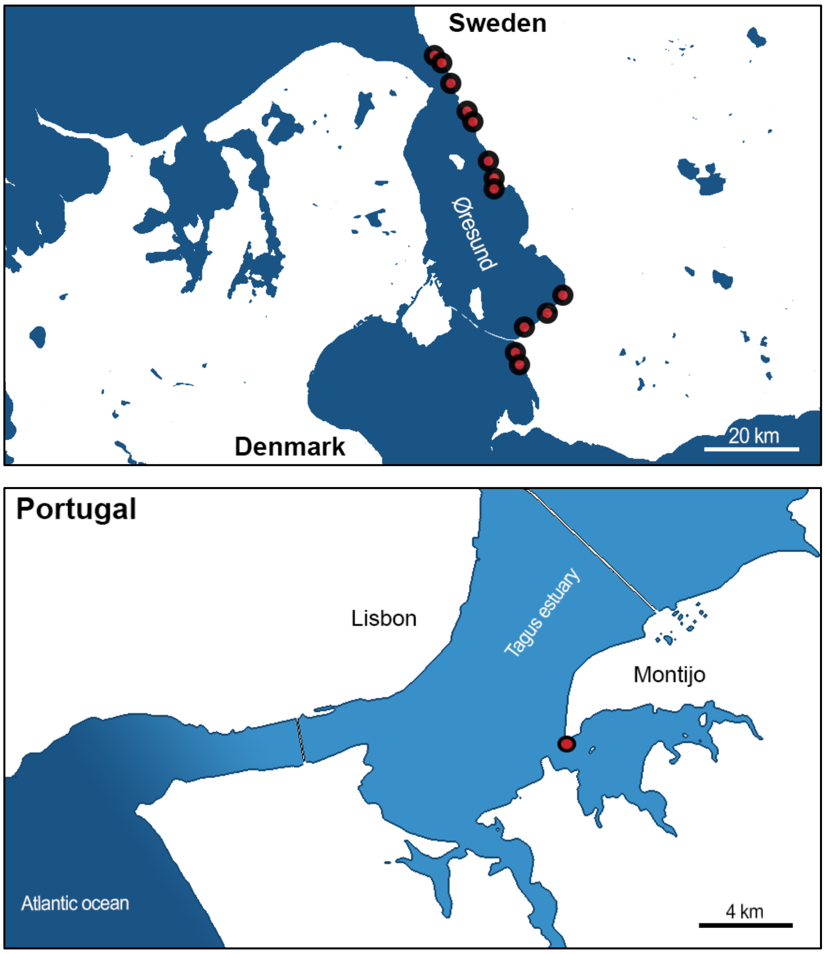

Figure 1. Recording locations: (1) the several marinas across Öresund strait (Sweden); and (2) the passive acoustic monitoring station in Tagus estuary (Portugal).

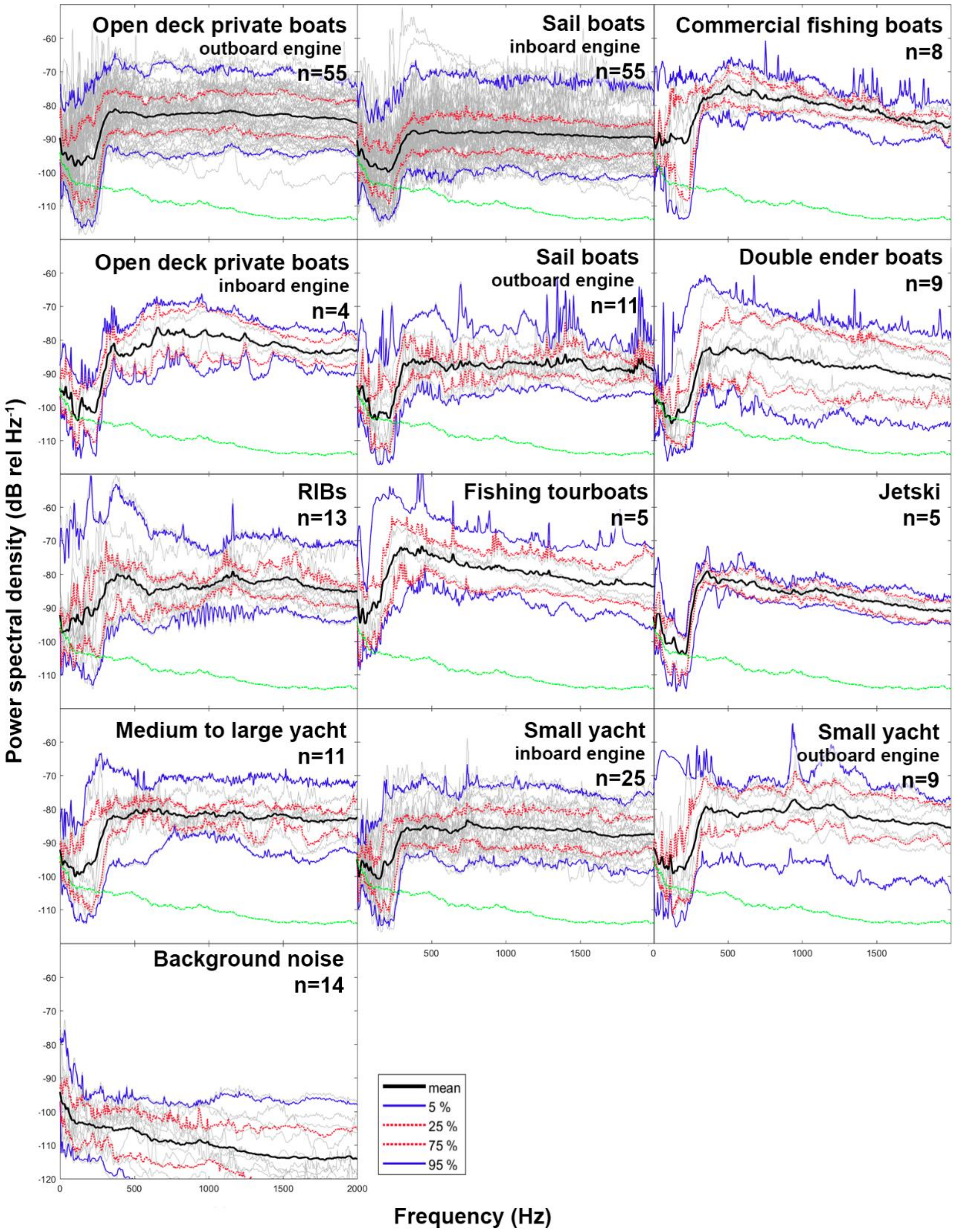

Figure 2. Power spectral density (PSD) of boat noises and background noise received levels for the full sampled period from 4 to 7 of July 2017 on several marinas of the Öresund strait (Sweden). The black line represents the mean power spectral density (averaging of dB values) and the blue and red lines depict 5, 25, 75 and 95 percentiles. The green line represents the mean power spectral density of the background noise. PSDs were calculated with the Welch's power spectral density estimate algorithm on MATLAB using a frequency bandwidth up to 2000 Hz (1024 point FFT). We defined boats as all small vessels for travelling on water, propelled by an engine.

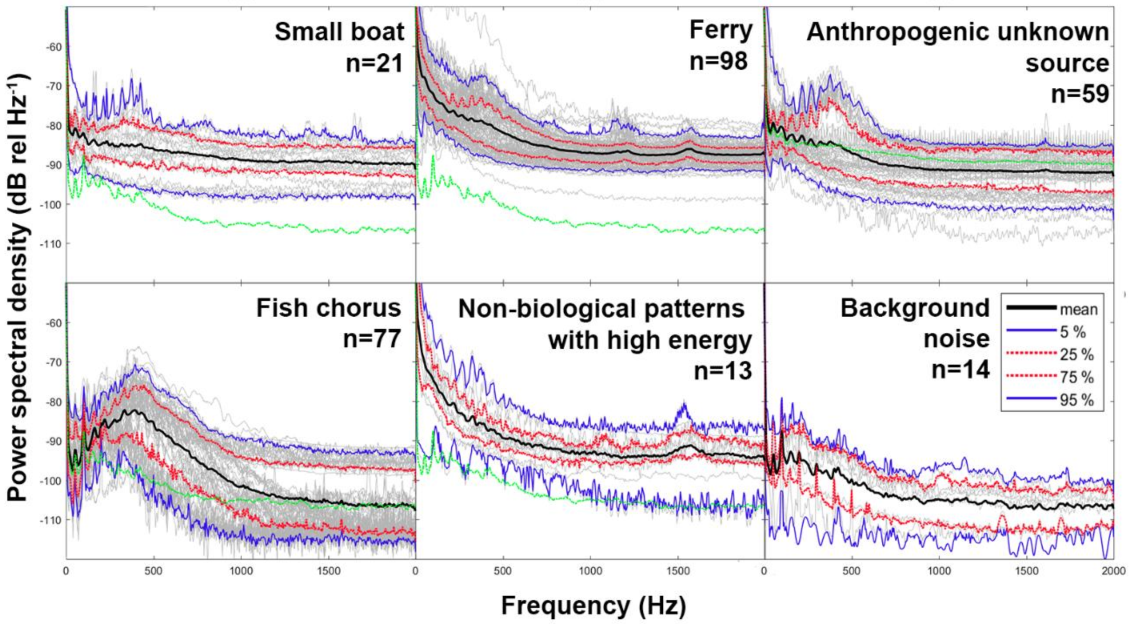

Figure 3. Power spectral density (PSD) of received levels of marine vessels noises, biological sounds and background noise measured as full bandwidth for the data set consisted of ca. 2 day round-the-clock recordings of the sounds from 15 to 16 of May 2017, in the Tagus estuary (Portugal). The black line represents mean power spectral density (averaging of dB values) with blue and red lines depicting 5, 25, 75 and 95 percentiles. The green line represents the mean power spectral density of the background noise. PSDs were calculated with the Welch's power spectral density estimate algorithm on MATLAB using a frequency bandwidth up to 2000 Hz (1024 point FFT). The term "vessels" was used to include boats, ferries and ships.

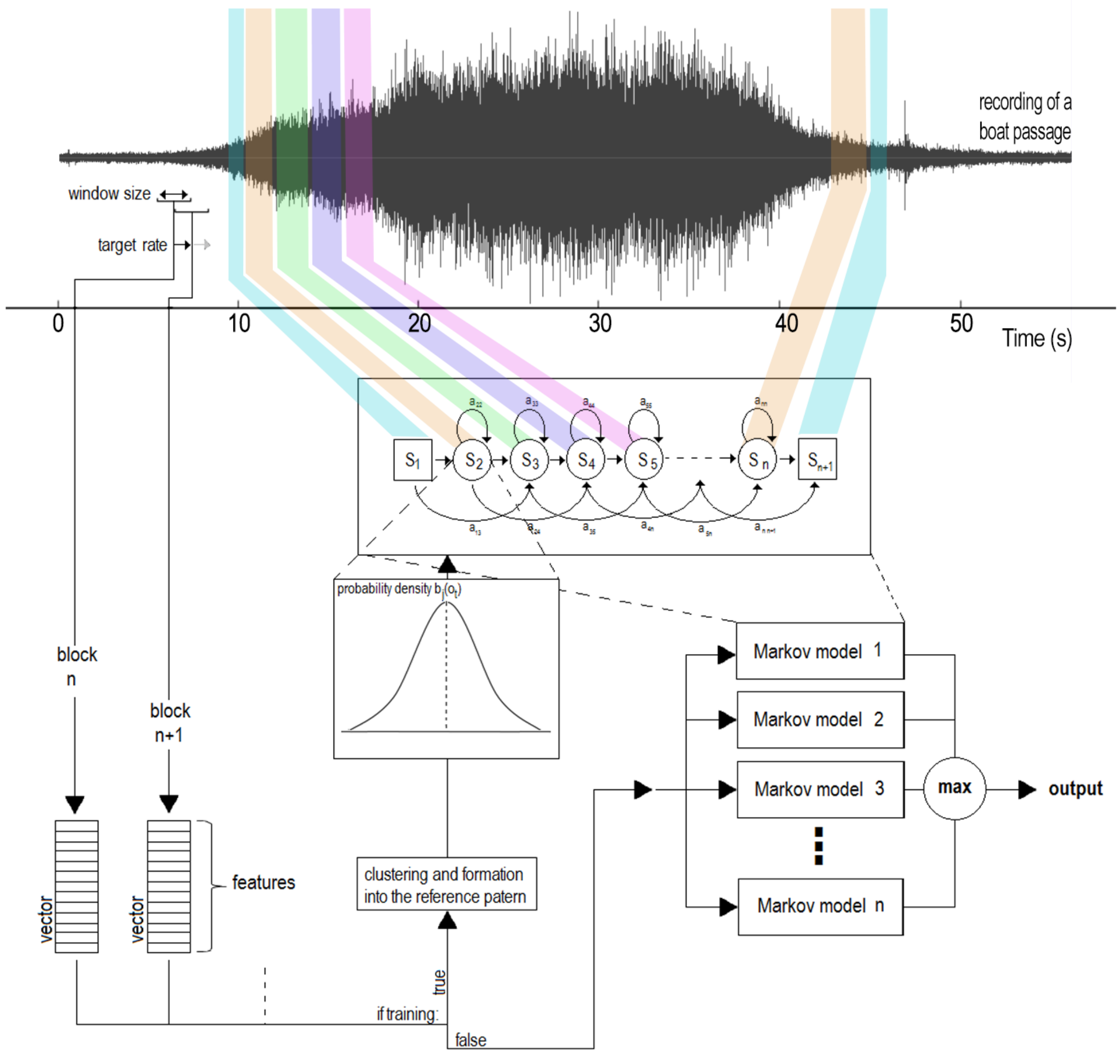

Figure 4. Workflow of the HMM recognition system using the HMM ToolKit (HTK, diagram based on Young et al., 2006). The use of Markov models for classification of acoustic signals in the time domain is naturally associated with linear topologies. Each state in a HMM can then be compared to a human language phoneme. Each word, as each phoneme, has an average expected duration that is directly related to the number of states. However, because we do not have a phoneme set for vessels noise, we assumed a window of 200 ms and a high number of states to represent the boat noises. The probability of sound being represented by each Markov model (representing each sound type) is calculated as the product of the transition probabilities and the output probabilities (extracted from the probability density of each state). However, in practice, only the observation sequence is known and the underlying state sequence is hidden. The signal represents an oscillogram of a boat noise.

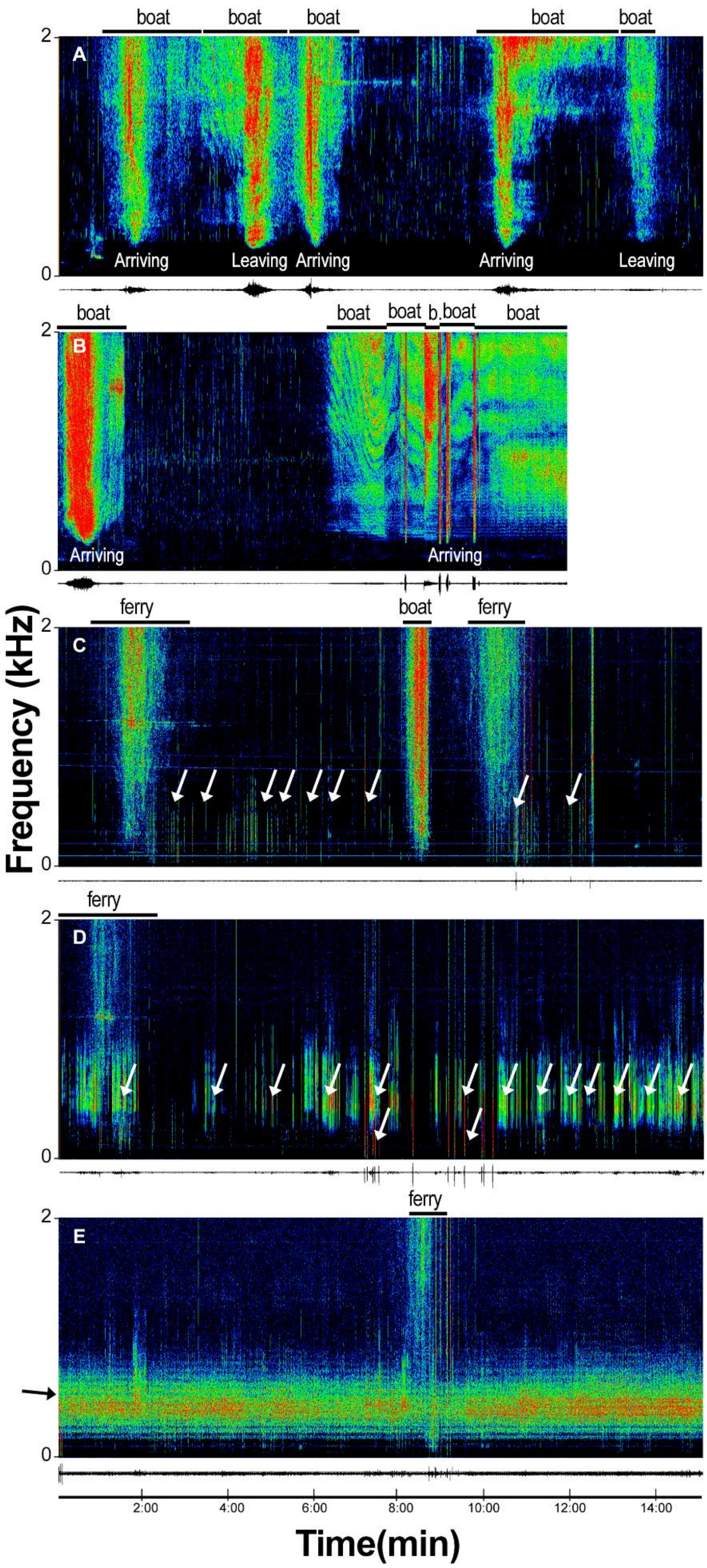

Figure 5. During the present study we recorded ca. 18 hours across Öresund strait (Sweden) and ca.144 hours in the Tagus estuary (Portugal). Spectrograms (FFT 1024 points) and oscillograms illustrate marine vessels noises. Horizontal black bars at the

top of each spectrogram show examples of the output given by the automatic recognition systems. (A,B) show the results of boat noise automatic recognition system using boats noises recorded at several marinas across Öresund strait (Sweden) using a 0-10000 Hz frequency bandwidth; (B) also represents an example of one boat noise segmented due to manoeuvres using the engine. (C,D,E) illustrates the results of the marine vessel recognition system using sounds recorded by a passive acoustic monitoring station in Tagus estuary (Portugal); Lloyd’s mirror effect was evident on most ferries’ recordings (see e.g. first ferry sound on C). We encountered choruses produced by fish species, namely Lusitanian toadfish (C and D; Amorim et al., 2008), and Meagre's series of isolated pulses (D; Pereira, 2019) and long grunts (E; Lagardère et Mariani, 2006). Arrows point to the presence of biological sounds.

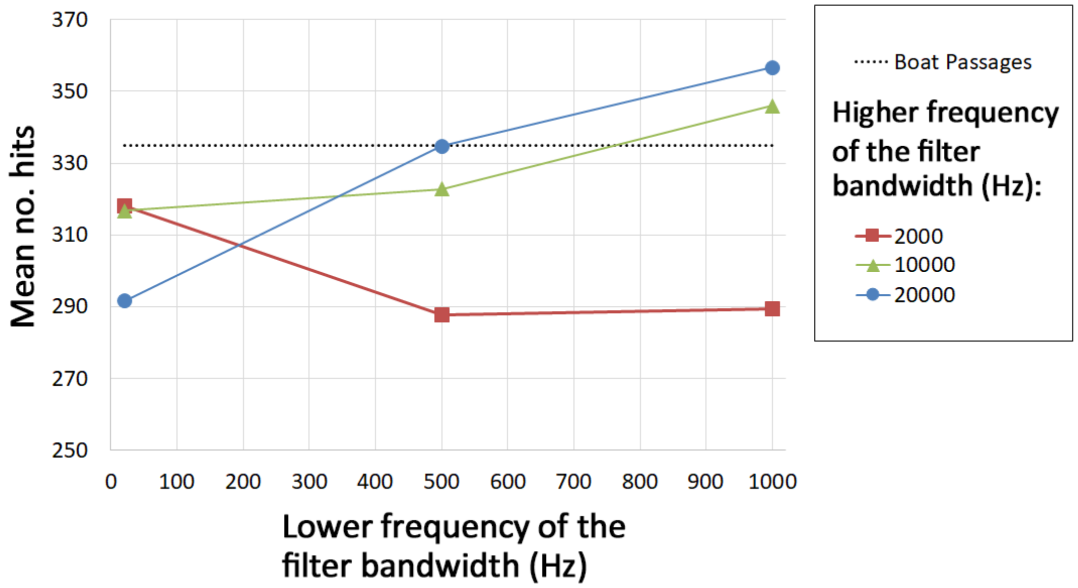

Figure 6. Mean number of hits of hidden Markov model recognition systems computed on the MFC with cepstrum, delta, acceleration coefficients, and 9 different frequency bandwidths for Öresund strait (Sweden). Each mean represents 100 iterations using 20 boat sounds randomly selected from the dataset in each training set. Overall depicted data consider circa 20 hours of continuous recordings. Boat passages represent the number of boats that passed by the entrance of the marina during the recorded period confirmed by visual observations.

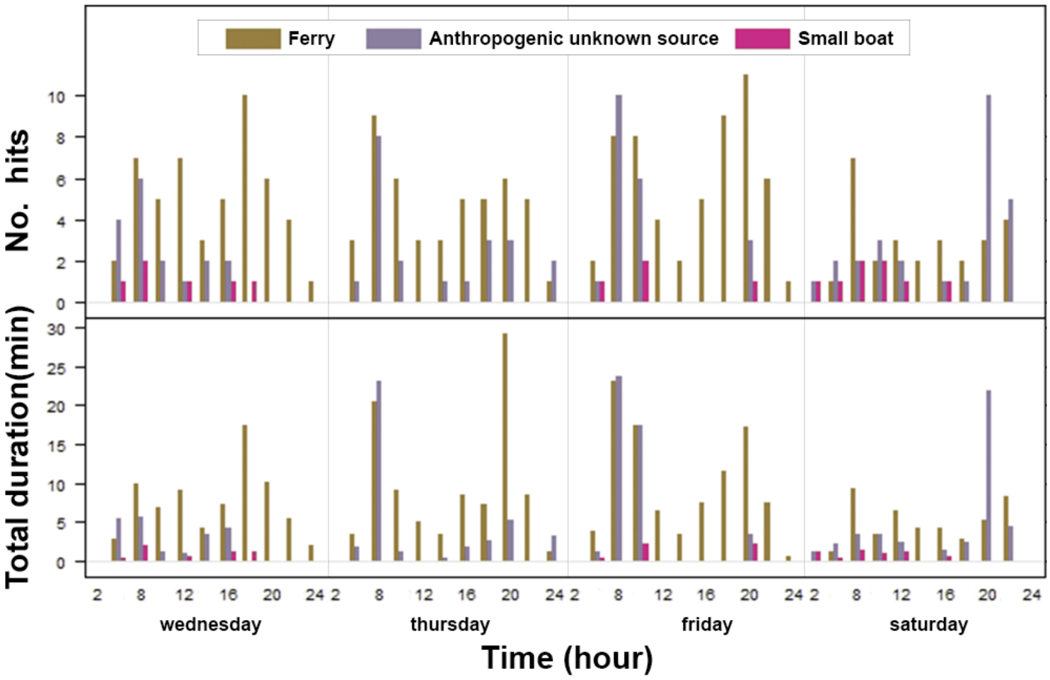

Figure 7. Presence of marine vessels at the passive acoustic monitoring station in the Tagus estuary (Portugal) from 17 to 20 of May 2017 (Wednesday to Saturday), estimated using the automatic recognition system: (A) shows the number of hits per 2 hours; (B) represents the total time per 2 hours where a marine vessel sound was detected. Each bar represents a 2 hour period. Note that a vessel that stays near the recording place and or change significantly the engine power could cause a higher number of hits and an overestimation of the number of vessels while the overlapping of noise from two different boats could cause an underestimation.

Table 3. Mean confusion matrix from the hidden Markov model classification computed on the MFC with cepstrum, delta and acceleration coefficients with a frequency bandwidth of 1200 Hz to 2000 Hz, for the Tagus estuary. The model was trained with 142 boat sounds from the first two recorded days, and tested with the remaining four days (a total of 96 hours with 286 boat sounds). 90,87 ± 8,17 % (mean identification rate of the four days ± SD) of tested sounds were correctly classified, with an accuracy with the same value.

	Predicted group membership				
Boat/Vessel noise type	Small boat	Ferry	AUS	False negative	identification rate (%)
Small boat	4	2	0	0	66,7
Ferry	0	40	2	0	95,2
Anthropogenic unknown source (AUS)	0	3	19	1	86,4
False positive	0	0	0		
Overall mean					**90,87 ± 8,17**