Canadian Journal of Fisheries and Aquatic Sciences

Discrimination between schools and submerged trees in reservoirs: A preliminary approach using narrowband and broadband acoustics

Journal:	<i>Canadian Journal of Fisheries and Aquatic Sciences</i>
Manuscript ID	cjfas-2021-0087.R2
Manuscript Type:	Article
Date Submitted by the Author:	25-Sep-2021
Complete List of Authors:	Blanluet, Arthur; The University of Queensland, The Mathematical Marine Ecology Lab; INRAE, Univ. Savoie Mont Blanc, INRAE, CARRETEL, Pôle R\&D ECLA (ECosystèmes LAcustres) OFB – INRAE – USMB Gastauer, Sven; Thünen-Institute of Sea Fisheries; University of California San Diego Scripps Institution of Oceanography Cattanéo, Franck; University of Applied Sciences and Arts Western Switzerland, Hepia Geneva, Land, Nature, Environment Institute Goulon, Chloé; INRAE, Univ. Savoie Mont Blanc, INRAE, CARRETEL, Pôle R\&D ECLA (ECosystèmes LAcustres) OFB – INRAE – USMB Grimardias, David; University of Applied Sciences and Arts Western Switzerland, Hepia Geneva, Land, Nature, Environment Institute Guillard, Jean; INRAE, Univ. Savoie Mont Blanc, INRAE, CARRETEL, Pôle R\&D ECLA (ECosystèmes LAcustres) OFB – INRAE – USMB
Keyword:	Fisheries acoustics, Immersed trees, FRESHWATER < Environment/Habitat, Broadband acoustic, Random forest
Is the invited manuscript for consideration in a Special Issue? :	Not applicable (regular submission)

SCHOLARONE™
Manuscripts

Discrimination between schools and submerged trees in reservoirs: A preliminary approach using narrowband and broadband acoustics

Arthur Blanluet^{1,2,3*} Sven Gastauer^{4,5} Franck Cattaneo⁶ Chloe Goulon^{1,2} David Grimardias⁶ Jean Guillard^{1,2}

¹ Univ. Savoie Mont Blanc, INRAE, CARTELE, 74200 Thonon-les-Bains

² Pôle R&D ECLA (ECosystèmes LAcustres) OFB – INRAE – USMB

³ The University of Queensland, The Mathematical Marine Ecology Lab School of Mathematics and Physics, Level 2, Physics Annexe, The University of Queensland, Saint Lucia, Brisbane, QLD, AUS 4072

⁴ Thünen-Institute of Sea Fisheries, Bremerhaven, Germany

⁵ Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, United States

⁶ University of Applied Sciences and Arts Western Switzerland - Hepia Geneva, Land, Nature, Environment Institute, Route de Presinge 150, CH-1254, Jussy, Switzerland

* Corresponding author: arthur.blanluet0@gmail.com

Abstract

With a growing demand for hydroelectric energy, the number of reservoirs is dramatically increasing worldwide. These new water bodies also present an opportunity for the development of fishing activities. However, these reservoirs are commonly impounded on uncut forests, resulting in many immersed trees. These trees hinder fish assessments by disrupting both gill netting and acoustic sampling. Immersed trees can easily be confused with fish schools on echograms. To overcome this issue, we developed a method to discriminate fish schools from immersed trees. A random forest algorithm was used to classify echo-traces at 120 and 200 kHz, recorded by an EK80 (SIMRAD) in narrowband (Continuous Wave) and in broadband mode (Frequency Modulated). We obtained a good discrimination rate between trees and schools, especially in broadband (90 % ratio of good classification). We demonstrate that it is possible to discriminate fish schools from immersed trees and thus facilitate the use of fisheries acoustics in reservoirs.

keywords Fisheries acoustics, immersed trees, freshwater, broadband acoustic, random forest

1 Introduction

Fisheries acoustics are now a standard method to estimate fish abundances in lakes and reservoirs and to describe their relationships with ecosystems [Rudstam et al., 2003, Parker-Stetter et al., 2009, Pollom and Rose, 2016, Draščík et al., 2017]. Fisheries acoustics can provide a non-intrusive and reliable estimate of fish biomass, size distribution and spatial distributions in a relatively short period [Rudstam et al., 2012]. Despite these advantages, fisheries acoustics remains under-used in reservoirs, notably in tropical environments [Tessier et al., 2016]. Recently, the number of artificial lakes and associated dams significantly increased globally, and is expected to continue to increase in the coming decades [Zarfl et al., 2015]. This increase is mainly caused by an increase in demand for energy and as a response to the transition towards sustainable energy sources. Besides the benefits in terms of energy generation, these artificial aquatic ecosystems can present an important source of animal proteins for the local population through fishing, most notably in Asia [Tessier et al., 2016, 2020].

The widely distributed presence of immersed trees in reservoirs could heavily impair the monitoring of fish stock [Coll et al., 2007, Frouzová et al., 2008], especially in tropical areas, where the forests were generally not cut down before the flooding. Immersed trees largely limit navigation, constrain sampling to daylight hours for security reasons [Tessier et al., 2020], hinder the use of gill nets to sample fish (entanglement risks), and make acoustic sampling challenging: immersed trees appear similar to fish schools on echograms (Figure 1) and could easily be confused with them [Cardenas et al., 1988, Coll et al., 2007]. This similarity between trees and schools dramatically increase the needed time to clean and process echograms during the data processing [Coll et al., 2007, Frouzová et al., 2008]; as well as being an important source of uncertainty in stock estimate within reservoirs [Tessier et al., 2020]. Finding a way to discriminate immersed trees from fish schools on acoustic data is critical to develop acoustic in reservoir lakes and improve the fish stock assessment precision.

One of the oldest applications of fisheries acoustics is the classification of different types of “echo-traces” (coherent acoustic objects on an echogram, ICES, 2000, Fernandes, 2009), e.g., extracting fish schools from planktonic layers [Madureira et al., 1993, Jurvelius et al., 2008] or distinguishing between different fish species [Scalabrin et al., 1996, Getabu et al., 2003, Logerwell and Wilson, 2004, Horne, 2000]. Despite the technological evolution of fisheries acoustic, this classification of acoustic targets remains challenging [Korneliussen et al., 2018].

Despite the development of several echo-traces discriminant methods [Korneliussen et al., 2018], few studies have aimed to discriminate fish from trees in reservoirs [Coll and Lebourges Dhaussy, 2005, Josse and Le Squin, 2013]. These studies, based on single frequency acoustics school parameters [Reid et al., 2000], did not yield satisfactory results [Josse and Le Squin, 2013]. Josse and Le Squin [2013], concluding that the processing of at least two frequencies might improve the discrimination. Trees and fish are expected to present different scattering properties with varying frequency responses based on their differences in material proprieties (acoustic impedance) (gaseous for fish, fluid like or rigid for trees [Stanton et al., 1994]).

Additionally, the recent development of broadband echo-sounders (Frequency Modulated signal, Demer et al. [2017]) presents an opportunity to use the broadband spectrum to discriminate between different targets [Bassett et al., 2018, Blanluet et al., 2019] and to develop new broadband school parameters to classify echo-traces.

Our study investigated how to discriminate between immersed trees and schools in lakes using bi-frequency and broadband parameters. To do so, we constructed two Random Forest models [Breiman, 2001] to discriminate fish schools from trees, using the energetic parameters of echo-traces at 120 and 200 kHz center frequencies. The first model used parameters extracted from Continuous Wave

Figure 1: 120 kHz Frequency Modulated (FM) echograms of (a) schools in Lake Geneva and of (b) immersed trees in Bimont reservoir. The corrected lakebed detection is represented by a thick black line.

(CW) acoustic data, and a second model used parameters extracted from Frequency Modulated (FM) acoustic data. The success of the CW and FM models are discussed, and the main parameters contributing to the discrimination power were identified.

Three main hypotheses were tested as discriminant factors between echo-traces: (i) Do immersed trees show a higher variance of the returned energy, expressed as the Volume backscattering strength (S_v , $dB \text{ re } 1m^{-1}$, MacLennan et al., 2002), than fish schools due to the high contrast between the energetic response of the trunk compared to the branches? (ii) Do immersed trees and fish schools present different broadband acoustics spectra due to differences in their material properties [Stanton et al., 1994]? (iii) Do the amplitude Probability Density Function (PDF, i.e. the echo's energy distribution within an echo-trace) of fish schools will be closer to Rayleigh distribution [Scalabrin et al., 1996, Chu and Stanton, 2010] than trees one? To answer these questions, we sampled a series of lakes to gather echo-traces of trees and fish schools in different environments.

2 Materials and Methods

2.1 Study sites

Data were acquired at daylight (when fish are expected to be schooling, while being more dispersed at night [Girard et al., 2020]) in a selection of peri-alpine and reservoir lakes (Table 1). The peri-alpine lakes were sampled to record fish school echo-traces in zones known to be rich in fish schools [Yule et al., 2013, Anneville et al., 2017] and without any immersed trees. Reservoirs were sampled to obtain tree echo-traces in areas known to be rich in immersed trees.

Table 1: Lake and reservoir characteristics (area, mean depth, max depth) and sampling date.

Lakes	surface (km^2)	mean depth (m)	max depth (m)	Sampling date
Hydropower reservoirs				
Chancy-Pougny	0.264	6.35	9.9	2019-09-26
Bimont	0.73	27	55	2020-01-28
Génissiat	0.4	5	20	2020-03-04
Peri-alpine lakes				
Lake Geneva	580	154	310	2019-10-11 & 2020-06-24
Lake Bourget	44.5	85	145	2019-12-04

The fish school data were acquired in two French peri-alpine lakes, Lakes Geneva and Bourget, regularly monitored in the Observatory of LAkes (OLA, <http://www6.inra.fr/soere-ola> ©OLA-IS, AnaEE-France, INRAE Thonon-les-Bains, developed by Eco-Informatics ORE INRAE Team), with well known fish population [Rimet et al., 2020]. The fish schools in the alpine lakes are dominated by juvenile perch (*Perca fluviatilis*) with the occasional roach (*Rutilus rutilus*) [Guillard et al., 2006, 2011]. No specific biological sampling has been carried out within the present study, but being part of a long term observatory project (OLA; Rimet et al., 2020), these ecosystems are well studied and regularly assessed.

Due to their abundance and our knowledge of these lakes' fish populations, we assumed that the observed fish schools were mainly composed of young-of-the-

115 year (YoY) perch. The perch sampled in October in Lake Geneva and in December
116 in Lake Bourget have reached their asymptotic size [Guillard et al., 2006], around
117 a mean total length of 95 and 85 mm, respectively for Lake Geneva and Bourget
118 [Frantz et al., 2018]. In the same way, the June fish school acquisitions in Lake
119 Geneva were assumed to be mainly composed of YoY perch, smaller in this pe-
120 riod of the year [Guillard et al., 2006], with an expected mean total length of 40 mm.

121
122 The immersed trees were sampled in two run-of-river hydropower reservoirs
123 on the Rhône River: the Chancy-Pougny and the Génissiat Reservoirs, and on
124 the Cause River: the Bimont Reservoir. The position and the condition of the
125 immersed trees on the Rhône River reservoirs are known from other surveys [Gri-
126 mardias et al., 2017, Cattaneo et al., 2021] and have been inspected visually for
127 the presence of branches reaching the surface.

128
129 The trees in the studied reservoirs on the Rhône River resulted from recent
130 landslides, presenting a great variety of orientation, condition and species (mainly
131 ash, oak, poplar and willow). The targeted acoustic sampling was performed on
132 trees visible from the surface, on immobile echo-traces with a high backscattering.
133 Only a few fish species, mainly young Cyprinids, are known to school in the Rhône
134 Reservoirs [Grimardias et al., 2017]. As the position of the trees were spotted by
135 the landslide mark and by the emerging branches, we assumed that a misidentifi-
136 cation of a fish school as a tree's echo-traces was unlikely.

137
138 In the Bimont Reservoir, the position and the condition of the immersed trees
139 were identified from a 2018 satellite photo of the empty lake (T. Tormos, personal
140 communication). The reservoir was flooded in 1952, and the immersed pines are
141 old and stand upright, with few branches. As the Bimont Reservoir was emp-
142 tied from 2017 to May 2019 for dam renovations, very few fish remained inside.
143 Therefore, based on similar clues related to the Rhône, it was assumed that a
144 misidentification of trees and fish schools echo-traces was unlikely.

2.2 Data acquisition

The acoustic recordings were performed from small boats equipped with a SIMRAD (Simrad Kongsberg Maritime AS, Horten, Norway) EK80 WBT-Mini and two multiplexed split-beam transducers (ES120-7C and ES200-7CD) mounted on a pole fixed on the boat side, oriented vertically down and kept at a depth of 0.5 m below the waterline. We will refer to the ensemble WBT-Mini/ES120-7C as the “120 kHz channel” and the ensemble WBT-Mini/ES200-7CD as the “200 kHz channel”. The two channels were set to ping sequentially at 10 Hz, with an acquisition threshold of -80 dB. The 3-dB beamwidth, pulse length, transmitting power and bandwidth limits for the CW or FM pulses are presented in Table 2. The boat speed varied between 0.5 and 5 $m.s^{-1}$, due to navigational constraints. At each sampling site, the first round of acquisitions was carried out in CW mode, followed by a second acquisition in FM mode covering roughly the same area. Standard sphere calibrations using a 22 mm and a 38.1 mm tungsten carbide (WC) with 6% cobalt binder spheres, according to the standard protocol of Demer et al. [2015] and the manufacturer’s manual, were performed in November 2019 in Lake Geneva and in March 2020 in Lake Annecy.

Table 2: Echosounders settings.

Transducer	ES120-7C	ES120-7C	ES200-7CD	ES200-7CD
Signal	CW	FM	CW	FM
Ramping ^a	NA	Fast	NA	Fast
Frequency range (kHz)	120	95-150	200	170-260
Beamwidth (°)	7	8.8° at 95 kHz, 5.6° at 150 kHz ^c	7	8.3° at 170 kHz, 5.4° at 260 kHz ^c
Pulse duration (ms)	0.256 ^b	1.024 ^c	0.256 ^b	1.024 ^c
Transmit power (W)	75	75	75	75
Sphere diameter (mm) ^d	22	22	38.1	22

^a see Demer et al. [2017] for the ramping definition

^b Godlewska et al. [2011]

^c Demer et al. [2017]

^d the 22-mm sphere was used for the FM calibrations and the 120-kHz CW mode as the nulls in the spectrum were less constraining than the 38.1-mm sphere for the FM calibration. The 38.1-mm sphere was used for the 200-kHz CW mode because of the null in the spectrum at 200 kHz with the 22-mm sphere.

2.3 Data post-processing

After the conversion of the acoustic data from the “.RAW” format to the “.HAC” format [ICES, 2005] using the HERMES software [Trenkel et al., 2009, Doray et al., 2018], acoustic data were pre-processed through the MOVIES3D software [Trenkel et al., 2009]. A Fast Fourier Transform (FFT) [Demer et al., 2017] with a 5-kHz window size was applied to the FM data (resulting in an effective temporal resolution of 0.15 m). This spectral resolution in FM mode was chosen as a trade-off between a stable broadband spectrum and a good spatial resolution [Blanluet et al., 2019]. Before working at the echo-trace scale, the acoustic data was echo-integrated (EI) into 0.5 m of depth by two horizontal ping bins, in FM and CW mode. This echo-integration aimed to compensate for the small shift between the two channels due to the sequential pinging and to reduce the backscattering variability while keeping a good resolution. Lakebed detections were obtained automatically and corrected manually to improve the separation between trees or schools and the lake bed (black lines in Figure 1).

2.4 Echo-traces extraction

Echo-traces were defined as structures observed within 12 or more contiguous EI cells above an S_v threshold of -62 dB on the 120 kHz and 200 kHz channels. These thresholds were chosen to better distinguish the echo-traces from the background (Figure 2) and were based on our empirical knowledge of the expected trees [Coll et al., 2007] and fish schools [Guillard et al., 2006, Grimardias et al., 2017] backscattering properties. Each identified echo-trace was labeled manually as immersed trees or fish schools in their corresponding data set.

2.5 Data analysis

According to Josse and Le Squin [2013], morphological and positional parameters were not sufficiently informative to discriminate trees from schools. Thus, based on fish school parameters available [Reid et al., 2000, Gastauer et al., 2017], we

Figure 2: Echo-integration and echo-traces extraction of a Lake Bourget FM dataset subset. (a): 120 and 200 kHz frequencies MOVIES3D echograms with -62 dB threshold; (b): 120 and 200 kHz frequencies echo-integrated echograms (0.5 m height and 2 pings width cells); (c): echo-traces extraction from the background.

selected a set of six (CW mode) or seven (FM mode) schools energetic parameters [Reid et al., 2000], for both frequencies 120 and 200 kHz.

We classed these energetic parameters in three categories, that each corresponding at one of the tested hypotheses. The mean backscattering volume (120/200 kHz mean S_v , $dB \text{ re } 1m^{-1}$), the difference between maximum and minimum backscattering volume (120/200 kHz max S_v - min S_v , $dB \text{ re } 1m^{-1}$) and the backscattering volume standard deviation (120/200 kHz S_v standard deviation, $dB \text{ re } 1m^{-1}$), were selected to investigate the variance difference hypothesis between tree and fish school echo-traces. The mean difference between 120 kHz and 200 kHz backscattering volume ($\Delta MVBS$, $dB \text{ re } 1m^{-1}$) and the median frequency spectrum slope parameters (120/200 kHz median frequency spectrum slope, $s_{m120kHz}$ and $s_{m200kHz}$, calculated in FM mode only) were selected to investigate the difference in frequency response between tree and fish school echo-traces. The skewness, an indicator of the S_v distribution symmetry [Lawson, 2001, Gastauer et al., 2017] (120/200 kHz S_v skewness), and the kurtosis, an indicator of the S_v distribution tail length [Lawson, 2001] (120/200 kHz S_v kurtosis), were selected to investigate if the amplitude PDF (Probability Density Function) of fish school and tree echo-traces could be differentiated. As the selected parameters were designed for CW mode, we calculated them in FM mode on an average in natural scale over the 120 or 200 kHz bandwidth [Blanluet et al., 2019]. The median frequency spectrum slope parameters (120/200 kHz median frequency spectrum slope, $s_{m120kHz}$ and $s_{m200kHz}$) were calculated in FM mode only, through a linear regression on the echo-traces EI cells median frequency spectrum at 120 and 200 kHz.

2.6 Statistical analyses

Various statistical classification methods have been used to discriminate between echo-traces [Korneliussen et al., 2018], as classification trees [Fernandes et al., 2006, Fernandes, 2009, Escobar-Flores et al., 2019] or random forest [D'Elia et al. [2014], Fallon et al. [2016], Proud et al. [2020]. The use of ensemble tree methods

(repetition of regression/classification tree method), as Random Forest or boosted trees [Breiman, 2001], obtains high classification rates on fisheries acoustics data and are increasingly used [Fernandes, 2009, D'Elia et al., 2014, Fallon et al., 2016, Proud et al., 2020], notably because this method is recognized for its ability to deal with small sample sizes and high-dimensional feature spaces [Cutler et al., 2007, Biau and Scornet, 2016].

Random Forest (RF) models [Breiman, 2001, Cutler et al., 2007] were used to build the classification algorithms. A “forest” of classification/regression trees, each using a random (training) subset of the available parameters to classify the dataset, was built. The resulting classifier used the mean values of this forest of decision trees to perform a class prediction on the test data subset, composed of the data not contained within the training subset.

A separate RF model was built for both CW and FM data. Following the standard RF protocol [Breiman, 2001], the respective datasets were randomly split into a training dataset (75% of the echo-traces) and a test dataset (25% of the echo-traces). The R [R Core Team, 2020] packages “caret” and “ranger” [Wright and Ziegler, 2017, Kuhn, 2019] were used to build the RF models. Two tuning parameters were set for each RF: the number of parameters selected randomly was at each tree node (mtry), and the number of trees to build (ntrees). These tuning parameters were set independently for each RF model by cross-validation, using the accuracy metric (% of good classification) [Breiman, 2001]. A multiple (three times) tenfold cross-validation [Stone, 1974, Breiman, 2001, Proud et al., 2020] has been performed after selection of the tuning parameters to assess the models training accuracy and kappa statistic (proportion of the classification results beyond the results expected to obtain by chance alone, Biau and Scornet, 2016). The contribution of each parameter to the models' powers was evaluated by the "Mean Decrease Impurity" value [Biau and Scornet, 2016]. This indicator estimates the total (averaged over all the trees) decrease in impurity obtained by splitting the respective tree on the given variable.

3 Results

3.1 Echo-traces extraction

A total of 104 echo-traces were isolated in CW (43 trees and 61 schools) and 117 echo-traces in FM (53 trees and 64 schools). Details of the number of echo-traces by lakes are presented in the Table 3.

Table 3: Number of echo-traces by lake and by acquisition mode (Continuous Wave (CW) and Frequency Modulated (FM))

	CW echo-traces	FM echo-traces
Hydropower reservoirs (trees)		
Chancy-Pougny	16	16
Bimont	19	23
Génissiat	8	14
Peri-alpine lakes (schools)		
Lake Geneva (September)	15	5
Lake Bourget	33	45
Lake Geneva (June)	13	14

3.2 Parameter distributions

The selected parameter boxplots of echo-traces for each lake are presented in Figures 3 and 4 for the CW and FM modes, respectively. The mean depth (a) and the echo-trace height (b) are also presented in the Figures 3 and 4 in an indicative manner. All parameters present a high variability from one lake to another lake, even for the same echo-trace types (immersed trees or fish schools).

In CW mode, no clear tendency can be shown, and thus no single parameter was able to discriminate directly between trees and schools. Mean dB differences at 200 and 120 kHz were low for school echo-traces ($\Delta MVBS_{school}$: 0.63 dB, with a 2.5 dB s.d., at the exception of the September Geneva's school echo-traces, with a $\Delta MVBS$ of -4.1 dB, s.d.: 0.8 dB); when dB differences were on average greater for tree echo-traces ($\Delta MVBS_{tree}$: -3.6 dB with a 1.7 dB s.d.), Figure 3 (k). Note that these values are expressed in dB, a logarithmic scale. 0.6 dB equates

Figure 3: Each echo-traces parameter boxplots by lake in CW mode. Boxes show the median and interquartile ranges; whiskers show the maximum and minimum values (excluding outliers); points represent outliers; diamonds represent means.

Figure 4: Each echo-traces parameter boxplots by lake in FM mode. Boxes show the median and interquartile ranges; whiskers show the maximum and minimum values (excluding outliers); points represent outliers; diamonds represent means.

approximately 1.1 times the backscatter volume in a linear scale and -3.6 equates to 2.3 times less acoustic energy scattered at the given frequency.

The Bimont reservoir tree echo-trace parameters generally yielded results closer to the fish school echo-trace parameters (Figure 3 (a) to (f)). In FM mode, the Bimont reservoir tree echo-trace parameters were also generally closer to fish schools echo-trace parameters (Figure 4, (a) to (f)). The $\Delta MVBS$ parameter revealed less separating power than in CW mode ($\Delta MVBS_{school}$: -4.8 dB and $\Delta MVBS_{tree}$: -4.3 dB, Figure 4 (m)).

The extracted broadband spectra for each FM echo-traces were normalized by Z-score and were presented in Figure 5. No obvious discriminant features could be extracted from the spectral noise for trees or schools. At 120 kHz, the median spectrum for trees presented a slightly decreasing slope ($s_{m120kHz}$ schools : -0.075, s.d. 0.076), while this slope was slightly flatter for schools ($s_{m120kHz}$ schools : 0.022, s.d. 0.048).

3.3 Random forest models

For the CW model as for the FM one, the default values of mtry (4) and ntrees (500) gave the best model fit, evaluated by model accuracy. The other mtry and ntrees parameter values tested (mtry: 2–6 and ntrees: 200–2500) showed no improvement in accuracy (Figure 6).

The random forest model tenfold cross-validation for the CW and FM model resulted in accuracy values of 0.86 and 0.87 with kappa values of 0.73 and 0.74, respectively.

(b)

Figure 5: Echo-traces (schools in (a) and trees in (b)) normalized $S_v(f)$ median spectrum in FM mode at 120 kHz (left) and 200 kHz (right). The mean spectrum for each lake are represented by the thicker line of the corresponding color.

Figure 6: Model accuracy of the random forest in function of the number of trees (ntrees) and the number of parameters selected at each tree nodes (mtry), for the CW modes (a) and the FM mode (b).

3.4 Random forest predictions

The application of the CW RF model on the CW test dataset gave a 65.38 % classification ratio, while the application of the FM RF model on the FM test dataset gave a 90 % classification ratio.

The prediction confusion matrices of the two models are presented in Table 4. The two models presented an acceptable prediction error for the schools (respectively 100% for CW and 77% for FM), the CW model displayed a poor prediction for trees (respectively 55% for CW and 95% for FM).

Table 4: Results of the confusion matrix of the random forest predictive models in CW and FM mode on the test data sets.

	CW		FM	
	Predicted tree	Predicted school	Predicted tree	Predicted school
Labeled tree	11 (55%)	9 (45%)	20 (95%)	1 (5%)
Labeled school	0 (0%)	6 (100%)	2 (22%)	7 (78%)

3.5 Parameters importance

The evaluation of the importance [Breiman, 2001, Biau and Scornet, 2016] of each school parameter (Figure 7) revealed a different importance pattern for the FM and the CW model. $\Delta MVBS$ is the clear key parameter (100% relative variable importance) in the CW model, with a small contribution attributed to the other parameters. In FM, key variables were the $s_{m120kHz}$ and the $\Delta MVBS$ parameters (82% and 100% relative variable importance, respectively). Four other parameters contributed over 50%: the 120 kHz skewness (53%), the $s_{m200kHz}$ (52.7%), and the 120/200 kHz S_v standard deviation (respectively 51% for the 120 kHz and 48% for the 200 kHz).

Figure 7: Relative variable importance (conditional variable importance normalized between 0 and 1) of the echo-traces parameters used to build the CW (red) and FM (blue) Random Forest models.

4 Discussion

This paper aimed to test a method to select parameters to discriminate echo-traces of trees and fish schools in reservoir lakes, using random forest models. These echo-traces were obtained by standard echosounders operated in FM and CW modes, with center frequencies at 120 and 200 kHz. Such a discrimination is not possible using a single frequency sounder [Josse and Le Squin, 2013], as demonstrated in previous studies. We tested two approaches using CW and FM sound datasets.

The CW model's classification result was not sufficient (65.38 %) to efficiently discriminate trees from schools, but the FM model displayed a good classification ratio on the test dataset (90 %). These results demonstrate, for the first time at the best of our knowledge, the possibility to discriminate trees from fish schools in reservoir lakes using broadband acoustics.

The results highlight the benefits of the new parameters added by using broadband acoustics. It shows that the choice of parameters was critical to obtain good classification results. Acoustic parameters are classed in three categories, morphological, positional and energetic Reid et al. [2000]. Following the conclusion of Josse and Le Squin [2013], the morphological and positional parameters were not considered enough informative to discriminate trees from schools and to be included in our analysis. Due to the high diversity of shapes and positions of trees and schools, these parameters are likely to be difficult to generalize [Horne, 2000, Coll and Lebourges Dhaussy, 2005]. However, the distance from the lake bottom to echo-traces could be seen as a good parameter, as it is expected that trees are associated with the lake bottom. Nevertheless, schools could also be regularly associated with the lake bottom and immersed trees regularly show no connection with the bottom due to horizontal branches, and/or when the tree center is not in the beam [Coll and Lebourges Dhaussy, 2005].

Therefore we only considered energetic parameters in this work, selected to test our three hypotheses on how to discriminate trees from fish schools. The mean S_v , S_v standard deviation and $\min S_v - \max S_v$ were selected on the hypothesis that variance of the backscattering echoes were higher in trees. However, even if trees present a higher S_v variance than schools (Figure 4 (e) and (f)) and that the 120 and 200 kHz S_v standard deviation parameters substantially participate to the FM model (respectively 51 and 48%), these parameters are not discriminant by their own. The skewness and kurtosis parameters were selected on the hypothesis that the PDFs of fish schools will be closer to a Rayleigh function [Scalabrin et al., 1996, Chu and Stanton, 2010] than trees. The 120 kHz skewness participates substantially to the FM model (53%), but like the standard deviation parameters, it is not discriminant by their own.

The high importance of the $\Delta MVBS$ for the CW and the FM model, and the $S_{m120kHz}$ parameter for the FM model, confirm that frequency spectrum differences between trees and fish schools are the key parameters to discriminate trees from schools in our models. Even if the slope difference between trees' and schools' 120 kHz median spectrum is small (around 0.05), it is the most important parameter in the FM random forest models. These differences in the slope between tree and school 120 kHz spectrum is probably due to the non-random distribution of echoes from trees resulting from the material properties and the complex shape of the trees (L. Berger, personal communication): the dense trunk will result in a higher backscattering amplitude, when the surrounding smaller and less dense branches, will result in less pronounced backscatter. Furthermore, trees are motionless from one ping to another, compared to moving schools with a random distribution of echoes from one ping to another [Scalabrin et al., 1996]. In broadband, higher frequencies present a smaller beam than the lower frequencies for a given band [Demer et al., 2017], implying that the probability of missing the center of the tree will be higher for the higher part of the band. The detected backscattering differences detected between trees and fish schools seem, at these frequencies, likely not primarily driven by their respective material properties and

acoustic impedance, as first hypothesized, but by their spatial organization (the distribution of the elementary particles of each object in the water column).

Most considered parameters significantly contributed to the power of the FM model. This fact strengthens the relevance of ensemble tree methods, such as random forest, to discriminate echo-traces. Confirming observations by D'Elia et al. [2014] and Proud et al. [2020], identifying different fish species groups.

Immobile trees are expected to show a coherent phase response, while mobile fish schools are expected to show an incoherent phase response (L. Berger, personal communication). If the spectrum slope was a way to highlight the spatial organization discriminant characteristic, phase responses [?] could be, in a future work, an interesting way to discriminate trees from schools, in FM as well in CW mode.

Simultaneous use of side-scan sonars [Bollinger and Kline, 2017] with echosounders could be another interesting way to discriminate trees from schools. Portable recreational side-scans could probably deliver enough resolution to highlight the details of the tree branch structure and identify trees. Thus, side-scan sonars could be used as “decision support tools” [Kaeser and Litts, 2010] to remove trees from the acoustic dataset during the data cleaning step before a stock assessment. However, the recorded volume being different, the user will have to manually set the echo-traces, which can be time-consuming.

We observed high inter-lake variability in the echotrace parameters (Figures 3 and 4). Building a reference dataset with a variety of different lakes (with different fish and tree species) will be necessary steps to ensure the efficiency of the model in a variety of conditions. Additionally, this method could be tested in tropical reservoir lakes that present immersed trees and fish schools simultaneously, as Manantali (Mali, Coll et al., 2007) or Nam Theun 2 (Lao PDR, Tessier et al., 2020).

Horizontally-oriented echosounders are frequently used in reservoirs, as these lakes are often shallow and/or the fish are commonly present near the surface [Kubecka and Wittingerova, 1998, Goulon et al., 2018]. Our method could potentially be adapted to horizontal pinging, but this type of survey mainly use single echo detection of individual fish to assess fish stock, and not fish school echos [Kubecka and Wittingerova, 1998, Knudsen and Sægrov, 2002]. In addition, we will need to take account of the high directivity of fish acoustic backscattering in horizontal acquisition [Tušer et al., 2009].

In lakes, 120 and 200 kHz are the ones of the main commonly used frequencies [Draštík et al., 2017] and are convenient due to the small size of the transducer. However, using a lower frequency (as the 70 kHz) could potentially help to obtain a better discrimination power as the small branches will possibly be in the Rayleigh scattering regime [Simmonds and MacLennan, 2005]. However, the larger size of low frequency transducers (18/38 kHz) [Mouget et al., 2019] could hinder their use in lakes or in reservoirs, where small craft are more commonly used [Tessier et al., 2016].

The development of an automatic method to discriminate submerged trees and fish schools, notably using broadband acoustics, is a new subject explored in this article. We showed that it is possible to obtain an improved classification rate when using broadband acoustics. Larger datasets, including tropical reservoirs, are necessary to obtain a robust discrimination tool that could be used routinely. In this way, hydroacoustics, a non-intrusive tool already used in numerous lakes can be used to monitor fish populations in reservoirs with more confidence.

Acknowledgements

The authors want to thank the IRD staff from the UMR LEMAR and especially Jeremie Habasque for the lending of an ES200-7CD transducer, as well as Tonny Algrøy from SIMRAD Kongsberg for the lending of the corresponding license for WBT. We additionally want to thank Julien Dublon and Thierry Tormos for their knowledge of the Bimont Reservoir and their help during the field survey, as well as Laurent Berger for his help concerning the use of MOVIES3D software on broadband data. Finally, we gratefully thank Jean-Christophe Hustache for his priceless help in the sampling preparation and the fieldwork, without whom nothing would have been possible. This work was supported by Pole ECLA (OFB - INRAE - USMB) and had support from AnaEE-France and Observatory of LAkes (OLA), boat and technical facilities. This work is associated to Bi-O-Rhône project, supported and co-financed by the FEDER fund (Fonds Européens de Développement Régional) Interreg France-Switzerland.

References

- Orlane Anneville, Camille Vogel, Jérémy Lobry, and Jean Guillard. Fish communities in the Anthropocene: detecting drivers of changes in the deep peri-alpine Lake Geneva. *Inland Waters*, 7(1):65–76, January 2017. ISSN 2044-2041, 2044-205X. doi: 10.1080/20442041.2017.1294350. URL <https://www.tandfonline.com/doi/full/10.1080/20442041.2017.1294350>.
- Christopher Bassett, Alex De Robertis, Christopher D Wilson, and Handling editor: Purnima Ratilal. Broadband echosounder measurements of the frequency response of fishes and euphausiids in the Gulf of Alaska. *ICES Journal of Marine Science*, 75(3):1131–1142, 2018. doi: 10.1093/icesjms/fsx204. URL <http://dx.doi.org/10.1093/icesjms/fsx204>.
- Gérard Biau and Erwan Scornet. A random forest guided tour. *TEST*, 25(2):197–227, June 2016. ISSN 1133-0686, 1863-8260. doi: 10.1007/s11749-016-0481-7. URL <http://link.springer.com/10.1007/s11749-016-0481-7>.
- Arthur Blanluet, Mathieu Doray, Laurent Berger, Jean-Baptiste Romagnan, Naig Le Bouffant, Sigrid Lehuta, and Pierre Petitgas. Characterization of sound scattering layers in the Bay of Biscay using broadband acoustics, nets and video. *PLOS ONE*, 14(10):e0223618, October 2019. ISSN 1932-6203. doi: 10.1371/journal.pone.0223618. URL <http://dx.plos.org/10.1371/journal.pone.0223618>.
- Michael A. Bollinger and Richard J. Kline. Validating Sidescan Sonar as a Fish Survey Tool over Artificial Reefs. *Journal of Coastal Research*, 336:1397–1407, November 2017. ISSN 0749-0208, 1551-5036. doi: 10.2112/JCOASTRES-D-16-00174.1. URL <http://www.bioone.org/doi/10.2112/JCOASTRES-D-16-00174.1>.
- Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001. Publisher: Springer.
- J Cardenas, D Petit, and D Novoa. Prospección hidroacústica del Lago Guri

(Venezuela). Ensayo preliminar para la cuantificación y determinación de la distribución de la biomasa íctica. *Memoria de la Sociedad de Ciencias Naturales La Salle*, 130:77–90, 1988.

Franck Cattaneo, Jean Guillard, Seydina Diouf, Jane O'Rourke, and David Grismardias. Mitigation of ecological impacts on fish of large reservoir sediment management through controlled flushing – The case of the Verbois dam (Rhône River, Switzerland). *Science of The Total Environment*, 756:144053, February 2021. ISSN 00489697. doi: 10.1016/j.scitotenv.2020.144053. URL <https://linkinghub.elsevier.com/retrieve/pii/S0048969720375847>.

Dezhang Chu and Timothy K Stanton. Statistics of Echoes From a Directional Sonar Beam Insonifying Finite Numbers of Single Scatterers and Patches of Scatterers. *IEEE Journal of Oceanic Engineering*, 35(2):267–277, April 2010. ISSN 0364-9059, 1558-1691. doi: 10.1109/JOE.2009.2037988. URL <http://ieeexplore.ieee.org/document/5439764/>.

Céline Coll and Anne Lebourges Dhaussy. How to recognize the fish in the trees ? Methodological reflections for the fish population assessment in an artificial reservoir, 2005. URL https://www.researchgate.net/publication/215664216_How_recognize_the_fish_in_the_trees_Methodological_reflections_for_the_fish_population_assessment_in_an_artificial_reservoir. Presentation, WGFAST 2005.

Céline Coll, Luis Tito de Morais, Raymond Laë, Anne Lebourges-Dhaussy, Monique Simier, Jean Guillard, Erwan Josse, Jean-Marc Ecoutin, Jean-Jacques Albaret, Jean Raffray, and Justin Kantoussan. Use and limits of three methods for assessing fish size spectra and fish abundance in two tropical man-made lakes. *Fisheries Research*, 83(2-3):306–318, February 2007. ISSN 01657836. doi: 10.1016/j.fishres.2006.10.005. URL <https://linkinghub.elsevier.com/retrieve/pii/S0165783606003675>.

D. Richard Cutler, Thomas C. Edwards, Karen H. Beard, Adele Cutler, Kyle T. Hess, Jacob Gibson, and Joshua J. Lawler. Random Forest For Classification

In Ecology. *Ecology*, 88(11):2783–2792, November 2007. ISSN 0012-9658.
doi: 10.1890/07-0539.1. URL <http://doi.wiley.com/10.1890/07-0539.1>.

M. D’Elia, B. Patti, A. Bonanno, I. Fontana, G. Giacalone, G. Basilone, and P.G. Fernandes. Analysis of backscatter properties and application of classification procedures for the identification of small pelagic fish species in the Central Mediterranean. *Fisheries Research*, 149:33–42, January 2014. ISSN 01657836. doi: 10.1016/j.fishres.2013.08.006. URL <https://linkinghub.elsevier.com/retrieve/pii/S0165783613001872>.

David A. Demer, Laurent Berger, Matteo Bernasconi, Bethke Eckhard, Kevin Boswell, Dezhang Chu, Reka Domokos, Adam Dunford, Sascha Fässler, Stéphane Gauthier, Lawrence T. Hufnagle, J Michael Jech, Naig Le Bouffant, Anne Lebourges Dhaussy, Xavier Lurton, Gavin J. Macaulay, Yannick Perrot, Tim E. Ryan, Sandra Parker-Stetter, Sarah Stienessen, Thomas Weber, and Neal Williamson. Calibration of acoustic instruments. Technical Report 326, ICES, Denmark, May 2015. URL [http://ices.dk/sites/pub/Publication%20Reports/Cooperative%20Research%20Report%20\(CRR\)/crr326/CRR326.pdf](http://ices.dk/sites/pub/Publication%20Reports/Cooperative%20Research%20Report%20(CRR)/crr326/CRR326.pdf).

David A. Demer, Lars N. Andersen, Chris Bassett, Laurent Berger, Dezhang Chu, Jeff Condiotty, George Cutter Jr., Briony Hutton, Rolf J. Korneliussen, Naig Le Bouffant, Gavin J. Macaulay, William L Michaels, David Murfin, Armin Potitzer, Josiah S. Renfree, Thomas S. Sessions, Kevin L. Stierhoff, and Charles Thompson. USA–Norway EK80 Workshop Report: Evaluation of a wideband echosounder for fisheries and marine ecosystem science. Cooperative Research Report 336, ICES, NOAA’s Southwest Fisheries Science Center in San Diego, California, USA, April 2017.

Mathieu Doray, Pierre Petitgas, Jean Baptiste Romagnan, Martin Huret, Erwan Duhamel, Christine Dupuy, Jérôme Spitz, Matthieu Authier, Florence Sanchez, Laurent Berger, Ghislain Doremus, Paul Bourriau, Patrick Grellier, and Jacques Massé. The PELGAS survey: ship-based integrated monitoring of the Bay of Biscay pelagic ecosystem. *Progress in Oceanography*, 166:

533 15–29, 2018. ISSN 0079-6611. doi: 10.1016/j.pocean.2017.09.015. URL
534 <https://www.sciencedirect.com/science/article/pii/S0079661117300939>.

535 Vladislav Draščík, Malgorzata Godlewska, Helge Balk, Peter Clabburn, Jan
536 Kubečka, Emma Morrissey, Jon Hateley, Ian J. Winfield, Tomáš Mrkvička, and
537 Jean Guillard. Fish hydroacoustic survey standardization: A step forward
538 based on comparisons of methods and systems from vertical surveys of a
539 large deep lake: Fish hydroacoustic survey standardization. *Limnology and*
540 *Oceanography: Methods*, 15(10):836–846, October 2017. ISSN 15415856. doi:
541 10.1002/lom3.10202. URL <http://doi.wiley.com/10.1002/lom3.10202>.

542 Pablo Christian Escobar-Flores, Yoann Lacroix, and Richard Lyell O'Driscoll.
543 Acoustic Assessment of the Micronekton Community on the Chatham Rise,
544 New Zealand, Using a Semi-Automated Approach. *Frontiers in Marine Sci-*
545 *ence*, 6, August 2019. ISSN 2296-7745. doi: 10.3389/fmars.2019.00507. URL
546 <https://www.frontiersin.org/article/10.3389/fmars.2019.00507/full>.

547 Niall G. Fallon, Sophie Fielding, and Paul G. Fernandes. Classification of Southern
548 Ocean krill and icefish echoes using random forests. *ICES Journal of Marine*
549 *Science*, 73(8):1998–2008, September 2016. ISSN 1095-9289, 1054-3139. doi:
550 10.1093/icesjms/fsw057. URL [https://academic.oup.com/icesjms/article/](https://academic.oup.com/icesjms/article/73/8/1998/2198397)
551 [73/8/1998/2198397](https://academic.oup.com/icesjms/article/73/8/1998/2198397).

552 Paul G. Fernandes. Classification trees for species identification of fish-school
553 echotraces. *ICES Journal of Marine Science*, 66(6):1073–1080, July 2009.
554 ISSN 1095-9289, 1054-3139. doi: 10.1093/icesjms/fsp060. URL [https://](https://academic.oup.com/icesjms/article/66/6/1073/691837)
555 academic.oup.com/icesjms/article/66/6/1073/691837.

556 PG Fernandes, RJ Korneliussen, A Lebourges-Dhaussy, J Masse, M Iglesias,
557 N Diner, E Ona, T Knutsen, J Gajate, and R Ponce. The SIMFAMI project:
558 species identification methods from acoustic multifrequency information. Final
559 Report to the EC No. Final Report to the EC 2054, 2006.

560 Amélie Frantz, Marie-Elodie Perga, and Jean Guillard. Parasitic versus nutritional
561 regulation of natural fish populations. *Ecology and Evolution*, 8(17):8713–8725,

September 2018. ISSN 20457758. doi: 10.1002/ece3.4391. URL <https://onlinelibrary.wiley.com/doi/10.1002/ece3.4391>.

Jaroslava Frouzová, Jan Kubečka, Marie Prchalová, and V Draštik. Use of hydroacoustics in tropical freshwater ecology: a pilot study of reservoirs and lakes in tropical Asia. *Aquatic ecosystem and development: comparative Asian perspectives*. Backhuys Publishers, Leiden, The Netherlands, pages 195–202, 2008.

Sven Gastauer, Ben Scoulding, and Miles Parsons. An Unsupervised Acoustic Description of Fish Schools and the Seabed in Three Fishing Regions Within the Northern Demersal Scalefish Fishery (NDSF, Western Australia). *Acoustics Australia*, 45(2):363–380, August 2017. ISSN 0814-6039, 1839-2571. doi: 10.1007/s40857-017-0100-0. URL <http://link.springer.com/10.1007/s40857-017-0100-0>.

A Getabu, Rhoda Tumwebaze, and D. N. MacLennan. Spatial distribution and temporal changes in the fish populations of Lake Victoria. *Aquatic Living Resources*, 16(3):159–165, July 2003. ISSN 09907440. doi: 10.1016/S0990-7440(03)00008-1. URL <https://linkinghub.elsevier.com/retrieve/pii/S0990744003000081>.

Michaël Girard, Chloé Goulon, Anne Tessier, Pascal Vonlanthen, and Jean Guillard. Comparisons of day-time and night-time hydroacoustic surveys in temperate lakes. *Aquatic Living Resources*, 33:9, 2020. ISSN 1765-2952. doi: 10.1051/alr/2020011. URL <https://www.alr-journal.org/10.1051/alr/2020011>.

Małgorzata Godlewska, Michel Colon, Adam Jóźwik, and Jean Guillard. How pulse lengths impact fish stock estimations during hydroacoustic measurements at 70 kHz. *Aquatic Living Resources*, 24(1):71–78, January 2011. ISSN 0990-7440, 1765-2952. doi: 10.1051/alr/2011104. URL <http://www.alr-journal.org/10.1051/alr/2011104>.

C. Goulon, S. Westrelin, V. Samedy, R. Roy, J. Guillard, and C. Argillier. Complementarity of two high-resolution spatiotemporal methods (hydroacoustics

and acoustic telemetry) for assessing fish distribution in a reservoir. *Hydroécologie Appliquée*, 20:57–84, May 2018. ISSN 1147-9213, 1958-556X. doi: 10.1051/hydro/2017001. URL <https://www.hydroecologie.org/10.1051/hydro/2017001>.

David Grimardias, Jean Guillard, and Franck Cattaneo. Drawdown flushing of a hydroelectric reservoir on the Rhône River: Impacts on the fish community and implications for the sediment management. *Journal of Environmental Management*, 197:239–249, July 2017. ISSN 03014797. doi: 10.1016/j.jenvman.2017.03.096. URL <https://linkinghub.elsevier.com/retrieve/pii/S0301479717303213>.

Jean Guillard, Marie-Elodie Perga, Michel Colon, and Nadine Angeli. Hydroacoustic assessment of young-of-year perch, *Perca fluviatilis*, population dynamics in an oligotrophic lake (Lake Annecy, France). *Fisheries Management and Ecology*, 13(5):319–327, October 2006. ISSN 0969-997X, 1365-2400. doi: 10.1111/j.1365-2400.2006.00508.x. URL <http://doi.wiley.com/10.1111/j.1365-2400.2006.00508.x>.

Jean Guillard, Paul. G. Fernandes, Thomas Laloë, and Patrice Brehmer. Three-dimensional internal spatial structure of young-of-the-year pelagic freshwater fish provides evidence for the identification of fish school species: Species identification from 3D structure. *Limnology and Oceanography: Methods*, 9(7):322–328, July 2011. ISSN 15415856. doi: 10.4319/lom.2011.9.322. URL <http://doi.wiley.com/10.4319/lom.2011.9.322>.

John K. Horne. Acoustic approaches to remote species identification: a review. *Fisheries Oceanography*, 9(4):356–371, December 2000. ISSN 1054-6006, 1365-2419. doi: 10.1046/j.1365-2419.2000.00143.x. URL <http://doi.wiley.com/10.1046/j.1365-2419.2000.00143.x>.

ICES. Report on Echo Trace Classification. 2000. doi: 10.17895/ICES.PUB.5371. URL <http://www.ices.dk/sites/pub/PublicationReports/Forms/DispForm.aspx?ID=35681>. Publisher: ICES.

ICES. Description of the ICES HAC standard data exchange format, version 1.60. Rapport de recherches collectives 278, ICES, Copenhagen, 2005. OCLC: 255405946.

Julie Josse and Sandrine Le Squin. Discrimination : Applications aux écotraces de milieux lacustres. Rapport de stage, IRD, Brest, 2013.

Juha Jurvelius, Frank Reier Knudsen, Helge Balk, Timo J. Marjomäki, Heikki Peltonen, Jouni Taskinen, Antti Tuomaala, and Markku Viljanen. Echo-sounding can discriminate between fish and macroinvertebrates in fresh water. *Freshwater Biology*, 53(5):912–923, May 2008. ISSN 0046-5070, 1365-2427. doi: 10.1111/j.1365-2427.2007.01944.x. URL <http://doi.wiley.com/10.1111/j.1365-2427.2007.01944.x>.

Adam J. Kaeser and Thomas L. Litts. A Novel Technique for Mapping Habitat in Navigable Streams Using Low-cost Side Scan Sonar. *Fisheries*, 35(4):163–174, April 2010. ISSN 0363-2415, 1548-8446. doi: 10.1577/1548-8446-35.4.163. URL <http://doi.wiley.com/10.1577/1548-8446-35.4.163>.

Frank Reier Knudsen and Harald Sægvog. Benefits from horizontal beaming during acoustic survey: application to three Norwegian lakes. *Fisheries Research*, 56(2):205–211, May 2002. ISSN 01657836. doi: 10.1016/S0165-7836(01)00318-6. URL <https://linkinghub.elsevier.com/retrieve/pii/S0165783601003186>.

Rolf Korneliussen, Laurent Berger, Fabio Campanella, Dezhang Chu, David A. Demer, Alex De Robertis, Reka Domokos, Mathieu Doray, Sophie Fielding, Sascha Fässler, Stéphane Gauthier, Sven Gastauer, John K. Horne, Briony Hutton, Frederico Iriarte, J Michael Jech, Rudy J. Kloser, Gareth L. Lawson, Anne Lebourges-Dhaussy, Ian McQuinn, Marian Peña, Ben Scoulding, Serdar Sakinan, Matthias Schaber, C Taylor, and Charles Thompson. Acoustic target classification. 2018. doi: 10.17895/ices.pub.4567. URL <http://www.ices.dk/sites/pub/PublicationReports/Forms/DispForm.aspx?ID=34895>.

Jan Kubecka and Monika Wittingerova. Horizontal beaming as a crucial component of acoustic fish stock assessment in freshwater reservoirs. *Fish-*

648 *eries Research*, 35(1-2):99–106, May 1998. ISSN 01657836. doi: 10.1016/
649 S0165-7836(98)00064-2. URL [https://linkinghub.elsevier.com/retrieve/
650 pii/S0165783698000642](https://linkinghub.elsevier.com/retrieve/pii/S0165783698000642).

651 M. Kuhn. caret: Classification and Regression Training, 2019.

652 G Lawson. Species identification of pelagic fish schools on the South African
653 continental shelf using acoustic descriptors and ancillary information. *ICES
654 Journal of Marine Science*, 58(1):275–287, February 2001. ISSN 10543139.
655 doi: 10.1006/jmsc.2000.1009. URL [https://academic.oup.com/icesjms/
656 article-lookup/doi/10.1006/jmsc.2000.1009](https://academic.oup.com/icesjms/article-lookup/doi/10.1006/jmsc.2000.1009).

657 Elizabeth A. Logerwell and Christopher D. Wilson. Species discrimination of
658 fish using frequency-dependent acoustic backscatter. *ICES Journal of Ma-
659 rine Science*, 61(6):1004–1013, January 2004. ISSN 1095-9289, 1054-3139.
660 doi: 10.1016/j.icesjms.2004.04.004. URL [https://academic.oup.com/icesjms/
661 article/61/6/1004/675855](https://academic.oup.com/icesjms/article/61/6/1004/675855).

662 David N. MacLennan, Paul. G. Fernandes, and John Dalen. A consistent ap-
663 proach to definitions and symbols in fisheries acoustics. *ICES Journal of Ma-
664 rine Science*, 59(2):365–369, April 2002. ISSN 10543139. doi: 10.1006/jmsc.
665 2001.1158. URL [https://academic.oup.com/icesjms/article-lookup/doi/
666 10.1006/jmsc.2001.1158](https://academic.oup.com/icesjms/article-lookup/doi/10.1006/jmsc.2001.1158).

667 Lauro S.P. Madureira, Inigo Everson, and Eugene J. Murphy. Interpretation of
668 acoustic data at two frequencies to discriminate between Antarctic krill (*Eu-
669 phausia superba* Dana) and other scatterers. *Journal of Plankton Research*,
670 15(7):787–802, 1993. ISSN 0142-7873, 1464-3774. doi: 10.1093/plankt/15.7.
671 787. URL [https://academic.oup.com/plankt/article-lookup/doi/10.1093/
672 plankt/15.7.787](https://academic.oup.com/plankt/article-lookup/doi/10.1093/plankt/15.7.787).

673 Anne Mouget, Chloé Goulon, Thomas Axenrot, Helge Balk, Anne Lebourges-
674 Dhaussy, Malgorzata Godlewska, and Jean Guillard. Including 38 kHz
675 in the standardization protocol for hydroacoustic fish surveys in temperate

lakes. *Remote Sensing in Ecology and Conservation*, 5(4):332–345, December 2019. ISSN 2056-3485, 2056-3485. doi: 10.1002/rse2.112. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/rse2.112>.

Sandra L Parker-Stetter, Lars G Rudstam, Patrick J. Sullivan, and David M. Warner. Standard operating procedures for fisheries acoustic surveys in the Great Lakes. *Great Lakes Fish. Comm. Spec. Pub.*, 2009. ISSN 1090-1051.

Riley A. Pollom and George A. Rose. A global review of the spatial, taxonomic, and temporal scope of freshwater fisheries hydroacoustics research. *Environmental Reviews*, 24(3):333–347, September 2016. ISSN 1181-8700, 1208-6053. doi: 10.1139/er-2016-0017. URL <http://www.nrcresearchpress.com/doi/10.1139/er-2016-0017>.

Roland Proud, Richard Mangeni-Sande, Robert J Kayanda, Martin J Cox, Chrisphine Nyamweya, Collins Ongore, Vianny Natugonza, Inigo Everson, Mboni Elison, Laura Hobbs, Benedicto Boniphace Kashindye, Enock W Mlaponi, Anthony Taabu-Munyaho, Venny M Mwainge, Esther Kagoya, Antonio Pegado, Evarist Nduwayesu, and Andrew S Brierley. Automated classification of schools of the silver cyprinid *Rastrineobola argentea* in Lake Victoria acoustic survey data using random forests. *ICES Journal of Marine Science*, May 2020. ISSN 1054-3139, 1095-9289. doi: 10.1093/icesjms/fsaa052. URL <https://academic.oup.com/icesjms/advance-article/doi/10.1093/icesjms/fsaa052/5835268>.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL <https://www.R-project.org/>.

D Reid, C Scalabrin, P Petitgas, J Masse, R Aukland, P Carrera, and S Georgarakos. Standard protocols for the analysis of school based data from echo sounder surveys. *Fisheries Research*, 47(2-3):125–136, July 2000. ISSN 01657836. doi: 10.1016/S0165-7836(00)00164-8. URL <https://linkinghub.elsevier.com/retrieve/pii/S0165783600001648>.

Frédéric Rimet, Orlane Anneville, Denis Barbet, Cécile Chardon, Laura Crépin, Isabelle Domaizon, Jean-Marcel Dorioz, Laurent Espinat, Victor Frossard, Jean Guillard, Chloé Goulon, Valérie Hamelet, Jean-Christophe Hustache, Stéphan Jacquet, Leslie Lainé, Bernard Montuelle, Pascal Perney, Philippe Quetin, Serena Rasconi, Antoine Schellenberger, Viet Tran-Khac, and Ghislaine Monet. The Observatory on LAkes (OLA) database: Sixty years of environmental data accessible to the public. *Journal of Limnology*, 79(2):164–178, 2020. ISSN 1723-8633, 1129-5767. doi: 10.4081/jlimnol.2020.1944. URL <https://jlimnol.it/index.php/jlimnol/article/view/jlimnol.2020.1944>.

LARS G Rudstam, J MICHAEL Jech, SANDRA L Parker-Stetter, JOHN K Horne, PATRICK J Sullivan, and DORAN M Mason. Fisheries acoustics. *third edition* (AV Zale, DL Parrish, TM Sutton, Eds.) (2012), page 40, 2012.

L.G Rudstam, S.L Parker, D.W Einhouse, L.D Witzel, D.M Warner, J.L Stritzel, D.L Parrish, and P.J Sullivan. Application of in situ target-strength estimations in lakes: examples from rainbow-smelt surveys in Lakes Erie and Champlain. *ICES Journal of Marine Science*, 60(3):500–507, January 2003. ISSN 1095-9289, 1054-3139. doi: 10.1016/S1054-3139(03)00046-8. URL <https://academic.oup.com/icesjms/article/60/3/500/658346>.

C. Scalabrin, N. Diner, A. Weill, A. Hillion, and M. C. Mouchot. Narrowband acoustic identification of monospecific fish shoals. *ICES Journal of Marine Science*, 53(2):181–188, April 1996. ISSN 10543139. doi: 10.1006/jmsc.1996.0020. URL <https://academic.oup.com/icesjms/article-lookup/doi/10.1006/jmsc.1996.0020>.

E. John Simmonds and D. N. MacLennan. *Fisheries acoustics: theory and practice*. Number 10 in Fish and aquatic resources series. Blackwell Science, Oxford ; Ames, Iowa, 2 edition, 2005. ISBN 978-0-632-05994-2.

T Stanton, Peter H. Wiebe, Dezhang Chu, Mark C. Benfield, Lori Scalon, Linda V. Martin, and Robert L. Eastwood. On acoustic estimates of zooplankton biomass. *ICES Journal of Marine Science*, 51(4):505–512, November

1994. ISSN 10543139. doi: 10.1006/jmsc.1994.1051. URL <http://icesjms.oxfordjournals.org/cgi/doi/10.1006/jmsc.1994.1051>.

M. Stone. Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36(2):111–133, January 1974. ISSN 00359246. doi: 10.1111/j.2517-6161.1974.tb00994.x. URL <http://doi.wiley.com/10.1111/j.2517-6161.1974.tb00994.x>.

Anne Tessier, Stephane Descoux, Raymond Lae, Maud Cottet, Pierre Guedant, and Jean Guillard. Fish Assemblages in Large Tropical Reservoirs: Overview of Fish Population Monitoring Methods. *Reviews in Fisheries Science & Aquaculture*, 24(2):160–177, April 2016. ISSN 2330-8249, 2330-8257. doi: 10.1080/23308249.2015.1112766. URL <https://www.tandfonline.com/doi/full/10.1080/23308249.2015.1112766>.

Anne Tessier, Maud Cottet, Kaoboun Kue, Vincent Chanudet, Stéphane Descoux, and Jean Guillard. Low input of offshore areas to fisheries in a large tropical reservoir in Lao PDR. *Limnology*, 21(1):73–86, January 2020. ISSN 1439-8621, 1439-863X. doi: 10.1007/s10201-019-00583-1. URL <http://link.springer.com/10.1007/s10201-019-00583-1>.

Verena M. Trenkel, Laurent Berger, Sébastien Bourguignon, Mathieu Doray, Ronan Fablet, Jacques Massé, Valérie Mazauric, Cyrille Poncelet, Gael Quemener, Carla Scalabrin, and Hector Villalobos. Overview of recent progress in fisheries acoustics made by Ifremer with examples from the Bay of Biscay. *Aquatic Living Resources*, 22(4):433–445, October 2009. ISSN 0990-7440, 1765-2952. doi: 10.1051/alr/2009027. URL <http://www.alr-journal.org/10.1051/alr/2009027>.

Michal Tušer, Jan Kubečka, Jaroslava Frouzová, and Oldřich Jarolím. Fish orientation along the longitudinal profile of the Římov reservoir during daytime: Consequences for horizontal acoustic surveys. *Fisheries Research*, 96(1):23–29, February 2009. ISSN 01657836. doi: 10.1016/j.fishres.2008.09.022. URL <https://linkinghub.elsevier.com/retrieve/pii/S0165783608002774>.

Marvin N. Wright and Andreas Ziegler. **ranger** : A Fast Implementation of Random Forests for High Dimensional Data in *C++* and *R*. *Journal of Statistical Software*, 77(1), 2017. ISSN 1548-7660. doi: 10.18637/jss.v077.i01. URL <http://www.jstatsoft.org/v77/i01/>.

Daniel L. Yule, Lori M. Evrard, Sébastien Cachera, Michel Colon, and Jean Guillard. Comparing two fish sampling standards over time: largely congruent results but with caveats. *Freshwater Biology*, 58(10):2074–2088, October 2013. ISSN 00465070. doi: 10.1111/fwb.12192. URL <http://doi.wiley.com/10.1111/fwb.12192>.

Christiane Zarfl, Alexander E. Lumsdon, Jürgen Berlekamp, Laura Tydecks, and Klement Tockner. A global boom in hydropower dam construction. *Aquatic Sciences*, 77(1):161–170, January 2015. ISSN 1015-1621, 1420-9055. doi: 10.1007/s00027-014-0377-0. URL <http://link.springer.com/10.1007/s00027-014-0377-0>.