Classification trees for species identification of fish-school echotraces

Paul G. Fernandes

Fernandes, P. G. 2009. Classification trees for species identification of fish-school echotraces. – ICES Journal of Marine Science, 66: 1073 – 1080.

Acoustic surveys provide valuable information on the abundance and distribution of many fish species, but are particularly effective for schooling pelagic fish of commercial importance. However, despite recent advances in multifrequency processing, the technique still requires some subjective judgement when allocating the acoustic data, fish-school echotraces, to particular species—the so-called "scrutiny process". This is assisted by "ground truth" trawling and operator experience of relating trawl data to echotraces of particular fish schools. In this paper, a method to identify species based on "classification trees" is applied to data from a component of the International North Sea Herring Acoustic Survey. Classification trees may be considered as a variant of decision trees, and have properties that are intuitive to biologists, because they are similar to the keys used for the biological identification of species. The method described here incorporates a multifrequency fish-school filter, image analysis to isolate fish-school echotraces, and finally, a classification-tree system using the multifrequency information from the ground-truthed echotraces that can be translated into a processing tool for objective species allocation. The classification-tree system is compared with the traditional method of expert-based scrutiny. Unlike the latter, however, a measure of uncertainty is attributed to the classification-tree approach and this could be propagated through the acoustic-survey estimation procedure as a component of the total uncertainty in the abundance estimate.

Keywords: acoustics, classification trees, fish, ground truth, herring, mackerel, multifrequency, species identification.

Received 20 August 2008; accepted 2 January 2009; advance access publication 8 April 2009.

P. G. Fernandes: FRS Marine Laboratory, PO Box 101, 375 Victoria Road, Torry, Aberdeen AB11 9DB, Scotland, UK tel: +44 1224 295403; fax: +44 1224 295511; e-mail fernandespg@marlab.ac.uk.

Introduction

Acoustic surveys using quantitative scientific echosounders to determine the abundance and distribution of fish and other marine fauna are becoming increasingly important for the management of marine resources (Simmonds and MacLennan, 2005). With a move towards ecosystem-based fishery management, demands for marine-resource monitoring will go beyond the traditional target species to include other components of the marine environment. There is, therefore, not only a need for acoustic surveys to be more accurate and precise, but also to direct efforts towards an increasing array of targets. In such a context, it is clear why MacLennan and Holliday (1996) proclaimed that species identification is "the grand challenge of fisheries and plankton acoustics", and Horne (2000) in his review of remote-species identification identified this as the "Holy Grail" for acoustic researchers.

A scientific echosounder measures echo returns from sound pulses transmitted downwards into the water column. These echoes are arranged to form "echograms", which are two-dimensional (depth and distance) representations of the positions of any objects that scatter sound (Figure 1). Many of these objects, such as the seabed, fish schools, and plankton aggregations, are large enough to form coherent objects known as "echotraces" [see Reid *et al.* (2000) for a review of echotrace classification]. These are, however, just echotraces, not fish. Although the echotraces contain a great deal of information, such as their size, location, and echo intensity, their composition (i.e. the species) is unknown. Usually, expert knowledge is used to associate characteristic echotraces with specific species; fishers have used such classification techniques for decades (Cushing, 1963). Acoustic surveys also rely on such skills, but the interpretation or "scrutiny" (e.g. see Appendix 6 of ICES, 1998) is aided by sampling the echotraces using alternative means, usually a pelagic trawl. This may be termed "ground truthing", but as McClatchie *et al.* (2000) rightly point out, the alternative methods may not always provide an adequate standard for comparison and, therefore, may not always represent the "truth". Sampling with a pelagic trawl, however, is still generally perceived as the best means of identifying the echotraces of fish schools (Simmonds and MacLennan, 2005), notwithstanding the inherent limitations that make this method less than ideal (Bethke *et al.*, 1999).

Following on from subjective echotrace identification, many studies have attempted to classify and identify echotraces based on the information provided by the echosounder; see Horne (2000) for a review. These studies range from the investigation of single-frequency echotrace characteristics (e.g. Rose and Leggett, 1988; Scalabrin *et al.*, 1996) to combinations of synchronous measurements made at more than one echosounder frequency (multifrequency) (e.g. Madureira *et al.*, 1993; Kloser *et al.*, 2002), to the use of broad- or wide-band echosounders employing a sweep of frequencies (e.g. Simmonds *et al.*, 1996; Lundgren and Nielsen, 2008). Multifrequency studies on fish schools are now facilitated by the wide availability of affordable powerful computers and software developments (Higginbottom *et al.*, 2000), which

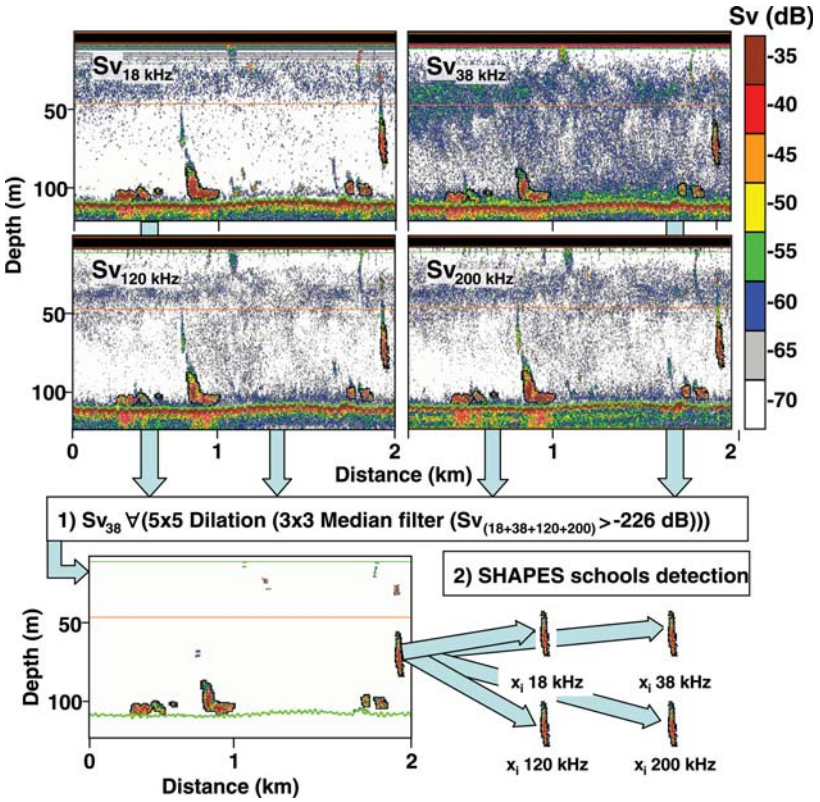

Figure 1. Multifrequency, fish-school, filter algorithm. The top four panels illustrate raw-data echograms at 18, 38, 120, and 200 kHz (−70 dB threshold, according to the colour scale on the right hand side). The bottom left panel illustrates the effect of the multifrequency filter described in the middle text box (1). The resulting 38-kHz masked echogram (bottom left panel) is then subject to image analysis and extraction techniques, as defined in text box (2) using the SHAPES algorithm to isolate echotrace characteristics (echotrace *x*, with characteristic *i*, at frequency *n*). The orange line indicates the upper depth limit of the "sampol" (see definition later). At depths greater than this line, echotraces were believed to have been sampled by the pelagic trawl; therefore, these were isolated and labelled, and the data associated with them (x_i at *n* kHz) extracted.

allow for sophisticated digital-image processing (LeFeuvre *et al.*, 2000; Korneliussen and Ona, 2002). As a result of such developments, the work on fish-species identification using multifrequency data has expanded (e.g. Kang *et al.*, 2002; Brierley *et al.*, 2004; Gauthier and Horne, 2004; Korneliussen and Ona, 2004). Typically, the acoustic data are subject to algorithms that incorporate a number of processes, such as data correction (Korneliussen *et al.*, 2008), image analysis (Barange, 1994), and other data manipulations that isolate echotraces with particular (multifrequency) characteristics (Kloser *et al.*, 2002; Korneliussen and Ona, 2003).

Various statistical techniques have been used to classify the isolated echotrace objects. These include principal-component analyses and discriminant-function analysis (Nero and Magnuson, 1989; Vray *et al.*, 1990; Scalabrin *et al.*, 1996; Lawson *et al.*, 2001); artificial neural networks (Haralabous and Georgakarakos, 1996; Simmonds *et al.*, 1996); nearest-neighbour analyses (Richards *et al.*, 1991); *k*-mean clustering (Tegowski *et al.*, 2003); mixture models (Fleischman and Burwen, 2003); or comparisons of several of these methods (Woodd-Walker *et al.*, 2003; Hutin, *et al.*, 2005). These methods, although useful research exercises, have limited applicability in acoustic-survey practice, because they are either good classifiers, but difficult to apply (e.g. neural networks), or less powerful classifiers that can only be applied with some difficulty, but with lower performance and/or statistical assumptions that could be violated (e.g. DFA; see ICES, 2000).

One potentially suitable statistical method that has not yet been considered is the "classification tree" (Breiman *et al.*, 1984). This is rather surprising, because this method seems ideally suited to classification problems in fisheries acoustics; in fact, one of the first studies to use this technique was based on recognizing ship classes from their radar profiles (Hooper and Lucero, 1976), which is analogous to the case of fisheries acoustics. Classification trees are used to predict membership of objects in the classes of a categorical-dependent variable from their measurements on one or more predictor variables. In fisheries acoustics, the objects are the isolated fish-school echotraces, the class is the fish species, and the predictor variables are drawn from the list of morphometric, positional, and energetic parameters, as defined by Reid *et al.* (2000). Classification trees have been used widely to analyse ecological data (De'ath, 2007) and have many desirable properties that are suited to such data. They can take account of non-linear relationships between variables, high-order interactions, missing values, and lack of balance, and they deliver easy graphical interpretations of complex results (De'ath and Fabricius, 2000).

In this paper, a number of the techniques described above are combined, ultimately in a classification tree, with the objective

of identifying the major fish-school components of an acoustic survey of the North Sea. Classification trees are proposed as a robust, intuitive, and practicable method for objectively classifying subjects based on the available evidence, largely trawl samples, which also deliver some measure of the uncertainty ascribed to the process. Although the principal objective of this survey was to determine the abundance and distribution of North Sea herring (ICES, 2007), the classification tree also serves to identify the echotraces of other fish species and, equally important, the echotraces that are not herring. This is of great value, not only for the precision of the herring abundance estimate, but also, in due consideration of the ecosystem approach, for examining other components of the pelagic realm. The performance of the classification tree is determined both by the standard methods of comparing training and test datasets and by comparison with the expert-based scrutiny method.

Material and methods

Data collection

Data were taken from the Scottish component of the International North Sea Herring Acoustic Survey, conducted in July 2007 on the FRV "Scotia". This vessel is responsible for the northwestern sector that covers part of ICES Division IVa. Further details of survey procedures can be found in the survey reports (ICES, 2007). Acoustic data were collected during daylight with a Simrad EK60 scientific multifrequency echosounder, with four split-beam transducers operating at 18, 38, 120, and 200 kHz. The beam widths were all 7°, except for the 18-kHz transducer, which has a beam width of 11°. The transducers were mounted close together on the vessel's drop keel, which protruded 3 m below the hull during surveying. The acoustic data consisted of volume-backscattering strengths (VBS, in dB) at the four frequencies, collected simultaneously every second, which is equivalent to 5.1 m distance travelled at a survey speed of 10 knots. Here 1 ms pulse durations were used at all frequencies. All transducers were calibrated using a 38.1 mm, tungsten-carbide, standard target (Foote *et al.*, 1987) near the start of the survey.

As in any classification exercise, there is a requirement for a learning sample of measurement data on cases where the classification is known. In this case, the survey provided a learning sample by combining synchronous "ground truth" (biological) data from trawl catches. A pelagic trawl was directed at the observed echotraces to (i) obtain biological samples of herring to determine fish sizes for the conversion to abundance using the established *TS*–length equations, and ages to split the abundance estimate by age; and (ii) aid the scrutiny process for species identification. The speed of the vessel while trawling was 4–5.5 knots (2–2.8 m s^{-1}), at which time the trawl had a 15 m vertical opening and 20 m horizontal opening. Trawl depths varied between 75 and 175 m. The net had a 20 mm mesh in the codend. The trawl catches provided information on catch (species) composition in numbers, latitude and longitude, start and end times of the haul, as well as other parameters, such as the trawl depth and a description of the fished echotraces. In total, 48 trawls were conducted, of which 33 provided a sufficient sample to be considered representative (>30 individuals of any one species).

These trawling operations took place after completing a 15 min (2.5 nautical miles or 4.6 km) elementary distance sampling unit (EDSU; Simmonds and MacLennan, 2005), and involved temporarily interrupting the survey track to trawl on the observed echotraces. During this activity, additional acoustic data were collected that were not used in the survey estimate. Echotraces from the latter data can then be said to be "ground-truthed" and, together with acoustic data from the EDSU immediately before trawling, form the known classified echotrace cases (i.e. the learning sample). Acoustic data were, therefore, available from (i) trawling operations at 4 knots (Trawl dataset); and (ii) the survey at normal speed (10 knots), immediately before trawling (PreTrawl dataset).

Data analysis

Post-processing of the acoustic data was done using Sonardata Echoview software (Higginbottom *et al.*, 2000). This integrated the acoustic data with the ship's positional data, obtained from the GPS according to a synchronous time stamp. Echotraces of potential fish schools were isolated from the echogram in two steps (Figure 1). First, a fish-school filter was applied as summarized in the equation in box (1) in Figure 1. This filter, developed in the SIMFAMI project (Fernandes *et al.*, 2006), was implemented as an algorithm in Sonardata Echoview's virtual-variables module (Higginbottom *et al.*, 2000). Starting with the individual echograms, it removes VBS samples below the detected seabed and above 12 m (the nearfield), then sums all VBS samples across the four frequencies and applies a threshold (−226 dB) to the summed output. This threshold was determined empirically by inspection to achieve the "cleanest" echogram. The rationale behind this component of the algorithm is based on theoretical backscattering properties of many objects in the ocean (see Figure 1 of Lavery *et al.*, 2007). At the frequencies used here, many objects are generally in the Rayleigh or resonant regions, where backscattering varies significantly with frequency. Conversely, fish display geometric backscattering at these frequencies; therefore, their backscatter is more consistent with frequency. Summing, then thresholding by frequency, therefore, retains only those objects that persist on all frequencies, e.g. fish schools. The algorithm then applies convolution kernels to remove individual (small) samples with a median filter and augment larger collections of samples (echotraces) with a dilation filter. Finally, the resulting virtual echogram masks the original 38-kHz echogram with the threshold of −70 dB applied during the survey.

Step 2 (Figure 1) involved detecting the remaining echotraces on the masked 38-kHz echogram using the SHAPES algorithm in Echoview (Barange, 1994). The detection parameters used were a minimum total school length of 10 m, a minimum school height of 1 m, a minimum candidate length of 5 m, a minimum candidate height of 1 m, a vertical-linking distance of 2 m, and a maximum horizontal-linking distance of 15 m. An example of the resulting echogram from these two steps is given in Figure 1. The detected regions (echotrace boundaries) could then be used to isolate statistics of the VBS associated with the echotraces at the four frequencies.

At this stage, with 38-kHz data filtered to display only fish echotraces, the traditional scrutiny methods labelled each detected echotrace with a species category, for example, "definitely herring" labels were given to echotraces deemed by the experienced operator to be representative of either "classic" herring schools, that is to say, big, dense pillar or narrow tear shapes close to the seabed, or dense schools, close to trawl hauls that caught more than 85% herring. Then, "probably herring" labels were given to echotraces that have a similar appearance to the former category,

but were not actually fished, while "possibly herring" labels were given to those echotraces that have some relevant characteristics, but where the operator was not sure that they were herring. Finally, "surface herring" labels were given to echotraces close to the surface with the characteristics of herring schools. The standard survey-scrutiny practice involves an experienced operator labelling all the echotraces that have been isolated by the two-step procedure described in Figure 1.

As an alternative to this experienced-based scrutiny, a classification tree was built using the rpart package (Therneau and Atkinson, 1997) in the R statistical software language (R Development Core Team, 2008). The tree was built using the learning sample described above. The trawl-identified echotraces, isolated using the two-step procedure described in Figure 1, were used as the training set, and the echotraces observed before trawling were used as the testing set. In keeping with the procedures of this and many other acoustic surveys, where trawling is used to confirm the identity of echotraces, but acoustic data during trawling are not used for abundance estimation, it was considered prudent to base the construction of the tree on the echotraces with the greatest certainty (i.e. trawl-based), then to test the tree on those echotraces lacking identification by trawling (i.e. the pre-trawl acoustic data). In both cases, positional information was used to construct polygons to delimit the areas on the echograms sampled by the trawl. These sample polygons were designated as "sampols". All the isolated echotraces that fell within these samples were then selected as candidate objects to be classified.

The "true" classification of these objects was then based on the species composition in the 33 trawl hauls. Where >90% of the catch was one species, the echotraces were classed as that species (23 trawl hauls). In the other ten cases, the proportion of any one species was <90%; some expert-based scrutiny was then done to determine which echotraces could be ascribed to which species, based on the available evidence, e.g. trawl-sample proportion, dB-difference profile, echotrace shape, and intensity. This latter procedure was not ideal, but necessary because, apart from herring, it was rare that other species comprised >90% of the catch. There were also isolated echotraces for which no catch information was available, but it was clear that these belonged to unidentified scattering layers of unknown plankton (Mair *et al.*, 2005). These echotraces were classed as "layer". Overall, this procedure produced 716 schools in the training set (594 herring, 48 layer, 38 mackerel, and 36 myctophid). Of these, 163 were from the ten trawls where no single species comprised >90% of the catch. There were 738 schools in the test set (697 herring, 19 layer, 17 mackerel, and 5 myctophid). Of these, 226 were from the ten "less than 90% trawls", as above.

The predictor variables for the tree were chosen from a suite of 33 positional, morphometric, and energetic variables available for each frequency, analogous to the list produced by Reid *et al.* (2000). In addition to these, several differences in mean volume-backscattering strengths (MVBS) or "dB difference" variables (Fässler *et al.*, 2007) were added: $\Delta dB18 = MVBS_{38kHz} - MVBS_{18kHz}$; $\Delta dB120 = MVBS_{38kHz} - MVBS_{120kHz}$; and $\Delta dB200 = MVBS_{38kHz} - MVBS_{200kHz}$. Many of the exported variables were correlated or were derivations of each other (e.g. the area-backscattering coefficient and the nautical-area-scattering strength), so not all were useful for classification. Preliminary data exploration identified the variables most often selected in the construction of a tree (see below), so the first tree was constructed with the following 24 variables: mean $depth_{38kHz}$; mean $height_{38kHz}$; $length_{38kHz}$; $perimeter_{38kHz}$; $area_{38kHz}$; MVBS at 18, 38, 120, and 200 kHz; maximum VBS at 18, 28, 120, and 200 kHz; minimum VBS at 18, 28, 120, and 200 kHz; coefficient of variation of VBS at 18, 28, 120, and 200 kHz; ΔdB18; ΔdB120; and ΔdB200.

A classification tree is built by recursive partitioning (De'ath and Fabricius, 2000). The data in the training set were split into two groups, based on a binary threshold value for a particular variable. Then the variable and threshold that "best" split the data into two groups were chosen. The "best" split was determined by a measure of maximal impurity reduction (Therneau and Atkinson, 1997), effectively that which best isolated the subsequent classes into the most "pure" categories. This process was repeated on the remaining subgroups, and repeated until no improvement could be made to the partitioning (i.e. all classes were accounted for). The resulting tree was then tested by applying its classification rules to the test dataset. The test dataset's "true" classification was then compared with the tree-based classification in a confusion matrix.

Initially (treatment 1), all the data in the training set were used to build the tree. However, in the data-exploration process, it was found that the structure and size of the tree varied with slight changes in the input data, e.g. different numbers of herring schools. This was attributed to variability in the "truth" as interpreted from the trawl hauls. To account for this, a "forest" of 10 000 trees was built, each from the same set of input variables, but differing through random selections of the objects classed as herring in the training set. A second treatment then selected random sets of 43 herring schools, approximately equal in number to those of the other classes, for the 10 000 trees. This number of schools was chosen to produce a "balanced" dataset, which is often desirable in clustering statistics. However, the disadvantage with the latter selection is the very small number of schools, when so many more are available. De'ath (2007) defines a "small" dataset as one with fewer than 1000 observations, and ultimately this leads to poorer predictions. Therefore, a third treatment was based on using two-thirds of the herring schools ($n = 594$). This proportion, and much of the "forest" philosophy, was chosen based on the description of "cross-validation" by De'ath and Fabricius (2000), where random subsets (they suggested one-half to two-thirds of all data) are selected to build a sequence of trees, and the tree with the smallest predicted error is then selected. Of these 10 000 trees, the resulting confusion matrices were compared and one tree was chosen based on the best

Table 1. Confusion matrix for the classification tree in Figure 2 as applied to the test dataset from the survey done by FRV "Scotia" during the 2007 International North Sea Herring Survey.

Tree test	Herring	Layer	Mackerel	Mycto	Class. rate
Herring	654	0	42	1	0.94
Layer	0	19	0	0	1.00
Mackerel	4	0	13	0	0.76
Mycto	3	0	1	1	0.20
Total correct					0.93
Average correct					0.73
Mackerel and herring correct					0.90

Layer, planktonic scattering; Mycto, myctophids; Class. Rate, classification rate, i.e. the number of objects classified correctly divided by the total number of objects in a class; Total correct, total number of objects classified correctly/total number of objects; Average correct, sum of individual classification rates/number of object classes (i.e. 4).

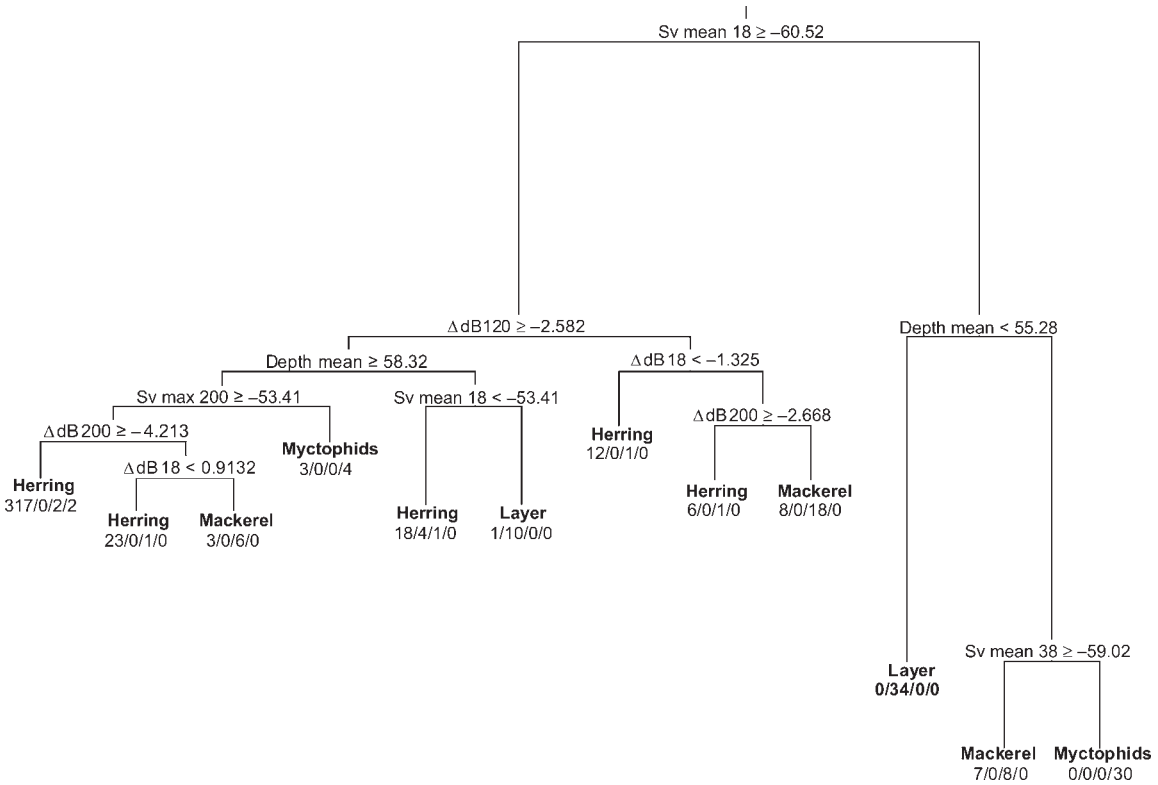

Figure 2. Classification tree for the identification of echotraces from the survey done by the FRV "Scotia" during the 2007 International North Sea Herring Survey, using a training set of 520 echotraces identified as herring, mackerel, and myctophids using a pelagic trawl; and a "layer" of planktonic scattering identified by expert-based scrutiny. Underneath the designated class at the terminal node (in bold) are the numbers of each class assigned to that node (Herring/Layer/Mackerel/Myctophids). Splitting variables are mean depth (m), Sv mean *n* (MVBS at *n* kHz, dB), Δ*n* (MVBS 38 kHz – MVBS *n* kHz), and Sv max (maximum VBS, dB).

average classification rate. This was calculated from the confusion matrix as the simple average of the individual species-classification rates (final column in Table 1), and it differs from the total classification rate, which is weighted towards the more dominant herring category.

Results

The classification tree using all the available data in the training set is not presented here. When applying it to the test dataset, it achieved a successful classification rate of 96.6% for herring, 47.1% for mackerel, 0% for myctophids, and 100% for objects classed as "layer". Overall, this equated to a total successful classification rate of 95%. Because this is dominated by the large number of herring schools, it is more meaningful to calculate the average classification rate across the important classes, 61% in this case. This is not particularly good, but when the myctophid class is omitted, which was based on one haul and very few schools, the average successful classification rate increased to 81%.

Trees with the best average classification rates were then selected from the forest of 10 000. The best tree to arise from the balanced dataset (treatment 2) resulted in an average classification rate of 69%, which was better than that of the original tree using all the available data. However, an even better tree was found in the forest built from a random selection of 67% of the available herring schools (treatment 3). This tree (Figure 2) had a total classification rate of 93%, similar to the initial tree, but an average classification rate of 73%, (cf. 61% for the original tree). Omitting the myctophid class produced an average classification rate of 90%, compared with 81% for the original tree. The confusion matrix for the final tree is given in Table 1.

The final tree, based on the third treatment (67% of herring schools), was then applied to all the survey data, to compare the tree-based system with the expert-based scrutiny. The resulting confusion matrix is given in Table 2. This reveals some interesting results. First, the experts had not identified the layer as a class, because it was not meant to be used in any subsequent assessment.

Table 2. Confusion matrix for the classification tree in Figure 2 as applied to the scrutinized data from the 2007 North Sea pelagic survey.

Tree scrutiny	Herring	Layer	Mackerel	Mycto	Agree. rate
Definitely herring	1 043	0	58	12	0.94
Probably herring	1 845	52	46	5	0.95
Possibly herring	0	6	0	0	0.00
Surface herring	474	8	16	1	0.95
Mackerel	22	3	5	0	0.17
UnID	469	42	48	4	
?	31	2	2	3	
Total agreed					0.80
Mackerel and herring agreed					0.75

Herring, classed as either definitely, probably, possibly, or surface (see text for an explanation of these terms); Layer, planktonic scattering; Mytco, myctophids; Agree. Rate, agreement rate; UnilD, unidentified fish schools; ?, unidentified echotraces; Total agreed, total number of objects classified in agreement/total number of objects.

The tree and the experts had >94% agreement for all herring schools. The exception to this was the surface-herring component. In this instance, there were no trawl samples of the surface-herring schools; therefore, this component was not built into the tree. Not surprisingly, the surface herring was interpreted by the tree as being the "layer", because this would be in shallow water, which is a characteristic of the "layer" category. In all, 83% of the unidentified fish schools were classed as herring in the tree. Second, for mackerel there was some disagreement. Only 17% of the mackerel schools were classified as mackerel by the tree, whereas 73% were classified as herring.

Discussion

As herring schools predominated by a large margin, a successful classification rate for herring was somewhat inevitable. Nevertheless, the tree serves to illustrate the important characteristics that define what is deemed to be a herring school—be it by objective trawl data or by the subjective scrutiny process. The final tree (Figure 2) has characteristics that one might expect to find to separate the various classes. Clearly, the layer occurred in shallow water, predominantly above the fish schools, shallower than 55 m (Figure 1). This accords with various observations of high scattering close to the surface, attributed to unidentified planktonic scatterers by Mair *et al.* (2005) and, in another case, to turbulence associated with internal waves (Warren *et al.*, 2003). Thereafter, the tree is dominated by splits that use the energetic variables and, specifically when splitting herring from mackerel, the dB-difference variables. The two most interesting splits that separate herring and mackerel are the 38–120 and 38–200 kHz dB differences. These are set at values that closely mirror those found by other studies that characterize the frequency response of mackerel schools (Korneliussen and Ona, 2004; Gorska *et al.*, 2005). These authors reported that backscattering by mackerel at 200 kHz is greater than that at 38 kHz, by an amount very similar to that indicated by the tree (i.e. ~3 dB). The same dB-difference for herring is, although variable, generally much less (Fässler *et al.*, 2007). The variability in frequency response, specifically among herring schools, however, means that discrimination only based on this dB-difference would be poor. In this study, there were many herring schools with 38–200 kHz dB differences of −3 dB or less. Therefore, it is not surprising that the tree has a more complex structure incorporating other variables, such as the maximum VBS at 200 kHz and dB-differences between 38 and 18 kHz, 38 and 120 kHz, and 38 and 200 kHz.

The myctophid class was based on a single trawl haul and the experts failed to identify it, because this category was outside their range of experience (Table 2). Currently, myctophids are not required to be assessed, and their schools were not particularly evident on the 38 kHz echogram, which is the one inspected most during the scrutiny process. They were more evident, incidentally, on the 120 kHz echogram.

The methods used here, in common with many classification applications, assume that the training set comprises a set of schools whose class is known. This is very difficult to achieve for an acoustic survey, because of the difficulty in obtaining proper ground-truth data on mobile organisms underwater (McClatchie *et al.*, 2000). Even when a sampling device can be directed towards a particular target, there is no guarantee that the selective properties of the device will represent the true composition of that target. Mesh-size selectivity (Bethke *et al.*, 1999), or behavioural effects (Misund and Aglen, 1992), may confound the catch data. To some degree, this can be mitigated by careful selection of trawl datasets, but such "scrutiny" invokes subjectivity that compromises the ambitions of objectivity of the approach. Examination of netsonde data can reveal if the intended targets were actually caught, for example.

An important factor concerns the coincidence of data. The acoustic and trawl samples were assumed to be derived from the same space and time. This is not always the case. Fish approached by a trawl often dive to avoid it (Suuronen *et al.*, 1997). To catch fish, therefore, the trawl position needs to be altered, usually to greater depth, with the guidance of the netsonde. In addition, the samples may not actually reflect the true location of the trawl catch, which is taken some distance behind the vessel. Because of such difficulties, it is almost impossible to obtain a pure ground-truth sample, so the assumption of pure catches is not necessarily achievable. In the present study, entirely pure catches were not a major assumption; instead, some variability in the ground truth was accepted and accounted for by building a forest of trees constructed from random subsets of the (uncertain) ground-truth data.

In this regard, the approach here differs slightly from many previous studies that based the identification on pure catches of the target species. Lawson *et al.* (2001) reported that pelagic fish schools off South Africa could be identified to species in 88.3% of all "known" cases; Brierley *et al.* (1998) reported a correct classification rate of 77% for Antarctic krill; and LeFeuvre *et al.* (2000) classified 99% of cod aggregations in Newfoundland correctly. Although these studies are subject to the same caveats as those described above, the reported success is often high when dealing with such a summary statistic. Testing a classification method in areas where two species "can be easily separated" (Vray *et al.*, 1990) may be informative to characterize echoes in those areas, but is not much of a test of a useful system to identify fish where the species cannot be easily separated. The results obtained here have achieved classification rates similar to the studies described above, though in a less challenging environment. These rates may be propagated through to estimates of error in fish abundances, as measures of uncertainty in the species-identification process. However, because they are based on individual echotraces, the rates would need to be weighted in some way by the nautical-area-scattering coefficient attributed to each echotrace. It may be useful, for example, to distinguish large herring schools from small ones, to ensure that those contributing most to the abundance estimate are identified separately. This would require further objective scrutiny of the training set, perhaps based on the current method of apportioning the data into "definitely herring" and "probably herring" schools.

An important advantage of the classification tree is its ability to incorporate this as a standard software tool for acoustic surveys. The tree illustrated in Figure 2 could be built into the Echoview virtual-variables module to classify echotraces accepted by the fish-school filter described in Figure 1. This would also allow real-time classification using Echoview's live-viewing feature.

The current study was almost certainly compromised by the lack of data on species other than herring. The principal objective of the survey from which these data were taken was to obtain an abundance index-at-age for herring. To do this effectively, many biological samples of herring are required, to (i) determine the length distribution so that the acoustic data may be converted to abundance using TS functions, and (ii) build an age–length key. Therefore, the primary need is to fish on herring, giving less

opportunity to sample other species. To have a greater range of species available, it was deemed necessary to allocate some of the echotraces based on less certain trawl information, that is to say, trawls where the species composition was $<90\%$. This introduces some subjectivity into what is supposed to be an objective process, but was the best compromise under these circumstances. As such, the classification tree is actually a mix of (i) an objective classification scheme based on that portion of the ground-truth data that is correct; and (ii) a description of the expert interpretation in a quantitative manner. The latter is probably another use of the tree. One could envisage, for example, basing the identification process purely on the subjective, expert-based scrutiny. Then, instead of having a set of words describing the current scrutiny process as outlined earlier, one would have a decision tree that sets out how each class was defined in numerical terms. This would allow the (current) subjective scrutiny process to be repeatable (e.g. by alternative operators), to have some measure of the uncertainty, and to be defined in numerical terms.

So far, the classification-tree system has not really been "tested" to the extent that it could be in, say, a more multispecies environment, such as the Bay of Biscay (Scalabrin *et al.*, 1996). Future studies should examine trawl samples from a series of similar surveys, perhaps going several years back in time, to establish a good number of alternative identified targets. Once a larger database of schools is available, a tree could be built with enough information to cope with the range of objects that need to be classified. In the meantime, developments in tree-based methods continue with the introduction of bagged trees and random forests (Breiman, 2001), which use principles similar to the random-subsetting approach described here, and boosted trees (De'ath, 2007), which use a sequential-weighting procedure to improve the tree. The possibilities for statistical refinement are considerable, but just as important is a careful selection and examination of the multifrequency data comprising the training set (Korneliussen *et al.*, 2008) and its interpretation (McClatchie *et al.*, 2000).

Conclusions

Classification trees are proposed as methods to identify echotraces of fish schools based largely on their multifrequency energetics and other features. When combined with image-analysis techniques that filter much of the irrelevant scattering in the water column, they can provide a tool for decision-making that is accountable, easily interpreted, and implementable via commercially available software. The method also ascribes a level of certainty that assesses the quality of the decision-making process. In the example presented here, using data from the International North Sea Herring Acoustic Survey, the method was compromised by the limited quality of data from known classes. In common with many classification applications, the tree method would be improved with the provision of better and more extensive ground-truth information.

Acknowledgements

Thanks are due to Aileen Shanks for suggesting the use of classification trees. Some of the initial work was done with support from the European Commission's Fifth Framework Programme (SIMFAMI project; Grant No. Q5RS-2001–02054). The author is indebted to colleagues at the Marine Laboratory, Aberdeen, who provided data and advice at various stages of this work, particularly John Simmonds, Phil Copland, and Stephen Keltz. The manuscript was improved by the need to consider and react to four rather diverse comments by referees and the guest editors.

References

Barange, M. 1994. Acoustic identification, classification and structure of biological patchiness on the edge of the Agulhas Bank and its relationship to frontal features. South African Journal of Marine Science, 14: 333–347.

Bethke, E., Arrhenius, F., Cardinale, M., and Håkansson, N. 1999. Comparison of the selectivity of three pelagic sampling trawls in a hydroacoustic survey. Fisheries Research, 44: 15–23.

Breiman, L. 2001. Random forests. Machine Learning, 45: 5–32.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. 1984. Classification and Regression Trees. Chapman and Hall, Boca Raton, FL.

Brierley, A. S., Axelsen, B. E., and Lynam, C. P. 2004. Pinging down the food web: multi-frequency acoustic discrimination of jellyfish and fish. ICES Document CM 2004/R: 6. 21 pp.

Brierley, A. S., Ward, P., Watkins, J. L., and Goss, C. 1998. Acoustic discrimination of Southern Ocean zooplankton. Deep Sea Research II, 45: 1155–1173.

Cushing, D. H. 1963. The Uses of Echo Sounding for Fishermen. HMSO, London.

De'ath, G. 2007. Boosted trees for ecological modelling and prediction. Ecology, 88: 243–251.

De'ath, G., and Fabricius, K. E. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology, 81: 3178–3192.

Fässler, S. M. M., Santos, R., García-Núñez, N., and Fernandes, P. G. 2007. Multifrequency backscattering properties of herring (*Clupea harengus*) and Norway pout (*Trisopterus esmarkii*). Canadian Journal of Fisheries and Aquatic Sciences, 64: 362–374.

Fernandes, P. G., Korneliussen, R. J., Lebourges-Dhaussy, A., Masse, J., Iglesias, M., Diner, N., Ona, E., *et al.* 2006. The SIMFAMI Project: Species Identification Methods from Acoustic Multifrequency Information. Final Report to the EC, Q5RS-2001–02054.

Fleischman, S. J., and Burwen, D. L. 2003. Mixture models for the species apportionment of hydroacoustic data, with echo-envelope length as the discriminatory variable. ICES Journal of Marine Science, 60: 592–598.

Foote, K. G., Knudsen, H. P., Vestnes, G., MacLennan, D. N., and Simmonds, E. J. 1987. Calibration of acoustic instruments for fish density estimation: a practical guide. ICES Cooperative Research Report, 144. 57 pp.

Gauthier, S., and Horne, J. K. 2004. Potential acoustic discrimination within boreal fish assemblages. ICES Journal of Marine Science, 61: 836–845.

Gorska, N., Ona, E., and Korneliussen, R. 2005. Acoustic backscattering by Atlantic mackerel as being representative of fish that lack a swimbladder. Backscattering by individual fish. ICES Journal of Marine Science, 62: 984–995.

Haralabous, J., and Georgakarakos, S. 1996. Artificial neural networks as a tool for species identification of fish schools. ICES Journal of Marine Science, 53: 173–180.

Higginbottom, I. R., Pauly, T. J., and Heatley, D. C. 2000. Virtual echograms for visualization and post-processing of multiple-frequency echosounder data. *In* Fifth European Conference on Underwater Acoustics, ECUA 2000, pp. 1497–1502. Ed. by M. E. Zakharia. eesd@cec.eu.int.

Hooper, R., and Lucero, A. 1976. Radar profile classification: a feasibility study. Technical Report Technology Service Corporation, Santa Monica, CA.

Horne, J. 2000. Acoustic approaches to remote species identification: a review. Fisheries Oceanography, 9: 356–371.

Hutin, E., Simard, Y., and Archambault, P. 2005. Acoustic detection of a scallop bed from a single-beam echosounder in the St. Lawrence. ICES Journal of Marine Science, 62: 966–983.

ICES. 1998. Report of the Study Group on Echo Trace Classification. ICES Document CM 1998/B: 1. 58 pp.

ICES. 2000. Report on Echo Trace Classification. ICES Cooperative Research Report, 238. 107 pp.

ICES. 2007. Report of the Planning Group for Herring Surveys (PGHERS), 22–26 January, 2007, Charlottenlund, Denmark. ICES Document CM 2007/LRC: 01. 187 pp.

Kang, M., Furusawa, M., and Miyashita, K. 2002. Effective and accurate use of difference in mean volume backscattering strength to identify fish and plankton. ICES Journal of Marine Science, 59: 794–804.

Kloser, R., Ryan, T., Sakov, P., Williams, A., and Koslow, J. A. 2002. Species identification in deep water using multiple acoustic frequencies. Canadian Journal of Fisheries and Aquatic Sciences, 59: 1065–1077.

Korneliussen, R. J., Diner, N., Ona, E., Berger, L., and Fernandes, P. G. 2008. Proposals for the collection of multifrequency acoustic data. ICES Journal of Marine Science, 65: 982–994.

Korneliussen, R. J., and Ona, E. 2002. An operational system for processing and visualizing multi-frequency acoustic data. ICES Journal of Marine Science, 59: 293–313.

Korneliussen, R. J., and Ona, E. 2003. Synthetic echograms generated from the relative frequency response. ICES Journal of Marine Science, 60: 636–640.

Korneliussen, R. J., and Ona, E. 2004. Verified acoustic identification of Atlantic mackerel. ICES Document CM 2004/R: 20. 14 pp.

Lavery, A. C., Wiebe, P. H., Stanton, T. K., Lawson, G. L., Benfield, M. C., and Copley, N. 2007. Determining dominant scatterers of sound in mixed zooplankton populations. Journal of the Acoustical Society of America, 122: 3304–3326.

Lawson, G. L., Barange, M., and Fréon, P. 2001. Species identification of pelagic fish schools on the South African continental shelf using acoustic descriptors and ancillary information. ICES Journal of Marine Science, 58: 275–287.

LeFeuvre, P., Rose, G. A., Gosine, R., Hale, R., Pearson, W., and Khan, R. 2000. Acoustic species identification in the Northwest Atlantic using digital image processing. Fisheries Research, 47: 137–147.

Lundgren, B., and Nielsen, J. R. 2008. A method for the possible species discrimination of juvenile gadoids by broad-bandwidth backscattering spectra vs. angle of incidence. ICES Journal of Marine Science, 65: 581–593.

MacLennan, D. N., and Holliday, D. V. 1996. Fisheries and plankton acoustics: past, present, and future. ICES Journal of Marine Science, 53: 513–516.

Madureira, L. S. P., Everson, I., and Murphy, E. J. 1993. Interpretation of acoustic data at two frequencies to discriminate between Antarctic krill (*Euphausia superba* Dana) and other scatterers. Journal of Plankton Research, 15: 787–802.

Mair, A. M., Fernandes, P. G., Lebourges-Dhaussy, A., and Brierley, A. 2005. An investigation into the zooplankton composition of a prominent 38-kHz scattering layer in the North Sea. Journal of Plankton Research, 27: 623–633.

McClatchie, S., Thorne, R. E., Grimes, P., and Hanchet, S. 2000. Ground truth and target identification for fisheries acoustics. Fisheries Research, 47: 173–191.

Misund, O. A., and Aglen, A. 1992. Swimming behaviour of fish schools in the North Sea during acoustic surveying and pelagic trawl sampling. ICES Journal of Marine Science, 49: 325–334.

Nero, R. W., and Magnuson, J. J. 1989. Characterization of patches along transects using high-resolution 70-kHz integrated acoustics data. Canadian Journal of Fisheries and Aquatic Sciences, 46: 2056–2064.

R Development Core Team. 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Reid, D. G., Scalabrin, C., Petitgas, P., Masse, J., Aukland, R., Carrera, P., and Georgakarakos, S. 2000. Standard protocols for the analysis of school based data from echosounder surveys. Fisheries Research, 47: 125–136.

Richards, L. J., Kieser, R., Mulligan, T. J., and Candy, J. R. 1991. Classification of fish assemblages based on echo integration surveys. Canadian Journal of Fisheries and Aquatic Sciences, 48: 1264–1272.

Rose, G. A., and Leggett, W. C. 1988. Hydroacoustical signal classification of fish schools by species. Canadian Journal of Fisheries and Aquatic Sciences, 45: 597–607.

Scalabrin, C., Diner, N., Weill, A., Hillion, A., and Mouchot, M. C. 1996. Narrowband acoustic identification of monospecific fish shoals. ICES Journal of Marine Science, 53: 181–188.

Simmonds, E. J., Armstrong, F., and Copland, P. J. 1996. Species identification using wideband backscatter with neural network and discriminant analysis. ICES Journal of Marine Science, 53: 189–195.

Simmonds, E. J., and MacLennan, D. N. 2005. Fisheries Acoustics: Theory and Practice, 2nd edn. Blackwell Publishing, Oxford.

Suuronen, P., Lehtonen, E., and Wallace, J. 1997. Avoidance and escape behaviour by herring encountering midwater trawls. Fisheries Research, 29: 13–24.

Tegowski, J., Gorska, N., and Klusek, Z. 2003. Statistical analysis of acoustic echoes from underwater meadows in the eutrophic Puck Bay (southern Baltic Sea). Aquatic Living Resources, 16: 215–221.

Therneau, T. M., and Atkinson, E. J. 1997. An introduction to recursive partitioning using the RPART routines. http://mayoresearch.mayo.edu/mayo/research/biostat/splusfunctions.cfm.

Vray, D., Gimenez, G., and Person, R. 1990. Attempt at classification of echo-sounder signals based on the linear discrimination function of Fisher. Rapports et Procès-Verbaux des Réunions du Conseil International pour l'Exploration de la Mer, 189: 388–393.

Warren, J. D., Stanton, T. K., Wiebe, P. H., and Seim, H. E. 2003. Inference of biological and physical parameters in an internal wave using multiple-frequency, acoustic-scattering data. ICES Journal of Marine Science, 60: 1033–1046.

Woodd-Walker, R. S., Watkins, J. L., and Brierley, A. S. 2003. Identification of Southern Ocean acoustic targets using aggregation backscatter and shape characteristics. ICES Journal of Marine Science, 60: 641–649.

doi:10.1093/icesjms/fsp060