Classification of Southern Ocean krill and icefish echoes using Random Forests

Published in ICES Journal of Marine Science (2016) - DOI:10.1093/icesjms/fsw057

Niall G Fallon[1] (corresponding author), Sophie Fielding[2], Paul G Fernandes[1]

[1]Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen, AB24 2TZ, UK; [2]British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge, CB3 0ET, UK

Email - n.fallon@abdn.ac.uk, phone - +44 (0)1224 272648

Abstract

Target identification remains a challenge for acoustic surveys of marine fauna. Antarctic krill, *Euphausia superba*, are typically identified through a combination of expert scrutiny of echograms and analysis of differences in mean volume backscattering strengths (S_V; dB re ^{1} m^{-1}) measured at two or more echosounder frequencies. For commonly used frequencies, however, the differences for krill are similar to those for many co-occurring fish species that do not possess swim bladders. At South Georgia, South Atlantic, one species in particular, mackerel icefish, *Champsocephalus gunnari*, forms pelagic aggregations, which can be difficult to distinguish acoustically from large krill layers. Mackerel icefish are currently surveyed using bottom-trawls, but the resultant estimates of abundance may be biased because of the species' semi-pelagic distribution. An acoustic estimate of the pelagic component of the population could indicate the magnitude of this bias, but first a reliable target identification method is required. To address this, random forests were generated using acoustic and net sample data collected during surveys. The final random forest classified krill, icefish, and mixed aggregations of weak scattering fish species with an overall estimated accuracy of 95%. Minimum S_V, mean aggregation depth (m), mean distance from the seabed (m) and geographic positional data were most important to the accuracy of the random forest. Time-of-day and the difference between S_V at 120 kHz ($S_{V\ 120}$) and that at 38 kHz ($S_{V\ 38}$) were also important. The random forest classification resulted in significantly higher estimates of backscatter apportioned to krill when compared to widely applied identification methods based on fixed and variable ranges of $S_{V\ 120}$-$S_{V\ 38}$. These results suggest that krill density is underestimated when those S_V-differencing methods are used for target identification. Random forests are an objective means for target identification, and could enhance the utility of incidentally collected acoustic data.

Key words: acoustics, target identification, fish survey, South Georgia

Introduction

Mackerel icefish, *Champsocephalus gunnari*, hereafter "icefish", is a semi-pelagic finfish occurring across shelf areas in the Southern Ocean (Kock, 2005a). The population at South Georgia, South Atlantic, is the target of a commercial pelagic trawl fishery constrained by quotas of 1000 to 5000 tonnes per season, in recent years (Barnes *et al.*, 2011; CCAMLR, 2014). Icefish are assessed using bottom trawl surveys which may yield biased estimates of abundance as a result of limited availability to the sampling method due to pelagic feeding migrations undertaken by the species (Fallon *et al.*, 2015; Hill *et al.*, 2005, 2012). Adaptive acoustic-trawl surveys (Everson *et al.*, 1996), or other implementations of combined acoustic-trawl survey (Kotwicki *et al.*, 2013; McQuinn *et al.*, 2005) have the potential to address this issue.

The hypothesis that bias in icefish abundance estimates results from their vertical distribution can be explored using data from an echosounder. Acoustic data can be collected concurrently with bottom trawling (Bez *et al.*, 2007) to estimate the density of fish which are unavailable to the trawl (e.g. Aglen *et al.*, 1999). However, to incorporate acoustic estimates into the assessment of the population, backscatter from icefish must first be identified (Horne, 2000). When attributing acoustic data to species, a number of spatial scales can be considered (e.g. that of the school; the elementary distance sampling unit, EDSU; or the region of interest; Reid *et al.*, 2000). Distinguishing between groups of objects with different scattering properties (e.g. fish or plankton with or without gas inclusions) is often achievable using data processing on an EDSU or regional scale (Korneliussen *et al.*, 2009; Madureira *et al.*, 1993). This typically involves resampling acoustic data across some range of depth and distance or time, followed by classification according to assumptions regarding scattering properties of the group or groups of interest (Hewitt *et al.*, 2004; Madureira *et al.*, 1993). Assumptions are based on the backscatter versus frequency, the frequency response, of the target organism. This is a function of its orientation relative to the incident sound wave, the incidence angle, as well as its size and composition (Korneliussen and Ona, 2003). Classification may also depend on the target location and depth, associated seabed type, or other distributional co-variates (Reid *et al.*, 2000). However, organism aggregations are often geometrically complex, and resampling methods can degrade identifying characteristics (Reid and Simmonds, 1993). A school-level analysis preserves finer spatial-scale information, which could improve classification accuracy, and avoid any problems which might arise from several different target types occurring in a single EDSU.

Although the acoustic scattering properties of icefish need further study, information can be inferred from physical characteristics, which will aid in the identification of candidate echoes. Icefish lack swim bladders, so the frequency response could be similar to that of mackerel (Korneliussen, 2010): dominated by a flesh component at lower frequencies (e.g. 38 kHz), and by a bone component at higher frequencies (e.g. 120 kHz; see Gorska *et al.*, 2007). The flesh component should be relatively frequency independent across the typical operating frequencies (38-200 kHz) and may vary according to factors such as temperature and individual condition. The bone component would be characterised by a rising frequency response, peaking at ~200 kHz, varying with fish orientation (Gorska *et al.*, 2005; Korneliussen, 2010). The frequency response of icefish schools may therefore be low and flat at lower echosounder frequencies (38-100 kHz) relative to 120 and 200 kHz (Gorska *et al.*, 2007). Krill (*Euphausia superba*), icefish, and much of the South Georgia groundfish assemblage have similar frequency responses across commonly used frequencies (i.e. 38, 120 and 200 kHz), and therefore may be indistinguishable on an echosounder display (Collins *et al.*, 2008; Kock and Kellermann, 1991; Kock, 2005a; Lavery *et al.*, 2007). When such similarities exist, non-acoustic characteristics may be more important to accurate classification (Reid and Simmonds, 1993). Therefore, the data processing and analysis should incorporate all available variables.

Ideally, an objective target identification method should be applied due to the extensive training required for an operator to consistently and objectively identify a given species (Fernandes, 2009; Horne, 2000). In the Southern Ocean, Antarctic krill density and distribution is routinely estimated using the difference in volume backscattering strength (S_V; dB re 1 m^{-1}) measured at multiple frequencies (CCAMLR, 2010; Madureira *et al.*, 1993). Initially, a constant range of S_V measured at 120 kHz ($S_{V\ 120}$) minus S_V measured at 38 kHz ($S_{V\ 38}$) was used (Hewitt *et al.*, 2004; Madureira *et al.*, 1993). This has changed to include variable ranges of differences between $S_{V\ 38}$, $S_{V\ 120}$, and $S_{V\ 200}$ (Fielding *et al.*, 2014; Reiss *et al.*, 2008). However, these methods are typically applied at the EDSU level and may not differentiate well between species at the school level (Lawson *et al.*, 2008). The latter may require additional classification rules regarding target location, depth, or time-of-day. Woodd-Walker *et al.* (2003) compared an S_V-difference method with school-level classification of plankton using discriminant function analysis (DFA) and artificial neural networks (ANN). Although reasonable classification results were attained for krill, classifications for other groups in the analysis had higher error rates. In addition, the DFA required some transformation of variables to account for non-normality, and a simplified ANN had to be used because only a small training dataset was available. Tree-based methods (e.g. classification trees, bagged trees, random forests; Breiman, 2001; Hastie *et al.*, 2009) have also been explored as a means for acoustic target identification, and have yielded promising results albeit in a small number of case studies (D'Elia *et al.*, 2014; Fernandes, 2009).

The objective of this study is to explore and develop a random forest (RF) method for discrimination between weak acoustic-scattering organisms at South Georgia. Given the many varied properties of trawl-verified echoes collected during fish surveys, a RF approach is employed to distinguish between three classes of echoes (krill, mackerel icefish, and mixed groundfish). Tree-based classification methods do not require variables to be linear, and can be used to process large, high-dimensional datasets efficiently. In addition, RF classification accuracy is not affected by correlations or interactions between variables (James *et al.*, 2013). Further to the development of the method, the RF algorithm is tested against fixed and variable S_V-difference approaches (Fielding *et al.*, 2014; Madureira *et al.*, 1993) to compare outcomes. The intention of this comparison is to explore whether the alternative methods may overestimate the amount of backscatter attributable to krill due to the inclusion of backscatter from all weak scatterers, including icefish.

Methods

Data Sources

Data were from South Georgia groundfish surveys, conducted during the Austral summers of 2004-2006, and the Austral winter of 2007 aboard the Fisheries Patrol/Research Vessel (FPRV) *Dorada* (Figure 1). The surveys followed a stratified design across five areas (Mitchell *et al.*, 2010), in which icefish density (kg km^{-2}) is estimated for two depth strata, 50-200 m and > 200 m (generally < 300 m) using demersal trawl data (FP-120 trawl net; Pilling and Parkes, 1995). At the end of each of these surveys, a small number of pelagic hauls (International Young Gadoid Pelagic Trawl) targeted krill swarms and pelagic aggregations of icefish. During these surveys, echosounders (Simrad EK500) collected $S_{V\ 38}$ and $S_{V\ 120}$ following synchronized 1.0-ms pulse transmissions every 2.170 s. The echosounders were calibrated using a standard 38.1-mm

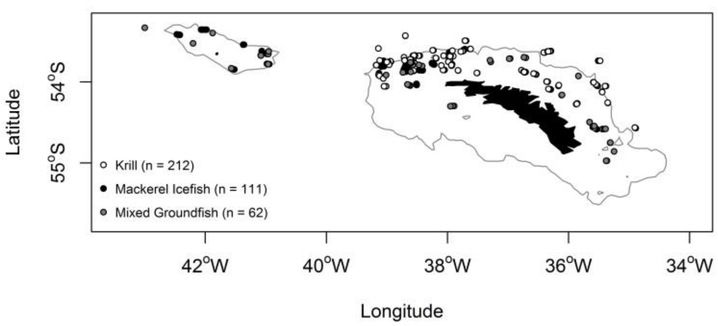

Figure 1. Locations of trawl-verified echoes used in the training dataset for generating random forests.

diameter tungsten-carbide sphere (Foote *et al.*, 1987), during each survey, at Husvik, Stromness Bay. In the Austral summers of 2010-2013, krill abundance was surveyed on the South Georgia shelf aboard the Royal Research Ship (RRS) *James Clark Ross* (*JCR*). During these surveys, $S_{V\ 38}$, $S_{V\ 120}$, and $S_{V\ 200}$ (Simrad EK60) were collected following 1.024-ms pulse transmissions every 2 s. The scatterers of interest were sampled using a Rectangular Midwater Trawl (RMT8; Fielding *et al.*, 2014). The echosounders were calibrated using copper spheres (Foote *et al.*, 1987), during each survey, at Stromness Harbour (see Fielding *et al.*, 2014 and supplementary material Table S1 for more details).

Post-processing of Echosounder Data

The echosounder data were post-processed using commercial software (Echoview, Sonardata; Higginbottom *et al.*, 2000). Aboard FPRV Dorada, the transmit power for the 120-kHz pulses was 1000 W instead of the recommended 250 W (Korneliussen *et al.*, 2008), which likely caused nonlinear distortion in the collected data. A nonlinearity correction factor was thus applied to the $S_{V\ 120}$ data to compensate for nonlinear distortion. The correction factor was derived as the simulated ratio of S_V corrected for nonlinear attenuation to measured S_V, where finite amplitude effects were assumed to be influential during both echosounder calibration and survey data collection due to high transmit power (see Lunde and Pedersen, 2012, and Pedersen, 2006, for further details). A multifrequency threshold (similar to that used in Fernandes, 2009) was applied to the S_V data as a series of virtual echograms (Higginbottom *et al.*, 2000) to remove data outside of animal aggregations from the analysis for the sole putrpose of improving on the single frequency threshold normally required for “school” detection using the Shapes algorithm (Coetzee, 2000). Single frequency S_V data, thresholded at -70 dB, were summed across all available frequencies (ICES, 2015). Thresholds for these virtual echograms, determined empirically to retain schools and eliminate non-school echoes, were -135 dB for $S_{V\ 38} + S_{V\ 120}$ and -240 dB for $S_{V\ 38} + S_{V\ 120} + S_{V\ 200}$. A 5x5 median convolution kernel, giving each pixel in the acoustic data matrix the median value of the surrounding set of 5x5 pixels, was then applied to remove single target observations and noise spikes (Fielding *et al.*, 2014). A 7x7 dilation convolution kernel (giving the maximum value in each 7x7 set of pixels) was then applied to the summed S_V data to mitigate any removal of data within schools by the other filtering steps. Finally a bitmap was used to mask the S_V data, removing data outside of schools from the analysis and retaining data assumed to originate from aggregations of organisms.

The SHAPES school detection algorithm (Barange, 1994) was then applied to the virtual echograms arising from the image analysis steps described above. The SHAPES parameters were: minimum total school length = 5 m; minimum school height = 1 m; minimum candidate length = 5 m, minimum candidate height = 1 m, maximum vertical-linking distance = 5 m, and maximum horizontal-linking distance = 20 m. The school polygons defined by the algorithm were then used to compile variables associated with each school, to serve as a training dataset for the purpose of classification. Echoes were assigned one of the following categories according to trawl composition data, assuming that the composition of echoes is represented by the complementary evidence collected by trawl: "Krill" schools, or swarms, were 100% *Euphausia* spp., almost exclusively *Euphausia superba*; "Mackerel Icefish" schools were >85% *C. gunnari*; and "Mix" were mixed aggregations of groundfish without swim bladders, consisting of <85% of any single fish species. Aggregations including fish possessing swim bladders (e.g. myctophid species such as *Electrona carlsbergi*) were excluded from the analysis. The inclusion of "Mix" was necessary to represent the wide assemblage of weak scattering species present in the area, to avoid misclassification of backscatter as "Mackerel Icefish" which occupies an overlapping location-depth niche.

The Random Forest Algorithm

All of the variables exported from the acoustic data were evaluated for collinearity, to identify superfluous variables that might be discarded. The final vector of variables (*p*) consisted of: mean $S_{V\ 120}$, maximum $S_{V\ 120}$, minimum $S_{V\ 120}$, standard deviation of $S_{V\ 120}$, $S_{V\ 120}$ skewness, mean height of school (m), mean aggregation depth (m), mean distance from seabed (m), latitude at centre of school, longitude at centre of school, corrected school length (m), corrected school thickness (m), corrected school perimeter (m), corrected school area (m), attack angle (°; Diner, 2001), image compactness (a ratio of the perimeter to the area of a school), corrected mean amplitude (m^2 m^3), horizontal roughness coefficient (Nero and Magnuson, 1989), $S_{V\ 120} - S_{V\ 38}$, time-of-day, and estimated school volume assuming a cylindrical shape (m^3). An RF was then generated using this training dataset (Breiman, 2001). Each tree within a RF was generated by recursive partitioning of the data, using the best splitting variable from a vector *m* randomly selected from *p* to partition the data at each node on the b^{th} tree (T_b), where *m* was of length $2 \times \sqrt{p}$. Vectors (*m*) of length $\sqrt{p}$ and $\sqrt{p}/2$ were also tested, but resulted in higher error rates. Nodes were split until they reached a specified minimum number of echoes (n_{min}) of n=1. The RF was then used to make predictions according to:

$$\hat{C}^{B}_{rf}(x) = \text{mode}\left\{\hat{C}_b(x)\right\}^{B}_{1} \qquad (1)$$

where $\hat{C}_b(x)$ is the classification prediction of the b^{th} tree in the ensemble of B = 1 x 10^4 trees, and $\hat{C}^{B}_{rf}(x)$ is the prediction of the RF.

Out-of-bag (OOB) error estimates were inspected as a means of cross-validation of prediction accuracy (Breiman, 2001; Hastie *et al.*, 2009). In addition to the RF generated using all available variables, RFs were generated using acoustically derived variables only (to explore how well the method might be generalised to other regions in the Southern Ocean), and using variables from schools around the main South Georgia shelf only (i.e. excluding Shag Rocks where krill data was not collected).

Confusion matrices were generated from OOB classifications, providing both overall and class-specific estimates of generalisation error. The kappa statistic (κ; Cohen, 1960) was used to measure classification performance by indicating the proportion of classification agreement beyond that expected to occur by chance. Variable importance was examined to assess the ranked importance of each variable to classification accuracy. The two typical measures of variable importance were calculated for the RF: mean decrease in accuracy, and the mean decrease in Gini Importance Index (left and right panels, respectively, in supplementary material Figure S1; Breiman, 2001). The first gives a measure of the decrease in prediction accuracy when the best node splitting variable is randomly permuted for all variables in *p*. The mean decrease in accuracy across all trees gives a measure of variable importance (Breiman, 2001). Secondly, the Gini Impurity Criterion (GIC) is a measure of the rate of misclassification of randomly chosen elements of a given node when classified according to the distribution of classes in its daughter node. The sum of decreases in the GIC for each variable across all trees results in a Gini Importance Index (GII). As these two measures may be biased by correlated variables (Strobl *et al.*, 2008), a third measure of conditional variable importance was calculated to verify their validity (Figure 2). The RF analyses were implemented in the R software environment using the “randomForest” and “party” packages (Liaw and Wiener, 2002; R Development Core Team, 2015; Strobl *et al.*, 2009).

Comparison of Methods

Other methods for krill identification were also used to apportion backscatter to weak scatterers, i.e. krill, icefish, and other fish species without swim bladders (Madureira *et al.*, 1993). Acoustic data collected during the course of the 2006 South Georgia groundfish survey was resampled to mean values within 5-m vertical by 100-m horizontal data bins (Demer, 2004; Fielding *et al.*, 2014). It was then assumed that resampled values of $S_{V\ 120} - S_{V\ 38}$ which fell within the range of 2-12 dB represented bins in which weak scattering targets which might be classified as krill would be found (as applied in Fielding *et al.*, 2014 & Woodd-Walker *et al.*, 2003). This method was also applied using a wider range, 2-16 dB (Watkins and Brierley, 2002). A third range, 0.37-12 dB, was also tested, based on the values used in the application of the variable window method (Fielding *et al.*, 2014), although the accuracy of this approach would likely be improved with the availability of additional frequencies. Data was then integrated from 12 m below the transducer to 0.5 m above the echosounder-detected seabed to give nautical area scattering coefficient (s_A; m^2 nmi^{-2}) values per 1-nmi EDSU. The derived s_A would be classified as krill within the integration volume, according to Madureira *et al.* (1993) and Fielding *et al.*, (2014), but that energy could have been reflected by many weak scatterers. The 2006 survey data was also classified using the above RF method. Integration over each region defined by the SHAPES algorithm gave s_A apportioned to each RF classification group for each EDSU. As $S_{V\ 200}$ data was not available in all datasets, it was not considered in this part of the analysis.

Results

Trawl-verified echoes across the three classification categories exhibited a range of variability in morphological, spatial (both vertical and horizontal) and acoustic properties. Krill echoes are the most highly studied of the three classes, and are known to exhibit temporal and spatial variability across a number of descriptors, including estimated density and echo morphology (Klevjer *et al.*, 2010; Tarling *et al.*, 2009). Krill echoes verified in the trawl data were broadly similar to those described elsewhere. Krill were most often found in discrete, dense swarms, which were relatively easily visually identifiable, given

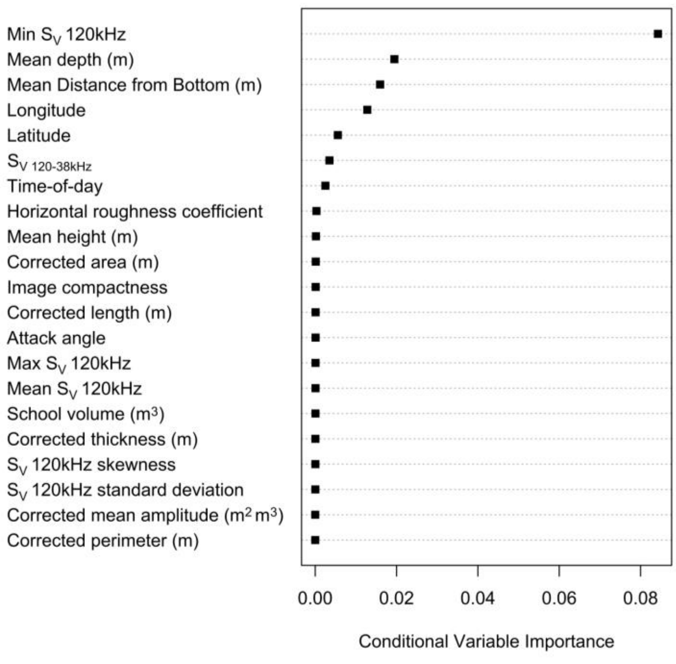

Figure 2. Conditional variable importance plot for the random forest using the full training dataset.

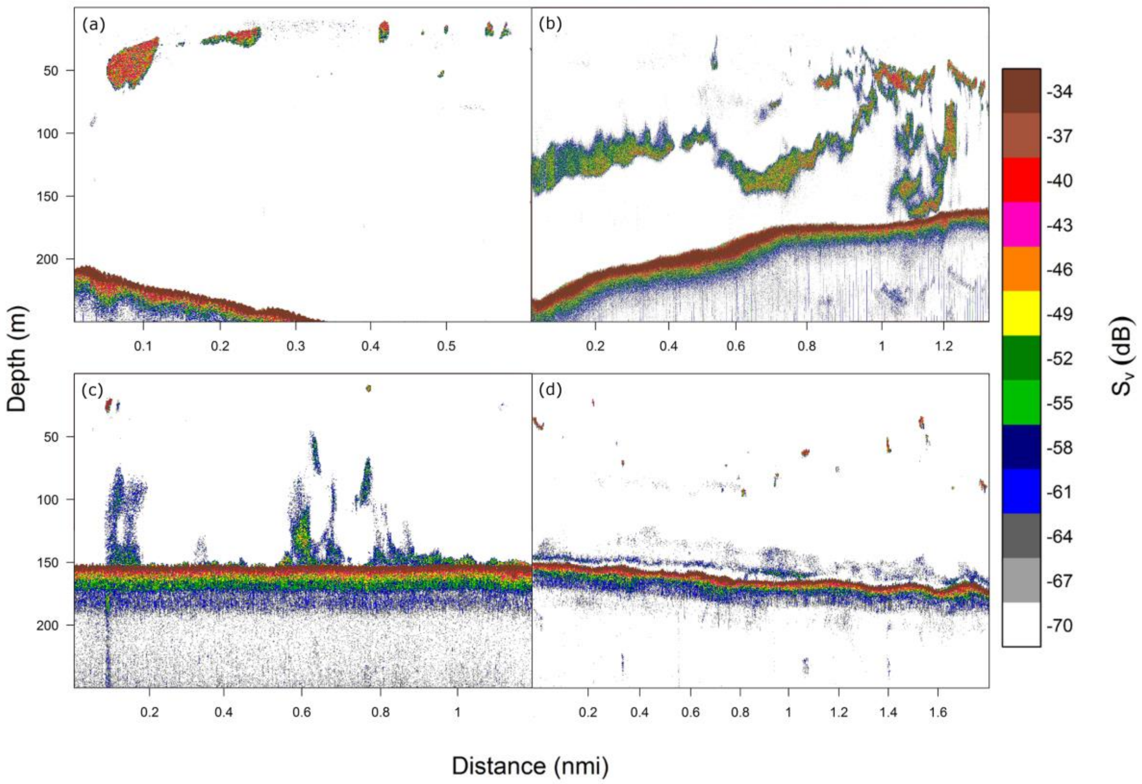

Figure 3. (a) Krill (*Euphausia superba*) echo from the JR245 research cruise. Echoes such as this, discrete dense backscatter formations in a relatively shallow position in the water column, are typically easy to distinguish as krill; (b) Krill (*Euphausia superba*) echo from the JR245 research cruise. Large, dynamic echoes were less typical of krill and more difficult to visually distinguish from pelagic icefish echoes; (c) Echo from the 2006 South Georgia groundfish survey. Pelagic trawl catches targeting dynamic echoes extending up to ~150 m from the seabed included only mackerel icefish. Krill was caught when the dense echoes <50m below the surface were targeted during a separate trawling event; (d) Echo from the 2006 South Georgia groundfish survey. Relatively weak backscattering close to the seabed, such as this, was typical of mixed groundfish trawls, although more extensive and dynamic aggregations were observed in a minority of cases. Targeting this aggregation with a demersal trawl yielded a catch comprised of mostly humped rockcod (*Gobionotothen gibberifrons*), blackfin icefish (*Chaenocephalus aceratus*) and South Georgia icefish (*Pseudochaenichthys georgianus*). All echoes were generated from 120 kHz S_V data thresholded at -70 dB.

some experience (Figure 3a). However, more dynamic and patchy echoes were also observed, which could be mistakenly associated with other weakly scattering organisms (Figure 3b). This ambiguity is exemplified by a trawl-verified echo from the 2006 groundfish survey (Figure 3c), where a monospecific catch of icefish was obtained from the scatterers rising as much as 100 m above the seabed. Krill was caught during a separate haul targeting the relatively small dense scatterer aggregations <50 m below the surface. The fish and krill echoes in this example are difficult to visually distinguish with certainty (e.g. Figure 3b). Mixed groundfish typically formed more diffuse aggregations extending <20 m from the seabed (Figure 3d), but were also observed to

Table 1. Confusion matrix for random forest generated using the full trawl-verified dataset, with class-specific estimates of generalisation error.

	Predicted			Class Generalisation Error
Actual	Mackerel Icefish	Krill	Mix	
Mackerel Icefish	104	3	3	5.45%
Krill	6	206	3	3.73%
Mix	4	1	57	8.06%

form denser, more extensive echoes in some cases.

A value of κ = 0.92, 95% confidence interval ±0.04, was calculated from the RF confusion matrix (Table 1), where values of κ > 0.75 are considered an indication of an excellent classifier (Fielding and Bell, 1997). The total OOB estimate of error rate (i.e. the ratio of the sum of misclassified echoes from each category to the total number of samples) gave an estimate of overall prediction accuracy for the full RF of 95.08%. The top seven variables in order of importance for both indices were identical, although the order was different (supplementary material Figure S2 shows an alternative means of visualising the contribution of each variable to classification; Welling *et al.*, 2015). The most important variable using each metric was the minimum $S_{V\ 120}$ (dB). The next four most important variables were those pertaining to position, depth and time-of-day. The remaining variables related to measures of the acoustic and geometric properties of echoes whose order of importance varied in each case. The order of importance suggests that the use of acoustic descriptors alone is not a comprehensive basis for target identification. It is noteworthy that the distributions of $S_{V\ 120} - S_{V\ 38}$ values exhibited substantial overlap across all three groups (Figure 4), although Kolmogorov-Smirnov tests detected significant differences between them ($p < 0.05$). Crucially, the fixed 2-12 dB range, which designates backscatter as krill in the Madureira *et al.* (1993) method, only accounted for approximately 61% of the trawl-verified krill echoes. The RF models using only acoustically derived variables and South Georgia shelf data, proved similarly effective, with estimated generalisation accuracies of 88 % and 97 %, and κ values of 0.84±0.05 and 0.94±0.04 respectively (see also supplementary material Tables S2 & S3).

The spatial distributions of s_A classified as krill by each method were in broad agreement (Figures 5 & 6a). Variability in the spatial distribution of s_A was similar in both cases, with relatively larger values occurring to the northwest and east of the South Georgia

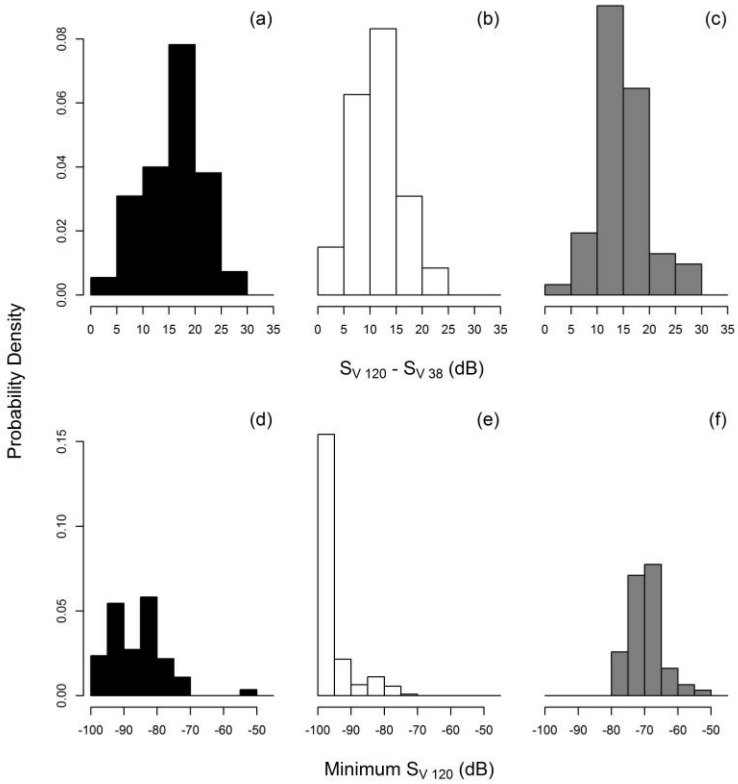

Figure 4. Distributions of school-level $S_{V\ 120} - S_{V\ 38}$ (dB) values (a-c) and minimum $S_{V\ 120}$ (dB) values (d-f) from trawl-verified echoes for mackerel icefish (*black*), krill (*white*), and mixed groundfish (*grey*).

shelf, as well as to the west of Shag Rocks. Due to the school-based nature of the RF method, only aggregations above some background density were detected, and so there are several EDSUs associated with this method where no krill was detected. The corresponding EDSUs from the other methods often contained low densities. Overall, however, the s_A per EDSU attributed to krill using the RF method were significantly higher than those resulting from both the fixed 2-12 dB range (Wilcoxon signed rank test, V = 1815149, $p < 0.05$), and the variable 0.37-12 dB range (Wilcoxon signed rank test, V = 1833645, $p < 0.05$). The fixed 2-16 dB range resulted in significantly higher s_A per EDSU than the RF (Wilcoxon signed rank test, V = 1944337, $p < 0.05$). Relatively small amounts of s_A were attributed to icefish using the RF, mainly to the northwest of the South Georgia shelf and the east of Shag Rocks (Figure 6b). These correspond to areas where the commercial fishery mainly operates, as well as where the highest densities of icefish are typically recorded during groundfish surveys (Main *et al.*, 2008). s_A attributed to mixed groundfish aggregations by the RF method were fairly uniformly distributed across the South Georgia shelf, with some small amounts at Shag Rocks (Figure 6c). This pattern is again in agreement with groundfish survey observations of the benthic assemblage. Of the RF-assigned backscatter, 93% of icefish s_A and 62% of mixed groundfish s_A was above the 6m mean headline height of the bottom trawl (Parkes, 1991).

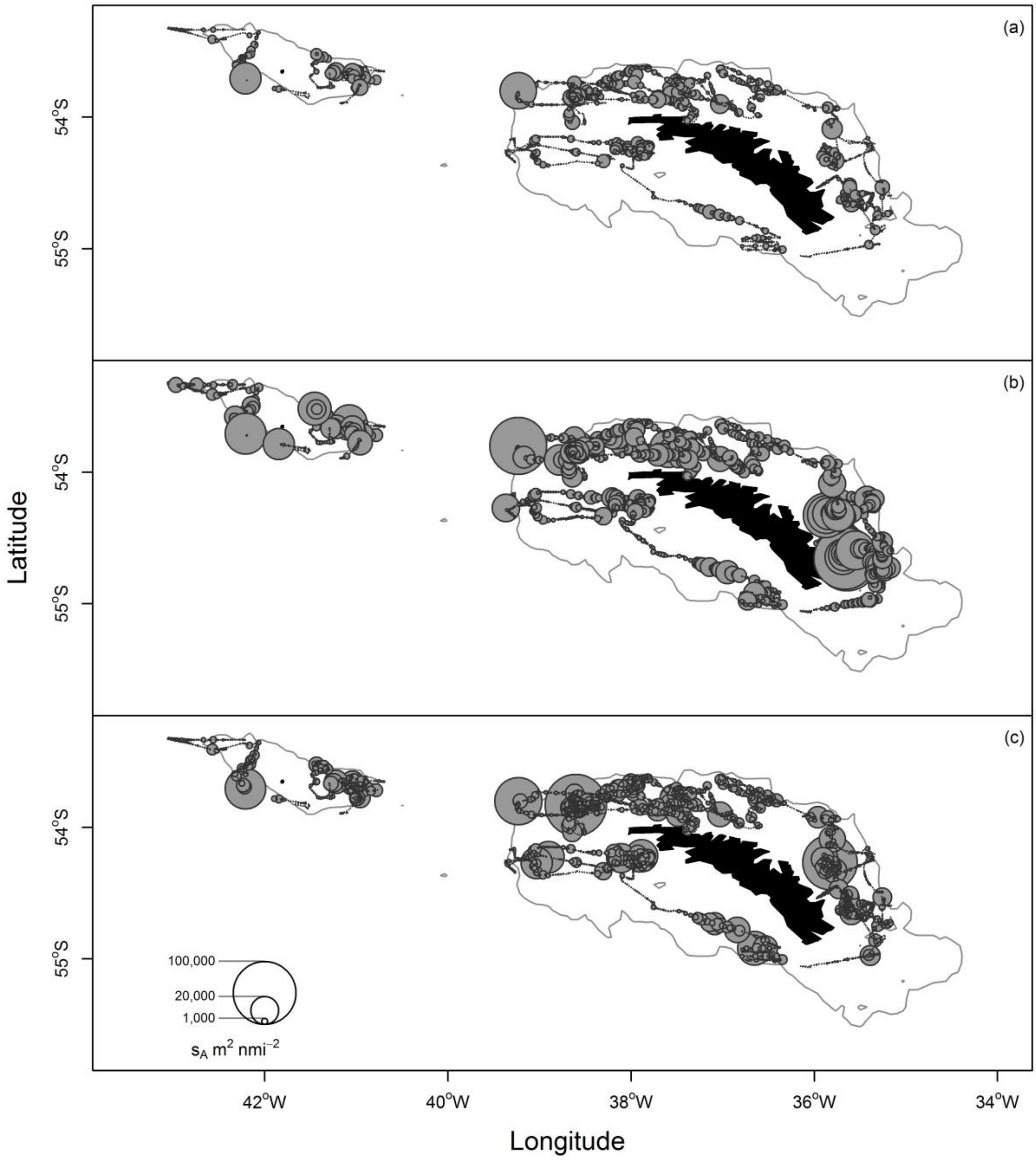

Figure 5. Spatial distributions of s_A (m^2 nmi^{-2}) for krill using: (a) the 2-12 dB fixed window method, (b) the 2-16 dB fixed window method, (c) the variable window method.

Discussion and Conclusion

Random Forest (RF) models classify echoes on the basis of their empirically observable attributes while making few assumptions after the data has been collected. RF models may be improved with the addition of new data to the training dataset, and selection of variables according to the particular attributes of the species being classified (Genuer *et al.*, 2010). Expert knowledge can thus be incorporated via case-specific variable selection. Relative to other

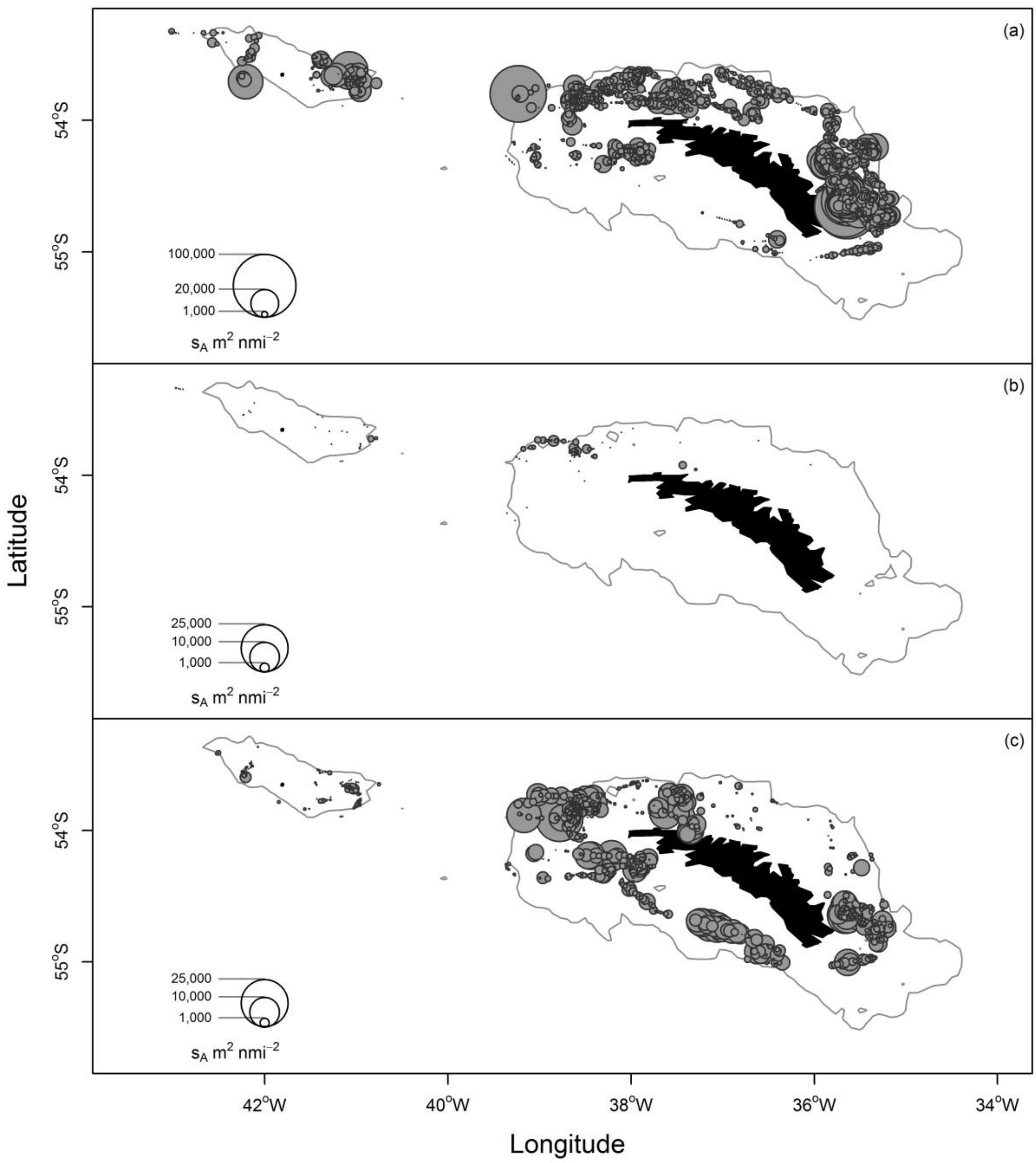

Figure 6. Spatial distributions of s_A (m^2 nmi^{-2}) for: (a) krill, (b) icefish, and (c) mixed groundfish, where s_A was classified using the random forest method trained on the full dataset.

methods, RF models are also simple to implement, and accept variables with diverse statistical properties (Hastie *et al.*, 2009). For identifying icefish in the water column, the RF in this study had an estimated 94% accuracy, and an overall prediction accuracy higher than other methods (D'Elia *et al.*, 2014; Woodd-Walker *et al.*, 2003). Accepting the need to develop reliable target strength models for icefish, the method presented here could be used in the quantification of any bottom trawl sampling bias, and may be integrated into survey analyses that inform the icefish assessment. The RF method was preconditioned on schools, and so, unlike the S_V-difference methods, it did not function in the detection and classification of backscatter below a given density, e.g. that which is observed in some dispersed krill layers (Watkins and Murray, 1998). However, the fact that krill s_A as defined by the RF method was still significantly higher than that from the fixed 2-12 dB method illustrates that excluding those diffuse layers from the analysis may not substantially bias density estimates, and that the majority of krill biomass is contained in swarms (Fielding *et al.*, 2014).

Although this study was motivated by the investigation of the pelagic component of the icefish stock, there is also potential for acoustic data collected during groundfish surveys to supplement other analyses, such as the Western core box (WCB) krill survey and ecosystem modelling. In order to provide more accurate data on the various pelagic scatterers detected during groundfish surveys, some effort should be allocated to the collection of krill data around Shag Rocks. In reality, krill is not absent from the Shag Rocks shelf, but as longitude has a relatively strong influence in the RF the lack of training data in that area might have biased classification. Another inherent challenge will be in dealing with the non-systematic nature of this incidentally collected data, which should be surmountable through the specification of spatially explicit models. Provided the above issues are addressed, species- or assemblage-level acoustic indices could give valuable insight into uncertainties regarding the composition of the pelagic ecosystem at South Georgia. For example, given that the majority of s_A attributed to the various fish species was recorded above the bottom trawl headline height, with a greater understanding of bottom trawl catchability discrepancies between survey-based abundance estimates and estimated piscivore food requirements (Hill *et al.*, 2012) might be explained.

The methods using fixed and variable ranges of $S_{V\,120} - S_{V\,38}$ may provide inaccurate estimates of krill backscatter, but not only because they include echoes from other zooplankton. Echoes from fish without swimbladders may also be erroneously classified as krill. This is because the distributions of school-level $S_{V\,120} - S_{V\,38}$ for krill, icefish, and mixed fish overlap (Figure 4). Conversely, krill backscatter may be underestimated because only a portion of the $S_{V\,120} - S_{V\,38}$ values measured from krill swarms were included in the S_V-difference ranges assumed for krill. For example, some haul-verified krill swarms had $S_{V\,120} - S_{V\,38}$ values >12 dB, which Madureira *et al.* (1993) defined as non-krill zooplankton. Therefore, a 2-12 dB range of $S_{V\,120} - S_{V\,38}$ alone is unlikely to account for all krill backscatter, and may include backscatter from other zooplankton and fish species. Similarly, the wider 2-16 dB range may result in significantly higher s_A than the RF method due to the inclusion of non-krill echoes (Watkins and Brierley, 2002). This exemplifies a trade-off in the EDSU-level approach; an excessively conservative $S_{V\ 120}$ – $S_{V\ 38}$ range excludes both non-krill targets and some krill echoes, whereas a wider range includes most types of weak scatterers.

Minimum $S_{V\ 120}$ was the most important predictor variable in the RF. A wider range of minimum $S_{V\ 120}$ was observed across icefish echoes than from those of mixed groundfish aggregations, with minimum values in both categories being generally higher than those of krill swarms. In the case of fish schools, minimum $S_{V\ 120}$ is perhaps most likely to be a function of orientation, with lower values recorded for icefish which spends more time swimming vertically in the water column than other species (Kock, 2005b). Many species included in the mixed aggregation category also have a larger mean body size (Kock and Kellermann, 1991), which could account for the generally higher minimum $S_{V\ 120}$ values. It is apparent (Figure 4) that a large portion of minimum $S_{V\ 120}$ values in krill echoes were between -95 and -100 dB, clearly distinguishing it from the other categories. A single 40 mm krill per m^3 at a near horizontal orientation has an approximate $S_{V\ 120}$ = -70 dB (Lawson *et al.*, 2006, 2008), and so values of $S_{V\ 120}$ = -100 dB would most likely represent a discontinuity in density within the swarm under those assumptions. At fine scales, krill within swarms have been shown to exhibit measurable levels of uniformity in terms of their orientation (Kubilius *et al.*, 2015). Most typically they assume a near horizontal position, particularly when actively swimming (Demer and Conti, 2005; Lawson *et al.*, 2006), but are assumed to vary in orientation across swarms. It was thus posited that these minimum S_V samples between -95 and -100 dB could either represent vacuoles or variability in krill orientation within dense swarms, but are perhaps most likely observed due to low density regions where krill are oriented vertically, minimising their profile in the acoustic beam.

Including a “mixed groundfish” category was necessary, as a sufficient number of trawl-verified echoes were not available to subset the data any further. Operator intervention was thus required to verify some RF classifications. For instance, the yellowfin notothen, *Patagonotothen guntheri*, another weak scattering species, forms dense pelagic feeding aggregations around Shag Rocks (Collins *et al.*, 2008). If monospecific aggregations such as this are known to occur then it is preferable to include a corresponding class in the RF method. However, few trawl-verified echoes were available for *P. guntheri* in this case, and so further scrutiny was essential for verification of some RF classifications. It is also apparent from Table 1 that the dataset was not balanced in terms of the number of observations on each group, which can affect the interpretation of results. For example, if echoes designated as “krill” were to make up ~5% of observations in the confusion matrix of a binary classifier, 95% accuracy could be achieved by labelling all schools as “mackerel icefish” (Fielding and Bell, 1997).

The properties of echoes considered in this analysis exhibited variability, non-linearity, interaction, and collinearity. Accordingly, classification of echoes at the level of the school is complex. Compiling a training dataset that adequately represents the distributions of those variables of interest can be a significant hurdle to reliable classification (Woodd-Walker *et al.*, 2003). This should be considered when choosing which approach to adopt to a given echo classification problem, and emphasises that the choice of a method is sometimes as dependent on the properties and quality of the available data as it is on the question being addressed (Reid *et al.*, 2000). Indeed, there are situations where considering the data at broader spatial scales (i.e. EDSU-level analysis) is more appropriate (Reid *et al.*, 2000). This can reduce or eliminate the need for training data entirely, with the caveat that more generalised assumptions will need to be accepted regarding the acoustic properties of the target species. To that end, EDSU-level analyses have been developed which can provide more accurate classification than the fixed S_V-difference method applied in this study (Fielding *et al.*, 2014). However, the loss of fine-scale detail of individual schools makes accurate classification beyond broad categories (e.g. weak scatterers) challenging.

Supplementary Material

Supplementary figures and tables are available at ICESJMS online.

Acknowledgements

The authors wish to thank the crews, fishermen and scientists who conducted the various surveys from which data were obtained. This work was supported by the Government of South Georgia and South Sandwich Islands. Additional logistical support provided by The South Atlantic Environmental Research Institute, with thanks to Paul Brickle. Paul Fernandes receives funding from the MASTS pooling initiative (The Marine Alliance for Science and Technology for Scotland), and their support is gratefully acknowledged. MASTS is funded by the Scottish Funding Council (grant reference HR09011) and contributing institutions. Sophie Fielding is funded by the Natural Environment Research Council, and data was provided from the British Antarctic Survey Ecosystems Long-term Monitoring and Surveys programme as part of the BAS Polar Science for Planet Earth Programme. The authors also wish to thank the anonymous referees for their helpful suggestions on an earlier version of this manuscript.

References

Aglen, A., Engås, A., Huse, I., Michalsen, K., and Stensholt, B. K. 1999. How vertical fish distribution may affect survey results. ICES Journal of Marine Science, **56**: 345–360.

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

Barnes, D. K. A., Collins, M. A., Brickle, P., Fretwell, P., Griffiths, H. J., Herbert, D., Hogg, O. T., et al. 2011.

The need to implement the Convention on Biological Diversity at the high latitude site, South Georgia. Antarctic Science, **23**: 323–331.

Bez, N., Reid, D., and Neville, S. 2007. Acoustic data collected during and between bottom trawl stations: consistency and common trends. Canadian journal of Fisheries and Aquatic Science, **180**: 166–180.

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

CCAMLR. 2010. Report of the fifth meeting of the subgroup on acoustic survey and analysis methods.

CCAMLR. 2014. Report of the Working Group on Fish Stock Assessment. In CCAMLR, Report of the Thirty-Third Meeting of the Scientific Committee, Annex 7. Hobart, Australia. 276-367 pp.

Coetzee, J. 2000. Use of a shoal analysis and patch estimation system (SHAPES) to characterise sardine schools. Aquatic Living Resources, **13**: 1–10.

Cohen, J. 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, **20**: 37–46.

Collins, M. A., Shreeve, R. S., Fielding, S., and Thurston, M. H. 2008. Distribution, growth, diet and foraging behaviour of the yellow-fin notothen *Patagonotothen guntheri* (Norman) on the Shag Rocks shelf (Southern Ocean). Journal of Fish Biology, **72**: 271–286.

D'Elia, M., Patti, B., Bonanno, a., Fontana, I., Giacalone, G., Basilone, G., and Fernandes, P. G. 2014. 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. Elsevier B.V.

Demer, D. A. 2004. An estimate of error for the CCAMLR 2000 survey estimate of krill biomass. Deep-Sea Research Part II: Topical Studies in Oceanography, **51**: 1237–1251.

Demer, D. A., and Conti, S. G. 2005. New target-strength model indicates more krill in the Southern Ocean. ICES Journal of Marine Science, **62**: 25–32.

Diner, N. 2001. Correction on school geometry and density: Approach based on acoustic image simulation. Aquatic Living Resources, **14**: 211–222.

Everson, I., Bravington, M., and Goss, C. 1996. A combined acoustic and trawl survey for efficiently estimating fish abundance. Fisheries research, **26**: 75–91.

Fallon, N. G., Collins, M. A., Marshall, C. T., and Fernandes, P. G. 2015. Assessing consistency of fish survey data: uncertainties in the estimation of mackerel icefish (*Champsocephalus gunnari*) abundance at South Georgia. Polar Biology.

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

Fielding, A., and Bell, J. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental conservation, **24**: 38–49.

Fielding, S., Watkins, J. L., Trathan, P. N., Enderlein, P., Waluda, C. M., Stowasser, G., Tarling, G. A., et al. 2014. Interannual variability in Antarctic krill (*Euphausia superba*) density at south georgia, Southern Ocean: 1997-2013. ICES Journal of Marine Science, **71**: 2578–2588.

Foote, K., Knudsen, H., and Vestnes, G. 1987. Calibration of acoustic instruments for fish density estimation: a practical guide. ICES Cooperative Research Report: 57.

Genuer, R., Poggi, J., and Tuleau-Malot, C. 2010. Variable selection using random forests. Pattern Recognition Letters, **31**: 2225–2236.

Gorska, N., Korneliussen, R. J., and Ona, E. 2007. Acoustic backscatter by schools of adult Atlantic mackerel. ICES Journal of Marine Science, **64**: 1145–1151.

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.

Hastie, T., Tibshirani, R., and Friedman, J. 2009. The Elements of Statistical Learning. Springer-Verlag, London. 763 pp.

Hewitt, R. P., Watkins, J., Naganobu, M., Sushin, V., Brierley, A. S., Demer, D., Kasatkina, S., et al. 2004. Biomass of Antarctic krill in the Scotia Sea in January/February 2000 and its use in revising an estimate of precautionary yield. Deep-Sea Research Part II: Topical Studies in Oceanography, **51**: 1215–1236.

Higginbottom, I., Pauly, T. J., and Heatley, D. C. 2000. Virtual echograms for visualization and post-processing of multiple-frequency echosounder data. In Proceedings of the Fifth European Conference on Underwater Acoustics, ECUA 2000, pp. 1497–1502. Ed. by M. E. Zhakharia, P. Cheveret, and P. Dubail. Office for Official Publications of the European Communities, Luxembourg, Lyon, France.

Hill, S. L., Keeble, K., Atkinson, A., and Murphy, E. J. 2012. A foodweb model to explore uncertainties in the South Georgia shelf pelagic ecosystem. Deep Sea Research Part II: Topical Studies in Oceanography, **59-60**: 237–252. Elsevier.

Hill, S., Reid, K., and North, A. 2005. Recruitment of mackerel icefish (*Champsocephalus gunnari*) at South Georgia indicated by predator diets and its relationship with sea surface temperature. Canadian Journal of Fisheries and Aquatic Sciences, **27**: 2530–2537.

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

ICES. 2015. Manual for International Pelagic Surveys (IPS). Series of ICES Survey Protocols SISP 9 – IPS. 92 pp.

James, G., Witten, D., Hastie, T., and Tibshirani, R. 2013. An Introduction to Statistical Learning. Springer Texts in Statistics. Springer New York, New York, NY. 441 pp.

Klevjer, T. A., Tarling, G. A., and Fielding, S. 2010. Swarm characteristics of Antarctic krill *Euphausia superba* relative to the proximity of land during summer in the Scotia Sea. Marine Ecology Progress Series, **409**: 157–170.

Kock, K., and Kellermann, A. 1991. Reproduction in Antarctic notothenioid fish. Antarctic Science, **3**: 125–150.

Kock, K.-H. 2005a. Antarctic icefishes (Channichthyidae): a unique family of fishes. A review, Part I. Polar Biology, **28**: 862–895.

Kock, K.-H. 2005b. Antarctic icefishes (Channichthyidae): a unique family of fishes. A review, Part II. Polar Biology, **28**: 897–909.

Korneliussen, R. J. 2010. The acoustic identification of Atlantic mackerel. ICES Journal of Marine Science, **67**: 1749–1758.

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., Heggelund, Y., Eliassen, I. K., and Johansen, G. O. 2009. Acoustic species identification of schooling fish. ICES Journal of Marine Science, **66**: 1111–1118.

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

Kotwicki, S., De Robertis, A., Ianelli, J. N., Punt, A. E., Horne, J. K., and Jech, J. M. 2013. Combining bottom trawl and acoustic data to model acoustic dead zone correction and bottom trawl efficiency parameters for semipelagic species. Canadian Journal of Fisheries and Aquatic Sciences, **70**: 208–219.

Kubilius, R., Ona, E., and Calise, L. 2015. Measuring in situ krill tilt orientation by stereo photogrammetry: examples for *Euphausia superba* and Meganyctiphanes norvegica. ICES Journal of Marine Science, 72: 2494–2505.

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. The Journal of the Acoustical Society of America, **122**: 3304–3326.

Lawson, G. L., Wiebe, P. H., Ashjian, C. J., Chu, D., and Stanton, T. K. 2006. Improved parameterization of Antarctic krill target strength models. The Journal of the Acoustical Society of America, **119**: 232–242.

Lawson, G. L., Wiebe, P. H., Ashjian, C. J., and Stanton, T. K. 2008. Euphausiid distribution along the Western Antarctic Peninsula-Part A: Distribution of euphausiid aggregations and biomass, and associations with environmental features. Deep-Sea Research Part II: Topical Studies in Oceanography, **55**: 432–454.

Liaw, A., and Wiener, M. 2002. Classification and Regression by randomForest. R news, **2**: 18–22.

Lunde, P., and Pedersen, A. O. 2012. Sonar and power budget equations for backscattering of finite amplitude sound waves, with implications in fisheries acoustics for abundance estimation of marine resources. Proceedings of the 35th Scandinavian Symposium on Physical Acoustics, Geilo 29 January – 1 February: 1–11.

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

Main, C. E., Collins, M. A., Mitchell, R., and Belchier, M. 2008. Identifying patterns in the diet of mackerel icefish (*Champsocephalus gunnari*) at South Georgia using bootstrapped confidence intervals of a dietary index. Polar Biology, **32**: 569–581.

McQuinn, I., Simard, Y., Stroud, T., Beaulieu, J., and Walsh, S. 2005. An adaptive, integrated 'acoustic-trawl' survey design for Atlantic cod (*Gadus morhua*) with estimation of the acoustic and trawl dead zones. ICES Journal of Marine Science, **62**: 93–106.

Mitchell, R. E., Belchier, M., Gregory, S., Kenny, L., Nelson, J., Brown, J., and Featherstone, L. 2010. Results of the groundfish survey carried out in CCAMLR sub-area 48 . 3 in January 2010. CCAMLR WG-FSA-10/38: 38.

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

Parkes, G. 1991. The UK fish stock assessment survey bottom trawl for South Georgia. CCAMLR WG-FSA-91/16.

Pedersen, A. 2006. Effects of nonlinear sounds propagation in fisheries research. University of Bergen. 308 pp.

Pilling, G., and Parkes, G. 1995. Performance and geometry of the FP-120 trawl used during UK fish stock assessment surveys around South Georgia, Subarea 48.3. CCAMLR Science, **2**: 51–69.

R Development Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Reid, D. G., and Simmonds, E. J. 1993. Image Analysis Techniques for the Study of Fish School Structure from Acoustic Survey Data.

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

Reiss, C. S., Cossio, A. M., Loeb, V., and Demer, D. A. 2008. Variations in the biomass of Antarctic krill (*Euphausia superba*) around the South Shetland Islands, 1996–2006. ICES Journal of Marine Science: Journal du Conseil, **65**: 497–508.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., and Zeileis, A. 2008. Conditional Variable Importance for Random Forests. BMC Bioinformatics, **9**: 307.

Strobl, C., Hothorn, T., and Zeileis, A. 2009. Party on! R Journal, **1**: 14–17.

Tarling, G. A., Klevjer, T., Fielding, S., Watkins, J., Atkinson, A., Murphy, E., Korb, R., et al. 2009. Variability and predictability of Antarctic krill swarm structure. Deep Sea Research Part I: Oceanographic Research Papers, **56**: 1994–2012. Elsevier.

Watkins, J. L., and Brierley, A. S. 2002. Verification of the acoustic techniques used to identify Antarctic krill. ICES Journal of Marine Science, **59**: 1326–1336.

Watkins, J. L., and Murray, a. W. a. 1998. Layers of Antarctic krill, *Euphausia superba* : are they just long krill swarms? Marine Biology, **131**: 237–247.

Welling, S. H., Clemmensen, L. K. H., Buckley, S. T., Hovgaard, L., Brockhoff, P. B., and Refsgaard, H. H. F. 2015. In silico modelling of permeation enhancement potency in Caco-2 monolayers based on molecular descriptors and random forest. European Journal of Pharmaceutics and Biopharmaceutics, **94**: 152–159.

Woodd-Walker, R., 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, **3139**: 641–649.