ICES ASC 2004 Theme Session R: New developments in fisheries acoustics. Paper R12

A dual frequency algorithm for the identification of sandeel school echotraces

Alicia Mosteiro[1], Paul G. Fernandes[2], Eric Armstrong[2] and Simon Greenstreet[2]

Mosteiro, A., Fernandes, P.G., Armstrong, F. and Greenstreet S.P.R. (2004). A dual frequency algorithm for the identification of sandeel school echotraces. *ICES CM2004/**R**:12*, 13 pp.

Sandeels are not only targets of a significant industrial fishery, but are also considered to be a vital trophic component of the North Sea ecosystem. At present there is no satisfactory survey method to sample sandeels which produces a global absolute abundance estimate. Acoustic surveys have been carried out, but suffer from an inability to consistently identify sandeel echo traces in an objective manner. As sandeels lack a swimbladder, their acoustic properties are very different to other fish species which occur adjacent to them. We report on the development of a dual frequency algorithm which aims to identify echotraces of sandeel schools based on the observed difference in acoustic scattering at 38 and 120 kHz. The algorithm also includes noise reduction and plankton filtering components. Multifrequency algorithms such as these will no doubt result in acoustic surveys being used more widely for fishery independent surveys with direct benefits for an ecosystems approach to fisheries management.

Keywords: multi-frequency, acoustics, sandeel, surveys.

[1]Mosteiro, A.: Malta Centre for Fisheries Sciences, Fort San Lucjan, Qattenza Rd, M`Xlokk, BBG06, Malta [tel: +356 21655525, fax: +356 21659380, e-mail: alicia.mosteiro@gov.mt]. [2]Fernandes P.G, Armstrong, F. and Greenstreet, S.P.R.: FRS Marine Laboratory, PO Box 101, 375 Victoria Road, Aberdeen, Scotland AB11 9DB [tel: +44 1224 295403, fax: +44 1224 295511, e-mail: fernandespg@marlab.ac.uk].

Introduction

Sandeels (*Ammodytes spp.*) are not only targets of significant industrial fisheries (ICES, 2004), but are also considered to be a vital trophic component of many marine ecosystems, comprising large parts of the diet of seabirds (Furness and Tasker, 2000), marine mammals (Tollit *et al.,* 1997) and other piscivorous fish (Greenstreet *et al.,* 1998). There is, therefore, a need to determine the abundance and distribution of sandeels for the purposes of fishery management, especially if consideration is to be given to management that is ecosystem-based (Pikitch *et al.,* 2004). At present there is no satisfactory survey method to sample sandeels which produces a global absolute abundance estimate.

Acoustic surveys are used in fisheries science to assess the abundance, distribution, and behaviour, of fish, plankton, and other marine organisms (MacLennan and Simmonds, 1992). Within ICES, there are currently over 20 fish stocks for which acoustic estimates are carried out. Most of these are pelagic (midwater) species such as herring (*Clupea harengus*), sardine (*Sardina pilchardus*), and anchovy (*Engraulis encrasicolus*). Fish density is estimated by the technique of echocounting or by echointegration (MacLennan and Forbes 1984). These techniques rely on net sampling of the echotraces to identify them: so-called ground truthing (McClatchie *et al.*, 2000; Everson *et al.*, 1996). However, net sampling cannot achieve the spatial or temporal coverage that is comparable to that of the acoustic data (Rose & Legget, 1988). Most of the echotraces, therefore, are not ground-truthed, so their identification is based on visual interpretation or scrutiny (Reid *et al.,* 1998). The latter technique, although subjective and specialised, is well established: Cushing (1957) classified different shapes in echograms as spots, comets, plumes, etc., in an attempt to identify different species.

Developments in signal processing led to the study of morphometric or energetic school parameters, to extract characteristics that could discriminate between species, using them as inputs to pattern recognition algorithms (e.g., Rose & Legget, 1988). Such techniques were able to identify up to 98% of fish shoals if space and time scales were reduced and seasonal variation was included in the image processing (Lefeuvre *et al*, 2000). Other studies of echotrace classification based on single frequency echosounders were carried out by Lee *et al.* (1990); Vray *et al.* (1990); Richards *et al.* (1991); Lu and Lee (1994); Barange (1994); Scalabrin *et al.* (1996); Haralabous and Georgakarakos (1996); Coetze (2000); and Lawson *et al.* (2001).

On the basis of the theoretical dependence on frequency of acoustic scattering from marine organisms (Holliday, 1992), wideband (or broadband) and multiple-frequency echosounders have been used to characterise echotraces from fish schools. The difference in mean volume backscattering strength (MVBS) at two or more frequencies, has been used to separate fish and plankton (Madureira *et al.*, 1993; Barange, 1994; Brierley *et al.*, 2001; Kang *et al.*, 2002). The relative scattering contribution at a number of discrete acoustic frequencies, has been used to produce synthetic combined-frequency echograms which isolate fish species into categories, based on their relative frequency response (Korneliussen and Ona, 2002; 2003). Broadband echo-sounders provide a more continuous spectral signature which may also be species specific (Zakharia *et al.*, 1996; Simmonds *et al.,* 1996).

The aim of this work is to provide a tool for automatically identifying sandeels schools using multiple frequency acoustic data. The tool consists of an algorithm which applies various image analysis and mathematical operators to echosounder data simultaneously collected at 38 and 120 kHz. Noise reduction and plankton filtering components were included to isolate fish school echotraces. Efforts were then focussed on determining the range of differences in fish school scattering at 38 and 120 kHz that would discriminate sandeel from other co-occurring species (mainly clupeoids). The algorithm was developed using data that had been ground-truthed using pelagic trawls such that an alternative measurement of the fish species composition was known. The algorithm was also compared with the more traditional visual scrutiny method on a new sandeel survey dataset.

Materials and methods

The method relies on the analysis of acoustic data that has been 'ground-truthed' using a pelagic trawl. The input data therefore consist of paired acoustic and trawl datasets. The acoustic data consist of echogram sections immediately prior to trawling, which have subsequently been sampled with an alternative method (hauls were carried out at the same location where the fish were previously detected acoustically). Each pair of acoustic and trawl data, referenced by the trawl haul number is henceforth referred to as a 'paired dataset'. Acoustic data for identification was collected at 10 knots (normal speed for the acoustic surveys).

Data collection

Data were taken from a number of spring and summer acoustic surveys in the North Sea (*FRV Scotia*), Firth of Forth, and Moray Firth (*FRV Clupea*) off the coast of Scotland, UK. Further details of survey procedures can be found in individual survey reports (e.g., Simmonds *et al.,* 2001).

A total of 28 trawl hauls were taken from the survey data (Table 1). 21 of them were taken by the *Clupea* using a pelagic trawl with 6 m opening and 6 mm mesh in the codend. The remaining seven *Scotia* net samples were taken with a pelagic trawl of 12 metres vertical opening and 20 mm mesh in the codend. The trawl catches provide information on catch composition in numbers, latitude and longitude, time of start and end for the haul, as well as other parameters such as depth trawled and a description of the fished echotraces. Sixteen trawl hauls were composed exclusively of sandeel, five were an abundant mixture of sandeel and clupeoid and the rest were different mixtures of species in a variety of numbers. The trawling speed was approximately 4 knots.

Table 1. Details of the trawl sample data used to develop the sandeel identification algorithm, arranged by cruise and date. Catch composition in number.

Ship Month Year	Date	Haul no.	Sandeel *Ammodytes spp.*	Herring *Clupea harengus*	Sprat *Sprattus sprattus*	Gadidae	Others	Sandeel proportion
Clupea	21	020	21	3494	5	1		0.01
June	23	023	15648	10992	246			0.58
1997	24	025	64	15472	16			0.00
	25	027	17	18528	2464			0.00
Clupea	2	082	9951					1.00
July	3	083	887			11		0.99
1999	3	084	5780			8		1.00
	4	085	28341		2		108	1.00
	5	087	22949	85200	480			0.21
Clupea	15	056	2	1230	220	20	10	0.00
June	16	058	18159	2	40	9		1.00
2000	16	059	136634		78	42		1.00
	18	060	65548			28		1.00
	19	061	16788	4		60		1.00
Clupea	5	220	31499	17500		56		0.64
June	5	221	94557					1.00
2001	6	222	99040				19	1.00
	7	224	33081					1.00
	8	225	253	955				0.21
	9	226	109668	2750				0.98
	9	227	6869	70	3648	22	12	0.65
Scotia	5	200	O group meshed		171	9		0.00
June	6	201	1440					1.00
2000	6	203	1597					1.00
	6	205	4	134		49626	4	0.00
	7	208	3128					1.00
Scotia	6	222	105			1		0.99
July 2000	8	229	37		7	5		0.76

Acoustic data were collected using a Simrad EK500 echosounder operating hull-mounted (*Scotia*), and towed body (*Clupea*), transducers with 7° beamwidths. For this study, only data at 38 and 120 kHz were considered. The echosounder was configured to ping simultaneously at each frequency once a second with a pulse length of 1 and 0.3 ms at 38 and 120 kHz respectively. The performance of the echosounder was monitored using standard target calibration techniques (Fernandes and Simmonds, 1996). The acoustic data were collected from 03:00 to 23:00 hrs.

The raw acoustic data consisted of echogram "Q" telegrams collected from the EK500 (Bodholt *et al.*, 1989). These telegrams consist of time-stamped digitised volume backscattering strengths (VBS). Each pixel on the echogram therefore corresponds to a VBS (symbol, S_v; unit, dB re $1m^{-1}$). Other telegrams collected include seabed depth and geographic location (latitude and longitude). The data were logged from the echosounder to a PC with SonarData's Echolog software (SonarData Pty Ltd., GPO Box 1387 Hobart, Tasmania, Australia) for all years, except for 1997 and 1998 when Simrad BI500 was still in use for logging onto Unix platforms. SonarData Echoview software was used for the analysis of the echosounder data.

Output from the algorithm was compared to results from a previously analysed sandeel cruise performed aboard FRV *Clupea* from 14 - 17 June 2003. This cruise was not included in the data set used to develop the algorithm. Nautical area scaterring coefficients (NASCs) were derived from 1.5 km equivalent distance sampling units (EDSUs) using the algorithm. Visual scrutiny (manual identification) was carried out on the same EDSUs to determine NASCs based on previous experience and data supplied from a number of trawls carried out during the survey. Biomass estimates for the survey based on the

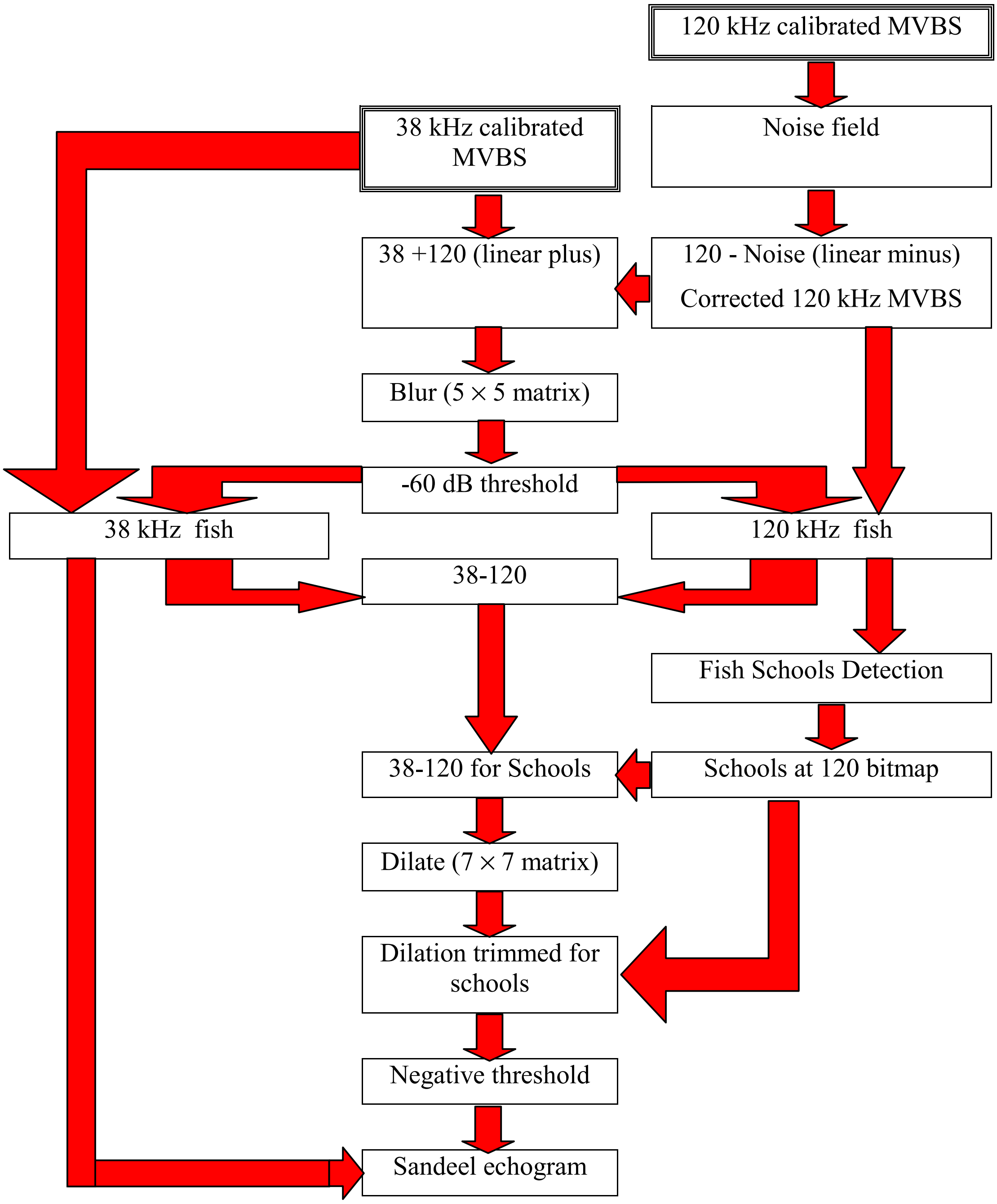

Figure 1. Schematic flow diagram of processing steps in the sandeel identification algorithm.

two analyses were carried out using the Marine Laboratory Integration and Analysis software package (MILAP).

Algorithm description

Echoview facilitates the analyses of multifrequency echograms by synchronising the depth and time of each VBS from each echogram frequency. A number of mathematical functions can then be applied as 'operators' on the data (e.g. "plus" = summation of each synchronised pixel in two frequencies) to produce '*virtual*' echograms. In the explanation that follows, italics are used to illustrate the names of the *virtual* echograms produced at each step (Figure 1). Calibration parameters and any processing steps, including all mathematical operators, were stored in Echoview (EV) files for each trawl haul. Bottom exclusion (0.5 m above the bottom) and surface (at a fixed depth) lines were created specifically for each EV file to avoid including bottom pixels and the surface tranmsit pulse.

The higher frequency 120 kHz data are vulnerable to depth dependent noise which can have adverse effects on multifrequency signal processing (Kang *et al.,* 2002; Korneliussen and Ona, 2002). These data were, therefore, filtered for noise, using methods analogous to those described in Watkins and Brierley (1996): a *noise field* was created and subtracted (in the linear domain) from the raw 120 kHz data to produce a *120-Noise (linear minus) corrected 120 kHz MVBS echogram*.

Plankton usually have lower levels of frequency specific backscattered energy than fish (with or without swimbladders) and their distribution pattern is quite diffuse which contrasts with the common schooling behaviour of small pelagic fish. Three techniques were combined to filter out plankton: thresholding, multifrequency processing and image analysis. A minimum MVBS threshold of –82 dB was applied to both frequencies, eliminating the weak and diffuse plankton echoes. The two frequencies were then summed to produce a *38+120 (linear plus)* echogram. A blurring process was then performed on the data to smooth the edges of the schools and to account for losses due to any mismatches when synchronising pixels from the two frequencies. The result is a *Blur (5x5 matrix)* echogram. The blurring employed a 5*5 convolution matrix which was found to keep the shape and size of the schools: the matrix kernel was as follows:

$$\begin{matrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{matrix}$$

After blurring, a threshold of –60 dB was applied (‘range’ operator), to eliminate plankton from the blurred echogram (*-60 dB threshold* echogram). The thresholded echogram is a boolean variable (true for values above threshold and false for the rest) which was then used as a mask for both the 38 & 120 kHz echograms to produce plankton extracted *38 kHz fish* and *120 kHz fish* echograms.

Echoview has an image analysis facility, the ‘Schools Module’, for detecting schools (defined as regions on the echogram) and measuring their characteristics (morphometric or energetic parameters). This is based on the SHAPES fish school image analysis algorithm (Barange, 1994). The set of parameters for school detection chosen were: a minimum school length of 15 m; a minimum school height of 3 m; a minimum connected length of 8m; minimum connected height of 1.5 m; maximum vertical linking distance of 2 m; and a maximum horizontal linking distance of 15 m. Schools were then detected in the plankton extracted 120 kHz echogram (*120 kHz fish*), and a schools bitmap echogram was produced (*Schools at 120 bitmap*).

Preliminary examination of paired datasets consisting of monospecific sandeel schools revealed that, in the majority of cases, sandeel schools have stronger MVBS values at 120 kHz than at 38 kHz. Other species present (mainly herring) were usually stronger at 38 than 120 kHz. It would, therefore, seem logical to use the difference in echo return energy levels from the two frequencies as a means of separating sandeels from other various species. A non-linear subtraction of the two frequencies was therefore carried out producing a dB difference (ΔdB) echogram (*38-120*). The ΔdB echogram was then masked with the fish school bitmap to produce a *38-120 for schools* echogram. After this a dilation (Reid and Simmonds, 1993) was applied to smooth the ΔdB data producing a *Dilate (7x7 matrix)* echogram. This was trimmed by applying the schools bitmap producing a *Dilation trimmed for schools* echogram. A threshold was then applied to separate sandeels (negative dB difference) from other fish (positive ΔdB) producing a *negative threshold* bitmap echogram with only values below the threshold. This ΔdB threshold (ΔdB_t) was set to –3 dB (see below). Finally, the latter bitmap was used as a mask on the 38 kHz fish echogram to provide a *sandeel echogram* at 38 kHz which is the integrating frequency. The complete sandeel algorithm is described by the flow chart in Figure 1.

The ΔdB_t was chosen by applying the similarity of identification index ($\hat{s}_{id}$) as described in (Fernandes and Stewart 2004). $\hat{s}_{id}$ is a numerical value which expresses how closely on average, the species composition, as determined by the acoustic identification algorithm, compares with the species composition from an alternative source (e.g. pelagic trawl). Thus, it is not a “probability” or percentage, but an expression of relative difference in the identification of fish between two independent methods. An exact match of the species composition of the algorithm derived proportion and the proportion derived from the alternative sampler would result in an $\hat{s}_{id}$ of 1.000. $\hat{s}_{id}$ can only be applied to isolated schools (echograms without plankton), and only when two species groups dominate the catch, which is the case in most of the hauls used in this study. The algorithms initial value of $\Delta dB_t = -0.1$ dB gave an $\hat{s}_{id}$ of 0.421. The value of ΔdB_t was then adjusted between –15 and +15 dB to determine a ΔdB_t giving the maximum mean $\hat{s}_{id}$. The maximum value of mean $\hat{s}_{id}$ (0.83) was obtained for a ΔdB_t of –3 dB.

Results

Preliminary observation of the 16 monospecific sandeel paired datasets showed that echotrace patterns during daylight were predominantly pillar-shaped (tall and thin), and occurred in midwater. The highest acoustic values were detected in the lower parts (deeper) of the echotraces. Often echotrace shapes tend to be more diffuse and spread-out, towards the surface, with the lowest acoustic values detected at the top. This behaviour contrasts with that of herring which form very dense schools, generally near the bottom, during the daylight hours, and more diffuse patterns throughout the water column during hours of darkness (Mackinson, 1999). Sandeel schools were found mainly in shallow areas (about 50 m depth) corresponding to sandy-muddy banks where they burrow. Average sandeel size was 9.5 cm, with a range of 2.5 to 15.4 cm.

The rest of the paired datasets’ hauls were composed of different mixtures of sandeel and other species, where herring, sprat and small gadoids were the most important in numbers. Sandeels were present in very different percentages in these paired datasets, ranging from 0.001% to 98%.

School Detection

The algorithm developed to extract non-fish targets performed extremely well for both frequencies, even given the fact that plankton reflected better at 120 kHz in the majority of cases. The thresholding value of –60 in the *38 kHz fish* and *120 kHz fish* provided a preliminary extraction of plankton. Several convolution matrices were tested for the subsequent blur but the 5*5 kernel described above provided the best results maintaining shape of the schools and smoothing their borders. The threshold value of –60 dB was found, by iterative observation, to perform very well in removing all smoothed edges and diffuse points. The output echograms showed distinctive fish marks aggregated according to different distribution patterns in the water column.

Most schools were correctly detected with the ‘schools module’ algorithm. The detection of the schools was essential for the creation of bitmaps. The school detection parameters were chosen by iteration after successive observations of the outputs on the echogram. Changes in the parameters led to unsatisfactory detection of schools in the dataset. In some cases some very small schools were neglected by the algorithm. Nevertheless, this bias was sacrificed to the efficiency of the identification and did not represent a high percentage in numbers or acoustic returns. This process, however, could noonetheless be subjected to the same optimisation as that for ΔdB_t using the s_{id}.

Sandeel Identification Algorithm

The frequency distributions of the dB differences (MVBS 38 kHz-MVBS 120 kHz) for sandeel schools were centred on negative values but still included a percentage of positive values. The frequency distribution of the ΔdBs for other species overlapped that of sandeel but their values were mainly distributed in the positive range. This is reflected in the analysis of sid precision (Fernandes and Stewart, 2004): according to the bootstrap analysis, 95% of the maximum mean sid were obtained across a broad spectrum of ΔdB_t, between +1 and –5 dB.

Examples of the implementation and validation of the algorithm are given in Figures 2-6. The total numbers of fish caught in the echogram section illustrated are given in each of the figure legends.

The algorithm seems to work well in cases where there are dense well defined schools that are not mixed in with very much plankton (e.g. Fig. 2 and Fig. 3). In one particularly notable case (Fig. 4) a haul of gadoids was taken from a number of distinct, large schools which had the inverse signature to sandeel: i.e. their ΔdB was positive (as evidenced from the red marks in the ΔdB algorithm in Fig. 4). The resulting sandeel echogram was blank, as expected.

Problems arise when the schools are either diffuse or very small in the horizontal dimension (thin echotraces). In the latter cases the ΔdB is not well established perhaps due to beamwidth effects and or transducer placement (see Korneliussen *et al.,* 2004). Diffuse schools are difficult to pick out from plankton clouds and so the algorithm suffers from the application of plankton extraction (which in itself is a similar process) and the inability of the schools detection algorithm to pick out diffuse echotraces. One example of this is given in Figure 5. In this case a number of diffuse echotraces were detected close to the surface. These have not been picked up by the schools detection algorithm. The few that are, do get correctly allocated to sandeel, but many remain undetected. Note that this haul caught mostly herring in the layer just below the sandeel one and that no sandeel were isolated in that layer. The algorithm is, therefore, reasonably conservative (likely to produce an underestimate of abundance) by nature of it dealing only with well defined schools.

In another difficult example (Fig. 6) the problem of thin echotraces arises. In this case a mixture of herring and sandeel does produce a mixture from the algorithm but in some cases the mixture is within the same school. This is difficult to deal with because it implies a stratified mixture (top of the school to one species, bottom to another), which may not reflect reality.

The algorithms output (NASC) was compared with the more traditional visual scrutiny method (Fig. 7). In terms of presence or absence of sandeel, the visual scrutiny agreed with the algorithm in 217 cases out of 327 (66.4%) and tended to estimate lower NASCs where there was agreement. On a number of occasions (three in particular), the algorithm positively identified sandeels that, manually, had been assumed to be herring due to their high density. These gave substantial VBSs within the school at 38 kHz which were assumed in the visual scrutiny, to be too large to attribute to sandeels. Examination of the ΔdB clearly shows a stronger return at 120 kHz than at 38 kHz from this shoal.

The total sandeel biomass estimate for the survey based on the visual scrutiny was 42,550 t. The algorithm provided an estimate of 62,080 t. This difference can be almost entirely attributed to the large NASC at 360 n.mi. (Fig. 8): this contributed to a value of 20,840 t in the area which was counted as herring in the visual scrutiny.

Discussion

There is no generally accepted method for assessing sandeel biomass that takes into account their pelagic phase. Acoustics could be used, but problems with the discrimination of the signals, e.g., overlap with other species, occur. Knowledge of the shape of schools and their location in the water column provides a means of identifying species; however, it is much more subjective and is not precise since some species can show the same distribution patterns.

To date there have been no studies on the identification of sandeel using the dB difference (ΔdB) method. Some other investigations have been carried out on mackerel (Korneliussen and Ona, 2002). A common characteristic of these two species is their lack of a swimbladder, which is likely to make the discrimination between them a difficult task as they are both schooling poelagic species with low target strengths (Armstrong and Edwards, 1985; Armstrong, 1986; Edwards *et al*, 1984).

Fernandes & Stewart (2004) have developed an identification algorithm for mackerel based on the same approach but in their case have included a third frequency (200 kHz). The addition of a higher frequency (>120 kHz) could also improve sandeel identification. Korneliussen and Ona, (2002) used a categorization method based on combined-frequency echograms to quantify the relative contribution of each frequency to the total acoustic backscattering (though a rate of success for comparisons was not provided). However an acoustic identification algorithm must take into account the presence of other co-occurring species that need to be discriminated from the target one. It is likely that the inclusion of more frequencies in the algorithm will allow a multi-species discrimination, provided that different species have different relative frequency responses (Korneliussen and Ona, 2002).

When using two frequencies, care should be taken in interpreting high rates of correct classification where other species are present, as some of the backscattered sound may be derived

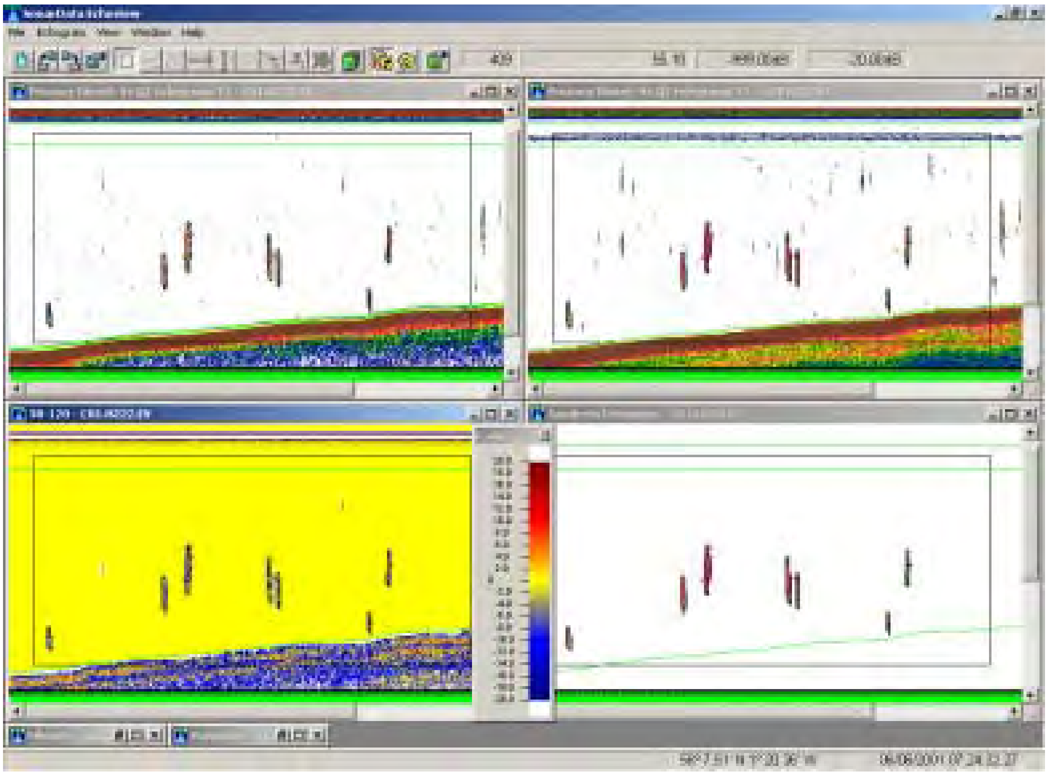

Figure 2. Implementation of the algorithm to an echogram taken in the North Sea in June 2001; depth is approx. 50 m. The figure shows four echogram panels of the same space, differing only in processing: top left = 38 kHz; top right = 120 kHz; bottom left = 38-120 kHz dB difference according to the scale at bottom centre (blue = negative; yellow = 0; red = positive; bottom right = sandeel echogram. Catch from the area defined by the black box = 99040 sandeel and 19 misc others.

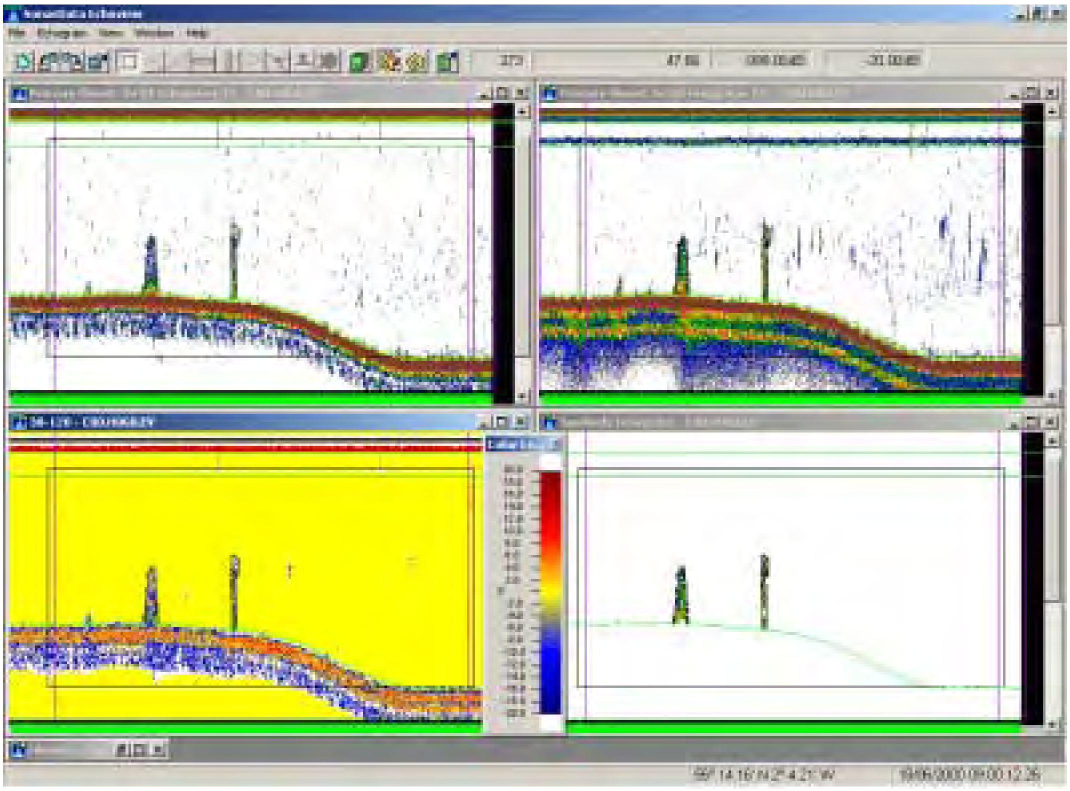

Figure 3. Implementation of the algorithm to an echogram taken in the North Sea in June 2000; depth is approx. 40 m. The figure shows four echogram panels of the same space, differing only in processing: top left = 38 kHz; top right = 120 kHz; bottom left = 38-120 kHz dB difference according to the scale at bottom centre (blue = negative; yellow = 0; red = positive; bottom right = sandeel echogram. Catch from the area defined by the black box = 136634 sandeel, 78 sprat and 42 gadoids.

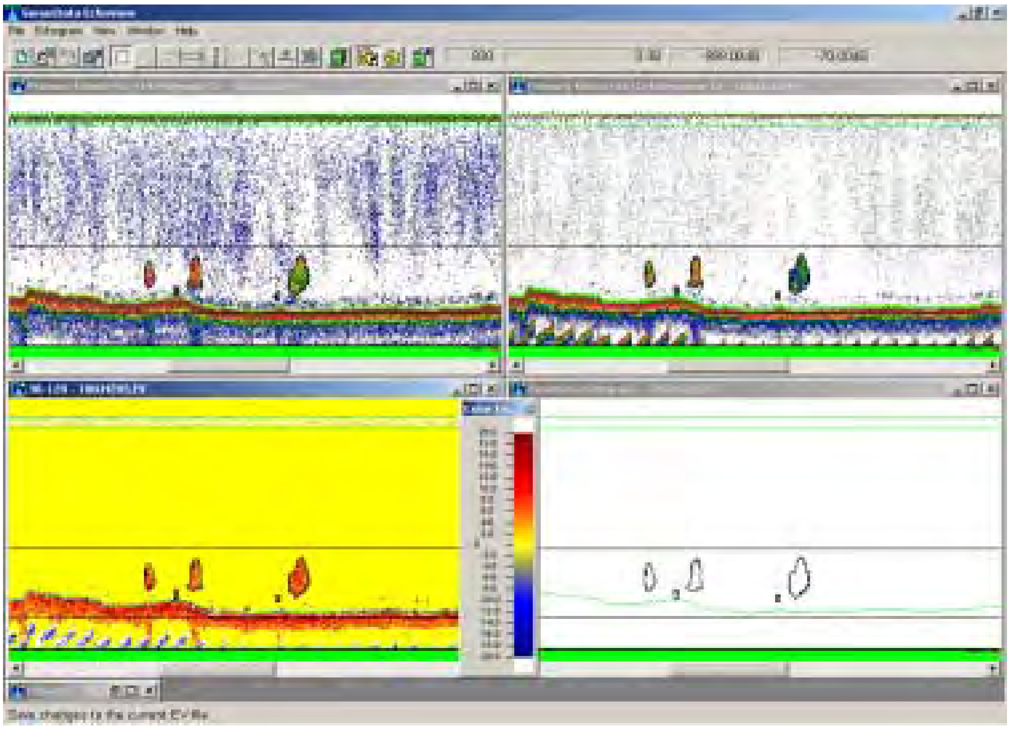

Figure 4. Implementation of the algorithm to an echogram taken in the North Sea in June 2000; depth is approx. 80 m. The figure shows four echogram panels of the same space, differing only in processing: top left = 38 kHz; top right = 120 kHz; bottom left = 38-120 kHz dB difference according to the scale at bottom centre (blue = negative; yellow = 0; red = positive; bottom right = sandeel echogram. Catch from the area defined by the black box = 4 sandeel, 49626 gadoids, 134 herring & 4 misc. others.

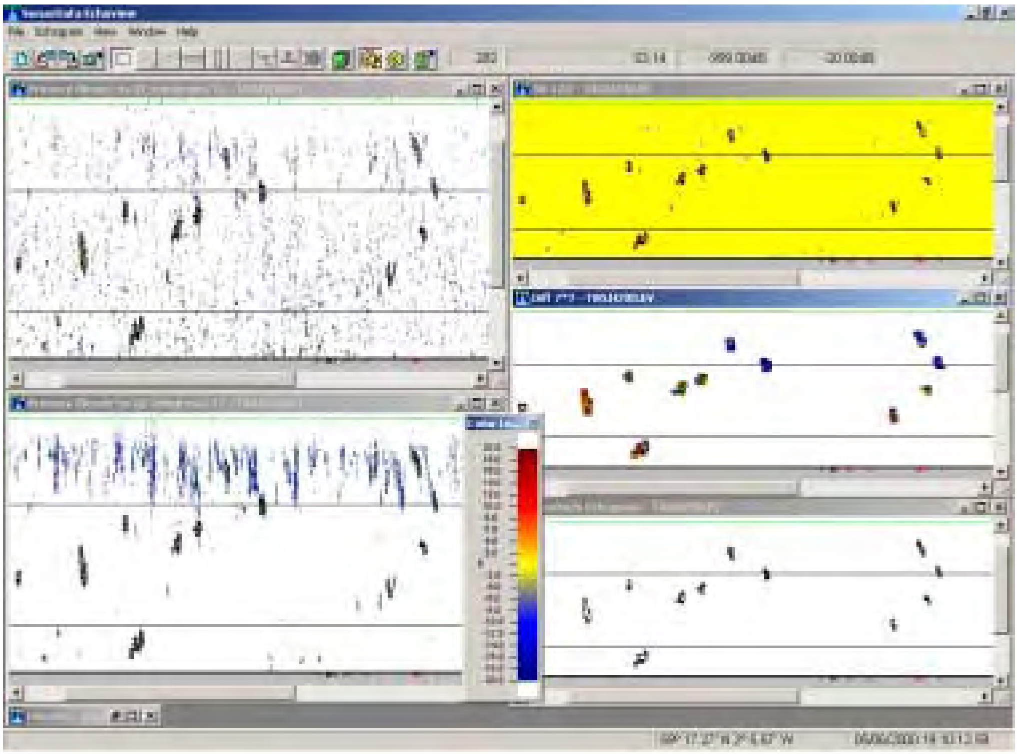

Figure 5. Implementation of the algorithm to an echogram taken in the North Sea in June 2000; depth to the top black line is 32 m. The figure shows five echogram panels of the same space, differing only in processing: top left = 38 kHz; bottom left = 120 kHz; top right = 38-120 kHz dB difference according to the scale at bottom centre (blue = negative; yellow = 0; red = positive; middle right = blurred dB difference (shown to augment the small schools); bottom right = sandeel echogram. Catch from the area defined by the black box = 171 herring, meshed sandeel & 9 gadoids.

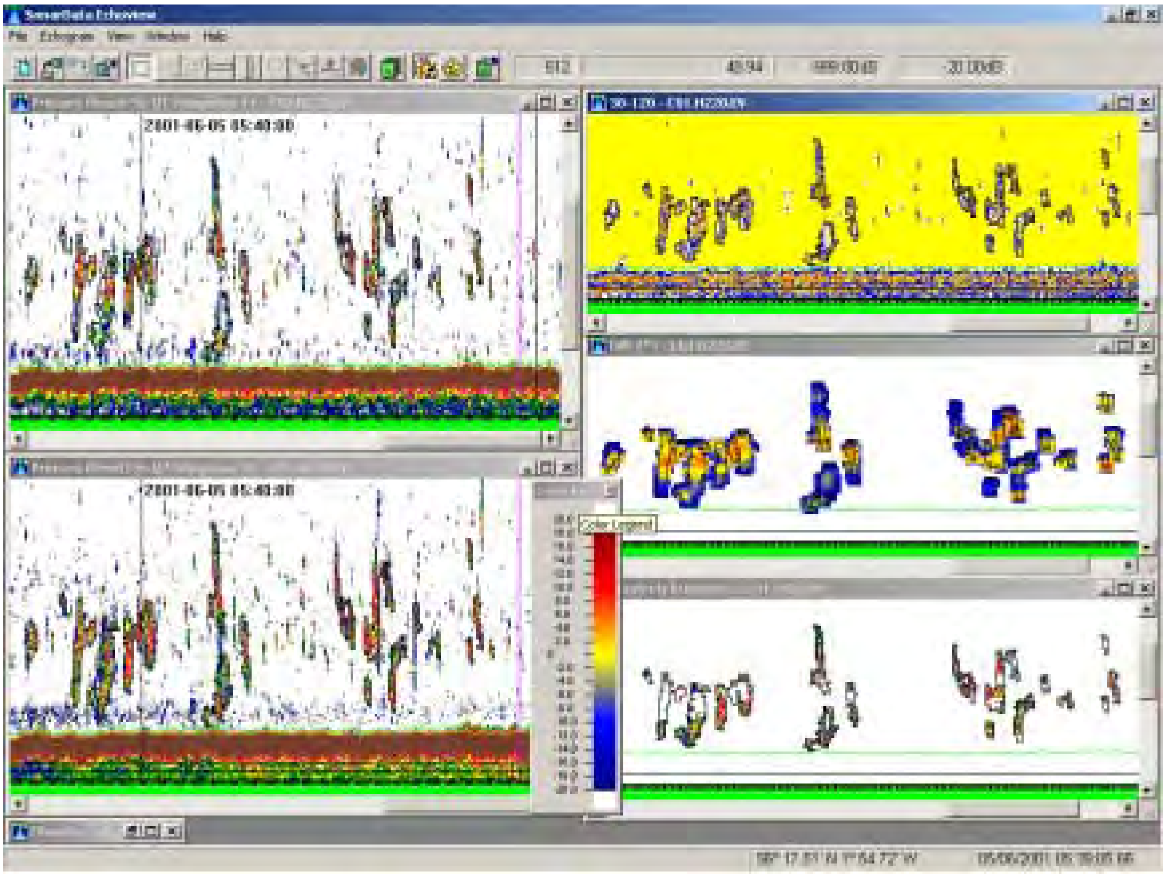

Figure 6. Implementation of the algorithm to an echogram taken in the North Sea in June 2000; depth is approx 45 m. The figure shows five echogram panels of the same space, differing only in processing: top left = 38 kHz; bottom left = 120 kHz; top right = 38-120 kHz dB difference according to the scale at bottom centre (blue = negative; yellow = 0; red = positive; middle right = blurred dB difference (shown to augment the dB difference split inside schools); bottom right = sandeel echogram. Catch from the area defined by the black box = 17500 herring, 31499 sandeel & 56 gadoids.

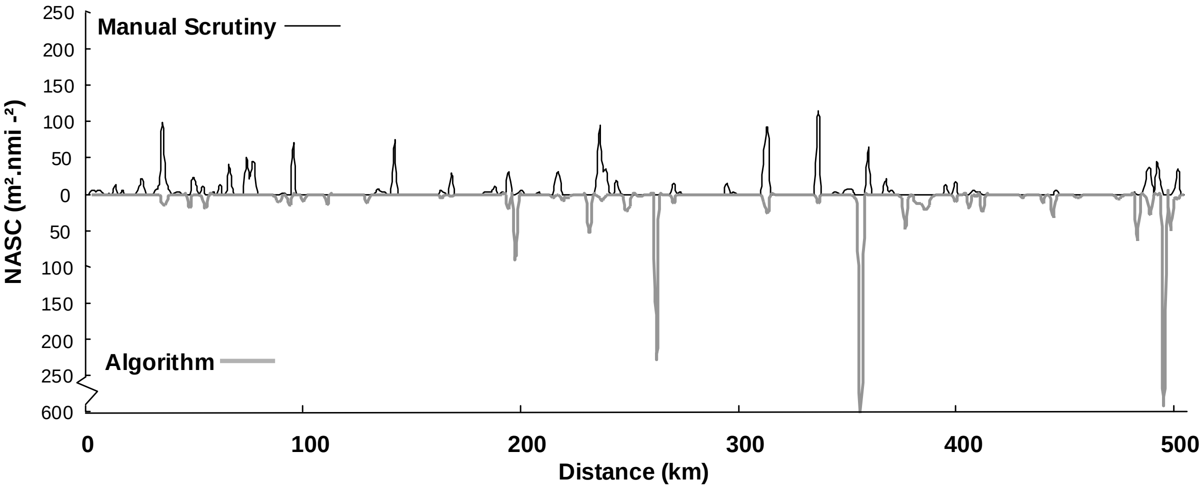

Figure 7. Comparison of NASCs attributed to sandeel from a survey off the Firth of Forth on FRV Clupea June 2003. The values on the scale in the upper portion (thin black line) are derived from a visual scrutiny of the acoustic data: the values in the lower portion (thick grey line) are derived from an the sandeel identifcation algorithm. The large NASC calculated by the algorithm at 360 nmi is 667 m^2.nmi^{-2}.

from the other species in the same range of ΔdB, resulting in misclassification of the signals.

In this study, the presence of other “acoustically similar” species like mackerel was very rare and always in insignificant proportions of the catch; therefore, the results of the developed algorithm should be tested in situations where, for example, mackerel is present. The ΔdB technique has also been used to separate fish and micronekton from plankton taxa (Kang *et al.,* 2002). Three common macroplankters were separated on the basis of their relative echo-strengths at 38 and 120 kHz (Madureira *et al*, 1993). Brierley *et al* (1998) went further, demonstrating that the same technique could discriminate in situ between several zooplanktonic taxa, which sometimes have similar morphological characteristics and have overlapping length-frequency distributions. This was achieved by using discriminant function analysis of differences between mean volume backscattering strength at 38, 120 and 200 kHz giving a correct classification rate of 77% overall. They found that the difference between backscattering at 120 and 38 kHz caused by krill was dependent on its size. Theoretically, sandeel, not having a swimbladder should reflect sound better at higher frequencies, than at lower frequencies. This was found to be the case in the data used for this study but a few exceptions were found; though, it was not possible to establish if different sandeel size resulted in a different range of ΔdBs.

The method for calculating the dB threshold value for the negative mask was different to that used by Brierley *et al.,* (1998). The latter study used a regression analysis of MVBS at 120 kHz versus MVBS at 38 kHz to determine the ΔdB threshold, as described by Madureira *et al.,* (1993). In this study, since sandeel have no swimbladder and all the rest of co-occurring species were swimbladdered, the ΔdB range was remarkably different for the sandeels with regard to the other species present. As a consequence, the threshold value was found to be -3 dB by investigating the s_{id} (Fernandes and Stewart, 2004). Assumptions were made, however, as to the ability of the trawl to catch the identified schools detected in the echogram: this is likely to be a source of error.

Another possible source of error could come from the fact that the net is not an exact sampling device and may not catch the species in their correct proportions. These two facts indicate how important it is to keep a good record of the trawl track and any observations made during trawling (shape of the schools, position in the water column and intensities at the different frequencies used).

The comparative analysis of algorithm output and visual scrutiny showed some significant differences. The algorithm picks out shoals of sandeels which were considered by the scrutiny process to be some other species. The algorithm also excluded traces which were firmly believed to contain sandeels; particularly the lighter scattered traces, which do not form the distinct echotraces required by the algorithm.

In the few cases where the algorithm picked out dense schools with large dB differences characteristic of sandeel, it would seem prudent to assign the echotraces to sandeel in the visual scrutiny process. This would raise the estimate for the survey to a figure that is very close to the biomass calculated using the algorithm. Over the whole survey some gains and losses were made which to some extent cancel each other out. The reason for these differences give indications as to where improvements in the algorithm can be made.

Difficulties in discriminating sandeel schools using the algorithm could be ascribed to the existence of positive values in their ΔdB (38-120 kHz) distributions, which overlap with other swimbladdered species and could therefore misclassify some schools. A suggestion for future development could be to develop a tool (application of the School module on Echoview) that will calculate the mean Sv difference value within the school and then apply this calculated value to each one of the pixels of the school. In this way, sandeel schools, mainly negative but with some positive points within, would be transformed into a completely negative school for which the values of its pixels would be the mean ΔdB. The same could work for other species (mainly swimbladdered) that would be transformed into completely positive schools. Afterwards, the simple application of two masks, one negative and the other positive, would in theory, improve the classification of sandeel schools.

Improvements in the detection of small schools is another key issue for future research since there is a need for detecting thinner sandeel schools while rejecting similar marks from other species (e.g.high plankton concentrations).

The ultimate applications are a more precise stock assessment for commercial species, and an improvement of ecological and population studies (Fernandes *et al*, 2002). Nevertheless, there is still a need to optimize the extraction of information for practical use on large-scale surveys (Korneliussen and Ona, 2002).

Conclusions

Sandeels reflect sound more intensely at 120 kHz than at 38 kHz; many other fish have the opposite trait (higher or equal at 38 kHz). This provided the basis for the development of an algorithm based on the difference in intensity between these two frequencies. The dual frequency algorithm developed for the extraction of plankton performed well, and was particularly useful for isolating fish schools.

Different combinations of plankton extraction, subtraction of virtual echograms (ΔdB) and masking processes, with various threshold values, affected the performance of the algorithm. The final identification algorithm that was developed consisted of the following steps: plankton extraction from calibrated and noise extracted echograms; subtraction 38-120 kHz; school detection; dilation (7*7); schools trimming; negative mask for sandeel (threshold = -3 dB); and application to the 38 kHz echogram for echo integration.

There is scope for improving the algorithm by means of adding other frequencies to the process, and also in applying the algorithm to other species of commercial importance.

Acknowledgements

This work started as a master project supported by the private foundation Pedro Barrie de la Maza, which is therefore thanked for the financial contribution to the realisation of this study. This document was also made with support from the European Commission's Fifth Framework Programme (SIMFAMI project; Grant No. Q5RS-2001-02054). We thank as well, our colleagues of Marine Laboratory, Aberdeen, who provided any kind of help for the purpose of this work.

References

Armstrong, F., 1986. Target strength of sandeels. International Council for the Exploration of the Sea CM 1986/**B**:5.

Armstrong, F. and Edwards, J. I., 1985. Target strength of sandeels. International Council for the Exploration of the Sea CM 1985/**B**:20.

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.

Brierley, A. S., Warde, P. et al., 1998. Acoustic discrimination of Southern Ocean zooplankton. Deep-Sea Research II **45**(7): 1155-1173.

Bodholt, H., Ness, H. and Solli, H., 1989. A new echo-sounder system. Proceedings of the Institute of Acoustics.**11**(3): 123-130.

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

Cushing, D.H., 1957. The interpretation of Echo Traces. Fishery Investigation. Series II, Volume XXI, Number 3. London. Her Majesty`s Stationery Office. Craig, R. E., 1984. Fisheries Acoustics. Introduction, Rapports et Proces-verbaux des Reunions du Conseil International pour l´Exploration de la Mer. **184**.

Edwards, J. I., Armstrong, F. et al., 1984. Herring mackerel and sprat target strength experiments with behavioural observations. International Council for the Exploration of the Sea CM **B**(34).

Everson, I., Bravington, M. et al., 1996. A combined acoustic and trawl survey for efficiently estimating fish abundance. Fisheries Research **26**: 75-91.

Fernandes, P. G. and Simmonds, E. J., 1996. Practical approaches to account for receiver delay and the TVG start time in the calibration of SIMRAD EK500. International Council for the Exploration of the Sea CM 1996 **B: 17**.

Fernandes, P. G., Gerlotto, F., et al., 2002. Acoustic applications in fisheries science: the ICES contribution. ICES Marine Science Symposia **215**.

Fernandes, P.G. and Stewart, M. (2004). Determining the quality of a multifrequency acoustic identification algorithm. *ICES CM 2004/**R**:11.* pp.

Furness, R. W. and Tasker, M. L., 2000. Seabird-fishery interactions: Quantifying the sensitivity of seabirds to reductions in sandeel abundance, and identification of key areas for sensitive seabirds in the North Sea. Marine Ecology Progress Series. vol 202, pp 253-264.

Greenstreet, S.P.R., McMillan, J.A. and Armstrong, F. (1998). Seasonal variation in the importance of pelagic fish in the diet of piscivorous fish in the Moray Firth, NE Scotland: a response to variation in prey abundance? *ICES*

Haralabous, J. and Georgakaratos, S., 1996. Artificial Neural Networks as a tool for species identification of fish schools. ICES Journal of Marine Science **53**: 173-180.

Holliday, D. V., 1992. Zooplankton acoustics. Oceanography of the Indian Ocean, pp. 733-740. Ed. by B. N. Desai. Oxford-IBH, New Delhi. 772 pp

ICES (2004). Report of the Working Group on the assessment of Demersal Stocks in the North Sea and Skagerrak. *ICES CM 2004/**ACFM**:07.* 1-555 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 J. Mar. Sci.* **59**: 794-804.

Korneliussen, R. J. and Ona, E., 2002. An operational system for extraction of plankton and fish from mixed recordings. ICES Journal of Marine Science 59: 293-313.

Lawson, G. L., Barange, M. et al., 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. et al., 2000. Acoustic species identification in the Northwest Atlantic using digital image processing. Fisheries Research **47**: 137-147.

Mackinson, S., 1999. Variation in structure and distribution of pre-spawning Pacific herring shoals in two regions of British Columbia. Journal of Fish Biology **55**: 972-989.

MacLennan, D. N. and Forbes, S. T., 1984. Fisheries acoustics: a review of general principles, Rapports et Proces-verbaux des Reunions du Conseil International pour l´Exploration de la Mer. **184**.

MacLennan, D. N. and Simmonds, E. J., 1992. Fisheries Acoustics. London, Chapman and Hall.

Madureira, L. S. P., Everson, I. et al., 1993. Interpretation of acoustic data at two frequencies to discriminate between Antartic Krill (Eupausia superba Dana) and other scatterers. Journal of Plankton Research **15**(7): 787-802.

McClatchie, S., Thorne, R. E. et al., 2000. Ground truth and target identification for fisheries acoustics. Fisheries Research **47**: 173-191.

Pikitch, E.K., Santora, C., Babcock, E.A., Bakun, A., Bonfil, R., Conover, D.O., Dayton, P., Doukakis, P., Fluharty, D., Heneman, B., Houde, E.D., Link, J., Livingston, P.A., Mangel, M., McAllister, M.K., Pope, J. and Sainsbury, K.J. (2004). Ecosystem-based fishery management. *Science* **305**: 346-347.

Reid, D. and Simmonds, E.J. (1993). Image analysis techniques for the study of fish school structure from acoustic survey data. *Can. J. Fish. Aquat. Sci.* **50**: 1264-1272.

Reid, D.G., Fernandes, P.G., Bethke, E., A., C., Goetze, E., Hakansson, N., Pedersen, J., Staehr, K.J., Simmonds, E.J., Toresen, R. and Tortensen, E. (1998). On Visual Scrutiny of echograms for acoustic stock estimation. *ICES CM 1998/**J**:3.* pp.

Rose, G. A. and Leggett, W. C., 1988. Hydroacoustic Signal Classification of Fish Schools by Species. Canadian Journal of Fisheries and Aquatic Science **45**: 597-604.

Scalabrin, C., Diner, N. et al., 1996. Narrowband acoustic identification of monospecific fish shoals. ICES Journal of Marine Science **53**: 181-188.

Simmonds, E.J., *et al*, 2001. North Sea ICES Herring Acoustic Survey Report Marine Laboratory, Aberdeen.

Tollit, D.J., Greenstreet, S.P.R. and Thompson, P.M. (1997). Prey selection by harbour seals *Phoca vitulina* in relation to variations in prey abundance. *Canadian Journal of Zoology* **75**: 1508-1518

Watkins, J.L. and Brierley, A.S. (1996). A post-processing technique to remove background noise from echo-integration data. *ICES J. Mar. Sci.* **53**: 339-344.

Zakharia, M. E., Magand, F. et al., 1996. Wideband sounder for fish species identification at sea. ICES Journal of Marine Science **53**: 203-208.