Available online at www.sciencedirect.com
SCIENCE DIRECT®
Aquatic Living Resources 16 (2003) 307–312
Aquatic Living Resources
www.elsevier.com/locate/aquliv

Fish schooling behaviour in the northwest North Sea: interspecific associations measured by acoustic survey

Doug J. Beare [a,*], David G. Reid [a], Eddie McKenzie [b]

[a] FRS, Marine Laboratory, Victoria Road, Torry, Aberdeen AB11 9DB, Scotland, UK
[b] Department of Statistics and Modelling Science, University of Strathclyde, Glasgow G1 1XH, Scotland UK

Accepted 3 March 2003

Abstract

This study investigates whether pelagic fish schools of different species or groupings (e.g. herring, "surface herring", "gadoids", mackerel, sprat and sandeels) were positively or negatively associated with each other in time and space. To do this, statistical models were fitted to pre-processed acoustic fisheries data to reveal how pelagic school prevalence varied with respect to spatial (latitude and longitude) and temporal (time of day) information. The model outputs, which take the form of probabilities fitted to the presence or absence of schools, were then used to calculate correlation coefficients, which are useful for measuring association between pairs of variables. Results depended upon the specific species pairs under investigation. Herring and "surface herring" were, for example, very generally sympatrically associated with each other in both space and time, while herring and gadoid schools, on the other hand, had allopatric distributions.

Keywords: Fisheries; Herring; Gadoids; Acoustics; Inter-specific associations; Correlation coefficients; Statistical models

1. Introduction

The Marine Laboratory, Aberdeen conducts annual acoustic surveys of the Orkney/Shetland area during July each year (Bailey et al., 1998). The principal objective of the surveys is to estimate the total biomass of herring (*Clupea harengus*) for use as an index in the annual assessment of the stock (Anonymous, 2001), although the data have also been used to answer questions relating to various aspects of herring biology (Bailey et al., 1998; Maravelias and Reid, 1995, 1997; Maravelias et al., 1996; Reid and Maravelias, 2001). However, data relating to the abundance of other schooling species, also collected during the surveys, have not been analysed until now. In this paper, we address the shortfall by demonstrating how statistical models can be used to expose how school prevalence among different pelagic species groups varies with respect to spatial (latitude, longitude) and temporal (time of day) information (Beare et al., 2002). The baseline data output by these statistical models are then incorporated into an assessment of whether the schools of different species groups are positively, or negatively, associated with each other in space and time. The specific hypothesis addressed by this study can be summarised as follows:

"At any specific point in space, and at any time of day, do schools of species group A avoid, or tend to congregate with, schools of species group B?"

2. Materials and methods

2.1. Acoustic data

The investigation used acoustic survey data collected during six surveys carried out by the FRS Marine Laboratory, Aberdeen during July in 1991 and 1993-1997 (Fig. 1). The data were collected by a Simrad EK500 38-kHz echosounder and stored using the B1500 format, which were then transformed into matrices, each number corresponding to an individual calibrated back-scattering strength sample from a single echosounder transmission. Data summarising the school characteristics were then extracted using image-processing software (ImagePro Plus, Media Cybernetics), combining image filtering algorithms with interactive decisions by the operator (Reid and Simmonds, 1993). The detection threshold for schools was set at -60 dB, providing the best effective beam angle for the school volume backscatter (Sv) of the schools retained (Reid et al., 2000). After thresholding, a single morphological filter pass eliminated small objects and clarified larger ones. Remaining objects detected were classified as schools, and then separated into species

* Corresponding author.
E-mail address: d.beare@marlab.ac.uk (D.J. Beare).

doi:10.1016/S0990-7440(03)00041-X

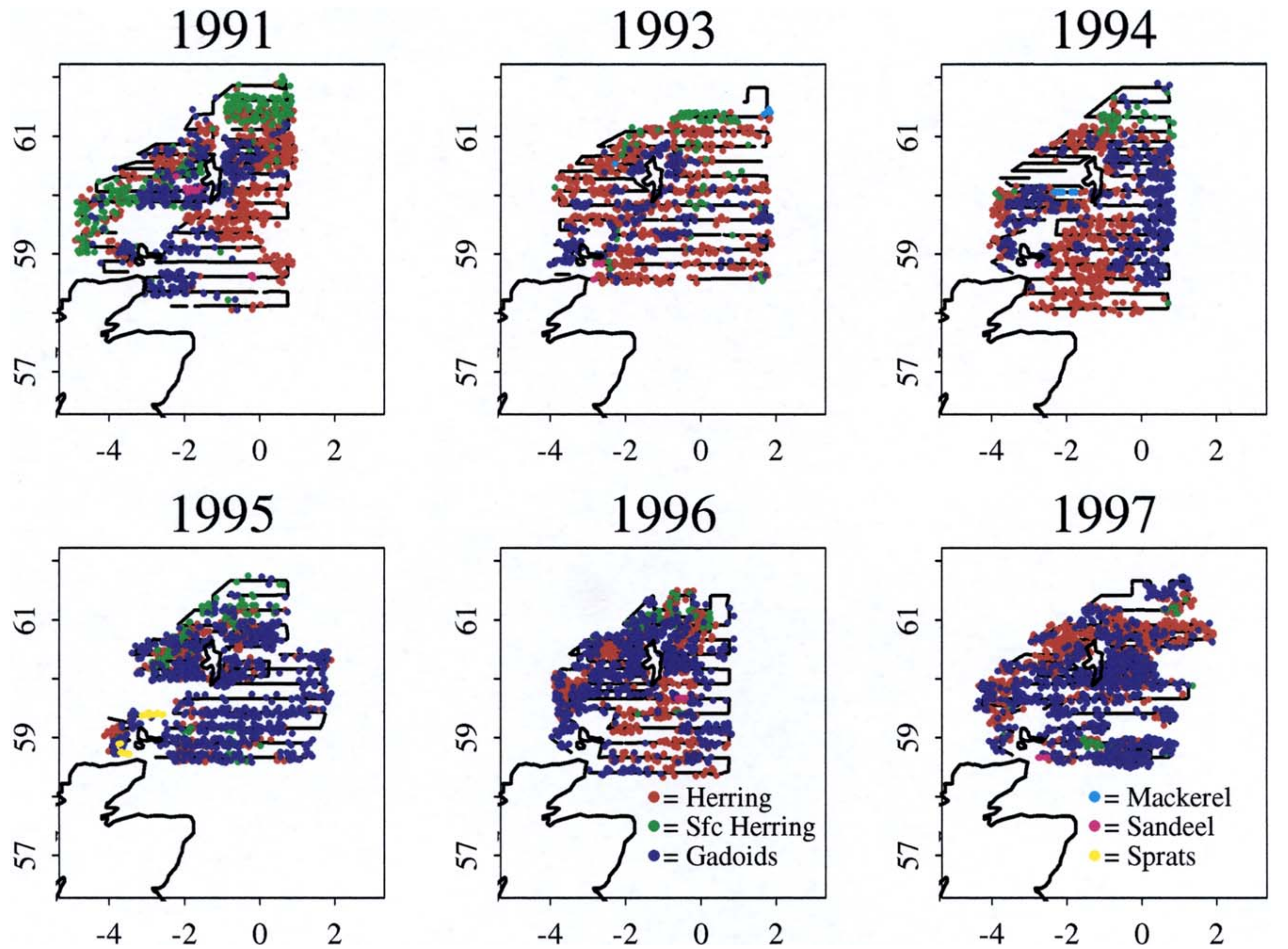

Fig. 1. July acoustic survey data northwest of Scotland. The solid black line represents the survey track; the dots, positive recordings of herring (red), surface herring (green), gadoids (blue), mackerel (turquoise), sandeel (pink), and sprats (yellow) (NB, random noise was added to each observation in both the *x* and *y* dimensions to reflect sample density).

groups. The following taxonomic groups were identified: herring, "surface herring", "gadoids", mackerel, sandeels and sprats. School classification to taxonomic group was carried out according to the standard procedure used for the ICES International Surveys (Simmonds et al., 1992). The allocation of schools to species principally involves the use of ground truth data from pelagic trawling. During the course of an average survey 20-50 trawl stations will be carried out opportunistically on the largest echotraces. During subsequent analyses, the results of these trawl hauls, and the experience of the operator will be used to assign each echotrace to a taxonomic group. The process is subjective to some degree, although it has been shown to be relatively robust and reasonably operator independent (Reid et al., 1998). It was generally possible during these surveys to identify schools of pelagic fish, e.g. herring, mackerel (*Scomber scombrus*), sandeel (*Ammodytes* sp.) and sprats (*Sprattus sprattus*). Most such schools sampled were monospecific, or occasionally mixtures of herring and sprat, or herring and mackerel. Herring schools were generally found, either as small, sharp echotraces near the surface (<20 m), or as relatively high energy marks within 50 m of the seabed. Mackerel schools were identified as large, mid water traces with a weaker backscatter than similar sized herring schools, and they also tended to have a higher backscatter at high frequencies (120 and 200 kHz). Sandeel schools were large and irregular in shape. Sprat schools were similar to the deeper type of herring schools, but could be separated using the trawl data. The "gadoid" category represents a species assemblage made up of Norway pout (*Trisopterus esmarkii*), whiting (*Merlangius merlangus*), haddock (*Melanogrammus aeglefinus*) or saithe (*Pollachius virens*), with smaller numbers of other gadoid species. The fish are usually young (1-2 years), small (maximum 20 cm) and found in discrete schools close to the seabed.

2.2. Statistical analyses

The acoustic data were collected continually along each transect during the surveys (Fig. 1). The transects were then divided into one nautical mile elementary distance sampling units (EDSUs), which minimises the degree of spatial autocorrelation between successive samples (Simmonds et al., 1992). Three species pairs were then selected, based on the amount of available data (e.g. herring and surface herring; herring and gadoids; gadoids and surface herring). For convenience, the species from each pair were denoted as A and B. We then categorised each EDSU with respect to each schooling pair (A, B) as follows: (the corresponding probability is also shown in each case)

- Category 0: neither species A nor B recorded in the EDSU (P_{00});
- Category 1: only species A recorded in the EDSU (P_{A0});
- Category 2: only species B recorded in the EDSU (P_{B0});
- Category 3: both A and B recorded in the EDSU (P_{AB}).

(Note: mackerel, sandeels and sprats could not be examined in such detail due to data-sparsity, and for these groups we noted only whether or not they were recorded in each EDSU with no attention paid to whether other species were present).

For each pair (A, B), the observed EDSUs are a sample from a multinomial distribution such that the probability of observing n_i occurrences of category i for i = 0, 1, 2, 3 in an EDSU is given by

$$P_{00}^{\;n_0} P_{A0}^{\;n_1} P_{B0}^{\;n_2} P_{AB}^{\;n_3} \quad (1)$$

The four probabilities are modelled as functions of predictive variables known for each EDSU. In this case, these are latitude, longitude and time of day. We write the vector of these values for the kth observed EDSU as z_k. Modelling a multinomial distribution can be achieved in various ways (Cox, 1989) but one simple approach involves rewriting Eq. (1) above in terms of nested conditional events. In this case, we obtain

$$\prod_{i=1}^{3} p_i^{\;n_i}(1-p_i)^{n_0+\ldots+n_{i-1}} \quad (2)$$

where $p_3 = P_{AB}$, $p_2 = P_{B0}/(1 - P_{AB})$ and $p_1 = P_{A0}/(1 - P_{AB} - P_{B0})$. Thus, we have reduced the multinomial probability to three binomial probabilities, which can be modelled separately, i.e. we can estimate each of (p_1, p_2, p_3) as functions of the predictor variables z_k. Here, we use generalised additive models (GAMs) for each binomial (Hastie and Tibshirani, 1990). It is straightforward then to recover the original probabilities (P_{00}, P_{A0}, P_{B0}, P_{AB}) as functions of the predictor variables.

Specific details of model selection protocols are described elsewhere in the literature (Augustin et al., 1998; Borchers et al., 1997a,b). Once a GAM was selected for each survey, spatial grids at approximately a 10th of a degree latitude and 5th of a degree longitude for nine different times of day (04:00, 06:00, 08:00, 10:00, 12:00, 14:00, 16:00, 18:00, 21:00 h) were constructed. Parameters from the fitted models were then used to interpolate over the grids producing spatio-temporally (time of day in this context) resolved surfaces, together with the appropriate standard errors.

2.3. Measuring association

The probability values, output by the statistical models (GAMs) were then used to calculate correlation coefficients between each possible pair of species groups according to the following expression:

$$r = \frac{P_{AB} - (P_A \cdot P_B)}{\sqrt{P_A(1-P_A)\,P_B(1-P_B)}} \quad (3)$$

where P_A (or P_B), for example, equals the probability of recording pelagic species group A (or B) within each EDSU whether or not the other species is also recorded, i.e. mathematically, $P_A = P_{A0} + P_{AB}$, and similarly, $P_B = P_{B0} + P_{AB}$. All of the probabilities were estimated, for specific locations and times of day, using the GAMs described above. The correlation coefficient was used here as a measure of the 'association' in time and space between the schools of different species; a high positive value indicating positive association, i.e. a tendency for the two species to be present or absent simultaneously. Negative association or a high negative correlation coefficient is, therefore, a tendency for only one species to be present in an EDSU at a particular time and location.

Note that for binary data, as here, the lower limit of the correlation coefficient r is given, not by -1, but by the greater of $-\sqrt{P_A P_B/(1-P_A)(1-P_B)}$ and its reciprocal (Cox, 1989). Thus, since P_A and P_B are functions of location and time we have divided the negative correlations by the magnitude of the appropriate lower limit so that the output may be interpreted in the usual way.

3. Results

3.1. Prevalence of schools in relation to location and time of day

The data for positive recordings of each species group (Fig. 1) indicate that their spatial distributions have varied considerably between the survey years (July 1991, 1993-1997 (Beare et al., 2002)). Herring schools, for example, were more concentrated around Shetland in 1991, but much more sparsely distributed throughout the survey area in 1993 (Fig. 1). Similarly, gadoid schools were recorded predominantly in the shallow areas between Orkney and Shetland, and around their coasts in 1991 and 1995 (Fig. 1), whereas in 1994 gadoids were more commonly observed in the east of the study area (Fig. 1).

GAMs were fitted to estimate the probabilities as smooth functions of latitude, longitude and hour of day. As an example, the most appropriate model for the 1991 herring data had the functional form: Lo(longitude, latitude) + Lo(time of day), where Lo denotes a locally weighted regression smoother, LOWESS (Cleveland and Devlin, 1988). In scientific terms, the selection of this model implies the assumption that the shape, and not just the level, of herring school prevalence from east to west (longitude) depended crucially on the distance north or south (latitude). Similarly, the selection of time of day independently from the locational information (latitude and longitude), indicates that herring school prevalence has the same shape of diel aggregation/dis-

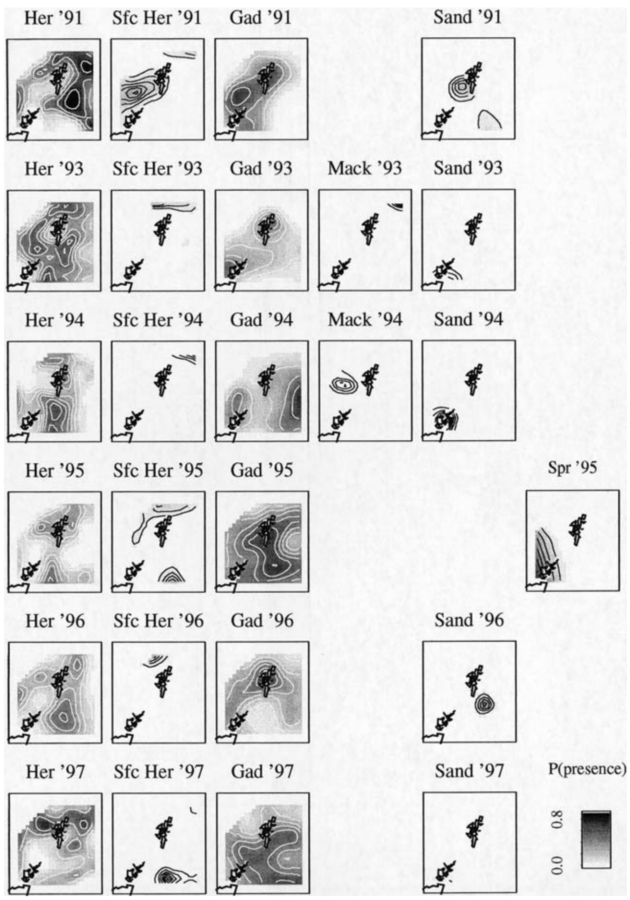

Fig. 2. Prevalence of herring (column 1), surface herring (column 2), gadoids (column 3), mackerel (column 4), sandeel (column 5), and sprats (column 6) estimated for each survey at 10:00 h (July 1991 and 1993-1997) using GAMs.

aggregation behaviour (Pitcher, 1993) across the study region (Fig. 1), and only the overall level (corresponding to abundance) changes according to location. After similar exploration and experimentation, the same, or at least very similar, model formulations were selected to describe the data for the other species groupings and survey combinations.

The spatial patterns output by the selected GAMs for each species grouping and survey year combination are plotted in Fig. 2 . The first column of Fig. 2 represents the probability of recording a herring school given there were no surface herring, during each survey. Herring spatial distributions were fairly similar between years, and high probabilities of occurrence were typical west of Orkney and Shetland, along the continental shelf edge, and also in the Fair Isle Channel area to the southeast of Shetland. Herring were always rare (1991, 1993–1997) over the relatively shallow ridge between Orkney and Shetland (Fig. 2).

In spite of the evidence for a degree of inter-annual stability in spatial distribution, there were also notable differences between the six surveys. In 1991, for example, the chance of recording herring schools was high west of Shetland, but low in the same area between 1994 and 1996 (Fig. 2). Herring schools recorded on, or very near the surface, were comparatively rare, but most likely to be encountered to the north and west of Shetland. The spatial distribution of gadoids was highly variable between surveys, in terms of both shape and level (Fig. 2). In July 1991 and 1993, for example, gadoids were most common over the relatively shallow ridge between Orkney and Shetland, whereas by 1994 they were more abundant east of Shetland (Fig. 2). Mackerel schools were seen west of Shetland during July 1993 and 1994, although they are rarely recorded on acoustic survey, because of a low target strength caused by their lack of a swimbladder (Foote et al., 1987). Sandeel and sprat schools were also comparatively unusual (Fig. 2), but characteristic of the shallower areas around Orkney, Shetland, and mainland Scotland.

Since the effect of time of day on school prevalence was estimated independently from the spatial information (see Table 1), its shape is the same across the entire study area, only the overall level being able to change with location. The chance of recording herring schools was highest during the day between 05:00 and 19:00 h in all 6 years surveyed. Gadoid schools were also most commonly recorded during daylight, although they varied in detail between surveys. Mackerel, sandeel and sprat schools were recorded rarely throughout the surveys, and probabilities were correspondingly low. Mackerel and sprat were more likely to be detected in the afternoon, while sandeel schools were more commonly observed in the morning between 06:00 and 10:00 h.

3.2. Spatial associations for a fixed time of day

In order to examine the association between the different species groups, correlation coefficients, *r*, were calculated using probabilities output by the GAMs described above. Since there was a clear inter-annual, or between survey, effect in the data (see Figs. 1 and 2) correlations were examined separately for each year. Unfortunately mackerel, sandeel and sprat were omitted from this part of the study because of data sparsity caused, either by their failure to be detected by the echosounder, or by their actual relative rarity (e.g. Fig. 2) in the survey area.

Correlation coefficients at 10:00 h, between herring and surface herring for all six surveys are plotted in the first column of Fig. 3, followed by the association between herring and gadoids (column 2, Fig. 3), and finally between surface herring and gadoids (column 3, Fig. 3). Correlations were of mixed sign between herring and surface herring. There were positive associations along the continental shelf north of Shetland, where herring abundance is usually high (Fig. 2), and generally negative associations in the shallow areas where the gadoids are most prevalent. There were large, negative correlations in space between herring and gadoid schools around Orkney and Shetland, although there was some evidence of positive association between herring and gadoids at the periphery of the study region (Fig. 3). The association between surface herring and gadoids was also generally negative in space and varied considerably between surveys (Fig. 3).

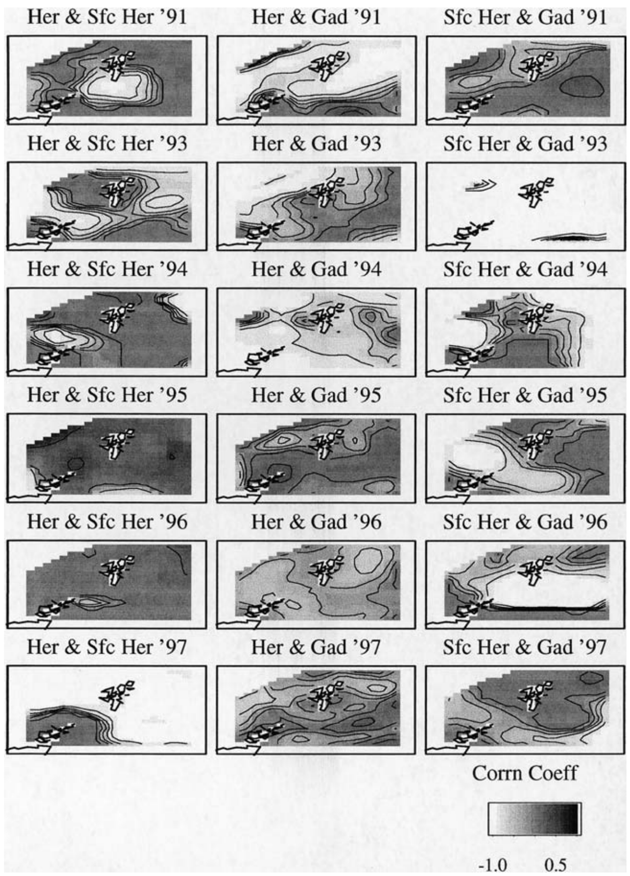

Fig. 3. Spatial variation in association between three species pairs (e.g. herring, surface herring and gadoids) at 10:00 h for each of the six July acoustic surveys. Units are correlation coefficients calculated using Eq. (3).

3.3. *Temporal (time of day) associations for a fixed location*

Correlation coefficients calculated for the 1991 survey are displayed as a function of time of day, for an arbitrarily selected location west of Shetland, in Fig. 4. The horizontal line (Fig. 4) is a zero correlation, below which association is negative and vice versa. The correlation coefficient estimated between herring and surface herring remained negative throughout the day, suggesting fairly stable negative associations (Fig. 4). Herring and gadoids, however, were much more strongly, negatively associated with each other (e.g. Fig. 4). There was a pronounced time of day effect with the herring and gadoid schools least negatively associated with each other at noon.

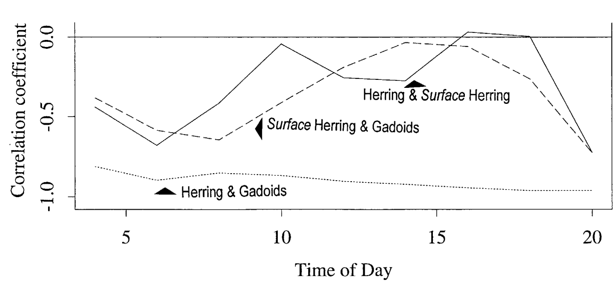

Fig. 4. Daily variation in association between the three species pairs (e.g. herring and surface herring; herring and gadoids; surface herring and gadoids) at an arbitrary location west of Shetland during the 1991 survey. Units are correlation coefficients computed using Eq. (3).

4. Discussion

During the study, we have demonstrated the potential value of acoustic data for exploring the spatial interactions between fish assemblages according to their taxonomic groupings. The use of statistical models has further enabled us to summarise spatial and temporal (time of day) distributions of the various schooling groups. In addition, the probability of positive observations for pairs (e.g. herring and gadoids) of species groups both being recorded, within each EDSU at the same time of day, could be modelled, allowing unusually detailed examination of relative associations between the groups in space/time contexts.

The usual approach to such problems is to calculate Pearson rank correlation coefficient between abundances of, say, species A and species B. But for these data, however, that approach could not be taken because there is only one observation available in each EDSU, i.e. one count of species group A and one of species group B. Binary or presence/absence models were used because most of the data (>80%) are actually zero, i.e. any particular EDSU is most likely to have no schools present at all. Therefore, any costs due to the binary simplification, in terms of lost information at higher levels of school abundance, are compensated for by a statistically robust protocol for dealing with zeros, which ultimately allows reasonable model selection to take place. Moreover, the four probabilities, estimated at each EDSU, also then enable an estimation of 'association' (spatio-temporally resolved correlation coefficients), which would not be possible otherwise because of the data-sparsity alluded to above. A further advantage of our method is that correlation or association is examined at a much higher level of detail than could be achieved using more conventional methodologies. The chance of finding only species A in an EDSU, only species B, both species A and B together, and neither species A or B can be estimated here, and there are thus four different possibilities for correlation between any species pair instead of two.

The schools of most of the species groups studied here were most likely to be recorded by echosounder during the daytime period. This may not be because there are necessarily more fish present, but because of regular diel transitions between densely packed school formations in daylight, and individual, or small groups of, fish in darkness. Schooling is believed to be, primarily, a predator avoidance behaviour (Pitcher, 1993) in that an individual fish is less likely to be eaten in a large aggregation than when alone. Schools are often depicted as breaking up at night as the behaviour is largely in response to visual cues. Interestingly, in this study, sprat schools were most abundant in the evening, which is contrary to typical behaviour evidenced, for example, by herring. It should be noted, however, that sprats are not particularly common in most of the survey area.

The mainly positive associations observed (Fig. 3) between the two categories of herring schools along the continental shelf edge was contrary to expectations. In a previous study, Reid and Maravelias (2001) showed that, over short spatial scales, herring schools related differently to topography and substrate depending on their relative position in the water column. Schools close to the seabed showed good correlation with particular substrates and bottom features, while no such relationship could be seen for the surface schools. The finding of the present work indicates that the two herring school categories are positively associated with each other where herring numbers are high, and for the purposes of stock estimation, could be handled together, using the same biological data (e.g. length frequency and weight-length relationships).

Herring and gadoid schools were more strongly, negatively associated with each other than herring and surface herring (Figs. 3 and 4) suggesting that herring and gadoids occupy different environmental niches. An alternative possibility is that the adult herring may be actively predating on the small gadoids, which make up this assemblage. Again, from a survey analytical perspective, the negative association between the herring and the gadoids is useful, as it suggests that they tend not to be co-located, and so echotraces are less likely to be wrongly allocated. A problem in interpreting this observation is that the "gadoid" category comprises a number of species (predominantly whiting, haddock, Norway pout and saithe, but including others). Additionally, these schools or aggregations would vary in terms of species composition across the survey area. Unfortunately, the survey is not designed to differentiate gadoid species, and this would be difficult to achieve retrospectively, as the trawls were principally conducted to identify pelagic species and differentiate those from the gadoid assemblage.

In conclusion, the present study has shown that the school distribution of a pelagic species such as herring can be shown to have a relationship with the distribution of other fish species, occupying the same general area. In this case, the associations were mostly negative and rather weakly positive (Figs. 3 and 4), suggesting that the schools of pelagic fish occupy space (Fig. 3) and time of day (Fig. 4) in different ways. This information can be added to the corpus of data on other biotic and abiotic correlates with herring distribution presented in previous papers (Bailey et al., 1998; Maravelias and Reid, 1995, 1997; Maravelias et al., 1996). The combined picture from all these studies is one of a complex interaction between the spatial distribution of the herring and many aspects of their environment. In addition, the study has shown that the different types of herring school—surface and deep—are positively associated in space and time of day, and can be analysed as a single population.

Acknowledgements

We would like to thank the European Union for partially funding this work within the CLUSTER project (FAIR CT96.1799) and our colleagues at the Marine Laboratory, Aberdeen who participated in the herring acoustic surveys. We would also like to acknowledge Mr. Richard Aukland for developing the latest version of our image processing system, and for extracting the "school" database used in this study.

References

Anonymous, 2001. Report of the herring assessment working group for the area south of 62N. ICES CM 2001/ACFM:12, ICES, Hamburg.

Augustin, N.H., Borchers, D.L., Clarke, E.D., Buckland, S.T., Walsh, M., 1998. Spatiotemporal modelling for the annual egg production methods of stock assessment using generalized additive models. Can. J. Fish. Aquat. Sci. 55, 2608–2621.

Bailey, M.C., Maravelias, C.D., Simmonds, E.J., 1998. Changes in the spatial distribution of autumn spawning herring (*Clupea harengus* L.) derived from annual acoustic surveys during the period 1984-1996. ICES J. Mar. Sci. 55, 545–555.

Beare, D.J., Reid, D.G., Petitgas, P., 2002. Spatio-temporal patterns in herring (*Clupea harengus* L.) school abundance and size in the NW North Sea: modelling space time dependencies to allow examination of the impact of local school size. ICES J. Mar. Sci. 59, 469–479.

Borchers, D.L., Buckland, S.T., Priede, I.G., Ahmadi, S., 1997. Improving the precision of the daily egg production method using generalized additive models. Can. J. Fish. Aquat. Sci. 54, 2727–2742.

Borchers, D.L., Richardson, A., Motos, L., 1997. Modelling the spatial distribution of fish eggs using generalized additive models. Ozeanografika 1, 103–120.

Cleveland, W.S., Devlin, S.J., 1988. Locally-weighted regression: an approach to regression analysis by local fitting. J. Am. Stat. Assoc. 83, 596–610.

Cox, D.R., 1989. Analysis of Binary Data. Chapman and Hall,, London.

Foote, K.G., Knudsen, H.P., Vestnes, G., MacLennan, D.N., Simmonds, E.J., 1987. Calibration of acoustic instruments for fish density estimation: a practical guide. ICES Coop. Res. Rep., 144.

Hastie, T., Tibshirani, R., 1990. Generalized additive models. Chapman and Hall, London.

Maravelias, C.D., Reid, D.G., 1995. Relationship between herring (*Clupea harengus*, L.) distribution and sea surface salinity and temperature in the northern North Sea. Sci. Mar. 59, 427–438.

Maravelias, C.D., Reid, D.G., 1997. Identifying the effects of oceanographic features and zooplankton on prespawning herring abundances using generalized additive models. Mar. Ecol. Prog. Ser. 147, 1–9.

Maravelias, C.D., Reid, D.G., Simmonds, E.J., Haralabous, J., 1996. Spatial analysis and mapping of acoustic survey data in the presence of high local variability: geostatistical application to the North Sea herring. Can. J. Fish. Aquat. Sci. 53 (7), 1497–1505.

Pitcher, T.J., 1993. Ecology, behaviour and exploitation in fish shoals. Nord. Semin. -Arb. Rapp. 572, 53–56.

Reid, D.G., Maravelias, C.D., 2001. Relationships between herring school distribution and seabed substrate derived from RoxAnn. ICES J. Mar. Sci. 58, 1161–1173.

Reid, D.G., Scalabin, C., Petitgas, P., Messe, J., Aukland, R., Carrerra, P., Georgakarakos, S., 2000. Standard protocols for the analysis of school based data from echosounder surveys. Fish. Res. 47, 125–136.

Reid, D.G., Simmonds, E.J., 1993. Image analysis techniques for the study of fish school structure from acoustic survey data. Can. J. Fish. Aquat. Sci. 50, 886–893.

Reid, D.G., Fernandes, P.G., Bethke, E., Coopers, A., Goetze, E., Makansson, N., Pedersen, J., Staehr, K.J., Simmonds, E.J., Toresen, R., Tortensen, E., 1998. On visual scrutiny of echograms for acoustic stock estimation. ICESCM/J, 3.

Simmonds, E.J., Williamson, N.J., Gerlotto, F., Aglen, A., 1992. Acoustic survey design and analysis procedure: a comprehensive review of current practice. ICES Coop. Res. Rep., 187.