ICES Journal of Marine Science, 58: 1150–1160. 2001
doi:10.1006/jmsc.2001.1130, available online at http://www.idealibrary.com on IDEAL®

# On the relation between schools, clusters of schools, and abundance in pelagic fish stocks

P. Petitgas, D. Reid, P. Carrera, M. Iglesias,
S. Georgakarakos, B. Liorzou, and J. Massé

Petitgas, P., Reid, D., Carrera, P., Iglesias, M., Georgakarakos, S., Liorzou, B., and Massé, J. 2001. On the relation between schools, clusters of schools, and abundance in pelagic fish stocks. – ICES Journal of Marine Research, 58: 1150–1160.

Small pelagic fish are known to aggregate into schools and clusters of schools. It is commonly assumed that the number of such schools and clusters, as well as their size and densities, will vary with the stock abundance. We have carried out a PCA based meta-analysis, using series of acoustic survey data from five different locations in Europe to examine this assumption. The study concluded that there was no discernible relationship between stock abundance and the number of schools seen, or on the clustering of those schools. The study also showed that the number and structure of the school clusters was strongly correlated with the number of schools seen. An increased number of schools in an area tended to be linked with denser clusters (more schools per kilometre) and a higher occupation of the survey area by those clusters. There was also a weaker tendency to find more clusters. It is not clear whether these relationships and the absence of a link to abundance are due to density independence in aggregation patterns or whether such density dependence is only functional at relatively low stock abundance levels.



Keywords: small pelagics, acoustic surveys, schools, clusters, aggregation, density dependence, PCA with passive variables, meta-analysis.

Received 17 April 2001; accepted 1 August 2001.

*P. Petitgas and J. Massé: IFREMER, Laboratoire Ecologie Halieutique, rue de l'ile d'Yeu, BP 21105, F-44311 cedex 03, Nantes, France; tel: +33 2 40 37 41 63; e-mail: Pierre.Petitgas@ifremer.fr and Jacques.Masse@ifremer.fr. D. Reid: Fisheries Research Services, Marine Laboratory Aberdeen, Victoria Road, Aberdeen, AB11 9DB, Scotland, UK; tel: +44 1224 295363; e-mail: reiddg@marlab.ac.uk. P. Carrera: Instituto Español de Oceanografía, Centro Oceanografico de A Coruña, PO Box 130, 15080 A Coruña, Spain; tel: +34 81 22 90 77; e-mail: pablo.carrera@co.ieo.es. M. Iglesias: Instituto Español de Oceanografia, Centro Oceanografico, Apdo. 291, Palma de Mallorca; tel: +34 71 40 15 61; e-mail: magdalena.iglesias@ba.ieo.es. S. Georgekarakos: Institute of Marine Biology of Crete, PO Box 2214, Main Port of Heraklion, GR-71003 Heraklion, Greece; tel: +30 81 24 18 82; e-mail: stratis@imbc.gr. B. Liorzou: IFREMER, 1 rue Vilar, F-34200, Sete; e-mail: Bernard.Liorzou@ifremer.fr. Correspondence to P. Petitgas.*

## Introduction

Density dependence of spatial distribution in marine fish stocks is an important subject for fisheries scientists because it may affect both stock catchability and the results of assessment surveys. Paloheimo and Dickie (1964) first raised the issue of density-dependent catchability in pelagic fish and suggested that pelagic fish may reduce their spatial extension to maintain constant density within schools. MacCall (1990) explained changes in spatial distribution with abundance using density-dependent habitat selection and competition: the "basin effect". Range-collapse together with abundance-collapse has been observed in stocks and different statistical tools have been proposed to identify how local densities are related with global abundance (Myers and Stokes, 1989; Swain and Sinclair, 1994; Petitgas, 1998). All the above are largely concerned with the overall area occupied by the stock, rather than the spatial organisation of that stock. The relationships between the schools, their spatial distribution, and the stock abundance have attracted much less attention. An understanding of this would be needed in modelling the interaction between fish aggregation pattern and fishing strategy. Different modelling approaches have been undertaken (Mangel and Beder, 1985; Petitgas and Laloë, 1998; Maury and Gascuel, 1999) but they were all based on theoretical scenarios. For a review of these scenarios together with their potential consequences on fishing catches see Fréon and Misund (1999). The EU funded CLUSTER project (1997–1999) was dedicated to filling this perceived gap by characterising schools and clusters of schools in European pelagic fish stocks and analysing how these characteristics varied with overall population abundance. In essence, the project set out to determine the pattern of spatial occupation in a variety of different pelagic stocks across a range of stock-state scenarios.

We can regard fish aggregation at three distinct scales: fish are aggregated into schools, schools are aggregated into clusters, and clusters are aggregated into stock units. The organisation of schools in clusters of schools has already been described in many pelagic stocks (Cram and Hampton 1976; MacLennan and MacKenzie, 1988; Swartzman *et al*., 1994; Petitgas and Lévénez, 1996; Mackinson *et al*., 1999) but the details of these school clusters and their spatial organisation have not been studied. In this study we have grouped the schools into clusters and have estimated parameters to characterise both scales of spatial organisation. We then examined how the school parameters and the cluster parameters inter-related and how both related to the overall population abundance; i.e. how the different scales in the spatial organisation interacted.

The object of this paper then is to determine the relationship between the spatial organisation of the stock and the state of that stock, i.e. its abundance. Intuitively, and most straightforward, if stock biomass increased one might expect to encounter more schools, or larger, denser schools. Given that there may be a higher scale of aggregation, i.e. at the cluster level, one might expect more clusters, or larger, denser clusters or that the stock may occupy a larger area. The converse, of smaller, less dense schools and clusters would be expected as stock biomass decreased. First, interactions between school and cluster parameters were analysed using Principal Component Analysis (PCA; Lebart *et al*., 1995). Then, to understand how the interactions were related to population abundance, population abundance was positioned in the factorial space as a passive variable. Because we intended to work with a set of parameters describing different aspects and scales in the spatial organisation in the population, a large number of surveys (i.e. population observations) were needed. Therefore we have used a meta-analysis approach by pooling together different surveys on a range of different pelagic stocks. Such an approach is now common in fisheries science (Myers and Mertz, 1998) because it allows the highlighting of common general features of fish stock populations.

# Materials and methods

## Fisheries surveys and school echo-traces

Historical acoustic fisheries assessment surveys were analysed. For this study 26 surveys were available which were carried out in the period 1991–1997 in the following European areas: (i) the Aegean Sea: The Gulf of Thessaloniki, Greece; (ii) the western Mediterranean: the Gulf of Lions, France and the Catalan Sea, Spain; (iii) Western European platform: Spanish and French coasts of the Bay of Biscay; and (iv) the northern North Sea: Orkney/Shetland area, Scotland, UK.

Details of the surveys are summarised in Table 1 and Figure 1. All stocks were of small pelagic species. The target species were anchovy (*Engraulis encrasicholus*) and sardine (*Sardina pilchardus*) in all areas except the North Sea where it was herring (*Clupea harengus*). Non-target species were also present in the survey data including mackerel (*Scomber scombrus*), horse-mackerel (*Trachurus trachurus*), blue-whiting (*Micromesistius poutassou*) and norway pout (*Trisopterus esmarki*). Digital echo-sounders now permit digital echograms to be saved as sequences of pixels or as images. Image-analysis algorithms have been applied on digital echograms to identify school echo-traces and estimate parameters of their characterisation and classification. A review of different procedures is given in are report by ICES (2000). All surveys were analysed and school databases extracted using the same acoustic processing threshold of $-60$ dB. Daytime school echo-traces for each survey were extracted and characterised using different image-analysis softwares (Reid and Simmonds, 1993; Georgakarakos and Paterakis, 1993; Weill *et al*., 1993) in the different areas. Schools were not identified to species except for the Orkney-Shetland areas. This was because, apart from the North Sea, identification of schools to species using only the school echo-trace parameters at the 38 kHz frequency was considered to be too imprecise (Scalabrin *et al*., 1996). All schools for all species were considered for all surveys except for the Orkney-Shetland surveys for which only herring schools were considered as these surveys provide assessments for herring only.

## Choice of school parameters

Reid *et al*. (2000) proposed a list of parameters for the characterisation of schools and clusters. These parameters can be grouped into three categories, morphological (length, height, area, perimeter, outline rugosity, etc.), energetic (acoustic back-scattering energy, density, internal variability, etc.) and positional (latitude, longitude, time, depth from surface, altitude from bottom). Standard definitions and notations were proposed in ICES (2000) for school parameters and these have

Table 1. Surveys, areas, and species considered in the meta-analysis. Surveys 96A and 96B of IMBC are repeated surveys. Each survey in the Tables and Figures is labelled using the code and year. So, the Scottish, North Sea survey in 1991 would be A91 and so on.

| Institutes | Code | Survey years | Survey location (area) | Target species | Non-target species |
|---|---|---|---|---|---|
| FRS-MLA, Aberdeen, Scotland | A | 1991, 93, 94, 95, 96, 97 | North Sea: Orkney/Shetland, (81 520 nmi$^2$) | Herring<br>Sprat | Gadoids and mackerel |
| IFREMER, Nantes, France | N | 1991, 92, 94, 97 | Biscay: SW France (20 214 nmi$^2$) | Anchovy<br>Sardine | Sprat, Mackerel spp., Horse Mackerel spp., and Gadoids |
| IEO, A Coruña, Spain | C | 1992, 93, 95, 96, 97 | Biscay: Cantabrian Sea North Spain (25 609 nmi$^2$) | Anchovy<br>Sardine | Mackerel spp., Horse Mackerel spp., and Gadoids |
| IEO, Palma de Mallorca, Spain | P | 1992, 93, 95, 96 | Mediterranean: Catalonian Sea, Spain (9877 nmi$^2$) | Anchovy<br>Sardine | Sardinella spp. |
| IFREMER, Sète, France | S | 1993, 95, 96, 97 | Mediterranean: Gulf of Lions (5439 nmi$^2$) | Anchovy<br>Sardine | Sprat, Mackerel spp., Horse Mackerel spp., and Sardinella spp. |
| IMBC, Heraklion, Greece | H | 1996A, 96B, 97 | Aegean Sea: Gulf of Thessaloniki (1555 nmi$^2$) | Anchovy<br>Sardine | Horse Mackerel spp. and Gadoids |

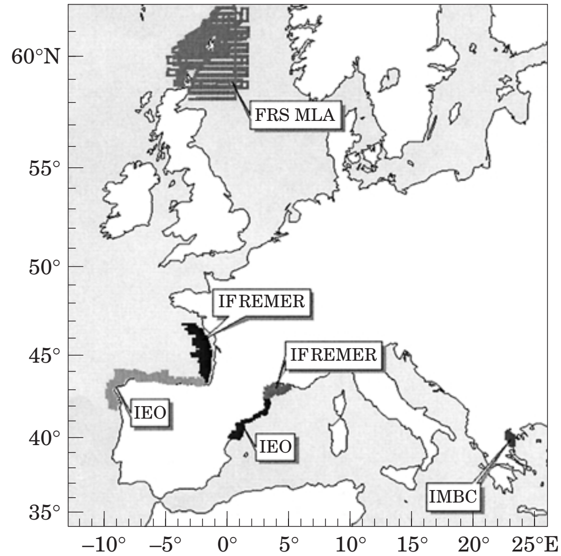


Figure 1. Map of the areas surveyed by the different institutes who collaborated in the present study. FRS-MLA: Marine Laboratory Aberdeen, Scotland, UK. IFREMER: Institut Français de Recherche pour l'Exploitation de la Mer, France. IEO: Instituto Español de Oceanografía, Spain. IMBC: Institute of Marine Biology of Crete, Greece. See Table 1 for details on the acoustic survey series.

been used here. For species identification purposes, Haralabous and Georgakarakos (1996) have shown the importance of including many or all of these descriptors. However for this analysis we have selected a reduced parameter set from each category to characterise the schools. The school parameters selected for this study were: (i) SMsca: School cross-sectional area ($m^2$); (ii) SEtot: School acoustic back-scattered total energy as an index of school biomass ($m^2$ $nmi^{-2}$); and (iii) SPdep: School depth, i.e. the distance between sea surface and school centre (m).

School echo-traces are actually a distorted image of cross sections of real schools. Therefore it is necessary to correct school echo-trace parameters to estimate school parameters. An algorithm was developed by Diner (ICES, 2000) to make these corrections and we used it for all the surveys in this analysis in the estimation of the school area. More details on the protocol used for the surveys are given in Petitgas *et al.* (1998).

## Grouping schools in clusters of schools

We considered schools as being discrete events in space (i.e. a point process). The distance to the next-neighbour (NND) was computed for each school along the surveyed acoustic transects. We then attempted to group schools into clusters based on these distances. Essentially this involved setting a threshold NND beyond which the next school was taken to be in a different cluster (i.e. maximum distance between schools in a cluster). There is little behavioural knowledge with which to define such threshold distance *a priori* and historically this has always been done using an empirical approach. Basically the researcher defines a distance based on his observations. Here we used a statistical approach based on a multi-criteria algorithm to define the threshold distance. The algorithm divides the spatial point process of school occurrence into clusters of schools that have similar internal characteristics. The reference conceptual Point Process model was the Matern cluster process (Stoyan and Stoyan, 1994) although the algorithm did not explicitly estimate the parameters of such model.

For a given threshold distance schools were grouped into clusters and the following parameters were then estimated: (i) the number of clusters; (ii) their length; (iii) the number of schools per unit-cluster-length; (iv) the number of solitary schools; and (v) the homogeneity of the spatial distribution of schools within a cluster.

A range of thresholds were considered and one retained which minimised the following empirical criteria: (i) not too many clusters; (ii) not too many solitary schools; (iii) homogeneous spatial distribution within clusters; and (iv) high $r^2$ for the regression of the number of schools in clusters on the length of clusters.

Schools can occur at different depths in the water column. For the purposes of this analysis the school locations were vertically collapsed to two dimensions (latitude and longitude), with no reference to depth. The distance between schools was computed as the horizontal distance between their centres.

## Choice of cluster parameters

Cluster parameters were chosen to characterise the spatial organisation at this scale. They were grouped into two categories: first, the occupation of space by clusters and second, the occupation of clusters by schools.

Two parameters were selected to characterise the occupation of space by clusters: (i) Nclu: the number of clusters; and (ii) Lclu/Lsur: the ratio of the summed cluster lengths to the summed transect lengths.

Two parameters were selected to characterise the occupation of clusters by schools: (i) Slop: the slope of the regression between the number of schools and cluster lengths. The regression was forced to pass through the origin thus the slope estimates the average number of schools per km in the clusters; and (ii) Lsch/Lclu: the ratio of the summed school lengths to the summed cluster lengths.

Cluster length was defined as the distance between centres of the first and last school in the cluster. The

Table 2. Matrix of parameter values characterising schools and clusters (active variables, i.e. matrix Z in the text) and of population parameters to be explained (passive variables, i.e. matrix $Z^+$ in the text). SMcsa=School area ($m^2$), SEtot=School acoustic backscatter ($m^2$ $nmi^{-2}$), SPdep=School depth (m), Nclu=Number of clusters, Lclu/Lsur=summed cluster lengths/survey length, Slop=average school number per km in the clusters, Lsch/Lclu=average ratio for summed school lengths/cluster length. Abu=survey biomass estimate (thousand tonnes), Ntot=survey estimate of total school number. Codes for surveys are detailed in Table 1. The variable TD is the threshold distance (m) defining the maximum distance between two schools in a cluster.

| | | Active variables | | | | | | | Passive variables | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Schools | | | Clusters | | | | Population | |
| Code | TD | SMcsa | SEtot | SPdep | Nclu | Lclu/Lsur | Slop | Lsch/Lclu | Ntot | Abu |
| A91 | 2 947 | 30.0 | 194.8 | 127.1 | 164 | 0.145 | 3.092 | 0.073 | 135 050 | 1 259 |
| A93 | 6 004 | 47.9 | 261.9 | 102.7 | 150 | 0.217 | 1.807 | 0.053 | 116 596 | 865 |
| A94 | 4 803 | 45.7 | 201.0 | 105.0 | 137 | 0.195 | 2.409 | 0.056 | 130 856 | 740 |
| A95 | 4 603 | 64.6 | 486.6 | 109.6 | 73 | 0.112 | 2.261 | 0.073 | 79 408 | 797 |
| A96 | 2 802 | 56.7 | 536.9 | 122.0 | 170 | 0.188 | 3.784 | 0.089 | 166 086 | 1 376 |
| A97 | 3 923 | 58.2 | 261.4 | 127.0 | 159 | 0.207 | 2.772 | 0.073 | 157 418 | 1 480 |
| H96a | 410 | 21.2 | 44.2 | 22.6 | 33 | 0.147 | 18.7 | 0.706 | 54 487 | 97 |
| H96b | 420 | 28.5 | 202.4 | 27.6 | 57 | 0.177 | 8.5 | 0.368 | 10 134 | 86 |
| H97 | 1 140 | 30.2 | 95.3 | 12.1 | 109 | 0.443 | 5.3 | 0.211 | 13 280 | 50 |
| P92 | 2 356 | 46.7 | 250.5 | 46.7 | 27 | 0.114 | 1.729 | 0.053 | 8 063 | 178 |
| P93 | 2 178 | 24.0 | 161.2 | 28.3 | 42 | 0.18 | 2.79 | 0.056 | 18 700 | 147 |
| P95 | 2 889 | 37.3 | 151.0 | 40.9 | 28 | 0.156 | 2.015 | 0.043 | 11 315 | 100 |
| P96 | 3 000 | 32.4 | 113.0 | 37.9 | 56 | 0.277 | 2.874 | 0.052 | 24 358 | 93 |
| S93 | 2 211 | 20.8 | 49.3 | 20.3 | 55 | 0.417 | 2.79 | 0.059 | 20 334 | 179 |
| S95 | 1 945 | 3.3 | 28.8 | 21.0 | 62 | 0.534 | 4.16 | 0.067 | 34 885 | 122 |
| S96 | 1 945 | 2.1 | 16.4 | 18.3 | 61 | 0.636 | 8.4 | 0.061 | 69 211 | 97 |
| S97 | 2 000 | 6.8 | 23.5 | 12.8 | 59 | 0.542 | 3.2 | 0.048 | 28 916 | 93 |
| C92 | 4 111 | 105.2 | 152.0 | 70.5 | 64 | 0.126 | 0.92 | 0.05 | 52 878 | 318 |
| C93 | 2 778 | 79.5 | 179.1 | 51.6 | 81 | 0.076 | 1.569 | 0.068 | 56 918 | 792 |
| C95 | 2 806 | 72.3 | 209.3 | 60.9 | 53 | 0.167 | 1.439 | 0.068 | 71 938 | 125 |
| C96 | 2 111 | 41.8 | 156.4 | 48.2 | 44 | 0.085 | 4.238 | 0.098 | 9 398 | 326 |
| C97 | 2 522 | 77.5 | 206.8 | 52.1 | 46 | 0.052 | 1.735 | 0.081 | 55 688 | 457 |
| N91 | 3 334 | 30.5 | 123.7 | 46.8 | 60 | 0.495 | 2.36 | 0.061 | 63 716 | 431 |
| N92 | 2 080 | 19.0 | 84.9 | 31.2 | 46 | 0.581 | 5.21 | 0.091 | 165 565 | 361 |
| N94 | 4 445 | 20.2 | 107.1 | 32.0 | 52 | 0.327 | 1.77 | 0.03 | 31 754 | 213 |
| N97 | 3 334 | 9.9 | 115.7 | 30.9 | 79 | 0.532 | 2.66 | 0.028 | 81 951 | 682 |

average for each parameter was computed for each survey.

## PCA with added passive variables

From the above steps we produced a table, table Z (see Table 2), comprising n=26 surveys (lines) and p=7 parameters (columns) characterising the spatial organisation (schools and clusters of schools). The aim was then to analyse the relationship between: (i) Abu: Population abundance: the survey biomass estimate in thousand tonnes; (ii) Ntot: The total school number: the survey estimate of the average number of schools per km multiplied by the area surveyed in $km^2$; and (iii) and the seven school and cluster parameters.

Because the number of observations is limited (26 surveys), we restricted the list of school and cluster parameters.

Principal Component Analysis (PCA) allows the summarising of the multi-variate correlations using factors of a limited number by grouping and ordering the many relations between the parameters (Lebart *et al*., 1995). PCA was thus used on the correlation matrix of the seven spatial structure parameters to analyse their linear interactions. The technique of passive variables (Lebart *et al*., 1995) was then used to analyse the correlation between population abundance, school number, and the set of spatial structure parameters. Population abundance and school number were projected on the factors resulting from the PCA.

In the factorial space the proximity between the position of the parameters is interpreted in terms of correlation. Let $Cjj'$ denote the correlation coefficient between parameter j and j′, the distance between point variables Pj and Pj′ is $2(1 - Cjj')$. Consider the vector for parameter j between the point origin O and the point

Table 3. Correlation coefficients between variables. The matrix corresponding to the active variables is $X^tX$ on which PCA was performed. The two columns for the passive variables were appended. SMcsa=School area, SEtot=School energy, SPdep=School depth, Nclu=Number of clusters, Lclu/Lsur=summed cluster lengths/survey length, Slop=average school number per km in the clusters, Lsch/Lclu=average of summed school lengths/cluster length. Abu=survey biomass estimate, Ntot=survey estimate of total school number.

| | Active variables | | | | | | | Passive variables | |
|---|---|---|---|---|---|---|---|---|---|
| | SMcsa | SEtot | SPdep | Nclu | Lclu/Lsur | Slop | Lsch/Lclu | Abu | Ntot |
| SMcsa | 1.000 | 0.511 | 0.385 | −0.202 | −0.129 | −0.610 | −0.293 | 0.003 | −0.354 |
| SEtot | 0.511 | 1.000 | 0.212 | −0.478 | −0.586 | −0.311 | 0.073 | 0.228 | −0.687 |
| SPdep | 0.385 | 0.212 | 1.000 | −0.101 | −0.173 | −0.090 | 0.173 | 0.354 | 0.056 |
| Nclu | −0.202 | −0.478 | −0.101 | 1.000 | 0.600 | 0.030 | −0.203 | −0.002 | 0.314 |
| Lclu/Lsur | −0.129 | −0.586 | −0.173 | 0.600 | 1.000 | 0.219 | −0.136 | −0.414 | 0.529 |
| Slop | −0.610 | −0.311 | −0.090 | 0.030 | 0.219 | 1.000 | 0.730 | 0.092 | 0.620 |
| Lsch/Lclu | −0.293 | 0.073 | 0.173 | −0.203 | −0.136 | 0.730 | 1.000 | 0.368 | 0.347 |
| Abu | 0.003 | 0.228 | 0.354 | −0.002 | −0.414 | 0.092 | 0.368 | 1.000 | 0.067 |
| Ntot | −0.354 | −0.687 | 0.056 | 0.314 | 0.529 | 0.620 | 0.347 | 0.067 | 1.000 |

variable Pj, OPj, the cosine between OPj and OPj′ is Cjj′. Therefore if one projects in the factorial space a supplementary parameter which has not served in the principal component analysis, its position allows the interpretation of how the added passive parameter correlates with the interaction of the active parameters.

Let X denote the matrix such that $X^tX$ is the correlation matrix of the parameters. The diagonalisation of $X^tX$ gives the eigen vectors $\mu_\alpha$ corresponding to the eigen values $\lambda_\alpha$:

$$X^t \times \mu_\alpha = \lambda_\alpha \, \mu_\alpha$$

In the factorial space of the parameters the coordinates of the point- surveys on the principal axis α are the components of the vector:

$$\psi_\alpha = X \, \mu_\alpha$$

In the factorial space of the observations the coordinates of the active point- parameters on the principal axis α are the components of the vector:

$$\varphi_\alpha = X^tX \, \mu_\alpha/\sqrt{\lambda_\alpha} = \mu_\alpha \, \sqrt{\lambda_\alpha}$$

Let $X^+$ denote the matrix made of $p^+$ columns of added passive variables and n lines. In the factorial space of the observations the coordinates of the passive added point-parameters on the principal axis α are the components of the vector:

$$\varphi^+_\alpha = X^{+t} \, X \, \mu_\alpha/\sqrt{\lambda_\alpha}$$

More details can be found in Lebart *et al.* (1995).

Now from table Z we deduce matrix X. The different areas and stocks have different mean values for their parameters. In order to pool all surveys in a single meta-analysis we first computed standardised residuals for each area. For survey i, parameter j and area a, the residual was:

$$r_{ija} = (z_{ija} - m_{ja})/(\sigma_{ja} \sqrt{n_a})$$

where z is the parameter value, m its mean for the surveys of the area, σ its standard deviation for the surveys of the area and n the number of surveys for the area. Let $Y_a$ denote the matrix of residuals with element $r_{ija}$. We considered matrix Y made by appending the sub matrices $Y_a$. The matrix X considered was $X = Y/\sqrt{6}$ as we have six areas (see Table 1). The correlation matrix analysed by PCA was $X^tX = Y^tY/6$, the average correlation matrix for all areas.

## Results

Table 2 presents the average parameter values estimated for the schools and the clusters. It also includes the survey abundance estimate (biomass) and the total number of schools. These can be denoted as matrices Z and $Z^+$. Table 2 also includes the threshold NND allowing the grouping of schools in clusters although this parameter is not included in the PCA analysis. It is worth noting that there often appears to be one survey in each area for which one parameter is very different from the other surveys in that area. For instance, the 1995 North Sea survey (A95) found much fewer schools than the other North Sea surveys. The overall correlation matrix ($X^tX$ ) is given in Table 3. The Principal Component Analysis of this matrix allows the description of the structure of the linear interactions between the school and cluster parameters.

Table 4 gives the “percent of variance” explained by the principal components. The first three components

Table 4. Cumulative percent of variance explained by the principal components.

| | Comp.1 | Comp.2 | Comp.3 | Comp.4 | Comp.5 | Comp.6 | Comp.7 |
|---|---|---|---|---|---|---|---|
| %var | 0.377 | 0.653 | 0.807 | 0.886 | 0.952 | 0.982 | 1.000 |

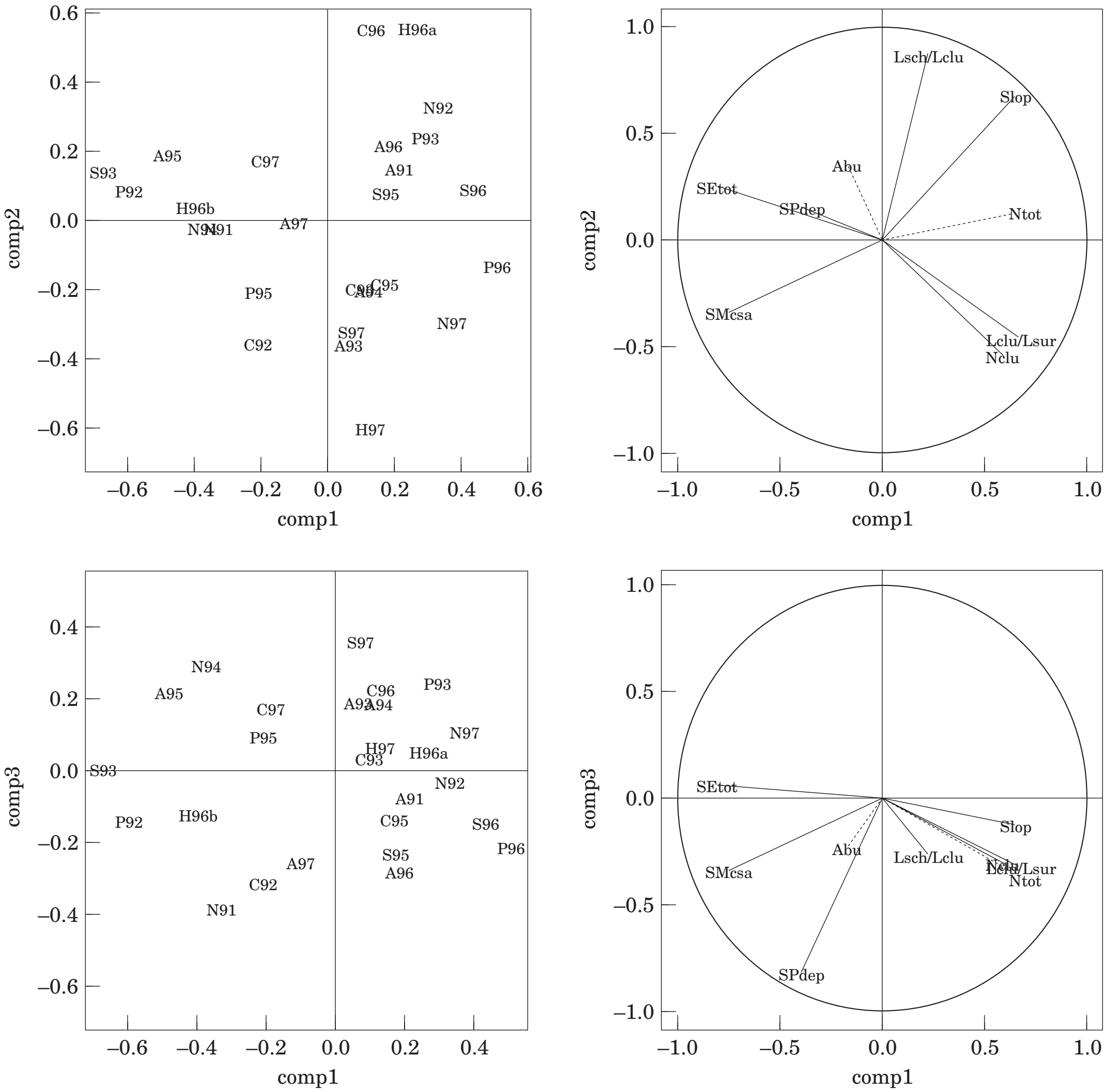


Figure 2. A representation in the first factorial plane of the survey observations (left) and of the parameters (right). Parameters are represented in their circle of correlation. Active parameters in the PCA are: SMcsa=School area ($m^2$), SEtot=School acoustic backscatter ($m^2\ nmi^{-2}$), SPdep=School depth (m), Nclu=Number of clusters, Lclu/Lsur=summed cluster lengths/survey length, Slop=average school number per km in the clusters, Lsch/Lclu=summed school lengths/cluster length, averaged for all clusters. Passive parameters in the PCA are: Abu=survey biomass estimate (thousand tonnes), Ntot=survey estimate of total school number.

accounted for 80% of the variance of matrix X. Figure 2 shows the results of the PCA in the factorial planes (1,2) and (1,3). The first component was defined by the opposition between the school parameters (SEtot, SPdep, and SMcsa) and the cluster parameters (particularly Lclu/Lsur and Nclu) and represented 38% of total variance. When the average school size was larger (high SMcsa and SEtot), there were fewer clusters and these occupied less of the overall surveyed area. The index of school biomass (SEtot) was the parameter most correlated with the first component. Component 2 was defined by the occupation of clusters by schools (principally Lsch/Lclu but also Slop). This component represented 28% of the total variance. Component 3 was largely defined by the vertical position of the schools and represented 15% of the variance.

The total school number (Ntot) was correlated with component 1 (Figure 2 and Table 3) and was positioned between the occupation of the area by clusters (Lclu/Lsur and Nclu) and the occupation of clusters by schools (Lsch/Lclu and Slop). It was diametrically opposite the school descriptors and in particular school biomass index (SEtot). So when we have more schools, we would expect: (i) smaller and lower biomass schools; (ii) more clusters; (iii) clusters to occupy more of the survey area; and (iv) more densely packed schools within clusters.

The abundance (Abu) showed only a weak relationship to any of the other variables (Figure 2 and Table 3) the correlation coefficients being less than 0.4. Also the abundance (Abu) showed no relationship with the number of schools (Ntot) (coefficient of correlation of 0.07). On Figure 2 in the factorial space of the three first components the position of abundance could be interpreted as standing a little closer to the school descriptors, and in opposition to the number of clusters and their occupation of space (Lclu/Lsurv).

To investigate in more detail the relationship between school size and abundance we examined the proportion of "rich" schools in each area. A rich school was defined as one with a value greater than that of the 0.9 quartile in each area. Hence for each survey area we pooled all the schools and computed the value corresponding to the 0.9 quantile. Then we computed the percentage of schools in each individual year that was greater than this value. Figures 3 and 4 show the percentage of "rich" schools plotted by survey and year against the average school biomass (SEtot) and population abundance (Abu). These are presented as deviations from the area mean. So a positive value represents a year in which the average school biomass was above the mean for that area. Figure 3 shows that there is a strong link (correlation coefficient of 0.816) between average school biomass and the presence of more rich schools. Figure 4 shows a weaker link (correlation coefficient of 0.394) between rich schools and the stock abundance.

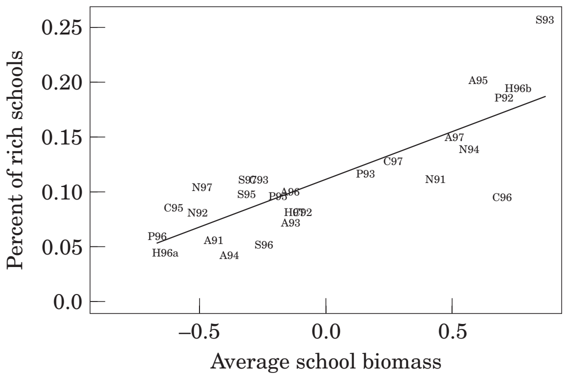


Figure 3. The relationship between the average school biomass and the proportion of rich schools. Average school biomass is expressed in residuals relative to each area mean. Rich schools are defined as schools with biomass greater than a threshold (the 90th centile of all schools pooled together for each area). Labels represent the survey codes (See Table 1 for details of codes). The regression is $y=0.112+0.088x$. The $r^2$ is 0.67.

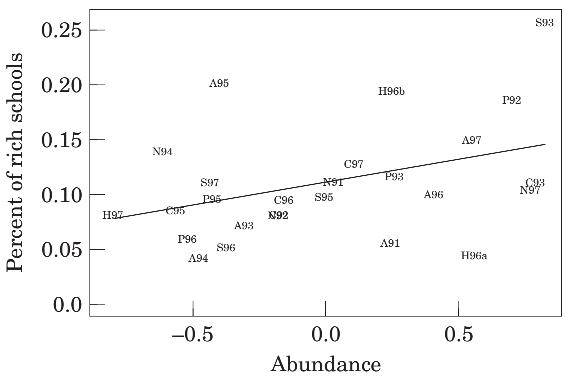


Figure 4. The relationship between stock abundance and the proportion of rich schools. Stock abundance is expressed in residuals relative to each area mean. Rich schools are defined as schools with biomass greater than a threshold (the 90th centile of all schools pooled together for each area). Labels represent the survey codes (See Table 1 for details of codes). The regression is $y=0.112+0.0424x$. The $r^2$ is 0.16.

# Discussion

## Number of schools and stock abundance

A key result from this meta-analytical approach is that there was no discernible relationship between school number and stock abundance. Aukland and Reid (1998) also found such an absence for North Sea herring. More fish did not mean more schools but more fish resulted in some schools being larger in biomass, as shown in Figure 4. This suggests that stock size was not the primary factor driving schooling aggregation. The pattern of schooling aggregation was driven by other factors which could be of an environmental and behavioural nature. Indeed the number of schools and the school parameters have been shown to vary under a variety of such factors, e.g. time of day (Beare *et al.*, 2000; Fréon and Misund, 1999), distance to coast (Petitgas and Lévénez, 1996), heterogeneity of the sea bed (Reid, 2000), oceanographic conditions (Swartzman *et al.*, 1994), predator, prey, and fish community (Swartzman, 1991; Schneider, 1989; Massé *et al.*, 1996) and seasonal behavioural change in feeding, spawning, and wintering (Fréon and Misund, 1999).

## Number of schools and clusters

The number of schools found on a survey is clearly linked to the pattern of spatial organisation in clusters. In the correlation matrix school number showed higher correlation with the occupation of the survey area by clusters and the density of schools (number of schools per kilometre) in those clusters. School number showed weaker correlation with the number of clusters and the area occupied by schools within those clusters. Along the first principal component the number of schools (Ntot) was strongly linked with the number of clusters, the occupation of the survey area by clusters and the density of schools in those clusters. Our interpretation of these observations is that when there are more schools the strongest effect is that the density of schools in the clusters increases (Slop). At the same time the proportion of the survey track occupied by the clusters increases (Lclu/Lsur). To a lesser extent the number of clusters (Nclu) also increases. There is little evidence that the schools are occupying more of the area of the clusters (Lsch/Lclu).

## School size and number of schools

There was a strong negative link between school number (Ntot) and school area (SMcsa) along the first principal component. So more schools means a smaller average school size. This fits well with the previous observation that school number shows a weak link with the occupation of cluster area (Lsch/Lclu).

## School size and stock abundance

There was evidence that higher stock biomass was associated with some schools being larger in biomass but the number of schools and their pattern of aggregation in clusters was not related to stock size. This suggests that when the stock level was high, there was a slight tendency to find a larger average for the school biomass. The average value itself may increase as a result of all schools generally being bigger or because the proportion of "rich" schools increases. In other words any additional biomass may contribute to a *pro rata* increase in size of all schools or it may preferentially increase the number of large schools. Fréon and Misund (1999) reviewed the variation in school size across a variety of stocks, ecosystems, years and seasons. They identified no general rule and, in particular, no fixed value of school biomass which could be used to define a rich school. Therefore we used the simple statistical approach described above. The percent of rich schools was well correlated to the average school biomass suggesting that the variation in the mean school biomass can be attributed to the variation in the proportion of rich schools. There was also a relationship between the proportion of rich schools and the overall abundance. So when the stock abundance was high in a particular year there were more rich schools. This may partly explain why an increase in abundance was not linked to an increase in school number in the analysis. What appears to happen is that, in situations of high stock abundance, we get larger schools and we get more very large schools, rather than simply more schools. Petitgas and Lévénez (1996) showed a similar relationship between the proportion of rich schools and biomass in small pelagic fish in Senegalese waters.

## Clusters and stock abundance

The stock abundance showed only slight relationships with any of the parameters describing the occupation of space in clusters of schools. This may now not be surprising given the lack of a link between abundance and number of schools, and the existence of a link between number of schools and their spatial aggregation pattern. Petitgas and Samb (1998) also found no relationship between abundance and cluster parameters (number and length) in Senegalese waters.

## Implications of the present findings

The implication of these findings is that the aggregation pattern is density-independent, i.e. independent of the biomass of the stock, at least in the small pelagic stocks studied and for the range of variation observed in the stock sizes. Another possibility is that the stocks studied here do not include a wide enough range of stock size scenarios to allow density dependence to be seen acting. In the present study the limits of stock size are best known for North Sea herring where the stock varied between 500 000 and 1 000 000 tonnes over the period under review when its historical low was at less than 100 000 tonnes.

The occupation of space is related to the notion of habitat. The present results suggest that the "basin effect" based on density dependent habitat selection and intra-specific competition (McCall, 1990) which relates stock size to spatial organisation, would act only for extreme values of abundance, high and low. For intermediate values of stock size fish abundance would not be the limiting factor and thus the spatial organisation of the stock would be more dependent on environmental parameters than on straight abundance. One consequence of this is that because fishermen exploit aggregations, the catchability of the stock would largely be dependent on the environmentally induced aggregation pattern.

This can be considered as analogous to the search for meaningful stock recruit relationships. It is self evident that if we have no fish there will be no recruitment, and here, no schools or clusters. When the stock is reasonably healthy we often get substantial variation in recruitment which links poorly to the stock abundance. Equally in this case we have no link between abundance and aggregation parameters. This may explain why the literature provides a variety of situations: e.g. range collapse at low stock size (Myers and Stokes, 1989; Swain and Sinclair 1994; Petitgas, 1998) and density independence of spatial pattern (Swain and Morin, 1996; Petitgas and Lévénez, 1996).

Other confounding factors may include our inability to desegregate into species (except for the North Sea herring data). The surveys may not in all cases cover the whole range of the stocks and so effects may be taking place beyond the boundaries. Second, schools were grouped in clusters based on a threshold distance and this value plays a determinant role. Most commonly in ecology the threshold distance for clustering is set empirically. The operator chooses a distance based on his observations. Soria *et al*. (1998) used a fixed threshold distance of 1 km. Here we used a variable threshold based on the features in the data, particularly internal cluster homogeneity. Either approach will be an approximation. Schools are dynamic objects which interact with one another. Any single threshold is unlikely to capture this process fully. Finally, in our approach, empty segments inside clusters would be counted as occupied area (area occupied by clusters), while areas counted as empty would include only those empty segments between the clusters. This also may impact the relation between occupied stock area and stock size.

A conceptual model can be proposed from our findings to allow the modelling of the relationship between fish capture and the stock spatial aggregation pattern. First, envisage a scenario of a few schools and a few clusters in an area. What happens if we add more schools to the picture? From our results it would appear that the extra schools would mostly join the existing clusters, making them denser (more schools per kilometre) and longer (more of the survey track occupied by clusters). At the same time some of the schools go to make up new clusters and so the number of clusters increases with school number. However, the correlation here is much lower than for cluster density and length so this effect would be less important than the "strengthening" of the original clusters. There was a strong negative link between school number (Ntot) and school area (SMcsa). So more schools means smaller schools. In this conceptual model the extra schools added tend to be smaller, so the area occupied within the clusters does not vary as much as would be expected if the school size was not related to school number. Clusters can be interpreted as groups of schools sharing the same habitat and interacting with each other more than with other schools.

## Conclusion

To our knowledge this study is the first attempt in the fisheries literature to estimate from survey data parameters characterising various levels in the spatial organisation of pelagic fish and investigate the relationships between schools, school clusters, and population parameters. School parameters were anti-correlated with cluster parameters meaning that larger schools meant smaller clusters occupying less space and *vice versa*. The total number of schools in the area was well correlated with the schooling aggregation pattern: more schools meant smaller schools organised in more denser clusters occupying more space. In contrast population abundance showed less correlation either with school or cluster parameters. Further, population abundance was not correlated with total school number. We interpreted this result by the fact that population abundance was not seen at extreme values for the years studied and, therefore, would not have been a forcing parameter on the spatial aggregation of schools. The schooling organisation being one component of catchability, catchability would not have been density dependent for the years studied here. An important implication is that stock catchability could vary for other causes than abundance at intermediate abundance levels.

## Acknowledgements

We would like to acknowledge the European Union for joint funding of the CLUSTER program (FAIR CT 96.1799). Particular thanks go to the colleagues who have replayed entire surveys and extracted school echotraces for constituting the school data base: Mr R. Aukland (Aberdeen), Ms N. Schreiber (Nantes), Mr P. Beillois (Nantes), Dr R. Muiño (A Coruña). We thank the audience of theme session K "Incorporation of External Factors in Marine Resource Surveys" at the ICES Annual Science Conference, Bruges, Belgium (Sept 2000), and particularly Odd Nakken for discussing some aspects of the present analysis. Dr J. Rivoirard and an anonymous referee are thanked for their constructive review of the manuscript.

## References

Aukland, R., and Reid, D. 1998. The impact of changing stock size on the aggregative behaviour of North Sea herring. ICES CM 1998/J: 2, 19 pp.

Beare, D., Reid, D., Petitgas, P., Carrera, P., Georgakarakos, S., Haralambous, J., Iglesias, M., Liorzou, B., Massé, J., and Muiño, R. 2000. Spatio-temporal patterns in pelagic fish school abundance and size: a study of pelagic fish aggregation using acoustic surveys from Senegal to Shetland. ICES CM 2000/K: 03, 29 pp.

Cram, D., and Hampton, J. 1976. A proposed aerial/acoustic strategy for pelagic fish stock assessment. Journal du Conseil International pour l'Exploration de la Mer, 37: 91–97.

Fréon, P., and Misund, O. 1999. Dynamics of pelagic fish distribution and behaviour: effects on fisheries and stock assessment. Fishing News Books, Blackwell Science Ltd, Oxford. 348 pp.

Georgakarakos, S., and Paterakis, G. 1993. "School": a software for fish school identification. ICES CM 1993/B: 8, 14 pp.

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

ICES. 2000. Report on Echotrace Classification. Ed. by D. Reid. ICES Cooperative Research Report, No. 238, 107 pp.

Lebart, L., Morineau, A., and Piron, M. 1995. Statistique exploratoire et multidimensionelle. Dunod, Paris. 439 pp.

MacCall, A. 1990. Dynamic geography of marine fish populations. University of Washington Press, Seattle. 153 pp.

Mackinson, S., Nøttestad, L., Guénette, S., Pitcher, T., Misund, O., and Fernö, A. 1999. Cross-scale observations on distribution and behaviour dynamics of ocean feeding Norwegian spring-spawning herring. ICES Journal of Marine Science, 56: 613–626.

MacLennan, D., and MacKenzie, I. 1988. Precision of acoustic fish stock estimates. Canadian Journal of Fisheries and Aquatic Sciences, 45: 605–616.

Mangel, M., and Beder, J. 1985. Search and stock depletion: theory and applications. Canadian Journal of Fisheries and Aquatic Sciences, 42: 150–163.

Massé, J., Koutsikopoulos, C., and Patty, W. 1996. The structure and spatial distribution of pelagic fish schools in multispecies clusters: an acoustic study. ICES Journal of Marine Science, 53: 155–160.

Maury, O., and Gascuel, D. 1999. SHADYS ("simulateur halieutique de dynamiques spatiales"), a GIS based numerical model of fisheries, example application: the study of a marine protected area. Aquatic Living Resources, 12: 77–88.

Myers, R., and Mertz, G. 1998. Reducing uncertainty in the biological basis of fisheries management by meta-analysis of data from many populations: a synthesis. Fisheries Research, 37: 51–60.

Myers, R., and Stokes, K. 1989. Density dependent habitat utilization of groundfish and the optimisation of research surveys. ICES CM 1989/D: 15, 11 pp.

Paloheimo, J., and Dickie, L. 1964. Abundance and fishing success. Rapport et Procès-Verbaux des Réunions du Conseil International pour l'Exploration de la Mer, 155: 152–163.

Petitgas, P. 1998. Biomass dependent dynamics of fish spatial distributions characterised by geostatistical aggregation curves. ICES Journal of Marine Science, 55: 443–453.

Petitgas, P., Diner, N., Georgakarakos, S., Reid, D., Aukland, R., Massé, J., Scalabrin, C., Iglesias, M., Muino, R., and Carrera, P. 1998. Sensitivity analysis of school parameters to compare schools from different surveys: a review of the task standardization of the EC FAIR program CLUSTER. ICES CM 1998/J: 23, 18 pp.

Petitgas, P., and Laloë, F. 1998. Représentation simulée d'un processus de recherche et de capture de bancs de poissons répartis selon un processus ponctuel marqué. *In* Biométrie et Halieutique, pp. 21–26. Ed. by C. Duby, J. P. Gouet, and F. Laloë. Société Française de Biométrie N°15, Paris. 128 pp.

Petitgas, P., and Lévénez, J. J. 1996. Spatial organisation of pelagic fish: echogram structure, spatio-temporal condition and biomass in senegalese waters. ICES Journal of Marine Science, 53: 147–154.

Petitgas, P., and Samb, B. 1998. On the clustered occurrence of fish schools along acoutic survey transcets and its relation with population abundance. ICES CM 1998/J: 5, 20 pp.

Reid, D. 2000. The relationship of herring school size to seabed structure and local school abundance in the NW North Sea. ICES CM 2000/K: 29, 15 pp.

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

Reid, D., and Simmonds, E. 1993. Image analysis techniques for the study of fish school structure from acoustic survey data. Canadian Journal of Fisheries and Aquatic Sciences, 50: 1264–1272.

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

Schneider, D. 1989. Identifying the spatial scale of density-dependent interaction of predators with schooling fish in the southern Labrador current. Journal of Fish Biology, 35(Supplement A): 109–115.

Soria, M., Petitgas, P., and Bahri, T. 1998. On the size of fish schools and clusters: a spatial analysis of multibeam sonar images in the Mediterranean Sea. ICES CM 1998/J: 8, 12 pp.

Stoyan, D., and Stoyan, H. 1994. Fractals random shapes and point fields. John Wiley and Sons, NY. 389 pp.

Swain, D., and Morin, R. 1996. Relationships between geographic distribution and abundance of American plaice in the southern Gulf of Saint Laurence. Canadian Journal of Fisheries and Aquatic Sciences, 53: 106–119.

Swain, D., and Sinclair, A. 1994. Fish distribution and catchability: what is the appropriate measure of distribution? Canadian Journal of Fisheries and Aquatic Sciences, 51: 1046–1054.

Swartzman, G. 1991. Fish school formation and maintenance: a random encounter model. Ecological Modelling, 56: 63–80.

Swartzman, G., Stuetzle, W., Kulman, K., and Powojowski, M. 1994. Relating the spatial distribution of pollack schools in the Bering Sea to environmental factors. ICES Journal of Marine Science, 51: 481–492.

Weill, A., Scalabrin, C., and Diner, N. 1993. MoviesB: an acoustic detection description software: application to shoal species classification. Aquatic Living Resources, 6: 255–267.