International Council
for the Exploration of the Sea

Fish Capture
Committee
C.M. 1993/B:8

"SCHOOL": A SOFTWARE FOR FISH SCHOOL IDENTIFICATION

by
S. Georgakarakos and G. Paterakis

Institute of Marine Biology of Crete (IMBC)
Department of Hydroacoustics
P.O.Box 2214; GR-71003 Heraklion Crete, Greece

1. INTRODUCTION

The majority of pelagic species that we are able to recognise during the hydroacoustic surveys, occurs in schools during the day. Their echograms showing vertical cross-sections of the aggregations have been studied for possible application to the:

a. Improvement of stock assessment and debiasing techniques (Ertugrul and Smith 1982)
b. Species identification (Azzali 1982) and
c. Behaviour analysis (Gerlotto and Freon 1988)

Especially in a multispecies environment, the most important source of error to the biomass estimation is caused by species identification methods using echogram scrutinising techniques based on the concurrent trawling data.

Therefore, the development of indirect new identification techniques based on acoustic information coming from the shape and strength of the reflected echo signal could be extremely useful in reducing this error and accelerating routine survey work.

The first attempts of this kind of realisation may be separated into two steps:

1. Classification of echograms, where the samples are gathered in "acoustic populations" (Gerlotto and Marchall 1987) aiming not to recognise each species but to split up the integrated values according to a set of parameters, and

2. Identification of the target species where a set of acoustic and non acoustic parameters is extracted from the echo signal in order to apply discriminant statistical methods (Diner et al. 89).

Also some additional tools could improve the extraction of the school descriptors and accelerate the species identification procedures:

One is the use of multibeam techniques that allows both integration of the school energy and in situ TS estimation of single targets bordering on the school.

The second improvement is the system integration of echogram analysis and species identification in one expert system working on-line during routine surveys (Haralabous and Georgakarakos 1993).

2. METHODS

Software development

"SCHOOL" is a software package developed in IMBC that enables the translation of the echo signal into a 2 D image, the recognition of school aggregations and the extraction of quantitative descriptors from each school.

Up to now SCHOOL is able to process data acquired using Biosonics echo Signal Processor (ESP), but also additional routines are under development to make the software compatible with the SIMRAD EK 500 system.

The software is IBM-PC compatible and includes the following routines:

- -SCROLL (echogram visualisation)
- -EDIT (posteriori replaying for chosen pings)
- -FILTER (contiguity and threshold filters)
- -EXTRACT (extraction of school descriptors)
- -STATS (statistical pre-processing)

The extracted parameters are analysed using commercial software (SYSTAT, STATGRAPHICS) or statistical routines developed in IMBC.

Echogram visualisation. The formation of the echogram is based on a set of elements (pixels) with a resolution equal to the pulse duration on the vertical axis and to one ping interval on the horizontal axis. Each element is presented by the sum of voltage square (svs), the ping number (i) and the depth interval in the water column (j).

The area (a) of each element is defined as the product of the horizontal distance (d) between two successive pings in meters and the vertical distance (h) between two successive integration layers:

$$a = d \cdot h$$

In the calculations of the geometrical parameters of the school, the variables d and h act as distance units for x and y axis respectively. Integrated values (svs) from the TVG corrected and calibrated signal are displayed on screen on a logarithmic basis. This logarithmic contrast enhancement was useful in increasing the contrast between elements that have very low or very high energy value and in scaling the data within the range of the display device.

Echogram Replaying. After echogram visualisation both integration and TS data are matched to each other by using ping numbers and depth interval counters.

The software allows the user to change the different settings (contiguity level, threshold energy, bottom recognition) concerning school identification and form a header file containing information gained from the positioning equipment (GPS) and system clock (e.g. latitude, longitude, log, time, date) and the hydroacoustic parameters.

Thresholding and contiguity filters. Three procedures have been developed to perform school recognition. The first procedure applied on the input data matrix is echo integration thresholding in order to cut-off very low biomass concentration. The second procedure is based on an algorithm that detects contiguous elements along the same ping or /and contiguous elements from one ping to the next, in an i x j matrix of pixels. Elements that fulfil this continuity test are considered as belonging to the same aggregation.

The meaning of the third procedure is to distinguish between low and high energy concentration of biomass. Such a procedure would be useful if it were necessary to remove plankton aggregations from the analysis of schools. The software is able to recognise school aggregations from other echoes, define their boundaries and apply geometrical corrections to eliminate the beam pattern effects. The user is able to define in an interactive way the input settings of the program (e.g. threshold of pixel energy , level of horizontal and vertical contiguity and school mean energy threshold).

Data Acquisition

Both real and simulated data have been used in order to test performance and sensitivity of the system. Data obtained from hydroacoustic surveys in Thermaikos Gulf in 1991 and 1992 have been analysed by using both SCHOOL software and spreadsheet tools for comparison and debugging purposes.

Data collection at sea was performed using Biosonics dual beam equipment operated at 120 kHz. The pulse duration was 0.5 to 1.0 msec and integration was carried out over 1m water column per one transmission. Collected data were stored in analog (DAT) and digital (removable hard disks) form.

3. RESULTS AND DISCUSSION

Up to now 3420 schools have been encountered, digitised and analysed with this software. More than 90% of these schools belong to the three most common species in this area:

- Sardine, *Sardina pilchardus*
- Anchovy, *Engraulis encrasicolus*
- Horse Mackerel, *Trachurus mediterraneous*

Selection of school parameters

Special algorithms were developed in order to quantify the different school parameters. More than 30 parameters could be calculated by the SCHOOL software. They may be classified in three groups: morphological, energetic and spatio-temporal parameters.

Morphological parameters contain information related to the size, shape and homogeneity of the aggregation. Energetic parameters are statistical quantities which describe the dispersion pattern of the fish density inside the school and spatio-temporal parameters which locate the school in space and time (Table 1).

Since some of the extracted parameters are highly correlated, Principal Component Analysis (pca) has been used in order to reduce the number of the variables in the data set. Data were at first normalised and submitted to standard pca routines (Statgraphics, STSC ver. 5.22) Over 79% of data variability can be explained by the first three factors.

The first component seems to be representative of the morphological descriptors (ELEM, PERI, RMAX), the second component is a combination of the bathymetric position (BMAX, ALT) and the later describes the energy level of the school (SMEAN, SCV).

Parameters extracted using the software package "SCHOOL" have also been submitted to standard Discriminant Function Analysis (DFA) and to a simulated artificial neural network in order to associate the target species with a set of school descriptors (Haralabous and Georgakarakos 1993). These results are indicative that by using the above descriptors on a neural network we are able to realise species identification.

To obtain this approach both developments are designed to include DLL capabilities, which could facilitate the on line data flow of the school descriptors to the neural network program.

4. REFERENCES

Azzali, M. 1982. Regarding the possibility of relating echo signals features to classes of marine organisms: tests carried out in the north and middle Adriatic Sea. Contrib. 23, Symp. Fish. Acoustics, Bergen, 21-24 June, 1982 .

Ertugrul, M.G. and Smith, B.V. 1982. Multiple scattering effects in fish abundance estimation. Dep. Memo. Electron. Elect. Eng. Univ. Birmingham, (492)

Gerlotto, F. and Freon, P. 1988. School structure observed in some clupeid species. ICES/FAST Work Groups, Ostend, Belgium, 19-22 April 1988.

Gerlotto, F. and Marchall, E. 1987. The concept of Acoustic Populations: its use for analysis the Results of Acoustic Cruises. Internat. Symp. Fish. Acoust., June 22-26, 1987, Seattle, Washington, USA.

Haralabous, J. and Georgakarakos, S. 1993. Fish-School species identification using a neural network. ICES 1993/B:9 Fish Capture Committee.

Diner, N. et al. INES-MOVIES: a new acoustic data acquisition and processing system. ICES CM 1989/B:45, Fish Capture Committee.

Table 1 SCHOOL BASE

No	Descriptors	Full name	From No		Units
1	ID	School ID			-
2	SPE	Species ID / Species Composition			-
		Morphological			
3	ELEM	Number of elements (pixels)/school			-
4	H	Height of school			m
5	L	Length of school			m
6	PING1	N° of the first ping of school			-
7	PING2	N° of the last ping of school			-
8	AREA	Area of school			m^2
9	ELON	Elongation of school	4, 5		-
10	FRA	Fractal Dimension	8, 11		-
11	PERI	Perimeter of school			m
12	RMIN	Minimum Radius of Perimeter			m
13	RMAX	Maximum Radius of Perimeter			m
14	RMEAN	Mean Radius of Perimeter			m
15	RVAR	Variance of Radius of perimeter			-
16	RCV	Coeff.of variation of Radius of perimeter	14,15		-
17	CIRC	Circularity of school	8, 11		-
18	RECT	Rectangularity of school	4, 5, 8		-
		Energetic			
19	SVST	Total SVS of school			V^2
20	SVS	Mean per element SVS of school	3, 19		V^2
21	SMIN	Minimum SVS of school			V^2
22	SMAX	Maximum SVS of school			V^2
23	TSS	Mean TS of single fish			-
24	SCV	Coeff. of variation of SVS	3, 19, 23		-
25	CROWD	Mean Crowding of school	3, 19, 23		-
26	PATCH	Patchiness of school	3, 19, 23		-
27	IOD	Index of Dispersion	3, 19, 23		-
28	K	Clumping Coefficient	3, 19, 23		-
		Spatio-temporal			
29	DATE	Date			-
30	TIME	Time			-
31	AMIN	Minimum Altitude			m
32	AMAX	Maximum Altitude			m
33	AREL	Relative Altitude	4, 31, 34		-
34	BOT	Mean Bottom Depth			m
35	BOT1	Bottom of first ping			m
36	BOT2	Bottom of last ping			m
37	DMIN	Minimum Depth of school			m
38	DMAX	Max Depth of school			m
39	DMEAN	Mean Depth of school			m
40	LON	Longtitude			°
41	LAN	Latitude			°