ICES Journal of Marine Science, 53: 203–208. 1996

Wideband sounder for fish species identification at sea

Manell E. Zakharia, François Magand,
François Hetroit, and Noël Diner

Zakharia M. E., Magand, F., Hetroit, F., and Diner, N. 1996. Wideband sounder for fish species identification at sea. – ICES Journal of Marine Science 53: 203–208.

A brief description is given of both a wideband echo-sounder and data acquisition at sea. Experiments were conducted in the Bay of Biscay at various seasons for four years. Fish species were identified by trawling. Only echoes associated with trawl catches that were monospecific were used in the classification analyses. Species discrimination was based only on the spectral signature of the echoes and did not take into account the characteristics of the school shape. A modelling of the power spectrum of the echo was used to limit the spectral signature to a reduced set of parameters that could be used for classification using a neural network. Thirty-six different monospecific schools, including about 900 echoes, were processed. Three species were considered: sardine (*Sardina pilchardus*), anchovy (*Engraulis encrasicolus*), and horse mackerel (*Trachurus trachurus*). Classification performance had a success rate as high as 75%.

© 1996 International Council for the Exploration of the Sea

Key words: fish species classification, neural network, spectral analysis, wideband sounder.

M. E. Zakharia and F. Magand, F. Hetroit: CPE Lyon, LISA/LASSSO, Laboratoire d'Acoustique, Systèmes, Signaux et SOnar, 25 rue du Plat, 69288 Lyon Cedex 02, France. N. Diner: IFREMER Centre de Brest, BP 70, 29263 Plouzane, France. Correspondence to Zakharia [tel: +33 72 32 50 74, fax: +33 78 37 80 34].

Introduction

Species classification at sea is a great challenge for both fish-stock estimation and commercial fishing. Two different approaches have been used to achieve this goal. The first is based on the characterization of fish school "shape" and is very similar to a "fisherman's approach" when interpreting sounder recordings (Scalabrin *et al.*, 1992). In such a case, classification is based mainly on group behaviour rather than on the acoustical response of fish aggregations. It becomes less effective when dealing with several species mixed in the same school. The approach developed in this paper is a complementary one based on echo analysis of a single ping. It makes use of a new generation of wideband sounders and is based on the investigation of the variation of individual fish echoes with frequency. Such a "frequency signature" can be used to set up a species classification analysis.

A prototype sounder was designed and constructed for this purpose. Several experiments were conducted in the Bay of Biscay where echoes of schooling fish were collected over a period of 4 years (1991–1994). During the acoustic surveys, trawling was performed for fish species identification. Only echoes associated with trawl catches that were monospecific were selected for the classification analyses.

The classification only uses the spectral signature of the echoes. No information on the echo energy or on the geometrical characteristics of school is used. It is part of a global approach for species classification that combines both detailed analyses (spectral signature) and global analyses (school shape).

Wideband sounder description

A prototype wideband sounder, operated on a frequency range of 2 octaves (20 kHz to 80 kHz) was specially designed and constructed for this study (Zakharia *et al.*, 1989).

Transducer and power amplifiers

The transducer is composed of a square matrix of 100 identical elements, wired in five concentric rings labelled T1 to T5 (Fig. 1). Each ring is associated with an independent transmitting and receiving channel with a power amplifier (1 kW/ring), a transmitting–receiving switch and a preamplifier. As the system is a fully wideband one (and not a multiple frequency one), no

1054–3139/96/020203+06 $18.00/0
© 1996 International Council for the Exploration of the Sea

Downloaded from http://icesjms.oxfordjournals.org/ at Ifremer, BibliothÃ¨queLa PÃ©rouse on November 25, 2015

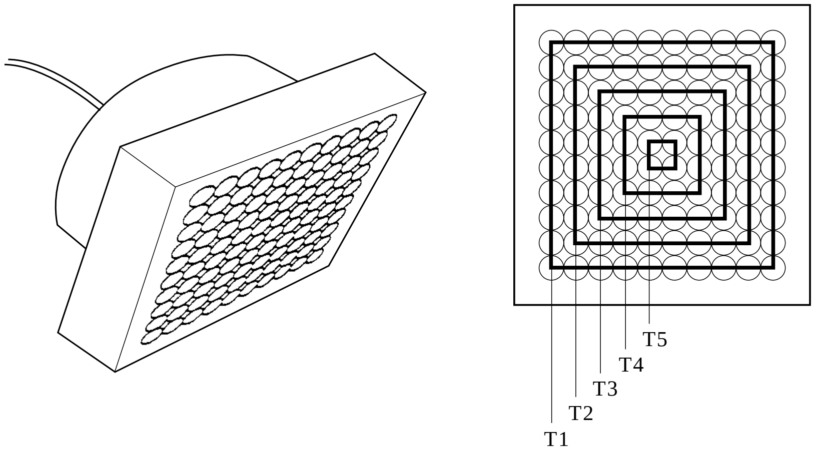

Figure 1. Transducer element arrangements.

Table 1. Main characteristics of the wideband sounder.

Frequency range	Bandwidth B (kHz)	Rings used	Beam aperture (− 3 dB)	Peak level re. 1 μPa at 1 m
Low	20–40	T1 to T5	10° to 5°	214 dB
Medium	40–60	T3-T4-T5	9° to 6°	215 dB
High	60–80	T4-T5	9° to 7°	207 dB

impedance matching (resonant device) was used. The whole system has been designed to be as flexible as possible. In particular, several signals can be computed, synthesized, and transmitted independently. Frequency weighting can then be applied either to the transmitters and/or to the receivers to maintain the directivity pattern and the transmitted level constant over the whole frequency range (Lardies and Guilhot 1987; Bel *et al.*, 1992). This fine improvement is still under development and, as a first step, beam width control is presently achieved by ring switching: the number of rings used depends on the frequency range of the transmitted signals, as shown in Table 1.

The available bandwidth has thus been split into three sub-bands as shown in Table 1 (low-, medium- and high-frequency ranges). Aside from reducing beam width variations, this splitting helps in reducing the complexity of digital signal processing by producing a limited bandwidth of 20 kHz. A modulation bandwidth of 20 kHz is then used with various carrier frequencies (30 kHz, 50 kHz, and 70 kHz).

Nine signals have been computed, stored, and made available for transmission in the standard operation mode: three different values of the duration (1 ms, 5 ms, and 10 ms) for each frequency range (20–40 kHz, 40–60 kHz, and 60–80 kHz).

Transmitted signals

The transmitted signals are chirps with linear period modulation and ‘‘bell-shaped’’ envelopes, as shown in Figure 2a and b (top traces). The transmitted waveforms have been selected for several reasons, mainly the smooth behaviour of their envelope in both the time and frequency domains, i.e. no side lobes (Altes, 1976).

Digital processing is achieved in the baseband after frequency shifting. The design of signals, both carrier frequency and bandwidth, results from a compromise between two major constraints: processing in real-time and reduction of beam width variations.

On-board processing

Echo processing is carried out at both the analogue and digital level: analogue (nevertheless digitally controlled) processing is separate for each frequency range and digital processing is common to all frequency ranges.

The analogue processing is, in fact, a digitally controlled hybrid technique, consisting of bandpass filters (in the frequency ranges indicated in Table 1) and digitally controlled time varied gains (TVG: 10, 20, 30, or 40 log $r+2\alpha r$). The absorption α is computed for the central frequency of each band and assumed to be constant over each frequency sub-range. The maximum TVG error at

Downloaded from http://icesjms.oxfordjournals.org/ at Ifremer, BibliothÃ¨queLa PÃ©rouse on November 25, 2015

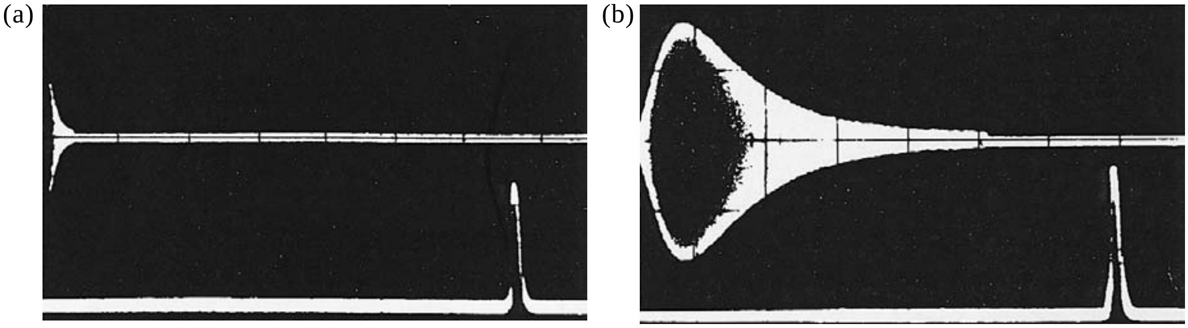

Figure 2. Real-time pulse compression, oscillogram examples. Top traces = transmitted signal, bottom traces = processing output, horizontal scale: 2 ms/division. (a) Short low-frequency signal (B = 20–40 kHz, T = 1 ms), (b) long high-frequency signal (B = 60–80 kHz, T = 10 ms).

100 m is less than 0.5 dB. The TVG is followed by a frequency shifting (corresponding to the carrier frequency). Data are then digitally filtered in the baseband: 0 to 20 kHz.

As a wideband approach is considered and wideband receivers are used, noise level is increased and so matched filtering is needed to increase the signal-to-noise ratio (Van Trees, 1971). When using transmitted chirps, the duration of the filter output is proportional to the inverse of the bandwidth and its magnitude is proportional to the transmitted signal duration (i.e. the longer the signal duration, the higher the signal-to-noise ratio at the output of the matched filter). Matched filtering is thus well suited for noise reduction without degrading spatial resolution.

This filtering is achieved digitally, in real-time, using special purpose hardware. Figure 2a and b (bottom traces) shows some examples of this real-time processing, the envelope of the output of the matched filter. Thanks to pulse compression (Fig. 2a, b) to the particular design of the transmitted chirps and to the frequency range division between appropriate sub-bands, the output of the matched filter does not depend on the selected signal. Figure 2 clearly illustrates this independence.

After pulse compression, the time resolution is proportional only to the inverse of the bandwidth, so the echogram resolution will be the same for all the transmitted signals after processing: about 7.5 cm (at −6 dB). This is meant to facilitate data interpretation by an operator without in-depth knowledge of signal processing.

Storage

Signals were stored during the experiments using either analogue recording (analogue tapes were replayed at the laboratory, digitized, and stored) or on-board digital storage.

In order to display and store the sounder echoes in real-time (at sea or from playbacks) special software was developed: NEPTUNE. This software uses a standard PC associated with commercial digitizing board (multiple channels).

While the raw echoes (sampled at 400 kHz) are digitized and stored on a hard disk, the echo envelope is computed and displayed in false colours on a scrolling screen (as in standard digital sounders). Both global view and fine details are thus available in parallel.

In the digital playback mode, each ping can be extracted, gated, displayed (digital oscilloscope mode), and stored. These individual echoes can then be processed separately.

Sea experiments

The wideband transducer and its associated electronics were installed on a tow fish and were linked to the processing and storage unit by a 50 m cable. A delta-shaped towed body was used (maximum towing speed: 3 knots). Wideband data on schooling fish were recorded mainly during the day time, in the Bay of Biscay, during several cruises (on RV ''Thalassa'') at different seasons over a period of 4 years (1991–1994). Both wideband and narrow band (38 kHz) echo-sounders were run simultaneously, using alternate trigger pulses. The medium frequency range was most often used, since it corresponds to the highest acoustic level. The time duration of the transmitted signal was 1 ms or 5 ms, depending on the bottom depth (as more energy is needed for deeper water).

Trawl monitoring (as well as a netsonde) was used in order to ensure matching between acoustical and trawling data. In the first experiments trawling was carried out just after passing a school and matching was achieved through the hull-mounted narrow-band sounder. In the later ones, wideband sounding and trawling were performed simultaneously.

After selecting the echoes corresponding to monospecific catches, only three different species remained available for setting up the echo database: sardine

Downloaded from http://icesjms.oxfordjournals.org/ at Ifremer, BibliothÃ¨queLa PÃ©rouse on November 25, 2015

Table 2. Details of signal database obtained from sea experiments (only school echoes corresponding to monospecific catches have been selected).

	Fish species	No. of schools	No. of pings
Class 1	Sardine	15	270
Class 2	Anchovy	10	154
Class 2	Horse mackerel	21	465

(*Sardina pilchardus*), anchovy (*Engraulis encrasicolus*), and horse mackerel (*Trachurus trachurus*).

A detailed description of the data collected (number of schools, number of ping for each species) is given in Table 2.

Spectral Signature

The selected portion of every ping was considered as a single realization associated with the corresponding fish species. Modelling techniques were then applied to the power spectrum of the echoes in order to reduce the information contained in each echo to a set of parameters that should be able to describe the ''resonant'' behaviour of the spectrum that was commonly observed.

Such a parametric analysis is widely used in signal processing problems where resonance phenomena occur, such as speech recognition or data transmission, as it can lead to a complete description of the power spectrum using only a reduced set of relevant spectral parameters, the so-called AR (Auto-Regressive) coefficients (Kay, 1987).

This numerical signal model deals with the discrete time sequence $S(kT_e)$ ($k=0 \ldots N$), obtained by sampling the continuous time signal S(t) at a sampling period T_e. The model assumes that any sample S(k) can be estimated by a linear combination of the p previous samples and a realization of a random process N(k), whose standard deviation is σ. The time sequence can be modelled by the following relation (Kay, 1987):

$$S(k) = -\sum_{l=1}^{p} a_l S(k-l) + N(k)$$

The coefficients a_l are the AR coefficients associated with the time sequence S(k) and p is the order of the model. The parameter extraction methods and the determination of the optimal order p have been widely discussed in the signal-processing literature (Kay, 1987).

This time modelling can, of course, be associated with a spectral model. The modelled power spectrum can then be expressed as follows:

$$|\hat{s}(\omega)|^2 = \frac{\sigma^2}{\left|\prod_{l=1}^{p}(1-p_l \exp(-i\omega t))\right|^2}$$

where the $(p_k)_{k=1}, \ldots,$ p coefficients are the roots of the following complex polynomial:

$$1 + \sum_{l=1}^{p} a_l z^l = 0$$

$\omega = 2\pi f$ is the angular frequency (f: frequency) and a_l are the AR coefficients.

The expression of $|\hat{s}(\omega)|^2$ clearly shows that the AR modelling corresponds, in fact, to modelling the resonance existing in the power spectrum.

When few resonances are considered, the model order (p) is commonly chosen as twice the number of resonances. However, this rule of thumb cannot be applied when a large number of resonances have to be considered. Objective criteria must then be used to optimize the order (Kay, 1987).

During a previous study on size discrimination using spectral signatures (Degoul *et al.*, 1989; Magand and Zakharia, 1992), it was found that the classification performance was improved by considering a set of parameters obtained from the AR parameters, the so-called cepstral coefficients (Gray and Markel, 1976). These parameters (commonly used in speech recognition) can be used to compute a ''spectral distance'' between two signals. The cepstral parameters c_l can be estimated from the AR parameters a_l in a recursive manner using the following relations:

$$c_0 = \ln(\sigma^2)$$

$$c_1 = a_1$$

$$c_l = -a_l - \sum_{j=1}^{l-1} \frac{l-j}{l} c_{l-j} a_j \quad \text{for } l = 2 \ldots p$$

$$c_l = -\sum_{j=1}^{p} \frac{l-j}{l} c_{l-j} a_j \quad \text{for } l > p$$

As the values of the c_l coefficients decrease rapidly as n tends to infinity, the choice of cepstral coefficient order q is less critical than in the AR case. Both AR and cepstral coefficients can be used as input vectors for the classification.

Classification and performance

Classification method

The aim of the classification process is to discriminate fish species using the spectral signature provided by the AR and/or cepstral coefficients. This task has been performed using a supervised neural network (multilayer neural network) (Wasserman, 1989) composed of several neuronal layers interconnected from the input to the output so that the ''state information'' can propagate forward only from one layer to the next. The number of neurones in the input layer is the dimension of input vector and the number of neurones in the output layer is the number of fish species.

Downloaded from http://icesjms.oxfordjournals.org/ at Ifremer, BibliothÃ¨queLa PÃ©rouse on November 25, 2015

Table 3. Confusion matrix (diagonals in bold correspond to correct classification) for different sizes of the learning set (A: 154 echoes, B: 70 echoes).

	Estimated species					
	Configuration A			Configuration B		
Species	Sardine	Anchovy	Horse mackerel	Sardine	Anchovy	Horse mackerel
Sardine	**73%**	12%	15%	**59%**	17%	24%
Anchovy	19%	**64%**	17%	33%	**44%**	23%
Horse mackerel	18%	8%	**74%**	28%	13%	**59%**

The classification process is split into two parts: training phase and test (or generalization) phase. During the training phase, the network optimizes its internal parameters (synaptic weights) using known input and output vectors (i.e. spectral parameters vector and its associated fish species). The training data set is obtained by randomly selecting echoes from among the total data set. The generalization phase consists in providing the network with the spectral signature vector of all the echoes and estimates the species by choosing the output neurone with the highest ‘‘activation level’’.

The number of intermediate layers (so-called ‘‘hidden’’ layers) and the number of neurones per layer result from a compromise between performance, complexity, and flexibility: on the one hand, if those numbers are too high, the network will require too many training examples and the generalization will be poor; on the other hand, if those numbers are too small, the network will not be able to learn the training examples. After several preliminary trials on both the network structure and the AR model, the following compromise was reached: 10 AR coefficients, 30 cepstral coefficients, 2 hidden layers.

As the number of available echoes is different for each species (Table 2), several sizes for the training set were considered for network training (configuration A and B).

The classification performance is presented in Table 3 as a confusion matrix for both configurations. The columns are the species for input echoes and the rows the species provided by the network. The diagonal values of each matrix give the average percentages of correct classifications. The other values are the percentages of wrong classifications (or confusion).

Configuration A

All anchovy echoes have been included in the learning set (154 echoes), and the same number of learning patterns have been used for the other two classes. It turns out that species discrimination is quite good, since over 70% of success has been obtained for sardine and horse mackerel. In this case, the anchovy discrimination rate is less significant, since no new data are available for network testing.

Configuration B

Only 45% of anchovy echoes have been included in the learning set (70 echoes). Anchovy recognition is then far less efficient and the confusion with the other species increases. Results, however, are quite good for sardine and horse mackerel, although few data were used to train the network (70 echoes: 26% of sardine echoes and 15% of horse mackerel echoes).

Conclusion and perspectives

The results of the classification clearly show the feasibility of fish-species classification at sea using the spectral signature of the echoes over a wide frequency range. As for every classification task, the learning phase, the quality and the abundance of the learning data are key issues. Our effort is now directed toward the continuation of building up a large data bank of wideband echoes for a great variety of experimental conditions (season, time, place, species, . . .).

Several improvements are under development or are already achieved both for the sounder design and the trawling aspects. For the sounder, systematic digitizing during surveys, and automatic school searching in the digital playback mode are planned as well as accurate control of the directivity pattern over the whole frequency range in order to reduce the influence of the fish position on its spectral signature.

Trawling performance will be improved by the use of a multi-compartment trawl that separates several fish schools trawled sequentially. This will reduce the number of trawl operations and will effectively allow ‘‘continuous’’ trawling.

Finally, it will be worth investigating the combination of both the spatial (school shape) and the spectral classification methods in order to take advantage of their complementary advantages.

Downloaded from http://icesjms.oxfordjournals.org/ at Ifremer, BibliothÃ¨queLa PÃ©rouse on November 25, 2015

Acknowledgements

This work was supported by the Commission of the European Communities, FAR (fisheries and aquaculture) programme, BIOMASS project.

References

Altes, R. A. 1976. Sonar for generalized target description and its similarity to animal echo location. Journal of the Acoustical Society of America, 59: 97–105.

Bel, R., Zakharia, M. E., and Blazère, O. 1992. Sonar arrays with constant directivity over a wide frequency range, pp. 725–728. 1st European Conference on Underwater Acoustics. Elsevier Applied Science, London.

Degoul, P., Flandrin, P., Gache, N., and Zakharia, M. 1989. Fish echoes classification via auto-regressive modelling, pp. 161–168. Proceedings of the Institute of Acoustics, Volume 11, Loughborough, UK.

Gray, A. H. and Markel, J. D. 1976. Distance measures for speech processing. IEEE Trans. on Acoustics, Speech and Signal Processing, 24: 380–391.

Kay, S. M. 1987. Modern spectral estimation: theory and applications. Prentice Hall Inc., Englewood Cliffs, N.J. 543 pp.

Lardies, J. and Guilhot, J. P. 1987. Realisation of a broadband constant beamwidth end-fire line array. Acoustic Letters, 10: 122–127.

Magand, F. and Zakharia, M. E. 1992. Fish echo classification using an a-priori physical model: advantages and limitations, pp. 133–136. 1st European Conference on Underwater Acoustics. Elsevier Applied Science, London.

Scalabrin, C., Weill, A., and Diner, N. 1992. The structure of multidimensional data from acoustic detection of fish schools, pp. 141–146. 1st European Conference on Underwater Acoustics. Elsevier Applied Science, London.

Van Trees, H. L. 1971. Detection, estimation and modulation theory. Part III: Radar-sonar signal processing and gaussian signal in noise, pp. 244–247. John Wiley and Sons Inc. 626 pp.

Wasserman, P. D. 1989. Neural computing, theory and practice, pp. 17–59. Van Nostrand Reinhold, New York. 230 pp.

Zakharia, M., Corgiatti, J. P., Joly, F., and Person, R. 1989. Wide-band sounder for fisheries. Proceedings of the Institute of Acoustics, 11: 274–281.

Downloaded from http://icesjms.oxfordjournals.org/ at Ifremer, BibliothÃ¨queLa PÃ©rouse on November 25, 2015