DEVELOPMENT OF AN OPEN SOURCE SOFTWARE TO TRANSFORM ACOUSTIC DATA FOR CLASSIFICATION OF BENTHIC HABITATS

NOELA SÁNCHEZ-CARNERO[1], CRISTINA BERNÁRDEZ[1], MAREK MOSZYŃSKI[2], JUAN FREIRE[1]

[1]Grupo de Recursos Marinos y Pesquerías
Universidade da Coruña
Campus da Zapateira s/n, E-15071 A Coruña, Spain
noesanchez@udc.es, cbernardez@udc.es, jfreire@udc.es
[2]Gdansk University of Technology
Narutowicza 11/12, 80-952 Gdańsk, Poland
marmo@eti.pg.gda.pl

Knowledge and detailed mapping of benthic habitats is a key tool for the development of strategies for coastal ecosystem management. However, existing methods are used as a "black box" where the user cannot know the equations and criteria used in the acoustic corrections, nor inspect and modify the statistical and computational protocols used too.

The RMyP is developing an open protocol for marine habitat classification, from the field techniques and sampling details to large data manipulation techniques and statistical data analyses. Several echosounders will be operated from small boats, with a resulting low cost. Computing and statistical methods will be made open and free to allow for their posterior use, modification and improvement by other researchers.

A fundamental point in the design of an open protocol is the independence from commercial software. Thus, a software is being built which uses the raw data from the echosounder as an input, generating energy matrixes with the corrected values (transmission loss and depth adjustment) as an output, for their further statistical analysis.

INTRODUCTION

Knowledge of benthic habitats plays a key role in the development of strategies for sustainable exploitation of marine resources (Botsford et al. 1997, Vallega 1999). Acoustic surveys have been widely used in seabed classification studies (Brown et al. 2002, Caddell 1998, Greenstreet et al. 1997, Legendre et al. 2002; Tuck et al. 1998), but a unified and standardized methodology has not been developed. Our ultimate goal is to develop an open protocol, using a combination of acoustic tools, statistical data analysis and groundtruthing for the identification, description, classification and mapping of marine benthic habitats.

Nowadays specifically designed commercial software is available for the classification of benthic habitats (QTC, RoxAnn, etc.). Some software desgined for fishery resources assessment implements also specific modules for bottom classification (Biosonics, Echoview/Sonardata). A review of these programs can be found in Hamilton, L.J. in 2001.

RoxAnn and QTC are the most used packages, but both products regard their procedures and algorithms as proprietary. In this way, two problems arise: the software and accompanying information give out little information about the procedures applied to raw acoustic to identify bottom types (this information is also protected by the companies' patents). (Clarke, P.A. & Hamilton, L.J.), and as a consequence there is no possibility to modify, adapt or improve the procedures.

At present, users must select one of these solutions and build their classification in a partially blind process, following non-transparent criteria according to the chosen software. Our aim is to be able to know, understand and control every data transformation, as well as the criteria used in the capture of the data, making it possible to assess the performance of each step taken towards the final objective, and to be able to modify any of these criteria to improve the accuracy and efficiency of our protocol.

There are two main types of acoustical bottom classification systems, one using multiple echo energy methods, and the other using a first echo shape approach. (L. J. Hamilton, Acoustic seabed classification systems).

The aim of this project is the production of a software of the first group (ECOSONS 1.0), using the first echo of the 38 and 200 kHz frequencies and the second echo of the 38 kHz frequency signal. The protocol will allow the user to understand and decide upon the parameters being used to capture and transform the data. As an open source software, it will as well be open for customization, modification and improvement by other researchers.

The ECOSONS package has been developed using Perl language. The administrator version can work on Windows or Linux platforms, allowing the user to access through the Internet or an intranet.

1. METHODOLOGY

<u>Field work</u>

Three acoustic surveys were carried out in the NW Spain and Mexican Pacific coast. The three areas are of a hign fishing interest, and heterogeneous regarding habitat types (sandy, rocky, seaweed bottoms, etc). Surveys were performed using a portable EA400P echo-sounder (Simrad, Norway). This is an hydrographic echo-sounder working simultaneously at 38 and 200 kHz. During the surveys, parameters were adjusted to enhance performance for future data analyses. Small fishing boats were used, having the transducer mounted on the side using a portable device. The surveys consisted of a series of transect grids, mostly perpendicular to the coastline, from depth 15 to 150 m. Distance between contiguous transects was set according to maximum depth. Positions were recorded into the sounder files using a GPS signal input.

Acoustic data were recorded in raw files that contained information on the echo-sounder settings, position and backscatter energy values of each ping.

Postprocessing

Scientific echosounders continuously record backscatter signals from each ping sent by the transducer. These signals are discretized by the echosounder with no possible intervention of the user in this process, integrating the energy values at certain known intervals (constant or variable, depending on the echo-sounder). Each energy value represents a portion of the sampled water column.

These discrete data are stored in files (raw files) along with the configuration settings used in the recording (pulse length, power, etc.) for each ping. Raw files are not the same but similar in a number of commercial echo-sounders, and thus a detailed knowledge of their structure is fundamental to both avoid dependence on commercial software as an intermediate step and to build a flexible methodology, transferable to different types of echo-sounders.

In the EA400P raw file, the energy values are originally compressed to allow easy storage of data. As a consequence, the real energy value EL (dB relative to 1W) is obtained by decompressing the EL_c value obtained from the *raw* file. The decompression process the proposed software performs is further described in another section in this paper (see Data transformation and generation of output files). The EA400P echosounder produces a second type of file, *out*, that stores GPS position and bottom depth for each ping.

ECOSONS user must start by creating a “study” and including in it every surveyed surveyed transect. For each of these transects the corresponding *raw* and *out* files can then be uploaded for processing. Once the transect library is created, the software extracts the heading information stored in the .out files. The second step is the decompressing of the raw energy values. Using the raw file(s) corresponding to each transect as an input, the software extracts the energy values that need to be corrected to obtain energy values depth independent, in a similar way as other commercial software for viewing, quantifying or editing echo-sounder data.

As a final step, it links geographic locations to each ping and selects the energy values corresponding to the bottom. This is done by reading the *out* file and using the bottom depth value of each ping to select and save the energy values corresponding to the bottom in the first and following reflections (first and second bottoms). Since the user sets the cutting parameters (linear surface extension and depth, both in meters), the number of bottom values selected depends on the interval between consecutive energy values. As a result, a matrix containing energy values for the first and second bottom at 38 kHz and for the first bottom at 200 kHz of each ping (as well as the GPS position obtained from the .out file) is generated and stored in a MySQL database accessible through the Internet. This matrix can be exported to any statistical analysis package such as R.

Complementarily, the software extracts two more data bases, the first one being a report containing the settings of the survey and any complementary data that the user may ask for (gain, equivalent beam angle, beam width along ship, beam width athwart ship, transmit power, pulse length, band width, sample interval, absorption coefficient and G.G.M. The second one is a data base of bathymetry containing georeferenced bottom depth values, allowing the user to easily create a bathymetric cartography using any GIS.

It is important to note that the process can be restarted at any intermediate point where the user can decide on the settings, to allow flexible testing of the performance of the software at any stage.

Data transformation and generation of output files

A single *out* file can contain data referring to several .raw files so the first step in the software performance is to identify the associations among these files.

In a first step the .out file is processed resulting in a table where the GPS values for each ping are stored (after linear interpolation from initial records), in addition to the bottom depth value for both frequencies and the second bottom value for 38 kHz.

Subsequently, the raw file is processed resulting in a table with the EL values (after decompression) corresponding to the customized section of the bottom (for each frequency) and its reflections (for 38 kHz frequency).

The EL values stored in the raw files are scaled using the brand criteria consisting in:

$$EL = EL_c \frac{10\log(2)}{256} \qquad (1)$$

These values of EL are dB/1W, so they are power values (J/s).

Once the matrix with the real **EL** values is obtained, and since the final statistical analysis will analyze the data with no regards to the effects of depth and travelling characteristics of the wave, it is necessary to apply a series of algorithms to the energy values, with the aim of standardizing to correct the following effects: spreading, absorption and depth effect (time adjustment and power adjustment).

To eliminate these effects the following equation is used:

$$EL_f = EL + 30\log r + 2\alpha r \qquad (2)$$

Where $\mathbf{El_f}$ is the energy value corrected for the depth effect, $\mathbf{r_0}$ is the reference depth and α is the absorption (Equation Ainslie and McColm,1998).

2. STATISTICAL ANALYSIS OF ACOUSTICAL DATA AND GEOGRAPHICAL COHERENCE ALGORITHM

The energy value matrix will be used as direct input for statistical analysis to identify, describe and classify benthic habitat types. Each one of the intervals selected from the raw file would be treated as a variable. The first step in the statistical procedure will consist of a reduction of the dimensionality of the matrix using principal component analysis. Once the new variables, principal components, are defined, grouping of acoustically similar observations will be performed using non-hierarchical cluster analysis. Data georeferencing will allow for rapid mapping of the observations, already classified, and to analyze their geographic coherence. Statistical algorithms will be implemented to restrict the groupings obtained by the cluster algorithm to be spatially contiguous.

3. DATA VALIDATION

The resulting cluster will be validated to corroborate their correspondence with distinct habitat types. For this purpose, observations will be carried out using several methods such as sediment analysis (granulometry), direct observations using remotely operated vehicles equipped with video and photographic systems and SCUBA diving transects.

4. APPLICATION

The characteristics of each step of the procedure presented here allow the application of the same protocol, software and data analysis methods to other projects using different echosounders. Both a detailed description of the field, postprocessing and statistical procedures and the original software will be available in publications and the Internet for easy use by other users.

ACKNOWLEGMENTS

We would especially like to thank EnLaRed SL for their help and patience during the long process of development of the corrections and its integration in the software. Helge Balk for his encouraging help and for providing helpful tools for the understanding of raw files and Alicia Pallas who helped in the English translation of this manuscript..

Financial support for this study was given by the research grants ECOSOST (REN2003-04816) supported by the Spanish Ministerio de Ciencia y Tecnología and CONNECT (CTM2006-09043/MAR) supported by the Spanish Ministerio de Educación y Ciencia.

REFERENCES

[1] L.W. Botsford, J.C. Castilla, CH. Peterson, The management of fisheries and marine ecosystems, Science 277:509-515, 1997.

[2] A. Vallega, Fundamentals of integrated coastal management, Dordrecht: Kluwer Academic Publishers,1999.

[3] C.J. Brown, K.M. Cooper, W.J. Meadows, D.S. Limpenny, H.L. Rees, Small-scale mapping of sea-bed assemblages in the eastern English Channel using sidescan sonar and remote sampling techniques, Estuar. Coast. Shef Sci. 54:263-278, 2002.

[4] S.E. Caddell, Application of an acoustic sea floor classification system for benthic habitat assessment, J. Shell. Res. 17:1459-1461, 1998.

[5] S. Greenstreet, I. Tuck, G. Grewar, E. Armstrong, D.G. Reid, P.J. Wright, An assessment of the acoustic survey technique, RoxAnn, as a means of mapping seabed habitat. ICES J Mar Sci 54:939-959, 1997.

[6] P. Legendre, K.E. Ellingsen, E. Bjørnbom, P. Casgrain, Acoustic seabed classification: improved statistical method, Can. J. Fish. Aquat. Sci. 59: 1085–1089, 2002.

[7] I. Tuck, S.J. Hall, M.R. Robertson, E. Armstrong, D.J. Basford, Effects of physical trawling distubance in a previously unfished sheltered Scottish sea loch, Mar Ecol Prog Ser 162:227-242, 1998.

[8] L.J. Hamilton, Acoustic Seabed Classification Systems, DSTO Aeronautical and Maritime Research Laboratory, 2001.

[9] P.A Clarke, L.J. Hamilton, The ABCS Program for the Analysis of Echo Sounder Returns for Acoustic Bottom Classification, DSTO Aeronautical and Maritime Research Laboratory, 1999.

[10] M.A. Ainslie, J. G. McColm, A simplified formula for viscous and chemical absorption in sea water, Journal of the Acoustic Society of America, 103 (3): 1971-1672, 1998.