



**Ifremer**

Ifremer/RBE/EMH

Mathieu Doray

....  
12 14 – R.INT RBE/EMH 2014-XX

---

# WGACEGG2014 EchoR training course

## Introduction

[EchoR](#) is a suite of R codes aiming at:

- handling pre-processed fisheries acoustics data collected during sea surveys;
- computing standard ec(h)osystemic indicators based on those fisheries acoustics data. These indicators include:
  - biomass estimates per fish species and elementary sampling distance units (ESDU) ;
  - biomass-at-length estimates per fish species and ESDU ;
  - biomass-at-age estimates per fish species and ESDU;
  - biomass estimates per fish species and post-stratification regions;
  - synthetic spatial (Woillez et al. 2007) and population dynamics indicators can also be computed based on per ESDU data;
  - Actually, EchoR comprises plotting routines that allow to produce the standard grid maps used to exchange data within the WGACEGG working group.

Methods for acoustic fish biomass assessment implemented in EchoR are described in Simmonds and MacLennan (2005) and Doray et al. (2010).

Package sources and support can be found at: <https://forge.ifremer.fr/plugins/mediawiki/wiki/echor/index.php/Accueil>

The objectives of the WGACEGG2014 EchoR training were:

- To provide an overview of the methods implemented in the package to all users, and discuss everyone practice, based on the EchoR framework;
- to present the new package new features to more experienced users. The new features are essentially related to biomass estimation per age and post-stratification region;
- to demonstrate the package capabilities with a demo dataset included in the package, and eventually with users data.

## Training highlights

A total of 14 people from France, Spain, Portugal, Italy, Greece, Croatia and Turkey took the training, from 09:00 to 19:00 on 16/11/2014. Some trainees from Portugal, Basque Country and Greece already use EchoR in their institution. The other have discovered the software.

In IPMA in Portugal, the anchovy and sardine biomass estimations per ESDU computed with EchoR were almost identical to those computed with previous homemade dedicated softwares.

The presentation of EchoR built-in methods raised interesting issues on how fisheries acoustic data should be processed to derive fish biomass estimates. The main questions (and answers) are summarized below.

## Why bothering with echotypes?

Echotype is like prose, all fisheries acoustician use it, sometimes unconsciously.

An echotype can be defined as a recurrent pattern that can be observed on the echograms of a given survey, and be associated with a species, or a group of species. An echotype can be as general as “all the echoes that are supposed to be produced by fish in my survey” or as specific as “the typical school shape of a given species and/or size or age class”, which I am sure to be able to separate from other echoes. Echotypes must hence be defined prior to scrutinizing. They simply represent the different categories to which fish (or other species of interest) NASC will be classified into during the scrutinizing process. The echotype definition in EchoR is quite flexible: an echotype is defined as the combination of:

- the name(s) of the species it refers to,
- an optionnal size categories (one or several),
- a depth stratum, where the echotype is supposed to always appear.

The depth strata classically corresponds to the depth layers that are well sampled by the survey fishing gears (Figure 1). In Figure 1, the surface echotype will be exclusively associated to surface trawl hauls, if it is associated to the “surface” depth stratum. The other echotypes will be exclusively associated to near-bottom trawl hauls, that can possibly sample the species they comprise. They will be defined as belonging to the “near bottom” depth stratum.

If one can just extract fish NASC from other NASC (plankton etc...) while scrutinizing a survey, one can simply define an echotype that comprises all fish species in the area. If trawl hauls are only performed efficiently near the bottom, this global fish echotype will be defined only in the area near the bottom that is well sampled. Fish NASC (and biomass) will in this case be split between species solely based on the trawl catches composition.

On the other hand, if some fish species form schools that have so specific features that allows to separate them from other echotraces while scrutinizing, it is advisable to create a specific, single species echotype to treat the specific school of the species independantly. This is especially true if the schools of the recognizable species are dense (such as sardine for example). Creating a specific echotype for the species will reduce the risk of allocating some NASC of the dense, well recognizable species schools to other species, which form less dense and less recognizable schools (cf. echotype 4 in Figure 1). This bias for instance happens when the trawl catch composition does not perfectly represent the real species composition in the schools, whereas all species belong to a single echotype,.

In EchoR, fish biomass estimates are first computed by echotype, and therefore split between the species comprising the echotypes. All the target species must then appear in the echotypes. If they do not, they will not be assessed.

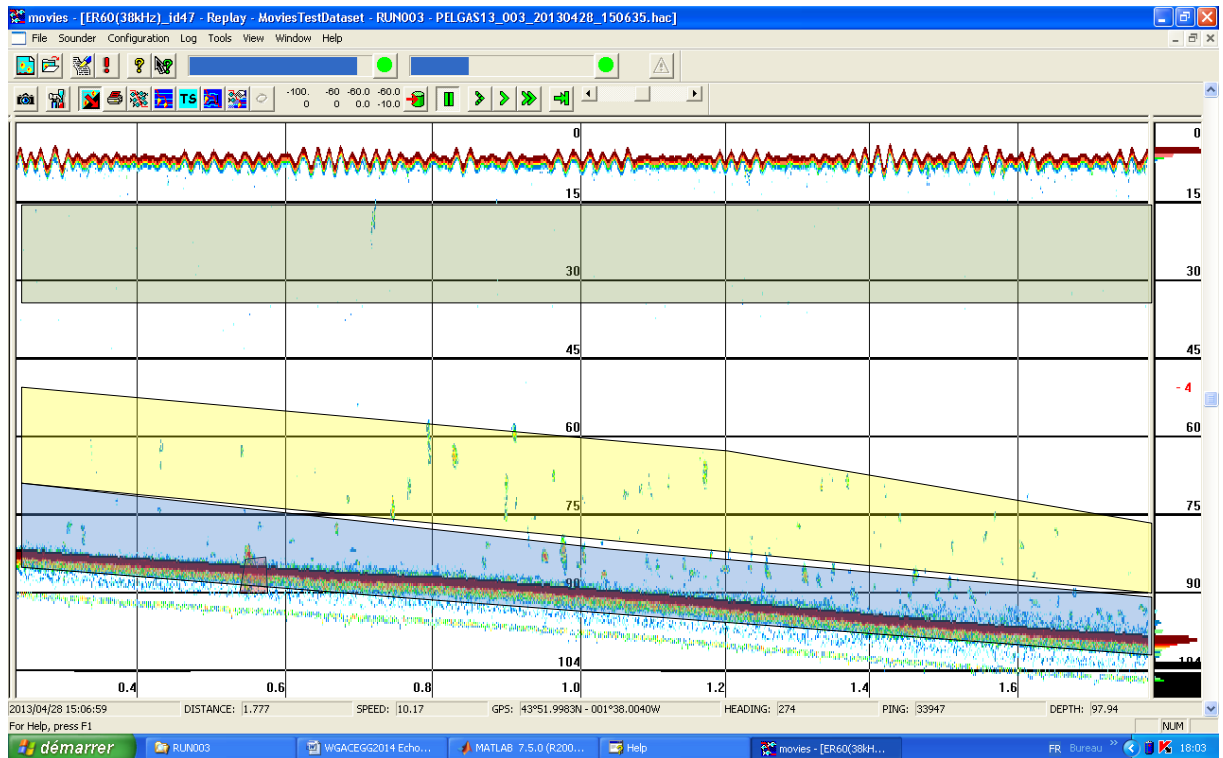


Figure 1: echotypes and depth stratum examples. A total of 4 different echotypes have been used to scrutinize this echogram: i) a surface echogram (green polygon) that comprises all the school echotraces and species observed between 10 and 30 m on the echogram; ii) an echotype comprising all the schools at least than 10 m above the bottom (blue polygon), iii) an echotype comprising the schools located at more than 10 m above the bottom (yellow polygon), and iv) an echotype comprising only the species that produces the dense school that has been isolated from the rest in the red polygon.

## How to handle rare species in EchoR?

A subject related to the definition of echotypes is the handling of rare species. Echotypes must comprise the species of interest, but what to do with the rare species that appear from time to time in the catches?

One cannot simply remove them from the catches, as they were present in the sea and hence have contributed to the fish NASC that was recorded.

One option consists in including all the species found in the catches in the echotype definition, even the rare ones. Every species will in this case be assessed. This method has the merit of simplicity, except perhaps during the echotype definition process, if the species list is very long. It has however several drawbacks:

- the processing time might increase significantly, as all species will be processed;
- the results files and plots are more complicated to interpret, as all species are included;
- the procedures that produces results per size and age class will probably not work, as length and/or age distributions will be missing for rare species;
- the absence of specific TS for the rare species might bias some results, especially if some rare individuals are very big.

A procedure is available in EchoR to mitigate some of these drawbacks: the “complementation”.

The idea is to pool all rare (or non-target species) into a fake species called “complement” (species code: “COMP-LEM”). This fake species will be given constant mean length, weight and TS parameter to ensure that the TS values used to compute its fake biomass will not be extreme. Only the rare species total weight and number per haul will be retained and summed over each haul to compute the “complement” proportion in the hauls, that will be further used to derive the  $X_E$  scaling factor for each haul and species.

## **But why the heck should I define size categories?**

Several size categories have eventually to be defined for a given species, if its size distribution displays several modes. In fact, the TS per species and trawl haul used for biomass assessment are computed based on the mean length provided in the “TotalSample” fishing data. If the species length distribution is plurimodal in a haul, the mean length value, and therefore the TS, will not be representative of the real fish size distribution, and the biomass results will be biased. This is why it is advised to define several unimodal size categories in the case of species with plurimodal size distribution, to ensure that the mean length per species, size category and haul will be accurately computed. This can be done automatically using the “sizecat” function on at-length data (subsample input date). The function checks if the global size distributions of the species are unimodal. If not, it defines 2 size modes and a break value, which is used to update the belongings to size categories in the “CATEG” columns of the total and sub samples datasets. By default, the two size categories defined by the “sizecat” functions are: “small” (or “no size category”): size code “0”) or “big” (size code “G”).

## **How to compute standard “grid maps” with EchoR?**

The functions for producing standard WGACEGG grid maps have been included in EchoR. Grid maps can be computed and plotted using the “gridNmap” front-end function. See the help page and examples for details.