

GFCM – Workshop on task 1 – September 2010

**Workshop on data collection
methods
(applied to all segments of the
fleet and their coherence with the
requirements of GFCM Task 1)**

Institute for Economic Research in Fishery and Aquaculture (IREPA)



IREPA onlus

GFCM – Workshop on task 1 – September 2010

**The Italian fisheries statistical
system
and practical difficulties
encountered in the
transmission of task 1 data**

Institute for Economic Research in Fishery and Aquaculture (IREPA)



IREPA onlus

General data on the Italian Fishery Sector

Shelf area (to 200 m)	301.270 Km
Length of coastline	7.456 Km
Landing point	800
Fishing fleet (active vessels)	13.335
Number of crewmen	29.349
Volume of landing	216.567 tonnes
Value of landing	1.082 mln euro
Days at sea	1.588.502
Total Costs	783 mln euro
Average volume/vessel	16 tonnes
Average value/vessel	80.000 euro
Average day/vessel	118



IREPA onlus

3

Italian Monitoring System for the collection of fishery data

The Italian national programme for the collection of fishery data is developed according to the Data Collection Legal Framework (Council Regulation No 199/2008, Commission Regulation No 665/2008 and Commission Decision 93/2010)

Considering the characteristics of the Italian fishery (high number of species caught, spreading of the fleet along the coastline, vast number of landing points available), and considering the ineffectiveness of logbook system, the most efficient method to collect data was considered a sample survey.

The survey - a **multivariate sample survey** – aimed to achieve an evaluation of:

- Total landings and prices per species
- Fishing effort and activity
- Landings per unit of effort
- Socio-economic data



Compliance with GFCM-TASK 1

4

Methodology

- **The Universe:** Official National Archive of fishing vessels (ALP)
- **Sampling units:** vessels belonging to the ALP
- **Sampling design:** Stratified Sampling

STRATIFICATION VARIABLES

1. *Regions-GSA* (Geographical stratification)
2. *Fishing system* (Technical stratification)
3. *Length overall (LOA)* (Dimensional stratification)



Total number of strata: 201

5

Methodology

Sample allocation across strata:

Bethel method



Sample size: 1.500 units

Coverage degree: 12%

Sampling error per species: 3,5%

6

Geographical stratification



Italian GSAs:

- GSA 9: Ligurian and North Tyrrhenian Sea;
- GSA 10: South and Central Tyrrhenian Sea;
- GSA 11: Sardinia;
- GSA 16: South of Sicily;
- GSA 17: Northern Adriatic Sea;
- GSA 18: Southern Adriatic Sea;
- GSA 19: Western Ionian Sea.



Thecnical and dimensional stratification

Fishing system	LOA class	Number of vessels
Bottom otter trawl	6 classes	2.601
Pelagic pair trawl	3 classes	152
Beam trawl	3 classes	74
Purse seine	6 classes	302
Hydraulic dredge	1 class	703
Longlines	3 classes	196
Small scale fisheries	3 classes	8.820
Polyvalent passive	2 classes	440
Polyvalent	2 classes	47
Total number		13.335



Collection of information: the data collection network

The data collection is conducted by *data collectors*, people with strong relationship with fisheries sector (category associations and service centre of fisheries). Being part of the sector, they can easily contact the owners and can be easily present during the landing phase.

The network is selected and trained by IREPA and is annually reviewed.

The list of vessels, randomly selected, is submitted to the data collectors, who contact the operators to notify them of their inclusion in the sample survey and to establish an interview plan and to request permission to use data of sales notes.

After gathering and recording information, the data collectors proceed to computerise the data using a software designed by IREPA



IREPA onlus

9

Software for data entry: landing data



IREPA onlus

10

Software for data entry: economic data



Collection of information: the strenghts of survey

his methodology of obtaining data overcomes the difficulty of acquiring data from the fish markets and other official sources or from statements from operators. These difficulties, caused by the nature of the sector and outlined here, can be summarised briefly under the following aspects:

- the division of the landings in the sale phase; when landed, part of the product is taken directly by wholesalers, fishmongers or restaurants under pre-defined agreements, while another quota of the product goes to the fish market.
- the different or erroneous names attributed to fish species, owing to the difficulty of precisely identifying similar species not distinguished at a commercial level, and by the different local names used.
- the inherent underestimation in information on the landings and prices declared in the invoices of the operators for fiscal reasons



The Italian Data Collection System

part II

Statistical aspects of the of the survey



Statistical aspects of the of the survey

The statistical methodology provides guidelines for the selection of the sample and data collection. However, it is necessary to evaluate the computational aspects.

Fortunately you can make use of algorithms and software already available and of high quality.

All these procedures are readily available on the web, thus easily adaptable to situations of each country.



Statistical aspects of the of the survey

From a statistical point of view, the steps to manage the survey are:

- 1- Stratify the population
- 2 - Define the optimum sample allocation in each strata, that is: how many sample units must be interviewed in each strata?
- 3 - Select (**randomly**) the units of the population that will be sampled
- 4 – Collect the data and store them into a database
- 5 - Obtain estimates for the population from the information collected in the sample



Statistical aspects of the of the survey

In the Italian case, the following algorithms are used to execute the survey:

Bethel algorithm: a mathematical algorithm to achieve the optimum sample allocation in a multivariate sample survey

Hanurav-Vijayan algorithm: sampling units are selected according to probability proportional to size methodology

Horvitz-Thompson and **Sen-Yates-Grundy estimator:** for each collected variable, the total estimate and the coefficient of variation are obtained according to these two formula.

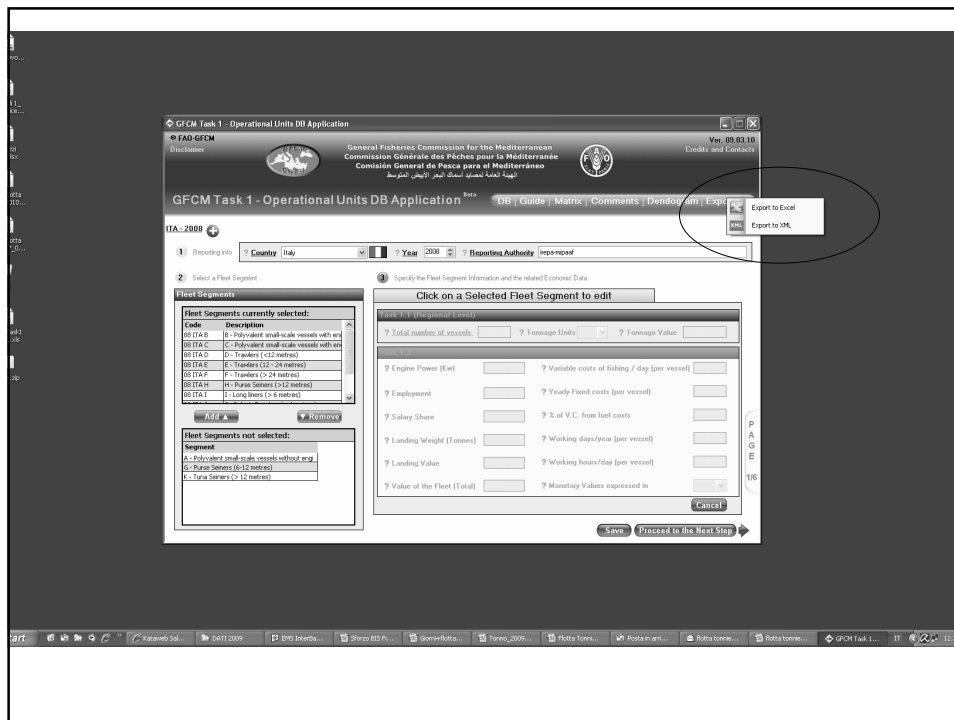


Practical difficulties

- GFCM Task 1 - Operational Units DB Application - Upload of data in the system:
 - Manual, not possible through exportation from other databases. XML schema compatible only with Excel datasets!!!! **Task1ExportDS.xsd**
 - a very high number of records
- Solution:
 - An XML format of final data required by Task 1, compatible with all possible data bases structures
 - Not necessary to use the tool for transmission of data in case data are already organized in databases

Conceptual problems

- Different levels of aggregation considered all together:
 - Segments
 - GSA
 - Gears
 - Target species
 - Species



Practical difficulties

■ Timing

- *For the data regarding economic variables (task 1.3), Italy will not be able to present the data of year $n-1$ by the month of May of year n but will only be able to present the data of the year $n-2$.*
- *Economic data for the year n will be available by the end of September of the year $n+1$*

- GFCM Task 1 - Operational Units DB

Application does not require information on:

- Methodologies used to collect data
- Quality of data reported

- GFCM Task 1 - Operational Units DB

Application, it is essential a guide with clear instructions and definitions

- 2008 data have been sent but not complete, problems of transmission

- For the future: it is essential to improve the interface to transmit data, otherwise it will not be possible to send complete data, as required by GFCM task 1