MPLEMENTATION OF A NATIONAL OBSERVATORY FOR MONITORING TECHNO-ECONOMIC DATA OF THE ITALIAN FLEET AND THE EVALUATION OF SOCIO-ECONOMIC PARAMETERS[1]

1.	Data collection and estimates of economic parameters concerning the Italian fishing fleet is produced by IREPA through a National Observatory, which dates back to the early 80's. Other sources of statistical information also exist; in particular the National Institute of Statistics (ISTAT) is involved in the production of estimates concerning some of the main figures of the Italian fishery sector.

2.	In 2000 Italy started a process of rationalisation and harmonisation of the existing surveys on the fishery sector. This process was aimed at solving one of the most important problems debated in the national and international working groups that is the availability of different and disagreed statistical sources for the same phenomenon. The practical outcome of this process was the definition of a sample survey on the catches and the relative values whose objective is to satisfy the EU legislation and, more in general, the national and international information needs. The methodology of the new survey has been developed by IREPA Onlus, in collaboration with ISTAT (the national statistical institute). In future, the Ministry for Agricultural and Forestry Policy will be responsible for the surveys and for the publication of statistics. In two years, this new approach shall replace the surveys currently carried out by ISTAT and by IREPA and it shall represent the only official statistical source on the sector. The proposed sample survey is described in the following pages.

## 1.	AIMS OF THE NEW MONITORING SYSTEM

3.	The statistical survey's aim is to gather information on the most significant parameters of the fishery sector.

4.	To this aim the existing observatory will be improved and it will consist of three main modules:

Module of evaluation of fishing effort and activity.

Module of evaluation of landings and prices by group of species.

Module of evaluation of economic and social data.

5.	The survey is based on a unique sample. More then 750 vessels are monitored each week and elementary data are later expanded to the universe (the whole Italian fleet) using statistical sampling procedures.

---

[1] The observatory has been implemented by IREPA Onlus – Istituto Ricerche Economiche Pesca e Acquacoltura, and it is partially funded by the FIFG programme, under the technical assistance measure.

6.   It is worth underlining that the research programme on systematic monitoring of fishery indicators in Italy has targeted and still targets an evaluation of economic and management features of fisheries. It does not aim to estimate and assess biological resources.

DESCRIPTION OF THE SURVEY

Assessment of the sampling universe and of the list

7.   Data collection concerning the fishery sector in Italy is very complex due to the high number of species caught, the spreading of the fleet along the 8000 Km of coastline and the vast number of landing points available (estimated at around 800).

8.   The National Fleet Register (Archivio Licenze di Pesca – ALP-) constitutes the list from which the sampling units are extracted where all Italian fishing vessels are included. The Fleet Register is held at the General Directorate for Fisheries and Aquaculture of the Ministry of Agricultural and Forestry Policies (Direzione Generale Pesca del Ministero delle Politiche Agricole e Forestali).

9.   Therefore, the sampling basis used for the survey (2001) is the Fleet Register updated to June 2000. In this archive the total number of fishing vessels in Italy is slightly higher than 18 thousand units. Twenty-seven vessels fish beyond the straits (the so-called oceanic fleet). Considering the specific activity of the latter vessels, they have been excluded from the sampling basis and data referring to this segment are gathered on the basis of taxable property, by means of agreements with ship owners and their category association (Federpesca). Tuna fishing vessels (associated in the Associazione Produttori Tonnieri Salerno) are also excluded from the sampling basis and data concerning their activity are provided directly by the Association for its member vessels.

The questionnaire and the choice of the data collectors

10.   Sample data are recorded by means of three specific questionnaires:

An annual questionnaire to record technical, dimensional and vessel – management information on the sample units and relevant socio-economic aspects (number of shipowners, their ages, their property quotas and relationships between them).

A quarterly questionnaire to record data on fixed and variable costs, and on social aspects of property and crew.

A weekly questionnaire to record information reporting activity such as fishing time and area, average number of crewmembers, gears used, quantities, prices and revenues – as per species or group of species – and trade channel for sales.

11.   The questionnaire for the survey has been designed so that the questions' sequence can be defined as "funnel-shaped". The first part of the questionnaire concerns general information such as the name of the boat, the gears employed, the days of activity at sea, the days of bad weather, non-fishing days for rest, the total number of hours, the average crew and the distance of the fishing area from the coast. The second part is meant to gather the survey's target information such as the species caught (quantity, quality, average prices, destination).

12.   The questions are structured according of the characteristics of the phenomenon and the degree of knowledge. In other words, there is no need to choose between open-response questions and fixed-response ones. In particular, an exhaustive list of the species for which quantity and prices are

required has been prepared, and the data collectors' duty is to report the individual species caught and their prices.

13.   Other important aspects of the questionnaire's design, such as the use of the language, the formulation of the questions, the correct reporting of information, are handled directly by the data collectors, that therefore represent a filter between the people interviewed and the data processing unit. It is also to be noted that the input of data for the single vessel is fully computerised; the software, specifically designed for the survey's objectives, is logically structured and also includes crosscheck programmes to avoid partial or inconsistent filling of the questionnaire.

14.   In brief, the most important annual, monthly and weekly informations recorded are the following:

Annual information

Name	Gross registered tonnage (GRT)
Maritime district where the boat has been registered, (coastal area/sector)	Gross tonnage (GT) based on London Convention (Reg. EC 2930/86)
First year of service (therefore, age)	Horsepower (kW)
Authorised fishing gears	Engine make, location and type of propeller
Maritime district from where the ship departed for fishing	Communication engine
Maritime district where the product is landed	Navigation engine
Type of association and year of its creation	Fish location engine
Number of ship-owners, their ages, their property quotas and relationships between them	Conservation equipment
Type of association and year of its creation	Employment contract used
Length overall and length between perpendiculars	

Quarterly information

Name	Fish transport cost
Month	Other running cost
Maritime district where the boat has been registered (coastal area/sector)	Labour share, wages and social insurance
Fuel (total and unit value)	Ordinary maintenance
Cost of nets	Extraordinary vessel maintenance
Cost of bait	Extraordinary hull maintenance
Cordage and ropes	Extraordinary engine maintenance
Food	Vessel insurance
Boxes and ice	Tax and other fiscal costs
Commercialisation costs	Bank charges
Other running costs	Other vessel costs

Weekly information

Name	Non fishing days for rest, repair and other
Week	Hulls
Maritime district where the boat has been registered	Average time (in hours) for each single trip
Engine used	Minimum and maximum fishing area's distance perpendicular to coast line
Average crew	Maritime district from where the ship departs
Fishing days	Maritime district where the product is sold
Total hours at sea (navigation and fishing)	For each single species or group of species landed: quantity, prices, income and commercial channel (wholesaler, fish market, retail dealer, others).
Non fishing days for bad weather	

15.   Data collectors belong to the productive or management fishery sectors. This is certainly an innovative element for official surveys. As matter of fact, data collectors are usually external to the phenomenon that is being examined and, moreover, they are often part of some public structure, in order to avoid possible influences due to personal interests. On the basis of the experience acquired in this field, it has been demonstrated that it is essential to have data collectors belonging to the fishery productive chain in order to obtain correct and timely data. Obviously, periodic inspections in the various areas are carried out in order to check the data collectors' work.

Monitoring system - survey points

Sampling design: Single Stage Stratified Sampling

16.   The sampling design is based on a single stage sample stratified over two variables. Stratification is carried out in order to create strata as homogeneous as possible, using characteristics correlated to the target variables.

17.   The maritime regions of the Italian coast represent the first stratification variable from an administrative point of view. The sampling design considers only 13 of these 15 maritime regions, since there are no enrolment offices in Basilicata and, due to the small size of Molise's fleet (0.3% of the total number of Italian fishing vessels), the latter has been aggregated to Abruzzo.

18.   The second stratification variable is the fishing system used by the vessel; the fishing fleet is therefore divided in groups and each fishing vessel belongs to one of these groups according to the fishing system it uses. For the present surveys the following systems are considered: trawlers, dredges, drift-nets, long-line, multipurpose and small scale fishery. The identification of the fishing system follows the fishing systems actually present in the Italian fleet, while taking into account criteria of consistency with the segmentation considered under the MAPG IV (Multi Annual Guidelines Programme).

19.   Furthermore, for some systems (trawler and multipurpose) and for some regions (Abruzzo, Marche, Puglia, Sicilia and Veneto) another stratification is carried out on the basis of the dimensional parameter GT (Gross Tonnage). This in order to obtain more homogeneous study domains and to take into account the differences between the coastal or short range trawler and the high sea trawler.

20.   Moreover, for dredges registered in Veneto, Emilia Romagna and Marche regions, another stratification has taken place, based on the marine compartments to which they belong. This decision has been taken in order to consider the peculiar situation of this fishing system that, due to the actual national regulation which allows for a self management approach, at the compartment level by the Bivalve Molluscs Management Consortiums (Consorzi di Gestione dei Molluschi bivalvi).

21.   Finally, those vessels authorised to fish tuna among other species represent an additional specific stratum. This stratum is identified in reference to the 7/2/2000 ministerial decree regarding the "determination of individual fishing quotas for blue-fin tuna for the year 2000", according to which a complete list of the units operating on this target species is reported. The vessels of the Associazione dei Produttori Tonnieri di Salerno do not belong to this segment and, as said before, they are not sampled and belong to a specific stratum, since they fish exclusively tuna.

22.   The final number of strata or domains from which the overall sample is extracted amounts to 70. To the latter, the two strata that are out of the sample (Oceanic fleet and Associazione dei Produttori Tonnieri di Salerno) need to be added.

Sample size and allocation across strata

23.   The size of the sample is assessed on the basis of the evaluation of the sampling error.

24.   In particular, this implies the specification of the estimates' reliability, which consists in setting the value of the average square error; furthermore, since correct or approximately correct estimators are used, the values of the estimates' variances are determined.

25.   In the case of single stage stratified sampling, and in the hypothesis of extracting the sampling units with equal probability and without re-pooling, the sampling variance of the estimate $\hat{Y}$ of the total Y is described by the following expression:

$$V(\hat{Y}) = \sum_{h=1}^{H} N_h^2 \frac{N_h - n_h}{N_h} \frac{S_h^2}{n_h} \quad ,$$

given these definitions:

H   represents the number of strata in the population

$\hat{Y}$ represents the estimated total of Y for the population

N_h represents the total number of sampling units in the *h*th stratum

n_h represents the total number of sampling units measured in the *h*th stratum

S_h^2 represents the variance of Y for the *h*th stratum

26.   Therefore, for a given population, variance varies both as a function of the overall size of the sample n and, for a set value of n, as a function of the numerosities $n_1, \ldots, n_h, \ldots, n_H$, with the imposed constraint that their sum must be equal to n.

27.   Among the various ways of determining the sampling sizes of the H strata, the Neyman criterion has been used instead of the proportional one. The Neyman method is based on the criterion by which a different percentage of elements is drawn from each stratum in order to minimise the value of $V(\hat{Y})$.

28.   Nevertheless, the Neyman method can be applied only in the case of a single target variable, otherwise we would obtain a different sample size for each variable considered. Since the present survey is multivariate, that is, the variables investigated are more than one, to calculate the sample size we use the Bethel method, which is the application of Neyman's method to the multivariate case. The approach used by this method is to transform the analysis into a linear programming model that allows to identify the sample size and the allocation across strata, minimising the variances of all variables simultaneously (see also Bethel, 1989).

29.   The optimal allocation across strata for multi-scope studies has been solved by Bethel using the Kuhn-Tucker theorem and then deriving the expressions for the optimal allocation in terms of the LaGrange multipliers.

30.   In order to apply this method, a pre-estimate of the S^2_h variances is required; in other words, the variances of the target variables of the survey must be known. For this purpose, the results of a sampling survey conducted in 1999 on the production data of more than 400 vessels have been used. It was decided that the variables to be considered to calculate the sampling size, must be those for which the variance is highest; catches by species and by region were chosen. So, to apply the Bethel model monthly estimates of the total catches by species and by region must be known.

31.   The Bethel method has been applied to the data available for 1999, with a procedure specifically implemented on SAS basis.

32.   In a first phase, since the sample size is inversely proportional to the error level, three different levels of sampling, with three different levels of maximum acceptable error, have been identified. However, the final choice has been the lowest level of the maximum acceptable error (average sampling error of 5%, on a confidence interval of 95%), with a total sample size of 440 units and a degree of coverage of 2.4%.

33.   The final size has been obtained by applying the Bethel procedure with a constraint of minimum size per stratum of 4 units (with the exception of the stratum Campania-mechanised dredge; the numerosity considered for this stratum is the one that derived from the application of Bethel's algorithm).

34.   The sample units have been extracted by applying the PPS method (proportional to the size method). Each unit has a different probability to be sampled and this probability is proportional to the following measure:

$$P_i = \frac{LFT_i}{LFT_h}$$

where:

i = a generic vessel

h = stratum

LFT= overall length.

Among the different methods to extract a sample, the Hanurav algorithm was chosen.

Expansion factors[2]

35.   In brief, to pass from the sampling data to the overall estimates, the weight applied to each elementary data is the following:

$$k_{hi} = \frac{1}{\pi_{hi}} = \frac{1}{n_h \dfrac{LFT_i}{LFT_h}} = \frac{LFT_h}{n_h LFT_i}$$

where:

π_{hi}: the probability of the unit i to be selected

n_h: sample size for the stratum h

LFT: cumulative vessel length.

36.   In the case of non-responses, the initial weights k_{hi} are adjusted on the basis of data referring to the responses (r_h) and the non-responses (s_h) of the sample (n_h). The method consists in multiplying the initial weights (k_{hi}) by a factor (d_h) equivalent to:

$$d_h = \frac{r_h + s_h}{r_h}$$

37.   The resulting weights (v_{hi}), that are called base weights, because they are finalised to the calibration of the sum of weights of the different levels of the population and to the elimination of the bias caused by different non-response rates among strata. In our specific case, the base weights are:

$$v_{hi} = \frac{r_h + s_h}{r_h} \frac{LFT_h}{n_h LFT_i}$$

38.   The hypothesis underlying this method is that the total number of non-responses does not have an impact for homogeneous groups of statistical units.

39.   It is demonstrated that, for homogeneous groups of responses (response homogeneity group – RHG), the estimator

$$\hat{Y}_{(r)h,RHG} = \sum_{i=1}^{n_h} v_{hi} \, y_{(r)hi}$$

is unbiased.

---

[2] For greater detail, the reader is invited to refer to IREPA: "Osservatorio economico sulle strutture produttive della pesca marittima in Italia". 2000. Ed. F. Angeli, methodological annex

Non-sampling errors

40.   The final step of the survey is the check on the elementary data to eliminate non-sampling errors. This has been achieved by means of computer programmes implemented to correct the erroneous values and to permit statistical data analysis. These programmes are mainly based on graphical analysis of elementary data.

41.   Other standard interconnected computer programmes were added to support procedures for controlling, filing, correcting and elaborating data. These are able to facilitate the process of assessment, transmission and diffusion of statistical information.

3.   RICA SOFTWARE

42.   Specific software has been developed to conduct the survey.

43.   The software is divided into the following sections:

Data collectors' software to be filled and transmitted.

Software for data processing: queries on specific groups of elementary data.

Software for the production of the final tables: checking and correction of the elementary data, application of the expansion factors.

44.   Software is developed on Delphi language. Data bank is structured on Interbase module.

45.   Specific statistical software (Windows Statsoft) is used to treat and to analyse data for scientific purposes.