Ministero delle politiche agricole alimentari e forestali

DIPARTIMENTO DELLE POLITICHE EUROPEE ED INTERNAZIONALI

DIREZIONE GENERALE DELLA PESCA MARITTIMA

E DELL'ACQUACOLTURA

Italian National Program

Under Council Regulation (EC) N° 199/2008 and Commission Regulation (EC) N° 665/2008

National Program 2011-2013
Version December 2010

TO BE APPLIED IN 2014 – 2016 ACCORDING TO COMMISSION IMPLEMENTING DECISION of 30.8.2013 extending the national programmes for the collection of primary biological, technical, environmental and socio-economic data in the fisheries sector for the period 2011-2013 to the period 2014-2016

I. General framework

The present document represents the Italian national programme for the collection of fishery data for the years 2011 and 2013, according to the legal Community framework put in place in 2008/2009 with the adoption of a Council Regulation[1], a Commission Regulation[2] and the two Commission Decisions[3] laying down the detailed rules of application.

In accordance with chapter II of the Annex of the Commission Decision 93/2010, this national program comprises of the following modules:

(1) Module of evaluation of the fishing sector:
   (a) Section for the collection of economic variables
   (b) Section for the collection of biological variables
   (c) Section for the collection of transversal variables
   (d) Section for research surveys at sea

(2) Module of evaluation of the economic situation of the aquaculture and processing industry sectors:
   (a) Section for the collection of economic data for the aquaculture sector
   (b) Section for the collection of economic data for the processing industry

(3) Module of evaluation of the effects of the fishing sector on the marine ecosystem

(4) Module for management and use of the data covered by the data collection framework

In implementing the new DCF framework, continuity with data and time series collected under the previous DCF will be assured. A particular attention will be given to the regional approach and compliance with RCMMed&BS will be assured.

[1] Council Regulation (EC) No 199/2008 of 25 February 2008 concerning the establishment of a Community framework for the collection, management and use of data in the fisheries sector and support for scientific advice regarding the Common Fisheries Policy.

[2] Commission Regulation (EC) No 665/2008 of 14 July 2008 laying down detailed rules for the application of Council Regulation (EC) No 199/2008 concerning the establishment of a Community framework for the collection, management and use of data in the fisheries sector and support for scientific advice regarding the Common Fisheries Policy.

[3] Commission Decision (2008/949/EC) adopting a multi annual Community programme pursuant to Council Regulation (EC) No 199/2008 establishing a Community framework for the collection, management and use of data in the fisheries sector and support for scientific advice regarding the Common Fisheries Policy.

[3] Commission Decision (2010/93/EC) adopting a multiannual Community programme for the collection, management and use of data in the fisheries sector for the period 2011-2013.

II. Organisation of the National Programme

II.A National organisation and co-ordination

Two annual national meetings for co-ordination will be organised each year, according to article 3 of Commission Regulation (EC) No 665/2008. Research institutes taking part to the implementation of the national program will participate to the meetings. Moreover, an ad hoc scientific Committee manages the National program for the gathering of fishery data. This Committee is composed of the national correspondent and of other member's experts in the fields of biology, economy and statistics. The scientific Committee is entrusted with the task of co-ordinating the data collection program, of monitoring the entire process, of proposing adjustments aimed at improving the methodologies of data gathering and endorsing the pilot studies to be conducted within the present national program.

II.B International co-ordination

Standard table II.B.1 gives an overview of the international co-ordination meetings (Planning Groups, Study Groups, Regional Co-ordination Meetings, etc.) and the workshops that will be attended.

The annual list of eligible meetings for 2011 and onwards has not yet been provided by the EC, therefore the column 'Eligible under DCF' of table II.B.1 is empty and the list of workshop is not complete.

II.C Regional co-ordination

Italy will participate in the following Regional Co-ordination Meetings:

- RCM for the Mediterranean and Black Seas (RCMMed&BS).

- RCM for the "Other regions".

In different sections of the present National program, it is written that some of the proposed actions need to be discussed and agreed in the RCMMed&BS, according to Article 5 of Reg. 199/2008.

Therefore, an amendment of the present NP will be presented in case the RCMMed&BS will not approve or agree on some of the planned actions or in case the RCMMed&BS will decide to carry out sharing activities in order to increase the accuracy, effectiveness and cost efficiency of data collection.

III Module of evaluation of the fishing sector

III.A General description of the fishing sector

The national fleet consists of about 15 thousands vessels, of which more than 9 thousands are classified in the segment of passive gears less then 12 meters (source: official vessel register December 2008).

The official vessel register classifies 18 vessels as fishery outside the Mediterranean Sea; 10 of these vessels since the beginning of 2009 are inactive. 7 operate around Mauritania (Area FAO 34.1; 34.3) waters and 1 in the Guinea Conakry (Area FAO 34.3). All these vessels operate with trawls nets targetting demersal species.

Another vessel, operated in the Indian Ocean (Area FAO 51) catching large pelagic species (Albacares and Skipjack tunas) with purse seine, is not more under Italian flag.

The Italian Mediterranean fleet is characterised by a strong multispecifity and multigear activity. The fishing sector appears highly fragmented in many regions along the coast and there are many large structural and technical differences in vessels from different geographical areas. The fleet is classified into the following segments: bottom trawlers, purse seiners, midwater pair trawlers, dredges, polyvalent vessels (using a combination of passive and mobile gears), vessels using passive gears and long liners.

The bottom trawlers represent the most important segment of the Italian fleet in terms of production (37% of total landings). The main target species are shrimps, hakes, mullets, nephrops, and cuttlefishes.

Small-scale fishery is the most relevant Italian fleet segment in terms of number of vessels, representing 66% of national total. Small scale fishery accounts for about a quarter of the national value of landings. Even if average incomes are low, these vessels represent an important economic resource in some geographical areas with a high level of dependence on fishery.

Table III.A.1 gives an overview of the Italian fisheries covered by this Program.

III.B Economic variables

SUPRA-REGION: Mediterranean Sea and Black Sea

III.B.1 Data acquisition (Mediterranean Sea)

The parameters to be evaluated for analysis of the economic situation of the sector are those reported in Appendix VI of Commission Decision 2010/93/EC.

This section describes the data sources and methodologies to estimate all the variables listed in Appendix VI excluding those classified as transversal variables.

The following transversal variables:

- Gross value of landings
- Fleet (number, mean LOA, mean vessel's tonnage, mean vessel's power, mean age)
- Days at sea

- Value of landings per species
- Average price per species,

are discussed in chapter III.F.

Each economic variable will be estimated for each group of vessels as defined in Appendix III of Commission Decision 2010/93/EC

In addition to variables listed in Appendix VI of Commission Decision 2010/93/EC, environmental indicators to measure the effects of fisheries on the marine ecosystem will be calculated. In particular, within this section of the NP, the methodology to calculate the "fuel efficiency of fish capture" (indicator 9 of Appendix XIII of Commission Decision 2010/93/EC) is described. This indicator is calculated as the ratio between value of landings and cost of fuel, by quarter and by metier. The collection of value of landings by metier is described under section III.F.3. The quarterly cost of fuel by metier is desciber in the following paragraphs

b) Type of data collection

Standard table III.B.3 illustrates the different types of data collection schemes that will be used for different segments and different variable.

Both Probability Sample Survey and Census will be used. In particular, the following variables will be estimated through a Probability Sample Survey:

1. Income
    Income from leasing out quota or other fishing rights
    Direct subsidies
    Other income
2. Personnel costs
    Wages and salaries of crew
    Imputed value of unpaid labour
3. Energy costs
4. Repair and maintenance costs
5. Other operational costs
    Variable costs
    Non-variable costs
    Lease/rental payments for quota or other fishing rights
6. Investments
    Investments in physical capital
7. Financial position
8. Effort
    Energy consumption

The sample survey is continuous in character and has a reference period of one year. The target population of the survey comprises the Italian fishing fleet and the list is based on the

Vessel Register kept at the Directorate-General of Fisheries and Acquaculture of the Ministry of Agricultural and Forestry Policies. It includes vessels < 12 meters.

In order to get reliable estimates, and on the basis of previous surveys, it is estimated that the sample should inlcude at least 1200 vessels (see Annex I for details on the definition of the sample size).

The survey methodology and the questionnaire are reported in Annex I.

The following variables will be estimated according to a Census (in particular, the methodology that will be used is that suggested by the study on "evaluation of the capital value, investments and capital costs in the fisheries sector" (No FISH/2005/03):

- Capital costs
- Annual depreciation
- Capital value
- Value of physical capital: depreciated replacement value
- Value of physical capital: depreciated historical value
- Value of quota and other fishing rights

The primary data sources for the calculation of capital costs and capital value will be the national fleet register.

Employment (in terms of engaged crew, FTE National, FTE harmonised) will be estimated according to a Census (in particular, the methodology that will be used is that suggested by the study on "calculation of labour including FTE (full-time equivalent) in fisheries" (No FISH/2005/14).

The estimation of "number of fishing enterprises/units" will follow a census approach.

c) Target and frame population

The population is all vessels in the Community Fishing Fleet Register on 1 January. Inactive vessels will be considered as a separate segment. For inactive vessels only capital value, fleet and capacity will collected as required by the DCF.

The frame population of the survey comprises the Italian fishing fleet and the list is based on the Vessel Register kept at the Directorate-General of Fisheries and Acquaculture of the Ministry of Agricultural and Forestry Policies.

Table III.B.1 gives information on (i) the population nos. by fleet segment, (ii) the planned sampling levels and sample rates, and (iii) the sampling method that will be used. The fleet segments in table III.B.1 correspond to those listed in Appendix III (MP) of the DCF and the 'Total population nos.' refers to the official fleet register at the 1st of January 2010.

Standard Table III.B.2 reports the segments that have been clustered. Clusters are named after the biggest segment in terms of number of vessels.

Clustering of fleet segments follows the recommendations given by SGECA 09-02 in particular, only segments similar to other segments have been clustered for sampling purposes. The approach to determine similarity is based on statistical analysis that verifies the homogeneity of the cludtered segments. These statistical analyses are reported in Annex I.

8 clusters have been necessary in order to design the sampling plan and to report economic variables. Of these clusters, 3 take place with segments of more than 10 vessels. The first case is that of the dredgers. The clustering is justified by the fact that: Dredgers are based almost exclusively in central-north Adriatic cost. Vessels are very specialised targeting only clams and smooth-callista (Venus gallina and Callistachione) and they are homogenous in terms of

size, gears and fishing practises. Vessels have an average LOA of 13.4 meters and 81% of them belong to the class 12-18 meters. Therefore, the split into the class <12 m and > 12 meters is not statistical reliable for this segment. Moreover, it is demonstrated that revenues are not correlated with the LOA of the vessels (see graph) and this fact proves the high level of homogeneity of the vessels.

The second case is that of Purse seiners 12-18 m. also in this case vessels are concentrated near the average value, as shown by the graph in annex I of NP . The clustered segment (purse seiners 12-18 m*) is composed by 143 vessels with an average LOA of 13.5 m. The clustered segment is highly homogenous from a statistical point of view.

The third case is that of vessels using polyvalent "passive" gears only 12 - 18 m *. The clustering of the 20 vessels >18 m into the lower LOA class is necessary in order to design the sampling plan. In fact these vessels are scattered along the Italian coast (GSA 9, GSA 10, GSA 11, GSA 17, and GSA 19). The average length of these 20 vessels is 19.6 meters, therefore very close to the upper limit of the 12-18 m class.

d) Data sources

Data will be collected by using a computerised questionnaire in the case of the following variable groups:

- income
- personnel costs
- energy costs
- repair and maintenance costs
- other operational costs
- investments

For other variable groups, a census will be used and data sources are questionnaires and additional information based on spefic knowledge. Detailes are given under paragraph “estimation”.

Information on data sources used to collect each variable per segment are provided in Standard table III.B.3. A copy of the questionnaire is provided in Annex I.

e) Sampling frame and allocation scheme

The probability sample survey carried out to estimate economic variables is a multivariate sample survey. The sample unit is the single vessel and this unit is selected from the Vessel Register.

The sampling is of a stratified nature in that the fishing vessels of the fleet are divided into homogenous groups based on suitable variables and independent samples are taken from each of these clusters. The elementary economic data are collected trough a specific questionnaire.

The optimum sample number per stratum is defined according to Bethel’s procedure (1989), the vessels are selected using PPS methodology (Probability Proportional to Size) and, to be more exact, using the algorithm of Hanurav-Vijayan. In each of these phases the data is elaborated using the R software language. For more details see Annex I.

III.B.2 Estimation

Variables estimated through a probability sample survey

As regards the variables estimated through a probability sample survey, the estimate of totals per stratum is obtained through the Horvitz-Thomson estimator, while the Sen-Yates-Grundy formula is used to estimate the relative sampling error. Finally, for the estimate of the variance of the population relative to each stratum, the formula of Chaudhuri is used. For more details see Annex I.

As regards the imputation of non-responses, a process of localization of errors is applied . The control procedure of the survey can be considered as *interactive graphic micro-editing of the univariate type*. The term interaction refers to the fact that, in the procedure of localization of errors, there are not only automatic phases but also phases which require human intervention to investigate the situation and to evaluate the effective presence of the error (therefore the human intervention regards the localization phase and not that of imputation). The control is mainly of the univariate type because the variables are controlled individually and only in rare cases are suspected relationships existing among them controlled, using suitable synthesis indexes. During the various phases wide use is made of graphic tools to visibly evaluate situations marked as errors. Finally the word micro-editing is used because the data is gathered in suitable domains of study within which the sampling units can be considered very homogenous. For each of these sets of data, suitable control functions are first calculated, and then, for each of them, certain rules of incompatibility are verified. In the case of activation of conditions of error that is in the case where the observed value does not belong to the region of acceptance those control functions are then observed individually for all the sampling units forming the single domain. Thus the sample unit, or units, responsible for the activation of conditions of error is localized for the entire domain of study and then the imputation of the erroneous data follow. For more details about the method of imputation of non-responses, see Annex I.

Capital value

The estimation of the capital value (GCS) will consist of three steps:

1. Specification of the composition of the active fleet by age

2. Estimation of price per unit of capacity (e.g. per GT)

3. Calculation of the value of each vintage (year of construction) of the fleet and either converting values of all vintages to current prices or to historic prices using price indices.

The specification of the composition of the active fleet by age will be done by processing the fleet register (which reports the year of construction for each vessel).

Regarding the estimation of the value per capacity unit, the current price (historical value) per unit of capital will be estimated using the RINA construction index. In order to apply the PIM (perpetual inventory method) and in absence of other possibilities, as initial benchmark of the gross capital stock, the RINA (Italian Naval Register) construction index will be used.

The RINA indexes are based on a survey undertaken by the Italian Naval Register in 1992. These indexes express the value of a GRT unit for different GRT classes and types of hull (Table B.1). However, they don't distinguish by fleet's segments and vessel's age.

GRT classes	wood	steel	Fiberglass
less than 5 GRT	13,428	-	11,362
from 6 to 10 GRT	12,395	13,428	9,296

from 11 to 20 GRT	10,846	8,263	7,747
from 21 to 50 GRT	8,263	9,296	-
from 51 to 100 GRT	7,230	8,263	-
over 100 GRT	6,714	9,296	-

Table B.1 - RINA indexes: Price per GRT unit, 1992 (Euro)

Hence, the year 1992 will be considered as base year for the application of PIM. Then, on the basis of the RINA indexes the total price per GRT unit for the vessel's vintage 1992 will be estimated.

The calculation of the value of each vintage will require the use of a price index per unit of capital. The year classes (vintage) included in the data set will be estimated on the basis of the annual price change of Italian producer index for heavy machines, which also includes boat constructions. This index lies between 0.1 and 1.3 per annum, with an average of 0.5% (*Source*: ISTAT)

Finally, the share in total investments of hull, engine, electronics and other equipment has been already estimated on the basis of a specific survey conducted in previous years on a sample of 62 fishermen (table B.2). As expected the share of equipment (electronics, engine and other equipments) tends to decrease with the increasing of the hull's dimension.

%	No Vessels	hull	engine	other equipment	electronics
LOA>=12	44	36%	38%	24%	2%
LOA<12	18	35%	38%	17%	10%
Trawlers	7	46%	25%	27%	2%
Passive gears	5	33%	39%	16%	12%
Total sample	62	36%	38%	23%	3%
General scheme		60%	20%	10%	10%

Table B.2 - Percentage composition of investments by main sub-segment and group of assets

Capital costs

The assumptions for the calculation of the capital costs are as follows:

The age schedule of the assets is the same of the general scheme of the capital study. Hence it will be assumed that engine is renovated every 10 years, electronics every 5 years, other equipment every 7 years while the hull is never renovated.

The depreciation rates are reported in the following table:

	General scheme
Hull	7%
Engine	25%
Electronics	50%
Other equipment	35%

Table B.3 - Depreciation rates by groups of assets

Ten year government bond yields will be used as a measure for long-term interest rates.

Concerning the fiscal depreciation scheme, the Italian tax scheme[4] provides the following rates (Table B.4) for the fishing industry:

[5] Ministerial Decree No. 31, January 1998.

	Italian tax scheme	Age schedule (renovation years)
Hull	12.5%	8
Engine	31.5%	3
Electronics	20%	5
Other equipment	31.5%	3

Table B.4 - Fiscal rates by groups of assets

Applying the depreciation rates to the gross capital stock by asset type will generate the depreciation costs of the current year.

Employment

The engaged crew is defined as the number of jobs on board, equal to the average number of persons working for and paid by the vessel. This includes temporary crew as well as rotation crew.

Full-time equivalent (FTE) National is based on the national reference level for FTE working hours of the crew members on board the vessel (excluding resting time) and the working hours onshore. If the annual working hours per crew member exceed the reference level, the FTE equals 1 per crew member. If not, the FTE equals the ratio between the hours worked and the reference level. Full-time equivalent (FTE) harmonised is based on a threshold of 2000 hours per FTE.

In order to trace the reality of labour input in fishing as closely as possible, the study proposes to follow a 'métier' approach, where 'métier' is defined as a "specific activity of a segment or the use of a specific fishing gear within a segment".

Data will be estimated for each combination fishing segment – metier identified. The estimation of the employment in terms of FTE will be made by assuming a FTE threshold (representing the standard working time for the fishing sector) expressed in terms of yearly hours per man.

The calculation of FTE will take place in the following steps:

a) Total man hours worked in each métier will be calculated on the basis of the fleet, days at sea, crew and average hours per man per day, using crew on board

b) Total man hours of the segment will amount to the sum over the métiers.

c) From the total man-hours, hours per man per year will be calculated.

d) Annual man hours for crew will be compared with the FTE threshold, from which the number of FTEs will follow.

Some assumptions have to be made in order to define the FTE threshold in terms of yearly hours for the Italian fishing sector.

Usually, and in particular by the Italian official statistics, the standard working time is fixed based on the collective national agreements referring to the sector where the employment in terms of FTE is requested. Taking into account that:

a) in the Italian fishery sector a standard working time does not exist[5];

b) considering the great differences among the various fishing technique and

c) referring to what happens in other sector (i.e. agriculture[6]),

[5] The national collective labour agreement for the personnel employed on the fishing vessels for the period 01.01.2005-31.12.2008 establishes, at art. 16, that in the fishery sector, giving the specificity of the fishing activity due to the great uncertainty of catches and weather and to the arrangement of catches and equipments, the working time cannot be defined *a-priori* and will depend on the specific needs of the fishing trip.

[6] The FTE threshold for the Italian agricultural sector has been set in 248 days. It has been estimated based on the 2000 data of the Survey on the economic performance of the agricultural enterprises, that counted the working time for each workplace.

the FTE threshold for the estimation of Italian employment data will be determined on the average days at sea per vessel over the last 6 years.
It is assumed that, giving the inexistence of rotation, on average, each crewman engaged on a vessel works for a number of days per year equal to the days at sea of the vessel he is employed on.
In this way the FTE threshold per year per crewman in terms of days ($Th_{days/year}$) is determined for each fishing segment. This can be written as:

$$Th_{days/year} = DaS\ /man = DaS\ /vessel$$

In order to have the FTE threshold in terms of hours ($Th_{hours/year}$), the FTE/days is multiplied times the "standard" daily working hours ($Th_{hours/day}$). In formula:

$$Th_{hours/year} = Th_{days/year} * Th_{hours/day}$$

This calculation will be repeated for each fishing segments. The FTE threshold used is the average of the FTE by fishing segments. It is equal to 1,600 hours per year. Therefore:
a) each crewman working annually 1,600 hours or more will be counted as one FTE.
b) each crewman working less than 1,600 hours per year will be counted as a percentage of an FTE according to the number of hours worked in relation to the threshold of 1,600 hours.
c) working time is the time spent on fishing and related activities on board or on shore. This means that working time is only a part of the duration of a fishing trip.
Regarding the calculation of the imputed value of the unpaid labour, an average labour cost will be estimated on the basis of national labour contracts and it will be imputed to the item "personnel cost"

Financial position

The financial position is defined as the % debt in relation to total capital value. It will be estimated trough a specific survey. The sample size will be defined in order to interview a number of vessels proportioned to the fleet composition by fishing segments and allowing having coverage of the total population of at least 5%.

The questionnaire will include questions on the level of debts, in their different forms (short, medium and long terms) and different sources. The same well trained data collectors - in charge of collecting data on revenues and costs – will be used. This should increase the response rate that has been quite low in previous year, especially for the artisanal fisherys.

Number of fishing enterprises/units

The number of fishing enterprises/units will be given by the following size categories:

1) owned vessel

2) 2-5 owned vessels

3) >5 owned vessels

The source will be the fleet register at 1st of January. The fleet register includes information on the owners of each recordered vessel. Shared ownership (involving more than one person) will be regarded as one unit.

Quarterly cost of fuel by metier

Quarterly cost of fuel by metier will be estimated considering a proportionality with the quarterly effort by metier.

Imputed value of unpaid labour

Imputed value of unpaid labour can be relevant in the case of small fishery enterprises managed at family levels. In this case, the average labour cost, estimated for “paid labour force” will be applied and imputed to the item “personnel cost” for enterprises declaring to have not dependent labour force.

III.B.3 Data quality evaluation

Table III.B.3 gives details on the sampling methods used (column 'Sampling strategy') and on the methods used to assure the quality of the collected data, for each parameter.

Information on data quality is given in terms of target precision levels in the case of random sample and in terms of coverage rate in the case of fixed panel. The estimation of the variance for the calculation of the CV in the case of statistical sample is explained in Annex I (methodology for the estimation of economic variables).

Methodologies for calculation of capital value are in line with those described in the report of the study N° FISH/2005/03, Evaluation of the capital value, investments and capital costs in the fisheries sector.
The data sources for estimation of variables in Appendix VI are homogenous, therefore data are consistent. This guarantees that economic indicators on profitability (gross cash flow, gross value added, net profit) are not biased by the use of different data sources.
For more details about the data quality evaluation, see Annex I.

III.B.4 Data presentation

In 2011, data referred to 2010 will be collected. They will be available by the end of September 2011. The same time lag will be applied for data referred to the years 2012 and 2013.

III.B.5 Regional coordination

Following the proposal of the 2009 RCMMed&BS and in order to ensure the comparability of data at regional level, clusters will be discussed and agreed by RCMs after the first year, i.e. in 2010.
There are several points that would get benefits from discussions at regional levels, however, in previous RCMMed&BS economic issues have not been treated extensively. In future RCMMed&BS Italy will propose to standardise the methodologies and the variable definition in order to increase the level of analysis in the Mediterranean area. A preparatory work at European level made by independent experts could facilitate RCMMed&BS discussions.

III.B.6 Derogations and non-conformities

No derogation is requested. All the vessel population will be considered.

SUPRA-REGION: Other Regions

III.B.1 Data acquisition (Other Regions)

The official vessel register classifies 18 vessels as fishery outside the Mediterranean Sea. One vessel operates in the Indian Ocean targeting tunas and the rest operate in the Atlantic, fishing demersal species.

Table III.B.1 gives a general outline of (i) the population nos. by fleet segment, (ii) the planned sampling levels and sample rates (columns 'Planned sample no.' and 'Planned sample rate'), and (iii) the sampling method(s) that will be used (column 'Sampling strategy').

III.B.2 Estimation (Other Regions)

The economic variables listed in Appendix VI will be collected through a census based on an elaboration of official balance sheets that, according to Italian law, have to be published and accessible to the public. The elaboration will allow the estimation of the the following data: income, personnel costs, energy costs, repair and maintenance costs, other operational costs, investments, financial position.

Capital costs and capital value will be estimated according to the methodology suggested by the study on "evaluation of the capital value, investments and capital costs in the fisheries sector" (No FISH/2005/03).

Employment (in terms of engaged crew, FTE National, FTE harmonised) will be estimated according to the methodology suggested by the study on "calculation of labour including FTE (full-time equivalent) in fisheries" (No FISH/2005/14).

The number of fishing enterprises/units will be given by the following size categories:

1) owned vessels

2) 2-5 owned vessels

3) >5 owned vessels

The source will be the fleet register at 1st of January.

The following transversal variables

- Gross value of landings
- Fleet (number, mean LOA, mean vessel's tonnage, mean vessel's power, mean age)
- Days at sea
- Value of landings per species
- Average price per species,

are discussed in chapter III.F.

III.B.3 Data quality evaluation (Other Regions)

Data will be collected trough a census, which attempts to collect data from all members of a population. This will include collection of data from administrative records that is balance sheets.

The variability of the estimates and the bias derived from non-responses will be assessed trough appropriate statistical methods. The variability achieved and the bias will be reported in the technical reports.

III.B.4 Data presentation (Other Regions)

In 2011, data referred to 2010 will be collected. They will be available by the end of September 2011. The same time lag will be applied for data referred to the years 2012 and 2013.

III.B.5 Regional coordination (Other Regions)

Italy will attend the first RCM for "Other regions" (Spain, March 2010) and related recommendetions will be considered.

III.B.6 Derogations and non-conformities (Other Regions)

No derogation requested.

III.C Biological - metier-related variables

REGION: Mediterranean Sea and Black Sea

III.C.1 Data acquisition (Mediterranean and black sea)

a) Codification and naming convention
Sampling will be performed in order to evaluate the quarterly length distribution of species in the catches, and the quarterly volume of discards.
Data will be collected by metier referred to level 6 of the matrix defined in Appendix IV (4) (EU Decisions 949/2008 and 93/2010) per quarter and per GSA (Level 4 Appendix I EU Decisions 949/2008 and 93/2010).

Must be highlited that, as reported by the PGMed 2009 and endorsed by the RCMMed&BS 2009: "*the definition at level 6 of nets, all trawl types, purse seines and lampara nets was defined in most cases at the minimum mesh sizes defined by Regulation 1967/2006, except for the relevant métiers of Black Sea.*" Following this issue the reference list of mètiers agreed at Regional level during the RCMMed&BS 2009 has been used for the selection of the metièr to sample. The following table (C.1) reports the namig convention for the Mediterranean and Black Sea mètier.

Metier coding	Metier naming
DRB_MOL_0_0_0	Boat dredge for molluscs
FPN_LPF_0_0_0	Stationary uncovered pound nets for large pelagic
FPO_DEF_0_0_0	Pot and Traps for demersal species
FYK_CAT_0_0_0	Fyke nets for eels
FYK_DEF_0_0_0	Fyke nets for demersal species
GND_DEF_0_0_0	Driftnets for demersal fish
GND_SPF_0_0_0	Driftnets for small pelagic fish
GNS_DEF_360-400_0_0*	Set gillnets for demersal fish
GNS_DEF_>=16_0_0	Set gillnets for demersal fish according meshsize regulation
GNS_SLP_>=16_0_0	Set gillnets for small and large pelagics according meshsize regulation
GTR_DEF_>=16_0_0	Set trammel nets for demersal species according meshsize regulation
LA_SLP_14_0_0	Lampara nets according meshsize regulation
LHP-LHM_FIF_0_0_0	Hand and Pole lines for finfish
LHP-LHM_CEP_0_0_0	Hand and Pole lines for cephalopods
LLD_LPF_0_0_0	Drifting longlines for large pelagic
LLS_DEF_0_0_0	Set longlines for demersal fish
LTL_LPF_0_0_0	Trolling lines for large pelagic
OTB_DEF_>=40_0_0	Bottom otter trawl for demersal species
OTB_DWS_>=40_0_0	Bottom otter trawl for deep water species
OTB_MDD_>=40_0_0	Bottom otter trawl for mixed demersal and deep water species
OTM_MPD_>=13_19_0_0*	Pelagic traw according meshsize regulation
OTM_MPD_>=20_0_0	Midwater otter trawl for mixed demersal and pelagic species
PS_LPF_14_0_0	Purse seine for large pelagic
PS_SPF_>=14_0_0	Purse seine for small pelagics according meshsize regulation
PTM_SPF_>=20_0_0	Pelagic pair trawl for small pelagic species
SB-SV_DEF_0_0_0	Beach and boat seines for demersal species
TBB_DEF_0_0_0	Beam trawl for demersal trawling
MISC	Miscellaenous metiers (defined at national level)

Table C.1 – List of finalised mètier at level 6 for the Mediterranean and Black Sea (RCMMed&BS, 2009).

b) Selection of mètier to sample

In order to identify the métiers to be sampled, the ranking system described in the DCF (EU Decisions 949/2008 and 93/2010) has been applied. For sampling purpose, only the major mètier will be considered. Official statistics (landings and effort data) have been used to apply the ranking system. The methodologies to collect these informations are reported in chapter III.F (transversal variables).

The ranking system has been performed at GSA level using as reference the average values of the years 2007 and 2008. For each GSA (Level 4 Appendix I EU Decisions 949/2008 and 93/2010), as reported from the SGRN 08-01, the metier cells have been first ranked according to their share in the total commercial landings (tons). Thereafter the shares have been cumulated, starting with the largest, until a cut-off level of 90% has been reached. All metiers belonging to the top 90% have been selected for sampling. The ranking has been repeated according to the total value of the commercial landings (euro) and repeated a third time according to the total effort (days at sea). For each GSA, mètier which not belong to the top 90% in terms of total effort, value or landing have been deleted from the sample plane. Selected métiers are reported in table III.C.1.

Overall, the results of the ranking system have identified, on a GSA level, 55 métiers[7] belonging to the top 90%. Metiers for which a derogation is requested are discussed under the Chapter "III.C.6 Derogations and non conformities".

c) Type of data collection

Sampling strategy in each GSA will be a mix of concurrent-at-sea and concurrent-at-landing site covering a one-year period.

The survey has been designed taking into account the spatial (GSA) and temporal (quarterly) variability in order to detect seasonal differences in the demographic structure and composition of the landings for different metier.

Standard table III.C.3 illustrates the different types of data collection strategy and the sampling schemes that will be used for the different mètier. Both Probability Sample Survey (B) and Non-Probability Sample Survey (C) will be used.

d) Target and frame population

The target population for the reference year will be the number of trips by metier of the previous year.

The frame population is a subsample of the target population: it will be a selection of fishing trips, mainly on spatial (GSA) and time stratification basis (quarterly) with measurements of the composition of the catch. The sampling frame is given in Table III_C_4.

Table III.C.3 gives information on the average total number of trips in the reference years (2007 and 2008) for the different mètier, the sampling scheme applied for the planned sampling strategy, and the expected no. of trips to be sampled.

e) Sampling stratification and allocation scheme

The sampling will be accomplished according to the methods of a *two-stage stratified random sampling* (Cochran, 1977) and carried out monthly with strata represented by a combination of geographical sub-areas (GSA) and metier.

The sampling unit belonging to the metier (primary unit) will be the fishing trip (secondary unit). The number of fishing days to be sampled has been defined proportionally to the effort (number of days at sea for each metier) and the landings. According to the Commission Decision, the minimum number of fishing days to be sampled will be at least 1 fishing day per month during the fishing season. Thus at least three samples will be collected in each quarter for each metier. The number of fishing-days that will be sampled by metier and GSA is reported in Table III.C.3.

The Italian fishing fleet is characterized by vessels of small or medium size that perform their activities mostly within the 12 miles. Fishing trips last about 12 hours on average. This is true for most of fleet in all geographical areas. Bigger trawlers (> 24 meters) mostly in the Adriatic Sea (GSA 17) and Sicily Channel (GSA 16); long liners > 24 meters mostly in Ionian Sea (GSA 19) and Sicily Channel (GSA 16), practice fishing trips of more than one day. All these vessels represent only 2% of the total fleet. In these cases, the biological sample will consider this specificity and the fishing trip will be sampled accordingly. Table III.C.4 summarise the sampling effort that have been adopted for the different fishing activities.

[7] 3 metiers have been selected for sampling the large pelagic species

For each species, the total number of planned specimens to be collected is reported in Table III.C.5. Such number is proportional to the landing obtained in the previous years.
Below is reported the detail, related to mètier variables acquisition, for the major groups of species:

- demersal and small pelagic species
- large pelagic species
- sharks

DEMERSAL AND SMALL PELAGIC SPECIES

A total of 51 métiers, at GSA level for the sampling program of demersal and small pelagic species, have been identified. The selected metier are reported in table III.C.1.

The bottom trawlers (OTB_DES_>=40_0_0 and OTB_MDD_>40_0_0) have been selected in all GSA. They represent the most important segment of the Italian fleet in terms of production (around 37% of total landings). Vessel, operating mainly between 50 and 400 m depth, target hake (*Merluccius merluccius*), mullets (*Mullus barbatus* and *Mullus surmuletus*), nephrops (*Nephorps norvegicus*), cuttlefishes (*Sepia officinalis*) and pink rose shrimp (*Parapenaeus longirostris*).

Bottom otter trawlers for deep water species, OTB_DWS_>40_0_0, are represented by trawlers whose target species (over 40% of total landing) are represented by red shrimps (*Aristeus antennatus* and *Aristaeomorpha foliacea*). Those metier, operating between 400 and 800 m depth, have been selected in two GSA: 16 (Sicilian Strait) and 19 (Ionian Sea).

TBB_DES_0_0_0 is represented by rapido trawls operating mainly in the northern and Central Adriatic (GSA 17), catching common sole (*Solea vulgaris*), common cuttlefish (*Sepia officinalis*) and different species of benthic organism (e.g. Gastropoda).

Boat dredge for molluscs (DRB_MOL_0_0_0) performed their activity in the Adriatic Sea (GSA 17 and GSA 18) mainly targeting striped venus (*Chamelea gallina)* and other molluscs *(i.e. Donax clams, Gastropoda).*

Pots and traps for demersal species (FPO_DES_0_0_0), have been selected in GSA 11 (Sardinia) and GSA 17 (northern and Central Adriatic). Landing is represented by cephalopods (e.g *Sepia officinalis* and *Octopus vulgaris*) crustaceans (e.g *Squilla mantis*) and fish (mainly of the family Serranidae and Sparidae)

Small-scale fishery (using a combination of passive and mobile gears), represent the most relevant Italian fleet segment in terms of number of vessels, representing an high percentage of national total. Small scale fishery accounts for about a quarter of the national value of landings. The metier selected (GNS_DEF_>=16_0_0; GTR_DES_>=16_0_0; LLS_DEF_0_0_0), operating in almost every Italian GSA, target different species of fish crustaceans and cephalopods.

Hand and Pole lines for cephalopods (LHP_LHM_CEP_0_0_0) has been selected by the ranking system in the southern Tyrrhenian Sea (GSA 10) and in the GSA 19 (Ionian Sea). Ommastrephidae (e.g. *Illex coindetii*; *Todaropsis eblane*) the Loliginidae (e.g. *Loligo vulgaris*) and *Octopus vulgaris* are the main target of the selected metier.

The Midwater pair trawlers (PTM_SPF_>=20_0_0) performed their activity mainly in the Adriatic Sea (GSA 17 and GSA 18) and in the Sicilian Strait (GSA 16). The activity is concentrated mainly on small pelagic species, such as anchovies (*Engraulis encrasicolus*) and sardine (*Sardina pilchardus*).

Small pelagic species (*Engraulis encrasicolus* and *Sardina pilchardus*) are the main target also for the last metiers selected: the Purse seine PS_SPF_>=14_0_0. This metier operates in almost every Italian GSA, with the exclusion of the Sardinia seas (GSA 11).

Driftnets for small pelagic species (GND_SPF_0_0_0) have been selected in the GSA 10 (South Tyrrhenian Sea) and in the GSA 19 (Ionian Sea). In the South Tyrrhenian Sea landing is represented by small pelagic species (i.e. *Engraulis encrasicolus*, *Trachurus trachurus*) and some demersal such as *Boops boops*. The activity of Driftnets for small pelagic species (GND_SPF_0_0_0) in the GSA 19 is concentrated mainly on small pelagic species, such as *Engraulis encrasicolus* and *Sardina pilcardus*.

Fyke nets for demersal species (FYK_DES_0_0_0) have been selected in the GSA 17 (North Adriatic). Landing is represented by fish of the family Atherinidae and Gobiidae and by crustacenas Brachyura and Caridea.

Beach and boat seines for demersal species (SB_SV_DES_0_0_0) have been selected in the GSA 9 (Ligurian, North and Central Tyrrhenian Sea). The fishing activity is seasonally and concentrated mainly on the transparent goby *Aphia minuta*

The planned specimens, to be used for estimating annually the length frequency distribution (LFD), will be collected per year during a total number of 1270 sampling fishing days.
For all species both in Group 1 and Group 2 length data will be collected following the concurrent sampling approach.

Concerning discards, the survey will cover all the trawlers activity:
- bottom otter trawl for demersal species OTB_DES_>=40_0_0
- bottom otter trawl for deep water species OTB_DWS_>40_0_0
- bottom otter trawl for mixed demersal and deep water species OTB_MDD_>40_0_0
- pelagic pair trawl for small pelagic species PTM_SPF_>=20_0_0
- beam trawl for demersal species TBB_DES_0_0_0

For the above selected mètiers, discards will be monitored for the species in Group 1 and Group 2, as defined in the Commission Decision Chapter III. Data on the quantities of the target discarded species will be collected directly on board. The length composition by species will also be sampled in order to derive (LFD) by species and fishing day.

Following the output of the RCMMed&BS 2009 discards for

- Boat dredge for molluscs DRB_MOL_0_0_0
- Pots and traps for demersal species FPO_DES_0_0_0
- Fyke nets for demersal species FYK_DES_0_0_0
- Hand and pole line for cephalopods LHP_LHM_CEP_0_0_0

can be considered negligible (see report RCMMed&BS 2009 section 3.3.4 "*Sampling of discards*"). These mètier are to be considered not mandatory for discards sampling and MS need not ask for derogations.

Concerning the discards evaluation for the other mètiers, see section "III.C.6 Derogations and non-conformities (Mediterranean and black sea)".

LARGE PELAGIC SPECIES

The mètiers selected for the sampling program of large pelagic species, *Thunnus thynnus*, *Thunnus alalunga*, *Sarda sarda*, *Xiphias gladius*, *Coryphaena hippurus* and *Tetrapturus belone* (Istiophoridae), are:

- Purse seine (PS_LPF_0_0_0)
- Drifting longlines (LLD_LPF_0_0_0)
- Miscellanea (MIS, Harppon and FADs)

Miscellanea for large pelagics species (mainly harpoon and FADs) have not been selected for any of the parameters utilized (effort, landing, value). However, those gears (harpoon and FADs), have been picked up due to their peculiarity in the catches of *C. hippurus* and *T. belone*.

The sampling programme will follow the regional sampling approach as agreed during the RCMMed&BS 2009.

The survey to estimate length distribution will cover a one-year period. The planned number of specimens (set up at Regional level), to be used for estimating the LFDs, will be collected during a total number of 200 sampling fishing days (see table III.C.5 for disaggregation of sampling intensities among different species). The number of samples required to achieve the required precision level has been set and will be reviewed annually at Regional level.
Methodologies, following the ICCAT recommendations, and sampling strategies, will be applied at national/regional scale and will be related to temporal and spatial scale of the different fishing activity involved.

Data will be collected following the ICCAT recommendation to level 7 (see also RCMMed&BS 2009 and 2010 agreement).

Data will be reported as:

LLD_LPF_0_0_0 SWO

LLD_LPF_0_0_0 BFT

LLD_LPF_0_0_0 ALB

Discards for large pelagic species will be evaluated for the following mètiers:

- Purse seine (PS_LPF_0_0_0)
- Drifting longlines (LLD_LPF_0_0_0).

Following the output of the PGMed 2010 (Lisbona, 1/5-3-2010) discards for Miscellanea gears (harpoon and FADs) can be considered negligible. This mètier can be considered not mandatory for discards sampling and MS need not ask for derogations.

SHARKS

The Communication from the Commission to the European Parliament and the Council of 5 February 2009 on a European Community Action Plan for the Conservation and Management

of Sharks (EC-APCMS, COM-2009-40) has recommended the collection of reliable and detailed species-specific quantitative and biological data concerning commercial fisheries involving catches of Chondrichthyans (hereinafter referred to as "sharks").

To follow this issue the collection of sharks' biological variables (i.e. length frequency distribution) will be associated to the Metier-related variables following the concurrent sampling approach.

Most of the proposed species in the new Appendix VII (actually the list presents 47 species of sharks and rays) of the Commission Decision 93/2010, are rare and with a sporadic and not confirmed presence in the Mediterranean area. During the period 2011-2013 all shark species will be collected concurrently for length. No stock-based sampling will be added if metier based sampling fails to provide the appropriate precision for length distributions.

In table III.C.5 the list of species to be sampled will be included together with the required precision level. However some cells were labelled as N/A (Not Applicable) since the minimum number of specimens to be sampled to achieve the required precision cannot be known '*a priori'* since no length samples are available for these species. Furthermore since most of the species are rare the number of samples recorded will be extremely low and hence the minimum number to achieve the precision target cannot be reached.

III.C.2 Estimation procedures (Mediterranean Sea and Black Sea)

The LFD obtained quarterly, by GSA, will be raised to the quarter production by species and mètier, according to the proportion between sampled and landed weight. Then the total LFD will be obtained summing up the estimates by quarter and metier (Annex II: Methodology for the estimation of demography of landings, for demersal and small pelagic species, related to the biological metier related variables).

The precision, in terms of coefficient of variation (CV) of the length frequency distributions (LFDs) will be estimated annually by metier and GSA (see paragraph III.C.3 and III.E.3 **"**Data quality Mediterranean and black sea")

The method developed by Vigneau and Mahevas (2004), which allows estimating the precision, in terms of coefficient of variation (CV) for each length class and for the whole LFD at metier level, will be adopted. The estimates of precision obtained for each metier will be summed up for the metier targeting the same stock
For discards see Annex III (Methodology for the estimation of discards - Biological metier related variables).

III.C.3 Data quality evaluation (Mediterranean and black sea)

CV for length of demersal and small pelagic species included both in Group 1 and in Group 2 list, will be calculated annually for each metier and per each GSA.

CV for length of large pelagic species included both in Group 1 and in Group 2 list, will be calculated annually at Regional level following the RCMMed&BS and PGMed recommendations.

Concerning all the 47 species of sharks, we will try to associate to the collection of metier related variables (length data) a precision target at least for the most common species landed

at National scale. However the precision target requested will be difficult to achieve since most of the species are rare and the number of samples recorded will be extremely low. RCMMed&BS 2009 noted that "*the sampling of sharks in the routine concurrent sampling schemes, poses a number of problems for certain metiers. The sampling of just a few shark individuals in these metiers, forces to largely increase the sampling effort, and decrease significantly the efficiency of the sampling for commercial species. It is also stressed that no precision target could be reached for Elasmobranches.*"

Annex IV (Methodology to estimate the precision level – Metièr related variables) describes in details the methodologies for calculation of precision levels.

Applicability of the COST project (Vigneau, 2008) and the output of the Workshop on the Implementation of the Common Open Source Tool (Nantes, 13-16 April 2010) will be verified.

III.C.4 Data presentation (Mediterranean and black sea)

The reference period will be one year. The results pertaining to the period January – December 2011 will be ready by the end of June 2012. Some preliminary data (referring to the metier related variables collected during the first and second quarter of 2011) will be produced by the end of April 2012. The same time lag will be applied for data referred to the years 2012 and 2013.

Potential sources of bias may be:

- No access to some vessels; thus the sampling population could not be the same as the target population, causing potential bias. Mitigation measures include the documentation of non-accessibility to such vessels to be able to account for it.
- Difficulty to obtain monthly data; this problem could arise due to weather conditions. It is very difficult to have mitigation measures for this issue. However, whenever vessels operate in months where fishing due to bad weather is very limited, there is an increased effort to sample the boats.

III.C.5 Regional co-ordination (Mediterranean and black sea)

Activities to be carried out in cooperation with other European Mediterranean countries will be discussed during the RCMMed&BS.

In particular, regional sampling program for large pelagics species will be annually monitored and evaluated.

In the following table, recommendations from RCM and responsive actions are reported:

Recommendation	Responsive actions
RCMMed&BS recommends MS to follow the sampling intensities as detailed in table 3.3.3.c for bluefin tuna and in tables 3.4.2.a-d for other main large pelagics, and to achieve the required minimum number of samples as stated by the Decision EC 949/08. RCM recommends also that the data gathered in 2009 should be provided to the PGMed meeting of 2010, to	In PGMED 2010, the number of length samples for all large pelagics was revised and the Mediterranean countries agreed to sample regionally for length. Italy will follow this Regional agreement.

allow PGMed to calculate the precision level and the minimum number of samples required to achieve the required precision, MS to adjust their NP 2011-2013 accordingly.

III.C.6 Derogations and non-conformities (Mediterranean and black sea)

The results of the ranking system have identified, on a GSA level, 55 métiers[8] belonging to the top 90%. Derogation is requested for the following métier:

- Hand and pole line for cephalopods (LHP_LHM_CEP_0_0_0) in the GSA 19

This métier has not been selected both for landing and for economic values. It has been picked up only for the contribution in terms of effort. The low incidence in terms of fishing days and catches, the seasonality and the high fragmentation of these activities along the Italian coasts, and the low importance also in terms of economic values of the catches, would result a difficult and expensive exercise.
However, as suggested by SGRN in 2011, Italy will conduct a pilot study to investigate the fishing behaviour of this metier (LHP_LHM_CEP_0_0_0) in the GSA 19

Derogation is also requested for the collection of discards data related to some métier. The métier that the discards sampling program will not cover are:

Set gillents for demersal fish	GNS_DEF_>=16_0_0
Drifnets for small pelagic fish	GND_SPF_0_0_0
Set trammel nets for demersal species	GTR_DES_>=16_0_0
Purse seine for small pelagic fish	PS_SPF_>=14_0_0
Beach and boat seine for demersal species	SB_SV_DES_0_0_0
Set longlines for demersal fish	LLS_DEF_0_0_0
Set gillnets for small and large pelagic species	GNS_SLP_>=16_0_0

As suggested by SGRN, in 2011, Italy will conduct a pilot study to investigate the discards behaviour of these metiers (even not selected by the ranking system).

The object of the study will be to estimate the discard, in weight and number of the species G1 and G2 in the Appendix VII (Commission Decision 93/2010 and 949/2008).

In the pilot study on the discard behaviour Italy will follow the output of the RCMMed&BS 2010 (Varna, Bulgaria 2010):

Recommendation	Responsive actions
RCMMed&BS recommends that Table 7 is used to provide justification for not sampling certain métiers. This justification could be	The RCMMed&BS reviewed the work conducted by the PGMed 2010 in order to create a regional view of the

[8] 3 metiers have been selected for sampling the large pelagic species

based in the discards behaviour or in the non selection of métier in the regional ranking system	métier important to sample for discards and on providing scientific justification for not sampling certain métiers. RCMMed&BS was agreed with the outputs and the integration made by the PGMed 2010, and recommends to strictly following the proposed table made up in the PGMed sampling métiers for discards. Italy will follow this regional agreement.

REGION: Other regions where fisheries are operated by EU vessels and managed by RFMO's to which the Community is contracting party or observer (e.g. ICCAT, IOTC, CECAF)

III.C.1 Selection of métiers to sample and data acquisition (Other regions)

Italian fleet operating outside community waters consist of 18 vessel (official registered) corresponding to around 0.2% of the total fleet (vessel number), 5.7% in terms of GT and 2% in terms of kW.

10 of the above mentioned vessels since the beginning of 2009 are inactive. 7 operate around Mauritania (Area FAO 34.1 e 34.3) waters and 1 in the Guinea Conakry (Area FAO 34.3). All vessels operate with trawls nets targetting demersal species.

Another vessel, which operated in the Indian Ocean (Area FAO 51) catching large pelagic species (Albacares and Skipjack tunas) with purse seine, is not more under Italian flag.

According to the decision for the application of the Regulation (CE) 199/2008 the collection of statistics on catches (landings / fishing log books) and on length frequencies of the species, caught by fleet operated outside community waters, will be carried out. Up to know no biological sampling has been implemented for these vessels, they only have a national obligation to fulfil logbooks.

Italy participated to the first RCM for “Other regions” (RCM for long distance fisheries, Madrid, March 2010) and related recommendations will be considered.

III.D Biological - Recreational fisheries

REGION: Mediterranean Sea and Black Sea

III.D.1 Data acquisition (Mediterranean Sea and Black Sea)

Collection of data from recreational fisheries will cover *Thunnus thynnus* (bluefin tuna) and *Anguilla anguilla* (eel) (species listed in Appendix IV, 4 EU Decision 93/2010).

Concerning *A. anguilla*, information on the distribution and stock management have greatly increased during 2009, since the recovery plan foreseen by Council Regulation (EC) 1100/2007 has been prepared and sent to the Commission for discussion and eventual endorsement; information on eel stock, eel fishing on the different life stages, general situation of the stock have been very useful also for the preparation of the pilot study includes in the previous National Program. Further developments regarding data collection on eel recreational fisheries shall possibly issue from the definitive application of the Italian Eel Management Plan under Regulation 1100/2007. The Plan has been submitted to the approval of the Commission September 30rd, 2009. The Plan has been considered eligible by the Commission, and is currently undergoing the scientific evaluation from ICES.

a) Type of data collection

Recreational fisheries of bluefin tuna

Data collection on recreational and sport fisheries of bluefin tuna from 2011 to 2013 will follow the methodologies suggested by the pilot study previously carried out within the 2004 National Program.

Applied methodologies will assure the estimate of total catches and total fleet, as well as biological information on length composition of catches.

The main basis of the data collection program is the overall census of recreational and sport fisheries that will cover all the Italian coast line and islands that means more than 800 ports and landing sites. Starting from this census, information on fleet and type of activities and an estimate of catches per administrative region will be obtained.

The census will take place in the last two months of each year and it will require about 30 days of activity.

Recreational fisheries of eels

Data collection on recreational fisheries of eel from 2011 to 2013 will follow the methodologies suggested by the pilot study previously carried out within the 2009/2010 National Program.

The pilot study approached the problem of lack of information on eel recreational fishing, through an extensive review of available information and “on the field” data; bibliographic references have been reviewed, legislative framework for eel recreational fishing has been detailed, and data from the Fishing offices of the local Administrations (Province) have been elaborated. Data from “Hycthyological maps” for information on eel distribution have also been considered.

On the basis of all the available information collected, the pilot study has been carried out through interviews to an extensive sample of fishermen of the main fishermen association and the local sections of the association (present in over 60 Italian local administrative bodies, the “Province”).

The study has clarified the distribution of the activity, apparently concentrated only in some regions and not widely distributed, either because not traditionally interesting or because the species is scarcely abundant or not present.

b) Target and frame population

Recreational fisheries of bluefin tuna

The target and frame population are the same: represented by all vessels of recreational and sport fisheries. The main basis of the data collection program is the overall census of recreational and sport fisheries that will cover all the Italian coast line and islands that means more than 800 ports and landing sites.

Recreational fisheries of eels

The target population are the main fishermen association and the local sections of the association (present in over 60 Italian local administrative bodies, the "Province"). The frame population is represented by a sample of the population members randomly selected.

c) Data sources

Recreational fisheries of bluefin tuna

Tournaments of sport fishery will be monitored directly through collaboration with the national clubs (FIPSAS, EFSA, Big Game Italia) that are committed to providing all the data on the tournaments organised by them. In these tournaments, length measurements will also be collected.

Recreational and sport fisheries will also be monitored in one port for each coast side where this type of fishery is practised. Interviews, through forms specifically prepared (see the Table D.1 below), will be carried out directly from local fishermen in order to acquire information on catch composition and seasonal distribution of the catches.

Official declarations of blue fin tuna catches will also be monitored through an analysis of declarations delivered to the port authorities and the authorisation required at the beginning of the season.

Recreational fisheries of eels

Methodology foresees the implementation of interviews through forms specifically prepared. The sample will be statistically chosen according to the relative importance of the different regions. Sample data will be raised to the universe of the Italian fishermen practising eel fishing through statistical techniques. Interviews, according to what requested by Reg 199/2008, will give information on the distribution and consistency of the activity, different gears used and on reasonable estimate of the amount of catches.

Table D.1: Form for the interviews (Tuna-Recreational fishery)

BASE PORT												
OWNER NAME	**BOAT NAME**						**LENGHT**				**HP**	
NUMBER OF FISHING TRIPS												
MONTH	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEPT	OCT	NOV	DEC
DRIFTING												
TROLL												
OTHER (SPECIFY)												
NOTES												
CATCHES												
MONTH	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEPT	OCT	NOV	DEC
CATCHES IN NUMBER												
Other species (specify)												
Other species (specify)												
Other species (specify)												
Other species (specify)												
Other species (specify)												
CATCHES IN WEIGHT												
Other species (specify)												
Other species (specify)												
Other species (specify)												
Other species (specify)												
Other species (specify)												
NOTES												
codes other species: ALB = Albacore; **MSP** = istiophoridae; **SWO** = swordfish **LTA** = Atlantic black skipjack; **FRZ** = Frigate tuna ; **DOL** = dolphin fish; **SHK** = shark **BON** = bonito; **OTH** = other species												

III.D.2 Data quality

Recreational fisheries of bluefin tuna

As required by the regulation, data related to annual estimates of the catches in volumes will lead to a precision of level 1. Data on capacity will be collected exhaustively.

Continuity and consistency of the series collected is assured by the methodology applied that will be the same as in previous programs.

Recreational fisheries of eels

As required by the regulation, data related to annual estimates of the catches in volumes will lead to a precision of level 1. A statistical survey started in 2010 and only from this year precision levels will be calculated.

III.D.3 Regional co-ordination

The intention of the RCMMed&BS is to adopt for recreational fisheries an approach in order to better define useful data collection to carry out and to avoid unnecessary works whether this activity not really being a real targeted practice. The opportunities of sampling recreational fisheries may differ quite substantially between countries and also may change rapidly in the short term (i.e. season closure for BFT, eel fishing banned because of environmental reasons).

Italy will consider all the future recommendations on recreational fisheries, monitoring eel and BFT, coming out from the RCMMed&BS.

III.D.4 Data presentation

Both for bluefin tuna and for eel the reference period will be one year. The results pertaining to the period January–December 2011 will be ready by the end of June 2012. The same time lag will be applied for data referred to the years 2012 and 2013.

III.E Biological - stock-related variables

REGION: Mediterranean Sea and Black Seas

III.E.1 Data acquisition (Mediterranean Sea and Black Sea)

a) Selection of stocks to sample

Biological sampling, for the acquisition of stock related variables (i.e. age, weight, sex and maturity) will be carried out by scientific institutions designated by the Italian Administration separately for each GFCM Geographical Sub-area (GSA) and by major groups of species (demersal and small pelagic species, large pelagic species and sharks).

Within its assigned area, each scientific institution will be required to carry out biological sampling and data analysis and processing, and to forward the results to the Administration in accordance with a standard general procedure laid down for all Italian areas.

For all species included in the sampling scheme, stock variables will be collected according to the requirements of the Appendix VII (i.e. on an annual or three-annual basis) of the EU Decision 93/2010.

Table III.E.1 identify which stocks are going to be included in sampling scheme. Below is reported the detail, related to stock variables acquisition, for the major groups of species:

- demersal and small pelagic species
- large pelagic species
- sharks

DEMERSAL AND SMALL PELAGIC SPECIES

Sampling scheme for demersal and small pelagic species will consider all Group 1 species plus species of Group 2. In detail it will consider 17 teleosteans (i.e. *Merluccius merluccius, Mullus barbatus, Mullus surmuletus, Solea vulgaris, Engraulis encrasicolus, Sardina pilchardus, Trachurus trachurus, T. mediterraneus, Lophius boudegassa, L. piscatorius, Boops boops, Spicara smaris, Pagellus erythrinus, Micromesistius poutassou, Eutrigla gurnardus, Trigla lucerna* and *Scomber* spp.), 6 cephalopods (i.e. *Sepia officinalis, Octopus vulgaris, Eledone cirrhosa, E. moschata, Loligo vulgaris* and the *Illex coindetii*) and 6 crustaceans species (i.e. *Aristeus antennatus, Aristaeomorpha foliacea, Parapenaeus longirostris, Nephrops norvegicus, Penaeus keraturus* and *Squilla mantis*).

For all demersal and small pelagic species included in the sampling scheme, stock variables (age, weight, sex and maturity) will be collected according to the requirements of the Appendix VII of the EU Decision 93/2010 (i.e. on an annual or three-annual basis). An overview of the stocks included in the sampling scheme is presented in Table III.E.1. For this group of species, landing values and % share in EU landing, have been obtained from the PGMed report (Sete, 2-6 March 2009)

LARGE PELAGIC SPECIES

Evaluation of the stock variables (age, weight, sex and maturity) for large pelagic species will be evaluated every three years following the Appendix VII of the Commission Decision 93/2010.

The sampling program, including large pelagic species both Group 1 (*Xiphias gladius*, *Thunnus thynnus*) and Group 2 (*Thunnus alalunga*, *Sarda sarda*, *Coryphaena hippurus*), will follow the regional sampling approach as agreed during the RCMMed&BS 2009 and PGMed 2010 (Lisbon, 1-5 March). Table III.E.1 reports the landing values and % share in EU landing obtained from the PGMed report (Sete, 2-6 March 2009).

SHARKS

Evaluation of the stock variables (age, weight, sex and maturity), from fisheries based sampling data for shark's species, will be evaluated every three years only for *Raja clavata* (landing for this species is around 328 tons). All the other sharks' species are less than 200 tons of national landings (see section III.E.5 "Deroagtion and non Conformities"). Table III.E.1 reports the National landing values for some sharks species for the 2008.

Although we are not obliged to collect stock related variables from fisheries based sampling, since the landings of all the others shark species are less than 200 tons, data (i.e. weight, sex and maturity) will be collected anyway during scientific survey (i.e. for the Mediterranean the MEDITS and the MEDIAS surveys).

b) Type of data collection

Sampling strategy in each GSA will be a mix of concurrent-at-sea and concurrent-at-landing site covering a one-year period.

Strata, as for Biological-Metier-related variables of the NP, are represented by a combination of geographical sub-areas (GSA) and métier. The sampling unit is the fishing day (corresponding to the fishing trip). Stock related variables (sex, weight, and maturity) will be collected during different fishing days taking into account the spatial (GSA) and temporal (quarterly) variability in order to detect seasonal differences in the demographic structure and composition of the species examined.

Data will be collected mainly from commercial fisheries (i.e. landing, on board sampling). Following each stratum, both "Probability Sample Survey (B)" and "Non-Probability Sample Survey (C)" will be used.
Survey data will be used (if the survey is conducted in the correct period) to integrate information on growth and reproduction parameters.

c) Target and frame population

The population will be defined as the total number of fish landed. The frame and target population are the same. They refer to the stocks within the geographical boundaries (GSA).

Italy will collect biological data on the species above mentioned following the temporal stratification (yearly or three-annually) reported in the Appendix VII (Commission Decision 93/2010 and 949/2008). Following the SGNR suggestions, Tables III_E_1, III_E_2 and III_E_3 have been modified accordingly.

The population is the universe of total trips carried out by commercial vessels over a period of one year.

The frame population of the sampling survey does not exist because there is not an exhaustive list of all trips that will be carried out for each metier. The logbook cannot be considered a source to compile such list because only a part of the trips are registered in the logbook.

However, the sampling scheme defines the total number of fishing trips to be sampled for each metier and these trips are selected randomly on the basis of a calendar which assigns a consecutive number to each operative fishing day.

d) Sampling stratification and allocation scheme

The sampling will be accomplished according to the methods of a *two-stage stratified random sampling* (Cochran, 1977) and carried out monthly with strata represented by a combination of geographical sub-areas (GSA) and metier.

The sampling unit belonging to the metier (primary unit) will be the fishing trip (secondary unit. The number of fishing days to be sampled has been defined proportionally to the effort (number of days at sea for each metier) and the landings. According to the Commission Decision, the minimum number of fishing days to be sampled will be at least 1 fishing day per month during the fishing season. Thus at least three samples will be collected in each quarter for each metier.

For each species, the total number of planned specimens to be collected is reported in Table III.E.3. Below is reported the detail, related to stock variables, for the major groups of species:

- demersal and small pelagic species
- large pelagic species
- sharks

DEMERSAL AND SMALL PELAGICS SPECIES

Table III.E.2 gives an overview of the long-term sampling strategy for age, weight, sex ratio and maturity, with respect to selected stock. Planned sampling for the estimation of biological parameters is also reported in table III.E.3. Number of planned specimens is proportional to the landing obtained in the previous years (2007-2008).

Biological parameters (i.e. length/weight relationship; length at age; length at maturity; sex ratio) for demersal and pelagic species will be estimated both from commercial fisheries (i.e. landing, on board sampling) and surveys. All the parameters collected, sex ratio, maturity and weight will be referred to length.

LARGE PELAGIC SPECIES

Table III.E.2 gives an overview of the long-term sampling strategy for age, weight, sex ratio and maturity, with respect to selected stock.

Stock variables (age, sex, weight, maturity) for large pelagic species will be collected from commercial fisheries (i.e. landing, on board sampling). All the parameters collected, sex ratio, maturity and weight will be refered to length.

The sampling intensity for the estimation of biological parameters (i.e. length/weight relationship; length at age; length at maturity; sex ratio) and the number of specimens to be sampled (see table III.E.3 - from the PGMed analysis, 2010), including both species in Group 1 and in Group 2 list, has been agreed on a Mediterranean-wide basis (RCMMed&BS, 2009).

Methodologies, following the ICCAT recommendations, and sampling strategies, will be applied at national/regional scale and will be related to temporal and spatial scale of the different fishing activity involved.

SHARKS

Concerning sharks, as already stated above, we are not obliged to collect stock related variables, since the landings of almost all sharks species are less than 200 tons. However data

related to weight, sex and maturity will be collected during scientific survey (i.e. for the Mediterranean the MEDITS and the MEDIAS surveys). For the species *Raja clavata* data will be collected also during commercial fisheries (Table III.E.2 and Table III.E.3).

III.E.2 Estimation procedures (Mediterranean Sea and Black Sea)

Methods that will be used for estimating the stock related variables (e.g. sex, maturity and weight) and the related parameters (e.g. growth and natural mortality, sex ratio, gonad-somatic index, length/age at maturity) are reported in the Annex V (Methodology to estimate the stock related variables and the growth and reproduction parameters).

III.E.3 Data quality evaluation (Mediterranean Sea and Black Sea)

The estimation of precision levels will be performed according to the methods described in Annex VI (Methodology to estimate the precision level – Stock related variables) and to the requirement of the Appendix VII of the EU Dec. 93/2010 (i.e. on an annual or three-annual basis).

CV for age of demersal and small pelagic in G1 species will be calculated annually per each GSA and for all metiers combined. CV for age of demersal and small pelagic in G2 species will be calculated every three years per each GSA and for all metiers combined.

CV for age of large pelagic species included both in Group 1 and in Group 2 lists, will be calculated every three years at Regional level following the RCMMed&BS and PGMed recommendations.

Precision levels of other biological parameters (i.e. length/weight relationship; length at age; length at maturity; sex ratio) for demersal and small pelagic in G1 species will be calculated annually per each GSA and for all mètiers combined.

Precision levels of other biological parameters (i.e. length/weight relationship; length at age; length at maturity; sex ratio) for demersal and small pelagic species in G2 will be calculated every three years per each GSA and for all mètiers combined.

Precision levels of other biological parameters (i.e. length/weight relationship; length at age; length at maturity; sex ratio) for large pelagic species, included both in G1 and in G2 list, will be calculated every three years at Regional level following the RCMMed&BS and PGMed recommendations .

Concerning the sharks species (*Raja clavata*), the precision level associated to the collection of stock variables (i.e. sex, age, weight and maturity), will be calculated every three years. As stated during the RCMMed&BS 2009 *"It is also stressed that no precision target could be reached for Elasmobranches. Therefore, no minimum number or sampling strategy should be associated to the collection of all the "sharks" species reported in the new Appendix VII."*

Applicability of the COST project (Vigneau, 2008) and the output of the Workshop on the Implementation of the Common Open Source Tool (Nantes, 13-16 April 2010) will be verified.

III.E.4 Regional co-ordination (Mediterranean Sea and Black Sea)

Activities to be carried out in cooperation with other European Mediterranean countries will be discussed during the RCMMed&BS.
In particular, regional sampling program for large pelagics species will be annually monitored and evaluated.

III.E.5 Derogations and non-conformities (Mediterranean Sea and Black Sea)

Derogation is asked for Mugilidae, for which was already given derogation in 2009. It is a complex of species and it is mostly caught in lagoons. SGRN have found the reason for not sampling this species group for biological parameters acceptable and has suggested that the derogation to be prolonged.

All sharks' species, with the only exception of *Raja clavata*, are less than 200 tons. Following the Commission Decisions 949/2008 and 93/2010 which stated that "*a Member State may exclude the estimation of the stock related variables for stocks for which the landing is less than 200 tonnes*", derogation is requested for the collection of the "Stock variables" related to sharks.

Italy recalled that, if Mediterranean MSs do not provide landing by species, and not by group of species or at genus level, the exemption rule fixed by DFC "*for stocks in the Mediterranean Sea, the landings by weight of a Mediterranean Member State for a species corresponding to less than 10 % of the total Community landings from the Mediterranean Sea* (Commission Decision 93/2010)" cannot be determined.
As recalled by the RCMMed&BS (October, 2009), we stress the importance of providing landings data by species, as required by the DCF (EC Decisions 949/2008 and 93/2010), and not by group of species (based also on the exercise "*Sampling for mixture of species in the landings*" carried out in 2008).

Other regions where fisheries are operated by EU vessels and managed by RFMO's to which the Community is contracting party or observer (e.g. ICCAT, IOTC, CECAF…)

III.E.5 Derogations and non-conformities (Other regions)

Considering that for stock related variable a self sampling is not feasible because biological experts are needed, derogation is required for the years 2011-2013. In any case, Italy has partecipated to the first RCM for "Other regions" (Spain, March 2010) and related recommendations will be considered.

III.E.6 Monitoring of commercial eel

The following paragraph describes the monitoring of commercial fisheries for eel and the activities planned for 2011-2013

Premise: general framework and on-going actions on eel in Italy
Eel (*Anguilla anguilla* L.) exploitation in Italy has a long standing tradition, and it concerns all continental stages, i.e. glass eel, yellow and migratory silver eel. The most distinctive exploitation pattern for eel in Italy is coastal lagoon fishery that yields most of yellow and silver eel extensive culture (a practice of sustained fishery). Inland eel fisheries are found in main rivers and lakes.

EC Regulation 1100/2007 requires MS to implement a national management plan for eel recovery. The Italian Eel Management Plan has been prepared by the Directorate General of Sea Fisheries and Aquaculture, Ministry of Agriculture, Food and Forestry, with the support of experts from the scientific community and representatives of the Regions. The work was in fact conducted jointly with the Regional administrations, with a series of meetings during the 2008/9, in order to coordinate activities. The final version of the document was submitted to the Commission on September 30rd, 2009, and is still under evaluation.

Within the Italian Eel Managent Plan (EMP), a preliminary outline of the situation in Italy for eel was given in order to establish a scenario of interventions articulated at the national and regional levels so to comply with the requirements of the Regulation 1100/2007. However, in the document it has been pointed out that the Management Plan should be seen as a document liable of updating and improvement, in relation to the fact that the management of eels in Europe for the restoration of the stock is a process still under development. In fact, the process of implementing the management plan, along with activities provided by other instruments (Directive 2000/60, the Habitats Directive, and Council Regulation 199/2008) will lead to a large influx of data, information and assistance.

The National Plan takes into account the complex reality of Italy. The eel is present in the lagoon waters and inland waters of all regions, but the density, population, the increase is very variable depending on the type of environment (lagoons, rivers, lakes) and thus the pattern of production that results is extremely heterogeneous. Heterogeneous is also the administrative competence on eel fisheries: sea fishing and fishing in river mouths are attributable Central Administration (Ministry of Agriculture, Food and Forestry - Directorate General for Fisheries and Maritime aquaculture), while jurisdiction on inland fisheries is attributable to the Regions, including fishing for eels, with the power conferred on January 15 n.11 Presidential Decrees 1972 and 24 July 1977, no 616. For all these reasons, the Eel National Plan for Anguilla in Italy is a joint Plan. A framework for action at national level for coastal waters and for those regions (total 11) who preferred to delegate to the Central Administration the eel management is envisaged. Conversely, 9 regions have developed specific Regional Plans. The decentralization of responsibilities is a factor that may partly result in a delay in the implementation of such measures, but the presence of an Eel Management Plan approved by the European Commission in Italy can be an effective part in support of the implementation process at various levels.

At the present moment, Italy is waiting for the outcome of the scientific evaluation and for the definitive answer of the Community. Possible changes could be necessary in the Plan to achieve the final approval.

<u>Ongoing actions within eel data collection 2009-2010 and activities for 2011-2013</u>

On the basis of what has been said in the previous paragraphs, it is clear that the eel management, as well as the eel data collection, are at the present moment not entirely definite, and shall possibly be amended, in relation to the outcome of ongoing actions.

In the Data Collection Program 2009-2010, a strong effort for the methodological setting up of eel commercial fisheries data collection has been underway, whose outcomes shall be available by the end of 2010. A number of problems arise when dealing with eel, whose fisheries is carried out nearly exclusively in inland waters, and hence under Regional control, and because of the heterogeneity of fishing typologies (lake fisheries, river fisheries, artisanal fishery in lagoons, sustained fishery in lagoons, glass eel fishery). One main problem for eel data collection relies in the identification of management units, and a list of authorities responsible, as required by Regulation 1100. These should require an inventory of individual

river basins, as designated under Directive 2000/60/EC. To meet this point, an analysis was carried out within the Italian EMP. In relation to the high fragmentation of responsibilities in Italy, the adoption of rules implementing the plans for the management of eels, if established on the basis of the river basins units o as specified in Article 2 (1) Regulation (EC) No 1100/2007, would have made difficult the effective coordination, implementation and monitoring of the measures. Italy therefore has availed itself of the the opportunity provided in the same art. 2 of the Regulation (under which in presence of adequate justification, a Member State may designate the whole or any administrative unit as the eel management unit) and has suggested the Regions as eel management units.

In the course of the Data Collection Program 2009/2010, the effort has concentrated, as mentioned above, on the setting up of an appropriate methodology. On the basis of a census of all eel commercial fisheries carried out in coastal and inland waters, i.e. coastal lagoons, lakes and rivers, a statistical procedure has been set up that has allowed to define the sample of fishermen to be used for evaluations. The methodology is currently being verified, and shall be applied for the year 2010 data collection.

For the period 2011-2013, on the basis of what has been described above, for each management unit (i.e. Region) and relevant fishing typology within the Unit (river, lake, lagoon, managed lagoon) the following basic parameters shall be collected, by direct interviews on the statistical sample:

A) Boats involved in eel fishing in each management unit in order to estimate fishing capacity.

To determine the fishing capacity, the following parameters will be evaluated:

- total number of boats,
- average value of tonnage or power engine in kW

B) Estimate of fishing effort.

The fishing effort will be evaluated by census of specific fishing techniques (type, possibly gear size and mesh size), and of fishing days dedicated to eel.

C) Evaluation of landings by life stage (G, Y, S) of all eel catches.

All data will be recorded specifically for each life stage. The feasibility of the application of logbooks to be entrusted to a subsample of fishermen in selected sites is currently under evaluation.

The program for 2011-2013 shall also consider:

D) Biological samplings of catches are foreseen to evaluate their composition by age and by length, with the aim to obtain information on the demographic structure of the eel local stocks and on their growth, sexuality and mortality.

Samplings shall be carried out on a management unit/tipology basis, and shall foresee length, weight and livery stage assessment (yellow/silver) on a sample of significant size, on a yearly basis. Age estimation by otoliths examination and sex ratio evaluation by histological examination will be performed on the third year.

Biological samplings schedule shall possibly undergo modifications, by taking into account coordination within international actions (ICES Working Group on intercalibration for Age Reading for European and American Eel, Ices Study Group on Anguillid Eels in Saline Waters, ICES/EIFAC Working Group on Eels).

III.F Transversal variables

REGION: Mediterranean Sea and Black Seas

III.F.1 Capacity

III.F.1.1 Data acquisition

The following parameters will be given for capacity estimation:

- number of vessels

- GT, kW, Age (as defined in Regulation (EC) No 26/2004)

Parameters will be given annually, per fleet segments (Appendix III) and per supra-regions (Mediterranean & Black Seas and Other regions).

The basic data source will be the fleet register at the 1st of January.

Further analyses are necessary to identify the prevalent fishing technique as required by Appendix III of the DCF. According to Commission Decision, the dominance criteria shall be used to allocate each vessel to a segment based on the number of fishing days used with each gear. If a fishing gear is used by more than the sum of all the others (i.e. a vessel spends more than 50% of its fishing time using that gear), the vessel shall be allocated to that segment. If not, the vessel shall be allocated to the following fleet segment:

- "Vessels using Polyvalent active gears" if it only uses active gears;

- "Vessels using Polyvalent passive gears" if it only uses passive gears;

- "Vessels using active and passive gears"

More than 70% of the Italian fishing-vessel licences allow the use of more than one fishing system. In these cases the existence or otherwise of actual polyvalent activity have to be verified.

In order to get this information on the prevalent fishing activity, field surveys have been carried out periodically since the implementation of the DCF and will be updated every quarter. In particular, the following activities will be implemented:

- analysis of information on available logbooks
- quarterly analysis of the vessel register in order to verify fishing techniques
- comparison with previous defined structures of the fleet
- for vessels that did not move to another ports, the fishing technique is confirmed
- for vessels moved to another port (around 150 vessels, each quarter), information on the fishing technique will be verified
- all new records of the vessel register (around 70 vessels, each quarter) will be checked through interviews.

This survey involves all the vessels in the fleet register, including those less than 12 meters.

III.F.1.2 Data Quality evaluation

The procedure illustrated in the previous paragraph guarantees the quality of the data collected. Information on the prevalent fishing technique is gathered extensively for the whole fleet. A data collection network covering the whole Italian coast is used. This network covers the whole Italian coast. Data collectors are people working in the fishery sector and well integrated with the vessels' owners and the crew from whom they get the required information.

A specific database has been implemented in previous years to manage all the information regarding the fleet. This database allows for validation and cross checking of data related to each single vessel in the fleet.

III.F.2 Effort

III.F.2.1 Data acquisition

Effort will be estimated according to the variables and disaggregation levels listed in Appendix VIII.

According to RCMMed&BS recommendations (based on SGRN and PGMED suggestions) effort data will be collected only for metiers selected by the ranking approach.

The major problem in respecting the required disaggregation is given by the disaggregation per metier (Level 6 of Appendix IV, 4). Compulsory information for all the fleet is not available. In Italy, there is no obligation for vessels to indicate the type of activity they practice along the year. Each vessel can use all the gears indicated in the licence. More than 70% of the Italian fishing-vessel licences allow the use of more than one fishing system, and therefore gears. This framework is also more complex if we consider the high number (24) of existing métiers actually practised by Italian vessels, with differences in seasonality and geographical areas.

In order to estimate fishing effort per metier and GSA, different data sources will be used:

- logbook information, when available. Logbook is compulsory in the Mediterranean only for a minority of fleet (vessels of overall length exceeding 10 metres) and will contain production information only where vessels retain on board quantities exceeding 15 kg live-weight equivalent of the species included in a specific list (Appendix VII of Regulation (EC) No 2737/1999).
- field survey to detect the prevalent fishing activity (see paragraph II.F.1.1)
- sample survey to estimate the monthly distribution of activity by métiers.

The sample survey is based on a panel of around 1500 vessels (10% of the fleet), including the small scale (vessels < 12 m). Survey takes place every week on a continuous basis. Data on fishing effort, vessel activity and fishing area are recorded by gear and species using purposely formulated questionnaires. Results for each area, by month and by metier are obtained by applying raising factors to the sampled data (see Annex VII for details on the methodology).

III.F.2.2 Data quality evaluation

Elementary and aggregated data will be checked to verify their reliability. Consistency among different variables will also analyse.

Precision levels will be calculated in terms of CV at the level of fleet segments and per GSA (see Annex VII).

III.F.2.3 Data presentation

Effort data will be available with a time lag of 6 months (for instance data referred to January 2011 will be available not before July 2011).

III.F.2.4 Regional coordination

In the next RCMMed&BS, Italy will propose to discuss the methodologies used by different countries, especially for vessels < 12 m, in order to standardise them and increase the level of comparability.

III.F.2.5 Derogations and non-conformities

Regarding the variable "number of fishing operations" for the estimation of effort of purse seines, Italy asks for derogation, as already required in the previous NP (2009-2010). This information is not required in the logbooks and it is also very difficult to be obtained through questionnaires. Fisherman does not record this information because it does not affect in any way nor the production or the operating costs. Moreover, we think that this parameter is not useful for the estimation of fishing effort of purse seiners. Productivity of this gear, and then the fishing effort on resource, does not depend on the number of sets. And this is true in any of several types of purse seines (white fish, blue fish, and tuna). The productivity of a single operation (fishing haul) has no regularity, which is what happens in the trawl activity. Therefore the number of fishing operations does not give useful information for effort estimation. Number of hours fished would have been more reasonable (but it is not required by the regulation).

III.F.3 Landings

III.F.3.1 Data acquisition

Landings will be estimated according to the variables and disaggregation levels listed in Appendix VIII (EU Dec 93/2010).

Conversion factors will not be applied to landing-weight-based quantities as all species are landed ungutted. Conversion factors could be necessary only for marginal share of landings. For these species, quantities will be converted to live weight by the FAO and Eurostat conversion factors (see table III.F.3)[9].

Commercial landings will be assessed on the basis of a sampling procedure. In particular, monitoring of the activities of fishing vessels less than 10 meters requires the definition of a specific sample program (EU Regulation 2847/93[10], article 8, paragraph 3). The national

[9] FAO Fisheries Circular No 847 rev. 1.

[10] Council Regulation (EC) no 2847/93 of 12 October 1993 establishing a control system applicable to the common fisheries policy

program takes into account this provision by integrating the ongoing sample survey Italy is carrying on in order to fulfil Regulation (EC) No 1921/06[11].

The methodology of the survey for estimation of landings per species was approved by Eurostat during the meeting of the working group "fishery statistics" held in Luxembourg, 18-19 February 2002.

The survey will be based on a sample of around 1500 vessels that is about 10% of the total fleet. The sample is stratified according to fishing segments and geographical areas. Elementary data are collected through questionnaires filled by data collectors, which are scattered along the Italian coast. Survey takes place every week on a continuous basis. Data on landings (weight and prices) and fishing area are recorded by gear and species. This allows post-stratifying the final estimates according to the metiers level (Level 6, Appendix IV, 4). Annual average prices per species are calculated using weighted averages.

The full methodology is reported in Annex VII.

The sample survey will cover the entire Italian fleet which operates in the Mediterranean Sea. Landing data of vessels fishing beyond the straits ("Other regions") will be recorded on a census basis and taken from the compulsory documentation.

Data related to BFT landings (including tuna farming) will be provided by a specific data collection implemented by the national administration in accordance with ICCAT procedures and, as such, will be excluded from the sample survey.

III.F.3.2 Data quality evaluation

Precision levels will be calculated in terms of CV at the level of fleet segments and per GSA. The sample size (1500 vessels, 10% of the fleet) has been fixed according to this precision target. Estimates of variance to calculate CV are reported in Annex VII.

III.F.3.3 Data presentation

Landings data will be available with a time lag of 6 months (for instance data referred to January 2011 will be available not before July 2011).

III.F.3.4 Regional co-ordination

The RCMMed&BS, together with the PGMED (planing group for the Mediterranean) has implemented a landings database to be updated every year. Italy encourages this initiative and suggests discussing methodologies used to estimates landings per species in the Mediterranean in order to better compare final statistics.

III.F.3.5 Derogations and non-conformities

No derogation required.

[11] Regulation (EC) no 1921/2006 of the European parliament and of the council of 18 December 2006 on the submission of statistical data on landings of fishery products in Member States and repealing Council Regulation (EEC) No 1382/91

III.G Research surveys at sea

REGION: Mediterranean Sea and Black Seas

III.G.1 Planned surveys

According to Appendix IX of DCF, the Italian National Program covering 2011-2013 includes two surveys:

- MEDITs, MeDiterranean International bottom Trawl survey
- MEDIAS, Pan-Mediterranean pelagic survey

MEDITs

The Medits programme aims to conduct co-ordinated surveys from bottom trawling in the Mediterranean (Fig. 1). This survey derives from a EU project started in 1994 at European Mediterranean level (Bertrand, et al., 2002).

The basic protocol (Medits, Instruction manual 2007), common to all the Mediterranean partners, includes the design of the survey, the sampling gear (feature and handling), the information to be collected, and the management of the data as far as the production of common standardized analyses of the data.

The challenge of Medits survey, as for other scientific trawl-surveys, is to provide data useful for describing and quantifying changes in the fish populations, through indices of demography, mortality, spatial occupation, biological traits, thus contributing to the development of assessment and management advice tools.

The Italian data collection program foresees the continuation of the Medits survey, principally in the perspective of obtaining information comparable among the various Italian areas and with other Mediterranean countries. Since the beginning of the survey (1994) Medits produce, for a pool of target species (benthic and demersal): abundance indices by species (in number of individuals and biomass per square km; i.e. N/km^2 and kg/km^2) and length frequency distribution (splitted by sex and maturity stages) by depth macrostratum (shelf and slope) and geographical sub-area (FAO/GFCM Geographical sub-areas, GSA).

The working zone is defined as the totality of the trawlable areas off the Italian coasts from 10 to 800 m depth (on the continental shelves and along the upper slopes). These limits have been adopted to cover at best the distribution areas of the main exploited - or potentially exploitable - species, considering the administrative and technical constraints of the project.

The stations have been distributed in each GSA applying a stratified sampling scheme with random drawing inside each stratum.

Since 1994, one survey has been carried out every year, during the spring and the beginning of summer. The duration of the hauls is fixed to 30 minutes on depths less than 200 m and 60 minutes on more important depths. The same positions were visited each year. A total of about 750 hauls was carried out during each annual survey.

Following table (G.1) shows the breakdown of Medits hauls in the seven Geographical sub-areas in which the Italian fleet predominantly operates.

GFCM Geographical sub areas (GSA)	*N. of hauls planned*
9 Ligurian Sea/Upper Tyrrhenian Sea,	120
10 Lower Tyrrhenian Sea	70
11 Sardinian	100
16 Sicilian Strait	120
17 Upper and Middle Adriatic	180
18 Lower Adriatic	90
19 W. Ionian Sea	70
Total	***750***

Table G.1 – Number of Medits hauls planned in each GSA

A list of common target species (including fish, molluscs and crustaceans) was established at the beginning of the Medits project (1994) with reference to their commercial production, their accessibility to a bottom trawl and their potential interest as biological indicators in the different areas. The short reference list of species defined at the beginning of the programme included thirty species. It was enlarged to 38 species, plus all Selachians, during the following years.

In the following table (G.2), the target species for the Medits surveys 2011-2013 are reported.

Table G.2 – List of reference species for the Medits survey (Medits, Instruction manual 2007)

Scientific name	**Date**[1]	**CODE**	**English**
Aspitrigla cuculus	1998	ASPI CUC	Red gurnard
Boops boops	2006	BOOPBOO	Bogue
Citharus linguatula	1994	CITH MAC	Spotted flounder
Eutrigla gurnardus	1994	EUTR GUR	Grey gurnard
Galeus melastomus	1998	GALU MEL	Blackmouth catshark
Helicolenus dactylopterus	1994	HELI DAC	Rockfish
Lepidorhombus boscii	1994	LEPM BOS	Four-spotted megrim
Lophius budegassa	1994	LOPH BUD	Black-bellied angler
Lophius piscatorius	1994	LOPH PIS	Angler
Merluccius merluccius	1994	MERL MER	European hake
Micromesistius poutassou	1994	MICM POU	Blue whiting
Mullus barbatus	1994	MULL BAR	Red mullet
Mullus surmuletus	1994	MULL SUR	Striped red mullet
Pagellus acarne	1994	PAGE ACA	Axillary seabream
Pagellus bogaraveo	1994	PAGE BOG	Blackspot seabream
Pagellus erythrinus	1994	PAGE ERY	Common pandora
Sparus pagrus	> 1996	SPAR PAG	Common seabream
Phycis blennoides	1994	PHYI BLE	Greater forkbeard
Raja clavata	1994	RAJA CLA	Thornback ray
Scyliorhinus canicula	1998	SCYO CAN	Smallspotted catshark
Solea vulgaris	1994	SOLE VUL	Common sole
Spicara flexuosa	1994	SPIC FLE	Picarel
Spicara smaris	1998	SPIC SMA	Picarel
Trachurus mediterraneus	1994	TRAC MED	Mediterranean horse mackerel
Trachurus trachurus	1994	TRAC TRA	Atlantic horse mackerel
Trigla lucerna	2006	TRIGLUC	Tub gurnard
Trigloporus lastoviza	1998	TRIP LAS	Streaked gurnard
Trisopterus minutus capelanus	1994	TRIS CAP	Poor-cod
Zeus faber	1994	ZEUS FAB	John dory
Selacians[2]	2006		
Aristaeomorpha foliacea	1994	ARIS FOL	Giant red shrimp
Aristeus antennatus	1994	ARIT ANT	Blue and red shrimp
Nephrops norvegicus	1994	NEPR NOR	Norway lobster

Parapenaeus longirostris	1994	PAPE LON	Deep-water pink shrimp
Eledone cirrhosa	1994	ELED CIR	Horned octopus
Eledone moschata	1997	ELED MOS	Musky octopus
Illex coindetii	1994	ILLE COI	Broadtail squid
Loligo vulgaris	1994	LOLI VUL	European squid
Octopus vulgaris	1994	OCTO VUL	Common octopus
Sepia officinalis	1994	SEPI OFF	Common cuttlefish

1 Year in which the species was introduced in the list (or removed if the year is preceded by >)

2 It is recommended to carry out the observations referring to this list to all the selacian species in the GSAs where it is technically possible. To allow coherent analyses of the results, it is highlighted that the decision to enlarge or not biological observations on selacians must be applied consistently during all the surveys.

Ref. Common names: Fischer W., M.L. Bauchot, M. Schneider (rédacteurs), 1987. Fiches FAO d'identification des espèces pour les besoins de la pêche. (Révision 1). Méditerranée et Mer Noire Zone de pêche 37. Rome, FAO, vol 1 et 2, 1530 p.

Once collected, the data are stored in computer files by the teams in charge of the survey. Four standard exchange formats (in ASCII) including normalized coding are defined.

The four standard exchange formats are:

- Type A: Characteristics of each haul;
- Type B: Catch of each haul in weight, number and number by sex;
- Type C: Biological parameters for the species in the reference list - length, sex, maturity;
- Type D: Temperature data plus the relative abundance indices (in kg/km^2 and in number of individuals /km^2) as well as length frequency distributions by species and strata

Specific software was written for an automatic checking of the data (Souplet, 1996a and 1996b).

Data are processed and analysed in order to estimate the production of biomass and relative abundance indices (in kg/km^2 and in number of individuals/km^2) as well as length frequency distributions by species and strata. These analyses are carried out using statistical methods approved by the Steering Committee, and included in a specific software (see References above Souplet, 1996a and 1996b). Basic results are presented on standardized media (tables, figures and maps), including visualization of interannual variations.

To preserve the unitary international approach, the Medits project will, as at present, have a coordinator chosen at international level from the national coordinators of the participating countries, as well as a national coordinator and individual officers responsible for each of the seven zones. Working together with his counterparts in the other countries, the national coordinator will preserve the unitary project framework. There is also the willingness to work together with researchers from other countries of the East side of the Adriatic. The national coordinator and zone officers will retain their links with the relevant Italian Administration to ensure that the results are supplied within the time scale and in the form stipulated.

At national level the primary data (the four standard exchanges format Type A; Type B; Type C: Type D) plus the relative abundance indices (in kg/km^2 and in number of individuals /km^2) as well as length frequency distributions by species and strata are stored in a common database. The database is hosted by the Italian Ministry of Agriculture Forestry Polices.

At international level aggregated data are used during the annual Medits international working group. For the time being a web site organized and hosted by Ifremer (https://www.ifremer.fr/medits/index.html) to make available a set of population indices from the Medits surveys carried out by the institute involved, has been built. The indices related to abundance, biomass and length size, in the site are the following:

- for all the selected species: natural logarithm of abundance, total biomass in the area, average individual weight in the population;

- for the species for which individual length is collected: mean length in the population, length at the fifth percentile of the length distribution, length at the twenty-fifth percentile of the length distribution, length at the seventy-fifth percentile of the length distribution, length at the ninety- fifth percentile of the length distribution, sampling variance of length.

All the calculation has been made with the R-SUFI software. All the reference information as well as the R-SUFI software package is available in the Medits website.

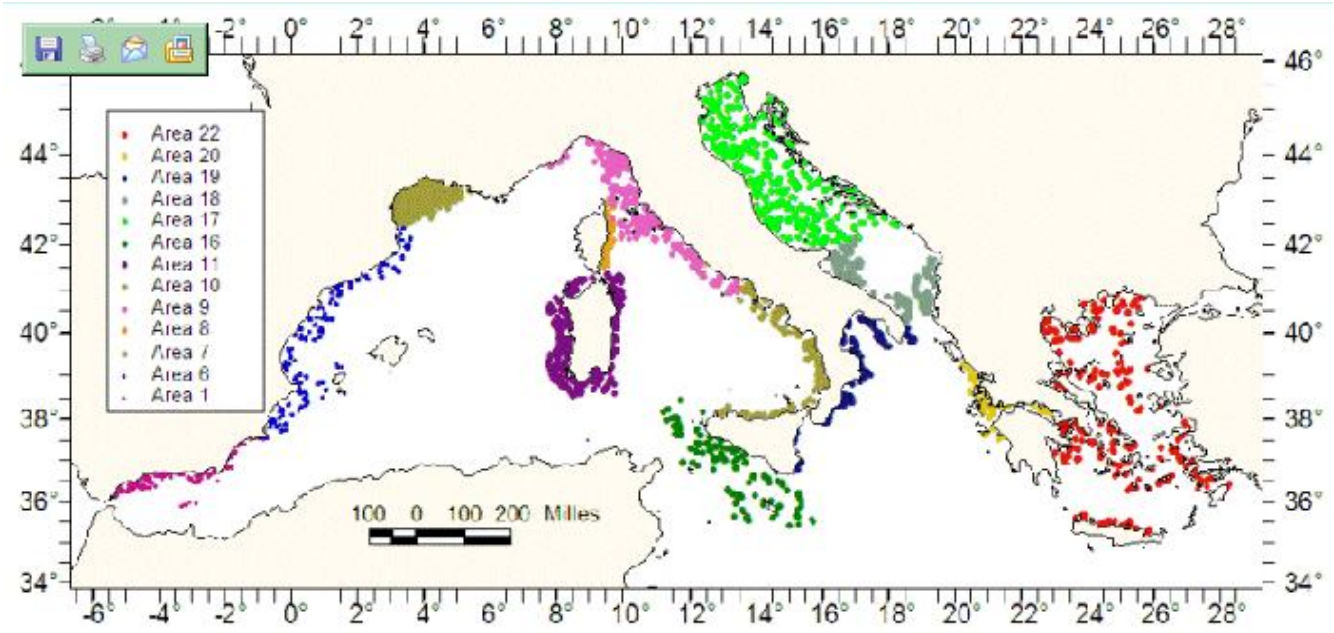


Fig.1 - Localization of the Medits hauls in the Mediterranean and around Italy

MEDIAS

The MEDIAS acoustic surveys on small pelagic fish target anchovy (*Engraulis encrasicolus*) and sardine (*Sardina pilchardus*) covering a series of areas in the Mediterranean EU MS (Spain, France, Italy, Malta, Slovenia and Greece) with a standardised methodology. The aim is to gain knowledge of biomass levels and spatial distribution of small pelagic fishes covering the major stocks of these species in the Mediterranean European Union waters. Italy is responsible for the performance of three cruises which include also territorial waters of Slovenia and Malta:

a) in the Adriatic Sea (GSA 17 and 18; FAO sub areas 37.2.1 and 37.2.2) – Fig. 2;

b) in the Sicilian Channel (GSA 15 and 16; FAO sub area 37.2.2) – Fig. 2;

c) in the Tyrrhenian Sea (GSA 9 and 10; FAO sub area 37.1.3) – Fig. 2. (see Paragraph "Modifications in the surveys")

The investigated area is limited by the Mid-line with Croatia in the northern and central Adriatic part and by the continental shelf area (200 m depth) in southern Adriatic Sea, in the Sicilian Channel (Sicily and Malta) and in the Tyrrhenian Sea.

In northern and central Adriatic Sea (GSA 17) Slovenian waters will be covered by MEDIAS with R/V G. Dallaporta, with a team involving scientists from ISMAR-CNR (Istituto di Scienze Marine – Italy) and FRIS (Fishery Research Institute of Slovenia), whilst the eastern half of GSA 17 will be covered by the Institute of Oceanography and Fishery of Split with its own research vessel: methodologies of survey performance and data analysis will be agreed through a coordination in the ambit of ADRIAMED FAO regional programme and based on the MEDIAS protocol.

In southern Adriatic Sea (GSA 18), if ADRIAMED funds will be available, Montenegro and Albania continental shelf will be covered by R/V G. Dallaporta, with a team involving scientists from ISMAR-CNR (Istituto di Scienze Marine – Italy), IBM (Institute of Marine Biology - Montenegro) and University of Tirana (Albania).

In the Sicilian Channel (GSA 16) Maltese waters will be covered by MEDIAS with a team involving scientists from IAMC-CNR (Istituto per l'Ambiente Marino Costiero – Italy) and MCFS (Malta Centre for Fishery Sciences – Malta).

The study area is covered following systematic parallel grids for a total of about 2100 nautical miles identifying an area of about 15000 square nautical miles in the western Adriatic Sea (Fig. 2), for a total of about 900 nautical miles identifying an area of about 4500 square nautical miles in the Sicilian Channel (Fig. 2), and for a total of about 2000 nautical miles in the Tyrrhenian Sea.

All surveys are performed in summer – early autumn following the internationally agreed MEDIAS protocol. Acoustic data acquisition is done by means of SIMRAD scientific echo-sounder at the frequencies of 38, 120, 200 kHz using split-beam transducers. 38 kHz is the frequency used for assessment, while 120 and 200 kHz are the complementary frequencies. Biomass estimation (t/nm^2) and its spatial distribution are calculated with the standard echo-integration method improved by split-beam and multi-frequency technology. Research Vessels belonging to CNR (National Research Council) will be utilised in the three acoustic surveys.

This research is established with a multidisciplinary approach, in fact, synoptically with the acoustic data acquisition, net samplings on small pelagic fish are performed by means of a pelagic trawl with the aim of determining species and size composition of the pelagic biomass (fish sampling is required to collect representative samples of the population from a qualitative point of view and not a quantitative point of view as is the case of demersal

surveys); length frequency distribution of all the caught fish species will be recorded; age samples of *E. encrasicolus* and *S. pilchardus* will be collected and analysed; CTD oceanographic data (temperature, salinity, fluorescence and dissolved oxygen) will be recorded.

The abundance indices that will be estimated and will be provided in the report should include both NASC (independent from Target Strength equations) and Biomass estimations according to the following: Total fish NASC per Elementary Distance Sampling Unit (EDSU); Biomass per EDSU per target species; Numbers per EDSU per target species; Number/age/Target species and per length class; Biomass/age/Target species and per length class.
In addition the maps of spatial distribution of the target species, anchovy and sardine, will be provided in the report: Point maps of total fish NASC; Point maps of target species in NASC/mile and biomass/mile; pie charts showing catch compositions of the pelagic trawls.

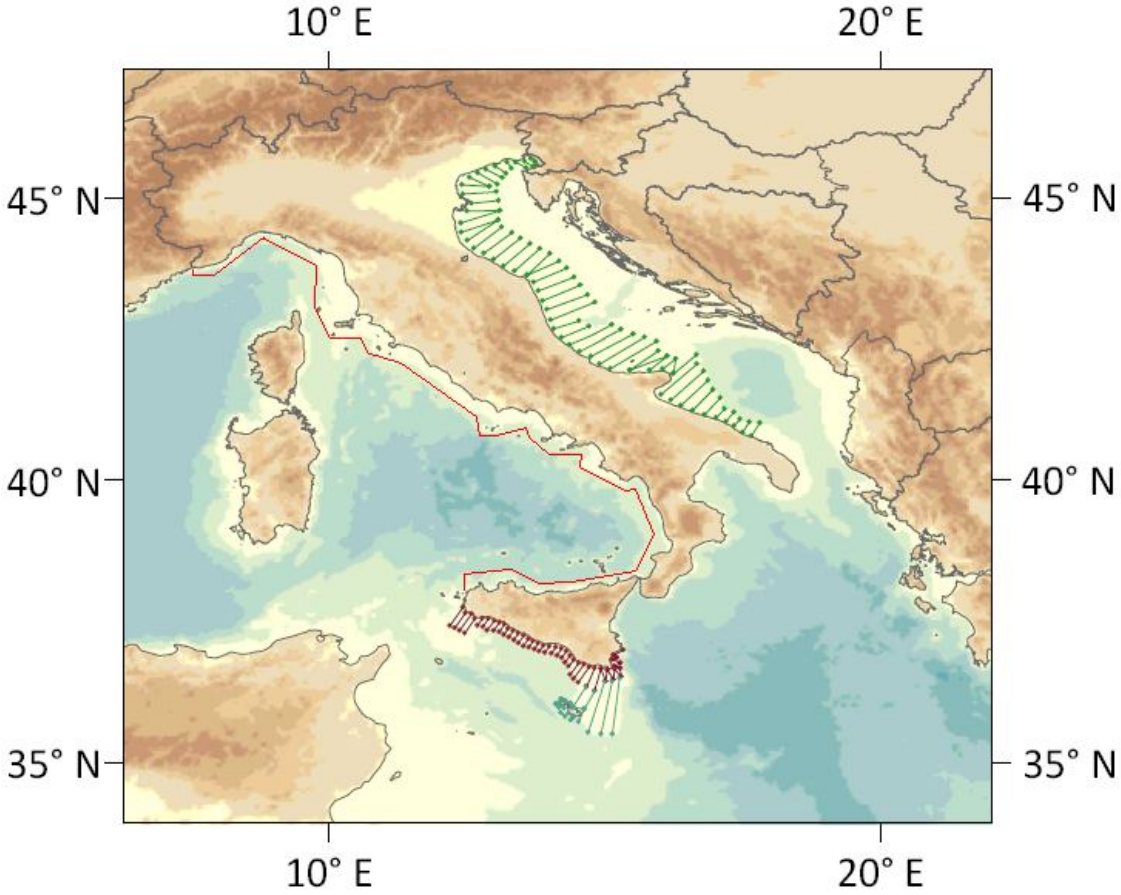


Fig. 2 – Transect grid of the MEDIAS acoustic surveys in the Adriatic Sea (Italy and Slovenia) and in the Strait of Sicily (Italy and Malta); in the Tyrrhenian Sea only the area to be covered is defined by the red poly-line.

III.G.2 Modifications in the surveys

MEDITs survey will be carried out accordingly to the international protocols. No change in the design of the survey is foreseen.

During the RCMMed&BS 2009 has been evaluated the possibility of carry out a second Medits survey in autumn. The use of two complementary data sets would permit more accurate estimates of life history parameters such as mortality and growth. In addition it would provide valuable records for the estimation of stock-recruitment relationships. At the same time, spatial occupation of the different components of the stocks would be better assessed. The gain in accuracy would, in turn, make also more robust the evaluation of changes in the population indicators and of the input parameters for population and community modelling.

During the next RCMMed&BS (Varna, 2010) and the Medits Woking Group (Split, 2010) will be evaluated the possibility to carry out a second survey in autumn and related recommendations will be considered.

MEDIAS survey will be carried out following the methodologies and protocols agreed by the international steering committee (1st meeting of the steering committee for the coordination and harmonization of the 5 ongoing acoustic surveys in the Mediterranean, 25/26 February 2008, Athens (Greece) and 2nd meeting, 1/3 June 2009, Palma de Majorca (Spain)).

The extension of the MEDIAS survey (in the GSA 9 and GSA 10) has been discussed within the overall revision of surveys (SGRN 10-03). Concerning the activity that should be carried out in the GSA 9 and GSA 10, Italy will wait for the output of the SGRN 10-03.

III.G.3 Data presentation

The reference period will be one year. For both surveys (MEDITS and MEDIAS), the results pertaining to the period January –December 2011 will be ready by the end of June 2012.

Some preliminary data (referring to technical part of both surveys MEDITS and MEDIAS) will be produced by the end of April 2012. The same time lag will be applied for data referred to the years 2012 and 2013.

III.G.4 Regional co-ordination

Activities to be carried out in cooperation with other European Mediterranean countries will be discussed and annually monitored during the RCMMed&BS and during both the annual MEDITS and MEDIAS working group.

III.G.5 Derogations and non-conformities

No derogation is requested

IV Module of the evaluation of the economic situation of the aquaculture and processing industry

REGION: Mediterranean Sea and Black Seas

IV.A Collection of data concerning the aquaculture

IV.A.1 General description of the aquaculture sector

Total number of aquaculture firms for 2008 is 826, 422 of which (51,1 % of the total) for fish production, 5 (0.6 %) for crustacean production and 399 (48.3 %) for molluscs production. 890 farms correspond to this firms. (864 active and 26 non active). 452 farms (50,8% of the total), are dedicated to fish production (443 active and 9 non active). Among active farms 334 utilize freshwater and 109 salt water (57 sea water and 52 brackish water). 5 farms have both salt water and freshwater. 5 farms (0,6%) produce crustaceans.433 farms (48,7%) produce molluscs (416 are active and 17 non active), and among active ones 190 have sea water and 226 brackishwater .

	fish		crustacean		mollusc		**TOTAL**
	n°	%	n°	%	n°	%	**n°**
firms	422	51,1	5	0,6	399	48,3	**826**
farms	452	50,8	5	0,6	433	48,7	**890**
	active	non-active	active	non-active	active	non-active	
	443	**9**	**5**	**0**	**416**	**17**	**890**
seawater	57	3	0	0	190	10	**260**
brackishwater	52	0	3	0	226	7	**288**
salt water	109	3	3	0	416	17	**548**
freshwater	334	6	2	0	0	0	**342**

A general overview of aquaculture activity is reported in tab IV.A.1. Total national production for 2007 is 180.928,8 tons, 49.149,7 t fishes, 131.777,6 molluscs and 1,5 t crustaceans..

	production (tonns)			**TOTAL**
	seawater	brackishwater	freshwater	
Fish	12.437,0	4.193,2	32.519,5	**49.149,7**
Crustacean		1,5		**1,5**
Mollusc	70.074,4	61.703,2		**131.177,6**
TOTAL	**82.511,4**	**65.897,9**	**32.519,5**	

According to the origin of the water utilized, 32.519,5 t come from freshwater and 148.409,3 from salt water, (82.511,4 from sea water and 6.5897,9 from brackish water).

It is worth to mention that , besides production, it is necessary to consider that it exists as well a production for on-growing (5.373,5 t) ; general total production is therefore 186302,2 t.

Molluscs culture is mainly missels culture (*Mytilus galloprovincialis*) (56,6%), then clams (*Ruditapes decussatus, R. philippinarum*) (43,4%) and oysters (*Crassostrea spp.*), with very low quantities. Among fishes main species is trout (29945 t, that is 61% of the total) particularly rainbow trout (*Oncorhynchus mykiss*). Then follow the two main species in the Mediterranean, sea bass (*Dicentrarchus labrax*) 15% and sea bream (*Sparus aurata*) 12%. Eel production represented 1.5 % of the total, common carp only 0,3 %. On total national fish production (data from IREPA) aquaculture represents 51,1% in terms of volume, respect marine fishing 48,9%, whereas in terms of value aquaculture represents the 35,4% of the total respect to marine fishing (64,6%).

IV.A.2 Data acquisition

(a) Definition of variables

The methodology proposed by this program will be consistent with that applied to produced statistics required by Regulation (EC) No 762/2008.

Regulation (EC) No 762/2008 on the submission by Member States of statistics on aquaculture and repealing Council Regulation (EC) No 788/96, obliges MS to submit to the Commission statistics on (a) the annual production (volume and unit value) of aquaculture; (b) the annual input (volume and unit value) to capture-based aquaculture; (c) the annual production of hatcheries and nurseries; (d) the structure of the aquaculture sector.

The DCR regulation requires information on the economic aspects (costs, capital, employment, …) of the aquaculture sector.

Even if the information and therefore the methodologies required by the two regulations are different, statistics have to be consistent and estimations on total production must derive from the same source.

Economical data will be collected according to what foreseen in Appendix X of Commission decision 2008/949/CE. Parameters will be collected on annual basis and according to productive segments of Appendix XI. Statistical unit will be the firm, clearly identifiable from a legal point of view according definition Eurostat NACE Code 05.02 “Fish Farming”.

(b) Type of data collection

All the economical parameters will be collected, no the basis of what foreseen in Appendix XII of Commission Decision 2008/949/CE, with the system “Probability Sample Survey” in which the sample will be randomly selected from the universe of aquaculture firms.

(c) Target and frame population

On the basis of the most recent National aquaculture data, Table IV.A.2 has been prepared. In the different segment foreseen in Appendix XI, 754 firms have been classified. Sample has been choosen according to what already prepared in the previous National programmes.

Sample has been defined starting from the total of aquaculture firms 2008 , classified according to the productive segments of the Table IV.A.2. The definition has followed the Bethel's procedure.

(d) Data sources

Economical data to be monitored are those reported in Table IV.A.3. with a specific form.

(e) Sampling stratification and allocation scheme

Information will be collected through the choice of a representative sample per single productive segment, with a random selection. Once selected the firm sample per productive segment it will also be possible to substitute a single firm if necessary.. Substitutions will be reported at the end of the national programme.

IV.A.3 Estimation

The optimum sample number per stratum is defined according to Bethel's procedure (1989). Then, for each collected variable, to obtain the estimates of the totals per stratum, the Horvitz-Thompson formula is used, derived for the particular case of the simple random sampling without replacement. According to this particular estimator, the variance and the CV are calculated to evaluate the precision level . See Annex I for more details.

As regards the imputation of non-responses, there is a process of localization of errors . The control procedure of the survey can be considered as interactive graphic micro-editing of the univariate type. The term interaction refers to the fact that, in the procedure of localization of errors, there are not only automatic phases but also phases which require human intervention to investigate the situation and to evaluate the effective presence of the error (therefore the human intervention regards the localization phase and not that of imputation). The control is mainly of the univariate type because the variables are controlled individually and only in rare cases are suspected relationships existing among them controlled, using suitable synthesis indexes. During the various phases wide use is made of graphic tools to visibly evaluate situations marked as errors. Finally the word micro-editing is used because the data is gathered in suitable domains of study within which the sampling units can be considered very homogenous. For each of these sets of data, suitable control functions are first calculated, and then, for each of them, certain rules of incompatibility are verified. In the case of activation of conditions of error, that is in the case where the observed value does not belong to the region of acceptance, those control functions are then observed individually for all the sampling units forming the single domain. Thus the sample unit, or units, responsible for the activation of conditions of error is localized for the entire domain of study and then the imputation of the erroneous data follows. For more details about the method of imputation of non-responses, see Annex I.

IV.A.4 Data quality evaluation

Table IV.A.3 gives further details on the accuracy indicators to assure the quality of the collected data, for each parameter.

Information on data quality is given in terms of target precision levels in the case of random sample and in terms of coverage rate in the case of fixed panel. The estimation of the variance for the calculation of the CV is explained in Annex I.

IV.A.5 Presentation

Data collected and elaborated will be available at the conclusion of the National programme 2011-2013 after the due control and elaboration procedures.

IV.A.6 Regional coordination

No regional co-ordination is foreseen

IV.A.7 Derogations and non-conformities

No derogation required

IV.B. Collection of data concerning the processing industry

REGION: Mediterranean Sea and Black Seas

IV.B.1 Data acquisition

(a) Definition of variables

The parameters that will be collected are those listed in Appendix XII of Commission Decision 2008/949/EC. This Appendix specifies the definition to be applied for most variables. For those variable not provided by definition the following will be considered:

- Other income: the amount of income non imputable to sales of products (turnover) and not being subsidies, i.e. "variation in stocks of finished goods and in work in progress", "work performed by the undertaking for its own purposes and capitalized", "other operating income" as defined in the IV Council Directive 78/660/EEC, art. 23, item 2 and 3 and in the Commission Regulation (EC) No. 2700/98, containing definition for Structural Business Statistics (SBS), item 42 31 0.
- Imputed value of unpaid labour can be relevant in the case of small and medium enterprises managed at family levels. In this case, the average labour cost, estimated for "paid labour force" will be applied and imputed to the item "personnel cost" for enterprises declaring to have not dependent labour force.
- Financial costs will be will be estimated referring to the IV Council Directive 78/660/EEC, art. 23, item C.13 ("*Interest payable and similar charges, with a separate indication of those concerning affiliated undertakings"*).
- Extraordinary costs: this parameter will be estimated referring to the IV Council Directive 78/660/EEC, art. 23, item 17.
- Debts: parameter will be estimated referring to the IV Council Directive 78/660/EEC article 9, Passive C. and to the Commission Regulation (EC) No. 2700/98, containing definition for Structural Business Statistics (SBS), item 43 21 0.

As far as collection of data on FTE, the study FISH/2005/14 is not considered appropriate. It is, indeed, focused on the estimation of FTE for the fishing fleet at a metier level, by using data on the time (hours) spent at sea in fishing operation. FTE for the fish processing sectors, as well the breakdown of employment by gender, will be estimated by taking into account previous studies on the collection of socio-economic data of the sector.

(b)Type of data collection

Data will be collect by mean of a Probability Sample Survey, as reported in the standard table IV.B.1 and IV.B.2.

Two type of data collection will be carried out. The first on capital companies (Ltd., Plc. or cooperatives) and the second on partnerships (limited, ordinary or general partnerships) and individual firms because of the different type of commitment they have in publishing the balance sheets. Capital companies have to publish balance sheets every year while this is not mandatory for the other type of firms.

For capital companies the survey will be carried out by making queries on the official data stored by the national Chamber of Commerce. For partnerships the survey will be carried out through postal technique and/or personal interviews. In the second case, partnerships will have the possibility to send the filled questionnaire by fax or by e-mail. In the case of personal interviews, data collectors will be prepared through an informative meeting. Official company accounts (where published) will also used to cross-checked the information gathered from the questionnaires.

(c)Target and frame population

The target population is that defined in the DCF, that is all the companies and partnerships whose activity fall under the NACE 10.20 codification: "Processing and preserving of fish, crustaceans and molluscs".

The frame population coincides with the target population as it is represented by all the firms present in the registers of the national Chamber of Commerce as carrying out the NACE 10.20 activity. The population is represented by all the enterprises declaring the NACE 10.20 activity at the date of writing that is February 2010. The sampling strategy is planned on this population size. Each programming year the register will be interrogated in order to update the population size (this can change following the activity ceasing or bankruptcy of some enterprises). The planned sample rate will be respected, independently by the population size. Technical report will describe eventual changes in the population size.

The target (as well the frame population) is made up of 780 units as reported in the standard table IV.B.1.

(d) Data sources

The data source is represented by the register of the national Chamber of Commerce. This register provides all the necessary information about the population units (registration number for queries on balance sheets and contact information for sending questionnaires).

As far as enterprises carrying out fish processing not as a main activity, the official and more reliable archive used for sampling (Chamber of Commerce) does not allow, at the moment, to distinguish between main and secondary activity.

(e) Sampling stratification and allocation scheme

Type of sampling strategy

The sampling units will be selected by mean of a simple random sampling.

Further stratification

The population will be stratified in order to get more precise estimates, also depending on the methodology of data collection. The stratification will be based on the firm's typology or type legal entity that is capital companies and other type of firms (partnerships, individual firms, etc…).

Determination of sample size

According to the past experience (previous national data collection for the fish processing sector) a coverage rate of about 15% is deemed to be appropriate to get precise estimates.

The size of the sample in each stratum is defined by mean of the methodology based on the proportional allocation, where the size of the sample is taken in proportion to the size of the stratum.

Standard table IV.B.1 gives a general outline of the planned sampling no. and sample rates.

IV.B.2 Estimation

Estimation methods from sample to population

To obtain an estimate of totals from sample to population for each stratum, the Horvitz-Thomson estimator for the simple random sampling will be used.

Imputation of non responses/Non-responses adjustments

In order to face the problem of non-responses, the initial weights assigned to the sample units will be adjusted based on data referring to respondents and non-respondent of the selected sample.

Data coming from different data sources will be checked against sectorial studies in order to ensure the consistency.

IV.B.3 Data quality evaluation

According to the new guidelines for the submission of multiannual national programmes, table IV.B.2 gives further details on the sampling methods used and on the methods used to assure the quality of the collected data, for each parameter.

Information on data quality is given in terms of target precision levels in the case of random sample and in terms of coverage rate in the case of fixed panel. The estimation of the variance for the calculation of the CV in the case of statistical sample is explained in annex I (methodology for the estimation of economic variables for the fleet).

IV.B.4 Data presentation

In the second half of 2011, data referred to 2010 will be collected (the deadline to publish financial accounts is June of each year for accounts of the previous year). Aggregated and validated data will be available by December 2011. In the same way data referred to 2011 will be collected in 2012 and delivered by December 2012. Finally, data referred to 2012 will be collected in 2013 and delivered by December 2013.

The time lag between the date of availability and the reference year is one year.

IV.B.5 Regional coordination

No regional co-ordination is foreseen.

IV.B.6 Derogations and non-conformities

No derogation requested.

V Module of evaluation of the effects of the fishing sector on the marine ecosystem

REGION: Mediterranean Sea and Black Seas

Environmental indicators listed in Appendix XIII of the DCF will be estimated. The following tables reports the different data sources that will be used for each indicator:

Code specification	Indicator	Definition	Source
1	Conservation status of fish species	Indicator of biodiversity to be used for synthesizing, assessing and reporting trends in the biodiversity of vulnerable fish species	MEDITs/MEDIAS
2	Proportion of large fish	Indicator for the proportion of large fish by weight in the assemblage, reflecting the size structure and life history composition of the fish community.	MEDITs/MEDIAS
3	Mean maximum length of fishes	Indicator for the life history composition of the fish community	MEDITs/MEDIAS
4	Size at maturation of exploited fish species	Indicator of the potential "genetic effects" on a population	MEDITs/Biological sampling of catches
5	Distribution of fishing activities	Indicator of the spatial extent of fishing activity. It would be reported in conjunction with the indicator for 'Aggregation of fishing activity'.	VMS data Vessel Register
6	Aggregation of fishing activities	Indicator of the extent to which fishing activity is aggregated. It would be reported in conjunction with the indicator for 'Distribution of fishing activity'.	VMS data Vessel Register
7	Areas not impacted by mobile bottom gears	Indicator of the area of seabed that has not been impacted by mobile bottom fishing gears in the last year. It responds to changes in the distribution of bottom fishing activity resulting from catch controls, effort controls or technical measures (including MPA established in support of conservation legislation) and to the development of any other human activities that displace fishing activity (e.g. wind farms).	VMS data Vessel Register Other sources
8	Discarding rates of commercially exploited species	Indicator of the rate of discarding of commercially exploited species in relation to landings.	Biological sampling of catches
9	Fuel efficiency of fish capture	Indicator of the relationship between fuel consumption and the value of landed catch. It will provide information on trends in the fuel efficiency of different fisheries.	Economic data, see chapter III B of the NP

The MEDITS and MEDIAS surveys will be used to estimates ecosystem indicators from 1 to 4 listed in Appendix XIII.

Indicator 1 – "Conservation status of fish species - Indicator of biodiversity to be used for synthesizing, assessing and reporting trends in the biodiversity of vulnerable fish species"

To evaluate this indicator will be used historical series from MEDITS and MEDIAS surveys.

As a proxy will be used the R-SUFI routine for the two following indicators:

1.1) Total abundance in number and weight – it is calculated as abundance index in number and weight of the total of fish, cephalopods, crustaceans decapods, stomatopods and selachians in the community.

1.2) Diversity index Δ (Hulbert, 1971) – used as a measure of eveness and interpretable as the probability that two individuals taken randomly from a community belong to different species

Indicator	Data input	Formula
1.1.) Total abundance in number (N) and weight (B)	Catch in N and B per haul k in the stratum j $y_{k,j}$ Trawled area $a_{k,j}$ Area of the stratum A_j	$Y = \sum_j Y_{i,j} = \sum_j A_j \sum_{k=1}^{n_j} \sum_i^n y_{ikj} \Big/ \sum_{k=1}^{n_j} a_{k,j}$ $Var(Y) = \sum_j \frac{A_j^{\,2}}{n_j - 1} \sum_{k=1}^{n_j} \left(\frac{\sum_i y_{i,kj}}{a_{k,j}} - \frac{\sum_{k=1}^{n_j} \sum_i^n y_{i,,}}{\sum_{k=1}^{n_j} a_{k,j}} \right)^2$
1.2) Diversity Δ	N_i	$\Delta = \frac{N}{N-1} \left[1 - \sum_{i=1}^{n} \left(\frac{N_i}{N} \right)^2 \right]$ $Var[\Delta] \approx \sum_i^n Var[N_i] \left(\frac{2N_i}{N^2} - \sum_i^n \frac{2N_i^2}{N^3} \right)$

Indicator 2 – "Proportion of large fish - Indicator for the proportion of large fish by weight in the assemblage, reflecting the size structure and life history composition of the fish community"

To evaluate this indicator will be used historical series from MEDITS and MEDIAS surveys.

Will be used the p*large* indicator calculated trough the R-SUFI routine. This will give the proportion, based on the biomass, of specimens bigger than a certain size.

ICES has identified a threshold of 40 cm TL, whereas in our case will be evaluated and compared the method proposed by Rochet et al., 2004 and tested on 4 different size range (15, 20, 25 e 30 cm TL).

Indicator	Data input	Formula
2) Proportion of fish bigger than a certain determined thresold l_{big}	$y_l(t)$ catches per l $y(t)$ total catch (measured species) determined thresold *lbig*	$p_{\text{large}}(t) = \sum_{l>lbig} y_l(t) \Big/ y(t)$ $Var[p_{\text{large}}] = \frac{p_{\text{large}}(1 - p_{\text{large}})}{y(t)}$

Indicator 3 – "Mean maximum length of fishes - Indicator for the life history composition of the fish community"

Length at the ninety- fifth percentile of the length distribution $L_{0.95}$ - As a proxy will be used the R-SUFI routine that will estimate the Length at the ninety- fifth percentile for each species measured. The mean maximum length in the community will be considered as the mean length of the bigger fish in the community (Shin et al., 2005).

Indicator	Data input	Formula
3) Mean length of the community at 95 percentile ($l_{0.95}$)	- Length at 95 percentile - $L_{q,i}$ per species - S number of species measured in the survey - are considered the measured species with a temporal historic series	$l_q = \sum_{i=1}^{S} L_{q,i} \Big/ S$ $Var[l_q] = \sum_{i=1}^{S} Var[L_{q,i}]$

Indicator 4 – "Size at maturation of exploited fish species - Indicator of the potential "genetic effects" on a population"

To evaluate this indicator will be used historical series from MEDITS surveys.

Length at maturity L_{50} – This parameter will be calculated trough the R-SUFI routine that will estimate the maturity length at 50% (l_{50});

Indicator	Data input	Formula
4) The length on which 50% of the population *i* is mature in the year *t* (L_{50})	$N_{m,i}(t)$	1) Estimate of the probabilty to be mature p_l in function of the body length *l* in the year *t*: $\ell(p_{l,t}) = \log\left(\frac{p_{l,t}}{1-p_{l,t}}\right) = \mu + a_t + b_t l + \varepsilon$ 2) Estiamte of L_{50} as: $L_{50} = \frac{\ell(0.5) - \mu - a_t}{b_t}$

Ecosystem indicators from 5 to 7 listed in Appendix XIII (Distribution of fishing activities, Aggregation of fishing activities, Areas not impacted by mobile bottom gears) will be estimated using VMS data. They were proposed to monitor the impact of fisheries on the ecosystem as described by the spatial extent of fishing activity. These are:

Indicator 5 – "Distribution of fishing activities"

Indicator of the spatial extent of fishing activity. It would be reported in conjunction with indicator 6. It would be based on the total area of grids (3 km x 3 km) within which VMS records were obtained, each month.

Indicator 6 – "Aggregation of fishing activities"

Indicator of the extent to which fishing activity is aggregated. It would be reported in conjunction with the indicator for 'Distribution of fishing activities'. It would be based on the total area of grids (3 km x 3 km) within which 90% of VMS records were obtained, each month.

Indicator 7 – "Areas not impacted by mobile bottom gears"

Indicator of the area of seabed that has not been impacted by mobile bottom fishing gears in the last year. It responds to changes in the distribution of bottom fishing activity resulting from catch controls, effort controls or technical measures (including MPA established in support of conservation

The use of VMS apparatus started mainly as a control tool for the application of the Common Fsheries Policy (CFP), for the conservation and sustainable exploitation of fisheries' resources. The so-called "blue box", i.e. an on-board system allowing a fishing vessel to be tracked remotely, is one of the measures that the European Commission adopted under Regulation (EC) No. 2371/2002). The Regulation provides that, beginning on 1st of January 2005, fishing vessels with an overall length exceeding 15 m "shall have installed on board a functioning system which allows detection and identification of that vessel by remote monitoring systems". Subsequent Regulation (EC) No. 2244/2003 lays down more detailed provisions. VMS data represent a useful tool for management purposes as well, and are fully compatible with graphic rendering and interpretation through the GIS.

The VMS database contains a huge amount of data, and a careful and long job of data "cleaning" will be necessary, since quality of data is imperative for its use in resource management. It will be necessary as well the separation between signals coming from the vessel during navigation and those transmitted during fishing activity, and their selection according to different fishing system. This data will be crossed with other sources of data:

- **Vessel Register**, containing information about Boat characteristic (OL, GT, KW, fishing license...);
- **Logbook**, containing vessel ID and information on the gear used. This allows the identification of the vessel to métier level 4. Logbook information usually includes retained catches of the main commercial species on a trip-by-trip basis and at the scale of an ICES rectangle. This information allows the identification of the vessel to métier level 5 or 6.

VMS and logbook information needs to be processed in order to better describe fishing activity for use in the three pressure indicators. Different methods are used to:

• Identify fishing activity;

• Create fishing tracks;

• Define métiers.

At first, it is necessary to distinguish fishing activity from other activities (e.g., steaming). This is possible by using both information data about boat activity (the bluebox code) and vessel speed. These two sources of information have to be combined in different ways for the different metiers.

The time frequency of VMS position returns for Italian vessels is variable, ranging from 20 minutes to 2 hours. According to the Report of the Working Group on the Ecosystem Effects of Fishing Activities (WGECO), frequency of VMS signals should be interpolated at low time intervals (< 0.5 h), so that it is likely that many fewer unaccounted movements have been undertaken and joining points by straight lines is reasonable. In this way, VMS signals of Italian fleet will be interpolated using Kochanek–Bartels cubic splines.

Métier level 4 describes gear types. Level 5 describes trawl with species (i.e., otter trawl and plaice), while métier level 6 is a specific description based on the catch. The DGMARE request indicates that activity information is required disaggregated to métier level 6. This level is based on the catch composition. However, level 6 (mesh size) corresponds to level 5 for what concern trawlers (for the bottom and pelagic trawl the minimum mesh size is 40 and 20 mm respectively, Reg. 1967/06/CE). Regarding the small scale fishery, all the different mesh sizes have been aggregated at level 5. According to the WGECO, the key information would be about the vessels and its gear, so down to métier level 5. If this is agreed, it should significantly reduce the data volume needed to address the request and have no impact on the conclusions.

The approach used to define métiers at level 5 consists in a quantitative analysis of logbook data, which were directly combined with information gathered from VMS. The results will permit to describe the distribution of activity of fishing boats (over 15 metres long), linking descriptive with “operative” information. All the calculation has been made with the R software, in agreement with the other experiences carried out for indicators 1-4.

Definition of indicator 5, in particular, foresees the selection of signals during fishing activity, and can be defined according to the main different fishing system. With this purpose, exercises and case studies were already performed and utilized for activities of technical assistance to the Directorate of Fisheries, as well as in framework of the project “Construction of a GIS supporting management processes in maritime fishing and aquaculture sectors in the framework of EFF (European Fishing Fund)”. This exercises made possible the identification of the geographical (using GSA as a reference) and temporal distribution of fishing boats considered.

Definition of indicator 6 needs the elaboration of “frequency areas” for signals coming from different kind of fishing systems, and their assemblage according to their frequencies in the different geographical areas (GSA as a reference), and this could fit with the definition of indicator 6.

Last indicator (N. 7 “Areas not impacted by mobile bottom gears”) can only partly derived from the elaborations of number 5. It foresees the use and the crossing among different sources of data, as the areas subject to a kind of total or partial protection (either geographical or temporal), or dedicated to different activities. The presence of the majority of this data already stored in a GIS will facilitate their use according to the specific goal.

It has to be noted, as a final note, that the definition of all the three indicators will be covered by the VMS data only for the fishing fleet over 15 metres long.

Indicator 8 - "Discarding rates of commercially exploited species - Indicator of the rate of discarding of commercially exploited species in relation to landings"

Regarding the indicator of state of the population, it will be represented by the discarding rates of commercially exploited species in relation to their landing. Biological samples, to be carried out under the National Program, will provide sampling on board of commercial vessels (i.e. trawlers).

During the biological sampling quantity of the target species discarded will be estimated. This indicator, although not giving a direct estimate of the state of exploitation of the resource, will allow to determine whether sampling is performed optimally integrating the information derived from other indicators.

Indices of abundance in weight and number for commercial landing and discards of the target species (SIBM, 2005) will be the requested data.

The rate of discard per metier and target species, will be expressed as "ratio estimator" and corresponding variance (Cochran, 1977). This rate will be calculated seasonally and annually. For each year and for each metier will be also calculated a "composite indicator" (a *geometric* mean of annual rates of individual species).

The high rate of discard of commercial species is considered an indicator of lack in the harmonization between fishing gear and minimum sizes of fish. This could be an indicator of fishing pressure in areas where juveniles are concentrated, or of a less selective fishing activity with respect to market demands.

8) Rate of discard of commercial species	Discard (D) and landing (L) in weight per trip j, vessel k, metier m and species i n= number of sampled vessels N= total number of vessels per metier .	Mean rate of discard (R) will be calculated as $\hat{R} = \frac{\sum_{k=1}^{n} D_k}{\sum_{k=1}^{n} (L_k)}$ Variance: $Var[\hat{R}] = \frac{(1-f)}{n\bar{L}^2}\left(s_D{}^2 + \hat{R}^2 s_L{}^2 - 2\hat{R}s_{DL}\right)$ f =n/N; S_D= sampling variance of discards; S_L= sampling variance of landing and S_{DL}= sampling covariance

Indicator 9 – "Fuel efficiency of fish capture"

Ecosystem indicators 9 will be calculated using economic and transversal data. The methodology for estimation of this indicator is reported under chapter III B (economic data).

VI Module for management and use of the data

REGION: Mediterranean Sea and Black Seas

VI.A Management of the data

Under the provisions of the previous NP 2009-2010, an information system for the fishing sector in application of Reg. (EU) 199/2008 has been implemented.
This databank is directed toward the collection, storage, management and systemisation of data of an economic and biological nature pertaining to the sector, coming from different sources and intended for differentiated users. In particular, the information system allows different levels of access to meet the information needs of the Commission, the General Directorate of Maritime Fishing and Aquaculture, the same suppliers of the data and general use.

The three annual NP (2011-2013) will consider the following activities:

- update the system through the methodological and technological development of the current procedures for the collection and storage of data;

- the evolution of the existing system of registration of data of a biological nature, with the integration/supplement of other data on the fishing sector of an economic nature; implementation of the different typologies of information needs, and implementation of the different typologies for the feeding of the Database;

- the implementation of the informatics coding for the standardisation of the parameters for the management of the data;

- the evolution of a Web platform (organisation of the contents and services, surfing mechanisms, graphics) for the consultation of the data. It will make possible and more feasible to meet the many and diversified needs relating to the usability of the data by managing different degrees of access to the Website and providing for a set of processing functions on demand.

VI.B Use of the data

Data collected under the DCF will be primarily used to prepare data sets suitable to support the scientific analysis as a basis for advice to fishery management.

Thus, biological metier related variables will be estimated as well as the demographic characteristics of the stocks, complementing the information from fishery-dependent and fishery-independent data collected under DCF framework.

To meet this objective data related to the demographic structure of landings and discards (when available) as well as data obtained from the surveys (MEDITS trawl survey and MEDIAS) will be used to gather estimates of the main parameters that are relevant for the population dynamics of the target species and to assist the stock assessment processes.

The basic parameters supporting this process are: indices of abundance, the sex ratio, the von Bertalanffy growth pattern, the proportion of mature fish by size/age, the recruitment pattern, the structure of the landing/catches by length and age.

This kind of information can be used to assist the scientific analysis based on population modelling applying different methods (virtual population analysis, length cohort analysis, forward dynamic pool models, composite models, etc.), that as a result are necessary to estimate reference points (Caddy, 1999) (e.g. . $F_{0.1}$, F_{max}, F_{MSY}, $SPR_{\%}$, exploitation rate E, Z_{mbp}). All this could implement a precautionary approach to fishery management, that given the level of uncertainty in the underlying process, would benefit of complementary data and approaches.

In addition, as the Common Fisheries Policy (CFP) needs the progressive implementation of an ecosystem-based approach to fisheries management (EBFM), a suite of indicators will be estimated (see chapter V for details) with the aim of supporting advice trough the analysis of the state of the ecosystem, the description of the pressures affecting the ecosystem, and the communication of trends to a non-specialist audience. In order to put the analysis in the context of "Pressure, state and response" (Garcia et al., 2000) focus will be given on the indicators that describe the state of components and attributes of the ecosystem, but also on indicators related to the impact of fishery on the target stocks and on the ecosystem. All this could establish the basis for supporting advice through, for instance, multiple criteria decision analysis approach (e.g. Leung, 2006).

Finally, collected data will be used to run bio-economic models. In the last 10 years, the approach using computer-based simulation models has been increasingly adopted to indicate and predict the effects of management measures on fisheries from biological, economic and social points of view. The aim of using bio-economic models is to be able to explore options through a comparison of the expected performance of candidate management and assessment strategies relative to the management objectives. As the impact on fish stocks becomes greater, as evidenced in the majority of EU fisheries, so does the need for robust and reliable simulation approaches with which to provide confident management advice.

EU and national research programmes have supported many research projects concerning fisheries bio-economic modelling where the objective was to produce a bio-economic model suitable for the analysis of management strategies of Mediterranean fisheries.

In the context of the National Program, different existing bio-economic models will be compared and evaluated. The validation of selected models will be performed in terms of verification of their predictive capacity and adaptability to different areas and fisheries. This will consist of two steps:

- validation of the model structure. In order to use a bio-economic model to simulate a scenario within a different fishery, verification of the model structure needs to be undertaken to be sure that it reflects its features;

- validation of the model results, to test how well a model is able to predict the real behaviour of the system modelled.

After the validation phase, sets of data for bio-economic modelling will be prepared to be used as input in the bio-economic models and corresponding scientific analysis will be performed.

VII Follow-up of STECF recommendations

SGECA-SGRN recommendations 09_01	Actions
ON DEROGATIONS FOR DISCARDS AND LANDING SAMPLING According to DCF, all metiers selected by the ranking system should be sampled for landings and discard. There are provisions in DCF to reduce the sampling burden (e.g. merging; international co-operation). SGRN recognize that it will take time for MSs to adjust to the new DCF. SGRN would point out that derogations can be granted if reaching for target precision level implies excessive costs or if they are fully documented and scientifically proven. No such documentation has been provided by MSs. SGRN consider that derogations can only be granted if the level of discard is statistically proven and supported by documentation.	Italy will sample all metier selected by the ranking system with the exception requested under section III.E.5 "Derogation and Conformities". During the RCMMed&BS 2009 the attention has been focused on identifying the key metiers important to sample for discards and on providing scientific justification for not sampling certain metiers. The metiers of the regional reference list were analysed by the group to define whether the discarding levels for G1, G2 or G3 species are significant or could justify derogation. In the proposed table (see RCMMed&BS report 2009) where a 'no' answer is provided, the metier is not mandatory for discards sampling and MS need not ask for derogations. Where a 'YES' answer is provided, discarding level is assessed as significant and concerned metier must be sampled adequately and every year for discards. For some metiers, references should be provided (in order to request derogation) and this should be done at national level.
ON METIER SAMPLING – MINIMUM REQUIREMENTS According to the DCF, at least 12 trips per metier per year should be carried out for landings and 8 for discard. For several MSs, the planned number of trips is lower than indicated above. However, SGRN notes that some fisheries are only seasonal and therefore the minimum requirement is likely to be less than 12 trips per year. Even if it is less than 12 trips, the MS have to allocate their sampling in proportion to the seasonality of fishery.	In the present NP, the number of fishing days to be sampled has been defined proportionally to the effort (number of days at sea for each metier) and the landings. The sampling unit belonging to the metier (primary unit) will be the fishing trip (secondary unit). According to the Commission Decision, the minimum number of fishing days to be sampled will be at least 1 fishing day per month during the fishing season. Thus at least three samples will be collected in each quarter for each metier.
ON USE OF SURVEY LENGTH AND AGE DATA SGRN notes that some MS are using both surveys and landings sources for sampling age composition. SGRN interpretation of the DCF (Section B2-3.1) is that length composition or age composition from surveys cannot be applied to derive length and age composition of landings. SGRN recognize that the DCF does not forbid the use of age reading (i.e. age at length information) from surveys to support/complement age reading from landings for the construction of age length keys.	Stock related variables (sex, weight, and maturity) will be collected during different fishing days taking into account the spatial (GSA) and temporal (quarterly) variability in order to detect seasonal differences in the demographic structure and composition of the species examined. Data will be collected mainly from commercial fisheries (i.e. landing, on board sampling). Following each stratum, both "Probability Sample Survey (B)" and "Non-Probability Sample Survey (C)" will be used. Survey data will be used (if the survey is conducted in the correct period) to integrate information on growth and reproduction parameters.
ON NON SELECTED METIERS There is no documentation or explanation about the level of discard for all metiers that are not selected by ranking to allow SGRN for evaluating derogations for discard sampling of those metiers.	Level and importance of discards, has been defined at Regional Level (RCMMed&BS Final report 2009) for all the metier listed in the Appendix IV of the EU Decision 949/2008. See the first comment above.
ON THE ESTIMATION OF CAPITAL VALUE AND CAPITAL COSTS SGRN would point out that the capital value and capital costs should be estimated according to the PIM methodology proposed in the capital valuation report of study No FISH/2005/03 (2006. Evaluation of the capital value, investments and capital costs in the fisheries sector, 203 p.). The data and estimation	Italy will follow the PIM methodology in order to estimate capital value and capital costs. Data and estimations procedures are explained in section III.B of NP.

procedures should be explained in the national programme. In case this methodology cannot be applied, appropriate justifications should be given in the NP.	
ON SPECIES DEROGATIONS AND NON CONFORMITIES Species derogations can be granted on the basis of less than 200 tons in total landing, if there is not any different provision from relevant RFMO. However, when official references such as EUROSTAT reveal inconsistencies, the derogation cannot be evaluated on the basis of the EU total landing. MS should provide these data so the derogation can be evaluated.	Since 2007, following RCM recommendation (4th RCMMed Report - Cyprus, 2007), MS provided landings data of the species presented in Appendix VII of the Commission Decision 2008/949/EC. A common template (update every years during the PGMed meeting), collating all landings data per country, has been used as a reference for the selection of species to be included in the biological sampling. When PGMed collates all landings data per country some MSs do not provide landing by species but for genus or large group of species (i.e. cephalopods). In this case the exemption rule fixed by DFC "*for stocks in the Mediterranean Sea, the landings by weight of a Mediterranean Member State for a species corresponding to less than 10 % of the total Community landings from the Mediterranean Sea* (Commission Decision 93/2010)" cannot be determined. Italy stress the importance of providing landings data by species, as required by the DCF (EC Decisions 949/2008 and 93/2010), and not by group of species.
ON THE ESTIMATION OF EMPLOYMENT. SGRN would point out that the methodology for the estimation of employment (engaged crew and FTE) should be in accordance with the Study FISH/2005/14 (2006, Calculation of labour including full-time equivalent (FTE) in fisheries, 142 p.) and amended by the SGECA 07-01 report and should be explained in the national programmes.	Italy will follow the methodology included in the Study FISH/2005/14 in order to estimate capital value and capital costs. Data and estimations procedures are explained in section III.B of NP.
ON THE CONSISTENCY OF DIFFERENT ECONOMIC DATA SOURCES SGRN recognizes the necessity to use different data sources to collect economic variables. However, SGRN reminds MS that in this case the DCF requires Member States to ensure consistency and comparability of all economic variables when derived from different sources (e.g. surveys, fleet register, logbooks, sales notes). SGRN asks MS to explain in the national program how the consistency of information derived from different data sources has been checked.	The data sources for estimation of variables in Appendix VI are homogenous, therefore data are consistent. This guarantees that economic indicators on profitability (gross cash flow, gross value added, net profit) are not biased by the use of different data sources. For more details about the data quality evaluation, see Annex I.
ON THE COVERAGE OF POPULATION MS has to cover all the fleet and collect the required data also for inactive vessels. Most of the MS do not provide any information about inactive vessels in the relevant tables of the national programs.	The population is all vessels in the Community Fishing Fleet Register on 1 January. Inactive vessels will be considered as a separate segment. For inactive vessels only capital value, fleet and capacity will collected as required by the DCF.
ON THE AVAILABVILITY OF FINAL ESIMATES SGRN reminds MS to indicate in the national programs when the final validated data will be available, as required in the guidelines.	This information is included in the NP
ON THE ESTIMATION OF ECOSYSTEM INDICATOR – FUEL EFFICIENCY SGRN reminds MSs that according to the guidelines, they have to provide the method of estimation of fuel efficiency of fish capture in the Section III.B of the NP. Not all the MS provided this information in the section	This information is included in section III.B

mentioned. SGRN reiterates that MS to follow the guidelines.	
ON THE CALCULATION OF IMPUTED VALUE OF UNPAID LABOUR SGRN reminds MS that methods of calculation of imputed value of unpaid labour have to be explained in the national programs. This request refers to fishery, as well to aquaculture and processing industry. SGRN recognizes that there is no common methodology of calculation of imputed value of unpaid labour and suggests that this methodology be proposed by SGECA.	This information is included in the relevant sections of the NP.
ON TUNA CAGING ACTIVITY (TUNA FARMING AND TUNA FATTENING) – COLLECTION OF ECONOMIC DATA. At the moment, several member states involved in tuna farming activity included these plants among the aquaculture sector. Tuna caging plants have been supported by aquaculture subsidies and this activity is included by FAO among the aquaculture systems (wild-caught aquaculture). SGRN recommends that all MS having tuna farms shall include them among the aquaculture plants and carry out the data collection activities required. According to the ICCAT Rec.06-07, art.2c, the levels of biological sampling of bluefin tuna kept in cages have been defined. The obligation for the country where cages are located to ensure the sampling and cooperation agreements with all third countries where tuna cages are eventually located will be established. The very recent ICCAT Rec.08-05 clearly establish the responsibility for the data collection at the harvesting of the country concerned by the tuna farming or fattening activity. In the same ICCAT Rec.08-05, art.89, it is established that the ICCAT Regional Observer Programme shall monitor the harvest of bluefin tuna from 100% of the cages and that the data collection will be according to the ICCAT Rec.06-07, including the scientific work. SGRN notes that these provisions do not prevent MS to ensure that the required EC data collection is fully accomplished. SGRN supports the opinion that MS where cages are located shall ensure the data collection at the harvesting and that bilateral agreements are to be established with the Country of the flag fishing vessel(s) concerned.	In its 2008 report, RCMMed&BS identified the regional coordination for sampling large pelagic catches as a very important issue for task-sharing in the Mediterranean. The RCM carried on its work to propose a regional sampling plan for these species to include in the NP 2011-2013. Starting from the values of the total "MS production" a minimum number of samples, for all large pelagic species, have been calculated according to DCF regulation intensities. Furthermore the production of large pelagic has been separated by metier required and a minimum sample size has been set up at regional level both for metier and stock related variables. RCMMed&BS also addressed the problem of tuna transferred to tuna cages in non-member states. RCM reinforces the recommendations made by SGRN and EC/199/08 article 5 for Member states to sample the tuna transferred in non-member countries either themselves or through bi-lateral agreements and to sample according (both for length sampling and for stock related variables). Regional sampling program for large pelagic species (calculation of CV, number of samples to be collected for length and stock related variables) will be annually monitored and evaluated.
ON DATA COLLECTION ON SHARKS CAUGHT BY LARGE PELAGIC FISHING ACTIVITIES. All pelagic sharks caught by all fisheries directed to tuna and tuna-like species in the ICCAT convention area must be reported to ICCAT (Rec.04-10 and Rec.07-06 for all shark species concerned, including task I and task II data, followed by Rec. 08-07 specifically for the bigeye thresher shark, *Alopias superciliosus*). This fact implies that all the pelagic shark species shall be reported in terms of catch and possibly monitored, independently from the fact that they are target or by-catch species. The EC Data Collection Regulation impose the fully implementation of the obligations deriving from the various RFMOs.	The new Appendix VII of the Commission Decision 93/2010 has been discussed and analyzed at Regional Level. The RCMMed&BS pointed out about the necessity of clarify which biological variables should be sampled in each case of Appendix VII. RCM supports the idea to collect, as a first estimation, the metier based variables for these species (i.e. length structures of landings or of catches if sampling at sea). Additionally, RCM noted that the sampling of sharks in the routine concurrent sampling schemes, poses a number of problems for certain metiers. The sampling of just a few shark individuals in these metiers, forces to largely increase the sampling effort, and decrease significantly the efficiency of the sampling for commercial species. It is also stressed that no precision

SGRN remarks that the current list of species includes a category (Sharks-like *Selachii*) which is not allowing the distinction between pelagic and demersal shark species, creating a serious problem about the mandatory sampling of pelagic shark species. SGRN recommends that all MS concerned with large pelagic fisheries, make every effort to report catches of pelagic sharks at the species level and establish the proper sampling for the pelagic species to be reported to ICCAT or other RFMOs. SGRN would point out that this is in line with the new EU Shark Action plan.	target could be reached for Elasmobranches. Therefore, no minimum number or sampling strategy should be associated to the collection of all the "sharks" species reported in the new Appendix VII. Below is reported the recommendation of the RCMMed&BS 2009: RCMMed&BS 2009 recommends "*the MS to check the new list of sharks and decided at national level which species are presented in their catches. RCM Med&BS supports the idea to collect, as a first estimation, the metier based variables for these species (length frequency distribution). RCM Med&BS recommends SGRN to check again and validate the list for Elasmobranches and to clarify which variables should be sampled for each shark species. For the present the letter "T" (every three years) presents in correspondence of* Raja clavata *and* Raja miraletus *should be removed from the Appendix VII. RCM proposes also to delete "shark-like Selachii" from the list of species, if they are added independently in the table.* "
ON DATA COLLECTION FOR FISHING ACTIVITIES USING GEARS NOT LISTED AMONG THE RECOGNISED ONES. If a fishing activity is carried out by a MS by using a gear not officially listed and if this segment is relevant in term of catches or to improve the data used for the stock assessment of the target species concerned, than SGRN recommends that the related sampling shall be properly included in the NP, by using the general gear category and appropriate codification. SGRN recommends that the gear category to be used for the data transmission to the RFMO concerned should use an appropriate codification and encourage co-operation among relevant MS.	Not applicable for IT NP.

VIII List of derogations

List of requests for derogations:

Short title of derogation	NP Proposal section	Derogation approved or rejected[1]	Year of approval or rejection of past requests for derogations
Metièr: Hand and pole line for cephalopods (LHP_LHM_CEP_0_0_0) in the GSA 19	III.C		
Discards: Trammel net for demersal species (GTR_DES_>=16_0_0)	III.C	a	2009
Discards: Set gillnet for demersal fish (GNS_DEF_>=16_0_0)	III.C	a	2009
Discards: Purse seine for small pelagic fish (PS_SPF_>=14_0_0)	III.C	a	2009
Discards: Beach and boat seine for demersal species (SB_SV_DES_0_0_0)	III.C	a	2009
Discards: Set longline for demersal fish (LLS_DEF_0_0_0)	III.C	a	2009
Discards: Drifnets for small pelagic fish (GND_SPF_0_0_0)	III.C	a	2009
Stock: Mugilidae	III.E	a	2009
Stock: *Coryphaena equiselis*	III.E	a	2009
Stock: *Sparus aurata*	III.E	a	2009
Stock: *Dicentrarchus labrax*	III.E	a	2009
Stock related variables – Other regions	III.E	a	2009
Stock: Collection of "Stock variables" related to all shark species with the exception of *Raja clavata*	III.E		
Number of fishing operations for purse seiners	III.F (Effort)	a	2009

Insert 'a' for approved or 'r' for rejected

IX List of acronyms and abbreviations

ALP	Archivio Licenza Pesca – Vessel Register
ASIA	statistical archives of active companies
ATECO code	Codification for sectors of economic activities
CPUE	Catch per Unit of Effort
DCF	Data Collection Regulation
DGPA	Direzione Generale Pesca e Acquacoltura
GFCM	General Fishery Commission for the Mediterranean
GRT	Gross Registered Tonnage
GSA	Geographical Sub Areas
ICCAT	International Commission for the Conservation of Atlantic Tunas
ISTAT	National statistical institute
MEDIAS	Pan-Mediterranean pelagic survey
MEDITS	Mediterranean International Trawl Survey
MIPAF	Ministry of Agricultural and Forestry Policies
RINA	Registro Italiano Navale
SAC	Scientific Advisory Committee of GFCM

X Comments, suggestions and reflections

Concerning the new Appendix VII of (EU Decision 93/2010), which includes a list of all sharks species, the European Commission should clarify which biological variables should be sampled and if precision level should be associated to the collection of both metier and stock related variables. The EU Decision 93/2010 is not enough clear on this issue. We report the recommendation of the RCMMed&BS 2009 "*RCMMed&BS was critical with the (too large) proposed list for the Mediterranean since some of the proposed species are presumably not present in the supra-region. The group pointed out also about the necessity of clarify which biological variables should be sampled in each case of Appendix VII. RCM supports the idea to collect, as a first estimation, the metier based variables for these species (i.e. length structures of landings or of catches if sampling at sea). RCM supports the idea to collect, as a first estimation, the metier based variables for these species (i.e. length structures of landings or of catches if sampling at sea). Additionally, RCM noted that the sampling of sharks in the routine concurrent sampling schemes, poses a number of problems for certain metiers. The sampling of just a few shark individuals in these metiers, forces to largely increase the sampling effort, and decrease significantly the efficiency of the sampling for commercial species. It is also stressed that no precision target could be reached for Elasmobranches. Therefore, no minimum number or sampling strategy should be associated to the collection of all the "sharks" species reported in the new Appendix VII.*"

XI References

Baird J.W., 1983 - A method to select optimum numbers for aging in a stratified random approach. In Sampling commercial catches of marine fish and invertebrates. Edited by W.G. Doubleday and D. Rivard. Can. Spec. Publ. Fish. Aquat. Sci. 66: 161–164.

Bertrand J.A., L. Gil de Sola, C. Papaconstantinou, G. Relini and A. Souplet, 2002 - The General specifications of the Medits surveys. Scientia Marina, 66 (Supl.2): 9-17. Barcelona (Spain).

Caddy J.F., 1999 - Decising on precautionary management measures for a stock based on a suite of Limit Reference Points (LRPs) as a basis for a multi-LRP harvest law. NAFO Scientific Council Studies 32: 55-68.

Ciccotti E., 2005 - Data collection for the European Eel in Italy. In: Dekker W. (ed.) 2005 Report of the Workshop on National Data Collection for the European Eel, Sånga Säby (Stockholm, Sweden). 6 – 8 September 2005, 271-283.

Ciccotti E., 2007 - Report on the eel stock and fisheries in Italy 2007. In: Report of the 2007 Session of the Joint EIFAC/ICES Working Group on Eels Bordeaux. 3–7 September 2007, ICES CM 2007/ACFM:23, 499-508.

Dekker W. (ed.), 2005 - Report of the Workshop on National Data Collection for the European Eel, Sånga Säby (Stockholm, Sweden), 6 – 8 September 2005, 283 pp.

Fischer W., M.L. Bauchot M., Schneider (rédacteurs), 1987 - Fiches FAO d'identification des espèces pour les besoins de la pêche. (Révision 1). Méditerranée et Mer Noire Zone de pêche 37. Rome, FAO, vol 1 et 2, 1530 p.

Garcia S.M., Staples D.J., Chesson J., 2000 - The FAO guidelines for the development and use of indicators of sustainable development of marine capture fisheries and an Australian example of their application. Ocean and Coastal Management 43, 537-556.

Haddon M., 2001 - Modelling and Quantitative Methods in Fisheries. Chapman and Hall/CRC Press: Boca Raton, FL.

Hurlbert S.H., 1971 - The non-concept of species diversity: a critique and alternative parameters. Ecology 52, pp. 577–586.

IREPA 2007 - Osservatorio Economico sulle strutture produttive della pesca marittima in Italia, Franco Angeli.

Italian Technical Reports 2005 - National Data Collection Programs under Council Regulation (EC) N° 1543/2000 and Commission Regulation (EC) N° 1639/2001 as amended by Commission Regulation (EC) N° 1581/2004, 70pp.

Leung P.S., 2006 - Multiple-criteria decision making (MCDM) applications in fishery management. Int. J. Environmental Technology and Management 6 (1/2): 96-110.
Manly B.F.J., 1997 - Randomization, Bootstrap, and Monte Carlo Methods in Biology. 2nd edn. (Chapman and Hall: London.).

Medias 2008 – Report of the "1th Steering Committee Meeting of the MEDiterranean Acoustic Surveys (Medias)", Athens 25-26 February 2008.

Medias 2009 – Report of the "2nd Steering Committee Meeting of the MEDiterranean Acoustic Surveys (Medias)", Palma 1-3 June 2009.

Medits 2007 - International bottom trawl survey in the Mediterranean - Instruction manual. Version 5. Ifremer, Nantes. 60 p.

Pennington M., Burmeister L.M., Hjellvik V., 2002 - Assessing the precision of frequency distributions estimated from trawl-survey samples. Fish. Bull., 100: 74-80.

PGMed 2009 - Final Report of the "3th Meeting of the Mediterranean Planning Group for Methodological Development (PGMed)" Montpellier, 2-6 March 2009, 53 pp.

RCMMed&BS 2009 - Final Report of the "6th Regional Coordination Meeting for the Mediterranean and Black Sea" Venice, 13-16 October 2009.

Rochet M.J., V.M. Trenkel, J.A. Bertrand, J.C. Poulard, 2004 - R routines for survey based fisheries population and community indicators (R-SUFI). Ifremer, Nantes. Limited distribution.

SGRN-08-01 2008 - Commission Staff Working Document "Review of Guidelines for the new DCF" Sub-Group on Research Needs (SGRN), of the Scientific, technical and Economic Committee for Fisheries (STECF). 59pp.

SIBM 2005 - Biological Sampling of Discards – Final Report July 2006, 171 pp., in Italian.

Shin Y.J., Rochet M.J., Jennings S., Field J., Gislason H., 2005 - Using size-based indicators to evaluate the ecosystem effects of fishing. ICES Journal of Marine Science, 62, 384-396.

Souplet A. 1996a - Calculation of abundance indices and length frequencies in the Medits survey. In: J. Bertrand (ed.), *Campagne internationale de chalutage démersal en Méditerranée* - Medits. *Campagne 1995 Vol. III*. Rapport final de contrat CEEIFREMER-IEO-SIBM-NCMR (MED/93/020,018,006,004).

Souplet A. 1996b - Checkmed: a software to check the MEDITS data files. In: J. Bertrand (ed.), *Campagne internationale de chalutage démersal en Méditerranée* - Medits. *Campagne 1995 Vol. I,* pp. 21-26. Rapport final de contrat CEEIFREMER-IEO-SIBM-NCMR (MED/93/020,018,006,004).

Trenkel V.M., J. Rochet, B. Mésnil, 2007 - From model-based prescriptive advice to indicator-based interactive advice. ICES Journal of Marine Science 64: 768-774.

Vigneau J., Mahevas S., 2004 - Precision in catch at age data with regard to sampling design. Working document for the WKSMFD (Nantes 26-31/01/2004): 26 pp.

Vigneau J., 2008 - Common tool for raising and estimating properties of statistical estimates derived from the Data Collection Regulation (COST). European Commission, FISH/2006/15

– lot 2, Project no : SI2.467814. Title: Studies and Pilot projects for carrying out the common fisheries policy.

XII Annexes

Annex I: Methodology for the estimation of economic variables

1. The population and the questionnaire

The survey carried out to estimate economic variables is a multivariate sample survey. The sample unit is the single vessel and this unit is selected from the Vessel Register.

The sampling is of a stratified nature in that the fishing vessels of the fleet are divided into homogenous groups based on suitable variables and independent samples are taken from each of these clusters.

The elementary economic data are collected trough a specific questionnaire. The following table reports the list of primary cost headings. The single cost headings are subsequently aggregated in macro-headings in order to produce the estimates as required by the regulation:

Items in the economic questionnaire	
Gross wages	taxes (irpef, irpeg, irap)
social costs and pension contributions	MUD costs, CONAI, etc.
IRPEF - crewmembers	purchase engine
IPSEMA social security	purchase various tools
accounting and payroll	other fixed costs (1)
legal fees, notary	purchase new fishing gear
insurance on the vessel	purchase new networks
operating costs c / banking	purchase cables
interest expense on c / banking and loans	Shipping customs
expenditure services and trade union membership fee	repair networks (retai)
quay services (water and electricity)	diesel fuel costs
management and storage costs ashore	lubricants costs
travel expenses	expenses of board
stamp duty and motor vehicles insurance	work clothes for crew
maintenance vehicles	purchase bait
purchase vehicles	expenditure telephony board (mobile and CB)
Office supplies	expenditure on-board TV
Boat painting	various expenses for material endowment board
Costs for routine maintenance	Rights fish market
expenditure blacksmith and carpentry	wholesaler commissions
expenditure slipways	other commissions for sales
repair electrical system	porterage fish products
mechanical repair plant (engine)	costs for vehicles to transport fish
plumbing repairs (pumps)	Costs for ice
repair radar	crates and packing
repair refrigerators	other production costs
RINA inspection costs	other costs of sale
VHF tests	purchase new blubox

Costs health certificates	purchase new equipment (radar-sounder-sonar-plotter)
extinguishers	purchase new refrigeration plant
Health certificates	purchase new electrical system
veterinary service	purchase new hydraulic system (pumps)
renewal medicines	blubox quota
health certificates ICCAT	reimbursement fishing stop
VAT and other administrative charges	Revenues extra-fishing
expenditure demands and stamps with PA for bodies	other subsidies (mucilage-military servitude-de minimis)
renewal fishing licences	indebtedness for trust banking
tributes to maritime authorities	indebtedness for mutual
annual quota Chamber of Commerce	indebtedness for other loans
	indebtedness to advance from suppliers

The optimum sample number per stratum is defined according to Bethel's procedure (1989), the vessels are selected using PPS methodology (Probability Proportional to Size) and, to be more exact, using the algorithm of Hanurav-Vijayan. To obtain an estimate of totals per stratum the Horvitz-Thomson estimator is used, while the Sen-Yates-Grundy formula is used to estimate the relative sampling error. Finally this estimate phase or universe carry-over is preceded by a set of control and correction procedures of sample data to guarantee results with a determined level of quality. In each of these phases the data is elaborated using Software R and Statistic 6.

2 The estimation of the sample size: Bethel's procedure

Bethel's procedure (1989) is a mathematical algorithm to achieve the optimum sample allocation in a multivariate sample survey, that is to say the study of several subject variables which are also stratified.

The aim of Bethel's procedure is to ascertain the «minimum cost» of the sample, given the precision limits required for each stratum. The cost C is defined as:

$$C = c_0 + \sum_{h=1}^{H} c_h n_h \qquad (*)$$

where c_0 represents a fixed cost correlated with the organisation of the collection of data, c_h represents the costs of the sampling of a unit within the stratum h-th (h= 1...H), while n_h represents the number of units selected from within the h-th stratum.

Given that the sampling is stratified, the precision limits on the estimate can be expressed as follows[12]:

$$\mathrm{var}\left(\hat{Y}_j\right) = \sum_{h=1}^{H} N_h^{\,2} \left(1 - \frac{n_h}{N_h}\right) \frac{S_{hj}^{\,2}}{n_h} \le \tilde{v}_j^{\,2} \qquad \text{j=1 … J (**)}$$

[12] In Bethel's original article, the correction for finite populations was not considered, and therefore (considering the differences due to the fact that in the article the quantities to be estimated were averages and not totals) the formula (*) in in fact presented as: $\mathrm{var}\left(\hat{Y}_j\right) = \sum_{h=1}^{H} N_h^{\,2} \frac{S_{hj}^{\,2}}{n_h} \le \tilde{v}_j^{\,2}$

where $\hat{Y}_j$ represents the total for the j-th variable (j=1....J), S_{hj}^2 represents an estimate (or a hypothetical value) of the variance of the j-th variable within the h-th stratum and $\tilde{v}_j^2$ represents the threshold level (the limit), in absolute terms, for the value of the variance of the total estimator for the j-th variable.

This set of limited J can be equivalently expressed in an alternative form:

$$\sum_{h=1}^{H} N_h^2 \frac{S_{hj}^2}{n_h} \leq \tilde{v}_j^2 + \sum_{h=1}^{H} N_h S_{hj}^2 \quad \Leftrightarrow \quad \frac{\sum_{h=1}^{H} N_h^2 \frac{S_{hj}^2}{n_h}}{\tilde{v}_j^2 + \sum_{h=1}^{H} N_h S_{hj}^2} \leq 1 \quad \Leftrightarrow \quad \sum_{h=1}^{H} \left(\frac{N_h^2 S_{hj}^2}{\varepsilon_j^2 Y_j^2 + \sum_{h'=1}^{H} N_{h'} S_{h'j}^2} \cdot \frac{1}{n_h} \right) \leq 1$$

where $\hat{Y}_j$ represents the total estimated (or hypothesized) for the variable j-th, and ε_j represents the relative error (absolute error of the estimate divided by the value of the estimate) acceptable for the j-th variable.

Thus using a_{hj} to indicate the term on the left of the product in brackets of the last inequality, with x_h at a value of $1 / n_h$, all the last inequality can be expressed in the form:

$a_j' \underline{x} \leq 1 \quad j=1 \ldots J$

or, equivalently,

$A' \underline{x} \leq \underline{1}$

dove $\underset{(H \text{ x } J)}{A} = \left[a_{hj} \right]$ e $\underset{(H \text{ x } 1)}{\underline{x}}$ the vector of values $1/n_h$.

The whole problem of the minimum limit can be expressed in the following way:

$$\begin{cases} \min_{\underline{x}} g(\underline{x}) = \sum_{h=1}^{H} \frac{c_h}{x_h} \\ s.v.: A' \underline{x} \leq \underline{1} \end{cases}$$

Bethel demonstrated that this problem always has a solution, and that this corresponds to the following formula:

$$x_h^* = \frac{\sqrt{c_h}}{\sqrt{\sum_{j=1}^{J} \alpha_j^* a_{hj}} \cdot \sum_{k=1}^{H} \sqrt{c_k \sum_{j=1}^{J} \alpha_j^* \cdot a_{kj}}}$$

Where the α_j^* are suitable normalised constants (Lagrange multipliers), that is to say those for which $\sum_{j=1}^{J} \alpha_j^* = 1$.

2.1 The solution of the problem of the minimum: Chromy's algorithm

To solve the problem of the minimum limit, Bethel proposes the use of an algorithm which is neither particularly efficient nor easy to apply. At that time, in fact, another algorithm was already available, formulated by Chromy (1987) and also put forward in the same publication of Bethel, which made it easier to find a solution to the problem from the point of view of the development of the code and quicker in terms of elaboration time.

Once the initial values of α_j ,equivalent to1/J, are in place, this algorithm develops fundamentally in two steps, which are repeated continually until reaching an acceptable criteria of convergence.

1. Calculate: $$x_h\left(\underline{\alpha}^{(r-1)}\right)=\frac{\sqrt{c_h}}{\sqrt{\sum_{j=1}^{J}\alpha_j{}^{(r-1)}a_{hj}}\cdot\sum_{k=1}^{H}\sqrt{c_k\sum_{j=1}^{J}\alpha_j{}^{(r-1)}\cdot a_{kj}}}$$

2. Calculate: $$\alpha_j{}^{(r)}=\frac{\alpha_j{}^{(r-1)}\left[a_j{}'\underline{x}\left(\underline{\alpha}^{(r-1)}\right)\right]^2}{\sum_{k=1}^{J}\alpha_k{}^{(r-1)}\left[a_k{}'\underline{x}\left(\underline{\alpha}^{(r-1)}\right)\right]^2}$$

2.2 Input data for the procedure

As input data to start the procedure, the variance estimates for each stratum of variables being studied and the estimates of the totals are needed; these estimates are obtained from the data available, at the time of the analysis, for the most recent year (Income, Personnel costs, Energy costs, Repair and maintenance costs, Other operational costs).

3 The sample selection procedure: the algorithm of Hanurav-Vijayan

The sampling design adopted requires the extraction, without repetition, of the sampling units based on the PPS (probability proportional to size) method; in simpler terms, this sampling plan involves the extraction of various units with a first-order inclusion probability which is not constant, but is proportional to a suitably selected auxiliary variable. The use of such a sampling plan, and thus its use in place of simple random sampling, is justified by the intention of wanting to exploit the information given by the auxiliary variable. This auxiliary variable obviously must be noted for all units in the reference population, and must be «linked» to the unknown variable, the estimate of which is being attempted. This link, in statistical terms, is translated in «proportional relation» between the variable to be estimated and the noted auxiliary variable. The use of information supplied by the auxiliary variable aims to improve the estimate; put in other words, the «stronger» this proportional relation is, the smaller the variability of the estimator (or variance), and so the estimate is much more precise. In the theoretical situation limit of exact proportionality, the estimator would have zero variance and would assume ,in any sample, the exact total to estimate. In the case considered, the noted auxiliary variable is the LOA, the use of which as an accessory variable was preceded by an exploratory analysis, which confirmed the hypothesis of proportionality between the LOA on the one hand, and economic variables on the other (this obviously does not refer to an «exact» relationship between the variables).

The algorithm of Hanurav-Vijayan defines a series of steps to be taken to select a sample of a pre-defined number (n), without replacement, and with a non-uniform probability of each individual unit being included in the sample. By following this algorithm, a sample is obtained which has a series of properties, some of which are worthy of note:

i. $\pi_i = n\, X_i / X$, where π_i represents the inclusion probability (also called probability of inclusion of the first order) of the i-th unit, n indicates the pre-determined size of the sample, Xi represents the size of the noted variable (or «accessory» measure) from which the inclusion probability is calculated and X is the sum of the values X_i for i=1…N, where N is used to denote the size of the universe being sampled. This identity is «required for construction» and necessitates some special treatment in specific circumstances (considered further on).

ii. $\pi_{ij}>0$, where π_{ij} represents the probability (called of the second order) of the simultaneous presence of units i and j. The very fact of being able to determine these probabilities exactly and relatively simply, a consequence of the sampling procedure, is already a notable result which assures the existence of an unbiased estimate of the variance.

iii. $\pi_{ij} \leq \pi_i \pi_j$. This characteristic is notable because it guarantees a positive Sean-Yates-Grundy estimator of the variance of the total

iv. $\pi_{ij} - \pi_i \pi_j > \beta$, for β nor too close to 0. This property guarantees the stability of the Sean-Yates-Grundy total variance estimator.

The values π_i and π_{ij} (for i,j=1…N) satisfy the following two properties:

$$\sum_{i=1}^{N} \pi_i = n$$

$$\sum_{i=1}^{N} \sum_{j>i}^{N} \pi_{ij} = \frac{n\,(n-1)}{2}$$

It is interesting to note that the sum of the probabilities of the first order never equals 1 (unless the sample is composed of only a single unit).The same can be said for probabilities of the second order (unless the sample is composed of only 2 units).It is also to be noted how the application of the formula (i) can sometimes cause the inclusion probability of the first order to be more than 1. In this case corrections in the procedure of sample selection and the probabilities of inclusion must be applied. Specifically, the inclusion probability of the first order is assigned equal to 1, to the k units of which the probability results more than 1, and the n-k units within the entire population are selected, once the unit with the probability of 1 is excluded. It is clear that, once the units with a probability greater than 1 are «set aside» (or rather, selected with a probability of 1), should others with a inclusion probability greater than 1 appear within the remaining N-k, a gradual «setting aside» of these must be provided for, as for all other units, until a population of units with all the probabilities of being selected randomly in the first order inferior to 1, is obtained. Finally a sample is selected of (n-h) units among the (N-h) units of the entire population (where h ($\leq$n) represents the number of units «set aside» or «pre-sampled»).

For the description of the procedure, focus will be placed on the simple random sample. For a stratified sampling it is sufficient to apply the following procedure to every population sub-set.

It is thus presumed, without loss of generality, to have a population composed of N units, pre-ordered with respect to an accessory measure X_j (j=1….N) (in our case this measure was given by the value of the LOA). Thus $X_1 \leq X_2 \leq \ldots \leq X_N$. is obtained.

The following steps are then followed:

a whole number between 1 and n is chosen randomly with a probability

1. $\theta_i = n\,(p_{N-n+i+1} - p_{N-n+i})\,(S + i\,p_{N-n+1}) / S \qquad (i=1 \ldots n)$

   where $p_j = X_j / X$, $S = \sum_{j=1}^{N-n} p_j$ and $p_{N+1} = 1/n$ to ensure that $\sum_{i=1}^{n} \theta_i = 1$

2. If at step (1.) the value i is selected, the last (n-i) elements of the population are selected and the next step is used to obtain the remaining i.

3. New normalized measures are defined in place of p_j , which are then indicated as p_j^*:

$$p_j^*\,(i) = \begin{cases} \dfrac{p_j}{S + i\,p_{N-n+1}} & se\ j \leq N-n+1 \\ \dfrac{p_{N-n+1}}{S + i\,p_{N-n+1}} & se\ N-n+1 < j \leq N-n+i \end{cases}$$

The missing units are selected, in order, using for each selection (indicated by 1 (1=0…i-1), probability values proportional to $a_j(l, j_{l-1})$, where:

$$a_j(l, j_{l-1}) = \begin{cases} (i-l)\ p_j^*\ (i) & \text{se } j = j_{l-1}+1 \\ (i-l)\ p_j^*\ (i) \prod_{k=j_{l-1}+1}^{j-1}\left[1-(i-l-1)\ P_k^*\ (i)\right] & \text{se } j = j_{l-1}+2 \ldots N-n+l+1 \end{cases}$$

where j_l represents the position of the unit selected in the 1-th selection (1=0…i-1) and $P_k^*(i)$ is calculated as:

$$P_k^*(i) = p_k^*\ (i) \Big/ \sum_{h=k+1}^{N-n+i} p_k^*\ (i)$$

For construction, the first-order inclusion probability π_j for the randomly sampled units (thus excluding the «pre-selected» units with a inclusion probability greater than 1) are equal to:

$\pi_j = np_j$ (taking the number of «pre-selected» units to be equal to 0).

The probability of inclusion in the second order is instead equal to:

$$\pi_{ij} = \sum_{r=1}^{n} \theta_r K_{ij}^{(r)}$$

Where

$$K_{ij}^{(r)} = \begin{cases} 1 & \text{se } N-n+r < i \le N-1 \\ \dfrac{r\ p_{N-n+1}}{S+r\ p_{N-n+1}} & \text{se } N-n < i \le N-n+r \text{ e } j > N-n+r \\ \dfrac{r\ p_i}{S+r\ p_{N-n+1}} & \text{se } 0 < i \le N-n \text{ e } j > N-n+r \\ \pi_{ij}^{(r)} & \text{se } j \le N-n+r \end{cases}$$

and where

$$\pi_{ij}^{(r)} = n\,(n-1) P_i^*(r) \prod_{k=1}^{r-1}\left[1-P_k^*(r)\right].$$

4 Estimate of totals and calculation of relative sampling errors

For each variable the estimate of the total is obtained using the Horvitz-Thompson estimator:

$\hat{Y}_h = \sum_{i=1}^{n_h} \frac{y_{hi}}{\pi_{(h)i}}$, where h is the stratum index and i indicates the sampling unit; obviously, by summing all the strata, the estimate of the total of the variable Y is obtained:

$$\hat{Y} = \sum_{h=1}^{H} \hat{Y}_h = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \frac{y_{hi}}{\pi_{(h)i}}.$$

The estimate of the average , in analogy with the estimate of the totals, will be given by $\hat{\bar{Y}}_h = \frac{1}{N_h} \sum_{i=1}^{n_h} \frac{y_{hi}}{\pi_{(h)i}}$, for the single stratum h, by $\hat{\bar{Y}} = \frac{1}{N} \sum_{h=1}^{H} N_h \hat{\bar{Y}}_h = \frac{1}{N} \sum_{h=1}^{H} \sum_{i=1}^{n_h} \frac{y_{hi}}{\pi_{(h)i}} = \frac{\hat{Y}}{N}$ for the total of the variable Y.

For the estimate of the variance of the total the Sen-Yates-Grundy formula is used:

$$\hat{\sigma}^2(\hat{Y}_h) = \sum_{i=1}^{n_h} \sum_{j>i}^{n_h} \left(\frac{\pi_{(h)i}\pi_{(h)j}}{\pi_{(h)ij}} - 1 \right) \left(\frac{y_{hi}}{\pi_{(h)i}} - \frac{y_{hj}}{\pi_{(h)j}} \right)^2$$,for the single stratum h, while, having obtained the sample of H independent selection in each stratum , the total variance is obtained from the sum of the variances from within each single stratum:

$$\hat{\sigma}^2(\hat{Y}) = \sum_{h=1}^{H} \hat{\sigma}^2(\hat{Y}_h) = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j>i}^{n_h} \left(\frac{\pi_{(h)i}\pi_{(h)j}}{\pi_{(h)ij}} - 1 \right) \left(\frac{y_{hi}}{\pi_{(h)i}} - \frac{y_{hj}}{\pi_{(h)j}} \right)^2 .$$

The relationship between the estimate of the standard deviation of the total and the estimate of the total itself, provides the estimate of the sampling error committed ($\hat{\sigma}(\hat{Y}_h)/\hat{Y}_h$ or $\hat{\sigma}(\hat{Y})/\hat{Y}$ depending on whether or not reference is made to the single stratum).

For the estimate of the variance of the population relative to each stratum the formula of Chaudhuri is used:

$$\hat{\sigma}_h^2 = \frac{1}{2N_h^{\,2}} \sum_{i \neq j}^{n_h} \frac{\left(y_{hi} - y_{hj}\right)^2}{\pi_{(h)ij}} = \frac{1}{N^2} \sum_{i=1}^{n_h} \sum_{j>i}^{n_h} \frac{\left(y_{hi} - y_{hj}\right)^2}{\pi_{(h)ij}} .$$

This last value can be used as an input parameter for the procedure of Bethel.

5. Sampling errors

Sampling errors refer to those errors which are encountered in the estimate of a parameter of the universe because of the fact that not all the population, but only a sub-set of it (the sample), is the object of observation. To control this kind of error, reference was made to the procedure of Bethel (1989), as described above, to estimate the optimum sample allocation.

Through this algorithm it is possible to identify the sample size necessary to obtain, for the variables under study, estimates with pre-fixed sampling error levels. In our case the controlled variables are: Income, Personnel costs, Energy costs, Repair and maintenance costs, Other operational costs, and the error required for these is equal to 5%. It is obvious that, in general ,there will not be an exact coincidence between the pre-fixed error and the estimated error afterwards, mainly because it deals with «a sampling estimate of the sampling error» (thus it is also derived from the observation of a sub-set of the population). Moreover, the parameters of input required by the algorithm will be the totals and the variances for Income, Personnel costs, Energy costs, Repair and maintenance costs, Other operational costs, in the strata in which the population was sub-divided. Obviously, in the estimate phase of the sampling size, those values relating to the following year in which the survey will carried out, can only be hypothesized on the base of what had been observed in the year preceding the survey, exactly because it really deals with the same values which we intend to estimate. It is however evident from the data that the application of such a method allows the control of the sampling error around pre-fixed levels.

6 Non-sampling errors: control and localization

Non-sampling errors are those which are directly connected to the elementary data and are revealed as the difference between the value y_i of the variable Y, observed in the i-th unit, and the real value Y_i . It is thus obvious that, all other factors being equal, if the sampling error diminishes with the increase in sample size (annulling itself for census), this will not, in general, be true for the non-sampling error.

An initial important classification of this type of error makes the distinction between complete non-response (CNR) and partial non-response (PNR). In general, in every survey, for every sampling unit, responses are gathered from a fixed number Q of questions. CNR is taken to refer to a statistical unit which does not supply responses to any of the Q questions given, while PNR is taken to refer to a statistical unit for which no information is available for a sub-set of Q questions. In reality, the distinction between these two types of missing responses is not as evident as it may appear. Generally, in fact, a statistical unit for which only a sub-set of questions is available, may still however be considered an CNR if it deals with a number of data too low with respect to Q, or also in the case where the variables not noted are considered strategic and of fundamental importance to the survey. In speaking of «non-availability» of information, reference is made not only to the cases of missing values (obviously identified), but also to the case where the value of a response to a question does not correspond to the reality, actually observed in the unit. The methods of control and localization of errors aim to identify these values.

In general, the control procedure of the survey in question can be considered as interactive graphic micro-editing of the univariate type. The term interaction refers to the fact that, in the procedure of localization of errors, there are not only automatic phases but also phases which require human intervention to investigate the situation and to evaluate the effective presence of the error (therefore the human intervention regards the localization phase and not that of imputation). The control is mainly of the univariate type because the variables are controlled individually and only in rare cases are suspected relationships existing among them controlled, using suitable synthesis indexes. During the various phases wide use is made of graphic tools to visibly evaluate situations marked as errors. Finally the word micro-editing is used because the data is gathered in suitable domains of study within which the sampling units can be considered very homogenous. For each of these sets of data, suitable control functions are first calculated, and then, for each of them, certain rules of incompatibility are verified. In the case of activation of conditions of error that is in the case where the observed value does not belong to the region of acceptance those control functions are then observed individually for all the sampling units forming the single domain. Thus the sample unit, or units, responsible for the activation of conditions of error is localized for the entire domain of study and then the imputation of the erroneous data follow.
The control functions adopted for the localization of errors will then be daily costs (so, for example: other operational costs / days, personnel costs /days, energy costs /days) and the ratio between costs and revenues (other operational costs / revenues, personnel costs /revenues, energy costs /revenues, and os on). As already anticipated, it deals with micro-editing, and so the sample units are grouped into suitable domains of study, which, in the first phase of control, coincide with the same strata derived from the initial stratification of the universe (stratification based on varying geographical regions, fishing techniques, dimensional class of the vessel). Within each stratum the vessels can result very similar to each other and thus the dispersion of the functions of control only lightly dispersed; it would not be wrong, in general, to consider these dispersions close to normal distribution. The control procedure thus occurs in four levels for each species revealed and, in each phase, aggregations of sample units in ever wider domains of study are obtained. The aggregations of vessels in each of the four phases occurs, step by step, without taking into consideration one of the three variables of stratification. Thus, in the first level, all the strata obtained based on the classification of fishing techniques, geographical region of origin and on the base of dimensional class are observed. In the second level, vessels grouped according to geographical region and fishing methods are observed, while in the third the domains of study are obtained considering the sole variable of fishing techniques and, finally, at the last level, the total aggregate for Italy is considered. In each of these phases, the two functions of control aggregated for each stratum obtained are observed. A region of acceptance for them is set out, obtained by an analysis which is both transversal and longitudinal. Thus there will be an activation of the conditions of error by these functions of control (or a value outside the region of acceptance) in the case of a value «too far» from the same indicator observed in the preceding year of the survey. Such threshold limits for the definition of the regions of acceptance, in general, are obtained through the observation of historical series of data. Once therefore an anomalous situation for a domain of observation is revealed, by proceeding backwards in re-controlling the data in the various levels preceding that where the error was localized, eventually the analysis of daily costs for the single sample unit is reached. Experts are usually present at such phases of control and they evaluate, in each phase of activation of the rules of error, whether the data is effectively erroneous. It is possible, in fact, that in time the distribution of functions of control undergo variations in position and dispersion and that the state of error marked by the functions of control are therefore not actually due to the effective presence of anomalous data, but to a structural variation in its distribution (in such cases, therefore, it is necessary to reconsider the threshold limits of the region of acceptance). During such phases, in order to achieve a correct execution of the procedure of localization of errors, the help of graphic tools is fundamental. For a generic cost variable, the procedure ends when no states of error are indicated for any of the strata generated in the various levels.

7 Non-sampling errors: imputation of non-response errors

With the procedure of localization of errors (as regards the PNR) the identification of a set of data is reached for which the condition of error has been activated by at least one of the two functions of control: costs / days and costs / revenues.

Therefore, once the presence of an error has been ascertained, the imputation is done using deterministic methods. In particular, the imputation is done using the following formula:

$$\hat{c}(s,i) = \overline{c}(s) \cdot g(i)$$

where:

- $g(i)$ is the number of fishing days of the i-th vessel;
- $\overline{c}(s)$ is the average cost of the strata s (the i-th vessel belongs to strata s);
- $\hat{c}(s,i)$ is the imputed cost for the i-th vessel.

8 Clustering of segments

Table III.B.2 reports the segments that have been clustered. Clusters are named after the biggest segment in terms of number of vessels.

Clustering is necessary in order to design the sampling plan and to report economic variables. The economic sample is stratified by segments according to Appendix III and by geographical sub areas (GSA).

This double level of stratification of the population (technical and geographical) may generate very small strata that have to be grouped in order to get a statistical sample. When a strata is too small (less than 10 vessels) it is very difficult to randomly select a sample. At the same time, the sampling plan is subject to budget constraints and clustering of small segments is also necessary to reach cost efficiency.

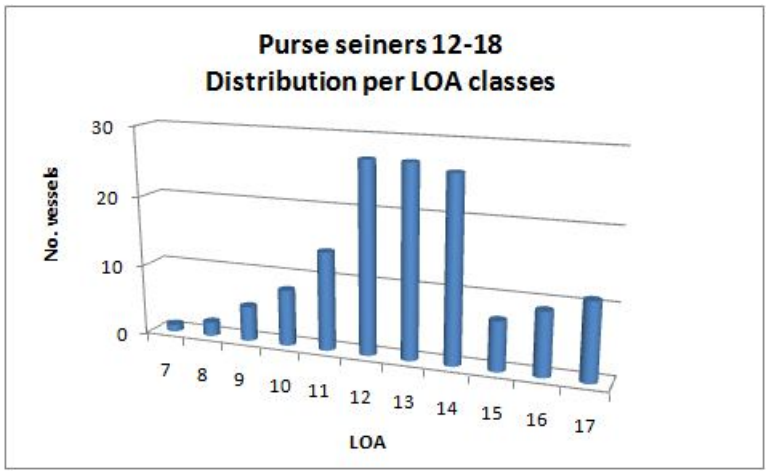
The proposed clustering also guarantees continuity in the time series.

In the following section, the scientific evidence justifying the clustering is explained for each clustering reported in table III.B.2.

Purse seiners 12-18 m*

Name of the clustered fleet segments	*Total number of vessels in the cluster*	*Fleet segments which have been clustered*	*No. Of vessels*
purse seiners 12-18 m*	143	purse seiners >-12<18 m	113
		purse seiners >-6<12 m	30

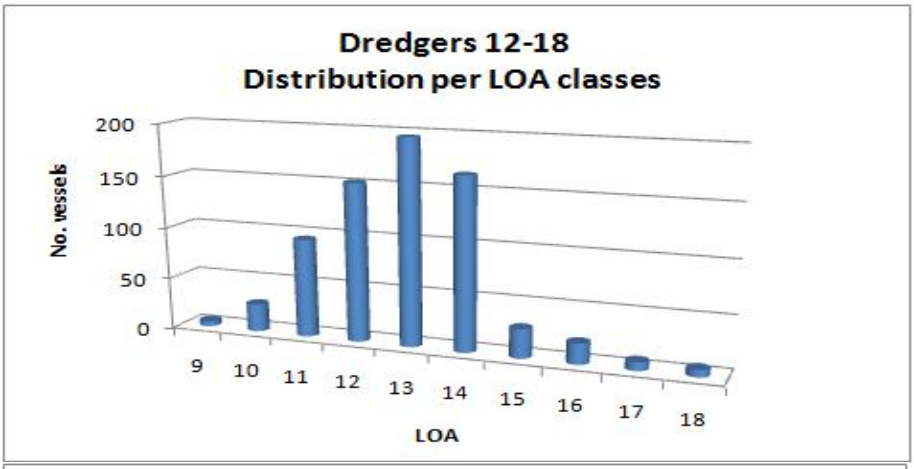
The clustered segment (purse seiners 12-18 m*) is composed by 143 vessels with an average LOA of 13.5 m. Vessels are concentrated near the average value, as shown by the graph. Therefore the clustered segment is homogenous from a statistical point of view.

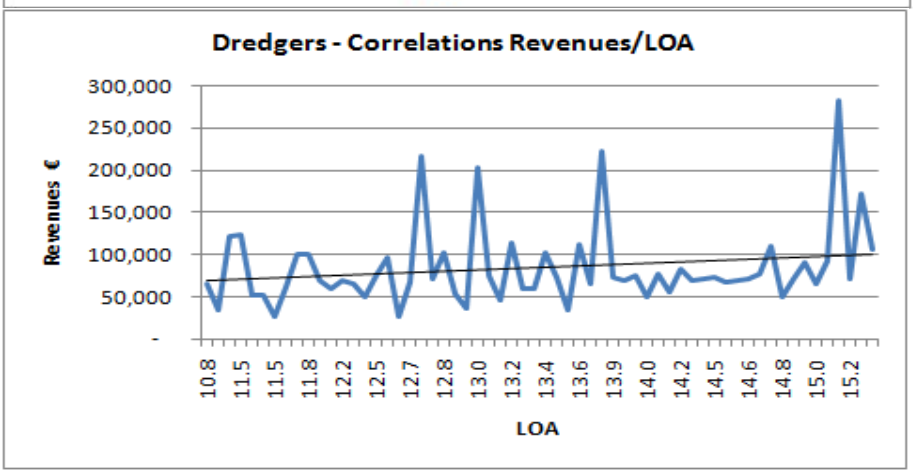


Dredgers >-12<18 m*

Name of the clustered fleet segments	*Total number of vessels in the cluster*	*Fleet segments which have been clustered*	*No. Of vessels*
dredgers 12<18 m*	703	dredgers >-6<12 m	127
		dredgers >-12<18 m	569
		dredgers >-18<24 m	7

Dredgers are based almost exclusively in central-north Adriatic cost. Vessels are very specialised targeting only clams and smooth-callista (*Venus gallina* and *Callistachione*) and they are homogenous in terms of size, gears and fishing practises. Vessels have an average LOA of 13.4 meters and 81% of them belong to the class 12-18 meters. Therefore, the split into the class <12 m and > 12 meters is not statistical reliable for this segment. Moreover, it is demonstrated that revenues are not correlated with the LOA of the vessels (see graph) and this fact proves the high level of homogeneity of the vessels.

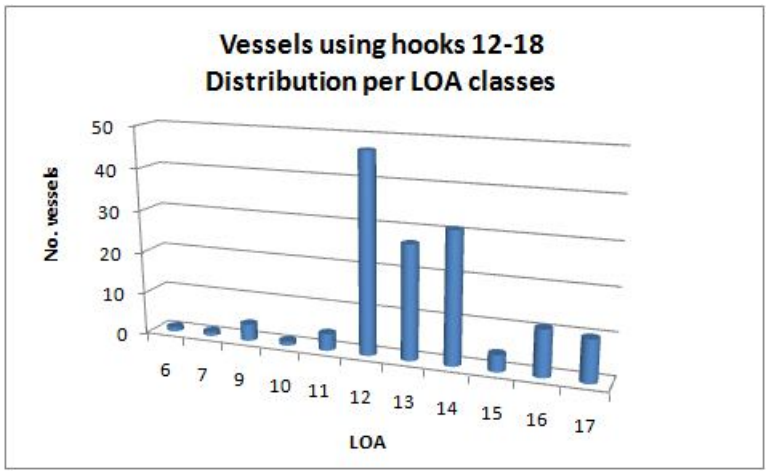




Vessels using hooks 12-18 m *

Name of the clustered fleet segments	*Total number of vessels in the cluster*	*Fleet segments which have been clustered*	*No. Of vessels*
vessels using hooks 12-18 m *	143	vessels using hooks >-6<12 m	11
		vessels using hooks >-12<18 m	132

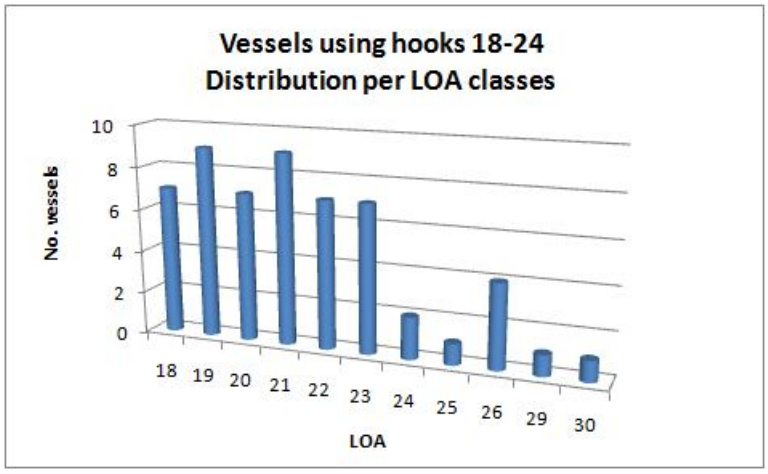
This is composed by 11 vessels and they operate in different areas (GSA 10, GSA 16, GSA 18, GSA 19). The sampling plan is stratified by area, therefore the rule that allows the clustering of segments with less than 10 vessels, is applied in each GSA. The distribution of the vessels per LOA classes shows a concentration around the average value of 13.8.



Vessels using hooks >-18<24 m*

Name of the clustered fleet segments	*Total number of vessels in the cluster*	*Fleet segments which have been clustered*	*No. Of vessels*
vessels using hooks 18-24 m*	53	vessels using hooks 18-24 m	44
		vessels using hooks 24-40 m	9

The class 24-40 m is composed by 9 vessels, 5 of which operate in GSA 19, 3 in GSA 16 and 1 in GSA 10. The average LOA of these vessels is 26.7 metres, therefore they are quite homogenous in terms of size with the vessels in the LOA class 18-24 m (whose average length is 21 meters). All these vessels operate in the same way, exploiting the same fishing grounds and targeting the same species.



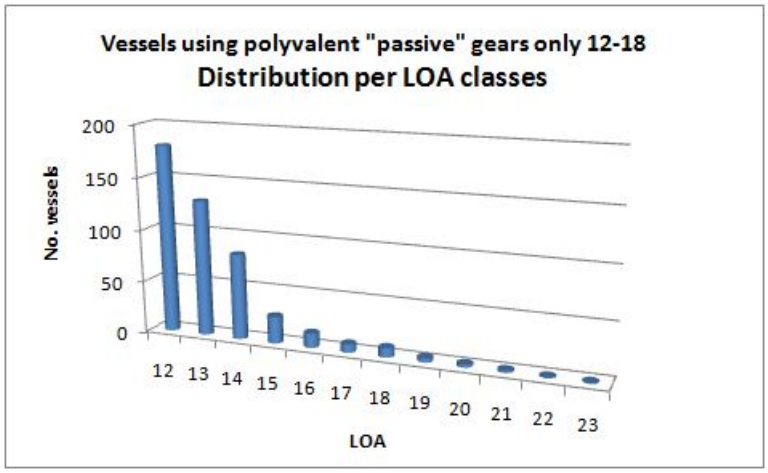
Vessels using polyvalent "passive" gears only >-12<18 m*

Name of the clustered fleet segments	*Total number of vessels in the cluster*	*Fleet segments which have been clustered*	*No. Of vessels*
vessels using polyvalent "passive" gears only >-12<18 m*	440	vessels using polyvalent "passive" gears only >-12<18 m	420
		vessels using polyvalent "passive" gears only >-18<24 m	20

The clustering of the 20 vessels >18 m into the lower LOA class is necessary in order to design the sampling plan. In fact these vessels are scattered along the Italian coast (GSA 9, GSA 10, GSA 11, GSA 17, and GSA 19).

The average length of these 20 vessels is 19.6 meters, therefore very close to the upper limit of the 12-18 m class.

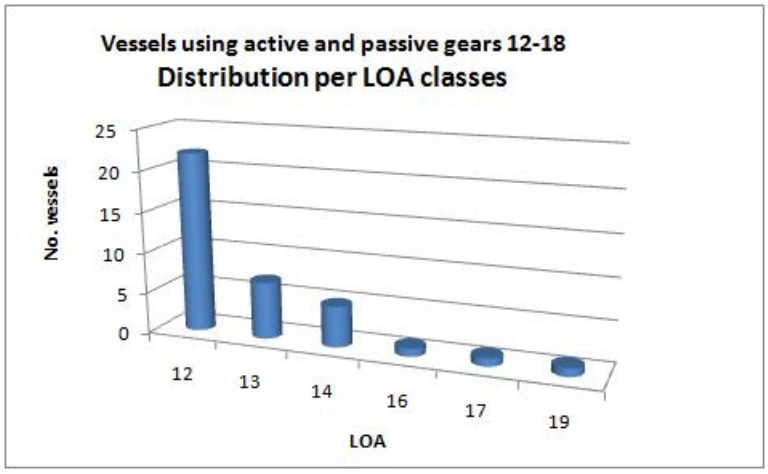
The graph shows the high concentration of the vessels on the left side of the distribution.



Vessels using active and passive gears >-12<18 m*

Name of the clustered fleet segments	*Total number of vessels in the cluster*	*Fleet segments which have been clustered*	*No. Of vessels*
vessels using active and passive gears >-12<18 m*	37	vessels using active and passive gears >-12<18 m	36
		vessels using active and passive gears >-18<24 m	1

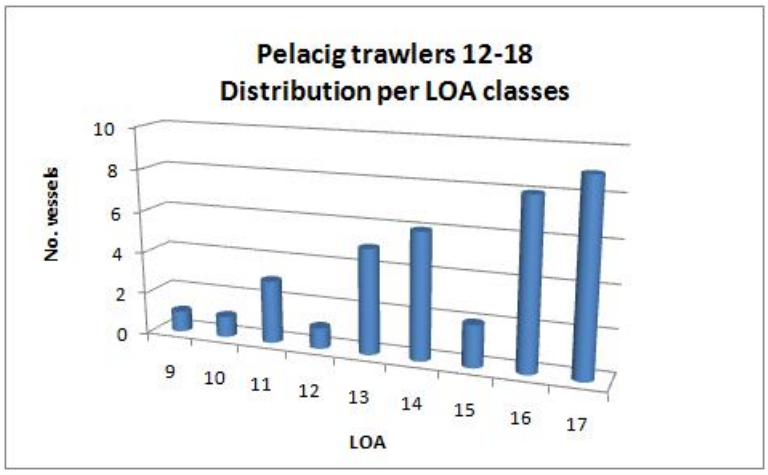
There is only one vessel in the class 18-24 meters. It is obvious that this vessel is grouped in the neighbouring class because it is impossible to get a statistical random sample from a stratum of only one vessel.



Pelagic trawlers >-12<18 m*

Name of the clustered fleet segments	*Total number of vessels in the cluster*	*Fleet segments which have been clustered*	*No. Of vessels*
pelagic trawlers >-12<18 m*	32	pelagic trawlers >-6<12 m	5
		pelagic trawlers >-12<18 m	27

The segment less than 12 meters is grouped in the higher class because this segment is too small to be randomly sampled. The graph shows the concentration of the vessels on the right side of the distribution.



Annex II: Methodology for the estimation of demography of landings (for demersal and small pelagic species) related to the biological metier related variables

Length frequency distribution (LFD)

The sampling will be accomplished according to the methods of a *two-stage stratified random sampling* (Cochran, 1977) and carried out monthly with strata represented by a combination of geographical sub-areas (GSA) and metier.

The sampling unit belonging to the metier (primary unit) will be the fishing trip (secondary unit), thus at least three samples will be collected in each quarter for each metier. The numbers of fishing trips in which the biological sampling of landings will be collected are proportional to the fishing activity of the fleet. The sampling units will be extracted with equal probability and without replacement, using the national commercial landings database for the selection process.
When possible the LFD obtained quarterly will be raised to the quarter production by species and metier, according to the proportion between sampled and landed weight. Then the total LFD will be obtained summing up the estimates by quarter and metier.
The precision, in terms of coefficient of variation (CV) of the length frequency distributions (LFDs) will be estimated annually or three annually (following Appendix VII Decision 93/2010) by metier and GSA.

The general framework is aiming at providing:

- information on the exploitation pattern of the target species disaggregated by metier;
- information on the demographic structure of the landings (size/age);
- information on the demographic structure of catches and estimates of discards for relevant metier;
- information on the growth.

To be consistent with the above objectives, the sampling scheme has been designed taking into account:

a) the fishing capacity, techniques and effort of the Italian fleet;

b) the spatial and temporal variability of the landings.

The biological metier-related variables regard the demography of landings, in terms of length and age structure, of those species that can be aged.

References

Cochran W.G. 1977. Sampling Techniques (3rd edition). John Wiley & Sons, N Y, USA, 428 pp.

Annex III: Methodology for the estimation of discards - Biological metier related variables

Discard sampling survey: quantity estimation by species

In the case of discard the biological metier-related variables regard:

- the evaluation of total discards by target species and metier
- the demography of the discarded fraction of catches, in terms of length and age structure of those species that can be aged.

Demography will be accomplished in case the discard will be higher than 10% in weight or 20% in number.

The sampling will be accomplished according to the methods of a *two-stage stratified random sampling* (Cochran, 1977) and carried out monthly with strata represented by quarters, metiers and GSA. The monitoring will be accomplished with observer onboard when possible; otherwise a self-sampling will be performed with checks at landing time.

Regarding the amount of discards, on the volume of landings, we will evaluate the discarded fraction of species commercially exploited using a ratio estimator (ratio estimator - Cochran, 1977).

This rate will be calculated seasonally and annually.

The ratio estimator of discard (R) is calculated as:

$$\hat{R} = \frac{\sum_{k=1}^{n} D_k}{\sum_{k=1}^{n} (L_k)}$$

with discard (D) and landing (L) in weight per trip *j*, vessel *k*, métier *m* and for species *i*.

n= number of vessels sampled and N= total number o vessel per metièr

Variance will be expressed as:

$$Var\left[\hat{R}\right] = \frac{(1-f)}{n\bar{L}^2}\left(s_D{}^2 + \hat{R}^2 s_L{}^2 - 2\hat{R}s_{DL}\right)$$

f =n/N; S_D= sampling variance of discards; S_L= sampling variance of landing and S_{DL}= sampling covariance

On the basis of the total landings (L_T) it will provides an estimate of the total deviation (T_D) using the formula:

$$D_T = \hat{R} \cdot L_T$$

Variance can be estimated as:

$$Var\left[D_T\right] = (1-f)\left(\frac{L_T^2}{L_C} \cdot Var\left[\hat{R}\right]\right)$$

with L_C= landing observed; f =L_C/L_T and Var R= rate of discards variances.

References

Cochran W.G. 1977. Sampling Techniques (3rd edition). John Wiley & Sons, N Y, USA, 428 pp.

Annex IV: Methodology to estimate the precision level – Metièr related variables

The precision, in terms of coefficient of variation (CV) of the length frequency distributions (LFDs) will be estimated annually or three annually (following Appendix VII Decision 949/2008) by metier and GSA.

The method developed by Vigneau and Mahevas (2004) will be adopted. It's allows to estimate the precision, in terms of coefficient of variation (CV) for each length class and for the whole LFD at metier level.

The CV for each length class will be estimated according to the following procedure:

- calculation for each length class of the DELTA^2 function. This is the square difference between the following quantities:
    1. the absolute number of individuals (d_j) for each class (d_{jkv}) in a given sample (v) of a stratum (k) and
    2. the product of the weight (landing in weight) of each sample (w_k) by the sum of items in each length class (LFD_j) divided by the summed weight of samples (w_{kv}) (sum of the landings)
- computation of variance for length class as sum of the DELTA^2 which will be used for CV estimation.

The main formulas used for CV estimation are:

- Number of landed fish for each length class j (D_j):

$$\hat{D}_j = \sum_{k=1}^{K} \frac{W_k}{\sum_{v=1}^{n_k} w_{kv}} \left(\sum_{v=1}^{n_k} d_{jkv} \right) = \sum_k W_k \frac{\sum_v d_{jkv}}{\sum_v w_{kv}}$$

where W is the total landing and n the number of samples.

- Variance for each length class j (Var D_j):

$$\mathrm{var}\left(\hat{D}_j\right) = \sum_k W_k^2 \,\mathrm{var}\left(\frac{\sum_v d_{jkv}}{\sum_v w_{kv}} \right)$$

- Coefficient of variation CV_j:

$$CV_j = \frac{\sqrt{\mathrm{var}\left(\hat{D}_j\right)}}{\hat{D}_j}$$

The estimator ($\hat{D}_j$) and the variance var ($\hat{D}_j$) and, consequently the CV, will be estimated first quarterly and then yearly for each metier. Method details, formulation and application are reported in Vigneau & Mahevas (2004).
Besides the CV in vector form (length class related), scalar CV will be furnished as weighted mean of the number of individual for each length class. This will be computed considering all the examined length classes (Mp tot), the length class contributing more than 2.5% to the total landing in number (Mp>2.5%) and the length class being 90% of the total landing in number (Mp 90%).

References

Vigneau J., Mahevas S. 2004. Precision in catch at age data with regard to sampling design. Working document for the WKSMFD (Nantes 26-31/01/2004): 26 pp.

Annex V: Methodology to estimate the stock related variables and the growth and reproduction parameters

Maturity

The attribution of the stages of maturity to the gonads is implemented trough the use of empirical macroscopic scale. The references scales are those contained in annexes VIII and IX of the Protocol "*MEDITS – handbook - Instruction manual* version 5 rev" (Medits, 2007).

Exceptions are represented by *Engraulis encrasicolus* and sardine *Sardina pilchardus* specimens sampled during the biological sampling of catches and acoustic survey MEDIAS. In this case will be adopted the maturity scales developed in WKSPMAT (ICES, 2008). Other exception is represented by the crustaceans *Aristeus antennatus*, *Aristaeomorpha foliacea*, *Parapenaeus longirostris* and *Nephrops norvegicus* for which will be adopted maturity scales developed in WKMSC (ICES, 2010).

Regarding the biological sampling of large pelagics, we will use the ICCAT references (http://www.iccat.int/Documents/SCRS/Manual/CH4/CH4_8-ENG.pdf).

Sex ratio

The sex ratio is calculated as the ratio between the number of females and number of males. This ratio can be estimated at the global level, expressed here as a single value, given the number of females in the total of individuals as:

$$Sr = \frac{\sum_{i=1}^{n} FE_i}{\sum_{i=1}^{n} (FE_i + MA_i)}$$

where FE are females, MA males and I indicates the different length classes. To test whether the deviation from an expected ratio of 0.5 is significant is possible to apply the χ^2test:

$$\chi^2 = \frac{\left(FE_{obs} - FE_{\exp}\right)^2}{FE_{\exp}} + \frac{\left(MA_{obs} - MA_{\exp}\right)^2}{MA_{\exp}}$$

If the expected sex ratio (Sr) is 0.5, and the number of samples is large enough, the variance of the proportions can be calculated as:

$$Var(Sr) = \sqrt{\frac{Sr * (1 - Sr)}{N_s}}$$

where Ns is the number of the observations.

Length-weight relationship

The length-weight relationship is described by the following equation:

$P = a L^b$

where *a* is a constant defined empirically and *b* the exponent. The exponent *b* has a value close to 3 with isometric growth. This ratio is calculated, either for separate and/or combined sex (M + F + indeterminate), on all the different segments and metier. The curves are calculated for the seasons and the whole year. The curve, traditionally, will be liberalized using a logarithmic transformation of data and the fit obtained by linear regression:

$L_n P = L_n a + b(L_n L)$

The parameters of length-weight relationship can also be calculated trough a nonlinear estimation by minimizing the sum of squared differences between observed and expected values. The coefficient of determination R^2 will provide an index of goodness of fit of the model and the standard error of estimate of the dependent variable y is:

$$e.s. = \sqrt{\frac{1}{n-2}} \cdot \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2$$

where n is the number of pairs in length (x_i) and weight (y_i).

Growth and Natural Mortality

The growth will be illustrated by graphs showing the progression of lengths with increasing age and described by the classical von Bertalanffy model (VBGF):

$Lt = L((1-\exp(-K(t-t_o)))$

where Lt is the length at age t, L it's the “asymptotic" length or maximum mean size that an individual would reach if it lived indefinitely; K is the growth rate, which represent the rate of average speed with which the individual reaches its maximum size or “asymptotic”. The value t_0 is defined as the theoretical age of an individual with a length of 0.

The natural mortality M will be estimated using two assumptions: constant or variable depending by the size.

In the first case can be calculated with the invariants of Beverton & Holt (Jensen, 1996):

$$M = c * K$$

giving to the constant *c* one of the values reported in the literature for fish and crustaceans. Alternatively, if you have information on the longevity of the species may use the nearest Hewitt and Hoenig (2005):

$$M = \frac{4.22}{t_{\max}}$$

where the maximum age (t_{max}) is based on direct estimates of growth on otoliths or is calculated from the Taylor (3 / K) approximation that correspond to age in which the cohort declined by 1% compared to its original size.

In the second scenario may be adopted a vector of M calculated, for example, according to Abella et al, (1997) or Chen & Watanabe (1989) that require essentially a set of growth parameters of von Bertalanffy and the relation size / length. The last of the methods mentioned require not only a pattern of growth coherent with the model of von Bertalanffy, but also that the value of t_0 will be negative and is maintained within certain limits.

Gonad-somatic index

The gonad-somatic index (gonad weight / gutted weight total * 100) is useful for defining the period of greatest gonad development of a population. For an accurate estimate of the evolution of their development, the gonads must be sampled at different times of the year (preferably on a monthly or, at most, quarterly basis).

Length at maturity or age at maturity (Lm or Tm)

The size (or age) at which individuals reach the development of the gonads may depend on several factors, but generally there is an age "biologically determined" to which at least part of individuals of a population reach in a given period of the year, the sexual maturity. The data, length and stage of maturity, are used to estimate the parameters of the classical maturity ogive, respectively for males and females, according to the model:

$$p_{m,l} = \frac{1}{1 + Exp\left(-g * \left(l_i - l_{m50\%}\right)\right)}$$

where $p_{m,l}$ indicates the proportion of mature individuals in length class l_i, g is the parameter of curvature and $l_{m50\%}$ represents the length at which 50% of individuals is sexually mature (see SAMED 2002).

Parameters of the maturity ogive are calculated as non-linear estimation of the best fit obtained by minimizing the sum of the deviation between observed and expected values.

References

Abella A., Ria M., Rustighi C., Cecchi E., Serena F., 2006 – Time series analysis of commercial landings of five species in the port of Viareggio. Biol. Mar. Medit., 13 (1): 19-29.

Chen S., Watanabe S., 1989 - Age Dependence of Natural Mortality Coefficient in Fish Population Dynamics. *Nippon Suisan Gakkaishi*, 55(2): 205-208.

ICES, 2008 - Report of the "Workshop on Small Pelagics (*Sardina pilchardus, Engraulis encrasicolus*) maturity stages (WKSPMAT)", 10-14 November 2008 Mazara del Vallo, Italy, 82pp.

ICES, 2010. Report of the Workshop on crustaceans (*Aristeus antennatus, Aristaeomorpha foliacea, Parapenaeus longirostris, Nephrops norvegicus*) maturity stages (WKMSC), 19-23 October 2009, Messina, Italy. ICES CM 2009/ACOM:46. 77 pp.

Jensen A.L. 1996 – Beverton and Holt life history invariants result from optimal trade-off of reproduction and survival. *Can. J. Fish. Aquat. Sci.*, 53: 820-822.

Medits 2007 - International bottom trawl survey in the Mediterranean - Instruction manual. Version 5. Ifremer, Nantes. 60 p.

Rochet M.J., V.M. Trenkel, J.A. Bertrand, J.C. Poulard, 2004 - R routines for survey based fisheries population and community indicators (R-SUFI). Ifremer, Nantes. Limited distribution.

SAMED 2002 - Stock Assessment in the Mediterranean. Final Report EU Project n° 99/047.

Annex VI: Methodology to estimate the precision level – Stock related variables

Other biological parameters: Age

The estimate of the individual fish age will be organised in age-length-keys (ALK). Age-length key will be constructed for the aged target species and for metier combined**.**

The precision of the age length keys expressed in terms of coefficient of variation (CV) will be estimated for each age group according to the method proposed by Baird (1983).

The total number of individuals for a given age group is calculated as:

$$N = \sum_i N_i * p_i$$

where

N_i= number of individuals for length class *i*;

p_i= proportion of individuals of a given age group for length class *i*;

$N_i p_i$ =number of individuals for length class *i* belonging to a given age group;

$n_i p_i$=number of individuals whose otolith were read for ageing for length class *i* belonging to a given age group.

Variance for each length class *i* is calculate according to Gulland (1966) as:

$$Var(N_i \cdot p_i) = N_i^2 Var(p_i) + p_i^2 Var(N_i)$$

The second term of the above equation is related to the variability associated with the length measurement and can b considered negligible, thus assuming that age groups are distributed by length according to a binomial function we have:

$$\text{var}(p_i) = \frac{p_i \cdot (1 - p_i)}{n_i}$$

with

n_i= number of individuals "read" for length class *i*, i.e. all the fish whose age was estimated in length class *i*.

The variance of total individuals of a given age group is calculated as sum of variance for each length class in which there are individuals of the age group as:

$$\text{var}(N) = \sum_{i=1}^{L} N_i^2 \text{var}(p_i)$$

where *L* are the length class in which individuals of a given age group are found.

Finally CV for a given age group is calculated as:

$$CV = \frac{\sqrt{\text{var}(N)}}{\sum_{i=1}^{L} N_i p_i}$$

Other biological parameters: sex, weight and maturity

The analytical approach to calculate empirically the coefficients of precision of biological variables (maturity, sex and weight) and their variances follow two sample designs:

a) random sampling in the case that the variables are expressed per length class and assuming that the information observed in the sample, directly reflects the population;

b) the two-stage stratified random sampling by length, where the sample of measured individuals to provide a length distribution representative of the population will further sampled (sub sample) to extract individuals which gather information relating to biological variables under consideration (eg 10 fish per length class).

The standard formulas for estimating the mean and variance of the average of one of the biological variables examined (maturity, sex ratio or weight) per length class are:

$$\bar{x}_l = \frac{1}{n_l} \sum_{l_i = l} x_l$$

$$var\left(\bar{x}_l\right) = \frac{1}{n_l\left(n_l - 1\right)} \sum_{l_i = l} \left(x_l - \bar{x}_l\right)^2$$

where x is the maturity (0 immature, 1 mature), or the sex (0 male, 1 female) or weight and n_l is the number of individuals in the length class *l*.

In the case of the estimates by age rather than length class, the estimate must take into account the stratified sampling for the appropriate length and should be used statistic "weights" (w).

These weights are formulated so that for each age group the sum of the weights is equal to the number of individuals for whom it was detected biological parameters.

The weighted mean and its variance are expressed as:

$$\bar{x}_a^w = \frac{1}{\sum w_i} \sum_{a_i = a} w_i x_i$$

$$var\left(\bar{x}_a^w\right) = 1 - \sum w_i \left(\sum w_i - 1\right) \sum_{a_i = a} w_i \left(x_i - \bar{x}_a^w\right)^2$$

The statistical weights for each age group are calculated using the following procedure:

1) for each individual, which has been detected a biological parameter, first of all is defined a raising factor:

$r_l = n_l / m_l$

where n_l is the total number of individuals measured in the length class *l*, and m_l is the number of subsamples in the same class of fish length in which has been collected the biological parameters.

2) then will be calculated the sum of the raising factors for each age group as:

$$R_a = \sum_{a_i = a} r_l$$

where a_i denotes the age of the individual.

3) then will be assigned the statistical weight to the individual *i* in the length class *l* and in the age group using the formula:

$wi=m_a r_l/R_a$

where m_a d is the number of individuals at that age which has been found biological parameter to be examined
For each age group, the sum of statistical weights will be equal to the number of individuals on which the biological parameter has been collected.

Concerning the precision, related to the relationship between length and biological parameters (weight, maturity, sex and age), of the models used, it will be expressed in terms of CV calculated as standard error of the estimate or the precise value of the parameter.

References

Baird J.W., 1983 - A method to select optimum numbers for aging in a stratified random approach. In Sampling commercial catches of marine fish and invertebrates. Edited by W.G. Doubleday and D. Rivard. Can. Spec. Publ. Fish. Aquat. Sci. 66: 161–164.

Annex VII: Methodology for the estimation of transversal variables

The survey is a multivariate sample survey in that it is characterised by various goals which are also stratified in order to obtain monthly landings per species (Reg. CEE No.1921/06). Along with the related data of landings by species, the variables needed to estimate the fishing effort are also obtained.

The target population (universe) is thus made up of the set landings on the territory independently of whether the vessels belong to the Italian fishing fleet or not. Landings by vessels from other countries is however presumed non – existent and thus the observed population can be defined as a set of Italian fishing vessels. The sample set unit is therefore the single vessel and this unit was selected from the list of the Vessel Register. All vessels regularly supplied with a Fishing Licence are registered with the Vessel Register. Finally, the framework of the survey based on the individual fishing vessel as the sample unit makes it possible to observe and estimate other variables such as the fishing effort.

The sampling is of a stratified nature in that the fishing vessels of the fleet are divided into homogenous groups based on suitable variables and independent samples are taken from each of these clusters. The optimum sample number per stratum is defined according to Bethel's procedure (1989), the vessels are selected using PPS methodology (Probability Proportional to Size) and, to be more exact, using the algorithm of Hanurav-Vijayan. To obtain an estimate of totals per stratum the Horvitz-Thomson estimator is used, while the Sen-Yates-Grundy formula is used to estimate the relative sampling error. Finally this estimate phase or universe carry-over is preceded by a set of control and correction procedures of sample data to guarantee results with a determined level of quality. In each of these phases the data is elaborated using Software R and Statistic 6.

1. Collection of data: the data collection network and the questionnaire

The data collectors are people from within the production and management field of the fisheries sector The professional categories currently used most frequently as data collectors are the following: biologists, ship owners, ex-fishermen, business consultants. Since they form part of the sector, they can easily contact the owners of vessels and are usually present when the fish is landed.

The usual procedure for obtaining data is divided in two phases.

The first phase is at the moment when the fish is landed, usually between 10.00pm and 2.00am, where the number of crates of fish by species is obtained. In this sense it is important to note that the data collectors are experts in the taxonomy of fisheries products and, in some cases, are highly qualified in marine biology.

The second phase consists of the interview. Thanks to the trust developed over time, the natural mistrust of the operators (captains and fishermen on vessels being surveyed) has been overcome and further information on the levels of production, prices, fishing areas and activity (days and hours of fishing and time in the hold) is obtained.

This 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.

2. The questionnaire

The transcription of data for each individual vessel is completely computerised through a software programme elaborated for the specific aims of the gathered data and is structured in a sequential manner with filters aimed at avoiding partial or inconsistent of the compilation of the questionnaires.

The order of questions can be defined as «funnel-shaped» in that they start with general questions to then lead to the gathering of target information of the survey. The first part contains general information such as:

- The name of the vessel (in code)

- Gears used
- Days of activity at sea
- Rest days
- Bad weather days
- Total number of hours
- Number of fishing trips

The second part obtains information relating to:

- Species caught
- Quality
- Quantity (kgs)
- Average prices or revenue
- Destination (market, wholesale, retail, other)

3. Organisation of the collection of information

Following the random extraction of vessels to be studied, the list is submitted to the network of data collectors, from the relevant area, who then 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 present in the sale notes. In this phase, rights from the regulations on the respect of privacy are waived

Periodically, at least once a year, training courses are organised for the data collectors to update them on possible new developments in software and also to further their knowledge on aspects linked to the identification of fish species. In the latter case, the professionalism and experience of marine biology experts in the fisheries sector, constantly involved in surveys on the evaluation of fisheries resources, are used.

4. Stratification

In order to obtain the best possible division of the fleet, the following criteria are used:

1. Geographical area of the registration of the vessels
2. Technical segmentation based on the fishing gears used most frequently
3. Vessels size

6. The estimation of the sample size: Bethel's procedure

Bethel's procedure (1989) is a mathematical algorithm to achieve the optimum sample allocation in a multivariate sample survey, that is to say the study of several subject variables which are also stratified.

The aim of Bethel's procedure is to ascertain the «minimum cost» of the sample, given the precision limits required for each stratum. The cost C is defined as:

$$C = c_0 + \sum_{h=1}^{H} c_h n_h \qquad (*)$$

where c_0 represents a fixed cost correlated with the organisation of the collection of data, c_h represents the costs of the sampling of a unit within the stratum h-th (h= 1...H), while n_h represents the number of units selected from within the h-th stratum.

Given that the sampling is stratified, the precision limits on the estimate can be expressed as follows[13]:

$$\mathrm{var}\left(\hat{Y}_j\right) = \sum_{h=1}^{H} N_h^2 \left(1 - \frac{n_h}{N_h}\right) \frac{S_{hj}^2}{n_h} \leq \tilde{v}_j^2 \qquad \text{j=1 … J (**)}$$

where $\hat{Y}_j$ represents the total for the j-th variable (j=1....J), S_{hj}^2 represents an estimate (or a hypothetical value) of the variance of the j-th variable within the h-th stratum and $\tilde{v}_j^2$ represents the threshold level (the limit), in absolute terms, for the value of the variance of the total estimator for the j-th variable.
This set of limited J can be equivalently expressed in an alternative form:

$$\sum_{h=1}^{H} N_h^2 \frac{S_{hj}^2}{n_h} \leq \tilde{v}_j^2 + \sum_{h=1}^{H} N_h S_{hj}^2 \;\Leftrightarrow\; \frac{\sum_{h=1}^{H} N_h^2 \frac{S_{hj}^2}{n_h}}{\tilde{v}_j^2 + \sum_{h=1}^{H} N_h S_{hj}^2} \leq 1 \;\Leftrightarrow\; \sum_{h=1}^{H} \left(\frac{N_h^2 S_{hj}^2}{\varepsilon_j^2 Y_j^2 + \sum_{h'=1}^{H} N_{h'} S_{h'j}^2} \cdot \frac{1}{n_h} \right) \leq 1$$

where $\hat{Y}_j$ represents the total estimated (or hypothesized) for the variable j-th, and ε_j represents the relative error (absolute error of the estimate divided by the value of the estimate) acceptable for the j-th variable.
Thus using a_{hj} to indicate the term on the left of the product in brackets of the last inequality, with x_h at a value of $1 / n_h$, all the last inequality can be expressed in the form:

$a_j' \underline{x} \leq 1 \quad j=1 \ldots J$

or, equivalently,

$A' \underline{x} \leq \underline{1}$

dove $\underset{(H \times J)}{A} = [a_{hj}]$ e $\underset{(H \times 1)}{\underline{x}}$ the vector of values $1/n_h$.

The whole problem of the minimum limit can be expressed in the following way:

$$\begin{cases} \min_{\underline{x}} g(\underline{x}) = \sum_{h=1}^{H} \frac{c_h}{x_h} \\ s.v.: A' \underline{x} \leq \underline{1} \end{cases}$$

Bethel demonstrated that this problem always has a solution, and that this corresponds to the following formula:

---

[13] In Bethel's original article, the correction for finite populations was not considered, and therefore (considering the differences due to the fact that in the article the quantities to be estimated were averages and not totals) the formula (*) in in fact presented as: $\mathrm{var}\left(\hat{Y}_j\right) = \sum_{h=1}^{H} N_h^2 \frac{S_{hj}^2}{n_h} \leq \tilde{v}_j^2$

$$x_h^* = \frac{\sqrt{c_h}}{\sqrt{\sum_{j=1}^{J} \alpha_j^* a_{hj}} \cdot \sum_{k=1}^{H} \sqrt{c_k \sum_{j=1}^{J} \alpha_j^* \cdot a_{kj}}}$$

Where the α_j^* are suitable normalised constants (Lagrange multipliers), that is to say those for which $\sum_{j=1}^{J} \alpha_j^* = 1 \cdot$

As input data to start the procedure, the variance estimates for each stratum of variables being studied and the estimates of the totals are needed; these estimates are obtained from the data available, at the time of the analysis, for the most recent year. For the sample of the year 2010 the landings and earnings of the 12 species of fish most representative of the national production, which form more than 75% of the total, were considered: Anchovies, Sardines, European Hake, Striped Mullet, Red Mullet, Other Fish, Clams, Squid, Octopus, Horned and musky octopuses, Deepwater rose shrimp, Norway lobster and Spottail mantis squillid.

7. The sample selection procedure: the algorithm of Hanurav-Vijayan

The sampling design adopted requires the extraction, without repetition, of the sampling units based on the PPS (probability proportional to size) method; in simpler terms, this sampling plan involves the extraction of various units with a first-order inclusion probability which is not constant, but is proportional to a suitably selected auxiliary variable. The use of such a sampling plan, and thus its use in place of simple random sampling, is justified by the intention of wanting to exploit the information given by the auxiliary variable. This auxiliary variable obviously must be noted for all units in the reference population, and must be «linked» to the unknown variable, the estimate of which is being attempted. This link, in statistical terms, is translated in «proportional relation» between the variable to be estimated and the noted auxiliary variable. The use of information supplied by the auxiliary variable aims to improve the estimate; put in other words, the «stronger» this proportional relation is, the smaller the variability of the estimator (or variance), and so the estimate is much more precise. In the theoretical situation limit of exact proportionality, the estimator would have zero variance and would assume ,in any sample, the exact total to estimate. In the case considered, the noted auxiliary variable is the LOA, the use of which as an accessory variable was preceded by an exploratory analysis, which confirmed the hypothesis of proportionality between the LOA on the one hand, and the quantity fished and revenue on the other (this obviously does not refer to an «exact» relationship between the variables).

The algorithm of Hanurav-Vijayan defines a series of steps to be taken to select a sample of a pre-defined number (n), without replacement, and with a non-uniform probability of each individual unit being included in the sample. By following this algorithm, a sample is obtained which has a series of properties, some of which are worthy of note:

v. $\pi_i = n X_i / X$, where π_i represents the inclusion probability (also called probability of inclusion of the first order) of the i-th unit, n indicates the pre-determined size of the sample, Xi represents the size of the noted variable (or «accessory» measure) from which the inclusion probability is calculated and X is the sum of the values X_i for i=1…N, where N is used to denote the size of the universe being sampled. This identity is «required for construction» and necessitates some special treatment in specific circumstances (considered further on).

vi. $\pi_{ij} > 0$, where π_{ij} represents the probability (called of the second order) of the simultaneous presence of units i and j. The very fact of being able to determine these probabilities exactly and relatively simply, a consequence of the sampling procedure, is already a notable result which assures the existence of an unbiased estimate of the variance.

vii. $\pi_{ij} \leq \pi_i \pi_j$. This characteristic is notable because it guarantees a positive Sean-Yates-Grundy estimator of the variance of the total

viii. $\pi_{ij} - \pi_i \pi_j > \beta$, for β nor too close to 0. This property guarantees the stability of the Sean-Yates-Grundy total variance estimator.

The values π_i and π_{ij} (for i,j=1…N) satisfy the following two properties:

$$\sum_{i=1}^{N}\pi_i = n$$

$$\sum_{i=1}^{N}\sum_{j>i}^{N}\pi_{ij} = \frac{n\,(n-1)}{2}$$

It is interesting to note that the sum of the probabilities of the first order never equals 1 (unless the sample is composed of only a single unit).The same can be said for probabilities of the second order (unless the sample is composed of only 2 units).It is also to be noted how the application of the formula (i) can sometimes cause the inclusion probability of the first order to be more than 1. In this case corrections in the procedure of sample selection and the probabilities of inclusion must be applied. Specifically, the inclusion probability of the first order is assigned equal to 1, to the k units of which the probability results more than 1, and the n-k units within the entire population are selected, once the unit with the probability of 1 is excluded. It is clear that, once the units with a probability greater than 1 are «set aside» (or rather, selected with a probability of 1), should others with a inclusion probability greater than 1 appear within the remaining N-k, a gradual «setting aside» of these must be provided for, as for all other units, until a population of units with all the probabilities of being selected randomly in the first order inferior to 1, is obtained. Finally a sample is selected of (n-h) units among the (N-h) units of the entire population (where h (≤n) represents the number of units «set aside» or «pre-sampled»).

8. Estimate of totals and calculation of relative sampling errors

For each variable the estimate of the total is obtained using the Horvitz-Thompson estimator:

$\hat{Y}_h = \sum_{i=1}^{n_h}\frac{y_{hi}}{\pi_{(h)i}}$, where h is the stratum index and i indicates the sampling unit; obviously, by summing all the strata, the estimate of the total of the variable Y is obtained:

$$\hat{Y} = \sum_{h=1}^{H}\hat{Y}_h = \sum_{h=1}^{H}\sum_{i=1}^{n_h}\frac{y_{hi}}{\pi_{(h)i}}.$$

The estimate of the average , in analogy with the estimate of the totals, will be given by $\hat{\bar{Y}}_h = \frac{1}{N_h}\sum_{i=1}^{n_h}\frac{y_{hi}}{\pi_{(h)i}}$, for the single stratum h, by $\hat{\bar{Y}} = \frac{1}{N}\sum_{h=1}^{H}N_h\hat{\bar{Y}}_h = \frac{1}{N}\sum_{h=1}^{H}\sum_{i=1}^{n_h}\frac{y_{hi}}{\pi_{(h)i}} = \frac{\hat{Y}}{N}$ for the total of the variable Y.

For the estimate of the variance of the total the Sen-Yates-Grundy formula is used:

$\hat{\sigma}^2(\hat{Y}_h) = \sum_{i=1}^{n_h}\sum_{j>i}^{n_h}\left(\frac{\pi_{(h)i}\pi_{(h)j}}{\pi_{(h)ij}} - 1\right)\left(\frac{y_{hi}}{\pi_{(h)i}} - \frac{y_{hj}}{\pi_{(h)j}}\right)^2$,for the single stratum h, while, having obtained the sample of H independent selection in each stratum , the total variance is obtained from the sum of the variances from within each single stratum:

$$\hat{\sigma}^2(\hat{Y}) = \sum_{h=1}^{H}\hat{\sigma}^2(\hat{Y}_h) = \sum_{h=1}^{H}\sum_{i=1}^{n_h}\sum_{j>i}^{n_h}\left(\frac{\pi_{(h)i}\pi_{(h)j}}{\pi_{(h)ij}} - 1\right)\left(\frac{y_{hi}}{\pi_{(h)i}} - \frac{y_{hj}}{\pi_{(h)j}}\right)^2.$$

The relationship between the estimate of the standard deviation of the total and the estimate of the total itself, provides the estimate of the sampling error committed ($\hat{\sigma}(\hat{Y}_h)/\hat{Y}_h$ or $\hat{\sigma}(\hat{Y})/\hat{Y}$ depending on whether or not reference is made to the single stratum).

For the estimate of the variance of the population relative to each stratum the formula of Chaudhuri is used:

$$\hat{\sigma}_h^2 = \frac{1}{2N_h^{\,2}}\sum_{i\neq j}^{n_h}\frac{\left(y_{hi}-y_{hj}\right)^2}{\pi_{(h)ij}} = \frac{1}{N^2}\sum_{i=1}^{n_h}\sum_{j>i}^{n_h}\frac{\left(y_{hi}-y_{hj}\right)^2}{\pi_{(h)ij}}.$$

This last value can be used as an input parameter for the procedure of Bethel.

9. Procedure for controlling and correcting data

The phase of controlling and correcting data consists in identifying and treating errors present in the data gathered in the survey, in the aim of guaranteeing final results with specified levels of quality. For the creation of software to manage data, errors which characterize the survey were classified as sampling and non-sampling errors.

9.1 Sampling errors

Sampling errors refer to those errors which are encountered in the estimate of a parameter of the universe because of the fact that not all the population, but only a sub-set of it (the sample), is the object of observation. To control this kind of error, reference was made to the procedure of Bethel (1989), as described above, to estimate the optimum sample allocation.

Through this algorithm it is possible to identify the sample size necessary to obtain, for the variables under study, estimates with pre-fixed sampling error levels. In our case the variables of interest are the total landings and the revenues observed in Italy for the main groups of harvested species, and the error generally required for these is equal to 3.5%. For the year of the survey, therefore, first the optimum sample size (based on the error limit required), and then the sampling error itself was estimated for each month. It is obvious that, in general ,there will not be an exact coincidence between the pre-fixed error and the estimated error afterwards, mainly because it deals with «a sampling estimate of the sampling error» (thus it is also derived from the observation of a sub-set of the population). Moreover, the parameters of input required by the algorithm will be the totals and the variances for landings and revenues of the species controlled, in the strata in which the population was sub-divided. Obviously, in the estimate phase of the sampling size, those values relating to the following year in which the survey will carried out, can only be hypothesized on the base of what had been observed in the year preceding the survey, exactly because it really deals with the same values which we intend to estimate. It is however evident from the data that the application of such a method allows the control of the sampling error around pre-fixed levels.

Sampling error (cv) estimated for the landings and revenues from the main species caught in 2008

Species	***Landings error***	***Revenues error***
Anchovies	0.0504	0.0510
Sardines	0.1222	0.1810
Squid	0.0389	0.0325
Deepwater rose shrimp	0.0527	0.0495
Other Fish	0.0299	0.0306
European Hake	0.0243	0.0278
Red Mullet	0.0434	0.0308
Striped Mullet	0.0936	0.0664
Norway lobster	0.0431	0.0369
Octopus	0.0354	0.0363
Clams	0.0186	0.0268

9.2 Non-sampling errors: control and localization

Non-sampling errors are those which are directly connected to the elementary data and are revealed as the difference between the value y_i of the variable Y, observed in the i-th unit, and the real value Y_i . It is thus obvious that, all other factors being equal, if the sampling error diminishes with the increase in sample size (annulling itself for census), this will not, in general, be true for the non-sampling error.

An initial important classification of this type of error makes the distinction between complete non-response (CNR) and partial non-response (PNR). In general, in every survey, for every sampling unit, responses are gathered from a fixed number Q of questions. CNR is taken to refer to a statistical unit which does not supply responses to any of the Q questions given, while PNR is taken to refer to a statistical unit for which no information is available for a sub-set of Q questions. In reality, the distinction between these two types of missing responses is not as evident as it may appear. Generally, in fact, a statistical unit for which only a sub-set of questions is available, may still however be considered an CNR if it deals with a number of data too low with respect to Q, or also in the case where the variables not noted are considered strategic and of fundamental importance to the survey. In speaking of «non-availability» of information, reference is made not only to the cases of missing values (obviously identified), but also to the case where the value of a response to a question does not correspond to the reality, actually observed in the unit. The methods of control and localization of errors aim to identify these values.

In general, the control procedure of the survey in question can be considered as interactive graphic micro-editing of the univariate type. The term interaction refers to the fact that, in the procedure of localization of errors, there are not only automatic phases but also phases which require human intervention to investigate the situation and to evaluate the effective presence of the error (therefore the human intervention regards the localization phase and not that of imputation). The control is mainly of the univariate type because the variables are controlled individually and only in rare cases are suspected relationships existing among them controlled, using suitable synthesis indexes. During the various phases wide use is made of graphic tools to visibly evaluate situations marked as errors. Finally the word micro-editing is used because the data is gathered in suitable domains of study within which the sampling units can be considered very homogenous. For each of these sets of data, suitable control functions are first calculated, and then, for each of them, certain rules of incompatibility are verified. In the case of activation of conditions of error, that is in the case where the observed value does not belong to the region of acceptance, those control functions are then observed individually for all the sampling units forming the single domain. Thus the sample unit, or units, responsible for the activation of conditions of error is localized for the entire domain of study and then the imputation of the erroneous data follows. The collection of information on the subject of the study occurs constantly over the year and is based on the concept of the fishing trip (defined ad hoc for the collection of data): with the time interval measured in days and at least equal to one, during which fishing activity takes place. The data collector thus indicates the starting and ending date of the fishing trip, the days of activity, or days where fishing activity effectively take place, the species caught, the quantity in kg for each species caught, the revenue in euro from the sale of each species caught. It is important to note that the length of the fishing trip (obtained from the difference between the starting and ending dates of the fishing trip) will always be greater or equal to the days of activity. Since the estimates at a universe level are obtained by taking a month as the time scale of reference, the procedures of control and localization of errors were also developed on a monthly basis.

The variables obtained are thus days, landings and revenues. The control functions adopted for the localization of errors will then be daily landings (landing/day) and prices (revenue/landing), totaled for the month considered. As already anticipated, it deals with micro-editing, and so the sample units are grouped into suitable domains of study, which, in the first phase of control, coincide with the same strata derived from the initial stratification of the universe (stratification based on varying geographical regions, fishing techniques, dimensional class of the vessel). Within each stratum the vessels can result very similar to each other and thus the dispersion of the functions of control only lightly dispersed; it would not be wrong, in general, to consider these dispersions close to normal distribution. The control procedure thus occurs in four levels for each species revealed and, in each phase, aggregations of sample units in ever wider domains of study are obtained. The aggregations of vessels in each of the four phases occurs, step by step, without taking into consideration one of the three variables of stratification. Thus, in the first level, all the strata obtained based on the classification of fishing techniques, geographical region of origin and on the base of dimensional class are observed. In the second level, vessels grouped according to geographical region and fishing methods are observed, while in the third the domains of study are obtained considering the sole variable of fishing techniques and, finally, at the last level, the total aggregate for Italy is considered. In each of these phases, and for each species, the two functions of control aggregated for each stratum obtained (total landing of the stratum/total days of the stratum and total revenue of the stratum/ total landing of the stratum) are observed. A region of acceptance for them is set out, obtained by an analysis which is both transversal and longitudinal. Thus there will be an activation of the conditions of error by these functions of control (or a value outside the region of acceptance) in the case of a value «too far» from the same indicator observed in the preceding month of the survey and/or in the same month of the preceding year.

Such threshold limits for the definition of the regions of acceptance, in general, are obtained through the observation of historical series of data. Once therefore an anomalous situation for a domain of observation is revealed, by proceeding backwards in re-controlling the data in the various levels preceding that where the error was localized, eventually the analysis of daily landings and prices for the single sample unit is reached. Experts are usually present at such phases of control and they evaluate, in each phase of activation of the rules of error, whether the data is effectively erroneous. It is possible, in fact, that in time the distribution of functions of control undergo variations in position and dispersion and that the state of error marked by the functions of control are therefore not actually due to the effective presence of anomalous data, but to a structural variation in its distribution (in such cases, therefore, it is necessary to reconsider the threshold limits of the region of acceptance). During such phases, in order to achieve a correct execution of the procedure of localization of errors, the help of graphic tools is fundamental. For a generic species, the procedure ends when no states of error are indicated for any of the strata generated in the various levels.

9.3 Non-sampling errors: imputation of non-response errors

With the procedure of localization of errors (as regards the PNR) the identification of a set of data is reached for which the condition of error has been activated by at least one of the two functions of control: landings / days and revenues / landings (prices). Therefore once the presence of an error has been ascertained, it is necessary to define which of the variables covered by the function of control are not exact (if necessary, both of the variables) to be able to proceed correctly to the imputation.

Among the variables investigated, a sort of hierarchy of reliability is used which becomes a reference point in the different phases of the procedure. The first control is effected on the daily landings: when confronted with erroneous data, marked on the basis of controls carried out in the different levels previously described, the distribution of the fishing days of the vessels belonging to the same stratum of the suspected sample unit is controlled. It is important to note, however, that when confronted with a situation of error marked by the function of control, the possibility that it concerns the day instead of the landings is considered unlikely, and, consequently, the region of acceptance for evaluating the presence of errors in such a variable is quite wide. Moreover, the fishing days of a vessel, relative to a fishing trip, is information which is easily reconstructed through a second interview with the sample unit and is generally the path taken to reconstruct this variable. If it is difficult to obtain this information through a second interview, the method of imputation, subsequently described, is applied.

Therefore, summarizing, if following the activation of the conditions of error caused by the relationship between landings and days, the variable day is considered correct, imputation is applied directly to the landings. On the contrary, if the days are considered to be erroneous, these will then be reconstructed through a second interview or method of imputation and it is again verified whether the function of control falls within its own region of acceptance. If it does not, the variable landing is then estimated.

Once the control of daily productivity (landings/ days) is finished, it is obvious that the variables days and landings are considered correct. Going on, therefore, to the control of prices, imputation of revenues is carried out when a probable error is shown.

For the real and proper process of imputation, reference is made to deterministic methods, and in particular, deal with imputation of historical series with values «adjusted» to take into account changes in trends. In particular, it is necessary to distinguish between the case where the sample unit which shows the erroneous data is also present in the disclosure of the previous year and that where this does not occur.

In the first case, when the «behaviour» of the vessel is known for the previous year, the following formulae are used:

$$k_1 = \frac{\bar{g}(str(i),t)}{\bar{g}(str(i),t-12)}$$

$$k_2(s) = \frac{p(s,t)}{p(s,t-12)}$$

$$\hat{g}(i,t) = g(i,t-12)\cdot k_1$$

$$\hat{c}(s,i,t) = \frac{c(s,i,t-12)}{g(i,t-12)}\cdot \hat{g}(i,t) = c(s,i,t-12)\cdot k_1$$

$$\hat{r}(s,i,t) = p(s,i,t-12)\cdot k_2(s)\cdot \hat{c}(s,i,t) = \ldots = r(s,i,t-12)\cdot k_1 \cdot k_2(s)$$

$\bar{g}(str(i),t)$ is the average number of fishing days in the stratum of the i-th vessel, for the month, t, of the current year. This average is calculated considering only the data for vessels which do not show errors. Thus, k_1 represents the index of variation of fishing days between the month t and the same month in the previous year.

$p(s,t)$ indicates the price for the s-th species, in the month t, for the aggregate of Italy; this value is obtained considering only the data for vessels which do not show errors. Thus k_2 represents the index of global variation for the price of species s.

g(i,t-12) indicates the days of activity for the i-th vessel in the same month of analysis, but for the previous year.

$c(s,i,t-12)$, $r(s,i,t-12)$, $p(s,i,t-12)$, represents the landings, revenues and prices (revenue over landing) for species s, for vessel i in the month of analysis, relative to the year preceding the analysis.

Finally, $\hat{g}(i,t)$, $\hat{c}(s,i,t)$ e $\hat{r}(s,i,t)$ represent the estimates of days, landings and revenues (vessel i, species s, current month t).

It is interesting to note that this estimate procedure makes the days and landings vary on the basis of the general course of the stratum of fishing activity, (leaving the daily productivity of the vessel unvaried), and makes the revenues vary on the basis of the variation of fishing activity in the stratum of reference and on the basis of the general variation of the price of species for the entire aggregate of Italy.

The procedure just described assumes that the vessel with erroneous data is present in the survey for the previous year. When this is not so, a different estimate procedure is used:

$$k_1(s) = \frac{\dfrac{\sum_{h=1}^{t-1} c(s,i,h)}{t-1}}{\dfrac{\sum_{j\neq i=1}^{n}\sum_{h=1}^{t-1} c(s,j,h)}{(t-1)(n-1)}}; \qquad k_2(s) = \frac{\dfrac{\sum_{h=1}^{t-1} r(s,i,h)}{t-1}}{\dfrac{\sum_{j\neq i=1}^{n}\sum_{h=1}^{t-1} r(s,j,h)}{(t-1)(n-1)}};$$

$$\hat{g}(i,t) = \frac{\sum_{j\neq i=1}^{n} g(j,t)}{n-1}; \qquad \hat{c}(s,i,t) = k_1(s)\frac{\sum_{j\neq i=1}^{n} c(s,j,t)}{n-1}; \qquad \hat{r}(s,i,t) = k_2(s)\frac{\sum_{j\neq i=1}^{n} r(s,j,t)}{n-1}.$$

The variables assume the same meaning as previously described.

It is evident that, in this second case, k_1 and k_2 represent measures of «distance» of the vessel from the rest of the stratum. This distance is then used to extract the missing data.

These two procedures were derived in a context of estimating PNRs. According to the case, that is, depending on whether the vessel is present or not in the past year, the same procedure is then used for the estimate of the complete non-responses. This means therefore that the formulae just described are applied for all species s, relative to the sample unit which represents the CNR.

Annex VIII: Methodology for the estimation of aquaculture variables and estimation of the CV

The estimation of the sample size: Bethel's procedure

Bethel's procedure (1989) is a mathematical algorithm to achieve the optimum sample allocation in a multivariate sample survey, that is to say the study of several subject variables which are also stratified. The aim of Bethel's procedure is to ascertain the "minimum cost" of the sample, given the precision limits required for each stratum. The cost C is defined as:

$$C = c_0 + \sum_{h=1}^{H} c_h n_h \qquad (*)$$

where c_0 represents a fixed cost correlated with the organisation of the collection of data,

c_h represents the costs of the sampling of a unit within the stratum h-th (h= 1...H),

while n_h represents the number of units selected from within the h-th stratum.

Given that the sampling is stratified, the precision limits on the estimate can be expressed as follows[14]:

$$\mathrm{var}\left(\hat{Y}_j\right) = \sum_{h=1}^{H} N_h^{\,2} \left(1 - \frac{n_h}{N_h}\right) \frac{S_{hj}^{\,2}}{n_h} \le \tilde{v}_j^{\,2} \qquad \text{j=1 … J (**)}$$

where $\hat{Y}_j$ represents the total for the j-th variable (j=1....J),

$S_{hj}^{\,2}$ represents an estimate (or a hypothetical value) of the variance of the j-th variable within the h-th stratum and

$\tilde{v}_j^{\,2}$ represents the threshold level (the limit), in absolute terms, for the value of the variance of the total estimator for the j-th variable.

This set of limited J can be equivalently expressed in an alternative form:

$$\sum_{h=1}^{H} N_h^{\,2} \frac{S_{hj}^{\,2}}{n_h} \le \tilde{v}_j^{\,2} + \sum_{h=1}^{H} N_h S_{hj}^{\,2} \;\Leftrightarrow\; \frac{\sum_{h=1}^{H} N_h^{\,2} \frac{S_{hj}^{\,2}}{n_h}}{\tilde{v}_j^{\,2} + \sum_{h=1}^{H} N_h S_{hj}^{\,2}} \le 1 \;\Leftrightarrow\; \sum_{h=1}^{H} \left(\frac{N_h^{\,2} S_{hj}^{\,2}}{\varepsilon_j^{\,2} Y_j^{\,2} + \sum_{h'=1}^{H} N_{h'} S_{h'j}^{\,2}} \cdot \frac{1}{n_h} \right) \le 1$$

where $\hat{Y}_j$ represents the total estimated (or hypothesized) for the variable j-th,

[14] In Bethel's original article, the correction for finite populations was not considered, and therefore (considering the differences due to the fact that in the article the quantities to be estimated were averages and not totals) the formula (*) in in fact presented as: $\mathrm{var}\left(\hat{Y}_j\right) = \sum_{h=1}^{H} N_h^{\,2} \frac{S_{hj}^{\,2}}{n_h} \le \tilde{v}_j^{\,2}$

and ε_j represents the relative error (absolute error of the estimate divided by the value of the estimate) acceptable for the j-th variable.

Thus using a_{hj} to indicate the term on the left of the product in brackets of the last inequality, with x_h at a value of $1 / n_h$, all the last inequality can be expressed in the form:

$a_j' \underline{x} \leq 1 \quad j=1 \ldots J$

or, equivalently,

$A' \underline{x} \leq \underline{1}$

dove $\underset{(H \text{ x } J)}{A} = \lfloor a_{hj} \rfloor$ e $\underset{(H \text{ x } 1)}{\underline{x}}$ the vector of values $1/n_h$.

The whole problem of the minimum limit can be expressed in the following way:

$$\begin{cases} \min_{\underline{x}} g(\underline{x}) = \sum_{h=1}^{H} \dfrac{c_h}{x_h} \\ s.v.: A' \underline{x} \leq \underline{1} \end{cases}$$

Bethel demonstrated that this problem always has a solution, and that this corresponds to the following formula:

$$x_h^* = \frac{\sqrt{c_h}}{\sqrt{\sum_{j=1}^{J} \alpha_j^* \, a_{hj}} \cdot \sum_{k=1}^{H} \sqrt{c_k \sum_{j=1}^{J} \alpha_j^* \cdot a_{kj}}}$$

Where the α_j^* are suitable normalised constants (Lagrange multipliers), that is to say those for which $\sum_{j=1}^{J} \alpha_j^* = 1.$

To solve the problem of the minimum limit, Bethel proposes the use of an algorithm which is neither particularly efficient nor easy to apply. At that time, in fact, another algorithm was already available, formulated by Chromy (1987) and also put forward in the same publication of Bethel, which made it easier to find a solution to the problem from the point of view of the development of the code and quicker in terms of elaboration time.

Once the initial values of α_j ,equivalent to1/J, are in place, this algorithm develops fundamentally in two steps, which are repeated continually until reaching an acceptable criteria of convergence.

1. Calculate: $x_h\left(\underline{\alpha}^{(r-1)}\right) = \dfrac{\sqrt{c_h}}{\sqrt{\sum_{j=1}^{J} \alpha_j^{(r-1)} a_{hj}} \cdot \sum_{k=1}^{H} \sqrt{c_k \sum_{j=1}^{J} \alpha_j^{(r-1)} \cdot a_{kj}}}$

2. Calculate: $\alpha_j^{(r)} = \dfrac{\alpha_j^{(r-1)}\left[a_j' \underline{x}\left(\underline{\alpha}^{(r-1)}\right)\right]^2}{\sum_{k=1}^{J}\alpha_k^{(r-1)}\left[a_k' \underline{x}\left(\underline{\alpha}^{(r-1)}\right)\right]^2}$

2. Estimate of totals and calculation of relative sampling errors

For each variable the estimate of the total is obtained using the Horvitz-Thompson estimator:

$\hat{Y}_h = \sum_{i=1}^{n_h} \dfrac{y_{hi}}{\pi_{(h)i}}$, where h is the stratum index and i indicates the sampling unit.

In our particular case, that is the *simple random sampling without replacement*, the derived formula is:

$$\hat{Y}_h = \frac{N_h}{n_h}\sum_{i=1}^{n_h} y_{hi} \quad .$$

For the estimate of the variance of the total the Sen-Yates-Grundy formula is used:

$\hat{\sigma}^2(\hat{Y}_h) = \sum_{i=1}^{n_h}\sum_{j>i}^{n_h}\left(\dfrac{\pi_{(h)i}\pi_{(h)j}}{\pi_{(h)ij}} - 1\right)\left(\dfrac{y_{hi}}{\pi_{(h)i}} - \dfrac{y_{hj}}{\pi_{(h)j}}\right)^2$, for the single stratum h, while, having obtained the sample of H independent selection in each stratum , the total variance is obtained from the sum of the variances from within each single stratum.

In the particular case of the *simple random sampling without replacement*, the derived formula is:

$$\hat{\sigma}^2(\hat{Y}_h) = \frac{N_h^{\ 2}}{n_h}\frac{N_h - n_h}{n_h}\sum_{i=1}^{n_h}(y_{hi} - \bar{y}_h)^2 \quad .$$

The precision level, that is the Coefficient of Variation, is obtained by:

$$CV_h = \frac{\sqrt{\hat{\sigma}^2(\hat{Y}_h)}}{\hat{Y}_h}.$$

3. Procedure for controlling and correcting data

The phase of controlling and correcting data consists in identifying and treating errors present in the data gathered in the survey, in the aim of guaranteeing final results with specified levels of quality. For the creation of software to manage data, errors which characterize the survey were classified as sampling and non-sampling errors.

3.1 Sampling errors

Sampling errors refer to those errors which are encountered in the estimate of a parameter of the universe because of the fact that not all the population, but only a sub-set of it (the sample), is the object of observation. To control this kind of error, reference was made to the procedure of Bethel (1989), as described above, to estimate the optimum sample allocation. Through this algorithm it is possible to identify the sample size necessary to obtain, for the variables under study, estimates with pre-fixed sampling error levels. For the year of the survey, therefore, first the optimum sample size (based on the error limit required), and then the sampling error itself was estimated for the reference year. It is obvious that, in general ,there will not be an exact coincidence between the pre-fixed error and the estimated error afterwards, mainly because it deals with «a sampling estimate of the sampling error» (thus it is also derived from the observation of a sub-set of the population). Moreover, the parameters of input required by the algorithm will be the totals and the variances for each collected variable, in the strata in which the population was sub-divided. Obviously, in the estimate phase of the sampling size, those values relating to the following year in which the survey will carried out, can only be

hypothesized on the base of what had been observed in the year preceding the survey, exactly because it really deals with the same values which we intend to estimate.

3.2 Non-sampling errors

Non-sampling errors are those which are directly connected to the elementary data and are revealed as the difference between the value y_i of the variable Y, observed in the i-th unit, and the real value Y_i . It is thus obvious that, all other factors being equal, if the sampling error diminishes with the increase in sample size (annulling itself for census), this will not, in general, be true for the non-sampling error. An initial important classification of this type of error makes the distinction between complete non-response (CNR) and partial non-response (PNR). In general, in every survey, for every sampling unit, responses are gathered from a fixed number Q of questions. CNR is taken to refer to a statistical unit which does not supply responses to any of the Q questions given, while PNR is taken to refer to a statistical unit for which no information is available for a sub-set of Q questions. In reality, the distinction between these two types of missing responses is not as evident as it may appear. Generally, in fact, a statistical unit for which only a sub-set of questions is available, may still however be considered an CNR if it deals with a number of data too low with respect to Q, or also in the case where the variables not noted are considered strategic and of fundamental importance to the survey. In speaking of «non-availability» of information, reference is made not only to the cases of missing values (obviously identified), but also to the case where the value of a response to a question does not correspond to the reality, actually observed in the unit. The methods of control and localization of errors aim to identify these values.

In general, the control procedure of the survey in question can be considered as interactive graphic micro-editing of the univariate type. The term interaction refers to the fact that, in the procedure of localization of errors, there are not only automatic phases but also phases which require human intervention to investigate the situation and to evaluate the effective presence of the error (therefore the human intervention regards the localization phase and not that of imputation). The control is mainly of the univariate type because the variables are controlled individually and only in rare cases are suspected relationships existing among them controlled, using suitable synthesis indexes. During the various phases wide use is made of graphic tools to visibly evaluate situations marked as errors. Finally the word micro-editing is used because the data is gathered in suitable domains of study within which the sampling units can be considered very homogenous. For each of these sets of data, suitable control functions are first calculated, and then, for each of them, certain rules of incompatibility are verified. In the case of activation of conditions of error, that is in the case where the observed value does not belong to the region of acceptance, those control functions are then observed individually for all the sampling units forming the single domain. Thus the sample unit, or units, responsible for the activation of conditions of error is localized for the entire domain of study and then the imputation of the erroneous data follows. Considering that the number of activation of conditions of error is very low, the imputation of non-response error is generally obtained by re-interviewing the unit.

Annex IX - Provisional budget for 2014

BUDGET PREVISIONALEPROGRAMMA NAZIONALE RACCOLTA DATI ANNUALITA' 2014

Modulo	Sezione	Importo (Euro)
MEDITERRANEAN SEA AND BLACK SEA		
A - Evaluation of the Fishing Sector		4.498.132,00
	A1-T Transversal Variables	1.028.457,00
	A2 - Biological Variables	2.096.871,25
	A3 - Research Surveys at sea	1.372.803,75
OTHER REGIONS		
A - Evaluation of the Fishing Sector		21.456,75
	A1-T Transversal Variables	0,00
	A2 - Biological Variables	21.456,75
	A3 - Research Surveys at sea	0,00
SUPRAREGIONAL COSTS		
A1-E Evaluation of the Fishing Sector - Economic Variables		1.020.237,00
Costs related with Data Management and Data Use (A1-T, A1-E, A2, A3)		2.725.399,25
B - Evaluation of the Economic Situation of the Aquaculture and Processing Industry Sectors		412.991,00
	B1 - Aquaculture Sector	274.927,00
	B2 - Processing Industry	138.064,00
C - Evaluation of the Effects of the Fishing Sector on the Marine Ecosystem		333.806,75
TOTAL		**9.012.022,75**