TSDV: a GIS tool for inspecting trawl survey data

Fortunati L.[(1)], *Garofalo G.*[(2)], *Demontis R.*[(1)]

[(1)] CNR - Istituto CNUCE - Via S. Maria 36, 56126 Pisa (Italy)

[(2)] CNR - Istituto di Tecnologia della Pesca e del Pescato - Via L. Vaccara 61, 90126 Mazara del Vallo (Italy)

1. Introduction

Trawl surveys generally produce a vast quantity of data for several reasons:

- the frequency of the surveys;
- the geographical size of the area under study;
- the number of hauls made in the area;
- the number of parameters to acquire for each haul;
- the number of species to identify;
- the number of parameters acquired for each species.

Normally, such data is analysed with tools which don't consider the geographical position of individual data and which produce statistical or sythentic information that is not geographically correlated.

The Trawl Survey Data Viewer (TSDV) is based on a geographical information system (GIS) which can manage and analyse data produced in trawl surveys by referring them to the respective geographical positions, thus producing efficient summaries, both in tabular and graphic form. The TSDV makes available a set of functions with which users can operate dynamically on the set of data, by selecting from the various species and the parameters that characterise the data.

The TSDV is a tool for preliminary analyses of data produced by the surveys. The results obtained are a solid starting point for a more detailed study focusing on the knowledge of the distribution, the abundance and the population dynamics of demersal species.

The TSDV is based on ArcView GIS 3.0 software by ESRI and was developed within the framework of the FIGIS project of the European Community. The current version makes use of data from surveys in the Gulf of Castellammare in Sicily from 1993-94 and it can be generalised to historical data relating to any geographical context.

2. The data

The TSDV operates on data acquired both during the hauls and afterwards while analysing samples in the laboratory. The former data are technical and geographical, the latter are biological. Originally, all these data were acquired in tabular form; the geographical data were then inserted into a GIS and organised into layers of geographical elements (points, lines and polygons).

Besides the data produced by the surveys, accessory data are also used in order to get a better description of the environment of the surveys. Specifically, data regarding the bathymetry have been exploited along with data on the coastline and on the subdivision of the area of study into Elementary Sampling Units (ESUs) within which the hauls were conducted.

2.1 Fish sampling

The trawl surveys in the Gulf of Castellammare were carried out between autumn 1993 and summer 1994, one for each season. For each survey 68 hauls were scheduled (34 diurnal and 34 nocturnal) of 30 minutes each. The hauls were selected on the basis of random strata sampling, depth being the variable of stratification:

- the following three bathymetric strata were chosen: 0-50 m. (stratum A), 50-100 m. (stratum B), 100-200 m. (stratum C);
- each stratum was subdivided into ESUs areas;
- the number of hauls per stratum was calculated proportionally to the area of each stratum;
- the distribution of the hauls (one per ESU) was thus determined by the ESUs selected randomly.

A total of 245 valid hauls in the four surveys were found and 130 different species were identified and divided into three groups taking into consideration both the commercial importance and the quantity of the area of study (D'Anna, Badamenti et al, 1994).

The TSDV operates on a part of the existing set of data. Considering the high number of species, only the species belonging to the first and second groups were taken into consideration. In the first group are the 12 most important species and for them data were acquired on their length, weight, sex and maturity. For the 24 species in the second group only the length and weight were considered.

The 36 species are present in the database both individually and grouped into categories: fish, crustaceans and cephalopods. In the latter case, there is only data on total weight and the total number of individuals.
For all the species and groups of species, the raw data are processed and synthesised producing the abundance indices, in number (No./km^2) and weight (gr./km^2), calculated with the swept area method. Moreover, for the single species the density (weight and number) was also calculated for each class of values of the length and, where present, of the sex and maturity.
For the various attributes, the following classification was used:

- sex male, female, undetermined, common to all species;
- length up to four classes for each species;
- maturity up to six classes plus undetermined (the scale of maturity depends on the species).

2.2 Geographical data and anciliary data

The TSDV operates with geographical data relating to the Gulf of Castellammare. The data are indentified in a reference system of geographical coordinates (longitude, latititude) expressed in decimal degrees (DD). All the geographical data are thus consistent with the above reference system and have an accuracy of localisation x,y which depends on the acquisition system and the source of geographical data.
The geographical data used by the TSDV are in two sets and refer to:

- *location of the hauls:* describing the individual hauls carried out in the various surveys. Each is represented synthetically by a start and end point of the haul, and is indentified by a progressive number. These data were acquired during the survey with a geographical positioning system (aboard the ship)**.**
- *location of the samples*: describing the location point of the samples (station) related to the various hauls. The sample acquired during a haul is referred geographically to the central point of the segment that unites the start and end points of the haul. This information is thus derived from the first set of data.

The anciliary data used by the TSDV are acquired from maps and regard:

- *bathymetry*: obtained by digitalising the map "Da Capo Rama a Marsala" (Istituto Idrografico della Marina, 1983), Mercator projection, 1:100.000 scale;
- *coast line*: obtained in the same way as the bathymetry;
- *toponyms*: limited to the ports affected by the study;
- *ESUs:* subdivision of the bathymetric strata 0-50, 50-100, 100-200 m. into elementary sampling areas.

All these data, composed of geographical elements, are organized into shape files.

2.3 Tabular data and their organisation

The data that describe the surveys are organised into a set of files, subdivided into two categories that respectively contain:

- *the identification of the species*: 36 species and three groups of species are listed. For 12 species the length, maturity and sex are defined, for the other 24 only the length, while no attribute is defined for the three groups of species. In turn, each attribute is defined in the number of classes and in the respective definition:
  1. *species group*: 3 classes;
  2. *length*: up to a maximum of four classes belonging to each species;
  3. *sex*: 3 classes common to all species;
  4. *maturity*: 7 classes belonging to each species;
- *the characteristics of each sample*: each species is described in a file containing the characteristics of the various samples. For each haul are reported the values (in terms of unit/km^2) of the weight and number of the elements belonging to individual classes of each attribute.

3. System functions

The TSDV is designed to display data of trawl surveys in a synthetic, selective and interactive way, thus enabling the user to select the various parameters that specify both the characteristics of the survey along with those of the species to be studied. These functionalities are available for operating on data previously described and can be activated via buttons on the upper part of the user interface.

3.1 Types of selection

The TSDV allows the following functions:

- *layout of the geographical base reference;*
- *survey selection (season);*
- *haul selection (diurnal, nocturnal, daily);*
- *selection of the species or group of species (36 species and three groups of species);*
- *selection of the attibrute (sex/length/maturity);*
- *selection of the class/es of the attribute (specifications of the attribute);*
- *selection of the variable (number/weight of the elements).*

Functions are selected via a series of buttons that activate windows containing the list of possible selections. For some parameters (survey, haul, species, attribute, variable) a single element in the list can be selected, for others (classes) selection can be multiple.

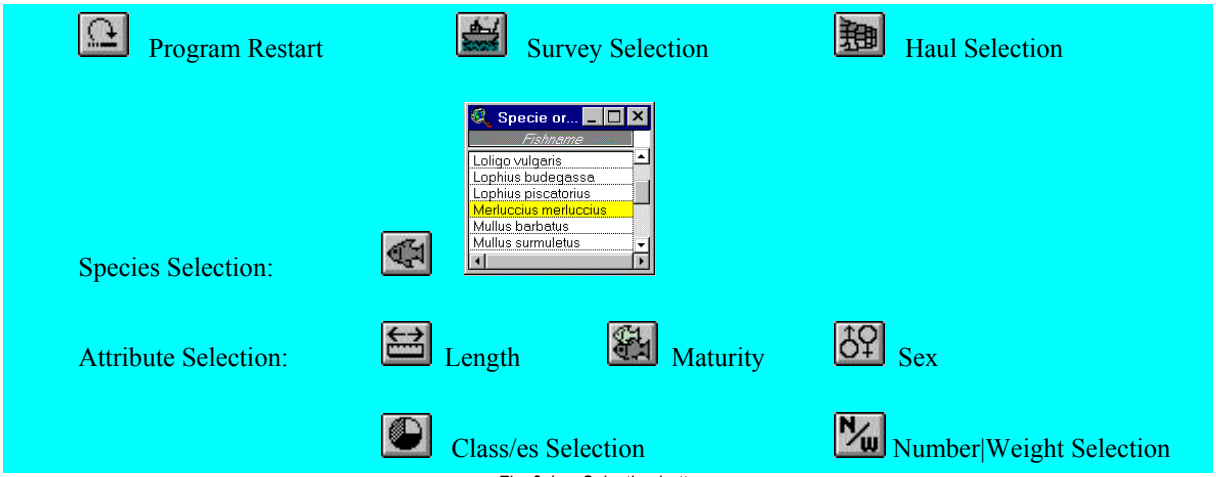

Fig. 3.1a - Selection buttons

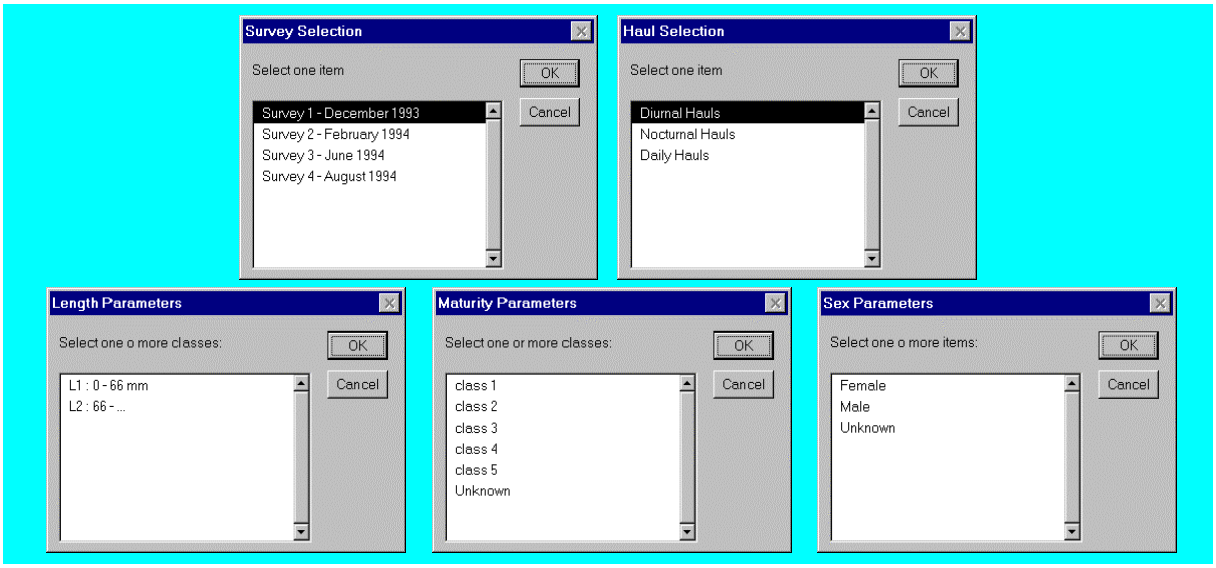

Fig. 3.1b - Windows for single and multiple parameter selection

Each of the selection functions, if activated by the respective button, produces a different result depending on the status which the system is in. In turn, the status of the system depends on the set of selections carried out and still active.

3.2 Representation types

The TSDV allows the following representations to be obtained:

a) coast line, bathymetry, toponyms: common reference base to all the maps;
b) location of the ESUs;
c) location of the hauls of each survey (diurnal, nocturnal, daily);
d) distribution of the total weight or the total number of the elements of a species, in relation to the hauls in a survey: representation via a colour scale applied to the ESU;
e) distribution of the total weight or the total number of the elements of an attribute (sex, length, maturity) of a species, relative to the hauls in a survey: representation via a colour scale applied to the ESU;
f) distribution of the weight or of the number of elements of one or more classes of an attribute (sex, length, maturity) of a species, in relation to the hauls in a survey: representation via proportional symbols or pie charts, or histograms, applied to the location point of the samples.

All the representations are carried out in the UTM reference system, with metric coordinates.

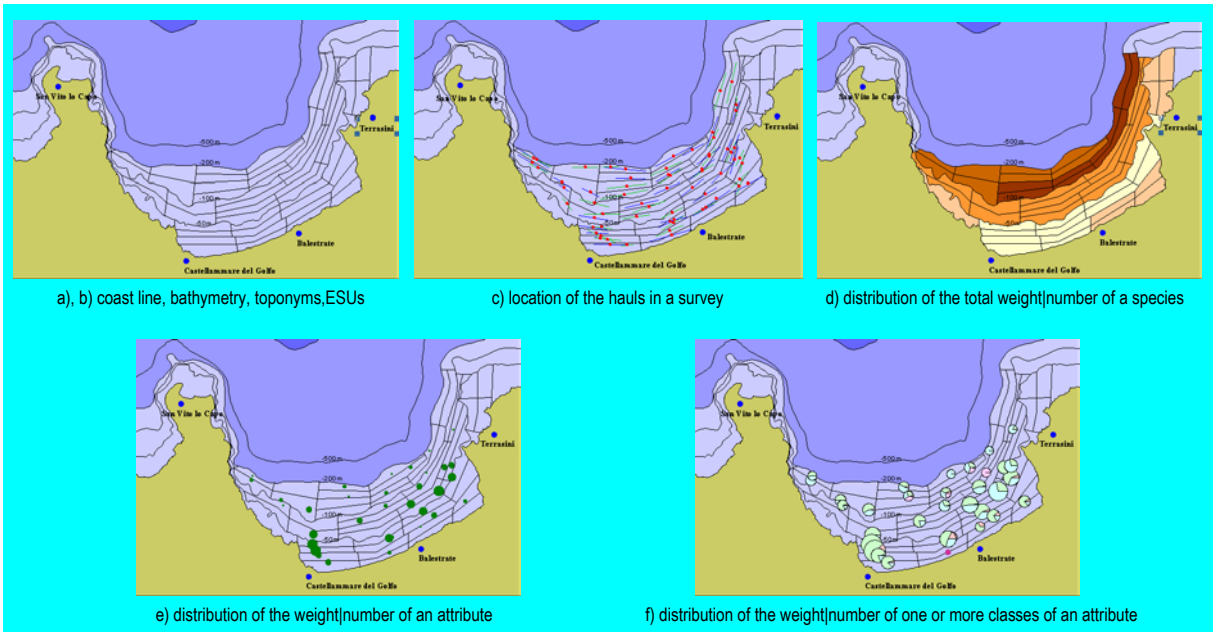

Fig. 3.2 - Representation types

4. TSDV Architecture

The TSDV was developed using ArcView GIS 3.0 by ESRI. This application consists of a set of scripts in Avenue language, which carry out specific functions.

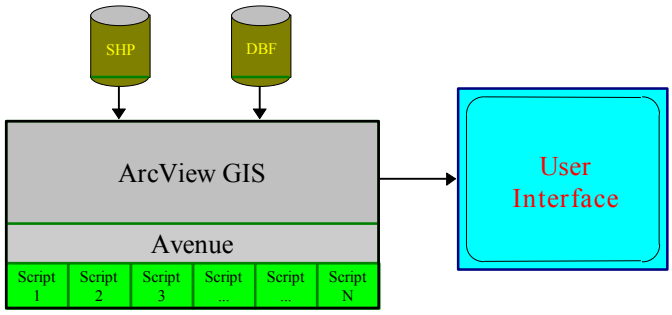

Fig. 4 - Sistem architecture

In input the system has two sets of data (geographic and tabular) and produces in output a set of representations in accordance with the parameters selected by the user.
The user interface is based on the standard schema of the ArcView system and includes the documents: *Views*, *Tables*, *Charts, Layouts* and *Script*.

- ***Views***
  When the system is activated the "Trawl Survey Data View" is defined which containts the base layers: bathymetry, coast line and toponyms, with the relating legends already set. On this View all the various representations are built dynamically and interactively. A series of buttons, specific to the application, are defined in the Button Bar associated with the View, each of which corresponds to a specific representation function.
- ***Tables***
  The system uses tables dynamically as a function of the selections made. Only one table containining the species identification is permanently loaded.
- ***Charts***
  Not used.
- ***Layouts***
  When the system is activated a layout is defined for printing the current View. The layout contains the following elements:
  • TSDV and FIGIS logos;
  • title of the layout;
  • representation window of the current view with a scale of 1:125000;
  • legend of the symbols used in the view;

- scale bar;
- indication of geographical north.

- ***Script***
  When the system is activated all the scripts of the TSDV application are defined.

4.1 Organisation and structure of the system

The system is designed to operate essentially in two different ways depending on whether the choice of parameters is made completely or not. Specifically, two parameters drive the behaviour of the system in one or the other mode: the attribute and the class of the attribute. If these two parameters are not selected, depending on what other selections have been made, the system produces, not only the general representations (ESUs and hauls in a survey), but also the total distribution of the species selected. However, if all the parameters have been set, reselecting one of these produces a new representation of the distribution, for the active parameters of the species selected. These two modes of functioning of the program are described in more detail below.

Incomplete selection of the parameters

Figure 4.1 shows the functionalities of the system when not all the selections have been set: in particular when the selection of the attribute and of the class are not active. In this status, the survey, the haul, the species and the variable, considered in a decreasing hierarchical order, can be selected; each of these selections produces a representation which also depends on the active selections of the hierarchical parameters at a lower level.

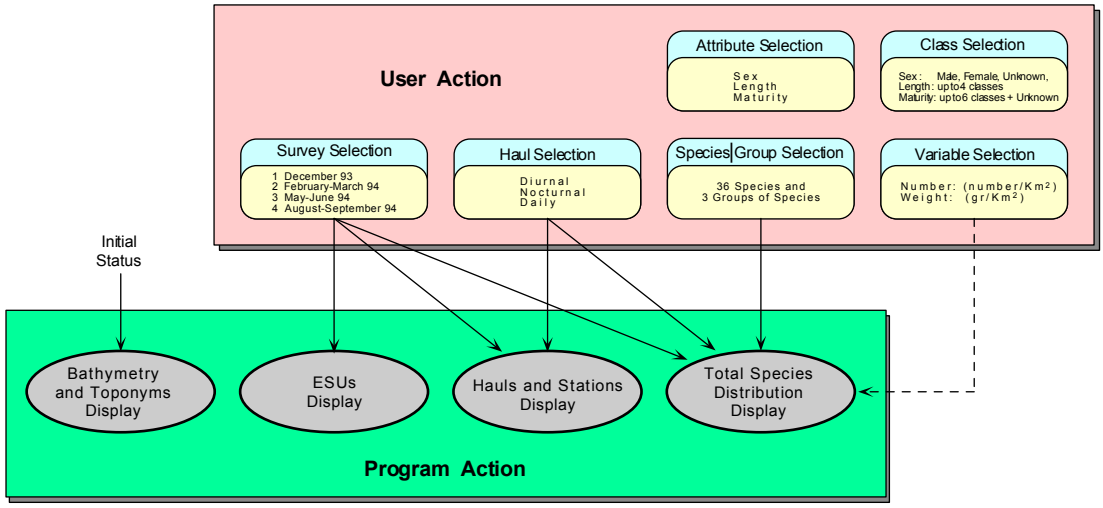

Fig. 4.1 - Program actions in case of incomplete parameters selection

When the system is activated there are no active selections, thus the View consists of the bathymetry and the toponyms:

- Selection of the survey: this leads to an overlaid view:
  1) of the ESUs (if the other parameters have not been selected);
  2) of the hauls (if these have been selected);
  3) of the total distribution of the species (if one of these has been selected) for the number|weight variable active (weight by default).
- Selection of the haul: if the survey has been selected, there is an overlaid view:
  1) of the hauls: for daily hauls there is a distinct representation of the diurnal and nocturnal ones;
  2) of the total distribution of the species (if one of these has been selected) for the number|weight variable active (weight by default).
- Selection of the species: if the survey and the hauls have been selected there is an overlaid view of:
  1) of the total distribution of the species for the number|weight variable active (weight by default); this representation is carried out by correlating each haul to the corresponding ESU and applying a scale of colours (proportional to the value of the variable) to the polygons which represent the ESUs.
- Selection of the varibale: a choice between number and weight can be made at any time (weight by default) with different results depending on whether the species selected is (a) or (b):
  a) allows the display of the total distribution of the species in terms of the number|weight variable selected;
  b) sets only the selection of the variable.

Complete selection of the parameters

When all the parameters have been set, there is a representation of the distribution of the active species with the relating attributes and classes selected. A new choice of any of the parameters involved leads to the production of a new representation of the distribution of the species.

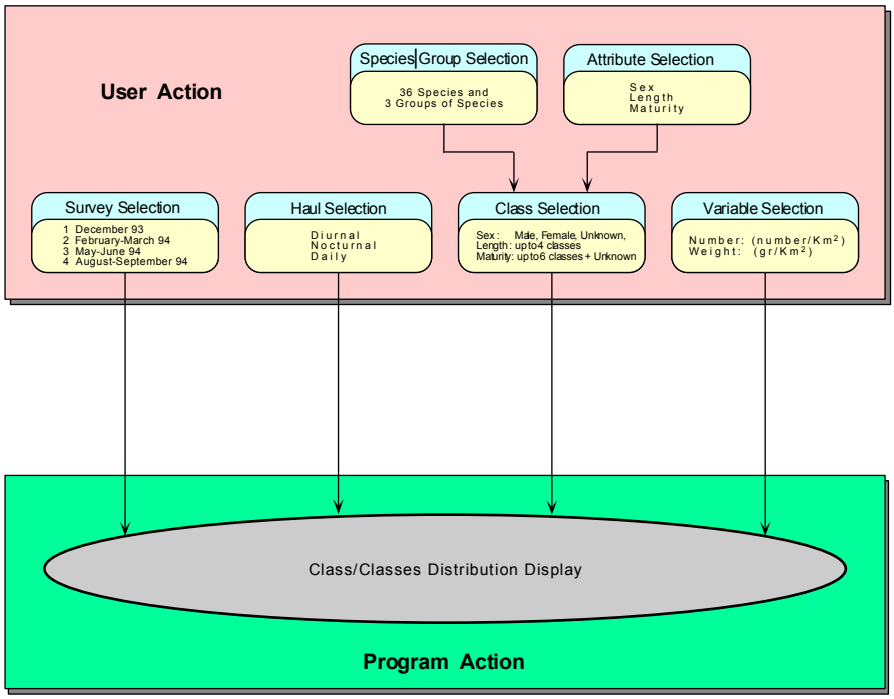

Fig. 4.2 - Program actions in the case of complete parameters selection

4.2 Organisation and structure of the data

There are two types of data that the system manages: goegraphical (SHP) and tabular (DBF).

- Geographical data: consisting of shape files of point, lines and polygons. The following datasets are part of this type of data:
  - bathymetric strata 0-50, 50-100, 100-200, 200-500 m. and dry land (polygons);
  - location of ports (points);
  - definition of the ESUs areas (polygons);
  - location of the hauls (lines);
  - location of the sampling stations (points);
- Tabular data: consisting of files of tables containing information indirectly correlated to geographical elements:
  - list of the species and of the groups of species;
  - biological data of the species;

Acknowledgements

This paper was based on data obtained from the fishery survey cruises carried out in the Gulf of Castellammare by the I.T.P.P.-C.N.R. with the financial support of EU-DGXIV. We would like to thank Dr. G. D'Anna, Dr. F. Badalamenti and Dr. C. Pipitone, who kindly made the data available.

References

D'Anna G., Badalamenti F. et al. (1994)
Structure of the populations and assessment of the biomass of the coastal demersal resources in the Gulf of Castellammare
Final Report - Project MED92/011, EU-DGXIV.

APPENDIX - USER INTERFACE

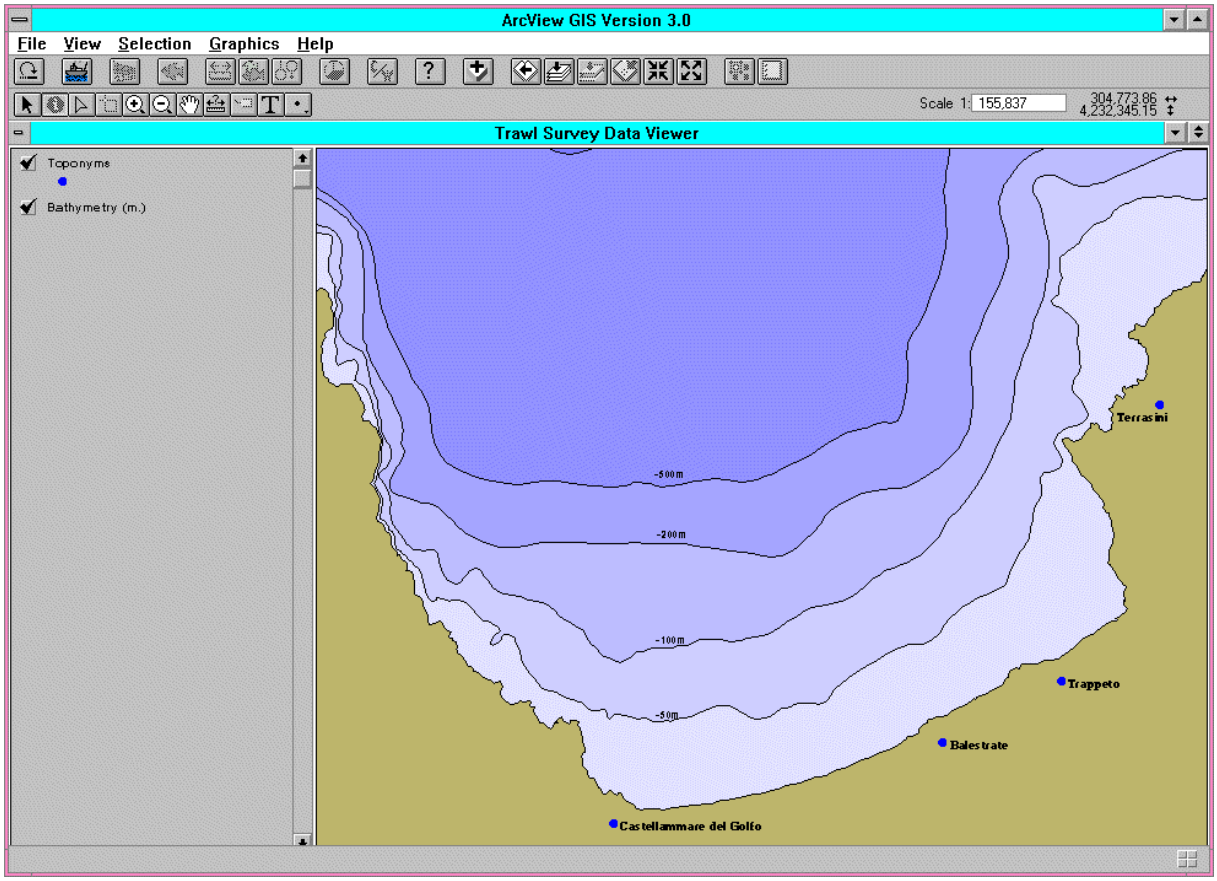

Fig. A1 - Bathymetry and toponyms

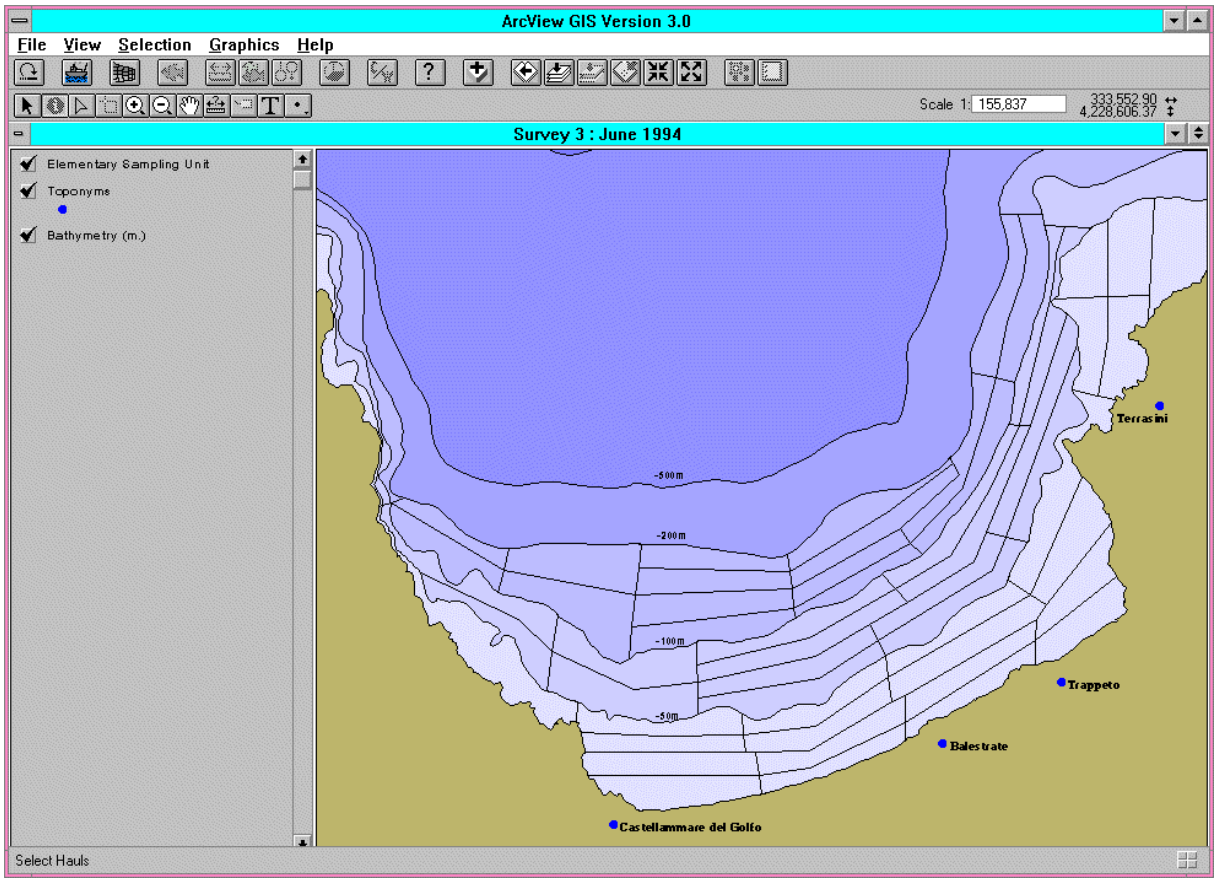

Fig. A2 - Elementary Sampling Units

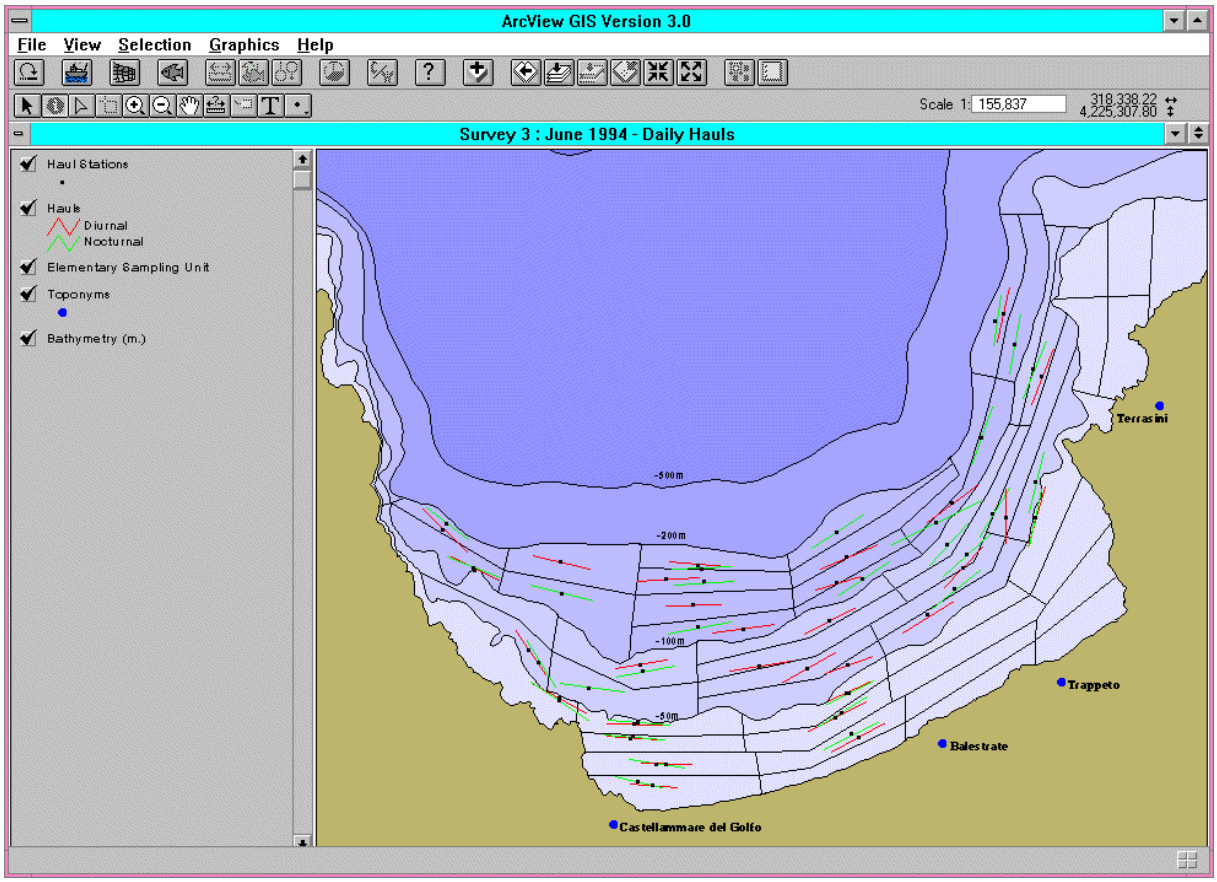

Fig. A3 - Haul distribution of a survey

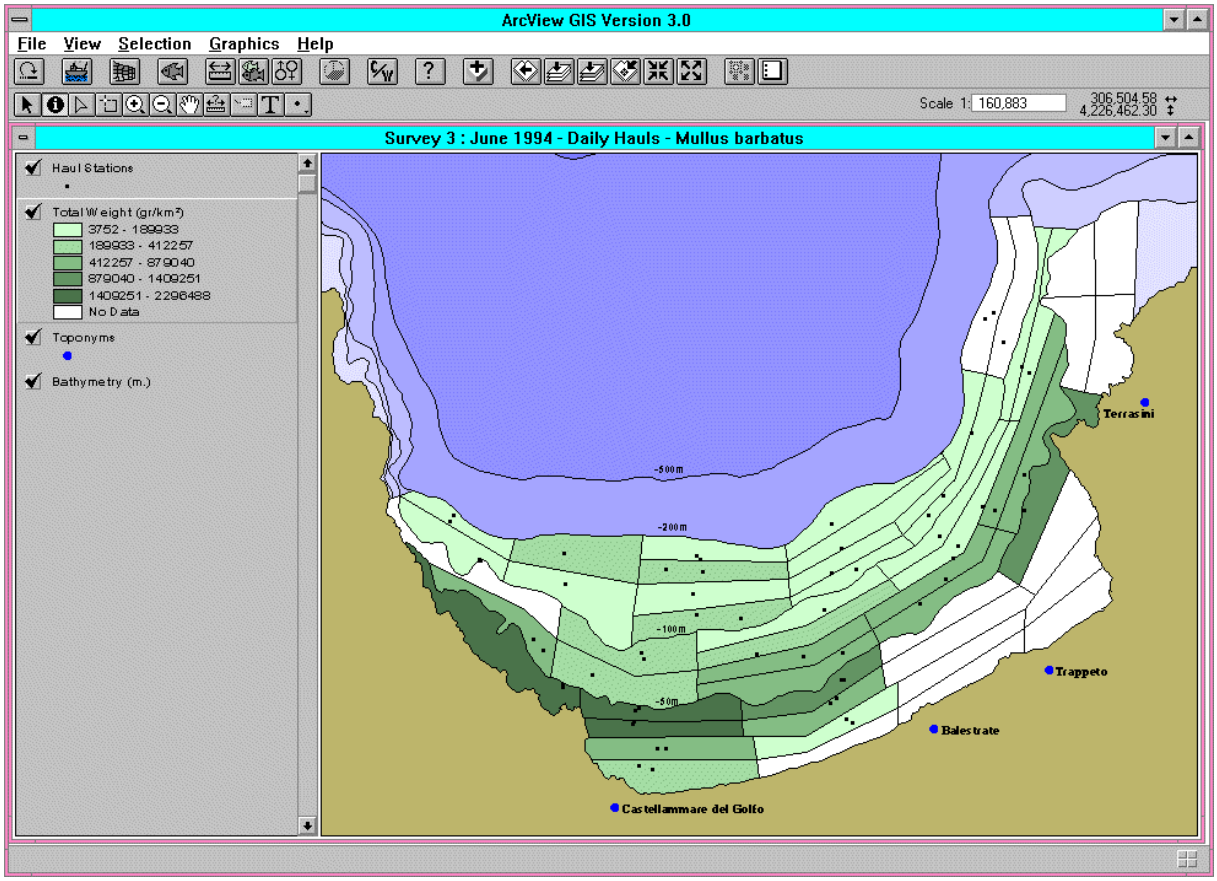

Fig. A4 - Total catch distribution of a species by ESU

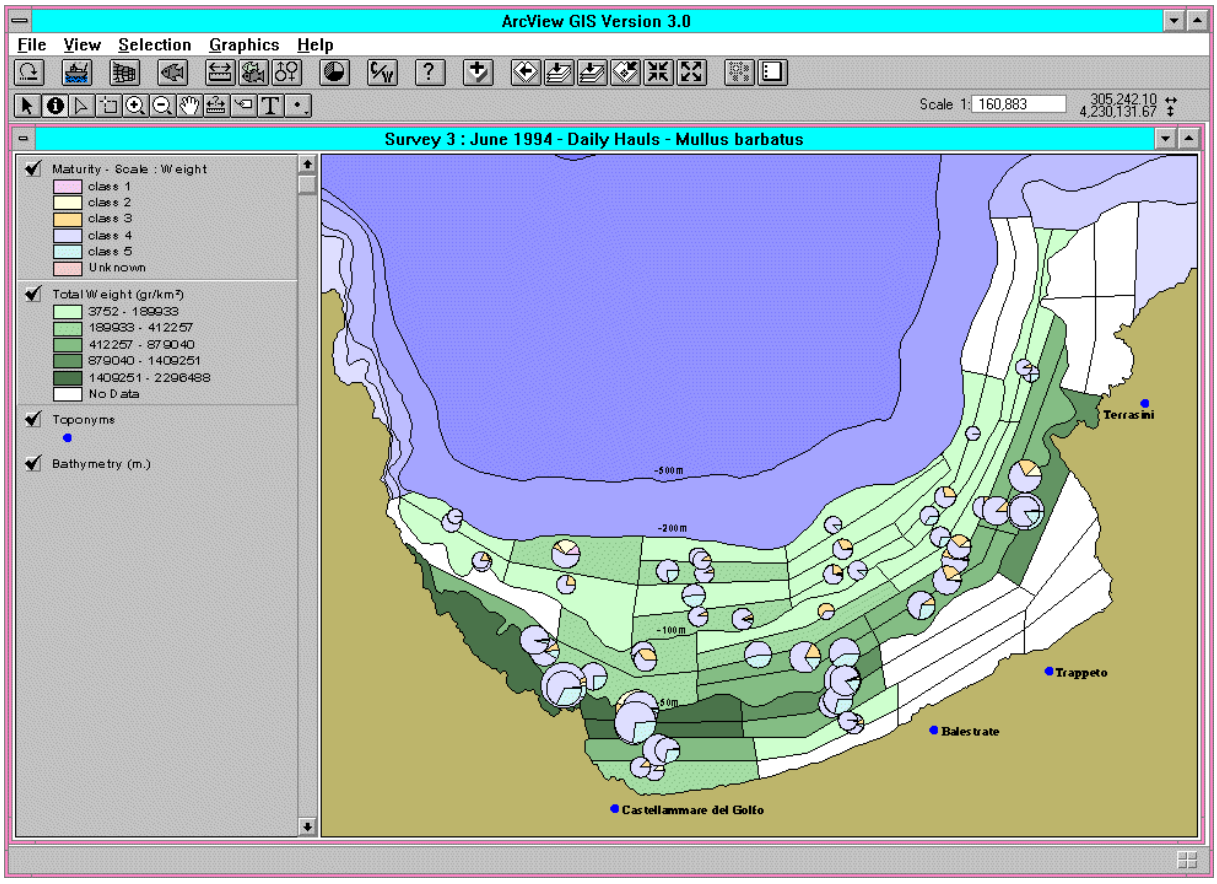

Fig. A5 - Catch distribution of a species by maturity classes

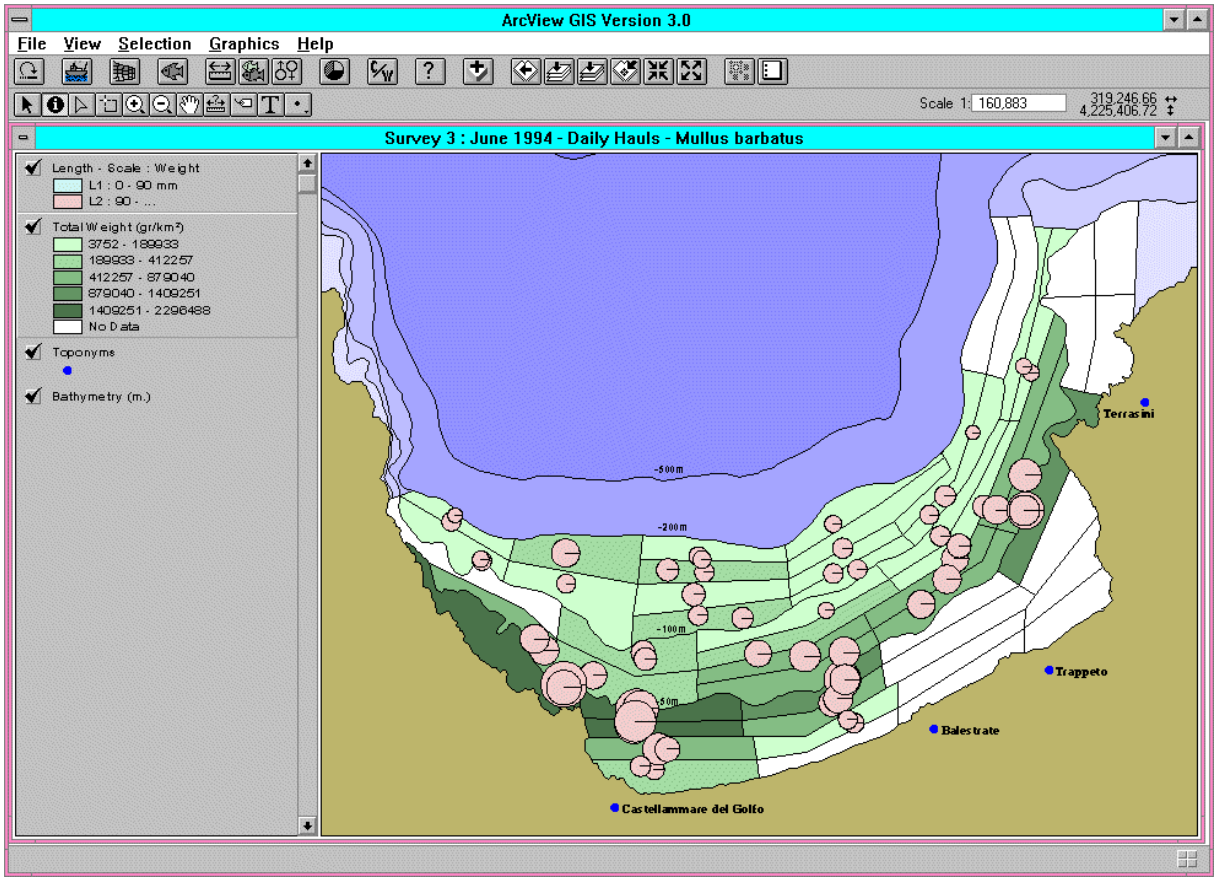

Fig. A6 - Catch distribution of a species by length classes

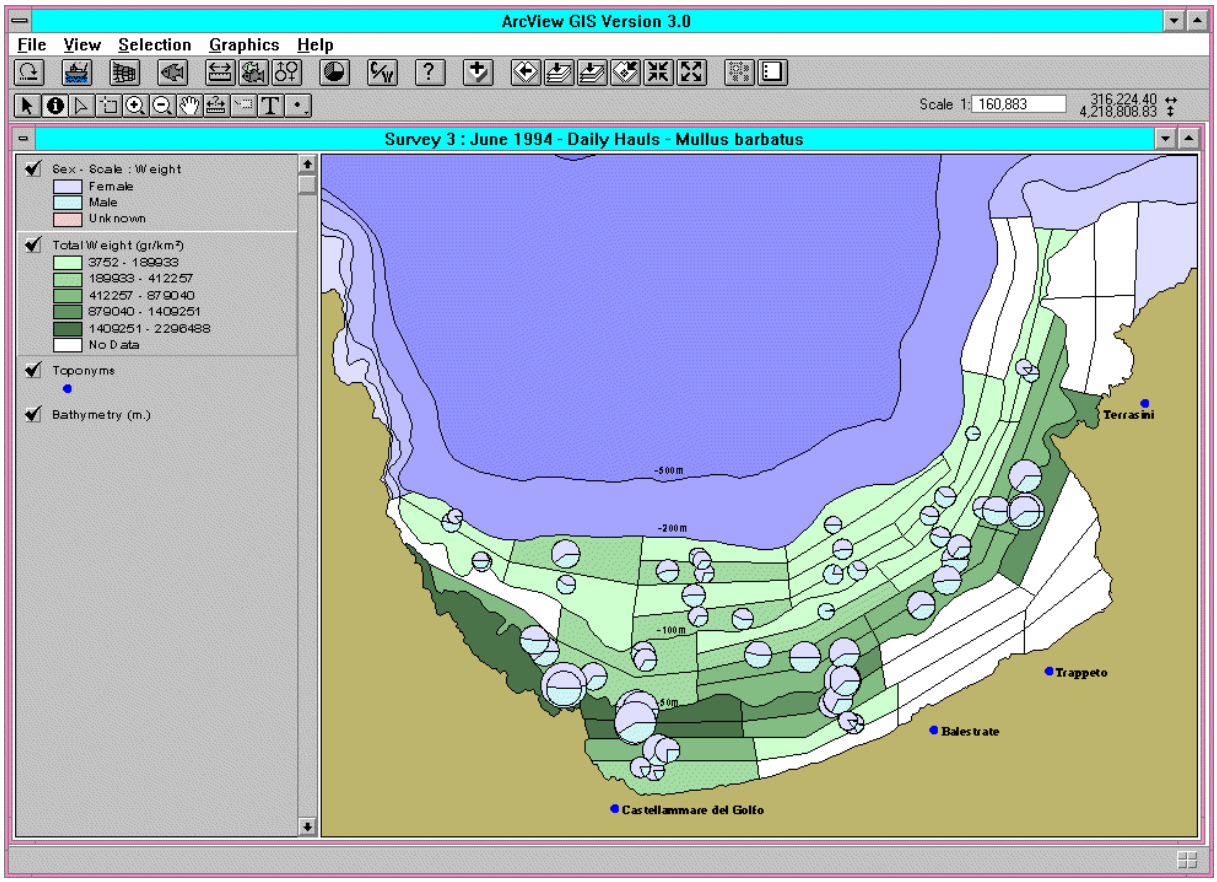

Fig. A7 - Catch distribution of a species by sex classes

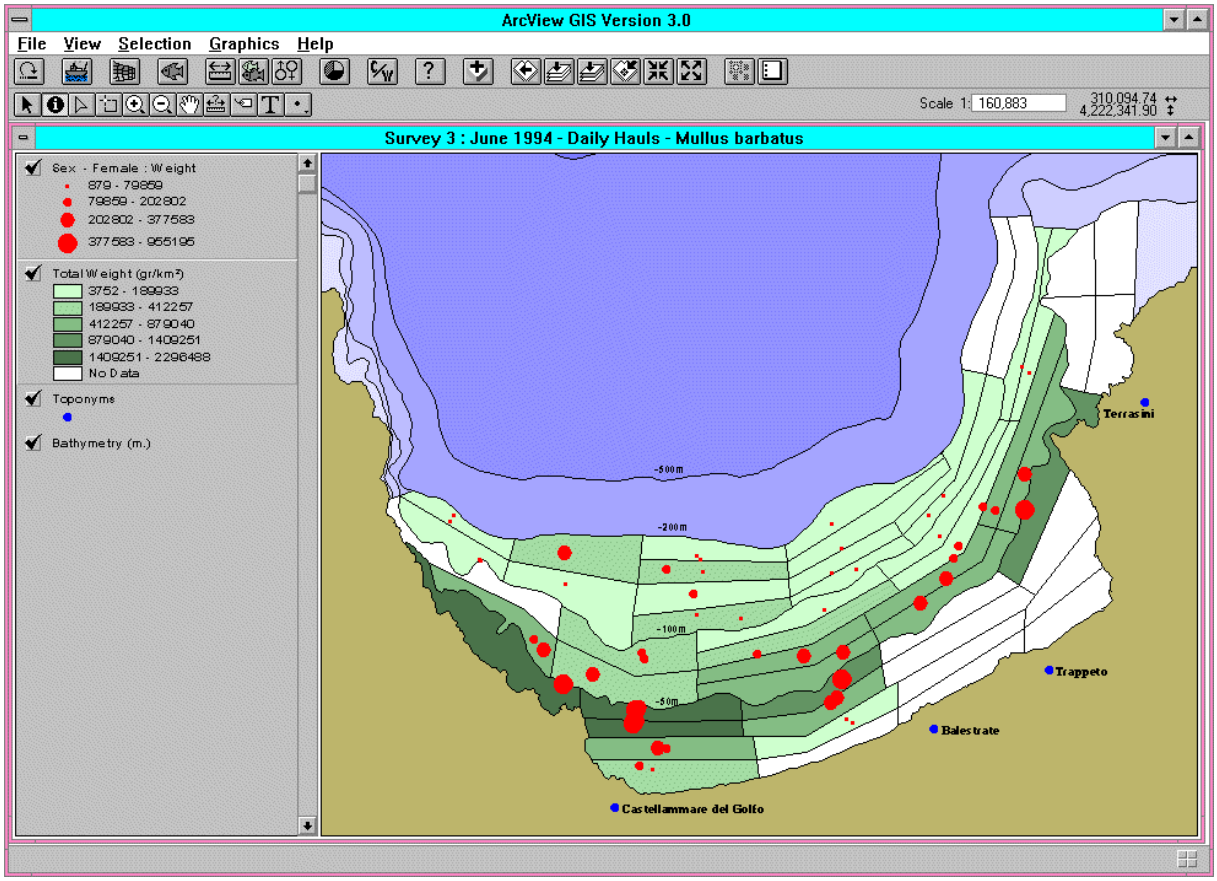

Fig. A8 - Catch distribution of a species by one sex class

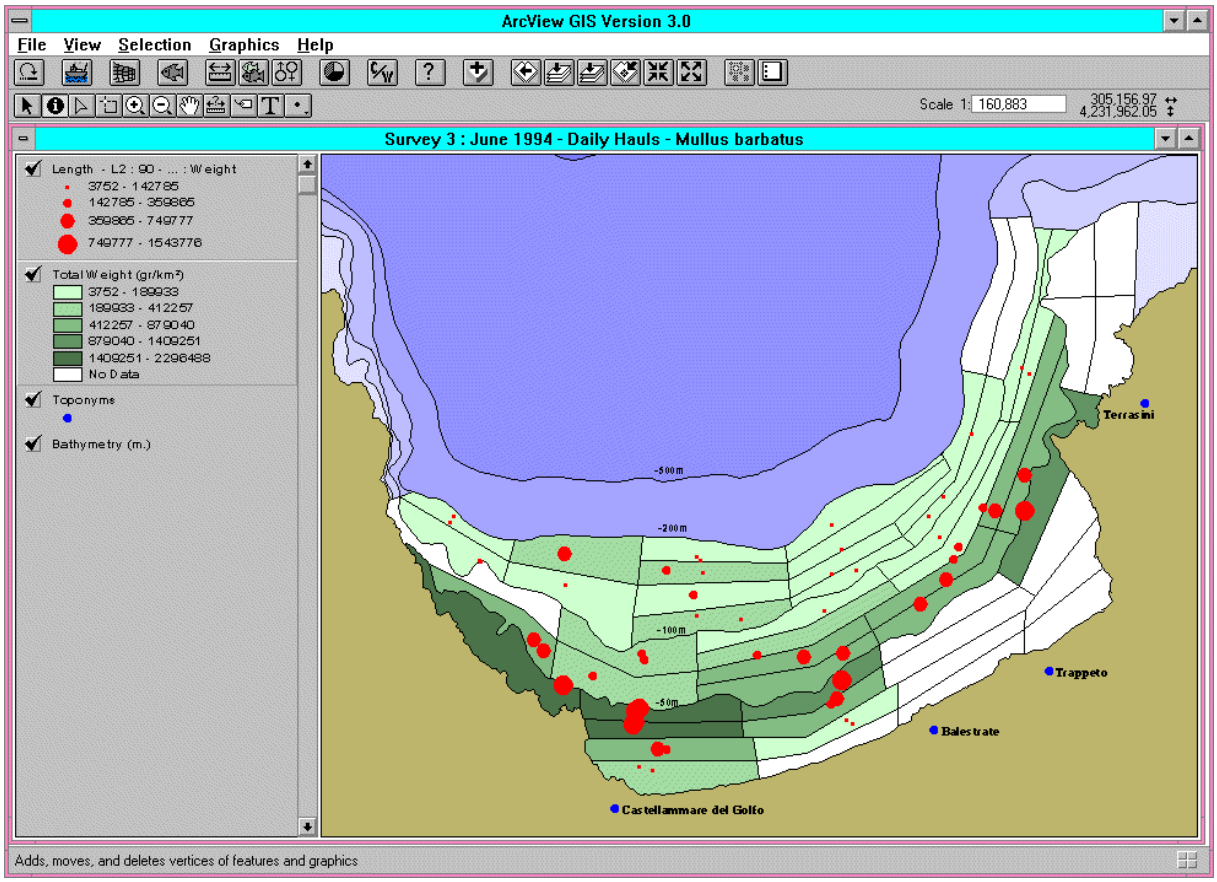

Fig. A9 - Catch distribution of a species by one length class

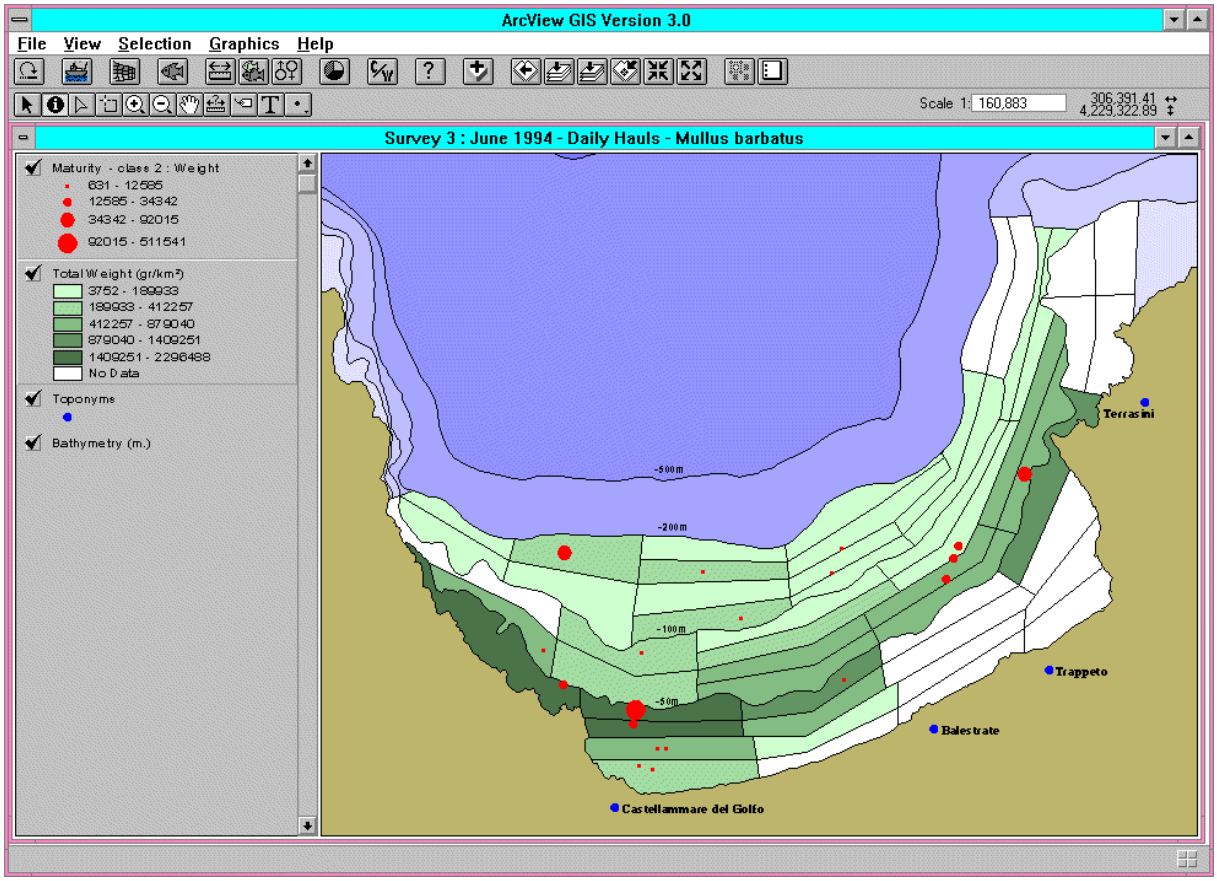

Fig. A10 - Catch distribution of a species by one maturity class