ANALYSIS OF TRAWL SURVEY DATA USING MEDITS FILE FORMAT

Walter Zupa, Isabella Bitetto, Pierluigi Carbonara, Loredana Casciaro and Maria Teresa Spedicato

COISPA Tecnologia&Ricerca, Stazione Sperimentale per lo Studio delle Risorse del Mare, Bari, Italy

1. Introduction

Historically, trawl-survey data have mainly been used to:

- estimate the pattern of the fishery resources distribution;
- provide indices for the tuning of VPA in several advisory bodies.

In the last decade trawl surveys have gained more attention as a primary tool for providing "*per se*" useful information to the assessment process (e.g. Cotter, 2009a).

In the Mediterranean, trawl survey data have been considered for long time the most important and sometimes the lone source of reliable information, for assessment purposes and for understanding the structure of communities and ecosystem (e.g. SAMED, 2002; Cotter, 2009).

MEDITS trawl survey (Bertrand et al., 2002) and other trawl surveys in the Mediterranean region have collected, for a range of target species, a variety of information related to:

- ✓ standardised indices (abundance and biomass) of the whole population or life stages;
- ✓ indicators of population demography and structure;
- ✓ spatial occupation indices;
- ✓ community indicators;

providing comparable information among the various geographical areas. In some cases results have also been exploited in several projects for example for the identification of nursery and spawning grounds (e.g Lembo et al., 2010; Giannoulaki et al., 2013) and the stability or change in biodiversity (e.g. Granger et al., 2015).

2. The indicators to be analysed

Monitoring indicators from scientific surveys offers the possibility, for managing fish stocks, that are either complementary or alternative to well-established methods of stock assessment involving the modelling of numbers at different ages (Cotter et al. 2009a).

The following population state-indicators have been selected to initiate an exercise with trawl survey data from populations of red shrimps in the eastern and central Mediterranean sub-regions:

1. number of positive hauls to the species;
2. mean biomass index (kg/km^2),
3. mean abundance index (number/km^2),
4. the inverse of mean abundance Coefficient of Variation (CV)
5. mean individual weight (MIW)
6. sex-ratio
7. index of recruits (number/km^2)
8. index of spawners (number/km^2)
9. length at 95° percentile (L0.95).

1. **Number of positive hauls to the species**. It is assumed that the size of the spreading area of a stock is mainly dependent on its abundance. A trend analysis can provide insight into the evolution of the occupied area in the medium term. This indicator can be also affected by environmental changes.

It is computed as: *Positive hauls * Total hauls^{-1}*100*

2. **Mean biomass index (kg/km^2)**. This index measures the total biomass of a species per unit area. Changes (decrease) in these indices can be caused by an excessive fishing pressure.
The index is calculated as (Souplet, 1996):

$$I = \sum_{i=1}^{N} W_i \bar{x}_i$$

where I is the index, W_i is the weight of the stratum i, and x_i is given by:

$$\bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{i,j}}{\sum_{j=1}^{n_i} A_{i,j}}$$

where $x_{i,j}$ is the weight of the individuals in the haul j of the stratum i and $A_{i,j}$ is the area trawled in the haul j of the stratum i; n_i is the number of hauls in the stratum i. within stratum variance is calculated as:

$$S_{x_i}{}^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} A_{i,j} \left(\frac{x_{i,j}}{A_{i,j}} - \bar{x}_i \right)^2$$

and the variance in the survey area as:

$$\mathrm{var}(I) = \sum_{i=1}^{N} \frac{W_i^2 S_{x_i}{}^2}{\sum_{j=1}^{n_i} A_{i,j}} (1 - f_i)$$

where f_i is the ratio between the area trawled in the stratum i and the stratum area (finite population correction factor, generally negligible).

In the case of a random stratified sampling in which the number of the hauls in each stratum is not proportional to the strata surface, the variance is calculated according to the following formula (Cochran, 1953):

$$var(I) = \frac{1-f}{n} \sum W_i S_i^2 + \frac{1}{n^2} \sum (1 - W_i) S_i^2$$

Where W_i is the weight of the stratum i, S^2 is the variance within the stratum, n is the sum of the swept areas and f is the ratio between the area trawled and the surface area

3. **Mean abundance index (number/km^2)**. Abundance indices, like the biomass ones, can change (decrease) for the effect of an excessive fishing pressure, however more than the biomass ones, are likely to be substantially affected by large recruitment pulses in the stock, particularly if numbers of adults are low. For the analysis of time series of such index the ln transformation of the variable *ln(x)+1* is usually applied.
For computation see mean biomass index.

4. **Inverse of mean abundance Coefficient of Variation (CV).** The reciprocal of the coefficient of variation could be seen as a descriptor of the stability of the variable under investigation (higher is the metrics, more stable is the variable).
It is the square root of the variance, as obtained for the biomass index, divided by the mean abundance index.

5. **Mean individual weight (MIW)**. Mean Individual Weight (MIW) is generally considered an indicator that synthesizes the structure of the population (Piet and Jennings, 2005) and its

changes in time are likely linked to changes in fishing pressure, though it can be also influenced by the recruitment peaks. This influence is expected to be less pronounced if older individuals in the population are well represented. Mean weight is particularly useful for those species caught in the trawl surveys for which no data on individual size is collected.
It is computed as the ratio between overall biomass by haul i and overall number of individuals by haul i: $Bi//Ni$

6. **Sex-ratio**. Sex ratio provides information on the distribution of female and male individuals present in a population. It represents the proportion of females in a population and indicates the level of sex dominance (Adebiyi, 2013). Generally this is a peculiar trait of the population. The sex-ratio, as the proportion of the females on the overall number of individuals, can be considered correlated with the stock productivity and renewal.
It is computed as:

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

where FE are females, MA males, i is the haul and n the total number of hauls.
Variance is estimated as:

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

7. **Index of recruits (number/km²).** Recruits are often measured as the individuals belonging to the first component of the length frequency distributions, or as the individuals of the first age class, according to the recruitment mode, population structure and species. Thresholds to split the recruits from the whole population index can be also obtained from different areas or from literature.
For the indices calculation see mean biomass.

8. **Index of spawners (number/km²).** As individuals in spawning phase are not always intercepted by surveys, spawners can be approximated using the indices of adult individuals, i.e. those larger than the size at first maturity.
For the indices calculation see mean biomass.

7. **Length at the 95° percentile (L0.95).** The different percentiles of a length frequency distribution (LFD) are expected to respond differently to fishing, recruitment pulses, and loss of spawning stock. A large percentile of the population length distribution (L0.95) is an indices of the numbers of adult, older fish. It is assumed to be negatively (decrease) affected by an excessive fishing pressure.
It is computed from the standardised LFD that is:

$$Fq_{j,l} = \frac{fq_{j,l}}{A_j} \qquad , \forall j, \forall l,$$

where $fq_{j,l}$ is the number of individuals in the length class l from the haul j standardised to the km^2, and A_j is the surface trawled in the haul j.
The length at the 95° percentile (L0.95) is computed as:

$$L_{q,i} = l_{q,i} \left| \frac{\sum_{l=1}^{l_q} y_{l,i}}{y_i} = q \right.$$

Where l is the length class corresponding to the 95° percentile (q = 0.95) for the species i, and $y_{l,i}$ is the value of the catch for the length class l.

The variance of the length at the 95° percentile is computed as:

$$Var[L_{q,i}] = \frac{q(1-q)}{y_i\left(y_{l_{q,i}}/y_i\right)^2}$$

Among the general methodological considerations it is worth mentioning that the variance of each indicator is generally high and the statistical power for detecting trends is low for indicator series <10 years (Nicholson & Jennings, 2004). Thus results from short series should be considered with caution.

3. Dataset preparation

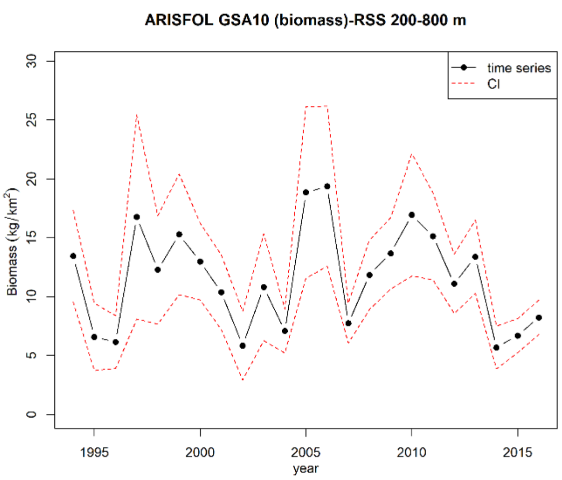
The dataset for the indices standardization is prepared using the structure of files like MEDITS TA, TB and TC (MEDITS Handbook, 2017).

Files are merged for the subsequent analyses.

4. Analysis of temporal trends

In presence of a short trawl survey time series the significance of the trend can be estimated using nonparametric statistical tests as Spearman rho (Cotter, 2009b). Spearman's rho is the product-moment correlation between the ranks of paired data. To test for trend, one member of the pair is the time of observation, the other is the observed variable. This is also known as the Hotelling-Pabst test. It appears to be sensitive to monotonic trends.

GAM models, as in the Intersection Union Test developed by Trenkel and Rochet (2009), can be applied for longer time series. This method estimates the direction of recent changes making use of first and second derivatives of smoothed indicator time series and the position of the most recent years with respect to the full time series. The first derivative, which is the local slope (tangent) at each point of a function, here each year, describes the dynamics of the indicator changes. In contrast the second derivative describes the changes in the slope. A positive second derivative indicates that the slope is increasing while a negative second derivative means that the slope is decreasing. The location where the second derivative is zero is called a change point as at this point the dynamics change from accelerating to decelerating. The slope will be zero when either a maximum or minimum is reached.



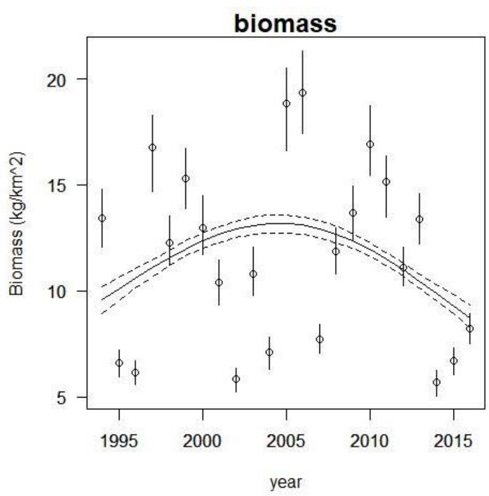


Fig. 3.1 Biomass index of *A. foliacea* in the depth range 200-800 (left) and an example of application of intersection-union test (right) (data from the demo).

5. Introduction to the spatial analysis

Placing the above described indices in the spatial dimension is very useful for several objectives, as the knowledge of the distribution and abundance of the species at local and regional spatial scale, the identification of areas with certain peculiarities, the localization of sensitive life stage of the population, etc… These objectives are also complementary to the management of fisheries resources at spatial scale, especially when combining the information on the spatial distribution of the fleets and the related fishing effort.

The objectives of the spatial analyses can be achieved using more or less complex modelling approach at spatial scale. For the aim of this exercise we decided to keep the approach simple, with results valid under certain assumptions. The approach consists indeed in placing in the spatial dimension the indices (a selection among the 9 proposed) computed as average of a recent time span (10 years).

For this exercise GFCM grid is considered as the better available trade-off for the spatial scale to be used in the different steps of the analysis, especially considering the subsequent possible overlap with maps of the fishing effort spatial distribution. This despite the fact that not any unit in the grid can actually encompass the spatial distribution of all the life stages of a given species, especially when this distribution is depth related.

Mean biomass and mean abundance

The mean abundance (likewise the mean biomass) in the GFCM grid cells (see an example in figure 4.1) Dcell is calculated as the average of the standardized numbers of individuals (N/km^2) over a certain number of years (for example the last 10 years):

$$\overline{D}_{cell} = \frac{\sum_{y,h} N/km^2{}_{y,h}}{n}$$

where *n* is the count of the combinations year-haul in all the last 10 years. The variance of the mean abundance in the cells is calculated as:

$$Var_{celli} = \frac{1}{n-1}\sum_{y,h}\left(N/km^2{}_{y,h} - \overline{D}_{cell}\right)^2$$

The CV is calculated as the ratio between the standard deviation of the mean annual value by haul and year (numerator) and the mean biomass (or abundance) in the cell (denominator).

$$CV = \frac{\sqrt{Var_{cell_i}}}{\overline{D}_{cell_i}}$$

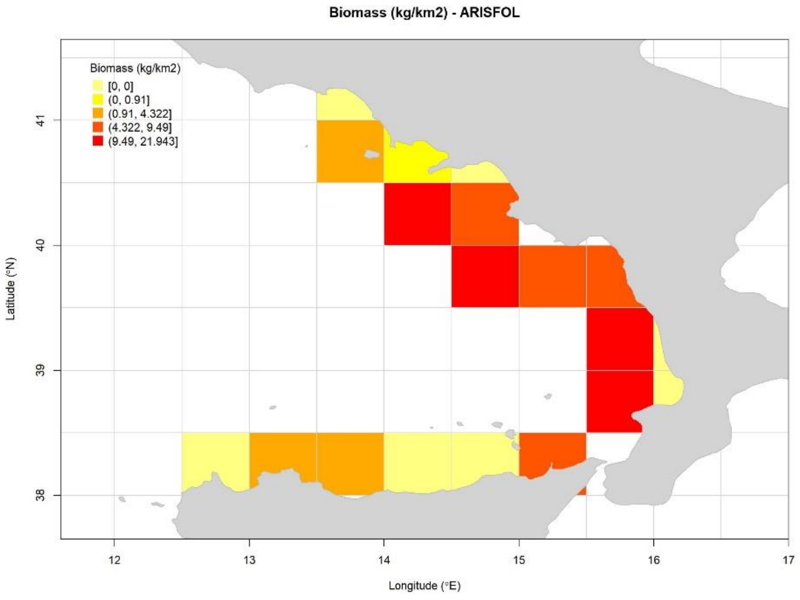


Fig. 4.1 Spatial pattern (by GFCM statistical rectangle) of biomass indices of *A. foliacea* (data from the demo).

Mean individual weight

For each GFCM cell the mean individual weight (Figure 4.2) is calculated by year y and haul h as ratio between the total weight W in the haul and the total number N in the haul (from the MEDITS samples data) as follows:

$$\overline{MIW}_{cell} = \frac{\sum_{y,h} \frac{W_{y,h}}{N_{y,h}}}{n}$$

where n is the count of the combinations year-haul in all the last 10 years.

The variance of the MIW in the *cells* is calculated using the following formula:

$$Var_{celli} = \frac{1}{n-1} \sum_{y,h} \left(\frac{W_{y,h}}{N_{y,h}} - \overline{MIW}_{cell} \right)^2$$

Then, the Coefficient of Variation (CV) of the MIW is calculated as:

$$CV = \frac{\sqrt{Var_{cell_i}}}{\overline{MIW}_{cell_i}}$$

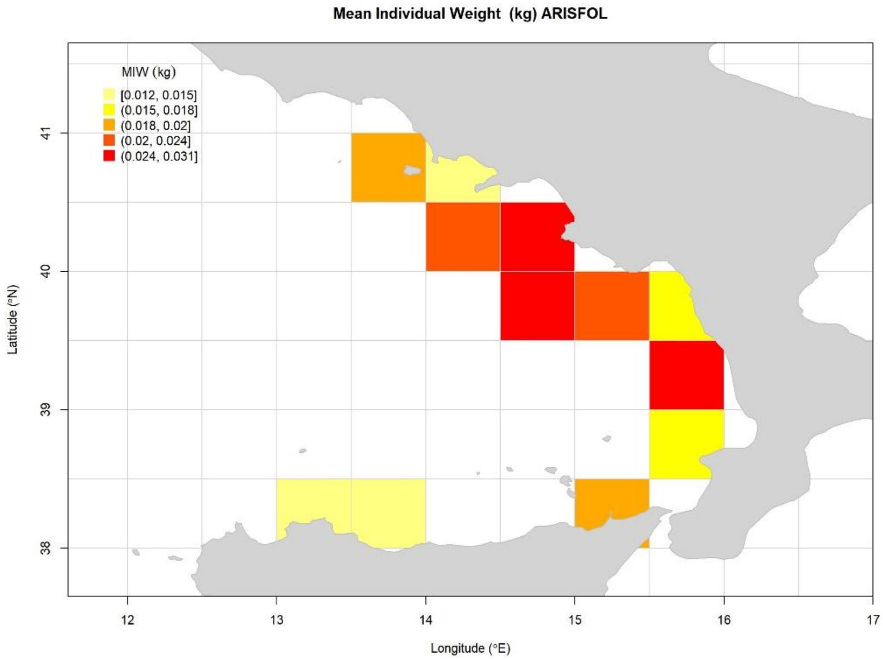


Fig. 4.2 Spatial pattern (by GFCM statistical rectangle) of mean individual weight of *A. foliacea*.

Sex ratio

The sex ratio in each GFCM cell (Figure 4.3) is calculated as the ratio between the sum of the standardized number of females and the sum of the standardized number of males and females over the hauls of the last 10 years:

$$SR_{cell} = \frac{\sum_{y,h} N_F / km^2{}_{y,h}}{\sum_{y,h} (N_F / km^2{}_{y,h} + N_M / km^2{}_{y,h})}$$

where N_F and N_M are the standardized number of the females and of males in the haul h and year y.

The variance of the sex ratio in the *cell* is calculated using the following formula:

$$Var_{cell_i} = \frac{1}{n-1} \sum_{y,h} \left(\frac{N_F / km^2{}_{y,h}}{N_F / km^2{}_{y,h} + N_M / km^2{}_{y,h}} - SR_{cell} \right)^2$$

where n is the count of the combinations year-haul in all the last pooled 10 years.

The CV is calculated as the ratio between the standard deviation of the sex ratio by haul and year to the sex ratio in the cell.

$$CV = \frac{\sqrt{Var_{cell_i}}}{SR_{cell_i}}$$

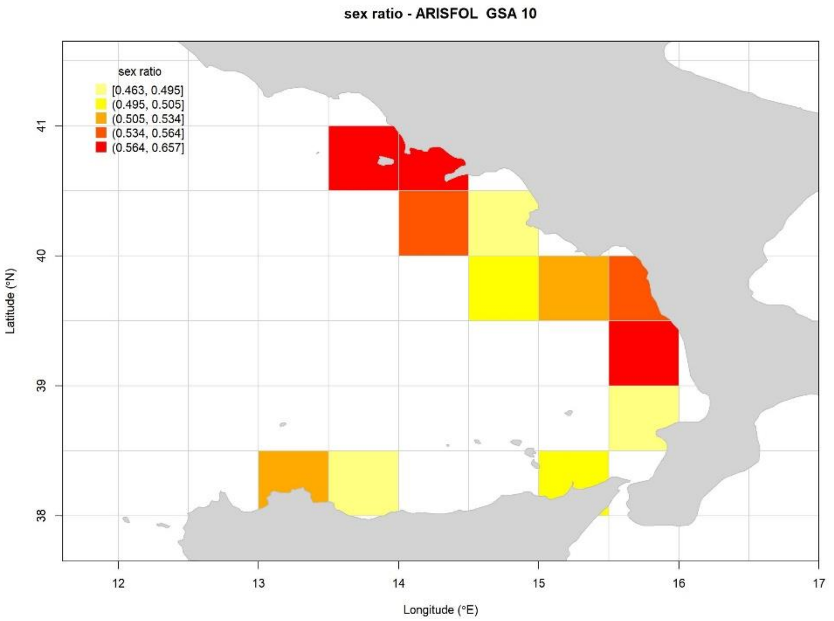


Fig. 4.3 Spatial pattern (by GFCM statistical rectangle) of sex ratio of *A. foliacea* (data from the demo).

The inverse of mean abundance Coefficient of Variation (CV)
The inverse of the coefficient of variation of the mean abundance by GFCM cell is plotted.

Recruits' and Spawners' indices
Plotting the bubble plots of the indices of abundance by year can be used to approximate the identification of areas - not using modelling - where critical life stages (juveniles and spawners) are more concentrated, using simple maps (Figure 4.4).

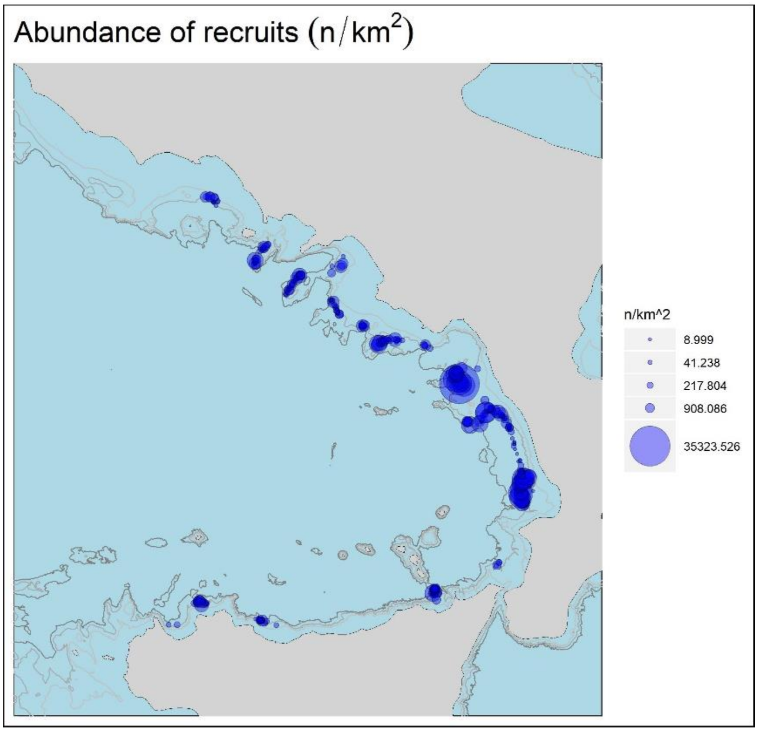


Fig. 4.4 Bubble plot of the number of recruits in the time series (data from the demo).

References

AA. VV. MEDITS-Handbook. Version n. 9, 2017, MEDITS Working Group : 106 pp. http://www.sibm.it/MEDITS%202011/principaledownload.htm

Adebiyi F.A. (2013). The sex ratio, gonadosomatic index, stages of gonadal development and fecundity of sompat grunt Pomadasys jubelini (Cuvier, 1830). Pakistan J. Zool., vol. 45(1), pp. 41-46.

Bertrand J.A., De Sola L.G., Papaconstantinou C., Relini G., Souplet A. 2002. The general specifications of the MEDITS surveys. Sci. Mar. 66: 9-17.

Cochran W. J. (1953). Sampling techniques. John Wiley & Sons, Inc., New York.

Cotter J., Petitgas P., Abella A., Apostolaki P., Mesnil B., Politou C-Y., Rivoirard J., Rochet M.-J., Spedicato M.T., Trenkel V.M., Woillez M. (2009a) - Towards an ecosystem approach to fisheries management (EAFM) when trawl surveys provide the main source of information. Aquat. Living Resour. 22, 243–254. DOI: 10.1051/alr/2009025.

Cotter J. (2009) A selection of nonparametric statistical methods for assessing trends in trawl survey indicators as part of an ecosystem approach to fisheries management (EAFM). Aquat. Living Resour. 22, 173–185.

Giannoulaki M., A. Belluscio, F. Colloca, S. Fraschetti, M. Scardi, C. Smith, P. Panayotidis, V. Valavanis M.T. Spedicato (edited by) (2013). Mediterranean Sensitive Habitats. DG MARE Specific Contract SI2.600741, Final Report, 557 p.

Granger V., Fromentin J.M., Bez N., Relini G., Meynard C., Gaertner J.C., Maiorano P., Garcia Ruiz C., Follesa C., Gristina M., Peristeraki P., Brind'Amour A., Carbonara P., Charilaou C., Esteban A., Jadaud A., Joksimovic A., Kallianiotis A., Kolitari J., Manfredi C., Massuti E., Mifsud R., Quetglas T., Refes W., Sbrana M., Vrgoc N., Spedicato M.T., Mérigot B. (2015) - Large spatio-temporal monitoring highlights shift in Mediterranean fish diversity hotspots. Progress in Oceanography 130: 65-74.

Lembo G. (coord.) (2010) – Identification and localization of main nursery areas of demersal species in the Italian sea - NURSERY. Final report (VI Plan MIPAAF), Italian Biology Association– S.I.B.M., Genova: pag. 1-119.

Nicholson M. D., and Jennings, S. 2004. Testing candidate indicators to support ecosystem-based management: the power of monitoring surveys to detect temporal trends in fish community metrics. ICES Journal of Marine Science, 61: 35-42.

Piet G.J. and Jennings S. (2005). Response of potential fish community indicators to fishing, ICES Journal of Marine Science, V.62-2, 214:225. Available on line at: http://icesjms.oxfordjournals.org/content/62/2/214.full.

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

Souplet A. (1996). Calculation of abundance indices and length frequencies in the MEDITS survey. In: J. A. Bertrand et al. (eds), Campagne internationale du chalutage démersal en Méditerraneé. Campagne 1995. EU Final Report, Vol. III.

Trenkel V.M., Rochet M.J., 2009. Intersection–union tests for characterising recent changes in smoothed indicator time series. Ecological Indicators, Volume 9, Issue 4: 732-739