ORIGINAL ARTICLE

Validation of macroscopic maturity stages according to microscopic histological examination for European anchovy

Rosalia Ferreri[1], Gualtiero Basilone[1], Marta D'Elia[1], Anna Traina[1], Francisco Saborido-Rey[2] & Salvatore Mazzola[1]

1 Istituto per l'Ambiente Marino Costiero (IAMC-CNR) U.O. di Torretta Granitola. Torretta Granitola, fraz. di Campobello di Mazara (TP), Italy
2 Institute of Marine Research (IIM-CSIC) Fisheries Ecology Research Group, Vigo, Spain

Abstract

The identification and classification of macroscopic maturity stages plays a key role in the assessment of small pelagic fishery resources. The main scientific international commissions strongly recommend standardizing methodologies across countries and scientists. Unfortunately, there is still a great deal of uncertainty concerning macroscopic identification, which remains to be validated. The current paper analyses reproductive data of European anchovy (*Engraulis encrasicolus* L. 1758), collected during three summer surveys (2001, 2005 and 2006) in the Strait of Sicily, to evaluate the uncertainty in the macroscopic maturity stage identification and the reliability of the macroscopic adopted scale. On board the survey vessels, the maturity stage of each fish was determined macroscopically by means of an adopted maturity scale subdivided in six stages. Later, at the laboratory, the gonads were prepared for histological examination. The histological slides were analysed, finally assigning the six maturity stages for macroscopic examinations. A correspondence table was obtained with the proportion and number of matches between the two methods. The results highlight critical aspects in the ascription of macroscopic maturity stages, particularly for the present research aim. Different recommendations were evaluated depending on the scope of the study conducted on maturity (*e.g.* daily egg production, fecundity and maturity ogive computation). The most interesting results concern the misclassification of stage IV and stages III and V (the most abundant), which confirms their macroscopic similarity. Although the results are based on a small number of samples, the advantages and disadvantages of macroscopic and histological methods are discussed with the aim to increase the accuracy of correct identification and to standardize macroscopic maturity ascription criteria.

Problem

The small pelagic species, especially sardines and anchovies, represent an important resource worldwide. In particular, the European anchovy (*Engraulis encrasicolus*) is widely distributed along the European Atlantic coast from South Africa to North Atlantic, and over the whole of the Mediterranean and Black Sea. This species was the one most fished by the Italian fleet in the last 4 years (IREPA – *Istituto Ricerche Economiche per la Pesca e l'Acquacoltura*, 2008). These resources undergo wide interannual biomass fluctuation mainly due to the effects of environmental variability on recruitment. Sustainable management and exploitation of fish resources are linked to the Stock Reproductive Potential concept (SPR; Trippel 1999). The value of reproducible potential for stock assessment appears to be very important for several commercially important demersal or semi-demer-

sal species (Murua & Saborido-Rey 2003; Murua *et al.* 2003).

In recent decades, the refinement of knowledge on the growth and development of fish oocytes (Yanamoto & Yoshoka 1964; Wallace & Selman 1981) and reproductive biology (Begenal 1973; Hunter & Goldberg 1980) has permitted more precise definitions of fish fecundity. The Daily Egg Production Method (DEPM) was developed in the late 1970s by the Coastal Division of the Southwest Fisheries Center, La Jolla, CA, USA (Parker 1980) as a response to the growing need to devise a suitable direct method for the assessment of Northern anchovy (*Engraulis mordax*), an indeterminate spawner with pelagic eggs (Hunter & Goldberg 1980). In the case of indeterminate spawners, the DEPM is based on the number of oocytes released per fish in each spawning event (batch fecundity) and the proportion of females reproducing daily (spawning fraction). Estimation of spawning fraction for Northern anchovy is possible because of the identification of post-ovulatory follicles (POFs; cellular remnants in the ovary after ovulation; Hunter & Goldberg 1980). Estimates of relative fecundity (batch fecundity divided by mean female weight) and spawning fraction give the number of eggs released per unit weight of mature females, which, multiplied by the sex ratio in the population expressed in weight, provides daily fecundity (number of eggs released per unit weight of stock). In the 1980s, applications rapidly extended to several important anchovy stocks worldwide. DEPM revisions in the 1990s (Alheit 1993; Hunter & Lo 1997) and 2000s (Stratoudakis *et al.* 2006) confirmed its wide scope and potential, its robustness in estimating spawning biomass and indicated possible challenges for better future applications.

In multiple-spawning fishes the determination of macroscopic maturity stages is difficult to achieve without the support of microscopic examination. Because of its subjectivity and variability, macroscopic examination of ova-ries plays a key role in the assessment of fishery resources. The identification and classification of maturity stages are used for the determination of spawning period according to different geographical and environmental areas and for studying the relationship between length at maturity and fishery exploitation (Picquelle & Hewitt 1983; Armstrong *et al.* 1988; Pérez *et al.* 1989; Millan 1999). Several international commissions for fisheries studies (*i.e.* ICES – International Council for the Exploration of the Sea, PICES – North Pacific Marine Science Organization, GFCM, – General Fisheries Commission for the Mediterranean) strongly recommended standardizing methodologies across countries and researchers. In the present study the gonads maturity data from the Strait of Sicily European anchovy (*E. encrasicolus* L. 1758) were examined. The results of macroscopic and microscopic methods were compared to verify the correspondence or the discordance of these two classification methods. Critical aspects in the ascription of macroscopic maturity stages were investigated. An improvement in macroscopic classification performance during the three study years can be seen. The obtained results underline the importance of adopting the histological analysis when the investigation required the highest possible accuracy.

Material and Methods

The anchovy samples were collected during the peak spawning period in the Strait of Sicily (July–August), onboard a research vessel equipped with a mid-water pelagic trawl (experimental design). DEPM surveys were made by IAMC – CNR during 2001, 2005 and 2006. From each trawl a subsample of 50 specimens was randomly collected and processed, measuring total and standard length (±1 mm) and total and somatic weight (±0.1 g). The sex was determined and the fish gonads were macroscopically classified (Fig. 1).
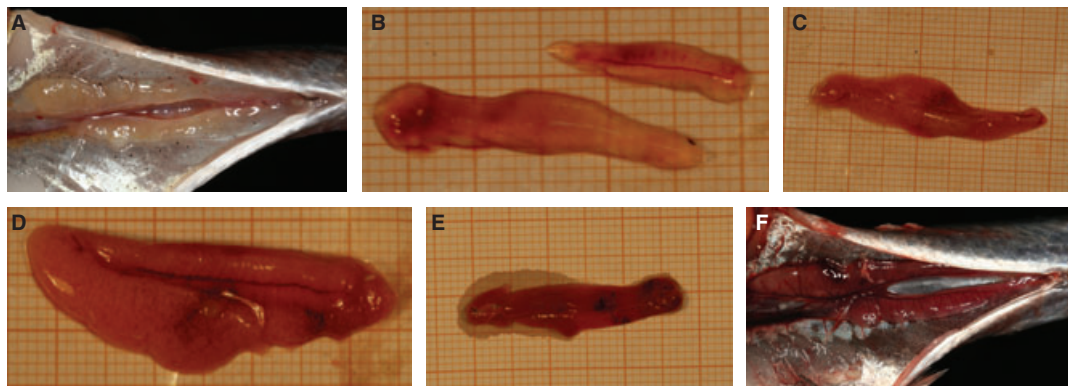

Fig. 1. Six macroscopic maturity stages of ovary. A: immature ovary in body cavity. B: developing ovary. C: imminent spawning. D: spawning. E: partial post-spawning. F: spent ovary in body cavity.

Table 1. Description of macroscopic and microscopic characteristics of each maturity stage.

stage	stage name	macroscopic characteristics	microscopic characteristics
I	Immature or rest	Invisible or very small ovaries (cord shaped), translucent or slightly coloured (when resting)	In the case of immature females, all the oocytes in the ovary were in primary growth stage and oocytes well packaged. Resting females might contain late atretic oocytes
II	Developing	Wider ovaries occupying 1/4 to 1/3 of body cavity; pinkish or yellow colour. Visible oocytes are not present	Occurrence of cortical alveoli and no signs of advanced spawning process such as thick ovary wall, high vas cularization of gonad
III	Imminent spawning	Ovaries occupying 3/4 to almost fitting body cavity; opaque with yellow or orange colour. Opaque oocytes are visible.	Occurrence of vitellogenic oocytes, but post-ovulatory follicles are not present and some oocytes in nucleus migration phase could be observed
IV	Spawning	Large ovaries occupying the full body cavity; fully or partially translucent with gelatinous aspect. Hyaline oocytes are visible	This stage starts with the nucleus migration phase fol lowing by oocyte hydration, without post-ovulatory follicles
V	Partial post-spawning	Size from 1/2 to 3/4 of abdominal cavity; not-turgid ovaries with haemorrhagic zones. Blood coloured	Post-ovulatory follicles are observed together with vitellogenic oocytes in different stages. Signs of advanced spawning process could be detected. Atresia could also be present.
VI	Spent	Reddish ovary shrunked; Size <2/3 of abdominal cavity Flaccid ovary. Some small opaque oocytes	High level of POF and atresia, disorganization of ovary structures, numerous blood vessels, absence of yolked oocyte groups

The maturity stage for each fish was determined using the six-degree macroscopic maturity scale shown in Table 1 (Holden & Raitt 1974, modified for six stages; ICES 2004, modified by authors). The ovaries were fixed in 4% buffered formalin. For species of the genus *Engraulis* it is possible to exploit the method of Harris Hematoxylin and Eosin (H & E). The preparation of samples with the H & E method includes five steps (Hunter & Macewicz 1985): (i) fixation; (ii) infiltration; (iii) sectioning; (iv) mounting; (v) staining. The microscopic data were also arranged according to the six maturity stages from the macroscopic scale (Fig. 2). The six maturity stages could be split into immature (stages I and II) and mature (stages III to VI).

The operator was the same during the study period. For samples collected during the acoustic survey the identification of the six microscopic stages was based on the characteristics of the most advanced oocyte stage (West 1990). The microscopic and macroscopic maturity data were compared to assess the correspondence levels achievable by an advanced operator using macroscopic examination. The proportion of a robust classification and its variability was also evaluated.

Results

The comparison results contain the following observations: stage I is not widely represented in the samples (around 3%), most probably because the surveys were conducted during the peak anchovy spawning season. Stages I and II represent immature ovaries. The main staging mistake resulted in the misclassification of stage V as stage I in 6/14 examined ovaries (43%, Table 3).

Macroscopic examination correctly classified stage II in only 7/39 ovaries (18%, Table 3). However, the bulk of misclassification occurred between stages II and V (44%, Table 3).

Stage III is the most abundant in the samples (Table 3), and shows the highest matching proportion between micro- and macroscopic classifications (Table 3). After a rather low matching rate in 2001, the correspondence between the two classification methods

Table 2. The percentage of correspondence between macroscopic (MACRO) and microscopic (MICRO) examination in three study years; the number of specimens is given in parentheses.

survey	MACRO	I	II	III	IV	V	VI	III + V
2001	I	0	–	–	–	100 (1)	–	
	II	–	0	–	–	–	–	
	III	–	9 (2)	35 (8)	39 (9)	17 (4)	–	52
	IV	–	–	24 (5)	24 (5)	52 (11)	–	
	V	–	–	55 (6)	18 (2)	27 (3)	–	82
	VI	–	–	100 (2)	–	–	0	
2005	I	0	–	–	–	100 (1)	–	
	II	33 (1)	0	67 (2)	–	–	–	
	III	–	5 (8)	71 (108)	6 (10)	18 (28)	–	89
	IV	–	–	24 (5)	24 (5)	52 (11)	–	
	V	–	6 (8)	50 (69)	4 (5)	39 (54)	1 (1)	89
	VI	11 (1)	23 (2)	33 (3)	–	33 (3)	0	
2006	I	15 (2)	15 (2)	23 (3)	–	39 (5)	8 (1)	
	II	17 (6)	19 (7)	8 (3)	6 (2)	47 (17)	3 (1)	
	III	–	1 (1)	66 (57)	5 (4)	28 (24)	–	94
	IV	–	–	62 (5)	38 (3)	–	–	
	V	2 (1)	2 (1)	41 (25)	–	52 (25)	3 (2)	93
	VI	33 (1)	–	–	–	33 (1)	33 (1)	

Table 3. The overall percentage of correspondence between macroscopic (MACRO) and histological examinations (MICRO) for each maturity stage in the whole study period (2001, 2005 and 2006); the number of specimens is given in parentheses.

MACRO	MICRO						
	I	II	III	IV	V	VI	III + V
I	14 (2)	14 (2)	21 (3)	–	43 (6)	7 (1)	
II	18 (7)	18 (7)	13 (5)	5 (2)	44 (17)	3 (1)	
III	–	4 (9)	69 (165)	6 (14)	22 (52)	–	91
IV	–	–	34 (10)	28 (8)	38 (11)	–	
V	1 (1)	5 (9)	48 (94)	3 (5)	43 (85)	2 (3)	91
VI	17 (2)	17 (2)	25 (3)	–	33 (4)	8 (1)	

increased about 70% in 2005 and 2006 (Table 2). The most common misclassification occurred from stage III to stage V because these stages are different for only a short period after spawning, when POFs have not yet been reabsorbed. In the multiple-spawner species there is a continuous transition from stage III to V and from V to III during the spawning season. Macroscopically, these two stages are very similar and several external factors can make stage III appear as stage V (*i.e.* trawl time, trawl type, presence of damaging object with the sample). The percentage of stage V that was classified as stage III macroscopically was 22%, (52/240 samples) (Table 3). When the two stages were merged, the per-

centage of agreement between macroscopic and microscopic classifications increased from 52% and 82% in 2001 (Table 2) and 89% in 2005 (Table 2), to over 90% in 2006 (Table 2).

Stage IV (hydrated) is recognizable only during the daily peak of spawning, for few hours. That explains the low presence of these stages in the samples (Table 3), although it is well known from the literature that this stage is easily recognized by its evident attributes (Fig. 1; Holden & Raitt 1974 modified for six stages; ICES 2004). The bulk of ovaries (21 of 29), macroscopically identified as stage IV, were microscopically assigned to stage III (10, Table 3) and stage V (11, Table 3), 34% and 38%, respectively (Table 3). The correspondence was weak for stage IV in all surveys (Table 2).

During the study period, the correspondence between macroscopic and microscopic maturity classifications increased mainly for stages IV and V.

The abundance and percentage of correspondence was highest for stage III, followed by stage V (Table 3, Fig. 3). The bulk of misclassification was between these stages: 94/197 checked ovaries (Table 3), equal to 48% (Table 3).

The percentage of correct classification for stage VI was very low (8%, Table 3) because it was confused mainly with stage V (4/12 ovaries; Table 3). Analysing data by year, it is clear that misclassification is always present,
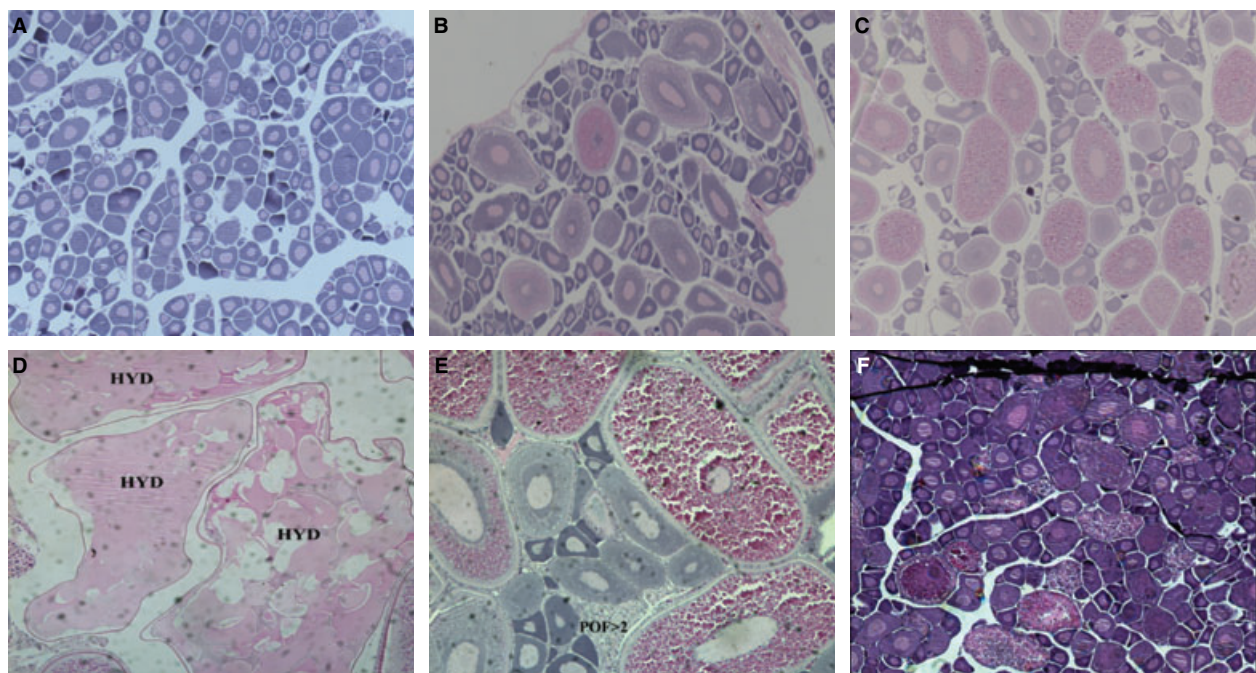

Fig. 2. Six microscopic maturity stages of ovary. A: immature oocytes. B: developing ovary with oocytes in nucleus migration. C: imminent spawning ovary with vitellogenic oocytes. D: spawning ovary with hydrated oocytes (HYD). E: partial post-spawning ovary with old POF (POF >2). F: spent ovary.
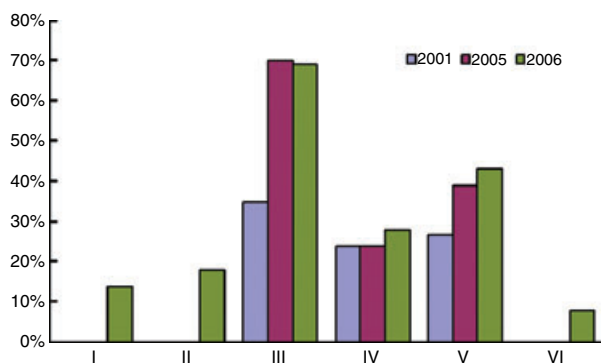
Fig. 3. Bar chart of percent correspondence between macroscopic and microscopic classifications.

despite an improvement during 2006: however, there is a correspondence between macro- and microscopic classification (Fig. 3), even if low: only 1/12 samples (Table 3).

Discussion

Generally, it was observed that for anchovy, as for many other species, it is not possible to distinguish immature and resting females macroscopically (Trippel & Morgan 1997; Saborido-Rey & Junquera 1998; Domínguez-Petit 2007), In both stages the oocytes are not visible. Only by means of histological analyses is it possible to identify each stage accurately. This misclassification has an impact on the estimation of the mature proportion of the stock because resting females have already contributed to the spawning biomass of that year and are macroscopically considered immature.

Stages I and II are not widely represented in the samples, probably due to the sampling periods, which occurred during the peak of the anchovy spawning season. The large misclassification of stages I and II with stage V is remarkable, especially when the aim of analysis is to discriminate mature and immature specimens.

By comparing the results with literature data (ICES, 2004), it is possible to see that the misclassification between III and V stages is a recurring problem and that the match rate between macroscopic and microscopic classification increases when the two stages are merged. This is mainly because these stages differ for only a short time after spawning, when the POFs are not yet reabsorbed, which, in the studied area, with high mean water temperature, represents a very fast process (a few hours). In the multi-spawner species, the change between stages III and V is a continuous process during the spawning season until the last batch has been spawned. During the study period, there is an increasing proportion of correct ascription, but it appears lower than the values reported in the literature, *e.g.* the anchovies from Portugal (100%

matching in Stage III; ICES, 2004). In any case, there appears to be a clear need for a training period onboard the fishing vessels to be able to discriminate the maturity stages of very fresh gonads. Freezing, formalin or alcohol preservation is known to produce deep alterations at a macroscopic level, mainly in the colour and consistency of ovaries. As reported in Lasker (1980) the gonad tissues have to be sampled within 2 h of the catch, before degradation becomes significant. Bearing in mind that Lasker used the hydrated stage for histological preparation, some hours more could be a reasonable time within which to carry out the macroscopic examination.

Further interesting comments concern stage IV. Despite this stage being generally easily recognizable by visual inspection of fresh individuals, the bulk of ovaries macroscopically identified as stage IV became stage III and V when later microscopically examined. Stage IV could easily become stage V when the spawning begins. So if only a few eggs are spawned, the macroscopic aspect does not differ compared to the entire hydrated oocyte. However, it should be noted that the percentages for stage IV are based on only a few observations, and for this reason their representativeness is proportionally low.

During the whole study period, stage VI was rarely correctly identified. The misclassification between stage VI and stages I and II is particularly problematic, because stage VI represents mature fish and stages I and II immature fish. In addition, the proportions of stage VI are based on a small number of observations.

Maturity ogives should only be based on data collected during the peak of the spawning season, taking into consideration geographical variation, as it is impossible to distinguish immature and resting females macroscopically. The proportion of resting females during the peak of the spawning season is lower than during the rest of the year. If possible, a gonad subsampling for histological analysis should be carried out to obtain a correction factor to reduce the impact of misclassification between immature and resting females.

Furthermore, the samples should be analysed as soon as caught, when the ovaries are in their best condition, allowing the right macroscopic maturity stage to be assigned more reliably. However, samples are generally analysed after a preservation period in ice or formalin, which changes the colour and texture. Also, frozen gonads are not appropriate for histological examination. When fecundity has to be estimated for stock assessment purposes, it is necessary to validate the macroscopic classification with a microscopic one, to reduce mistakes.

The results from macroscopic and microscopic maturity classifications could be used for several applications. In the case of the estimation of size at first maturity of the population (L_{50}) the macroscopic examination would

be enough to discriminate immature from mature specimens, as follows: immature = stage I and II; mature = III to VI. Histological analysis would not be necessary if macroscopic classification were applied, especially when a high number of individuals have to be processed. Also, in the latter case, a higher level of accuracy may be achieved when a representative subsampling, for further histological examinations, can be done. When macro- and microscopic examinations are compared, it is possible to build a table, as in the present study (Table 3), where the correspondence proportion for each macroscopic maturity stage can be used as a correction factor for all the sampled specimens classified macroscopically. Finally, because maturity ogives are generally correlated to the length of the fish, a correction table has to be obtained for size classes. From the present study the misclassification between stage I *versus* V, stage II *versus* V and stage VI *versus* I and II, has to be corrected to avoid a significant source of mistake. With regard to the fecundity estimation, the misclassification rate from macroscopic examination is very large when determining consistent variations in batch fecundity (F) and in spawning fraction (S) estimation; consequently, there is a greater uncertainty in the stock assessment abundance estimations (*e.g.* DEPM applications). As known from the literature (Hunter & Macewicz 1985) and as supported by the present data, in multiple-spawning fishes, histological analysis is essential when the DEPM is applied.

During the three study years there was an improvement in the macroscopic classification performance of the operator: clearly, experience helps to improve the ability to recognize different stages; however, it is also true that the level of misclassification in the last year was high enough to compromise the estimation of the DEPM parameter.

Finally, the interpretation of maturity scales and determination of maturity stages may vary considerably among people and labs because no exchanges or standardization meetings are conducted regularly among countries where similar resources are studied. Sometimes, when datasets from different countries are compared, this variability may be larger than the misclassification errors. The obtained results confirm the importance of common and standardized protocols for the identification and macroscopic classification of maturity stages among scientists from different country. As a future objective, to overcome this problem, we suggest that common training and intercalibration exercises be extended as much as possible to all the labs that study fish fecundity.

The correspondence analysis between macroscopic and microscopic classifications underlined the weakness of macroscopic classification even when conducted by the same operator with the best sampling conditions (direct on board) and following clear maturity scale criteria

widely accepted by the scientific community. Unfortunately, when the above recommendations about intercalibration exercises and samples freshness, are followed the misclassification proportion remains highly significant. The reasons for the discrepancies between macroscopic and histological observations are mainly as follows: operator subjectivity; unclear distinction between the descriptions of maturity stages in the reference tables; daytime sampling; the level of fish injury due to the trawl; and the elapsed time from the catch to examination of the samples. Therefore, histological preparation appears to be the only way to obtain an accurate classification (*i.e.* DEPM applications, fecundity studies, *etc.*). When the study aims do not require such high standards, it is possible to carry out the macroscopic examinations with the scale used here with satisfying results (*i.e.* maturity ogives computation).

References

Alheit J. (1993) Use of the daily egg production method for estimating biomass of clupeoid fishes: a review and evaluation. *Bulletin of Marine Science*, **53**, 750–767.

Armstrong M., Shelton P., Hampton I., Jolly G., Melo Y. (1988) Egg production estimates of anchovy biomass in the Southern Benguela system. *CalCOFI Reports*, **XXIX**, 137–157.

Begenal T.B. (1973) Fish fecundity and its relations with stock and recruitment. *Rapports and Proces-Verbeaux des Reunions, Conseil International pour l'Exploration de la Mer*, **164**, 186–198.

Domínguez-Petit R. (2007) *Study of Reproductive Potential of Merluccius Merluccius in the Galician Shelf*. Doctoral Thesis. University of Vigo, Vigo, Spain: 253 pp.

Holden M.J., Raitt D.F.S. (1974) Manual of fisheries science. 2. Methods of resource investigation and their application. *FAO Fisheries Technical Paper*, 115, Rev. 1: 211 pp.

Hunter J.R., Goldberg S.R. (1980) Spawning incidence and batch fecundity in Northern Anchovy (Engraulis nordax). *Fishery Bulletin US*, **78**, 811–816.

Hunter J.R., Lo N.C.H. (1997) The daily egg production method of biomass estimation: some problems and potential improvements. *Ozeanografika*, **2**, 41–69.

Hunter J.R., Macewicz B. (1985) Measurement of spawning frequency in multiple spawning fishes. In: Lasker R. (Ed.), *An Egg Production Method for Estimating Spawning Biomass of Pelagic Fish: Application to the Northern Anchovy, Engraulis Mordax. NOAA Technical Report NMFS*, **36**, 79–93.

ICES (2004) The DEPM estimation of spawning-stock biomass for sardine and anchovy. *Rapport des Recherches Collectives*, Vol. 268, ICES, Pasara: 95 pp.

Lasker R. (1980) An Egg Production Method for Estimating Spawning Biomass of Pelagic Fish: Application to the North-

ern Anchovy, *Engraulis mordax. NOAA Technical Rep. NMFS*, **36**, 59–62.

Millan M. (1999) Reproductive characteristics and condition status of anchovy *Engraulis encrasicolus* L. from the Bay of Cadiz (SW Spain). *Fisheries Research*, **41**, 73–86.

Murua H., Saborido-Rey F. (2003) Female reproductive strategies of marine fish species of the North Atlantic. *Journal of Northwest Atlantic Fishery Science*, **33**, 23–31.

Murua H., Kraus G., Saborido-Rey F., Witthames P., Thorsen A., Junquera S. (2003) Procedures to estimate fecundity of marine fish species in relation to their reproductive strategy. *Journal of Northwest Atlantic Fishery Science*, **33**, 33–54.

Parker K. (1980) A direct method for estimating northern anchovy, Engraulis mordax, spawning biomass. *Fishery Bulletin US*, **78**, 541–544.

Pérez N., García A., Lo N.C.H., Franco C. (1989) The egg production method applied to the spawning biomass estimation of sardine (*Sardina pilchardus* W) in the North-Atlantic Spanish coast. *ICES CM Documents 1989*, **H23**, 1–20.

Picquelle S.J., Hewitt R.P. (1983) The Northern Anchovy spawning biomass for the 1982 and 1983 California fishing season. *CalCOFI Report*, **24**, 16–28.

Saborido-Rey F., Junquera S. (1998) Histological assessment of variations in sexual maturity of cod (Gadus morhua L.) at the Flemish Cap (north-west Atlantic). *ICES Journal of Marine Science*, **55**, 515–521.

Stratoudakis Y., Bernal M., Ganias K., Uriate A. (2006) The daily egg production method: recent advances, current application and future challenges. *Fish and Fisheries*, **7**, 35–57.

Trippel E.A. (1999) Estimation of stock reproductive potential: history and challenges for Canadian Atlantic gadoid stock assessments. *Journal of Northwest Atlantic Fishery Science*, **25**, 61–81.

Trippel E.A., Morgan M.J. (1997) Skewed sex ratios in spawning shoals of Atlantic cod (Gadus morhua). *Oceanographic Literature Review*, **44**, 511 pp.

Wallace A.R., Selman K. (1981) Cellular and dynamic aspects of oocyte growth in teleosts. *American Zoologist*, **21**, 325–343.

West G. (1990) Methods of assessing ovarian development in fishes: a review. *Australian Journal of Freshwater Research*, **41**, 199–222.

Yanamoto K., Yoshoka H. (1964) Rhythm of development in the oocytes of medaka, *Oryzias latipes. Bulletin of the Faculty of Fisheries, Hokkaido University*, **15**, 5–19.