

A revision of daily egg production estimation methods, with application to Atlanto-Iberian sardine. 1. Daily spawning synchronicity and estimates of egg mortality

Miguel Bernal^{1*†}, Yorgos Stratoudakis², Simon Wood³, Leire Ibaibarriaga⁴, Andres Uriarte⁵, Luis Valdés⁶, and David Borchers⁷

¹Instituto Español de Oceanografía (IEO), Centro Oceanográfico de Cádiz, Puerto Pesquero, Muelle de Levante s/n, Apartado 2609, 11006 Cádiz, Spain

²Instituto Nacional de Recursos Biológicos (INRB/IPIMAR), Avenida de Brasília 1449-006, Lisboa, Portugal

³Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK

⁴AZTI-Tecnalia, Marine Research Unit, Txatxarramendi ugarte a z/g, 48395 Sukarrieta, Basque Country, Spain

⁵AZTI-Tecnalia, Marine Research Unit, Herrera Kaia Portualdea z/g, 20110 Pasaia, Basque Country, Spain

⁶Intergovernmental Oceanographic Commission of UNESCO, 1 rue Miollis, 75732 Paris cedex 15, France

⁷CREEM, University of St Andrews, The Observatory, Buchanan Gardens, St Andrews, Fife KY16 9LZ, Scotland, UK

*Corresponding Author: tel: +34 932 309500; fax: +34 932 309555; e-mail: miguel.bernal@cd.iao.es.

†Present address: Institute of Marine and Coastal Sciences, University of Rutgers, 71 Dudley Road, New Brunswick, NJ 08901, USA.

Bernal, M., Stratoudakis, Y., Wood, S., Ibaibarriaga, L., Uriarte, A., Valdés, L., and Borchers, D. 2011. A revision of daily egg production estimation methods, with application to Atlanto-Iberian sardine. 1. Daily spawning synchronicity and estimates of egg mortality. – ICES Journal of Marine Science, 68: 519–527.

Received 1 June 2010; accepted 2 January 2011.

Assumptions of daily spawning synchronicity and estimation of egg mortality for Atlanto-Iberian sardine are revised in the context of daily estimators of egg production. An extensive database of ichthyoplankton surveys from 1985 to 2008, aggregated at different levels, is used, a set of mortality models is derived, and a comparison among them is carried out using standard statistical techniques. Analysis of the database shows Atlanto-Iberian sardine to be a late-evening spawner, in agreement with previous knowledge, but suggests a lengthier daily period of spawning. Comparison among the set of mortality models used suggests that estimates of mortality from single surveys for Atlanto-Iberian sardine are often unreliable, and that the mean estimates obtained after aggregating data from various years are statistically significant, more robust, and in line with existing information gathered from the literature.

Keywords: egg production, mortality, *Sardina pilchardus*, sardine, spawning synchronicity.

Introduction

Estimates of egg production are used worldwide to evaluate the spawning-stock biomass of a variety of fish species (see, for example, the review in Stratoudakis *et al.*, 2006), especially small pelagic fish, but also demersal fish (Armstrong *et al.*, 2001; Jackson and Cheng, 2001). For small pelagic fish, egg production methods together with acoustic estimates of biomass (Simmonds and MacLennan, 2005) are the most commonly used fishery-independent methods of evaluation employed (see comparison in Hampton, 1996, and a recent review in Barange *et al.*, 2009).

There are various families of egg production methods, and differences in their application to stocks and species around the world (Stratoudakis *et al.*, 2006). Around the Iberian Peninsula, sardine (*Sardina pilchardus*) is one of the main resources of marine animal protein, and is currently assessed by Virtual Population Analysis (ICES, 2007; Barange *et al.*, 2009), using both acoustic and daily egg production method (DEPM) estimates of biomass as tuning indices. The DEPM was developed in the late 1970s to assess Californian anchovy (Parker, 1980; Lasker, 1985) and was adapted for Atlanto-Iberian sardine in the late 1980s

(Miranda *et al.*, 1990; Cunha *et al.*, 1992; García *et al.*, 1992); it has been used routinely for assessment since then (ICES, 2004).

In the DEPM, estimates of egg age and mortality are required to convert egg abundance into egg production (i.e. abundance-at-age 0). The basic model relating observed egg abundance to production is a variation of a simple population growth model developed in the early ecological literature (e.g. Verhulst, 1839; Lotka, 1925):

$$D_a = D_0 e^{(-ma)}, \quad (1)$$

where D_a is the density of eggs at age a , D_0 the rate of egg production (i.e. the number of eggs produced per unit of time, generally day, and per unit of area), and m the rate of mortality. Equation (1) represents a basic mortality model in which the abundance of eggs decreases exponentially with age, at a constant or mean exponential rate m . A slightly modified version of this equation, which is also known as a “mortality curve”, is one of the pillars of the DEPM (Parker, 1980; Stratoudakis *et al.*, 2006). Estimates of both D_0 and m are obtained by fitting the mortality curve to the sampled density of eggs classified into cohorts. To fit Equation (1), it is

necessary to have (i) estimates of egg age, and (ii) criteria to aggregate the eggs into classes for which abundance and mean age can be estimated. Estimates of age are obtained using information on the rates of egg development, which are obtained from incubation experiments run under controlled environmental conditions (e.g. [Miranda et al., 1990](#)). These estimates are refined using information on daily spawning synchronicity when this is available ([Lo, 1985](#)). A normal distribution assumption is often used in traditional applications of daily spawning distribution for small pelagic fish (e.g. [Lo, 1985](#)). The appropriate means of aggregating eggs for fish that spawn synchronously is to group the eggs into daily cohorts ([Lo, 1985](#)), because in that case, the eggs released on a single day can be distinguished easily from those released on other days.

A number of problems with the assumptions of both the general DEPM method and its application to Atlanto-Iberian sardine and anchovy have recently been identified ([ICES, 2004](#); [Stratoudakis et al., 2006](#)). Spurious positive estimates of egg mortality (i.e. “natality”) can arise from the data collected during egg production surveys ([McGarvey and Kinloch, 2001](#); [Somarakis et al., 2004](#)), either as a result of sampling problems or a low signal-to-noise ratio. Also, bias in mortality estimates may arise from a variety of problems affecting the age tails of the dataset, i.e. the densities of very young and very old eggs, used to estimate the mortality curve ([Stratoudakis et al., 2006](#)). In general, the underlying assumptions on fixed daily spawning synchronicity and constant rates of daily egg production and mortality may fail, and may, for a given survey of egg production, lead to unrealistic or biased estimates, e.g. as suggested by [Parker \(1980\)](#).

In this paper, a statistical analysis of all available large-scale ichthyoplankton surveys for Atlanto-Iberian sardine is carried out to check the assumptions underlying the mortality-curve model, with the objective of obtaining reliable estimates of mortality. Daily spawning synchronicity is analysed using both survey data and back-computed birthdates, and a recent method of age determination is used to assign ages to staged eggs. Survey data are aggregated at different spatial and temporal scales, to determine the minimum quantity of data required to obtain reliable, i.e. statistically significant and biologically plausible, average estimates of mortality. The robustness of the final estimates to the assumptions taken in this analysis is investigated, and estimated confidence intervals of mortality are provided. These estimates are then used in a separate paper ([Bernal et al., 2011](#)) to review the whole time-series of estimates of egg production of Atlanto-Iberian sardine.

Material and methods

Most of the data used for estimating egg mortality in Atlantic waters of the Iberian Peninsula were compiled within the European project SARDYN (EU QLRT-2001-00818) and used to analyse Atlanto-Iberian sardine spawning areas by [Bernal et al. \(2007\)](#). The data, for the years 1985–2005, cover the area from the Strait of Gibraltar to the boundary between the Spanish and French continental shelves (see Figure 1 in [Bernal et al., 2011](#)). In addition to these data, results from a combined Spanish and Portuguese ichthyoplankton survey carried out in 2008 ([ICES, 2009](#)) were used. Data from the Armorican Shelf available within the SARDYN project were not used in the analysis, because they did not cover the temporal range of interest (1985–2005). The data used for the work herein were stratified as described below, following the recommendations of [Bernal et al. \(2007\)](#): (i) three spatial strata, south (from the Strait of

Gibraltar to Cape San Vicente), west (from Cape San Vicente to Rio Miño, the northern border of Spain and Portugal), and north (from Rio Miño to the border between the Spanish and French continental shelves); and (ii) two temporal strata, 1985–1994 and 1995–2008. The spatial strata were selected to represent three spawning nuclei, and the two temporal strata to represent two periods with different extents of occupancy of the shelf ([Bernal et al., 2007](#)). All surveys within these limits for which data on egg abundance by stage and temperature are available were used to estimate mortality. To compare with earlier estimates, and to illustrate the lack of significance of mortality estimates obtained from a single survey, mortality estimates are also provided by year and stratum (see below).

Egg age and mortality

Daily spawning synchronicity

The distribution of spawning events during the day (“daily spawning frequency”) is inferred from the distribution of sampling times at which stage I eggs (the first stage of sardine embryonic development after fertilization, with a duration of a few hours; [Miranda et al., 1990](#)) were observed in the samples. Spawning times for stage I eggs are estimated using the sampling times and average stage I age at the surface temperature of each station, i.e. spawning time = sampling time – mean age ([Bernal et al., 2001](#)). Unfertilized or non-viable eggs can appear in the data as young eggs (early stages) but do not develop to later stages, so can create some bias in the distribution of spawning times estimated ([Motos, 1994](#)), so to reduce this potential bias, the distribution of spawning times derived from egg-sampling distribution is also compared with a few available samples of running females. Running females are fish with both hydrated oocytes and very recent post-ovulatory follicles, and are therefore good indicators of spawning times ([Ganias et al., 2003](#)), although they are usually scarce in the surveys.

Once the daily frequency distribution of spawning times had been obtained, a probability density function (pdf) was fitted to these data by maximum likelihood. Two different distributions were fitted to the data: (i) a normal distribution (similar to the assumption of [Lo, 1985](#)), and (ii) a lognormal distribution to account for skewed distribution of spawning events, such as that found by [Motos \(1994\)](#). The distributions fitted to the data were then used as priors in estimating age, as explained below.

Estimation of age and cohort abundance

A recently published multinomial model of sardine egg development ([Bernal et al., 2008](#)) was used to relate egg stage and age for the sampled temperatures. Egg age determination was performed using the Bayesian method proposed in [ICES \(2004\)](#) and briefly described in [Ibaibarriaga et al. \(2007\)](#) and [Bernal et al. \(2008\)](#). Essentially, ageing was performed by updating a prior distribution of ages derived from the distribution of spawning times and actual sampling times, using information from the multinomial model of egg development:

$$f(\text{age}, \text{temp}) = \frac{p(\text{stage}, \text{temp})\pi(\text{age})}{p(\text{stage})} \quad (2)$$

$$f(\text{age}, \text{temp}) \propto p(\text{stage}, \text{temp})\pi(\text{age}),$$

where $\pi(\text{age})$ is the prior distribution of ages and $p(\text{stage}|\text{age}, \text{temp})$ the pdf of stages given age and temperature ([Bernal et al.,](#)

2008). The prior distribution of ages was derived from a mixture of the distribution of spawning times, the sampling time, and the mortality rate. For each station, the distribution of spawning times was expressed in relation to the sampling time (i.e. the probability of spawning happening a hours before the sampling time), and multiplied by the probability of survival, estimated as the mortality rate times the elapsed time between spawning and the observation (i.e. the age):

$$\pi(\text{age}) \equiv f(\text{age}|\tau) \propto f(S = -a|\tau)e^{-ma}, \quad (3)$$

where τ indicates the sampling time, $f(S = -a|\tau)$ evaluates the probability of an egg being spawned a hours before sampling, and e^{-ma} is the probability of it surviving until age a (see Bernal *et al.*, 2001, for an example of the construction of a prior on spawning times).

Two models were used to describe the spawning-time distribution: (i) the spawning-time pdf fitted to the observed spawning times as described above, and (ii) a uniform distribution between 0 and 23.99 h, used to construct an uninformative prior on spawning distribution. The uninformative prior was also used to check whether the raw observed data had any information on the daily spawning cycle, and whether that information was coherent with the information extracted from the distribution of spawning times inferred from the stage I egg observed times. To test the effect of the prior used on the results, a posterior distribution of spawning times was obtained by transforming the posterior distribution of ages into the distribution of back-calculated spawning times (Bernal *et al.*, 2001).

To apply Equations (2) and (3), a mortality rate was required, together with the spawning synchronicity and egg development models. Here, an initial fixed mortality value was assumed for age determination, but an iterative process was implemented to update the appropriate mortality estimate obtained from the mortality curve model (see below).

The estimated peak spawning time from the previous section, i.e. the mode of the distribution chosen, was also used to define the daily cohorts. Each daily cohort included all the eggs released between 12 h before the peak spawning time and 12 h after it. Note that the exact age limits for each cohort will depend on sampling time. For example, for a station sampled at exactly the spawning peak, the age limits of the first cohort would be between 0 and 12 h (because only the eggs released up to 12 h before sampling belong to that cohort), the second cohort from 12 to 36 h old, etc. For a station sampled 12 h after the spawning peak, the first cohort encompasses eggs with ages from 0 to 24 h old, the second between 24 and 48 h old, etc. Once the age limits for each cohort at a given station have been defined, estimates of cohort abundance and mean age can be obtained. First, the posterior distribution of ages for each stage, given station temperature and sampling time, is multiplied by the abundance of eggs in the respective stage, and the weighted posteriors of ages by stage (weights being the abundance of eggs in each stage) are summed. Then, cohort abundance is estimated by integrating the posterior distribution over the limits of each cohort (rounded to the closest integer), and the mean age estimated as the average of the combined posterior of ages within each cohort limit. Finally, once cohort abundance and mean cohort age for all stations has been obtained, the data are used to fit the mortality curve, as explained below.

Mortality estimation

To obtain mortality estimates, Equation (1) was reformulated as a generalized linear model (GLM) following the ideas described in ICES (2004), and extended to accommodate strata and environmental variables. The expected number of eggs of age a , obtained by sampling an effective area (Efarea), could be written as follows:

$$E[N_a] = g^{-1}(\text{offset}(\log(\text{Efarea})) + \log(D_0) - ma). \quad (4)$$

Here $E[N_a]$ is the expected number of eggs in a cohort of mean age a , D_0 the rate of egg production, m the mortality rate, and g^{-1} indicates the inverse of the link function that relates the linear predictor [the right of Equation (4)] and the response, N_a (McCullagh and Nelder, 1989). The offset term contains a variable that is not included in the fitting procedure, which in this case is the logarithm of the effective area, and which accounts for differences in volume or sampling depths between stations (ICES, 2004; Stratoudakis *et al.*, 2006). Equation (4) was fitted to the observed abundance-by-cohort data, assuming a negative binomial distribution, with an unknown dispersion parameter and using the iterative procedure described in Venables and Ripley (2002) and weights proportional to the area represented by each station (ICES, 2004).

Equation (4) was fitted to the joint data from the entire database using the statistical language R (Ihaka and Gentleman, 1996), then extended to allow both egg production and mortality to vary between spatial and temporal strata, and also with temperature, as suggested by Pepin (1991). A general full model of egg production and mortality for the entire dataset is then formulated as

$$\begin{aligned} E[N_a] = & g^{-1}(\text{offset}(\log(\text{Efarea})) + \text{Sstrata} \\ & + \text{Tstrata} + \text{Sstrata}:\text{Tstrata} + \text{Temp} \\ & + \text{Sstrata}:\text{Temp} + \text{Tstrata}:\text{Temp} \\ & + \text{age} + \text{Sstrata}:\text{age} + \text{Tstrata}:\text{age} \\ & + \text{Temp}:\text{age}). \end{aligned} \quad (5)$$

Here both egg production and mortality are allowed to be a function of the strata (both temporal, Tstrata , and spatial, Sstrata) and temperature, as well as their first-order interactions (e.g. $\text{Tstrata}:\text{Sstrata}$). Terms in which “age” is involved indicate mortality terms, and the rest of the terms affect egg production. Stepwise backward model selection was carried out from this model, using the likelihood ratio test. At each step, the term with least significance ($<5\%$) was dropped, and this procedure was repeated until dropping terms led to no improvement. A comparison with Akaike information criterion (AIC) profiles of the model selection procedure was also performed. In addition to the model obtained as explained above, a similar procedure was carried out from a full model without temperature, to check whether there were differences between a pure spatio-temporal model and the final model selected when temperature was also used.

To avoid bias in the mortality model caused by sampling bias at the extremes of the data used for fitting the mortality curve (as described in Stratoudakis *et al.*, 2006), the age range used for estimating the mortality model was restricted. A lower limit was set to exclude cohorts that had not yet been completely spawned, and the upper limit was set to exclude cohorts in which hatching had already started. The exact procedure is as follows. First, for the lower limit, the first cohort from stations whose sampling time

fell within the daily spawning period were excluded from the analysis (as in Lo, 1985). Second, for the upper limit and for each stratum, the age limit was set to that where 5% of the eggs would hatch if the incubation temperature was assumed to be the 95th quantile of observed temperatures within that stratum.

To illustrate the problems that can arise when single surveys are used to estimate mortality, estimates of mortality by year and for each of the spatial strata were also obtained. Those estimates were obtained by a modified version of Equation (5) in which T_{strata} is redefined as year and temperature is excluded from the analysis. The data available for each estimate were therefore similar to those used to report national estimates of egg production and mortality (because the three regional strata used coincide with the limits used for reporting national results to ICES; see ICES, 2004), although the age limits used here, i.e. the limits on the tails of the mortality curve described above, are generally more restrictive than those applied previously (see, for example, ICES, 2004).

Iterative estimation of parameters

To obtain the required estimates for Equations (2) and (3), an iterative procedure was implemented. First an initial estimate of mortality was assumed (i), in this case a mortality value of 0.01 h^{-1} . Then, estimates of cohort abundance and cohort mean age were obtained (ii) using the current estimate of mortality, as explained above. Third, estimates of cohort abundance and mean age were used to update the estimate of mortality (iii), using the full model described in Equation (5), with the age limits as stated above. Finally, steps (i)–(iii) were repeated until the estimate of mortality converged.

To analyse the convergence of the mortality estimates, the absolute differences between original and updated mortality estimates by strata in steps (i) and (iii), respectively, were summed. Convergence criteria were set up to combined differences < 0.001 .

Robustness test

The robustness of the mortality estimates was tested against variations in the age range allowed and against variations in the surveys used to fit the model. First, the minimum and maximum age allowed in the data used to fit the model was modified one at a time over ranges of 0–13 h (for minimum age) and 40–120 h (for maximum age). Also, to test the influence of each survey in the final estimate, estimates of mortality were obtained by subtracting the data for a given survey from the dataset, then re-estimating the mortality without that survey, for each survey, i.e. as in a jackknife procedure with surveys as the sampling unit.

Results

The distribution of spawning events derived from stage I eggs shows a peak between 18:00 and 20:00 UTC, with a lesser probability of spawning from 21:00 to 15:00 (Figure 1a and b). The distribution of spawning times is skewed, with a long tail to the right of the peak. When a normal distribution is fitted to the data (mean = 21.5 h, s.d. = 4 h), the fitted peak spawning time is displaced to the right of the peak observed (Figure 1a). However, when a three-parameter lognormal distribution is fitted to the data (offset = 12 h, equivalent mean = 21 h, equivalent s.d. = 4 h), the model adequately represents both the spawning peak (mode at 18.5 h) and the skewness of the observations (Figure 1b). When comparing the distribution of spawning times inferred from the eggs with that inferred by the few

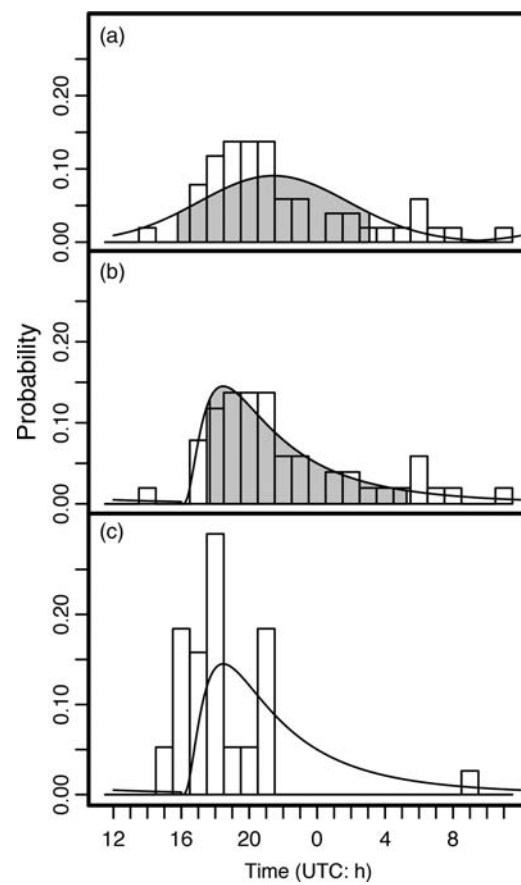


Figure 1. Distribution of spawning times estimated from (a) and (b) observed stage I eggs, and from (c) running females (females with both hydrated oocytes and recent postovulatory follicles in the ovaries). Superimposed curves represent (a) normal (mean = 21.5, s.d. = 4) and (b) and (c) lognormal (offset = 12 h, mean = 1.55, s.d. = 0.81) distributions fitted to the data. Shaded areas in (a) and (b) represent the area between the 10th and 90th percentiles for the normal and lognormal distributions, respectively.

running females observed (Figure 1c), the running females show a spawning time concentrated near the peak, with few observations on the tail. As a compromise to avoid potential bias attributable to the presence of unfertilized eggs in the samples (as described in Motos, 1994), a lognormal distribution, truncated to include 80% of the pdf area (i.e. the area between the 10% and 90% quartiles) was selected as best describing the daily spawning frequency (see grey area in Figure 1b).

The posterior distribution of spawning times obtained using a non-informative prior shows a peak spawning time located between 22:00 and 02:00 UTC (Figure 2a), slightly later than the mode of the lognormal distribution although with greater dispersion and no clear skewness. When the posterior is estimated by cohort, the first cohort shows a clearer daily spawning peak, located between 23:00 and 03:00 UTC, whereas the second cohort shows a flatter pdf, with a peak slightly displaced earlier (Figure 2b). The peak spawning time of the third cohort is not clearly defined. The height of the curves is proportional to the abundance of eggs found in each cohort, following a clearly decreasing trend along successive cohorts. When lognormal spawning-time distribution is used as a prior, both the general

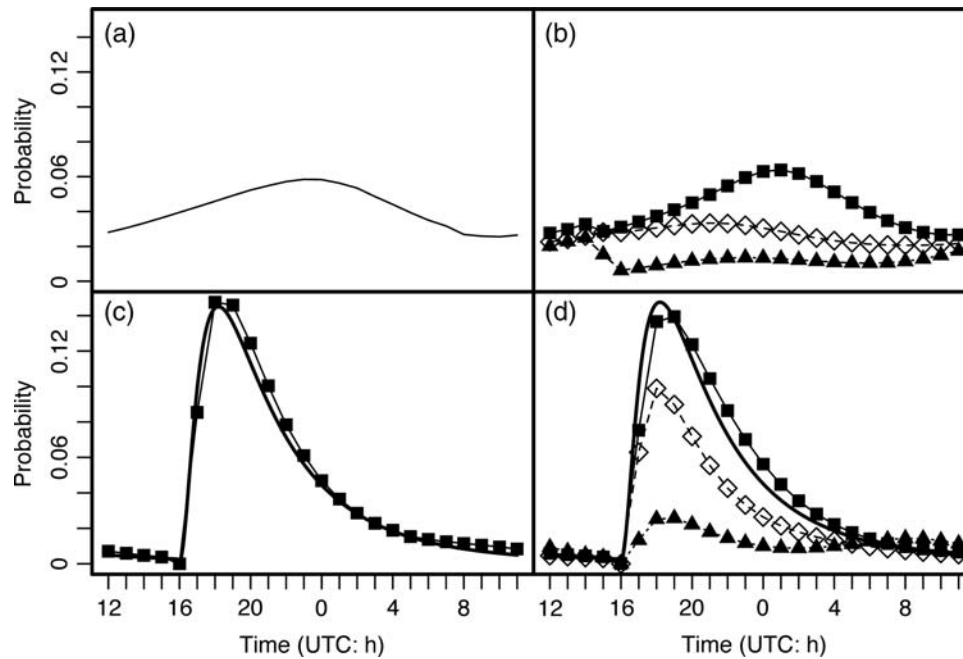


Figure 2. Prior (thick black line) and posterior (line and dots) distribution of spawning times under different model scenarios. (a) Flat prior on ages (not shown), peak spawning of posterior around 23 h. (b) Flat prior and posterior by cohorts (continuous line and filled squares, first cohort; dashed line and open diamonds, second cohort; dashed and dotted line and filled triangles, third cohort), with the heights of the posteriors weighted by abundance. (c) Lognormal prior distribution with equivalent mean of 21 and s.d. of 4 (heavy black line) and posterior distribution after ageing (thin line and filled squares). (d) Prior distribution as in (c) and posterior distribution of the different cohorts (continuous line and filled squares, first cohort; dashed line and open diamonds, second cohort; dashed and dotted line and filled triangles, third cohort).

and the by-cohort posterior distributions are forced to be similar to the prior (Figure 2c and d).

Backward stepwise model selection carried out from the full model shown in Equation (5) discards interactions between age and temporal and spatial strata (Table 1), resulting in a model in which mortality is estimated by a general term and an interaction with temperature (Table 2). When temperature is not used in the initial full model, the final model selected includes spatial strata only (model results not shown; estimated mortality shown in Figure 3). The resulting mortality values for the final model with temperature (for the range of observed temperatures) and for that with spatial strata alone are within the range of mortalities reported by Pepin (1991; see Figure 3). In contrast, when mortality is estimated by year and spatial strata independently, only 8 of 21 mortality estimates are significantly different from zero (results not shown), illustrating the absence of reliability in estimates from a single survey.

Deviance residuals for the selected model do not show clear patterns when plotted against fitted values and temperature, but some zero-inflated data have not been captured by the distribution used (Figure 4).

Figure 5 shows the results of the robustness test carried out to evaluate the effects of changes in the minimum and maximum age allowed in the model, respectively, as well as the influence of individual surveys in mortality estimation. Increasing the minimum age produces an increment in the magnitude of both egg production and mortality (on a negative scale), which is obvious in the first 2 h, when cohorts are probably not fully recruited, i.e. there is ongoing spawning so they are not yet all available to the

Table 1. Model selection from the full model indicated in Equation (5).

Term	d.f.	AIC	LRT	Pr(Chi)
Full model	14	27066		
Sstrata:age	-2	27065	3.14	0.21
Tstrata:age	-1	27065	2.07	0.15

d.f. indicates degrees of freedom of the full model, and thereafter the reduction for each term deletion. AIC is Akaike Information Criterion, LRT the likelihood ratio test, and Pr(Chi) the probability of the null hypothesis (H_0 : new model = full model).

Table 2. Fitted parameters of the final mortality model.

Variable	Estimate	s.e.	z-value	Pr(> z)
Sstrata1	1.173	1.177	0.997	0.319
Sstrata2	7.518	0.963	7.803	0.000
Sstrata3	-0.203	0.889	-0.228	0.820
Tstrata1	3.855	0.928	4.155	0.000
Temp	0.395	0.064	6.170	0.000
Age	0.049	0.017	2.812	0.005
Sstrata2:Tstrata1	-0.406	0.161	-2.521	0.012
Sstrata1:Temp	-0.089	0.081	-1.097	0.273
Sstrata2:Temp	-0.525	0.067	-7.820	0.000
Tstrata1:Temp	-0.303	0.067	-4.561	0.000
Temp:age	-0.005	0.001	-3.926	0.000

The z-value indicates the value of the z-statistics used to test the significance, and Pr(>|z|) the probability of the null hypothesis (H_0 : parameter does not differ from zero).

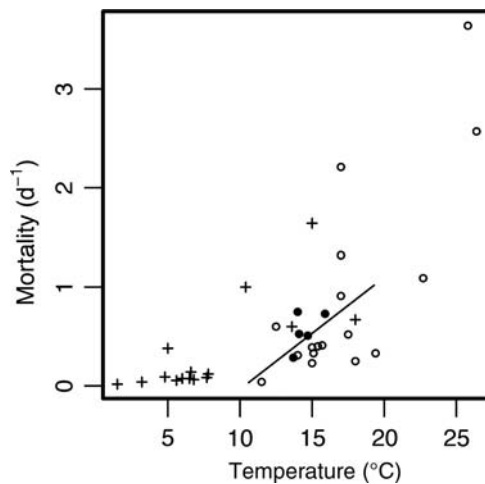


Figure 3. Comparison of the mortality estimates for Atlanto-Iberian sardine with those reported in [Pepin \(1991\)](#). Open circles indicate clupeoid species reported in [Pepin \(1991\)](#), and plus signs represent other species (mainly plaice, mackerel, and sole). Dots indicate the mortality estimates obtained using the model separated by strata, and the straight line is the fitted temperature-dependent mortality model.

net; [Lo, 1985](#)). In relation to the maximum age allowed, both egg production and mortality reveal a generally flat profile, i.e. no effect attributable to a change in the maximum age allowed, when the maximum age varies between 50 and 100 h, although the confidence intervals increase below 65 h maximum age. Similarly, the effect of removing individual surveys seems to have a minor impact, with only a few surveys demonstrating some influence on estimation and whose exclusion seems to increase the estimate of mortality slightly.

Discussion

Sardine spawning shows a daily pattern, with spawning mainly in the late evening, and a spawning peak around 18:00–19:00 UTC. These results are generally consistent with previous Atlanto-Iberian sardine studies based on analysis of the distributions of recently spawned eggs ([ICES, 2002](#)), the distribution of different ovary stages around spawning, e.g. the distribution of hydrated oocytes and early post-ovulatory follicles ([Pérez et al., 1992](#); [Zwolinski et al., 2001](#)), back-estimated spawning times using estimated ages and observed sampling times ([Bernal et al., 2001](#)), and *in situ* observations of spawning school formation ([Zwolinski et al., 2006](#)). Such spawning behaviour is also consistent with findings in other DEPM applications worldwide, in which a daily spawning synchronicity pattern has been found for both sardine and anchovy (e.g. [Lasker, 1985](#); [Ganias et al., 2003](#); [Somarakis et al., 2004, 2006](#)).

The distribution of spawning times extracted from the analysis presented here is, however, wider than usually assumed for Atlanto-Iberian sardine (see [Cunha et al., 1992](#); [García et al., 1992](#); [ICES, 2002](#)) or for sardine elsewhere ([Somarakis et al., 2006](#)). Nevertheless, previous estimates of the daily spawning cycle of sardine are often not very precise (e.g. see [Pérez et al., 1992](#)), or for other areas may even be based entirely on assumptions, as in some of the applications discussed in [Somarakis et al. \(2004\)](#), whereas the estimates obtained here are based on a large and heterogeneous dataset that includes both eggs and

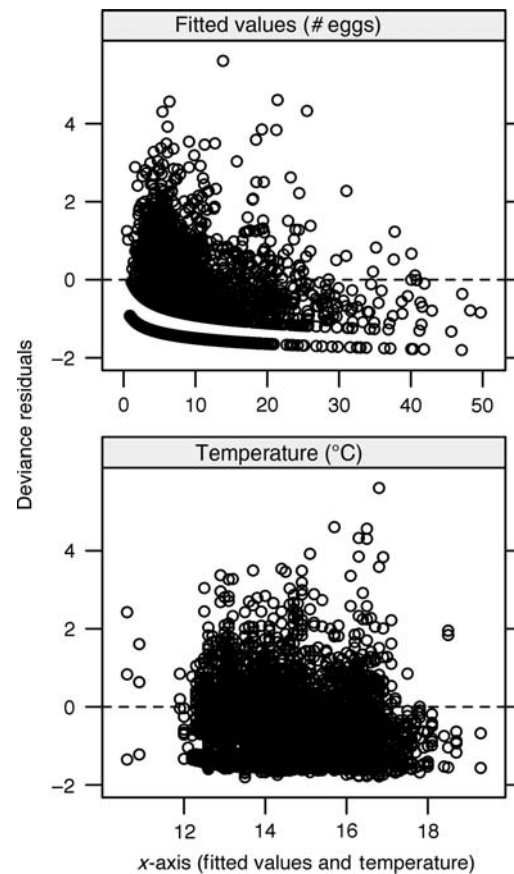


Figure 4. Deviance residuals of the final mortality curve model vs. fitted values (top panel, log scale) and temperature (bottom panel). The horizontal dashed line represents a residual value equal to zero.

running females. Some differences were noted, however, between the distributions of spawning times estimated from stage I eggs alone and from the posterior distribution of spawning times from all stages when no previous information on spawning was available. In the latter case, a late evening peak is still obvious, but it shifted to later in the day, although the information was less precise (i.e. a flatter pdf than that obtained from the distribution of stage I eggs alone). This reflects the use of different cohorts, the aggregation of data from different spatial and temporal strata, and the combined distribution of ages for a given stage ([Bernal et al., 2008](#)). The use of the distribution of spawning times from stage I eggs alone is, however, justified, because it is still coherent with the general picture provided by all stages and with the information provided by the few running females available in the data, and provides an improvement in the precision of both ageing and mortality estimation ([Lo, 1985](#)). Also, the use of a lognormal distribution allows for some probability of spawning happening until late hours of the night, which allows the accommodation of observations of late spawners that may appear in the data without reducing the precision of the overall analysis. As mentioned above, other sources of information can be used to construct a prior on spawning times, such as eggs at other development stages (e.g. stage II eggs), or evidence in the ovary of recent or imminent spawning, but in any of those cases, a detailed knowledge of the relationship between these events and the true spawning time is required.

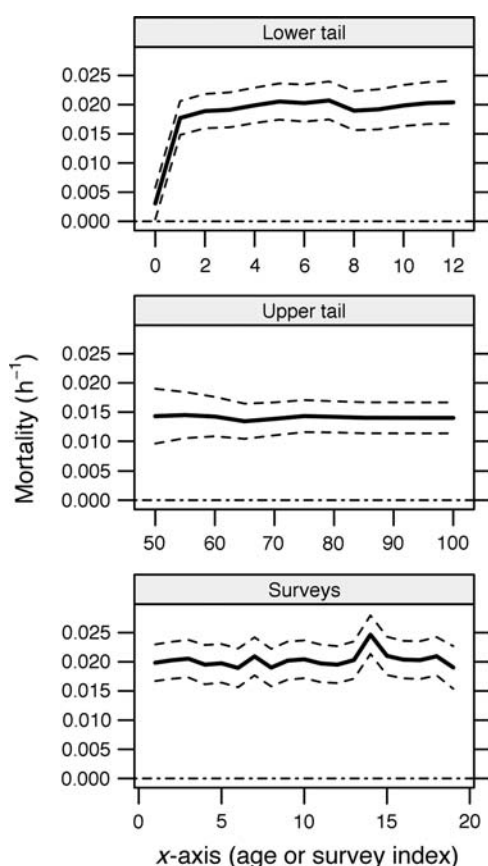


Figure 5. Effects of changing the minimum (top panel) and maximum (middle panel) age used in the mortality-curve fitting procedure, as well as excluding each single survey from the estimation process (bottom panel). The solid black line represents the average mortality (averaged over the temperature values in the survey), and the dashed lines the confidence intervals.

Mortality estimates obtained from analysis of the aggregated Atlanto-Iberian sardine data used in this study are also in accord with previous mortality estimates for the region (e.g. *García et al., 1992*) and also for other species with pelagic eggs (*Pepin, 1991*). The advantage is that, in contrast to previous DEPM applications in the area (*ICES, 2002*) and to previous criticism of the reliability of mortality estimates (*Bulman et al., 1999; Bunn et al., 2000; McGarvey and Kinloch, 2001*), the estimates obtained here are statistically significant, plausible, and robust to the assumptions used in their estimation. Some of the benefits of using aggregated data to estimate mortality were elaborated by *Smith (1973)*, and the use of a general GLM framework to choose how best to aggregate the data provides a sound and reliable statistical procedure that also allows estimation of the precision of the estimates. Robustness of the final estimates to changes in the age limits is greater than that anticipated by *Stratoudakis et al. (2006)*, with changes in the lower and upper age limits having only a minor effect on the final estimates. This is probably because the model used here aggregates all the data, and only uses temperature as a covariate of mortality rates, hence reducing the number of parameters to estimate in relation to any stratification scheme and increasing the available degrees of freedom for the model. A relation between egg mortality and temperature has already been reported by *Pepin (1991)* and others when

aggregating data from different species and ecosystems, but here can be observed by aggregating a large spatial and temporal domain for the same species. Temperature effects on reproductive capacity, egg development, or other physiological rates have been repeatedly reported for marine organisms (e.g. as in *Ottersen et al., 2001*, and references therein). Indirect temperature-driven effects that influence mortality include changes in ecosystem productivity and in foodweb energy transfer (see *Behrenfeld et al., 2006*). The relative influence between direct, physiologically related mortality drivers and indirect, ecosystem-related mortality drivers remains unknown.

Although the analysis provided here results in robust and significant estimates of mortality, a number of shortcomings need to be highlighted. The data used for the analysis do not provide a uniform coverage of the different areas and different periods of the time-series. In the first decade of observations (1985–1995), data were scarce off the northern Iberian Peninsula, and even more scarce off the western coast, whereas no data at all covering the southern nucleus were available. The lack of basic data makes interpretation of spatio-temporal changes in mortality through time in the different areas difficult. Also, interannual variability in mortality can influence the results. However, interannual variability in mortality is somehow represented by the temperature effect, with years of higher temperature yielding high rates of mortality. Also, when the mortality estimates obtained here are used in any application that requires annual rates of mortality (e.g. estimating the daily egg production, see *Bernal et al., 2011*), the interannual variability can be included in the estimation process, by taking into account the residual distribution. Mortality estimates for surveys carried out outside the temporal range of the dataset used here, i.e. in subsequent years, can be estimated by reanalysing new data, or better by using a Bayesian procedure, treating the results found here as a prior and letting the new data update the distribution. Finally, the precision of the estimates obtained here may have been influenced by the wider daily range in spawning times, although no effect on bias is expected, owing to the robustness of the estimates to variation in the lower age limit, which is generally established based on the duration of the spawning cycle. An alternative means of obtaining robust and reliable mortality estimates is to use non-feeding larvae, to increase the amount of data available to fit the mortality curve (as in *Lo et al., 2005*), but such data are not always available, and careful tests on whether the same mortality applies for eggs and larvae for a given species in a specific ecosystem have to be performed first (*Lo et al., 2005*).

Because of their robustness, the estimates of mortality obtained are good candidates for improving the reliability of the time-series of egg production estimates for Atlanto-Iberian sardine. To do that, however, an egg production model that can accommodate mortality estimates external to the estimation procedure is required. This model is developed in *Bernal et al. (2011)*, and makes use of the estimates here to revise the estimates of egg production for Atlanto-Iberian sardine.

Acknowledgements

We acknowledge the work carried out in the various research institutes (IEO, IPIMAR, and AZTI) in providing the data for the analysis carried out here, as well as the crew of the research vessels used to collect the data. Early surveys used in this paper were funded by each institute and national research funding bodies, but most of the recent surveys through the EU Data

Collection Regulation. Data collection was carried out under the auspices of the ICES study and working groups SGSBSA and WGACEGG, and the EU project SARDYN. The manuscript benefited from discussion generated in EC Project 99/080 (Using environmental variables with improved DEPM methods to consolidate the series of sardine and anchovy estimates), within the Research Unit for Wildlife Population Assessment (RUWPA) at St Andrews University, and within SGSBSA and WGACEGG, and we thank the participants of all those groups. Pierre Pepin and two anonymous referees provided us valued comments on earlier versions, though the work presented is part of a doctoral dissertation by the first author, who was partially funded by an EU Marie Curie research grant. Finally, MB acknowledges the Instituto Español de Oceanografía for allowing him time to dedicate himself fully to finalizing his doctorate and the manuscripts produced therein.

References

- Armstrong, M. J., Connolly, P., Nash, R. D. M., Pawson, M. G., Alesworth, E., Coulahan, P. J., Dickey-Collas, M., *et al.* 2001. An application of the annual egg production method to estimate the spawning biomass of cod (*Gadus morhua* L.), plaice (*Pleuronectes platessa* L.) and sole (*Solea solea* L.) in the Irish Sea. *ICES Journal of Marine Science*, 58: 183–203.
- Barange, M., Bernal, M., Cercole, M. C., Cubillos, L., Cunningham, C. L., Daskalov, G. M., De Oliveira, J. A. A., *et al.* 2009. Current trends in the assessment and management of small pelagic fish stocks. In *Climate Change and Small Pelagic Fish*, pp. 191–255. Ed. by D. M. Checkley, C. Roy, J. Alheit, and Y. Oozeki. Cambridge University Press, Cambridge, UK. xx + 372 pp.
- Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., McClain, C. R., Sarmiento, J. L., Feldman, G. C., Milligan, A. J., *et al.* 2006. Climate-driven trends in contemporary ocean productivity. *Nature*, 444: 752–755.
- Bernal, M., Borchers, D. L., Valdés, L., Lago de Lanzós, A., and Buckland, S. T. 2001. A new ageing method for eggs of fish species with daily spawning synchronicity. *Canadian Journal of Fisheries and Aquatic Sciences*, 58: 2330–2340.
- Bernal, M., Ibaibarriaga, L., Lago de Lanzós, A., Lonergan, M., Hernández, C., Franco, C., Rasines, I., *et al.* 2008. Using multinomial models to analyse data from sardine egg incubation experiments; a review of advances in fish egg incubation analysis techniques. *ICES Journal of Marine Science*, 65: 51–59.
- Bernal, M., Stratoudakis, Y., Coombs, S., Angelico, M., Lago de Lanzós, A., Porteiro, C., Sagarmínaga, Y., *et al.* 2007. Sardine spawning off the European Atlantic coast: characterization of and spatio-temporal variability in spawning habitat. *Progress in Oceanography*, 74: 210–227.
- Bernal, M., Stratoudakis, Y., Wood, S., Ibaibarriaga, L., Valdés, L., and Borchers, D. 2011. A revision of daily egg production estimation methods, with application to Atlanto-Iberian sardine. 2. Spatially and environmentally explicit estimates of egg production. *ICES Journal of Marine Science*, 68: 528–536.
- Bulman, C. M., Koslow, J. A., and Haskard, K. A. 1999. Estimation of the spawning stock biomass of blue grenadier (*Macruronus novaezelandiae*) off western Tasmania based upon the annual egg production method. *Marine and Freshwater Research*, 50: 197–207.
- Bunn, N. A., Fox, C. J., and Webb, T. 2000. A literature review of studies of fish egg mortality: implications for the estimation of spawning stock biomass by the annual egg production method. Technical Report Cefas Lowestoft, 111. 37 pp.
- Cunha, M. E., Figueiredo, I., Farinha, A., and Santos, M. 1992. Estimation of sardine spawning biomass off Portugal, by the daily egg production method. *Boletín del Instituto Español de Oceanografía*, 8: 139–153.
- Ganias, K., Somarakis, S., Machias, A., and Theodorou, A. J. 2003. Evaluation of spawning frequency in a Mediterranean sardine population (*Sardina pilchardus*). *Marine Biology*, 142: 1169–1179.
- García, A., Pérez, N., Lo, N. C. H., Lago de Lanzós, A., and Solá, A. 1992. The egg production method applied to the spawning biomass estimation of sardine, *Sardina pilchardus* (Walb.), on the North Atlantic Spanish coast. *Boletín del Instituto Español de Oceanografía*, 8: 123–138.
- Hampton, I. 1996. Acoustic and egg-production estimates of South African anchovy biomass over a decade: comparisons, accuracy, and utility. *ICES Journal of Marine Science*, 53: 493–500.
- Ibaibarriaga, L., Bernal, M., Motos, L., Uriarte, A., Borchers, D. L., Lonergan, M., and Wood, S. 2007. Estimation of development properties of stage-classified biological processes using multinomial models: a case study of Bay of Biscay anchovy (*Engraulis encrasicolus* L.) egg development. *Canadian Journal of Fisheries and Aquatic Sciences*, 64: 539–553.
- ICES. 2002. Report of the Study Group on the Estimation of Spawning Stock Biomass of Sardine and Anchovy. ICES Document CM 2002/G: 01. 59 pp.
- ICES. 2004. The DEPM estimation of spawning-stock biomass for sardine and anchovy. ICES Cooperative Research Report, 268. 91 pp.
- ICES. 2007. Report of the Working Group on the Assessment of Mackerel, Horse Mackerel, Sardine and Anchovy. ICES Document CM 2007/ACFM: 31. 725 pp.
- ICES. 2009. Report of the Working Group on Acoustic and Egg Surveys for Sardine and Anchovy in ICES Areas VIII and IX (WGACEGG). ICES Document CM 2009/LRC: 20. 177 pp.
- Ihaka, R., and Gentleman, R. 1996. R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5: 299–314.
- Jackson, G., and Cheng, Y. W. 2001. Parameter estimation with egg production surveys to estimate snapper, *Pagrus auratus*, biomass in Shark Bay, Western Australia. *Journal of Agricultural, Biological, and Environmental Statistics*, 6: 243–257.
- Lasker, R. (Ed). 1985. An egg production method for estimating spawning biomass of pelagic fish: application to the northern anchovy (*Engraulis mordax*). NOAA Technical Report, NMFS 36. 99 pp.
- Lo, N. C. H. 1985. A model for temperature-dependent northern anchovy egg development and an automated procedure for the assignment of age to staged eggs. NOAA Technical Report, NMFS 36: 43–50.
- Lo, N. C. H., Macewicz, B. J., and Griffith, D. A. 2005. Spawning biomass of Pacific sardine (*Sardinops sagax*), from 1994–2004 off California. *CalCOFI Report*, 46. 93 pp.
- Lotka, A. J. 1925. *Elements of Physical Biology*. Williams and Wilkins, Baltimore, MD. 495 pp.
- McCullagh, P., and Nelder, J. A. 1989. *Generalized Linear Models*. Chapman and Hall, London. 532 pp.
- McGarvey, R., and Kinloch, M. A. 2001. An analysis of the sensitivity of stock biomass estimates derived from the daily egg production method (DEPM) to uncertainty in egg mortality rates. *Fisheries Research*, 49: 303–307.
- Miranda, A., Cal, R. M., and Iglesias, J. 1990. Effect of temperature on the development of eggs and larvae of sardine *Sardina pilchardus* Walbaum in captivity. *Journal of Experimental Marine Biology and Ecology*, 140: 69–77.
- Motos, L. 1994. Estimación de la biomasa desovante de la población de anchoa del Golfo de Vizcaya, *Engraulis encrasicolus*, a partir de su producción de huevos. Bases metodológicas y aplicación. PhD thesis, Universidad del País Vasco. 240 pp.
- Ottersen, G., Planque, B., Belgrano, A., Post, E., Reid, P. C., and Stenseth, N. Ch. 2001. Ecological effects of the North Atlantic oscillation. *Oecologia*, 128: 1–14.

- Parker, K. 1980. A direct method for estimating northern anchovy, *Engraulis mordax*, spawning biomass. Fishery Bulletin US, 78: 541–544.
- Pepin, P. 1991. Effect of temperature and size on development, mortality, and survival rates of the pelagic early life history stages of marine fish. Canadian Journal of Fisheries and Aquatic Sciences, 48: 503–518.
- Pérez, N., Figueiredo, I., and Macewicz, B. 1992. The spawning frequency of pilchard, *Sardina pilchardus* (Walb.), on the Atlantic Iberian coast. Boletín Instituto Español de Oceanografía, 8: 175–189.
- Simmonds, E. J., and MacLennan, D. N. 2005. Fisheries Acoustics: Theory and Practice, 2nd edn. Wiley-Blackwell, Oxford. 437 pp.
- Smith, P. E. 1973. The mortality and dispersal of sardine eggs and larvae. Rapports et Procès-verbaux des Réunions du Conseil International pour l'Exploration de la Mer, 164: 282–292.
- Somarakis, S., Ganas, K., Siapatis, A., Koutsikopoulos, C., Machias, A., and Papaconstantinou, C. 2006. Spawning habitat and daily egg production of sardine (*Sardina pilchardus*) in the eastern Mediterranean. Fisheries Oceanography, 15: 281–292.
- Somarakis, S., Palomera, I., Garcia, A., Quintanilla, L., Koutsikopoulos, C., Uriarte, A., and Motos, L. 2004. Daily egg production of anchovy in European waters. ICES Journal of Marine Science, 61: 944–958.
- Stratoudakis, Y., Bernal, M., Ganas, K., and Uriarte, A. 2006. The daily egg production methods: recent advances, current applications and future challenges. Fish and Fisheries, 7: 35–57.
- Venables, W. N., and Ripley, B. D. 2002. Modern Applied Statistics with S, 4th edn. Statistics and Computing, Springer, Dordrecht. 495 pp.
- Verhulst, P. F. 1839. Notice sur la loi que la population suit dans son accroissement. Correspondence in Mathematics and Physics, 10: 113–121.
- Zwolinski, J., Mason, E., Oliveira, P. B., and Stratoudakis, Y. 2006. Fine-scale distribution of sardine (*Sardina pilchardus*) eggs and adults during a spawning event. Journal of Sea Research, 56: 294–304.
- Zwolinski, J., Stratoudakis, Y., and Sares, E. 2001. Intra-annual variation in the batch fecundity of sardine off Portugal. Journal of Fish Biology, 58: 1633–1645.