Re-analysis of mean daily egg production in jack mackerel

J. Lyle, K. Hartmann, C. Buxton & C. Gardner
Institute for Marine and Antarctic Studies

Background

Given the level of interest in the jack mackerel (*Trachurus declivis*) daily egg production method (DEPM) assessment reported by Neira (2011) and, in particular, focus on the estimation of mean daily egg production (P_0), IMAS have undertaken a re-analysis of the reported egg density data using a range of alternative model fitting methods suggested in the literature. Specifically, we have followed Ward *et al*. (2011) who compared several methods and their application in Australian sardine (*Sardinops sagax*) assessments.

Methods

Neira (2011) fitted an exponential function to the daily egg abundance-at-age data using two methods, namely the least squares non-linear regression (NLS) model (Seber and Wild, 1989; Lo *et al*., 2005) and a generalized linear model (GLM) using a negative binomial error distribution (Cubillos *et al*., 2007; Neira and Lyle, 2011). In the absence of a temperature egg development model for jack mackerel, Neira (2011) used models for horse mackerel (*T. trachurus*) from the north-east Atlantic (Cunha *et al*., 2008) and a deterministic stage-to-age model obtained for Chilean mackerel (*T. murphyi*) (Sepulveda *et al*. 2009) to estimate egg age.

In addition to repeating the original analysis using the methods applied by Neira we have also run four GLMs and the log-linear method (with bias correction) as described by Ward *et al*. (2011). For this report we have used the same notation provided by Ward *et al*. (2011) to describe the alternative GLMs. In summary, GLM 1 assumes a Gaussian distribution with a log link function; GLM 2 assumes a Quasi distribution with a log link function and variance proportional to the mean; GLM 3 assumes a Quasi distribution with log link function and variance proportional to the mean squared; and GLM 4 assumes a Quasi distribution with identity link function and variance proportional to the mean but is applied to ln-transformed egg density data. As estimates of P_0 for this last method (GLM 4) are negatively biased, a bias correction factor equivalent to that applied to the log-linear model is applied.

In undertaking the analyses we have followed Neira (2011) by excluding 'extreme cohorts' but have also re-analysed the data using all available eggs. Extreme cohorts refer to eggs with assigned ages of <4 h old and those with >98% probability of being hatched at the mean station temperature. Exclusion of these extreme cohorts follows DEPM protocols associated with addressing biases caused by under-sampling of newly spawned eggs near peak spawning, and/or the reduced probability of encountering small, highly concentrated patches of newly spawned eggs, as well as the high probability of under-sampling eggs close to hatching age.

Results

Mean daily egg production values based on the various models and ageing methods for the partial data set (extreme cohorts excluded) and full data set (extreme cohorts included) are presented in Tables 1 and 2, respectively. In relation to repeatability, we have been able to replicate Neira's P_0 estimates for the NLS and GLM (negative binomial error distribution) methods precisely (Table 1).

Neira's GLM (negative binomial error distribution) estimates compare favourably with those for GLMs 1-3 (ranging between 81-105% of the values for these models) but are substantially greater (by factors of between 4.7 and 7.7 times) than estimates based on GLM 4 ~~and the log-linear model~~ (Table 1). Similarly, the NLS method as applied by Neira resulted in substantially higher P_0's than those produced by the log-linear model (5.5 to 5.9 times greater). If all other DEPM input parameters are not varied, then the implication of any variation in P_0 is linear, i.e. an increase or decrease in P_0 by a factor of two will have the effect of the resultant biomass estimate being doubled or halved.

Table 1: Mean daily egg production values based on alternative model fitting methods (refer Neira 2011 and Ward *et al*., 2011) and alternative temperature-dependent egg incubation models using 2002 jack mackerel egg data from southern New South Wales (**extreme cohorts excluded**). P_0 adjusted is calculated by multiplying P_0 by spawning area/DEPM survey area (ie 21,327/23,934 km^2).

	Ageing method			
	Sepulveda *et al.* 2009 (*T. murphyi*)		**Cunha *et al.* 2008 (*T. trachurus*)**	
Model fitting method	**P_0**	**P_0 adjusted**	**P_0**	**P_0 adjusted**
NLS*	5.54	4.93	3.77	3.36
GLM (negative binomial error)*	4.40	3.92	4.27	3.80
GLM1	5.45	4.86	3.78	3.37
GLM2	4.49	4.00	3.82	3.40
GLM3	4.32	3.85	4.05	3.61
GLM4	0.70	0.62	0.55	0.49
Log-linear (with bias correction)	0.94	0.84	0.69	0.61

*Methods reported by Neira 2011

The implications of including all egg data in the analyses can be evaluated by comparing results in Table 1 with corresponding results in Table 2. The inclusion of extreme cohorts had mixed outcomes on P_0 estimates, with higher values for the GLM (negative binomial error distribution), GLM 3 and GLM 4 (based on Cunha *et al*~~.),~~ .), while the remainder had lower P_0 estimates (61-96% of corresponding values when extreme cohorts are excluded). The GLM (negative binomial error distribution) appeared to be the most sensitive of all methods to the inclusion or exclusion of the extreme cohorts.

Table 2: Weighted egg production values based on alternative model fitting methods (refer Neira 2011 and Ward *et al*., 2011) and alternative temperature-dependent egg incubation models using 2002 jack mackerel egg data from southern New South Wales (**extreme cohorts included**). P_0 adjusted is calculated by multiplying P_0 by spawning area/DEPM survey area (ie 21,327/23,934 km^2).

	Ageing method			
	Sepulveda *et al.* 2009 (*T. murphyi*)		**Cunha *et al.* 2008 (*T. trachurus*)**	
Model fitting method	**P_0**	**P_0 adjusted**	**P_0**	**P_0 adjusted**
NLS	3.38	3.01	3.17	2.82
GLM (negative binomial error)	13.33	11.88	6.01	5.35
GLM1	3.41	3.04	3.20	2.85
GLM2	3.98	3.55	3.60	3.21
GLM3	6.19	5.52	4.64	4.13
GLM4	0.43	0.38	0.68	0.61
Log-linear (with bias correction)	0.72	0.64	0.66	0.59

Discussion and recommendation

This analysis has highlighted how the choice of model to describe the egg mortality relationship can significantly influence the estimation of P_0. The advantages and disadvantages of each model approach will depend on the nature of the input data, noting that egg density by age data tend to be over-dispersed and model fits are invariably poor regardless of whether an exponential model (non-linear least squares regression), log-linear model or GLMs are used (e.g. Ward *et al*. 2011). This is certainly the case for jack mackerel; estimates of P_0 reported by Neira (2011) had coefficients of variation ranging between 0.59 and 0.91 indicating that the reported estimates were imprecise (a point emphasised by Neira, 2011). Furthermore, it should be noted that the probability distributions for P_0 are typically positively skewed (results not shown here) with greater agreement on the lower bound of P_0 than on the best estimate. That is, there is greater agreement between methods about the lowest plausible biomass than there is about the best estimate.

Ward *et al.* (2011) concluded that the log-linear model should be used for Australian sardine because it fitted their egg density data better and resulted in more precautionary (lower) estimates of P_0 than either the NLS and GLM approaches (although their GLM 4 did result in lower estimates of P_0 than the log-linear model). Ward *et al.* (2011) have had the benefit of a decade of DEPM assessments and understand the Australian sardine stocks well. By contrast, research into each of the key Small Pelagic Fishery (SPF) species (blue mackerel, jack mackerel and redbait) is in its infancy and issues such as the preferred statistical approaches for estimating each of the DEPM input parameters require further consideration.

In the case of the SPF, the SPF Harvest Strategy dictates that the recommended biological catch (RBC) should not exceed a prescribed percentage of the best available spawning biomass estimate (the percentage values determined by the age of the estimate). A shortcoming of the current harvest strategy is that it does not explicitly consider the effect of differing uncertainties between studies and analytical methods.

As scrutiny of any future DEPM assessments in the SPF is likely to be intense, we consider that there would be merit in the SPF Resource Assessment Group (SPFRAG) undertaking a review of existing data on the SPF species along with available DEPM literature to develop recommendations on the preferred method(s) to be used to estimate each of the DEPM input parameters. Development of an *a-priori* preferred approach would be of value in assisting the SPFRAG, South East Management Advisory Committee and Australian Fisheries Management Authority in the application of future DEPM assessments when recommending RBCs (and total allowable catches) and would reduce the risk of different parties 'cherry picking' estimates to suit particular agendas.

References

Cubillos, L.A., Ruiz, P., Claramunt, G., Gacitua, S., Nunez, S., Castro, L.R., Riquelme, K., Alarcon, C., Oyarzun, C. and Sepulveda, A. (2007). Spawning, daily egg production, and spawning stock biomass estimation for common sardine (*Strangomera bentincki*) and anchovy (*Engraulis ringens*) off central southern Chile in 2002. *Fisheries Research* **86**: 228-240.

Cunha, M. E., Vendrell, C. and Goncalves, P. (2008). Experimental study of the dependence of embryonic development of *Trachurus trachurus* eggs on temperature. *ICES Journal of Marine Science* **65**: 17-24.

Lo, N.C.H., Macewicz B.J. and Griffith D.A. (2005). Spawning biomass of Pacific sardine (*Sardinops sargax*), from 1994-2004 off California. CalCOFI Report 46, 93-112.

Neira, F.J. (2011) Application of daily egg production to estimate biomass of jack mackerel, *Trachurus declivis* – a key fish species in the pelagic ecosystem of south-eastern Australia. IMAS Report, 42p.

Neira, F.J. and Lyle, J.M. (2011). DEPM-based spawning biomass of *Emmelichthys nitidus* (Emmelichthyidae) to underpin a developing mid-water trawl fishery in south-eastern Australia. *Fisheries Research* **110**, 236-243.

G.A.F Seber and C.J. Wild. (1989) Nonlinear Regression. New York: John Wiley and Sons.

Sepulveda, A., Ruiz, P., Alarcón Muñoz, R., Núñez, S., Bahamonde, R., Córdova, J., Chong, J., Rebolledo, H., Aguayo, M. and Bustos, L. (2009). Condición biológica del jurel en alta mar, año 2006. Informe Final FIP 2006-05. Instituto de Investigación Pesquera, Talcahuano, Chile. 291 pp. (http://www.fip.cl/FIP/Archivos/pdf/informes/inffinal%202006-05.pdf).

Ward,T.M, Burch,P., McLeay, L.J., and Ivey, A.R. (2011): Use of the Daily Egg Production Method for Stock Assessment of Sardine, *Sardinops sagax*; Lessons Learned over a Decade of Application off Southern Australia, *Reviews in Fisheries Science*, **19**, 1-20.

Disclaimer