4th Meeting of the Scientific Committee

The Hague, Kingdom of the Netherlands
10 - 15 October 2016

SC-04-JM-06

Assessing the Parsimony in the Jack mackerel stock assessment model (JJM)

Cristian Canales

Assessing the parsimony in the jack mackerel stock assessment model (JJM)

Cristian Canales
cristian.canales.r@pucv.cl
Chile

1. Background

In science, it is often accepted that when two theories on an equal basis have the same consequences, the simplest theory is more likely to be correct between the two. This concept is called principle of parsimony or “Ockham’s razor”. In statistics, this also refers to the advantages of a model to explain the variability of a set of data through the combination of predictors as simple as possible.

Several assumptions have been considered in the jack mackerel stock assessment to explain the data. One of the most important is the annual variability of the age-specific selectivity pattern per fleet. Indeed, at the first SC meeting held in La Jolla, San Diego, USA in 2013, it was decided to include annual variability in the selectivity patterns of the fleets to mitigate, in part, the high level of correlation observed in the age and size composition residuals of catches/surveys. On that basis, the number of the model parameters increased nearly 3 times without further analysis on its implications regarding the statistical support. A similar reduction result in the correlations in the pattern of residuals can be achieved reducing the sample sizes (Francis, 2012).

This document includes an analysis of comparative assessment considering the model used until 2015 and its performance based on the reduction of the parameter number through the use of the selectivity blocks, i.e. maintaining constant the selectivity of each one of the fleets by year blocks.

2. Materials and methods

Configuration of annual selectivity of the jack mackerel assessment model was compared in respect of a reduced model, in which selectivity patterns are modeled by year blocks. Identification of the first years of these blocks was conducted by inspection regarding the magnitude of the annual residuals of the average age/size of the catches.

The Akaike Information Criterion (AIC) (Akaike, 1974) and the Schwarz-Bayesian Information Criterion (Schwartz, 1978) were used as a comparison statistic among models. These indicators have demonstrated to be consistent with meeting the fundamental criteria for the selection of models: goodness of fit, parsimony, and objectivity (Langitoto et al., 2000).

$$AIC = 2\mathrm{L}_{max} + 2p$$
$$BIC = 2\mathrm{L}_{max} + p\log(n)$$

where L_{max} is the maximum of the total log-likelihood function negative of the data (penalties or a priori distributions are not considered), p represents the number of parameters and n the number of observations.

3. Results

Identification of the years that define the selectivity blocks was conducted by inspection in respect of the model fit (a reduced model version) to the average age/size of the catches per fleet. A total of 6 selectivity blocks were identified for the fishery of the North of Chile (F1) (1970-1986; 1987-1988; 1989-1992; 1993-2012; 2013-2014 and 2015), and 4 blocks in the fishery of the center-south area of Chile (F2) (1970-1983; 1984-1996; 1997-2000; 2001-2015). On the other hand, 7 blocks defined by 1970-1981; 1982-1997; 1998-2001; 2002-2004; 2005-2006; 2007-2009; 2010-2015 were considered for Peru's fishery (F3). For the international fleet operating off Chile (F4), 1970-1979; 1980-2001; 2002-2005; 2006-2013 y 2014-2015 were considered.

Selectivities by annual blocks regarding the model with annual variability of selectivity show, for instance, that the general pattern of change is reproduced properly (**Figure 1, 2**), while in the current model, annual changes would not be justified such as, those of the period 1993-2013 in the northern area of Chile (F1) given their low inter-annual variability. Likewise and while spawning biomass between models does not show significant differences, the growing trend of jack mackerel in the reduced model is more pronounced (**Figure 3**).

Regarding the parsimony level and considering the performance only of the likelihood of the model data, that is, excluding the value of a priori distributions and penalties, the results show that, statistically, the reduced model is more efficient since the AIC (Akaike, 1974) and BIC (Schwarz, 1978) criteria were significantly lower in respect of the base model (**Table 1**). It is important to note that both metrics meet the selection criterion of the models: goodness of fit, parsimony, and objectivity (Langitoto et al., 2000).

The previous result indicates that the best fit of the base model (lowest log-likelihood) to the data is mainly reached by the higher number of parameters (1,387 vs 484) and not necessarily by a model describing the exploitation process in a better manner.

4. Discussion

The better performance of the reduced model highlights from the results obtained. This situation tends to reduce the parameter correlation produced in large-scale models. In this sense and although there were no significant differences in terms of the performance variables such as virgin biomass and population scale, it is concluded that the current assessment model would not justify the use of variable selectivity by year but to prefer the use of temporal blocks, reducing significantly the number of parameters (parsimony) without a large loss in the fit goodness of the model.

Additionally, the implicit effect of correlation in the residuals of the fit of age/size compositions of the catches (Francis, 2011) could be addressed by the reduction in the effective sample size.

5. References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. AC-19(6), 716-723.

Francis, C. 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68: 1124–1138 (2011). doi:10.1139/F2011-025.

Langitoto, S., D.B. Sampson, D.B., Yin, Y., 2000. Application of statistical model selection criteria to the Stock Synthesis assessment program. Can. J. Fish. Aquat. Sci. 57: 1784-1793.

Schwarz, G. E., 1978. Estimating the dimension of a model». Annals of Statistics 6 (2): 461–464. doi:10.1214/aos/1176344136.

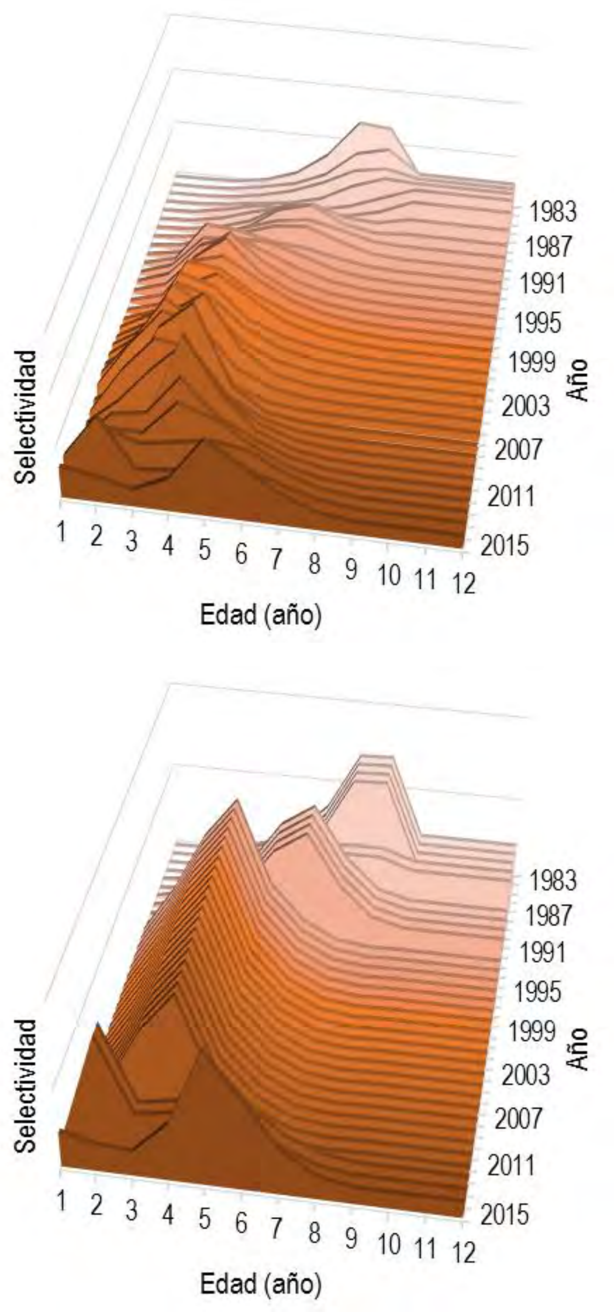

Figure 1. Selectivity for northern Chilean fleet (F1) and for two model configuration. Upper panel is the base model and panel below is the reduced model.

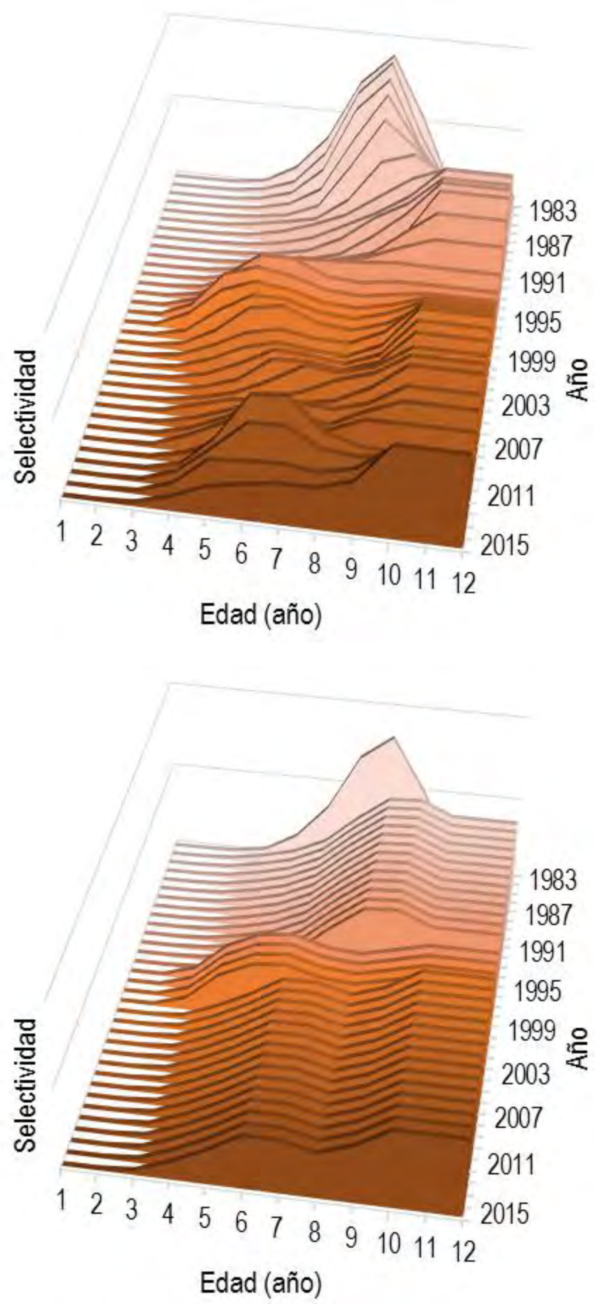

Figure 2. Selectivity for southern Chilean fleet (F2) and for two model configuration. Upper panel is the base model and panel below is the reduced model.

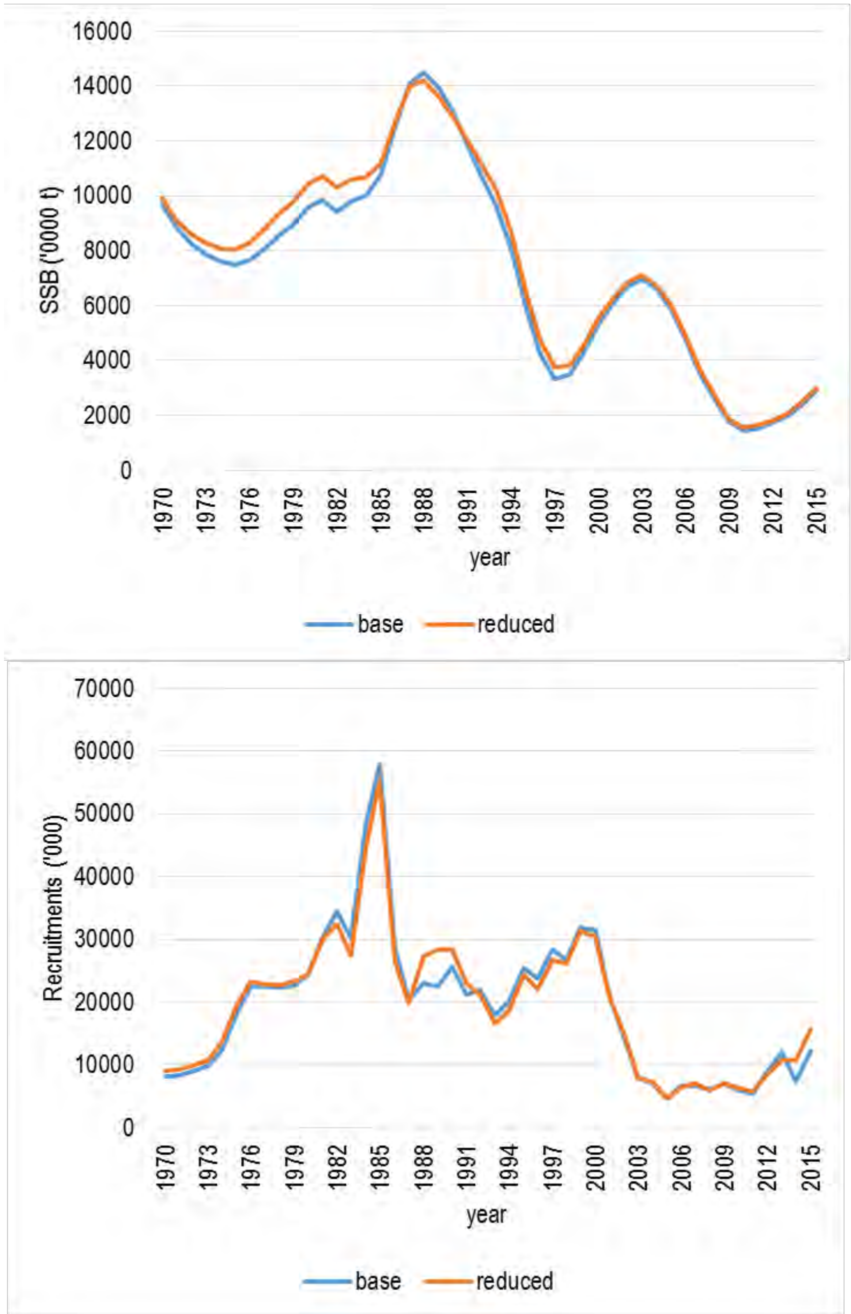

Figure 3. Spawning biomass (SSB) (upper panel) and recruitments (below) of Jack mackerel for two model configurations.

Table 1.

Marginal log likelihood (LL) components and main performance variables of model JJM 2015 for two selectivity configurations. The base model considers annual variability in selectivities, while the reduced model considers selectivity by blocks. SSB is the spawning biomass

	Base model	Reduced model
LL total (data)	1,433	1,729
LL Catch	1	3
LL Age comps in catch	488	600
LL length comps in catch	435	511
LL Abundance indices	288	388
LL Age comps in surveys	221	227
N. parameters	1,387	484
AIC	5,639	4,427
BIC (*)	14,151	7,397
SSB_0 (000' t)	19,470	19,627
F/Fmsy	0.69	0.59
SSB/SSBmsy	0.49	0.53
MSY (000' t)	1,586	1,591

(*) base on 3,418 observations