Stock Reduction Analysis using catch at length data

Catarina Wor, Roberto Licandeo, Brett van Poorten, Carl Walters

March 8, 2017

Abstract

Last thing to be written

1 Introduction

Modern stock assessments typically attempt to fit population dynamics models to catch at age and catch at length data, in hopes of extracting information from these data about age/size vulnerability and fishing mortality patterns (Methot and Wetzel, 2013; Hilborn and Walters, 1992). In cases where age data are lacking, models like MULTIFAN-CL attempt to obtain estimates of vulnerability and fishing mortality only from size distribution data (Fournier et al., 1998). Combined with a few assumptions regarding the structure and variability in length at age, this procedure can even be used to attempt to recover information about changes in body growth patterns if there is a strong age-class signal in the length frequency data (Fournier et al., 1998). Some assessment methods attempt to put aside the length frequency data, by converting these data to age compositions using age-from-length tables, perhaps using iterative methods to estimate proportions of fish at age for each length interval (Kimura and Chikuni, 1987). It is typical for assessment results from length-based assessment models to show substantial deviations between predicted and observed length distributions of catches, reflecting both sampling variation in the length composition data and incorrect assumptions about stability of growth and vulnerability patterns (reference).

The vulnerability process is the combination of two processes: selectivity of the fishing gear and availability of the fished population in the area being fished (Beverton and Holt, 1957). Both processes can vary over time and therefore modify the resulting selectivity. Although selectivity process can often be directly measured through gear experiments, availability is generally harder to measure as it depends on the distribution of the exploited population. Changes in availability occur when changes in the spatial distribution of the fleet in relation to the fished population happen. Availability changes can be caused by fish movement, changes in fish distribution, and by changes in fleet distribution. Changes in vulnerability are not uncommon (Sampson and Scott, 2012) but are usually difficult to track overtime. This difficulty is associated with the inability to distinguish between changes in fishing mortality and changes in vulnerability in most age and length based stock assessment methods. For this reason, many assessment methods rely on ad hoc parametric vulnerability models that may or may not include changes over time (Maunder et al., 2014). If misspecified, such models might lead to severe bias in fishing mortality estimates, which could result in misleading management advice (Martell and Stewart, 2014).

Here we suggest an alternative approach to assessment modeling that begins by assuming that the assessment model should exactly reproduce the observed catch at length distribution. This is similar to the classical assumption in virtual population analysis that reconstructed numbers at age should exactly match observed catch at age data (Hilborn and Walters, 1992). This assumption is also analogous to the suggestion by Schnute (1994) that statistical catch at age models might best be run in a“conditioned on catch” format by subtracting observed catches at age from modeled numbers at age in estimation of numbers at age over time. The suggested approach may have two key advantages over statistical catch at age and/or catch at length models: (1) it does not require estimation of age or size vulnerability schedules, and (2) catch at length data are commonly available for every year, even when age composition sampling has not been conducted.

We demonstrate the performance of this model with a simulation-evaluation analysis and apply it to real fisheries data from the Chilean jack mackerel and Pacific Hake fisheries.

2 Methods

In this section we describe the stock reduction analysis with catch at length data (Length-SRA), describe the simulation analysis and scenarios used to test the model and provide a description of the real data used to illustrate the model applicability.

2.1 Stock reduction analysis with catch at length data

The stock reduction analysis described here starts by calculating , the proportions of individual at length for each age class (Table 1. The calculation of such proportions relies on three main assumptions regarding the distribution of length at age: (1) The mean length at age follows a von Bertalanffy growth curve (eq.T1.5), (2)

Table 1: Age at Length

Variable definition
$P_{l\|a}$ = Matrix of proportions of lengh at age
$z1_{a,l}$ = Normalized Z score for lower limit length bins
$z2_{a,l}$ = Normalized Z score for upper limit length bins
$b1_l$ = Lower limit of length bins
$b2_l$ = upper limit of length bins
$\bar{L}_a$ = Mean length at age
σ_{L_a} = Standard deviation of lengh at age
L_{inf} = Maximum average length
k = rate of approach to L_{inf}
t_o = Theoretical time in which length of individuals is zero
cvL = Coefficient of variation for length curve (T1.1)

$$P_{l|a} = \int_{z1_{a,l}}^{z2_{a,l}} \mathcal{N}(0,1) \tag{T1.2}$$

$$z1_{a,l} = \frac{b1_l - \bar{L}_a}{\sigma_{L_a}} \tag{T1.3}$$

$$z2_{a,l} = \frac{b2_l - \bar{L}_a}{\sigma_{L_a}} \tag{T1.4}$$

$$\bar{L}_a = L_{inf} \cdot (1 - \exp^{(-k\cdot(a-t_o))}) \tag{T1.5}$$

$$\sigma_{L_a} = \bar{L}_a \cdot cvL \tag{T1.6}$$

The length at age is normally distributed (eqs. T1.2 -T1.4) and (3) The standard deviations of the length at age distributions is given by the product of the mean length at age and a constant CV (eq.T1.6).

The proportions of length at age is used to convert the length based quantities into age based quantities which are used to propagate the age structured population dynamics forward (Table 2). We assume that recruitment follows a Beverton & Holt type recruitment curve (eq. T2.1), that harvesting occurs over a short, discrete season in each time step (year or shorter time period) and that natural survival rate is stable over time (eqs. T2.1-T2.5). Differences in the population dynamics equations in the initial year as well as incidence functions are shown in Table 2.

Table 2: Population dynamics

Variable definition
$N_{a,t}$ = Numbers of fish at age and time
SB_t = Spawning biomass at time t
a_{rec}, b_{rec} = Beverton & Holt stock recruitment parameters
κ = Goodyear recruitment compensation ratio
wt = Normally distributed recruitment deviations
S_a = Survival rate at age
$U_{a,t}$ = Exploitation rate at age and time
$U_{l,t}$ = Exploitation rate at length and time
$C_{l,t}$ = Catch at length and time
$N_{l,t}$ = Numbers at length and time
syr = Initial year of data
a_o = Age of recruitment
ϕ_e = Unfished average spawning biomass per recruit

$$N_{a,t>syr} = \begin{cases} \frac{a_{rec} \cdot SB_{t-1}}{1+b_{rec} \cdot SB_{t-1}} \cdot e^{wt}, & a = a_o \\ N_{a-1,t-1} \cdot S_{a-1} \cdot (1 - U_{a-1,t-1}), & 1 < a < A \\ \frac{N_{a-1,t-1} \cdot S_{a-1} \cdot (1-U_{a-1,t-1})}{1-S_A \cdot 1 - Ua,t}, & a = A \end{cases} \tag{T2.1}$$

$$U_{a,t} = \sum_a (P_{l|a} \cdot U_{l,t}) \tag{T2.2}$$

$$U_{l,t} = \frac{C_{l,t}}{N_{l,t}} \tag{T2.3}$$

$$N_{l,t} = \sum_a (P_{l|a} \cdot N_{a,t}) \tag{T2.4}$$

$$SB_t = \sum_a (fec_a \cdot w_a \cdot N_{a,t}) \tag{T2.5}$$

Initial year

$$N_{a_o,syr} = R_{init} * \cdot e^{wt} \tag{T2.6}$$

$$a_{rec} = \frac{\kappa}{\phi_e} \tag{T2.7}$$

$$b_{rec} = \frac{\kappa - 1}{R_o \cdot \phi_e} \tag{T2.8}$$

$$\phi_e = \sum_a lx_a \tag{T2.9}$$

$$lx_a = \begin{cases} 1, & a = a_o \\ lx_{a-1} \cdot S_{a-1}, & 1 < a < A \\ \frac{lx_{a-1} \cdot S_{a-1}}{1-S_A}, & a = A \end{cases} \tag{T2.10}$$

Table 3: Operating model dynamics

Variable definition
$U_{l,t}$ = Exploitation rate at length and time
$sel_{l,t}$ = fishing selectivity at length and time
U_t = annual maximum exploitation rate
$C_{l,t}$ = Catch at length and time
$N_{l,t}$ = Numbers at length and time
I_t = Index of abundance at time t
VB_t = Biomass that is vulnerable to the survey at time t
q = catchability coefficient

$$U_{l,t} = U_t \cdot sel_{l,t} \tag{T3.1}$$

$$C_{l,t} = N_{l,t} \cdot U_{l,t} \cdot P_{l|a} \tag{T3.2}$$

$$sel_{l,t} = \frac{1}{1-g} \cdot \left(\frac{1-g}{g}\right)^{g} \cdot \frac{e^{a \cdot g \cdot (b-l)}}{1+e^{a \cdot (b-l)}} \tag{T3.3}$$

$$I_t = q \cdot VB_t \tag{T3.4}$$

The model estimates three main parameters: the average unexploited recruitment R_0, the recruitment compensation ratio κ and the recruitment in the initial year R_{init}. In addition, the recruitment deviations are estimated for all cohorts observed in the model, that is, the number of recruitment deviations is equal to the number of years in the time series plus the number of age classes greater that recruitment age. The parameters of the model are estimated with two likelihood components: Index of abundance and Recruitment deviations, both are assumed to be lognormally distributed with fixed variances.

In order to perform a simulation evaluation of the Length-SRA under various scenarios we used the same model dynamics described in Table 2 as an operating model. However we modified the model population dynamics to include time varying selectivity and a maximum annual exploitation rate (eq. T3.1) as well as observation and process errors. Selectivity in the operating model was computed with the three parameter selectivity equation described by Thompson (1994) (eq. T3.3). The observation error in the operating model included lognormal error in the index of abundance and logistic multivariate error in the catch numbers at length. Recruitment deviations were assumed to be lognormally distributed.

2.2 Simulation evaluation scenarios

We considered a total of six different scenarios in the simulation evaluation runs. Three different historical exploitation rate trajectories were used: contrast, one way trip and U ramp. In the contrast scenario the exploitation rate(U) starts low and increases up to $U > U_{msy}$ and then decreases until $U = U_{msy}$. In the one way trip scenario U increased through time until $U = 2 \cdot U_{msy}$. In the U ramp scenario, U increases steadily until $U = U_{msy}$ and remains constant thereafter. In addition to the exploitation rate scenarios, we considered two selectivity scenarios: constant and time varying selectivity. In the constant selectivity scenario, selectivity was assumed to follow a sigmoid shape. In the time varying selectivity scenario, the selectivity curve was assumed to vary every year, progressively changing form a dome shaped curve to sigmoid and back to dome shaped. The complete list of scenarios and the code used for them is presented in Table 4.

All simulation runs had 30 years of data and we used 200 simulation runs for each scenario. We evaluated the distribution of the % relative error ($\frac{esimated-simulated}{simulated} \cdot 100$) for the main parameter estimates (R_0, R_{init} and κ) and for four derived quantities (Depletion: $\frac{B_t}{B_0}$, MSY, U_{MSY} and q).

Table 4: Simulation-estimation scenarios

Scenario Code	Selectivity	U trajectory
CC	constant	contrast
CO	constant	one way trip
CR	constant	U ramp
VC	time-varying	contrast
VO	time-varying	one way trip
VR	time-varying	U ramp

2.3 Real data examples

Two species were chosen to illustrate the application of the Length-SRA to real datasets: Chilean jack mackerel and Pacific hake. Both species are believed to be subject to time varying selectivity. In the case of Pacific hake, time varying selectivity is believed to be associated with cohort targeting and fleet spatial distribution. The population is know to have spasmodic recruitment, with high recruitment events occurring once or twice every decade (Ressler et al., 2008). Pacific hake tends to segregate by size during their annual migration(Ressler et al., 2008), allowing the fishing fleet to target the strong cohorts by changing the spatial distribution of fishing effort as the cohort ages.

In the case of Jack mackerel (...) Have to come up with good justification for JM.

3 Results

3.1 Simulation-evaluation

The simulation-evaluation showed that the parameters R_0 and R_{init} tend to be underestimated but with very low bias. The median % errors for those parameters were within the ± 10% interval, with the exception of the U ramp scenarios (CR and VR) in which the median % error for R_o and R_{init} was as high as 25% (Figure 1). The parameter κ was also underestimated with higher median % error. Once again, the U ramp scenarios (CR and VR) resulted in the highest bias (Figure 1).

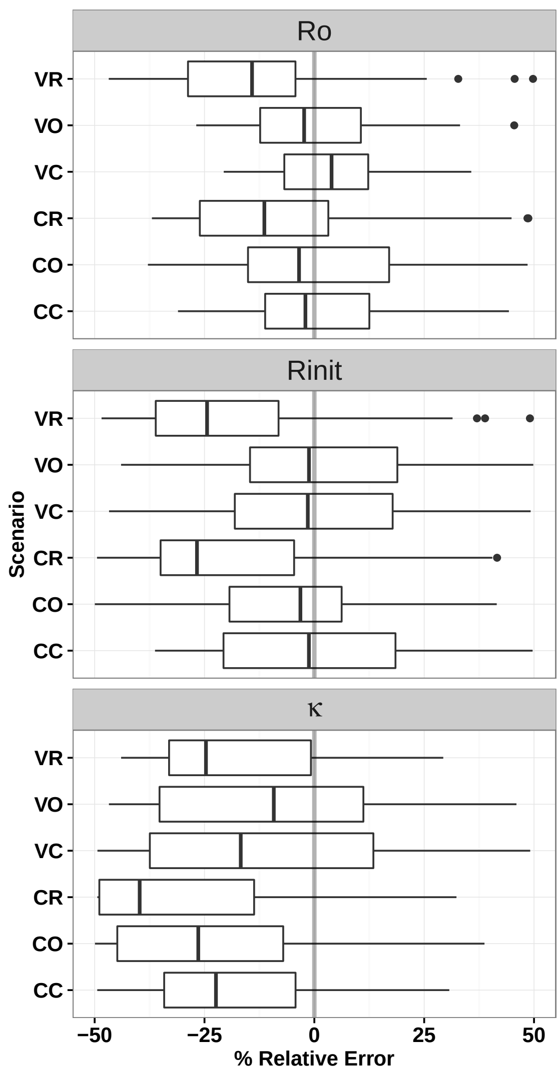

Figure 1: Relative error (%) for main parameters for all scenarios considered in the simulation-evaluation.

In relation to the derived quantities, the Length-SRA tended to underestimate both depletion and MSY estimates with median % errors ranging between -26% and -2% . Once again the U ramp scenarios (CR and VR) yielded the worst bias (Figure 2). The estimates for U_{MSY} showed very low (<5%) median % errors for all scenarios except for the CC scenario (Contrast and constant selectivity) which showed a 13% median % error (Figure 2). The estimates of q tended to be underestimated for the U ramp scenarios an d overestimated for the remaining scenarios, with median % error ranging between -3% and 10% (Figure 2).

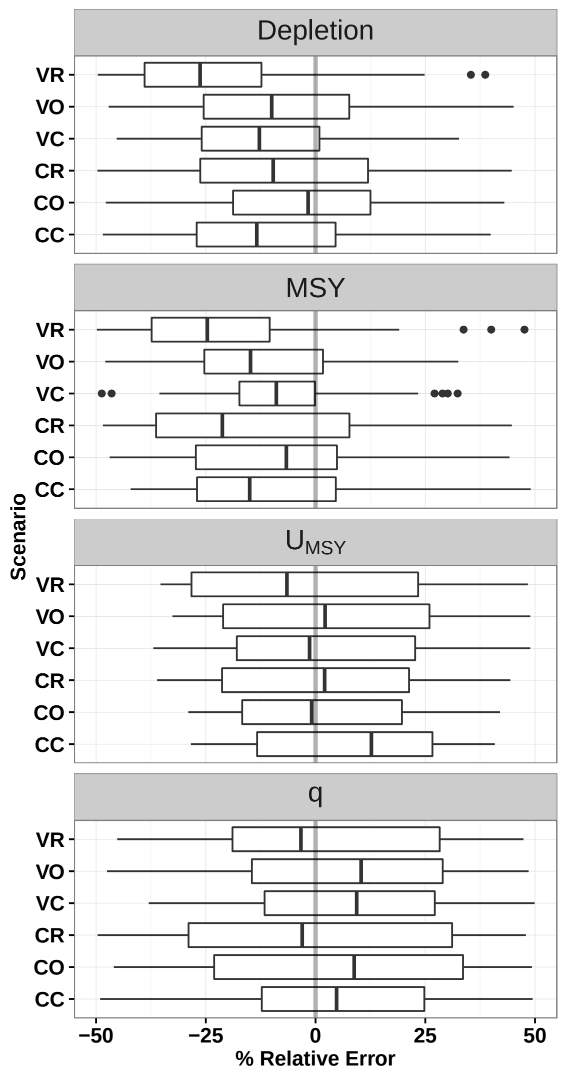

Figure 2: Relative error (%) for main parameters for all scenarios considered in the simulation-evaluation.

3.2 Real data examples

show figures with real

4 Discussion

Does the model work?

assumption regarding Umsy -¿ changes with selectivity

Time- varying growth might render the model less useful. Suggest estimating cohort-specific growth curves or implementing the density dependent functions shown in multifan-CL.

5 Acknowledgments

We would like to thank Allan Hicks and NOAA for the provision of Pacific hake data and ????? for the provision of Jack mackerel data.

References

Beverton, R. J. H. and Holt, S. J. (1957). *On the Dynamics of Exploited Fish Populations*, volume 19 of *Investment series*. U.K. Ministry of Agriculture and Fisheries, London. Google-Books-ID: BqbnCAAAQBAJ.

Fournier, D. A., Hampton, J., and Sibert, J. R. (1998). MULTIFAN-CL: A length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, Thunnus alalunga. *Canadian Journal of Fisheries and Aquatic Sciences*, 55(9):2105–2116.

Hilborn, R. and Walters, C. J. (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty/Book and Disk*. Springer Science & Business Media.

Kimura, D. K. and Chikuni, S. (1987). Mixtures of Empirical Distributions: An Iterative Application of the Age- Length Key. *Biometrics*, 43(1):23–35.

Martell, S. and Stewart, I. (2014). Towards defining good practices for modeling time-varying selectivity. *Fisheries Research*, 158:84–95.

Maunder, M. N., Crone, P. R., Valero, J. L., and Semmens, B. X. (2014). Selectivity: Theory, estimation, and application in fishery stock assessment models. *Fisheries Research*, 158:1–4.

Methot, R. D. and Wetzel, C. R. (2013). Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*, 142:86–99.

Ressler, P. H., Holmes, J. A., Fleischer, G. W., Thomas, R. E., and Cooke, K. C. (2007(2008) 2008). Pacific hake, Merluccius productus, autecology: A timely review. *U S National Marine Fisheries Service Marine Fisheries Review*, 69(1-4). ZOOREC:ZOOR14601004724.

Sampson, D. B. and Scott, R. D. (2012). An exploration of the shapes and stability of population–selection curves. *Fish and Fisheries*, 13(1):89–104.

Schnute, J. T. (1994). A General Framework for Developing Sequential Fisheries Models. *Canadian Journal of Fisheries and Aquatic Sciences*, 51(8):1676–1688.

Thompson, G. G. (1994). Confounding of gear selectivity and the natural mortality rate in cases where the former is a nonmonotone function of age. *Canadian Journal of Fisheries and Aquatic Sciences*, 51(12):2654–2664.