Stock Reduction Analysis using catch at length data

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

May 17, 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 (Hilborn and Walters, 1992).

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. Fish movement, changes in fish distribution, and by changes in fleet distribution, can all cause availability changes.Changes in vulnerability are not uncommon (Sampson and Scott, 2012) but are usually difficult to track overtime. This difficulty is associated with an 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 approach follows the dynamics of an age structred stock reduction analysis (SRA) (Walters et al., 2006) which follows a "conditioned on catch" format, subtracting observed catches at age from modeled numbers at age in estimation of numbers at age over time. This assumption is analogous 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). 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 Peruvian jack mackerel (*Trachurus murphyi*) and Pacific Hake (*Merluccius productus*) 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 given each age class ($P_{l|a}$ - eqs. T2.1-T2.5). The calculation of such proportions (eq. T2.1) relies on four main assumptions regarding the distribution of length at age: (1) The mean length at age follows a von Bertalanffy growth curve (eq.T2.4), (2) The length at age is normally distributed (eqs. T2.1 -T2.3), (3) The standard deviation of the length at age distributions is defined (e.g. eq.T2.5) and (4) $P_{l|a}$ is constant for all lengths equal or greater than L (eq.T2.3).

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 - Population dynamics). We assume that recruitment follows a Beverton & Holt type recruitment curve (eq. T2.6), 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. T2.6-T2.10). Differences in the computation of recruitment in the initial year (eq. T2.11) as well as incidence functions (eqs. T2.12-T2.15) are shown in Table 2 - Initial year and incidence functions.

The model estimates three main parameters: average unexploited recruitment R_0, recruitment compensation ratio κ and recruitment in the initial year R_{init}. In addition, the recruitment deviations w_t 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 objective function (eq. T3.8) is composed of a negative log likelihood component, two penalties and a prior component for the recruitment compensation ratio κ. The negative log likelihood component minimizes the differences between the predicted and observed index of abundance (eq. T3.1). We assume that such differences are lognormally distributed (eqs. T3.3-T3.4) and use the conditional maximum likelihood estimator described by Walters and Ludwig (1994) to estimate the catchability coefficient(eq. T3.2). A lognormal penalty is added to the negative log-likelihood function to constrain annual recruitment residuals so that the estimates have mean of zero and fixed standard deviation σ_R (eq. T3.5). An additional penalty was implemented in order to prevent estimated values of exploitation rate at length greater than one ($U_{l,t} > 1$) (eq. T3.6) . Lastly, a normal prior for $log(\kappa)$ was included in the objective function (eq. T3.7).

2.2 Simulation evaluation

In order to perform a simulation evaluation of the Length-SRA under various scenarios we used the same model dynamics described in Table 2 - Population dynamics, as an operating model. However we modified the model population dynamics to control annual exploitation rate (eq. T2.17), time varying selectivity (eq. T2.19), and observation and process errors. Selectivity in the operating model was computed with the three parameter selectivity equation described by Thompson (1994) (eq. T2.19). 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.

Table 1: Indexes, variable definition and values used in simulation-evaluation

Symbol	Value	Description
l	$\{l_o,...,L\}$	Central point of length bin, $L = 15$
a	$\{a_o,...,A\}$	Age-class, $A = 10$
t	$\{1,...,T\}$	Annual time step, $T = 50$
a_o	21	First age or age of recruitment
l_o	1	Central point of first length bin
$init$	21	Annual time step in which data starts to be reported
Distribution of length given age		
L_{inf}	10	Maximum average length
k	0.3	Rate of approach to L_{inf}
t_o	-0.1	Theoretical time in which length of individuals is zero
cv_l	0.08	Coefficient of variation for length at age curve
$P_{l\|a}$		Matrix of proportions of length at age
Φ		Standard normal distribution
$z1_{a,l}$		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 length at age
Population dynamics		
R_o	100	Average unfished recruitment
κ	10	Goodyear recruitment compensation ratio
S	0.7	Natural survival
σ_{rec}	0.4	standard deviation for recruitment deviations
w_t	$\mathcal{N}(0,\sigma_R)$	Recruitment deviations for years {init-A+1,...,T}
$N_{a,t}$		Numbers of fish at age and time
SB_t		Spawning biomass at time
mat_a		Proportion of mature individuals at age
a_{rec}, b_{rec}		Beverton & Holt stock recruitment parameters
R_{init}		Recruitment in the year data starts to be reported
$U_{a,init}$		Exploitation rate at age before data starts being reported.
$U_{a,t}$		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
lx_a		Unfished survivorship at age
ϕ_e		Unfished average spawning biomass per recruit
Operating model		
$sel_{l,t}$		Fishing selectivity at length and time
g,a,b		Parameters for selectivity function
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
VB_t		Biomass that is vulnerable to the survey at time
q	1.0	catchability coefficient

Table 2: population dynamics for Length-SRA and operating model

	Distribution of length given age
(T2.1)	$P_{l\|a} = \int_{z1_{a,l}}^{z2_{a,l}} \Phi(z) dz$
(T2.2)	$z1_{a,l} = \frac{b1_l - \bar{L}_a}{\sigma_{L_a}}$
(T2.3)	$z2_{a,l} = \begin{cases} \frac{b2_l - \bar{L}_a}{\sigma_{L_a}} & l < L \\ 1.0 & l = L \end{cases}$
(T2.4)	$\bar{L}_a = L_{inf} \cdot (1 - \exp^{(-k \cdot (a - t_o))})$
(T2.5)	$\sigma_{L_a} = \bar{L}_a \cdot cv_l$
	Population dynamics
(T2.6)	$N_{a,t>init} = \begin{cases} \frac{a_{rec} \cdot SB_{t-1}}{1 + b_{rec} \cdot SB_{t-1}} \cdot e^{w_t}, & a = a_o \\ N_{a-1,t-1} \cdot S \cdot (1 - U_{a-1,t-1}), & a_o < a < A \\ \frac{N_{a-1,t-1} \cdot S \cdot (1 - U_{a-1,t-1})}{1 - S \cdot (1 - Ua,t)}, & a = A \end{cases}$
(T2.7)	$U_{a,t} = \sum_a (P_{l\|a} \cdot U_{l,t})$
(T2.8)	$U_{l,t} = \frac{C_{l,t}}{N_{l,t}}$
(T2.9)	$N_{l,t} = \sum_a (P_{l\|a} \cdot N_{a,t})$
(T2.10)	$SB_t = \sum_a (mat_a \cdot w_a \cdot N_{a,t})$
	Initial year and incidence functions
(T2.11)	$N_{a,init} = \begin{cases} R_{init} \cdot e^{w_{init}} & a = a_o \\ N_{a-1,init} \cdot S \cdot (1 - U_{a-1,init}) \cdot e^{w_{init-a+1}}, & a_o < a < A \\ \frac{N_{a-1,init} \cdot S \cdot (1 - U_{a-1,init})}{1 - S \cdot (1 - Ua,init)} \cdot e^{w_{init-a+1}}, & a = A \end{cases}$
(T2.12)	$a_{rec} = \frac{\kappa}{\phi_e}$
(T2.13)	$b_{rec} = \frac{\kappa - 1}{R_o \cdot \phi_e}$
(T2.14)	$\phi_e = \sum_a lx_a$
(T2.15)	$lx_a = \begin{cases} 1, & a = 1 \\ lx_{a-1} \cdot S, & 1 < a < A \\ \frac{lx_{a-1} \cdot S}{1 - S}, & a = A \end{cases}$
	Operating model
(T2.16)	$N_{a,t=1} = lx_a * R_o$
(T2.17)	$U_{l,t} = U_t \cdot sel_{l,t}$
(T2.18)	$C_{l,t} = N_{l,t} \cdot U_{l,t} \cdot P_{l\|a}$
(T2.19)	$sel_{l,t} = \frac{1}{1-g} \cdot \left(\frac{1-g}{g}\right)^g \cdot \frac{e^{a \cdot g \cdot (b-l)}}{1 + e^{a \cdot (b-l)}}$
(T2.20)	$I_t = q \cdot VB_t \cdot e^{(\mathcal{N}(0, \sigma_{I_t})}$

Table 3: Likelihood functions and penalties

Likelihood	
(T3.1)	$Z_t = log(I_t) - log(VB_t)$
(T3.2)	$q = e^{\bar{Z}}$
(T3.3)	$Zstat = Z_t - \bar{Z}$
(T3.4)	$\mathscr{L}_{I_t} = -logL(Zstat \sim \mathcal{N}(0, \sigma_{I_t}))$
Penalties	
(T3.5)	$pen(w_t) = -logL(w_t \sim \mathcal{N}(0, \sigma_R))$
(T3.6)	$pen(U) = \frac{\sum^t \sum^l U_{l,t}}{T}$
Priors	
(T3.7)	$prior(log(\kappa)) = \mathcal{N}(log(10), 0.9)$
Objective function	
(T3.8)	$Obj = \mathscr{L}_{I_t} + pen(w_t) + pen(U) + prior(log(\kappa))$

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_t) starts low and increases beyond U_{msy} and then decreases until $U_t = U_{msy}$. In the one way trip scenario U increased through time until $U = 2 \cdot U_{msy}$. In the U ramp scenario, U_t increases steadily until $U_t = U_{msy}$ 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 acronym used for them is presented in Table 4.

All simulation runs had 30 years of data and we used 100 simulation runs for each scenario. We evaluated the distribution of the % relative error ($\frac{esimated-simulated}{simulated}$) 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.

The Pacific hake fishery is believed to exhibit time varying selectivity due to cohort targeting and annual changes 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.

The movement pattern of jack mackerel is not as well known, although fish appear to move between spawning and feeding areas (Gerlotto et al., 2012). Variability in selectivity patterns for the jack mackerel fishery are believed to be associated both with evolution of fleet capacity and gear utilization and with compression and expansion of the species range associated with abundance changes (Gerlotto et al., 2012).

3 Results

3.1 Simulation-evaluation

Even though the parameter estimates were not entirely unbiased, the amount of bias was relatively low, with median % error being always less than 25% for all parameters. The parameters R_0 and R_{init} were overestimated for all scenarios with the exception of the U ramp scenarios (CR and VR) (Figure 1). The median % errors for R_0 ranged from -9.1% to 15.4%. For R_{init} the median % errors varied from -22.5% to 13.3%. The parameter κ was overestimated for all scenarios (Figure 1) with median % errors for κ ranged from 11.3 % to 18.9%

In relation to the derived quantities, the Length-SRA tended to underestimate both depletion and MSY estimates with median % errors ranging between -34.9% and -3.5% . For both depletion and MSY quantities, the U ramp scenarios (CR and VR) yielded the highest bias (Figure 2). The estimates for U_{MSY} showed very low (<4%) median % errors for all scenarios (Figure 2). The estimates of q were overestimated for all scenarios but were particularly biased for the the U ramp scenarios (CR and VR) (Figure 2).

The simulation-evaluation exercise showed that the Length-SRA model is able to track selectivity changes over time relatively well. There is a tendency to underestimate selectivity for lower lengths and overestimate it for higher lengths across all scenarios, However, this pattern is particularly prominent for the U ramp scenarios (Figure 3).

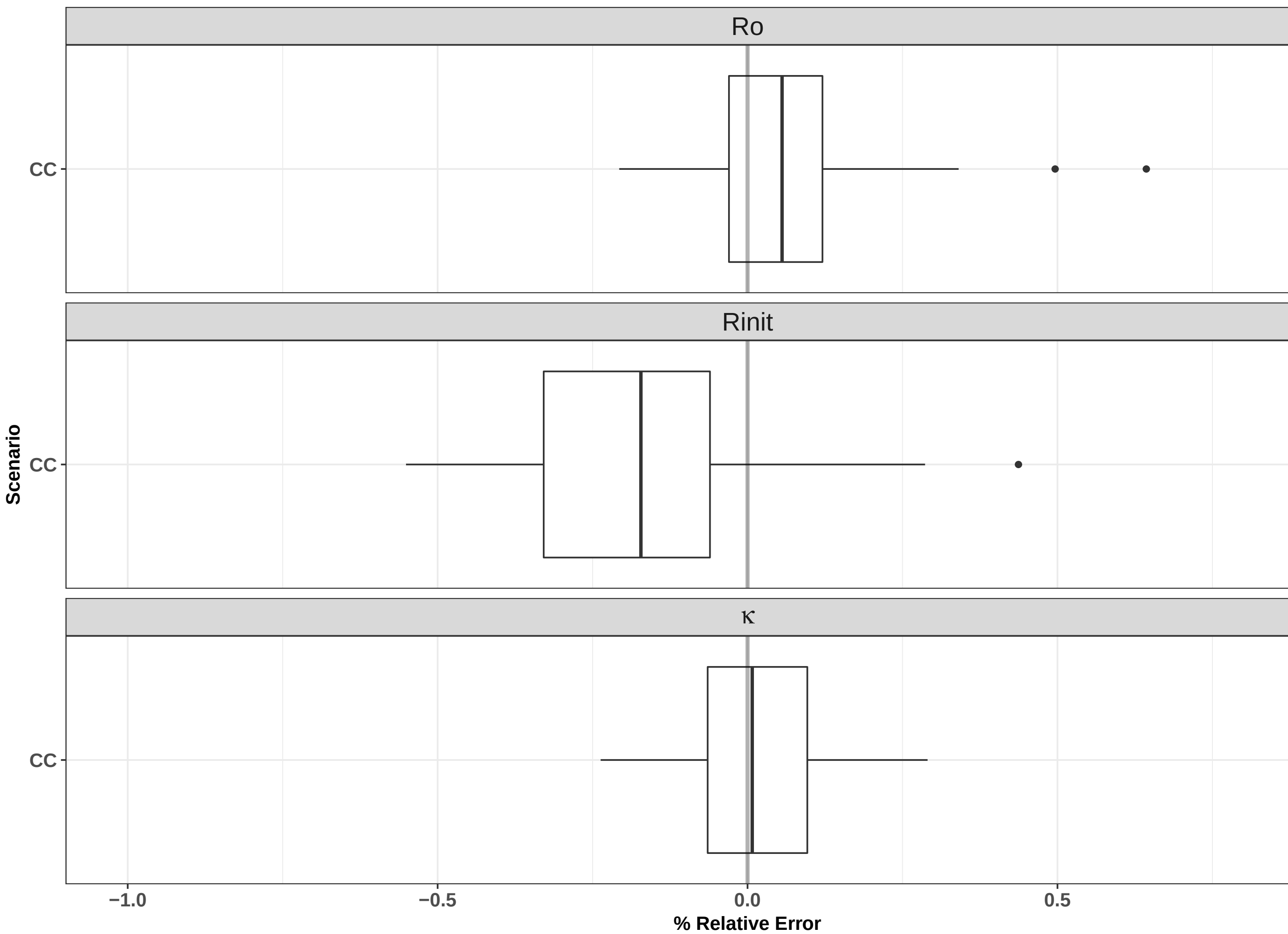

Figure 1: Relative error (%) for main parameters for all scenarios considered in the simulation-evaluation.

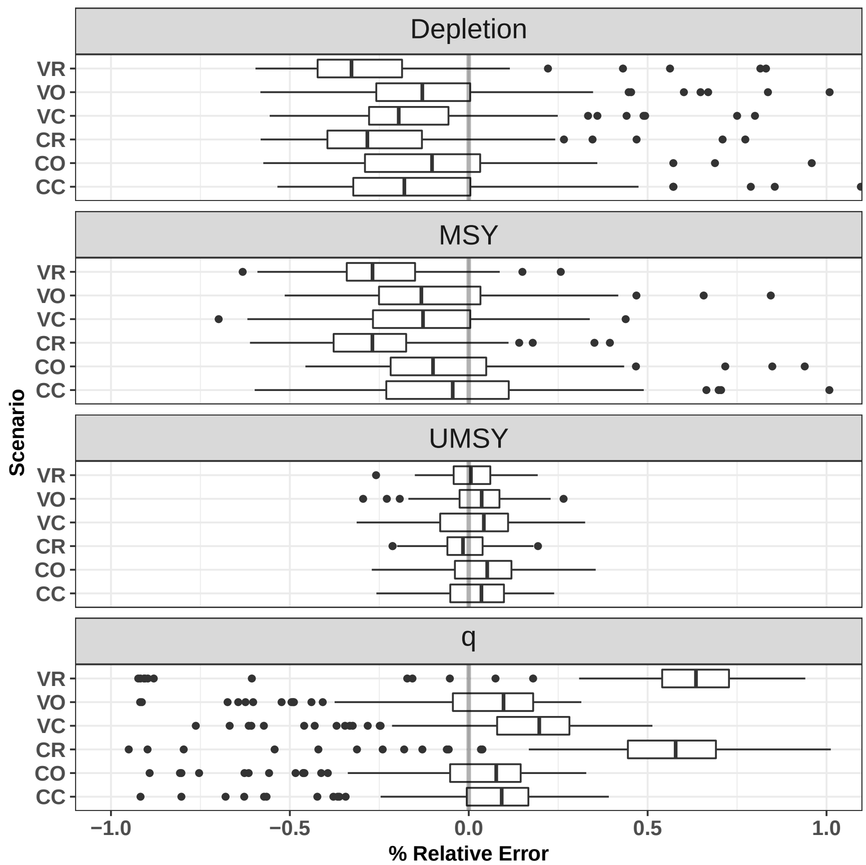

Figure 2: Relative error (%) for main parameters for all scenarios considered in the simulation-evaluation.

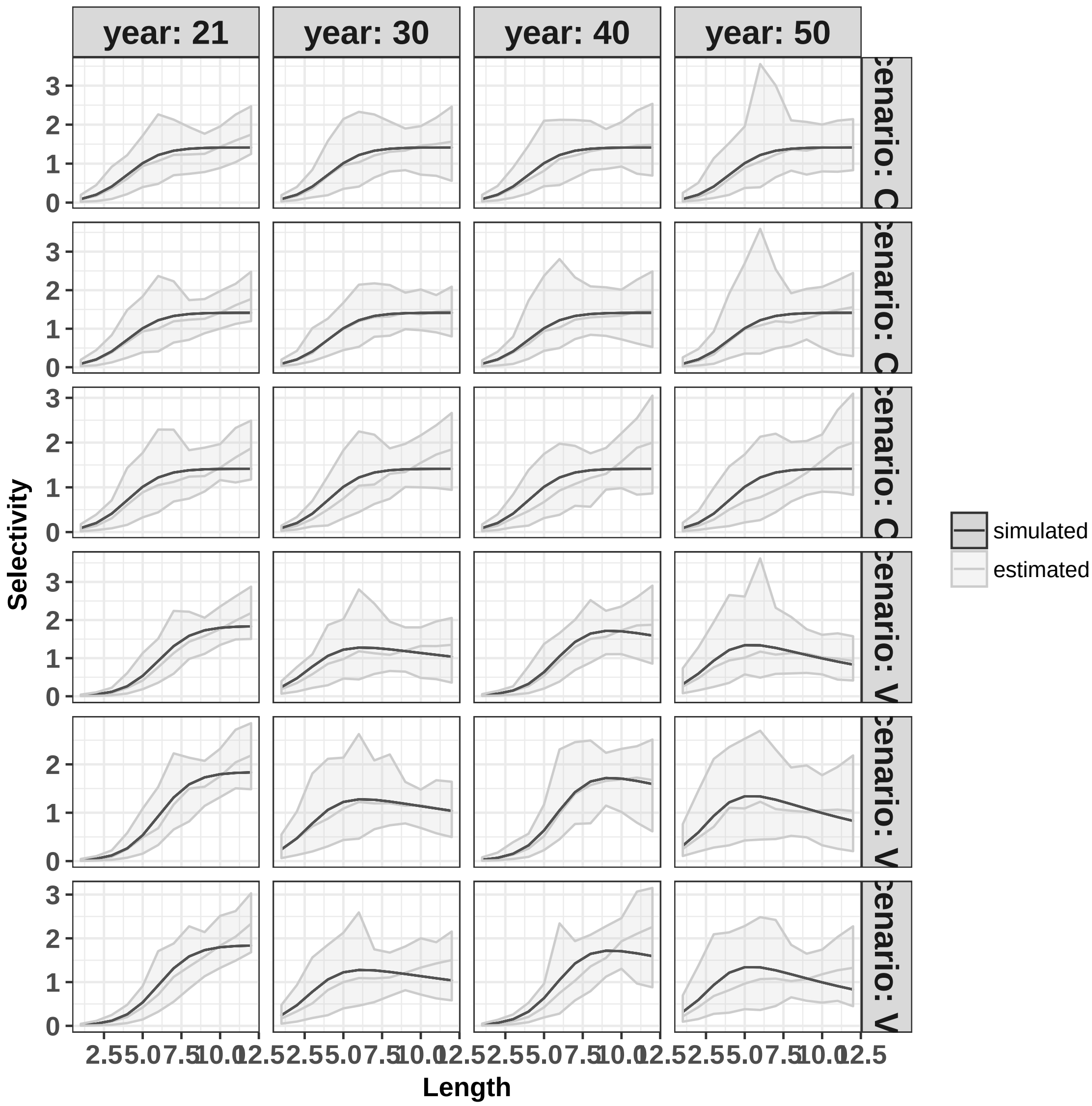

Figure 3: Simulated and realized selectivity estimates for a set of years within simulation-evaluation time series.

3.2 Real data examples

show figures with real

4 Discussion

We present a length-based stock reduction analysis that allows monitoring of time varying selectivity. In this model catch at length is assumed to be known without error and exploitation rate at length is derived directly from the estimates of numbers at length. This fact is important because it allows the mode to bypass the requirement for the estimation of a selectivity ogive, as is required in other length based models (Sullivan et al., 1990; Fournier et al., 1998).

Selectivity parameters can be particularly hard to estimate, especially it is changing over time and

main conclusions: Length-SRA underestimates kappa productivity parameter

Management quantities Depletion and MSY are underestimated - the model tends to produce conservative benchmarks.

UMSY estimates had low bias- good?

U ramp scenario yield very bad results - lack of information in the time series.

Model is able to track selectivity over time, - good

How does it compare to other length based models

Does the model work? Could it produce useful management advice? My nee dto evaluate in a closed loop frameworks to be certain.

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.

Further testing of this model in a closed-loop simulation set up would provide more insight into how

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.

Gerlotto, F., Gutiérrez, M., and Bertrand, A. (2012). Insight on population structure of the Chilean jack mackerel (*Trachurus murphyi*). *Aquatic Living Resources*, 25(4):341–355.

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.

Sullivan, P. J., Lai, H.-L., and Gallucci, V. F. (1990). A Catch-at-Length Analysis that Incorporates a Stochastic Model of Growth. *Canadian Journal of Fisheries and Aquatic Sciences*, 47(1):184–198.

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.

Walters, C. and Ludwig, D. (1994). Calculation of Bayes posterior probability distributions for key population parameters. *Canadian Journal of Fisheries and Aquatic Sciences*, 51(3):713–722.

Walters, C. J., Martell, S. J., and Korman, J. (2006). A stochastic approach to stock reduction analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, 63(1):212–223.