Performance of catch-based and length-based methods in data-limited fisheries

Journal:	*ICES Journal of Marine Science*
Manuscript ID	ICESJMS-2018-484.R1
Manuscript Types:	Original Article
Date Submitted by the Author:	19-Feb-2019
Complete List of Authors:	Pons, Maite; University of Washington, School of Aquatic and Fishery Sciences Cope, Jason; Northwest Fisheries Science Center, National Oceanic and Atmospheric Administration Fisheries Kell, Laurence; Imperial College London Department of Life Sciences, Centre for Environmental Policy
Keyword:	data-limited assessment methods, simulation testing, depletion, life-history, harvest rates

Performance of catch-based and length-based methods in data-limited fisheries

Maite Pons [1], Jason M. Cope [1,2], and Laurence T. Kell [3]

[1] School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, Washington 98195-5020, USA

[2] National Oceanic and Atmospheric Administration Fisheries, Northwest Fisheries Science Center, Fishery Resource Analysis and Monitoring Division, 2725 Montlake Boulevard East, Seattle, Washington 98112-2097, USA.

[3] Centre for Environmental Policy, Imperial College London, London SW7 1NE, UK.

Abstract

Despite the adoption of the Precautionary Approach to Fisheries the quality of data available for many small scale fisheries and bycatch species is insufficient to allow the application of conventional stock assessments methods. In such situations two main approaches are used, catch-only and length-only based methods. It is difficult to compare these methods as they provide different estimates of stock status. Therefore, we used exploitation rate as a common measure for comparisons using simulated populations with contrasting life histories, under different harvest scenarios and final depletion levels. The methods evaluated were Catch-MSY, State-Space Catch-Only Model (SSCOM), Depletion Based Stock Reduction Analysis (DBSRA), Simple Stock Synthesis (SSS), Length Based Spawning Potential Ratio (LBSPR) and Length-Based Integrated Mixed Effects (LIME). In general, results were less biased for the highly depleted stocks and more biased for longer-lived species. Length-based models performed

as well as catch-based methods in many scenarios. For fisheries that are still developing and time series of catch are unavailable, obtaining length data could provide estimates of stock status for use in management. While, if time series of catch are available and catch limits want to be enforce, catch-based methods such as DBSRA and SSS showed promise over other catch-based methods.~~The quality of data for many small scale fisheries and by caught species around the world is insufficient to use conventional stock assessments methods. Therefore, recently mMany methods have therefore been developed to assess fisheries when data and resources are limited. For example, catch-based models can be used when only total catch data are available and length-based models when only samples of the length composition are taken from the catch. Here, we evaluated the performance of both catch-based and length-based models, using simulation testing to estimate the exploitation status of species with contrasting life histories under different harvest scenariosrelative stock abundance. allFor unassessed fisheries where reconstructing time series of catch is possible and fishery removals can be managed, catch-based methods such as Depletion Based Stock Reduction Analysis (DBSRA) and Simple Stock Synthesis (SSS) seemed to be show promisethe best approach to assess stocks. For fisheries that are still developing and time series of catch are not available, obtaining length-composition data could give a good approximation of the exploitation stock status of the stocks. In many of the scenarios tested, length-based models such as Length Based Spawning Potential Ratio (LBSPR) performed as well as catch-based methods.~~

Keywords: data-limited assessment methods, simulation testing, depletion, life-history, harvest rates

Introduction

In many cases, ~~T~~the provision of fisheries management advice ~~requires~~ involves the assessment of stock status relative to reference points, the prediction of the response of a stock to management, and checking that those predictions are consistent with reality (Kell *et al.*, 2016). Simulation testing is therefore a key tool for checking that stock assessment methods are robust, particularly where data and knowledge are limited and a variety of simplifying assumptions need to be made. We therefore use an operating model (OM) to represents the main sources of uncertainty and to generate data for use in data limited stock assessment methods.

Major commercial species usually have substantial sets of data that can be integrated in ~~to inform~~ complex stock assessments models (e.g. Methot and Wetzel 2013); ~~this~~ these data may include ~~includes long~~ time series of total removals, catch-at-length or -age data, relative abundance indices, fishing effort, tag recoveries ~~size and/or age composition,~~ and information on life-history parameters. Most of the datasets required for ~~these~~ such stock assessments are unavailable ~~however,~~ for most small-scale fisheries around the world. Fisheries and stocks lacking ~~these~~ such multiple data types are commonly known as “data-poor” or “data-limited” fisheries (Costello *et al.*, 2012; Dowling *et al.*, 2015). Recently, many data-limited approaches have been developed to meet an increase demand for science-based fisheries management for unassessed fisheries, stocks and species -where resources are limited (Wetzel and Punt 2011; Costello *et al.* 2012; Dowling *et al.* 2015, 2016; Chrysafi and Kuparinen 2016; Rosenberg *et al.* 2017).

Assessing stocks using only catch and life history data started many years ago with the development of Stock Reduction Analysis, SRA (Kimura and Tagart 1982; Kimura *et al.* 1984)~~;~~. Since then, ~~the~~ this method has been extended to estimate productivity and reconstruct historical abundance trends by making assumptions about final biomass relative to unfished or initial biomass (i.e., stock depletion; Thorson and Cope 2015). SRA has subsequently been further extended to incorporate stochastic variability in population dynamics (Stochastic-SRA; Walters *et al.* 2006), a flexible shape for the production function (Depletion-Based SRA; Dick and MacCall 2011), prior information regarding resilience and population abundance at the start of the catch time series (Catch-MSY; Martell and Froese 2013; Froese *et al.* 2017), and age-structured population dynamics (Simple Stock Synthesis; Cope 2013). Despite these differences, this family of catch-only models shares a common dependence upon prior assumptions about final stock depletion. Simulation testing indicates that these methods perform well only when assumptions regarding final relative abundance are met. Also, some methods might perform differently under different stock depletion levels (i.e. highly depleted or slightly depleted stocks, Walters *et al.* 2006) or under different harvest history or catch trends. ~~and although they might be appropriate to predict sustainable catch or biomass, but not to reconstruct abundance time series (Carruthers *et al.*, 2012; Wetzel and Punt, 2015).~~

For many small-scale fisheries, obtaining reliable information on historical total catch is difficult, while ~~collecting~~ sampling of length measurements from ~~samples of~~ the catch is easier. Mean-length mortality estimators (Beverton and Holt, 1957) assume that fishing mortality directly influences the mean length of the catch under equilibrium conditions. This basic method is extended by l~~L~~ength-based spawning potential ratio (LBSPR, Hordyk *et al.* 2015a) and length-based Integrated Mixed Effects (LIME, Rudd and Thorson 2017) models~~, have recently been developed~~. These allow~~ing~~ the estimation of instantaneous fishing mortality (*F*) and spawning potential ratio (SPR) when basic biological parameters are known. SPR is the proportion of the unfished reproductive potential per recruit under a given level of fishing pressure (Goodyear

1993). Both methods have the same data-requirements, but LIME does not assume equilibrium conditions; the mixed-effects aspect of LIME extends length-based methods by estimating changes in recruitment and separating it from fishing mortality over time (Rudd and Thorson, 2018).

It is good practice to simulation test the performance of assessment methods before applying them in practice~~There is increasing interest in developing new methodologies approaches to quantitatively assess data-limited fisheries to manage them and prevent overfishing. Usually, these assessment method performances are tested using simulation experimentation~~ (Cope, 2008).~~-~~ Carruthers *et al.* (2016) used a closed-loop simulation approach to compare a range of management procedures for setting catch limits in data-limited fisheries. They found that data-limited methods using observations of stock depletion offer the best overall performance across life history types, data quality and autocorrelation in recruitment strength. However, these management procedures are based on setting catch limits and were designed for use in data-limited fisheries for which annual catch data are available, sometimes together with a relative abundance index (~~delay-difference stock assessment,~~ Carruthers *et al.* 2014). In many data poor fisheries, measuring total removals is difficult, as is enforcing catch limits. Recently, Hordyk *et al.* (2015b) tested some harvest strategies using a simulation approach to assess the utility of LBSPR as a tool for management in data-limited fisheries using an effort-based harvest control rule. They found that the LBSPR assessment model with an iterative effort-based harvest control rule can be used to rebuild an overfished stock back to sustainable levels or fish down a stock to the target SPR. However, no

~~No~~ studies, ~~however,~~ have compared the performance of both, length-based and catch-based methods either ~~as estimation models~~ in a closed-loop simulation or as assessment models to estimate stock status. A~~using the same simulated populations, a~~ possible reason is because finding a common metric between catch-based and length-based stock status metrics is difficult, as the former measures overfishing and the latter stock depletion. ~~.~~ ~~Therefore,~~

In the present study, we use simulation testing where an operating model (OM) is used to represent~~s~~ the main sources of uncertainty and to generate data for use in data-~~-~~limited stock assessment methods. We ~~then used~~use a common metric for comparison across models, namely exploitation or harvest rates, to evaluate the performance of catch-based and length-based models using simulation testing. We considered ~~to estimate exploitation status for~~ three ~~different~~ fish stocks with contrasting life history strategies under three different ~~possible~~ ~~exploitation~~harvest ~~scenarios~~trends, and three levels of final stock depletion.

Methods

Three different scenarios of times series of 20 years of harvest or ~~exploitation rate (*U*)~~catch trends ~~scenarios~~ were considered~~, which~~ corresponding to ~~historical~~ fishing mortality histories commonly seen in many fisheries. In ~~the first~~harvest scenario 1, ~~*U*~~ catch increases until it reaches a maximum and start declining afterwards; this is a classic example where management measures ~~were~~ are implemented to reduce fishing pressure or where catch decreases because the population is heavily depleted and hence less productive. ~~The second~~ Harvest scenario 2 assumes that ~~*U*~~ catch increases and remains constant after reaching a maximum; ~~this could be~~foe example due to implementation of catch or effort limits~~ for example~~. ~~The third~~ Harvest scenario 3 has constantly increasing ~~U~~catches, which would occur for fisheries that are still developing, so effort and catches continue to increase (Figure 1a).

In addition, ~~three population~~ life history types of varying longevity and somatic growth rate were simulated: (i) a short-lived fast-growth species, ~~pacific~~ Pacific chub mackerel, *Scomber* *japonicus*, (ii) a medium-lived medium-growth fish, albacore tuna, *Thunnus alalunga*, and (ii) a longer-lived slow-growth species, canary rockfish, *Sebastes pinniger* (Table 1). Finally, three different depletion levels were considered, from heavily fished (depletion = 0.2), sustainable fished (depletion = 0.4), to slightly fished (depletion = 0.6).

OMs were generated using a factorial design encompassing 27 scenarios: three harvest scenarios (Scenarios 1 to 3, Figure 1a); three contrasting life histories (chub mackerel, albacore tuna and canary rockfish, Table 1); and three different final stock depletion levels (0.2, 0.4 and 0.6). ~~In order to simulate different exploitation histories in each run, we added a random observation error to the catch of 0.1 around the observed catch value in each year.~~

Operating model specifications

The OM was implemented using Stock Synthesis (SS) Version 3.30.10 (Methot and Wetzel 2013~~(Methot and Wetzel, 2013)(Methot and Wetzel, 2013)(Methot and Wetzel, 2013)~~) in order to simulate age structured populations and ensure consistency between the assumptions, simulated stock dynamics and the pseudo data generated. SS assumes that the absolute level of catch is known well enough to allow the model to calculate the level of fishing intensity needed to obtain that level of catch conditioned on the model estimates of age-specific population abundance and selectivity (Methot and Wetzel, 2013). Fishing intensity in SS is estimated to match the observed catch; the harvest rate is therefore the total annual catch divided by the total abundance of the exploited biomass. As outlined above three different catch histories were considered namely harvest scenario 1) that increases until it reaches a maximum and start declining afterwards; harvest scenario 2) that increases and remains constant after reaching a maximum and harvest scenario 3) that is constantly increasing. ~~Different catch histories thus lead~~

~~to different exploitation histories, and the different scenarios could affect the performance of a data-limited methods.~~ For each OM (N=27) 100 catch trends data-sets were simulated with a random observation error of 10% around the observed catch value in each year. Total biomass (*B*) and catches were then extracted from the OM to calculate the harvest rates (*U*) as *B*/catch per year (Figures S1 to S3).

~~In order to simulate different exploitation histories in each run, we added a random observation error to the catch of 0.1 around the observed catch value in each year.~~

To condition the OM published life history parameter values (Table 1) reported in formal stock assessments were used (~~ICCAT, 2014;~~ Crone and Hill, 2015 for the short-lived Pacific chub mackerel; ICCAT, 2014 for the medium-lived albacore tuna; and Thorson and Wetzel, 2015 for long-lived canary rockfish). Each population was assumed to be targeted in a single area, by only one fleet with a selectivity pattern (parameters in Table 1) that was logistic and constant through time. This controls for any complications arising from having multiple fleets and allow focus to be on method performance, not model specification.

Each simulated population began at ~~an~~ the unfished biomass level and all catch scenarios terminate at the specified ~~same stock~~ depletion level. No indices of abundance were included, but ~~we defined~~ a final stock depletion level was defined, implemented through the use of a survey index in SS equal to 1 at the beginning of the time series and 0.2, 0.4 or 0.6 in the last year depending on the depletion level scenario considered~~, so spawning biomass (SB) in the last year is 0.4 SB_0~~ (Cope, 2013). ~~Each simulated population began at an unfished biomass and all catch scenarios terminate at the same stock depletion level. In each OM, all parameters were fixed, except R_0 (which allow the model to set the absolute biomass of the population), and annual lognormal recruitment deviations were assumed (Table 1).~~ A Beverton–Holt spawner–recruit function ~~was assumed~~ (Beverton and Holt, 1957) and annual lognormal recruitment deviations were assumed (parameters in Table 1).

To ~~simulate~~ simulate catch length frequencies from each simulation, the expected numbers at length data from the population by year was extracted from each SS simulation. ~~age-length conversion matrix output from SS was used to assign a length distribution to each age. Summing within a length bin then gave the length distribution of the catch.~~ Length bins were defined based in their assessments at every 2 cm ~~from 30 to 150 cm for Albacore,~~ from 12 to 76 cm for the short-lived Pacific ~~Chub~~ chub ~~Mackerel~~mackerel (Crone and Hill, 2015), from 30 to 150 cm for the medium-lived albacore tuna (ICCAT, 2014) and from 8 to 60 cm for the long-lived ~~Canary~~ canary ~~Rockfish~~rockfish (Thorson and Wetzel, 2015). To obtain samples of the length composition of the catch to use in the length-based assessment models, 1000 fish/year were drawn using a multinomial distribution from the ~~catch~~ numbers at length by year, using the probability of being caught (selectivity) at each length bin in each year.

~~Oms were generated using a factorial design encompassing 9 factors; namely (i) three scenarios of harvest rates (Figure 1 to 3) and (ii) life history with three levels (Table 1). For each OM 100 data-sets were simulated for harvest rate (U), total biomass (B) and SPR (Figure 1 to 3).~~

Comparing methods outputs

One of the challenges when comparing catch-based and length-based methods is they produce different model outputs. Catch-only models estimate total and/or spawning stock biomass and sustainable catches, whereas length-based models estimate exploitation and transient SPR, which it can be used to infer stock status. These are fundamentally different measures of the population. The performance metric was then defined as the error relative to the OM (RE), where RE = $(U_{Method}-U_{OM})/U_{OM}$. This allows the measure of uncertainty, in both bias and precision, in the methods under each scenario, and is used as a standardized metric of model performance. Bias in this study is how far, on average, the performance measure from each estimation model is from the OM. Imprecision is related to the variability around central tendency.

We used the exploitation rate (*U*) as a common measure for comparisons between each data limited method and the OM. For catch-only approaches it was defined as catch/biomass; for the length-based models, the estimated fishing mortality (*F*) was transformed to an exploration rate via $U = 1 - e\,(-F)$. In addition, we presented the average RE across the last 5 years of the time series and not along the entire time series of data, since we are interested in the estimation of the current exploitation rates. Different studies have shown that catch based methods might be appropriate to predict sustainable catch or biomass at the end of the time series, but not to reconstruct a biomass time series (Carruthers *et al.*, 2012; Wetzel and Punt, 2015).

Estimation models

Each catch or length-based method evaluated ~~is described~~ are summarized ~~in detail~~ below:~~.~~

Catch-based data-limited methods

Catch-MSY (CMSY; Martell and Froese 2013). ~~It is a~~A SRA approach with a Schaefer biomass dynamic model. ~~As input data, it requires~~Inputs are a time series of removals, priors for ~~ranges of~~ the population rate of increase at low population size (*r*), ~~and~~ carrying capacity (*K*), and possible ranges of relative stock sizes in the final year ~~of the time series~~(stock depletion) (Table 1). ~~Probable ranges for~~Values of *r* and *K* are filtered with a Monte Carlo approach to detect ‘viable’ *r*-*K* pairs. A parameter pair is considered ‘viable’ if the corresponding biomass trajectories calculated from a production model are compatible with the observed catches, so that the population abundance never falls below 0, and is compatible with the prior assumption of relative biomass (i.e., stock depletion; Martell and Froese 2013). ~~The~~ *r*-*K* pairs are drawn from uniform prior distributions and the Bernoulli distribution is used as the likelihood function for accepting each *r*-*K* pair. CMSY uses catch and productivity to estimate MSY. However, here we used the modified version of CMSY by Rosenberg *et al.* (2017) to extract biomass trends from all viable *r*-*K* pairs. T~~hen t~~he biomass trajectory is calculated as the median of all viable biomass trajectories generated during the Monte Carlo process. ~~We used t~~The R package *datalimited* version 0.1.0 (Anderson *et al.* 2016) available at https://github.com/datalimited/datalimited was used.~~.~~

State-space catch-only model (SSCOM). ~~This is a hierarchical model based on a coupled harvest-dynamics model.~~ The model is a Bayesian state-space model that integrates across three stochastic functional forms: variation in effort, population dynamics and fishing efficiency (Thorson *et al.*, 2013). SSCOM can reconstruct biomass time series from catch data whenever fishing mortality follows semi-predictable dynamics over time. The different types of population and effort dynamics can be extracted from the same catch stream using nonlinear models for population-dynamics as a function of biomass and linear models for effort dynamics as a function of log-scaled biomass for example. ~~We used t~~The package *datalimited* version 0.1.0 (Anderson *et al.*, 2016) ~~to run this mo~~was useddel. ~~We modified t~~The code was extended to extract biomass trajectories and to use a lognormal distribution for depletion (Table 1). However, the effort dynamic priors were set as in Anderson *et al.* (2017). Using this modified version of SSCOM, the required inputs are priors for *r*, *K*, and final stock depletion (Table 1).

Depletion based stock reduction analysis (DBSRA). DBSRA (Dick and MacCall, 2011) modifies the SRA approach as it uses Monte Carlo draws from four parameter distributions (M, F_{MSY}/M, B_{MSY}/B_0 and *depletion*) while using age at maturity (A_{mat}) to separate the biomass into immature and mature biomass (fishery selectivity is also assumed to have an identical pattern to the age-at-maturity ogive). It uses a delay-difference production model with a time lag for recruitment and mortality as:

$$B_{t\underline{+1}} = B_{t-\cancel{1}} + \cancel{P_t}\underline{P}\,(B_{t-Amat}) \cancel{+ (1 - e^{-M})(B_{t-Amat} - B_{t-1})} - C_{t-\cancel{1}}$$

where B_t is the biomass at the start of the year t, ~~M is the instantaneous rate of natural mortality, and P_t~~ P (B_{t-Amat}) is the latent annual production based on a function of adult biomass in year t-$Amat$ and C_t is the catch in year t. ~~.~~Biomass in the first year (B_0) is assumed equal to K. The package *fishmethods* version 1.10-3 was used to perform this analysis (Nelson, 2017). For DBSRA we used the age at maturity ($Amat$) and natural mortality (M) as fixed inputs, and three priors: for final stock depletion, F_{MSY}/M and B_{MSY}/B_0 (distributions in Table 1). Each of these is assigned a distribution from which the Monte Carlo draws are taken.

Simple Stock Synthesis (SSS). This method is based on the ~~Stock Synthesis~~SS package (Methot and Wetzel, 2013). The approach uses the ~~Stock Synthesis~~SS framework by fixing all parameters in the model except for initial recruitment (lnR_0), which is the only estimated parameter. It also sets up an artificial index of abundance that represents the relative stock biomass. Thus the first value of the index is always 1, and the final year value represents the percent of the population left in that year. The values of ~~Like DBSRA, SSS uses Monte Carlo draws of M and,~~ steepness (h)~~,~~ and ~~initial recruitment (R_0)~~ the final year of the ~~while fitting to an artificial~~ abundance survey are all randomly drawn from specified distribution using a Monte Carlo approach ~~representing stock depletion~~ (Cope, 2013), and lnR_0 is then estimated. ~~All fixed~~

~~values are drawn from prior distributions to represent uncertainty in model-derived outputs.~~ The code for loading the ~~sss~~SSS library in R and user instructions~~running~~ ~~SSS~~ can be found at https://github.com/shcaba/SSS. Benefits of this approach is that it retains the same model structure of the data-rich stock assessments, but still allows for flexibility in a variety of parameter and model specifications, if desired. It also shares similarity in model structure to the OM, with growth being specified at the OM level. While selectivity could have also been matched to the OM, it was instead assumed to be equal to maturity, a shared assumption with DB-SRA. The input priors used for SSS were: relative stock status and steepness.

Length-based data-limited methods

Length based spawning potential ratio (LBSPR). In LBSPR, SPR in an exploited population is a function of the ratio of fishing mortality to natural mortality (F/M), selectivity and the two life history ratios $M/$~~K~~ k and L_m/L_∞; k is the von Bertalanffy growth coefficient, L_m is the size of maturity and L_∞ is asymptotic size (Hordyk *et al.*, 2015a). The inputs to LBSPR are: M/k, L_∞, the variability of length-at-age (CVL_∞), which is normally assumed to be around 10%; and length at maturity specified in terms of L_{50} and L_{95} (the size at which 50% and 95% of a population matures). Given the assumed values for ~~the~~ $M/$~~K~~ k and L_∞ ~~parameters~~ and length composition data are from an exploited stock~~,~~ ~~the~~ LBSPR model uses maximum likelihood methods to estimate the selectivity ogive, which is assumed to be of a logistic ~~curve~~ form defined by the selectivity-at-length parameters S_{50} and S_{95} (the size at which 50% and 95% of a population is retained by the fishing gear), and the relative fishing mortality (F/M), and these are used to calculate SPR (Hordyk *et al.*, 2015 a,b). Estimates of SPR are primarily determined by the length of fish relative to L_{50} and ~~and~~ L_∞, ~~,~~but it also depends on ~~Dependent on the selectivity, if a reasonable proportion of fish in a sample attain sizes approaching L_∞, a high estimate of SPR~~

~~will be derived.~~ life history parameters such as fecundity-at-age/length and selectivity. LBSPR is an equilibrium based method with the following assumptions: (i) asymptotic selectivity, (ii) growth is adequately described by the von Bertalanffy equation, (iii) a single growth curve can be used to describe both sexes which have equal catchability, (iv) length at-age is normally distributed, (v) rates of natural mortality are constant across adult age classes, (vi) recruitment is constant over time, and (vii) growth rates remain constant across the cohorts within a stock (Hordyk *et al.*, 2015a). Analyses were conducted using LBSPR package version 0.1.2 in R (Hordyk 2017). We used the Rauch-Tung-Striebel smoother function available in the LBSPR package to smooth out the multi-year estimates of F.

Length-based integrated mixed effects (LIME). ~~The model uses l~~Length data and biological information are used to estimate *F* and SPR. LIME has the same data-requirements as LBSPR, but does not assume equilibrium conditions; the mixed effects aspect of LIME extends length-based methods by estimating changes in recruitment and fishing mortality over time (Rudd and Thorson, 2018). LIME uses automatic differentiation and Laplace approximations as implemented in Template Model Builder (TMB; Kristensen *et al*. 2016) to calculate the marginal likelihood for the mixed-effects. All other assumptions are the same as LBSPR but LIME estimates one selectivity curve for the entire time series of length data while LBSPR estimates one selectivity curve for each year since each time step estimation in LBSPR is independent (Hordyk *et al.*, 2015a). The inputs to LIME are: M, k, L_{∞}, t_0, CVL_{∞}, L_{50}, L_{95}, h and the parameters of the length-weight relationship a and b (Table 1).

~~*Comparing methods outputs*~~

~~One of the challenges when comparing catch-based and length-based methods is they produce different model outputs. Catch-only models estimate total and/or spawning stock~~

~~biomass and sustainable catches, whereas length-based models estimate exploitation and transient SPR, which is similar to relative stock statusstock depletion. These are fundamentally different measures of the population. We therefore used the exploitation rate (U) as a common measure. For catch-only approaches it was defined as catch/biomass; for the length-based models, the estimated fishing mortality (F) was transformed to an exploration rate via $U = 1 - e(-F)$. TGiven the challenge of having different metrics, the performance metric was then defined ast he of a method can be compared to the OM and described as the Relative eError relative to the OM (RE), where RE = (U_{Method}estimated-U_{OM}true)/U_{OM}true. This allows the measure of uncertainty, in both bias and precision, in the methods under each scenario, and is used as a standardized metric of model performance. Bias in this study is how far, on average, the performance measure from each estimation model is from the true valueOM. Imprecision is related to the variability around that estimated average valuecentral tendency. We used as a performance the exploitation rate (U) catch/biomass for the catch-based models and, for the length-based models, to scale the estimated fishing mortality (F) between 0 and 1 we used this transformation: $U = 1 - e(-F)$. In addition, we presented the average RE across a period of time equal to the generation length of each species (see Table 1).~~ a

RESULTS

~~*Differences among harvest rate scenarios*~~

Large variability is seen in model performance a~~mong~~cross life histories, harvest scenarios and depletion levels (Figure 1). In Table 2 and 3 scenarios are sorted by catch-based and length-based model types, respectively~~.~~~~, so, we will describe the results for each species separately later.~~ Figure 1 can be used as a decision figure to search for the best model, i.e. to identify the less bias and more precise method based on the life history of the species, harvest trends, and knowledge of final depletion. For example, all assessment methods perform best at high depletion levels (0.2) and CMSY and SSCOM tend to perform poorly.

It was to be expected that the various catch-based models considered in this study to perform differently since they have different model structure and assumptions. In general, c~~In general, c~~atch-based models were ~~more~~ less biased and more ~~less~~ precise when ~~there was no contrast in the time series of catch data~~stocks were more depleted (i.e. using a prior centered around 0.2, Figure 1) ~~(Scenario 2 and 3, Figure 1 to 3)~~.

SSS performed best in most cases estimating unbiased exploitation rates across different scenarios of harvest trends, final stock depletion and life histories. However, it tended to underestimate harvest rates by 50% for the short-lived species in most of the scenarios considered. Although DBSRA also performed very well, it tended to underestimated harvest rates in general (Table 2).

CMSY was the most bias of the catch-based models tested overestimating harvest rates in particular for stocks that are slightly fished (i.e. using a prior centered around 0.6). CMSY produced non-biased estimations of MSY in some cases such as for highly depleted medium-lived species or sustainable fished short lived species (Table S1).

SSCOM was less biased than CMSY in most scenarios, but less precise than any other catch-based models, showing the broadest range of RE in particular when the stocks where slightly depleted. Also, it was highly and positive biased for the harvest scenario 1 where catch decreased at the end of the time series (Table 2, Figure 1). In the case of CMSY and SSCOM multiple modes were seen this suggests that the estimates are unstable and if either method used to provide advice large changes could be estimated between years which may not reflect actual stock status.

In many cases, length models gave a less biased estimation than catch-based models (Figure 1). However, LBSPR in general underestimated harvest rates and always estimated a lower harvest rate than LIME. LBSPR was more biased in harvest scenario 3 where fishing intensity keeps increasing at the end of the time series although, length-based models in general were not highly influenced by different catch trends. LIME, on the other hand, overestimated harvest rates in most scenarios considered. Both, LBSPR and LIME were less~~er~~ biased for the medium-lived species (Figure 1b) and more biased for the long-lived species (Figure 1d). LIME was less biased for the short-lived species than LBSPR and sometimes, it was less biased when the stocks where highly depleted (Table 3). LIME also showed multiple modes for long-lived stocks which may suggest poor convergence.

_~~For example, for the fast-growth mackerel, the median RE for CMSY in Scenario 1 was 1.3 (range: 1.1 to 1.5), but it increased to 3.2 in Scenario 3 (range: 2.1 to 4.2). However, LIME showed similar patterns in exploitation rates in both bias and precision for each species across scenarios and LBSPR seems to be more positive biased in Scenario 1 than in the Scenarios 2 and 3 for all species (Figure 4).~~

~~Catch-based models~~Short lived species

For the short lived life history, DBSRA and SSS were the least biased and most precise for the catch based methods. SSS however, was positively biased when harvest rates were decreasing at the end of the time series and DBSRA was negatively biased (i.e. Scenario 1) when the population was more depleted (i.e. depletion centered at 0.2 and 0.4). DBSRA showed to be sensitive to different catch trends but not so much to different final depletion levels (Figure 1b).

CMYS was always positively and highly biased, although the mean RE was lower for highly depleted stocks. SSCOM was less biased and imprecise in harvest scenario 3 where the catch keeps increasing through time and in harvest scenario 2 but only when the stock was highly depleted.

Between the two length-based models, LIME showed a better performance than LBSPR. LBSPR always underestimated harvest rates. LIME was unbiased for the highly depleted stocks (i.e. depletion = 0.2) under any harvest rate trend but, it overestimated harvest rates under the other scenarios in particular for slightly depleted stocks (i.e. depletion = 0.6, Figure 1b).

Medium lived species

Both catch-based and length-based models showed more precision as the stocks were more depleted. Among the catch-based methods, SSS was the most precise and less biased followed by DBSRA. DBSRA was slightly negatively biased for populations more depleted (i.e. depletion values around 0.2). CMSY was highly positive biased but less biased for highly depleted stocks. SSCOM was more precise for the harvest scenario 3 and more precise in the harvest scenario 2 when the depletion centered around 0.4 (Figure 1c).

Between the length-based models, LBSPR showed a very good performance in both bias and imprecision in the harvest scenarios 1 and 2 under the three depletion levels. LBSPR underestimated harvest rates in harvest scenario 3 where the catch keeps increasing. LIME showed a good performance in general but slightly positive biased in the harvest scenario 1 where the catch decreases at the end of the time series, in particular for the lesser depleted stocks (Figure 1c).

Long lived species

Catch-based and length-based methods, as well as for the other life histories, showed more precision (less variability in RE) as the stocks were more depleted. Among the catch-based methods, SSS was the most precise and less biased in harvest scenario 1. However, SSS underestimated harvest rates in scenarios 2 and 3 where the catch history is~~catch keeps~~ constant or increases at the end of the time series. In particular, for the long-lived canary rockfish a less biased and more precise estimations where observed for scenario 1 with depletion levels centered around 0.2. DBSRA was negatively biased in all cases. CMSY in harvest scenario 2 and 3 ~~and in particular~~performed relatively best compared to itself in this species relative to the others species ~~for CMSY this was the best performance in comparison to the other species~~ (Figure 1d).

LIME was highly imprecise and in general positively biased, except for the harvest scenario 1 when the stock was heavy depleted. LBSPR also showed its best performance under the same scenario. Length-based models showed to be highly imprecise for long-lived species (Figure 1c).

~~We expected that the various catch-based models considered in this study to perform differently since they have different model structure and assumptions. SSS and DBSRA were, in general, less biased and more precise than other catch-based models, in particular when comparing with both catch-MSY models. CMSY showed the worst performance among all catch-based models tested in this study. CMSY presented the highest RE for mackerel in Scenario 3 (median RE=3.1) and the lowest RE for Albacore in Scenario 1 (median RE=0.49). Even in this case, harvest rates were estimated to be 50% higher than the truth. CMSY was less biased when the catch time series had an increasing followed by a decreasing trend (Scenario1)~~

~~and highly biased when there was no contrast in the time series of catch (Scenario 2 and 3) for the three species (Figure 4).~~

~~SSCOM in general was less biased than CMSY and in some cases less biased than DBSRA or SSS (i.e. for mackerel). However, SSCOM was less precise than any other catch-based model, showing a broader range of RE in most of the cases (Figure 4).~~

~~SSS estimated unbiased exploitation rates across different scenarios of harvest trends and life histories, except for the case of the fast-growth mackerel. It was positively biased in those cases; the harvest rates were estimated to be 66% higher than the true values (median RE=0.66) in Scenario 1 and around 38% in Scenario 2 (median RE=0.38) and 30% Scenario 3 (median RE=0.30). For this species, DBSRA and SSCOM was less biased in the three scenarios. SSS was the less biased estimation method for Albacore and Rockfish for the three scenarios. The median RE range between -0.02 in Scenario 3 to 0.15 in Scenario 1 for Albacore, and between -0.09 in Scenario 3 and 0.15 in Scenario 1 for Rockfish. DBSRA seemed to be more sensitive to the different trends in harvest rates for the medium-growth Albacore (Figure 4).~~

~~In summary, among catch-based models, CMSY was the least precise and most positively biased, particularly in Scenarios 2 and 3 for the three species. SSCOM was also less precise but less biased than the CMYS and DBSRA in general. The method that performed best in terms of bias and precision was SSS most of the time (Figure 4). The age-structured aspect of SSS has also been shown elsewhere to be better suited for slower life histories (Wetzel and Punt 2015).~~

~~*Length-based models*~~

~~In some cases, length models gave a less biased estimation than catch-based models. LBSPR was generally less biased for slow-growth species like rockfish and highly positively~~

~~biased for fast-growth species like mackerel. LIME was positively biased in all cases and highly imprecise in general but in particular for the slow-growth species (Figure 4).~~

~~LIME did not converge in many cases, however, between 32% of the times in Scenario 1 and 9% in Scenario 3 for Albacore, between 27% in Scenario 1 and 4% in Scenario 2 for Mackerel, and between 61% in Scenario 1 and 67% in Scenario 3 for Rockfish. In all cases, runs that did not converge were not included. For slow-growth species LIME had more difficulties in converging and in this case also showed the greatest imprecision in estimates of harvest rates (Figure 4).~~

~~In summary, between the 2 length-based models, LBSPR was more precise than LIME in general. Both showed similar performance for slow-growth and medium-growth species and very different and opposite performance for fast-growth species. For short lived species LIME was less biased. On the other hand, for slow growth species LBSPR was less biased (Figure 4).~~

~~In Scenarios 2 and 3, where there is no contrast in the time series of catch, length-based methods performed better than CMSY. But, in general, all data-limited models tested here performed worse for the fast-growth species (Figure 4).~~

Discussion

Simulation studies commonly ~~use~~ make different model ~~specifications~~ assumptions from those of the methods being tested to allow robustness to be evaluated; in some cases, however, often the same population model is used for simulation and estimation, i.e. self-testing. Using the same model structure for simulation and estimation could result in optimistic results that might not be true under many scenarios (Francis, 2012). For example, it is not possible to explore the robustness to model structure and assumptions when the model used for simulation and estimation is the same. If a method performs poorly, however, when the assumptions in the OM are the same as the assessment method then it is unlikely to perform well in practice.

Our approach evaluated multiple data-limited assessment methods that assume different population dynamics, uncertainty and fishing effort dynamics. It is to be expected, due to these differences, that the various methods would performed differently. Rosenberg *et al.* (2017) used four catch-based data limited models and found that frequently models disagreed about population status estimations with no model showing good performance across all fish stocks, i.e. high precision and low bias, and the performance of the models~~-~~ depended on the scenario considered. When scenarios are chosen to represent ~~uncertainty about the~~specific fishing dynamics ~~this infers that the method may not be robust and when a~~ or ~~scenario~~ to represents a particularly stock or fishery then it may be difficult to draw any general conclusions. This is why we chose scenarios ~~were chosen~~ to represent different harvest ~~rates~~trends, depletion levels and species with contrasting life histories. ~~It was~~We found that model performance is highly dependent on the life history of the species of concern, ~~the dynamic of the population~~catch trends and the fishing intensity. ~~This is why scenarios were chosen to represent different harvest rates and species with contrasting life histories.~~

In general, catch-based models were less biased and more precise ~~when there was contrast (e.g., an increase in the catch and then a decrease) in the time series of catch data~~for stocks highly depleted compared to the length-based models. Walters *et al.* (2006) suggested that for SRA, stocks that have experienced extensive historical depletion gains precision from a high rate of rejected parameter draws. Moreover, ~~In Scenario 3, where catch is still increasing, it is very difficult to have a good estimate of the carrying capacity *K*. So,~~ SSS and DBSRA, which use priors in steepness or F_{MSY}/M and B_{MSY}/B_0, respectively, performed better than the models that only relied in *r* and *K* such as CMSY and SSCOM, even ~~with~~ when priors for depletion were centered in the true values. Length-based models, on the other hand, were not ~~dependent~~ highly influenced by~~on the harvest rate~~ catch trends. This is not surprising for LBSPR since in equilibrium conditions, the estimates are snapshots of the population and independent in each time step. For LIME, which is not an equilibrium model, Rudd and Thorson (2017) also did not find strong differences for alternative fishing mortality scenarios.

The choice of a “best model” ~~also depend~~depended on the life history characteristics and the biological information that is available. SSS seems to be the least biased catch-based model. However, unlike other catch-based models, age and growth estimates are needed in SSS to define age structure and remove catch according to age-/size-based selectivity patterns (Cope, 2013). SSS has the same structure as SS and this might be the reason why is the model that performed the best in most cases ~~since the simulation and estimation models have the same parametrization~~ (Francis, 2012). It is always good to sue more than one model with different structures and input parameters.

DBSRA and SSS performed very similar in ~~some~~ many cases ~~(Albacore Scenario 1, Mackerel Scenario 2 and 3)~~. In structure, both models are very similar, however there are a few notable differences between the population dynamics models used in DBSRA and SSS that could explain the different results found here for the other species (Wetzel and Punt, 2015). The underlying population dynamics model in SSS is fully age-structured whereas DBSRA uses a delay-difference model based on a biomass production function. When information about growth parameters or age structure of the entire population is unknown, but the age at maturity can be inferred, DBSRA can be used instead of~~d~~ SSS producing very similar unbiased and precise results. When possible, both should be used and compared, with model averaging ~~could be~~ consideration to combine the results (Hoeting *et al.* 1999)~~ed~~.

SSCOM was less biased in some cases, but always less precise than DBSRA. ~~We found that SSCOM and DBSRA performed similarly in terms of bias. In the SSCOM model, a prior in depletion is not needed, but it can be included as we did in this study. Thorson *et al.* (2013) explored the effect of specifying a prior on final depletion and compared the results with DBSRA. They suggested that using a strong prior on final depletion in SSCOM would result in similar performance to DBSRA.~~ Both, DBSRA and SSCOM approximate biomass dynamics using a production function expressed as exploitable biomass (which is equivalent to spawning biomass given selectivity and maturity curves are assumed identical), and both assume that biomass starts at average unfished biomass. However, DBSRA uses deterministic biomass dynamics and uses an asymmetric production function (Dick and MacCall, 2011), while SSCOM has stochastic biomass dynamics and uses a Schaefer production function (Thorson *et al.*, 2013), so it unsurprising they did not performed exactly the same. ~~SSCOM was less biased in some cases, but always less precise than DBSRA.~~ Specifying other priors in SSCOM in future studies, for example for effort-dynamics, could increase its precision. In addition, this method might be more appropriate for stocks with longer time series of catch. Pons (2018) found a better performance for SSCOM using the same species but with a longer time series (~ 80 years of catch data).

CMSY performed very poorly in all scenarios overestimating harvest rates, even when given a prior for depletion close to the true value. A key point of the CMSY is the ability to define a reasonable prior range for the parameters of the Schaefer model in particular *K*. In our case, we have arbitrarily chosen ~~50~~ 100 times the maximum catch as the upper bound for *K*~~.~~

~~However, in the Scenario 3, in a developing fishery, or a fishery that has a continuous increase in catch, it will be more difficult to define the upper bound of *K* because the maximum potential has yet to be reached~~based on ~~(~~Martell and Froese~~,~~ (2013)~~, thus limiting the performance of these methods under this scenario. However, they also performed poorly in Scenario 1 and 2, in particular for long-lived and short-lived species.~~ . Other *K* values could be explored in future studies to see if this improves the outcomes, but it remains a very difficult parameter to specify. ~~For example, Martell and Froese (2013) used maximum catch multiplied by 100.~~ Rosenberg et al. (2017) and Free et al. (2017) found that CMSY was the one that performed second best and better than SSCOM in their scenarios. One of the differences with our study is that they considered a uniform prior for depletion in SSCOM and we considered a Lognormal prior centered around the true value, but it is apparent that method performance is sensitive to a variety of scenarios.

LIME was highly imprecise for the long-lived species. Rudd and Thorson (2017) also found that LIME is imprecise and biased for long-lived species. The model is trying to track cohorts through the length data to estimate recruitment deviations and this is likely difficult for long-lived species when time series of length data are short or much of the population is found near the asymptotic size. However, for the short and medium-lived species LIME ~~worked better~~performance improved.

LBSPR was also highly biased for long-lived species. Hordyk et al. (2015a) suggested that LBSPR relies on detecting the signal of fishing mortality in the right-hand side of the length composition. Consequently, fishing is not likely to have a visible impact on the length composition until fishing mortality is very high. This is why LBSPR performed better for long-lived species under harvest scenario 1 when final depletion centered at 0.2.

LIME and LBSPR can be used in conjunction as a diagnostic tool to see if variations in recruitment or fishing mortality can be predicted by LIME and how the results of the multiple assessment types would vary in light of those possibly violated assumptions. Pons et al. (2019) used both length assessment methods to estimate current stock status of small scombrids in the Atlantic Ocean. For some highly variable species, both models gave very different stock status outcomes (i.e. bonito in the Atlantic Ocean). However, LIME can be used to complement LBSPR analyzing both process and observation uncertainty.

~~For long-lived species, recruitment variability does not affect the length composition of the catch as much as for short lived species. This is why LBSPR performed pretty well for long-lived species and it was highly biased for short-lived species. LIME however, performed better for short-lived species than LBSPR being able to capture changes in the length composition due to recruitment variability.~~ In general, all catch-based and length-based methods seems to perform worse for ~~the faster~~long-lived life history types. The length of the time series for the long-lived canary rockfish is probably too short in comparison to its maximum age of 64 years to capture the true dynamic of the population and the response to different harvest rates. For an 80 years' time series of catch and length data, and under a depletion level of 0.4, Pons (2018) found a better performance for LBSPR in comparison to LIME and for SSS among different catch-based methods.

The present study does not look at parameter misspecification, but correct specification (unbiased) in the life histories parameters and known catch histories. With that level of information, the length-based models ~~like LBSPR~~ showed sometimes better performance than ~~the~~ some catch-based models, as the latter were more sensitive to the catch history scenarios and depletion levels. Length-based models are no sensitive to catch trends ~~, and the length based methods were~~ since they are able to integrate the catch scenarios into the length compositions, but in some cases LIME performed better when the stocks are highly depleted.

~~*Conclusions*~~*Recommendations*

For unassessed fisheries where data are limited, but reconstructing time series of catch is possible, catch-based methods SSS or DBSRA ~~provided the most~~performed most reliably in the tests here, but this performance hinges on the proper specification of the input parameters such as~~e outputs to for management~~ stock depletion~~. However, to apply SSS and DBSRA, not only catch data is needed, these methods also require extensive prior information, such as growth, maturity~~, productivity, maturity and possibly even growth parameters ~~F_{MSY}/M (Zhou *et al.*, 2012)~~ ~~and B_{MSY}/B_0 (Thorson *et al.*, 2012) parameters~~. Misspecification in any of these parameters however, is likely to occur for data-limited stocks. ~~Some~~W~~w~~hile -meta-analyse~~i~~s ~~studies such as Zhou *et al.* (2012) and Thorson *et al.* (2012) can~~may offer some starting values for certain parameters (e.g., Myers 2001, Thorson *et al.* 2012, and Zhou *et al.* 2012), other inputs remain difficult to specify (Chrysafi and Cope *in review*)~~give you prior estimates of F_{MSY}/M and B_{MSY}/B_0 for different group of species, respectively~~. ~~When this prior information is not available, SSCOM could be use with a good prior for depletion. CMSY could also be considered, but with caution, because it has been proved here that they can be highly biased and influenced by catch trends and uncertainty in *K*.~~

For fisheries ~~that are still developing,~~ where the time series of catch are unavailable, getting length-composition data could give a good approximation of the status of the stock, in particular for short and medium~~- to long-~~lived species. It has been shown here that, in some cases, that length-based models can give the same or less biased estimates of exploitation status than catch-based models~~, though it must be recognized that these estimates of fishing mortality are aggregated over all years of the fishery~~.

Giving recommendations on which models can be applied to estimate exploitation intensity in different fisheries is challenging because it is very dependent on data availability, trends in fishing intensity and the biology of the species. I~~t is recommended~~f possible, simulation studies ~~to~~ testing different data-~~-~~limited ~~assessments~~methods with OMs based on the focus species ~~considered~~ and the dynamic of the fishery can greatly inform~~before determining~~ which method~~one is most~~could be most appropriate. Likewise, decision support tool such as FishPath (Dowling et al. 2015) can also help one weight the input requirements and assumptions to help identify the most appropriate methods given data and life history. ~~However, b~~Based on the~~our~~ OMs used in this study, we can conclude that for short and medium-~~-~~lived species when only catch data is available, both SSS and DBSRA ~~should be~~could be considered and results combined ~~used as complementary methods~~. In some scenarios they perform similarly, unbiased and precise, and in others when SSS tended to overestimates harvest rates and DBSRA tended to underestimates it. An ensemble of both appears to be a good option for short lives species. When only length data is available LIME showed to be less biased being able to capture changes in recruitment and fishing mortality better than LBSPR for short-lived species.

Commented [CJ1]: Looks like autocorrect gone wild. Not sure what is should be.

For medium-lived species, w~~when only catch data is available also both SSS and DBSRA can be use. W~~hen only length data is available, LBSPR and LIME can be applied as complementary models, but LBSPR is more appropriate when fishing intensity is decreasing and LIME when fishing intensity is increasing.

For long-lived species it is necessary to have longer time series of data to draw more conclusions. However, Pons (2018) recommended SSS and LBSPR when long time series (i.e.

80 years) of data are available for a species that lives more than 60 to see changes in fishing intensity.

If both, catch and length data are available, models that integrates both data should be considered. LIME for example allows for the inclusion of catch data as well as an index of abundance. Moreover, integrated assessment models like SS can also be implemented that uses catch as well as length information.

Finally, for the scenarios analyzed here and for the specific life-histories considered it is not recommended to use the method CMSY to estimate exploitation rates even with a good estimate~~pretty good estimation~~ of stock depletion. However, this methods, as it was created for originally, produces unbiased estimations of MSY, in particular for short and medium-lived and highly depleted species (Table S1).

Future directions

There has been an emerging field of ~~catch methods and~~ ensemble~~assemble~~ of catch-based methods to ~~estimate~~ improve ~~global~~ stock status estimates (Costello *et al.*, 2012; Anderson *et al.*, 2017; Rosenberg *et al.*, 2017). Combining estimates from different methods in a consistent reproducible manner may provide more stability in the advice for managers. The super-ensemble method based on catch-only methods ~~published by (~~Anderson *et al.* ~~(~~2017) allows for weighting individual models based on their accuracy. In their study, some models had different assumptions about uncertainty and the dynamics of fishing effort, but all assumed the same population dynamics. A new super-ensemble method that includes models that also assume different population dynamics could be developed in the future based on our results. It is important, however, that the behavior of the models in the ensemble are well understood, i.e.

their bias, precision and convergence properties. ~~Combining estimates from different methods in a consistent reproducible manner may provide more stability in the advice for managers.~~

Dowling et. al, (2019) in a review of data limited methods noted the dangers in the indiscriminate use of generic methods and recommended obtaining better data, using care in acknowledging and interpreting uncertainties, developing harvest strategies that are robust to the higher levels of uncertainty and tailoring them to the specific species' and fisheries' data and context. Therefore, ~~The best performing~~different methods ~~can then~~ should be tested using a management strategy evaluation (MSE) to specify Management Procedures (MP) that can help ensure robust and sustainable fisheries management. Where a MP is the combination of pre-defined data, together with an algorithm to which such data are input to provide a value for a management control measure. This must include evaluation of the robustness of the methods to misspecification of input parameters and the benefits of improving knowledge on them.

This study provides a way of conditioning OMs and generating pseudo data for use by the MP. The importance of considering assessment methods as part of a MP is that a method that provides biased estimates with high precision may be better for setting management regulations than an unbiased but imprecise estimator. Also if a method only provides estimates of exploitation level or MSY then management controls may be different, i.e. based on a total allowable catch (TAC) or effort. MSE also allows another approach to exploring ~~broad range of~~ uncertainty, since traditional stock assessment and advice based upon it, mainly considers measurement and process error when uncertainty about the actual dynamics has a larger impact on achieving management objectives (Punt, 2008).

MSE can also be used to identify the impact of uncertainty, by evaluating the robustness of stock assessment methods and advice to the misspecification of input parameters. For example final depletion is a key parameter for catch based methods, however, it is extremely difficult to estimate in data limited situations. Zhou *et al*. (2017) used the RAM Legacy database to derive priors for depletion for data-poor stock assessment methods, the framework developed could be used to evaluating the benefits of improving knowledge on depletion using such an approach.

Acknowledgments

We want to acknowledge Trevor Branch, Victor Restrepo, Dave Fluharty, Jim Ianelli and Ray Hilborn for their significant comments in the early stages of this study ~~The Walton Family Foundation generously funded MP. LK participation was possible due to the MyDas project, funded by the Irish Exchequer and the European Maritime and Fisheries Fund (EMFF)~~. ~~In addition, w~~We ~~would like to~~also want to acknowledge Merrill Rudd, Chantel Wetzel, Ian Taylor and James Thorson for their help with SS, r4ss, and other packages and models used in this study.

Funding

We acknowledge funding by the Walton Family Foundation (#2018-51) and International Seafood Sustainability Foundation (ISSF) for MP. LK's involvement was funded through the MyDas project under Marine Biodiversity Scheme which is financed by the Irish government and the European Maritime & Fisheries Fund as part of the EMFF Operational Programme for 2014-2020.

REFERENCES

Anderson, S. C., Afflerbach, Jamie Cooper, A. B., Dickey-Collas, M., Jensen, O. P., Kleisner, K. M., Longo, C., Osio, G. C., *et al.* 2016. datalimited: Stock Assessment Methods for Data-limited Fisheries. R package version 0.1.0. https://github.com/datalimited/datalimited.

Anderson, S. C., Cooper, A. B., Jensen, O. P., Minto, C., Thorson, J. T., Walsh, J. C., Afflerbach, J., *et al.* 2017. Improving estimates of population status and trend with superensemble models. Fish and Fisheries, 18: 732–741.

Beverton, R. J. H., and Holt, S. J. 1957. On the Dynamics of Exploited Fish Populations. Fishery Investigations (Great Britain, Ministry of Agriculture, Fisheries, and Food), London. 536 pp.

Carruthers, T. R., Walters, C. J., and McAllister, M. K. 2012. Evaluating methods that classify fisheries stock status using only fisheries catch data. Fisheries Research, 119: 66–79.

Carruthers, T. R., Punt, A. E., Walters, C. J., MacCall, A., McAllister, M. K., Dick, E. J., and Cope, J. 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fisheries Research, 153: 48–68.

Carruthers, T. R., Kell, L. T., Butterworth, D. D. S., Maunder, M. N., Geromont, H. F., Walters, C., McAllister, M. K., *et al.* 2016. Performance review of simple management procedures. ICES Journal of Marine Science, 73: 464–482.

Chrysafi, A., and J.M. Cope. in review. Testing methods of determining relative stock abundance priors when setting catch recommendations using data-limited approaches. Fisheries Research.

722 Chrysafi, A., and Kuparinen, A. 2016. Assessing abundance of populations with limited data: Lessons learned from data-poor fisheries stock assessment. Environmental Reviews, 1: 1–44.

725 Cope, J. M. 2008. Issues and Advances in Data-Limited Stock Assessment: Experimentation through Simulation. Dissertation - Doctor of Philosophy degree: 1–235.

727 Cope, J. M. 2013. Implementing a statistical catch-at-age model (Stock Synthesis) as a tool for deriving overfishing limits in data-limited situations. Fisheries Research, 142: 3–14.

729 Costello, C., Ovando, D., Hilborn, R., Gaines, S. D., Deschenes, O., and Lester, S. E. 2012. Status and solutions for the world's unassessed fisheries. Science (New York, N.Y.), 338: 517–20.

732 Crone, P. R., and Hill, K. T. 2015. Pacific mackerel (*Scomber japonicus*) stock assessment for USA management in the 2015-16 fishing year. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, Oregon 97220, USA: 131.

735 Dick, E. J., and MacCall, A. D. 2011. Depletion-Based Stock Reduction Analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. Fisheries Research, 110: 331–341.

738 Dowling, N. A., Dichmont, C. M., Haddon, M., Smith, D. C., Smith, A. D. M., and Sainsbury, K. 2015. Empirical harvest strategies for data-poor fisheries: A review of the literature. Fisheries Research, 171: 141–153.

741 Dowling, N., Wilson, J., Rudd, M., Babcock, E., Caillaux, M., Cope, J., Dougherty, D., Fujita,

742 R., Gedamke, T., Gleason, M., Guttierrez, M., Hordyk, A., Maina, G., Mous, P., Ovando, D., Parma, A., Prince, J., Revenga, C., Rude, J., Szuwalski, C., Valencia, S., and Victor, S., 2016. FishPath: A Decision Support System for Assessing and Managing Data- and Capacity- Limited Fisheries, in: Quinn II, T., Armstrong, J., Baker, M., Heifetz, J., Witherell, D. (Eds.), Assessing and Managing Data-Limited Fish Stocks. Alaska Sea Grant, University of Alaska Fairbansk. https://doi.org/10.4027/amdlfs.2016.03

748 Dowling, N. A., Smith, A. D. M., Smith, D. C., Parma, A. M., Dichmont, C. M., Sainsbury K., Wilson, J. R., Dougherty, D. T. and Cope, J. M. 2019. Generic solutions for data-limited fishery assessments are not so simple. Fish and Fisheries, 20: 174–188.

751 Francis, C. 2012. Reply to 'The reliability of estimates of natural mortality from stock assessment models'. Fisheries Research, 119–120: 133–134.

753 Froese, R., Demirel, N., Coro, G., Kleisner, K. M., and Winker, H. 2017. Estimating fisheries reference points from catch and resilience. Fish and Fisheries, 18: 506–526.

755 Hoeting, J.A., Madigan, D., Raftery, A.E., and Volinsky, C.T. 1999. Bayesian model averaging: a tutorial. Statatistical Science, 14: 382–401.

757 Hordyk, A., Ono, K., Valencia, S., Loneragan, N., and Prince, J. 2015a. A novel length-based empirical estimation method of spawning potential ratio (SPR), and tests of its performance, for small-scale, data-poor fisheries, 72: 217–231.

760 Hordyk, A. R., Loneragan, N. R., and Prince, J. D. 2015b. An evaluation of an iterative harvest strategy for data-poor fisheries using the length-based spawning potential ratio assessment methodology. Fisheries Research, 171: 20–32.

ICCAT. 2014. Report of the 2013 ICCAT North and South Atlantic Albacore stock assessment meeting. Collective Volume of Scientific Papers ICCAT, 70: 830–995.

Kell, L. T., Kimoto, A., and Kitakado, T. 2016. Evaluation of the prediction skill of stock assessment using hindcasting. Fisheries Research, 183: 119–127.

Kimura, D.K., and Tagart, J.V. 1982. Stock reduction analysis, another solution to the catch equations. Canadian Journal of Fisheries and Aquatic Sciences. 39: 1467–1472.

Kimura, D.K., Balsiger, J.W., and Ito, D.H. 1984. Generalized stock reduction analysis. Canadian Journal of Fisheries and Aquatic Sciences. 41: 1325–1333.

Martell, S., and Froese, R. 2013. A simple method for estimating MSY from catch and resilience. Fish and Fisheries, 14: 504–514.

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.

Methot, R. D., Wetzel, C. R., and Taylor, I. 2018. Stock Synthesis User Manual. V3.30.10: 1–186.

Myers, R.A., 2001. Stock and recruitment: generalizations about maximum reproductive rate, density dependence, and variability using meta-analytic approaches. ICES Journal of Marine Science 58: 937.

Nelson, G. A. 2017. fishmethods: Fishery Science Methods and Models in R. R package version 1.10-3.: https://CRAN.R-project.org/package=fishmethods.

Pons, M. 2018. Stock Status and Management in Tuna Fisheries: From Data-rich to Data-poor.

University of Washington, ProQuest Dissertations Publishing, 2018: 10846859. 186 pp. https://search.proquest.com/docview/2126689812? accountid¼14784 (last accessed 15 November 2018).

Pons, M., Kell, L., Rudd, M. B., Cope, J. M., and Lucena Fredou, F. 2019. Performance of length-based data-limited methods in a multifleet context: application to small tunas, mackerels, and bonitos in the Atlantic Ocean. ICES Journal of Marine Science, doi:10.1093/icesjms/fsz004

Punt, A. 2008. Refocusing stock assessment in support of policy evaluation. Fisheries for Global Welfare and Environment, 5th World Fisheries Congress: 139–152.

Rosenberg, A. A., Kleisner, K. M., Afflerbach, J., Anderson, S. C., Dickey-Collas, M., Cooper, A. B., Fogarty, M. J., *et al.* 2017. Applying a New Ensemble Approach to Estimating Stock Status of Marine Fisheries Around the World. Conservation Letters, 00: 1–9.

Rudd, M. B., and Thorson, J. T. 2018. Accounting for variable recruitment and fishing mortality in length-based stock assessments for data-limited fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 75: 1019–1035.

Thorson, J. T., Cope, J. M., Branch, T. A., Jensen, O. P., and Walters, C. J. 2012. Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. Canadian Journal of Fisheries and Aquatic Sciences, 69: 1556–1568.

Thorson, J. T., Minto, C., Minte-Vera, C. V., Kleisner, K. M., Longo, C., and Jacobson, L. 2013. A new role for effort dynamics in the theory of harvested populations and data-poor stock assessment. Canadian Journal of Fisheries and Aquatic Sciences, 70: 1829–1844.

804 Thorson, J. T., and Cope, J. M. 2015. Catch curve stock-reduction analysis: An alternative solution to the catch equations. Fisheries Research, 171: 33–41.

806 Thorson, J. T., and Wetzel, C. 2015. The status of canary rockfish (Sebastes pinniger) in the California Current in 2015. Pacific Fishery Management Council: 22.

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

810 Wetzel, C. R., and Punt, A. E. 2011. Model performance for the determination of appropriate harvest levels in the case of data-poor stocks. Fisheries Research, 110: 342–355.

812 Wetzel, C. R., and Punt, A. E. 2015. Evaluating the performance of data-moderate and catch-only assessment methods for U.S. west coast groundfish. Fisheries Research, 171: 170–187.

814 Zhou, S., Yin, S., Thorson, J. T., Smith, A. D. M., Fuller, M., and Walters, C. J. 2012. Linking fishing mortality reference points to life history traits: an empirical study. Canadian Journal of Fisheries and Aquatic Sciences, 69: 1292–1301.

817 Zhou, S., Punt, A.E., Ye, Y., Ellis, N., Dichmont, C.M., Haddon, M., Smith, D.C. and Smith, A.D. 2017. Estimating stock depletion level from patterns of catch history. Fish and Fisheries. 18: 742–751.

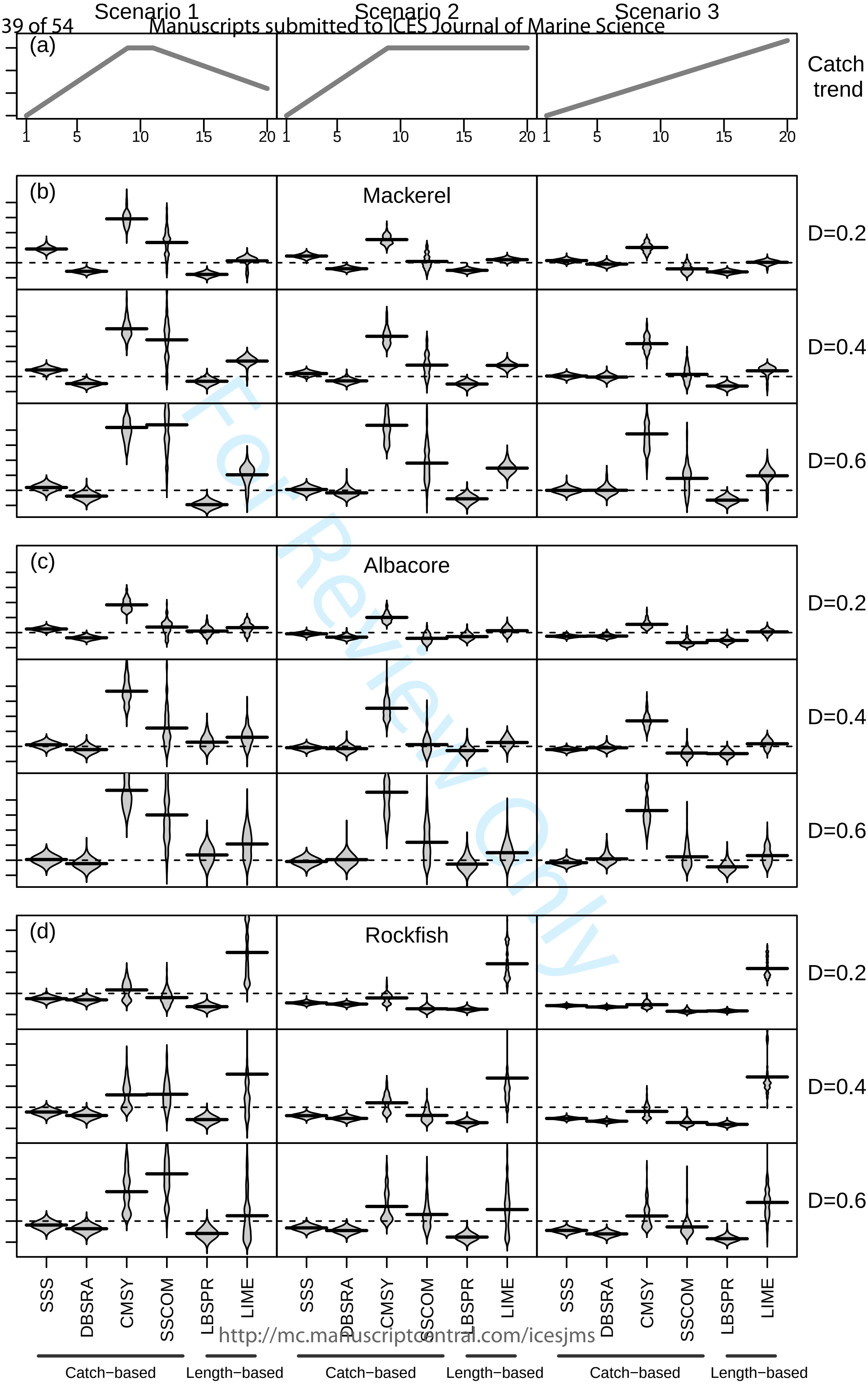
FIGURE CAPTIONS

~~Figure 1. Time series of catch, harvest rate (*U*), biomass (*B*) and spawning potential ratio (SPR) for each simulated Pacific Chub Mackerel population for the three harvest rate Scenarios tested. The color lines represent the median value for all runs.~~

~~Figure 2. Time series of catch, harvest rate (*U*), biomass (*B*) and spawning potential ratio (SPR) for each simulated North Atlantic Albacore population for the three harvest rate Scenarios tested. The color lines represent the median value for all runs.~~

~~Figure 3. Time series of catch, harvest rate (*U*), biomass (*B*) and spawning potential ratio (SPR) for each simulated Canary rockfish population for the three harvest rate Scenarios tested. The color lines represent the median value for all runs.~~

Figure ~~4~~1. Exploitation rate relative error for all the catch-based ~~(light colors)~~ and length-based ~~(dark colors)~~ models considered under the ~~three harvest~~ different harvest and depletion scenarios for the ~~three~~ life histories~~species~~. ~~First row: Scenario 1 – ramp shape harvest rate. Second row: Scenario 2 – constant harvest rate. Third row: Scenario 3 – increasing harvest rate.~~



Scenario 1
Scenario 2
Scenario 3
(a)
Catch trend
(b)
Mackerel
(c)
Albacore
(d)
Rockfish
D=0.2
D=0.4
D=0.6
SSS
DBSRA
CMSY
SSCOM
LBSPR
LIME
Catch−based
Length−based

Table 1. Life history information and priors for the three species used in the study. Notation: *Lognormal* (μ, σ²); Uniform *U* (a, b). Priors for *K* were Uniform between the maximum catch in the time series and ~~50~~ 100 times the maximum catch. ~~* Generation length (GL) calculated by Stock Synthesis. * For CMSY the depletion priors were Uniform centered in the true value with a minimum of *true - 0.1* and a maximum of *true + 0.1*.~~

Operating model inputs	**Symbol**	**Pacific Chub Mackerel**	**Albacore tuna**	**Canary Rockfish**
Maximum age	Age_{max}	12	15	64
~~Generation length~~Age at 50% maturity (years) ~~*~~	~~GL~~A_{mat}	~~2~~3	5	16
Length were 50% of the fish are mature (FL cm)	L_{50}	29	90	55
Length were 95% of the fish are mature (FL cm)	L_{95}	34	100	57
Length-weight scaling parameter	a	2.73×10^{-6}	1.34×10^{-5}	1.80×10^{-5}
Length-weight allometric parameter	b	3.444	3.107	3.094
Von Bertalanffy Brody growth coefficient (1/years)	k	0.40	0.21	0.14
Von Bertalanffy asymptotic length (cm)	L_∞	38.2	122.2	60.0
Theoretical age at length=0	t_0	-0.6	-1.3	-1.9
Coefficient of variation length at age for all ages	CVL	0.1	0.1	0.1
Natural mortality (1/years)	M	0.60	0.30	0.05
~~Relationship between *M* and *k*~~	~~M/k~~	~~1.50~~	~~1.40~~	~~0.35~~
Steepness	h	0.5	0.9	0.8
Selectivity at 50% (cm)	S_{50}	~~18~~25	60	~~42~~45
Selectivity at 95% (cm)	S_{95}	~~25~~30	75	~~47~~50
~~Depletion~~	~~XB_0~~	~~0.4~~	~~0.4~~	~~0.4~~
Survey or depletion standard error	σ_S	0.01	0.01	0.01
Observation error in catch	σ_C	0.1	0.1	0.1
Recruitment variations	σ_R	0.3	0.4	0.5
Estimation models prior distributions				
Depletion (used for all catch-based models)	XB_0	*Lognormal* (true, 0.1) *	*Lognormal* (true, 0.1) *	*Lognormal* (true, 0.1) *
Carrying capacity (used for CMSY and SSCOM)	K	*U* (max(catch), max(catch)~~x50~~x100)	*U* (max(catch), max(catch)~~x50~~x100)	*U* (max(catch), max(catch)~~x50~~x100)
Population rate of increase (used for CMSY and SSCOM)	r	*U* (0.8, 1.2)	*U* (0.2, 0.6)	*U* (0.05, 0.4)
Steepness (used for SSS)	h	*Normal* (0.5,0.1)	*Normal* (0.9,0.1)	*Normal* (0.8,0.1)
Vulnerability (used for DBSRA)	F_{MSY}/M	*U* (0, 2)	*U* (0, 2)	*U* (0, 2)

Compensation (used for DBSRA)	B_{MSY}/B_0	U (0, 1)	U (0, 1)	U (0, 1)

* For CMSY the depletion priors were Uniform centered in the true value with a minimum of *true - 0.1* and a maximum of *true + 0.1*.

Table 2 Bias measured as median absolute relative error for the catch-based models by harvest ra

Models	SSS (average absolute RE = 0.26)				DBSRA (average absolute RE = C		
Scenarios	Life-history	Harvest trend	Final depletion	RE	Life-history	Harvest trend	Final depletion
1	Short-lived	Scenario 3	0.6	-0.02	Medium-lived	Scenario 3	0.6
2	Short-lived	Scenario 3	0.4	0.03	Medium-lived	Scenario 2	0.6
3	Medium-lived	Scenario 1	0.6	0.04	Short-lived	Scenario 3	0.4
4	Short-lived	Scenario 2	0.6	0.06	Short-lived	Scenario 3	0.6
5	Medium-lived	Scenario 2	0.2	-0.08	Medium-lived	Scenario 3	0.4
6	Medium-lived	Scenario 2	0.6	-0.08	Short-lived	Scenario 3	0.2
7	Medium-lived	Scenario 2	0.4	-0.09	Medium-lived	Scenario 2	0.4
8	Medium-lived	Scenario 1	0.4	0.11	Short-lived	Scenario 2	0.6
9	Short-lived	Scenario 3	0.2	0.14	Medium-lived	Scenario 1	0.4
10	Medium-lived	Scenario 3	0.6	-0.17	Medium-lived	Scenario 3	0.2
11	Short-lived	Scenario 1	0.6	0.19	Medium-lived	Scenario 1	0.6
12	Long-lived	Scenario 1	0.6	-0.19	Short-lived	Scenario 2	0.4
13	Short-lived	Scenario 2	0.4	0.20	Long-lived	Scenario 1	0.2
14	Medium-lived	Scenario 3	0.4	-0.21	Medium-lived	Scenario 2	0.2
15	Long-lived	Scenario 1	0.4	-0.23	Long-lived	Scenario 1	0.6
16	Medium-lived	Scenario 1	0.2	0.23	Medium-lived	Scenario 1	0.2
17	Medium-lived	Scenario 3	0.2	-0.24	Long-lived	Scenario 1	0.4
18	Long-lived	Scenario 1	0.2	-0.25	Short-lived	Scenario 2	0.2
19	Long-lived	Scenario 2	0.6	-0.33	Short-lived	Scenario 1	0.6
20	Long-lived	Scenario 2	0.4	-0.40	Long-lived	Scenario 2	0.6
21	Short-lived	Scenario 1	0.4	0.44	Short-lived	Scenario 1	0.4
22	Short-lived	Scenario 2	0.2	0.44	Long-lived	Scenario 2	0.2
23	Long-lived	Scenario 2	0.2	-0.45	Long-lived	Scenario 2	0.4
24	Long-lived	Scenario 3	0.6	-0.45	Short-lived	Scenario 1	0.2
25	Long-lived	Scenario 3	0.4	-0.54	Long-lived	Scenario 3	0.6
26	Long-lived	Scenario 3	0.2	-0.58	Long-lived	Scenario 3	0.2
27	Short-lived	Scenario 1	0.2	0.90	Long-lived	Scenario 3	0.4

ates scenarios, species and final depletion levels.

).32)	SSCOM (average absolute RE = 0.93)				CMSY (average absolute RE = 1			
RE	Life-history	Harvest trend	Final depletion	RE	Life-history	Harvest trend	Final depletion
0.01	Short-lived	Scenario 2	0.2	0.00	Long-lived	Scenario 3	0.6
-0.01	Short-lived	Scenario 3	0.4	0.06	Long-lived	Scenario 2	0.2
-0.05	Medium-lived	Scenario 2	0.4	0.06	Long-lived	Scenario 1	0.2
-0.06	Medium-lived	Scenario 3	0.6	0.07	Long-lived	Scenario 3	0.4
-0.09	Long-lived	Scenario 2	0.6	0.19	Long-lived	Scenario 2	0.4
-0.09	Long-lived	Scenario 1	0.2	-0.25	Long-lived	Scenario 2	0.6
-0.16	Long-lived	Scenario 3	0.6	-0.37	Medium-lived	Scenario 3	0.2
-0.21	Medium-lived	Scenario 1	0.2	0.40	Long-lived	Scenario 3	0.2
-0.24	Medium-lived	Scenario 2	0.2	-0.41	Long-lived	Scenario 1	0.4
-0.24	Short-lived	Scenario 3	0.2	-0.43	Medium-lived	Scenario 2	0.2
-0.25	Long-lived	Scenario 2	0.4	-0.44	Short-lived	Scenario 3	0.2
-0.29	Medium-lived	Scenario 3	0.4	-0.47	Long-lived	Scenario 1	0.6
-0.31	Long-lived	Scenario 1	0.4	0.56	Short-lived	Scenario 2	0.2
-0.32	Short-lived	Scenario 2	0.4	0.70	Medium-lived	Scenario 3	0.4
-0.36	Medium-lived	Scenario 3	0.2	-0.72	Medium-lived	Scenario 1	0.2
-0.36	Long-lived	Scenario 2	0.2	-0.74	Short-lived	Scenario 3	0.4
-0.40	Long-lived	Scenario 3	0.4	-0.75	Medium-lived	Scenario 2	0.4
-0.40	Medium-lived	Scenario 1	0.4	0.84	Short-lived	Scenario 2	0.4
-0.41	Long-lived	Scenario 3	0.2	-0.85	Short-lived	Scenario 1	0.2
-0.44	Short-lived	Scenario 3	0.6	0.88	Medium-lived	Scenario 3	0.6
-0.49	Medium-lived	Scenario 2	0.6	0.96	Short-lived	Scenario 1	0.4
-0.51	Short-lived	Scenario 1	0.2	1.49	Short-lived	Scenario 3	0.6
-0.54	Short-lived	Scenario 2	0.6	1.69	Medium-lived	Scenario 1	0.4
-0.58	Long-lived	Scenario 1	0.6	2.16	Short-lived	Scenario 1	0.6
-0.61	Short-lived	Scenario 1	0.4	2.63	Short-lived	Scenario 2	0.6
-0.64	Medium-lived	Scenario 1	0.6	2.70	Medium-lived	Scenario 1	0.6
-0.66	Short-lived	Scenario 1	0.6	4.31	Medium-lived	Scenario 2	0.6

.89)
RE
0.07
-0.18
0.21
-0.23
0.27
0.33
0.49
-0.53
0.58
0.98
1.00
1.26
1.48
1.66
1.78
2.18
2.38
2.62
2.90
2.99
3.07
3.48
3.63
4.05
4.13
4.30
4.33

Table 3	Bias measured as median absolute relative error for the length-based models by harves						
Models	LBSPR (average absolute RE = 0.56)				LIME (average absolute RE = 0.		
Scenarios	Life-history	Harvest trend	Final depletion	RE	Life-history	Harvest trend	Final depletion
1	Medium-lived	Scenario 1	0.2	0.09	Short-lived	Scenario 3	0.2
2	Medium-lived	Scenario 1	0.4	0.23	Long-lived	Scenario 1	0.6
3	Medium-lived	Scenario 2	0.4	-0.28	Medium-lived	Scenario 3	0.2
4	Medium-lived	Scenario 2	0.6	-0.28	Medium-lived	Scenario 2	0.2
5	Medium-lived	Scenario 2	0.2	-0.30	Medium-lived	Scenario 3	0.4
6	Short-lived	Scenario 1	0.4	-0.31	Short-lived	Scenario 2	0.2
7	Medium-lived	Scenario 1	0.6	0.40	Short-lived	Scenario 1	0.2
8	Medium-lived	Scenario 3	0.4	-0.49	Medium-lived	Scenario 3	0.6
9	Short-lived	Scenario 2	0.4	-0.50	Medium-lived	Scenario 2	0.4
10	Medium-lived	Scenario 3	0.6	-0.50	Medium-lived	Scenario 1	0.2
11	Short-lived	Scenario 2	0.2	-0.51	Medium-lived	Scenario 2	0.6
12	Medium-lived	Scenario 3	0.2	-0.52	Short-lived	Scenario 3	0.4
13	Short-lived	Scenario 2	0.6	-0.56	Medium-lived	Scenario 1	0.4
14	Long-lived	Scenario 1	0.4	-0.60	Long-lived	Scenario 2	0.6
15	Short-lived	Scenario 3	0.2	-0.61	Long-lived	Scenario 1	0.4
16	Long-lived	Scenario 1	0.6	-0.61	Short-lived	Scenario 2	0.4
17	Long-lived	Scenario 1	0.2	-0.63	Long-lived	Scenario 3	0.6
18	Short-lived	Scenario 3	0.4	-0.65	Long-lived	Scenario 2	0.4
19	Short-lived	Scenario 3	0.6	-0.67	Long-lived	Scenario 3	0.2
20	Long-lived	Scenario 2	0.4	-0.74	Short-lived	Scenario 1	0.4
21	Long-lived	Scenario 2	0.2	-0.75	Short-lived	Scenario 3	0.6
22	Long-lived	Scenario 2	0.6	-0.77	Medium-lived	Scenario 1	0.6
23	Short-lived	Scenario 1	0.2	-0.78	Long-lived	Scenario 2	0.2
24	Long-lived	Scenario 3	0.4	-0.81	Long-lived	Scenario 3	0.4
25	Long-lived	Scenario 3	0.2	-0.83	Short-lived	Scenario 1	0.6
26	Long-lived	Scenario 3	0.6	-0.86	Short-lived	Scenario 2	0.6
27	Short-lived	Scenario 1	0.6	-0.99	Long-lived	Scenario 1	0.2

t rates scen
.65)
RE
0.05
-0.05
0.07
0.13
0.18
0.20
0.22
0.24
0.27
0.35
0.36
0.47
0.53
0.55
0.58
0.73
0.89
0.91
0.98
1.06
1.08
1.09
1.11
1.15
1.19
1.45
1.76

Supplementary information

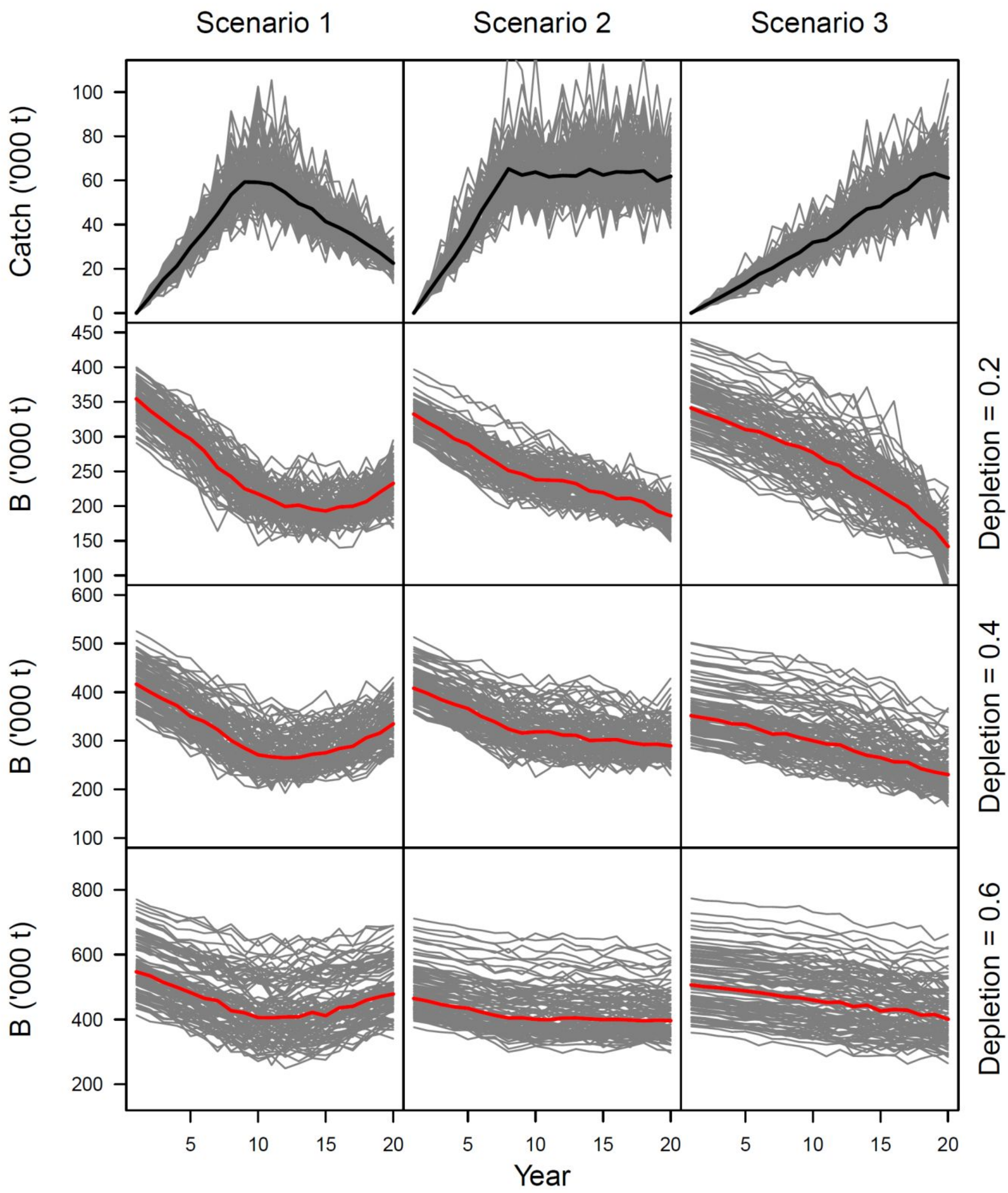


Figure S1. Time series of catch and biomass (*B*) for each simulated short-lived chub mackerel population for the three harvest rate and depletion levels scenarios tested. The color lines represent the median value for all runs.

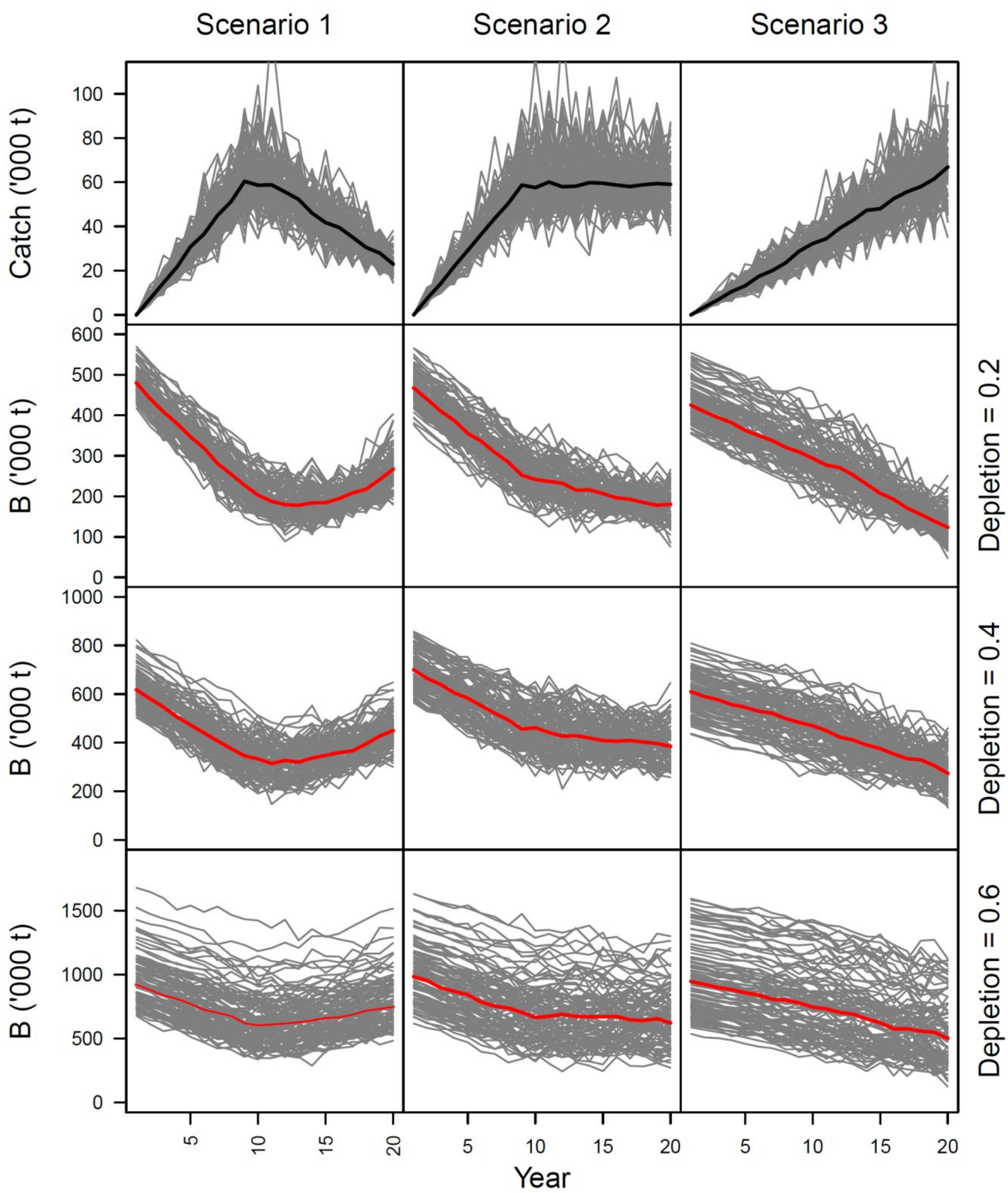


Figure S2. Time series of catch and biomass (*B*) for each simulated medium-lived albacore tuna population for the three harvest rate and depletion levels scenarios tested. The color lines represent the median value for all runs.

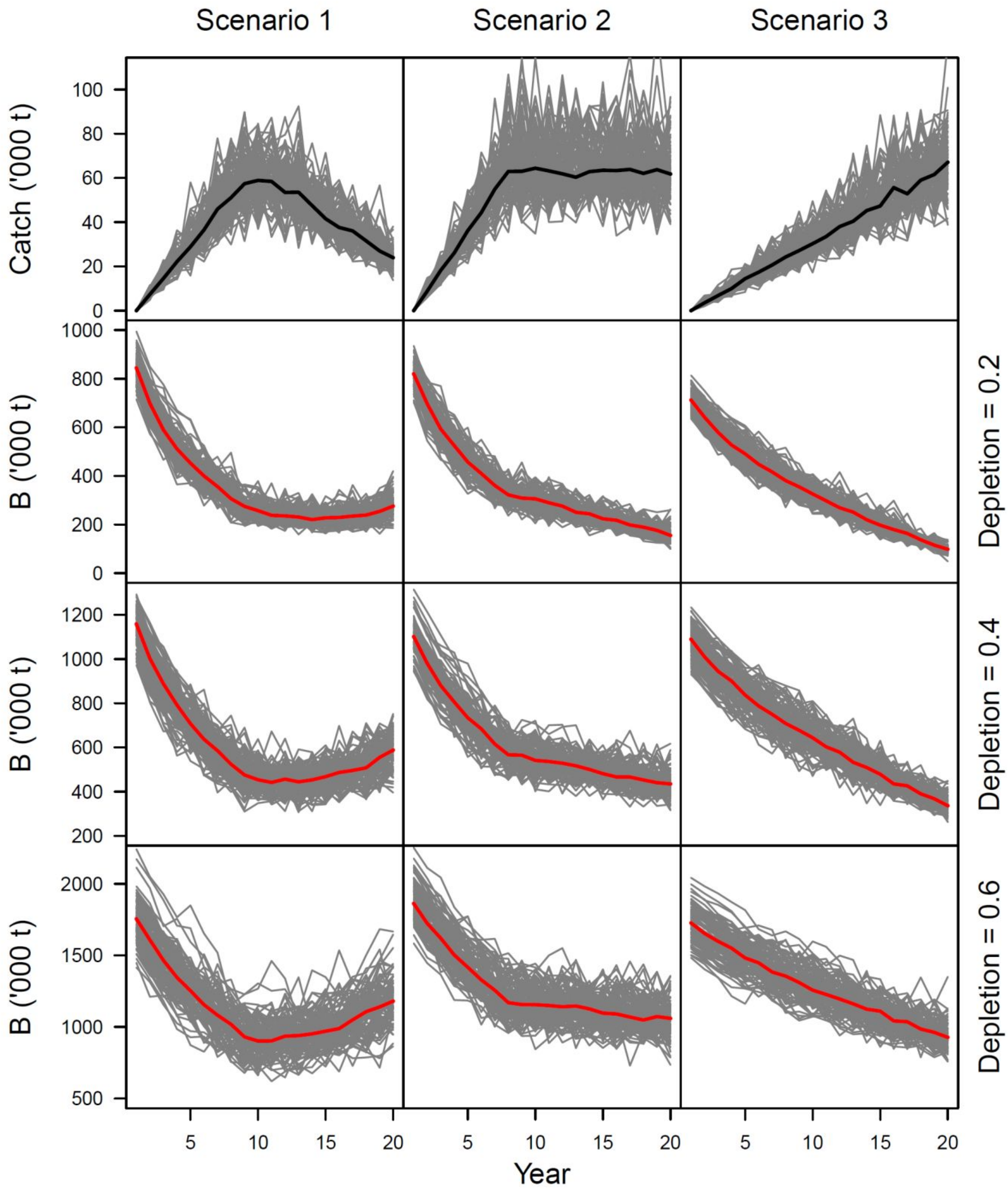


Figure S3. Time series of catch and biomass (*B*) for each simulated long-lived canary rockfish population for the three harvest rate and depletion levels scenarios tested. The color lines represent the median value for all runs.

Table S1. Mean and standard deviation of relative error (RE) between the maximum sustainable yield (MSY) estimated by the operating model (OM) and the MSY estimated by the CMSY method. Values are proportions. Positive values mean that the MSY was overestimated and negative values that it was underestimated.

Species	Final depletion levels / Harvest Scenario	Mean			Standard deviation		
		0.2	0.4	0.6	0.2	0.4	0.6
Mackerel	Scenario 1	0.37	0.18	-0.13	0.08	0.09	0.15
	Scenario 2	0.28	0.03	-0.22	0.08	0.09	0.13
	Scenario 3	0.33	0.05	-0.26	0.11	0.11	0.14
Albacore	Scenario 1	-0.07	-0.18	-0.46	0.02	0.03	0.02
	Scenario 2	-0.25	-0.42	-0.42	0.02	0.02	0.02
	Scenario 3	-0.12	-0.24	-0.43	0.02	0.02	0.02
Rockfish	Scenario 1	1.61	1.13	0.80	0.27	0.31	0.32
	Scenario 2	1.67	1.27	0.86	0.26	0.33	0.34
	Scenario 3	1.77	1.29	0.94	0.33	0.45	0.45