Fisheries Research 223 (2020) 105452

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Blood from a stone: Performance of catch-only methods in estimating stock biomass status

Christopher M. Free[a,b,*], Olaf P. Jensen[a], Sean C. Anderson[c], Nicolas L. Gutierrez[d], Kristin M. Kleisner[e], Catherine Longo[f], Cóilín Minto[g], Giacomo Chato Osio[h], Jessica C. Walsh[i,j]

[a] Department of Marine & Coastal Sciences, Rutgers University, New Brunswick, NJ, USA
[b] Bren School of Environmental Science and Management, University of California, Santa Barbara, Santa Barbara, CA, USA
[c] Pacific Biological Station, Fisheries & Oceans Canada, Nanaimo, BC, Canada
[d] Fisheries & Aquaculture Department, Food & Agriculture Organization, Rome, Italy
[e] Environmental Defense Fund, Boston, MA, USA
[f] Marine Stewardship Council, London, UK
[g] Marine & Freshwater Research Centre, Galway-Mayo Institute of Technology, Galway, Ireland
[h] European Commission, DG Joint Research Center, Directorate D – Sustainable Resources, Ispra, VA, Italy
[i] School of Biological Sciences, Monash University, Clayton, Victoria, Australia
[j] School of Resource & Environmental Management, Simon Fraser University, Burnaby, BC, Canada

ARTICLE INFO

Handled by A.E. Punt

Keywords:
Stock assessment
Stock status
Data-limited fisheries
Data-poor fisheries
Catch-only methods
Ensemble methods

ABSTRACT

Demand for data-limited stock assessment methods is increasing, and new methods are being developed rapidly. One class of these methods requires only catch time series and, in some cases, information about life history or fishery characteristics, to estimate stock status. These catch-only methods (COMs) range from statistical models trained on data-rich stocks to mechanistic population models that make assumptions about changes in fishing effort. We review 11 COMs, comparing performance through application to data-rich stocks and simulated fisheries. The catch-only methods evaluated here produce imprecise and biased estimates of B/B_{MSY}, especially for stocks that are lightly exploited. They were also generally poor classifiers of stock status. While no method performed best across all stocks, ensembles of multiple COMs generally performed better than individual COMs. We advocate for testing new COMs using this common platform. We also caution that performance in estimating stock status is not sufficient for gauging the usefulness of COMs in managing fisheries. Greater use of management strategy evaluation is needed before COMs can be considered a reliable tool for management.

1. Introduction

While many stocks in developed parts of the world have comprehensive stock assessments that take into account factors such as life history, age, and abundance trends (Ricard et al., 2012), the majority of global stocks remain unassessed (Costello et al., 2012). The dearth of formal assessments is due to several factors, including a lack of resources for data collection and evaluation. Although this problem may be more prevalent in developing regions and regions with high species diversity, it is also an issue in developed countries for stocks with small population size or low economic value (Neubauer et al., 2018; Thorson and Cope, 2015). Changes to national and international fisheries legislation have required assessment of many stocks not previously assessed (e.g., the reauthorization of the Magnuson–Stevens Fishery Conservation and Management Act in the U.S. in 2006 and the reform of the Common Fisheries Policy in the E.U. in 2013). More broadly, international commitments to the UN Sustainable Development Goals (e.g., Goal 14 requires stocks to be restored to MSY-levels), implies a need to understand the status of more of the world's stocks. In the U.S., Europe, and Australia, where many stocks have time series of catch (i.e., landings plus discards), many new methods for assessing the "catch-only" family of data-limited fisheries have been developed and adopted (Anderson et al., 2017; GFCM, 2017; Newman et al., 2015; Zhou et al., 2016).

Catch-only methods (COMs) are data-limited stock assessment methods that rely primarily on time series of catch or landings to estimate stock biomass status (e.g., B/B_{MSY} or depletion) and other common fisheries reference points and quantities. Some catch-only

* Corresponding author at: 2400 Bren Hall, Bren School of Environmental Science and Management, University of California, Santa Barbara, Santa Barbara, CA, 93106, USA.
E-mail address: cfree14@gmail.com (C.M. Free).

https://doi.org/10.1016/j.fishres.2019.105452
Received 30 September 2019; Received in revised form 19 November 2019; Accepted 20 November 2019

methods use mechanistic population models to estimate stock status by recreating stock history from catch data, assumptions about fishing effort, and informative priors on depletion and demographic parameters (e.g., Froese et al., 2017; Martell and Froese, 2013; Thorson et al., 2013). Others use empirical models trained on data-rich stocks to predict status from catch data and auxiliary information such as location, life history, or fishery characteristics (Berkson et al., 2011; Costello et al., 2012; Free et al., 2017; Thorson et al., 2012; Zhou et al., 2017). Recently, ensemble approaches have been used to combine individual method predictions in an attempt to further improve status estimates (e.g., Anderson et al., 2017).

Although there is growing interest in using catch-only methods to assess data-limited fisheries, there has been debate over the wisdom of these methods (Pauly et al., 2013). This debate has largely centered on the use of simple rules that assume trends in catch are indicative of trends in abundance (e.g., Worm et al., 2006) and present overly pessimistic views of global fisheries status (Branch et al., 2011). For example, studies that identify stocks as collapsed when their catch falls below 10 % of their historic maximum suggest that 24–36 % of global fisheries are collapsed (e.g., Pauly, 2008, 2007; Worm et al., 2006) while studies that identify collapse using biomass estimates from stock assessments indicate that only 8–14 % of global fisheries are collapsed (e.g., Branch et al., 2011; Hutchings et al., 2010; Worm et al., 2009). In reality, many factors besides abundance influence catch levels, including management regulations, taxonomy changes, exclusion of distant water fleets, shifting market and fuel prices, natural disasters, and civil war (Branch et al., 2011). Recent catch-only methods take more sophisticated approaches that attempt to account for some of these complexities and may therefore provide better estimates of population status than naïve interpretations of catch trends.

Understanding how well catch-only methods estimate stock status is necessary before they can be used in management. Here, we review 11 catch-only stock assessment methods and develop a framework for testing and comparing the performance of these and future catch-only methods. This framework involves applying the methods to both simulated and real-world fish stocks and quantifying their ability to predict both continuous (i.e., B/B_{MSY}) and categorical (i.e., under, over, or fully exploited) stock status using multiple performance metrics. We also provide recommendations for managers seeking to assess catch-only fisheries and identify opportunities for continuing to improve catch-only stock assessment methods.

2. Taxonomy of catch-only methods

Catch-only methods for estimating stock status can be broadly categorized as taking a graphical, empirical, mechanistic, or ensemble approach (Table 1). Graphical approaches use simple rules about fisheries development theory to classify stocks into development categories (Froese and Kesner-Reyes, 2002; Kleisner et al., 2013). Empirical approaches use information from assessed stocks to derive relationships between stock status and catch data, often with auxiliary information such as location, life history, or characteristics of the fishery (Berkson et al., 2011; Costello et al., 2012; Free et al., 2017; Thorson et al., 2012; Zhou et al., 2017). Thus, empirical approaches attempt to "learn from experience" (Hilborn and Liermann, 1998) by using the dynamics of data-rich stocks to interpret the catch data of data-limited stocks. They estimate a single reference point, generally B/B_{MSY}, but sometimes depletion (e.g., Zhou et al., 2017), and are typically justified through cross-validation where the COM is tested on stocks in a test data set of stocks assessed through more robust methods. These empirical approaches rely on statistical associations in real-world data, e.g., differences in average stock status among regions. Therefore, they are not typically evaluated using simulation testing.

By contrast, mechanistic approaches postulate an underlying population dynamics model and can be divided into those that fit only a population dynamics model (Froese et al., 2017; Martell and Froese, 2013; Zhou et al., 2018) or explicitly fit a coupled model of population and fishing effort dynamics (Thorson et al., 2013; Vasconcellos and Cochrane, 2005). That is, these models attempt to explain an observed change in catch through a combination of change in abundance and change in fishing effort. Both types require assumptions about fishing effort, and benefit from priors on quantities such as initial and final year depletion and model parameters such as intrinsic growth rate (r) or carrying capacity (K). Mechanistic approaches estimate a wide range of fisheries quantities including r and K and reference points including B/B_{MSY}, F/F_{MSY}, B_{MSY}, F_{MSY}, and MSY. They are often validated through a combination of simulation testing and comparison to more robust, data-rich assessment methods.

Ensemble models, which take the average or weighted average of several methods, have been proposed to provide more accurate and less biased estimates than individual methods. Most recently, super-ensembles, which attempt to achieve the best overall predictions by harnessing the strengths of many individual methods through calibration via a statistical model, have improved estimation performance in some circumstances compared to simple ensembles (Anderson et al., 2017).

The 11 methods evaluated here span graphical, empirical, mechanistic, and ensemble approaches (Table 1). We did not evaluate DCAC (MacCall, 2009) or DB-SRA (Dick and MacCall, 2011), which are based heavily on catch data, but focus on estimating overfishing limits rather than stock status and require more intensive life history information (e.g., B_{MSY}/B_0, F_{MSY}/M, age-at-maturity) and complete catch time series (DB-SRA only), and expect informed estimates of depletion. While the methods evaluated here include built-in procedures for establishing vague depletion priors, assumed depletions for DCAC and DB-SRA are generally set using expert judgement.

3. Catch-only method descriptions

3.1. Stock status plots

The first catch-only assessment methods used simple rules to graphically infer stock status from catch time series. (Csirke and Sharp, 1984) illustrated how a time series of landings could be used to classify a fishery's state of development into six phases: (1) predevelopment, (2) growth, (3) full exploitation, (4) overexploitation, (5) collapse, and (6) recovery. This classification is based on the relationship between abundance, fishing effort, and catch in each phase. It assumes that abundance decreases as catch and effort increase during the predevelopment and growth phases. As the fishery becomes fully exploited and moves into overexploitation, the relationship between abundance and catch breaks down until high fishing effort reduces abundance, after which catches decline (Hilborn and Walters, 1992). Later, (Grainger and Garcia, 1996) used catch trends of the 200 top-landed species globally to assign stocks to the development phases defined by (Csirke and Sharp, 1984) and presented the first description of global marine fisheries status.

These plots are called "stock status plots" (SSPs) and there have been several iterations of the approach (e.g., Froese and Kesner-Reyes, 2002; Garcia et al., 2005; Kleisner et al., 2013; Pauly, 2008). Generally, they assign status based on the ratio of current catch to the maximum catch with the development phases occurring before the year of maximum catch and the overexploitation, collapse, and recovery phases occurring after the year of maximum catch. Kleisner et al. (2013) present a detailed accounting of these methods and the development of SSP algorithms. SSPs have been widely used but have been criticized for being negatively biased and providing overly pessimistic conclusions regarding stock status (Branch et al., 2011; Carruthers et al., 2012). Anderson et al. (2012) attempted to overcome some of these deficiencies by smoothing the catch time series, categorizing fisheries as developed within three years of the peak catch, and using biological reference points to calibrate the thresholds. Kleisner et al. (2013) also

Table 1
Catch-only stock assessment methods.

	Method	References	Data input/output	Brief description
Graphical approaches				
1	**SSP-2002** Stock status plots	Froese and Kesner-Reyes (2002)	In: Catch Out: Development status	Uses simple rules that compare the current year's catch to the maximum catch to estimate status
2	**SSP-2013** Updated stock status plots	Kleisner and Pauly (2011) Kleisner et al. (2013)	In: Catch Out: Development status	Updates the rules of SSP-2002 to include a recovery phase in its estimation of status
Empirical approaches				
3	**rORCS** Refined ORCS approach	Berkson et al. (2011); Free et al. (2017)	In: Catch, 12 questions Out: Exploitation status, OFL	Uses a boosted classification tree model trained on the RAMLDB to predict status from 12 stock- and fishery-related predictors
4	**mPRM**[a] Modified panel regression model	Costello et al. (2012); Anderson et al. (2017)	In: Catch, taxonomic group Out: B/B_{MSY}	Uses a panel regression model trained on the RAMLDB to predict status from characteristics of the catch time series and taxonomic group
5	**zBRT** Catch-only boosted regression trees	Zhou et al. (2017)	In: Catch Out: Saturation	Uses a boosted regression tree model trained on the RAMLDB to predict status from 38 catch history statistics
Mechanistic approaches				
6	**cMSY-2013**[a] Catch-MSY	Martell and Froese (2013); Rosenberg et al. (2014)	In: Catch, resilience Out: B/B_{MSY}, MSY, B, B_{MSY}	Uses a stock reduction analysis with priors for *r*, *K*, and initial/final year depletion derived from resilience and maximum catch to estimate status
7	**cMSY-2017** Updated catch-MSY (cMSY)	Froese et al. (2017)	In: Catch, resilience Out: B/B_{MSY}, F/F_{MSY}, B, F, B_{MSY}, F_{MSY}, MSY	Updates the cMSY-2013 stock reduction analysis with a new algorithm for identifying probable r-k pairs in its estimation of status
8	**OCOM** Optimized catch-only model	Zhou et al. (2018)	In: Catch, natural mortality (M) Out: Saturation, MSY	Uses a stock reduction analysis with priors for *r* and final year depletion derived from *M* and saturation from zBRT to estimate status
9	**COM-SIR**[a] Catch-only-model with sampling-importance-resampling	Vasconcellos and Cochrane (2005); Rosenberg et al. (2014)	In: Catch, resilience Out: B/B_{MSY}	Uses a coupled harvest-dynamics model fit using a sampling-importance-resampling algorithm to estimate status
10	**SSCOM**[a] State-space catch-only model	Thorson et al. (2013)	In: Catch, resilience Out: B/B_{MSY}	Uses a coupled harvest-dynamics model fit using a Bayesian hierarchical state-space framework to estimate status
Ensemble approaches				
11	**Superensemble** Random forest superensemble model	Anderson et al. (2017)	In: Catch, B/B_{MSY} from 4 COMs Out: B/B_{MSY}	Uses a random forest model trained on simulated stocks to predict status from the B/B_{MSY} predictions of 4 COMs and two spectral properties of the catch time series

[a] Used as a predictor in the Anderson et al. (2017) superensemble model.

calculated the plots based on smoothed catches and updated the algorithm to account for stock rebuilding to further reduce biases.

3.2. Refined ORCS approach

The "Only Reliable Catch Stocks" (ORCS) Working Group approach estimates stock status (i.e., under, fully, or overexploited) as the unweighted mean of 14 categorical stock- and fishery-related predictors (Berkson et al., 2011; SAFMC, 2013, 2012). Free et al. (2017) refined the approach (rORCS) using a boosted classification tree model trained on data-rich stocks in the RAM Legacy Database to classify stock status from 12 of these predictors, the most important of which were the ex-vessel price, status of the assessed stocks in the fishery, targeting intensity, discard rate, and occurrence in the catch (Free et al., 2017). While these Likert-scale predictors can generally be provided by experts familiar with the fishery, the rORCS model can impute values for missing predictors and can thus be used to assess even the most data-poor fisheries. The rORCS approach also includes a step for estimating the overfishing limit as the product of a historical catch statistic and a scalar based on stock status and risk policy. The refined ORCS approach performed significantly better than the original approach and six other catch-only methods when evaluated on a small, independent test dataset of data-rich stocks (Free et al., 2017). The original approach was used to explore catch limits for 20 stocks in the U.S. Southeast (SAFMC, 2013, 2012) and the refined approach has been used to assess the status of chub mackerel and recommend a corresponding overfishing limit in the U.S. Mid-Atlantic (MAFMC, 2019).

3.3. Modified panel regression model

Costello et al. (2012) used a panel regression model (mPRM) trained on data-rich stocks in the RAM Legacy Database to estimate B/B_{MSY} using characteristics of the stock (e.g., species category, age-at-maturity, max length, etc.) and statistical properties of the catch time series (e.g., years to max catch, slope in catch in first six years, etc.). The approach has subsequently been modified by condensing the species categories into three broad life-history categories (i.e., small pelagic, large pelagic, demersal) and removing the maximum catch predictor (Anderson et al., 2017; Rosenberg et al., 2014). These modifications were made to standardize the method with other catch-only methods and to allow it to be tested on simulated data. The original model was used to assess the status of > 1,700 marine fisheries (Costello et al., 2012) and the modified model has been used as a predictor of stock status in the superensemble model described below (Anderson et al., 2017).

3.4. Catch-only boosted regression tree model

Zhou et al. (2017) used boosted regression tree models (zBRT) trained on data-rich stocks in the RAM Legacy Database to estimate saturation, the ratio of current biomass to the unfished biomass ($B/B_0 \approx 0.5^*B/B_{MSY}$), from 56 catch history statistics, the most important of which were linear regression coefficients for the whole catch time series, the subseries before and after the maximum catch, and recent years. Ultimately, saturation was estimated as the average of the values predicted by two reduced and bias-corrected BRT models (8 and 38 predictors each). The zBRT method was conceptualized as an advancement of the SSP-2002, SSP-2013, and mPRM methods, which also estimate status from patterns in catch history, and Zhou et al. (2017) show that the method is a better predictor of status than these methods when applied to the dataset used for model training. Estimates of saturation from the zBRT method may therefore be useful for deriving depletion (i.e., 1 – saturation) priors for more advanced stock reduction analyses such as DB-SRA (Dick and MacCall, 2011), DCAC (MacCall, 2009), and catch-MSY (described below). Indeed, Zhou et al. (2018) developed their own catch-only stock reduction analysis (described below) that uses their zBRT method to set a prior for final year depletion.

3.5. Catch-MSY methods

Catch-MSY methods (Froese et al., 2017; Martell and Froese, 2013) are widely used catch-only stock reduction analyses. Stock reduction analyses reconstruct historical abundance and exploitation rates by simulating biomass trajectories that could produce the observed catch time series given informative priors on initial and final year depletion and stock dynamics such as carrying capacity, *K*, or intrinsic growth rate, *r*, in the Schaefer model (Schaefer, 1954). The original catch-MSY method (cMSY-2013; Martell and Froese, 2013) establishes priors for *r* based on population resilience, *K* based on maximum catch (e.g., between C_{max} and 100^*C_{max}), and initial and final year depletion based on the ratio of initial and final year catch to the maximum catch. It then estimates 'viable' pairs of *r* and *K* (i.e., pairs that do not result in extinction or final year depletion outside the bounds of the prior) and uses the geometric mean of these pairs to estimate MSY. The method was modified by Rosenberg et al. (2014) to generate biomass trends from all viable *r*-*K* pairs and produce an estimate of B/B_{MSY} from the median trend. Froese et al. (2017) expanded the method (cMSY-2017) to estimate biomass, exploitation rates, and MSY and related reference points. They also introduced a procedure to identify the most probable *r*-*K* pairs that aims to address the tendency for production models to overestimate productivity at very low stock sizes. The catch-MSY methods have been used in the assessment of global (Costello et al., 2016) and regional fisheries (Froese et al., 2018) and the assessment of Western Pacific stocks (De Oliveira, 2015), East China Sea stocks (Zhang et al., 2018), Indian Ocean stocks (IOTC, 2017, 2016, 2015), and Atlantic shortfin mako shark (Winker et al., 2017), among many others.

3.6. Optimized catch-only model

The optimized catch-only model (OCOM) employs a stock reduction analysis using priors for intrinsic growth rate (*r*) and final year depletion derived from natural mortality and saturation estimated by the zBRT method, respectively (Zhou et al., 2018). The stock reduction analysis uses a Schaefer biomass dynamics model and an algorithm for identifying feasible parameter combinations to estimate biological quantities such as *r*, *K*, annual biomass, and depletion as well as management quantities such as MSY, B_{MSY}, and F_{MSY}. OCOM was conceptualized as an advancement of the catch-MSY stock reduction analyses because it uses an optimization algorithm rather than a stochastic "thread the needle" approach (Walters et al., 2006) as well as a more informed depletion prior. Zhou et al. (2018) test the OCOM method on 14 Australian fish stocks assessed using Stock Synthesis (Methot and Wetzel, 2013) and show that OCOM estimates biological and management quantities comparable to those from full assessments. OCOM has been used by the Indian Ocean Tuna Commission for assessment of several tuna stocks (IOTC, 2017, 2016, 2015).

3.7. Catch-only model with sampling-importance-resampling

COM-SIR (catch-only-model with sampling-importance-resampling) is a coupled harvest-dynamics model (Vasconcellos and Cochrane, 2005) in which biomass and harvest (i.e., fishing effort) dynamics are assumed to follow Schaefer and logistic models, respectively. The model is fit using a sampling-importance-resampling algorithm. COM-SIR was tested on Atlantic yellowfin tuna (*Thunnus albacares*) and Namibian hake (*Merluccius capensis*) stocks using data from years when these stocks were unregulated. The model performed poorly for yellowfin tuna but produced B/B_{MSY} and F/F_{MSY} estimates similar to those estimated by data-rich assessments for Namibian hake. Although we are unaware of other examples of COM-SIR having been used to assess the status of data-limited stocks, it has been used as a predictor of stock status in the superensemble model described below (Anderson et al., 2017).

3.8. State-space catch-only model

The state-space catch-only model (SSCOM) is a hierarchical model that, similar to COM-SIR, is based on a coupled harvest-dynamics model (Thorson et al., 2013). SSCOM estimates unobserved dynamics in both fishing effort and the fished population based on a catch time series and priors on *r*, the maximum rate of increase of fishing effort, and the magnitude of various forms of stochasticity. The model is fit in a Bayesian state-space framework to integrate across three forms of stochasticity: variation in effort, population dynamics, and fishing efficiency. The model was validated via simulation testing and application to eight assessed US West Coast groundfish stocks (Thorson et al., 2013). The model recovered coupled population and effort dynamics for simulated stocks and yielded results comparable to data-rich assessments for assessed stocks exhibiting roller coaster catch (i.e., increasing then decreasing catch or a "two-way trip"). Although we are unaware of other examples of SSCOM having been used to assess the status of data-limited stocks, it has been used as a predictor of stock status in the superensemble model described below (Anderson et al., 2017).

3.9. Superensemble models

The catch-only methods described above have strengths and weaknesses — they perform better under some conditions than others — the superensemble model attempts to harness the strengths of multiple method to derive the best overall prediction (Anderson et al., 2017). While basic ensemble models take an average or weighted average of multiple method predictions, a *super*ensemble goes a step further by calibrating individual method predictions through a regression model fitted to data with known or trusted properties (Krishnamurti et al., 1999). This "training" dataset might be a simulated dataset with known status or a set of stocks in which stock status is already well estimated. This process lets the superensemble exploit the covariance between individual methods, allows for nonlinear weightings of individual method predictions, and allows those weightings to be a function of covariates. These additional covariates could be, for example, exploitation patterns, life history characteristics, or statistical properties of the catch time series. For instance, if method A and method B disagree on the status of a stock, a superensemble model might choose to weight method A more heavily because it tends to perform better for species with a similar life history.

Anderson et al. (2017) showed that a random forest superensemble comprised of mPRM, cMSY-2013, COM-SIR, SSCOM, and two covariates describing the spectral properties (i.e., variation in the frequency domain) of the catch time series outperformed any one method in terms of bias, accuracy, and correlation across multiple stocks. This was true when tested on a simulated dataset and when tested on empirical data from the RAM Legacy Database. Recently, Rosenberg et al. (2017) applied the same superensemble approach to assess global fish stocks status with catch data from the FAO. Beyond these two applications, the superensemble approach could be used to combine predictions from any of the methods outlined in this paper. The approach is limited mainly by the accuracy of the individual methods and the degree to which the training or "trusted" dataset is representative of the dataset to which they are applied.

4. Performance evaluation methods

4.1. Overview

We evaluated the performance of the 11 catch-only methods by testing them on catch data from both real and simulated fish stocks. Although the status of real fish stocks is not known without error, they provide a potentially more representative basis for testing as they reflect all of the uncertainties (e.g., observation and process errors) inherent in actual fisheries data. Simulated fish stocks offer greater sample size, the ability to compare predicted status against known status (without error), and the opportunity to compare performance among stocks of varying traits. We evaluated the performance of all methods categorically (i.e., how well does each method classify a stock as under, fully, or overexploited) and eight methods continuously (i.e., how well does each method estimate B/B_{MSY}). Because these metrics are widely applicable and these datasets are publicly available, they provide a framework for comparing the performance of future data-limited assessment methods. We provide an R package ("datalimited2″) for calculating these metrics and creating the plots presented here: https://github.com/cfree14/datalimited2

4.2. Test stocks

We applied the 11 methods to real fish stocks in the RAM Legacy Stock Assessment Database (RAMLDB v. 2.95; Ricard et al., 2012) and simulated fish stocks from Rosenberg et al. (2014). The RAMLDB is a global database of catch data and stock assessment output, including reference points and time series of biomass and fishing mortality. We evaluated 175 of the 193 RAMLDB stocks used by Free et al. (2017) with catch time series ≥20 years long after trimming years of zero catch from the beginning of the time series. The methods converged on status predictions for 135 stocks (**Table S1**) including stocks that are under ($B/B_{MSY} > 1.5$; 37.8 % of stocks), fully ($0.5 < B/B_{MSY} < 1.5$; 45.9 % of stocks), and overexploited ($B/B_{MSY} < 0.5$; 16.3 % of stocks) and representing a variety of taxa, regions, and management agencies (**Fig. S1A-E**). The simulated stocks in Rosenberg et al. (2014) represent a fully factorial dataset of simulated fisheries including three fish life histories, three levels of initial biomass depletion, four exploitation scenarios, two levels of recruitment variability, two levels of recruitment autocorrelation, and two levels of measurement error, with each combination of parameters run through ten stochastic iterations (**Table S2**). The methods converged on status predictions for 5491 of 5,760 stocks (**Table S1**) including stocks that are under (26.1 % of stocks), fully (59.2 % of stocks), and overexploited (14.6 % of stocks) (**Fig. S1F**). See the supplementary text for details on applying the catch-only methods.

4.3. Performance metrics

The 11 methods provide a mixture of continuous (i.e., B/B_{MSY} or saturation) and categorical (i.e., exploitation or development category) status predictions (Table 1). To compare performance categorically, we standardized status estimates to the following three exploitation categories: under, fully, and overexploited (see the supplementary text for details). To compare performance continuously, we converted estimates of saturation to B/B_{MSY} (i.e., B/B_{MSY} = 2*saturation, assuming Schaefer production dynamics).

We evaluated the continuous performance of the eight methods that estimate B/B_{MSY} or saturation by measuring each method's bias, accuracy, and Spearman's rank-order correlation. We measured bias as the median proportional error (MPE) which quantifies directional error (i.e., positively or negatively biased). We measured accuracy as the median absolute proportional error (MAPE) which quantifies overall error (i.e., error in any direction). Proportional error (PE) is calculated as $(\hat{\theta} - \theta)/|\theta|$, where $\hat{\theta}$ and θ represent predicted and "true" (or data-rich stock assessment) B/B_{MSY} values, respectively.

We evaluated the categorical performance of each method using both proportional agreement (accuracy) and Cohen's kappa. Cohen's kappa measures inter-rate agreement between categorical items and is

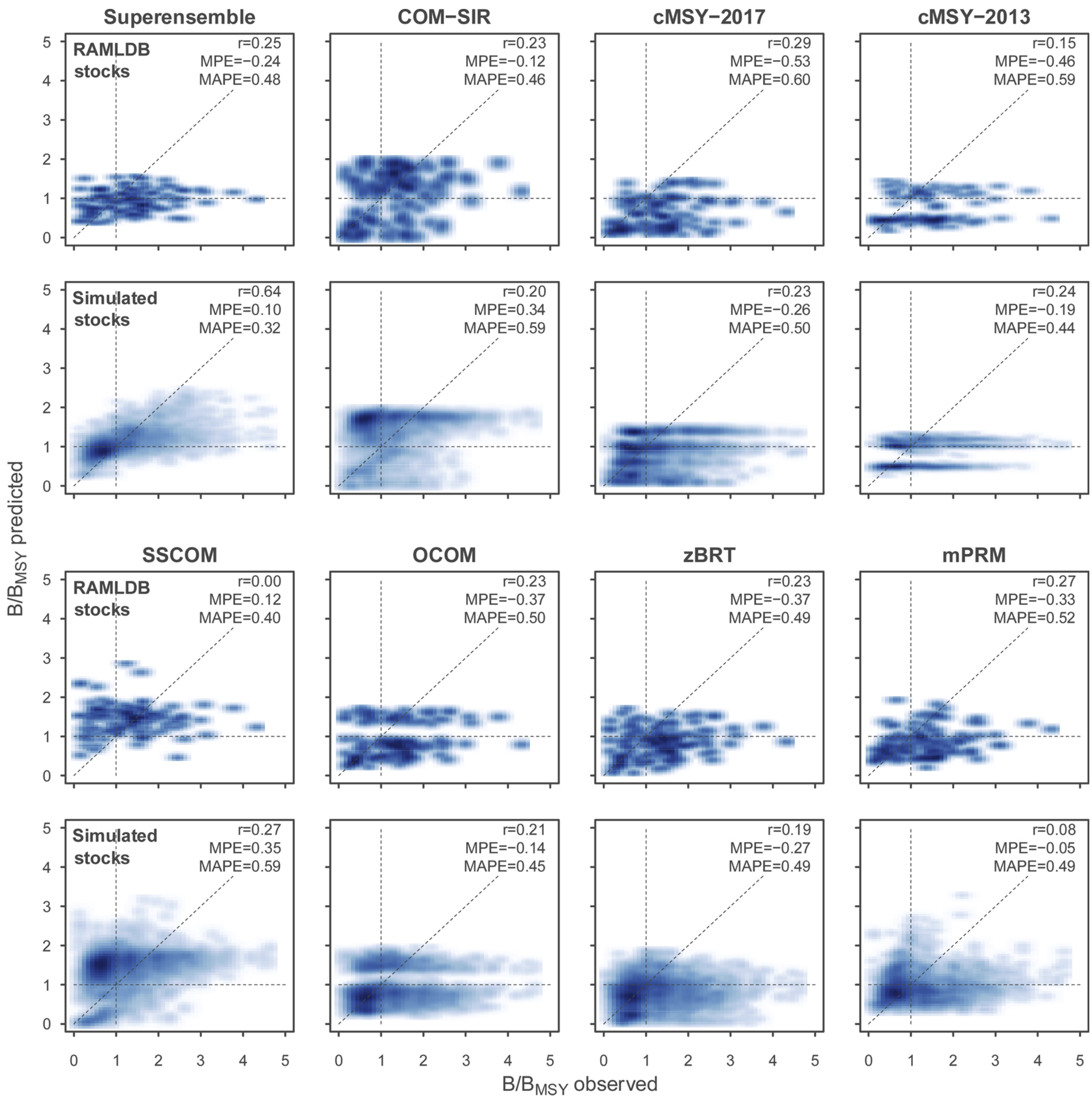

Fig. 1. Observed population status versus population status predicted by eight catch-only methods for the RAM Legacy Database stocks (n = 135) and Rosenberg et al. (2014) simulated stocks (n = 5,491). Points are binned for visual presentation: darker areas indicate areas with greater density of points. Median proportional error (MPE) and median absolute proportional error (MAPE) measure bias and accuracy, respectively.

more robust than simple proportional agreement because it takes into account the probability of agreement occurring by chance (Cohen, 1968). This is necessary for tests on both the RAMLDB and simulated stocks given the uneven distribution of exploitation categories in both datasets (many fully exploited stocks, few overexploited stocks; **Fig. S1**). For example, if a method misclassifies most overexploited stocks but correctly classifies most fully exploited stocks, it would still earn a high accuracy, but its kappa value would be appropriately penalized. Although there are no definitive rules for interpreting Cohen's kappa, general guidelines suggest that values > 0.70 are "excellent", 0.4-0.7 are "good", 0.2-0.4 are "fair", and < 0.2 are "poor" (Landis and Koch, 1977).

5. Performance results

The catch-only methods evaluated here produce imprecise and biased estimates of B/B_{MSY}, especially for stocks that are lightly exploited (Figs. 1 and 2). They are also "poor" classifiers of stock status with two exceptions (Fig. 2): (1) the superensemble was a "fair" classifier of status when tested on the simulated stocks (the dataset on which it was trained) and (2) the refined ORCS approach was a "good" classifier of status when tested on the RAMLDB stocks (the dataset on which it was trained). In general, the superensemble model was the best predictor of stock status. It produced the most accurate, most correlated, and least biased predictions of B/B_{MSY} when tested on the simulated stocks and was among the best predictors of B/B_{MSY} when tested on the RAMLDB stocks (Fig. 2). It performed consistently well across effort dynamics scenarios represented in the simulated stocks (Fig. 3): it was the best predictor of status for stocks experiencing constant fishing effort and it was among the best predictors of status for stocks experiencing roller coaster and biomass-coupled fishing effort. cMSY-2017, mPRM, and OCOM were better predictors of status for stocks experiencing increasing fishing effort (Fig. 3).

The generally poor performance of the evaluated catch-only

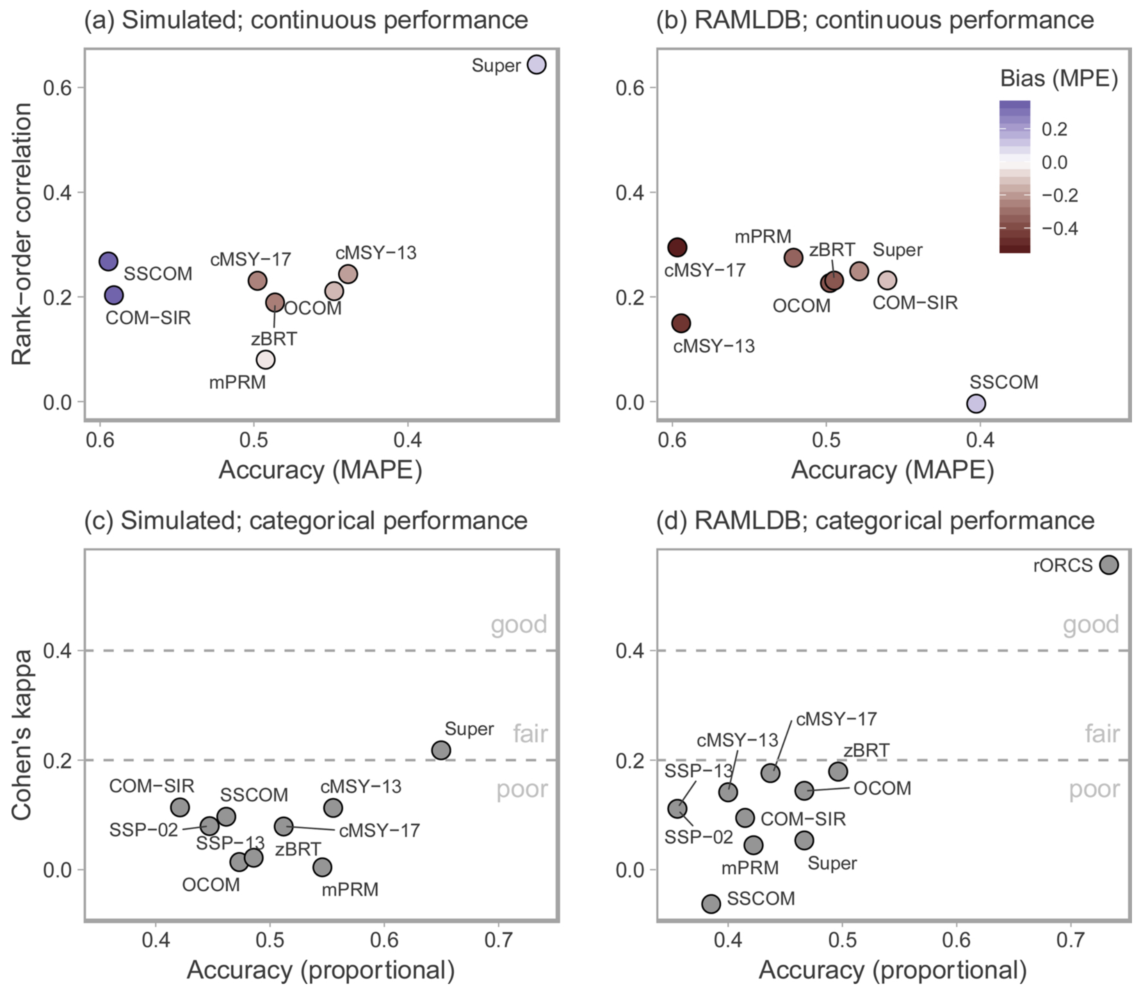

Fig. 2. The status estimation performance of catch-only methods tested on the Rosenberg et al. (2014) simulated stocks and RAM Legacy Database stocks. In the continuous performance plots (top row), the best performing methods are indicated by high rank-order correlation and high accuracy (low MAPE). The x-axes have been reversed so that the best performing methods appear in the top-right corners. In the categorical performance plots (bottom row), the best performing methods are indicated by high Cohen's kappa and high accuracy (top-right corner). Note: Data-rich statistical catch at-age models produced MAPEs of 0.07-0.33 (0.15 median) when tested in a similar simulation framework by Ono et al. (2015). These values are significantly lower (more accurate) than the MAPEs produced by the evaluated COMs.

methods arose through a variety of mechanisms. While the superensemble model made relatively unbiased predictions of status for overexploited stocks, it was negatively biased for lightly exploited stocks (Fig. 1). This pattern was more pronounced for the RAMLDB stocks than for the simulated stocks on which the superensemble model was trained. SSCOM and COM-SIR produced positively biased predictions of status, especially for overexploited stocks, when tested on both the simulated and RAMLDB stocks (Fig. 1). cMSY-13 and OCOM generated bimodal predictions of B/B_{MSY} for both the simulated and RAMLDB stocks (Fig. 1). Although patterns in performance revealed through testing on the simulated and RAMLDB stocks were generally shared, slight differences in the relative performance of the evaluated methods arose given that: (1) the efforts dynamics represented in the simulated stocks are factorially balanced while the effort dynamics represented in RAMLDB stocks are ad-hoc; and (2) some of the methods were trained on the simulated stock (i.e., the superensemble) or the RAMLDB stock (i.e., zBRT, mPRM, rORCS) datasets and perform better when tested on these datasets.

6. Discussion

Although all stock assessment methods, even data-intensive ones, are uncertain (Hilborn and Walters, 1992), the catch-only methods evaluated here produce especially imprecise and biased estimates of stock biomass status. Ono et al. (2015) provide a useful benchmark for comparing the accuracy of the catch-only methods evaluated here to the accuracy of more data-intensive assessment methods. They measured the accuracy of statistical catch-at-age (SCAA) models through testing on simulated life histories (cod-, flatfish-, and sardine-like species) and exploitation histories (constant, increasing, and roller coaster effort dynamics) highly similar to those implemented in our simulation framework (**Table S2**; Rosenberg et al., 2014). They found that SCAA models, given the highest quality and quantity of length and age composition data (similar to the data available in U.S. West Coast fisheries) were 9–16 % inaccurate (MAPE), which is significantly lower than the 32–60 % inaccuracy of the catch-only methods evaluated here. When provided lower qualities and quantities of data, the SCAA models were 7–33 % inaccurate, which is still considerably more accurate than the evaluated catch-only methods.

There is, however, opportunity to improve the performance and utility of catch-only assessment methods. While no method performed best across all stocks, the superensemble model generally produced better estimates of B/B_{MSY} than the individual catch-only methods, and exhibited accuracy (32 % inaccurate) comparable to statistical catch-at-age models provided low quality data (33 % inaccurate; Ono et al., 2015). The performance of the superensemble model could be further

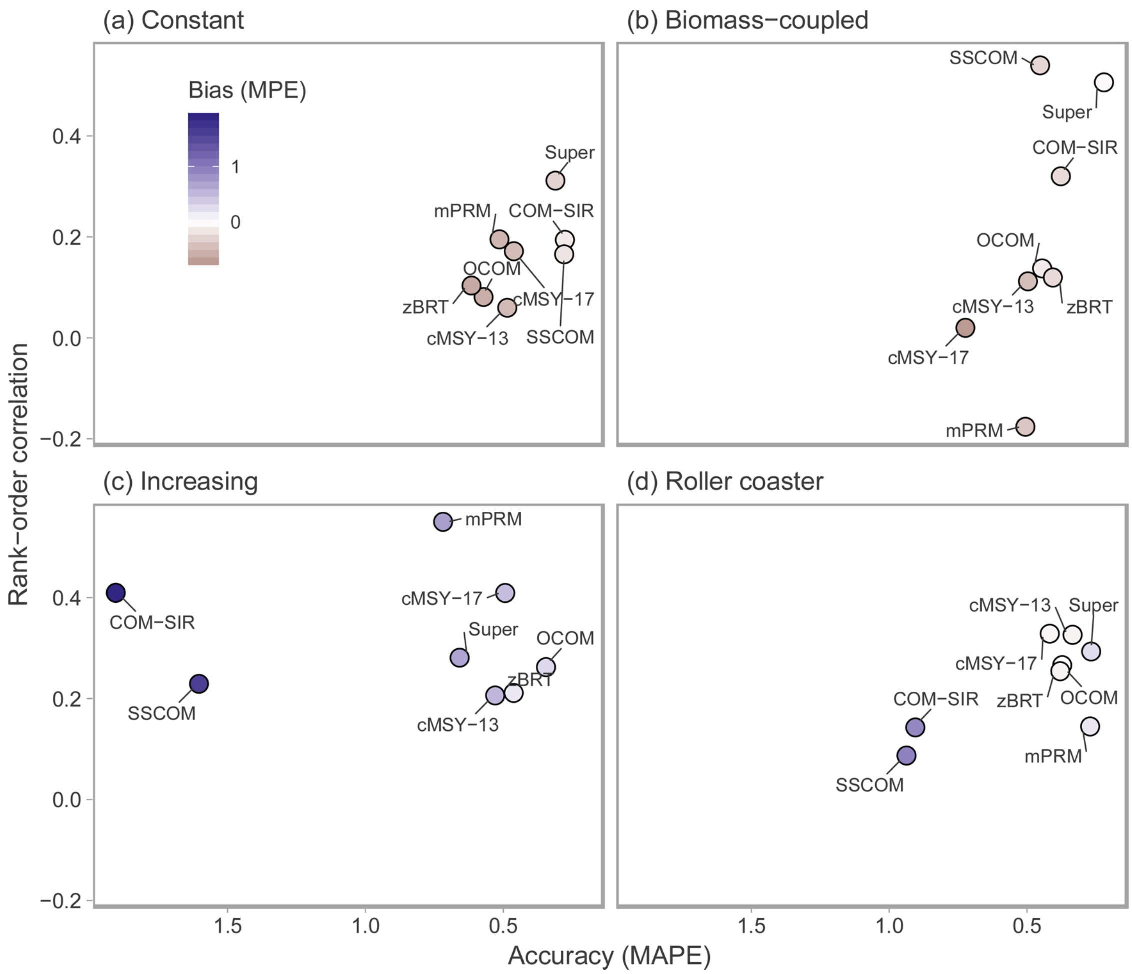

Fig. 3. The status estimation performance of catch-only methods tested on the Rosenberg et al. (2014) simulated stocks by effort dynamics scenario. The best performing methods are indicated by high rank-order correlation and high accuracy (low MAPE). The x-axes have been reversed so that the best performing methods appear in the top-right corners.

Note: Data-rich statistical catch at-age models produced MAPEs of 0.08-0.33 (0.15 median), 0.09-0.25 (0.16 median), and 0.07-0.21 (0.16 median) when tested on constant, increasing, and roller coaster effort dynamics in a similar simulation framework by Ono et al. (2015). These values are significantly lower (more accurate) than the MAPEs produced by the evaluated COMs.

improved by using the B/B_{MSY} predictions of additional catch-only methods as predictors, especially those that are not highly correlated with the methods already included in the model (Anderson et al., 2017). zBRT is a promising candidate because it yields reasonable predictions (e.g., not bimodal like those from OCOM) and its predictions are not highly correlated with the methods already included (**Fig. S3**). The performance of the superensemble model as a classifier may also be improved by explicitly training the model to predict categorical status. The sensitivity of model performance to effort history suggests that significant gains could be made by tailoring superensemble models to each of the Rosenberg et al. (2014) effort dynamics scenarios. The scenarios are quite general (i.e., constant, increasing, roller coaster, and biomass-coupled effort) and it is likely that an expert familiar with the fishery could accurately classify its effort history and select the corresponding effort-tailored model. Alternatively, superensemble models can be tailored to a specific fishery of interest by training a new model on fisheries simulated to reflect the target fishery (i.e., similar life history, exploitation history, depletion, etc.). R packages such as *FLR* (Kell et al., 2007) and *DLMtool* (Carruthers and Hordyk, 2018) make such an exercise widely accessible by simulating fisheries with user-specified characteristics.

Although the refined ORCS approach was by far the best classifier of stock status for the real stocks, it must undergo further evaluation before it can be unequivocally recommended. The training and testing of the approach on assessed stocks could be problematic as assessed (data-rich) stocks differ systematically from the unassessed (data-limited) stocks to which the method would be applied. For example, assessed fisheries generally target larger, slower growing, and higher trophic level species (Pinsky et al., 2011) and are generally higher volume, more valuable, and in better condition (Costello et al., 2012) than their unassessed counterparts. Consequently, it is possible that the refined ORCS approach is predictive for data-rich fisheries, but not data-limited ones. This problem is true of all empirical catch-only methods and testing these methods on data-rich fisheries with characteristics similar to data-limited fisheries is a necessary next step in their validation. Recently, Neubauer et al. (2018) made this possible by measuring the frequency of assessment of data-rich stocks in the United States and identifying the traits that determine assessment frequency. These results could be used in a “propensity score” framework (Rosenbaum and Rubin, 1983) to assess the bias of the data used to train and test empirical catch-only methods or to identify infrequently assessed, data-rich stocks (similar to unassessed, data-limited stocks) to test catch-only methods on.

Estimation of stock status is only the first step in the management process and status determinations alone do not guarantee, or even preclude (Dowling et al., 2015), the effective and sustainable management of fisheries. In addition, successful management requires the testing and setting of harvest control rules, often facilitated by management strategy evaluation (Punt et al., 2016). Management strategy evaluation involves developing an operating model that describes the dynamics of the fish population and fishery as well as the monitoring and implementation phases of management. The impact of the proposed management strategies (comprised of an assessment method and harvest control rules) on the operating model are evaluated by simulating possible outcomes in a feedback loop (Punt et al., 2016). Catch-only methods differ in accuracy and precision across various life history traits and fishery harvest dynamics and it is uncertain how these differences affect the performance of harvest control rules. Evaluating the performance of harvest control rules that have been informed by catch-only stock assessments is thus essential to ensure management objectives for data-limited fisheries can be met.

Past management strategy evaluations using simulated stocks with generic population and fishing dynamics have shown that the management performance of catch-only methods is difficult to generalize across all situations (Carruthers et al., 2014; Walsh et al., 2018). Catch-only methods coupled with cautious effort-based harvest control rules can result in improved B/B_{MSY} status and lower risks of overfishing, compared to unmanaged fishing practices, but with a substantial foregone yield compared to what might be achievable if status could be determined more precisely (Walsh et al., 2018). Therefore, tailored management strategy evaluations should be conducted for specific data-limited fisheries of interest prior to the implementation and uptake of catch-only assessment methods. These management strategy evaluations can also identify new sources of data than can improve certainty in assessments and facilitate the transition from data-limited to data-moderate fisheries.

In many cases, developing more sophisticated catch-only methods, tuning catch-only methods to specific fisheries, and/or conducting tailored management strategy evaluations may not be feasible due to capacity limitations (Bundy et al., 2016; Ricard et al., 2012) or economical due to small fishery size or commercial value (Neubauer et al., 2018). In capacity-limited systems, management could focus instead on implementing “primary fisheries management” (Cochrane et al., 2011), which uses best available science and precautionary principles to manage fisheries while also establishing or strengthening participatory co-management to incentivize sustainable stewardship. In all systems, procedure-based data-limited approaches that set catch targets or effort limits based on proxy indicators could be more constructive than data-limited methods for assessing status (Dowling et al., 2015). For example, Wiedenmann et al. (2019) showed that catch curve methods, which require the collection of additional catch-at-age data, estimate catch targets close to the overfishing limit and are not sensitive to historical exploitation rates. The supplementation of catch time series with data on catch length composition is also promising (Thorson and Cope, 2015) and could result in better management targets than purely length-based methods (reviewed by Chong et al., 2019) or the catch-only methods evaluated here.

On balance, existing catch-only methods have limited accuracy and a high probability of providing misleading information if used to guide fishery management (Walsh et al., 2018). Paradoxically, if these methods are used to justify new restrictions on harvest, the relationship between fishing effort, catch, and biomass is altered and catch-only methods applied to the resulting future catch time series may be even less accurate. That is, once catch is constrained by management, the information content of the catch time series is degraded. We thus recommend that catch-only methods be treated as a temporary stepping stone while data (e.g., size or age composition) are collected that will allow for more reliable methods to be applied.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank C. Grandin and three anonymous reviewers for valuable feedback on the manuscript. This research was funded by the Food and Agriculture Organization of the United Nations, Marine and Inland Fisheries Branch (FIAF).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.fishres.2019.105452.

References

Anderson, S.C., Branch, T.A., Ricard, D., Lotze, H.K., 2012. Assessing global marine fishery status with a revised dynamic catch-based method and stock-assessment reference points. ICES J. Mar. Sci. 69, 1491–1500. https://doi.org/10.1093/icesjms/fss105.
Anderson, S.C., Cooper, A.B., Jensen, O.P., Minto, C., Thorson, J.T., Walsh, J.C., Afflerbach, J., Dickey-Collas, M., Kleisner, K.M., Longo, C., Osio, G.C., Ovando, D., Mosqueira, I., Rosenberg, A.A., Selig, E.R., 2017. Improving estimates of population status and trend with superensemble models. Fish Fish. 24https://doi.org/10.1111/faf.12200. 173–10.
Berkson, J., Barbieri, L., Cadrin, S., Cass-Calay, S.L., Crone, P., Dorn, M., Friess, C., Kobayashi, D., Miller, T.J., Patrick, W.S., Pautzke, S., Ralston, S., Trianni, M., 2011. Calculating acceptable biological catch for stocks that have reliable catch data only (Only reliable catch stocks – ORCS). NOAA Tech. Memo. 1–56 NMFS-SEFSC-616.
Branch, T.A., Jensen, O.P., Ricard, D., Ye, Y., Hilborn, R., 2011. Contrasting global trends in marine fishery status obtained from catches and from stock assessments. Conserv. Biol. 25, 777–786. https://doi.org/10.1111/j.1523-1739.2011.01687.x.
Bundy, A., Chuenpagdee, R., Boldt, J.L., de Fatima Borges, M., Camara, M.L., Coll, M., Diallo, I., Fox, C., Fulton, E.A., Gazihan, A., Jarre, A., Jouffre, D., Kleisner, K.M., Knight, B., Link, J., Matiku, P.P., Masski, H., Moutopoulos, D.K., Piroddi, C., Raid, T., Sobrino, I., Tam, J., Thiao, D., Torres, M.A., Tsagarakis, K., van der Meeren, G.I., Shin, Y.-J., 2016. Strong fisheries management and governance positively impact ecosystem status. Fish Fish. 18, 412–439. https://doi.org/10.1111/faf.12184.
Carruthers, T., Hordyk, A., 2018. DLMtool: Data-Limited Methods Toolkit.
Carruthers, T.R., Punt, A.E., Walters, C.J., MacCall, A., McAllister, M.K., Dick, E.J., Cope, J., 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fish. Res. 153, 48–68. https://doi.org/10.1016/j.fishres.2013.12.014.
Carruthers, T.R., Walters, C.J., McAllister, M.K., 2012. Evaluating methods that classify fisheries stock status using only fisheries catch data. Fish. Res. 119–120, 66–79. https://doi.org/10.1016/j.fishres.2011.12.011.
Chong, L., Mildenberger, T.K., Rudd, M.B., Taylor, M.H., Cope, J.M., Branch, T.A., Wolff, M., Stäbler, M., 2019. Performance evaluation of data-limited, length-based stock assessment methods. ICES J. Mar. Sci. https://doi.org/10.1093/icesjms/fsz212.
Cochrane, K.L., Andrew, N.L., Parma, A.M., 2011. Primary fisheries management: a minimum requirement for provision of sustainable human benefits in small-scale fisheries. Fish Fish. 12, 275–288. https://doi.org/10.1111/j.1467-2979.2010.00392.x.
Cohen, J., 1968. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol. Bull. 70, 213–220. https://doi.org/10.1037/h0026256.
Costello, C., Ovando, D., Clavelle, T., Strauss, C.K., Hilborn, R., Melnychuk, M.C., Branch, T.A., Gaines, S.D., Szuwalski, C.S., Cabral, R.B., Rader, D.N., Leland, A., 2016. Global fishery prospects under contrasting management regimes. Proc. Natl. Acad. Sci. U. S. A. 113, 5125–5129. https://doi.org/10.1073/pnas.1520420113.
Costello, C., Ovando, D., Hilborn, R., Gaines, S.D., Deschenes, O., Lester, S.E., 2012. Status and solutions for the world’s unassessed fisheries. Science 338, 517–520. https://doi.org/10.1126/science.1223389.
Csirke, J., Sharp, G., 1984. In: Proceedings of the Expert Consultation to Examine Changes in Abundance and Species Composition of Neritic Fish Resources (No. FAO Fisheries Technical Paper 291). San José, Costa Rica.
De Oliveira, J.A.A., 2015. Developing Integrated Assessments for Data Poor Stocks (No. Cefas Contract C5386X). Cefas, Honolulu, HI.
Dick, E.J., MacCall, A.D., 2011. Depletion-based stock reduction analysis: a catch-based method for determining sustainable yields for data-poor fish stocks. Fish. Res. 110, 331–341. https://doi.org/10.1016/j.fishres.2011.05.007.
Dowling, N.A., Dichmont, C.M., Haddon, M., Smith, D.C., Smith, A.D.M., Sainsbury, K., 2015. Guidelines for developing formal harvest strategies for data-poor species and fisheries. Fish. Res. 171, 130–140. https://doi.org/10.1016/j.fishres.2014.09.013.
Free, C.M., Jensen, O.P., Wiedenmann, J., Deroba, J.J., 2017. The refined ORCS approach: a catch-based method for estimating stock status and catch limits for data-poor fish stocks. Fish. Res. 193, 60–70. https://doi.org/10.1016/j.fishres.2017.03.017.
Froese, R., Demirel, N., Coro, G., Kleisner, K.M., Winker, H., 2017. Estimating fisheries reference points from catch and resilience. Fish Fish. 18, 506–526. https://doi.org/10.1111/faf.12190.
Froese, R., Kesner-Reyes, K., 2002. Impact of fishing on the abundance of marine species. ICES CM 12, 1–12 2002/L.
Froese, R., Winker, H., Coro, G., Demirel, N., Tsikliras, A.C., Dimarchopoulou, D.,

Scarcella, G., Quaas, M., Matz-Lück, N., 2018. Status and rebuilding of European fisheries. Mar. Policy 93, 159–170. https://doi.org/10.1016/j.marpol.2018.04.018.
Garcia, S.M., de L. Moreno, I., Grainger, R.J.R., 2005. In: Garcia, S.M., de Leiva Moreno, I., Grainger, R.J.R. (Eds.), 2005. Global Trends in the State of Marine Resources. Review of the State of World Marine Resources, Rome, Italy.
GFCM, 2017. Working Group on Stock Assessment of Demersal Species (WGSAD). General Fisheries Commission for the Mediterranean (GFCM), Rome, Italy.
Grainger, R.J.R., Garcia, S.M., 1996. In: Grainger, R.J.R., Garcia, S.M. (Eds.), 1996. Chronicles of Marine Fishery Landings (1950–1994): Trend Analysis and Fisheries Potential. No. FAO Fisheries Technical Paper 359, Rome, Italy.
Hilborn, R., Liermann, M., 1998. Standing on the shoulders of giants: learning from experience in fisheries. Rev. Fish Biol. Fish. 8, 273–283. https://doi.org/10.1023/A:1008877912528.
Hilborn, R., Walters, C.J., 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Chapman and Hall, London.
Hutchings, J.A., Minto, C., Ricard, D., Baum, J.K., Jensen, O.P., 2010. Trends in the abundance of marine fishes. Can. J. Fish. Aquat. Sci. 67, 1205–1210. https://doi.org/10.1139/F10-081.
IOTC, 2017. Assessment of Indian Ocean Narrow-barred Spanish Mackerel (*Scomberomorus Commerson*) Using Data Limited Catch-based Methods.
IOTC, 2016. Ecological Risk Assessment (ERA) for Neritic Tunas in the IOTC Area of Competence. Indian Ocean Tuna Commission (IOTC).
IOTC, 2015. Assessment of Indian Ocean Kawakawa (*Euthynnus Affinis*) Using Data Poor Catch-based Methods. Indian Ocean Tuna Commission (IOTC).
Kell, L.T., Mosqueira, I., Grosjean, P., Fromentin, J.-M., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M.A., Poos, J.J., Scott, F., Scott, R.D., 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES J. Mar. Sci. 64, 640–646. https://doi.org/10.1093/icesjms/fsm012.
Kleisner, K., Pauly, D., 2011. Stock-status plots of fisheries for Regional Seas. In: In: Christensen, V., Lai, S., Palomares, M.L.D., Zeller, D., Pauly, D. (Eds.), The State of Biodiversity and Fisheries in Regional Seas. Fisheries Centre 602 Research Reports 19 University of British Columbia, Vancouver, Canada.
Kleisner, K., Zeller, D., Froese, R., Pauly, D., 2013. Using global catch data for inferences on the world's marine fisheries. Fish Fish. 14, 293–311. https://doi.org/10.1111/j.1467-2979.2012.00469.x.
Krishnamurti, T.N., Kishtawal, C.M., LaRow, T.E., Bachiochi, D.R., Zhang, Z., Williford, C.E., Gadgil, S., Surendran, S., 1999. Improved weather and seasonal climate forecasts from multimodel superensemble. Science 285, 1548–1550. https://doi.org/10.1126/science.285.5433.1548.
Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159–174. https://doi.org/10.2307/2529310.
MacCall, A.D., 2009. Depletion-corrected average catch: a simple formula for estimating sustainable yields in data-poor situations. ICES J. Mar. Sci. 66, 2267–2271.
MAFMC, 2019. Staff Recommendations for Final Action on the Chub Mackerel Amendment. Mid-Atlantic Fishery Management Council (MAFMC), Dover, DE.
Martell, S., Froese, R., 2013. A simple method for estimating MSY from catch and resilience. Fish Fish. 14, 504–514. https://doi.org/10.1111/j.1467-2979.2012.00485.x.
Methot, R.D., Wetzel, C.R., 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142, 86–99. https://doi.org/10.1016/j.fishres.2012.10.012.
Neubauer, P., Thorson, J.T., Melnychuk, M.C., Methot, R., Blackhart, K., 2018. Drivers and rates of stock assessments in the United States. PLoS One 13, e0196483–19. https://doi.org/10.1371/journal.pone.0196483.
Newman, D., Berkson, J., Suatoni, L., 2015. Current methods for setting catch limits for data-limited fish stocks in the United States. Fish. Res. 164, 86–93. https://doi.org/10.1016/j.fishres.2014.10.018.
Ono, K., Licandeo, R., Muradian, M.L., Cunningham, C.J., Anderson, S.C., Hurtado-Ferro, F., Johnson, K.F., McGilliard, C.R., Monnahan, C.C., Szuwalski, C.S., Valero, J.L., Vert-Pre, K.A., Whitten, A.R., Punt, A.E., 2015. The importance of length and age composition data in statistical age-structured models for marine species. ICES J. Mar. Sci. 72, 31–43. https://doi.org/10.1093/icesjms/fsu007.
Pauly, D., 2008. Global fisheries: a brief review. GJournal Biol. Res.-Thessalon. 9, 3–9.
Pauly, D., 2007. The sea around us project: documenting and communicating global fisheries impacts on marine ecosystems. AMBIO J. Hum. Environ. 36, 290–295. https://doi.org/10.1579/0044-7447(2007)36[290:TSAUPD]2.0.CO;2.
Pauly, D., Hilborn, R., Branch, T.A., 2013. Fisheries: Does catch reflect abundance? Nature 494, 303–306. https://doi.org/10.1038/494303a.
Pinsky, M.L., Jensen, O.P., Ricard, D., Palumbi, S.R., 2011. Unexpected patterns of fisheries collapse in the world's oceans. Proc. Natl. Acad. Sci. U. S. A. 108, 8317–8322. https://doi.org/10.1073/pnas.1015313108.
Punt, A.E., Butterworth, D.S., Oliveira, J.A.A.D., Haddon, M., 2016. Management strategy evaluation: best practices. Fish Fish. 17, 303–334. https://doi.org/10.1111/faf.12104.
Ricard, D., Minto, C., Jensen, O.P., Baum, J.K., 2012. Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish Fish. 13, 380–398. https://doi.org/10.1111/j.1467-2979.2011.00435.x.
Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70, 41–55. https://doi.org/10.1093/biomet/70.1.41.
Rosenberg, A.A., Fogarty, M.J., Cooper, A.B., 2014. Developing New Approaches to Global Stock Status Assessment and Fishery Production Potential of the Seas.
Rosenberg, A.A., Kleisner, K.M., Afflerbach, J., Anderson, S.C., Dickey-Collas, M., Cooper, A.B., Fogarty, M.J., Fulton, E.A., Gutiérrez, N.L., Hyde, K.J.W., Jardim, E., Jensen, O.P., Kristiansen, T., Longo, C., Minte-Vera, C.V., Minto, C., Mosqueira, I., Osio, G.C., Ovando, D., Selig, E.R., Thorson, J.T., Walsh, J.C., Ye, Y., 2017. Applying a new ensemble approach to estimating stock status of marine fisheries around the world. Conserv. Lett. 69, 1491–1499. https://doi.org/10.1111/conl.12363.
SAFMC, 2013. SAMFC SSC ORCS Workshop II. South Atlantic Fishery Management Council (SAFMC), North Charleston, SC.
SAFMC, 2012. SAFMC SSC ORCS Workshop Report. South Atlantic Fishery Management Council (SAFMC), North Charleston, SC.
Schaefer, M.B., 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Inter-Am. Trop. Tuna Comm. Bull. 1, 27–56.
Thorson, J.T., Branch, T.A., Jensen, O.P., Quinn, T., 2012. Using model-based inference to evaluate global fisheries status from landings, location, and life history data. J. Fish. Board Can. 69, 645–655. https://doi.org/10.1139/f2012-016.
Thorson, J.T., Cope, J.M., 2015. Catch curve stock-reduction analysis: an alternative solution to the catch equations. Fish. Res. 171, 33–41. https://doi.org/10.1016/j.fishres.2014.03.024.
Thorson, J.T., Minto, C., Minte-Vera, C.V., Kleisner, K.M., Longo, C., Jacobson, L., 2013. A new role for effort dynamics in the theory of harvested populations and data-poor stock assessment. J. Fish. Board Can. 70, 1829–1844. https://doi.org/10.1139/cjfas-2013-0280.
Vasconcellos, M., Cochrane, K., 2005. Overview of world status of data-limited fisheries: inferences from landings statistics. In: Kruse, G.H., Gallucci, V.F., Hay, D.E., Perry, R.I., Peterman, R.M., Shirley, T.C., Spencer, P.D., Wilson, B., Woodby, D. (Eds.), Fisheries Assessment and Management in Data-Limited Situations. Alaska Sea Grant College Program.
Walsh, J.C., Minto, C., Jardim, E., Anderson, S.C., Jensen, O.P., Afflerbach, J., Dickey-Collas, M., Kleisner, K.M., Longo, C., Osio, G.C., Selig, E.R., Thorson, J.T., Rudd, M.B., Papacostas, K.J., Kittinger, J.N., Rosenberg, A.A., Cooper, A.B., 2018. Trade-offs for data-limited fisheries when using harvest strategies based on catch-only models. Fish Fish. 4, e4570–17. https://doi.org/10.1111/faf.12316.
Walters, C.J., Martell, S.J.D., Korman, J., 2006. A stochastic approach to stock reduction analysis. Can. J. Fish. Aquat. Sci. 63, 212–223. https://doi.org/10.1139/f05-213.
Wiedenmann, J., Free, C.M., Jensen, O.P., 2019. Evaluating the performance of data-limited methods for setting catch targets through application to data-rich stocks_ A case study using Northeast U.S. Fish stocks. Fish. Res. 209, 129–142. https://doi.org/10.1016/j.fishres.2018.09.018.
Winker, H., Carvalho, F., Sharma, R., Parker, D., Kerwath, S., 2017. Initial results for north and south Atlantic shortfin mako (Isurus oxyrinchus) stock assessments using the Bayesian surplus production model JABBA and the catch-resilience method CMSY. Collect Vol Sci Pap ICCAT 74, 1836–1866.
Worm, B., Barbier, E., Beaumont, N., Duffy, J., Folke, C., Halpern, B., Jackson, J., Lotze, H., Micheli, F., Palumbi, S., 2006. Impacts of biodiversity loss on ocean ecosystem services. Science 314, 787–790.
Worm, B., Hilborn, R., Baum, J., Branch, T., Collie, J., Costello, C., Fogarty, M., Fulton, E., Hutchings, J., Jennings, S., 2009. Rebuilding global fisheries. Science 325, 578–585.
Zhang, K., Zhang, J., Xu, Y., Sun, M., Chen, Z., Yuan, M., 2018. Application of a catch-based method for stock assessment of three important fisheries in the East China Sea. Acta Oceanolog. Sin. 37, 102–109. https://doi.org/10.1007/s13131-018-1173-9.
Zhou, S., Chen, Z., Dichmont, C.M., Ellis, N., Haddon, M., Punt, A.E., Smith, A.D.M., Smith, D.C., Ye, Y., 2016. Catch-Based Methods for Data-Poor Fisheries. The Food and Agriculture Organization of the United Nations.
Zhou, S., Punt, A.E., Smith, A., of, Y.Y.I.J, 2018. An optimized catch-only assessment method for data poor fisheries. 2017. ICES J. Mar. 75, 964–976. https://doi.org/10.1093/icesjms/fsx226.
Zhou, S., Punt, A.E., Ye, Y., Ellis, N., Dichmont, C.M., Haddon, M., Smith, D.C., Smith, A.D., 2017. Estimating stock depletion level from patterns of catch history. Fish Fish. 18, 742–751. https://doi.org/10.1111/faf.12201.