Estimating stock depletion level from patterns of catch history

Journal:	<i>Fish and Fisheries</i>
Manuscript ID	Draft
Wiley - Manuscript type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Zhou, Shijie; CSIRO, Division of Marine and Atmospheric Research Punt, André; CSIRO, Oceans and Atmosphere Flagship Ye, Yimin; Food and Agriculture Organization of the United Nations, Fisheries and Aquaculture Ellis, Nick; CSIRO Dichmont, Catherine; CSIRO, Oceans and Atmosphere Flagship Haddon, Malcolm; CSIRO, Wealth from Oceans Smith, David; CSIRO, Marine Smith, Tony; CSIRO, Marine
Key terms:	catch-based, depletion, data-poor, stock status, unassessed stock, RAM legacy database
Abstract:	<p>The degree to which a stock is depleted is one of the most important quantities in fisheries management because it is used to quantify the success of management and to inform management responses. However, stock depletion is extremely difficult to estimate, particularly with limited data. Using the RAM Legacy database, we developed a boosted regression tree (BRT) model to correlate depletion with a range of predictors calculated from catch data, making the model usable for many fisheries worldwide. The most important predictors were found to be catch trends obtained from linear regressions of scaled catch on time, including regression coefficients for the whole catch time series, the sub-series before and after the maximum catch, and in recent years. Eight predictors explain about 80% of variation in depletion. There is a correlation of 0.5 between measured levels of depletion and the predictions of the BRT model. Predictions are less biased when the stock is fished down below half of the carrying capacity. The BRT model outperforms comparable existing catch-based methods, and could be used to provide priors for depletion for data-poor stock assessment methods, or used more directly to provide estimates of the probability that depletion is below a given threshold value.</p>

For Review Only

Click here to enter text.

Title: **Estimating stock depletion level from patterns of catch history**

Alternative 1: Catch-only boosted regression trees for estimating stock depletion

Alternative 2: Estimating stock depletion from catch history using boosted regression trees

Shijie Zhou^{a*}, André E. Punt^{b,c}, Yimin Ye^d, Nick Ellis^a, Cathy M. Dichmont^a, Malcolm Haddon^b, David C. Smith^b, Anthony D.M. Smith^b

^a*CSIRO Oceans and Atmosphere, P.O. Box 2583, Brisbane, QLD 4001, Australia*

^b*CSIRO Oceans and Atmosphere, Castray Esplanade, Hobart, TAS, 7001, Australia*

^c*School of Aquatic and Fishery Sciences, Box 355020, University of Washington, Seattle, WA 98195, USA*

^d*FAO Marine and Inland Fisheries Branch, Viale delle Terme di Caracalla, 00153 Rome, Italy*

*Correspondence: Tel: +61 7 38335968; Fax: +61 7 38335508; E-mail: Shijie.zhou@csiro.au

Running title: Estimating depletion from catch history

[Click here to enter text.](#)

19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42

Abstract

The degree to which a stock is depleted is one of the most important quantities in fisheries management because it is used to quantify the success of management and to inform management responses. However, stock depletion is extremely difficult to estimate, particularly with limited data. Using the RAM Legacy database, we developed a boosted regression tree (BRT) model to correlate depletion with a range of predictors calculated from catch data, making the model usable for many fisheries worldwide. The most important predictors were found to be catch trends obtained from linear regressions of scaled catch on time, including regression coefficients for the whole catch time series, the sub-series before and after the maximum catch, and in recent years. Eight predictors explain about 80% of variation in depletion. There is a correlation of 0.5 between measured levels of depletion and the predictions of the BRT model. Predictions are less biased when the stock is fished down below half of the carrying capacity. The BRT model outperforms comparable existing catch-based methods, and could be used to provide priors for depletion for data-poor stock assessment methods, or used more directly to provide estimates of the probability that depletion is below a given threshold value.

Keywords: catch-based, depletion, data-poor, stock status, unassessed stock, RAM legacy database, regression trees

Table of contents

- Abstract
- Introduction
- Methods and materials

[Click here to enter text.](#)

43	Data sources
44	Boosted regression trees (BRT)
45	Predictors for depletion
46	Modelling
47	Prediction
48	Distribution of predicted depletion
49	Stock status classification
50	Results
51	Model fitting
52	BRT prediction
53	Distribution of predicted depletion
54	Stock status classification
55	Comparison with panel regression approach and catch-based classification
56	Discussion
57	Acknowledgement
58	References

59

60 Introduction

61 The traditional approach to fisheries stock assessment assumes that there is a constant
 62 theoretical carrying capacity for each fish stock. This capacity (K) is the equilibrium biomass,
 63 in the absence of fishing, and is typically estimated, albeit with possibly large uncertainty,
 64 using quantitative stock assessment models. Fishing will reduce fish biomass below K .
 65 Depletion (d) for year y is defined here as one less the ratio of the biomass in year y (B_y) to K
 66 (i.e. $d_y = 1 - B_y/K$), such that depletion is zero when the stock is unfished and increases to a
 67 maximum of one if the stock is extirpated. The depletion status of global fish stocks has

[Click here to enter text.](#)

68 attracted substantial interest. For example, Grainger and Garcia (1996) developed a simple
69 catch trend analysis to evaluate global fisheries. This analysis was based on an idealised
70 model of an open access fishery. In the course of its development, a fishery can be considered
71 to pass through four phases: (1) undeveloped, (2) developing, (3) mature, and (4) senescent.
72 The method of Grainger and Garcia (1996) has been modified over the past two decades,
73 resulting in several catch-based methods that estimate stock status by comparing annual catch
74 to the historical maximum catch (e.g. Froese and Kesner-Reyes 2002; Pauly 2008; Kleisner
75 and Pauly 2011). This way to assess fisheries has, however, received widespread criticism
76 (Daan *et al.* 2011; Branch *et al.* 2011), leading to additional research (e.g. Anderson *et al.*
77 2012; Carruthers *et al.* 2012). While the value of using only catch data to classify fisheries
78 continues to be subject to active debate in the literature (e.g. Froese *et al.* 2012a; Pauly *et al.*
79 2013; Cook 2013), in many fisheries catch is the only data available.

80 An alternative method for estimating stock status was developed by Costello *et al.* (2012)
81 using catch data and additional information. The approach involved a panel regression model
82 (PRM) to fit data from the assessed fisheries in the RAM legacy database (RAMLD, a global
83 compilation of data on fish stock statistics; Ricard *et al.* 2012). A log-linear model linked
84 $\log(B_y/B_{msy})$ to a set of predictors, which included catch data, fishery development and
85 biological characteristics, a time trend, and fixed effects for species type. The regression
86 coefficient estimates were then used to predict B_y/B_{msy} for unassessed stocks. However, the
87 estimates from this approach had very high uncertainty and large biases for the stocks in the
88 RAMLD, over-predicting B_y/B_{msy} when the true stock size was less than B_{msy} and under-
89 predicting it when the true stock size was greater than B_{msy} (Supplementary material to
90 Costello *et al.* 2012; Costello *et al.* 2016). Furthermore, some life-history information used in
91 the PRM is often unavailable for many fisheries. The PRM approach was modified by
92 Rosenberg *et al.* (2014) so that it requires only a catch time series and broad life-history

[Click here to enter text.](#)

information that can be obtained for most fish stocks. This modified panel regression model (mPRM) uses the same catch-related variables as Costello *et al.* (2012), but only one source of life-history information: whether the species is a demersal fish, a small pelagic fish, or a large pelagic fish. The regression R^2 was low for these PRMs when applied to the RAMLD stocks (0.143 in Rosenberg *et al.* 2014 and from 0.21 to 0.45 in Costello *et al.* 2012 that used more life-history data).

Several recent data-poor stock assessment methods (e.g. MacCall 2009; Dick and MacCall 2011; Martell and Froese 2013) require priors on stock depletion. The main objective of this paper was to derive a prior on stock depletion for use in a catch-only assessment method (Zhou and Sharma 2013; IFOP 2014; Secretariat *et al.* 2015; IOTC 2016). We develop boosted regression tree (BRT) models to link stock depletion to catch history information using data-rich stocks that have been assessed using quantitative methods of stock assessment. The BRT model uses only information inferred from the catch time-series. It does not use data such as life history parameters or ecological information because of the potential prediction bias associated with these parameters. A bias may be associated with methods that use life history traits as predictors and fit to data for the stocks in RAMLD, because the stocks in the RAMLD tend to be well-managed, whereas many stocks to which catch-only methods may be applied are not well-managed. For example, PRMs estimate that the resilience of stocks in RAMLD is negatively correlated with B_y/B_{msy} , i.e., low productive stocks tend to have a high B_y/B_{msy} ratios (Costello *et al.* 2012, 2016), which is counterintuitive (also see discussion in Froese and Kesner-Reyes 2002). We anticipate that catch history contains unbiased information for determining depletion because fisheries removal, whether it is managed or not, will change biomass and hence depletion. Here, the quantity we attempt to model and predict is $S_y = B_y/K$, which we refer to as “stock saturation”. Because this continuous quantity can be converted into a categorical variable, the method can

Click here to enter text.

also be used for stock status classification, e.g., to assess whether the stock is developing, mature, or overfished.

Methods and materials

Data sources

The RAM Legacy database (<http://ramlegacy.org>) is considered the most quantitatively robust source of fishery data generally available (Ricard *et al.* 2012). It provides estimates of a range of biological and management quantities, including biomass, fishing mortality rates, catch, life history parameters, and biological reference points. Stock saturation can be derived from estimated biomass and associated reference points, e.g. virgin biomass B_0 , which we assume to be a proxy for K . The RAMLD contains several biomass-based reference points, including unfished biomass B_0 , unfished spawning stock biomass SSB_0 , biomass and the spawning biomass corresponding to MSY (B_{msy} and SSB_{msy}), biomass limit reference points, and precautionary biomasses. Some stocks have multiple reference point estimates, from which we derive relationships between the reference points:

$$B_{msy}/B_0 = 0.336 \text{ (SD} = 0.068, n = 31)$$

$$B_{lim}/B_0 = 0.247 \text{ (SD} = 0.162, n = 4)$$

$$SSB_{lim}/SSB_{pa} = 0.668 \text{ (SD} = 0.072, n = 28)$$

$$B_{msy}/B_{40\%} = 0.868 \text{ (SD} = 0.025, n = 11)$$

We assumed proxies for B_0 of $B_{msy}/0.35$, of $B_{lim}/0.20$, of $B_{pa}/0.25$, or of $2.5B_{40\%}$ for stocks that do not have values for B_0 . We used $B_{lim} = 0.20B_0$ instead of $0.25B_0$ because of the high variance of the mean, and because it is often assumed that $B_{lim} = 0.5B_{msy}$ and $B_{msy} = 0.40B_0$ (for example, see Rayns 2007). The ratio B_{pa}/B_{lim} has been estimated to lie between 1.39 and 1.64 (Cadima 2003), and the 95% confidence interval for B_{pa}/B_0 to be 0.28–0.44 (Froese and Proelß 2010). These results, together with our estimate from the RAMLD led us to adopt $B_0 =$

[Click here to enter text.](#)

$B_{pa}/0.25$. There are 386 stocks in RAMLD that have time series of catches. However, many stocks do not have any associated biomass reference points (e.g. virgin biomass). In addition, some stocks in the RAMLD have saturation levels, $S > 1$, indicating very low fishing intensity (or potential underestimation of B_0 or overestimation of B_y), which is not common for data poor stocks, and hence of little interest to this study. The final dataset of stocks with a catch history, a saturation level $S < 1.0$, and at least one biomass reference point, contains 191 stocks. The catch history durations range from 10 to 129 years, with a mean of 46 years, and the saturations range from 0.0003 to 0.982, with a mean of 0.367.

Boosted regression trees (BRT)

BRT is a relatively new technique that combines statistical analysis and machine learning (Ferrier *et al.* 2006; Elith *et al.* 2008; Elith and Leathwick 2016). The BRT approach differs from traditional regression methods in several ways. Using a training data set, the regression trees partition each X space into disjoint small regions and identify regions with the most consistent responses to the predictors. A model is then fit to the mean response for the observations in each region. The recursive partitioning process can accommodate all types of predictor variables, including numeric, binary, categorical, etc. Interactions among predictors are automatically included in the modelling process and model outcomes are insensitive to outliers and transformations of the explanatory variables. BRT uses cross validation to find the best regression trees by randomly dividing the data into a training set and a testing set. Boosting is used to improve model accuracy. This involves merging results from multiple models where many regression trees (typically ~1,000 – 2,000 for this paper) are fitted iteratively to the training data to optimize predictive performance. In common with other tree methods, BRT is nonparametric and nonlinear, i.e. there is no implicit assumption that the underlying relationships between the predictor variables and the dependent variable are linear, follow some specific non-linear link function, or are monotonic in nature. Thus, BRT

[Click here to enter text.](#)

methods are particularly well suited for data mining tasks, where often there is neither little *a priori* knowledge nor any coherent set of theories regarding which variables are related and how. However, tree-based methods may have difficulty in modelling smooth functions, and small changes in training data can result in different series of splits (Elith *et al.* 2008).

We use two R packages (R Core Team 2013) for the BRT analysis: *gbm* and *dismo*. The *dismo* package has functions that can be used to determine the optimal number of trees and the optimal number of predictors, which enhances the performance of *gbm* (see below).

Predictors for depletion

We have a training sample of 191 stocks from the RAMLD. This information is used to find a model to predict S for a new stock based on the X values for the stock. We examine 56 candidate predictors (Table 1). These predictors are based on scaled catch ($c_y = C_y/C_{\max}$, the annual catch in year y divided by the maximum historical catch), various linear regressions of scaled catch over time, intercept, slopes, and break point of a two-segmented regression of scaled catch on time, number of years in the catch history, number of years before the maximum catch divided by total number of years, mean scaled catch, etc. The linear regressions of scaled catch imply an increasing or decreasing catch trend during various time periods. The predictors based on these regressions include using all years, the years before or after catch reached its maximum value, the first 2 to 8 years, the last 2 to 8 years (so we exclude stocks with fewer than eight years of data), and segmented regressions that statistically break the trend into two linear segments (e.g. increasing and decreasing periods or two increasing periods of different slopes).

Modelling

We considered three possible response variables: S , $\log(S)$, and $\text{logit}(S)$, assuming a Gaussian distribution on the transformed data. Comparisons of the correlation coefficient between the

[Click here to enter text.](#)

fitted and the observed S and the deviance (the mean square error (MSE) for a Gaussian distribution in BRT) reveal similar results, but using S leads to the highest correlation and lowest deviance. Hence, S is used for all further analyses. The BRT method requires tuning several control parameters to achieve the best performance. We carried out the following explorations.

- *Tree complexity*: this parameter controls the interaction among predictors. We examined 15 models with tree complexity from 1 to 15, and compared the alternative models using model fit correlation coefficient, cross-validation correlation coefficient, deviance between fitted and observed S , and cross-validation deviance. The level of tree complexity is not well resolved using these criteria, but we found that a tree complexity < 5 has a relatively poor fit while a tree complexity of 7 generally has the highest correlation coefficient and lowest deviance for the training dataset.
- *Learning rate*: the learning rate is also referred to as the shrinkage parameter and determines the contribution of each tree to the growing model. We tested learning rates of 0.01, 0.005, and 0.001. Lower learning rates lead to more accurate predictions, at the expense of longer computation times. Nevertheless, we found the computation times acceptable and so we adopted a learning rate of 0.001 for further analysis.
- *Bagging fraction*: this is the fraction of the training set randomly selected to propose the next tree in the expansion. Bagging introduces randomness into the individual trees thus reducing their correlation. We tested 10 bagging fractions between 0.4 and 0.9, and found a bagging fraction of 0.75 generally has a lower deviance, so this rate is used in further analysis.

[Click here to enter text.](#)

- *Number of trees*: the function `gbm.step` in the `dismo` package can search for the best number of trees based on the change in cross-validated deviance so this function is applied throughout.
- *Number of predictors*: we use the function `gbm.simplify` in the `dismo` package to calculate the change in deviance after removing each predictor. This function assesses the potential to remove predictors using ten-fold cross-validation, a BRT internal procedure to validate the model (this validation should not be confused with leave-one-out cross-validation (LOOCV) below). This is conducted for each fold, removing the least influential predictor, and repeating this process for a set number of steps. After the removal of each predictor, the change in predictive deviance is computed relative to that obtained using all predictors. The optimal number of dropped predictors to attain the minimum deviance may vary between model runs. The final list of predictors is obtained by ranking the results of multiple model runs.

It is not uncommon that tree-based methods yield biased results that should be corrected (Zhang and Lu 2012; Song 2015). Plotting the fitted values against the observed values typically shows that smaller S are overestimated while larger S are underestimated. We tested several correction methods including linear regression, the residual adjustment method of Zhang and Lu (2012), and the residual rotation method of Song (2015). The results indicate that, although the more sophisticated methods (i.e., the residual adjustment method and the residual rotation method) could improve the fitted model well (increasing R^2 from 0.7 to 0.99), for prediction they do not perform any better than the simpler linear regression. Hence, linear regression is adopted here, in which the BRT-fitted S is adjusted using a linear regression of the observed on the fitted values. To correct the bias, we therefore first perform a linear regression:

$$\hat{S}_{fit} = a + bS_{obs} \quad (1)$$

[Click here to enter text.](#)

239 where $\hat{S}_{fit}$ is the BRT-fitted S , S_{obs} is the observed S in RAMLD, and a and b are the
 240 intercept and slope of the regression. We then obtain the bias-corrected estimate (constrained
 241 to be no smaller than zero) as:

$$242 \quad \hat{S}_{fit}^{bc} = \max[(\hat{S}_{fit} - a)/b, 0] \quad (2)$$

243 **Prediction**

244 Prediction is assessed using a LOOCV routine. Each stock is sequentially removed from the
 245 191 stocks used to fit the BRT models. The resulting BRT from the 190 stocks is then used to
 246 predict $\hat{S}_{pred,s}$ for the omitted stock s . This prediction is biased so the regression approach is
 247 used to correct the bias:

$$248 \quad \hat{S}_{pred,s}^{bc} = \max[(\hat{S}_{pred,s} - a)/b, 0] \quad (3)$$

249 where $\hat{S}_{pred,s}^{bc}$ is LOOCV-predicted and bias-corrected value for stock s , and parameters a and
 250 b are the regression coefficients from the 190 stocks. We use mean absolute error (MAE) for
 251 model evaluation:

$$252 \quad MAE = \frac{1}{n} \sum_{s=1}^{s=n} |\hat{S}_{pred,s}^{bc} - S_{obs,s}| \quad (4)$$

253 where $n = 191$, and $S_{obs,s}$ is the observed value of S for stock s in the RAMLD.

254 **Distribution of predicted depletion**

255 One of the key objectives of this study is to develop a prior for S for use in data poor
 256 methods. Since the predicted S has a skewed distribution biased toward overestimating S
 257 when the true S is low and underestimating S when the true S is large, we construct two
 258 skewed normal distributions, $f_1(\text{mean} = m_1, \text{standard deviation} = sd_1, \text{skewness} = \kappa_1)$ for $S \leq$
 259 0.5 and $f_2(m_2, sd_2, \kappa_2)$ for $S > 0.5$. The parameters of these distributions are estimated from

[Click here to enter text.](#)

the residuals between the bias-corrected predictions $\hat{S}_{pred}^{bc}$ and the S_{obs} . For any new stocks, if $\hat{S}_{pred}^{bc} \leq 0.5$ then the prior for S should be $f_1(m_1 + \hat{S}_{pred}^{bc}, sd_1, \kappa_1)$. Otherwise, it should be $f_2(m_2 + \hat{S}_{pred}^{bc}, sd_2, \kappa_2)$.

Stock status classification

Although quantifying stock status is not our key objective, the method can be used for this purpose. This requires dividing $\hat{S}_{pred}^{bc}$ values into several bins. We adopt similar thresholds to those used in other studies (e.g., Branch *et al.* 2011; Anderson *et al.* 2012): $\hat{S}_{pred}^{bc} > 0.75 =$ developing, $0.75 \geq \hat{S}_{pred}^{bc} > 0.25 =$ fully exploited, $0.25 \geq \hat{S}_{pred}^{bc} > 0.1 =$ overexploited, and $\hat{S}_{pred}^{bc} \leq 0.1 =$ collapsed. A 4×4 contingency table is produced to display the predicted frequency in each cell. The classification success rate is the proportion of stocks in a given category that are correctly assigned to that category. The overall prediction success rate is the total number of correct predictions divided by the total number of stocks (i.e. 191). Alternatively, we can calculate the number of stocks predicted to be in each category (regardless of their true status), compared with the true number of stock in the same category. This classification is informative for displaying the overall stock status in a region, for example the global fisheries status. This relative error (RE) is computed as $RE_g = (\hat{N}_g - N_g) / N_g$, where $\hat{N}_g$ is the number of stocks predicted to be in category g , and N_g is the actual number stocks in that category.

Comparison with panel regression and stock status classification methods

Rosenberg *et al.* (2014)'s modified panel regression models (mPRM Costello *et al.* 2012, 2016) estimate $\log(B_y/B_{msy})$ according to:

Click here to enter text.

$$\log\left(\frac{B}{B_{\text{msy}}}\right)_{ijt} = \alpha + \beta X_{ijt} + \gamma_j + \varepsilon_{ijt} \quad (5)$$

where i denotes fishery, j species type, t time, α is a constant, β are the regression coefficients for the variables X , γ_j is a species-type fixed effect, and ε_{ijt} is an error term. Rosenberg *et al.* used ten predictors derived from catch history in addition to a species type variable that categorized species into three groups: large pelagic fish, small pelagic fish, and demersal fish. The model was fitted to 166 stocks in the RAMLD. To compare the performance of the mPRM with our BRT model, we apply the same technique and the same predictors as in Rosenberg *et al.* (2014) to the 191 stocks that are analysed using the BRT model. However, one difference between our analysis and that of Rosenberg *et al.* (2014) is that the dependent variable used here is B_y/K instead of B_y/B_{msy} . Prediction bias for the mPRM is corrected using linear regression in the same way as for BRT. R^2 , visual comparison, the MAE and classification performance are used to compare the mPRN and the MAE.

We also compared stock status classification using mPRM and a catch-based classification method (Froese and Kesner-Reyes 2002; Zeller *et al.* 2009; Branch *et al.* 2011). For mPRM, the predicted $\hat{S}_{pred}^{bc}$ is assigned into one of the four categories as for the BRT. For catch-based classification, we use the criteria described in Froese and Kesner-Reyes (2002) and Branch *et al.* (2011). The time series of catches is divided into before and after the year of the maximum catch. In the first period, fisheries are classified as either developing ($C_y < 0.5 C_{\text{max}}$) or fully exploited ($C_y \geq 0.5 C_{\text{max}}$). In the second period, fisheries are classified as either fully exploited ($C_y \geq 0.5 C_{\text{max}}$), overexploited ($0.1-0.5 C_{\text{max}}$), or collapsed ($C_y < 0.1 C_{\text{max}}$).

[Click here to enter text.](#)

304 Results

305 Model fitting

306 Several of the 56 potential predictors are highly important contributors irrespective of the
 307 model structure and the specification of how the BRT method is implemented. The
 308 breakpoint of the trends and the regression slopes of catch trends generally have the highest
 309 influence on S . Other important predictors include number of years in the catch history,
 310 average catch, and the catch in recent years. With these predictors in the model, dropping
 311 additional predictors does not substantially increase the deviance. Several models with
 312 varying number of predictors perform very similarly. For example, the models with 8, 11, 19,
 313 35, and 38 predictors have the same mean deviances, but the variance increases with fewer
 314 predictors. The model that includes eight predictors (Table 1) has a training data correlation
 315 of 0.864 (which corresponds to R^2 in linear regression models) and a ten-fold cross-validation
 316 (i.e., model prediction) correlation of 0.497.

317 The BRT model fits the observed saturations fairly well, although it tends to underestimate S
 318 when stocks are not heavily fished ($S_{\text{obs}} > 0.6$). For the model with eight predictors, the bias-
 319 correction regression has parameters $a = 0.469$, $b = 0.199$, and $R^2 = 0.756$ (Figure 1).

320 BRT prediction

321 The $\hat{S}_{pred}^{bc}$ has a weaker correlation with S_{obs} than $\hat{S}_{fit}^{bc}$: the correlation coefficient between
 322 $\hat{S}_{pred}^{bc}$ and S_{obs} is 0.497 (compared to $r = 0.756$ for model fit) using the BRT model with eight
 323 predictors. The $\hat{S}_{pred}^{bc}$ is related linearly to S_{obs} when the $S_{\text{obs}} < 0.6$ (Figure 2). However, the
 324 model under-predicts S when the stocks are only lightly fished ($S_{\text{obs}} > 0.6$). The overall mean
 325 squared error of $\hat{S}_{pred}^{bc}$ is 0.231. The MAE between $\hat{S}_{pred}^{bc}$ and S_{obs} is 0.182.

[Click here to enter text.](#)

326 **Distribution of predicted depletion**

327 Two skewed normal distributions are used to represent the residuals (Figure 3). For $f_1(S \leq$
 328 $0.5)$, the three parameters (mean, standard deviation and skewness) are -0.072 , 0.189 , and
 329 0.763 ; for $f_2(S > 0.5)$, they are 0.179 , 0.223 , and 0.904 . Simulation of S using these
 330 distributions indicates better, but still imperfect, performance for $S_{\text{obs}} > 0.6$ (Figure 4).

331 **Stock status classification**

332 The success rate for predicting stock status ranges from 8% for the developing category to
 333 72% for the fully exploited category (Table 2). The overall success rate is 56%. The relative
 334 errors range from -38% in the developing category to 13% in the fully exploited category
 335 (Table 2). The mean absolute relative error is 20%. Thus, for example, 79 of the 109 stocks
 336 that were actually fully exploited were classified as such (72%). In contrast, 123 stocks were
 337 classified being fully-exploited, which is 13% more than the actual number of stocks actually
 338 classified this way.

339 **Comparison with panel regression approach and catch-based classification**

340 We use the same formulation (equation 5) and predictors as in Rosenberg *et al.* (2014) for
 341 modelling S . However, using a log-transformed dependent variable (i.e. $\log(B_y/K)$) results in
 342 a poor model fit, with an $R^2 = 0.11$. Hence, B_y/K is directly used as the dependent variable in
 343 the mPRM as in the BRT analysis. This model yields a $R^2 = 0.147$ (very close to 0.143 in
 344 Rosenberg *et al.* 2014). Without bias correction, the mPRM predicted $\hat{S}_{\text{pred}}$ shrinks to a
 345 narrow range around the mean S_{obs} , i.e., overestimating S when the true value is low, but
 346 underestimating it when it is actually high, as reported in other studies (Costello *et al.*, 2012,
 347 2016). With bias correction, the $\hat{S}_{\text{pred}}^{bc}$ for mPRN is more uncertain than that from the BRT
 348 prediction (Figure 5). The R^2 between $\hat{S}_{\text{pred}}^{bc}$ for mPRN and S_{obs} is 0.108, and the mean

[Click here to enter text.](#)

absolute error MAE is 0.403. The success rate for predicting stock status ranges from 5% for the overexploited category to 76% for the collapsed category (Table 2). The overall success rate is 35%. The relative error for mPRM ranges from -73% for the overexploited category to 269% for the developing category. The mean absolute RE is 141%.

The catch-based classification method tends to be more conservative than the BRT method, correctly classifying more stocks in the overexploited and collapsed categories, but fewer in the developing and fully exploited categories (Table 2). The success rate for predicting the status ranges from 0 for the developing category to 56% for the collapsed category. The overall success rate is 42%. The relative error of prediction ranges from -100% in the developing category to 91% in the overexploited category. The mean absolute RE is 76%.

Discussion

Estimating stock depletion is very difficult for the majority of fish stocks. Even for data-rich species, alternative assumptions and models may produce quite different estimates (e.g., Tuck 2013). Estimation of depletion is even more challenging for stocks that only have catch data. Encouragingly, our study demonstrates high correlation (average $R^2 = 0.81$) of the model fit using various aspects of catch history. Prediction is more accurate under certain conditions, i.e. when the stocks have been fished below 60% of unfished population size.

The use of various catch trends captured by linear regressions may have contributed to improved performance of the BRT relative to other methods. Parameters from regressions of scaled catch over the whole catch history, before and after the maximum catch, and from regressions that break the catch history into distinguishable periods always made significant contributions to the model. The amount of catch, whether it is managed or unmanaged, instantaneously changes stock biomass and possibly the ratio of B_y/K . the complete catch

[Click here to enter text.](#)

history would conceivably influence the stock status if this change in biomass is large enough to affect population growth.

Using various catch trends to infer stock depletion status is essentially based on the same theory that has received extensive discussion in the literature (Froese and Kesner-Reyes 2002; Zeller *et al.* 2009; Anderson *et al.* 2012; Froese *et al.* 2012b). The major difference between our approach and other catch-based methods is the use of various regression coefficients to convey catch trends instead of focusing on the catch in individual years. We find that there indeed exists a positive and significant correlation between scaled catch in the most recent year and depletion ($R^2 = 0.083$, not included in the Results), but its contribution to prediction power is weaker than many other catch trend indicators.

Similar to some other catch-based stock status classification methods (Froese and Kesner-Reyes 2002; Zeller *et al.* 2009; Anderson *et al.* 2012; Froese *et al.* 2012b), our approach is based only on catch history. Hence, the method can be applied to stocks that have catch data, avoiding additional information that may not be available.

Although BRT can predict B_y/K reasonably well for heavily fished stocks, the method performs poorly for lightly fished stocks. This is likely inevitable with any catch-based method simply because the average catch is not large enough to significantly impact the population dynamics. We have not found a way to effectively overcome this difficulty by using catch data. Despite the low prediction power for lightly fished stock, our approach is more accurate than other methods, such as the recently developed mPRM and catch-based stock classification.

Depletion level is an essential input parameter for several data-poor assessment methods. For example, the depletion-corrected average catch method (MacCall 2009) assumed a reasonable value such as $B_y/K = 0.5$. Similarly, the Depletion-Based Stock Reduction

[Click here to enter text.](#)

Analysis (DB-SRA, Dick and MacCall 2011) used a bounded beta distribution with $E[B_y/K] = 0.4$, $SD[B_y/K] = 0.1$ and upper and lower bounds of 0.99 and 0.01, respectively. The Catch-MSY method (Martell and Froese, 2013) assumed a B_y/K prior based on the last year's scaled catch: $C_y/C_{\max} > 0.5$ then $B_y/K \in [0.3, 0.7]$, and $C_y/C_{\max} \leq 0.5$ then $B_y/K \in [0.01, 0.4]$; although these particular ranges can be expanded. The method described here should provide a more objective and accurate depletion prior for these assessment models.

Acknowledgement

This project is funded by FAO and CSIRO.

References

- Anderson, S.C., Branch, T.A., Ricard, D. and Lotze, H.K. (2012) Assessing global marine fishery status with a revised dynamic catch-based method and stock-assessment reference points. *ICES Journal of Marine Science* **69**, 1491–1500.
- Branch, T. a, Jensen, O.P., Ricard, D., Ye, Y. and Hilborn, R. (2011) Contrasting global trends in marine fishery status obtained from catches and from stock assessments. *Conservation Biology* **25**, 777–786.
- Cadima, E.L. (2003) Fish stock assessment manual. FAO Fisheries Technical Paper 393. Food and Agriculture Organization of the United Nations, Rome.
- 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* **119-120**, 66–79.
- Cook, R.M. (2013) A comment on “What catch data can tell us about the status of global fisheries” (Froese et al. 2012). *Marine Biology* **160**, 1761–1763.

[Click here to enter text.](#)

- 419 Costello, C., Ovando, D., Clavelle, T., et al. (2016) Global fishery futures under contrasting
420 management regimes. *Proceedings of the National Academy of Sciences of the United*
421 *States of America*, in review.
- 422 Costello, C., Ovando, D., Hilborn, R., Gaines, S.D., Deschenes, O. and Lester, S.E. (2012)
423 Status and Solutions for the World's Unassessed Fisheries. *Science* **338**, 517–520.
- 424 Daan, N., Gislason, H., Pope, J.G. and Rice, J.C. (2011) Apocalypse in world fisheries? The
425 reports of their death are greatly exaggerated. *ICES Journal of Marine Science* **68**,
426 1375–1378.
- 427 Dick, E.J. and MacCall, A.D. (2011) Depletion-Based Stock Reduction Analysis: A catch-
428 based method for determining sustainable yields for data-poor fish stocks. *Fisheries*
429 *Research* **110**, 331–341.
- 430 Elith, J. and Leathwick, J. (2016) Boosted Regression Trees for ecological modeling. 1–22.
- 431 Elith, J., Leathwick, J.R. and Hastie, T. (2008) A working guide to boosted regression trees.
432 *Journal of Animal Ecology* **77**, 802–813.
- 433 Ferrier, S., Guisan, A., Elith, J., et al. (2006) Novel methods improve prediction of species '
434 distributions from occurrence data. **2**.
- 435 Froese, R. and Kesner-reyes, K. (2002) ICES CM 2002 / L : 12 Impact of Fishing on the
436 Abundance of Marine Species. 1–12.
- 437 Froese, R. and Proelß, A. (2010) Rebuilding fish stocks no later than 2015: will Europe meet
438 the deadline? *Fish and Fisheries* **11**, 194–202.
- 439 Froese, R., Zeller, D., Kleisner, K. and Pauly, D. (2012a) What catch data can tell us about
440 the status of global fisheries. *Marine Biology* **159**, 1283–1292.
- 441 Froese, R., Zeller, D., Kleisner, K. and Pauly, D. (2012b) What catch data can tell us about

[Click here to enter text.](#)

- 442 the status of global fisheries. *Marine Biology* **159**, 1283–1292.
- 443 IFOP (2014) Review of biological reference points in domestic fisheries. Report of the Third
444 International Workshop Held in Viña del Mar, Chile, August 2014. Instituto de Fomento
445 Pesquero.
- 446 IOTC (2016) Assessment of Indian Ocean longtail tuna (*Thunnus tonggol*) using data poor
447 catch-based methods. Indian Ocean Tuna Commission. IOTC–2016–WPNT06–17.
- 448 Kleisner, K. and Pauly, D. (2011) Stock-status plots for fisheries for Regional Seas. In: *The*
449 *State of Biodiversity and Fisheries in Regional Seas*. (eds V. Christensen, S. Lai, M.L.D.
450 Palomares, D. Zeller and D. Pauly). Fisheries Centre Research Reports 19(3). Fisheries
451 Centre, University of British Columbia, pp 37–40.
- 452 MacCall, a. D. (2009) Depletion-corrected average catch: a simple formula for estimating
453 sustainable yields in data-poor situations. *ICES Journal of Marine Science* **66**, 2267–
454 2271.
- 455 Martell, S. and Froese, R. (2013) A simple method for estimating MSY from catch and
456 resilience. *Fish and Fisheries* **14**, 504–514.
- 457 Pauly, D. (2008) Global fisheries: a brief review. *Journal of Biological Research-*
458 *Thessaloniki* **9**, 3–9.
- 459 Pauly, D., Hilborn, R. and Branch, T.A. (2013) Does catch reflect abundance? *Nature* **494**,
460 303–306.
- 461 R Core Team (2013) R: A language and environment for statistical computing. R Foundation
462 for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- 463 Rayns, N. (2007) The Australian government's harvest strategy policy. *ICES Journal of*
464 *Marine Science* **64**, 596–598.

[Click here to enter text.](#)

- 465 Ricard, D., Minto, C., Jensen, O.P. and Baum, J.K. (2012) Examining the knowledge base
466 and status of commercially exploited marine species with the RAM Legacy Stock
467 Assessment Database. *Fish and Fisheries* **13**, 380–398.
- 468 Rosenberg, A.A., Fogarty, M.J., Cooper, A.B., et al. (2014) *Developing new approaches to*
469 *global stock status assessment and fishery production potential of the seas*, (Vol. 1086).
470 Fisheries and Aquaculture Circular No. 1086. Rome, FAO.
- 471 Secretariat, I., Biology, B., Trends, C., Catch, O. and Method, O. (2015) Assessment of
472 Indian Ocean Indo-Pacific king mackerel (*Scomberomorus guttatus*) using data poor
473 catch-based methods. 1–24.
- 474 Song, J. (2015) Bias corrections for Random Forest in regression using residual rotation.
475 *Journal of the Korean Statistical Society* **44**, 321–326.
- 476 Tuck, G.N. (2013) *Stock Assessment for the Southern and Eastern Scalefish*.
- 477 Zeller, D., Cheung, W., Close, C. and Pauly, D. (2009) Trends in global marine fisheries - a
478 critical review. *Fisheries, trade and development*, 55–77.
- 479 Zhang, G. and Lu, Y. (2012) Bias-corrected random forests in regression Bias-corrected
480 random forests in regression. *Journal of Applied Statistics* **39**, 151–160.
- 481 Zhou, S. and Sharma, R. (2013) Stock assessment of two neritic tuna species in Indian
482 Ocean: kawakawa longtail tuna using catch-based stock reduction methods. Indian
483 Ocean Tuna Commission. IOTC–2013–WPNT03–25.

484

[Click here to enter text.](#)

485 Table 1. Variables included in the boosted regression trees. The number in parentheses is the
486 number of predictors of each type. The third columns shows the relative influence of the eight
487 predictors in the final model.

Description	Predictors	Influence (%)
Intercept a and slope b of scaled catch (2)	a.AllYr	–
	b.AllYr	12.64
Intercept a and slope b of scaled catch before the maximum catch (2)	a.BfMax	–
	b.BfMax	15.58
Intercept a and slope b of scaled catch after the maximum catch (2)	a.AfMax	–
	b.AfMax	10.63
Intercept a and slope b of scaled catch for the beginning y years (14)	a.BgY _{y}	–
	b.BgY _{y}	($y=3$) 11.64
Intercept a and slope b of scaled catch for the last y years (14)	a.LsY _{y}	–
	b.LsY _{y}	($y=1, 5$) 10.55, 10.53
Scaled catch ($C_y/C_{\max}$) in year y (first 7 and last 8 years are used) (15)	c_y	–
Intercept a and slopes b of the segmented regression of scaled catch (3)	a.seg	–
	b1.seg	–
	b2.seg	11.41
Breakpoint year in the two-segment linear regression of scaled catch	BreakPoint	17.01
Number of years of catches	nyr	–
No. years before the maximum catch divided by the total number of years	nyrToCmaxR	–
Mean scaled catch	c_{mean}	–

Click here to enter text.

Table 2. Stock status predicted by BRT, mPRM, and catch-based classification for the 191 stocks in the RAMLD. “Success rate” is the percentage of stocks in a category that are correctly assigned to that category, and the relative error RE is the ratio of the difference between the number of predicted stocks in a category and the observed number of stocks in that category divided by the observed number of stocks. The numbers in bold and parentheses are correct predictions, and “absolute” is mean absolute RE over the four categories.

	Observed number of stocks N_g	Predicted number of stocks				Success rate	RE
		Developing	Fully exploited	Overexploited	Collapsed		
Boosted regression trees							
Developing	13	1	12	0	0	8%	−38%
Fully exploited	109	6	79	21	3	72%	13%
Overexploited	44	1	25	14	4	32%	−9%
Collapsed	25	0	7	5	13	52%	−20%
Total or mean	191	$\widehat{N}_g = 8$	123	40	20	(107) 56%	(absolute) 20%
Modified panel regression model							
Developing	13	6	6	0	1	46%	269%
Fully exploited	109	34	40	9	26	37%	−44%
Overexploited	44	7	11	2	24	5%	−73%
Collapsed	25	1	4	1	19	76%	180%
Total or mean	191	$\widehat{N}_g = 48$	61	12	70	(67) 35%	(absolute) 141%
Catch-based classification							
Developing	13	0	5	6	2	0%	−100%
Fully exploited	109	0	46	47	16	42%	−41%
Overexploited	44	0	12	21	11	48%	91%
Collapsed	25	0	1	10	14	56%	72%
Total or mean	191	$\widehat{N}_g = 0$	64	84	43	(81) 42%	(absolute) 76%

[Click here to enter text.](#)

Figure 1. $\hat{S}_{fit}^{bc}$ versus S_{obs} for the 191 stocks. The error bars represent 95% confidence intervals. The straight line is the 1:1 line. The R^2 is 0.756.

[Click here to enter text.](#)

Figure 2. $\hat{S}_{pred}^{bc}$ versus S_{obs} . The whiskers extend to the most extreme data points.

[Click here to enter text.](#)

Figure 3. Residual distribution of the bias-corrected saturation prediction for $S_{\text{obs}} \leq 0.5$ (left panel) and $S_{\text{obs}} > 0.5$ (right panel). The lines are theoretical distributions derived from fitting a skewed normal distribution to the residuals.

[Click here to enter text.](#)

Figure 4. Construction of predicted S distribution from 100 simulations based on the $\hat{S}_{pred}^{bc}$ and the residual distribution in Fig. 4 for the use as a prior distribution.

[Click here to enter text.](#)

Figure 5. $\hat{S}_{pred}^{bc}$ from the mPRM (blank) and BRT (grey, reproduced from Fig. 2) methods. To show the full spread of the predictions the predicted values are not constrained.