Received: 9 August 2016 | Accepted: 1 December 2016

DOI: 10.1111/faf.12201

ORIGINAL ARTICLE

WILEY FISH and FISHERIES

Estimating stock depletion level from patterns of catch history

Shijie Zhou[1] | André E Punt[2,3] | Yimin Ye[4] | Nick Ellis[1] | Cathy M Dichmont[1] | Malcolm Haddon[2] | David C Smith[2] | Anthony DM Smith[2]

[1]CSIRO Oceans and Atmosphere, Brisbane, Qld, Australia

[2]CSIRO Oceans and Atmosphere, Hobart, Tas., Australia

[3]School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA, USA

[4]FAO Marine and Inland Fisheries Branch, Rome, Italy

Correspondence
Shijie Zhou, CSIRO Oceans and Atmosphere, Brisbane, Qld, Australia.
Email: Shijie.zhou@csiro.au

Funding information
FAO; CSIRO

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 subseries before and after the maximum catch, and in recent years. Eight predictors explain about 80% of variation in depletion. There is a correlation of .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 depletion estimators 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, data-poor, depletion, RAM legacy database, regression trees, stock status, unassessed stock

1 | INTRODUCTION

The traditional approach to fisheries stock assessment assumes that there is a constant theoretical carrying capacity for each fish stock. This capacity (K) is the equilibrium biomass, in the absence of fishing, and is typically estimated, albeit with possibly large uncertainty, using quantitative stock assessment models. Fishing will reduce fish biomass below K. Depletion (d) for year y is defined here as one less the ratio of the biomass in year y (B_y) to K (i.e. $d_y = 1-B_y/K$), such that depletion is zero when the stock is unfished and increases to a maximum of one if the stock is extirpated. The depletion status of global fish stocks has attracted substantial interest. For example, Grainger and Garcia (1996) developed a simple catch trend analysis to evaluate global fisheries. This analysis was based on an idealized model of an open access fishery. In the course of its development, a fishery can be considered to pass through four phases: (i) undeveloped, (ii) developing, (iii) mature and (iv) senescent. The method of Grainger and Garcia (1996) has been modified over the past two decades, resulting in several catch-based methods that estimate stock status by comparing annual catch to the historical maximum catch (e.g. Froese & Kesner-reyes, 2002; Kleisner & Pauly, 2011; Pauly, 2008). This way to assess fisheries has, however, received widespread criticism (Branch, Jensen, Ricard, Ye, & Hilborn, 2011; Daan, Gislason, Pope, & Rice, 2011), leading to additional research (e.g. Anderson, Branch, Ricard, & Lotze, 2012; Carruthers, Walters, & McAllister, 2012). While the value of using only catch data to classify fisheries continues to be subject to active debate in the literature (e.g. Cook, 2013; Froese, Zeller, Kleisner, & Pauly, 2012; Pauly, Hilborn, & Branch, 2013), in many fisheries, catches are the only data available.

An alternative method for estimating stock status was developed by Costello *et al.* (2012) using catch data and additional information. The approach involved a panel regression model (PRM) to fit data from the assessed fisheries in the RAM legacy database (RAMLD, a global compilation of data on fish stock statistics; Ricard, Minto, Jensen, & Baum, 2012). A log-linear model linked $\log(B_y/B_{msy})$ to a set of predictors, which included catch data, fishery development and biological characteristics, a time trend and fixed effects for species type. The regression coefficient estimates were then used to predict B_y/B_{msy} for unassessed stocks. However, the estimates from this approach had very high uncertainty and large biases for the stocks in the RAMLD, over-predicting B_y/B_{msy} when the true stock size was less than B_{msy} and under-predicting it when the true stock size was greater than B_{msy} (Supplementary material to Costello *et al.*, 2012, 2016). Furthermore, some life-history information used in the PRM is often unavailable for many fisheries. The PRM approach was modified by Rosenberg *et al.* (2014) so that it requires only a catch time series and broad life-history 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. Dick & MacCall, 2011; MacCall, 2009; Martell & Froese, 2013) require priors on stock depletion. The main objective of this study was to derive a prior on stock depletion for use in a catch-only assessment method (IFOP 2014; Martin & Robinson, 2016; Martin & Sharma, 2015; Zhou & Sharma, 2013). 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. 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}; that is, low productive stocks tend to have a high B_y/B_{msy} ratios (Costello *et al.*, 2012, 2016), which is counter-intuitive (also see discussion in Froese & 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 $S_e = B_e/K$, (e denotes the end of catch time series) which we refer to as "stock saturation." Because this continuous quantity can be converted into a categorical variable, the method can also be used for stock status classification, for example, to assess whether the stock is developing, mature or overfished.

2 | METHODS AND MATERIALS

2.1 | 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, for example 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 (B_{lim}) and precautionary biomasses (SSB_{pa}). Stock sizes below B_{lim} are considered to increase the risk of stock collapse. The precautionary reference points are used in advice given by ICES for fisheries management. SSB_{pa} is the spawning biomass at and above which there is a low probability that true SSB is so low that productivity is impaired (Piet & Rice, 2004). Some stocks have multiple reference point estimates, from which we derive relationships between the reference points:

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

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

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

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

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 & Proelß, 2010). These results together with our estimate from the RAMLD led us to adopt $B_0 = B_{pa}/0.25$. 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 exploited data-poor stocks and hence of little interest to this study. The final data set of stocks with a catch history, a saturation level $S < 1.0$ and at least one biomass reference point contains 191 stocks that are used in this study (the listed stocks and their saturation levels can be found in the supplementary materials). The catch history durations range from 10 to 129 years, with a mean of 46 years, and the saturations range from .0003 to .944, with a mean of 0.368 (Figure 1).

2.2 | Boosted regression trees (BRT)

BRT is a relatively new technique that combines statistical analysis and machine learning (Elith, Leathwick, & Hastie, 2008; Ferrier

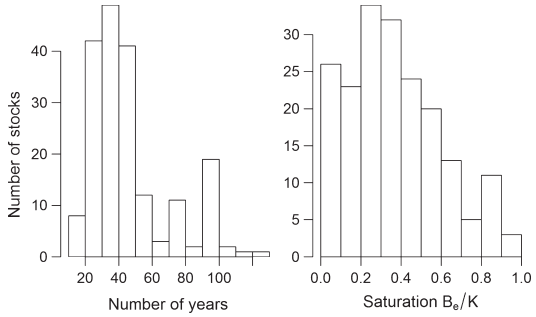

FIGURE 1 Distributions of length of catch history (number of years) and the "observed" stock saturation (B_e/K) for the stocks used in this study

et al., 2006). The BRT approach differs from traditional regression methods in several ways. First, it is a decision tree learning method, which can be either based on classification trees or regression trees depending on the variable of interest. In our case, this variable is continuous, that is B_e/K; that is, our BRT is based on regression trees. The goal is to develop a model that predicts the variable of interest based on several input variables (stock depletion from catch histories in our case). Each tree model has many interior nodes, each corresponding to one of the input variables (e.g. the slope of the whole catch history). Each node splits into two branches, based on whether a given input variable is greater or less than a value p (with p to be determined by "tree learning"). Each leaf represents a value of the target variable of interest given the values of the input variables represented by the branch from the root to the leaf (Figure 2). A tree can be "learned" by splitting each predictor into branches and repeating the sequential splitting process many times (called recursive partitioning). The recursion is completed when the subset stocks at a node have all the same saturation value, or when splitting no longer adds value to the predictions. Tree learning can accommodate all types of predictor variables, including numeric, binary and categorical. 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 boosting to improve accuracy of regression tree models. This involves merging results from multiple models where many regression trees (typically ~1,000–2,000 for this study) are fitted iteratively to the training data to optimize predictive performance. The best regression trees are found by cross-validation—randomly dividing the data into a training set and a testing set. In common with other tree methods, BRT is nonparametric and nonlinear; that is, there is no implicit assumption that the underlying relationships between the predictor variables and the dependent variable are linear, follow some specific nonlinear link function or are monotonic in nature. Thus, BRT 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 used 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.

2.3 | Predictors for stock saturation

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 predictors values for the stock. Choosing predictors is essential for predicting depletion. We expect that because fisheries removal, whether it is managed or not, immediately changes biomass, the trend of catch history can affect depletion. We focus on various catch trends represented by linear regressions of scaled catch ($c_y = C_y/C_{max}$, the annual catch in year y divided by the maximum historical catch) over time. The actual predictors are their intercepts, slopes and break point of a two-segmented regression of scaled catch on time. The linear regressions imply an increasing or decreasing catch trend during various centred 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) (Figure 3).

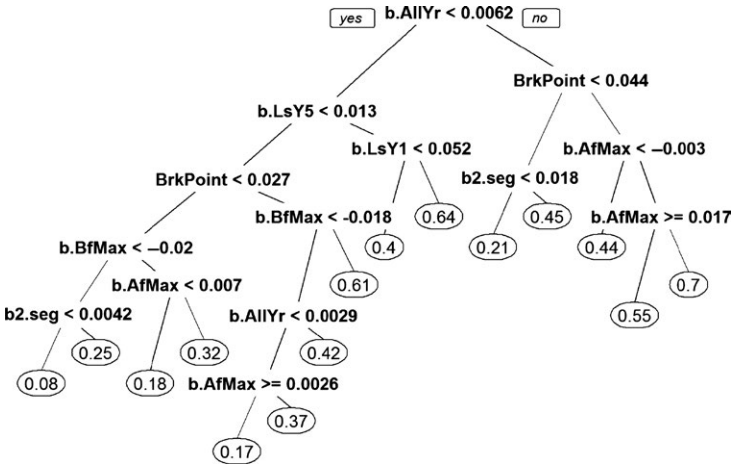

FIGURE 2 An example of BRT tree with eight predictors from Table 1. At each node are the predictor name and its splitting value. The leaves at the end of each branch are the response

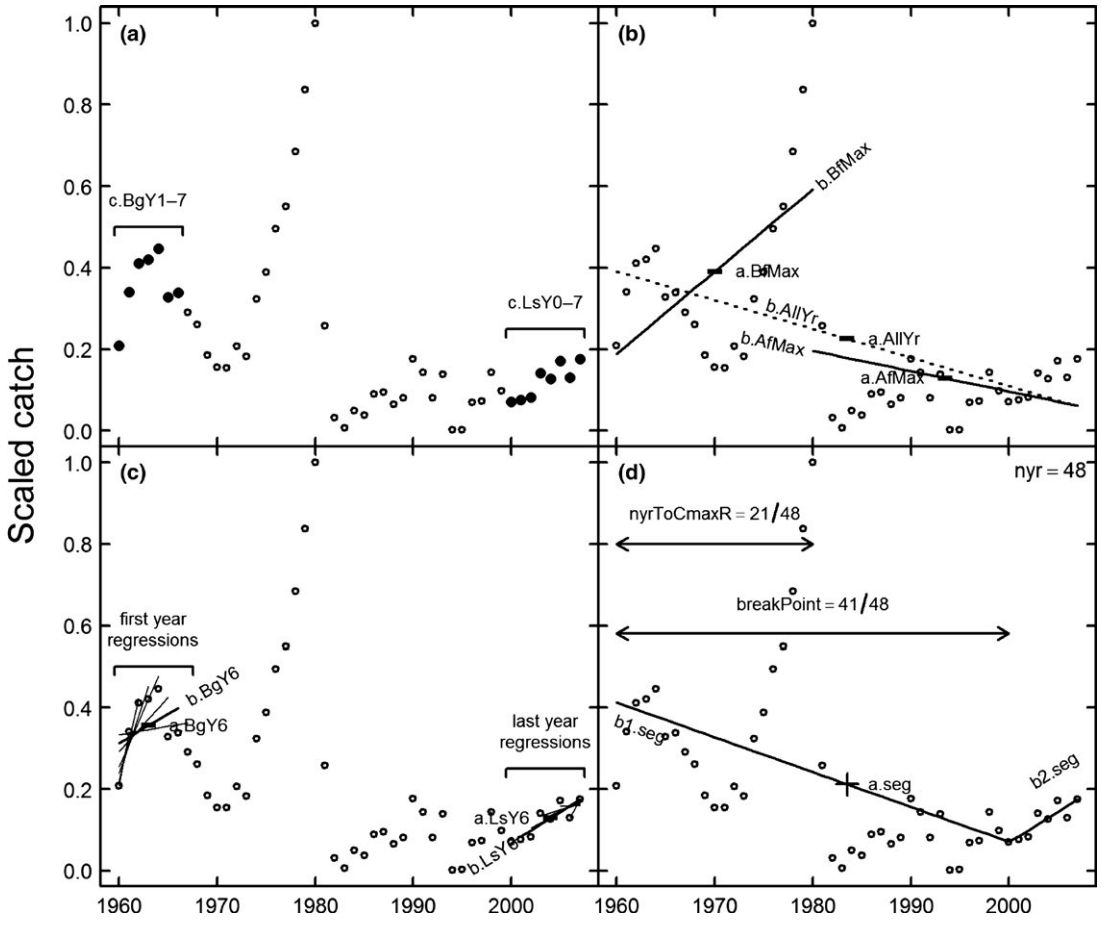

FIGURE 3 Example of predictors based on various catch patterns. (a) Scaled catch in the beginning 7 years and the last 7 years. (b) Regressions using all years (dashed line) and using years before and after peak catch (solid lines). (c) Regressions using the first 2 to 8 years, the last 2 to 8 years. (d) The total number of years, the ratio between the number years before the peak catch and the total number of years, and segmented regression (four parameters)

In addition, we also consider predictors that have been used by panel regression and stock status classification methods, including scaled catch, number of years in the catch history, number of years before the maximum catch divided by total number of years and mean scaled catch. As a result, a total of 56 candidate predictors are examined (Table 1).

2.4 | Modelling

We considered three possible response variables: *S*, log(*S*) and logit(*S*), assuming a Gaussian distribution on the transformed data. Comparisons of the correlation coefficient between the 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 (the sum of squared differences between predicted and observed values). 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.

1. *Tree complexity*: this parameter controls the interaction among predictors. BRT automatically models interactions between predictors as the response to one input variable depends on values of inputs higher in the tree. We examined 15 models with tree complexity from 1 to 15, and compared the alternative models using model fit correlation coefficient and cross-validation correlation coefficient between predicted and observed *S*. 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 about 6 or 7 generally has the highest correlation coefficients for the training data set (Table 2).
2. *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.
3. *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 0.45 and 0.90

TABLE 1 Variables included in the boosted regression trees. The number in parentheses is the number of predictors of each type. The third column shows the relative influence of the eight 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
Break point 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}	–

and found a bagging fraction of 0.75 generally has a lower deviance, so this rate is used in further analysis (Table 2).

4. *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. gbm.step divides the data into 10 subsets, fits a generalized boosted model (gbm) of increasing complexity (from tree number N_{tree} to N_{tree} plus learning step) and calculates the deviance at each step. After each of the 10 subsets is processed, the function identifies the optimal number of trees as the minimum holdout deviance from each of the 10-fold cross-validation.
5. *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 10-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 (Song, 2015; Zhang & Lu, 2012). Plotting the fitted values against the observed values typically shows that smaller *S* is overestimated, while larger *S* is 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 .7 to .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:

$$S_{fit} = a + bS_{obs} \quad (1)$$

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

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

2.5 | Prediction and model evaluation

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

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

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

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

TABLE 2 Sensitivity to number of trees and the bagging fraction. Corr: correlation between observed and model-fitted saturation. Corr.cv: cross-validation correlation between observed and model predicted saturation

N trees	1	2	3	4	5	6	7	8	9	10
Corr	.695	.813	.803	.861	.893	**.939**	.907	.883	.915	.907
Corr.cv	.413	.470	.397	.377	.428	**.487**	.415	.370	.443	.387
Bagging	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90
Corr	.870	.885	.881	.876	.888	.914	**.930**	.917	.930	.927
Corr.cv	.446	.436	.406	.411	.394	.427	**.510**	.421	.411	.393

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

2.6 | Distribution of predicted saturation

One of the key objectives of this study was to develop a prior for S for use in data-poor methods. As the predicted S has a skewed distribution biased towards overestimating S when the true S is low and underestimating S when the true S is large, we construct two skewed normal distributions, f_1(mean = m_1, standard deviation = sd_1, skewness = κ_1) for $S \le 0.5$ and $f_2(m_2, sd_2, \kappa_2)$ for $S > 0.5$. The parameters of these distributions are estimated from the residuals between the bias-corrected prediction S^{bc}_{pred} and the S_{obs}. For any new stocks, if $S^{bc}_{pred} \le 0.5$, then the prior for S should be $f_1(m_1+S^{bc}_{pred}, sd_1, \kappa_1)$. Otherwise, it should be $f_2(m_2+S^{bc}_{pred}, sd_2, \kappa_2)$.

2.7 | Stock status classification

Although quantifying stock status is not our key objective, the method can be used for this purpose. This requires dividing S^{bc}_{pred} values into several bins. We adopt similar thresholds to those used in other studies (Anderson *et al.*, 2012; Branch *et al.*, 2011): $S^{bc}_{pred} > 0.75$ = developing, $0.75 \ge S^{bc}_{pred} > 0.25$ = fully exploited, $0.25 \ge S^{bc}_{pred} > 0.1$ = over-exploited and $S^{bc}_{pred} \le 0.1$ = 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.

2.8 | 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:

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

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 10 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 Ref. 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_e/K instead of B_e/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 (Branch *et al.*, 2011; Froese & Kesner-Reyes, 2002; Zeller, Cheung, Close, & Pauly, 2009). For mPRM, the predicted $\hat{S}^{bc}_{pred}$ 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 ($C_y < 0.5\ C_{max}$) or fully exploited ($C_y \ge 0.5\ C_{max}$). In the second period, fisheries are classified as either fully exploited ($C_y \ge 0.5\ C_{max}$), over-exploited (0.1–0.5 C_{max}) or collapsed ($C_y < 0.1\ C_{max}$).

3 | RESULTS

3.1 | Model fitting

Several of the 56 potential predictors are highly important contributors irrespective of the model structure and the specification of how the BRT method is implemented. The break point of the trends and the regression slopes of catch trends generally have the highest influence on S. Other important predictors include number of years in the catch history, average catch and the catch in recent years. With these predictors in the model, dropping additional predictors does not substantially increase the deviance. Several models with varying number of predictors perform very similarly. For example, the models with 8, 11, 19, 35 and 38 predictors have the same mean deviances, but the variance increases with fewer predictors. The model that includes eight predictors (Table 1) has a training data correlation of .864 (which corresponds to R^2 in linear regression models) and a 10-fold cross-validation (i.e. model prediction) correlation of .497.

The BRT model fits the observed saturations fairly well, resulting in an oval-shaped scatter of points about the regression line, although it tends to underestimate S when stocks are not heavily fished ($S_{obs} > 0.6$). For the model with eight predictors, the bias-correction regression has parameters a = 0.469, b = 0.199 and R^2 = .756 (Figure 4).

3.2 | BRT prediction

The S^{bc}_{pred} has a weaker correlation with S_{obs} than $\hat{S}^{bc}_{pred}$: the correlation coefficient between S^{bc}_{pred} and S_{obs} is 0.497 (compared to correlation coefficient r = .756 for model fit) using the BRT model with eight predictors. The S^{bc}_{pred} is related linearly to S_{obs} when the $S_{obs} < 0.6$ (Figure 5). However, the model under-predicts S when the stocks are only lightly fished ($S_{obs} > 0.6$). The overall mean-squared error of S^{bc}_{pred} is 0.231. The MAE between S^{bc}_{pred} and S_{obs} is 0.182.

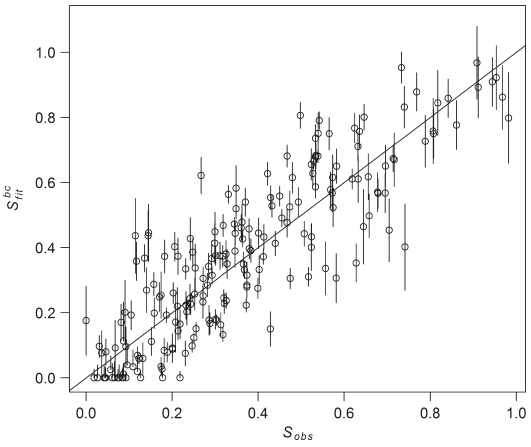

FIGURE 4 Model-fitted and bias-corrected saturation S^{bc}_{pred} versus observed saturation S_{obs} for the 191 stocks. The error bars represent 95% confidence intervals. The straight line is the 1:1 line. The R^2 is .756

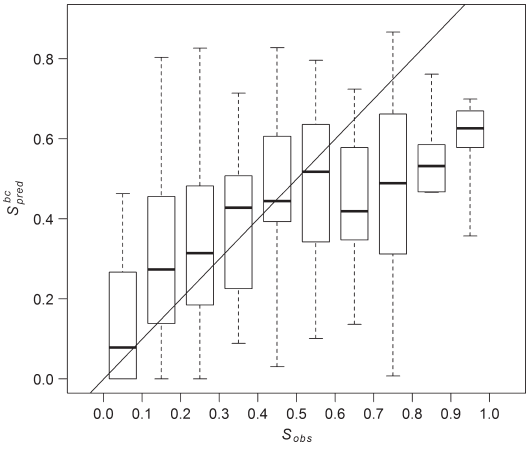

FIGURE 5 Predicted and bias-corrected saturation S^{bc}_{pred} versus observed saturation S_{obs}. The whiskers extend to the most extreme data points

3.3 | Distribution of predicted saturation

Two skewed normal distributions are used to represent the residuals (Figure 6). For $f_1(S \le 0.5)$, the three parameters (mean, standard deviation and skewness) are −0.072, 0.189 and 0.763; for $f_2(S > 0.5)$, they are 0.179, 0.223 and 0.904. Simulation of S using these distributions indicates better, but still imperfect, performance for $S_{obs} > 0.6$ (Figure 7).

3.4 | Stock status classification

The success rate for predicting stock status ranges from 8% for the developing category to 72% for the fully exploited category (Table 3). The overall success rate is 56%. The relative errors range from −38% in the developing category to 13% in the fully exploited category (Table 3). The mean absolute relative error is 20%. Thus, for example, 79 of the 109 stocks that were actually fully exploited were classified as such (72%). In contrast, 123 stocks were classified being fully exploited, which is 13% more than the actual number of stocks actually classified this way.

3.5 | Comparison with panel regression approach and catch-based classification

We use the same formulation (Equation 5) and predictors as in Rosenberg *et al.* (2014) for modelling S. However, using a log-transformed dependent variable (i.e. $\log(B_e/K)$) results in a poor model fit, with an R^2 = .11. Hence, B_e/K is directly used as the dependent variable in the mPRM as in the BRT analysis. This model yields a R^2 = .147 (very close to 0.143 in Rosenberg *et al.*, 2014). Without bias correction, the mPRM predicted $\hat{S}_{pred}$ shrinks to a narrow range around the mean S_{obs}, that is overestimating S when the true value is low, but underestimating it when it is actually high, as reported in other studies (Costello *et al.*, 2012, 2016). With bias correction, the S^{bc}_{pred} for mPRN is more uncertain than that from the BRT prediction (Figure 8). The R^2 between S^{bc}_{pred} for mPRN and S_{obs} is 0.108, and the ~~mean absolute error~~

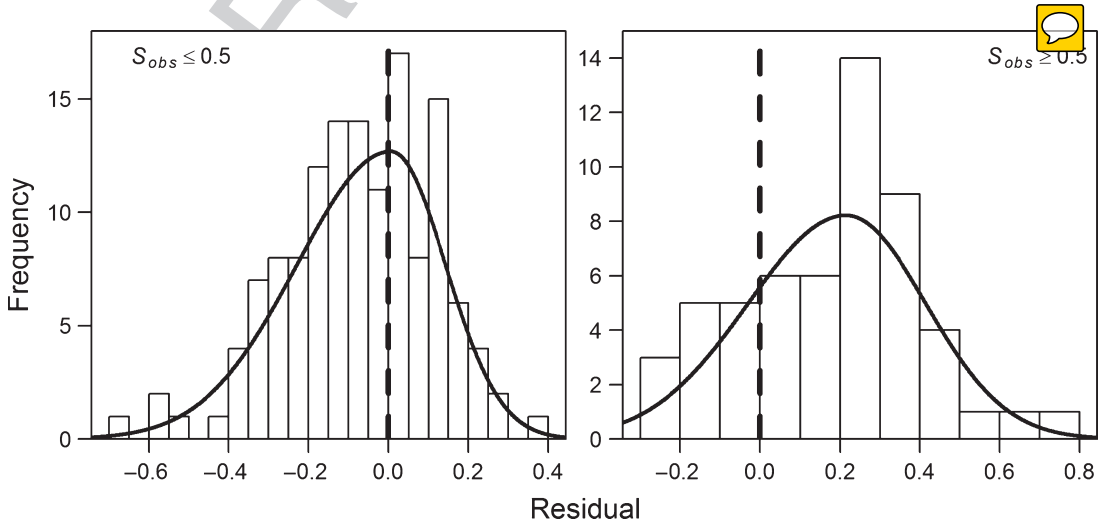

FIGURE 6 Residual distribution of the bias-corrected saturation prediction for $S_{obs} \le 0.5$ (left panel) and $S_{obs} > 0.5$ (right panel). The lines are theoretical distributions derived from fitting a skewed normal distribution to the residuals

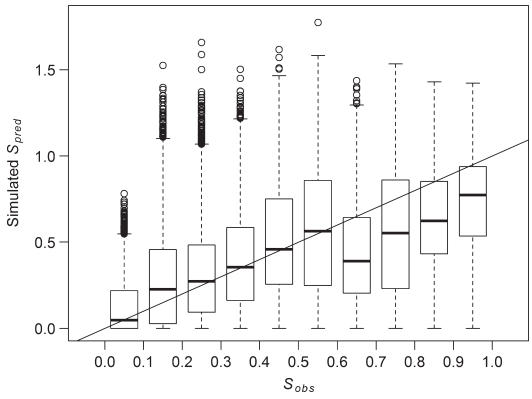

FIGURE 7 Construction of predicted *S* distribution from 100 simulations based on the S^{bc}_{pred} and the residual distribution in Figure 4 for the use as a prior distribution

MAE is 0.403. The success rate for predicting stock status ranges from 5% for the over-exploited category to 76% for the collapsed category (Table 3). The overall success rate is 35%. The relative error for mPRM ranges from −73% for the over-exploited 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 over-exploited and collapsed categories, but fewer in the developing and fully exploited categories (Table 3). 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 over-exploited category. The mean absolute RE is 76%.

4 | 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 = .81) of the model fit using various aspects of catch history. Prediction is more accurate under certain conditions, that is 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 fisheries removal, whether it is managed or unmanaged, instantaneously changes stock biomass and possibly the ratio of B_y/K. The

TABLE 3 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					
		Developing	Fully exploited	Over-exploited	Collapsed	Success rate	RE
Boosted regression trees							
Developing	13	**1**	12	0	0	8%	−38%
Fully exploited	109	6	**79**	21	3	72%	13%
Over-exploited	44	1	25	**14**	4	32%	−9%
Collapsed	25	0	7	5	**13**	52%	−20%
Total or mean	191	$\hat{N}_g = 8$	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%
Over-exploited	44	7	11	**2**	24	5%	−73%
Collapsed	25	1	4	1	**19**	76%	180%
Total or mean	191	$\hat{N}_g = 48$	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%
Over-exploited	44	0	12	**21**	11	48%	91%
Collapsed	25	0	1	10	**14**	56%	72%
Total or mean	191	$\hat{N}_g = 0$	64	84	43	(81) 42%	(absolute) 76%

COLOUR

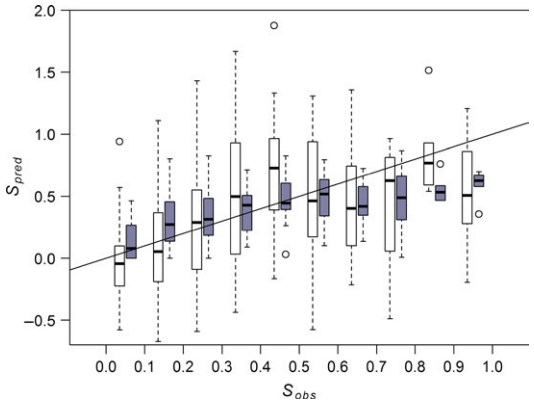

FIGURE 8 Predicted stock saturation from the mPRM (blank) and BRT (grey, reproduced from Fig. 5) methods shows the full spread of the predictions (i.e. the predicted values are not constrained)

complete catch 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 (Anderson *et al.*, 2012; Froese & Kesner-reyes, 2002; Froese, Zeller, Kleisner, & Pauly, 2012; Zeller *et al.*, 2009). 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 = .083, not included in the Results section), but its contribution to prediction power is weaker than many other catch trend indicators.

Similar to some other catch-based stock status classification methods (Anderson *et al.*, 2012; Froese & Kesner-reyes, 2002; Froese *et al.*, 2012; Zeller *et al.*, 2009), 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_e/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 stocks, our approach is more accurate than other methods, such as the recently developed mPRM and catch-based stock classification. For relatively heavily exploited stocks (S between 0.1 and 0.4), the BRT model tends to over-predict true S (overly optimistic) even after bias correction. Precaution is needed when accepting the predicted S for management advice of highly depleted individual stocks. In such cases, opinions from experts, if available, should be considered in conjunction with the result from this method.

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 $B_e/K = 0.5$. Similarly, the depletion-based stock reduction analysis (DB-SRA, Dick & MacCall, 2011) used a bounded beta distribution with $E[B_e/K] = 0.4$, $SD[B_e/K] = 0.1$ and upper and lower bounds of 0.99 and 0.01, respectively. The Catch-MSY method (Martell & Froese, 2013) assumed a B_e/K prior based on the last year's scaled catch: $C_e/C_{max} > 0.5$ then $B_e/K \in [0.3,0.7]$, and $C_e/C_{max} \leq 0.5$ then $B_e/K \in [0.01,0.4]$, 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

We thank Dr. N Dowling and two anonymous reviewers for their constructive comments on the earlier version of the manuscript. This project was funded by FAO and CSIRO.

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 Journal of Marine Science*, *69*, 1491–1500.

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. *Conservation Biology*, *25*, 777–786.

Cadima, E. L. (2003) Fish stock assessment manual. FAO Fisheries Technical Paper 393. Rome: Food and Agriculture Organization of the United Nations.

Carruthers, T. R., Walters, C. J., & 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.

Costello, C., Ovando, D., Clavelle, T., *et al.* (2016). Global fishery futures under contrasting management regimes. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 5125–5219.

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.

Daan, N., Gislason, H., Pope, J. G., & Rice, J. C. (2011). Apocalypse in world fisheries? The reports of their death are greatly exaggerated. *ICES Journal of Marine Science*, *68*, 1375–1378.

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. *Fisheries Research*, *110*, 331–341.

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*, 802–813.

Ferrier, S., Guisan, A., Elith, J., *et al.* (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*, 129–151.

Froese, R., & Kesner-reyes, K. (2002) Impact of fishing on the abundance of marine species. ICES CM 2002/L:12. 12p.

Froese, R., & Proelß, A. (2010). Rebuilding fish stocks no later than 2015: Will Europe meet the deadline? *Fish and Fisheries*, *11*, 194–202.

Froese, R., Zeller, D., Kleisner, K., & Pauly, D. (2012). What catch data can tell us about the status of global fisheries. *Marine Biology*, *159*, 1283–1292.

IFOP (2014) Review of biological reference points in domestic fisheries. Report of the Third International Workshop Held in Viña del Mar, Chile, August 2014. Instituto de Fomento Pesquero.

Kleisner, K., & Pauly, D. (2011) Stock-status plots for fisheries for Regional Seas. In: V. Christensen, S. Lai, M. L. D. Palomares, D. Zeller & D. Pauly (Eds.), *The State of biodiversity and fisheries in regional seas* (pp. 37–40). Fisheries Centre Research Reports 19(3). ~~Xxxxxx:~~ Fisheries Centre, University of British Columbia.

MacCall, a. D. (2009). Depletion-corrected average catch: A simple formula for estimating sustainable yields in data-poor situations. *ICES Journal of Marine Science*, *66*, 2267–2271.

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

Martin, S., & Robinson, J. (2016) Assessment of Indian Ocean longtail tuna (Thunnus tonggol) using data poor catch-based methods. Indian Ocean Tuna Commission. IOTC–2016–WPNT06–17.

Martin, S., & Sharma, R. (2015) Assessment of Indian Ocean Indo-Pacific king mackerel (Scomberomorus guttatus) using data poor catch-based methods. Indian Ocean Tuna Commission. IOTC–2015–WPNT05–21.

Pauly, D. (2008). Global fisheries: A brief review. *Journal of Biological Research-Thessaloniki*, *9*, 3–9.

Pauly, D., Hilborn, R., & Branch, T. A. (2013). Does catch reflect abundance? *Nature*, *494*, 303–306.

Piet, G. J., & Rice, J. C. (2004). Performance of precautionary reference points in providing management advice on North Sea fish stocks. *ICES Journal of Marine Science*, *61*, 1305–1312.

R Core Team (2013) *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Rayns, N. (2007). Short Communication The Australian government 's harvest strategy policy. *ICES Journal of Marine Science*, *64*, 596–598.

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 and Fisheries*, *13*, 380–398.

Rosenberg, A. A., Fogarty, M. J., & Cooper, A. B., *et al.* (2014) *Developing new approaches to global stock status assessment and fishery production potential of the seas*, (Vol. 1086). Fisheries and Aquaculture Circular No. 1086. Rome: FAO.

Song, J. (2015). Bias corrections for Random Forest in regression using residual rotation. *Journal of the Korean Statistical Society*, 44, 321–326.

Tuck, G. N. (2013) *Stock assessment for the Southern and Eastern Scalefish*. ~~Xxxxx: Xxxxx~~.

Zeller, D., Cheung, W., Close, C., & Pauly, D. (2009). Trends in global marine fisheries - a critical review. *Fisheries, Trade and Development*, ~~222~~, 55–77.

Zhang, G., & Lu, Y. (2012). Bias-corrected random forests in regression Bias-corrected random forests in regression. *Journal of Applied Statistics*, *39*, 151–160.

Zhou, S., & Sharma, R. (2013) Stock assessment of two neritic tuna species in Indian Ocean: kawakawa longtail tuna using catch-based stock reduction methods. Indian Ocean Tuna Commission. IOTC–2013–WPNT03–25.

How to cite this article: Zhou S, Punt AE, Ye Y, *et al.* Estimating stock depletion level from patterns of catch history. *Fish Fish*. 2016;00:1–10. doi: 10.1111/faf.12201

Author Query Form

Journal: FAF

Article: 12201

Dear Author,

During the copy-editing of your paper, the following queries arose. Please respond to these by marking up your proofs with the necessary changes/additions. Please write your answers on the query sheet if there is insufficient space on the page proofs. Please write clearly and follow the conventions shown on the attached corrections sheet. If returning the proof by fax do not write too close to the paper's edge. Please remember that illegible mark-ups may delay publication.

Many thanks for your assistance.

Query reference	Query	Remarks
1	**AUTHOR: There were three article titles provided in this proof, therefore we used one. Please check whether this is okay.**	OK
2	**AUTHOR: Please confirm that given names (red) and surnames/family names (green) have been identified correctly.**	Correct
3	**AUTHOR: Please check equations to ensure that they have been entered correctly.**	Correct
4	**AUTHOR: Please provide an appropriate table footnote to explain the bold values in Table 2.**	No need for bold
5	**AUTHOR: The journal style is to include up to 7 author names. If there are eight or more than eight authors, the journal style is to include the first six authors' names, then insert three ellipses, and add the last author's name. Therefore, please provide author names to conform to the journal style.**	This is the title of that paper
6	**AUTHOR: Please provide the publisher location for reference Kleisner and Pauly (2011).**	added.
7	**AUTHOR: Please provide the publisher name, publisher location for reference Tuck (2013).**	added.
8	**AUTHOR: Please provide the volume number for reference Zeller et al. (2009).**	added.

9	**AUTHOR: If you would like the figures in your article to appear as colour in print, please promptly post or courier the completed hard copy of the Colour Work Agreement Form (including payment information) to this mailing address:** **Customer Services (OPI)** **John Wiley & Sons Ltd** **European Distribution Centre** **New Era Estate, Oldlands Way** **Bognor Regis** **West Sussex** **PO22 9NQ** **The form and charge information can be found online at: http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1467-2979**	