ICES Journal of Marine Science Advance Access published June 17, 2012

ICES Journal of Marine Science

ICES CIEM International Council for the Exploration of the Sea Conseil International pour l'Exploration de la Mer

ICES Journal of Marine Science; doi:10.1093/icesjms/fss105

Assessing global marine fishery status with a revised dynamic catch-based method and stock-assessment reference points

Sean C. Anderson[1,2]*, Trevor A. Branch[3], Daniel Ricard[1], and Heike K. Lotze[1]

[1]*Department of Biology, Dalhousie University, 1355 Oxford Street, Halifax, NS B3H 4R2, Canada*
[2]*Department of Biological Sciences, Simon Fraser University, Burnaby, BC V5A 1S6, Canada*
[3]*School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195, USA*

**Corresponding Author: tel: +1 778 782 3989; e-mail: sean_anderson@sfu.ca*

Anderson, S. C., Branch, T. A., Ricard, D., and Lotze, H. K. Assessing global marine fishery status with a revised dynamic catch-based method and stock-assessment reference points. – Journal of Marine Science, doi:10.1093/icesjms/fss105.

Received 25 October 2011; accepted 9 May 2012.

The assessment of fishery status is essential for management, yet fishery-independent estimates of abundance are lacking for most fisheries. Methods exist to infer fishery status from catches, but the most commonly used method is biased towards classifying fisheries as overexploited or collapsed through time and does not account for still-developing fisheries. We introduce a revised method that overcomes these deficiencies by smoothing catch series iteratively, declaring fisheries developing within three years of peak catch, and calibrating thresholds to biological reference points. Compared with status obtained from stock-assessment reference points for 210 stocks, our approach provides a more realistic assessment than the original method, but cannot be perfect because catches are influenced by factors other than biomass. Applied to FAO catches, our method suggests in 2006 32% of global fisheries were developing, 27% fully exploited, 25% overexploited, and 16% collapsed or closed. Although less dire than previous assessments, this still indicates substantial numbers of overexploited stocks. Probably because median exploitation rate decreased since 1992, our catch-based results do not reflect recent stabilization of assessed-stock biomass. Whether this outlook also applies to unassessed stocks can only be revealed with increased or more representative collection of biomass- and exploitation-rate trends.

Keywords: biomass trends, catch data, collapses, fishery status, global fisheries, reference points, stock assessment.

Introduction

The ultimate goal of much of fisheries science is to assess the status of fish stocks. Underexploited stocks result in lost revenues (e.g. Hilborn and Walters, 1992) whereas overexploitation both decreases revenue (e.g. Roy, 1996; Kaiser and de Groot, 2000) and may result in stock collapse with resulting negative ecological impacts (e.g. Hutchings, 2000; Worm *et al.*, 2006). Ultimately, global food security depends on the proper assessment and subsequent management of fisheries (Smith *et al.*, 2010), which contribute on average 15.7% of animal protein for human consumption globally, and up to 50% for some developing island states (FAO, 2009).

Fisheries independent surveys (e.g. research trawl surveys) combined with formal stock assessment provide the most accurate picture of fisheries and population status (Hilborn and Walters, 1992). Unfortunately, for the majority of fisheries, independent surveys are unavailable due to limited resources, and catches are the only data available (e.g. Hilborn and Walters, 1992; Johannes, 1998; Worm *et al.*, 2009). Therefore, having a reasonable proxy for the status of fisheries from catches is vital since the assessed stocks only comprise a small proportion of world catches and are biased towards intensively-studied stocks in developed countries (Ricard *et al.*, 2011).

One widely used catch-based measure of fishery status is catch divided by historical maximum catch (Froese and Kesner-Reyes, 2002; Worm *et al.*, 2006; Pauly, 2007; Sumaila *et al.*, 2007; Pauly, 2008). This method built on and simplified earlier work by Grainger and Garcia (1996), and Garcia and Grainger (2005). Results from these studies using aggregated data suggest that many fisheries are overexploited or collapsed and that there are few fisheries left to develop. However, stock assessments suggest that globally, at least for fisheries that are most intensively managed and surveyed, fishery status may be stabilizing (Worm *et al.*, 2009; Hutchings *et al.*, 2010; Branch *et al.*, 2011).

Assessing a broader range of fisheries, still mostly from developed countries, the United Nations Food and Agriculture Organization found that the fraction of overexploited fisheries continues to increase, while the fraction of collapsed fisheries is currently at a low (FAO, 2010a).

Here we review two main issues related to the original catch-based method as proposed by Froese and Kesner-Reyes (2002), and introduce a robust-dynamic method that overcomes these particular issues. We then suggest revised catch-based thresholds for classifying aggregate fisheries status based on stock-assessment reference points. We compare the robust-dynamic method to fishery status derived from stock-assessments to evaluate classification performance before applying the revised method to global fishery catch data to estimate trends in global fishery status.

Material and Methods

The original catch-based method and its problems

The original catch-based method defines fishery status in hindsight based on the percentage of historical maximum catch and a given year's position with respect to the year of maximum catch (Figure 1a). Before maximum catch is reached, years in which catch is less than 10% of maximum catch are classified as "underdeveloped" and years in which catch is 10–50% of maximum catch as "developing". In this paper we have grouped "underdeveloped" and "developing" for the sake of simplicity as others have done (e.g. Pauly, 2008). Years in which catch is >50% of maximum catch, either before or after the peak, are classified as "fully exploited". After maximum catch has been reached, years in which catch is 10–50% of maximum catch are classified as "overexploited", and years in which catch is <10% of maximum catch are "collapsed".

Beyond the unavoidable issue that under many circumstances catch can have little relation to stock status (e.g. de Mutsert *et al.*, 2008; Branch *et al.*, 2011; Daan *et al.*, 2011), this method exhibits two logical and statistical shortcomings: (1) stochasticity in catch data means that fisheries are more likely to be classified as overexploited or collapsed later in the time-series, and (2) the method, by definition, does not allow for a fishery to be classified as overexploited or collapsed at the start of the time-series or developing at the end.

The problem of stochasticity

The inherent variability of catch data leads to an increased chance of being classified as "collapsed" and a decreased chance of being classified as "developing" over time if the percentage of historical maximum catch is used to estimate fishery status (Wilberg and Miller, 2007; Branch, 2008). Branch *et al.* (2011) compared status using the original catch-based method to status from stock-assessment biomass trends, and status from fisheries that were assessed by the FAO. They found that the original catch-based method was biased towards classifying more fisheries as overexploited or collapsed over time while classifying fewer fisheries as developing. Stock-assessment biomass trends indicated stabilization in the trend of collapsed fisheries since about 1990 and FAO status reports indicated a less severe decline. Recently, Carruthers *et al.* (2012) simulated a variety of fishing scenarios

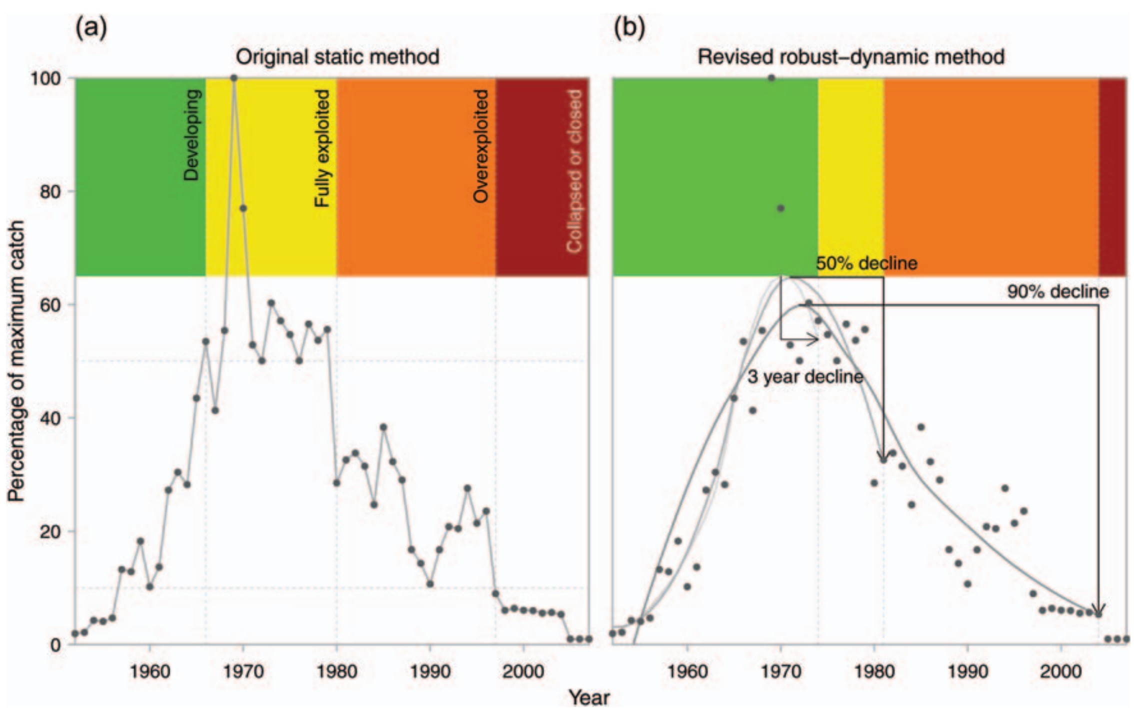

Figure 1. Methods of assessing aggregate fishery status from catch data: (a) the original catch-based method formalized by Froese and Kesner-Reyes (2002), and (b) our revised robust-dynamic method. Horizontal dashed lines in panel *a* mark the thresholds of 10% and 50% of maximum catch around which classification changes. In panel *b*, changes in classification are indicated with arrows and are based on a dynamically changing fit (loess smoother) to the catch data with time (grey solid lines). Vertical dashed lines in (a) and (b) denote changes in classification, and colours in the upper region indicate fishery status. See text for further details. Note: The thresholds for the robust-dynamic method shown in (b) were adjusted in all subsequent figures (59% for overexploited and 91% for collapsed) to better reflect fishery status from stock assessment reference points.

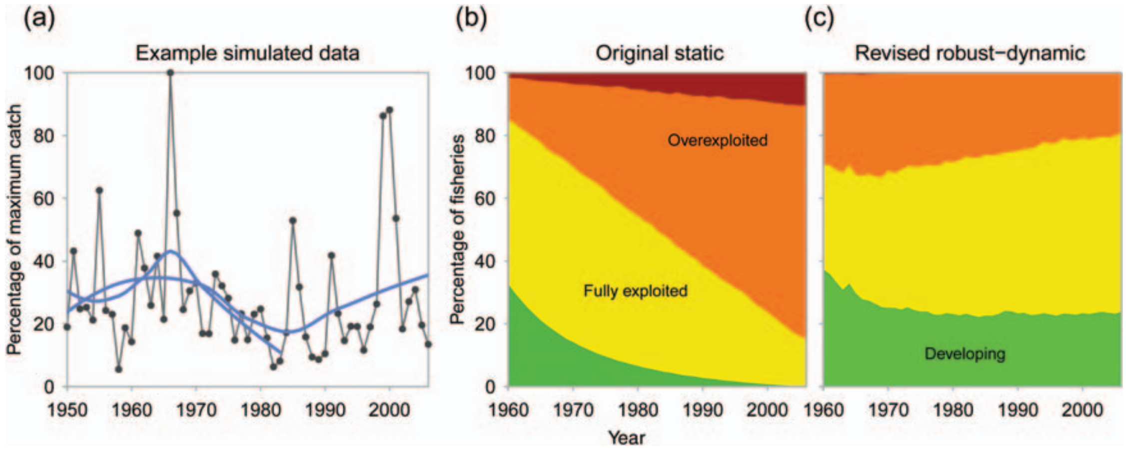

Figure 2. The effect of variability in catch data on fishery status assessment. (a) An example simulated stationary time-series with variability drawn from a log-normal distribution (coefficient of variation = 0.6) with first order autocorrelation ($\rho = 0.5$). See Branch *et al.* (2011) for details of the simulation. We generated and classified 10,000 of these series. Series were simulated from 1950–2006 and classification results are shown from 1960–2006 to establish enough initial years to fit the loess smoother. The blue lines indicate loess smooth functions fitted to the first half and the complete time-series. (b) Status of the simulated time-series using the original catch-based method. (c) Status of the simulated time-series using the robust-dynamic method.

and compared the status as assessed by the original catch-based method to that derived from two traditional stock assessment methods. They found that across this range of fisheries development trajectories, the method using the original catch-based approach was overly pessimistic about stock status. The maximum in a randomly generated time-series drawn from a log-normal distribution will continue to creep upwards asymptotically (see Wilberg and Miller, 2007). Using the original catch-based method, one year of large catch, even if erroneous, will classify all subsequent years as fully exploited, overexploited, or collapsed (Figure 2a, b). That single year of large catch could be caused by a spike in abundance, a spike in fishing effort, or a reporting error, among other factors. The cumulative chances of one of these events occurring increases asymptotically through time. Therefore, a revised method using a percentage of maximum catch as a reference point should account for the variability in catch data and reduce the effect of a creep in the maximum over time (see *Accounting for developing fisheries*).

The problem of developing fisheries

Using the original catch-based method, all fisheries must be classified as fully exploited, overexploited, or collapsed by the last year (Figure 2b; Figure 3a, b). Froese and Kesner-Reyes (2002) attempted to circumvent this by eliminating the last year of data after calculating the year of maximum catch, but this only slightly improves the situation. Many fisheries are still developing and have not yet peaked, and some fisheries have yet to develop (e.g. Sethi *et al.*, 2010; Anderson *et al.*, 2011a, 2011b). Although these new fisheries may not contribute strongly to global fisheries catch and revenue (Sethi *et al.*, 2010), a revised method should allow for fisheries to be classified as developing at the end of the time-series (see *Accounting for developing fisheries*).

The robust-dynamic method

In order to deal with the above outlined issues, we suggest: (i) smoothing the data iteratively to account for stochasticity, and (ii) declaring fisheries developing if within a certain number of

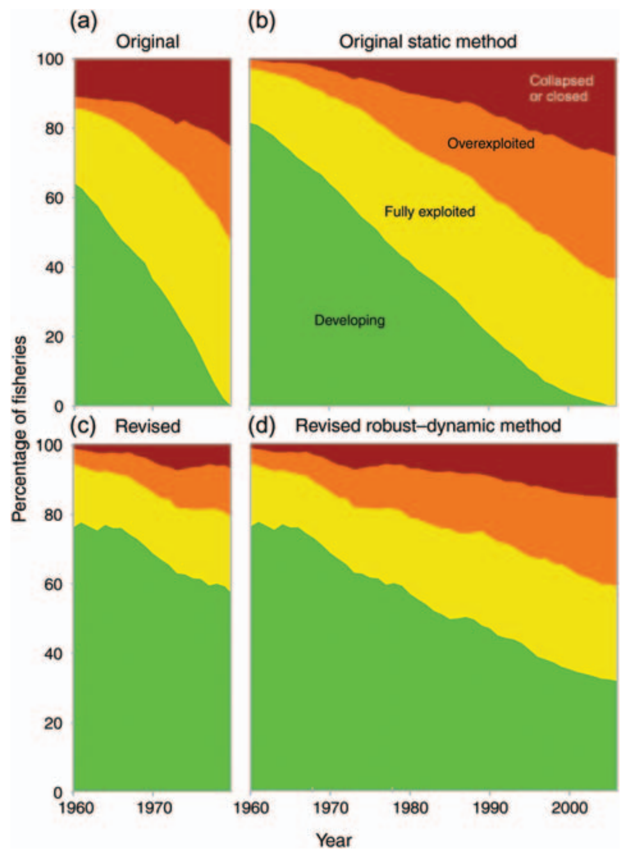

Figure 3. Global fishery status assessed from FAO global fishery capture production catch data using the original and revised classification methods. The original static method applied to catch data until (a) 1980 and (b) 2006, revealing changing estimates for the period of 1960–1980. (c, d) The robust-dynamic method applied to the same data as in panels a and b. The status estimates in panel c are the same as in panel d.

years after the peak in smoothed catch. We further suggest revising the thresholds for the classifications of fishery status based on stock-assessment reference points.

Accounting for stochasticity

We propose fitting a loess smoother to the catch data to reduce the variability and better reflect the underlying long-term trend (Figure 1b). Our loess smoother was fitted with the *loess* function in R (R Development Core Team, 2011) using a smoothing span of 67% of the data (the selection of this smoothing parameter is explained in *Calibrating to stock-assessment data*). We fitted the loess curves to log-transformed data, adding one tonne per year to account for years without catch, and transformed the fits back to the original scale before applying the thresholds. This effectively downweighed the influence of any one (or even a few) extreme data points, making the estimate more robust to outliers. Further, we fitted the curve dynamically, generating a new smoother with the addition of each subsequent year of catch data, for two reasons: First, this modification helps account for stochasticity. Since the degree of smoothing is based on a fixed percentage of the evaluated data and the number of years evaluated increases over time, the degree of smoothing will increase. This increased smoothing can ameliorate to some extent the upward "creep" of the maximum over time if the smoothing parameter is calibrated to the variability (Figure 2a, c). Second, the dynamic fitting of the smoother treats fisheries at the same stage of development equally regardless of their starting year (i.e. it gives them an equally robust smoothing). We are more likely to observe local peaks when evaluating fewer years of catch. With the dynamic fitting, we treat the early years of recently started fisheries and the early years of now established fisheries with equal levels of smoothing. Since many fisheries continue to develop (e.g. Sethi *et al.*, 2010; Anderson *et al.*, 2011a, 2011b) the original catch-based method could generate a time-based bias.

Accounting for developing fisheries

We also revise the method so that fisheries can be classified as developing in recent years. This is accomplished by classifying a fishery as developing until there are at least three years since the loess-smoothed maximum (Figure 1b). The choice of three years was based on the calibration described in the following section. This allows for a fishery to be re-classified as developing if the smoothed catch rises past the previous smoothed maximum. Our reasoning is that this represents the best estimate of fishery status based on catch at any point in time without knowledge of future data. This enables a less-biased classification through time. We required three years after the loess-smoothed maximum catch to ensure a maximum had been reached. However, the three years creates a delay in categorizing a fishery as fully exploited and is thus not a completely satisfactory solution. One alternative would be to define this threshold based on a change of slope of the smoothed-catch trajectory.

The revised threshold to transition from a developing to a fully exploited fishery also better reflects the underlying state of the fish stock. The term "fully exploited" alludes to the state of the fish stock itself (as opposed to "fully developed" which alludes to the state of the fishing intensity). To fish a stock down to B_{MSY} (biomass at maximum sustainable yield), catches must be greater than *MSY* (maximum sustainable yield) during the period when the fishery is developing. Therefore, a fully exploited fishery operating at or about *MSY* will have catches lower than the maximum catch.

Calibrating to stock-assessment data

The original thresholds were based on arbitrary thresholds of 10% and 50% of maximum catch, and not on relationships to stock reference points. We calibrated the catch thresholds based on fishery status as assessed by the ratio of *B* (biomass) to B_{MSY}. We used the same ratios as used by Branch *et al.* (2011): $\geq 1.5\ B/B_{MSY}$ = developing, $0.5 \leq B/B_{MSY} < 1.5$ = fully exploited, $0.2 \leq B/B_{MSY} < 0.5$ = overexploited, and $< 0.2\ B/B_{MSY}$ = collapsed. The threshold of 0.5 B/B_{MSY} to define overfished stocks is used in the United States and Australia (Rosenberg *et al.*, 2006; Hilborn and Stokes, 2010; Hutchings *et al.*, 2010). This also reflects the reality that if target biomass is B_{MSY}, environmental noise and implementation uncertainty will result in about equal probability of actual biomass being above or below B_{MSY}. As noted by Branch *et al.* (2011), the threshold of 0.2 B/B_{MSY} for the collapsed classification corresponds to 10% of unexploited biomass under a Schaefer (1954) model and was previously used in Worm *et al.* (2009) as a threshold for collapse.

We obtained B/B_{MSY} estimates (and U/U_{MSY}, exploitation rate scaled to exploitation rate at *MSY*, as used below) from the RAM Legacy Stock Assessment Database (Ricard *et al.*, 2011). These represent available data for a suite of assessed fisheries globally. Including only those fisheries for which biomass reference points were available or could be estimated, our dataset included 210 fisheries on 111 species. The stocks we evaluated are indicated in Table S1 (see the Supplementary material). We restricted the years of analysis to 1950 to 2006 where the majority of the data existed and the years overlapped the available FAO catch series. The biomass reference points were derived from a combination of B_{MSY} and SSB_{MSY} (spawning stock biomass at *MSY*) estimates from stock-assessment documents and Schaefer (1954) surplus production models. We preferentially choose SSB_{MSY}, B_{MSY}, and surplus production model-based estimates of *MSY*, in that order, to use reference points that were the most accurate and relevant to stock status. We note that these fisheries may present a biased picture of global fisheries given that they are primarily from developed countries with the most complete coverage from Europe and North America (Worm *et al.*, 2009; Ricard *et al.*, 2011); our analysis assumes that patterns of fisheries development are similar between developing and developed countries.

We compared the biomass-assessed fishery status to status based on catch trends for the same set of fisheries. We optimized the number of years after the smoothed peak in catch before a fishery was classified as fully exploited, the loess smoothness, and the overexploited and collapsed catch thresholds to maximize the intraclass correlation (Bartko, 1966; Gamer *et al.*, 2010) between biomass-assessed and catch-assessed status for the years 1960, 1970, 1980, 1990, and 2000. Here, the intraclass correlation coefficient measures the agreement between the two status assessments on a fishery-by-fishery basis, penalizing classifications that deviate further from each other. The best fits were found with: three years after the smoothed peak in catch before a fishery was classified as fully exploited, a loess smoother incorporating 67% of the data, and thresholds of <41% (overexploitation threshold), and <9% (collapsed threshold) of maximum dynamically-smoothed catch (Figure 4a, b, Figure S1). Based on these adjustments, the status of assessed fish stocks based on biomass and catch data from the RAM Legacy Stock Assessment Database, respectively, suggested 24% and 13% developing, 45% and 44%

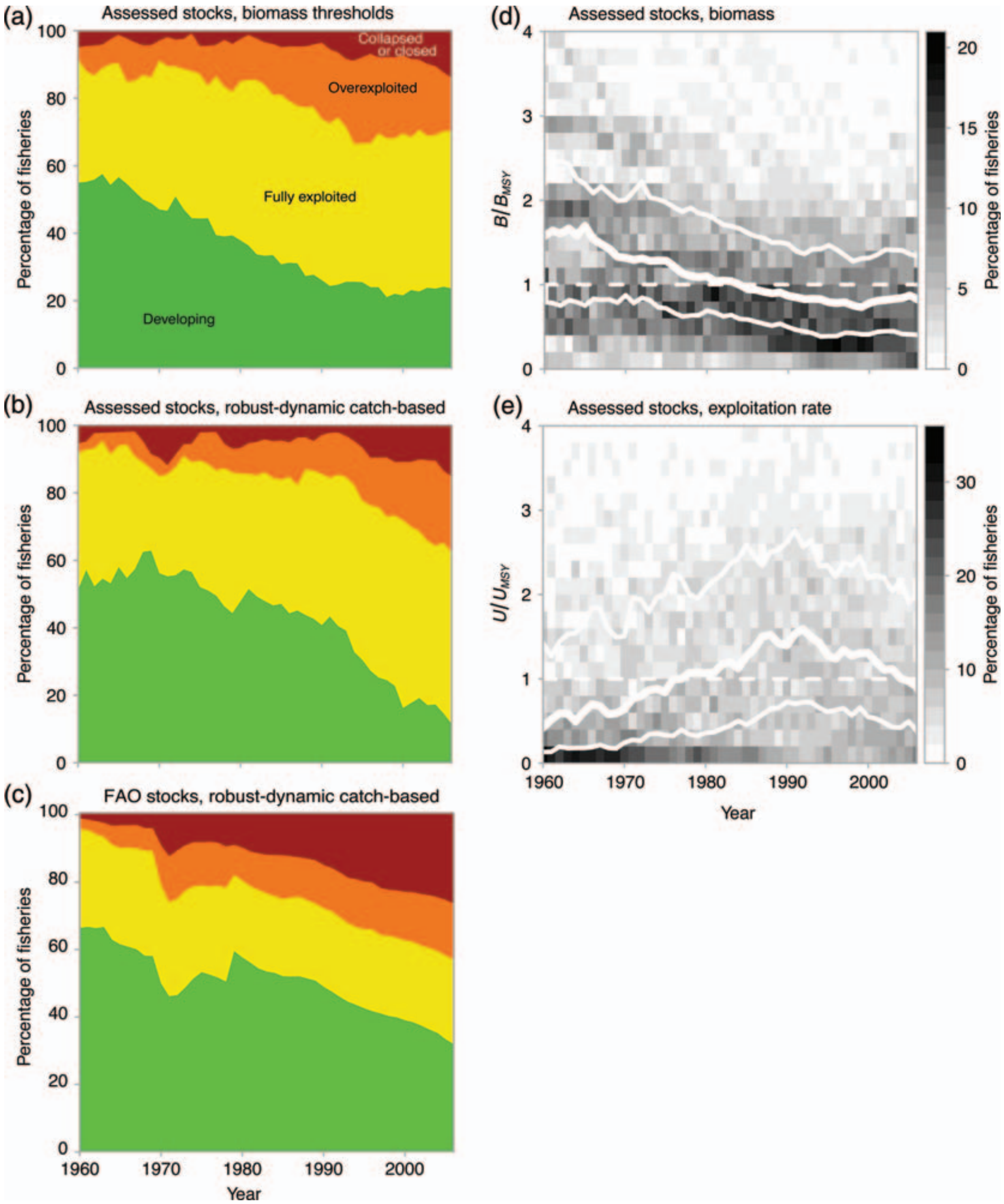

Figure 4. Global fishery status and trends using various approaches across two databases: the RAM Legacy Stock Assessment Database and the FAO global capture production database. Status of assessed stocks (n = 210) based on (a) biomass trends using biomass reference point thresholds (see text for definitions) and (b) catch trends using the robust-dynamic method. (c) Status of all stocks (n = 1475 after filtering; see methods) in the FAO database based on catch trends using the robust-dynamic method. Trends in global fisheries through time based on the distribution of (d) biomass and (e) exploitation rate U/U_{MSY} for assessed stocks. In panels d and e, thick white lines indicate median reference-point values, thinner white lines indicate upper and lower quartile values, and dashed white lines indicate values at which $B = B_{MSY}$ and $U = U_{MSY}$.

fully exploited, 17% and 27% overexploited and 13% and 16% collapsed in 2006 (Figure 4a, b).

Results

Evaluating the robust-dynamic method using stock-assessment data

We evaluated our robust-dynamic method by comparing the estimated fishery status to biomass and exploitation reference points derived from the RAM Legacy Stock Assessment Database.

Evaluating predictions on the spectrum of overfishing and overfished status

"Overfishing" typically refers to the act of exploiting a population beyond an exploitation-level reference point, and "overfished" to a population that has been driven below a biomass reference point by exploitation. Worm *et al.* (2009) showed the recent status of global fisheries along these U/U_{MSY} and B/B_{MSY} axes. Here, we overlay fishery status as estimated by our robust-dynamic method based on catches (with the 9% and 41% thresholds) on

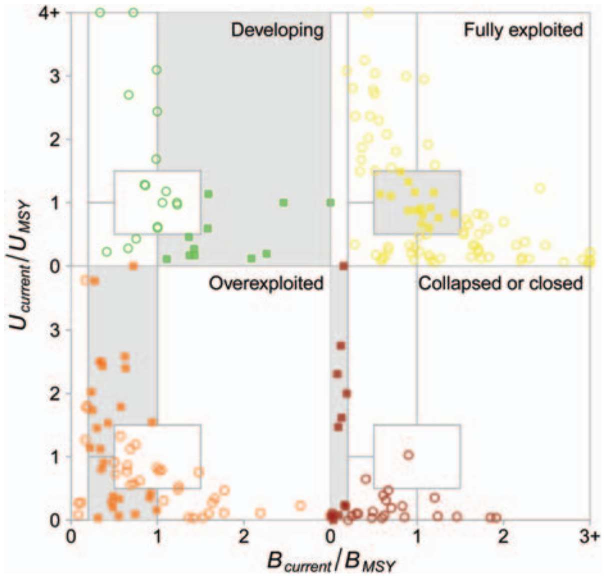

Figure 5. Exploitation rate and biomass of 210 stocks from the RAM Legacy Stock Assessment Database compared to their fishery status as assessed by catch using the robust-dynamic method. Current exploitation rate (U) and biomass (B) are scaled relative to maximum sustainable yield (MSY) values as documented in stock assessment documents or calculated from surplus production models. For each stock, a circle represents the most recent year with available B, U, and catch estimates (range = 2000–2008, median = 2006). The colour and panel indicate the status as assessed by catch (see Figure 1). Values of $B_{current}/B_{MSY} > 3$ and $U_{current}/U_{MSY} > 4$ have been set to 3 and 4, respectively, for graphical purposes. Grey backgrounds indicate expected classification regions based on $B_{current}/B_{MSY}$ and $U_{current}/U_{MSY}$ (thresholds modified from Carruthers *et al.*, 2012). Filled and open circles indicate correctly and incorrectly classified stocks, respectively. Stocks within the overexploited region and with $U_{current}/U_{MSY} < 1$ have the potential to rebuild past B_{MSY} at current exploitation rates and are not currently being overexploited.

an updated version of Figure 3B from Worm *et al.* (2009) (Figure 5). We observe that fisheries are often misclassified when exploitation rate is high and biomass is low (upper-left quadrants) or when exploitation rate is low and biomass is high (lower-right quadrants) (Figure 5). When observing only catch, high exploitation rate can mask low biomass and low exploitation rate can mask high biomass.

Given that our catch-based method assumes a relationship between smoothed percentage of maximum catch and B/B_{MSY}, we compared these two variables in the most recent year using the RAM Legacy Stock Assessment Database. We found a generally positive relationship with a high degree of scatter (Figure 6).

Evaluating aggregate performance

To assess the aggregate performance of the robust-dynamic method across all available fisheries and years (from the RAM Legacy Stock Assessment Database) we plotted the distribution of U/U_{MSY} and B/B_{MSY} within each fishery status category as assessed by their catch. We found a high degree of variability in the biomass and exploitation-rate reference points for the catch-based status classifications when looking across all fishery-year

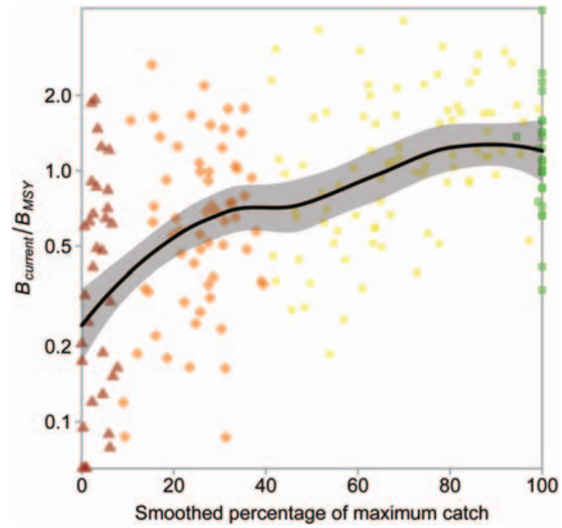

Figure 6. Relationship between $B_{current}/B_{MSY}$ and smoothed catch in the same year for stocks from the RAM Legacy Stock Assessment Database. Current biomass (B) is scaled relative to maximum sustainable yield (MSY) values as documented in stock assessment documents or calculated from surplus production models. The colour and symbol indicate the status as assessed by catch (red triangles: collapsed or closed, orange diamonds: overexploited, yellow squares: fully exploited, green circles: developing). The line represents a loess smoother (smoothing span = 0.75) fitted to log-transformed B/B_{MSY} and the shaded region represents a 95% confidence interval. Note the log-distributed y-axis.

combinations (Figure 7). However, overall we found that the median B/B_{MSY} decreased when progressing from developing to collapsed fisheries as assessed by catch (Figure 7a, from lower to upper panels). This highlights that the assessment of any individual fishery or year may not conform well to overall status; however, when aggregated across many fisheries and years, the general result may provide a reasonable estimate. Exploitation rate was lowest for fisheries classified by catch as developing and collapsed, and highest for fully exploited fisheries (Figure 7b).

Revised trends in global fishery status

Revised estimate from FAO catches

We applied the robust-dynamic method with revised thresholds (<41% for overexploited and <9% for collapsed fisheries) to the FAO (2010b) dataset of catches from global fishery capture production (Figure 4c). We extracted the FAO data for all marine fishes and invertebrates, excluding fisheries less than 15 years in duration and fisheries with less than 10,000 t of cumulative catch since 1950. We also excluded landings grouped as “not elsewhere included” since the size of these categories is partly a function of changing categorization accuracy through time. Our results were not materially affected by including these landings. We considered each taxon in each of the 18 FAO areas as a single stock. The resulting dataset contained 1475 stocks for 712 taxonomic groups.

As of 2006, in the FAO dataset, 32% of fisheries were classified as developing, 27% as fully exploited, 25% as overexploited, and 16% as collapsed. The resulting picture of fishery status is similar to that derived from applying the robust-dynamic

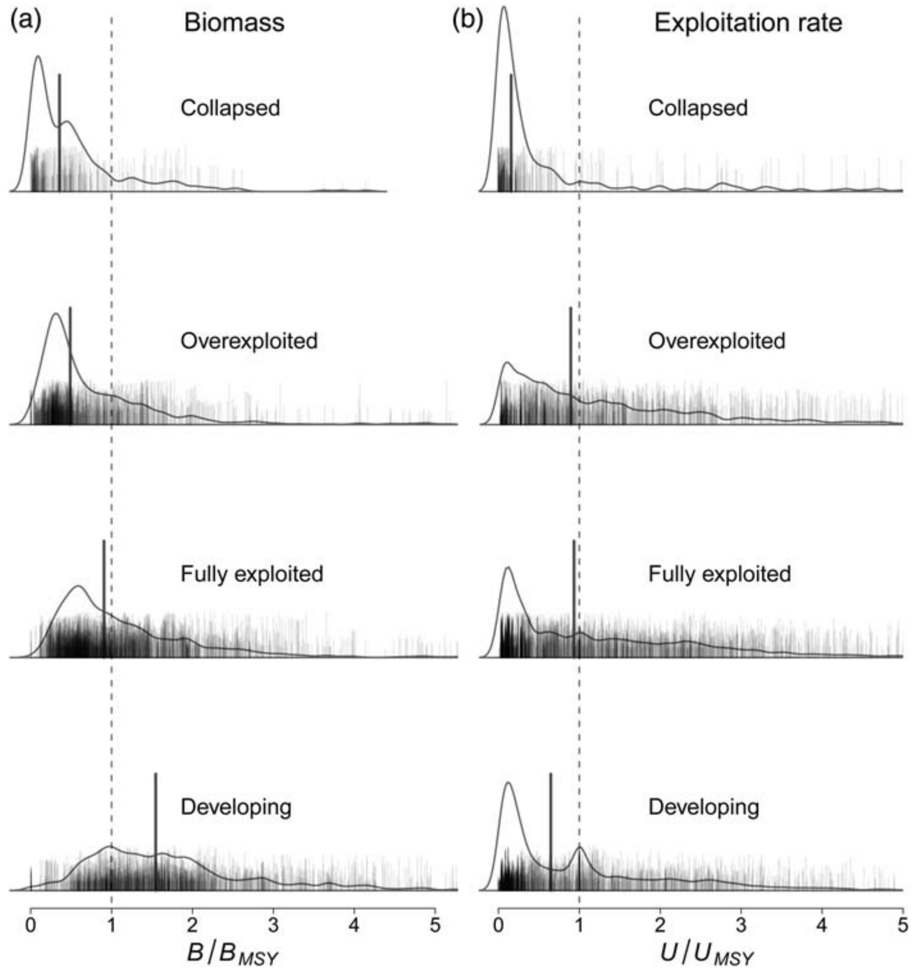

Figure 7. The aggregate distribution of reference points for (a) biomass (B/B_{MSY}) and (b) exploitation rate (U/U_{MSY}) compared to fishery status as assessed by catch using the robust-dynamic method. Short vertical segments represent reference point values (from stock assessments and surplus production models) for individual years from unique fisheries. The height of the segments reflects the length of time the catch series was evaluated to a maximum of 60 years. Tall vertical lines indicate median reference point values. Vertical dashed lines that traverse the panels a and b separate the regions above and below maximum sustainable yield (MSY). The smoothed "beans" (solid curved lines in each plot) represent a kernel density smoother of the distribution of reference points within each category (see Kampstra, 2008).

method to the catches of the assessed stocks in the RAM Legacy Stock Assessment Database (Figure 4b) with the exception that a higher percentage of fisheries are classified as developing in the FAO dataset and less are classified as fully exploited.

Trends in reference points for assessed fisheries

As a further method of evaluating the status of fisheries over time, we analyzed the temporal trend of the distribution of B/B_{MSY} and U/U_{MSY} (from the RAM Legacy Stock Assessment Database) (Figure 4d, e) and compared the results to the fishery status as derived by the classification based on biomass thresholds (Figure 4a) and from catches (Figure 4b, c). The median B/B_{MSY} remained above one until 1984, but declined steadily from around 1960 until the late 1990s when it may have begun to stabilize. As at 2006, the median B/B_{MSY} remained below one, and thus below the management target of B_{MSY} (Figure 4d). The median U/U_{MSY} increased steadily from 1960 until 1992 when it peaked at a value of 1.05 U/U_{MSY} and then declined and fell below one in 2001 (Figure 4e). The assessment of fishery status using biomass thresholds indicated that an increasing percentage of fisheries were becoming fully exploited, overexploited, and collapsed until the mid 1990s when the status of fisheries documented in the stock-assessment database began stabilizing (Figure 4a).

The robust-dynamic method, which was calibrated to have similar percentages of fisheries in the various categories as the biomass thresholds, differed somewhat in its estimate of the trend (Figure 4a, b). The robust-dynamic method indicated a more continual decline in overall status continuing through the 1990s and 2000s (Figure 4b) than the biomass-threshold method (Figure 4a). Given the patterns indicated in Figure 4d and e, this could be a result of the decrease in U/U_{MSY} since 1992, which (without increases in U_{MSY} or biomass) would result in decreased catches, creating the perception of a continuing decline in status using catch-based methods. This highlights the combined effect of changes in biomass and exploitation rate on catches.

Discussion

Proper fisheries management relies on an accurate assessment of fishery status, ideally based on fishery-independent estimates of fish abundance. Because such abundance estimates are lacking for most of the world's fisheries, assessing fishery status based on catch data has been often attempted but also much debated. Our robust-dynamic method addresses two major issues in the original static method: (i) the time-based bias due to stochasticity, and (ii) allowing fisheries to be classified as developing at the end of a time-series. The fishery status classification of our revised method is calibrated to stock-assessment reference points. Overall, our revised method performs better than the original method when comparing aggregate or average fishery status based on catch and assessment data. From a methodological perspective, our study highlights the need for comparing proposed status classification methods to biomass time trends, and demonstrates the utility of graphical approaches to this end.

Our robust-dynamic method improves on the original catch-based method in three main ways. (i) By dynamically fitting a loess smoother to the catch data, the upward creep in the maximum catch through time, and therefore the inherent decline in catch-based status, is largely counteracted. (ii) By requiring a peak in the smoothed catch followed by a period of decline, our method allows fisheries to remain classified as developing at the end of the time-series and to be re-classified as developing as further evidence accumulates. (iii) By calibrating the catch-based status thresholds to stock-assessment-based reference points, our estimates better reflect established interpretations of fisheries status. Despite these improvements, our robust-dynamic method is unable to account for inherent reasons why catch and biomass trends can differ. These reasons include management decisions such as reductions in allowable catches and area closures, data recording issues such as the splitting of higher taxonomic levels, and market changes such as declining demand for certain species (de Mutsert *et al.*, 2008; Branch *et al.*, 2011). Further, our method does not account for recovering fish stocks. While recovering stocks currently represent only 1% of the assessed stocks (FAO, 2010a), as more stocks are rebuilt in the future, detecting recovery will be of greater importance. Our method could be revised to detect recovery by applying the same criteria used to detect a peak in catch in reverse to detect low points of catch.

Our robust-dynamic method improves on previous approaches to estimating aggregate fishery status from catches (e.g. Froese and Kesner-Reyes, 2002) and can be easily applied to a suite of fisheries for which catches are the only data available. However, it remains unsuitable for assessing individual fisheries with reasonable certainty. We emphasize the difference between individual data-poor fishery stock assessment and estimating a long-term and unbiased aggregate measure of fishery status. With additional data, such as estimates of biomass reduction and natural mortality, methods such as depletion-corrected average catch (DCAC) (MacCall, 2009) or depletion-based stock reduction analysis (DB-SRA) (Dick and MacCall, 2010) can be used to estimate sustainable yields for individual data-poor fisheries from catches. Other approaches include pooling information from data-rich stocks to inform the assessment of data-poor fisheries (Punt *et al.*, 2011).

Our results suggest that although global fishery status may not be as dire as the original catch-based method suggests (Figure 3b) there is still cause for concern. Biomass thresholds indicate that the status of many assessed fisheries may be stabilizing (Worm *et al.*, 2009; Branch *et al.*, 2011, Figure 4a); however, 30% of the assessed stocks remain classified as overexploited or collapsed (Figure 4a). Our robust-dynamic method indicates a similar trend for all FAO stocks as the biomass-assessed stocks (Figure 4b, c), with the exception that a greater proportion of FAO stocks are classified as developing and fewer are classified as fully exploited.

Importantly, our method provides evidence of why catch-based global fishery status metrics may overestimate overexploitation since the early 1990s. As an illustration, our robust-dynamic method correctly indicates a decline in catch for the assessed stocks since the early 1990s but inferring a decline in stock status as shown in Figure 4b may be unwarranted when compared to the biomass-threshold assessment (Figure 4a). This may be caused by a reduction in exploitation rate instead of a decline in stock status (Figure 4d, e). The distribution of stock-assessment-based reference points indicates that, for assessed stocks, median U/U_{MSY} started to decline after 1992 (Figure 4e) but there was a seven-year lag before median B/B_{MSY} started to rise (Figure 4d). Although we have reduced global median U/U_{MSY} to below one for these assessed stocks (Figure 4e), median B/B_{MSY} remained below one (Figure 4d) as of 2006. Therefore, for these fisheries we would expect a reduction in catches due to U/U_{MSY} reduction rather than B/B_{MSY} decline, which catch-based classification methods would characterize as declining status. If exploitation rate has also been lowered in unassessed fish stocks, then the apparent deterioration in fishery status in recent years indicated by the robust-dynamic method for all FAO stocks (Figure 4c) may be exaggerated. However, given that the assessed fisheries represent the most intensively managed fisheries, primarily from developed nations (Worm *et al.*, 2009), it is unlikely exploitation has been reduced to the same degree elsewhere. An additional illustration of how exploitation level and biomass may interact to bias catch-based status assessments is evident in the 1970s. Around 1970, B/B_{MSY} began declining (Figure 4a, d) at the same time as exploitation levels increased (Figure 4e) but catch-assessed status did not begin to decline until after 1980 (Figure 4b). This may reflect the ability of fisheries to maintain catch levels to a point by fishing harder and longer, and improving efficiency.

Our study highlights the need for better allocation of resources to perform basic stock assessment for marine resource populations globally. In many situations, dealing with issues of overcapitalization, institutional capacity, poverty, and food needs may be more important to fishery success than estimating reference points (Hilborn, 2002); however, reference points can provide an important metric to gauge fishery status (e.g. Koeller, 2003; Worm *et al.*, 2009; Branch *et al.*, 2011). Currently, the RAM Legacy Stock Assessment Database (Ricard *et al.*, 2011) contains 210 stocks with both biomass and exploitation reference points. This represents only 20–25% of global catches and is a biased sample containing mostly intensively-managed fisheries in developed nations. Whereas gathering fisheries-independent quantitative data for many developing nations would be impractical, gathering detailed information from a small number of locations can provide information that can be applied elsewhere (Johannes, 1998). Better geographic, economic, and taxonomic representation of stocks in the RAM Legacy Stock Assessment Database would enable better estimates of global fishery status. Nonetheless, reliance on catch-based methods and other data-poor metrics to determine fisheries status will remain since many exploited stocks will not undergo proper stock assessment (Johannes, 1998).

We draw three main conclusions from our analyses. First, the graphical ground truthing methods demonstrated here are an important step to developing unbiased and accurate predictive methods and understanding their limitations. Second, our proposed method provides improved estimates of aggregate fishery status over current catch-based methods, but still represents an imperfect and coarse estimate when compared to estimates from stock-assessment-based reference points. Increased allocation of resources to collect basic biomass and exploitation rate trend data would enable more accurate assessment of global fishery status. Third, for the managed and assessed stocks, the median biomass trend has begun to stabilize in recent decades, coinciding with a decline in median exploitation rate; however, it remains unclear if this trend extends to all fisheries worldwide. In either case, many overexploited fish stocks persist and their recovery will largely depend on reductions in exploitation rate.

Supplementary material

Supplementary material with a table of assessed stocks from the RAM Legacy Stock Assessment Database and a figure of the optimization of the assessment parameters is available at the *ICESJMS* online version of the paper.

Acknowledgements

We thank J. S. Collie and J. Mills Flemming for helpful discussions. We also thank two anonymous reviewers for their thoughtful reviews, comments, and criticisms. Financial support is acknowledged from the Natural Sciences and Engineering Research Council of Canada (NSERC) with grants to SCA and HKL; the Sloan Foundation (Census of Marine Life, Future of Marine Animal Populations) with grants to HKL; and the School of Aquatic and Fisheries Sciences, University of Washington, and the U.S. National Science Foundation (NSF) Comparative Analysis of Marine Ecosystem Organization (CAMEO) program (grant 1041570) to TAB. Funding support for DR for construction of the RAM Legacy Stock Assessment Database came from NSERC grants to J. Hutchings, a Canadian Foundation for Innovation grant to HKL, from the Sloan Foundation (Census of Marine Life, Future of Marine Animal Populations), and from the Walton Family Foundation. This project began as part of a National Center for Ecological Analysis and Synthesis (NCEAS) working group funded by the National Science Foundation and the Gordon and Betty Moore Foundation.

References

Anderson, S. C., Flemming, J. M., Watson, R., and Lotze, H. K. 2011a. Rapid global expansion of invertebrate fisheries: Trends, drivers, and ecosystem effects. PLoS ONE, 6: e14735.

Anderson, S. C., Flemming, J. M., Watson, R., and Lotze, H. K. 2011b. Serial exploitation of global sea cucumber fisheries. Fish and Fisheries, 12: 317–339.

Bartko, J. J. 1966. The intraclass correlation coefficient as a measure of reliability. Psychological Reports, 19: 3–11.

Branch, T. 2008. Not all fisheries will be collapsed in 2048. Marine Policy, 32: 38–39.

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.

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.

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

de Mutsert, K., Cowan, J. H., Essington, T. E., and Hilborn, R. W. 2008. Reanalyses of Gulf of Mexico fisheries data: Landings can be misleading in assessments of fisheries and fisheries ecosystems. Proceedings of the National Academy of Sciences of the United States of America, 105: 2740–2744.

Dick, E. J., and MacCall, A. D. 2010. Estimates of sustainable yield for 50 data-poor stocks in the Pacific Coast groundfish fishery management plan. NOAA Technical Memorandum NMFS NOAA-TM-NMFS-SWFSC-460, National Oceanic and Atmospheric Administration, Santa Cruz, CA.

FAO. 2009. The State of World Fisheries and Aquaculture 2008. Food and Agriculture Organization of the United Nations, Rome, Italy.

FAO. 2010a. The State of World Fisheries and Aquaculture 2010. Food and Agriculture Organization of the United Nations, Rome, Italy.

FAO. 2010b. Fisheries commodities production and trade: Fisheries and Aquaculture Department, Food and Agriculture Organization of the United Nations. Electronic database. http://www.fao.org/fishery/statistics/software/fishstat/en (last accessed 1 October 2011).

Froese, R., and Kesner-Reyes, K. 2002. Impact of fishing on the abundance of marine species. ICES Council Meeting Report, CM 2002/L:12: 1–16.

Gamer, M., Lemon, J., and Singh, I. F. P. 2010. irr: Various Coefficients of Interrater Reliability and Agreement. R package version 0.83.

Garcia, S. M., and Grainger, R. J. R. 2005. Gloom and doom? The future of marine capture fisheries. Philosophical Transactions of the Royal Society of London. Series B: Biological, 360: 21–46.

Grainger, R. J. R., and Garcia, S. M. 1996. Chronicles of marine fishery landings (1950–1994): Trend analysis and fisheries potential. FAO Fisheries Technical Paper 359. 51 pp.

Hilborn, R. 2002. The dark side of reference points. Bulletin of Marine Sciences, 70: 403–408.

Hilborn, R., and Stokes, K. 2010. Defining overfished stocks: Have we lost the plot? Fisheries, 35: 113–120.

Hilborn, R., and Walters, C. J. (Ed.) 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics, and Uncertainty. Chapman and Hall, London. 570 pp.

Hutchings, J. A. 2000. Collapse and recovery of marine fishes. Nature, 406: 882–885.

Hutchings, J. A., Minto, C., Ricard, D., Baum, J. K., and Jensen, O. P. 2010. Trends in the abundance of marine fishes. Canadian Journal of Fisheries and Aquatic Sciences, 67: 1205–1210.

Johannes, R. E. 1998. The case for data-less marine resource management: examples from tropical nearshore finfisheries. Trends in Ecology & Evolution, 13: 243–246.

Kaiser, M. J., and de Groot, S. J., (Ed.) 2000. Effects of Fishing on Non-target Species and Habitats: Biological, Conservation and Socio-economic Issues. Blackwell Science, Oxford, UK.

Kampstra, P. 2008. Beanplot: A boxplot alternative for visual comparison of distributions. Journal of Statistical Software, 28: 1–9.

Koeller, P. 2003. The lighter side of reference points. Fisheries Research, 62: 1–6.

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.

Pauly, D. 2007. The Sea Around Us Project: Documenting and communicating global fisheries impacts on marine ecosystems. Ambio, 36: 290–295.

Pauly, D. 2008. Global fisheries: a brief review. Journal of Biological Research-Thessaloniki, 9: 3–9.

Punt, A. E., Smith, D. C., and Smith, A. D. M. 2011. Among-stock comparisons for improving stock assessments of data-poor stocks: the "Robin Hood" approach. ICES Journal of Marine Science, 68: 972–981.

R Development Core Team. 2011. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Ricard, D., Minto, C., Jensen, O. P., and Baum, J. K. 2011. Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish and Fisheries. doi:10.1111/j.1467-2979.2011.00435.x.

Rosenberg, A. A., Swasey, J. H., and Bowman, M. 2006. Rebuilding US fisheries: progress and problems. Frontiers in Ecology and the Environment, 4: 303–308.

Roy, N. 1996. The Atlantic Canada resource management catastrophe: What went wrong and what can we learn from it? Canadian Journal of Economics, 29: S139–S144.

Schaefer, M. B. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Inter-American Tropical Tuna Commission Bulletin, 1: 25–56.

Sethi, S. A., Branch, T. A., and Watson, R. 2010. Global fishery development patterns are driven by profit but not trophic level. Proceedings of the National Academy of Sciences of the United States of America, 107: 12163–12167.

Smith, M. D., Roheim, C. A., Crowder, L. B., Halpern, B. S., Turnipseed, M., Anderson, J. L., Asche, F., *et al.* 2010. Sustainability and global seafood. Science, 327: 784–786.

Sumaila, U. R., Khan, A., Watson, R., Munro, G., Zeller, D., Baron, N., and Pauly, D. 2007. The World Trade Organization and global fisheries sustainability. Fisheries Research, 88: 1–4.

Wilberg, M. J., and Miller, T. J. 2007. Comment on "Impacts of biodiversity loss on ocean ecosystem services". Science, 316: 1285.

Worm, B., Barbier, E. B., Beaumont, N., Duffy, J. E., Folke, C., Halpern, B. S., Jackson, J., *et al.* 2006. Impacts of biodiversity loss on ocean ecosystem services. Science, 314: 787–790.

Worm, B., Hilborn, R., Baum, J. K., Branch, T. A., Collie, J. S., Costello, C., Fogarty, M. J., *et al.* 2009. Rebuilding global fisheries. Science, 325: 578–585.

Handling editor: Sarah Kraak