FISHERIES MANAGEMENT IN THE FACE OF CAPACITY, DATA, AND CLIMATE CHALLENGES

By

CHRISTOPHER M. FREE

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Oceanography

Written under the direction of

Olaf Jensen

And approved by

____

____

____

____

New Brunswick, New Jersey

May 2018

ABSTRACT OF THE DISSERTATION

Fisheries management in the face of capacity, data, and climate challenges

By CHRISTOPHER M. FREE

Dissertation Director:
Olaf Jensen

The sustainable management of global fisheries is essential to addressing food and income security in the coming century. Although fisheries management has advanced significantly over the past few decades, a number of challenges still make the determination of sustainable catch limits difficult to impossible. Many fisheries remain unassessed due to a lack of capacity or lack of data to conduct stock assessments. Furthermore, even when catch limits can be determined, illegal, unreported, and unregulated fishing undermine their effectiveness. Finally, modern fisheries management is complicated by climate change, which is altering population dynamics through large-scale redistributions, changes in phenology, altered food availability, and habitat degradation. In my dissertation, I examine the manifestation of these three challenges – limited capacity, limited data, and climate change – in fisheries of small-, medium-, and large-scales, respectively.

In Chapter 1 (small-scale, limited capacity), I used a mixed-method approach to describe the extent, character, and motivations of illegal fishing in Lake Hovsgol National Park, Mongolia and its impact on the lake's fish populations, especially that of the endangered endemic Hovsgol grayling (*Thymallus nigrescens*). I show that illegal fishing threatens the Hovsgol grayling but also provides food and income for

park residents. An effective management system must therefore incorporate the needs of local people while also addressing the synergistic pressures of climate change, water pollution, increasing tourism, and invasive species.

In Chapter 2 (medium-scale, limited data), I evaluated the performance of the ORCS Working Group Approach to estimating stock status and overfishing limits for 'catch-only' fisheries. I show that the approach is a poor predictor of status and should not be used by managers. I subsequently refined the approach using a machine learning algorithm trained on data-rich stocks and show that the refined ORCS approach performs better than other widely used catch-only methods and can be used when data-moderate methods are not possible or appropriate.

In Chapter 3 (large-scale, climate change), I used surplus production models with monotonic temperature-dependence to measure the influence of sea surface temperature (SST) on the productivity of 190 global fish stocks. I show that ocean warming has significantly positively and negatively influenced the productivity of 20 and 14 stocks, respectively (34 total; 18% total). The influence of warming on a stock's productivity is determined by ecoregion, taxonomic family, life history, and exploitation history. Hindcasts of SST-dependent maximum sustainable yield indicate that MSY of assessed stocks decreased 12.4% from 1930 to 2010. These results show that we must adjust expectations for future food production from the ocean even as the global human population and demand for seafood grows.

Together, these chapters work to help fisheries management overcome challenges from capacity shortfalls, data limitations, and climate change.

Acknowledgements

This dissertation has been a collaborative effort and has been made possible through the work and support of many mentors, colleagues, and friends.

I am deeply grateful to have done my PhD with Olaf Jensen. Olaf is a wonderfully smart, funny, generous, and thoughtful person. He demonstrates through example what it means to be a thorough, curious, and conscientious researcher, teacher, and collaborator. Olaf has a great talent for attracting and extending opportunities and for making every moment, no matter how small, a teaching moment. I cannot imagine having had as diverse, exciting, and fulfilling of a graduate school experience working with anyone else and am truly grateful.

I have also been graced with an exceptional committee. John Wiedenmann has encouraged me to leave my comfort zone and to be thorough and precise while also pragmatic. Malin Pinsky has inspired me to be a true scholar, to read broadly, and to think about the big picture. Jon Deroba has challenged me to be a critical and skeptical scientist and has honed my scientific writing. I am grateful to have faced graduate school with such a smart, responsive, and thoughtful team of mentors.

I have been informally mentored by a number of wonderful postdocs. Kiva Oken has made difficult mathematics accessible. Doug Zemeckis has provided a fountain of excellent professional advice and helped ground my work in the real world. Becca Selden and Mike McCann have been amazing sounding boards.

My labmates have also been central to my graduate school experience. Talia Young is my model for becoming a conscientious, inclusive, community-driven scientist. I am thankful to Michael Roswell for too many things to count, the most important of which is our friendship. Nothing in our lab would get done without Anthony Vastano and Abigail Golden, Mattea Berglund, Jim Vasslides, and Joe Caracappa have been wonderfully fun and insightful labmates.

My ability to conduct my dissertation research has been enhanced by collaborations on a number of side projects. The following collaborators have strengthened my scientific capability: David Gill, Gabby Ahmadia, Megan Barnes, Mark Schulze, Andrea Johnson, Gustavo Pinelo Morales, and Susie Crate.

I have been lucky to conduct fieldwork in Mongolia and even luckier to make great Mongolian friends and colleagues. I am particularly grateful for the hard work, humor, and insight of B. Mendsaikhan (Mendee), Ganzorig Baitsaikhan (Ganzo), P. Tsogtsaikhan (Tsogo), M. Amaraa, and G. Jargalmaa. The Taimen Conservation Fund, Hovsgol Travel, and Mongol Ecology Center have also been vital to these projects.

My interest in quantitative ecology was cultivated by Matt Landis and Jimmy Grogan and I am deeply grateful for their friendship and mentorship, which is now over a decade old. Marcus Drymon mentored me through graduate school applications and opened opportunities for me to build a resume in fisheries science. My parents, brother, and friends have been a source of unwavering support.

Finally, I am grateful to have been funded by a NOAA Sea Grant Population Dynamics Fellowship, a Rutgers University Bevier Fellowship, and DMCS Graduate Assistantships throughout my dissertation. My research in Mongolia was funded by the Grayling Research Trust and American Center for Mongolian Studies and other research, travel, and scholarship expenditures were supported by funding from the Manasquan River Marlin & Tuna Club, NJ Sea Grant, and Rutgers University.

Table of Contents

List of tables

Chapter 1

Chapter 2

Chapter 3

List of illustrations

Introduction

Fisheries management has advanced significantly since Thomas Huxley, an English biologist, famously declared in 1883 that the "cod fishery… and probably all the great sea fisheries, are inexhaustible" (Huxley 1883). Indeed, in the mid-1990s, it became widely apparent that many fisheries were overexploiting stocks and that fisheries management systems in almost all countries required reform. Since then, many countries have implemented reforms ranging from the adoption of scientifically informed harvest policies to the restructuring of incentives to align profits and conservation (Hilborn & Ovando 2014). In many cases, these reforms have successfully reduced fishing effort to sustainable levels and prompted the rebuilding of overfished stocks (Worm et al. 2009; Ricard et al. 2012; Neubauer et al. 2013). These global meta-analyses demonstrate that modern fisheries management is capable of sustainably managing global fisheries.

Although the methods for developing scientifically informed catch limits for fisheries are well-established (Walters & Martell 2004), a number of challenges often make their implementation difficult to impossible. Many fisheries remain unassessed due to a lack of assessment capacity or a lack of data to conduct an assessment. In developing countries, only 5-20% of fish stocks are assessed and this fraction increases to only 10-50% in developed countries (Costello et al. 2012). Furthermore, even when catch limits can be calculated, illegal, unreported, and unregulated (IUU) fishing undermine their effectiveness (Agnew et al. 2009). Finally, modern fisheries management is complicated by climate change, which is altering population dynamics through large-scale redistributions (Cheung et. al

2010, 2013; Pinsky et al. 2013), changing phenology (Edwards & Richardson 2004), altered food availability (Boyce et al. 2014), and degraded habitat (Mora et al. 2013).

In this dissertation, I examine the manifestation of these challenges – limited capacity, limited data, and climate change – in fisheries of small-, medium-, and large-scales, respectively. In Chapter 1 (small-scale, limited capacity), I use a mixed-method approach to describe the extent, character, and motivations of illegal gillnet fishing in Lake Hovsgol National Park, Mongolia and its impact on the lake's fish populations, especially that of the endangered endemic Hovsgol grayling (*Thymallus nigrescens*). In Chapter 2 (medium-scale, limited data), I evaluate the ORCS Working Group approach to estimating stock status and overfishing limits for data-limited fisheries and develop a refined version of the approach to be used when other data-poor methods are not possible or appropriate. In Chapter 3 (large-scale, climate change), I use surplus production models with monotonic temperature-dependence to measure the influence of ocean warming on the productivity of 190 global fish stocks and hindcast their climate-driven maximum sustainable yield.

Together, these chapters work to help fisheries management overcome challenges from capacity shortfalls, data limitations, and climate change.

References

Agnew, D.J., Pearce, J., Pramod, G., Peatman, T., Watson, R., Beddington, J.R., Pitcher, T.J. (2009) Estimating the worldwide extent of illegal fishing. *PLoS One* 4(2): e4570.

Boyce, D.G., Dowd, M., Lewis, M.R., Worm, B. (2014) Estimating global chlorophyll changes over the past century. *Progress in Oceanography* 122(1): 163-173.

Cheung, W.W.L., Watson, R., Pauly, D. (2013) Signature of ocean warming in global fisheries catch. *Nature* 497(7449): 365–368.

Cheung, W.W., Lam, V.W., Sarmiento, J.L., Kearney, K., Watson, R.E.G., Zeller, D., Pauly, D. (2010) Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. *Global Change Biology* 16(1): 24-35.

Cheung, W.W., Sarmiento, J.L., Dunne, J., Frölicher, T.L., Lam, V.W., Palomares, M.D., Watson, R., Pauly, D. (2012) Shrinking of fishes exacerbates impacts of global ocean changes on marine ecosystems. *Nature Climate Change* 3(3): 254-258.

Costello, C., Ovando, D., Hilborn, R., Gains, S.D., Deschenes, O., Lester, S.E. (2012) Status and solutions for the world's unassessed fisheries. *Science* 338: 517-520.

Edwards, M., Richardson, A.J. (2004) Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature* 430: 881–884.

Hilborn, R., Ovando, D. (2014) Reflections on the success of traditional fisheries management. *ICES Journal of Marine Science* 71(5): 1040-1046.

Huxley, T. H. (1883) Inaugural Meeting of the Fishery Congress: Address by Professor Huxley. William Clowes and Sons: London, UK.

Mora, C., Wei, C.L., Rollo, A., Amaro, T., Baco, A.R., Billett, D., Bopp, L., Chen, Q., Collier, M., Danovaro, R., Gooday, A.J. (2013) Biotic and human vulnerability to projected changes in ocean biogeochemistry over the 21st century. *PLoS Biology* 11(10): e1001682.

Neubauer, P., Jensen, O.P., Hutchings, J.A., Baum, J.K. (2013) Resilience and recovery of overexploited marine populations. *Science* 340(6130): 347-349.

Pinsky, M.L., Worm, B., Fogarty, M.J., Sarmiento, J.L., Levin, S.A. (2013) Marine taxa track local climate velocities. *Science* 341(6151): 1239-1242.

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(4): 380-398.

Walters, C.J., Martell, S.J. (2004) Fisheries Ecology and Management. Princeton University Press: Princeton, NJ. 399 pp.

Worm, B., Hilborn, R., Baum, J.K., Branch, T.A., Collie, J.S., Costello, C., Jensen, O.P. (2009) Rebuilding global fisheries. *Science* 325(5940): 578-585.

Chapter 1: A mixed-method approach for quantifying illegal fishing and its impact on an endangered fish species*

Abstract

Illegal harvest is recognized as a widespread problem in natural resource management. The use of multiple methods for quantifying illegal harvest has been widely recommended yet infrequently applied. We used a mixed-method approach to evaluate the extent, character, and motivations of illegal gillnet fishing in Lake Hovsgol National Park, Mongolia and its impact on the lake's fish populations, especially that of the endangered endemic Hovsgol grayling (*Thymallus nigrescens*). Surveys for derelict fishing gear indicate that gillnet fishing is widespread and increasing and that fishers generally use 3-4 cm mesh gillnet. Interviews with resident herders and park rangers suggest that many residents fish for subsistence during the spring grayling spawning migration and that some residents fish commercially year-round. Interviewed herders and rangers generally agree that fish population sizes are decreasing but are divided on the causes and solutions. Biological monitoring indicates that the gillnet mesh sizes used by fishers efficiently target Hovsgol grayling. Of the five species sampled in the monitoring program, only burbot (*Lota lota*) showed a significant decrease in population abundance from 2009-2013. However, grayling, burbot, and roach (*Rutilus rutilus*) all showed

---

* Published as: Free CM, Jensen OP, Mendsaikhan B (2015) A mixed-method approach for quantifying illegal fishing and its impact on an endangered fish species. *PLoS One* 10(12): e0143960.

significant declines in average body size, suggesting a negative fishing impact. Data-poor stock assessment methods suggest that the fishing effort equivalent to each resident family fishing 50-m of gillnet 11-15 nights per year would be sufficient to overexploit the grayling population. Results from the derelict fishing gear survey and interviews suggest that this level of effort is not implausible. Overall, we demonstrate the ability for a mixed-method approach to effectively describe an illegal fishery and suggest that these methods be used to assess illegal fishing and its impacts in other protected areas.

Introduction

Illegal, unreported, and unregulated (IUU) fishing undermine efforts to sustainably manage fish stocks and threaten fish populations worldwide (Agnew et al. 2009). Managers must know as much as possible about the extent, character (e.g., gear types, target/bycatch species, timing, location), and motivations of illegal fishing to effectively develop and implement regulations. However, quantifying illegal fishing is inherently difficult: it is generally covert and significant incentives exist for informants to withhold information (Renzetti & Lee 1993). Furthermore, budget and human resource constraints often restrict efforts to monitor illegal resource use, especially in developing countries (James et al. 1999). There is a need to develop inexpensive yet informative methods for quantifying illegal fishing and its impacts.

Indirect observation, the use of signs of illegal activity as an indicator of non-compliance, has been commonly used to characterize illegal resource use in terrestrial systems (Gavin et al. 2010), but has been infrequently used in marine systems (Bergseth et al. 2013), and to our knowledge, has never been used in freshwater systems. In marine systems, dynamite blast craters (Guard & Masaiganah 1997; Crawford et al. 2004) and derelict fishing gear (Chiappone et al. 2004) have been used as indicators of illegal fishing, but have generally failed to quantitatively measure non-compliance (Bergseth et al. 2013). Most successful quantifications of illegal fishing compare the amount of derelict fishing gear inside and outside reserve boundaries (Cinner et al. 2005, 2006; McClanahan et al. 2006, 2009), but such comparisons are of little use in places without reserves or where the areas outside reserves are undesirable to fishers. The full capacity for indirect observation to reveal rich and quantitative information about illegal fishing remains unexplored.

Indirect observation offers several advantages over other approaches for assessing illegal fishing. It does not require large amounts of labor, specialized equipment, or training and can be recorded during routine enforcement patrols or biological surveys (Bleher et al. 2006). Repeated surveys can reveal spatial and temporal patterns of non-compliance (Chiappone et al. 2004; Cinner et al. 2005, 2006; Williamson et al. 2014) that can be compared to changes in fish communities to examine the effects of illegal fishing (Jachmann 2008). Although indirect observation generally cannot identify specific violators or motivations for non-compliance, they can contribute to a comprehensive understanding of non-compliance when combined with other methods, such as direct questioning (Cinner et al. 2005, 2006).

In this study, we used a mixed-method approach to evaluate the extent, character, and motivations of illegal gillnet fishing in Lake Hovsgol National Park (LHNP), Mongolia and its impact on the lake's fish populations, especially that of the endangered endemic Hovsgol grayling (*Thymallus nigrescens*). Despite the closure of the park to gillnet fishing in 1992, illegal fishing is known to persist (Ocock et al. 2006a,b). We used four complimentary methods to describe this fishery and evaluate its impacts: (1) surveys for derelict fishing gear, an indirect indicator of fishing activity, to evaluate how much illegal fishing is occurring, where illegal fishing is occurring, and what gear is being used; (2) interviews with herders living within the park and park rangers to validate and contextualize the results of the surveys for derelict fishing gear; (3) biological monitoring to identify fish species vulnerable to gillnet fishing and evaluate changes in population abundance potentially caused by fishing; and (4) data-poor stock assessment methods to estimate the effort required to overexploit the Hovsgol grayling population.

Overall, we demonstrate the ability for a mixed-method approach to describe an illegal gillnet fishery and suggest that these methods could be used to effectively and inexpensively assess illegal fishing and its impacts in other protected areas.

Methods

Study site

Lake Hovsgol (51°05'50"N, 100°30'E) is located in the mountains of northern Mongolia at the southern edge of the Siberian taiga forest. It is the 19th largest lake in the world by volume (480 km^3) and has a maximum depth of 262 m and surface area of 2,760 km^2 (Goulden et al. 2006). The lake was established as a National Park in 1992 and is mostly undeveloped. The majority of the resident population lives in two towns on the lakeshore: Hatgal (pop. 2,980) and Hankh (pop. 2,460; NSOM 2012). Tourist camps line the southwestern shore and herding families live intermittently along the lakeshore (Figure 1). Most of the park's ~35,000 annual visitors enter and remain in the southern portion of the park (MEGD 2013).

Lake Hovsgol has ten fish species, the most abundant of which, the Hovsgol grayling (*Thymallus nigrescens*), is endemic to the lake and is listed as endangered on the Mongolian Red List due to climate change and illegal fishing (Ocock et al. 2006a). Hovsgol grayling are more common in littoral areas than pelagic areas and are most abundant along the western shore (Ahrenstorff et al. 2012). A portion of the grayling population spawns in tributary streams in late spring while another portion spawns in the littoral in late summer (Sideleva 2006). The prevalence, fidelity, and success of these spawning strategies are unknown.

The sparse literature on Mongolian fisheries suggests that commercial fishing for Hovsgol grayling, lenok (*Brachymystax lenok*), roach (*Rutilus rutilus*), perch (*Perca fluviatilis*), and burbot (*Lota lota*) removed as much as 200–400 tons annually before the park was established (Dulmaa 1999; Supp. Table 1). Despite the ban on gillnet fishing, active gillnets are often observed and grayling and lenok are frequently sold in Hatgal and along the southwestern shore road. Recreational hook-and-line fishing is legal within the park and is regulated through permits and season and bag limits. Subsistence fishing during the spring spawning migration, though officially illegal, is generally tolerated.

Surveys for derelict fishing gear

We surveyed and collected derelict fishing gear at ten sites on the Lake Hovsgol shoreline in July 2013 and resurveyed six of these sites in July 2014 (Figure 1). Although fishing gear found in the 2013 surveys could represent several years of accumulation and even pre-date the ban on gillnet fishing, gear found in the 2014 resurveys must represent accumulation over the preceding year, since all gear was removed from these sites during the 2013 surveys. Sites were selected as part of a long-term fish monitoring study (Ahrenstorff et al. 2012); though non-random, they provide excellent spatial coverage and access to points and bays on all sides of the lake. In 2013, we censused 54.9 km of shoreline (10 sites, 13 transects, 0.4–8.5 km each, ~13% of total shoreline) for all anthropogenic debris, including derelict fishing gear, between the water and wrack lines (Free et al. 2014). In 2014, we recensused 31.9 km of the original transects (7 sites/transects, 1.3–8.3 km each) for derelict fishing gear only. Because transect widths were variable, we report linear (km^{-1}) rather than areal (km^{-2}) debris density. Derelict fishing gear was classified into the following gillnet categories: whole net, net fragment, float line, lead line, foam float, or bottle float (Supp. Figure 1); and hook-and-line categories: rod, monofilament, lure, or bobber. Bottles, string/rope, and stakes without mesh, floats, weights, or lines were not considered fishing gear. We weighed each item and measured the mesh size (knot to knot distance) of every whole gillnet or gillnet fragment.

Interviews with herders and rangers

The Rutgers University Internal Review Board (IRB) approved our interview protocol (Protocol E14-675) and all respondents gave informed verbal consent (written consent is problematic in former Soviet regions) as approved by the IRB.

We used a semi-structured questionnaire to interview ten herding families from three sites (Figure 1) about their fishing habits, fishing activity they observe, and status and conservation of fish in the lake (Appendix A). The first household at each site was selected opportunistically and additional households were recommended by this family. This “snowball sampling” method is commonly used to find respondents in isolated or hard-to-access groups (Atkinson & Flint 2001). We interviewed seven male and three female heads of household. Family and herd sizes ranged from 3-7 people and 4-630 animals, respectively.

We used a different semi-structured questionnaire to interview five park rangers, including the head ranger, from 5 of 17 ranger districts (Figure 1) about the frequency and character of illegal fishing, actions taken against illegal fishers, and status and conservation of fish in the lake (Appendix B). The interviewed rangers were male and had worked as rangers for 3-15 years. They were responsible for districts that varied in area (22-398 ha) and number of families (32-1,264 families).

Biological sampling, gillnet catch efficiency, and population trends

We used fish monitoring data to estimate catch rates for gillnet mesh sizes used by fishers and to evaluate changes in fish population abundance and body size.

The Rutgers University Animal Care and Facilities Committee approved our fish sampling protocol (Protocol 11-005). Permission to conduct field research (Permit 6/445) was granted by the Mongolian Ministry of Environment and Green Development (MEGD). In July 2009 and 2011-13, we set two monofilament horizontal gillnets at seven of the ten surveyed sites (Figure 1). Both gillnets were 2 m deep and 20 m long with 4 m panels of 2.54, 3.81, 5.08, 6.35, and 7.62 cm bar mesh. They were set at least 100 m apart, perpendicular to shore, using a stationary bottom set in water < 10 m deep, and were fished overnight (8.5–10.5 hr) at each location. Captured fish were identified and measured to the nearest millimeter in total length. Weights for fish without weight measurements were estimated using length-weight parameters derived from our data (Supp. Figure 2).

Vulnerability of fish to gillnets can vary depending on species, body size, and mesh size. We calculated catch-per-unit-of-effort (CPUE) for each gillnet panel in terms of count and biomass (#/kg 10 m^{-1} $night^{-1}$) to determine species-specific and overall catch rates for each mesh size. We also calculated the species-specific CPUE of each gillnet set in terms of count and biomass (#/kg $night^{-1}$) and used linear mixed effects models to examine changes in species-specific abundance from 2009-13 while accounting for sampling site as a random effect on the model intercepts. Decreases in body size can be a useful indicator of fishing impacts when changes in abundance cannot be accurately assessed (McClenachan 2009). Therefore, we also used linear mixed effects models to examine changes in species-specific body size (length/weight) from 2009-13. P-values were generated through likelihood ratio tests of the full models and null 'intercept only' models. All analyses were performed in R version 3.2.0 (R Core Team 2015) and mixed effects models were fit using the *lme4* package (Bates et al. 2015).

Potential population level impacts on Hovsgol grayling

We used methods commonly used in data-poor fisheries management to estimate the maximum sustainable yield (MSY) for Hovsgol grayling and evaluate the likelihood that illegal gillnet fishing could approach or exceed this threshold. Fishing at a rate greater than that which results in MSY is a common definition of overfishing (Hilborn & Walters 1992).

Meta-analyses have shown that fish life history traits can be used to estimate natural mortality rates (Kenchington 2014), which can in turn be used to estimate F_{MSY} (Zhou et al. 2012), the fishing mortality rate resulting in MSY. We estimated the Hovsgol grayling natural mortality rate (M) using three separate life history invariant approaches (Table 1) and applied the Zhou et al. (2012) method to estimate F_{MSY} as $0.87*M$. We used a length-converted catch curve analysis (Pauly & Morgan 1987) to calculate total mortality (total mortality = fishing mortality + natural mortality) to place an upper limit on possible natural mortality rates and estimate current fishing mortality rates. More details on the mortality estimation methods are provided in Appendix C.

We then calculated MSY for each F_{MSY} estimate using the Ahrenstorff et al. (2012) hydroacoustic biomass estimate for Hovsgol grayling (4.4 ± 0.9 kg ha^{-1}) and estimated the number of nights of gillnet fishing required to reach each MSY assuming fishers use 50-m gillnets with 2.54-cm mesh, the optimal mesh size for targeting grayling (~15 kg grayling night^{-1}; see *Section 3.4*). Finally, we estimated the number of fishers required to achieve each MSY assuming fishers use 50-m of gillnet 100 nights year^{-1}. These assumptions seem reasonable given the number of nets used by observed and self-reported fishers and reports that fishing continues throughout the winter (see *Sections 3.2* and *3.3*).

Results

Surveys for derelict fishing gear

A total of 220 (5.78 kg) and 281 (3.82 kg) pieces of derelict fishing gear were collected in the 2013 and 2014 surveys, respectively. Fishing gear comprised 25% of the total weight of plastic debris observed during the 2013 surveys (Free et al. 2014). Derelict gillnet material, the majority of fishing gear found in both years (Figure 2), was found in all but two 2013 transects and all 2014 transects (Figure 1). Foam floats were the most abundant gillnet debris items by count, likely due to their ability to separate from nets and disperse widely; gillnet fragments were the most abundant gillnet debris items by weight, likely due to their large size and heavy lead lines. Gillnet fragments ranged from 2-8 cm in mesh size with 3-4 cm mesh being the most common by both count and weight (Figure 2). All six active gillnets observed had 3.0 cm mesh. The density of derelict gillnet material varied among transects, but in both years, Site 7, the most remote and difficult to access site, had the lowest density of gillnet material and Site 10 (Har Us), the primary location of the spring spawning migration fishery, had the highest density of gillnet material. The density of derelict gillnet material in resurveyed sites was higher in 2014 than 2013 at all but Site 7 suggesting that illegal fishing may be increasing (Figure 1).

Interviews with herders

All of the interviewed herding families (n=10) reported fishing and observing others fishing (Supp. Table 4). Families on the eastern shore reported fishing with gillnets repeatedly throughout the year and during the spring grayling spawning migration. They also reported observing commercial gillnet fishers from Hatgal during the winter and during the spring spawning migration, and they reported finding enforcement ineffective. In contrast, families on the northwestern shore reported fishing with rods or by hand only once per spring spawning migration. They reported no commercial fishing activity and found enforcement effective. All of the families reported that Russian visitors fish recreationally year-round but especially in winter with ice fishing rods and gillnets (Supp. Table 4).

All of the families reported fishing primarily for Hovsgol grayling and primarily for household consumption; only one family from the eastern shore reported selling fish (Supp. Table 4). Families reported fishing primarily during the spawning migration because (1) grayling soup is healthy after the long winter; (2) fish are more abundant and easier to catch than any other time; (3) herders are too busy to fish, or they live away from the lake, the rest of the year; (4) cooking grayling soup interferes with milk production, their principal food source; and (5) eating grayling allows them to delay the slaughtering of herd animals until they have had time to fatten.

Nearly all of the interviewed herders stated that fish population sizes have decreased dramatically (Supp. Table 4). Many recalled that migrating fish were once so numerous that the rivers appeared to "be only fish and no water." Most of the herders also asserted that fish body sizes have decreased and that large lenok and burbot have become especially rare (Supp. Table 4). The herders stated that "local people should protect the lake and fish" but offered few concrete ideas for achieving this objective (Supp. Table 4).

Interviews with rangers

The rangers reported that recreational, commercial, and subsistence fishing all occur in LHNP (Supp. Table 5). The rangers agreed that the majority of recreational fishers are non-local Mongolians or foreigners who fish with rods primarily in summer but also through the ice in winter. The rangers reported that recreational fishers are generally permitted and compliant with the law. All but one ranger reported that local Mongolians use gillnets to target Hovsgol grayling and lenok for subsistence or commercial purposes (Supp. Table 5). The rangers reported that subsistence fishers fish almost exclusively at river mouths during the spring spawning migration and that commercial fishers come predominantly from Hatgal due to that town's proximity to the developed southwestern shore and the city of Mörön. The rangers asserted that the town of Hankh is too remote and undeveloped for commercial fishing to be viable. The rangers reported that commercial gillnet fishing occurs year-round and that fishing when the lake is freezing, thawing, or entirely frozen may even be preferred (Supp. Table 5).

The rangers were divided on the status of fish in the lake: three rangers reported that fish population sizes are decreasing and two rangers reported that they are increasing (Supp. Table 5). The rangers who reported fish population sizes to be decreasing reported that lenok have become especially rare. The majority of rangers reported that fish body sizes have not changed (Supp. Table 5). The rangers were also divided on the best approach to conservation. The head ranger asserted that the native Great Cormorant (*Phalacrocorax carbo*) population is the primary threat to fish and that their population must be controlled. Another ranger suggested that grayling die naturally after the spring spawning migration (an assertion that is not supported by the scientific literature) and that these migrations must therefore be prevented. The remaining rangers emphasized the importance of improved enforcement during the spawning migration (Supp. Table 5).

The rangers offered a detailed description of fishing at Har Us mineral spring (Site 10), the primary location of the spring grayling spawning migration fishery. Mineral springs are culturally important to Mongolians and visiting this spring in May-June is a longstanding social tradition. Rangers are instructed not to enforce the gillnet ban on fishers at Har Us during this time. The rangers reported that over 570 people visited the spring in 2013 and set a total of 60-100 nets per day with an average catch of 50-70 grayling per net. They estimated that 3,600 grayling were caught per day during peak migration (Jun 7-12) and 1,000-1,500 grayling per day from May 30-Jun 6 and Jun 13-24. Based on this report, we estimate that the Har Us fishery removes ~33,000 fish annually.

Gillnet catch efficiency and population trends

The 2.54-cm mesh in our survey gillnets maximized total nightly catch by numbers because it maximized the catch of the abundant Hovsgol grayling (Supp. Figure 3). The 3.81- and 5.08-cm mesh sizes showed similar catch rates and maximized total nightly catch by biomass because they maximized the catch of larger-bodied lenok and burbot (Figure 3); however, the median nightly catch biomass of the 2.54-cm mesh was comparable to those of the 3.08- and 5.81-cm mesh and the 2.54-cm mesh captured fish during every gillnet set, while the larger mesh sizes were often observed empty.

Analysis of the biological monitoring data identified significant reductions in body size for three species over the sampling period (2009-13), but a significant change in CPUE for only one species. Linear mixed effects regression on species-specific CPUE indicates that only burbot population abundance decreased significantly from 2009-13 (Figure 4; Supp. Figure 4). Linear mixed effects regression on body size indicates that grayling, roach, and burbot body size decreased significantly from 2009-13 (Figure 5; Supp. Figure 5). The abundance and body size of other species remained constant.

Potential population level impacts on Hovsgol grayling

Estimates of Hovsgol grayling natural mortality (*M*) ranged 0.25-0.37 (Table 1). A total mortality estimate of 0.42 (Supp. Figure 7) implies fishing moralities of 0.06-0.15, all of which are less than their associated F_{MSY} estimates (Table 1). The F_{MSY} estimates imply MSY values of ~255-331 metric tons yr^{-1}, which could be achieved in ~17,000-22,000 nights of fishing with 50-m optimal mesh gillnets (Table 1). Although these estimates seem large for a low-density resident population, they could be achieved by 170-220 fishers using 50-m of optimal mesh gillnet 100 nights year^{-1} (roughly twice per week). With an estimated permanent population of 5,440 in LHNP and average family size of 3.6 people (NSOM 2012), this effort could be attained if 11.3-14.6% of families participated in the fishery (Table 1). Alternatively, this effort could be attained if every family living in the park fished with 50-m of optimal mesh gillnet 11.3-14.6 nights per year.

Discussion

Knowledge of illegal fishing in Lake Hovsgol National Park (LHNP) has been anecdotal and limited in its usefulness to managers, but with a mixed-method approach, we have empirically described the extent, character, and motivations of illegal fishing and its potential impact on the lake's fish populations.

Our mixed-method approach reveals a fuller understanding of illegal fishing in LHNP than using a single method alone. Each method validates, contextualizes, and builds upon the others to construct a consistent story for a complex fishery: **(1) surveys for derelict fishing gear** quantitatively describe the extent, location, and methods of fishing: gillnet fishing is widespread and increasing and fishers generally use 3-4 cm mesh gillnet; **(2) interviews with herders and park rangers** contextualize these results by qualitatively describing the motivations of fishers, character of fishing, and status of fish in the lake: many residents gillnet fish for subsistence during the spring grayling spawning migration, some residents gillnet fish commercially year-round, and fish population sizes are decreasing; **(3) biological monitoring** documents the vulnerability of fish to gillnets as well as population-level trends in fish abundance and body size: the gillnet mesh sizes used by fishers efficiently target Hovsgol grayling and grayling, burbot, and roach exhibit negative population-level trends; and **(4) data-poor stock assessment analyses** demonstrate that plausible levels of fishing effort by Lake Hovsgol residents using gillnets have the capacity to result in overexploitation of the Hovsgol grayling population. Though seemingly intuitive, the use of multiple methods to quantify and characterize illegal resource use has been rare and should be more widely used by conservation scientists and resource managers (Gavin et al. 2010; Bergseth et al. 2013).

Our surveys for derelict fishing gear are an improvement to previous studies because we use repeated surveys to measure re-accumulation rates and biological monitoring data to evaluate the vulnerability of fish to the gear observed in surveys. The majority of studies have focused on comparing the density of derelict gear inside and outside marine reserves for quantifying non-compliance and fail to measure or report accumulation rates (e.g., Cinner et al. 2005, 2006; McClanahan et al. 2006). A few studies have measured the accumulation rates of derelict gear among habitat types to inform cleanup efforts but have not used the results to understand non-compliance (e.g., Donahue et al. 2001; Chiappone et al. 2004; Bauer et al. 2008). Only Williamson et al. (2014) and the present study have linked these objectives and used both the density and re-accumulation rate of derelict fishing gear to evaluate temporal and spatial trends in non-compliance. By measuring re-accumulation, we show not only that the observed gillnet was used recently and does not pre-date the ban on gillnet fishing, but also that gillnet fishing is becoming increasingly common. Neither Williamson et al. (2014) or our study properly control for the influence of habitat characteristics (e.g., shore/bottom cover or wind/wave exposure) on accumulation and future studies must consider these covariates when identifying hotspots of illegal fishing.

Although our interview method likely underestimates the rate of non-compliance (Soloman et al. 2007; Thomas et al. 2014), it provides a relative description of the frequency of illegal fishing and important information about the motivations for non-compliance, which cannot be gained using other respondent-based approaches (Gavin et al. 2010). The biases and limitations of direct questioning (DQ) can be reduced when researchers have long-standing relationships with the community (Mann 1995; Wolter & Preisendorfer 2013) or by interviewing multiple stakeholders (Mann 1995; Jupiter & Egli 2011). In our study, this likely contributes to the discrepancy in personal fishing habits reported by herders on the eastern and western shores. Whereas eastern shore herders, with whom we have long partnerships, reported frequent gillnet use, western shore herders reported fishing by hook and line or by hand only. Although this may reflect real geographic differences, it may also reflect social desirability bias (Fisher 1993), as western shore herders might be less comfortable revealing sensitive information to us. In our study, this bias is partially corrected by interviewing multiple stakeholders and by inquiring about observed illegal behavior (Mann 1995; Jupiter & Egli 2011). For example, herders were more likely than park rangers to characterize enforcement as ineffective and park rangers were more likely than herders to describe illegal fishing. Similarly, although some respondents were likely to underreport personal fishing, they may not be as likely to underreport observed fishing by others.

Because of these biases, recent papers promote the randomized response (RRT; Warner 1965) and item count techniques (ICT; Miller 1984) over DQ for quantifying non-compliance (Blank & Gavin 2009; St. John et al. 2010; Nuno & St. John 2014; Thomas et al. 2014), but we argue that DQ more easily and fully reveals the motivations for non-compliance (Gavin et al. 2010), which is essential information for successful management (Keane et al. 2008). RRT and ICT incentivize honest responses about illegal behavior by protecting anonymity and generally generate more accurate estimates of the proportion of the sample population engaging in illegal behavior (Soloman et al. 2007; Thomas et al. 2014); however, these approaches require large sample sizes and prevent researchers from implicitly discerning motivations for non-compliance by linking behaviors with covariates or from explicitly inquiring about the motivations for non-compliance (Nuno & St. John 2014). DQ, on the other hand, allows researchers to inquire about the motivations for non-compliance, importance of natural resources to culture or livelihood, and desire for changes to management rules. Managers must consider the socioeconomic functions of resource use and DQ should remain in the conservation science toolbox.

Although the population-level impacts observed in our biological monitoring data cannot necessarily be attributed to illegal fishing, they indicate the importance of improving fisheries management in LHNP, especially given the feasibility for gillnet fishers to overexploit the Hovsgol grayling population, as indicated by the data-poor stock assessment analysis. These calculations represent a simplification of population dynamics made necessary by the lack of time series of fishery removals or estimates of biological parameters needed for more complex data-poor assessment methods (Jensen et al. 2009). However, our indirect estimates of M for Hovsgol grayling are similar to direct estimates of M for Arctic grayling (*T. arcticus*), a close relative (0.29 average; Supp. Table 3). Furthermore, all of our MSY estimates indicate that overexploitation is possible even with only a small percentage of the population participating in the fishery using gillnets, an inexpensive and widely available fishing gear. The threat of overexploitation is not unrealistic given that grayling, as a taxonomic group, can be susceptible to anthropogenic influences as has been seen with the extirpation of many North American Arctic grayling populations in Montana and Wyoming (Northcote 1995). Salmonids are vulnerable to exploitation and other disruptions during their spring spawning migrations (Roberts & White 1992) and managers must carefully consider the value and impact of the spring spawning migration fishery.

The results of our mixed-method approach indicate that illegal fishing is a problem in Lake Hovsgol but that fish also serve an important socioeconomic function. An effective management system will need to incorporate the needs of local people as well as address the synergistic pressures of climate change, water pollution, increasing tourism, and invasive species on LHNP's fish populations. In the last 40 years, regional air temperatures have increased 2.1°C (Dagvadorj et al.

2009), a rate of warming more than three times faster than the global average (IPCC 2013), which has prompted the drying of many of Lake Hovsgol's previously reliable streams and loss of grayling spawning habitat (Ocock et al. 2006a,b). Increasing tourism may result in increased fishing pressure, habitat destruction, water pollution, and invasive species introductions without proper management. Lake Hovsgol is already heavily polluted with household trash and will only become more polluted with additional strains on its inadequate waste management system (Free et al. 2014). Although no invasive species have established to date, the successful introduction of a new fish or aquatic plant species could alter this otherwise intact ecosystem (Young et al. 2014).

Fishing, historically uncommon in Mongolia's pastoralist culture, may be gaining prevalence as a new source of food, income, or recreation, especially as climate change makes herding more difficult (Batima 2013) and urban Mongolians acquire more globalized tastes in food and leisure (Bruun & Odgaard 2013). At the same time, Mongolia aims to protect 30% of the country by 2030, more than doubling the area currently under protection (Myagmarsuren 2008). These trends forecast continued conflicts between economic and conservation objectives and the way in which these conflicts are resolved or ignored in the iconic LHNP could shape future protected area management in the country.

Acknowledgements

We are grateful for the devoted work of our many field assistants. We also thank Steven Alexander, Talia Young, and the members of the Jensen lab for their helpful feedback on the manuscript. The Taimen Conservation Fund, Mongol Ecology Center, Sweetwater Travel, and Hovsgol Travel provided field support and assistance. This work was supported by a NSF-IRES grant (OISE 1064843) to OPJ, a USAID-PEER grant (Project 98) to BM, and American Center for Mongolian Studies and Grayling Research Trust grants to CMF. CMF is funded by a NMFS-Sea Grant Population Dynamics Fellowship.

References

Agnew DJ, Pearce J, Pramod G, Peatman T, Watson R, Beddington JR, Pitcher TJ. Estimating the worldwide extent of illegal fishing. PLoS One. 2009; 4(2): e4570.

Ahrenstorff TD, Jensen OP, Weidel BC, Mendsaikhan B, Hrabik TR. Abundance, spatial distribution, and diet of endangered Hovsgol grayling (*Thymallus nigrescens*). Environ Biol Fishes. 2012; 94(2): 465-476.

Atkinson R, Flint J. Accessing hidden and hard-to-reach populations: Snowball research strategies. Soc Res Update. 2001; 33(1): 1-4.

Bates D, Maechler M, Bolker B, Walker S. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-8. 2015. Available: http://CRAN.R-project.org/package=lme4.

Batima P, Natsagdorj L, Batnasan N. Vulnerability of Mongolia's pastoralists to climate extremes and changes. In: Leary N, Conde C, Kulkarni J, Nyong A, Pulhin J, editors. Climate Change and Vulnerability and Adaptation. London, UK: Earthscan; 2013. pp. 67-87.

Bauer LJ, Kendall MS, Jeffrey CF. Incidence of marine debris and its relationships with benthic features in Gray's Reef National Marine Sanctuary, Southeast USA. Mar Pollut Bull. 2008; 56(3): 402-413.

Bergseth BJ, Russ GR, Cinner JE. Measuring and monitoring compliance in no-take marine reserves. Fish Fish. 2013; 16(2) 240-258.

Blank SG, Gavin MC. The randomized response technique as a tool for estimating non-compliance rates in fisheries: a case study of illegal red abalone (*Haliotis rufescens*) fishing in Northern California. Environ Conserv. 2009; 36(2): 112-119.

Bleher B, Uster D, Bergsdorf T. Assessment of threat status and management effectiveness in Kakamega Forest, Kenya. Biodivers Conserv. 2006; 15(4): 1159–1177.

Bruun O, Odgaard O. Mongolia in transition: old patterns, new challenges. Abingdon, UK: Nordic Institute of Asian Studies; 1996.

Chiappone M, Swanson DW, Miller SL, Dienes H. Spatial distribution of lost fishing gear on fished and protected offshore reefs in the Florida Keys National Marine Sanctuary. Caribb J Sci. 2004; 40(3): 312–326.

Cinner JE, Marnane MJ, McClanahan TR. Conservation and community benefits from traditional coral reef management at Ahus Island, Papua New Guinea. Conserv Biol. 2005; 19(6): 1714–1723.

Cinner JE, Marnane MJ, McClanahan TR, Almany GR. Periodic closures as adaptive coral reef management in the Indo-Pacific. Ecol Soc. 2006; 11(1): 31.

Crawford BR, Siahainenia A, Rotinsulu C, Sukmara A. Compliance and enforcement of community-based coastal resource management regulations in North Sulawesi, Indonesia. Coast Manage. 2004; 32(1): 39-50.

Dagvadorj D, Natsagdorj L, Dorjpurev J, Namkhainyam B. MARCC: Mongolia assessment report on climate change 2009. Ulaanbaatar, Mongolia: Ministry of Nature, Environment, and Tourism; 2009.

Donohue MJ, Boland RC, Sramek CM, Antonelis GA. Derelict fishing gear in the Northwestern Hawaiian Islands: diving surveys and debris removal in 1999 confirm threat to coral reef ecosystems. Mar Pollut Bull. 2001; 42(12): 1301-1312.

Dulmaa A. Fish and fisheries in Mongolia. In: Petr T, editor. Fish and fisheries at higher latitudes: Asia. FAO Fisheries Technical Paper 385. Rome, Italy: FAO; 1999.

Fisher RJ. Social desirability bias and the validity of indirect questioning. J Consum Res. 1993; 20: 303-315.

Free CM, Jensen OP, Mason SA, Eriksen M, Williamson NJ, Boldgiv B. High-levels of microplastic pollution in a large, remote, mountain lake. Mar Pollut Bull. 2014; 85(1): 156-163.

Gavin MC, Solomon JN, Blank SG. Measuring and monitoring illegal use of natural resources. Conserv Biol. 2010; 24(1): 89-100.

Goulden CE, Sitnikova T, Gelhaus J, Boldgiv B. The geology, biodiversity, and ecology of Lake Hovsgol (Mongolia). Leiden, Netherlands: Backhuys; 2006.

Guard M, Masaiganah M. Dynamite fishing in southern Tanzania, geographical variation, intensity of use and possible solutions. Mar Pollut Bull. 1997; 34(10): 758-762.

Gunderson D. Trade-off between reproductive effort and adult survival in oviparous and viviparous fishes. Can J Fish Aquat Sci. 1997; 54(5): 990-998.

Hilborn R, Walters CJ. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Dordrecht, Netherlands: Springer Science+Business Media; 1992.

IPCC. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Eds: Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM. Cambridge, UK: Cambridge University Press; 2013.

Jachmann H. Monitoring law-enforcement performance in nine protected areas in Ghana. Biol Conserv. 2008; 141(1): 89-99.

James AN, Gaston KJ, Balmford A. Balancing the Earth's accounts. Nature. 1999; 401(6751): 323-324.

Jensen OP, Gilroy DJ, Hogan Z, Allen BC, Hrabik TR, Weidel BC, Chandra S, Vander Zanden MJ. Evaluating recreational fisheries for an endangered species: a case study of taimen, *Hucho taimen*, in Mongolia. Can J Fish Aquat Sci. 2009; 66(10): 1707-1718.

Jupiter SD, Egli DP. Ecosystem-based management in Fiji: successes and challenges after five years of implementation. J Mar Biol. 2010; 2011: 940765.

Keane A, Jones JP, Edwards-Jones G, Milner-Gulland EJ. The sleeping policeman: understanding issues of enforcement and compliance in conservation. Anim Conserv. 2008; 11(2): 75-82.

Kenchington TJ. Natural mortality estimators for information-limited fisheries. Fish Fish. 2014; 15(4): 533-562.

Mann BQ. Quantification of illicit fish harvesting in the Lake St Lucia game reserve, South Africa. Biol Conserv. 1995; 74(2): 107-113.

McClanahan TR, Graham N, Wilson SK, Letourneur Y, Fisher R. Effects of fisheries closure size, age, and history of compliance on coral reef fish communities in the western Indian Ocean. Mar Ecol Prog Ser. 2009; 396: 99-109.

McClanahan TR, Marnane MJ, Cinner JE, Kiene WE. A comparison of marine protected areas and alternative approaches to coral-reef management. Curr Biol. 2006; 16(14): 1408-1413.

McClenachan L. Documenting loss of large trophy fish from the Florida Keys with historical photographs. Conserv Biol. 2009; 23(3): 636-643.

MEGD. Lake Hovsgol National Park: General Management Plan Foundation Document. Ulaanbaatar, Mongolia: Ministry of Environment and Green Development (MEGD); 2013.

Miller JD. A new survey technique for studying deviant behavior. Washington, D.C: George Washington University; 1984.

Myagmarsuren D. Special Protected Areas of Mongolia. Ulaanbaatar, Mongolia: Munkhyn Useg Co, Ltd; 2008.

Northcote TG. Comparative biology and management of Arctic and European grayling (Salmonidae, *Thymallus*). Rev Fish Biol Fish. 1995; 5(2): 141-194.

NSOM. Mongolian Statistical Yearbook 2012. Ulaanbaatar, Mongolia: National Statistical Office of Mongolia (NSOM); 2012.

Nuno A, St John FA. How to ask sensitive questions in conservation: A review of specialized questioning techniques. Biol Conserv. 2014; 189: 5-15.

Ocock J, Baasanjav G, Baillie JEM, Erdenebat M, Kottelat M, Mendsaikhan B, Smith K. Mongolian Red List of Fishes. London, UK: Zoological Society of London; 2006a.

Ocock J, Baasanjav G, Baillie JEM, Erdenebat M, Kottelat M, Mendsaikhan B, Smith K. Summary Conservation Action Plans for Mongolian Fishes. London, UK: Zoological Society of London; 2006b.

Pauly D, Morgan GR. Length-based methods in fisheries research. ICLARM Conference Proceedings 13. Manila, Philippines: International Center for Living Aquatic Resources Management; 1987.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria. 2015. Available: http://www.R-project.org/.

Renzetti CM, Lee RM. Researching sensitive topics. Newbury Park, CA: Sage Publications; 1993.
Roberts BC, White RG. Effects of angler wading on survival of trout eggs and pre-emergent fry. N Am J Fish Manage. 1992; 12(3): 450-459.

Sideleva VG. Fish fauna of Lake Hovsgol and Selenga River in comparison with ichthyofauna of Lake Baikal. In: Goulden CE, Sitnikova T, Gelhaus J, Boldgiv B, editors. The geology, biodiversity, and ecology of Lake Hovsgol (Mongolia). Leiden, Netherlands: Backhuys; 2006. pp 357-378.

Solomon J, Jacobson SK, Wald KD, Gavin M. Estimating illegal resource use at a Ugandan park with the Randomized Response Technique. Hum Dimens Wildl. 2007; 12(2): 75-88.

St John FAV, Edwards-Jones G, Gibbons JM, Jones JPG. Testing novel methods for assessing rule breaking in conservation. Biol Conserv. 2010; 143(4): 1025-1030.

Then AY, Hoenig JM, Hall NG, Hewitt DA. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES J Mar Sci. 2014; 72(1): 82-92.

Thomas AS, Gavin MC, Milfont TL. Estimating non-compliance among recreational fishers: Insights into factors affecting the usefulness of the Randomized Response and Item Count Techniques. Biol Conserv. 2014; 189: 24-32.

Tsogtsaikhan P, Mendsaikhan B, Jargalmaa G, Ganzorig B, Weidel BC, Filosa CM, Free CM, Young T, Jensen OP. Comparison of age and growth of lenok, *Brachymystax lenok* (Pallas 1773), and grayling, *Thymallus* spp., in lentic and lotic habitats of Northern Mongolia. In review in J Appl Ichthyol.

Warner SL. Randomized response: a survey technique for eliminating evasive answer bias. J Am Stat Assoc. 1965; 60(309) 63-69.

Williamson DH, Ceccarelli DM, Evans RD, Hill JK, Russ GR. Derelict fishing line provides a useful proxy for estimating levels of non-compliance with no-take marine reserves. PLoS One. 2014; 9(12): e114395.

Wolter F, Preisendörfer P. Asking sensitive questions an evaluation of the randomized response technique versus direct questioning using individual validation data. Sociol Methods Res. 2013; 42(3): 321-353.

Young T, Jensen OP, Weidel BC, Chandra S. Natural trophic variability in a large, oligotrophic, near-pristine lake. J Great Lakes Res. 2015; 41(2): 463–472.

Zhou S, Yin S, Thorson JT, Smith ADM, Fuller M. Linking fishing mortality reference points to life history traits: an empirical study. Can J Fish Aquat Sci. 2012; 69: 1292-1301.

Tables & Figures

Table 1. Natural mortality rates estimated by life history invariant methods and estimates of the effort required to exceed the sustainable harvest associated with each mortality rate.

Method	**Formula**[1]	**M**	**F**[2]	F_{MSY}[3]	**MSY (kg)**[4]	**# nights required**[5]	**# fishers required**[6]	**% families participating**[7]
$Hoenig_{nls}$ from Then et al. (2014)	$4.899 * t_{max}^{-0.916}$	0.37	0.06	0.32	330,869	22,058	220.6	14.6%
$Pauly_{nls-T}$ from Then et al. (2014)	$4.118 * K^{0.73} * Linf^{-0.33}$	0.27	0.15	0.24	255,285	17,019	170.2	11.3%
Gunderson (1997)	1.79 * GSI	0.30	0.12	0.26	279,557	18,637	186.4	12.3%

[1] See **Supp. Figure 6** for life history traits used in analysis.

[2] $F = Z - M$, where Z is 0.42 from the length-converted catch curve analysis (**Supp. Figure 7**).

[3] $F_{MSY} = 0.87 * M$, from Zhou et al. (2012).

[4] $MSY = (1 - exp(-F_{MSY})) * BIOMASS$, where Hovsgol grayling biomass is 1,214,400 kg based on Ahrenstorff et al. (2012).

[5] Number of nights required to reach MSY assuming fishers use 50-m of optimal mesh gillnet each night (15 kg grayling $night^{-1}$).

[6] Number of fishers required to reach MSY assuming each fisher uses 50-m of optimal mesh gillnet 100 nights per year.

[7] Percentage of families participating in the fishery assuming a resident population of 5,440 and average family size of 3.6 people per household (1,511 families; NSOM 2012).

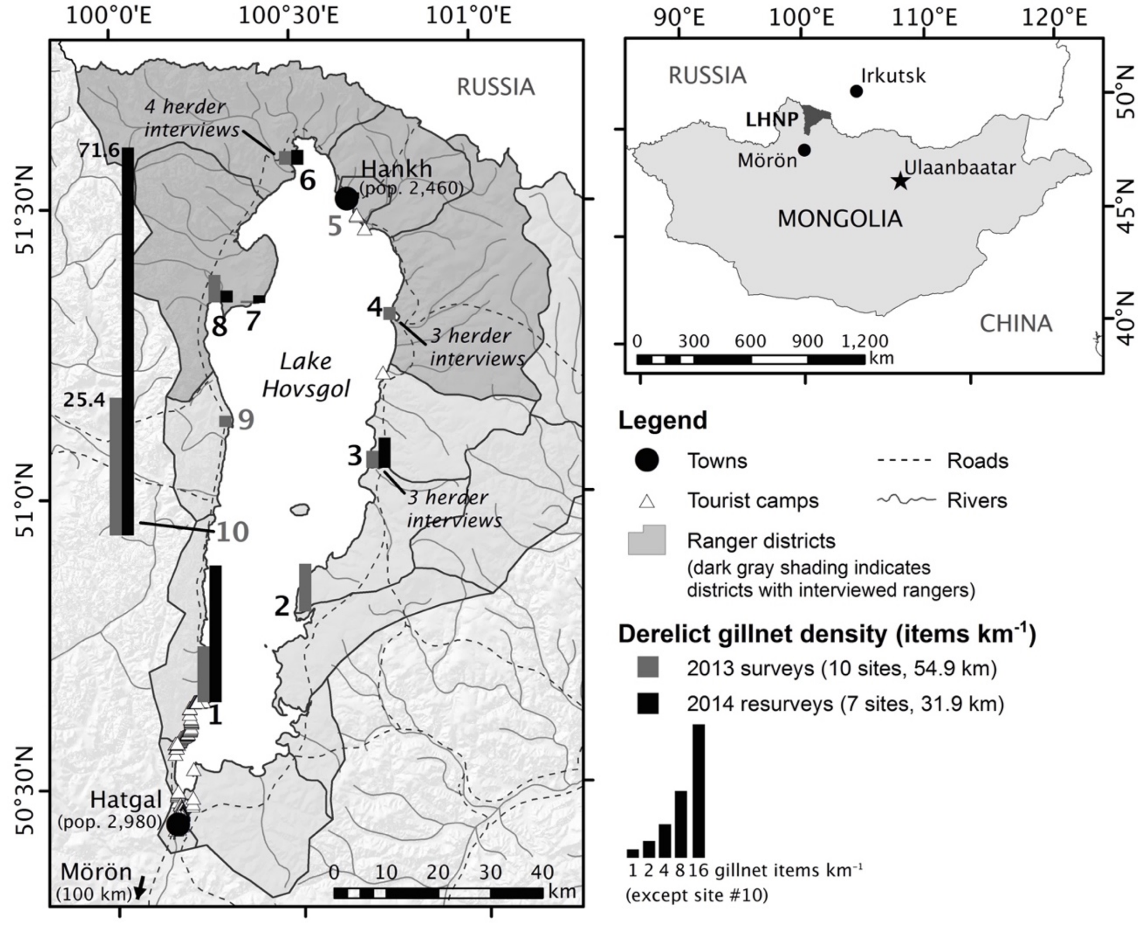


Figure 1. Location of shoreline surveys for derelict fishing gear, fish population monitoring sites, and interviews with park rangers and resident herders in Lake Hovsgol National Park (LHNP), Mongolia. Grey and black bars indicate the density (# km^{-1}) of derelict gillnet items observed in the 2013 (n=10) and 2014 (n=7) surveys, respectively (note different y-axis scale for Site 10). Black site numbers indicate the seven sites where fish population monitoring surveys were conducted in 2009 and 2011-13. Solid black lines indicate the park boundary and 17 ranger districts. Five rangers from five districts (dark grey; Hankh town limits represent one district) were interviewed. Herders were interviewed at Sites 3 (n=3), 4 (n=3), and 6 (n=4). Small white triangles indicate tourist camps, large black circles indicate town centers, dotted black lines indicate primitive roads, and solid gray lines indicate rivers and seasonal steams.

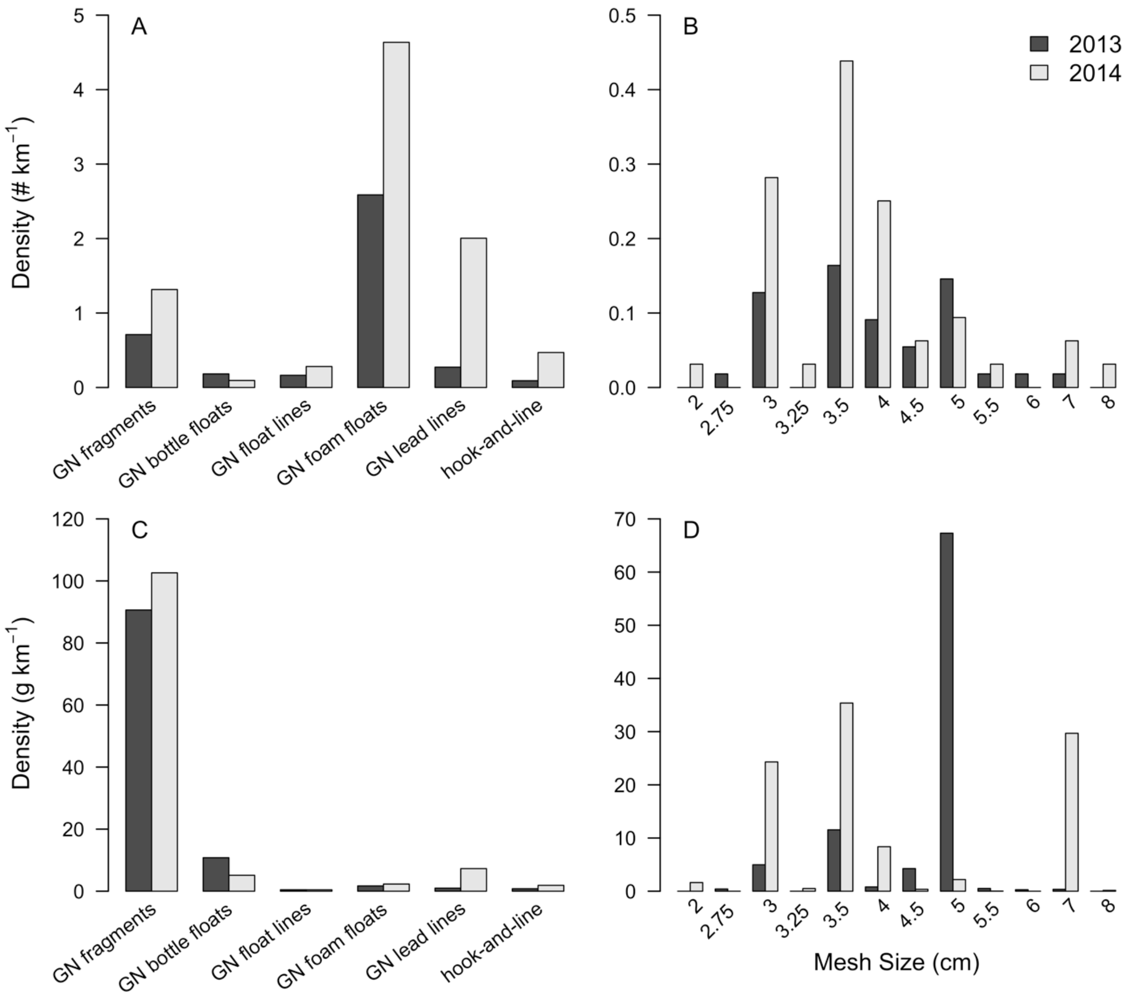


Figure 2. Average density of derelict fishing gear by category (GN=gillnet material) and derelict gillnet fragments by mesh size in count and weight among the 2013 (dark grey, n_{sites}=10, $n_{transects}$=14) and 2014 shoreline transects (light grey, n_{sites}/$n_{transects}$=7) weighted by transect length. Panels A and B indicate density in count (# km^{-1}) and Panels C and D indicate density in weight (g km^{-1}). Note variable y-axis scales.

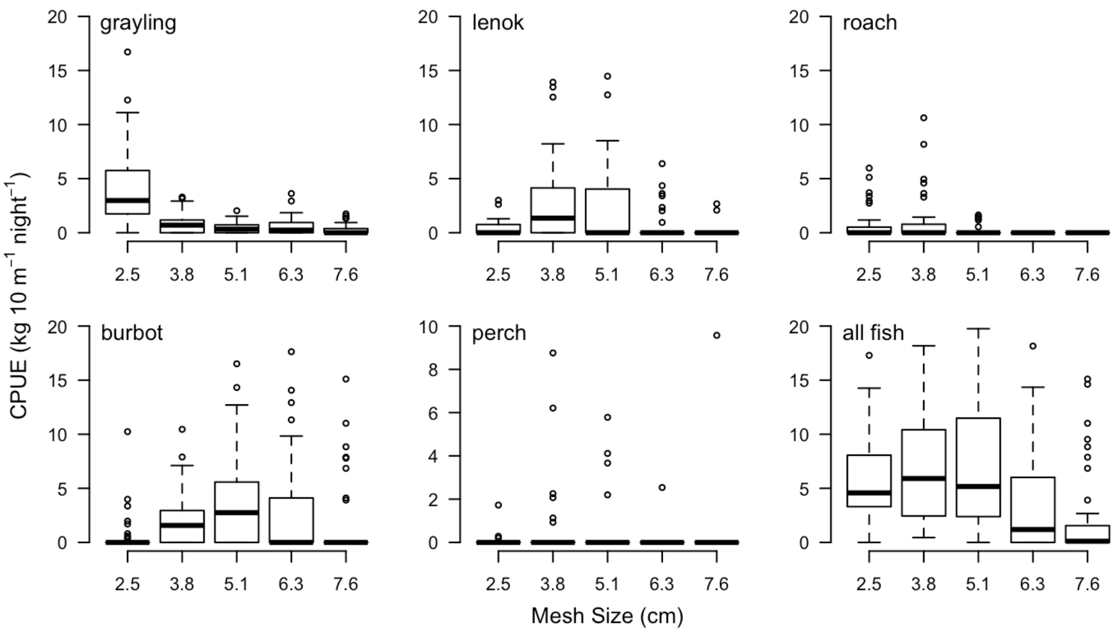


Figure 3. Catch-per-unit-of-effort (CPUE; kg 10 m^{-1} $night^{-1}$) by mesh size for the five most abundant species in gillnet catches and the sum of their weight from the two 5-panel sequential mesh gillnets used at seven sites in 2009 and 2011-2013 (14 sets yr^{-1}, 56 sets total). Boxplots indicate median (heavy black line), interquartile range (IQR; box), 1.5 times the IQR (whiskers), and extreme values (open circles). Note variable y-axis scales.

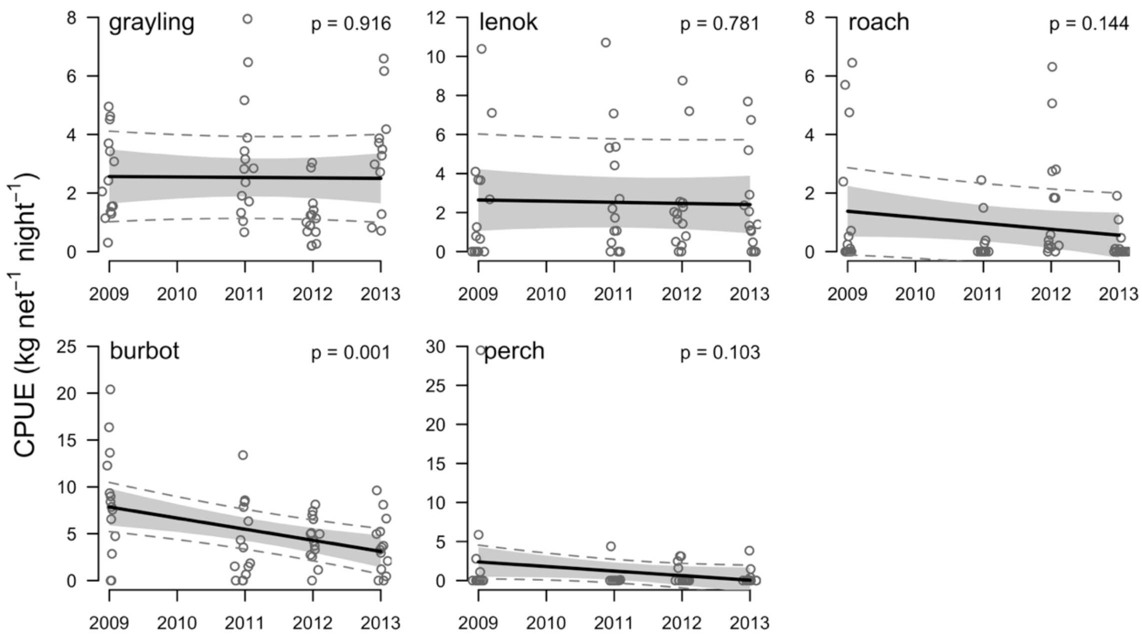


Figure 4. Trends in the abundance of the five most abundant fish species in gillnet catches from 2009-2013. Points indicate the CPUE (kg net^{-1} night^{-1}) of each 5-panel sequential mesh gillnet set (2 nets site^{-1} x 7 sites yr^{-1} = 14 sets yr^{-1}). Dark lines indicate linear mixed effects regressions fit to the catch data, gray shading indicates the confidence interval for each regression, and dashed lines indicate the prediction interval for the data. P-values are indicated in the upper right corner of each panel. Points are jittered around year for display. Note variable y-axis scales.

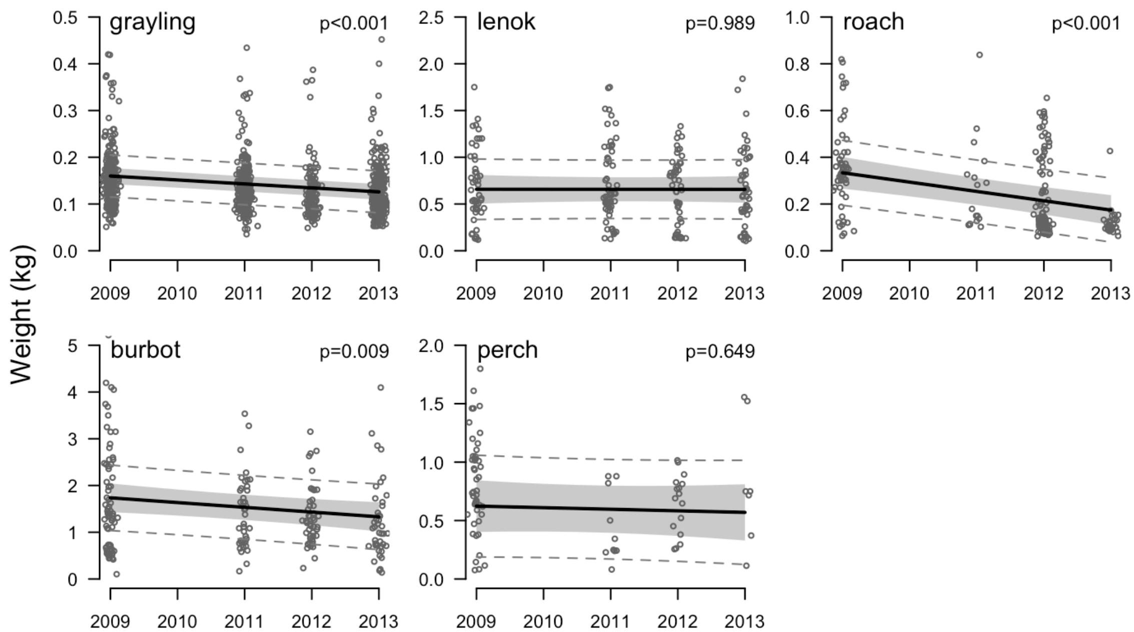


Figure 5. Trends in the body size of the five most abundant fish species in gillnet catches from 2009-2013. Points indicate the weight (kg) of every fish caught in gillnet sets that year (2 nets site^{-1} x 7 sites yr^{-1} = 14 sets yr^{-1}). Dark lines indicate linear mixed effects regressions fit to the catch data, gray shading indicates the confidence interval for each regression, and dashed lines indicate the prediction interval for the data. P-values are indicated in the upper right corner of each panel. Points are jittered around year for display. Note variable y-axis scales.

Supplemental Tables & Figures

Supp. Table 1. Large-bodied fish species in Lake Hovsgol, Mongolia and their historic catch*, market price†, and fine per illegally caught fish‡.

Family	Scientific Name	Common Name	Historic Catch (t yr^{-1})	Market Price (₮)	Fine Amount (₮)
Thymallidae	*Thymallus nigrescens*	Hovsgol grayling	100	1,500	5,000
Lotiidae	*Lota lota*	Burbot	50	15,000	13,000
Percidae	*Perca fluviatilis*	Eurasian perch	50	---	5000
Salmonidae	*Brachymystax lenok*	Lenok	24	---	17,000
Cyprinidae	*Rutilus rutilus*	Roach	---	---	5000

* From Dulmaa 1999.

† Mendsaikhan personal observation. ₮ denotes Mongolian tugrik. In July 2014, US$1 was approximately ₮1825.

‡ Established by Protocol #23 (2011) of the Mongolian Law on Hunting (2000) according to ecological evaluation.

Supp. Table 2. Life history invariant methods selected for estimating Hovsgol grayling natural mortality rate.

	Method	Formula*†	Used?	Comments
1	Hoenig 1983	$M \sim t_{max}$	YES	Used updated regression from Then et al. 2014 (based on 226 populations)
2	Pauly 1980	$M \sim L_{inf}/W_{inf}$, K, (T)	YES	Used updated regression from Then et al. 2014 (based on 218 populations)
3	Gunderson 1997	M ~ GSI	YES	Used regression from Gunderson 1997 (based on 28 populations)
4	Tanaka 1960 (rule of thumb)	$M \sim t_{max}$	NO	Does not perform better than Hoenig's estimator (Then et al. 2014)
5	Sekharan 1975	$M \sim t_{max}$	NO	Does not perform better than Hoenig's estimator (Then et al. 2014)
6	Griffiths & Harrod 2007	$M \sim L_{inf}/W_{inf}$, K	NO	Does not perform better than Pauly's estimator (Then et al. 2014)
7	Jensen 1996 - K	M ~ K	NO	Does not perform better than Pauly's estimator (Then et al. 2014)
8	Jensen 2001 - K, T	M ~ K, T	NO	Does not perform better than Pauly's estimator (Then et al. 2014)
9	Roff 1984 - L_{inf}, K, L_{mat}	$M \sim L_{inf}$, K, L_{mat}	NO	Requires L_{mat}
10	Charnov & Berrigan 1990	$M \sim t_{mat}$	NO	Requires validated t_{mat}
11	Rikhter & Efanov 1976 - t_{mat}	$M \sim t_{mat}$	NO	Requires validated t_{mat}
12	Jensen 1996 - t_{mat}	$M \sim t_{mat}$	NO	Requires validated t_{mat}
13	Roff 1984 - K, t_{mat}	M ~ K, t_{mat}	NO	Requires validated t_{mat}
14	Rikhter & Efanov 1976 - β, K, t_0, t_{mat}	M ~ β, K, t_0, t_{mat}	NO	Requires validated t_{mat}
15	Alverson & Carney 1975	M ~ K, t_{max}	NO	Requires data from unexploited era
16	Zhang & Megrey 2006	M ~ β, K, t_0, t_{mb}	NO	Requires data from unexploited era; requires validated t_{mb}; requires t_0 fit
17	Cubillos et al. 1999	M ~ K, t_0	NO	Requires t_0 fit [Tsogotsaikhan et al. (in review) fix t_0 at 0]
18	Frisk et al. 2001	$M \sim K/t_{max}$	NO	For elasmobranchs
19	Ralston 1987	M ~ K	NO	For Lutjanid snappers and Serranid groupers
20	Djabali et al. 1994	$M \sim L_{inf}/W_{inf}$, K	NO	For Mediterranean Sea fish; uses estimated M values; performs poorly
21	Alagaraja 1984	$M \sim L_{inf}$, K, t_0	NO	Dubious assumptions about age at L_{inf}; requires t_0 fit
22	Groeneveld 2000	$M \sim L_{inf}$, K, L_{mat}	NO	Estimates cannot be replicated; severe overestimation; requires L_{mat}

23	Lorenzen 1996	$M_w \sim w$	NO	M varies with individual weight
24	Peterson & Wroblewski 1984	$M_w \sim w$	NO	M varies with individual weight; includes non-fish animals
25	Ursin 1967	$M_w \sim w$	NO	M varies with individual weight; severe underestimation
26	Jennings & Dulvy 2008	$M_w \sim T, w$	NO	M varies with individual weight
27	Gislason et al. 2010	$M_L \sim L_{inf}, K, L$	NO	M varies with individual length
28	Chen & Watanabe 1989	$M_t \sim K, t_0, t_s, t$	NO	M varies with individual age; requires extreme assumptions; performs poorly
29	Bayliff 1967	$Z \sim t_{max}$	NO	Estimates Z; for Engraulidae (anchovies) with only 6 data points
30	Kenchington 2014	$Z \sim t_{max}, t_c, n_e$	NO	Estimates Z; problems with proof (Then et al. 2014)
31	Beverton & Holt 1959	$Z \sim L_{inf}, L_{mean}, L_{crit}$	NO	Estimates Z

* M = natural mortality rate; M_L = natural mortality rate at length L; M_t = natural mortality rate at age t; M_w = natural mortality rate at weight w; Z = total mortality rate

† β = exponent of the length/weight relationship; GSI = gonadosomatic index (wet ovary weight over wet body weight); K = parameter of von Bertalanffy growth curve; L = fish length; L_{crit} = minimum length of individuals fully represented in the catch; L_{inf} = asymptotic fish length; L_{mat} = length at reproductive maturity; L_{mean} = mean length of individuals between L_{crit} and the maximum length in the catch; n_e = effective sample size; t = fish age; T = mean environmental temperature; t_0 = parameter of the von Bertalanffy growth curve; t_c = youngest age fully represented in the catch; t_{mat} = age at reproductive maturity; t_{max} = maximum observed or assumed age; t_{mb} = age at which year-class achieves its maximum biomass in the absence of fishing; t_s = age at onset of senescence; w = fish weight; W_{inf} = asymptotic fish weight

Supp. Table 3. Arctic grayling (*Thymallus arcticus*) natural mortality rates reported in the literature.

Study	Location	Time period	M
Clark 1993	Fielding Lake, AK, USA	1986-1990 (4 yr)	0.24
Fleming 1995	Piledriver Slough, AK, USA	1993-1994 (1 yr)	0.27
Buzby & Deegan 2000	Kupurak River, AK, USA	1983-1998 (15 yr)	0.28
Clark 1992	Chena River, AK, USA	1979-1988 (9 yr)	0.31
Clark 1995	Chena River, AK, USA	1991-1994 (3 yr)	0.34
		Average:	**0.29**

Supp. Table 4. Responses of ten herding families interviewed about their personal fishing habits, fishing activity they observe, and status and conservation of fish in the lake.

Question	Responses (out of 10 respondents)	Response notes
Does your family fish?	yes, often (3) yes, sometimes (2) yes, rarely (5) no response (0)	Herding families on the northwestern shore rarely fish (~once during the spawning migrations). Herding families on the northeastern shore fish more often.
Why does your family fish?*	food (10) income (1) recreation (1) no response (0)	Herding families fish primarily for food. Only one family sells fish when the rare opportunity arises.
When does your family fish?*	during the spring spawning migration only (7) all-year but especially during the spring spawning migration (3) no response (0)	Herding families fish primarily during the spring spawning migration when grayling are easy to catch.
What gear does your family use to fish?*	rod (2) gillnet (4) hand catch (1) no response (3)	Herding families on the northwestern shore fish primarily with rods or by hand. Herding families on the northeastern shore fish primarily with gillnets.
What fish do you catch and eat?*	grayling (10) lenok (2) burbot (2) perch (1) no response (0)	Grayling are often used to make soup. Grayling soup is thought to be healthy after a long cold winter.
Who else do you see fishing?*	local Mongolians (8) Mongolians from Hatgal (6) foreigners (mainly Russians) (8) no response (0)	Local Mongolians fish for subsistence with rods and gillnets. Mongolians from Hatgal fish commercially with gillnets. Foreigners fish recreationally primarily with rods but sometimes with gillnets.

Why do they fish?*	food (9)	
	income (5)	
	recreation (7)	
	no response (0)	
What gear do they use to fish?*	rod (6)	
	gillnet (9)	
	no response (0)	
When do they fish?*	spring (10)	Local people fishing for subsistence fish primarily during the spring spawning migration. Tourists fish primarily during summer and winter. Commercial fishers fish year-round.
	summer (2)	
	fall (0)	
	winter (5)	
	no response (0)	
Have fish populations increased, decreased, or remained the same?	increased (0)	Grayling spring spawning migrations are shorter and less intense. Lenok and burbot have become especially rare.
	decreased (9)	
	remained the same (0)	
	doesn't know (1)	
	no response (0)	
Have fish body sizes increased, decreased, or remained the same?	increased (0)	Large lenok and burbot have become especially rare.
	decreased (6)	
	remained the same (1)	
	doesn't know (2)	
	no response (1)	
What should be done to protect fish?*	improve enforcement (4)	Few concrete ideas were provided. Herding families on northwestern shore find enforcement to be effective.
	control cormorant population (1)	
	local people should protect the lake (7)	
	no response (0)	

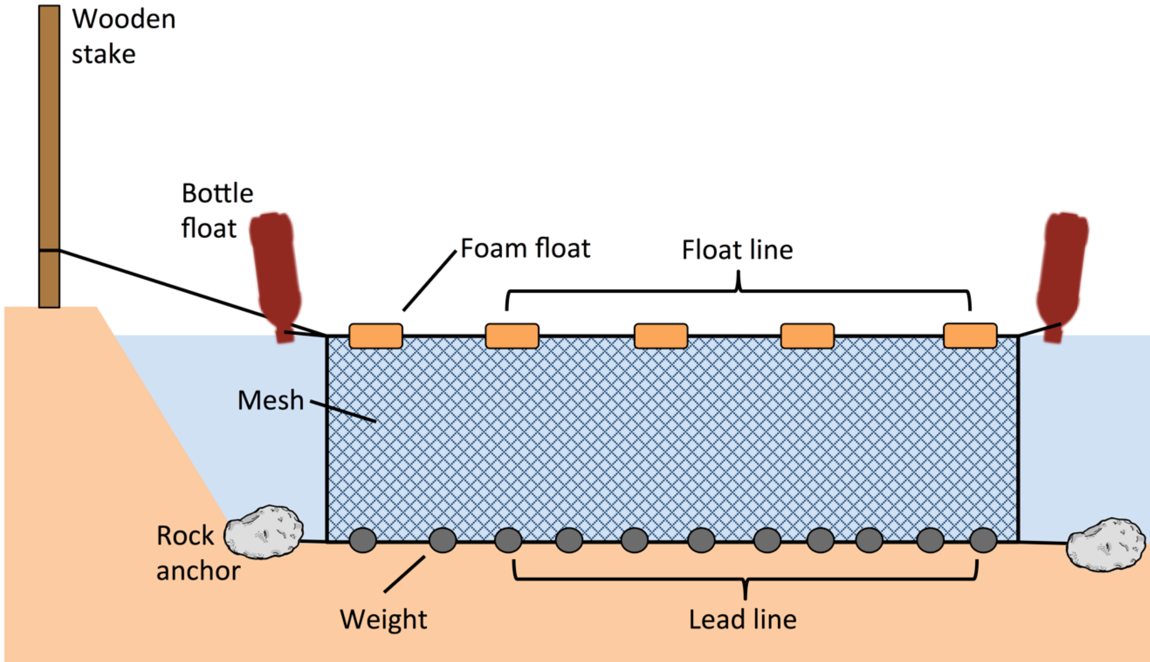
* Multiple responses possible for these questions.

Supp. Table 5. Responses of five park rangers interviewed about the frequency and character of illegal fishing, actions taken against illegal fishers, and status and conservation of fish in the lake.

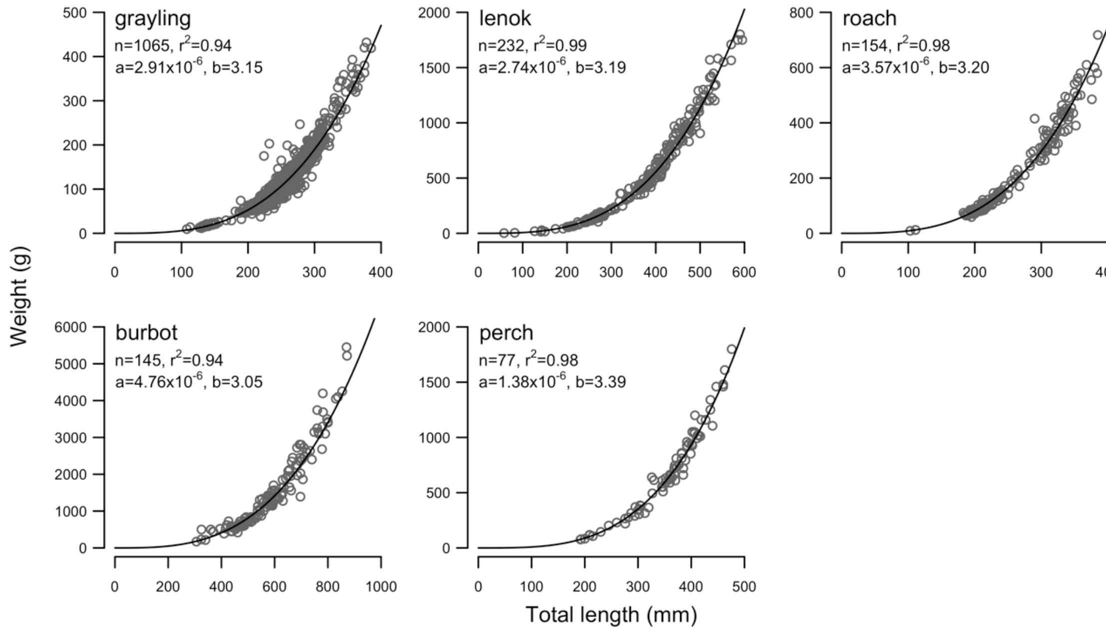
Question	Responses (out of 5 respondents)	Response notes
Who fishes?*	local Mongolians (5) non-local Mongolians (3) foreigners (3) no response (0)	Non-local Mongolians and foreigners fish recreationally. Local Mongolians fish for food and/or income. Commercial fishers come primarily from Hatgal.
Why do people fish?*	food (5) income (3) recreation (4) no response (0)	
What gear do people use?*	rods (4) nets (4) no response (0)	Recreational fishers use rods. Subsistence and commercial fishers use gillnets.
When do people fish?*	spring (3) summer (3) fall (3) winter (4) no response (0)	Recreational rod fishers fish all year-round but mostly in summer. Commercial and subsistence gillnet fishers fish mostly in the spring during the grayling migration and fall, winter, and spring when the lake is freezing or frozen.
What fish do people catch and eat?*	grayling (5) lenok (4) burbot (3) perch (1) no response (0) rod fishing, no (5) gillnet fishing, no (3)	Grayling and lenok are the primary targets of all types of fishing. Burbot are vulnerable to winter gillnet fishing and are targeted by commercial and subsistence fishers then. All fishers keep all of the fish they catch.

Is illegal fishing a significant problem in LHNP?*	gillnet fishing, yes (2) no response (0)	Most rod fishers are permitted and adhere to bag limits. Although 3 rangers say gillnet fishing is rare, 2 of these rangers report the confiscation of 60 gillnets in 3 years.
How do rangers enforce fishing regulations?*	fine (4) confiscate nets (2) no action necessary (1) no response (0)	Rangers report fining illegal fishers based on the quantity of their catch and confiscate their gillnets.
Have fish populations increased, decreased, or remained the same?	increased (2) decreased (3) remained the same (0) no response (0)	Most rangers report that fish populations have been decreasing and that lenok have become especially rare.
Have fish body sizes increased, decreased, or remained the same?	increased (0) decreased (1) remained the same (3) no response (1)	The majority of rangers report that fish body size has remained the same.
What should be done to protect fish?*	enforcement during spawn (2) research, education, enforcement (1) prevent upriver migration (1) control cormorant population (1) no response (0)	

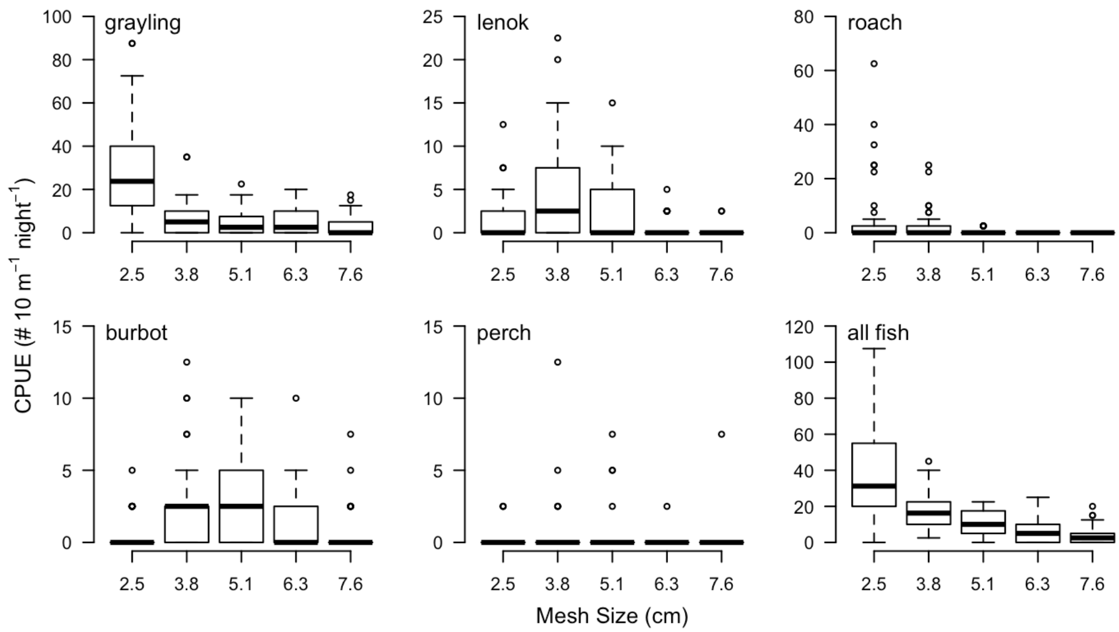
* Multiple responses possible for these questions.



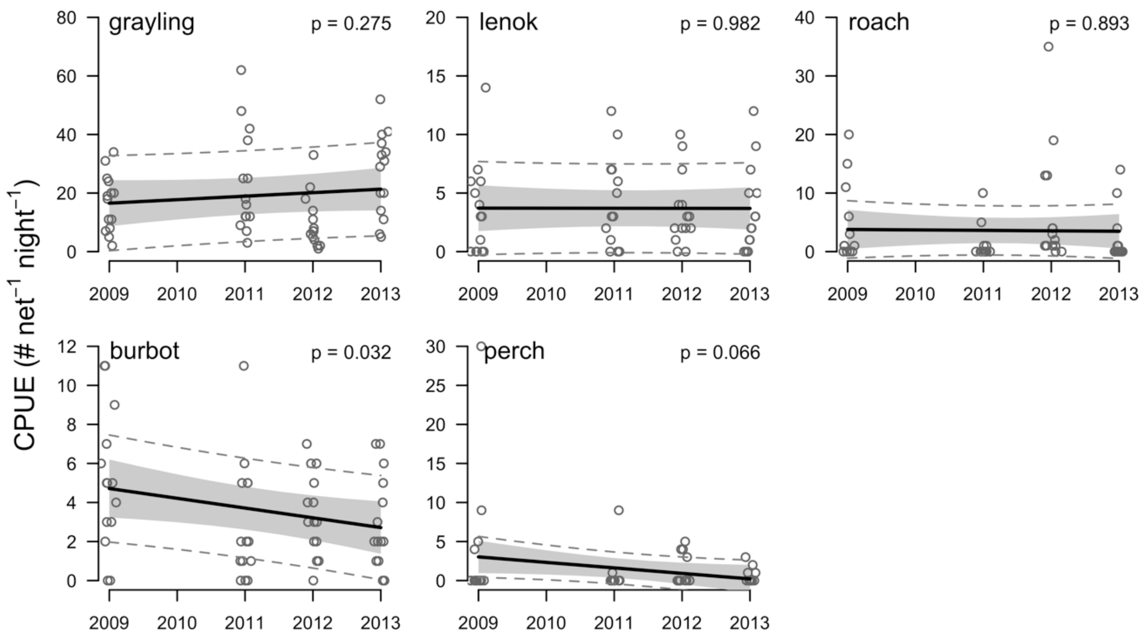
Supp. Figure 1. Diagram of a typical Mongolian horizontal gillnet and its components.



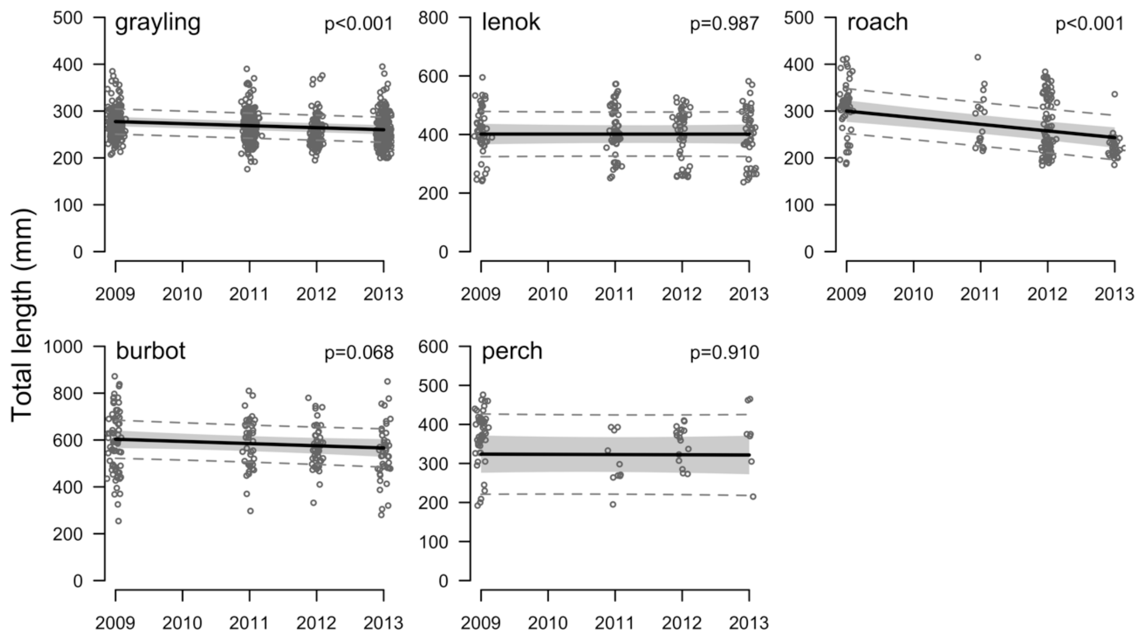
Supp. Figure 2. Length-weight relationships for the five most abundant fish species in gillnet catches in Lake Hovsgol. Note variable y-axis scales.



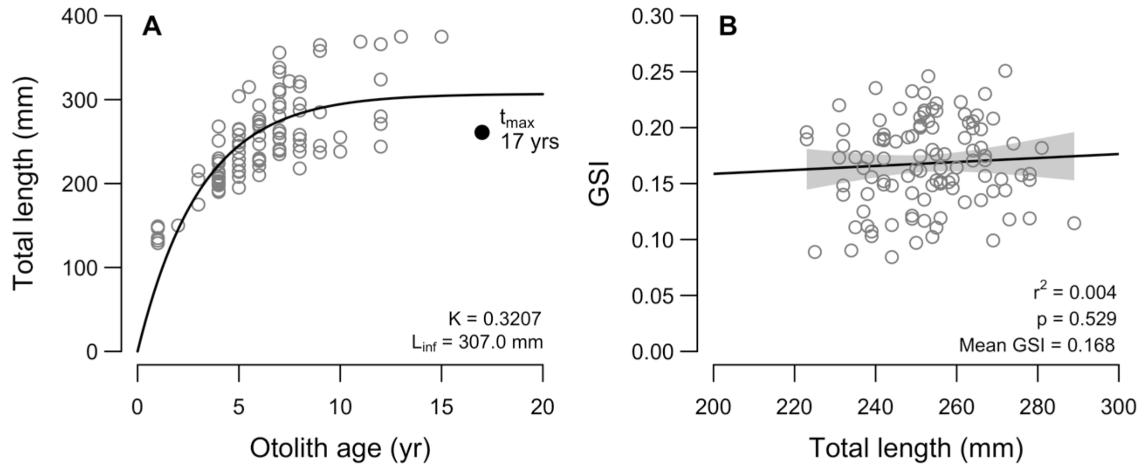
Supp. Figure 3. Catch-per-unit-of-effort (CPUE; # 10 m^{-1} $night^{-1}$) by mesh size for the five most abundant species in gillnet catches and the sum of their catch from the two 5-panel sequential mesh gillnets used at seven sites in 2009 and 2011-2013 (14 sets yr^{-1} x 4 yr = 56 sets total). Boxplots indicate median (heavy black line), interquartile range (IQR; box), 1.5 times the IQR (whiskers), and extreme values (open circles). Note variable y-axis scales.



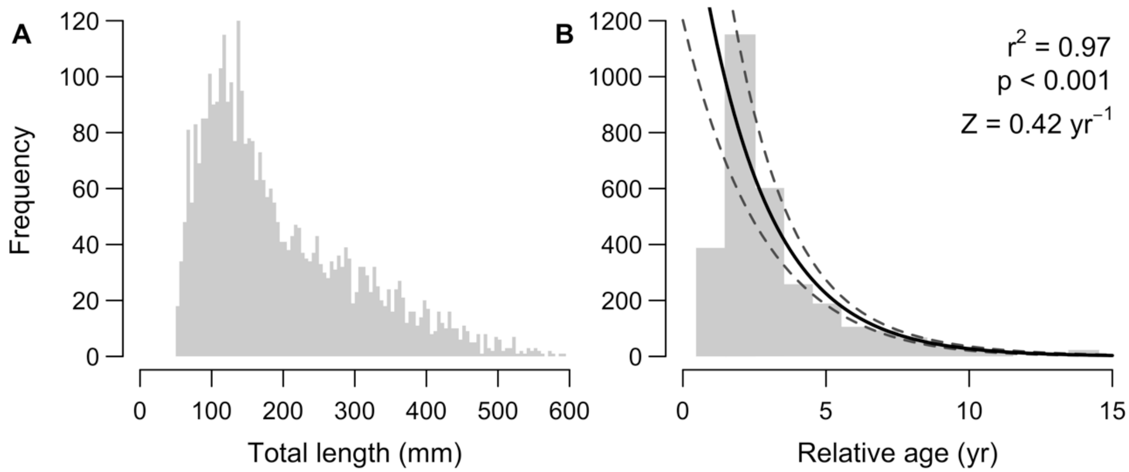
Supp. Figure 4. Trends in the abundance of the five most abundant fish species in gillnet catches from 2009-2013. Points indicate the CPUE (# net^{-1} night^{-1}) of each 5-panel sequential mesh gillnet set (2 nets site^{-1} x 7 sites yr^{-1} = 14 sets yr^{-1}). Dark lines indicate linear mixed effects regressions fit to the catch data, gray shading indicates the confidence interval for each regression, and dashed lines indicate the prediction interval for the data. Points are jittered around year for display. P-values are indicated in the upper right corner of each panel. Note variable y-axis scales.



Supp. Figure 5. Trends in the body size of the five most abundant fish species in gillnet catches from 2009-2013. Points indicate the total length (mm) of every fish caught in gillnet sets that year (2 nets site^{-1} x 7 sites yr^{-1} = 14 sets yr^{-1}). Dark lines indicate linear mixed effects regressions fit to the catch data, gray shading indicates the confidence interval for each regression, and dashed lines indicate the prediction interval for the data. P-values are indicated in the upper right corner of each panel. Points are jittered around year for display. Note variable y-axis scales.



Supp. Figure 6. Estimates of the life history characteristics used to calculate natural mortality (*M*) for Hovsgol grayling (see Table 1 for *M* estimation methods and results): (A) L_{inf}, *K*, and t_{max} were estimated from aged otoliths and a von Bertalanffy growth model (black line) fit through the observed age-size relationship and origin (Tsogotsaikhan et al. in review) and (B) *GSI* was estimated as the mean gonadosomatic index (GSI) for all observed grayling (Jensen, unpublished data). In (B) the black line indicates a linear regression fit and the grey shading indicates the confidence interval for the regression. Life history characteristics are marked and labeled in both panels.



Supp. Figure 7. The (A) length and (B) length-converted age structure of the Hovsgol grayling population. The length strucutre was observed in the Ahrenstorff et al. (2012) hydroacoustic surveys. In (B), the solid black line indicates a linear regression fit to the log-transformed trailing arm of the age structure. The dashed black lines indicate the confidence interval for the regression. Z is equal to the negative slope of the regression.

Appendix A: Herder interview questionnaire

Opening questions

1. Does your family eat fish?
2. How many times has your family eaten fish in the last month?
3. What types of fish does your family prefer to eat?
4. Does your family eat fish in all seasons?
5. How many people are in your family?
6. How many sheep, cows/yaks, goats, and horses does your family own?

Personal fishing habits

7. Does your family fish in Lake Hovsgol?
8. Does your family fish for recreation, food, or money?
    a. Do you keep all of the fish you catch? If not, which do you release?
9. How many times in the last month has your family fished?
10. For how many years has your family been fishing?
11. Who in your family fishes?
12. Where does your family go to fish?
13. How do you catch your fish?
    b. What type and how much net do you own?
    c. What type and how many rods do you own?
14. Where do you get your fishing equipment?
15. How many fish do you usually catch in a day of fishing?
16. What types of fish do you usually catch?

Observed fishing habits

17. Do you see other people fishing?
18. Are they mostly locals or foreigners?
19. Where do you see people fishing?
20. What type of gear do they use?
21. Do you know where you can buy this gear?
22. Do they fish for recreation, food, or money?

Fish market questions

23. Can you buy fish locally?
24. When and where can you buy fish?
25. What types and how much fish are usually available?
26. Does your family ever buy fish?
    a. What type of fish do you buy?
    b. How much fish and how frequently?
27. Are there any rules about fishing on Lake Hovsgol?

Fish population questions

28. Are fish more or less abundant than they used to be?
29. Are fish larger or smaller than they used to be?
30. What do you think should be done to protect the fish population?

Appendix B: Park ranger interview questionnaire

Opening Questions

1. How long have you been a ranger?
2. What is your district?
3. How long have you worked in this district?
4. How many families live in your district?
5. How many families in your district fish?

Observed Fishing Habits

6. When do you see people fishing?
7. Which season is the most active for fishing?
8. How many people do you see fishing in a month?
9. Are they mostly local, visiting Mongolians, or foreigners?
10. Where do you see people fishing?
11. What type of fishing equipment do they use?
12. What type of fish do they catch?
13. What type of fish do they keep?
14. What type of fish do they release?
15. Do they fish for recreation, food, or money?
16. How many fish do they catch in a day of fishing?

Law Enforcement Questions

17. Are fishermen complying with the law?
18. What do you do when you see people fishing illegally?
19. Why do you give a fine sometimes and not other times?

Fish Population Questions

20. Are fish more or less abundant than they used to be?
21. Are fish larger or smaller than they used to be?
22. What do you think should be done to protect the fish population?

Appendix C: Total and natural mortality estimation

Total mortality: length-converted catch curve analysis

We used a length-converted catch curve analysis (Pauly & Morgan 1987) to calculate the instantaneous total mortality (Z) rate where Z is the sum of the instantaneous natural (M) and fishing (F) mortality rates. First, we converted the length structure of Hovsgol grayling observed in the Ahrenstorff et al. (2012) hydroacoustic surveys into an age structure by (1) assigning observed fish to 5 cm length classes and (2) calculating the relative age of each fish based on the midpoint of its length class using the rearranged von Bertalanffy growth equation:

$$t_i = -\log\left(1 - \frac{L_i}{L_{inf}}\right)/K$$

where t_i is the mean age for the length class with midpoint L_i and L_{inf} is 307.0 mm and K is 0.3206 yr^{-1} from Tsogotsaikhan et al. (*in review*). A linear regression was fit to the log-transformed trailing arm of the resulting age structure where the negative slope of the regression is equal to the instantaneous total mortality rate (Z).

Natural mortality: life history invariant analysis

Natural mortality rate (M) is one of the most important parameters in fisheries population dynamics and management but can be difficult and expensive to estimate directly. As a result, many authors have developed simpler, though necessarily less reliable, methods for indirectly estimating M from life history traits such as maximum age, Von Bertalanffy growth parameters, and maturity/reproductive characteristics. Kenchington (2014) provides the best review of such life history invariant methods to date. We evaluated the appropriateness of 31 life history invariant methods for estimating the natural mortality rate of Hovsgol grayling based on (1) the availability of required life history traits and (2) the performance and biases of each method (Supp. Table 2). Ultimately, we decided to use three roughly independent estimators of natural mortality: Hoenig's (Hoenig 1983), Pauly's (Pauly 1980), and Gunderson's estimators (Gunderson 1997). A recent paper by Then et al. (2014) suggests that Hoenig's (t_{max}-based) and Pauly's estimators (growth-based) are the best in their respective classes and are independent of one another. Gunderson's estimator is GSI-based (GSI=gonadosomatic index) and is therefore expected to be independent of the other estimators. Although Then et al. (2014) suggest that the Hoenig estimator performs best of the t_{max}-and growth-based estimators and should be used alone, we consider three independent estimators to account for uncertainty in our life history trait estimates and uncertainty in the natural mortality estimates. See Supp. Table 2 for estimation methods considered, Table 1 for estimation methods used, and Supp. Figure 6 for the life history traits required by the selected methods. See the section below for a description of the data sources and methods used to calculate these required life history traits.

Natural mortality: life history trait data sources

The Von Bertalanffy growth parameters (L_{inf} and K with t_0 fixed at 0) and maximum age (t_{max}) were determined from 93 aged otoliths in Tsogotsaikhan et al. (*in review*) (Supp. Figure 6A). Thin sections through the core of the otoliths in the transverse plane were prepared and examined under a compound microscope (50x) using transmitted light. Alternating light and dark circuli, interpreted as annuli, were counted by two otolith readers. A von Bertalanffy growth model of length-at-age (L_t):

$$L_t = L_{inf} \times \left(1 - e^{-K(t-t_0)}\right)$$

was fit to the resulting age and length data with t_0 fixed at 0. L_{inf} (read "L-infinity") represents the average maximum size or asymptotic length and K represents the rate at which L_{inf} is approached. See Tsogotsaikhan et al. (*in review*) for more details.

The gonadosomatic index (*GSI*) was calculated for 106 grayling by dividing the wet ovary weight by the wet body weight (Jensen, unpublished data; Supp. Figure 6B). Linear regression suggests that *GSI* does not vary with body length (r^2=0.004, p=0.529) indicating that an average *GSI* value is representative of the entire grayling population.

We compared the indirect estimates of Hovsgol grayling natural mortality rate calculated here to direct estimates of natural mortality rate for Arctic grayling (*Thymallus arcticus*) from the literature to confirm realism (Supp. Table 3).

Appendix D: Supplemental references

Mongolia fish and fisheries references

Ahrenstorff TD, Jensen OP, Weidel BC, Mendsaikhan B, Hrabik TR. Abundance, spatial distribution, and diet of endangered Hovsgol grayling (*Thymallus nigrescens*). Environ Biol Fishes. 2012; 94(2): 465-476.

Dulmaa A. Fish and fisheries in Mongolia. In: Petr T, editor. Fish and fisheries at higher latitudes: Asia. FAO Fisheries Technical Paper 385. Rome, Italy: FAO; 1999.

Tsogtsaikhan P, Mendsaikhan B, Jargalmaa G, Ganzorig B, Weidel BC, Filosa CM, Free CM, Young T, Jensen OP. Comparison of age and growth of lenok, *Brachymystax lenok* (Pallas 1773), and grayling, *Thymallus* spp., in lentic and lotic habitats of Northern Mongolia. In review in J Appl Ichthyol.

Life history invariant methods references

Alagaraja K. Simple methods for estimation of parameters for assessing exploited fish stocks. Indian J Fish. 1984; 31(2): 177-208.

Alverson DL, Carney MJ. A graphic review of the growth and decay of population cohorts. ICES J Mar Sci. 1975; 36(2): 133-143.

Bayliff WH. Growth, mortality, and exploitation of the Engraulidae, with special reference to the anchoveta, *Cetengraulis mysticetus*, and the Colorado, *Anchoa naso*, in the eastern Pacific Ocean. Bulletin of the Inter-American Tropical Tuna Commission. 1967; 12(5): 365-432.

Beverton RJH, Holt SJ. A review of the lifespans and mortality rates of fish in nature, and their relation to growth and other physiological characteristics. In: Wolstenholme GEW, O'Connor M, editors. Ciba Foundation Symposium – the Lifespan of Animals (Colloquia on Ageing). Chichester, UK: John Wiley & Sons, Ltd; 1959. pp. 142–180.

Charnov EL, Berrigan D. Dimensionless numbers and life history evolution: age of maturity versus the adult lifespan. Evol Ecol. 1990; 4(3): 273-275.

Chen S, Watanabe S. Age dependence of natural mortality coefficient in fish population dynamics. Nippon Suisan Gakk. 1989; 55(2): 205-208.

Cubillos LA, Alarcón R, Brante A. Empirical estimates of natural mortality for the Chilean hake (*Merluccius gayi*): evaluation of precision. Fish Res. 1999; 42(1): 147-153.

Djabali F, Mehailia A, Koudil M, Brahmi B. A reassessment of equations for predicting natural mortality in Mediterranean teleosts. Naga ICLARM Q. 1994; 17(1): 33-34.

Frisk MG, Miller TJ, Fogarty MJ. Estimation and analysis of biological parameters in elasmobranch fishes: a comparative life history study. Can J Fish Aquat Sci. 2001; 58(5): 969-981.

Gislason H, Daan N, Rice JC, Pope JG. Size, growth, temperature and the natural mortality of marine fish. Fish Fish. 2010; 11(2): 149-158.

Griffiths D, Harrod C. Natural mortality, growth parameters, and environmental temperature in fishes revisited. Can J Fish Aquat Sci. 2007; 64(2): 249-255.

Groeneveld JC. Stock assessment, ecology and economics as criteria for choosing between trap and trawl fisheries for spiny lobster *Palinurus delagoae*. Fish Res. 2000; 48(2): 141-155.

Gunderson D. Trade-off between reproductive effort and adult survival in oviparous and viviparous fishes. Can J Fish Aquat Sci. 1997; 54(5): 990-998.

Hoenig JM. Empirical use of longevity data to estimate mortality-rates. Fish Bull. 1983; 82(1): 898-903.

Jennings S, Dulvy NK. Beverton and Holt's insights into life history theory: influence, application and future use. In: Payne AI, Cotter AJR, Potter ECE, editors. Advances in Fisheries Science: 50 Years on From Beverton and Holt. Oxford: Blackwell; 2008. pp. 434-450.

Jensen AL. Beverton and Holt life history invariants result from optimal trade-off of reproduction and survival. Can J Fish Aquat Sci. 1996; 53(4): 820-822.

Jensen AL. Comparison of theoretical derivations, simple linear regressions, multiple linear regression and principal components for analysis of fish mortality, growth and environmental temperature data. Environmetrics. 2001; 12: 591–598.

Lorenzen K. The relationship between body weight and natural mortality in juvenile and adult fish: A comparison of natural ecosystems and aquaculture. J Fish Biol. 1996; 49(4): 627–647.

Kenchington TJ. Natural mortality estimators for information-limited fisheries. Fish Fish. 2014; 15(4): 533-562.

Pauly D. On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. ICES J Mar Sci. 1980; 39(2): 175-192.

Pauly D, Morgan GR. Length-based methods in fisheries research. ICLARM Conference Proceedings 13. Manila, Philippines: International Center for Living Aquatic Resources Management; 1987.

Peterson I, Wroblewski JS. Mortality rate of fishes in the pelagic ecosystem. Can J Fish Aquat Sci. 1984; 41(7): 1117-1120.

Ralston S. Mortality rates of snappers and groupers. In: Polovina JJ, Ralston S, editors. Tropical Snappers and Groupers: Biology and Fisheries Management. Boulder, CO: Westview Press; 1987. pp. 375–404.

Richter VA, Efanov VN. On one of the approaches to estimation of natural mortality of fish populations. Kaliningrad, Russia: AtlantNIRO; 1977 (In Russian).

Roff DA. The evolution of life history parameters in teleosts. Can J Fish Aquat Sci. 1984; 41(6): 989-1000.

Sekharan KV. Estimates of the stocks of oil sardine and mackerel in the present fishing grounds off the west coast of India. Indian J Fish. 1975; 21(2): 177-182.

Tanaka S. Studies on the dynamics and the management of fish populations. Bull Tokai Reg Fish Res Lab. 1960; 28: 1-200 (In Japanese with English summary).

Ursin E. A mathematical model of some aspects of fish growth, respiration, and mortality. Can J Fish Aquat Sci. 1967; 24(11): 2355-2453.

Zhang CI, Megrey BA. A revised Alverson and Carney model for estimating the instantaneous rate of natural mortality. Trans Am Fish Soc. 2006; 135(3): 620-633.

Arctic grayling life history references

Buzby KM, Deegan LA. Inter-annual fidelity to summer feeding sites in Arctic grayling. Environ Biol Fishes. 2000; 59(3): 319-327.

Clark RA. Influence of stream flows and stock size on recruitment of Arctic grayling (*Thymallus arcticus*) in the Chena River, Alaska. Can J Fish Aquat Sci. 1992; 49(5): 1027-1034.

Clark RA. Stock assessment of Arctic grayling in Fielding Lake during 1991 and 1992. Fisheries Data Series No 93-2. Anchorage, AK: AK Department of Fish and Game; 1993.

Clark RA. Stock status and rehabilitation of Chena River Arctic grayling during 1994. Fisheries Data Series No 95-8. Anchorage, AK: AK Department of Fish and Game; 1995.

Fleming DF. Stock assessment of Arctic grayling in Piledriver Slough during 1994. Fisheries Data Series No 95-15. Anchorage, AK: AK Department of Fish and Game; 1995.

Chapter 2: The refined ORCS approach: a catch-based method for estimating stock status and catch limits for data-poor fish stocks*

Abstract

The 'Only Reliable Catch Stocks' (ORCS) Working Group approach to data-poor fisheries stock status and catch limit estimation has been used by U.S. fisheries managers but has yet to be fully evaluated. The ORCS approach estimates stock status using a fourteen question 'Table of Attributes' and the overfishing limit by multiplying a historical catch statistic by a scalar based on the estimated status. We evaluated the performance of the approach by applying it to 193 stocks with data-rich stock assessments and comparing its predictions of stock status with the assessment model estimates. The approach classified all but three stocks as fully exploited indicating that it is a poor predictor of status and should not be used by managers. We refined the original ORCS approach by: (1) developing a more predictive model of stock status using boosted classification trees and (2) identifying the historical catch statistics and scalars that best estimate overfishing limits using assessment model data. The refined ORCS approach correctly classified 74% of all stocks and 62% of overexploited stocks in a training dataset and 74% of all stocks and 50% of overexploited stocks in an independent test dataset. The refined approach performed better than other widely used catch-only methods. However, the overfishing limits estimated by the refined approach would further

---

* Published as: Free CM, Jensen OP, Wiedenmann J, Deroba JJ (2017) The refined ORCS approach: a catch-based method for estimating stock status and catch limits for data-poor fish stocks. *Fisheries Research* 193: 60-70.

deplete overexploited stocks without the use of conservative catch scalars to buffer against classification uncertainty. Conservative catch scalars can reduce the probability of overfishing below 50%, the U.S. legal maximum, but with concomitant increases in the probability and magnitude of underfishing. The refined ORCS approach may therefore be useful when other methods are not possible or appropriate and some risk of underfishing is acceptable.

Introduction

The majority of global fish stocks lack adequate data for estimating sustainable fishing levels using conventional stock assessment methods. In developing countries, only 5-20% of fish stocks are assessed and this fraction increases to only 10-50% in developed countries (Costello et al., 2012). In the United States, 30% of stocks are managed using conventional 'data-rich' assessment methods, while the remaining 11% and 59% of stocks are managed using 'data-moderate' and 'data-poor' methods, respectively (Newman et al., 2015). Data-rich stock assessment methods combine (1) total catch; (2) an index of relative abundance; and (3) other biological information to assess stock status and estimate sustainable fishing levels (Walters and Martell, 2004). Data-poor and data-moderate methods generally utilize only one and two of these data types, respectively, with total catch information often being the only data type available. Thus, data-poor methods are often synonymous with catch-only methods.

In 2006, the U.S. Magnuson-Stevens Fishery Conservation and Management Act was amended to require scientifically-derived annual catch limits (ACLs) that prevent overfishing for all federally managed fish stocks, including data-limited stocks (DOC, 2007). This mandate stimulated the revival of old data-limited methods (Gulland, 1971; Restrepo et al., 1998), development of new data-limited methods (MacCall, 2009; Dick and MacCall, 2011; Cope, 2013; Cope et al., 2013), and evaluation of the relative performance of these methods (Wetzel and Punt, 2011; Wiedenmann et al., 2013; Carruthers et al., 2014). In 2011, the 'Only Reliable Catch Stocks' (ORCS) Working Group (Berkson et al., 2011) convened to evaluate catch-only methods for ACL estimation and recommended the following hierarchy for determining ACLs for ORCS: (1) depletion-based stock reduction analysis (DB-SRA; Dick and MacCall, 2011) when a complete time series of annual catches is available (i.e., from the start of fishing to the present); (2) depletion-corrected average catch (DCAC; MacCall, 2009) when the stock exhibits low natural mortality rates (≤0.20 yr^{-1}); and (3) the new ORCS Working Group approach (hereafter called the 'ORCS approach') when neither DB-SRA or DCAC are possible or appropriate (Berkson et al., 2011; later modified by SAFMC, 2012, 2013).

The ORCS approach was designed to provide an ecological basis for the Restrepo et al. (1998) scalar approach. In both methods, the overfishing limit (OFL; the catch at F_{MSY}) is calculated by multiplying an expert-defined historical catch statistic (e.g., mean catch over the previous 10 years or median catch over the whole time series) by a scalar also based on expert judgment. In the ORCS approach, the choice of scalar is determined by stock status (i.e., under, fully, or overexploited), which is estimated as the mean score of fourteen stock- and fishery-related attributes (the 'Table of Attributes' or TOA; **Table 1**). The ORCS approach allows for considerable flexibility in its implementation, as scientists and managers can exercise expert judgement to: (1) estimate status using an arithmetic, geometric, or weighted mean of the Table of Attributes scores; (2) modify the Table of Attributes' estimate of status or the thresholds used to delineate status; and/or (3) choose appropriate catch statistics and scalars. While this flexibility and reliance on expert judgement could improve performance, it is necessary to adopt a more specific, albeit less inclusive, definition of the ORCS approach to validate the method and demonstrate its transferability.

The ORCS approach is widely applicable, but the ability of the Table of Attributes to correctly predict stock status has not been evaluated and the performance of only a limited range of potential catch statistics and scalars has been tested. In the only explicit evaluation of the ORCS approach to date, Wiedenmann et al. (2013) used management strategy evaluation to show that the default scalars used to estimate the OFL are too conservative for under (scalar=0.5) and fully (scalar=1.0) exploited stocks and too generous for overexploited (scalar=2.0) stocks when stock status is correctly classified. They also show that catch limits are unsustainable when stocks are incorrectly classified into less-depleted categories (e.g., an overexploited stock incorrectly classified as fully exploited). Evaluations of scalar-based methods similar to the ORCS approach have also been shown to result in overfishing, especially for already depleted stocks and stocks whose statuses have been incorrectly classified (Carruthers et al. 2014; ICES 2014, 2015, 2017). The sensitivity of management outcomes to status classification decisions makes the validation and refinement of the ORCS Table of Attributes' ability to estimate status necessary before the ORCS approach can be used to set catch-limits more widely.

The goals of the present study are to evaluate and refine the ORCS approach to data-poor catch limit estimation using stocks with data-rich stock assessments. We evaluate the original approach by applying it to data-rich stocks and comparing its predictions of status with the assessment model estimates. We refine the ORCS approach by: (1) developing a more predictive model of stock status that uses boosted classification trees to weight attributes by their relative importance, incorporate interactions between attributes, and account for non-linearity in attribute behavior; and (2) empirically identifying the best status-specific historical catch statistics and scalars for estimating overfishing limits using assessment model data. Finally, we evaluate the ability of the refined ORCS approach to estimate overfishing limits and compare the ability of the refined approach to estimate stock status to six other catch-only assessment methods.

Methods

Stock selection

We evaluated the ORCS approach to data-poor catch limit estimation by applying it to data-rich stocks with stock assessments based on underlying population dynamics models (generally statistical catch-at-age models, virtual population analyses, and production models) in the RAM Legacy Stock Assessment Database (RAMLDB v.2.95; Ricard et al., 2012). We used only stocks with assessments that estimate B_{MSY} internal to the model or estimate standard proxies used by the management agency (e.g., spawning potential ratio proxies common in the U.S. or B_0 proxies common in Australia). We excluded stocks whose assessments are considered particularly unreliable (n=2; 2002 Atlantic croaker and 2005 Atlantic herring). The resulting 193 stocks include underexploited (n=68), fully exploited (n=95), and overexploited (n=30) stocks representing a variety of taxa, geographic locations, and management agencies (Figure 1). The RAMLDB does not include the most up-to-date assessment for every stock. Therefore, data-rich statuses and answers to the Table of Attributes questions reflect the terminal year of the assessment in the RAMLDB.

Evaluation of the ORCS Table of Attributes

We estimated stock status using the expanded Table of Attributes developed by SAFMC (2012) with a few modifications to increase clarity and objectivity in the scoring process (Table 1; Supplementary Appendix A.1). We scored: *TOA #1 Status of assessed stocks in fishery* using U.S. Fisheries Management Plans and their foreign analogs to identify groups of stocks managed together and references from management agencies to determine the status of these stocks; *TOA #2 Refuge availability*, *#3 Behavior affecting capture*, *#4 Morphology affecting capture*, and *#11 Habitat loss* using information on the distribution, biology, and habitat of the taxa in FishBase (Froese and Pauly, 2016); *TOA #5 Discard rate*, *#6 Targeting intensity*, *#7 M compared to dominant species*, *#8 Occurrence in catch,* and *#14 Proportion of population protected* using information in the stock assessment documents; *TOA #9 Value* using ex-vessel price data from the Sea Around Us Project (Pauly and Zeller, 2015); and *TOA #10 Recent trend in catch*, *#12 effort*, and *#13 abundance index* using time series in the RAMLDB. Other technical sources (i.e., government reports or websites, peer-reviewed scientific papers, technical reports) were used when an attribute could not be scored using the principal reference. In some cases, attributes could not be scored due to a lack of data or applicability and were given an 'NA' value. Detailed information on the scoring process is available in Supplementary Appendices A.2 and A.3 and the scores and their justifications are available in Supplementary Appendix B. Estimated stock status was determined from the mean of the Table of Attributes scores with the following classifications provided by the original method: underexploited (<1.5), fully exploited (1.5–2.5), and overexploited (>2.5). This simplification of the broadly flexible ORCS approach is necessary for testing and validating the performance of the method on such a diverse and global array of stocks.

The ORCS approach has been thought to estimate both stock status (i.e., lightly, moderately, and heavily exploited; Berkson et al. 2011) and the risk of overexploitation (i.e., low, moderate, and high risk of overexploitation; SAFMC, 2012, 2013). Consequently, we evaluated the performance of the original approach using linear regression to assess the correlation between predicted status (mean Table of Attributes score) and the assessment's most recent estimates of (1) B/B_{MSY} as a proxy for stock status and (2) F/F_{MSY} as a proxy for overexploitation risk. We also assessed the ability of the original approach to correctly classify stock status using both percentage agreement (accuracy) and Cohen's kappa. Cohen's kappa measures inter-rate agreement between categorical items and is more robust than simple percentage agreement because it takes into account the probability of agreement occurring by chance (Cohen, 1968). This metric was preferred given the volume and ease of identifying fully exploited stocks compared to the paucity and difficulty of identifying overexploited stocks. If the method misclassifies most overexploited stocks but correctly classifies most fully exploited stocks, then it would still earn a high accuracy percentage, but it's kappa value would be appropriately penalized. Although there are no definitive rules for interpreting Cohen's kappa, general guidelines suggest that values >0.70 are 'excellent', 0.4-0.7 are 'good', 0.2-0.4 are 'fair', and <0.2 are 'poor' (Landis and Koch, 1977; Fleiss, 1981).

Refinement of the ORCS Table of Attributes

We refined the ORCS Table of Attributes using boosted classification trees (BCT) to weight attributes by their relative importance, incorporate interactions between attributes, and account for non-linearity in attribute behavior. Boosted classification trees combine classification and machine learning and offer predictive power superior to other modeling methods (Elith et al., 2008). Boosted classification trees can also accommodate missing values (i.e., NA scores) by imputing values from surrogate variables, which allowed the use of all scored stocks. The BCT analysis was performed using the *caret* (Kuhn, 2016) and *gbm* (Ridgeway, 2016) packages in R v.3.3.2 (R Core Team, 2016).

We trained the BCT model to estimate categorical status (i.e., under, fully, or overexploited) rather than continuous status (i.e., B/B_{MSY}) because (1) the ORCS approach was designed to use status categories and (2) stock assessment models exhibit more uncertainty in estimates of B/B_{MSY} than in more general status classifications. We trained the BCT model to estimate stock status rather than risk of overexploitation because (1) stock status is a more widely used metric and can be easily compared to other assessment methods and (2) F/F_{MSY} is an unsatisfying proxy for overexploitation risk because it can change rapidly and even sustained F/F_{MSY} values greater than 1.0 may not be "risky" over the short-term if B/B_{MSY} is high ($\gg$1.0). The BCT model attempts to determine stock status – whether a stock is under (B/B_{MSY}>1.5), fully (B/B_{MSY}=0.5–1.5), or overexploited (B/B_{MSY}<0.5) – from the TOA scores with a few modifications (Table 1): (1) we removed *TOA #2 Refuge availability* and *#4 Morphology affecting capture* because they lacked contrast (i.e., 97.9% and 100% of stocks were assigned scores of 3–>75% of habitat accessible and 2–Average susceptibility, respectively); (2) we used continuous rather than categorical price values for *TOA #9 Value* because these values are readily available to managers and continuous variables can increase predictive performance; and (3) we used all three categories for *TOA #10 Recent trend in catch* (i.e., 1=increasing, 2=stable, and 3=decreasing rather than the originally proposed options of 1.5=increasing/stable and 3=decreasing) because boosted classification trees can account for interactions between catch, effort, and abundance index trends.

We randomly divided the TOA scores into training (80% of data, n=155 stocks) and test (20% of data, n=38 stocks) datasets with stratification by stock status to ensure that both the test and training datasets included the same proportion of under, over, and fully exploited stocks. The training dataset was used to fit the BCT model, while the test dataset was used to provide an independent evaluation of the BCT model's predictive capacity. A grid search for the BCT model parameters that maximize Cohen's kappa using repeated 10-fold cross validation on the training dataset found the following optimal parameters: learning rate=0.001, interaction depth=2, number of trees=3000, and bag fraction=0.8 with multinomial error. Detailed information on model fitting is available in Supplementary Appendix A.4.

We evaluated the predictive performance of the BCT model by calculating the percentage agreement and Cohen's kappa for both the training and test datasets. For comparison, we evaluated the performance of six other catch-only methods for estimating status on stocks in the test dataset: SSP-2002 (Froese and Kesner-Reyes, 2002) and SSP-2013 (Kleisner et al., 2013), which estimate development status (e.g., undeveloped, developing, fully exploited), and CMSY (Martell and Froese, 2013), COM-SIR (Vasconcellos and Cochrane, 2005), SSCOM (Thorson et al., 2013), and mPRM (Costello et al., 2012), which estimate B/B_{MSY} (Table 2). The latter four methods were applied using the *datalimited* package in R (Anderson, 2016) based on the methods described in Rosenberg et al. (2014) and Anderson et al. (2017). Detailed information on implementing the alternative catch-only methods is available in Supplementary Appendix A.5.

Refinement of the historical catch statistics and scalars

The second step of the ORCS approach is to estimate the OFL as a factor of some historical catch statistic based on stock status; however, the original approach offers no formal recommendations on the choice of catch statistic and recommends simple catch scalars (i.e., 2.0, 1.0, 0.5 for under, fully, and overexploited stocks, respectively).

We identified the best status-specific historical catch statistics and scalars by comparing the most recent OFL (U_{MSY} x total biomass) to 24 historical catch statistics for the 105 stocks in the RAMLDB with the necessary information (i.e., U_{MSY}, total biomass time series, and catch/landings time series). The 24 historical catch statistics represent eight metrics (IQR, Winsorized, and arithmetic mean; 10th, 25th, 50th, 75th, and 90th percentiles) proposed in the original ORCS approach over three time periods (10 yr, 20 yr, whole time series). We used linear regression to assess the correlation between the OFL and each catch statistic and Akaike's Information Criterion (AIC) to rank the catch statistics within each status category. The best status-specific catch statistics were selected based on AIC ranking.

We calculated the ratio of the best status-specific catch statistic to the OFL for each stock based on its data-rich status estimate. We then calculated the 10th to 50th percentile of the observed ratios in each status category to evaluate as potential status-specific scalars for estimating the OFL. If stock status is correctly identified, the 50th percentile scalars should promote a 50% probability of overfishing (i.e., catch > OFL) in a given year, the U.S. legal maximum (DOC, 2016). Scalars more conservative than the median may be useful for buffering against classification uncertainty. Detailed information on calculating the OFL and the best status-specific historical catch statistics and scalars is available in Supplementary Appendices A.3 and A.6.

Evaluation of the refined ORCS approach

We evaluated ten potential refinements of the original ORCS approach. The first approach (the 'weighted 50th percentile scalar' approach) uses the BCT model to estimate the probability a stock is in each status category. It then estimates the OFL as the probability weighted average of the OFLs for each status category using the best status-specific catch statistics and 50th percentile scalars. The second approach (the 'unweighted 50th percentile scalar' approach) uses the BCT model to identify the most likely status category, then estimates the OFL using the best catch statistic and 50th percentile scalar for the category. The remaining eight approaches use the 45th-10th percentile scalars in the unweighted framework to examine the tradeoffs associated with using scalars more conservative than the median. We used the unweighted framework because preliminary analysis showed that the unweighted framework was superior to the weighted framework (Table 4; Figure 6). We evaluated the performance of these approaches by applying them to the 97 stocks ($n_{training}$=79, n_{test}=18) in the RAMLDB with the necessary information (i.e., B/B_{MSY}, U_{MSY}, total biomass time series, catch/landings time series) and calculated the percentage of stocks for which the predicted OFL exceeded the data-rich OFL estimate to use as a measure of the probability of overfishing. We also assessed the correlation between the OFLs predicted by the ORCS approach and those estimated by the data-rich assessments using linear regression.

Results

Evaluation of the ORCS Table of Attributes

Although most attributes exhibited good variation in scores, a few were dominated by a single score category (*TOA #2, #4*), omitted an entire score category (*TOA #3*), or underutilized a score category (*TOA #11, #14*) (Figure 2A). The original approach classified all but three stocks as fully exploited (Figure 2B). Although the approach correctly classified the U.S. Mid-Atlantic weakfish stock as overexploited (B/B_{MSY}=0.131 in 2008), it incorrectly classified the fully exploited U.S. Gulf of Maine haddock (B/B_{MSY}=0.585 in 2011) and New Zealand bluenose (B/B_{MSY}=0.658 in 2011) stocks as overexploited. In fact, there was no correlation between the statuses predicted by the ORCS approach and those estimated by the data-rich assessment models (Figure 2C), and a Cohen's kappa value of 0.0001 indicates 'poor' classification accuracy. There was a weak correlation between the overexploitation risks predicted by the ORCS approach and those estimated by the data-rich assessment models (Figure 2D; Supp. Figure 1)

Refinement of the ORCS Table of Attributes

The BCT model correctly classified 74% of stocks in the training dataset and yielded a Cohen's kappa of 0.56 indicating 'good' classification accuracy (Figure 3A). The model performed better on fully exploited stocks (89% correct) than either underexploited (58% correct) or overexploited (62% correct) stocks. The BCT model also correctly classified 74% of stocks in the independent test dataset and yielded a Cohen's kappa of 0.56 indicating 'good' classification accuracy (Figure 3B). The model still performed better for fully exploited stocks (79% correct) than underexploited (77% correct) or overexploited (50% correct) stocks in the test dataset. The nearly equivalent performance of the BCT model on the training and test datasets suggests that the model is not overfit, which is consistent with the flat model tuning curves (Supplementary Appendix A.4). In 60% of misclassifications, the correct classification was the second most probable status identified by the model and only one misclassification (U.S. S. Pacific Coast gopher rockfish – no remarkable scores to explain this outcome) was so egregious as to classify an underexploited stock as overexploited or vice versa (Figure 3B). The BCT model was a better predictor of stock status, in terms of both accuracy and Cohen's kappa, than the other six catch-only methods that we evaluated (Table 2; Supp. Tables 1 & 2).

The BCT model identified seven attributes that each contribute more than 5% of the total predictive power (percents indicate relative influence of an attribute on the classification of a stock): *TOA #9 Value* (33.5%), *#1 Status of assess stocks in fishery* (13.1%), *#6 Targeting intensity* (12.3%), *#5 Discard rate* (8.8%), *#8 Occurrence in catch* (8.5%), *#7 M compared to dominant species* (8.0%), and *#3 Behavior affecting capture* (7.3%; Figure 4A). The attribute marginal effects, the effect of each attribute when the other attributes are held constant, suggest that stocks are more likely to be: (1) <u>underexploited</u> if there is a low rate of overexploitation of other stocks in the fishery, the taxon is worth less than US$1.00 per pound, and the taxon does not exhibit any aggregation behavior; (2) <u>fully exploited</u> if the stock is occasionally or actively targeted, the taxon exhibits aggregation behavior, and the taxon is worth more than US$2.00 per pound; and (3) <u>overexploited</u> if there is a high rate of overexploitation of other stocks in the fishery, the taxon is worth more than US$1.00 per pound, and the taxon occurs sporadically in the catch (Figure 4B; Supp. Figure 2).

Refinement of the historical catch statistics and scalars

The 90th percentile catch over the whole time series was most highly correlated with the OFL for underexploited stocks and longer timeframe metrics generally performed better than shorter timeframe metrics (Table 3; Supp. Table 3). The 25th percentile catch over the previous 10 years performed best for fully exploited stocks with more central and shorter timeframe metrics generally performing better than higher percentile and longer timeframe metrics (Table 3; Supp. Table 3). The mean catch of the previous 20 years performed best for overexploited stocks but this correlation was driven by a single strong leverage point (S. Labrador/E. Newfoundland Atlantic cod, whose 20-year mean exceeded the current OFL by more than 5 times, considerably more than the other overexploited stocks) and may be spurious. The 10th percentile catch over the whole time series provided the second best correlation and is more appropriate for overexploited stocks whose catch limits must be significantly reduced to allow rebuilding under U.S. law (Table 3; Supp. Table 3). The median scalars for relating the best catch statistic to the OFL were 1.90, 2.16, and 1.56 for under, fully, and overexploited stocks, respectively (Table 3). Scalars more conservative than the median are provided in Table 3.

Evaluation of the refined ORCS approaches

The OFLs predicted by the ORCS approach and estimated by the data-rich assessment models were significantly correlated in all ten potential refined ORCS approaches (Table 4; Figure 5A-D). The ‘weighted 50th percentile scalar’ approach resulted in the underutilization (i.e., predicted OFL less than data-rich OFL) of 63% of underexploited stocks and overfishing (i.e., predicted OFL exceeds data-rich OFL) of 73% and 91% of fully and overexploited stocks, respectively (Figure 5E). The ‘unweighted 50th percentile scalar’ approach performed better, resulting in the underutilization of 54% of underexploited stocks and overfishing of 56% and 45% of fully and overexploited stocks, respectively (Figure 5F). The more conservative ‘unweighted 45th–10th percentile scalar’ approaches reduced the overfishing of overexploited stocks but increased the underexploitation of under and fully exploited stocks (Table 4; Figure 5G-H). The ‘unweighted 40th percentile scalars’ are the largest scalars to reduce the probability of overfishing below 50%, the U.S. legal maximum (DOC, 2016), in all three status categories (Table 4).

Discussion

Before being implemented, new stock assessment methods should be evaluated to validate their usefulness and transferability. Although the fully-flexible version of the original ORCS approach may produce useful status and catch limit estimates, it is challenging to validate because of its subjectivity. Therefore, we adopted a more specific, albeit less inclusive, definition of the ORCS approach for evaluation and refinement. Our results show that this interpretation of the ORCS approach is a poor predictor of stock status and should not be used for management decisions. The approach is heavily biased towards moderate classifications and classified all but three data-rich stocks as fully exploited. This result is not surprising given that all 20 stocks in the U.S. Southeast scored using the interpretation of the ORCS approach evaluated here were also classified as fully exploited, despite expert opinion that the stocks ranged from under to overexploited (SAFMC, 2012, 2013). The bias of the approach towards moderate classifications likely arises from: (1) an overrepresentation of moderate scores (in *TOA #4* notably and *#10, #12, #14* additionally) and (2) inappropriately wide threshold values for delineating status categories (1.75 and 2.25 might perform better). Furthermore, the non-linearity in the relative influence and marginal effects of the TOA attributes highlights the necessity of a weighting scheme. Although the original ORCS approach suggests that these adjustments can be made through expert judgement (Berkson et al. 2011), the refined ORCS approach presents an objective, transferable, and effective alternative.

The refined ORCS approach, which accounts for attribute importance, interactions, and non-linearity, is a better predictor of stock status than both the original ORCS approach and alternative catch-only methods. The refined approach correctly classified 73% (kappa=0.55, good) of the 37 stocks in the test dataset with a catch time series. In comparison, CMSY, which performed second best and also performed better than COM-SIR, SSCOM, and mPRM in Rosenberg et al.'s (2014) evaluation of these four methods, classified only 41% (kappa=0.15, poor) of these 37 data-rich test stocks correctly. The refined ORCS approach also outperformed SSP-2002 and SSP-2013, which have been shown to be poor and inherently pessimistic predictors of stock status (Branch et al., 2011; Carruthers et al., 2012), as well as mPRM, whose developers state that it should not be used to assess the status of individual stocks (Costello et al., 2012). Catch-based methods represent a class of widely used, but still controversial (Pauly et al., 2013), approaches to estimate status and the refined ORCS approach may be a useful alternative for estimating the status of data-poor stocks.

The refined ORCS approach also identifies catch statistics and scalars that estimate catch limits that prevent overfishing in accordance with U.S. legal mandates, suggesting that it can be used when data-moderate methods such as DB-SRA and DCAC are not possible or appropriate. Although the refined approach misclassifies many overexploited stocks, conservative catch scalars successfully buffer against classification uncertainty. The 40th percentile scalars produce the highest catches while reducing the probability of overfishing in all three status categories below 50%, the U.S. legal maximum (DOC, 2016); however, they also estimate OFLs more than five times the data-rich OFL for some stocks. More conservative catch scalars will further reduce the probability and magnitude of overfishing but will result in concomitant increases in the probability and magnitude of underfishing. Managers must therefore determine which catch scalars are most consistent with their risk policies. We provide a web tool for managers to implement the approach here: https://cfree.shinyapps.io/refined_orcs_approach/

The evaluation of the ORCS approach using data-rich stock assessments, while necessary because the ability of the approach to predict stock status cannot be evaluated through traditional simulation testing (Wiedenmann et al. 2013;

Carruthers et al. 2014), is somewhat problematic given the uncertainty in even the most sophisticated assessment models (Brooks and Deroba, 2015). For example, assessment model reference points (i.e., B/B_{MSY}, OFL) used to evaluate the performance of the ORCS approach and assessment model output (i.e., biomass and effort time series) used to score the ORCS Table of Attributes could both be incorrect. However, we took measures to eliminate the more uncertain assessments and we only used stock assessment output in the scoring of *TOA #12 Recent trend in effort* and *#13 Recent trend in abundance index*, which were both unimportant predictors of stock status. Furthermore, we trained the BCT model used in the refined approach to estimate categorical status (i.e., under, fully, or overexploited) rather than continuous status (i.e., B/B_{MSY}) because stock assessment models are generally more certain in status classifications than in precise B/B_{MSY} values. Finally, the ability of the refined ORCS approach to reproduce the conclusions of possibly incorrect but presumably better data-rich status determinations is still useful, especially given the recent success of data-rich assessment and management in rebuilding fisheries (Worm et al., 2009; Hilborn and Ovando, 2014).

The refinement of the ORCS approach through testing against data-rich stocks could also be problematic given the differences in the dynamics of data-poor and data-rich fisheries. Assessed (data-rich) fisheries generally target larger, slower growing, and higher trophic level species (Pinsky et al., 2011) and are higher volume, more valuable, and in better condition (Costello et al., 2012) than their unassessed (data-poor) counterparts. Consequently, it is possible that informative predictors of data-poor fisheries status could be uninformative or even trend opposite for data-rich fisheries. For example, in well-managed fisheries, decreasing catch could be the result of responsive management (Murawski et al., 2007) and increasing effort could indicate the sustainable development of a new or rebuilt fishery. Furthermore, the generally healthy status of data-rich stocks results in only a small sample of overexploited stocks (30 of 193 stocks, 15.5%) available for model training and testing. Thus, the model may have performed poorly at classifying overexploited stocks because of the limited number of overexploited stocks in the dataset.

The dynamics of the most important predictors of stock status in the BCT model are consistent with other studies and are likely conserved across data-poor and -rich fisheries. For example, the importance of ex-vessel price is not surprising given that fishery development is frequently driven by profits (Sethi et al., 2010). The importance of assessed stock status also makes intuitive sense (i.e., a stock in a generally well- or poorly-managed fishery is also likely to be well- or poorly-managed, respectively) and is similar to the region effect, which has been shown to be useful in discriminating stock status (Ricard et al., 2012; Thorson et al., 2012). The significant increase in overexploitation risk resulting from aggregation behavior is supported by emerging evidence that schooling, fast-lived fish may actually be more vulnerable to collapse than solitary, long-lived taxa due to high harvest rates lagging behind rapid changes in environment and productivity (Pinsky et al., 2011). The decrease in overexploitation risk with increasing occurrence in the catch opposes the predictions of the original Table of Attributes and suggests that rarity in the catch is indicative of a depleted stock rather than a lightly exploited one. Finally, recent trends in catch and effort, the attributes most likely to be confounding between data-poor and -rich fisheries, exert little predictive influence, increasing the likelihood that the refined ORCS approach is as predictive for data-poor stocks as it is for data-rich ones.

The refined ORCS approach also provides important guidance on the choice of historical catch statistics and scalars. Longer timeframe, higher percentile catch statistics perform best for underexploited stocks with light exploitation histories. Moderate timeframe, more central catch statistics perform best for fully exploited stocks where recent management has been effective in sustaining abundance and yield. Longer timeframe, lower percentile catch statistics perform best for overexploited stocks where recent catches have resulted in depletion. To consistently achieve a relatively low risk of overfishing, the catch scalars used to scale the historical catch statistic to the overfishing limit will have to be conservative to buffer against substantial classification uncertainty. This conclusion is especially true for data-poor stocks with uncertainty in their catch time series, such as the rarely caught snapper-grouper species in the U.S. Southeast which suffer from misidentification problems (SAFMC, 2013; Berkson et al., 2011). Although conservative scalars will effectively protect overexploited stocks, they will also result in forgone yield from under and fully exploited stocks.

The refined ORCS approach represents one step towards Berkson et al.'s (2011) recommendations for testing and improving the original ORCS approach but could under additional refinement and evaluation. The predictive performance of the approach could be improved by identifying new predictive attributes. For example, life history characteristics such as age at maturity, maximum age, maximum length, and trophic level and fishery characteristics such as time since development and exploitation history have all been shown to be useful in discriminating stock status (Sethi et al., 2010; Pinsky et al., 2011; Costello et al., 2012; Thorson et al., 2012; Neubauer et al., 2013) and could be incorporated into the refined TOA and BCT model. Furthermore, the performance of the status-specific historical catch statistics and scalars used in the refined approach should be tested through management strategy evaluation, such as in Wiedenmann et. al (2013), to determine whether they actually promote sustainable fishing levels. The development of simple data-limited decision support tools has been a central focus of recent fisheries management (Berkson and Thorson, 2015) and the refined ORCS approach provides an additional tool for managers faced with the legal mandates and data limitations of contemporary fisheries management.

Acknowledgements

This manuscript was greatly improved by the thoughtful reviews of Steve Cadrin and André Punt. We are also grateful to Emma Fuller, Doug Zemeckis, Kyle Shertzer, and Marcus Drymon for answering questions about U.S. fisheries, Jim Berkson for providing feedback on preliminary methods and results, Sean Anderson for help with the *datalimited* R package, and Michael Roswell and the members of the Jensen and Pinsky labs for their helpful feedback on the manuscript. This work was supported by a NMFS-Sea Grant Population Dynamics Fellowship to CMF and NJ Sea Grant Omnibus Grant to JW and OPJ. Thus, this publication (NJSG-17-913) is also the result of research sponsored by the New Jersey Sea Grant Consortium (NJSGC) with funds from the National Oceanic and Atmospheric Administration (NOAA) Office of Sea Grant, U.S. Department of Commerce, under NOAA grant numbers NA13OAR4170114 and NA14OAR4170085 and the NJSGC. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the NJSGC or the U.S. Department of Commerce.

References

Anderson, S.C., Afflerbach, J., Cooper, A.B., Dickey-Collas, M., Jensen, O.P., Kleisner, K.M., Longo, C., Osio, G.C., Ovando, D., Minte-Vera, C., Minto, C., Mosqueira, I., Rosenberg, A.A., Selig, E.R., Thorson, J.T., Walsh, J.C., 2016. datalimited: Stock assessment methods for data-limited fisheries. R package version 0.0.2. Available at: https://github.com/datalimited/datalimited

Anderson, S.C, Cooper, A.B, Jensen, O.P., Minto, C., Thorson, J.T., Walsh, J.C., Afflerbach, J, Dickey-Collas, M. Kleisner, K.M., Longo, C., Osio, G.C., Ovando, D., Mosqueira, I., Rosenberg, A.A., Selig, E.R., 2017. Improving estimates of population status and trend with superensemble models. Fish Fish. DOI: 10.1111/faf.12200.

Berkson, J., Barbieri, L., Cadrin, S., Cass-Calay, S., Crone, P., Dorn, M., Friess, C., Kobayashi, D., Miller, T.J., Patrick, W.S., Pautzke, S., Ralston, S., Trianni, M., 2011. Calculating Acceptable Biological Catch for stocks that have reliable catch data only (Only Reliable Catch Stocks – ORCS). NOAA Technical Memorandum NMFS-SEFSC-616. NMFS-SEFSC, Miami, FL.

Berkson, J., Thorson, J. T., 2015. The determination of data-poor catch limits in the United States: is there a better way? ICES J. Mar. Sci. 72(1), 237-242.

Branch, T.A., Jensen, O.P., Ricard, D., Yimin, Y., Hilborn, R., 2011. Contrasting global trends in marine fishery status obtained from catches and from stock assessments. Conserv. Biol. 25(4), 777–786.

Brooks, E.N., Deroba, J.J., 2015. When “data” are not data: the pitfalls of post hoc analyses that use stock assessment model output. Can. J. Fish. Aquat. Sci. 72(4), 634-641.

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

Carruthers, T.R., Walters, C.J., McAllister, M.K., 2012. Evaluating methods that classify fisheries stock status using only fisheries catch data. Fish. Res. 119-120, 66-79.

Cohen, J., 1968. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psychol. Bull. 70(4), 213-220.

Cope, J., Dick, E.J., MacCall, A., Monk, M., Soper, B., Wetzel, C., 2013. Data-moderate stock assessments for brown, China, copper, sharpchin, stripetail, and yellowtail rockfishes and English and rex soles in 2013. PFMC June 2013 Meeting Agenda Item F.5.a, Attachment 1.

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

Costello, C., Ovando, D., Hilborn, R., Gains, S.D., Deschenes, O., Lester, S.E., 2012. Status and solutions for the world’s unassessed fisheries. Science 338, 517-520.

Dick, E.J., MacCall, A.D., 2011. Depletion-based stock reduction analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. Fish. Res. 110, 331-341.

DOC, 2007. Magnuson-Stevens Fishery Conservation and Management Act (MSFCMA). U.S.A. Public Law 94-265. Department of Commerce (DOC), Washington, D.C.

DOC, 2016. Magnuson-Stevens Act Provisions; National Standard Guidelines; Final Rule. Federal Register 81(201): 71858-71904. Department of Commerce (DOC), Washington, D.C.

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted classification trees. J. Anim. Ecol. 77(4), 802-813.

Fleiss, J.L., 1973. Statistical methods for rates and proportions. John Wiley & Sons, New York, NY.

Froese, R., Kesner-Reyes, K., 2002. Impact of Fishing on the Abundance of Marine Species. ICES CM 2002/L:12.

Froese, R., Pauly, D., 2016. FishBase. Available at: www.fishbase.org

Gulland, J.A., 1971. The Fish Resources of the Ocean. Fishing News Books: West Byfleet, UK.

Hilborn, R., Ovando, D., 2014. Reflections on the success of traditional fisheries management. ICES J. Mar. Sci. 71(5), 1040-1046.

ICES, 2014. Report of the Workshop on the Development of Quantitative Assessment Methodologies based on LIFE-history traits, exploitation characteristics, and other relevant parameters for data-limited stocks (WKLIFE IV). ICES CM 2014/ACOM:54, Lisbon, Portugal.

ICES, 2015. Report of the Fifth Workshop on the Development of Quantitative Assessment Methodologies based on Life-history Traits, Exploitation Characteristics and other Relevant Parameters for Data-limited Stocks (WKLIFE V). ICES CM 2015/ACOM:56, Lisbon, Portugal.

ICES, 2017. Report of the ICES Workshop on the Development of Quantitative Assessment Methodologies based on Life-history traits, exploitation characteristics, and other relevant parameters for stocks in categories 3–6 (WKLIFEVI). ICES CM 2016/ACOM:59, Lisbon, Portugal.

Kleisner, K., Zeller, D., Froese, R., Pauly, D., 2013. Using global catch data for inferences on the world's marine fisheries. Fish Fish. 14(3), 293-311.

Kuhn, M., 2016. caret: Classification and Regression Training. R package version 6.0-71.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33(1), 159-174.

MacCall, A.D., 2009. Depletion-corrected average catch: A simple formula for estimating sustainable yields in data-poor situations. ICES J. Mar. Sci. 66, 2267-2271.

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

MPI, 2016. The Status of New Zealand's Fisheries 2015. Ministry for Primary Industries (MPI), Wellington, New Zealand.

Murawski, S., Methot, R., Tromble, G., 2007. Biodiversity loss in the ocean: how bad is it? Science 316(5829), 1281-1284.

Neubauer, P., Jensen, O.P., Hutchings, J.A., Baum, J.K., 2013. Resilience and recovery of overexploited marine populations. Science 340(6130), 347-349.

Newman, D., Berkson, J., Suatoni, L., 2015. Current methods for setting catch limits for data-limited fish stocks in the United States. Fish. Res. 164, 86-93.

Pauly, D., Hilborn, R., Branch, T.A., 2013. Fisheries: does catch reflect abundance? Nature 494(7437), 303-306.

Pauly D., Zeller D., 2015. Sea Around Us Concepts, Design and Data. Available at: www.seaaroundus.org

Pinsky, M.L., Jensen, O.P., Ricard, D., Palumbi, S.R., 2011. Unexpected patterns of fisheries collapse in the world's oceans. Proc. Natl. Acad. Sci. 108(20), 8317-8322.

R Core Team, 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria. Available at: http://www.R-project.org/

Restrepo, V.R., Thompson, G.G., Mace, P.M., Gabriel, W.L., Low, L.L., MacCall, A.D., Methot, R.D., Powers, J.E., Taylor, B.L., Wade, P.R., Witzig, J.F., 1998. Technical guidance on the use of precautionary approaches to implementing National Standard 1 of the Magnuson–Stevens Fishery Conservation and Management Act. NOAA Technical Memorandum, NMFS-F/ SPO-31.

Ricard, D., Minto, C., Jensen, O.P., Baum, J.K., 2012. Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish Fish. 13(4), 380-398.

Ridgeway, G., 2016. gbm: Generalized Boosted Classification Models. R package version 2.1.1.

Rosenberg, A.A., Fogarty, M.J., Cooper, A. B., Dickey-Collas, M., Fulton, E.A., Gutiérrez, N.L., Hyde, K.J.W., Kleisner, K.M., Longo, C., Minte-Vera, C.V., Minto, C., Mosqueira, I., Osio, G.C., Ovando, D., Selig, E.R., Thorson, J.T., Ye, Y., 2014. Developing new approaches to global stock status assessment and fishery production potential of the seas. FAO Fisheries and Aquaculture Circular, Rome, Italy.

SAFMC, 2012. SSC ORCS Workshop Report. South Atlantic Fishery Management Council-Scientific & Statistical Committee (SAFMC-SSC), North Charleston, SC.

SAFMC, 2013. SSC ORCS Workshop II Report. South Atlantic Fishery Management Council-Scientific & Statistical Committee (SAFMC-SSC), North Charleston, SC.

Sethi, S.A., Branch, T.A., Watson, R., 2010. Global fishery development patterns are driven by profit but not trophic level. Proc. Natl. Acad. Sci. 107(27), 12163-12167.

Thorson, J.T., Branch, T.A., Jensen, O.P., 2012. Using model-based inference to evaluate global fisheries status from landings, location, and life history data. Can. J. Fish. Aquat. Sci. 69(4), 645-655.

Thorson, J.T., Minto, C., Minte-Vera, C.V., Kleisner, K.M., Longo, C., 2013. A new role for effort dynamics in the theory of harvested populations and data-poor stock assessment. Can. J. Fish. Aquat. Sci. 70(12), 1829-1844.

Vasconcellos, M., Cochrane, K., 2005. Overview of world status of data-limited fisheries: inferences from landings statistics. In: Kruse, G.H., Gallucci, V.F., Hay, D.E., Perry, R.I., Peterman, R.M., Shirley, T.C., Spencer, P.D., Wilson, B., Woodby, D. (eds.): Fisheries Assessment and Management in Data-Limited Situations. Alaska Sea Grant College Program, University of Alaska Fairbanks, Fairbanks, AK, pp. 1–20.

Walters, C.J., Martell, S.J., 2004. Fisheries Ecology and Management. Princeton University Press, Princeton, NJ.

Wetzel, C.R., Punt, A.E., 2015. Evaluating the performance of data-moderate and catch-only assessment methods for US west coast groundfish. Fish. Res. 171, 170-187.

Wiedenmann, J., Wilberg, M.J., Miller, T.J., 2013. An evaluation of harvest control rules for data-poor fisheries. N. Am. J. Fish. Manag. 33(4), 845-860.

Worm, B., Hilborn, R., Baum, J.K., Branch, T.A., Collie, J.S., Costello, C., Fogarty, M.J., Fulton, E.A., Hutchings, J.A., Jennings, S., Jensen, O.P., 2009. Rebuilding global fisheries. Science 325(5940), 578-585.

Tables & Figures

Table 1. ORCS Table of Attributes (TOA; adapted from SAFMC, 2012).

#	Attribute	Stock status[1]		
		Underexploited (1)	**Fully exploited (2)**	**Overexploited (3)**
1	Status of assessed stocks in fishery[2]	<10% overfished	10-25% overfished	>25% overfished
2	Refuge availability *(not used in refined approach)*	<50% of habitat accessible	50-75% of habitat accessible	>75% of habitat accessible
3	Behavior affecting capture	-------------	No aggregation behavior	Exhibits aggregation behavior
4	Morphology affecting capture *(not used in refined approach)*	Low susceptibility	Average susceptibility	High susceptibility
5	Discard rate[2]	Discards <10% of catch	Discards 10-25% of catch	Discards >25% of catch
6	Targeting intensity	Not targeted	Occasionally targeted	Actively targeted
7	M compared to dominant species[3]	Higher mortality rate	Equivalent mortality rates	Lower mortality rate
8	Occurrence in catch	Sporadic *(in <10% of efforts)*	Common *(in 10-25% of efforts)*	Frequent *(in >25% of efforts)*
9	Value (US$/lb, 5-year mean) *(continuous value in refined approach)*	<$1 / lb	$1-$2.25 / lb	>$2.25 / lb
10	Recent trend in catch	Increasing last 5 years *(score=1.5 in original approach)*	Stable last 5 years *(score=1.5 in original approach)*	Decreasing last 5 years
11	Habitat loss[4]	No time in threatened habitats	Part time in threatened habitats *(full time in partially threatened habitats)*	Full time in threatened habitats
12	Recent trend in effort	Decreasing last 5 years	Stable last 5 years	Increasing last 5 years
13	Recent trend in abundance index	Increasing last 5 years	Stable last 5 years	Decreasing last 5 years

14	Proportion of population protected	Most of resource is protected *(size limits AND time/space closures)*	Some of resource is protected *(size limits OR time/space closures)*	None of resource is protected *(no size limits or time/space closures)*

[1] In the original ORCS approach, stock status is estimated as the mean of the TOA scores (<1.5=underexploited; 1.5-2.5=fully exploited; >2.5=overexploited).
[2] Replaced vague score descriptions in the original table with straightforward percentage thresholds. Note: the definition of overfishing varies by management agency (Supplementary Appendix A.2).
[3] Removed ambiguity of score descriptions in the original table and specified that M's must differ by >20% to be considered different. See Supplementary Appendix A.2 for definition of dominant species.
[4] Rephrased original attribute to be conceptually simpler and easier to score. See Supplementary Appendix A.2 for list of threatened habitats.

Table 2. Status classification performance of catch-only assessment methods applied to the 37 data-rich stocks with catch time series in the test dataset (COM-SIR and SSCOM converged for only 33 test stocks).

		37 stocks[1]		33 stocks[2]		
	Method	**Kappa**	**Accuracy**	**Kappa**	**Accuracy**	**Bias**
1	Refined ORCS approach	0.549	0.730	0.558	0.727	slightly optimistic
2	cMSY	0.148	0.405	0.205	0.455	pessimistic
3	SSP-2002	0.120	0.378	0.171	0.424	pessimistic
4	COM-SIR	-----	-----	0.114	0.424	optimistic
5	SSP-2013	0.041	0.270	0.052	0.303	toward extremes
6	mPRM	-0.015	0.405	-0.005	0.394	central
7	Original ORCS approach	-0.035	0.459	-0.040	0.424	central
8	SSCOM	-----	-----	-0.120	0.333	slightly central

[1] 13 underexploited, 18 fully exploited, and 6 overexploited stocks; See Supp. Table 1 for status classifications.
[2] 12 underexploited, 15 fully exploited, and 6 overexploited stocks; See Supp. Table 2 for status classifications.

Table 3. Best status-specific historical catch statistics and potential status-specific catch scalars for relating the best catch statistic to the overfishing limit (OFL).

				OFL scalars**								
Stock status	Best historical catch statistic*	r^2	n	50^{th}	45^{th}	40^{th}	35^{th}	30^{th}	25^{th}	20^{th}	15^{th}	10^{th}
Underexploited	90^{th} percentile, whole time series	0.91	45	1.90	1.78	1.62	1.53	1.41	1.34	1.29	1.11	0.88
Fully exploited	25^{th} percentile, previous 10 years	0.91	49	2.16	1.84	1.77	1.57	1.41	1.22	1.15	1.02	0.85
Overexploited	10^{th} percentile, whole time series	0.89	11	1.56	1.53	1.49	1.00	0.52	0.51	0.50	0.45	0.41

* See Supp. Table 3 for the ranking of potential catch statistics.
** The 50^{th} percentile scalars should promote a 50% probability of overfishing if stock status is correctly identified. The other, more conservative scalars may be useful for buffering against classification uncertainty.

Table 4. The percentage of stocks (n=97) whose predicted OFLs exceeded the data-rich OFL estimates under potential catch scalars with unweighted status predictions (maximum observed OFL_{ORCS}:$OFL_{data\text{-}rich}$ ratio shown in parentheses).

		OFL scalars								
Stock status	**n**	**50th**	**45th**	**40th**	**35th**	**30th**	**25th**	**20th**	**15th**	**10th**
Underexploited	41	46.3% (2.6)	41.5% (2.4)	39.0% (2.2)	34.1% (2.1)	26.8% (1.9)	19.5% (1.8)	17.1% (1.7)	9.8% (1.5)	4.9% (1.2)
Fully exploited	45	55.6% (4.8)	51.1% (4.1)	44.4% (4.0)	37.8% (3.5)	35.6% (3.2)	31.1% (2.7)	26.7% (2.6)	22.2% (2.3)	17.8% (1.9)
Overexploited	11	45.5% (8.7)	45.5% (7.4)	45.5% (7.2)	45.5% (6.3)	45.5% (5.7)	36.4% (4.9)	36.4% (4.6)	36.4% (4.1)	36.4% (3.4)
	r^2:	0.90	0.90	0.90	0.90	0.90	0.91	0.91	0.91	0.91
	RMSE, mt:	129,938	93,462	86,589	71,454	66,481	69,664	73,673	85,452	104,880

* See **Table 3** for the status-specific catch scalars and best historical catch statistics.

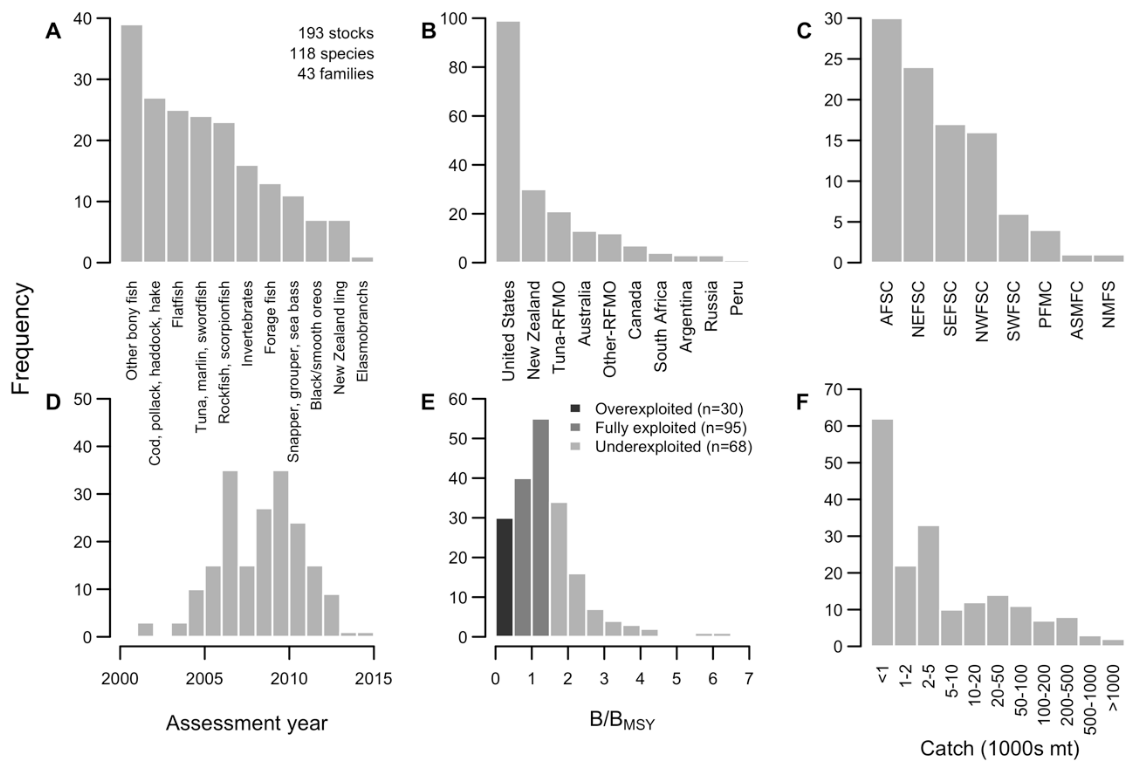


Figure 1. Demographics of the 193 data-rich stocks scored using the ORCS approach by: (A) taxonomic group; (B) managing country or multinational body; (C) U.S. assessment agency (U.S. stocks only; n=99, 51.3% of scored stocks); (D) assessment year; (E) stock status (B/B$_{MSY}$); and (F) fishery size (average annual catch over the most recent 5 years).

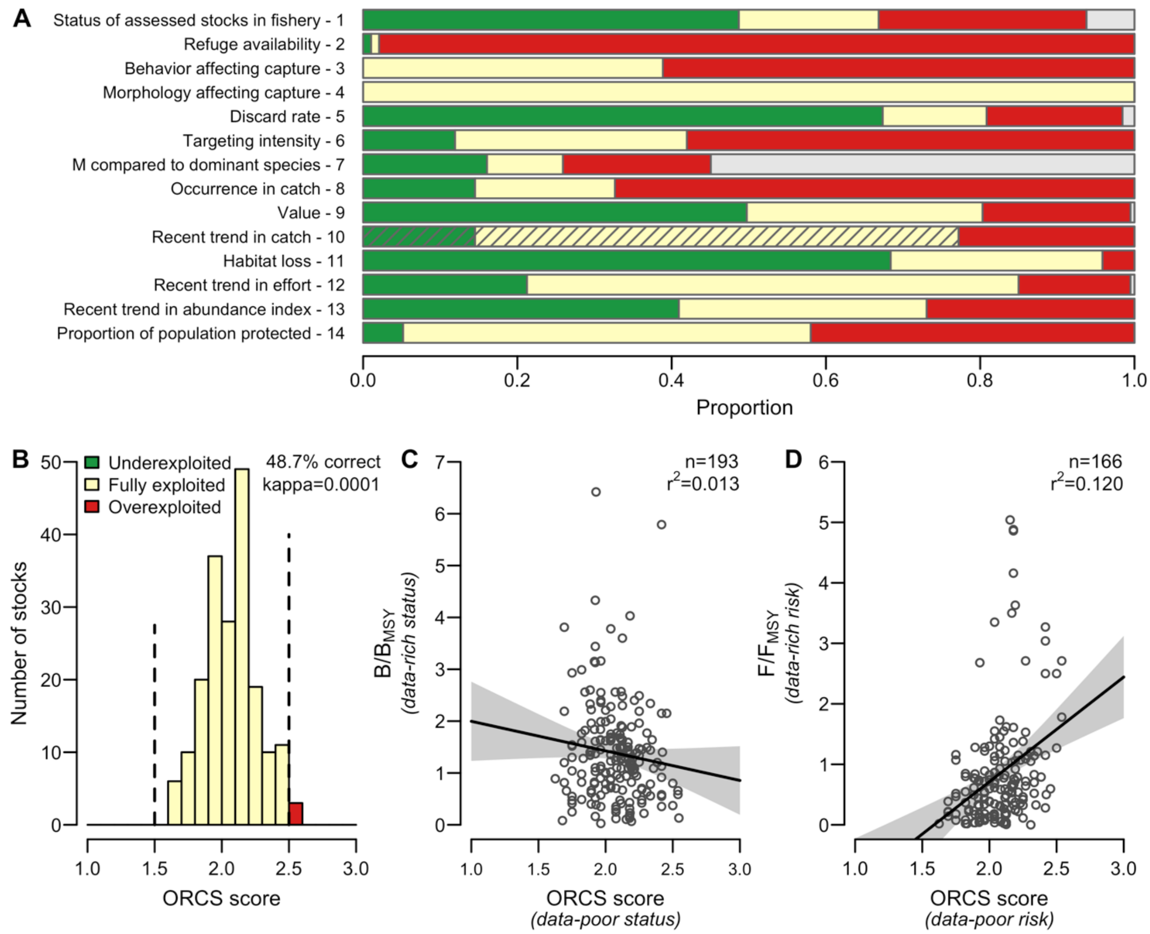


Figure 2. The distribution of (A) attribute scores and (B) overall scores for the 193 data-rich stocks scored using the original ORCS approach and (C) comparison of statuses and (D) risks predicted by the ORCS approach and estimated by data-rich assessment models. In (A), bars show the proportion of scores represented in each TOA attribute. For *TOA #10*, scores of 1 and 2 (hatched) are reassigned scores of 1.5 and only count towards the overall score if effort is stable (*TOA #12*, score=2). In some cases, attributes could not be scored due to a lack of data or applicability and were given an 'NA' value (grey shading). In (B), vertical lines indicate the threshold values (1.5 & 2.5) that separate under, fully, and overexploited stocks. In (C) and (D), the black lines indicate linear regressions fit to the data and the gray shading indicates the confidence intervals for the regressions.

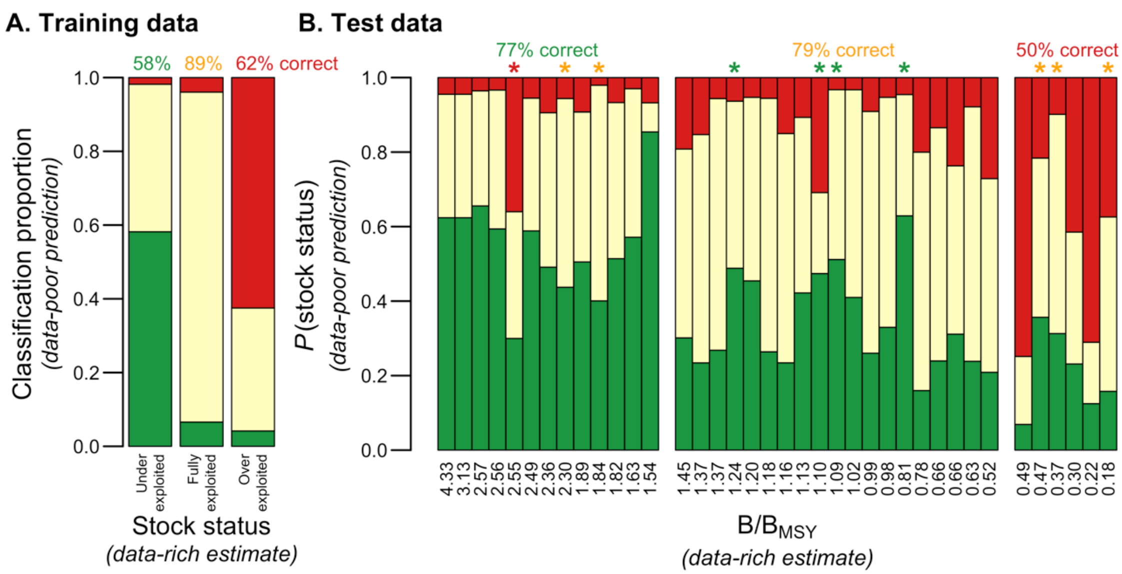


Figure 3. The performance of the boosted classification tree (BCT) model on the (A) training (n=155 stocks, 80% of data) and (B) test datasets (n=38 stocks, 20% of data). In (A), bars show the proportion of status predictions for each status category. Percentages indicate the proportion of correct classifications in each category (overall accuracy=74% and Cohen's kappa=0.56). In (B), bars show the probability that a stock is in each status category, where the highest probability category is the BCT model's prediction of stock status; stocks are grouped and sorted by B/B_{MSY} from the data-rich assessment model. Percentages indicate the proportion of correct classifications in each category; stars mark incorrectly classified stocks with colors indicating the direction of the misclassification (overall accuracy=74% and Cohen's kappa=0.56).

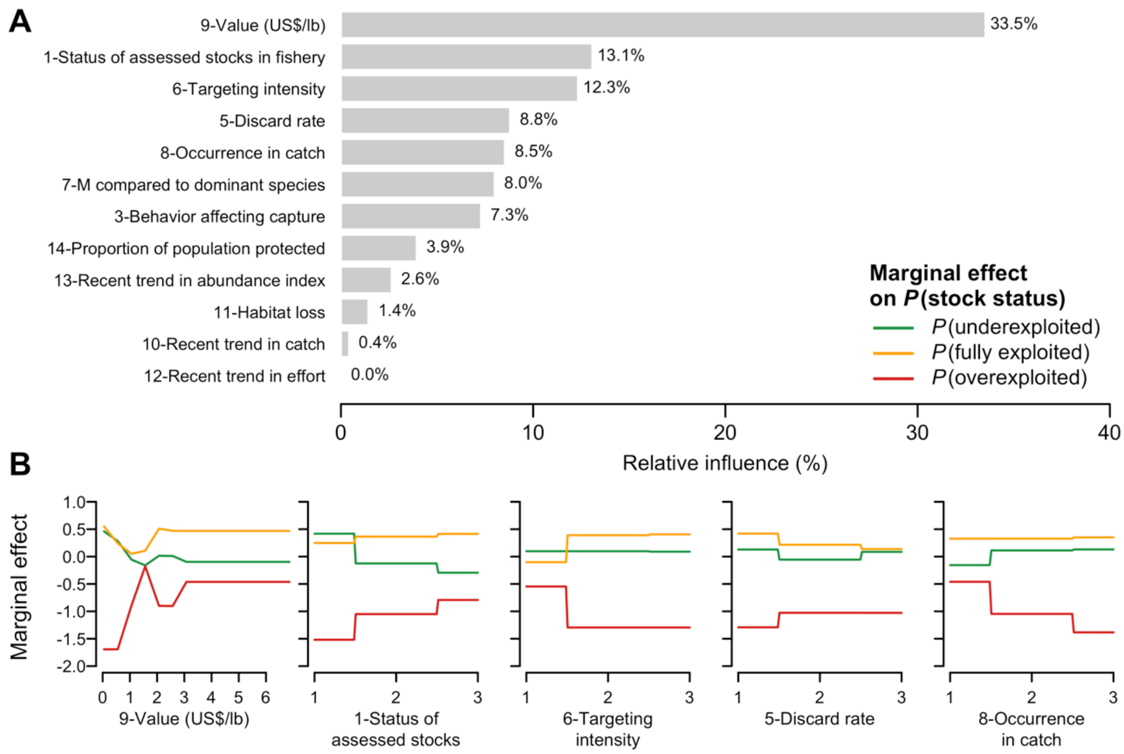


Figure 4. The (A) relative influence and (B) marginal effects of the five most important TOA attributes in the boosted classification tree model. In (B), lines represent the effect of each attribute on the probability that a stock is in each status category when the other attributes are held constant. *TOA #2 Refuge availability* and *#4 Morphology affecting capture* were omitted from the model due to lack of predictive power.

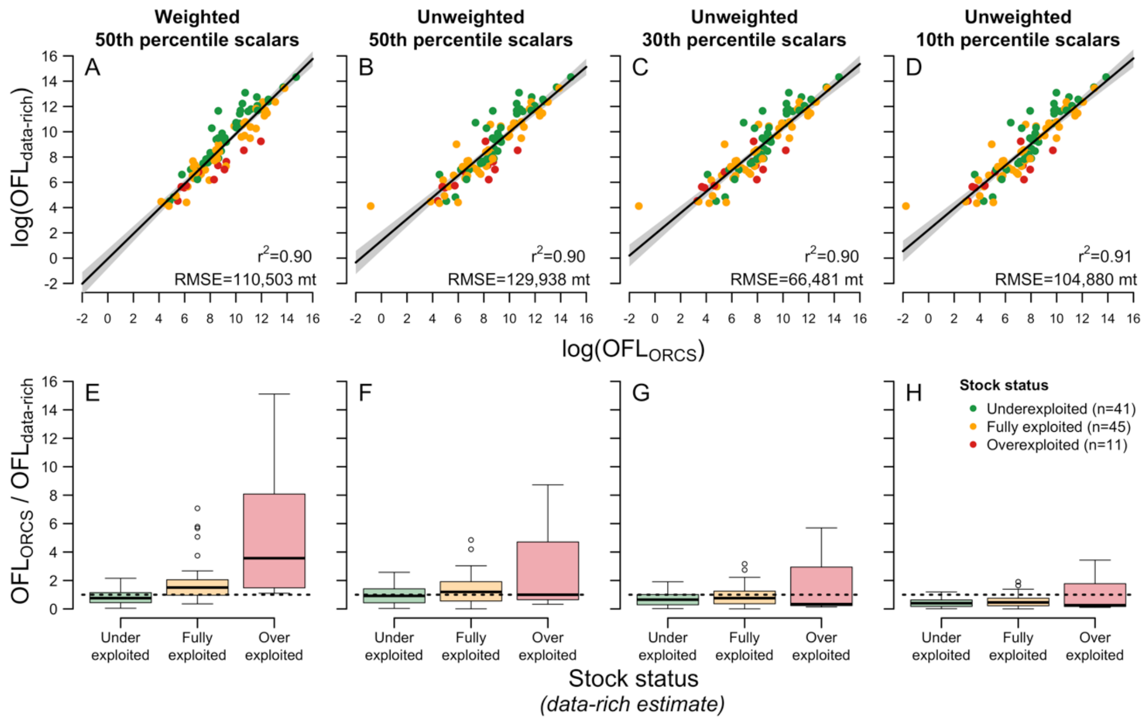


Figure 5. The (A-D) correlation between and (E-H) ratio of the overfishing limits (OFLs) predicted by the ORCS approach and estimated by data-rich assessment models for 97 stocks in four potential refined ORCS approaches. In (A-D), black lines indicate linear regressions fit to the untransformed data and the gray shading indicates the confidence interval for the regression. In (E-H), ratios were also calculated using the untransformed data. The dotted horizontal lines indicate perfect agreement between the ORCS predictions and the data-rich model estimates and boxplots indicate the median (heavy black line), interquartile range (IQR; box), 1.5 times the IQR (whiskers), and outliers.

Supplemental Tables & Figures

Supp. Table 1. Stock status predictions for the 37 data-rich stocks with catch time series in the test dataset (grey shading indicates correct predictions).

Method	Status	n(observed)	n(predicted) Underexploited	Fully exploited	Overexploited	Success rate (%)	Proportional error (%)
Refined ORCS approach	underexploited	13	10	2	1	76.9%	7.7%
	fully exploited	18	4	14	0	77.8%	5.6%
	overexploited	6	0	3	3	50.0%	-33.3%
	Total or mean:	37	14	19	4	(27) 73.0%	15.5% (absolute)
CMSY	underexploited	13	0	5	8	0.0%	-100.0%
	fully exploited	18	0	10	8	55.6%	-11.1%
	overexploited	6	0	1	5	83.3%	250.0%
	Total or mean:	37	0	16	21	(15) 40.5%	120.4% (absolute)
SSP-2002	underexploited	13	0	5	8	0.0%	-100.0%
	fully exploited	18	0	9	9	50.0%	-16.7%
	overexploited	6	0	1	5	83.3%	266.7%
	Total or mean:	37	0	15	22	(14) 37.8%	127.8% (absolute)
SSP-2013	underexploited	13	5	0	8	38.5%	15.4%
	fully exploited	18	9	0	9	0.0%	-100.0%
	overexploited	6	1	0	5	83.3%	266.7%
	Total or mean:	37	15	0	22	(10) 27.0%	127.4% (absolute)
mPRM	underexploited	13	1	11	1	7.7%	-76.9%
	fully exploited	18	2	12	4	66.7%	50.0%
	overexploited	6	0	4	2	33.3%	16.7%
	Total or mean:	37	3	27	7	(15) 40.5%	47.9% (absolute)
Original ORCS approach	underexploited	13	0	13	0	0.0%	-100.0%
	fully exploited	18	0	17	1	94.4%	100.0%
	overexploited	6	0	6	0	0.0%	-83.3%
	Total or mean:	37	0	36	1	(17) 45.9%	94.4% (absolute)

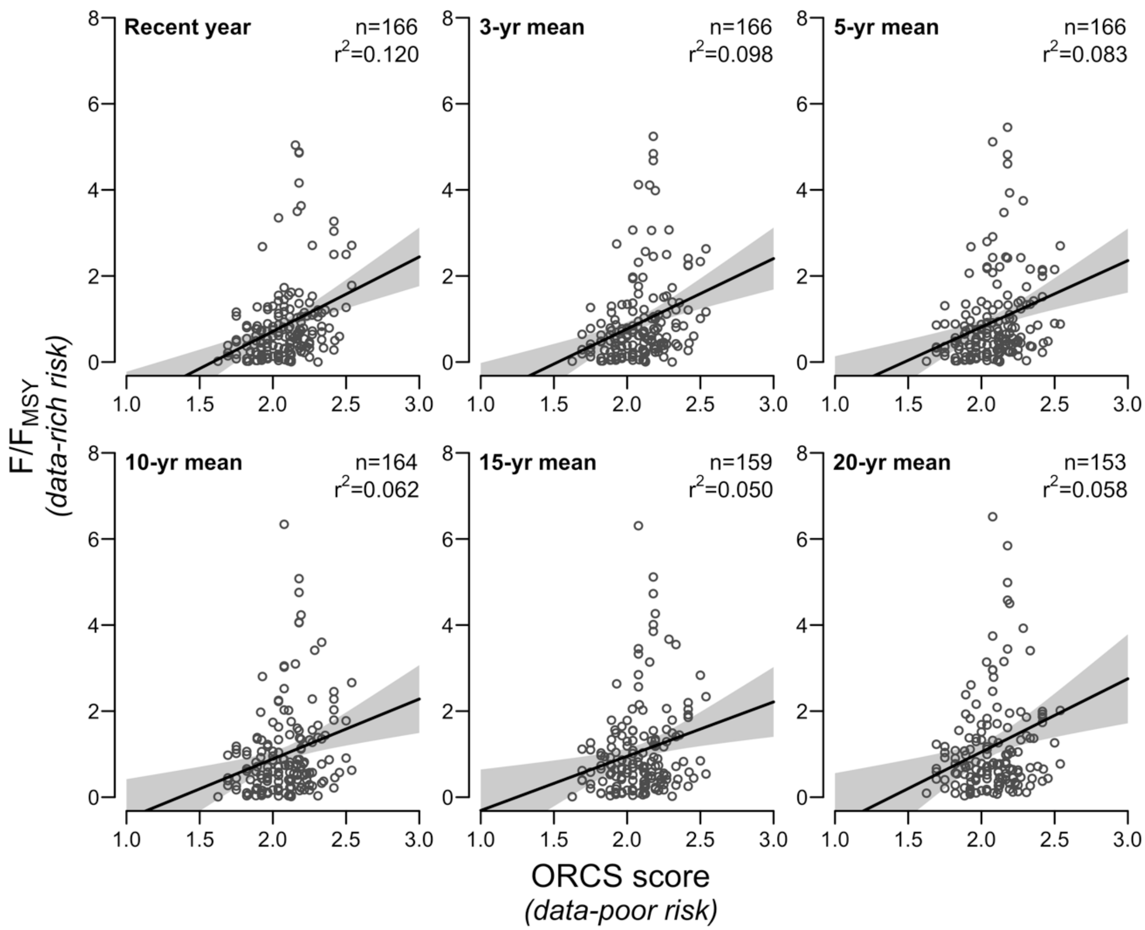
Supp. Table 2. Stock status predictions for the 33 data-rich stocks with catch time series in the test dataset converging for both COM-SIR and SSCOM (grey shading indicates correct predictions).

			n(predicted)			Success	Proportional
Method	Status	n(observed)	Underexploited	Fully exploited	Overexploited	rate (%)	error (%)
Refined ORCS approach	underexploited	12	10	1	1	83.3%	16.7%
	fully exploited	15	4	11	0	73.3%	0.0%
	overexploited	6	0	3	3	50.0%	-33.3%
	Total or mean:	33	14	15	4	(24) 72.7%	16.7% (absolute)
cMSY	underexploited	12	0	5	7	0.0%	-100.0%
	fully exploited	15	0	10	5	66.7%	6.7%
	overexploited	6	0	1	5	83.3%	183.3%
	Total or mean:	33	0	16	17	(15) 45.5%	96.7% (absolute)
SSP-2002	underexploited	12	0	5	7	0.0%	-100.0%
	fully exploited	15	0	9	6	60.0%	0.0%
	overexploited	6	0	1	5	83.3%	200.0%
	Total or mean:	33	0	15	18	(14) 42.4%	100.0% (absolute)
COM-SIR	underexploited	12	8	3	1	66.7%	41.7%
	fully exploited	15	7	4	4	26.7%	-40.0%
	overexploited	6	2	2	2	33.3%	16.7%
	Total or mean:	33	17	9	7	(14) 42.4%	32.8% (absolute)
SSP-2013	underexploited	12	5	0	7	41.7%	25.0%
	fully exploited	15	9	0	6	0.0%	-100.0%
	overexploited	6	1	0	5	83.3%	200.0%
	Total or mean:	33	15	0	18	(10) 30.3%	108.3% (absolute)
mPRM	underexploited	12	1	10	1	8.3%	-75.0%
	fully exploited	15	2	10	3	66.7%	60.0%
	overexploited	6	0	4	2	33.3%	0.0%
	Total or mean:	33	3	24	6	(13) 39.4%	45.0% (absolute)
Original ORCS approach	underexploited	12	0	12	0	0.0%	-100.0%
	fully exploited	15	0	14	1	93.3%	113.3%
	overexploited	6	0	6	0	0.0%	-83.3%
	Total or mean:	33	0	32	1	(14) 42.4%	98.9% (absolute)
SSCOM	underexploited	12	2	9	1	16.7%	-25.0%
	fully exploited	15	4	9	2	60.0%	40.0%
	overexploited	6	3	3	0	0.0%	-50.0%
	Total or mean:	33	9	21	3	(11) 33.3%	38.3% (absolute)

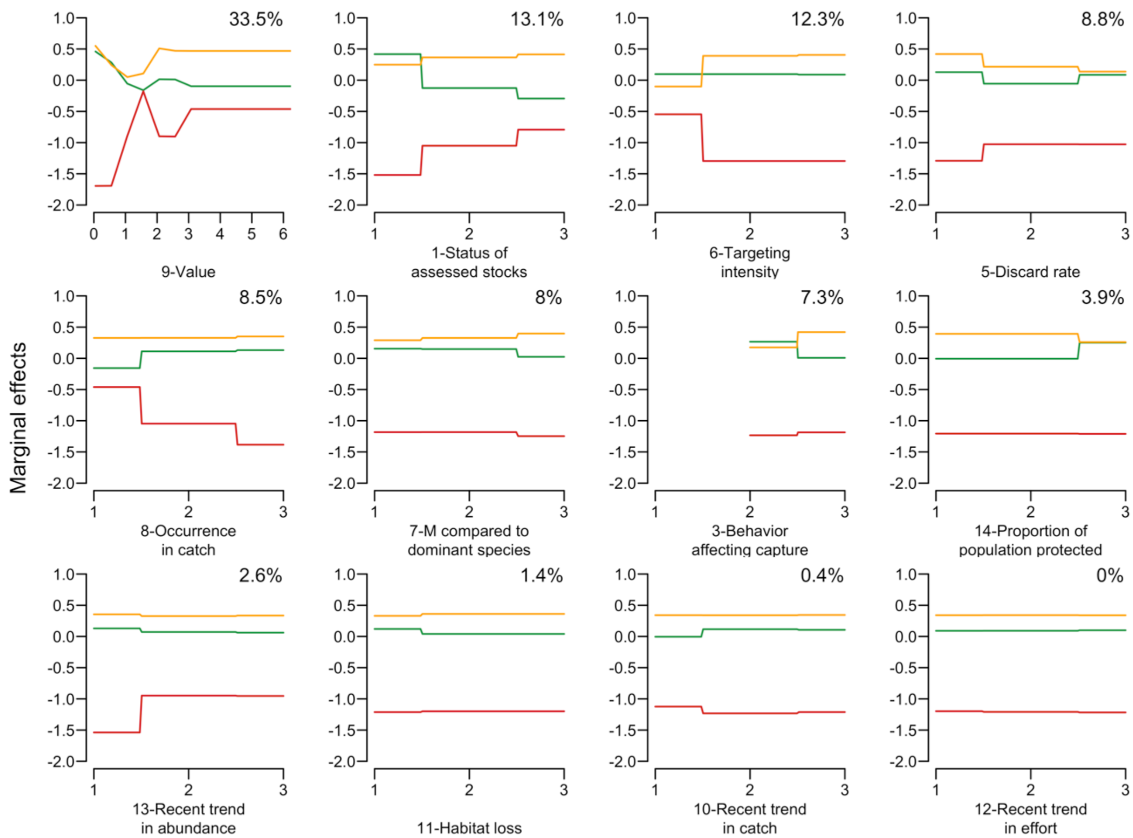
Supp. Table 3. Akaike's Information Criteria ranking of 24 potential historical catch statistics.

	All stocks				Underexploited stocks				Fully exploited stocks				Overexploited stocks			
#	Catch statistic	ΔAIC	Weight	r^2	Catch statistic	ΔAIC	Weight	r^2	Catch statistic	ΔAIC	Weight	r^2	Catch statistic	ΔAIC	Weight	r^2
1	25th perc. (10yr)	0.0	0.25	0.89	90th perc.	0.0	0.35	0.91	25th perc. (10yr)	0.0	0.39	0.91	Mean (20yr)	0.0	1.00	0.98
2	75th perc. (20yr)	0.4	0.20	0.89	75th perc.	1.2	0.19	0.91	10th perc. (10yr)	0.2	0.35	0.91	10th perc.	20.2	0.00	0.89
3	50th perc. (20yr)	0.4	0.20	0.89	50th perc. (20yr)	2.2	0.12	0.91	50th perc. (10yr)	2.5	0.12	0.91	90th perc. (20yr)	21.3	0.00	0.88
4	10th perc. (10yr)	0.6	0.18	0.89	IQR mean (20yr)	3.5	0.06	0.90	90th perc.	3.7	0.06	0.91	25th perc.	22.0	0.00	0.87
5	IQR mean (20yr)	3.2	0.05	0.89	75th perc. (20yr)	4.2	0.04	0.90	75th perc. (20yr)	4.7	0.04	0.90	75th perc.	22.1	0.00	0.87
6	Win. mean (20yr)	4.4	0.03	0.89	Win. mean (20yr)	4.4	0.04	0.90	IQR mean (10yr)	6.4	0.02	0.90	Win. mean	22.1	0.00	0.87
7	Mean (10yr)	4.5	0.03	0.89	25th perc. (10yr)	4.7	0.03	0.90	Mean (10yr)	7.5	0.01	0.90	IQR mean	22.3	0.00	0.87
8	50th perc. (10yr)	4.5	0.03	0.89	10th perc. (10yr)	5.4	0.02	0.90	Win. mean (10yr)	7.7	0.01	0.90	50th perc.	22.5	0.00	0.86
9	IQR mean (10yr)	5.7	0.01	0.89	Mean (20yr)	5.5	0.02	0.90	IQR mean (20yr)	9.3	0.00	0.89	90th perc.	22.5	0.00	0.86
10	Win. mean (10yr)	5.9	0.01	0.89	Mean	5.7	0.02	0.90	Win. mean (20yr)	10.1	0.00	0.89	Mean	22.6	0.00	0.86
11	Mean (20yr)	8.1	0.00	0.88	Mean (10yr)	5.8	0.02	0.90	50th perc. (20yr)	10.4	0.00	0.89	75th perc. (20yr)	31.3	0.00	0.70
12	90th perc. (10yr)	13.3	0.00	0.88	90th perc. (20yr)	6.1	0.02	0.90	10th perc. (20yr)	11.0	0.00	0.89	Win. mean (20yr)	39.4	0.00	0.37
13	25th perc. (20yr)	15.1	0.00	0.88	Win. mean (10yr)	6.5	0.01	0.90	Mean (20yr)	12.0	0.00	0.89	90th perc. (10yr)	41.0	0.00	0.27
14	75th perc. (10yr)	15.6	0.00	0.88	IQR mean (10yr)	6.5	0.01	0.90	25th perc. (20yr)	19.1	0.00	0.87	IQR mean (20yr)	41.4	0.00	0.25
15	10th perc. (20yr)	20.0	0.00	0.87	50th perc. (10yr)	6.6	0.01	0.90	75th perc.	21.4	0.00	0.86	50th perc. (20yr)	41.7	0.00	0.22
16	90th perc. (20yr)	25.1	0.00	0.87	90th perc. (10yr)	6.7	0.01	0.90	90th perc. (10yr)	23.0	0.00	0.86	75th perc. (10yr)	41.8	0.00	0.22
17	90th perc.	32.3	0.00	0.86	Win. mean	7.3	0.01	0.89	75th perc. (10yr)	25.6	0.00	0.85	Mean (10yr)	42.1	0.00	0.19
18	75th perc.	40.4	0.00	0.85	25th perc. (20yr)	7.5	0.01	0.89	90th perc. (20yr)	28.0	0.00	0.84	Win. mean (10yr)	42.3	0.00	0.18
19	Mean	88.4	0.00	0.77	75th perc. (10yr)	7.9	0.01	0.89	Mean	47.6	0.00	0.77	IQR mean (10yr)	42.3	0.00	0.18
20	Win. mean	96.0	0.00	0.76	10th perc. (20yr)	11.6	0.00	0.88	Win. mean	53.1	0.00	0.74	50th perc. (10yr)	42.3	0.00	0.18
21	IQR mean	109.3	0.00	0.73	IQR mean	13.0	0.00	0.88	IQR mean	56.2	0.00	0.72	25th perc. (10yr)	42.9	0.00	0.13
22	50th perc.	144.3	0.00	0.64	50th perc.	24.1	0.00	0.85	50th perc.	58.6	0.00	0.71	25th perc. (20yr)	43.0	0.00	0.12
23	25th perc.	210.2	0.00	0.37	25th perc.	38.6	0.00	0.79	25th perc.	95.3	0.00	0.38	10th perc. (10yr)	43.1	0.00	0.12
24	10th perc.	226.1	0.00	0.28	10th perc.	53.3	0.00	0.71	10th perc.	99.8	0.00	0.33	10th perc. (20yr)	43.2	0.00	0.11

* **8 statistics:** Interquartile (IQR), Winsorized (Win.), and arithmetic means; 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles; **3 time periods:** whole time series and last 10 and 20 years. Grey shading indicates models whose performances cannot be differentiated based on AIC.



Supp. Figure 1. The correlation between the mean ORCS Table of Attributes score and six metrics of overexploitation risk calculated from F/F_{MSY} values estimated in the data-rich assessment models. Black lines indicate linear regressions fit to the data and the gray shading indicates the confidence intervals for the regressions.



Supp. Figure 2. The marginal effects of the twelve TOA attributes in the boosted classification tree model. Lines represent the effect of each attribute on the probability that a stock is in each status category when the other attributes are held constant (green=underexploited, orange=fully exploited, red=overexploited). *TOA #2 Refuge availability* and *#4 Morphology affecting capture* were omitted from the model due to lack of predictive power. The relative influence of each attribute is shown in the upper right corner of each panel.

Appendix A: Modifications to the ORCS Table of Attributes

We estimated stock status using the expanded TOA developed by the SAFMC (2012, 2013) with a few modifications to increase clarity and objectivity in the scoring process:

TOA#1 Status of assessed stocks in fishery: We replaced the long and complex score descriptions with straightforward percentage thresholds.

TOA#5 Discard rate: We replaced the vague score descriptions with straightforward percentage thresholds and simplified the attribute to consider only the proportion of the catch discarded rather than the proportion discarded multiplied by the proportion of discards that die. The proportion discarded is generally more available than the proportion that die.

TOA#7 M compared to dominant species: We removed ambiguity in the difference between the score 1 and 2 descriptions (previously, both descriptions read *"M higher than or equal to M of dominant species"*; now, 1 reads *"M higher than M of dominant species"* and 2 reads *"M equal to M of dominant species"*). We specify that natural mortality rates must differ by 20% to be considered different.

TOA#8 Occurrence in catch: We simplified the language and replaced the vague score descriptions with straightforward percentage thresholds.

TOA#11 Habitat loss: We simplified the attribute to be conceptually simpler and more quantifiable. Rather than considering the rate of habitat loss (no loss, stable/declining, increasing), we consider the proportion of the taxa's life spent in threatened habitat (none of life, part of life, all of life). Data on the *rate* of regional and taxa specific habitat loss are difficult to find.

We made all scoring decisions within the context of the most recent 5 years. This is particularly relevant for *TOA #10*, *#12*, and *#13*, which examine recent trends in catch, fishing effort, and abundance index, respectively, but is also relevant for the *TOA #5 Discard rate*, *#8 Occurrence in catch*, *#14 Proportion of population protected*, etc.

We also renamed many of the original attributes for the sake of brevity and clarity.

Table 1. Renaming of the original TOA attributes.

#	New name	Original name
1	Status of assessed stocks in fishery	Overall fishery exploitation based on assessed stocks
2	Refuge availability	Presence of natural or managed refugia
3	Behavior affecting capture	Schooling, aggregation, or other behavior responses affecting capture
4	Morphology affecting capture	Morphological characteristics affecting capture
5	Discard rate	Discard mortality rate
6	Targeting intensity	Bycatch or actively targeted by the fishery
7	M compared to dominant species	Natural mortality compared to dominant species in the fishery
8	Occurrence in catch	Rarity
9	Value	Value or desirability
10	Recent trend in catch	Trend in catches (use only when effort is stable)
11	Habitat loss	Loss or alteration of habitat
12	Recent trend in effort	Fleet stability
13	Recent trend in abundance index	Fishery Independent CPUE
14	Proportion of population protected	Effectiveness of regulations (other than ACLs) to limit exploitation

Appendix B: ORCS Table of Attributes scoring guide

TOA #1 – Status of assessed stocks in fishery

We identified the proportion of assessed stocks in the fishery that are overfished by determining the status of all stocks within the relevant fisheries management unit. For United States fisheries, we identified stocks managed together by Fisheries Management Plans (FMPs) and used analogous management units for foreign fisheries (see **Table 2** below). The proportion of overfished stocks was calculated using only assessed stocks in the fishery. We used the definition of "overfished" used by the management agency. We identified the following thresholds for scoring:

1 – <10% of assessed stocks are overfished
2 – 10-25% of assessed stocks are overfished
3 – >25% of assessed stocks are overfished
NA – Target stock is the only stock in the fishery or stock statuses are unknown

Table 2. Management units used to identify groups of co-managed stocks.

Country	Assessments grouped by	Reference
United States	Fisheries Management Plan	US-FMC, 2016
New Zealand	Fisheries Group	NZ-MPI, 2016
Australia	Fishery	AFMA, 2016
South Africa	Fishery	DAFF, 2012
Tuna-RFMO	RFMO	Pons et al., 2016
Other-RFMO	Fishery	Many references
Canada	Integrated Fisheries Management Plan	DFO, 2016
Argentina	Fishery	INIDEP, 2016
Russia	Fishery	Sobolevskaya and Divovich, 2015

TOA #2 – Refuge availability

We scored this attribute by comparing maps of a taxa's range with maps of protected areas within its range (such maps are often available in the stock assessment). In nearly all cases, the range of the taxa was enormous relative to the protected areas. It was difficult to consider natural refugia because rocky outcrops and depth aren't true barriers to fishing when the resource is valuable or desired. We used the following percentage thresholds (though they are generally impossible to empirically quantify):

1 – <50% of habitat is accessible to fishing
2 – 50-75% of habitat is accessible to fishing
3 – >75% of habitat is accessible to fishing

TOA #3 – Behavior affecting capture

We reviewed the "biology of the species" portion of the relevant stock assessment, the FishBase profile of the species, and other resources to identify behavior that might affect the susceptibility of the taxa to capture. Only taxa exhibiting schooling, shoaling, or spawning aggregation behaviors targeted by fishermen were scored as being highly susceptible to capture. All other taxa (those not exhibiting these behaviors or those whose spawning aggregations are not targeted by fishermen) were scored as being moderately susceptible to capture. No stocks were scored as exhibiting low susceptibility to capture (what traits would even make a taxon unsusceptible to capture?)

1 – No examples – slimy eels, fast tuna, reclusive tilefish are all susceptible
2 – Don't exhibit schooling/shoaling/aggregation behavior
3 – Exhibit schooling/shoaling/aggregation behavior
NA – Schooling/shoaling/aggregation behavior unknown

TOA #4 – Morphology affecting capture

We reviewed the "biology of the species" portion of the relevant stock assessment, the FishBase profile of the species, and other resources to identify morphology that affect the susceptibility of the taxa to capture. We were unable to identify any taxa that exhibited morphology that would make them either unsusceptible or highly susceptible to capture. If the taxa are desired, the fishing gear/method is adapted to target the taxa despite its challenges. Everything was identified as being of average susceptibility to capture.

1 – No examples
2 – Everything
3 – No examples

TOA #5 – Discard rate

We determined the proportion of the catch discarded from the relevant stock assessment or other relevant resource and used the following percentage thresholds to assign scores:

1 – <10% of catch discarded
2 – 10-25% of catch discarded
3 – >25% of catch discarded
NA – Discard rates are unknown

TOA #6 – Targeting intensity

We reviewed the "history of the fishery" portion of the relevant stock assessment and other relevant resources to determine the targeting intensity of the fishery and used the following classifications to assign scores:

1 – Not targeted (bycatch / incidental catch only)

2 – Occasionally targeted (often part of multi-species catch)
3 – Actively targeted (directed fishery)
NA – Targeting intensity unknown

TOA #7 – M compared to dominant species

We reviewed the "history of the fishery" portion of the relevant stock assessment and other relevant resources to determine the proportion of the catch coming from different sectors in order to infer the other taxa represented in the fishery. If multiple taxa are represented in the fishery, we used a variety of sources to determine the dominant species by weight (not value). The natural mortality rates of the target and dominant species were determined from the relevant stock assessments. If the target species was dominant or if there were no other taxa in the fishery, the attribute could not be scored. The following classifications were used to assign scores when the attribute could be scored (note: natural mortality rates must differ by >10% to be considered different):

1 – M higher than M of dominant species
2 – M approximately equal to M of dominant species
3 – M lower than M of dominant species
NA (common for this attribute) – only taxa in fishery or is the dominant taxa in fishery or natural mortality rates are unknown
* Our scores were reviewed by experts familiar with regional fisheries for accuracy.

TOA #8 – Occurrence in catch

We use the percentage threshold guidelines listed below to assign scores. However, this data is not often available and decisions could be fairly subjective and non-quantitative.

1 – 0-10% of trawl tows, gillnet sets, trap pulls, etc.
2 – 10-25% of trawl tows, gillnet sets, trap pulls, etc.
3 – >25% of trawl tows, gillnet sets, trap pulls, etc.

NA – the relative occurrence of the taxa in the catch is unknown

TOA #9 – Value

We determined value (USD lb^{-1}) by deriving ex-vessel price for taxa by region from the Sea Around Us Project landings volume and value database (Sumaila et al., 2007; Pauly and Zeller, 2015). We used the average price from 2006-2010, the most recent 5 years with data, for scoring. For most stocks, appropriate regional prices could be tied to the stock. For highly migratory species like tuna, marlin, swordfish and stocks managed by a RFMO (e.g., Mediterranean or West African stocks), average values from the relevant countries were used. For the few stocks without price data in the database, we found ex-vessel prices in other references. We used the following thresholds to assign scores:

1 – <$1.00 lb^{-1}
2 – $1.00-2.25 lb^{-1}
3 – >$2.25 lb^{-1}
NA – Ex-vessel price is unknown

TOA #10 – Recent trend in catch

We identified the recent trend in catch for each scored stock using (1) annual catch time series in the RAM Legacy Stock Assessment Database or (2) figures and tables in the original stock assessment when catch time series were not included in the database. We used Theil-Sen regression to identify trends in catch in the most recent 5 years where a (1) significant positive slope indicates increasing catch, (2) significant negative slope indicates decreasing catch, and (3) non-significant slope indicates stable catch over the most recent 5 years. Theil-Sen regression fits a line to a set of points by identifying the median slope among lines through all possible point pairs and is insensitive to outliers and endpoints in short time series. Both trends in total catch (landings + discards) and landings were identified where possible and trends in total catch were used over trends in landings.

1 – Significant increase in catch in recent 5 years
2 – No significant change in catch in recent 5 years
3 – Significant decrease in catch in recent 5 years
NA – Catch data are not available

TOA #11 – Habitat loss

We classified taxa that reside in threatened estuary (Lotze et al., 2006), seagrass (Orth et al., 2006; Waycott et al., 2009), mangrove (Giri et al., 2010), or coral reef (Pandolfi et al., 2003, 2011) habitats for their whole lives or a portion of their lives as being at high and moderate risk of overexploitation. We classified taxa that spend their entire lives outside these threatened habitats as being at low risk of overexploitation. We classified taxa that spend the entirety of their lives in partially threatened inshore areas such as the intertidal zone or rocky reefs (Lotze et al., 2006; Rabalais et al., 2009) as being at moderate risk of overexploitation.

1 – No time in threatened habitats
2 – Part time in threatened habitats (or full time in partially threatened habitats)
3 – Full time in threatened habitats
NA – Habitat preferences are unknown

TOA #12 – Recent trend in effort

We identified the recent trend in fishing effort for each scored stock using fishing mortality rate estimates as a proxy for fishing effort. We analyzed (1) annual fishing mortality time series in the RAM Legacy Stock Assessment Database or (2) figures and tables in the original stock assessment when fishing mortality time series were not included in the database. We used Theil-Sen regression to identify trends in fishing mortality in the most recent 5 years where a (1) significant positive slope indicates increasing effort, (2) significant negative slope indicates decreasing effort, and (3) non-significant slope indicates stable effort over the most recent 5 years. Theil-Sen regression fits a line to a set of points by identifying the median slope among lines through all possible point pairs and is insensitive to outliers and endpoints in short time series. Both trends in fishing mortality rate (F) and exploitation rate (ER) were identified where possible and trends in fishing mortality rate were used over trends in exploitation rate.

1 – Significant decrease in fishing effort in recent 5 years
2 – No significant change in fishing effort in recent 5 years
3 – Significant increase in fishing effort in recent 5 years
NA – Effort data are not available

TOA #13 – Recent trend in abundance index

We identified the recent trend in fisheries independent CPUE for each scored stock using stock assessment model abundance estimates as a proxy for CPUE. We analyzed (1) annual abundance time series in the RAM Legacy Stock Assessment Database or (2) figures and tables in the original stock assessment when abundance time series were not included in the database. We used Theil-Sen regression to identify trends in the abundance index in the most recent 5 years where a (1) significant positive slope indicates increasing abundance index, (2) significant negative slope indicates decreasing abundance index, and (3) non-significant slope indicates stable abundance index over the most recent 5 years. Theil-Sen regression fits a line to a set of points by identifying the median slope among lines through all possible point pairs and is insensitive to outliers and endpoints in short time series. Trends in spawning stock biomass (SSB), total biomass (TB), number of individuals (TN), and number of recruits (R) were identified where possible and were preferentially used in the same order.

1 – Significant increase in abundance index in recent 5 years
2 – No significant change in abundance index in recent 5 years
3 – Significant decrease in abundance index in recent 5 years
NA – Survey data are not available

TOA #14 – Proportion of population protected

We determined whether the fishery was managed using (1) size limits, (2) protected areas, (3) seasonal closures, or (4) significant effort controls / gear restrictions. Fisheries employing no measures received a high risk score, one measure a moderate risk score, and size limits and one other measure a low risk score.

1 – Size limits AND (protected areas OR seasonal closures)
2 – Size limits OR protected areas OR seasonal closures
3 – No size limits, no protected areas, no seasonal closures
NA – Management regulations are unknown

Appendix C: Example application of the original and refined ORCS approaches

We illustrate the application of the ORCS approach using the US Southeast/Gulf of Mexico red porgy (*Pagrus pagrus*) stock. The scores and justifications for all 193 stocks are available in **Supplementary Appendix B**. In **Table 3**, we show the application of ORCS Table of Attributes and the status predicted by both the original and refined ORCS approaches compared to that of the data-rich assessment.

Table 3. ORCS Table of Attribute scores and justifications for US Southeast/Gulf of Mexico red porgy stock

	Attribute	Score	Justification *(source)*
1	Status of assessed stocks in fishery	3 - >25% overfished	SA Snapper-Grouper FMP = 14 other stocks: 3 overfished, 3 unknown, 3/11=27.3% overfished *(NOAA, 2015)*
2	Refuge availability	3 - >75% of habitat accessible	Wide range (South Atlantic / Gulf of Mexico) and few protected areas *(SAFMC, 2017)*
3	Behavior affecting capture	2 - Average susceptibility	No schooling/aggregation behavior *(FishBase, 2017)*
4	Morphology affecting capture	2 - Average susceptibility	91 cm max (7.7 kg max), 35 cm common, 26 cm mature *(FishBase, 2017)*
5	Discard rate	1 - Discards <10% of catch	5.1% discards; 2011: 18.55 mt discards / 359.75 mt total catch = 5.1% discards *(Table 16 in stock assessment)*
6	Targeting intensity	2 - Occasionally targeted	Targeted in multi-species snapper-grouper complex; catch is largely commercial, mostly lines and some trawl *(stock assessment)*
7	M compared to dominant species	2 - Equivalent mortality rates	M=0.225; Vermillion snapper (M=0.22) is dominant *(stock assessment)*
8	Occurrence in catch	2 - Common	Bag limit only met 12-22% of time *(Page 10 of stock assessment)*
9	Value	2 - $1-$2.25 / lb	2006-2010: $1.55 / lb in USA East Coast & Gulf of Mexico *(Sumuila et al., 2007)*
10	Recent trend in catch	2 - Stable last 5 years	Total landings stable in recent 5 years (*RAMLDB)*
11	Habitat loss	2 - Part time in threatened habitats	Benthopelagic (0-250m, 10-80m core), rock/rubble/sand bottoms, young in seagrass beds *(SAFMC, 2017)*
12	Recent trend in effort	2 - Stable last 5 years	F/ER stable in recent 5 years *(RAMLDB)*
13	Recent trend in abundance index	3 - Decreasing last 5 years	SSB/TB decreasing in recent 5 years *(RAMLDB)*
14	Proportion of population protected	2 - Some of resource protected	Commercial = size/season limits, fishery closed once quota is met; recreational = catch/size/trip/bag limits
	Original ORCS approach status:	**Fully exploited (2.143 mean score)**	
	Refined ORCS approach status:	**Overexploited (51.5% probability)**	
	Data-rich assessment status:	**Overexploited (B/B_{MSY} = 0.474)**	

The original ORCS approach incorrectly classified the US Southeast/Gulf of Mexico red porgy stock as fully exploited while the refined ORCS approach correctly classified the stock as overexploited. In the refined ORCS approach, the OFL (U_{MSY} x total biomass) is estimated for overexploited stocks by multiplying the 10th percentile of the whole catch time series by a scalar, where the choice of scalar is determined by the managing agency (see **Table 3** in the manuscript for potential catch scalars). For example, using the median catch scalar for overexploited stocks, the OFL for red porgy would be calculated as:

$$OFL = 10^{th}\ percentile\ whole\ time\ series * 1.56$$

$$OFL = 74.3\ mt * 1.56 = 115.9\ mt$$

The OFL estimated by the refined ORCS approach (115.9 mt) is less than that estimated from the data-rich assessment (282.7 mt) indicating that the refined approach would underutilize available biomass for the red porgy stock (**Figure 1**).

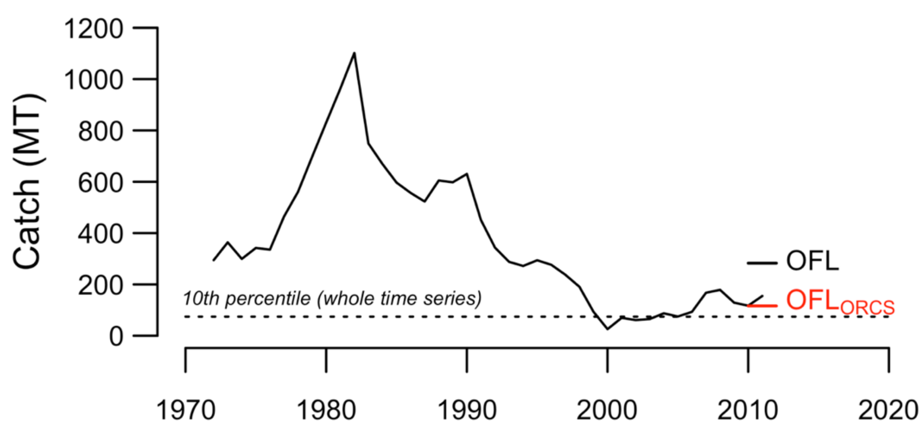


Figure 1. The US Southeast/Gulf of Mexico red porgy catch history showing the OFLs estimated for the terminal year by the refined ORCS approach (red; 115.9 mt) and from the data-rich assessment (black; 282.7). Because the refined ORCS approach classified the stock as overexploited, it estimates the OFL as the 10th percentile of the catch time series (dashed line) multiplied by a scalar, where the choice of scalar is determined by the managing agency. The OFL shown here was calculated using the median catch scalar for overexploited stocks.

Appendix D: Development of the boosted classification tree model

We refined the ORCS Table of Attributes using boosted classification trees (BCT) to weight attributes by their relative importance, incorporate interactions between attributes, and account for non-linearity in attribute behavior. The BCT analysis was performed using the *caret* (Kuhn, 2016) and *gbm* (Ridgeway, 2016) packages in R v.3.3.2 (R Core Team, 2016).

We used the *train* function in the *caret* package (Kuhn, 2016) to conduct a grid search for the BCT model parameters that maximize Cohen's kappa using repeated 10-fold cross validation on the training dataset. We optimized the standard BCT parameters – interaction depth, learning rate, and number of trees – but also optimized the bag fraction, which controls the proportion of the data used in each iteration (**Table 4**). Lower values promote stochasticity and increase predictive performance on large datasets where omitting data is not a problem. Higher values reduce stochasticity but give the model more data to learn from when working with small datasets (Natekin and Knoll, 2013). Because of the small size of our dataset, we evaluated bag fractions from 50% to 90%.

Table 4. Boosted classification tree (BCT) model parameters and tuning values.

BCT model parameter	Values evaluated
Interaction depth (a.k.a., tree complexity)	c(1, 2, 3)
Learning rate (a.k.a., shrinkage, step-size reduction)	c(0.001, 0.005, 0.0001)
Number of trees	seq(from=100, to=3500, by=100)
Bag fraction	seq(from=0.5, to=0.9, by=0.1)
Minimum number of observations in terminal nodes	10

We trained the model using this grid with both numeric (1s, 2s, and 3s are numbers) and categorical (1s, 2s, and 3s are factors) scores and with both unweighted observations and observations weighted by stock status. The weighted observation approach was designed to give more weight to the rare but important overexploited stocks in an attempt to improve predictions in this category (**Table 5**). This yielded the following four modeling approaches:

1. Numeric scores (unweighted)
2. Categorical scores (unweighted)
3. Numeric scores – observations weighted by stock status
4. Categorical scores – observations weighted by stock status

The unweighted numeric score approach generally performed better than the alternatives (**Figure 1**), presumably because numeric values provide valuable information on score order (factors are unordered) and because fitting the model by maximizing Cohen's kappa already does the work expected from weighting the stocks by the rarity of their status.

Table 5. Weights assigned to the data of stocks of each status in weighted model training.

Status	# of observations	Weight
Low	55	1.00 (55/55)
Moderate	76	0.72 (55/76)
High	24	2.29 (55/24)

Figure 2 shows the kappa and accuracy statistics for the best model tune in each modeling approach–bag fraction scenario. The numeric modeling approach with a 0.8 bag fraction was selected because it exhibited the highest mean and median values for both Cohen's kappa and accuracy on the training dataset.

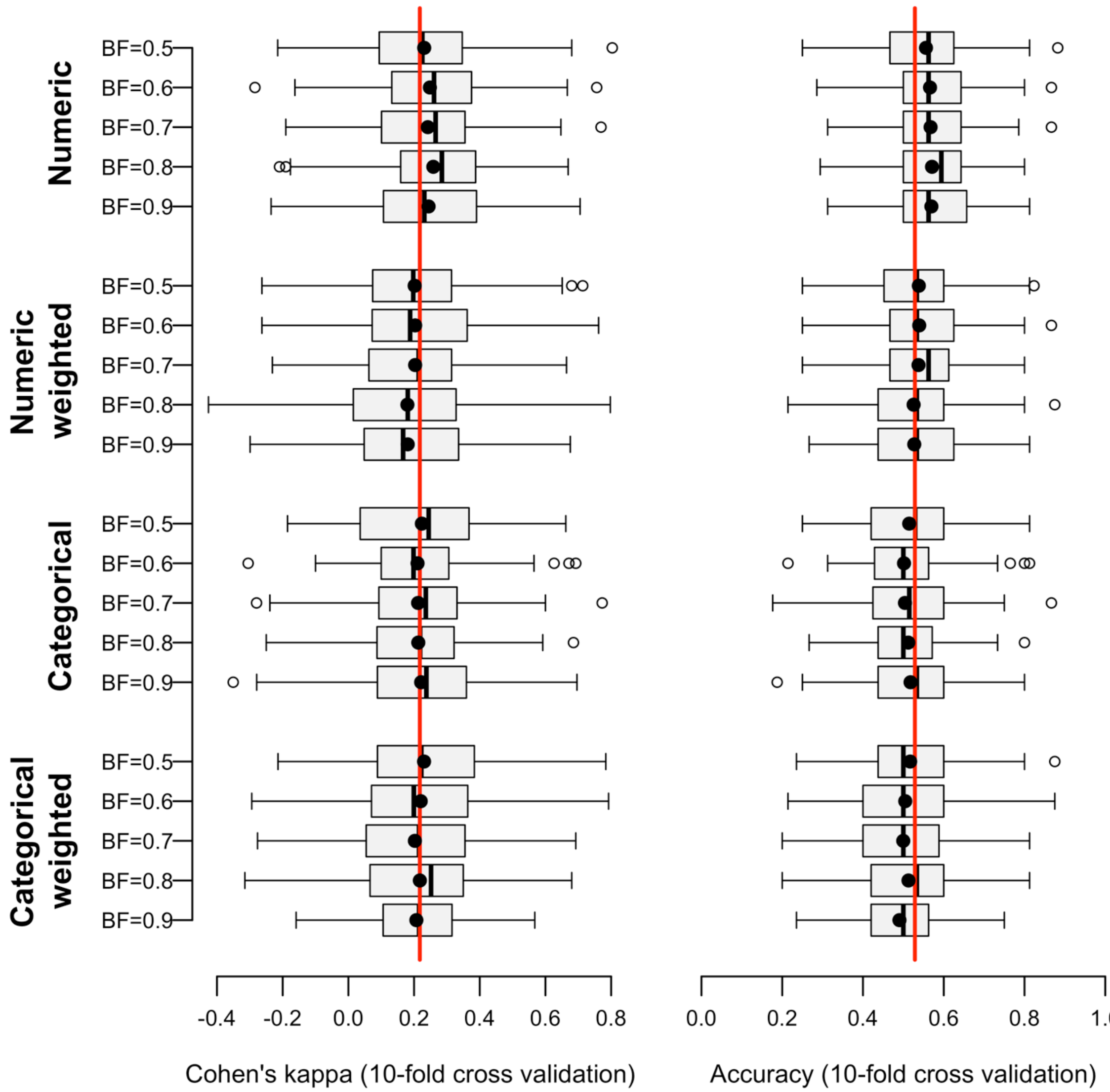


Figure 2. Cohen's kappa and accuracy statistics for the best model tune in each modeling approach–bag fraction scenario. Boxplots indicate the median (heavy black line), interquartile range (IQR; box), 1.5 times the IQR (whiskers), and outliers. Solid points indicate the mean. The red line indicates the mean value across all scenarios.

The mean kappa statistic for every parameter combination in the unweighted numeric modeling framework is shown in **Figure 3** and neatly illustrates the process for identifying the best model parameters. Ultimately, we used the model with learning rate=0.001, interaction depth=2, number of trees=3000, and bag fraction=0.8.

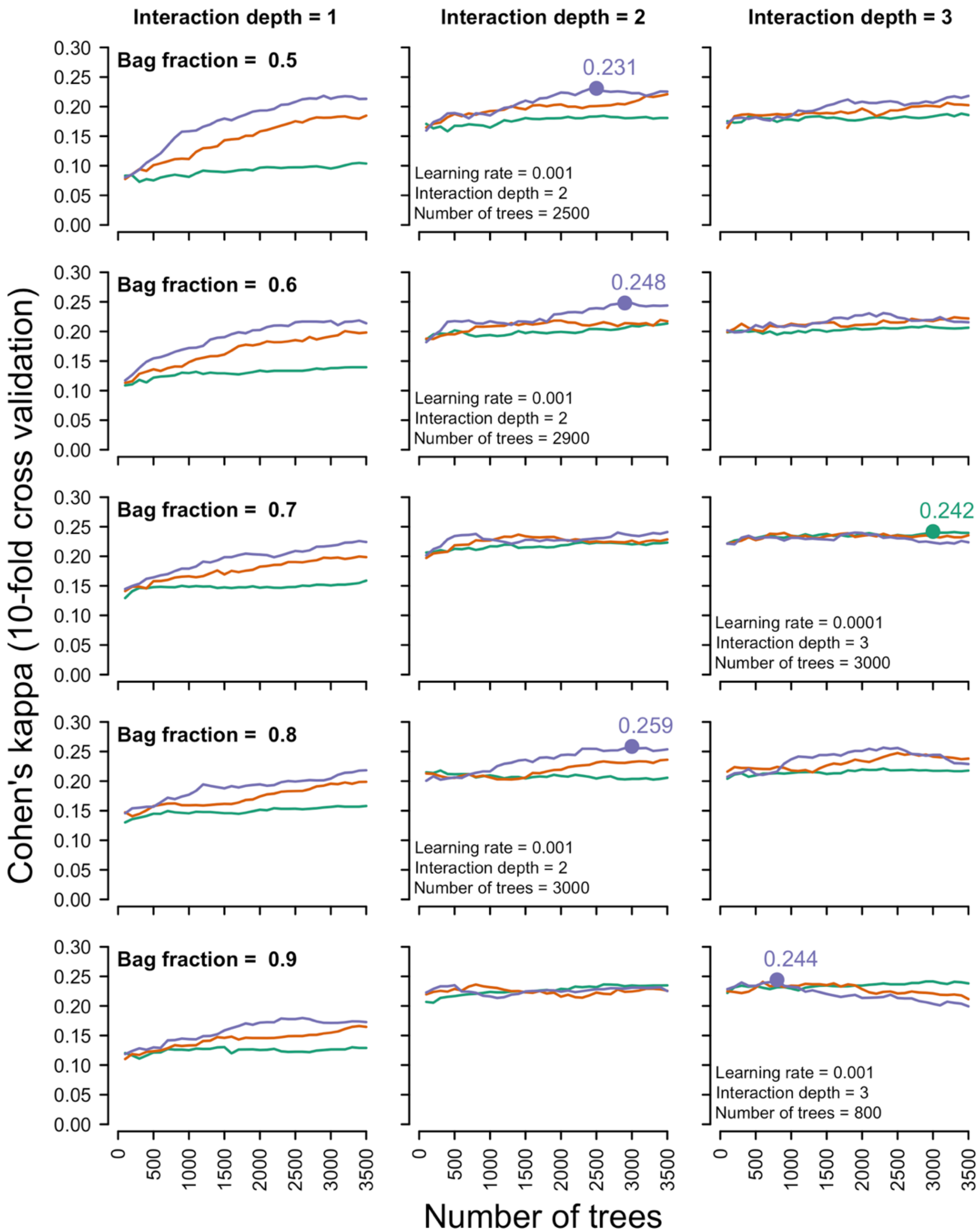


Figure 3. Mean Cohen's kappa for each combination of BCT model parameters. Lines indicate learning rate: purple (0.001), orange (0.0005), green (0.0001). The best combination of model parameters is the one that maximizes the kappa value. The best combination of parameters is labeled for each bag fraction scenario.

Appendix E: Application and performance of other catch-only methods

We compared the performance of the BCT model to six other catch-only methods for estimating status on the 37 stocks with catch time series in the test dataset: SSP-2002 (Froese and Kesner-Reyes, 2002) and SSP-2013 (Kleisner et al., 2013), which estimate development status (e.g., undeveloped, developing, fully exploited), and CMSY (Martell and Froese, 2013), COM-SIR (Vasconcellos and Cochrane, 2005), SSCOM (Thorson et al., 2013), and mPRM (Costello et al., 2012), which estimate B/B_{MSY}.

The two 'stock status plot' (SSP) methods evolved from the FAO's efforts to assess the status of global fisheries (Grainger and Garcia, 1996) and use simple rules to identify fisheries development status (Csirke and Sharp, 1984) from patterns in catch time series. The Froese and Kesner-Reyes (2002) method, SSP-2002, identifies a stock as 'undeveloped', 'developing', 'fully exploited', 'overfished', or 'collapsed' based on comparison of the target year's catch relative to the maximum year's catch (**Table 6**). The Kleisner et al. (2013) method, SSP-2013 (actually published first in Kleisner and Pauly, 2011), adds an additional 'rebuilding' category by considering the minimum catch occurring after the maximum catch ('post-maximum minimum') (**Table 7**). The tables illustrate how we mapped the SSP-2002 and SSP-2013 status categories into under, fully, and overexploited statuses.

Table 6. Criteria used to classify stock status in SSP-2002 (Froese and Kesner-Reyes, 2002).*

Stock status	SSP-2002 status	Criteria
Underexploited	Undeveloped	C_{curr} before C_{max} **AND** $C_{curr} < 0.1*C_{max}$
Underexploited	Developing	C_{curr} before C_{max} **AND** $0.1*C_{max} \leq C_{curr} \leq 0.5*C_{max}$
Fully exploited	Fully exploited	$C_{curr} > 0.5*C_{max}$
Overexploited	Overfished	C_{curr} after C_{max} **AND** $0.1*C_{max} \leq C_{curr} \leq 0.5*C_{max}$
Overexploited	Collapsed/closed	C_{curr} after C_{max} **AND** $C_{curr} < 0.1*C_{max}$

* C_{curr} = current catch; C_{max} = maximum catch

Table 7. Criteria used to classify stock status in SSP-2013 (Kleisner et al., 2013).*

Stock status	SSP-2013 status	Criteria
Underexploited	Developing	C_{curr} before C_{max} **AND** $C_{curr} \leq 0.5*C_{max}$ **OR** C_{max} in final year of time series
Fully exploited	Exploited	$C_{curr} > 0.5*C_{max}$
Overexploited	Overexploited	C_{curr} after C_{max} **AND** $0.1*C_{max} \leq C_{curr} \leq 0.5*C_{max}$
Overexploited	Collapsed	C_{curr} after C_{max} **AND** $C_{curr} < 0.1*C_{max}$
Overexploited	Rebuilding	C_{curr} after $C_{post\text{-}max\ min}$ **AND** $C_{post\text{-}max\ min}$ $0.1*C_{max}$ **AND** $0.1*C_{max} \leq C_{curr} \leq 0.5*C_{max}$

* C_{curr} = current catch; C_{max} = maximum catch; $C_{post\text{-}max\ min}$ = minimum catch after the maximum catch

The other four methods, CMSY (Martell and Froese, 2013), COM-SIR (Vasconcellos and Cochrane, 2005), SSCOM (Thorson et al., 2013), and mPRM (Costello et al., 2012), use catch data and basic life-history parameters to estimate B/B_{MSY}. We selected these methods because they can be applied to the vast majority of global fisheries, are established in the literature, have been extensively simulation tested (Rosenberg et al., 2014), and can be easily implemented using the *datalimited* package in R (Anderson et al., 2016, 2017).

1. **CMSY (catch-MSY)** implements a stock-reduction analysis with Schaefer biomass dynamics (Martell and Froese, 2013). It requires prior distributions on r and K as well as priors on the relative proportion of biomass at the

beginning and end of the time series compared to unfished biomass (depletion). The version of the model used in Rosenberg et al. (2014) and implemented in the *datalimited* package (Anderson et al., 2016) was modified from Martell and Froese (2013) to generate biomass trends from all viable r-K pairs and produce an estimate of B/B_{MSY} from the median trend.

2. **COM-SIR (catch-only-model with sampling-importance-resampling)** is a coupled harvest-dynamics model (Vasconcellos and Cochrane, 2005) in which biomass and harvest dynamics are assumed to follow Schaefer and logistic models, respectively. The model is fit using a sampling-importance-resampling algorithm (Rosenberg et al. 2014).
3. **SSCOM (state-space catch-only model)** is a hierarchical model also based on a coupled harvest-dynamics model (Thorson et al., 2013). SSCOM estimates unobserved dynamics in both population biomass and fishing effort based on a catch time series and priors on *r*, the maximum rate of increase of fishing effort, and the magnitude of three forms of stochasticity. The model is fit in a Bayesian state-space framework to integrate across three forms of stochasticity: variation in effort, population dynamics, and fishing efficiency (Thorson et al., 2013).
4. **mPRM (modified panel regression model)** is a modified version of the panel-regression model from Costello et al. (2012), which used the RAM Legacy Stock Assessment Database to predict B/BMSY from characteristics of the catch time series and stock. The implementation of the model in the *datalimited* package (Anderson et al., 2016) is modified from the original in

that it uses a different suite of life-history categories and removes the maximum catch predictor.

We applied CMSY using resilience categories from FishBase and 2 million iterations. We applied COM-SIR using resilience categories from FishBase and 4 million iterations. We applied mPRM using species categories from Anderson et al. (2016).

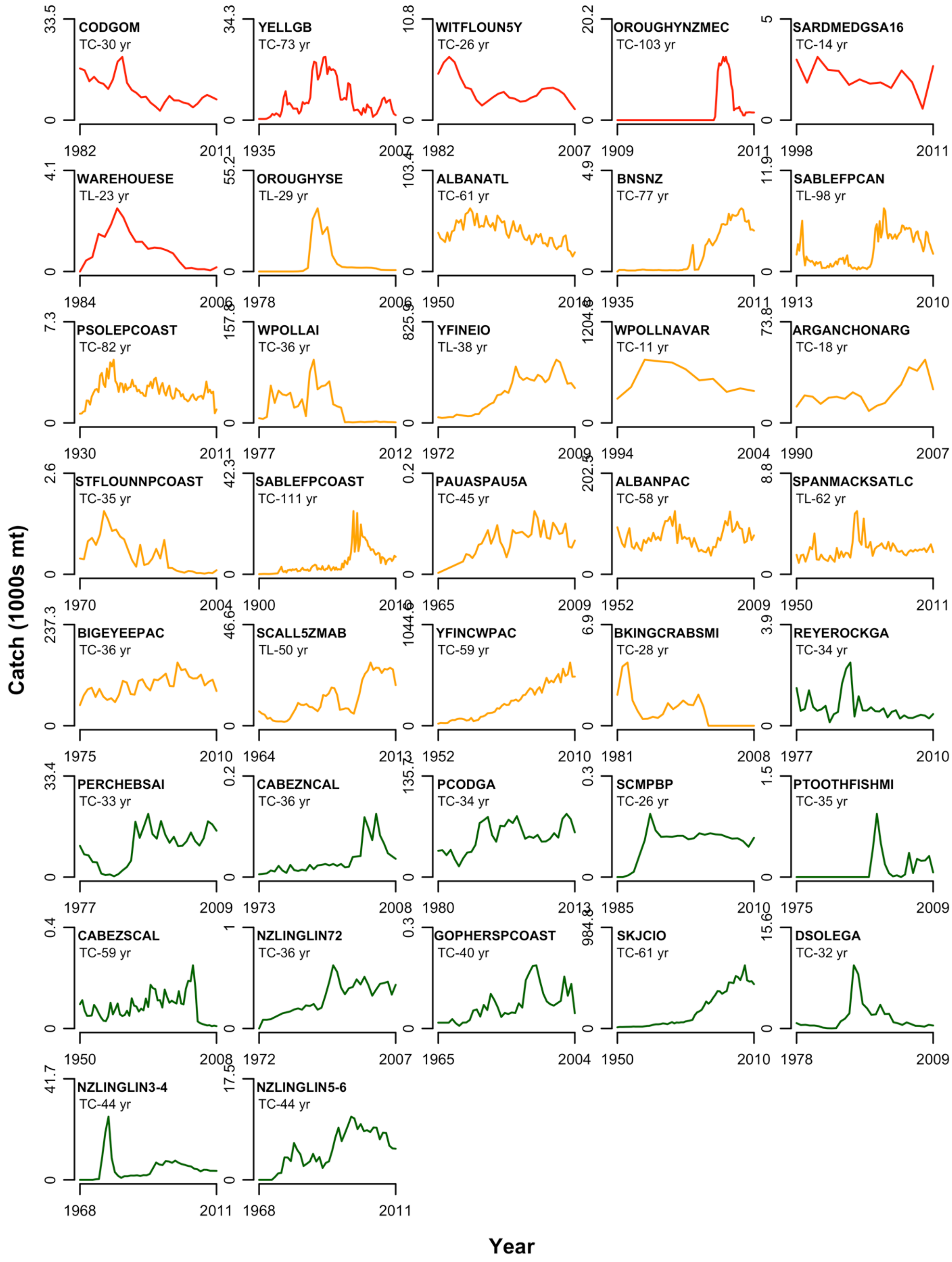


Figure 4. Catch time series for the 37 data-rich stocks with catch time series in the test dataset (TC=total catch, n=31; TL=total landings, n=6). These stocks were used to evaluate the status classification performance of the refined ORCS approach and four alternative catch-only methods. Stocks are arranged and colored by status (B/B_{MSY}).

Appendix F: Refinement of the historical catch statistics and scalars

We identified the best status-specific historical catch statistics and scalars by comparing the OFL in the terminal year to 24 potential historical catch statistics (*8 statistics x 3 time periods*) for the 105 stocks in the RAMLDB with the following information: (1) an estimate of U_{MSY} or an estimate of F_{MSY} and M if U_{MSY} is unavailable; (2) a time series of total biomass; (3) a time series of total catch or landings if total catch is unavailable; and B/B_{MSY} (stock status). We estimated U_{MSY} for stocks where U_{MSY} is unavailable using the following equation:

$$U_{MSY} = \left[\frac{F_{MSY}}{M + F_{MSY}}\right] * \left[1 - e^{-(M+F_{MSY})}\right]$$

and calculated the OFL as the product of U_{MSY} and the terminal year total biomass:

$$OFL = U_{MSY} * Total\ biomass$$

Table 8. Selection of stocks with usable U_{MSY} values.

RAM Legacy Database subset	# of stocks
All assessments	512
Assessments with U_{MSY}	114
Assessments without U_{MSY} but with F_{MSY} and M	64
Assessments with U_{MSY} provided and calculated	178
- minus 1 stock with an unrealistic U_{MSY} (1.47)	177

Table 9. Selection of stocks with usable U_{MSY} values and time series of total biomass and total catch (or total landings).

RAM Legacy Database subset	# of stocks
Assessments with usable U_{MSY} values (from above)	177
Assessments with U_{MSY} and TB and TC or TL time series in same units	128
Assessments with U_{MSY} and TB and TC or TL time series ≥ 10 yr	128
Assessments with U_{MSY} and TB and TC or TL time series ≥ 15 yr	125
Assessments with U_{MSY} and TB and TC or TL time series ≥ 20 yr	121
- minus 1 stock without an OFL in the terminal year of the catch time series	120
- minus 15 stocks without B/B_{MSY} values (stock statuses)	105
- minus 8 stocks not scored using the ORCS approach (b/c SSB_{MSY} in eggs, larvae, or gonads)	97

We calculated the following catch statistics over the whole time series, the most recent 10 years, and the most recent 20 years (*8 statistics x 3 time periods*):

- *Arithmetic mean*
- *Interquartile mean* – the mean of values in the interquartile range (25-75th percentile); less sensitive to outliers than the arithmetic mean
- *Winsorized mean* – the mean of the data with the upper and lower 25th percentile values replaced by the next largest value
- *10th, 25th, 50th, 75th, and 90th percentiles*

Appendix G: Supplemental references

ORCS approaches

Berkson, J., Barbieri, L., Cadrin, S., Cass-Calay, S., Crone, P., Dorn, M., Friess, C., Kobayashi, D., Miller, T.J., Patrick, W.S., Pautzke, S., Ralston, S., Trianni, M., 2011. Calculating Acceptable Biological Catch for stocks that have reliable catch data only (Only Reliable Catch Stocks – ORCS). NOAA Technical Memorandum NMFS-SEFSC-616. NMFS-SEFSC, Miami, FL.

SAFMC, 2012. SSC ORCS Workshop Report. South Atlantic Fishery Management Council-Scientific & Statistical Committee (SAFMC-SSC), North Charleston, SC.

SAFMC, 2013. SSC ORCS Workshop II Report. South Atlantic Fishery Management Council-Scientific & Statistical Committee (SAFMC-SSC), North Charleston, SC.

Databases used in scoring

Froese, R., Pauly, D., 2016. FishBase. Available: www.fishbase.org

Pauly D., Zeller D., 2015. Sea Around Us Concepts, Design and Data. Available: www.seaaroundus.org

Ricard, D., Minto, C., Jensen, O.P., Baum, J.K., 2012. Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish Fish. 13(4), 380-398.

Sumaila, U. R., Marsden, A. D., Watson, R., Pauly, D., 2007. A global ex-vessel fish price database: construction and applications. J. Bioecon. 9(1), 39-51.

Fisheries management units (used in scoring *TOA #1*)

AFMA, 2016. Fisheries. Australian Fisheries Management Authority (AFMA). Available: http://www.afma.gov.au/fisheries/

DAFF, 2012. Status of the South African Marine Fishery Resources 2012. Department of Agriculture, Forestry, and Fisheries (DAFF), Pretoria, South Africa.

DFO, 2016. Integrated Fisheries Management Plans. Fisheries and Oceans Canada (DFO). Available: http://www.dfo-mpo.gc.ca/fm-gp/peches-fisheries/ifmp-gmp/index-eng.htm

INIDEP, 2016. Argentinean Sea Fisheries (in Spanish). Argentina National Institute of Research and Fisheries Development. Available: http://www.inidep.edu.ar/pesquerias/ppales-pesquerias/

NZ-MPI, 2016. Our Fisheries: Fisheries Groups. New Zealand Ministry for Primary Industries (NZ-MPI). Available: http://fs.fish.govt.nz/Page.aspx?pk=45&tk=463

Pons, M., Branch, T. A., Melnychuk, M. C., Jensen, O. P., Brodziak, J., Fromentin, J. M., Harley, S. J., Haynie, A. C., Kell, L. T., Maunder, M. N., Parma, A. M., 2016. Effects of biological, economic and management factors on tuna and billfish stock status. Fish. Fish. 18(1), 1-21.

Sobolevskaya, A., Divovich, E., 2015. The Wall Street of fisheries: the Russian Far East, a catch reconstruction from 1950 to 2010. Working Paper #2015-45. Fisheries Centre, University of British Columbia, Vancouver, Canada.

US-FMC, 2016. Fishery Management Plans of the U.S. Regional Fishery Management Councils. Available: http://www.fisherycouncils.org/

Threatened habitats (used in scoring *TOA #11*)

Giri, C., Ochieng, E., Tieszen, L. L., Zhu, Z., Singh, A., Loveland, T., Masek, J., Duke, N., 2011. Status and distribution of mangrove forests of the world using earth observation satellite data. Glob. Ecol. Biogeogr. 20(1), 154-159.

Lotze, H. K., Lenihan, H. S., Bourque, B. J., Bradbury, R. H., Cooke, R. G., Kay, M. C., Kidwell, S. M., Kirby, M. X., Peterson, C. H., Jackson, J. B. C., 2006. Depletion, degradation, and recovery potential of estuaries and coastal seas. Science. 312(5781), 1806-1809.

Orth, R. J., Carruthers, T. J. B., Dennison, W. C., Duarte, C. M., Fourqurean, J. W., Heck, K. L., Hughes, A. R., Kendrick, G. A., Kenworthy, W. J., Olyarnik, S., Short, F. T., Waycott, M., Williams, S. L., 2006. A global crisis for seagrass ecosystems. Bioscience 56(12), 987-996.

Pandolfi, J. M., Bradbury, R. H., Sala, E., Hughes, T. P., Bjorndal, K. A., Cooke, R. G., McArdle, D., McClenachan, L., Newman, M. J. H., Paredes, G., Warner, R. R., Jackson, J. B. C., 2003. Global trajectories of the long-term decline of coral reef ecosystems. Science. 301(5635), 955-958.

Pandolfi, J. M., Connolly, S. R., Marshall, D. J., Cohen, A. L., 2011. Projecting coral reef futures under global warming and ocean acidification. Science. 333(6041), 418-422.

Rabalais, N. N., Turner, R. E., Díaz, R. J., Dubravko Justić, D., 2009. Global change and eutrophication of coastal waters. ICES J. Mar. Sci. 66(7), 1528-1537.

Waycott, M., Duarte, C. M., Carruthers, T. J., Orth, R. J., Dennison, W. C., Olyarnik, S., Calladine, A., Fourqurean, J. W., Heck, K. L., Hughes, A. R., Kendrick, G. A., 2009. Accelerating loss of seagrasses across the globe threatens coastal ecosystems. Proc. Natl. Acad. Sci. 106(30), 12377-12381.

Example TOA application (used in Supplementary Appendix A.3 Table 3)

NOAA, 2015. Status of Stocks 2015: Annual Report to Congress of the Status of U.S Fisheries. NOAA Fisheries.

SAFMC, 2017. Red porgy. South Atlantic Fishery Management Council (SAFMC). Available at: http://safmc.net/regulations/regulations-by-species/red-porgy/

SEFSC, 2012. Stock Assessment of Red Porgy off the Southeastern United States: SEDAR Update Assessment. Southeast Fisheries Science Center (SEFSC).

Boosted classification trees

Kuhn, M., 2016. caret: Classification and Regression Training. R package version 6.0-71.

Natekin, A., Knoll, A., 2013. Gradient boosting machines, a tutorial. Front. Neurorobot. 7(21): 1-21.

R Core Team, 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available: http://www.R-project.org/

Ridgeway, G., 2016. gbm: Generalized Boosted classification Models. R package version 2.1.1.

Other catch-only methods

Anderson, S.C., Afflerbach, J., Cooper, A.B., Dickey-Collas, M., Jensen, O.P., Kleisner, K.M., Longo, C., Osio, G.C., Ovando, D., Minte-Vera, C., Minto, C., Mosqueira, I., Rosenberg, A.A., Selig, E.R., Thorson, J.T., Walsh, J.C., 2016. datalimited: Stock assessment methods for data-limited fisheries. R package version 0.0.2. Available at: https://github.com/datalimited/datalimited

Anderson, S.C, Cooper, A.B, Jensen, O.P., Minto, C., Thorson, J.T., Walsh, J.C., Afflerbach, J, Dickey-Collas, M. Kleisner, K.M., Longo, C., Osio, G.C., Ovando, D., Mosqueira, I., Rosenberg, A.A., Selig, E.R., 2017. Improving estimates of population status and trend with superensemble models. Fish. Fish. DOI: 10.1111/faf.12200.

Costello, C., Ovando, D., Hilborn, R., Gains, S.D., Deschenes, O., Lester, S.E., 2012. Status and solutions for the world's unassessed fisheries. Science. 338, 517-520.

Csirke, J. and Sharp, G., 1984. Proceedings of the Expert Consultation to Examine Changes in Abundance and Species Composition of Neritic Fish Resources. FAO Fisheries Report No. 291, San Jose, Costa Rica.

Froese, R., Kesner-Reyes, K., 2002. Impact of Fishing on the Abundance of Marine Species. ICES CM 2002/L:12.

Grainger, R.J.R., Garcia, S.M., 1996. Chronicles of Marine Fishery Landings (1950–1994). Trend Analysis and Fisheries Potential. FAO Fisheries Technical Paper. FAO, Rome, Italy.

Kleisner, K., Pauly, D., 2011. Stock-status plots of fisheries for Regional Seas, in: Christensen, V., Lai, S., Palomares, M.L.D., Zeller, D., Pauly, D. (Eds.), The State of Biodiversity and Fisheries in Regional Seas. Fisheries Centre Research Reports 19(3). University of British Columbia, Vancouver, Canada, pp. 37-40.

Kleisner, K., Zeller, D., Froese, R., Pauly, D., 2013. Using global catch data for inferences on the world's marine fisheries. Fish. Fish. 14(3), 293-311.

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

Rosenberg, A.A., Fogarty, M.J., Cooper, A. B., Dickey-Collas, M., Fulton, E.A., Gutiérrez, N.L., Hyde, K.J.W., Kleisner, K.M., Longo, C., Minte-Vera, C.V., Minto, C., Mosqueira, I., Osio, G.C., Ovando, D., Selig, E.R., Thorson, J.T., Ye, Y., 2014. Developing new approaches to global stock status assessment and fishery production potential of the seas. FAO Fisheries and Aquaculture Circular, Rome, Italy.

Schaefer, M.B., 1954. Some aspects of the dynamics of populations important to the management of commercial marine fisheries. Bull. Inter-Am. Trop. Tuna. Comm. 1(2), 27-56.

Thorson, J.T., Minto, C., Minte-Vera, C.V., Kleisner, K.M., Longo, C., 2013. A new role for effort dynamics in the theory of harvested populations and data-poor stock assessment. Can. J. Fish. Aquat. Sci. 70(12), 1829-1844.

Vasconcellos, M., Cochrane, K., 2005. Overview of world status of data-limited fisheries: inferences from landings statistics, in: Kruse, G.H., Gallucci, V.F., Hay, D.E., Perry, R.I., Peterman, R.M., Shirley, T.C., Spencer, P.D., Wilson, B., Woodby, D. (Eds.), Fisheries Assessment and Management in Data-Limited Situations. Alaska Sea Grant College Program, University of Alaska Fairbanks, Fairbanks, AK, pp. 1-20.

Chapter 3: Influence of ocean warming on global marine fisheries productivity

Abstract

Marine fish and invertebrates are experiencing shifting distributions, changing phenology, and altered food availability and habitat quality as a result of climate change, but the net effect of these changes on global fisheries productivity remains unclear. In this study, we use surplus production models with monotonic temperature-dependence to measure the influence of sea surface temperature (SST) on the productivity of 235 fish stocks of 124 species in 38 large marine ecoregions. We found that ocean warming has significantly positively and negatively influenced the productivity of 18 and 19 stocks, respectively (37 total; 16% total). The influence of warming on a stock's productivity is determined by ecoregion, taxonomic family, life history, and exploitation history. Hindcasts of SST-dependent maximum sustainable yield indicate that MSY of assessed stocks decreased 5.6% from 1930 to 2010. The greatest SST-driven reductions in MSY occurred in the Sea of Japan, Kuroshio Current, East China Sea, North Sea, Iberian Coastal ecoregions while the greatest SST-driven gains in MSY occurred in the South Pacific Ocean, Canary Current, Indian Ocean, and North Pacific Ocean ecoregions. The model's assumption of monotonic rather than dome-shaped temperature dependence is optimistic and suggests that future climate-driven losses in MSY may be even larger.

Introduction

The growth of the world's human population and current levels of hunger in many parts of the world have raised concerns over food security in the future (Godfray et al. 2010). Currently, fisheries and aquaculture supply about 17% of global animal protein intake and support the livelihoods of approximately 12% of the world's population (FAO 2016). Human population growth is expected to be most profound in regions where fish provide most of the non-grain dietary protein (UN-DESA 2015). The extent to which marine fisheries will be able to keep pace with an increasing human population will depend on climate-driven changes to fisheries productivity and the adaptation of fisheries management systems to these changes.

Anticipating the net effect of climate change on marine fisheries is complicated because climate change affects a multitude of environmental variables that act across different levels of biological organization (Hollowed et al. 2013). Of these variables, temperature is arguably the most important because of its direct effect on marine organisms (Pörtner & Knust 2007; Pörtner & Farrell 2008) and its role in driving changes in stratification (Manabe & Stouffer 1993), primary production (Behrenfeld et al. 2006), and dissolved oxygen (Keeling et al. 2010). As a result of ocean warming, marine fish and invertebrates are experiencing large-scale redistributions (Perry et al. 2005; Pinsky et al. 2013), changing phenology and mismatches (Cushing 1990; Edwards & Richardson 2004), altered food availability (Boyce et al. 2010, 2014), and increased exposure to oxygen-depleted and acidic waters (Mora et al. 2013). However, the net effect of these changes on marine fisheries productivity remains poorly understood.

The first global-scale studies of climate-driven fisheries productivity linked bottom-up ecological models with climate models but failed to report uncertainty and were applied to broad species groups. For example, Cheung et al. (2010) coupled species distribution, primary productivity, and trophic transfer models to project shifts in catch potential for 1,066 species under two climate scenarios. They predicted that the distribution of global catch will shift dramatically but net productivity will remain the same. Blanchard et al. (2012) coupled physical-biogeochemical and food web models to project shifts in fish production for "demersal" and "pelagic" fish groups under two climate scenarios. They also predicted a dramatic redistribution of production but avoided statements regarding changes to net productivity. Although the agreement between these two studies is compelling, they are both limited in that they impose rather than detect a link between climate and productivity and that they propagate no uncertainty in their layered models and assumptions (Brander et al. 2013).

Recent studies have focused on empirical analyses of commercially important fish stocks but have only examined the influence of temperature in post-hoc analyses and have produced mixed results. For example, Britten et al. (2016) correlated time-varying trends in the recruitment capacity of 262 assessed fish stocks to trends in SST, chlorophyll, and overfishing history, and found SST change to be a non-significant driver of changing recruitment capacity. They also suggest that global recruitment capacity has declined at a rate of 3% per decade; however, this estimate gives equal weight to all stocks, regardless of size, and Szuwalski (2016)'s reanalysis shows that weighting by biomass eliminates the trend while weighting by catch reverses the trend entirely. Britten et al. (2017) conducted a similar analysis of time-varying surplus production for 276 assessed stocks, which they conceptually but not statistically attribute to changes in environment, and found no changes in net global productivity.

In this study, we use surplus production models with monotonic temperature-dependence to measure the influence of sea surface temperature (SST) on the productivity of 235 fish stocks and hindcast changes in their SST-driven maximum sustainable yield (MSY) from 1930-2010. This study is the first retrospective analysis of assessed global fish stocks to explicitly include temperature inside a population dynamics model and the first analysis to attribute SST experience using the actual stock boundaries. Furthermore, we measure the influence of temperature on MSY, the best metric for anticipating impacts of climate change on food security and livelihoods, while other studies have focused on recruitment (Britten et al. 2016; Szuwalski 2016), which is difficult to translate into food or dollars, or on "catch potential" defined as the mean of the five largest catches (Cheung et al. 2010), which is unlikely to be a sustainable quantity in moderately to heavily exploited fisheries. Thus, our study seeks to provide new insights into the past and future impacts of climate change on global marine fisheries productivity and implications for global food security and livelihoods.

Methods

Stock selection

We analyzed the non-salmon stocks in the RAM Legacy Stock Assessment Database (RAMLDB v3.8; Ricard et al. 2012) with time series of total biomass (metric tons) and catch or landings (metric tons; catch preferred) longer than 20 years after trimming data based on strong model assumptions (Supp. Table 1; Appendix C). We included 30 stocks assessed using biomass dynamics models that were judged to exhibit sufficient variability to be included in the analysis. We excluded 28 stocks that prevented model convergence because they (1) lacked periods of low exploitation and high biomass necessary to constraining carrying capacity or (2) exhibited population dynamics wildly divergent from stationary logistic population growth. The resulting 235 stocks represent a variety of taxa, life histories, and locations and approximately 33% of reported global catch (FAO 2016; 28 of 86 million metric tons in 2000).

Stock boundary delineation and SST time series

We estimated the sea surface temperatures (SST) experienced by each stock by mapping the boundary of the stock (i.e., the spatial domain of the stock assessment) and calculating the mean annual SST within this boundary using the COBE SST dataset (COBE v2; Ishii et al. 2005). The COBE dataset provides monthly SST on a globally complete 1°x1° grid from 1850-present based on an interpolation of in-situ and satellite-derived SST observations. We conducted sensitivity analyses using the ERSST and HadISST datasets to ensure that the results were not sensitive to the choice of SST dataset (Supp. Figure 1; Appendices A & D). Stock boundaries were delineated by either (1) merging the statistical/management areas used to define the assessment area; (2) digitizing the assessment area directly from the stock assessment; or (3) clipping the managing country's exclusive economic zone or the managing agency's area of competence to the geographical reference points provided in the stock assessment. In the USA and Australia, we used information on the geographic distribution of each species (i.e., essential fish habitat and modelled distribution, respectively) to further constrain stock boundaries (Appendix B).

Base SST-linked surplus production model

We modeled the influence of SST on productivity using a Pella-Tomlinson surplus production model with normal process uncertainty and multiplicative SST influence. Surplus production was calculated for each stock as the net change in biomass in the absence of harvest:

$$SP_{i,t} = B_{i,t+1} - B_{i,t} + C_{i,t}$$

where $SP_{i,t}$ is the surplus production for stock *i* over year *t*, $B_{i,t}$ and $B_{i,t+1}$ are the biomasses of stock *i* in years *t* and *t+1*, respectively, and $C_{i,t}$ is the catch for stock *i* removed between years *t* and *t+1*. We used a Pella-Tomlinson surplus production model (Pella & Tomlinson 1969) because it contains a shape parameter (p) that allows it to replicate either the Fox ($p \rightarrow 0$) or Schaefer (p=1) surplus production models (Schaefer 1954; Fox 1970). We extended the Pella-Tomlinson model to include SST influence and assumed normal process uncertainty:

$$SP_{i,t} = \frac{r_i}{p} B_{i,t} \left(1 - \left(\frac{B_{i,t}}{K_i}\right)^p\right) * \exp\left(SST_{i,t} * \theta_i\right) + \varepsilon_i$$

where $SP_{i,t}$ is the surplus production, $B_{i,t}$ is the biomass, $SST_{i,t}$ is the mean sea surface temperature for stock *i* in year *t* and r_i is the intrinsic rate of growth, K_i is the carrying capacity, θ_i is the influence of SST on productivity, and ε_i is normal process uncertainty, $N(0, \sigma^2_{P,i})$, for stock *i*. We used Akaike Information Criterion (AIC; Akaike 1974) to compare models using shape parameters (p) that maximize productivity at 50% (p=1), 45% (p=0.55), 40% (p=0.2), and 37% (p=0.01) of carrying capacity and selected the model with the lowest AIC score as the "base" model (Table 1). We evaluated these shape parameter values because 50% produces the symmetric Schaefer model, 40% is the meta-analytic mean (Thorson et al. 2012), and 37% is the asymptotic limit of this parameterization of the Pella-Tomlinson model.

We estimated SST influences, θ_i, as random effects:

$$\theta_i \sim N(\mu_{SST}, \sigma^2_{SST})$$

where μ_{SST} and σ_{SST} are the mean and standard deviation of the global distribution of SST influences (θ_i), respectively. θ_i < 0 means increasing SST reduces productivity at a given biomass and θ_i > 0 means increasing SST magnifies productivity at a given biomass. See Supp. Table 2 for a key to all model symbols.

To ease model fitting, we scaled biomass and production data to each stock's maximum biomass and centered SST data around each stock's mean SST. We also placed a likelihood penalty on carrying capacities greater than five times the observed maximum biomass to constrain unrealistic carrying capacities. We fit the model using maximum likelihood estimation in the *TMB* package (Kristensen et al. 2016; Template Model Builder) in R (R Core Team 2017).

Alternative SST-linked surplus production models

To determine whether taxonomy, geography, or stock assessment method structure SST influence, we used SST-linked surplus production models with hierarchical SST influence based on each of six "groups": (a) taxonomic order and family; (b) large marine ecoregion (LME; Spalding et al. 2007) and FAO major fishing area; and (c) generic and specific stock assessment method (Table 1; Supp. Table 3). These models were identical to the base model except that SST influence is estimated as a nested hierarchical random effect:

$$\theta_i \sim N(\mu_{G,j}, \sigma_G^2)$$

where SST influences (θ_i) for stocks in group *j* are drawn from a normal distribution with a group-specific mean ($\mu_{G,j}$) and group-wide standard deviation (σ_G). Group-specific means are drawn from a global normal distribution with mean (μ_{SST}) and standard deviation (σ_{SST}):

$$\mu_{G,j} \sim N(\mu_{SST}, \sigma_{SST}^2)$$

We compared the group models to the base model using AIC and judged a group to be a significant driver of SST influence if its model exhibited an AIC score more than two points lower the base model. The best or "final" SST-linked surplus production model was identified as the model producing the lowest AIC score.

We explored using SST-linked surplus production models with dome-shaped temperature dependence but these models failed to converge due to their inability to estimate species-specific thermal optima (see Appendix A for more details).

Model validation and simulation testing

We tested whether the final SST-linked surplus production model described population dynamics better than the standard surplus production model by competing the models with AIC. We tested whether the results of the final model were an artefact of model structure by decoupling the SST and productivity time series using three null models with simulated SST time series exhibiting: (1) the same mean, variance, autoregressive properties, and trend as the original time series; (2) the same mean, variance, and autoregressive properties as the original time series but without a trend; and (3) the same mean and variance as the original time series but without autocorrelation or a trend (Supp. Figure 3; Appendices A & F). The SST simulations were performed using the R package *forecast* (Hyndman 2017).

Data analysis and hindcasting global MSY

Because the influence of SST on productivity was estimated as a random effect, our estimates of SST influence cannot be considered independent and cannot undergo post-hoc analyses using formal statistical methods (i.e., formal hypothesis testing requires including explanatory variables inside the model as we did with taxonomy and geography). Therefore, we graphically evaluated whether SST influence is determined by: (1) <u>life history traits</u> such as growth rate, maximum age, and depth preference; (2) <u>stock characteristics</u> such as trend in biomass and fishing pressure; and (3) <u>thermal experience</u> such as mean SST, SST trend, or latitude. A list of evaluated explanatory variables and their sources is provided in Supp. Table 4.

We used the final model's estimates of p, r_i, K_i, and θ_i to hindcast SST-dependent maximum sustainable yield (MSY) from 1930-2010. We calculated MSY for stock i in year t as:

$$MSY_{i,t} = \frac{r_i * k_i}{(p+1)^{(p+1)/p}} * \exp(\hat{\theta}_i * \overline{SST}_{i,t})$$

where $\overline{SST}_{i,t}$ is $SST_{i,t}$ centered on the mean of the SST data used in model fitting and $\hat{\theta}_i$ is randomly drawn from a normal distribution described by the mean θ_i estimate and its standard error. We bootstrapped 10,000 MSY hindcasts for each stock to generate median MSY trends and confidence intervals. We assessed changes in MSY over the hindcast period using (1) Thiel-Sen regression slopes and (2) percent change in mean MSY from 1930-39 to 2001-2010. Theil-Sen regression, a form of robust regression, identifies the median slope of lines through all possible point pairs and is insensitive to outliers and endpoints in short time series. We limited the hindcast from 1930-2010 to minimize the extrapolation of MSY predictions to temperatures cooler or warmer than those used in model fitting (Supp. Figure 4) and explored the sensitivity of measures of MSY change to the selection of hindcast window (Supp. Figure 5).

Results

The SST-linked Schaefer surplus production model described population dynamics better than the standard Schafer surplus production model based on AIC (Table 1). The SST-linked Pella-Tomlinson production model with productivity maximized at 40% of carrying capacity described populated dynamics better than the SST-Schaefer model and Pella-Tomlinson models with other shape parameters and was selected as the "base" model (Table 1; Supp. Figure 2). Estimates of SST influence were not sensitive to choice of shape parameter (Supp. Figure 2). SST-linked models with hierarchy by large marine ecoregion, taxonomic family, and FAO major fishing area further improved model fit (Table 1; Supp. Figures 6&7). The SST-linked model with hierarchy by ecoregion, selected as the final model, estimated a wider range of SST influences and at a higher rate of significance than the three null models (Figure 1; Supp. Figures 9-12).

Although the global mean of the SST influence distribution was not significantly different from zero, the productivity of 18 and 19 stocks were estimated to be significantly positively and negatively influenced by warming, respectively (Figure 1; Supp. Table 5). In the final model, these influences were structured by ecoregion, with the Celtic-Biscay Shelf and North Sea showing significantly negative mean SST influences (Figure 2). The FAO area and taxonomic family models showed significantly negative mean SST influences for the Northeast Atlantic and Gadid family (codfishes), respectively (Figure 2). Fish with fast life histories (<20 yr max age) were especially sensitive, both positively and negatively, to warming and fish residing in deep water (>600 m) were particularly insensitive (Figure 3). Stocks experiencing intense overfishing and declining biomass were more likely to be negatively influenced by warming (Figure 3). The position of a stock within its species-specific thermal niche may also determine the influence of warming: Atlantic herring, Atlantic cod, and red rock lobster all showed negative relationships between SST influence and mean temperature experience (Figure 4).

Habitat, trophic level, latitude, and stock size did not appear to structure SST influence (Supp. Figures 8-11). Stock assessment method did not influence model results (Table 1; Supp. Figure 6-8).

Hindcasts of SST-dependent maximum sustainable yield indicate that MSY of assessed stocks decreased 5.6% (2.2 million metric tons) from an average of 41.7 million metric tons in 1930-39 to an average of 39.5 million metric tons in 2001-10 (Figure 5). At the LME-scale, SST-driven changes in MSY generally mirrored the mean SST influence of the LME, though change in large stocks sometimes neutralized or overrode the changes of many small stocks (e.g., on the SE US Cont. Shelf; Figure 6). The greatest SST-driven reductions in MSY occurred in the Sea of Japan, Kuroshio Current, East China Sea, North Sea, Iberian Coastal ecoregions while the greatest SST-driven gains in MSY occurred in the South Pacific Ocean, Canary Current, Indian Ocean, and North Pacific Ocean ecoregions (Figure 6; Supp. Table 6). The final model's estimates of MSY at average temperature are highly correlated with data-rich estimates and only 12.0% of stock-years between 1930-2010 required extrapolating outside temperatures seen by the model (Supp. Figure 3; 3.5% warmer, 8.5% cooler).

Discussion

This is the first study to show that climate change has resulted in a net loss in global marine fisheries productivity. This finding contradicts analyses of time-varying recruitment (Szuwalksi 2016) and productivity (Britten et al. 2017) that suggest that net productivity has not changed despite large-scale redistributions.

However, these studies were based on analyses of shifts in productivity over time series starting, on average, in 1971. Although these time series capture the period of ocean warming from 1970 to present (0.5°C of warming), they miss the period of warming from 1910-1945 (Karl et al. 2015), which limits their ability to describe long-term, climate-driven shifts in productivity. Our hindcasts of SST-dependent MSY extend back to 1930 and document a 5.6% decrease in marine fisheries productivity over 0.6°C of ocean warming. Limiting the hindcasts to 1970 would result in a dampened 3.3% decrease.

Even this decline in productivity is likely to be optimistic given our assumption of a monotonic influence of warming on production (i.e., warming is only good or only bad for productivity). In fact, the aerobic performance of individual fish is dome-shaped with regards to temperature (Pörtner & Knust 2007) and is likely to remain dome-shaped at the population-scale through cascading impacts on growth, mortality, and recruitment (Drinkwater et al. 2010). Thus, the stocks identified by our model as having been positively influenced by warming are unlikely to maintain productivity gains as continued warming (IPCC 2013; Raftery et al. 2017) drives these stocks past their thermal optima. Unfortunately, estimating thermal optima proved impossible given the lack of SST contrast when averaging SST experience over large spatial scales. Although our monotonic model cannot forecast productivity under out-of-sample warming, optimistic hindcasts imply that future losses in productivity will be greater than 12%.

Our results support suggestions that the impact of climate change on marine fisheries productivity varies among ecoregions (Blanchard et al. 2012; Britten et al. 2016) but presents a new map of climate “winners” and “losers”. Britten et al. (2016) and our estimates of ecoregion-scale SST influence are uncorrelated and in low agreement (**Supp. Figure 17**). On one hand, this is not surprising given that Britten et al. (2016) found no relationship between trend in ecoregion recruitment capacity and trend in ecoregion SST while we model this relationship explicitly. On the other hand, this is surprising given that we analyze the same data with conceptually similar methods. The stark differences between the results could indicate that (1) climate-induced effects on somatic growth and mortality are strong enough to yield divergent effects on production and recruitment or (2) Britten et al. (2016)’s recruitment potential trends are highly sensitive to the state of the initial or final recruitment regime (Szuwalksi 2016).

We also identify taxonomic family as an important driver of the influence of SST on productivity. However, taxonomy is collinear with geography and these results should be interpreted carefully. For example, all of the negatively influenced sandeel (Ammodytidae) stocks are found in the negatively influenced North Sea and most of the negatively influenced codfish (Gadidae) stocks are found in the negatively influenced Northeast Atlantic. Do environmental regime shifts in these regions drive the apparent response of these taxa or does the intrinsic vulnerability of these taxa drive the apparent response of these regions? Our methods cannot disentangle this complexity but finer-scale analyses suggest both are possible. The Northeast Atlantic has undergone large climate-driven shifts in primary productivity with cascading food web effects (Richardson & Schoeman 2004). This has been especially true in the North Sea where both forage fish (Clausen et al. 2017) and groundfish (Beaugrand et al. 2003) productivity have been reduced as a result of climate-induced changes to the zooplankton community. Alternatively, both Atlantic cod (Planque & Frédou 1999) and sandeel (Arnott & Ruxton 2002) recruitment are negatively correlated with warming temperatures.

We also present new evidence suggesting that fish with fast life histories (e.g., fast growth, early age at maturity, short lifespan, etc.) are more responsive to climate change than fish with slow life histories and that overfishing makes stocks more vulnerable to climate change. For example, Perry et al. (2005) showed that North Sea fish species shifting distributions in response to warming temperatures were smaller and matured earlier than non-shifting species. Similarly, we identified steep declines in the magnitude and significance of the influence of warming on productivity at 150 cm max length and 7 years old at maturity. We also show that stocks experiencing chronic and acute overfishing (F/F_{MSY} mean > 2) are significantly more likely to be negatively influenced by ocean warming. These results are consistent with the growing body of evidence that overfishing can magnify fluctuations due to environmental variability (Hsieh et al. 2006), reduce resilience to climate change (Planque et al. 2010), and interact with life history and climate variability to magnify sensitivity (Pinsky & Byler 2015).

Our results offer no support for hypotheses that demersal and pelagic species differ in their vulnerability to climate change but offers support for hypotheses that vulnerability to climate change varies by depth and position of a stock within its thermal niche. Rijnsdorp et al. (2009) hypothesize that pelagic species are more responsive to warming than demersal species due to higher mobility and lower fidelity and present evidence for this in the Northeast Atlantic. We found no evidence of this pattern on the global scale possibly because fish are rarely only pelagic or only demersal throughout their complex life histories and their climate-driven productivity is shaped by many life stages (Petitgas et al. 2013). We show support for Rijnsdorp et al. (2009)'s hypothesis that deep-water species are less sensitive to climate change due to more stable environmental conditions. However, these results could be spurious given our use of surface temperatures to describe the temperature experience of deep-water species despite evidence that trends in temperature and their impacts on fish are often depth-mediated (Thresher et al. 2007). Our results suggest that for species, like Atlantic cod, Atlantic herring, and red rock lobster, stocks at the warm end of their thermal range are more vulnerable to warming than stocks at the cool end of their thermal range. This pattern has been demonstrated for Atlantic cod recruitment (Drinkwater et al. 2005) but has rarely been demonstrated for other species or measures of productivity.

This study offers several advantages over other global-scale studies of climate-driven marine fisheries productivity: (1) it is the first analysis to attribute stock SST experience using the actual stock boundaries; (2) it is the first analysis to measure the influence of temperature on maximum sustainable yield, the best metric for anticipating impacts of climate change on food security and livelihoods; and (3) it is the first retrospective analysis of assessed global fish stocks to explicitly include temperature inside a population dynamics model. Our analysis also has several limitations. Although estimating dome-shaped SST dependence and allowing ocean warming to influence carrying capacity would both increase biologically realism, they were statistically infeasible. Furthermore, our study only evaluated the influence of changing SST on fisheries productivity when changing primary production (i.e., chlorophyll), dissolved oxygen, and pH are also influential (Sumaila et al. 2011). This was necessary given the need for long time series describing periods of low exploitation and high biomass when fitting surplus production models and the lack of globally complete, historic datasets for other environmental variables. Finally, the RAM Legacy Database presents a limited (e.g., v3.8=48% of reported catch in 2000) and non-random selection of global stocks. Although this analysis is representative of the dynamics of assessed stocks, it is not representative of global fisheries production.

This paper presents a sobering reality. As the world's human population and demand for seafood grows (FAO 2016), climate change is a driving a decline in marine fisheries productivity and sustainable catch potential. This study indicates that fisheries managers will need to adjust expectations as they begin ecosystem-based fisheries management seeking to manage fisheries in the face of climate change. Importantly, this study highlights the emerging fact that overfishing exacerbates the vulnerability of fish stocks to climate change. Thus, preventing overfishing is imperative as climate change extends recovery timelines (Britten et al. 2017) and developing stock assessment methods that account for reductions in productivity is essential as fish, fishermen, and fisheries managers move into a warmer world.

Acknowledgements

We are grateful for the hard work of the many scientists who have contributed to the RAM Legacy Stock Assessment Database and to C. Hofmann, S. Coffen-Smout, D. Keith, S. MacConnachie, and P. Neubauer for their contributions to the corresponding stock boundary database. We also thank M. Roswell and the members of the Jensen and Pinsky labs for valuable discussions and feedback while working on this paper.

References

Akaike, H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6): 716-723.

Anderson, C.N., Hsieh, C.H., Sandin, S.A., Hewitt, R., Hollowed, A., Beddington, J., May, R.M., Sugihara, G. (2008) Why fishing magnifies fluctuations in fish abundance. *Nature* 452(7189): 835-839.

Arnott, S.A., Ruxton, G.D. (2002) Sandeel recruitment in the North Sea: demographic, climatic and trophic effects. *Marine Ecology Progress Series* 238: 199-210.

Beaugrand, G., Brander, K.M., Lindley, J.A., Souissi, S., Reid, P.C. (2003) Plankton effect on cod recruitment in the North Sea. *Nature* 426(6967): 661-664.

Behrenfeld, M.J., O'Malley, R.T., Siegel, D.A., McClain, C.R., Sarmiento, J.L., Feldman, G.C., Milligan, A.J., Falkowski, P.G., Letelier, R.M., Boss, E.S. (2006) Climate-driven trends in contemporary ocean productivity. *Nature* 444(7120): 752-755.

Blanchard, J.L., Jennings, S., Holmes, R., Harle, J., Merino, G., Allen, J.I., Holt, J., Dulvy, N.K., Barange, M. (2012) Potential consequences of climate change for primary production and fish production in large marine ecosystems. *Philosophical Transactions of the Royal Society B* 367(1605): 2979-2989.

Boyce, D., Lewis, M., Worm, B. (2010) Global phytoplankton decline over the past century. *Nature* 466: 591-596.

Boyce, D.G., Dowd, M., Lewis, M.R., Worm, B. (2014) Estimating global chlorophyll changes over the past century. *Progress in Oceanography* 122(1): 163-173.

Brander, K., Neuheimer, A., Andersen, K.H., Hartvig, M. (2013) Overconfidence in model projections. *ICES Journal of Marine Science* 70(6): 1065-1068.

Britten, G.L., Dowd, M., Kanary, L., Worm, B. (2017) Extended fisheries recovery timelines in a changing environment. *Nature Communications* 8: 15325.

Britten, G.L., Dowd, M., Worm, B. (2016) Changing recruitment capacity in global fish stocks. *Proceedings of the National Academy of Sciences* 113(1): 134-139.

Cheung, W.W., Lam, V.W., Sarmiento, J.L., Kearney, K., Watson, R.E.G., Zeller, D., Pauly, D. (2010) Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. *Global Change Biology* 16(1): 24-35.

Clausen, L.W., Rindorf, A., Deurs, M., Dickey-Collas, M., Hintzen, N.T. (2017) Shifts in North Sea forage fish productivity and potential fisheries yield. *Journal of Applied Ecology*: 0.1111/1365-2664.13038.

Cushing, D. (1990) Plankton production and year-class strength in fish populations: an update of the match/mismatch hypothesis. *Advances in Marine Biology* 26: 249-293.

Drinkwater, K.F., Beaugrand, G., Kaeriyama, M., Kim, S., Ottersen, G., Perry, R.I., Pörtner, H.O., Polovina, J.J., Takasuka, A. (2010) On the processes linking climate to ecosystem changes. *Journal of Marine Systems* 79: 374-388.

Edwards, M., Richardson, A.J. (2004) Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature* 430: 881–884.

FAO (2016) The State of World Fisheries and Aquaculture 2016: Contributing to food security and nutrition for all. Food & Agriculture Organization (FAO): Rome.

Fox, W.W., Jr. (1970) An exponential surplus-yield model for optimizing exploited fish populations. *Transactions of the American Fisheries Society* 99(1): 80-88.

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., Toulmin, C. (2010) Food security: the challenge of feeding 9 billion people. *Science* 327(5967): 812-818.

Hollowed, A.B., Barange, M., Beamish, R.J., Brander, K., Cochrane, K., Drinkwater, K., Foreman, M.G., Hare, J.A., Holt, J., Ito, S.I., Kim, S. (2013) Projected impacts of climate change on marine fish and fisheries. *ICES Journal of Marine Science* 70(5): 1023-1037.

Hsieh, C-H, Reiss, C.S., Hunter, J.R., Beddington, J.R., May, R.M., Sugihara, G. (2006) Fishing elevates variability in the abundance of exploited species. *Nature* 443: 859-862.

Hyndman, R.J. (2017) forecast: Forecasting functions for time series and linear models. R package version 8.2. Available at: http://pkg.robjhyndman.com/forecast

IPCC (2013) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Editors: Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley. Cambridge University Press: Cambridge, UK and New York, NY.

Ishii, M., Shouji, A., Sugimoto, S., Matsumoto, T. (2005) Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe collection. *International Journal of Climatology* 25: 865-879.

Karl, T.R., Arguez, A., Huang, B., Lawrimore, J.H., McMahon, J.R., Menne, M.J., Peterson, T.C., Vose, R.S., Zhang, H.M. (2015) Possible artifacts of data biases in the recent global surface warming hiatus. *Science*: 348(6242): 1469-1472

Keeling, R. F., Körtzinger, A., Gruber, N. (2010) Ocean deoxygenation in a warming world. *Annual Reviews in Marine Science* 2: 199-229.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., Bell, B.M. (2016) TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software* 70(5): 1-21.

Manabe, S., Stouffer, R. J. (1993) Century-scale effects of increased atmospheric CO_2 on the ocean-atmosphere system. *Nature* 364: 215-218.

Mora, C., Wei, C.L., Rollo, A., Amaro, T., Baco, A.R., Billett, D., Bopp, L., Chen, Q., Collier, M., Danovaro, R., Gooday, A.J. (2013) Biotic and human vulnerability to projected changes in ocean biogeochemistry over the 21st century. *PLoS Biology* 11(10): e1001682.

Pella, J. J., Tomlinson, P.K. (1969) A generalized stock production model. *Bulletin of the Inter-American Tropical Tuna Commission* 3: 416-497.

Perry, A.L., Low, P.J., Ellis, J.R., Reynolds, J.D. (2005) Climate change and distribution shifts in marine fishes. *Science* 308(5730): 1912-1915.

Petitgas, P., Rijnsdorp, A.D., Dickey-Collas, M., Engelhard, G.H., Peck, M.A., Pinnegar, J.K., Drinkwater, K., Huret, M., Nash, R.D. (2013) Impacts of climate change on the complex life cycles of fish. *Fisheries Oceanography* 22(2): 121-139.

Pinsky, M.L., Byler, D. (2015) Fishing, fast growth and climate variability increase the risk of collapse. *Proceedings of the Royal Society B* 282(1813): 20151053.

Planque, B., Fromentin, J-M, Cury, P., Drinkwater, K.F., Jennings, S., Perry, R.I., Kifani, S. (2010) How does fishing alter marine populations and ecosystems sensitivity to climate? *Journal of Marine Systems* 79: 403-417.

Planque, B., Frédou, T. (1999) Temperature and the recruitment of Atlantic cod (*Gadus morhua*). *Canadian Journal of Fisheries and Aquatic Sciences* 56(11): 2069-2077.

Pörtner, H.O., Farrell, A.P. (2008) Physiology and climate change. *Science* 322(5902): 690-692.

Pörtner, H.O., Knust, R. (2007) Climate change affects marine fishes through the oxygen limitation of thermal tolerance. *Science* 315(5808): 95-97.

R Core Team (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: http:// www.R-project.org/

Raftery, A.E., Zimmer, A., Frierson, D.M., Startz, R., Liu, P. (2017) Less than 2°C warming by 2100 unlikely. *Nature Climate Change* 7(9): 637-641.

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(4): 380-398.

Richardson, A.J., Schoeman, D.S. (2004) Climate impact on plankton ecosystems in the Northeast Atlantic. *Science* 305(5690): 1609-1612.

Rijnsdorp, A.D., Peck, M.A., Engelhard, G.H., Möllmann, C., Pinnegar, J.K. (2009) Resolving the effect of climate change on fish populations. *ICES Journal of Marine Science* 66: 1570-1583.

Schaefer, M.B. (1954) Some aspects of the dynamics of populations important to the management of commercial marine fisheries. *Bulletin of the Inter-American Tropical Tuna Commission* 1(2): 27-56.

Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdana, Z.A., Finlayson, M.A.X., Halpern, B.S., Jorge, M.A., Lombana, A.L., Lourie, S.A., Martin, K.D. (2007) Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *BioScience* 57(7): 573-583.

Sumaila, U.R., Cheung, W.W., Lam, V.W., Pauly, D., Herrick, S. (2011) Climate change impacts on the biophysics and economics of world fisheries. *Nature Climate Change* 1(9): 449-456.

Szuwalski, C.S. (2016) Changing fisheries productivity and food security. *Proceedings of the National Academy of Sciences* 113(13): E1773-E1774.

Thresher, R.E., Koslow, J.A., Morison, A.K., Smith, D.C. (2007) Depth-mediated reversal of the effects of climate change on long-term growth rates of exploited marine fish. *Proceedings of the National Academy of Sciences* 104(18): 7461-7465.

UN-DESA (2015) World Urbanization Prospects: The 2014 Revision (ST/ESA/SER.A/366). United Nations, Department of Economic and Social Affairs, Population Division (UN-DESA): New York.

Tables & Figures

Table 1. AIC of candidate surplus production models (PT=Pella-Tomlinson).

Model	K	Likelihood	AIC	ΔAIC
Question 1. Does SST matter?				
SST-linked Schaefer model	707	-16537.7	-31661.5	0.0
Schaefer model **(standard model)**	705	-16444.8	-31479.5	182.0
Question 2. Does asymmetry matter?				
SST-linked PT model (MSY@40%K) **(base model)**	707	-16634.6	-31855.2	0.0
SST-linked PT model (MSY@45%K)	707	-16609.9	-31805.8	49.4
SST-linked PT model (MSY@37%K)	707	-16601.2	-31788.3	66.9
SST-linked Schaefer model	707	-16537.7	-31661.5	193.7
Question 3. Does taxonomy, geography, or assessment method matter?				
SST-linked PT model w/ hierarchy by LME **(final model)**	708	-16639.4	-31862.8	0.0
SST-linked PT model w/ hierarchy by family	708	-16637.2	-31858.5	4.3
SST-linked PT model w/ hierarchy by FAO area	708	-16636.9	-31857.9	4.9
SST-linked PT model w/ hierarchy by order	708	-16635.7	-31855.5	7.3
SST-linked PT model **(base model)**	707	-16634.6	-31855.2	7.6
SST-linked PT model w/ hierarchy by assessment method (specific)	708	-16635.2	-31854.5	8.3
SST-linked PT model w/ hierarchy by assessment method (generic)	708	-16634.6	-31853.2	9.6

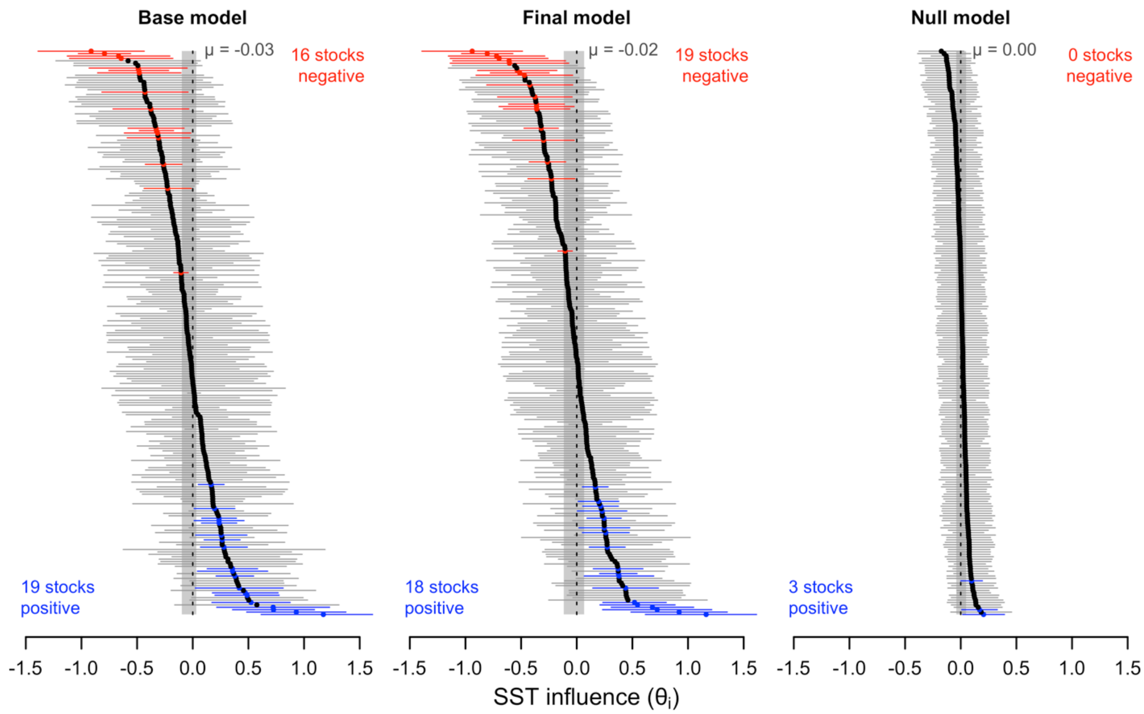


Figure 1. Distribution of SST influences estimated by the **(A)** base, **(B)** final, and **(C)** primary null models. Points show mean estimates and error bars show 95% confidence intervals. Significant positive and negative SST influences are shown in blue and red, respectively. The transparent rectangle indicates the 95% confidence interval for the global mean of the SST influences.

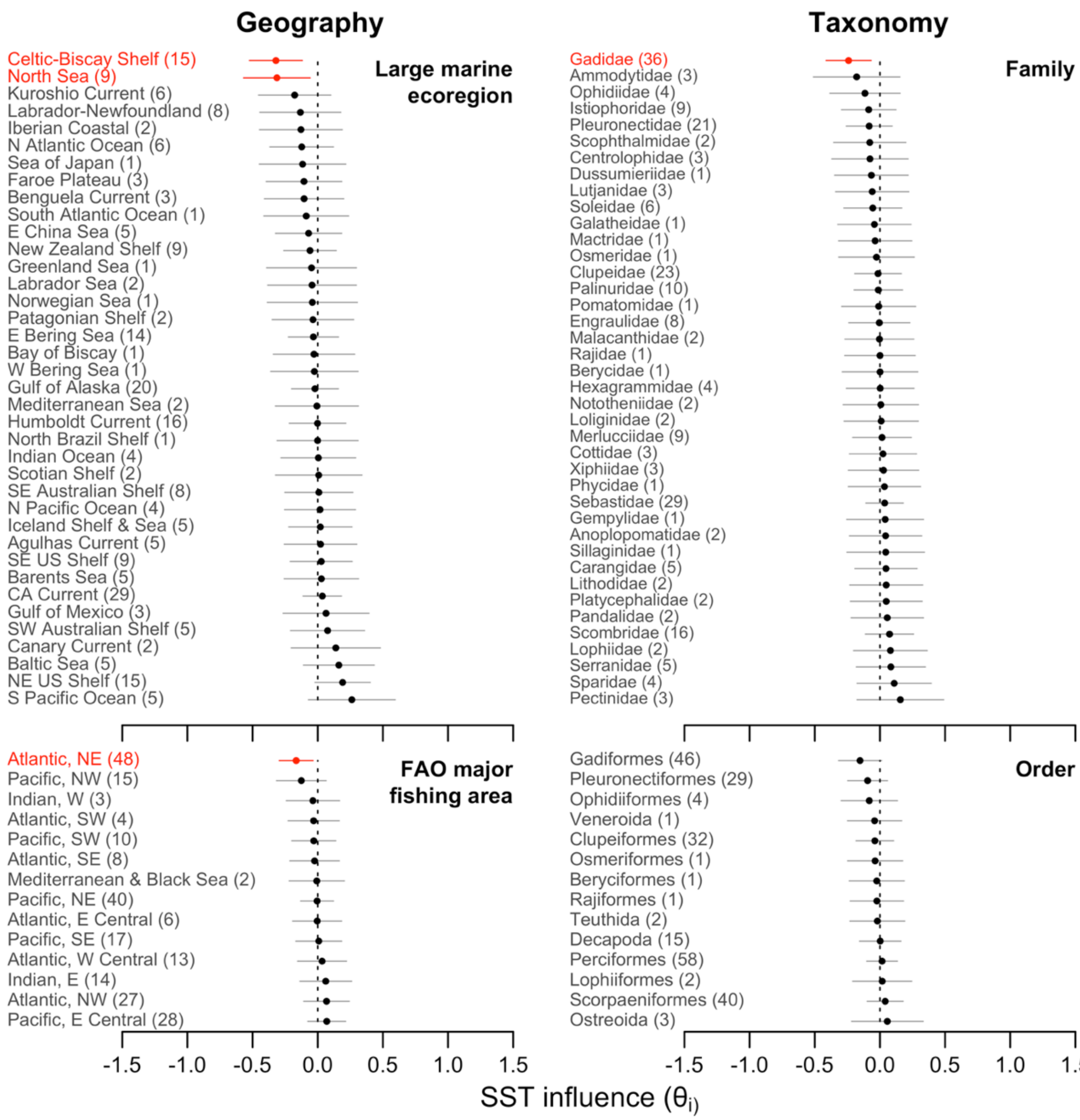


Figure 2. Mean of the SST influence distributions for geographic or taxonomic groups in models with hierarchy on SST influence by **(A)** large marine ecoregion, **(B)** FAO major fishing area, **(C)** taxonomic family, and **(D)** taxonomic order. Points show mean estimates and error bars show 95% confidence intervals. Significant positive and negative SST influences are shown in blue and red, respectively. All but the taxonomic order model had more support than the base model.

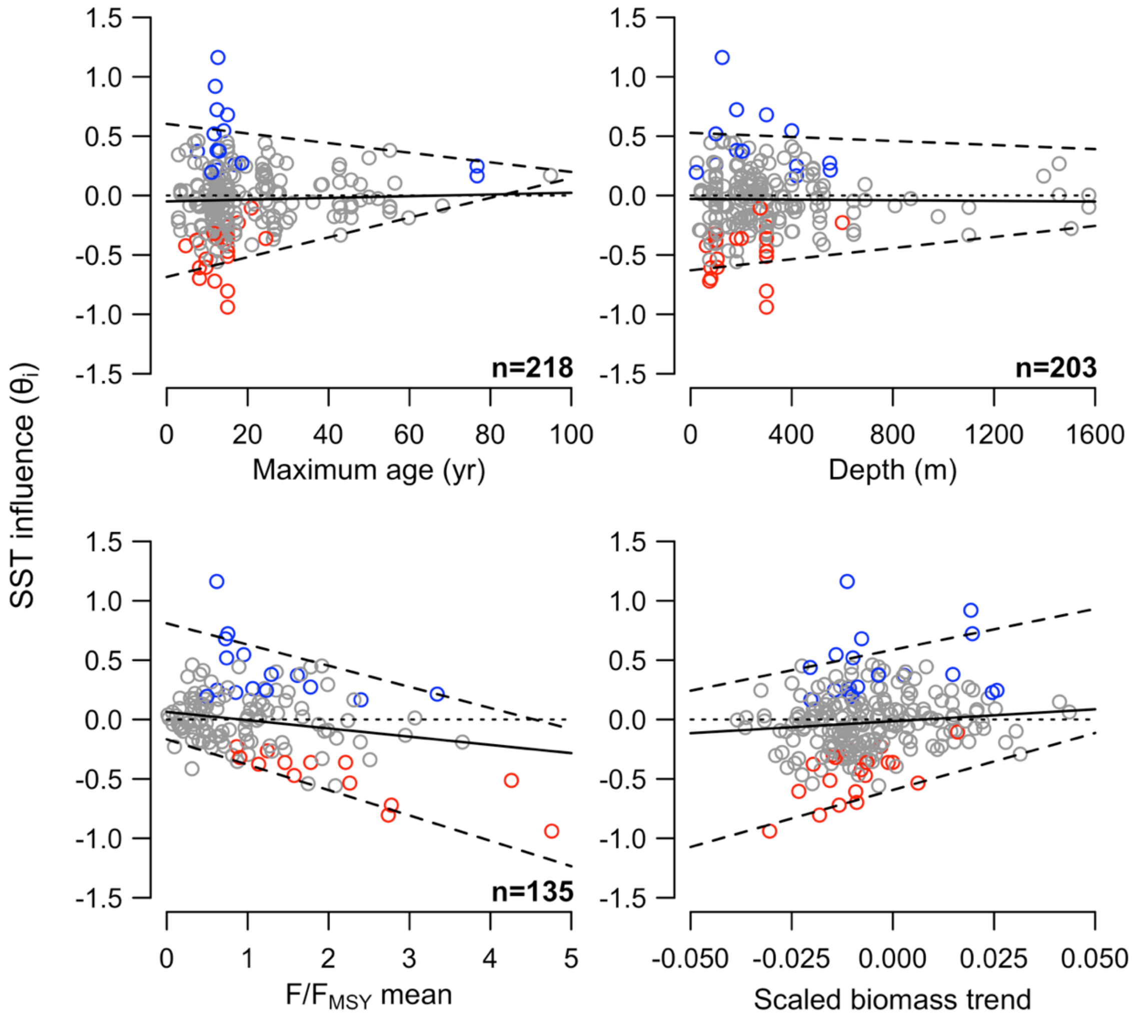


Figure 3. SST influence as a function of maximum age (T_{max}), depth, F/F_{MSY} mean, and scaled biomass trend. SST influences are colored by significance (blue=positive, red=negative, grey=non-significant). Solid lines show the 50th percentile quantile regression fit and dashed lines show the 2.5% and 97.5% quantile regression fits. Sample size is shown in the bottom-right corner if data were not available for all 235 stocks.

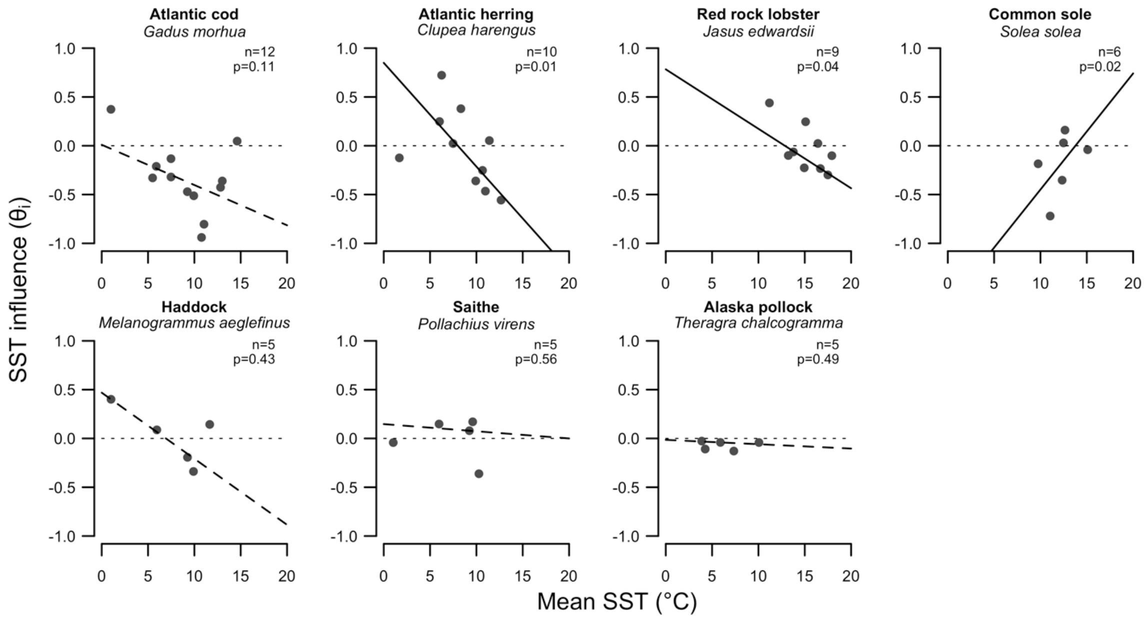


Figure 4. SST influence as a function of the mean temperature experienced by stocks of the same species for the seven species with ≥5 stocks in the analysis. Lines shows Theil-Sen regression fits with solid lines indicating regressions significant at the 0.10 level. Theil-Sen regression, a form of robust regression, identifies the median slope of lines through all possible point pairs and is insensitive to outliers and endpoints in small datasets.

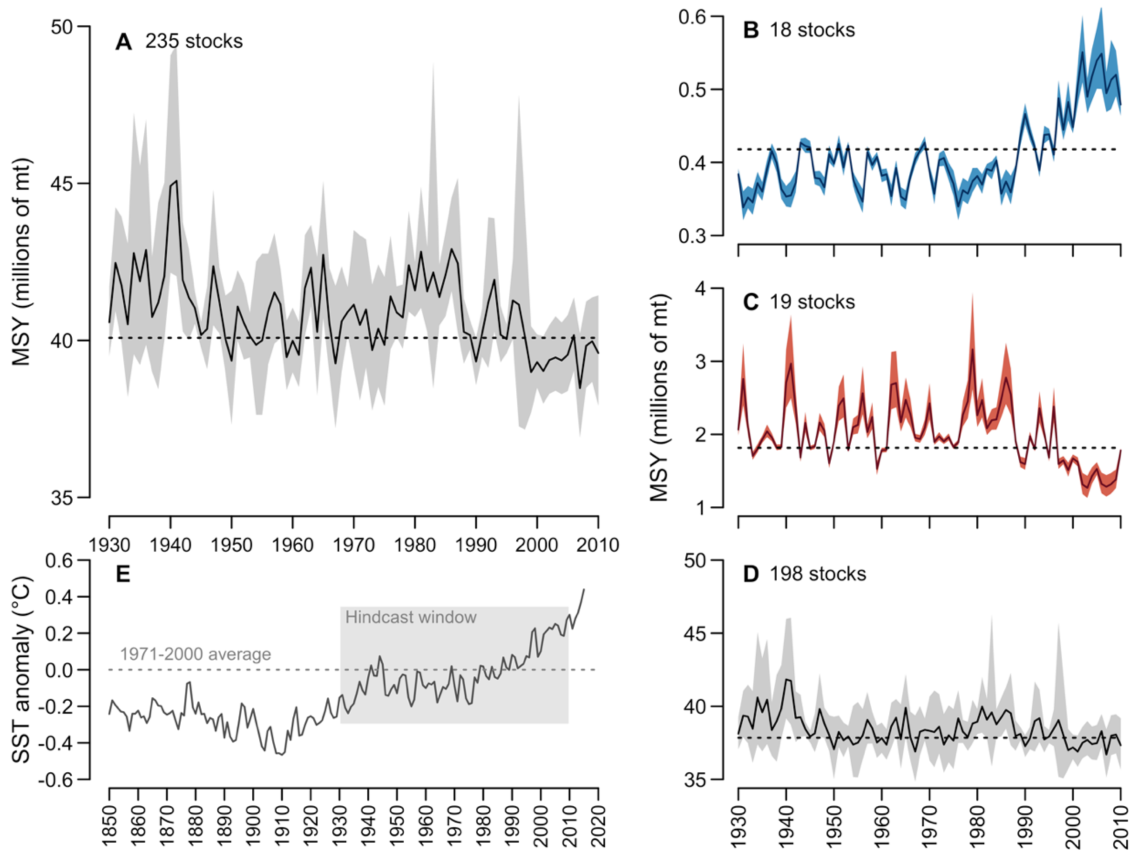


Figure 5. Hindcast of SST-dependent maximum sustainable yield (MSY) for **(A)** all stocks and for stock with **(B)** significant positive, **(C)** significant negative, and **(D)** non-significant influences of SST on productivity. Solid lines indicate the median MSY estimates, shading indicates the 95% confidence intervals, and horizontal dashed lines indicate the temperature-independent MSYs. Panel **(E)** shows the mean global SST anomaly from 1850-2015 based on the COBE dataset.

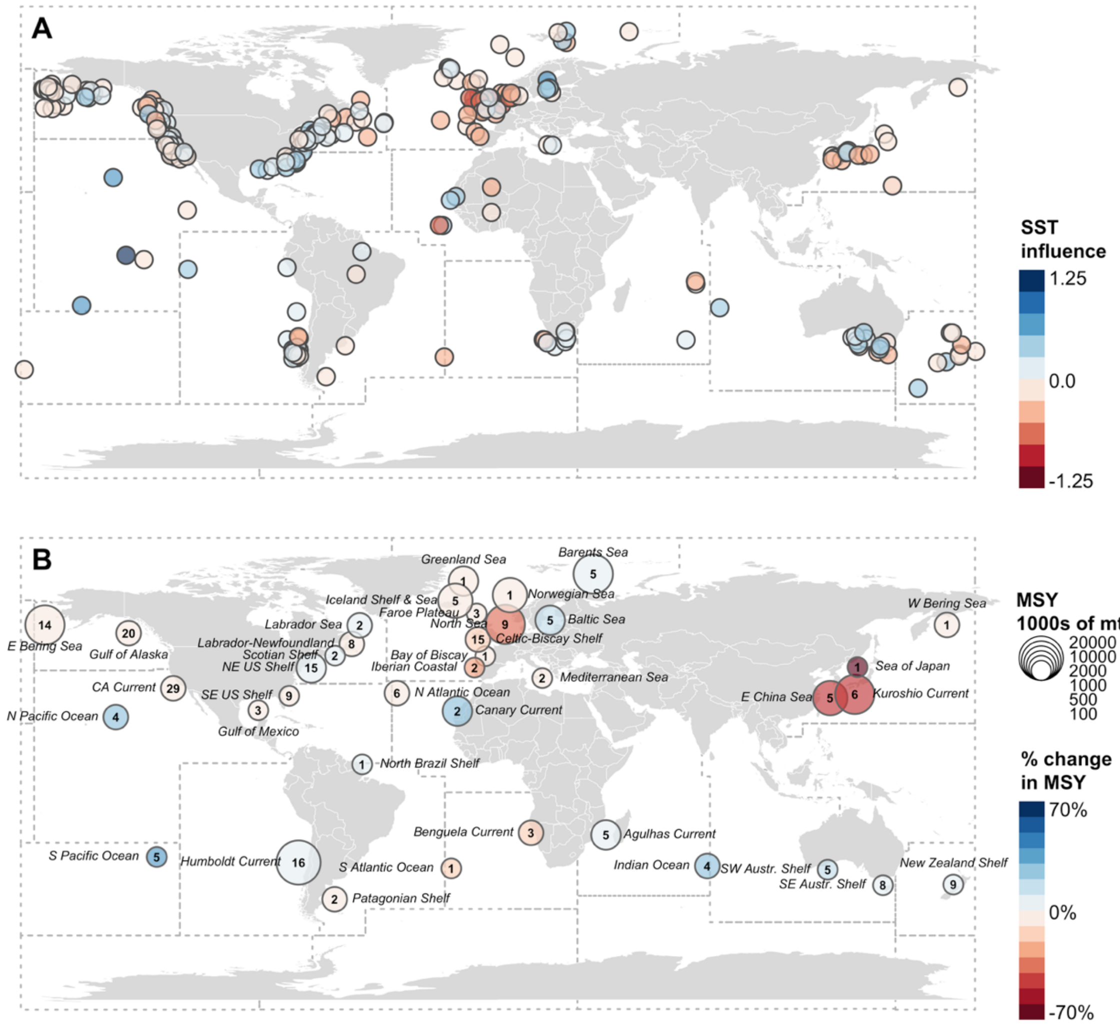


Figure 6. Maps showing the **(A)** global distribution of SST influences and **(B)** percent change in mean maximum sustainable yield (MSY) from 1930-39 to 2001-10 by ecoregion. In **(B)**, points are scaled to the 1930-39 mean and the number of stocks in the ecoregion is shown inside each point. In both plots, dashed lines indicate FAO major fishing areas.

Supplemental Tables & Figures

Supp. Table 1. RAM Legacy Database stocks used in analysis (TB = total biomass).

Condition	# of stocks
All RAMLDB stocks	1058
Not Pacific salmon stocks	685
Only stocks with TB/catch in metric tons	350
Only stocks with TB/catch time series ≥ 20 years	300
Removed 23 stocks with strong SP/SR relationships	277
Removed 9 stocks without 20 years of data after trimming	268
Removed 5 stocks without SST data (e.g., Seto Sea not covered by COBE)	263
Removed 28 stocks preventing model convergence	235

Supp. Table 2. Model symbols and their definitions.

Type	Symbol	Definition
Data	$C_{i,t}$	Catch for stock *i* in year *t*
Data	$SP_{i,t}$	Surplus production for stock *i* in year *t*
Data	$B_{i,t}$	Total biomass for stock *i* in year *t*
Data	$SST_{i,t}$	Sea surface temperature (SST) experienced by stock *i* in year *t*
Data	G_i	Group (taxonomic or geographic) for stock *i*
Derived	ε_i	Productivity process uncertainty for stock *i*
Parameter	r_i	Intrinsic rate of growth for stock *i*
Parameter	K_i	Carrying capacity for stock *i*
Parameter	θ_i	Influence of SST on productivity for stock *i*
Parameter	μ_{SST}	Mean of the distribution of SST influences (θ_i)
Parameter	σ_{SST}	Standard deviation of the distribution of SST influences (θ_i)
Parameter	$\mu_{G,j}$	Mean of the distribution of SST influences (θ_i) for group *j*
Parameter	σ_G	Standard deviation of the group-specific distributions of SST influences (θ_i)
Parameter	$\sigma_{P,i}$	Standard deviation of the productivity process uncertainty for stock *i*
Constant	p	Shape parameter: fixed at 1.00, 0.55, 0.20, or 0.01
Index	t	Year
Index	i	Stock
Index	j	Group (taxonomic or geographic)

Supp. Table 3. Stock assessment methods.

Assessment model	Number	Countries
Biomass dynamics model (n=30)		
BSPM: Bayesian surplus production model	10	Canada, Tuna-RFMO
ASPIC: Surplus production model	6	Tuna-RFMO, USA
Delay difference model	5	Canada, USA
ASPM: Age-structured surplus production model	4	South Africa
DPM: Dynamic production model	2	West Africa
qR: Surplus production model	2	Australia
LPM: Logistic production model	1	Canada
Integrated Analysis (n=57)		
SS3: Stock Synthesis v3.0 model	26	Australia, Europe, Tuna-RFMO, USA
SS2: Stock Synthesis v2.0 model	22	Australia, USA
SMS: Stochastic multi-species model	3	Europe
CASAL: C++ Algorithmic Stock Assessment Laboratory	2	New Zealand
IA: Integrated analysis	1	USA
JJM: Joint jack mackerel	1	Chile
SS1: Stock Synthesis v1.0 model	1	USA
SYM: Stochastic yield model	1	USA
Statistical catch at age model (n=55)		
AD-CAM: AD-Model Builder statistical catch-at-age model	20	Europe, South Africa, USA
SCA: Statistical catch-at-age model	8	Canada, Europe, Tuna-RFMO, USA
ASAP: Age Structured Assessment Program	6	USA
BAM: Beaufort assessment model	6	USA
ICA: Integrated catch-at-age analysis	5	Europe
TSA: State-space catch-at-age time series analysis	4	Canada, Europe
MULTIFAN-CL: Length-based, age/spatially-structured model	2	Tuna-RFMO
SAM: State-space assessment model	2	Europe

CSA: Catch-survey analysis (like a state space approach)	1	USA
SCALE: A statistical catch-at-length model	1	USA
Statistical catch at length model (n=3)		
AD-CAL: AD-Model Builder catch-at-length model	2	USA
LBA: Length-based analysis	1	USA
Survey index (n=5)		
Temporal indices derived from scientific survey data	3	USA
SURBA: Survey-based stock assessment method	2	Canada
Unknown (n=38)		
Unknown	27	Canada, Chile, Europe, Peru, South Africa, USA
MSLM: Multi-stock length-based model	7	New Zealand
SnapEst: SnapEst age- and length-based model	2	Australia
CapTool: Spreadsheet assessment model used for capelin	1	Europe
RYM: Replacement yield model	1	South Africa
Virtual population analysis (n=47)		
XSA: Extended survivor analysis	26	Argentina, Canada, Europe, Tuna-RFMO
VPA: Virtual population analysis	16	Argentina, Canada, Europe, Japan, Russia
ADAPT: Adaptive framework-virtual population analysis	1	Europe
B-ADAPT: ADAPT approach with year effects in a catch multiplier	1	Europe
FLXSA: FLR variant of extended survivor analysis	1	Europe
NFT-ADAPT: VPA/ADPAT version 2.3.2 NOAA Fisheries	1	Europe
SPA-ADAPT: Sequential population analysis / ADAPT	1	Canada

Supp. Table 4. Potential predictors of SST influence and their sources (percentage of stocks with predictor available shown in parenthesis when coverage is incomplete).

Variable	Source
SST experience	
SST average (°C)	HadISST + stock boundary database (1930-2010)
SST trend (°C/yr)	HadISST + stock boundary database (1930-2010)
Latitude (absolute value)	Centroid of the stock area (stock boundary database)
Stock characteristics	
Biomass average (MT)	RAM Legacy Database
Scaled biomass trend (scaled MT/yr)	RAM Legacy Database
Stock area (sq. km)	Stock boundary database
Time series length (year)	RAM Legacy Database
B/B_{MSY} average	RAM Legacy Database (52%)
F/F_{MSY} average	RAM Legacy Database (57%)
Geography	
Large Marine Ecoregion	Containing the centroid of the stock area
FAO Major Fishing Area	Containing the centroid of the stock area
Life history traits	
Taxonomy (family/order)	RAM Legacy Database (corrected for errors)
Natural mortality rate (M, 1/yr)	FishLife (finfish, 100%), SeaLifeBase (inverts, 19%)
Brody growth coefficient (K)	FishLife (finfish, 100%), SeaLifeBase (inverts, 100%)
Asymptotic maximum length (L_{inf}, cm)	FishLife (finfish, 100%), SeaLifeBase (inverts, 38%)
Asymptotic maximum mass (W_{inf}, kg)	FishLife (finfish, 100%), SeaLifeBase (inverts, 24%)
Length at maturity (L_{mat}, cm)	FishLife (finfish, 100%), SeaLifeBase (inverts, 0%)
Age at maturity (T_{mat}, yr)	FishLife (finfish, 100%), SeaLifeBase (inverts, 0%)
Maximum age (T_{max}, yr)	FishLife (finfish, 100%), SeaLifeBase (inverts, 19%)
Trophic level	FishBase (finfish, 93%), SeaLifeBase (inverts, 19%)
Habitat (e.g., demersal, pelagic, etc.)	FishBase (finfish, 99%), SeaLifeBase (inverts, 95%)
Depth (m)	FishBase (finfish, 95%), SeaLifeBase (inverts, 0%)

Supp. Table 5. Stocks whose productivity is significantly influenced by SST warming (sorted from most positive to most negative SST influence).

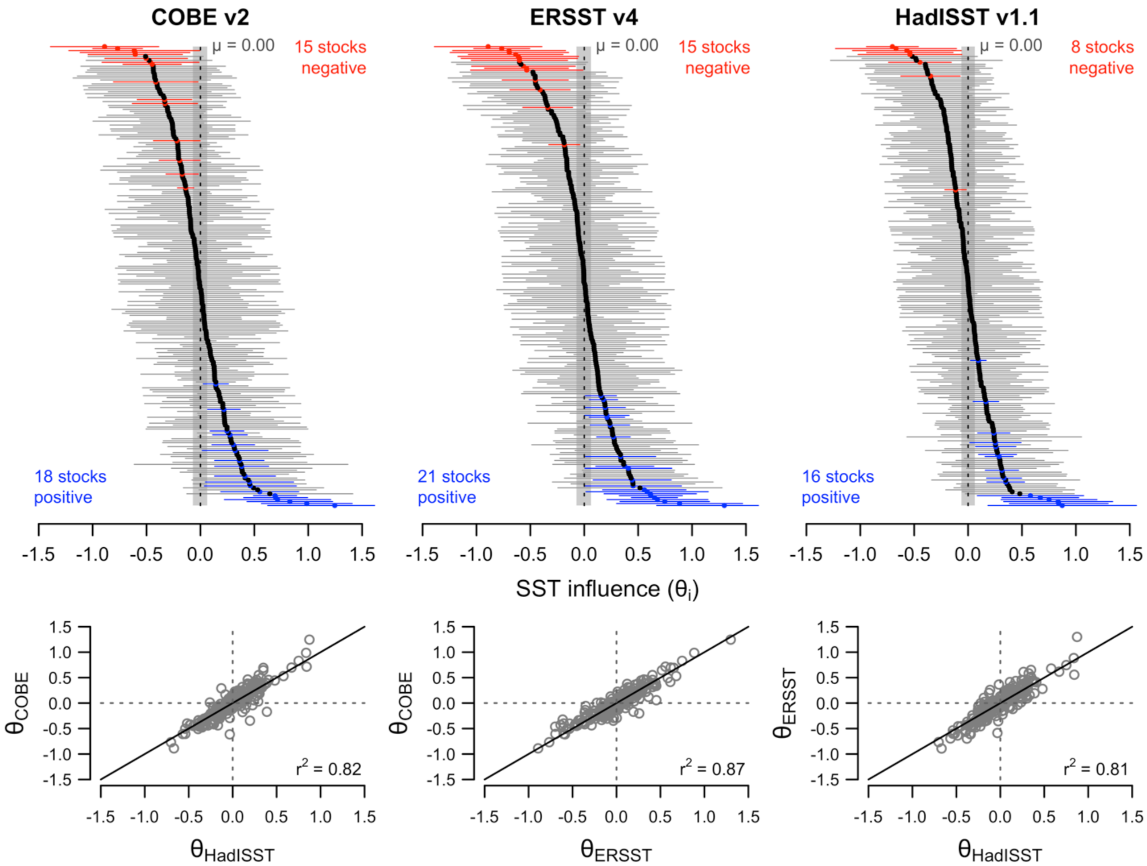
Stock id	**Species**	**Area**	**θ_i**
BIGEYEWPO	Bigeye tuna (*Thunnus obesus*)	Western Pacific Ocean	1.16
SCALLGB	Sea scallop (*Placopecten magellanicus*)	Georges Bank	0.92
HERR30	Atlantic herring (*Clupea harengus*)	Bothnian Sea	0.72
ALBANPAC	Albacore tuna (*Thunnus alalunga*)	North Pacific Ocean	0.68
MONKGOMNGB	Monkfish (*Lophius americanus*)	Gulf of Maine / Northern Georges Bank	0.55
STMARLINSWPO	Striped marlin (*Kajikia audax*)	Western Pacific Ocean	0.52
RROCKLOBSTERCRA7	Red rock lobster (*Jasus edwardsii*)	New Zealand Area CRA7	0.44
HERRRIGA	Atlantic herring (*Clupea harengus*)	Gulf of Riga East of Gotland	0.38
BSBASSMATLC	Black sea bass (*Centropristis striata*)	Mid-Atlantic Coast	0.37
TIGERFLATSE	Tiger flathead (*Platycephalus richardsoni*)	Southeast Australia	0.37
WHAKEGBGOM	White hake (*Urophycis tenuis*)	Gulf of Maine / Georges Bank	0.27
BMARLINATL	Blue marlin (*Makaira nigricans*)	Atlantic Ocean	0.26
CROCKWCVANISOGQCI	Canary rockfish (*Sebastes pinniger*)	WC Vancouver Island, Straight of Georgia, Queen Charlotte Islands	0.25
SCALL5ZMAB	Sea scallop (*Placopecten magellanicus*)	Georges Bank and Mid-Atlantic Bight	0.25
RKCRABBB	Red king crab (*Paralithodes camtschaticus*)	Bristol Bay	0.23
HAKENRTN	European hake (*Merluccius merluccius*)	IIIa-IV-VI-VII-VIIIabd	0.21
SPANMACKSATLC	Spanish mackerel (*Scomberomorus maculatus*)	Southern Atlantic coast	0.20
CROCKPCOAST	Canary rockfish (*Sebastes pinniger*)	Pacific Coast	0.17
KINGKLIPSA	Kingklip (*Genypterus capensis*)	South Africa	-0.10
DEEPCHAKESA	Deep-water cape hake (*Merluccius paradoxus*)	South Africa	-0.23
ALBANATL	Albacore tuna (*Thunnus alalunga*)	Northern Atlantic	-0.26

RROCKLOBSTERCRA2	Red rock lobster (*Jasus edwardsii*)	New Zealand Area CRA2	-0.30
STMARLINIO	Striped marlin (*Kajikia audax*)	Indian Ocean	-0.32
CODVIIek	Atlantic cod (*Gadus morhua*)	Celtic Sea	-0.36
HERRNS	Atlantic herring (*Clupea harengus*)	North Sea	-0.36
POLLNS-VI-IIIa	Saithe (*Pollachius virens*)	IIIa, VI and North Sea	-0.36
SAILEATL	Sailfish (*Istiophorus albicans*)	Eastern Atlantic	-0.38
RHERRTSST	Round herring (*Etrumeus sadina*)	Tsushima Strait	-0.42
CODFAPL	Atlantic cod (*Gadus morhua*)	Faroe Plateau	-0.47
CODNS	Atlantic cod (*Gadus morhua*)	North Sea	-0.51
WHITVIIek	Whiting (*Merlangius merlangus*)	Celtic Sea	-0.53
WHITVIa	Whiting (*Merlangius merlangus*)	West of Scotland	-0.60
SEELNSSA2	Sand eel (*Ammodytes marinus*)	North Sea	-0.61
SEELNSSA3	Sand eel (*Ammodytes marinus*)	North Sea	-0.70
SOLEIS	Common sole (*Solea solea*)	Irish Sea	-0.72
CODIS	Atlantic cod (*Gadus morhua*)	Irish Sea	-0.80
CODVIa	Atlantic cod (*Gadus morhua*)	West of Scotland	-0.94

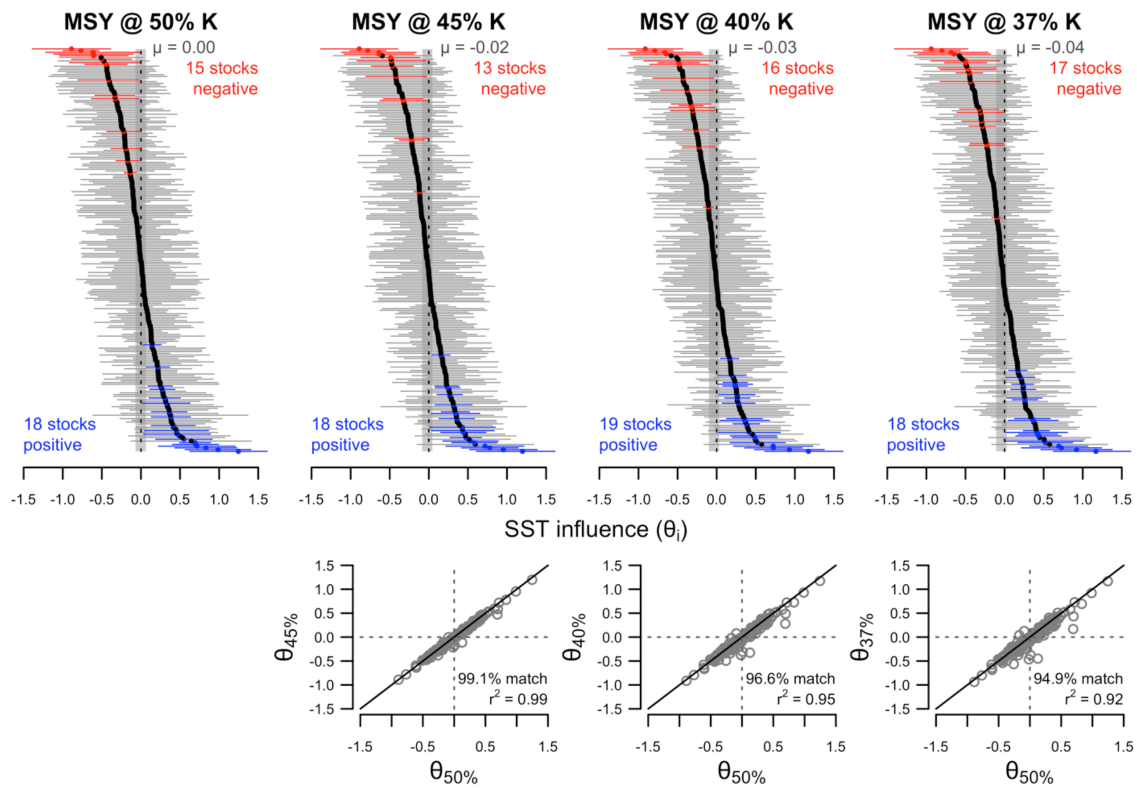
Supp. Table 6. Hindcasted changes in SST-dependent maximum sustainable yield (MSY) from 1930-2010 among ecoregions (LME=large marine ecoregion; HSA=high seas area; sorted by ascending percent difference).

Type	LME	# of stocks	Total MSY (1000s of mt)	Mean SST (°C)	SST trend (°C / decade)	SST influence (θ_i)	MSY change mt / decade	MSY change % difference
LME	Sea of Japan	1	30.0	12.3	0.134	-0.12	-1922.2	-66.8
LME	Kuroshio Current	6	3681.7	22.1	0.121	-0.18	-150431.8	-43.3
LME	East China Sea	5	1844.5	21.5	0.142	-0.07	-71994.8	-40.1
LME	North Sea	9	2603.7	9.8	0.082	-0.31	-92220.2	-38.2
LME	Iberian Coastal	2	2.6	16.2	0.085	-0.13	-74.9	-29.0
LME	Celtic-Biscay Shelf	15	294.4	12.5	0.064	-0.32	-3985.6	-17.0
LME	Benguela Current	3	157.4	19.6	0.079	-0.10	-2620.8	-16.2
HSA	South Atlantic Ocean	1	25.8	14.4	0.075	-0.09	-456.8	-13.4
LME	Labrador - Newfoundland	8	324.0	4.5	0.017	-0.13	-3237.9	-9.8
LME	Faroe Plateau	3	82.4	9.2	0.018	-0.11	90.4	-6.0
HSA	North Atlantic Ocean	6	364.4	20.4	0.023	-0.12	-1297.2	-4.0
LME	Greenland Sea	1	960.4	1.1	0.031	-0.05	-2020.3	-3.9
LME	East Bering Sea	14	2956.3	4.7	0.038	-0.03	-8938.1	-3.8
LME	Norwegian Sea	1	1943.5	6.8	0.040	-0.04	-3321.7	-3.7
LME	Patagonian Shelf	2	285.1	10.3	0.059	-0.04	-843.9	-3.7
HSA	Bay of Biscay	1	7.1	15.0	0.084	-0.03	-15.9	-2.7
LME	Southeast U.S. Continental Shelf	9	20.2	25.4	-0.115	0.03	-71.3	-2.1
LME	Gulf of Mexico	3	1.7	25.9	0.000	0.06	-6.4	-1.7
LME	California Current	29	278.2	16.5	0.034	0.04	-109.0	-1.6
LME	Gulf of Alaska	20	280.5	9.0	0.014	-0.02	-398.0	-1.2

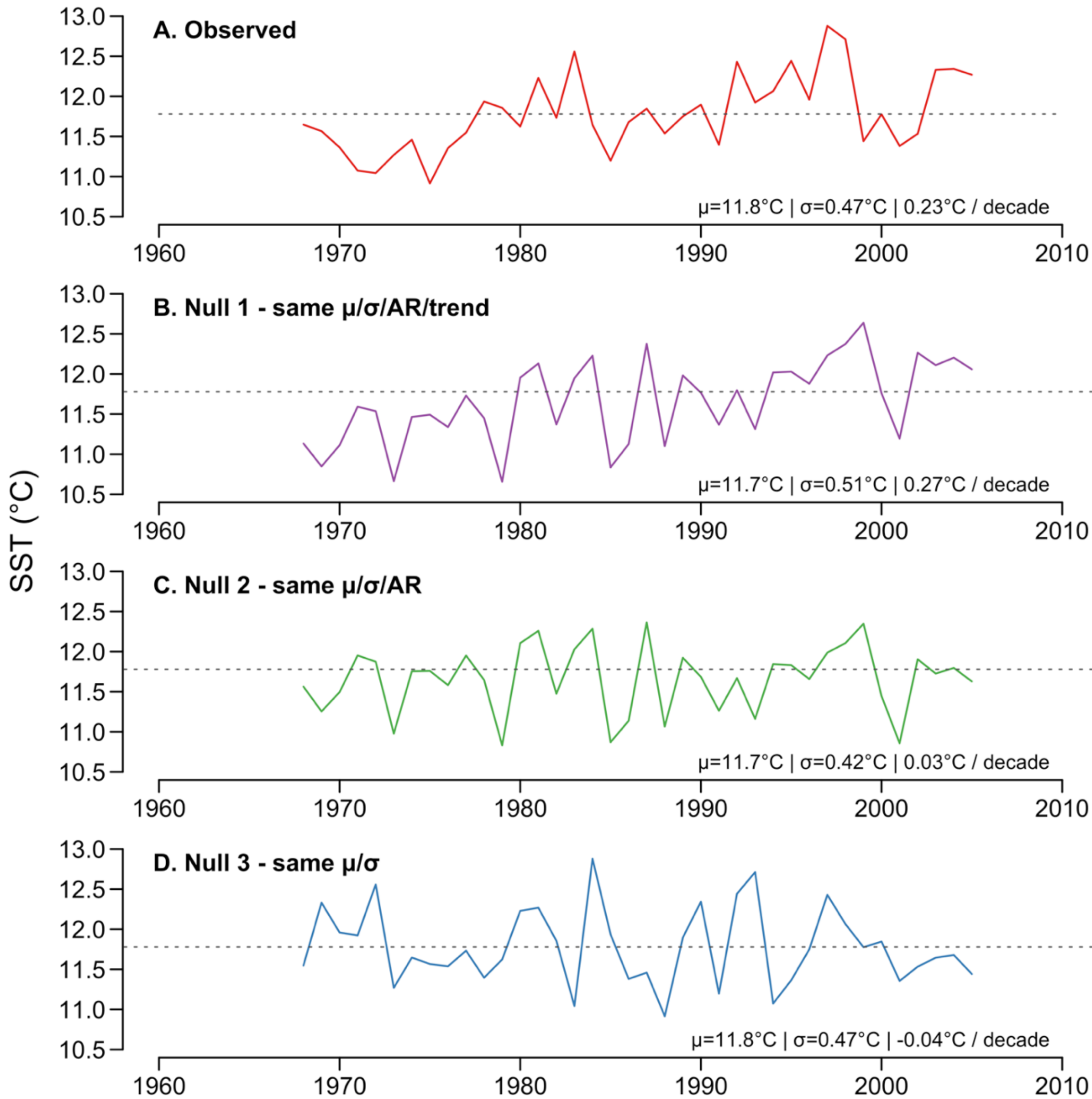
LME	Iceland Shelf and Sea	5	1300.3	4.6	0.031	0.02	-327.0	-1.1
LME	West Bering Sea	1	296.9	4.3	0.037	-0.03	-33.6	-0.8
LME	Mediterranean Sea	2	45.9	19.7	0.025	-0.01	-0.9	-0.1
LME	New Zealand Shelf	9	7.4	15.1	0.049	-0.06	7.2	0.0
LME	Scotian Shelf	2	51.7	7.2	0.074	0.01	11.4	0.3
LME	Southeast Australian Shelf	8	12.0	14.6	0.088	0.01	6.8	0.9
LME	Agulhas Current	5	760.4	25.0	0.076	0.02	1061.5	1.6
LME	North Brazil Shelf	1	34.2	27.7	0.022	0.00	65.5	2.9
LME	Northeast U.S. Continental Shelf	15	822.7	11.5	0.064	0.19	303.1	3.1
LME	Humboldt Current	16	15985.3	15.1	0.077	0.00	86036.1	3.5
LME	Barents Sea	5	2188.0	1.1	0.079	0.03	7360.7	4.3
HSA	Labrador Sea	2	204.2	4.2	0.042	-0.04	-301.6	5.9
LME	South West Australian Shelf	5	7.0	17.0	0.074	0.08	103.4	11.0
LME	Baltic Sea	5	877.0	7.5	0.088	0.16	9782.9	11.0
HSA	North Pacific Ocean	4	132.7	22.1	0.036	0.02	2961.8	20.6
HSA	Indian Ocean	4	398.2	15.3	0.066	0.00	12417.3	22.0
LME	Canary Current	2	714.8	21.0	0.066	0.14	15325.9	26.1
HSA	South Pacific Ocean	5	97.8	18.1	0.046	0.26	3777.6	31.4



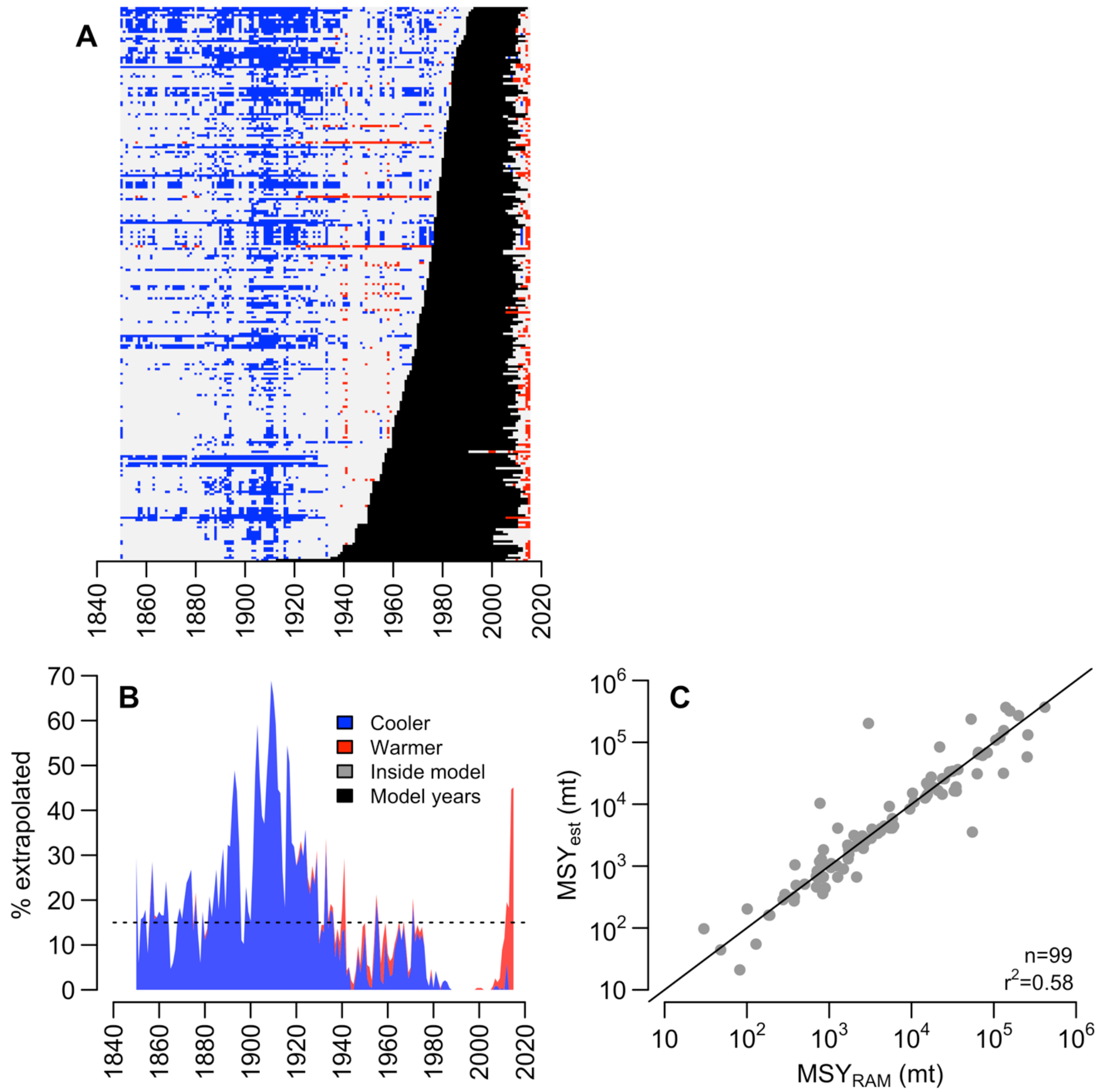
Supp. Figure 1. Comparison of SST influence estimates from the SST-linked Schaefer surplus production model using each of three SST datasets: COBE v2, ERSST v4, and HadISST v1.1. In the top panels, points show mean estimates and error bars show 95% confidence intervals. Significant positive and negative SST influences are shown in blue and red, respectively. The transparent rectangle indicates the 95% confidence interval for the global mean of the SST influences. In the bottom panels, the diagonal line is the one-to-one line for pairwise comparisons of SST influence estimates using the different SST datasets.



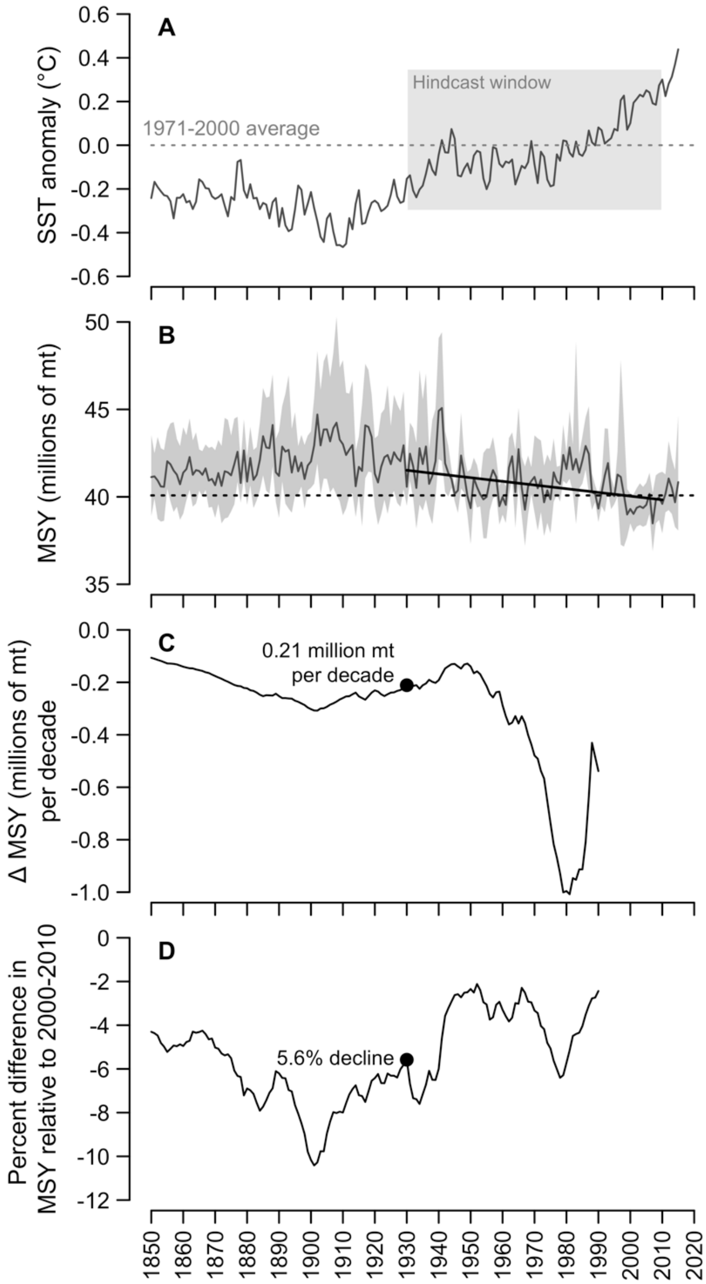
Supp. Figure 2. Comparison of SST influence estimates from the SST-linked Pella-Tomlinson surplus production model using four shape parameters: p=1.00 (MSY @ 50% K, Schaefer model), p=0.55 (MSY @ 45% K), p=0.20 (MSY @ 40% K), and p=0.01 (MSY @ 37% K). In the top panels, points show mean estimates and error bars show 95% confidence intervals. Significant positive and negative SST influences are shown in blue and red, respectively. The transparent rectangle indicates the 95% confidence interval for the global mean of the SST influences. In the bottom panels, the diagonal line is the one-to-one line for pairwise comparisons of SST influence between the symmetric Schaefer model (p=1.00) and the asymmetric Pella-Tomlinson models. The r^2 value and percent agreement in significance are shown in the bottom-right.



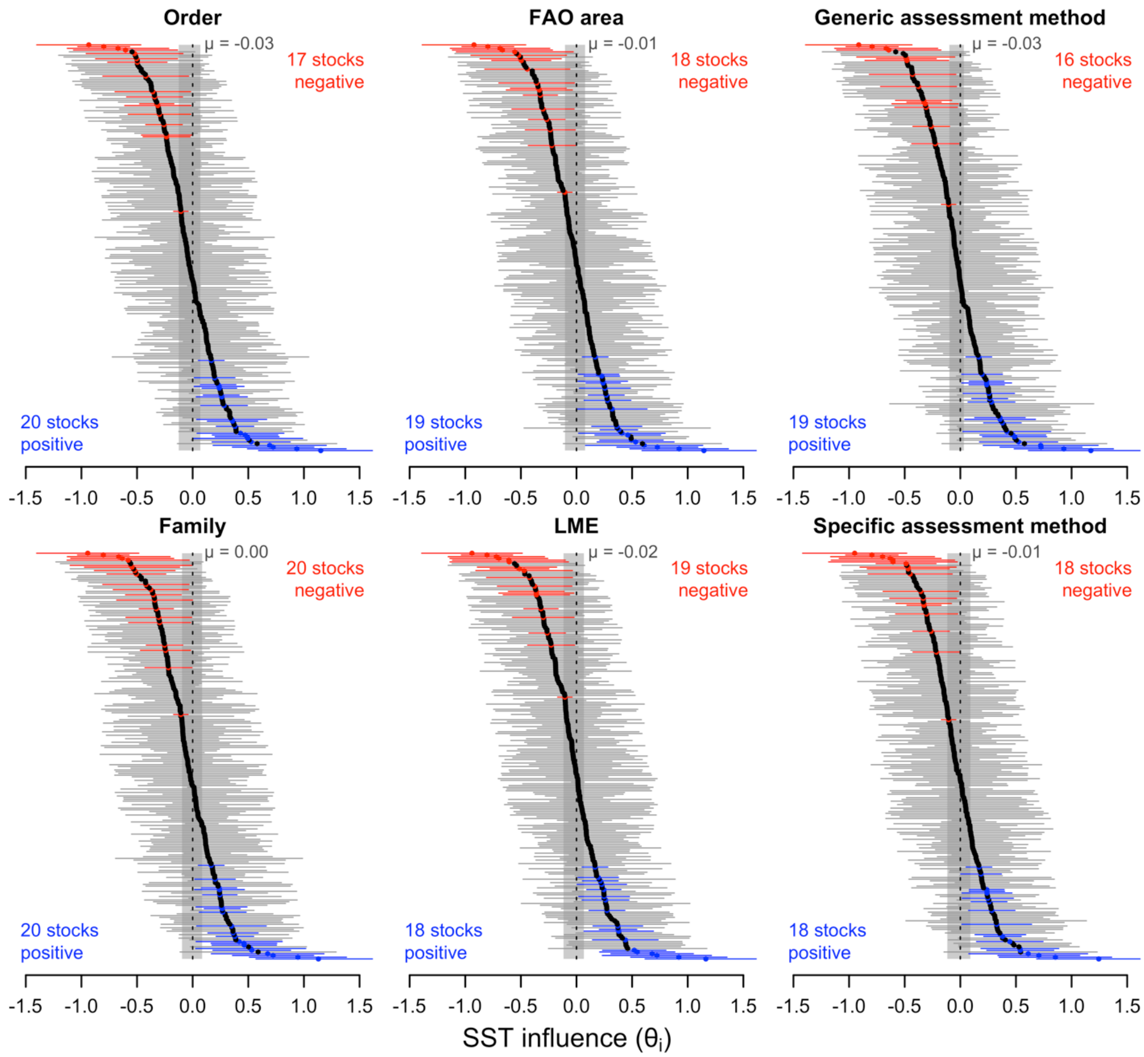
Supp. Figure 3. Example **(A)** observed and **(B-D)** simulated SST time series (US West Coast, Black rockfish). The simulated SST time series were used in the three null models.



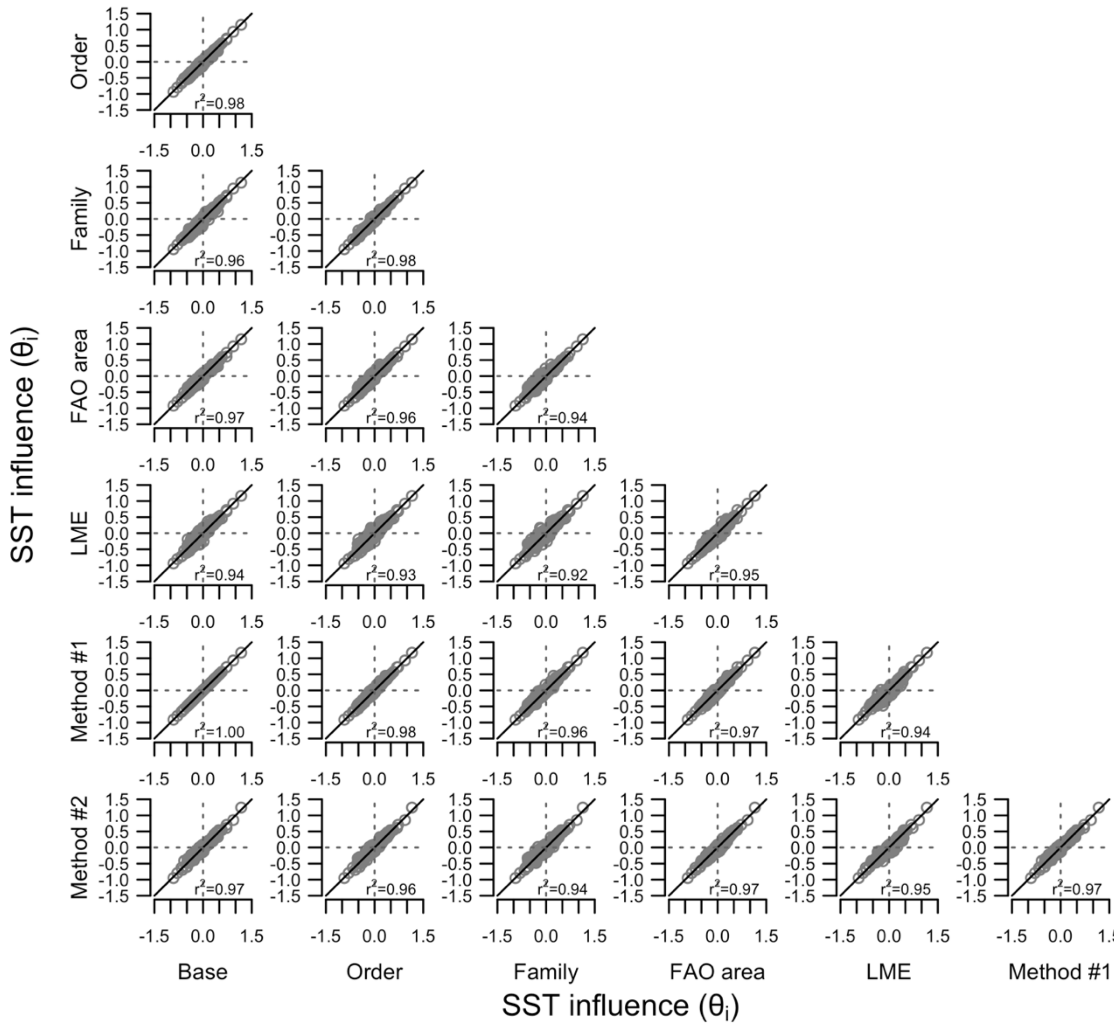
Supp. Figure 4. The **(A&B)** frequency of SST extrapolation by the hindcast model and **(C)** correlation between MSY estimates from the final model and data-rich stock assessments (diagonal line is the one-to-one line). In **(A)**, each row shows the SST experience of an individual stock where black years were used in model development, grey years experienced temperatures also experienced during model years, and blue and red years experienced temperatures cooler and warmer than those experienced during model years, respectively. In **(B)**, the blue and red shading show the percentage of years experiencing temperatures cooler and warmer than those experience during model years, respectively. The hindcast model generally extrapolates for fewer than 15% (dashed line) of years between 1930-2010.



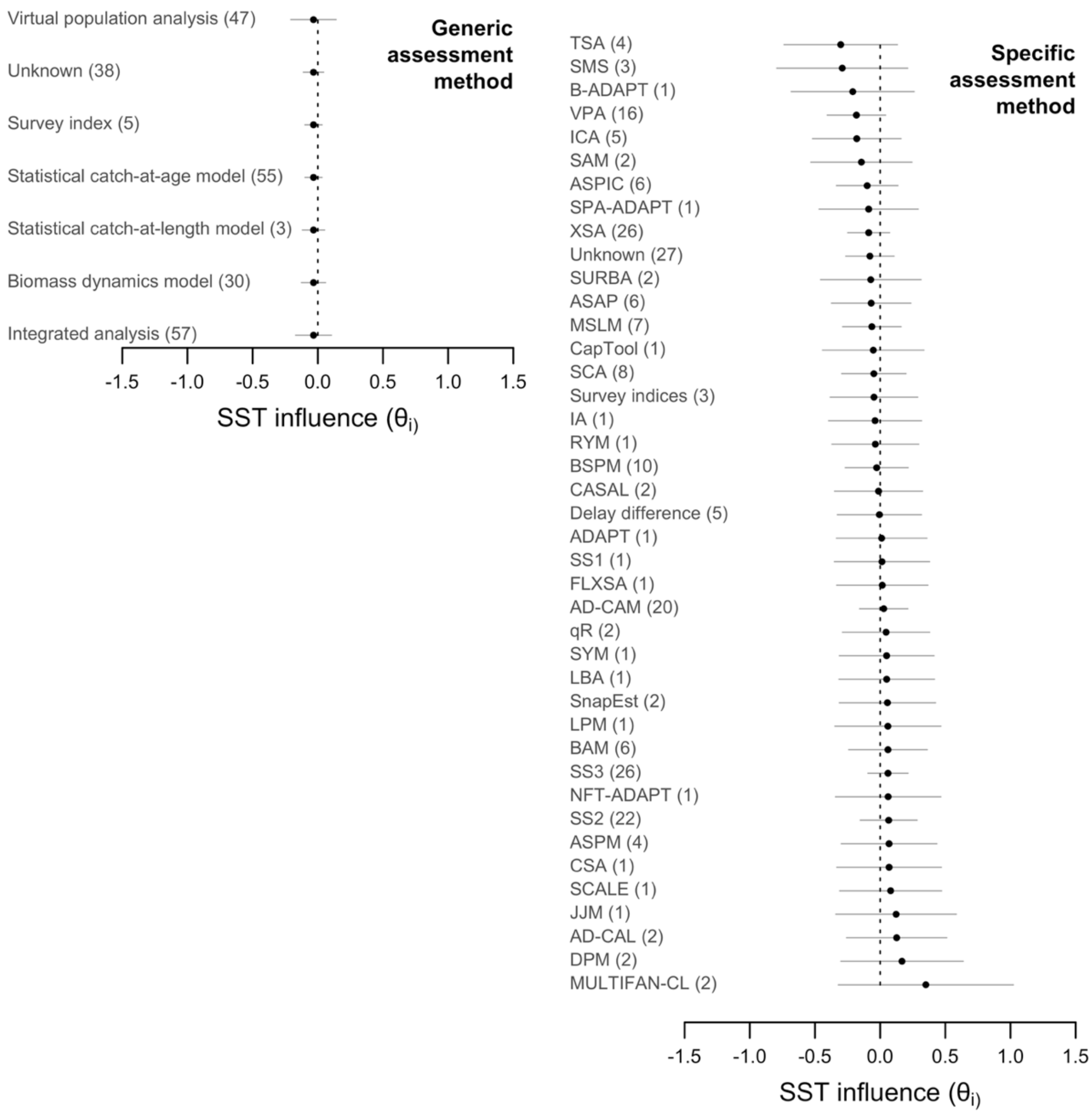
Supp. Figure 5. Sensitivity of hindcasted changes in MSY to the determination of the hindcast window. Time series showing **(A)** mean global SST anomaly, **(B)** hindcast of SST-dependent maximum sustainable yield (MSY) for all stocks included in the analysis, **(C)** Thiel-Sen regression slope when evaluating MSY trends beginning in each year from 1850-1990 and ending in 2010, and **(D)** percent difference in MSY when comparing the mean MSY over the 10 years following each year from 1850-1990 and the mean MSY from 2001-2010. In **(A)**, the grey shading indicates the hindcast window determined to minimize extrapolation to temperatures outside those included in the final model. In **(B)**, the dark line shows a Thiel-Sen regression fit to the MSY time series in the hindcast window. In **(C)** and **(D)**, the labeled points mark the measures of MSY change experienced over the hindcast window.



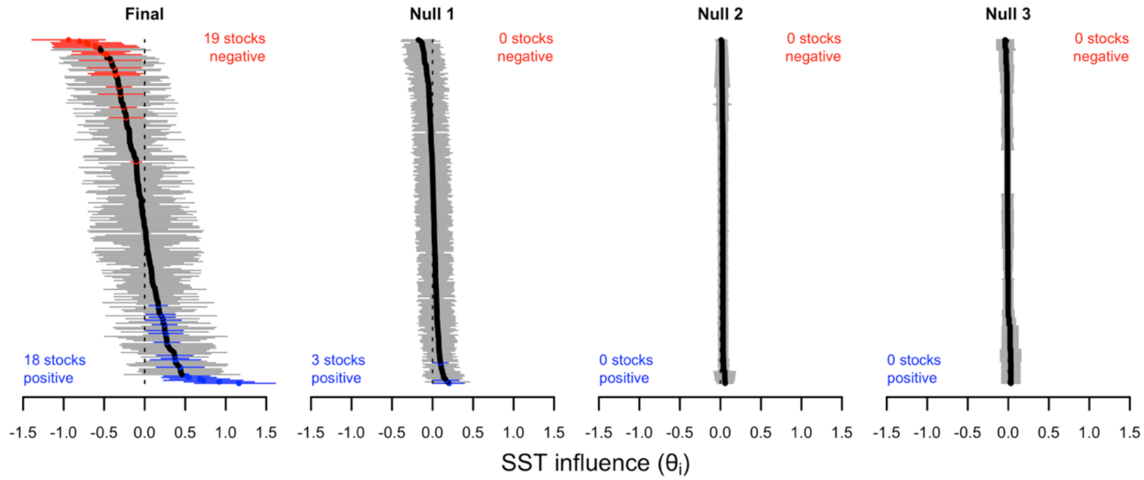
Supp. Figure 6. Distribution of SST influences estimated by the SST-linked Pella-Tomlinson surplus production models with hierarchy on SST influence by **(A)** taxonomic order and **(B)** taxonomic family, **(C)** FAO major fishing area and **(D)** large marine ecoregion (LME), and **(E)** generic and **(F)** specific stock assessment methods. Points show mean estimates and error bars show 95% confidence intervals. Significant positive and negative SST influences are shown in blue and red, respectively. The transparent rectangle indicates the 95% confidence interval for the global mean of the SST influences.



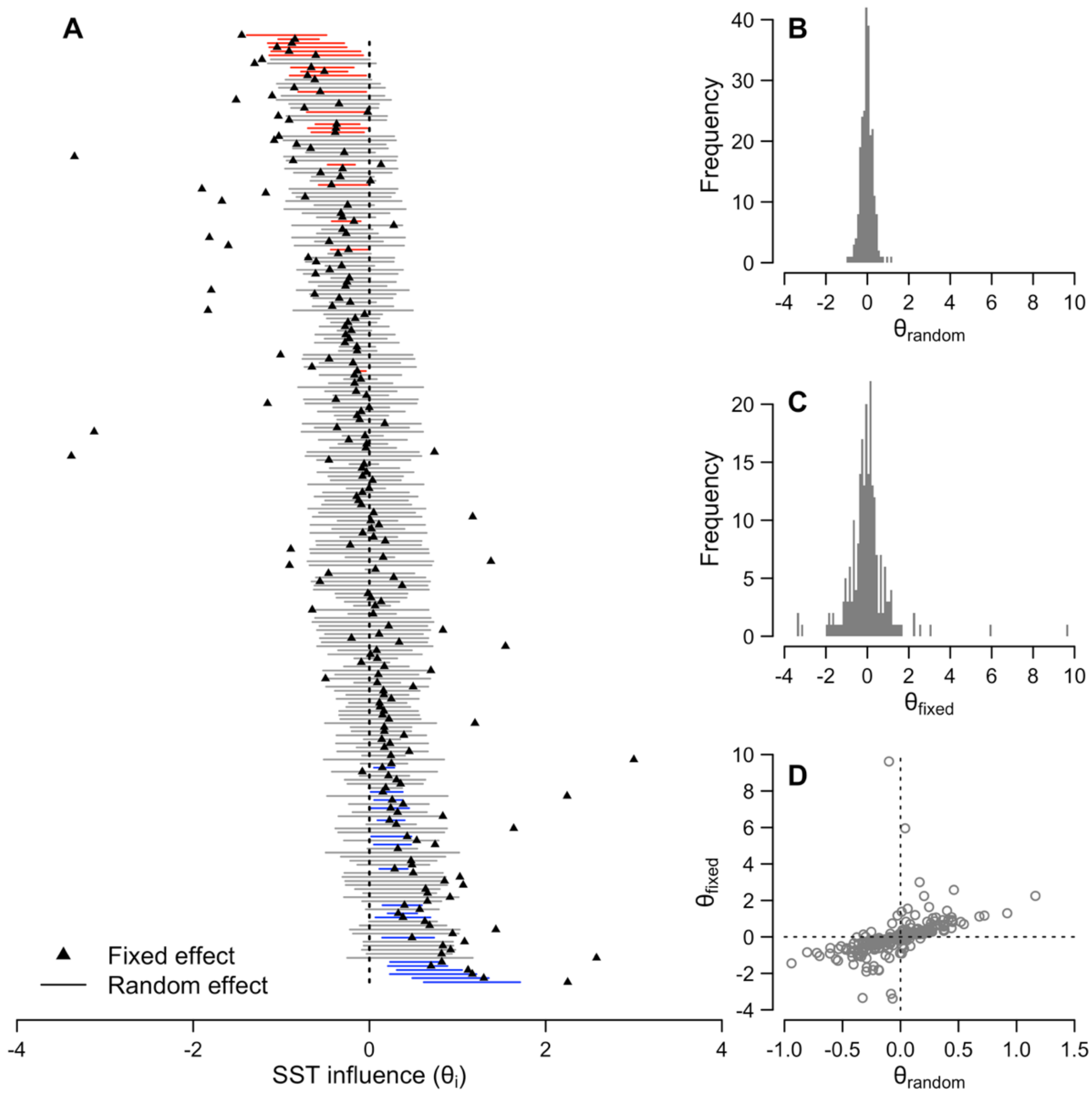
Supp. Figure 7. Correlation between SST influences estimated by the base model and six group models. Diagonal line is the one-to-one line.



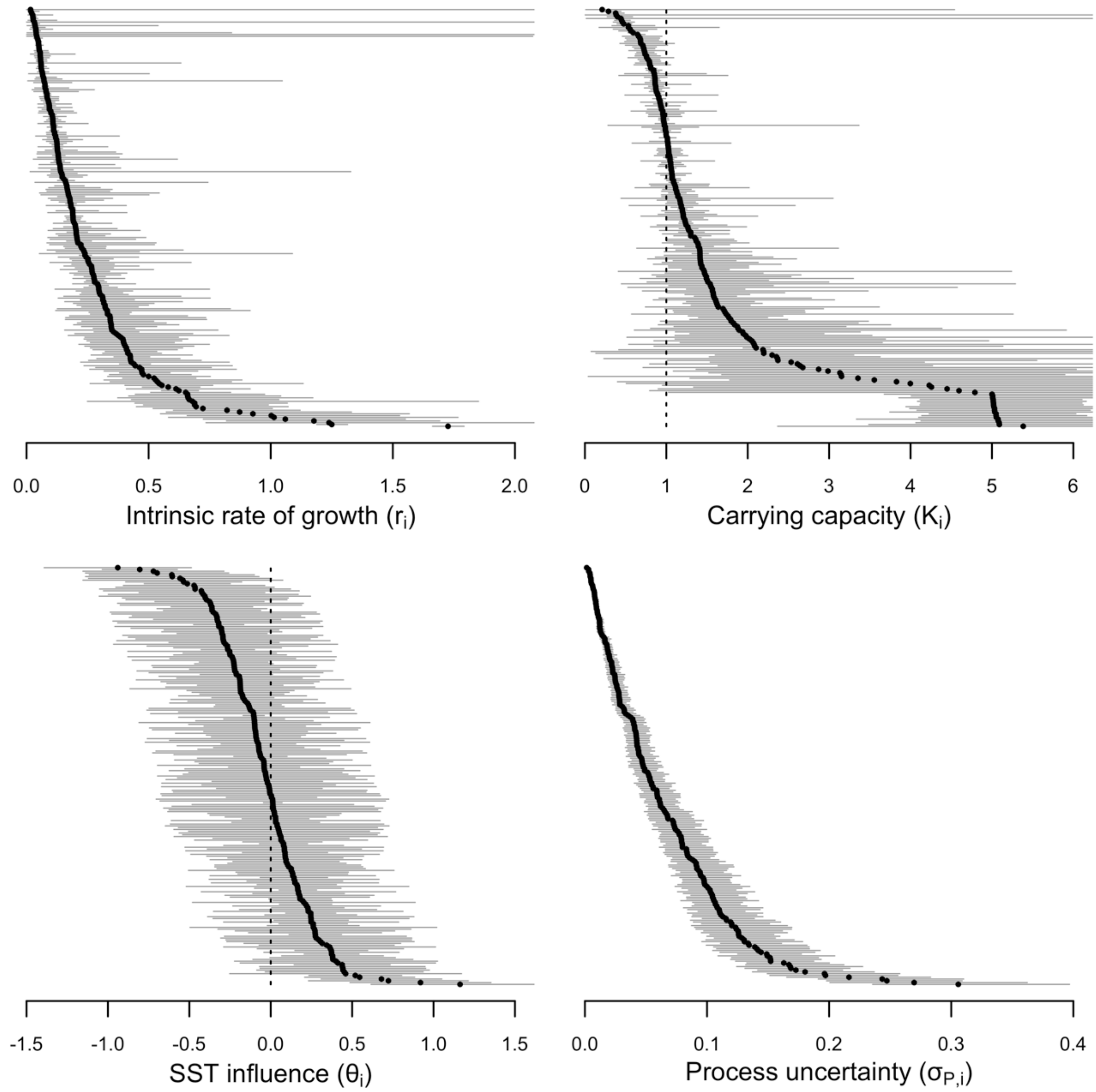
Supp. Figure 8. Mean of the SST influence distributions for assessment method groups in models with hierarchy on SST influence by **(A)** generic stock assessment method and **(B)** specific stock assessment method. Points show mean estimates and error bars show 95% confidence intervals. None of the SST influence means were significantly different from zero and neither of the models gained more support than the base model (see **Table 1**).



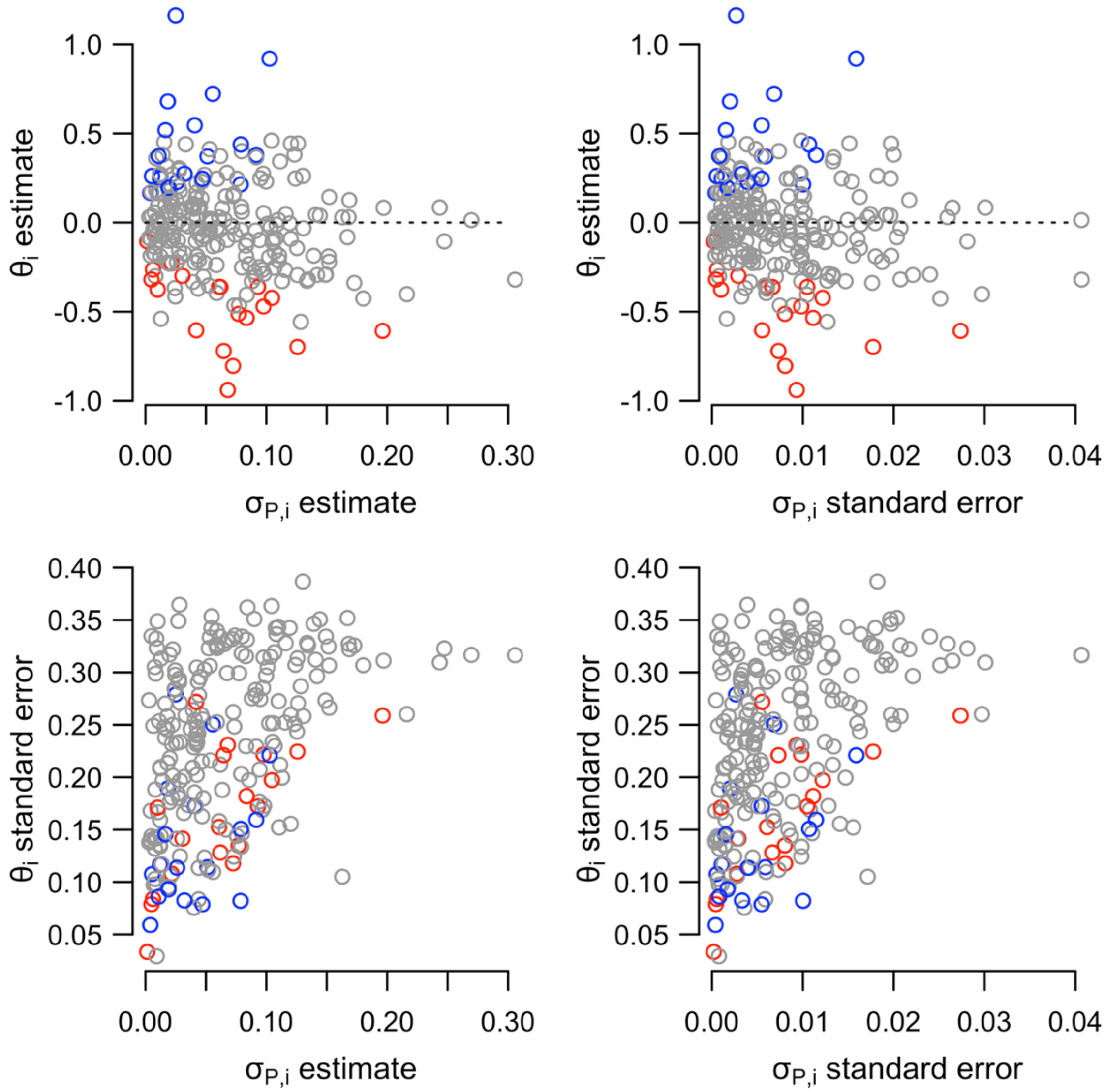
Supp. Figure 9. Distribution of SST influences estimated by the final model and three null models. Points show mean estimates and error bars show 95% confidence intervals. Significant positive and negative SST influences are shown in blue and red, respectively.



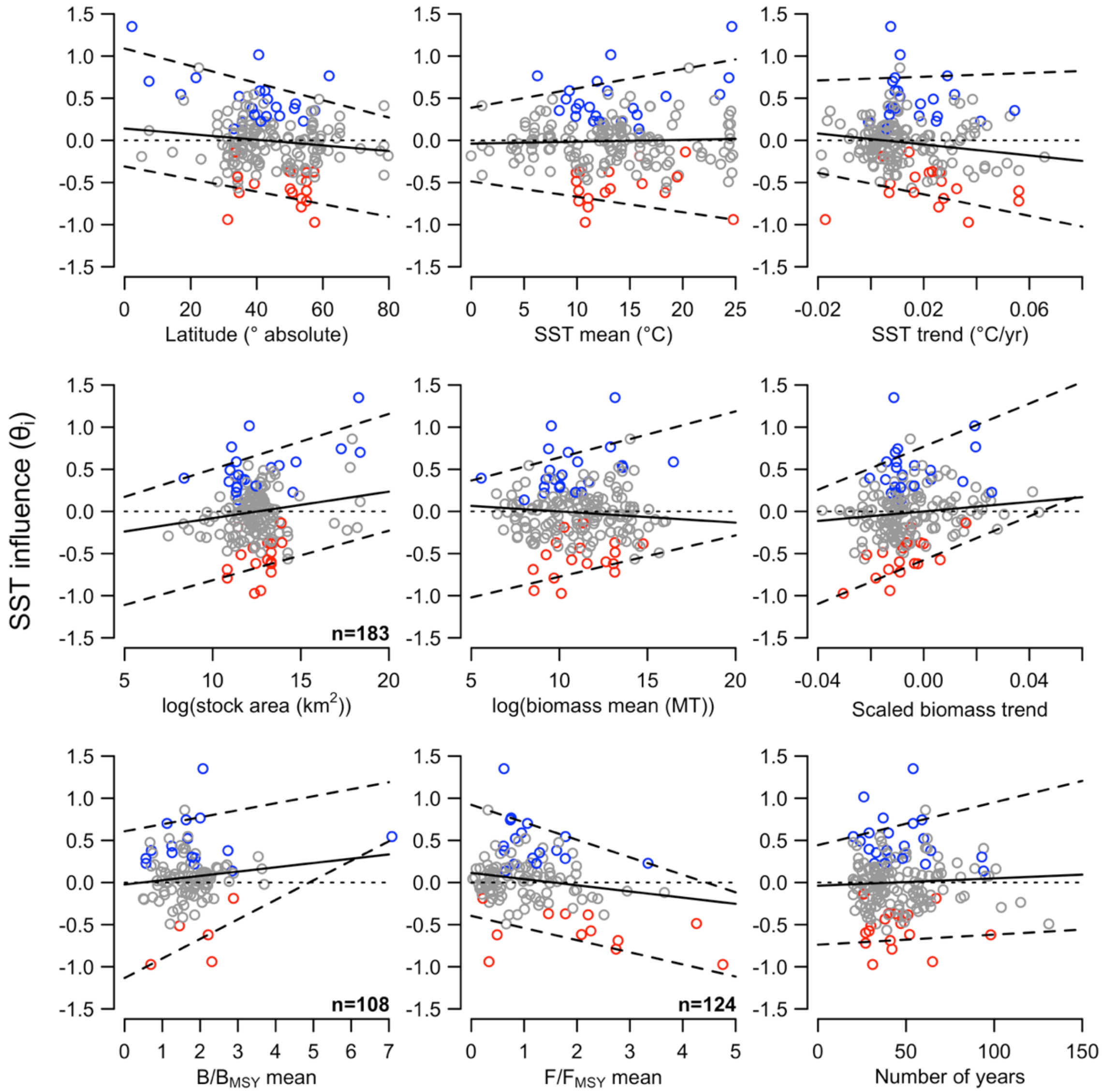
Supp. Figure 10. Comparison of SST influences estimated in a fixed effects framework with the random effects framework of the final model. Plots show **(A)** mean fixed effects estimates plotted over their corresponding random effects estimates (95% confidence interval), histograms of the **(B)** random and **(C)** fixed effects estimates, and **(D)** correlation between the random and fixed effects estimates.



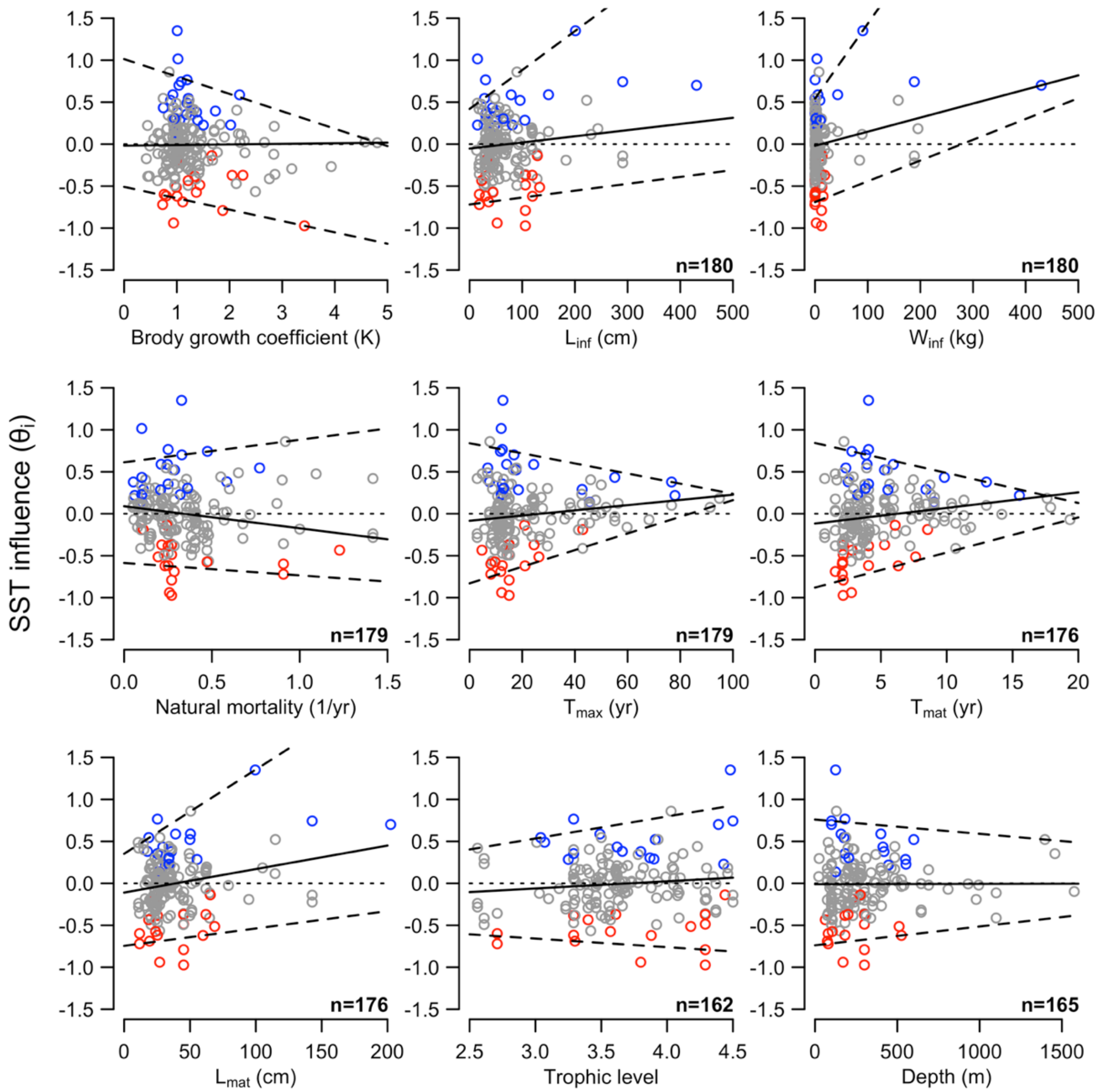
Supp. Figure 11. Distribution of intrinsic rate of growth (r_i), carrying capacity (K_i), SST influence (θ_i), and process uncertainty ($\sigma_{P,i}$) estimates from the final model. Points show mean estimates and lines show 95% confidence intervals. Carrying capacity is a multiple of the maximum observed biomass (e.g., a carrying capacity of 1, shown by the vertical dotted line, means that the carrying capacity is equivalent to the maximum observed biomass).



Supp. Figure 12. Correlation between the SST influence estimates and standard errors and the process uncertainty estimates and standard errors. Points are colored by significance of SST influence (blue=positive, red=negative, grey=non-significant).

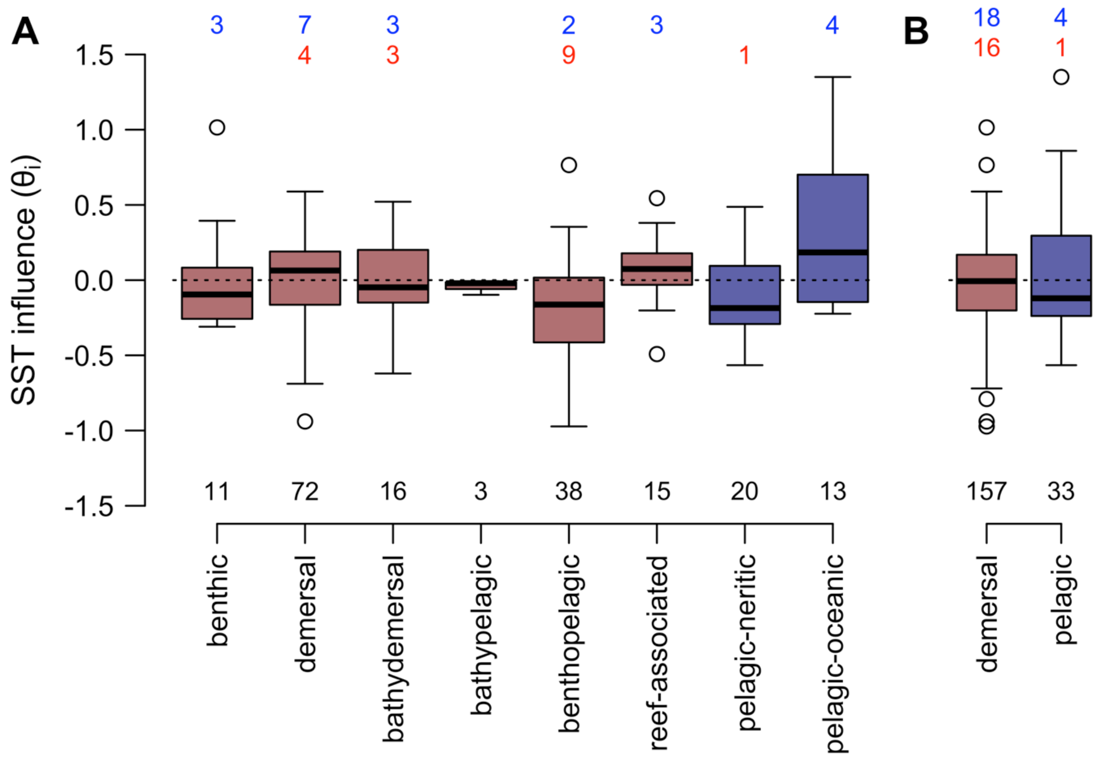


Supp. Figure 13. SST influence as a function of nine stock characteristics. SST influences are colored by significance (blue=positive, red=negative, grey=non-significant). Solid lines show the 50th percentile quantile regression fit and dashed lines show the 2.5% and 97.5% quantile regression fits. Sample size is shown in the bottom-right corner if data were not available for all 235 stocks.

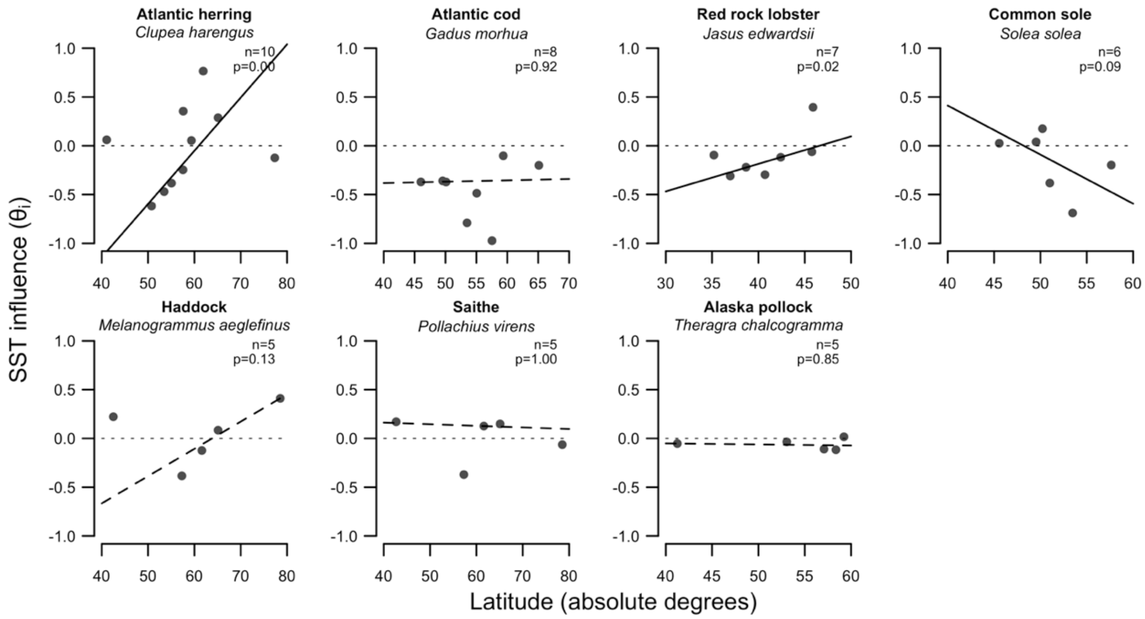


Supp. Figure 14. SST influence as a function of nine life history traits: Brody growth coefficient (K), asymptotic maximum length (L_{inf}), asymptotic maximum weight (W_{inf}), natural mortality (M), maximum age (T_{max}), age at maturity (T_{mat}), length at maturity (L_{mat}), trophic level, and median depth. SST influences are colored by significance (blue=positive, red=negative, grey=non-significant). Solid lines show the 50th percentile quantile regression fit and dashed lines show the 2.5% and

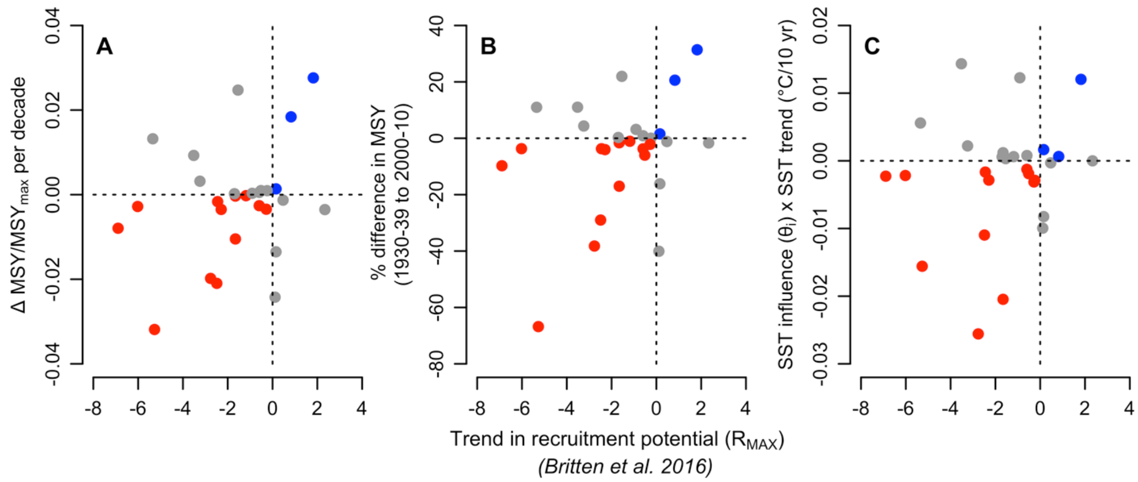
97.5% quantile regression fits. Sample size is shown in the bottom-right corner if data were not available for all 235 stocks.



Supp. Figure 15. Distribution of SST influence among **(A)** specific and (**B)** generic habitat types. Brown and blue boxplot shading corresponds to demersal and pelagic habitats, respectively. Black numbers indicate total number of stocks for each habitat type. Blue and red numbers show the number of stocks with a positive and negative SST influence, respectively.



Supp. Figure 16. SST influence as a function of the latitude of stocks of the same species for the seven species with ≥5 stocks in the analysis. Lines shows Theil-Sen regression fits with solid lines indicating regressions significant at the 0.10 level. Theil-Sen regression, a form of robust regression, identifies the median slope of lines through all possible point pairs and is insensitive to outliers and endpoints in small datasets.



Supp. Figure 17. Comparison of LME-scale changes in fisheries productivity estimated by Britten et al. (2016) and the present study. Britten et al. (2016) quantify the meta-analytic mean trend in recruitment potential (R_{MAX}). Comparable values derived from the present study are: **(A)** change in scaled MSY (MSY divided by maximum MSY) per decade from 1930-2010; **(B)** percent difference in mean MSY from 1930-39 to 2001-2010; and **(C)** the meta-analytic mean of the SST influences of stocks in an LME multiplied by the change in temperature from 1930-2010 in the LME. In both studies, negative and positive values represent a negative and positive change, respectively. Blue and red points indicate LMEs where both studies agree that change has positively and negatively impacted productivity, respectively. Grey points indicate LMEs in which the studies disagree on the direction of productivity change. The present study describes SST influence for ten LMEs not described in the Britten study (Bay of Biscay, Canary Current, Greenland Sea, Humboldt Current, Kuroshio Current, Labrador Sea, Mediterranean Sea, North Brazil Shelf, South Atlantic Ocean, West Bering Sea) and the Britten study describes SST influence on one LME not described in the present study (East-Central Australian Shelf).

Conclusions

Sustainable fisheries management is a global challenge requiring local solutions. I show that these solutions can be achieved by developing new, innovative interdisciplinary and quantitative methods. In Chapter 1, I show that a mixed-method approach can be used to quantify illegal fishing, its impacts on an endangered fish species, and its importance to the local community. The methods described here can be used to assess non-compliance and its impacts and motivations in protected areas around the world. In Chapter 2, I show that advanced analytic methods and information borrowed from data-rich stocks can be used to inform the management of data-poor stocks. I also provide a framework for comparing the performance of data-limited methods as new methods are developed. In Chapter 3, I show for the first time that climate change has resulted in a net decline in marine fisheries productivity and sustainable catch potential. Adapting fisheries management to account for shifts in productivity will require global innovations and local, regional, and national implementations of new policies. Together, these chapters work to help fisheries management overcome challenges from capacity shortfalls, data limitations, and climate change.