



Using a tier classification system to evaluate the quality of bycatch estimates from fisheries

Lee R. Benaka · Andrea N. Chan · Steven J. Kennelly · Noelle A. Olsen

Received: 18 November 2020 / Accepted: 24 June 2021 / Published online: 12 July 2021

© This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2021

Abstract Bycatch continues to be a challenge to sustainable fisheries management (The term “bycatch” in this paper covers discards and does not include retained incidental catch). Bycatch estimates can inform stock status determinations by improving understanding of fishing mortality, and help managers monitor the effectiveness of regulations. Assessments of the quality of bycatch estimation programs and procedures are necessary to evaluate the precision and limitations of their results over time. NOAA Fisheries experts used a “Tier Classification System” (TCS) to compare the quality of fish bycatch data and estimation methods for U.S. commercial fisheries in 2005 and 2015. The TCS included criteria related to data

adequacy and analytical approaches. A comparison of U.S. fishery tier scores demonstrated that most fisheries were classified into higher tiers in 2015 compared to 2005 due to factors including improved sampling design. In addition, this comparison identified region-specific trends (e.g., mostly improvements occurred for Alaska fisheries with more mixed results for Greater Atlantic fisheries). The improvements in bycatch data quality and estimation methods in the United States are a result of financial investments in observer programs by NOAA Fisheries and industry partners, as well as effective conservation measures implemented by regional fishery management councils and NOAA Fisheries. The TCS was also used to assess bycatch data and estimation methods in all of Australia’s fishery jurisdictions for the decade 2010–19, illustrating the international applicability of the method. Overall, Australian state fisheries scored lower than federally managed fisheries in both the United States and Australia, reflecting the fact that the latter fisheries tend to be larger (and more valuable) than those in state jurisdictions, with a larger investment in observer programs. A comparison of tier scores and estimates of discards by fishery may provide a useful input for decision-making processes regarding allocation of resources to improve bycatch monitoring.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11160-021-09670-y>.

L. R. Benaka (✉)
NOAA Fisheries, 1315 East-West Highway,
Silver Spring, MD, USA
e-mail: Lee.Benaka@noaa.gov

A. N. Chan
ECS Federal, LLC, NOAA Fisheries, Fairfax,
VA, USA

S. J. Kennelly
IC Independent Consulting, Cronulla, Australia

N. A. Olsen
Environmental Assessment Services, LLC, NOAA
Fisheries, Richland, WA, USA

Keywords Bycatch · Discards · Fisheries · Observer program · Data quality

Introduction

Bycatch, which for the purposes of this paper is defined as the discarded (live or dead) catch of living marine resources due to an encounter with fishing gear, continues to be a challenge to the sustainable management of fisheries throughout the world (Crowder and Murawski 1998; Hall et al. 2000; Lewison et al. 2004). Internationally, the term “bycatch” often encompasses retained incidental catch as well as discards; however, the analysis done in this paper covers discards and does not include retained incidental catch. Fishermen may choose to discard marine resources because the catch includes protected or undesired species, or individuals that may be outside the acceptable size or quality range (Eliassen et al. 2014). Fishermen may also choose to discard catch because of quota restrictions on target species. Bycatch can increase fish mortality if post-release mortality is high (e.g., Talwar et al. 2017); negatively affect industry profits (Patrick and Benaka 2013); and, when combined with overfishing, threaten food security (Srinivasan et al. 2010). Thus, many countries have identified bycatch reduction as a management objective for achieving sustainable fisheries, including:

- The United States, via the Magnuson-Stevens Fishery Conservation and Management Act, reauthorized in 2006 (Benaka et al. 2012);
- The European Union, via its Common Fisheries Policy that includes a Landing Obligation introduced in 2013 (Uhlmann et al. 2019);
- Kenya, via a trawling ban implemented in 2006 (Munga et al. 2012); and
- Australia, via its National Policy on Fisheries Bycatch, adopted in 1999 (Tuck et al. 2013).

Managers can more effectively minimize bycatch with more reliable bycatch estimates. Scientists have estimated bycatch regularly at a global level (Alverson et al. 1994; Kelleher 2005; Davies et al. 2009; Zeller et al. 2017; Perez Roda et al. 2019), revealing important differences in decadal trends in discards relative to catches. Multinational efforts to estimate bycatch have shown overall declines in recent decades (Zeller et al. 2017)—a promising result for fisheries management. However, global estimates are incomplete due to the lack of estimates of discards at the resolution of individual fisheries from many major

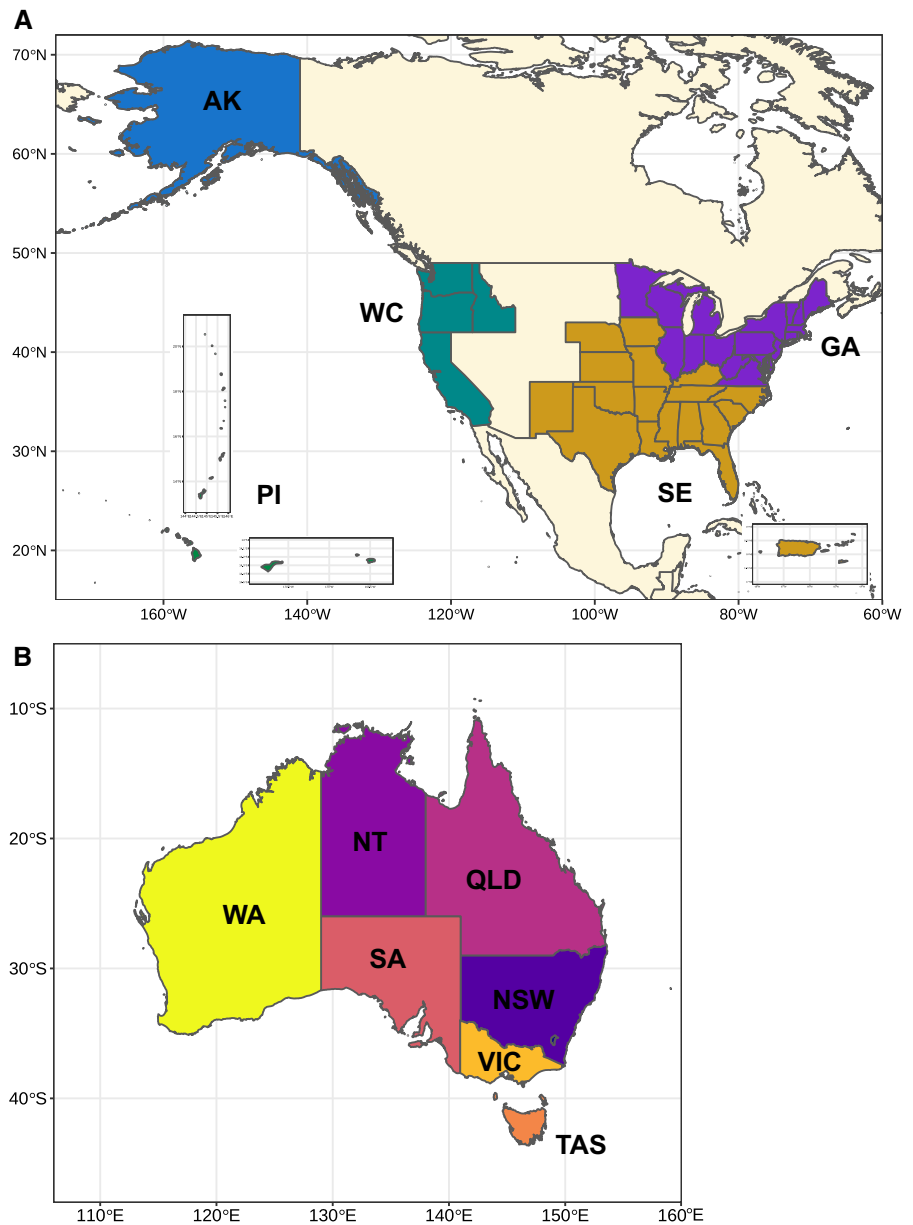
fishing countries (Perez Roda et al. 2019). Because the vast majority of discarding occurs within the exclusive economic zones of individual nations (Zeller et al. 2017), successful stewardship of living marine resources will benefit greatly from more comprehensive bycatch estimates at the national level (e.g., Harrington et al. 2005; NMFS 2011; Cahalan et al. 2016; Kennelly 2020; Savoca et al. 2020), as well as estimates for specific types of fisheries and species (e.g., Gilman et al. 2017; Melvin et al. 2019).

Bycatch estimates are essential to fisheries management, and fishery managers and scientists should assess the quality of bycatch estimation programs and procedures in order to understand the precision and limitations of such estimates, and to track improvements in them (or declines) over time. Regional level (Werner et al. 2019) and national (Desfosse et al. 2012) assessments have occurred to a limited extent in the United States. Regularly scheduled assessments will help scientists and managers identify data-deficient fisheries, improve estimation methodology, and allocate funds for bycatch monitoring programs when resources are limited (Desfosse et al. 2012). In addition, if managers in other countries adopt consistent assessments of the quality of bycatch data collection and estimates (e.g., Kennelly 2020), they will be able to compare methods among nations and help ensure additional confidence in the accuracy of those estimates where warranted.

NOAA Fisheries’ U.S. National Bycatch Report (NBR) First Edition documented bycatch estimates and estimation methods for commercial fisheries based on 2005 data (NMFS 2011; Brooke et al. 2012). This work included a standardized “Tier Classification System” (TCS) for evaluating the bycatch data sources and estimation methods for U.S. commercial fisheries (Desfosse et al. 2012). Although NOAA Fisheries developed the TCS for U.S. fisheries, analysts can also apply it internationally with minor modifications. Because the NBR First Edition (NMFS 2011) was a national compilation of bycatch estimates for resources managed by NOAA Fisheries, that report (and therefore the analyses in this paper) do not include state fisheries not managed by NOAA Fisheries.

This study illustrates the usefulness of the TCS to track changes in bycatch data collection and estimation methods over a 10-year period in the five NOAA fisheries management regions (Fig. 1a). In addition,

Fig. 1 Map of **a** NOAA Fisheries management regions and **b** Australian fishery jurisdictions included in this study. *AK* Alaska Region, *GA* Greater Atlantic Region, *WC* West Coast Region, *PI* Pacific Islands Region (insets show Guam, Northern Mariana Islands, and American Samoa), *SE* Southeast Region (inset shows Puerto Rico and the U.S. Virgin Islands), *NT* Northern Territory, *QLD* Queensland, *NSW* New South Wales, *TAS* Tasmania, *VIC* Victoria, *WA* Western Australia, *SA* South Australia. The eighth Australian jurisdiction is the Commonwealth government based in Canberra that manages mostly fisheries outside state and territory waters



the TCS was applied to fisheries in all Australian fishery jurisdictions to illustrate the international applicability of the system (Fig. 1b). NOAA Fisheries experts applied the TCS in 2005 for 158 fisheries and in 2015 for 130 fisheries (Table 1). Analysts consolidated some fisheries for the 2015 process to better reflect fishery names used by managers and eliminate inactive fisheries. Both analyses scored bycatch data and estimation processes in terms of fish (including crustaceans and molluscs) bycatch, marine mammal

bycatch, and other protected resources (including sea turtles and seabirds) bycatch. This paper focuses on the data and estimation process for fish bycatch because the data are available on a more consistent basis than for rarer protected species (for more information on the challenges of developing bycatch estimates for rare-event species such as sea turtles and seabirds, see Appendix 4 of NMFS 2004). For Australia, the TCS was used to evaluate bycatch estimates for all 307 commercial fisheries throughout

Table 1 Number of fisheries per NOAA Fisheries management region scored using the tier classification system in 2005 and 2015, and the number of fisheries comparisons completed in each region. The number of scored fisheries was generally lower for each region in 2015 due to consolidation of fisheries

Region	2005	2015	Number of fisheries comparisons
Alaska	27	17	27 ^a
Greater Atlantic	50	54	40 ^b
Pacific Islands	18	13	11
Southeast	27	23	20 ^a
West Coast	36	23	18
Total	158	130	114

^aDue to fishery consolidation, some 2015 fisheries were compared to multiple 2005 fisheries

^bSome 2005 fisheries were compared to multiple 2015 fisheries

Australia’s eight fishery jurisdictions as part of a project designed to develop a National Bycatch Reporting System for Australia (see Kennelly 2020 for additional information).

Methods

Regional teams of NOAA Fisheries bycatch experts used TCS criteria to evaluate regional fisheries based on available data and estimation procedures for 2005 and 2015. (NMFS 2011 and Desfosse et al. 2012 provide a detailed description of the criteria.) Table 1 lists the number of regional fisheries evaluated in 2005 and 2015, as well as the number of fisheries comparisons completed in each region. Only comparable fisheries scored in both 2005 and 2015 were included in the temporal comparison of tier scores. Evaluation team members from NOAA Fisheries Science Centers and Regional and Headquarters Offices met at national workshops to discuss scores and ensure the use of consistent approaches.

Fisheries were scored using a 72-point scale and then sorted into tiers ranging from 0 to 4 (Table 2). Tier classification criteria are broken into two general groups. The first group of criteria focus on adequacy of data to generate bycatch estimates with relatively low levels of uncertainty, and the sum of these criteria scores is called the “data score.” The second group of criteria focus on estimate quality, and the sum of these criteria scores is called the “analytical score.”

for evaluation purposes and elimination of inactive fisheries from the scoring process. Only comparable fisheries scored in both years were included in the temporal comparison of fishery score changes

Data adequacy, or data score, criteria include the following items and associated maximum scores:

- Longevity of observer program (5 points)
- Sampling frame (3 points)
- Sampling design (vessels/permits/licenses, trips, and hauls) (12 points)
- Design implementation (8 points)
- Data quality control (5 points)
- Industry bycatch data (2 points)
- Supplemental data (10 points)
- Database/information technology considerations (2 points)

Estimate quality, or analytical score, criteria include the following items and associated maximum scores:

- Assumptions identified, tested, and appropriate (10 points)
- Peer-reviewed/published (8 points)
- Statistical bias of estimators (factors considered include measures of association, cross validation, and overfitting) (4 points)
- Measures of uncertainty (3 points)

NOAA Fisheries updated scoring guidance for a few criteria for the 2015 assessment (Supplemental Table 1). These updates included changes to time-frames (e.g., for the length-of-observer program criterion) to reflect bycatch estimation practices through 2015 as opposed to through 2005. In addition, NOAA Fisheries modified the criterion regarding haul sampling design between 2005 and 2015 to create two

Table 2 Definitions of tiers and corresponding total point ranges for classifying fisheries according to the quality of bycatch data and the reliability of bycatch estimation methods

Tier number	Description	Points
4	Bycatch estimates are based on reliable observer program data collected on an annual basis for at least the past five years, with partial or complete sampling frames and negligible or nonexistent program design deficiencies	66–72
3	Observer program data were collected over the past five years, although not necessarily on an annual basis, and sampling frames were either partial or incomplete	49–65
2	Bycatch estimates typically were based on inconsistent or unreliable information (e.g., self-reported logbooks), but current or recent observer data may have been available for some of these fisheries	32–48
1	Observer data were not available or have not been collected during the past 10 years, or serious deficiencies or limitations in the design of the observer program were identified	1–31
0	Bycatch data collection programs have not been implemented	0

categories of random sampling schemes or probability-based sampling and remove the score for the census of hauls, because a census of hauls was neither realistic nor the best use of limited resources:

- No observer program, or sampling design does not support bycatch or total catch estimation (0 points).
- Opportunistic or haphazard sampling, including voluntary observer programs, to support bycatch or total catch estimation (1 point).
- Random sampling scheme or probability-based sampling with pilot/baseline observer coverage levels to support bycatch or total catch estimation (2 points).
- Random sampling scheme or probability-based sampling with adequate observer coverage levels to support bycatch or total catch estimation (3 points).
- Close to complete census of hauls with estimation (4 points).

For the design implementation criterion in the 2015 assessment, the TCS allowed scores of 0, 1, or 2, based on the level of possible bias. For 2005, the TCS only allowed scores of 0 or 2 (see Supplemental Table 1). NOAA Fisheries also updated the reference years used to determine scoring for industry bycatch data:

- No industry bycatch data available, or industry bycatch data are available but not used as a basis for bycatch estimates (0 points).
- Industry bycatch data available prior to 2005 and are used as a basis for bycatch estimates (1 point).

- Industry bycatch data available between 2005 and 2015 and are used as a basis for bycatch estimates, or, observer data are used as the basis of bycatch estimates (2 points).

The criterion regarding database/information technology did not reflect electronic monitoring or reporting considerations for either assessment.

NOAA Fisheries also slightly modified the measures of uncertainty criterion for the 2015 assessment. Guidelines provided to evaluation team members described this criterion as evaluating whether measures of uncertainty (e.g., coefficients of variation (CVs)) were calculated at various levels of sampling (vessel/permit/license, trip, and haul), in addition to the precision of the CV. For the 2005 review, scoring guidelines directed reviewers to award the top score of three points if measures of uncertainty were calculated at all levels of sampling. For the 2015 review, revised guidelines directed reviewers to provide the top score of three points if measures of uncertainty were calculated, but not necessarily at all levels of sampling, and the majority of estimates had low uncertainty. The guidelines suggested that a low level of uncertainty for fish bycatch estimates could be CVs of 0.3 or lower.

In designing a bycatch reporting system for Australia, Kennelly (2020) applied the various parts of the above scoring system to all fisheries jurisdictions that manage fisheries in the country, specifically, the states of New South Wales, Queensland, Tasmania, Victoria, Western Australia, South Australia, the Commonwealth government and the Northern Territory. In

order to adapt the U.S. TCS evaluation timeframes for use in Australia, Kennelly (2020) made minor changes to the longevity-of-observer-program criterion.

Finally, we compared tier scores to bycatch estimates for each fishery to further understand which U.S. and Australian fisheries had the greatest need for additional resources to improve bycatch data quality and reduce uncertainty levels related to bycatch estimates. The associated tier level provides a metric of confidence in the bycatch estimate, while the magnitude of bycatch indicates a need for monitoring and assessment.

Results

Scores provided by the U.S. evaluation team members showed a decrease in the percentage of tier 0 fisheries between 2005 and 2015, and an increase in the percentage of tier 4 fisheries between 2005 and 2015 (Fig. 2). The Alaska Region had the highest percentage of tier 4 fisheries, while the Southeast Region had the lowest percentage (0%). The Southeast Region also had the highest percentage of tier 0 fisheries, while the Alaska Region and Greater Atlantic Region both lacked fisheries with the lowest tier score. The total number of fisheries assessed per region also varied in 2015, with the highest number of fisheries in the Greater Atlantic Region ($n = 54$), followed by the West Coast and Southeast Regions which both had the same number of scored fisheries ($n = 23$). Finally, the Alaska and Pacific Islands Regions had the lowest number of scored fisheries ($n = 17$ and $n = 13$, respectively). Tier 0 scores decreased mostly due to the establishment of new pilot observer programs between 2005 and 2015, and tier 4 fisheries increased mostly due to the restructuring of regional observer programs to address bias issues and the introduction of new bycatch data quality controls. The following sections describe changes in tier classification scores in detail by each U.S. region between 2005 and 2015 for comparable fisheries scored in both years (Table 1).

Changes in Pacific Islands Region tier scores

For the Pacific Islands Region, the majority of fisheries tier classification scores stayed the same over the 10 years (Supplemental Table 2). However, four

fisheries, including the three major pelagic longline fisheries in the region, had increased tier scores in 2015 (Fig. 3c).

The recreational Hawai‘i bottomfish fishery was classified as tier 0 in 2005 and increased to tier 1 in 2015. This change was due in part to the availability of industry data or observer data for bycatch estimation between 2005 and 2015. The American Samoa pelagic longline fishery went from a tier score of 1 in 2005 to a score of 4 in 2015. Scores for each category increased from near zero to the maximum value for almost all scoring subsections due to the implementation of a comprehensive observer program providing approximately 20% coverage on an annual basis.

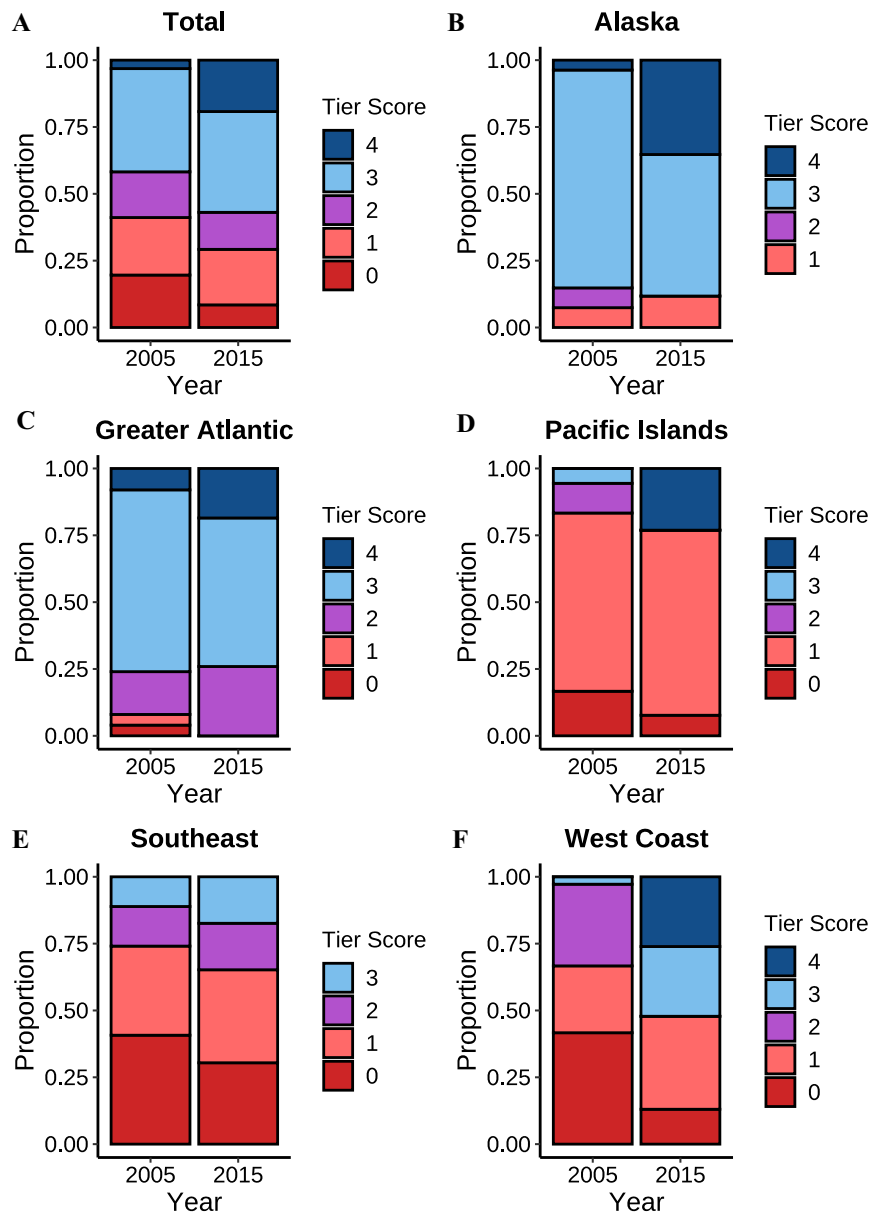
The Hawai‘i-based deep-set pelagic longline fishery for tuna, scored at tier 2 in 2005, increased to tier 4 in 2015. This increase was largely due to improved analytical approaches, such as identifying and testing assumptions to ensure they were appropriate. The tier score for the Hawai‘i-based shallow-set pelagic longline fishery for swordfish increased from a 3 to a 4 between 2005 and 2015. This increase, like that for the deep-set longline fishery, was largely due to improved analytical approaches, as well as using peer-reviewed sampling designs and analytical methods, and estimators with little to no statistical bias.

Changes in Alaska Region tier scores

In the Alaska Region, we compared fishery tier scores using the 2005 list of fisheries from the first edition of the NBR (NMFS 2011), and the 2015 list of fisheries from Table 5.1 in the NBR Update 3 (Benaka et al. 2019). In the NBR Update 3, some groundfish and Pacific halibut fisheries were consolidated to represent how fisheries are managed and prosecuted, rather than classifying fisheries based on factors such as the gear or area fished. Of the total fisheries analyzed, 11 were classified in a higher tier in 2015, while the remaining 16 scores did not change (Fig. 3a, Supplemental Table 3).

Two Gulf of Alaska (GOA) fisheries that had tier scores of 2 in 2005 increased to tier 3 in 2015. The GOA flatfish (shallow water flatfish) trawl fishery, which was included in the GOA non-pelagic trawl fishery for 2015, had improved scores for sampling frame and design, observer bias, and analytical methods. The GOA Pacific cod pot fishery had similar improvements in its bycatch estimation methods,

Fig. 2 Proportion of tier classification system scores across all of the NOAA Fisheries management regions **a** in 2005 (n = 158) and 2015 (n = 130) as well as in 2005 and 2015 for each of the NOAA Fisheries management regions **b–f**. Regional samples sizes in 2005 and 2015 are the same as in Table 1



although this fishery also had reduced scores due to the limited spatial coverage of its sampling design and increased observer bias.

Analysts classified two other GOA fisheries that had been scored as tier 3 fisheries in 2005 as tier 4 fisheries in 2015. The GOA sablefish trawl and rockfish trawl fisheries both improved their sampling frames and designs, as well as multiple aspects of their analytical approaches.

Analysts consolidated five separate 2005 fisheries into the Bering Sea/Aleutian Islands (BSAI) non-pollock trawl fishery for 2015, which increased the tier score from 3 to 4 in 2015. The BSAI rock sole trawl, BSAI flatfish group trawl, and BSAI Pacific Ocean perch trawl fisheries all had improved sampling frames and designs in 2015. They also obtained maximum scores for identifying, testing, and resolving all assumptions and using estimators with negligible statistical biases. The BSAI sablefish trawl fishery

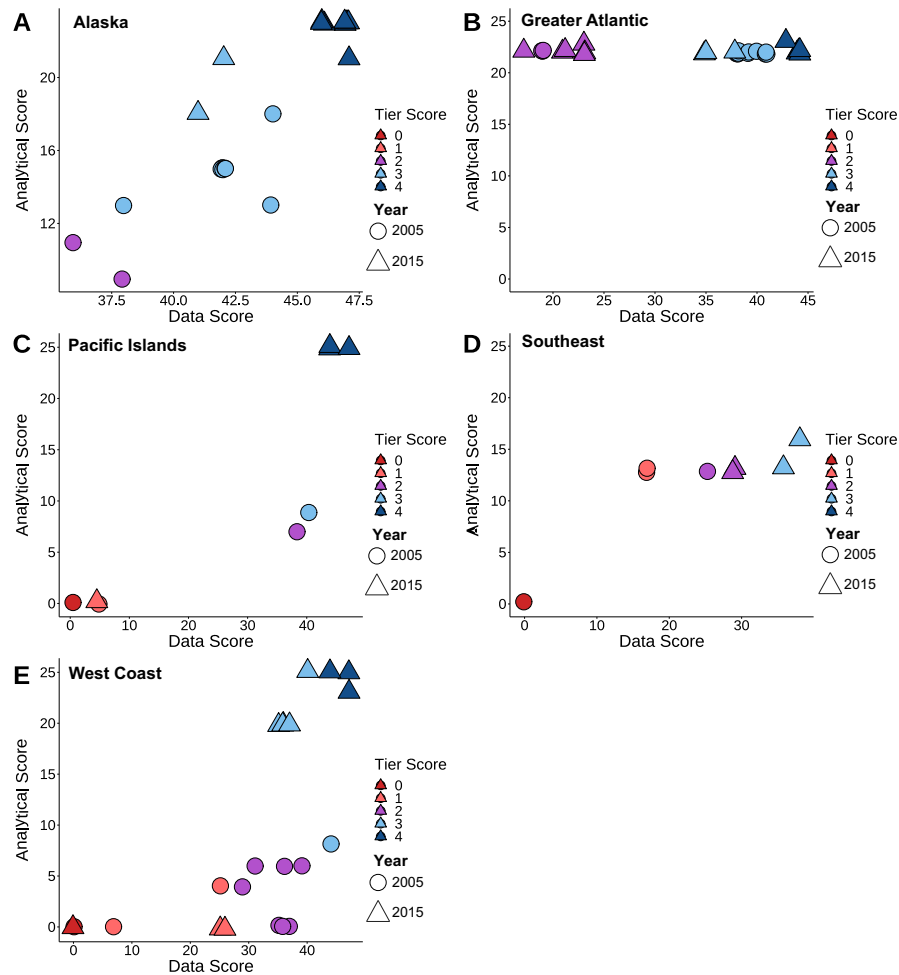


Fig. 3 Data, analytical, and tier score changes for US Regional fisheries in 2005 (circles) and 2015 (triangles) for **a** Alaska Region, **b** Greater Atlantic Region, **c** Pacific Islands Region,

d Southeast Region, and **e** West Coast Region. Within each region, fisheries with identical scores in 2005 and 2015 are not shown

achieved a higher score by making similar changes, and also by using an externally peer-reviewed analytical approach. Analysts also assigned the Aleutian Islands, Eastern Bering Sea Atka mackerel trawl fishery a higher score in 2015 due to an improved sampling design and similar improvements in the analytical approach.

Lastly, two BSAI fisheries that were consolidated into the BSAI catcher processor longline fishery in 2015 also increased from tier 3 to tier 4. Similar to the other BSAI fisheries, the BSAI Pacific cod longline and Greenland turbot longline employed complete sampling frames and sampled nearly all vessels, trips, and hauls, resulting in increased fishery tier scores in 2015.

Changes in West Coast Region tier scores

Due to changes in fishery classification over the 10-year period, we were only able to directly compare 18 fisheries between 2005 and 2015 for the West Coast region (Supplemental Table 4). Bycatch data and estimation methodology for these fisheries largely improved between 2005 and 2015, with nine fisheries achieving a higher tier classification in 2015 (Fig. 3e). However, analysts downgraded the scores of two fisheries in 2015 compared to 2005. The California swordfish harpoon fishery decreased from tier 1 to tier 0 due to a loss of supplemental data. The California small mesh drift gillnet fishery (mesh size > 3.5 in and < 14 in) decreased from tier 2 to tier 1 due to

lower scores for supplemental data as well as the loss of an observer program capable of random sampling.

The North Pacific albacore baitboat—pole and line fishery was scored as tier 0 in 2005, but increased to tier 1 in 2015. The implementation of an observer program in some form and increased availability of supplemental data largely contributed to the increased score.

Three fisheries had increased classification scores in 2015 mainly due to improvements in sampling design, design implementation, and analytical methods. Of these, the California halibut trawl had the highest score increase, moving from tier 1 to tier 3 due to the implementation of a complete sampling frame with a random sampling scheme. Both the (1) West Coast groundfish non-trawl gear: non-endorsed fixed gear fishery and the (2) California, Oregon nearshore rockfish fishery increased from tier 2 to tier 3 in 2015.

Substantial improvements in observer sampling designs and analytical approaches for two additional fisheries resulted in tier score increases from tier 2 to tier 4. The West Coast limited entry bottom trawl, groundfish bottom trawl fishery (renamed the West Coast limited entry bottom trawl fishery for 2015) and the California pelagic longline—deep set fishery (renamed the California pelagic longline fishery for 2015) both had considerably higher scores in 2015, in part due to improvements in analytical approaches such as the identification of critical assumptions. Extensive supplemental data were also available for the pelagic longline fishery in 2015, contributing to the increased tier classification.

Three additional fisheries increased one tier level, largely due to improvements in analytical methods. The West Coast groundfish non-trawl gear: limited entry sablefish-endorsed fixed gear fishery and the California/Oregon drift gillnet (mesh size > 14 inches) for swordfish and thresher shark fishery both increased from tier 2 to tier 3. The West Coast mid-water trawl for whiting, at-sea processing fishery obtained a tier increase from 3 to 4. All three fisheries used estimators with negligible statistical bias, and the majority of these estimates had low uncertainty, in addition to other analytical improvements for 2015.

Changes in Greater Atlantic Region tier scores

Similar to the West Coast fisheries, the Greater Atlantic fisheries evaluated in 2015 also experienced

significant fishery reorganization and some addition/subtraction of fisheries. Still, analysts scored 40 comparable fisheries in both 2005 and 2015, enabling the comparison of bycatch data collection and estimation methods (Supplemental Table 5). Of these, eight fisheries had a higher tier score in 2015 while six fisheries had a lower score (Fig. 3b). This region thus had the largest number of fisheries that had reduced quality of bycatch data and estimation methods over the 10 years.

All six fisheries that had lower scores in 2015 decreased from tier 3 to tier 2, predominantly due to the lack of an observer program that supports the estimation of bycatch and total catch. Fisheries classified in lower tiers included the Mid-Atlantic limited access scallop trawl (which was split into the Mid-Atlantic limited access open and closed area scallop trawl fisheries before 2015), Mid-Atlantic purse seine, New England Scottish seine, New England small-mesh gillnet, and New England shrimp trawl fisheries. The loss of industry bycatch data for bycatch estimation also contributed to declines in scores for all six fisheries.

The Mid-Atlantic lobster pot, Mid-Atlantic clam/quahog dredge, and New England clam/quahog dredge fisheries all scored higher in 2015, rising from tier 2 to tier 3. NOAA Fisheries implemented observer programs for all three fisheries on one or more occasions between 2010 and 2015, enabling random sampling or probability-based sampling with at least baseline observer coverage. Industry bycatch data also became available for use in bycatch estimates, resulting in higher scores.

Finally, analysts scored five fisheries as tier 4 in 2015 compared to tier 3 in 2005. However, the actual point increases were small, ranging from only two to five points. This group of fisheries included the Mid-Atlantic large-mesh gillnet, Mid-Atlantic extra-large mesh gillnet, New England limited access open area scallop dredge, Mid-Atlantic large-mesh otter trawl, and Mid-Atlantic small-mesh otter trawl fisheries. Modest improvements to sampling frame, design, and/or implementation resulted in slightly higher scores for these fisheries. For the New England limited access open area scallop dredge fishery, calculation of uncertainty measures increased in all sampling levels in 2015, slightly improving the total score.

Changes in Southeast Region tier scores

In the Southeast Region, analysts evaluated 20 fisheries in both 2005 and 2015 (Supplemental Table 6). Of these, four were classified in higher tiers for 2015, while none were classified in lower tiers (Fig. 3d). The largest tier increase occurred for the North Carolina coastal gillnet fishery, which was in tier 0 in 2005 and in tier 3 in 2015. This fishery was combined with the large coastal and small coastal shark aggregates (drift, strike, and bottom gillnet) fishery (a tier 3 fishery) to become the Southeastern Atlantic coastal gillnet fishery between 2005 and 2015. Implementation of an annual observer program contributed to the improved quality of bycatch data. The availability of extensive supplemental data and an improved analytical approach (such as identifying minor assumptions) also helped raise the score for this fishery.

Analysts combined the Gulf of Mexico marine shrimp butterfly nets and skimmer trawl fisheries into one fishery for the 2015 analysis, with a corresponding tier score increase from tier 1 to tier 2. A pilot observer program, which enabled random or probability-based sampling and comprehensive data quality control measures, largely led to the increased tier scores. However, observer bias also increased, causing the loss of one point each for both fisheries.

Lastly, the Gulf of Mexico shrimp trawl fishery was assigned to tier 3 in 2015—an increase from tier 2 in 2005. A complete sampling frame with representative spatial and temporal coverage, as well as implementation of a random or probability-based sampling scheme yielded an improved score for this fishery. Data quality control and database/IT considerations also improved during the 10-year period.

Application of the tier classification system to Australian fisheries

Unlike the U.S. analyses, the Australian application of the TCS has only been completed once, precluding a comparison of any changes in tiers for fisheries through time. Despite this, the Australian application of the method has yielded useful baseline information on bycatch data quality and estimation methods.

Of the 307 commercial fisheries/methods examined in Australian fishery jurisdictions, 266 featured some determination of discard estimates (Kennelly 2020). The remaining 41 fisheries were not examined because

discards were determined to be negligible due to the fisheries' highly selective harvesting methods (e.g., hand-gathering, spearing, etc.). Overall, the 266 fisheries/methods assessed resulted in the following scores: 198 fisheries at tier 1; 49 fisheries at tier 2; and 19 fisheries at tier 3 (Fig. 4). No fisheries received a tier score of 4. The majority of fisheries in all 8 jurisdictions were scored as tier 1. Commonwealth managed fisheries had the highest percentage of tier 3 fisheries, while the Northern Territory, New South Wales, Victoria, Queensland and Tasmania had no tier 3 fisheries.

Comparison of tier scores and reported estimates of bycatch levels in U.S. and Australian fisheries

Publicly available data on 2015 U.S. fish bycatch estimates by fishery were obtained from the National Bycatch Report Database System (<https://www.st.nmfs.noaa.gov/apex/f?p=243%3A101%3A12580571048861%3A%3A%3A%3A%3A>). Figure 5 shows each NBR fishery's estimated bycatch in pounds of dead fish, as well as its tier score for 2015. The database referenced above contains estimated bycatch weights, as well as fishery landings, for each of the NBR fisheries. (NBR fisheries for which bycatch was estimated partially or wholly as individuals as opposed to pounds are not included.) For U.S. fisheries, the majority of reported bycatch estimates are from fisheries with high-quality bycatch data and estimation methods (tiers 3 or 4). There are only three NBR fisheries with reported bycatch estimates in the Pacific Islands Region, and all of these fisheries have the highest tier score of 4. All fisheries with reported discards from the Alaska, West Coast, and Greater Atlantic Regions were scored as either tier 3 or tier 4. The Southeast Region was the only region with a Tier 2 fishery in Fig. 5. Notably, this fishery—the Southeastern Atlantic Shrimp Trawl fishery—had the second-highest level of estimated discards in the Southeast for 2015.

Figure 6 compares extrapolated discard weight estimates (in metric tonnes) by fishery to overall tier scores in each Australian jurisdiction, based on data published in Kennelly (2020). In contrast to the U.S. results, the fisheries with reported bycatch estimates range from tier 1 to tier 3 fisheries. All Commonwealth-managed fisheries that have significant levels of discards are also tier 3 fisheries. This is also the case

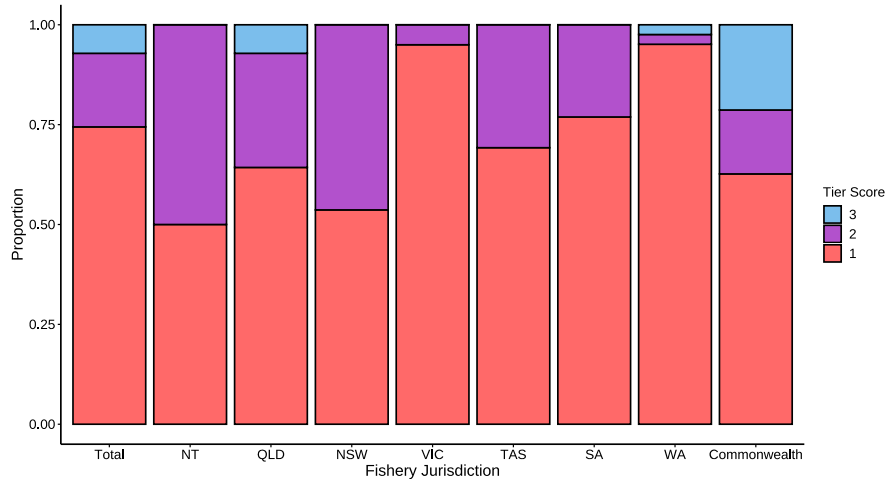


Fig. 4 Proportion of tier classification system scores for all of the scored Australian fisheries ($n = 266$) in each of Australia's fishery jurisdictions

for Queensland's (and Australia's) highest discarding fishery which has a tier score of 3. But for the other jurisdictions, despite quite high levels of discards, tier scores for their fisheries are only 1 or 2.

Discussion

The results of this study demonstrate that the TCS is a robust and replicable assessment to determine (1) the adequacy of data to generate bycatch estimates with relatively low levels of uncertainty and (2) the quality of estimation methods that support bycatch calculations. The work described here has shown how this system was applied by NOAA Fisheries experts to over 100 U.S. fisheries in 2005 and 2015. A temporal comparison of tier scores showed that data adequacy and estimation quality generally increased in all NOAA Fisheries' management regions between 2005 and 2015. In addition, the TCS was successfully applied to Australian fisheries in all eight jurisdictions, validating the international applicability of assessment for bycatch data and estimation quality.

The widespread improvements in U.S. fishery tier scores were expected because NOAA Fisheries and its industry partners have invested a great deal of resources to provide high-quality observer coverage, with total government and industry expenditures ranging from over \$40 million in 2005 to over \$70 million in 2017 (NMFS 2019). This increased

financial investment has led to increasingly sophisticated and comprehensive observer programs that have been implemented by Regional Fishery Management Councils and NOAA Fisheries during this period.

Although U.S. government and industry support for fishery monitoring is substantial, resources are limited, requiring some prioritization of observer coverage. In the Pacific Islands, three fisheries had improved tier scores in 2015, while the majority of fishery scores were unchanged from 2005 to 2015 (Figs. 2, 3c). NOAA Fisheries focuses its Pacific Islands observer resources on three longline fisheries that interact with protected species and represent the majority of the region's commercial landings. These same three longline fisheries are also the only fisheries with reported fish discard estimates in this region (Fig. 5). (For more information about Pacific Islands longline fisheries, see chapter 4 of WPRFMC 2009).

A similar pattern was observed in the Southeast Region, where only a few fisheries had higher tier scores in 2015 (Fig. 3d). NOAA Fisheries allocates the majority of its Southeast Region sea days to a highly migratory species pelagic longline fishery and large-scale shrimp trawl fishery that have higher levels of estimated protected species bycatch, and fewer sea days to reef fish and shark fisheries that have generally lower levels of protected species interactions (NMFS 2020a, b). Despite this uneven distribution of monitoring resources, NOAA Fisheries is able to estimate

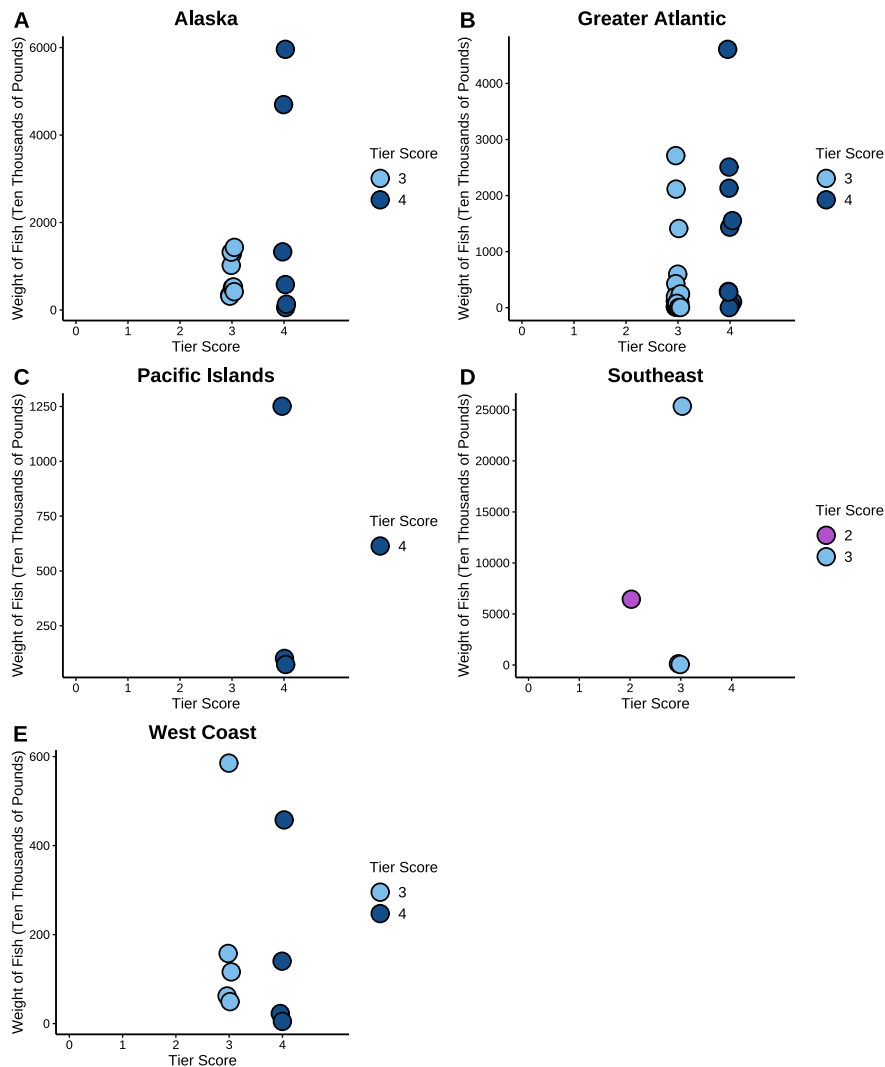


Fig. 5 Scatter plots of the weight of discards (in ten thousands of pounds, y-axis) for each U.S. fishery with bycatch estimates for 2015 in the National Bycatch Report (Benaka et al. 2019), with tier score on the x-axis. Tier score is shown by the symbol fill color. Fisheries from the **a** Alaska, **b** Greater Atlantic,

c Pacific Islands, **d** Southeast, and **e** West Coast Regions are included. National Bycatch Report fisheries for which bycatch was estimated partially or wholly as individuals as opposed to pounds are not included

bycatch for a variety of Southeast Region fisheries and gear types (Benaka et al. 2019).

Between 2005 and 2015, tier scores improved for 11 Alaska fisheries (Fig. 3a), with most fisheries scoring tier 3 or above (Fig. 2). The final rule to implement Amendment 86 to the Bering Sea and Aleutian Islands Fishery Management Plan (FMP) and Amendment 76 to the Gulf of Alaska FMP (77 FR 70062, November 21, 2012) greatly improved bycatch monitoring in this region. This rule provided the

framework for promulgation of annual deployment plans (ADPs) that describe how NOAA Fisheries intends to assign at-sea and shoreside fishery observers and electronic monitoring to North Pacific halibut and groundfish vessels and processing plants. The ADPs, which describe the science-driven method for observer deployment to support statistically reliable data collection, have enhanced coverage and sampling designs for selected Alaska fisheries. (For more information, see NMFS 2019).

West Coast fisheries, through the leadership of the Pacific Fishery Management Council and NOAA Fisheries' West Coast Groundfish Observer Program (WCGOP), greatly increased observer coverage between 2005 and 2015 (Figs. 2, 3e). This was accomplished through the transition of the limited entry trawl fleet, the largest groundfish fishery on the West Coast, to a catch shares management system in 2011 (75 FR 60867; October 1, 2010). The catch shares program requires participating vessels to carry an observer onboard 100% of the time the vessel is active in the catch share fishery. This increased observer coverage rate has resulted in enhanced bycatch estimate quality (Fig. 5) and the creation of a variety of data products that the WCGOP publishes on its website.¹

Although the Greater Atlantic Region had the highest number of fisheries with lower scores in 2015 compared to 2005 (Fig. 3b), all of the fisheries in this region had tier scores of 2 or above in 2015 (Fig. 2), demonstrating the overall high quality of bycatch data and estimation methods in this region. The Greater Atlantic Region strives to achieve a relatively high level of precision (30% CV) for fish bycatch estimates in its Northeast and Mid-Atlantic fisheries. However, NOAA Fisheries does not always have sufficient funding to implement a broad level of observer coverage that meets the 30% CV precision goal. In such cases, NOAA Fisheries may decrease observer coverage for particular fisheries, a factor that contributed in part to lower tier scores for some Greater Atlantic Region fisheries from 2005 to 2015. Nevertheless, this region has the highest number of fisheries with reported discard estimates (Fig. 5).

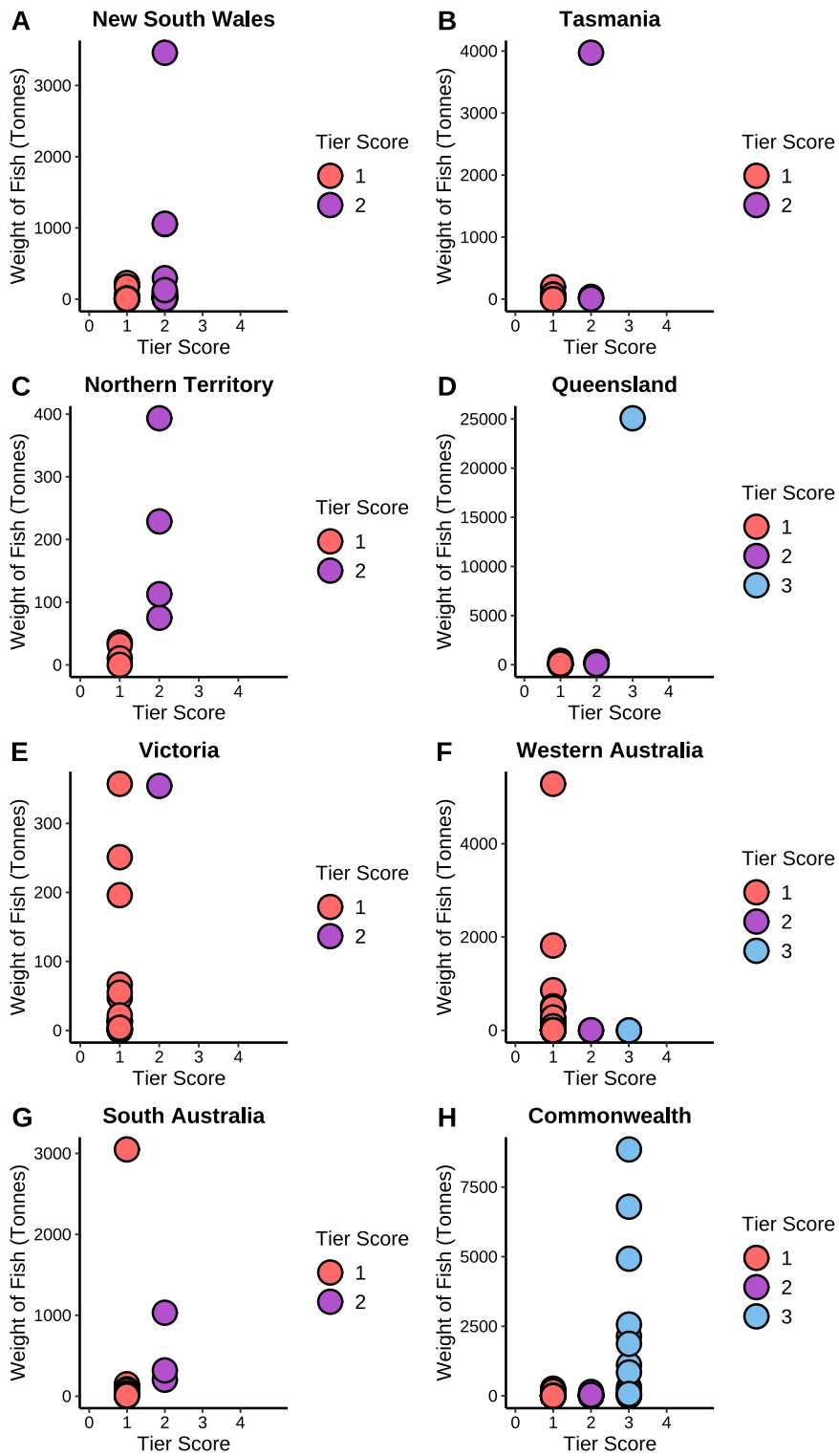
Some NOAA Fisheries observer programs and Science Centers have become increasingly interested in trying to determine whether “observer bias” exists, and if so, how that bias affects discard estimation. The TCS focuses on statistical bias to a relatively small extent, with a two-point difference in scoring for estimators with high statistical bias compared to estimators with negligible statistical bias. Future TCS evaluations should consider providing greater emphasis on bias due to its potential to influence, for example, fisheries stock assessments (see also Faunce and Barbeaux 2011).

¹ https://www.nwfsc.noaa.gov/research/divisions/fram/observation/data_products/sector_products.cfm

The Australian results of the application of the TCS were, for states, lower than the U.S. results and mainly reflect the far fewer (and mostly out-of-date) observer programs occurring in Australia's state jurisdictions. The exception is data for Australia's Commonwealth fisheries where quite comprehensive and long-running observer programs resulted in many tier 3 fisheries, making scores for this jurisdiction comparable to scores for U.S. federal fisheries. Such a result is understandable as U.S. and Australian federally managed fisheries tend to be larger (and more valuable) than those in state jurisdictions (NMFS 2020a, b; Steven et al. 2020, Kennelly 2020), which can justify a larger investment in federal government resources in observer programs. The main value of these tier metrics for Australia's fisheries, however, is in providing a baseline measure against which future scores can be compared, to allow one to gauge improvements (or diminishments) in information about the quality of bycatch estimates over time. In addition, these scores can guide future investments in bycatch monitoring programs for Australian fisheries, for example, in the oceanic prawn trawl fisheries that have high levels of discards (Kennelly 2020).

In Australia, standardized reporting of fishery discards and tracking of the quality of bycatch data collection and estimation methods is relatively new (Kennelly 2020). Therefore, discard estimates from fisheries with lower tier scores—particularly tier scores 1 and 2—should be interpreted with caution. Regardless, comparing the weight of discarded fish to the related fishery tier score can aid in prioritizing limited resources for bycatch monitoring efforts. Many fisheries in Australia have no estimated bycatch; therefore, increasing monitoring for these fisheries would not be a priority. But fisheries with low tier scores and high estimated discards (such as the many fisheries identified in Fig. 6 for New South Wales, Tasmania, the Northern Territory, Queensland, South Australia and Western Australia) could be priorities for improving bycatch monitoring and estimation methods.

The U.S. and Australian case studies described in this paper have demonstrated that the TCS has broad applicability as a tool for measuring the quality of bycatch estimates for fisheries. TCS results also could contribute, along with other inputs, to budget prioritization decisions related to fisheries monitoring. Although estimates of discards for all countries are



◀ **Fig. 6** Scatter plots of the weight of discards (in tonnes, y-axis) for each Australian fishery, with tier score on the x-axis. Tier score is also shown by the symbol fill color. Only fisheries with reported bycatch estimates in the **a** New South Wales, **b** Tasmania, **c** Northern Territory, **d** Queensland, **e** Victoria, **f** Western Australia, **g** South Australia, and **h** Commonwealth Jurisdictions are included

publicly available from the *Sea Around Us* project (Zeller et al. 2016) and the Food and Agriculture Organization of the United Nations (Perez-Roda et al. 2019), the completeness of the data used to produce those estimates is known to vary considerably. The TCS has strong potential to assist with such global efforts and, indeed, any country or jurisdiction that seeks to evaluate the adequacy and quality of its methods to estimate bycatch and track changes in performance over time.

Acknowledgements This work would not have been possible without the NOAA Fisheries experts who provided scores for each fishery, including Russell Brown, Susan Wigley, Paul Rago, Chris Tholke, Chris Orphanides, Elizabeth Scott-Denton, John Carlson, Randy Blankinship, Joseph Desfosse, Jon McVeigh, Kayleigh Somers, Vanessa Tuttle, Charles Villafana, Jennifer Mondragon, Cathy Tide, Keith Bigelow, and Aimee Hoover. The authors also recognize Samantha Brooke, Lisa Desfosse, and Bill Karp, who led the development of the Tier Classification System. The authors would like to thank all the hard-working and tenacious fisheries observers who provide the data needed to keep fisheries sustainable.

Author contributions LRB conceived and designed the study. ANC, SJK, and NAO analyzed the data. LRB, ANC, and SJK wrote the manuscript with input and edits from all authors.

Funding NOAA Fisheries provided funding support for this effort, and the Fisheries Research and Development Corporation (Projects 2015-208 and 2018-114) funded the Australian component.

Data availability Supporting data are available in the supplemental tables for U.S. fisheries and in Kennelly 2020 for Australian fisheries. The datasets generated during and/or analysed during the current study are available in the National Bycatch Report Database System, <https://www.st.nmfs.noaa.gov/apex/f?p=243:101:2918142625939>.

Declarations

Conflict of interest The authors declare no conflict of interest or competing interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use,

sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alverson DL, Greenberg MH, Pope JG, Murawski SA (1994) A global assessment of fisheries bycatch and discards. FAO Fisheries Technical Paper. No. 339. Rome, FAO
- Benaka LR, Bullock D, Hoover AL, Olsen NA (ed) (2019) U.S. national bycatch report first edition update 3. U.S. Dept. of Commer., NOAA Tech. Memo
- Benaka LR, Cimo LF, Jenkins LD (2012) Bycatch provisions in the reauthorized Magnuson-Stevens Act. *Mar Fish Rev* 74(2):1–12
- Brooke SG, Desfosse LL, Karp WA (2012) Estimating overall fish bycatch in U.S. commercial fisheries. *Mar Fish Rev* 74(3):1–5
- Cahalan J, Faunce C, Bonney J, Swanson R (2016) A field test of fisheries observer sampling methods for estimation of at-sea discards. *Fish Res* 174:219–233
- Crowder LB, Murawski SA (1998) Fisheries bycatch: implications for management. *Fisheries* 23(6):8–17
- Davies RWD, Cripps SJ, Nickson A, Porter G (2009) Defining and estimating global marine fisheries bycatch. *Mar Policy* 33:661–672
- Desfosse LL, Karp WA, Brooke SG (2012) Evaluating the quality of bycatch data and bycatch estimates among disparate fisheries. *Mar Fish Rev* 74(3):6–13
- Eliassen SQ, Papadopoulou K-N, Vassilopoulou V, Catchpole TL (2014) Socio-economic and institutional incentives influencing fishers' behavior in relation to fishing practices and discard. *ICES J Mar Sci* 71(5):1298–1307
- Faunce CH, Barbeaux SJ (2011) The frequency and quantity of Alaskan groundfish catcher-vessel landings made with and without an observer. *ICES J Mar Sci* 68(8):1757–1763
- Gilman E, Suuronen P, Chaloupka M (2017) Discards in global tuna fisheries. *Mar Ecol Prog Ser* 582:231–252
- Hall MA, Alverson DL, Metzuzals KI (2000) By-catch: problems and solutions. *Mar Pollut Bull* 41:1–6
- Harrington JM, Myers RA, Rosenberg AA (2005) Wasted fishery resources: discarded by-catch in the USA. *Fish Fish* 6(4):350–361
- Kelleher K (2005) Discards in the world's marine fisheries: an update. FAO Fisheries Technical Paper. No. 470. Rome, FAO
- Kennelly SJ (2020) Bycatch bekknown: methodology for jurisdictional reporting of fisheries discards – Using Australia

- as a case study. *Fish Fish.* <https://doi.org/10.1111/faf.12494>
- Lewison RL, Crowder LB, Read AJ, Freeman SA (2004) Understanding impacts of fisheries bycatch on marine megafauna. *Trends Ecol Evol* 19(11):598–604
- Melvin EF, Dietrich KS, Suryan RM, Fitzgerald SM (2019) Lessons from seabird conservation in Alaskan longline fisheries. *Conserv Biol* 33(4):842–852
- Munga C, Ndegwa S, Fulanda B, Manyala J, Kimani E, Ohtomi J, Vanreusel A (2012) Bottom shrimp trawling impacts on species distribution and fishery dynamics; Ungwana Bay fishery in Kenya before and after the 2006 trawl ban. *Fish Sci* 78:209–219
- NMFS (National Marine Fisheries Service) (2004) Evaluating bycatch: a national approach to standardized bycatch monitoring programs. U.S. Department of Commerce, NOAA Technical Memorandum. NMFS-F/SPO-66
- NMFS (2011) U.S. National Bycatch Report. In: Karp WA, Desfosse LL, Brooke SG (eds). U.S. Department of Commerce., NOAA Technical Memorandum. NMFS-F/SPO-117E
- NMFS (2019) 2020 annual deployment plan for observers and electronic monitoring in the groundfish and halibut fisheries off Alaska. National Oceanic and Atmospheric Administration, Juneau, Alaska
- NMFS (2020) National observer program FY 2018 annual report. NOAA Tech. Memo. NMFS-F/SPO-206
- NMFS (2020) Fisheries of the United States, 2018. U.S. Department of Commerce, NOAA Current Fishery Statistics No. 2018
- Patrick WS, Benaka LR (2013) Estimating the economic impacts of bycatch in U.S. commercial fisheries. *Mar Policy* 38:470–475
- Pérez Roda MA, Gilman E, Huntington T, Kennelly SJ, Suuronen P, Chaloupka M, Medley P (ed) (2019) A third assessment of global marine fisheries discards. FAO Technical Paper No. 633. Rome, FAO
- Savoca MS, Brodie S, Welch H, Hoover A, Benaka LR, Bograd SJ, Hazen EL (2020) Comprehensive bycatch assessment in US fisheries for prioritizing management. *Nat Sustain.* <https://doi.org/10.1038/s41893-020-0506-9>
- Srinivasan UT, Cheung WWL, Watson R, Sumalia UR (2010) Food security implications of global marine catch losses due to overfishing. *J Bioecon* 12:183–200
- Steven AH, Mobsby D, Curtotti R (2020) Australian Fisheries and Aquaculture Statistics 2018. Fisheries Research and Development Corporation project 2019-093, ABARES, Canberra, April. CC BY 4.0. <https://doi.org/10.25814/5de0959d55bab>
- Talwar B, Brooks EJ, Mandelman JW, Grubbs RD (2017) Stress, post-release mortality, and recovery of commonly discarded deep-sea sharks caught on longlines. *Mar Ecol Prog Ser* 582:147–161
- Tuck GN, Knuckey I, Klaer NL (2013) Informing the review of the Commonwealth Policy on Fisheries Bycatch through assessing trends in bycatch of key Commonwealth fisheries. Fisheries Research and Development Corporation final report 2012/046
- Uhlmann SS, Ulrich C, Kennelly SJ (eds) (2019) The European Landing Obligation: reducing discards in complex, multi-species and multi-jurisdictional fisheries. Springer, Cham
- Werner S, Demarest C, Thunberg E (2019) Assessing distributional impacts stemming from a methodological change in the Northeast multispecies fishery combined coverage monitoring program. U.S. Department of Commerce, Northeast Fisheries Science Center Reference Documents. 19-14
- WPRMFC (Western Pacific Regional Fishery Management Council) (2009) Fishery ecosystem plan for Pacific pelagic fisheries of the Western Pacific region. Western Pacific Regional Fishery Management Council, Honolulu, Hawaii
- Zeller D, Palomares MD, Tavakolie A et al (2016) Still catching attention: Sea Around Us reconstructed global catch data, their spatial expression and public accessibility. *Mar Policy* 70:145–152
- Zeller D, Cashion T, Palomares M, Pauly D (2017) Global marine fisheries discards: a synthesis of reconstructed data. *Fish Fish* 19(1):30–39

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.