A Review of the BC Option A Trawl Fishery Catch Monitoring Program

Howard McElderry* and Rick Stanley

Joint Authors, Retired and Operating as Private Consultants
howard.mcelderry@gmail.com
rick.stanley602@gmail.com
*Corresponding Author

2026

Canadian Industry Report of
Fisheries and Aquatic Sciences 299

Fisheries and Oceans Canada
Pêches et Océans Canada

Canada

Canadian Industry Report of Fisheries and Aquatic Sciences

Industry reports contain the results of research and development useful to industry for either immediate or future application. They are directed primarily toward individuals in the primary and secondary sectors of the fishing and marine industries. No restriction is placed on subject matter, and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Industry reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Industry reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-91 in this series were issued as Project Reports of the Industrial Development Branch, Technical Reports of the Industrial Development Branch, and Technical Reports of the Fisherman's Service Branch. Numbers 92-110 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Industry Reports. The current series name was changed with report number 111.

Rapport canadien à l'industrie sur les sciences halieutiques et aquatiques

Les rapports à l'industrie contiennent les résultats des activités de recherche et de développement qui peuvent être utiles à l'industrie pour des applications immédiates ou futures. Ils sont surtout destinés aux membres des secteurs primaire et secondaire de l'industrie des pêches et de la mer. Il n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports à l'industrie peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports à l'industrie sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques*.

Les rapports à l'industrie sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement d'origine dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 91 de cette série ont été publiés à titre de Rapports sur les travaux de la Direction du développement industriel, de Rapports techniques de la Direction du développement industriel, et de Rapports techniques de la Direction des services aux pêcheurs. Les numéros 92 à 110 sont parus à titre de Rapports à l'industrie du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 111.

Canadian Industry Report of
Fisheries and Aquatic Sciences 299

2026

A Review of the BC Option A Trawl Fishery Catch Monitoring Program

By

Howard McElderry* and Rick Stanley

Joint Authors, Retired and Operating as Private Consultants
howard.mcelderry@gmail.com
rick.stanley602@gmail.com
*Corresponding Author

© His Majesty the King in Right of Canada, as represented by the Minister of the Department of Fisheries and Oceans, 2026

This work is licensed under the Open Government Licence

Cat. No. Fs 97-14/299E-PDF ISBN 978-1-100-00981-0 ISSN 1488-5409

Correct citation for this publication:

McElderry, H. and Stanley, R. 2026. A Review of the BC Option A Trawl Fishery Catch Monitoring Program. Can. Ind. Rep. Fish. Aquat. Sci. 299: xxiii + 106 p.

Table of Contents

Table of Contents iii
List of Tables v
List of Figures vi
Glossary vii
Acronyms vii
Definitions viii
Abstract x
Résumé xi
Executive Summary xiii
Introduction 1
Project Purpose 1
Background 2
Fishery History 2
Management History 4
The Option A Trawl Monitoring Program 6
Overall TMP Design 7
TMP Administration 10
Hail System 11
Fisher Catch Reporting 11
Dockside Monitoring Program 13
Electronic Monitoring 13
The “Audit” 17
Data Architecture 25
Materials and Methods 27
Overall Approach 27
Background Materials 27
Stakeholder Input 28
Trawl Catch Data 28
Audit Scorecards and Trip Summary Reports 29
Statistical Analysis 29

Results....33
TMP Operational Performance....33
TMP Data Adequacy....37
Alignment of Released Estimates of Major Commercial Species....37
Release Estimation by Tow....44
Alignment of Retained Catch Estimates....44
Trends in Catch Alignment (FY2020–FY2025)....48
Estimation of Non-Commercial Releases and Species Richness....48
Catch Scoring Analysis....53
How the Number of Tests Affects Scores....55
Paired and Null Comparison Tests....57
Comparison Test Scoring....60
Discussion....66
TMP Design....66
TMP Data Adequacy....66
Released Catch....67
Retained Catch....67
Species Composition....68
TMP Performance....68
Operational Performance....68
Functional Performance....69
TMP Audit Methodology....74
Incentive Structures....79
Recommendations....81
Acknowledgements....86
References....87
Appendices....89
Appendix 1 - Fishery History....89
Appendix 2 - TMP Chronology of Major Changes to the TMP Since Implementation. 90
Appendix 3 – Example Audit Scorecard....95
Appendix 4 – Example Trip Summary Report....101

List of Tables

Table 1. Species Group Designations for FL and EM (or DMP) Catch Comparisons. 20
Table 2. Score Definitions by Test Area (see text for explanation). 21
Table 3. Example of Catch Scoring for a Trawl Fishing Trip. 23
Table 4. Schedule of Audit Outcomes, Outcome Definitions and Consequences. 25
Table 5. Summary of Audit Outcomes by Time Period. 33
Table 6. Summary of Category 4 Failures by Time Period. 34
Table 7. Number of Fishing Trips by Vessel, Summed by Audit Outcome (FY2025 data). 36
Table 8. Summary of Trips With Additional Review by Time Period. 37
Table 9. Regression Statistics for FL Releases vs. EM Releases by Species or Category. N represents individual comparison tests. 38
Table 10. Difference Between Total Releases (lbs) by EM and FL Relative to Total EM by Fishery Year. 43
Table 11. Estimated EM-FL Released Catch Delta Compared Against Total Catch by Species for FY2025. 43
Table 12. Regression Statistics for FL Retained to EM Retained. Halibut and spiny dogfish omitted because of too few observations. Salmon omitted because retained “weights” standardized to be 1lb in RVs. 45
Table 13. Relative Difference for Total EM Ret and Total FL Ret (lbs.) by Species by Fishery Year. Omitted halibut, and dogfish) from yearly averages owing to influence of low values (Ret EM estimates of <5,000 lbs (FY 2025/2026) and changing recording practices. 47
Table 14. List of Nominal Species Categories Recorded by EM Review for FY2025 (partial). 49
Table 15. List of the Number of Species and Species Categories Recorded Observers or EM by Fishery Year. 50
Table 16. List of Species Categories Reported in either but not Both ASO and EM Periods (data from DFO FOS database). 52
Table 17. EM-FL Delta for Released Catch by Weight Interval and Audit Outcome (FL used when EM-FL+). 54
Table 18. Distribution of EM-FL Deltas by Species Group and Audit Outcome (FL Used if EM=0). 55
Table 19. Distribution of EM-FL Deltas by Species Group and Audit Outcome (FL Used if EM=0). 56
Table 20. Average Number of Comparison Tests by Test Group and Trip Characteristic. 57
Table 21. Paired and Null Comparison Tests by Test Area and Audit Outcome. 58
Table 22. Species Comparison Test Symmetry in Relation to # Tests and Released Quantity..59
Table 23. Species Comparison Test Symmetry by Species Group. 60
Table 24. Summary of Test Score Patterns and Test Method. 61
Table 25. Score Distributions by Audit Category Using Matched Tests (All Tests Shown in Brackets, when Different). 63
Table 26. Null Comparison Tests with a Score of 10 by Test Area and Audit Outcome. 65

List of Figures

Figure 1. Spatial Footprint of Bottom Trawling in BC. Black Shows 2020-2025 Fishing Locations, Dark Grey Coral Sponge Closed Fishing Locations, Light Grey Historical Extent of the Fishery. 6
Figure 2. General Schematic of the TMP QA/QC System. 8
Figure 3. EM System Layout on a Typical Wetfish Trawl Vessel. 14
Figure 4. FishView InterpretTM Display Showing Temporal and Spatial Plots, Along with Synchronized Imagery. 16
Figure 5. Conceptual Diagram of the TMP Audit Process. 10 and 25% Sampling Rate for Species Level Catch for Wetfish and RT Vessels, respectively. Note that 100% for Camera Imagery Reflects Total Catch Estimates Made for All Fishing Events. 17
Figure 6. Simplified Example of Test Area Scoring Methods. 22
Figure 7. TMP Data System Architecture. 26
Figure 8. FL and EM Release Estimates All Species Combined (A), then Lingcod (B), Longnose skate (C), Big Skate (D), Halibut (E), and Spiny Dogfish (F). Sample Size, Regression Coefficient of Determinations, Intercept, and Slope are Shown within Each Plot. 39
Figure 9. FL and EM Release Estimates for Spiny dogfish (A) and Sablefish (B). Plots C and D Show the Delta (% Difference) in Relation to the EM Estimate, with Catch Values and < 200 lbs and ≥ 200 lbs Plotted Separately 41
Figure 10. Total FL Catch Versus Total EM Catch for Sampled Tows from Audit Scorecard Sample (FY2025 partial). 45
Figure 11. Retained Catch Comparisons, Showing All Species Combined (i.e., Trip Level DMP vs FL, 2A), All Retained Comparison Pairs (i.e., Set Level FL vs EM, 2B), Then for Longnose Skate (2C) and Big Skate (2D). 46
Figure 12. Cumulative Frequency of Catch Species for 2019-2020 (blue) Compared with 2021-2025 (red), Showing the Incidence of Occurrence by Species Reporting Events (upper graph) and by Species Weight (lower graph). 51
Figure 13. Example of T1-DMP Test Results Showing Null Comparison Tests with Score of 10 (red arrows). 64

Glossary

Acronyms

AMR:	Archipelago Marine Research Ltd.
ASO:	At-Sea Observer
C&P:	DFO Conservation and Protection Program
CGRCS:	Canadian Groundfish Research and Conservation Society
CSA:	Canadian Sablefish Association
DFO:	Department of Fisheries and Oceans (or Fisheries and Oceans Canada)
DMP:	Dockside Monitoring Program
EM:	Electronic Monitoring or EM Record for a Tow
ENGO:	Environmental Non-Governmental Organization
FOS:	DFO's Fishery Operations System, the platform housing fishery data
FL:	Fishing Log
GFIMP:	Groundfish Integrated Monitoring Program
GHLCMP:	Groundfish Hook-and-Line Catch Monitoring Program
GIS:	Geographic Information System
GMU:	Groundfish Management Unit (of DFO)
GPS:	Global Positioning System
Hails:	Fishers reporting the intention and timing of leaving (Hail-out) or returning (Hail-in) from fishing
IUU:	Illegal, Unreported and Unregulated fishing
ITQ:	Individual Transferable Quota
LSR:	Least Squares Regression
LOA:	Length Overall in feet for reference to vessel size
PHMA:	Pacific Halibut Management Association
REL:	Released catch
RET:	Retained catch
RT:	Receiving Tank Vessels, which spill fish into a receiving tank for sorting and processing below deck. Limited catch sorting occurs on the fishing deck.
Scuppers:	Gaps in the vessel's rails (deck level and above the water line) for deck drainage and releasing catch overboard.
SOC:	Species of Concern. Under Canada's Species at Risk Act (SARA), a species is listed as a Species of Concern (SOC) if it may become a threatened or an endangered species because of a combination of biological characteristics and identified threats.
TAC:	Total allowable catch in a quota management system
TMP:	BC Option A Trawl Monitoring Program
Wetfish:	Vessels that sort catch on deck, stow retained catch in refrigerated holds, and releases unwanted catch directly overboard. Usually smaller than RT vessels.

Definitions

Accountable/Responsible: Accountability is documenting all catch (all species) in a fishing logbook; responsibility reflects having ITQ to account for the mortality of all legal/marketable sized groundfish that are managed to quota.

Census/Sample: Census refers to a collection of 100% of the records of a data source, as opposed to Sample which refers to a partial subset of records.

Comparison Tests: FL, EM, and DMP catch estimates are compared as part of the Audit Scorecard process. There are five test areas, abbreviated as T1 to T5, reflecting the different species comparisons made with landed, retained and released catch. Each individual species (or species group) test is called a comparison pair, which is the FL estimate compared with EM (or DMP for landed catch). Comparison pairs are either two non-zero estimates, denoted as EM+FL+, or a non-zero estimate and a null value, denoted as FL+EM- or FL-EM+. These are also referred to as paired and null comparisons, respectively.

Data Gap: EM systems write a line of data at a set cadence of once per 10 seconds. Data gaps are any occurrence where time spacing is greater than 10 seconds.

Delta: Delta is used as a synonym of arithmetic difference between two numbers.

FY2025: Dates preceded with 'FY' denote the fiscal year for the trawl fishery (i.e., 21 February 2025 to 20 February 2026). References with FY for other years follows the same format.

Individual Transferable Quotas (ITQs): ITQs are a catch share fisheries management tool wherein regulators set a total allowable catch (TAC) for a species and allocate %-based, tradable, exclusive rights to individuals.

Precision: Precision is a measure of the degree to which repeated estimations are close in value, in other words, it is the "scatter" around an expected value. It is expressed in this document as R^2 the coefficient of determination. It ranges from 0.0-1.0. A perfect match in closeness is reflected with a value of $R^2=1.00$; a complete mismatch is $R^2=0$.

Reference Value: The Reference Value, also called a dependent variable, is the reference value when comparing EM and FL estimates. Usually, the reference value is the EM (or DMP) estimate, but FL is used in null comparison tests where EM (or DMP) is zero.

Scorecard: TMP produced record of alignment between FL, EM, and DMP for each trip. It is used to evaluate FL performance.

Slope/bias: Slope and bias characterize how one variable varies with another. In this document it usually quantifies how an FL estimate varies with the EM estimate. A slope of 1.0 indicates a 1:1 relationship. A slope of <1.0 indicates that FL underestimates the EM estimate, while a slope of >1.0 indicates that FL overestimates the EM estimate. Slope quantifies the degree of numerical bias. Bias is quantitative and does not imply deliberate modification of a result.

Species References: Throughout the report we use common names for the different species, with occasional references to the Latin names. In some cases, the same species is referred to by the official common or by the local vernacular name. For example, Arrowtooth flounder or turbot, Pacific halibut or halibut, and thornyheads or idiot rockfishes.

The larger Tables 14 and 16 of “nominal” species names, were taken directly from catch data (observer, EM reviewer, and FL) was recorded.

Abstract

McElderry, H. and Stanley, R. 2026. A Review of the BC Option A Trawl Fishery Catch Monitoring Program. Can. Ind. Rep. Fish. Aquat. Sci. 299: xxiii + 106 p.

This paper provides an evaluation of the current catch monitoring program in BC's Option A trawl fishery which adopted an Electronic Monitoring (EM) at-sea element in place of at-sea observers, during the pandemic in 2020. Although catches have been monitored in this fishery since the 1950s, we refer to the current version herein collectively as the Trawl Monitoring Program (TMP). Differing from pre 2020 monitoring with observers, the TMP is a fully integrated QA/QC system that creates a fishery catch data base using verified landings data (DMP), and skipper-reported at-sea catch (FL records), which are checked using EM. Following completion of a fishing trip, FL records are audited using the EM record, confirming time, location and total catch by set from 100% of the record, and randomly sampling fishing events (10% or 25% for wetfish and receiving tank fleets, respectively) to evaluate FL records for retained and released catch by species or species categories.

The program uses a comprehensive Audit Scorecard to produce a data quality evaluation across five test areas: FL to DMP, FL to EM for retained catch, FL to EM for released catch other than halibut and sablefish, FL to EM for released halibut, FL to EM for released sablefish. The Audit Scorecard outputs poorly distinguish catch complexity from FL data quality but has served the TMP well with important feedback to skippers that has driven improvements to FL data quality since inception. Although the EM element entrains its own sources of uncertainty, evaluation of the FL and EM record across more recent years of the program indicate that the FL record aligns sufficiently well with EM to provide reasonable confidence in the TMP's ability to generate adequate release estimates at broader scales, such as coastwide or area-specific totals especially for commonly caught species, including most quota species. Confidence in released catch estimates declines as the number of observations per category decreases, a limitation of the TMP that could be mitigated by higher sampling rates.

Future changes may include moving to 100% sampling for rarely encountered species. Several areas of improvement to the Scorecard methodology are identified, as well as other areas that would serve to strengthen the TMP. Nevertheless, the TMP should no longer be considered a pilot program, a change that would strengthen commitment and funding to bring about further improvements.

Résumé

McElderry, H. and Stanley, R. 2026. A Review of the BC Option A Trawl Fishery Catch Monitoring Program. Can. Ind. Rep. Fish. Aquat. Sci. 299: xxiii + 106 p.

Cet article présente une évaluation du programme actuel de suivi des prises dans la pêche au chalut de l'option A en Colombie-Britannique, qui a intégré un volet de surveillance électronique (SE) en mer en remplacement des observateurs à bord pendant la pandémie, en 2020. Bien que les prises de cette pêcherie soient surveillées depuis les années 1950, nous désignons collectivement la version actuelle du programme sous le nom de Programme de suivi de la pêche au chalut (PSPC).

Contrairement au système de surveillance effectué par des observateurs avant 2020, le PSPC est un système intégré d'assurance et de contrôle de la qualité (AQ/CQ) qui produit une base de données sur les prises de la pêcherie à partir des données vérifiées sur les débarquements (DMP) et des prises en mer déclarées par les capitaines (dossiers FL), lesquelles sont vérifiées au moyen de la SE. Après chaque sortie de pêche, les dossiers FL sont vérifiés à l'aide des données de la SE, ce qui permet de confirmer, pour 100 % des données enregistrées, l'heure, le lieu et la prise totale pour chaque trait. Un échantillon aléatoire des activités de pêche (10 % pour les flottilles de poisson humide et 25 % pour les flottilles utilisant des réservoirs de réception) est également analysé afin d'évaluer les dossiers FL concernant les prises conservées et rejetées, par espèce ou catégorie d'espèces.

Le programme utilise une fiche d'évaluation complète, ou **tableau de pointage de l'audit**, afin de produire une évaluation de la qualité des données dans cinq domaines : FL par rapport aux DMP; FL par rapport à la SE pour les prises conservées; FL par rapport à la SE pour les prises rejetées autres que le flétan et la morue charbonnière; FL par rapport à la SE pour le flétan rejeté; et FL par rapport à la SE pour la morue charbonnière rejetée.

Le tableau de pointage de l'audit distingue imparfaitement la complexité des prises de la qualité des données FL. Il s'est néanmoins révélé utile au PSPC en fournissant aux capitaines une rétroaction importante, laquelle a contribué à améliorer la qualité des données FL depuis la mise en œuvre du programme.

Bien que le volet de surveillance électronique comporte lui-même certaines sources d'incertitude, l'évaluation des dossiers FL et des données de SE au cours des années les plus récentes du programme indique que les dossiers FL concordent suffisamment bien avec les données de SE pour donner une confiance raisonnable dans la capacité du PSPC à produire des estimations adéquates des prises rejetées à des échelles plus larges, notamment pour l'ensemble de la côte ou pour des zones particulières, en particulier pour les espèces fréquemment capturées, y compris la plupart des espèces visées par des quotas.

La confiance accordée aux estimations des prises rejetées diminue à mesure que le nombre d'observations par catégorie diminue. Il s'agit d'une limite du PSPC qui pourrait être atténuée par l'augmentation des taux d'échantillonnage. Parmi les changements qui pourraient être apportés à l'avenir figure le recours à un échantillonnage de 100 % pour les espèces rarement rencontrées.

Plusieurs possibilités d'amélioration de la méthodologie du tableau de pointage sont relevées, ainsi que d'autres mesures qui permettraient de renforcer le PSPC. Néanmoins, le PSPC ne devrait plus être considéré comme un programme pilote. Cette évolution permettrait de renforcer l'engagement envers le programme et son financement, et ainsi de favoriser la mise en œuvre de nouvelles améliorations.

This French translation was made with the assistance of ChatGPT.

Executive Summary

Introduction

The BC Option A Trawl Monitoring Program (TMP) was established in 2020 following the suspension of the at-sea observer program during the COVID-19 pandemic. The TMP uses Electronic Monitoring (EM), Fisher Logbooks (FL), the Dockside Monitoring Program (DMP), and a Hail System as an integrated catch-monitoring and quality-assurance system. The program provides 100% EM coverage of fishing activity, while a sample of EM-recorded fishing events is reviewed to assess the quality of the fisher-reported catch record. The purpose of this review was to evaluate how well the TMP is functioning, whether its data are adequate for fishery management, and where changes could improve the program.

The review concludes that the TMP is operating effectively and that its catch data are adequate for the general management needs of the fishery. The program provides high levels of fleet coverage, produces data within required timelines, and has developed a catch-monitoring system that integrates EM with fisher reporting and verified dockside landings. The analysis also indicates that the alignment between FL and EM catch estimates has improved over time. For commonly caught species and species categories with sufficient observations, the data provide reasonable confidence in estimates at broader scales, including coastwide and area-specific totals.

At the same time, the review identifies limitations. Confidence decreases as the number of observations declines, particularly for species that occur infrequently or in small quantities. Released catch is the main area of uncertainty because, unlike retained catch, it cannot be verified through DMP. Species such as skates and other infrequently observed catch categories are more difficult to estimate, and coral and sponge observations are too limited to support quantitative estimates. The 10% minimum EM review rate for Wetfish vessels provides limited information for assessing individual-trip data quality and deterrence for these vessels which are usually smaller and employ open deck sorting.

The review also finds that the Audit Scorecard, although useful, does not always provide a reliable representation of fisher logbook data quality. The current scoring system converts differences between FL and EM estimates into ordinal scores and then averages those scores. This can cause individual large discrepancies to be masked by other high scores. Null comparison tests, which account for about 40% of comparison tests, can also increase scores without representing a direct comparison between two non-zero estimates. These issues do not appear to substantially alter the results at the whole-fishery level, but they can affect individual trip outcomes.

The report recommends that the TMP no longer be treated as a pilot program. It should be formally adopted as the ongoing monitoring system for the Option A trawl fishery, with continued

investment in its data systems, electronic logbook, audit methodology, sampling design, reporting, and incentive structure.

Program Context

The Option A trawl fishery is a multi-species fishery operating across a large area of British Columbia's offshore waters. It harvests more than 70 stocks and more than 50 commercially valuable species, as well as more than 100 non-commercial fish and invertebrate taxa. From 2021 through 2024, the fleet averaged approximately 40 vessels, 827 trips, 8,091 trawling events, and 1,680 hours of fishing effort per year. The fishery has declined in scale in recent years, largely because of the decline in offshore hake availability.

The monitoring requirements are complex because trawl gear captures multiple species and sizes at the same time. Retained catch by species is verified through the DMP when fish are landed. Released catch must be estimated at sea, making it the main source of uncertainty in the monitoring system. The TMP therefore relies on the FL as the complete at-sea catch record and uses sampled EM review to assess whether those records are sufficiently accurate.

The distinction between the roles of EM and FL is important. EM provides continuous video, GPS, and sensor records, but only a sample of the recorded catch is reviewed in detail. EM therefore does not normally provide a census of catch by species. Instead, the reviewed EM data serve as a benchmark for evaluating the FL, notwithstanding the fact that EM catch estimates are also estimated with uncertainty The Audit compares the two records and provides feedback intended to improve future FL reporting.

Operational Performance

The TMP has been operating continuously since April 2020 and is considered fully deployed and operational. The review found that the fleet has maintained high levels of EM data completeness and compliance with FL reporting requirements. The proportion of trips with complete EM data has generally remained in the mid-90% range, reaching 98% in FY2025. This represents a high level of monitoring coverage for a year-round fishery operating over a large geographic area with continuous video and sensor recording.

Audit processing is also meeting the program's timelines. In FY2025, 98% of trip reviews were completed within 10 calendar days, compared with a requirement of 10 business days. Approximately 40% were completed within five days. The review considers these timelines important because they include data collection, processing, analysis, and reporting after a fishing trip.

The most recent Category 4 failure rate was 2.4%. Category 4 outcomes indicate that monitoring specifications were not met. Earlier Category 4 failures included data gaps, differences in the definition of fishing events between EM and FL, and camera failures. The

fishing-event definition issue was largely resolved when EM and FL definitions were aligned. In FY2025, all 11 Category 4 failures were caused by data gaps.

The review also notes that the fleet's realized EM review rates were above the stated minimums. For Wetfish vessels, realized rates averaged 12.9% and 13.4% in the two most recent years, compared with a target of 10%. Refrigerated Tank (RT) vessel rates averaged 27.8% and 25.9%, compared with a target of 25%.

Overall, the operational results indicate that the TMP has moved beyond the implementation problems expected of a new program. The basic systems for EM, FL, DMP, data processing, auditing, and reporting are functioning together.

Catch-data Adequacy

The central finding of the review is that TMP catch data are adequate for general management purposes. The level of confidence depends on the species, the amount of catch, and the number of observations available.

The FY2025 analysis included 715 reviewed tows from 523 completed trips involving 21 vessels. Approximately 8,000 species-level weight estimates were available and were consolidated into about 5,000 comparison-test pairs, roughly half involving released catch. For released catch across all species, the FY2025 analysis found a slope of 1.18 between FL and EM estimates, with an R^2 of 0.79. This indicates that FL estimates tended, on average, to be about 18% higher than EM estimates, with some imprecision. Species-level results varied considerably. Halibut and spiny dogfish showed strong alignment, while longnose skate and especially big skate showed poor alignment. Hake also showed a tendency for FL estimates to exceed EM estimates.

These results should not be interpreted as evidence that the FL consistently overestimates catch for every species. The species-specific slopes vary, and some species, including arrowtooth flounder, sablefish, and pollock, showed FL estimates below EM estimates. The key point is that the degree of alignment depends on the species and the amount of information available.

The review concludes that broader-scale estimates are more reliable than individual tow-level estimates. For commonly caught species, sufficient observations provide reasonable confidence in coastwide or area-specific release totals. The significance of individual discrepancies is also reduced where most catch is retained and subsequently verified through DMP. However, tow-level discrepancies can matter for allocating catch to management areas and therefore for Individual Transferable Quota management. Most importantly, there is no evidence of large-scale unreported releases.

Confidence declines for species with fewer observations. Skates provide an example of this problem. Coral and sponge observations are too limited to support quantitative estimates; for

these categories, the data may provide information about presence or absence and location of catch but little basis for estimating total catch. For rarely observed species, the report notes that meaningful improvement could require 100% EM review.

Retained Catch

The review found relatively strong alignment between FL and EM retained-catch estimates for most species. Examples include arrowtooth flounder, lingcod, hake, sablefish, other flatfish, rockfish, and walleye pollock. Pacific cod estimation showed a tendency for FL to overestimate retained catch by approximately 30%, while big skate and longnose skate showed poorer alignment and lower precision.

However, retained catch is not the main concern for the monitoring system because DMP provides verified landed weights. The main purpose of comparing FL and EM retained catch is therefore to ensure that the FL can be used to allocate verified landed catch to the correct fishing areas and to support quota accounting.

The report notes that some skippers reportedly reallocate catch by tow after receiving DMP results. This is not permitted under the TMP. The report suggests that the prevalence of this practice could be investigated using time stamps from the FL data-entry system. It also notes that there appears to be less incentive to alter reported capture areas than in earlier periods. The review therefore recommends reducing the emphasis placed on retained-catch reporting and increasing the relative emphasis on released catch, where the FL is the main source of information.

Species Composition

The review examined whether the move from observers to EM had reduced the program's ability to identify species. The analysis found an approximate 18% decline in the number of recorded species categories associated with the transition from ASOP to EM. The number of categories has since remained relatively stable.

The report cautions that this result cannot be attributed solely to the monitoring method. The fishery itself changed between the two periods. Total catch and fishing effort declined, the contribution of RT vessels changed, fishing areas and gear characteristics changed, and thornyhead-directed fishing ceased in 2019. Observer identification capability also varied among individuals. A more detailed analysis would be required to isolate and estimate the effects of these factors.

The review concludes that both ASOP and TMP can produce comparable levels of species granularity, but the level of identification in FL records varies among vessels and trips. The current requirement to identify catch to the lowest possible taxonomic level is not sufficiently standardized. The report recommends simplifying the general species list and establishing more targeted data collection for species where greater identification detail is needed.

One option would be a sentinel approach in which selected vessels provide more detailed catch information. This would focus additional effort on species that occur in low volumes, are encountered infrequently, or are difficult to identify.

Fisher Catch Reporting

The transition to the TMP placed greater responsibility on skippers and crews. Under the previous observer system, observers produced the official at-sea catch record. Under the TMP, the FL became the official at-sea record and is subject to EM audit.

The report identifies significant practical challenges. Wetfish vessel crews generally consist of three or four people, and catch retrieval and sorting involve several activities occurring at the same time. Crew must manage fishing gear, sort retained and released catch, handle species subject to special requirements, and estimate quantities. RT vessels face a different challenge because much of the catch moves through processing systems over longer periods and must be accounted for throughout that process.

The report concludes that the catch-estimation process needs clearer procedures, tools, and thresholds. The current electronic FL application is based on an earlier observer logbook and was not designed specifically for the TMP. Stakeholders described it as difficult to use. The review recommends redesigning the application so that it automatically records information where possible, time-stamps entries, tracks corrections, and reduces unnecessary reporting requirements.

The report recommends placing less emphasis on retained catch, which is verified by DMP, and more emphasis on released catch. Clearer standards should define what species and quantities must be reported and to what level of detail.

Audit Scorecard Performance

The Audit is the main quality-control component of the TMP. It assesses whether FL data meet the standards required for use as the official at-sea catch record.

The review found that approximately 89% of trips received Pass or Category 1 outcomes, indicating that the majority of trips met the program’s acceptance criteria. The proportion of Category 2 and 3 outcomes has declined over time, indicating an improvement in FL data quality.

However, the review cautions against treating audit categories as a direct measure of skipper performance. More complex trips generate more comparison tests and more opportunities for discrepancies. The difference between Pass and Category 1 or 2 outcomes therefore reflects both catch complexity and data quality. A trip with more species and more released catch is more difficult to estimate than a trip with simple catch composition.

The report concludes that the Scorecard is useful as a communication and quality-control tool but should not be treated as a single numerical measure of overall data quality.

Problems with the Scoring Methodology

The main methodological concern is the conversion of catch differences into rank scores. The Scorecard uses scores of 0, 3, 5, 7, 8, 9, and 10. These are ordinal categories, not numerical measurements. Averaging them can therefore obscure important differences in the underlying data.

For example, two trips can have the same average score while having very different individual comparison results. A large discrepancy can be offset by several high scores, allowing an overall trip score to remain within an acceptable category. The review recommends that individual low scores not be allowed to disappear through averaging.

A second concern involves null comparison tests. These occur when one record reports catch and the other reports zero. They account for approximately 40% of comparison tests. Because the difference in a null test is automatically equal to the non-zero estimate, the scoring method can produce a high score simply because the amount of catch is small. The analysis found that 72% of null tests involved weights of 25 pounds or less, making a score of 10 common.

The report found that null tests can elevate average scores and obscure actual paired comparisons. At the fishery-wide level, the effect does not appear large enough to change the overall conclusions, but at the individual-trip level it can affect audit outcomes. The report therefore recommends changing the treatment of null tests rather than eliminating them. One proposed approach is to cap null-test scores at 8.

The review also identifies differences between weight-based (for smaller catches) and percentage-based scoring (for larger catches). Weight-based tests tend to produce high scores, while percentage-based tests tend to produce lower scores. The report recommends reviewing the tolerance ranges so that they reflect realistic catch-estimation performance.

For lingcod and sablefish, the current scoring system separately evaluates legal and sublegal releases. Because many observations are small, this can generate high scores without providing a useful assessment of total species catch. The report recommends assessing catch quantity separately from compliance with measurement-board requirements.

Sampling Rates and Additional Review

The TMP uses a target 10% EM review rate for events of Wetfish vessels and 25% for RT vessels. The report finds that the Wetfish rate produces too few reviewed events on many trips. A minimum of one reviewed tow can result in only a small number of comparison tests and provides limited information about trip-level data quality.

The report recommends maintaining the 10% Wetfish rate but increasing the minimum to two reviewed fishing events per trip. It also recommends using the same sampling approach for RT vessels rather than maintaining separate rates without a clear basis.

Additional review should become a standard, automatic response when initial results do not meet defined data-quality thresholds. It should be used to determine whether a discrepancy is attributable to the FL, EM interpretation, or both. Additional review should also contribute to the final Audit Scorecard rather than simply being appended to the original result.
The report further recommends targeted additional review for important species such as sablefish, halibut, and lingcod, particularly when large or infrequent catch events occur. Where EM review demonstrates that FL data are unacceptable for a species or an entire trip, the criteria for replacing FL data with EM estimates should be clearly defined.

Incentives and compliance

The TMP relies in part on incentives to encourage accurate FL reporting. Current consequences include additional review, delays, discussions with DFO, and additional processing costs. Interviews suggested that these measures, along with the desire to produce a good Scorecard result, influence compliance.

The review concludes, however, that the TMP does not have a sufficiently clear incentive structure or escalation pathway. The program should establish defined responses to data-quality problems. These could include education and follow-up monitoring, financial consequences, DFO compliance measures, and, in more serious cases, loss of FL reporting privileges or suspension of program participation.

The purpose should not simply be punishment. The incentive system should provide clear expectations, identify corrective actions, and create consistent consequences when requirements are not met.

Principal Findings

The review's principal findings can be summarized as follows:

1. **The TMP is operational -** The program has been operating for more than five years, and its core systems have stabilized. EM, FL, DMP, and Hail functions operate as an integrated monitoring system.

2. **Monitoring coverage is high - EM** records fishing activity continuously, and the Hail System provides an additional control that connects fishing activity to monitoring requirements.

3. **FL data have improved -** Audit results show a decline in Category 2 outcomes and an increase in Category 1 outcomes. This indicates improvement in the quality of fisher-reported catch records since the TMP began.

4. **The current system provides adequate data for general management -** The FL and EM records align sufficiently for reasonable confidence in broader-scale estimates of released catch, particularly for frequently caught species. There is no evidence of significant unreported releases.

5. **Released catch remains the main uncertainty -** Retained catch is supported by DMP verification. Released catch relies primarily on FL estimates, which are assessed using sampled EM observations.

6. **Confidence in the catch estimates varies by species –** Frequently caught species with many observations provide more reliable comparisons. Rare species and categories with few observations are more problematic.

7. **The current catch review design is not sufficient for all species -** Higher sampling rates, including 100% EM review in some cases, may be required for rare or management-important species, or other specific issues.

8. **The Audit Scorecard is useful but has methodological limitations -** The use of ordinal scores and averaging can obscure individual discrepancies. Null comparisons can also increase scores without representing matched catch estimates.

9. **Category 4 failures have declined -** The most recent Category 4 rate was 2.4%, and the remaining failures were caused by data gaps. Earlier problems involving FL and EM fishing-event definitions have largely been resolved.

10. **The FL application needs redesign -** The current application was adapted from the observer-era system and should be redesigned around the current TMP.

11. **The program should shift emphasis toward released catch -** The relative importance of retained-catch reporting should be reduced because DMP provides independent verification.

12. **The TMP should become permanent -** The program has moved beyond the conditions that justified its designation as a pilot and should be formally established as the ongoing monitoring system.

Priority Recommendations

The recommendations can be organized into a set of priorities for the next stage of the TMP:

1. **Establish the TMP as the permanent monitoring program -** The program should no longer be referred to or managed as a pilot. The review concludes that continued uncertainty about the program's future limits investment and development. Formal adoption would provide the basis for longer-term planning of EM systems, FL technology, data infrastructure, training, and quality assurance.

2. **Rebalance monitoring toward released catch -** The program should recognize that DMP already provides verified landed catch. Greater monitoring effort should therefore be directed toward released catch. The report recommends reducing the combined weighting of retained-catch test areas to no more than 20% and placing the remaining weight on released catch.

3. **Revise the Audit Scorecard -** The Scorecard should continue to provide a simple trip-level assessment but should be revised so that important individual discrepancies cannot be hidden by averaging. Low individual scores should trigger additional review. Null comparisons should be treated differently from matched comparisons. Weight- and percentage-based tolerance ranges should be evaluated using fleet-wide data.

4. **Increase minimum review sampling per trip -** The Wetfish minimum should increase from one reviewed event to two reviewed events per trip while retaining the 10% target. This would provide more information about individual-trip data quality without requiring a major increase in the overall sampling rate.

5. **Use targeted sampling -** Species with limited observations should receive additional monitoring when justified. Targeted (stratified) review should be considered for rare species, Species of Concern, and other categories where the existing sample does not provide sufficient information.

6. **Establish automatic additional-review thresholds -** The current case-by-case approach should be replaced or supplemented with defined triggers. Examples include low individual comparison scores, repeated discrepancies for the same species, or other predefined data-quality conditions.

7. **Redesign the electronic FL -** The application should be redesigned to reduce manual effort, capture information automatically where possible, record changes, and clarify reporting requirements. Minimum thresholds should be established for species reporting so that all vessels use comparable reporting practices.

8. **Improve audit reporting -** The Audit Scorecard and Trip Summary Report should be combined into a single feedback product that identifies the result, the underlying

discrepancies, and the corrective actions. The report should be accessible through a web-based system rather than relying only on static PDF reports.

9. **Improve program databases-** Data systems should be redesigned to support routine monitoring and analysis. The program should be able to produce fleet-wide and vessel-level performance summaries without reconstructing datasets for each analysis.

10. **Develop a Scorecard simulation model -** A simulation model should be developed to test proposed scoring changes before they are implemented. The model should be capable of varying species groupings, tolerance ranges, sampling rates, test-area weights, and other settings for previous FYs. This may require improving data systems to support multi-year analysis.

11. **Strengthen incentives -** The program should establish defined consequences for data-quality deficiencies and recognize areas of good performance. The consequences should be transparent, proportional, consistently applied and documented.

12. **Conduct regular program reviews -** A comprehensive review should occur approximately every five years. Between reviews, a dedicated quality-assurance program should monitor operational and data-quality indicators. The report emphasizes that no monitoring program can be assumed to remain effective without continued review and adjustment.

Overall Assessment

The review finds that the BC Option A Trawl Monitoring Program is functioning as a monitoring system and is producing data that are adequate for the general management of the fishery. The most important evidence supporting this conclusion is the alignment between FL and EM records for commonly caught species, the improvement in Audit outcomes over time, the high level of EM coverage, the integration of FL data with verified DMP landings, and the decline in operational failures.

The program is not without limitations. The principal uncertainty concerns released catch, particularly for species that occur infrequently or in small quantities. The EM review sample is not sufficient to provide the same level of confidence for all species. The Audit Scorecard provides useful feedback, but its current scoring methodology can obscure individual discrepancies. The FL application places unnecessary demands on crews and does not establish consistent thresholds for species reporting. The program also requires clearer rules for additional review, incentives, and escalation.

These issues do not require replacement of the TMP. They require modification of the system so that its resources are directed toward the information needs that matter most.
The most significant change should be a shift in emphasis from retained catch to released catch. DMP already provides verified retained-catch information, whereas released catch

depends on at-sea estimation. The Audit, sampling design, Scorecard weighting, and reporting requirements should reflect that difference.

The second major change should be to treat the Scorecard as a quality-control tool rather than as a single measure of skipper performance. The score should provide useful feedback, but individual discrepancies should remain visible and should trigger review where appropriate.
The third major change should be the development of a continuing quality-assurance process. The TMP has now generated enough operational history to support routine monitoring of its own performance. The program should not depend on occasional external reviews to identify problems.

Finally, the TMP should be formally established as the permanent monitoring program for the Option A trawl fishery. The pandemic conditions that led to its creation have ended, but the monitoring requirements of the fishery remain. The question is therefore no longer whether the program should be maintained as a temporary substitute for observers. The relevant question is how the existing system can continue to develop and be improved so that it continues to provide dependable, timely, accessible, and cost-effective information.

The review provides evidence that the TMP can meet that role. Its core design is functioning, the data are adequate for general management, and the program continues to improve since implementation. The next stage should focus on refining the system rather than redesigning it from the beginning.

The recommended approach is to retain the integrated EM, FL, DMP, and Hail framework; increase emphasis on released catch; improve the FL application; revise the Scorecard; strengthen targeted sampling; establish clearer follow-up procedures; improve data systems; and create a defined incentive structure. These changes would allow the TMP to build on the results achieved during its first five years while addressing the areas where the review found limitations.

The report's overall conclusion is therefore that the TMP should be retained, formally adopted, but subject to continued development. It has demonstrated that an EM-based monitoring system can provide the basic information required for management of the Option A trawl fishery when EM is integrated with fisher reporting, verified landings, quality assurance, and DFO oversight. Its future effectiveness will depend on continued investment, routine evaluation, and the willingness to revise procedures as the fishery and its information needs change.

This Executive Summary was compiled with the assistance of ChatGPT.

Introduction

Project Purpose

British Columbia's offshore groundfish trawl fishery (Option A Trawl) has for about seven decades maintained an evolving monitoring program starting initially with landing sales slips and voluntary fishing logs (FL) in the 1950s. This transitioned in 1993 to a validated dockside monitoring program (DMP) for landed weights using land-based observers for monitoring offloads. Fishing logs became mandatory in 1985. An at-sea observer program (ASOP) was first implemented in the early 1990s, then expanded to 100% coverage in 1996.

With the onset of the COVID pandemic in 2020, the ASOP was suspended and replaced shortly thereafter with an Electronic Monitoring-based solution. Prior to this change, the industry had been using EM for a few years in specific applications (e.g., directed fishing for thornyheads and hake) and already had EM in place for offshore hake fishing and as a supplement to Observers on Receiving Tank (RT) vessels. With the suspension of the ASOP, the trawl industry, DFO and the EM service provider worked quickly to implement a revised at-sea monitoring program, hereafter referred to as the Trawl Monitoring Program (TMP).

The TMP followed a similar design as was in place with the BC Groundfish Hook and Line Fisheries Monitoring Program (GHLCMP) (McElderry et al., 2003; 2007; Stanley et al., 2015; PHMA and CSA, 2023; DFO, 2024) wherein EM data are used to test ("Audit") the veracity of catch and effort data reported by skippers in their FL. Owing to the fishing method, quantity, and diversity of species in the catch, the TMP program is necessarily more complex than that for the GHLCMP and has been continually evolving and developing from April of 2020 to the present, as summarized in Appendix 2.

While the building blocks were in place for a 100% coverage program for the entire Option A trawl fishery, the pandemic emergency caused the EM element of the monitoring to be developed and tuned while it was being deployed. The revised TMP has now been in place for over five years, and the operational systems have stabilized, yet the program is still called Pilot EM Program. With the pandemic long over, continuing the path of an EM-dependent solution would seem inappropriate if this were considered an inferior option with the earlier ASOP. Alternatively, continuing to use an EM-based solution denies the program the credit it deserves if superior to ASOP but still regarded as a pilot program. The pilot status also inhibits investments in the program so long as its future is in limbo. With demersal mobile gear fisheries generating public scrutiny, it became even more important to ensure that the trawl fishery was effectively monitored. Recognizing the need to resolve this conundrum, Fisheries and Oceans Canada ("DFO") and the Canadian Groundfish Research and Conservation Society (CGRCS) pooled funds for a study to provide an evaluation of the current TMP. This report is a product of this evaluation.

It might seem a logical approach to conduct this study as a comparison between the current TMP with the pre-pandemic ASOP. However, this comparison is complicated by the fact that both EM and Observers provide estimates whose precision is not well-defined. Further, as will be shown, the trawl fishery has changed markedly such that comparing the ASOP years with the current TMP period would be confounded by each having very different underlying fishery characteristics. The strengths and weaknesses of EM and Observers have been extensively reviewed (e.g., McElderry, 2008, Michelin et al., 2018) and is not repeated here. Instead, the authors chose to evaluate the TMP as it currently functions, the approach being more a business analysis of operational systems and outputs of the TMP. This approach looks beyond the EM at-sea monitoring tool and considers the quality and level of its integration with other elements of the monitoring system (i.e., Hails, DMP, and FL). In our view, an effective program design includes: the monitoring tools and how they are deployed and integrated; data compilation, analyses, and reporting processes guiding the information outputs and uses; procedures, operational rules, and incentive systems setting the governance structure for program participants; and the overall organizational framework, blended in a way to best satisfy program requirements while achieving cost effectiveness. This perspective recognizes that all monitoring systems are imperfect and must be subject to an ongoing process of evaluation and evolution to continually adapt to a changing fishery and updating objectives while still providing cost effectiveness. An ASOP versus EM approach overlooks the design subtleties and would fail to examine key questions: How well does the TMP meet the monitoring requirements of the fishery now, and how well it is positioned to adapt to change?

DFO has finalized a national fishery monitoring policy to ensure that it has dependable, timely, and accessible information on fisheries to manage them sustainably and to ensure a unified approach for setting the type and degree of monitoring employed across fisheries nationwide (DFO 2019). This report presents an important contribution to assist the implementation of the policy: a unified framework to evaluate the quality and dependability (ability of an estimation process to achieve its intended objectives) of DFO's fishery monitoring programs (Allard and Benoit, 2019). In addition, and more specifically, a key objective of this project is to identify program strengths, weaknesses and gaps and to recommend changes and future research that could improve this program.

Background

Fishery History

The BC offshore trawl fishery dates to the 1940s and has evolved considerably to its present state. The trawl fishery is the largest harvester group by weight landed within the groundfish commercial sector, which also currently includes fixed gear fisheries (licence sectors) for lingcod, rockfish, sablefish, dogfish, and halibut. The trawl fishery is active over a large area of offshore waters throughout the year (Figure1) and ranks as BC's largest fishery in terms of catch quantities. There are several sources describing the fishery (DFO, 2025; Gislason, 2024). The fishery characteristics presented here are to provide context to the review.

Annual trawl landings have fluctuated between 60 and 130k tonnes (MT) over the years mostly owing to annual variability in hake landings (see Appendix 1). The fleet size has ranged from more than 50 vessels to fewer than 30 in recent years. During the 2021 to 2024 fishing years, the fleet averaged 40 vessels, 827 trips, 8,091 trawling events, and 1,680 hours of fishing effort. Fishing activity reported for 2025 represents about 75% of the season, and year-end projections suggest similar or less activity than for 2024. Nevertheless, the more recent levels of activity still translate to a significant monitoring program with over 600 landings and 5,000 fishing events projected for the FY2025 year.

The fishery harvests more than 70 stocks and 50 commercially valuable species including rockfishes, pollock, hake, flatfishes, cod, sablefish, and skates. During the last five years, landed catches from fishing trips targeting rockfishes, flatfishes, pollock, and “inside” hake (from waters between Vancouver Island and the mainland) have remained relatively stable (landed catch of 19k MT, 8k MT, 7k MT, 8k MT, respectively) while landings from offshore hake trips have declined precipitously from 61k MT to 3k MT. The decline in the availability of offshore hake is the principal cause for the decline in the fishery in recent years. Included in the catch are a large number (>100) of non-commercial species of fish and invertebrate taxa. Estimates of released catch (commercial and non-commercial species combined) range between 3% and 8% of total catch over the past decade (see Appendix 1).

The implementation of licence limitation in 1973 fixed the number of trawl licensed vessels to 142 although the number of active vessels is far fewer in each year. Each year, trawl licensed vessels choose between Option A, which is primarily a large vessel fleet accessing offshore areas of BC, and the much smaller Option B, which is primarily a small vessel, fresh fish fishery within the Strait of Georgia region. Options A and B are governed by separate conditions of licence.

Among the Option A fleet, the majority are 65-90 ft (LOA) and referred to as the ‘Wetfish’ fleet. These vessels carry a crew of 3–4, make 3–5-day trips, and deliver their fresh catch to shoreside processing/offloading locations (mostly Prince Rupert, Port Simpson, French Creek, Port Hardy, and Ucluelet). Wetfish vessels sort all catch on the fishing deck, either splitting or ramp-loading catch aboard at which point the retained portion is directed into refrigerated seawater holding tanks and the remainder released overboard through scuppers. A smaller component of the fleet, RT vessels, are larger in size (100-185 ft LOA) and often carry out some level of onboard catch processing, either producing totes of fresh iced fish or frozen blocks of headed-gutted-tailed product. These vessels are capable of longer trips, carry larger crews, and can operate continuously 24 hours a day. RT vessels typically ramp-load catch onto the trawl deck where limited sorting takes place (e.g., release of halibut and sablefish) with the remaining catch directed into a receiving tank where it is then processed on the factory deck. The configuration of fish processing operations may vary among RT vessels, but all use conveyor belt systems to move fish from the receiving tank to sorting and processing areas.

Management History

Management of the groundfish resources in BC is complex, mainly because of the multi-species, multi-stock makeup, with differing stock harvest capacities. Certain species are caught by multiple fleet sectors, each with differing access priorities and privileges. Furthermore, while the release of unwanted, unmarketable catch is common to many gear types, it tends to be more prevalent with trawl gear. Obtaining reliable estimates of released catch is important for understanding overall fishery impacts. With reliable DMP monitoring of unloading/landing operations in place, estimation of at-sea released catch is the most problematic and consequently the highest profile issue for the trawl fishery.

Prior to 1997, the trawl fishery operated under a variety of control measures, such as licence limitations (1973), time- and area-based landing, gear restrictions, sales slips, and others. Owing to misreporting in sales slips, beginning in 1993, the DMP became mandatory for all trawl landings where the sorting and weighing of landed catch was verified by independent observers paid for by industry, reducing the need to use processor-based sales slips for landings monitoring. However, because the sales slips were thought be useful for price data and their collection was congruent with other BC fisheries catch, which still relied entirely on sales slips these data continue to be recorded by DFO in different archival databases.

In 1997, the trawl fishery transitioned to an individual vessel quota (IVQ) management system for vessels opting (Option A) to primarily fish offshore areas, while the Option B inshore fleet continued under trip limit management. Along with IVQ management, the Option A fleet began carrying and funding the costs of Observers on all groundfish trips, except directed hake fishing, which carried EM. The ASOP was responsible for providing independent estimates of retained and released at-sea catch. These data were used as the official catch record, replacing the FL data record, even though skippers were still required to fill out and submit FL records.

In 2006, the remaining groundfish sectors, such as trap and longline, except for Option B trawl, transitioned to IVQ management. A consequence of this is that management became integrated for many species that overlap different groundfish licence and gear groups. Under this integrated management, fixed gear fisheries and Option B trawl were also subject to 100% at-sea monitoring but using EM technology rather than Observers. In 2010, EM technology was applied to the inshore hake fishery, employing a similar audit-style monitoring approach as used in hook-and-line gear fisheries.

At the turn of the century, the number of larger trawl vessels began to increase with the introduction of RT vessels. The first RT arrived in 2003, with a second and a third arriving in 2007 and 2009, respectively. EM-based monitoring on the RT fleet began in 2017. While these vessels carried observers under the ASOP, EM was incorporated to simultaneously monitor different areas of the vessel, as well as augment coverage for 24-hour operations.

As the fishery transitioned to IVQ management, the benefits accruing to quota holders carried individualized responsibilities. Quota holders became responsible for accounting for all legal/marketable sized groundfish managed under species and area TACs. Quota holders

became both **accountable** - having to provide verifiable means of catch documentation for all species - and **responsible** - having to carry or lease quotas for all harvested IVQ species.

Along with increased accountability for IVQ management, the trawl industry, DFO, and the ENGO community developed habitat protection measures, including a bottom trawl spatial footprint where bottom trawling was permitted (Wallace et al., 2016; Gale et al., 2022).

More verified fishing location information available from ASOP, and later from EM, enabled the definition and tracking of much more granular areas designated for bottom trawl fishing. With the availability of near real time (i.e., at trip end) tow location information, the spatial extent of the fishery was effectively monitored and controlled. Figure 1 depicts the spatial footprint of bottom trawl fishing in relation to historical fishing areas (light grey) and the location of coral/sponge closed areas. Note that published spatial data are resolved to a grid of 4km, while data used for spatial management, including sponge/coral reef closures, are resolved (i.e., verified by EM GPS data) to 2 nautical miles.

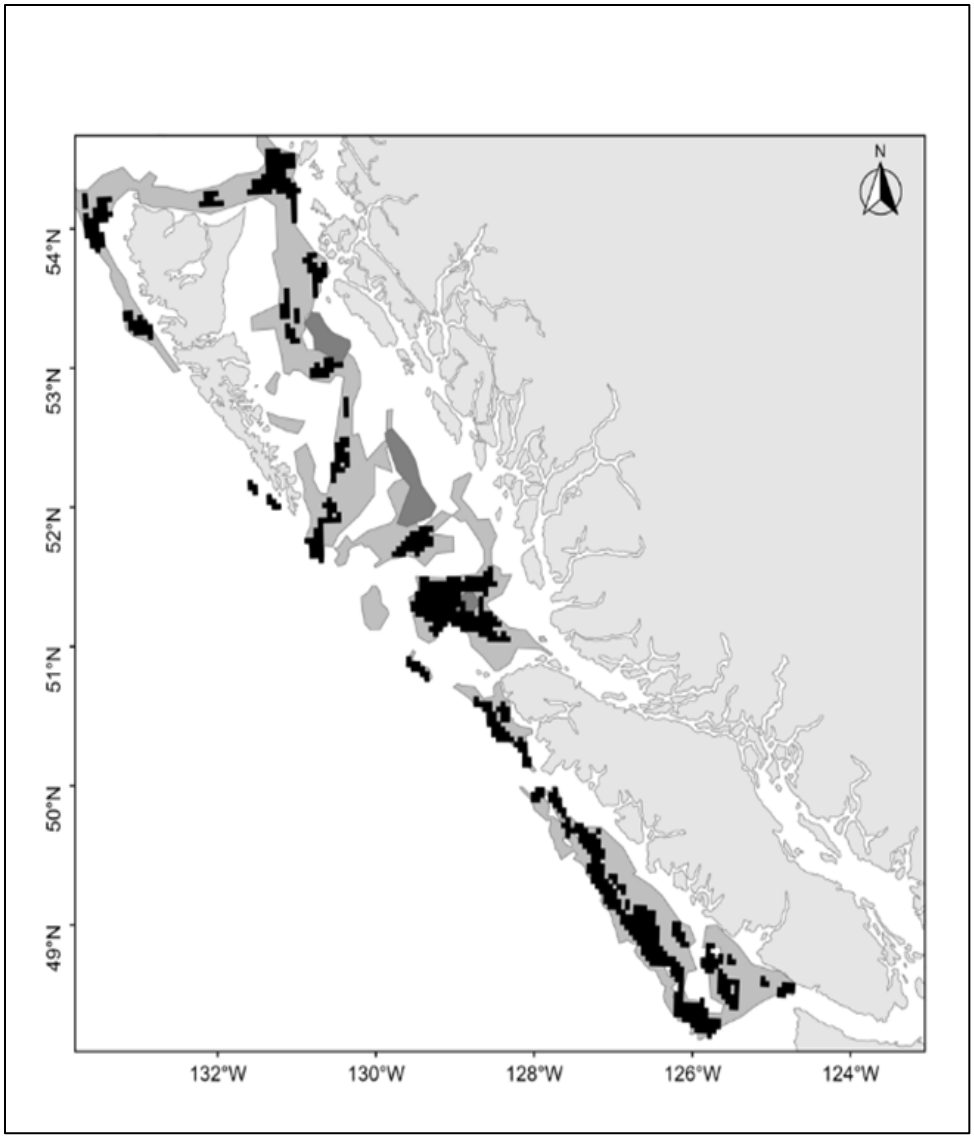


Figure 1. Spatial Footprint of Bottom Trawling in BC. Black Shows 2020-2025 Fishing Locations, Dark Grey Coral Sponge Closed Fishing Locations, Light Grey Historical Extent of the Fishery.

The Option A Trawl Monitoring Program

As mentioned earlier, the pandemic caused a dramatic change to Option A trawl fishery monitoring, with conflicting priorities of keeping the fishery open and meeting the monitoring objectives without ASO. In rapid fashion, the fishery incorporated EM in a similar way to the GHLCMP. The main difference in this incorporation is that the hook and line fishery usually catches fewer species, and fish in a slower, singulated (piece by piece) fashion. In contrast, trawl fisheries catch large volumes of mixed groundfish species and sizes at once, making the job of catch estimation more difficult, as well as less precise and less accurate. The TMP is described in the following sections. DFO provides a comprehensive document that describes the EM and catch analysis portions of the TMP (DFO, 2024). The intent herein is not to

duplicate that document but to describe the program in a way that is less technical and incorporates other elements of the TMP besides the EM component.

The TMP must meet several information needs and operational objectives (DFO, 2024 and 2025), including:

- Estimate retained catch by weight, species (management category), and catch location;
- Estimate released catch by weight, species (management category), and catch location;
- Estimate catch of Species of Concern by weight, species (management category), and catch location;
- Estimate Sponge/Coral by catch weight;
- Record fishing effort and location;
- Implement size measurement protocols for catches of sablefish, lingcod, and halibut;
- Implement biological sampling protocols, as requested;
- Provide evidence of adequate data quality (bias and precision);
- Ensure the TMP transparency;
- Ensure the TMP is cost-effective; and
- Ensure the TMP provides timely outputs.

Overall TMP Design

Central to the operation of the TMP, shown schematically in Figure 2, is a complex and comprehensive Quality Assurance (QA) program with its constituent Quality Control (QC) measures. The QA incorporates the design elements of all aspects of TMP operations, from equipment deployment and performance, through data capture and testing, to downstream communication, analysis, and reporting, including documents such as this one, to ensure that the information necessary for meeting the objectives is included and maintained in the program. The QCs within the overall QA package, in turn, are operational procedures designed to manage and safeguard specific elements of the system, covering a wide range of components such as hardware installation and placement, EM reviewer training, and other operational details.

The narrative approach adopted in this document is not to present a single, consolidated summary of all QCs. Instead, following Figure 2, individual QCs are introduced and elaborated as they are encountered as part of the broader discussion of TMP structure and performance. The principal exception to this approach is the element referred to herein as the “Audit,” which is viewed as the primary QC measure (Figure 5) with various subcomponents associated with evaluation of effort and catch estimation and therefore warrants focused treatment.

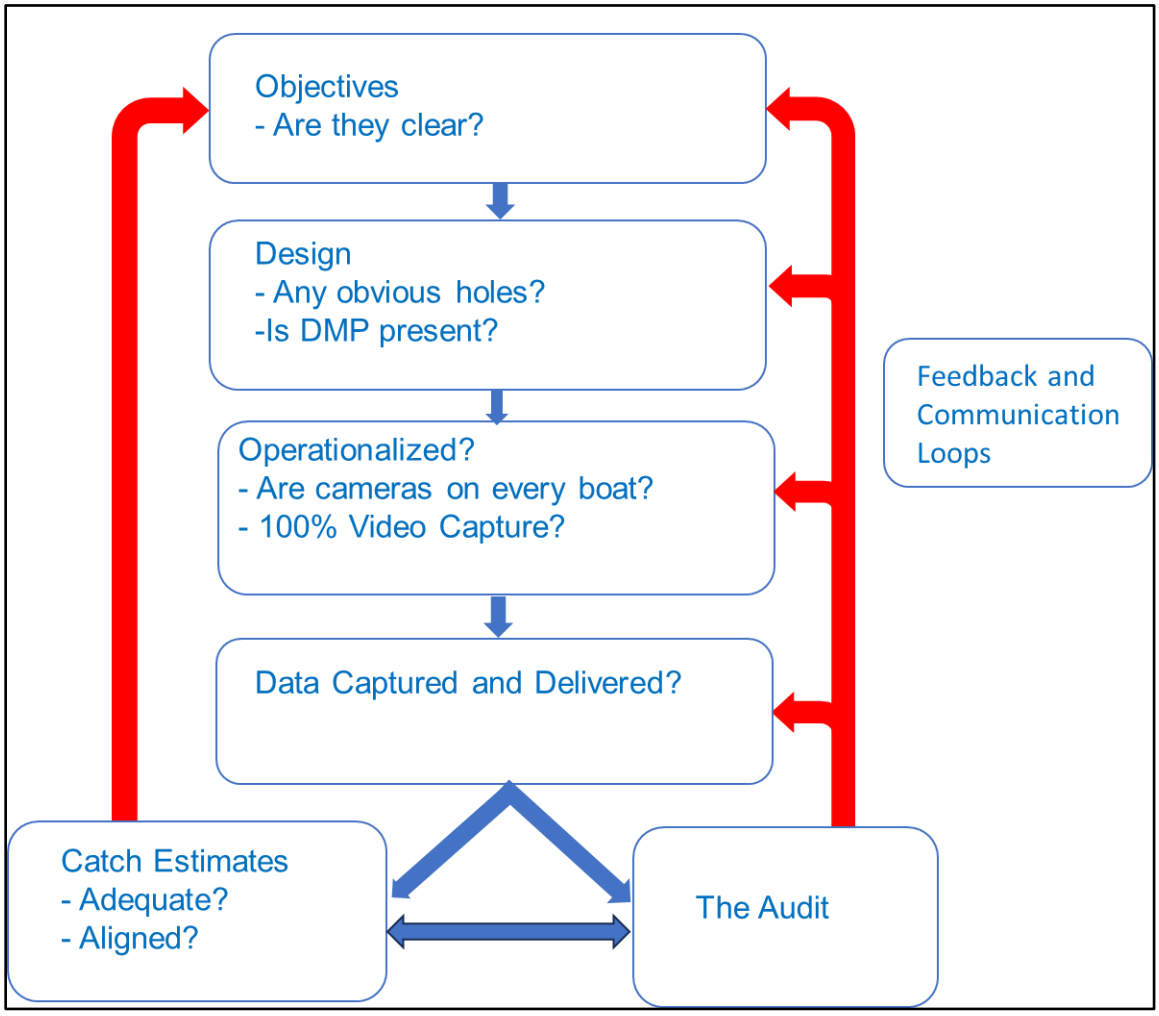


Figure 2. General Schematic of the TMP QA/QC System.

Verified landings provided through the DMP are a critical part of the TMP. The DMP has been in operation for over 30 years, and there is widespread trust in its information outputs. Hence, the review saw little reason to examine its operational efficacy in this study. These output data are used during the investigations of EM and FL data below. Details about the DMP and how program data are integrated into the TMP will be discussed in the next section.

Although the TMP achieves 100% video coverage of all retained and released catches (and effort tracking), practical considerations related to cost and efficiency led to the decision that only a random sample of the catch footage is reviewed - at least 10% of the tows for Wetfish vessels and at least 25% for RT vessels. This design choice fundamentally shifts the role of EM. Rather than providing a complete census of retained and released catch at sea, which would allow it to replace the TMP dependence on FL, EM becomes a sampled data source. Under a full-review EM model, total catch estimation could be simplified to combining DMP data with EM-derived estimates of releases, yielding a near-complete picture of fishery catch, with fishing logbooks (FL) playing a lesser but still critical role of providing the basis for allocating the catch to stock area for IVQ purposes and area-specific stock assessment.

In that hypothetical scenario, while EM estimates would still be subject to measurement error, EM would be considered superior to FL, and comparisons between the two would provide little additional value. However, once EM review is limited to a sample, the FL regains its role as the sole complete at-sea, or census, data source for the fishery, in particular for retained and released catch by area. Correspondingly, the sampled EM data take on a different but still critical function: they become the benchmark against which the credibility, or veracity, of FL data are considered. Consistent with the philosophy of the audit method, the events randomly selected for EM review should be considered more for random checking of the FL data rather than as a statistical sample used for catch estimation, although this has been done for the GHLCMP (Gravelle, 2004).

This shift necessitates the development and incorporation of a sophisticated QC within the TMP to examine and evaluate the alignment between sampled EM data and the corresponding FL data. In essence, this QC helps assess whether there is sufficient confidence that FL data can serve as an adequate complement to DMP data, standing in for the comprehensive coverage that a full EM review might otherwise provide. This shift is the same fundamental approach used in the GHLCMP (Stanley et al., 2015).

In this document, the specific QC and its Scorecard are designed to perform this comparison—or veracity check—of the FL component using sampled EM data, which is referred to as the "Audit." Because the Audit is central to establishing the credibility of FL data and, by extension, to the integrity of the TMP, it receives substantial attention in this review. The most complex and consequential QC element of the Audit is the Scorecard (see Appendix 3), which plays a critical role in evaluating performance, determining confidence in the overall system, and is the basis for providing feedback to skippers on the quality of the recordings in their logbooks.

The EM review is not designed to substitute EM estimates where individual FL estimates of individual catch by species are deemed inadequate, bearing in mind that only a small proportion of the tows are reviewed, although this occurs to a limited extent. Rather, the Audit is more forward-looking, providing an initial evaluation of the FL logs and corrective feedback, as well as the basis for implementing an incentive structure that encourages FL data quality improvement over time. Note that the sample approach only applies to verification of catches in contrast with fishing effort in time and space which is 100% reviewed. For example, checking for fishing in closed areas is 100% reviewed.

Thus, the EM data record is used to randomly spot check various catch estimates of the fishing trip as recorded in the FL. Consequently, skippers know there is a 100% chance that one or more of their fishing events from each trip will be selected for review, but do not know which. It is to be noted that there is no statistical basis for the sample rate; instead, the rate is determined by aligning the relative cost and benefits of the program to detection risk and failure consequences, and the 10% review rate continues to be considered appropriate in the GHLCMP.

The overall philosophy of this approach is to incentivize good FL recording practices in a way that promotes acceptable data records, or conversely, to disincentivize poor recording practices so that those who provide substandard FL records face greater data processing costs, time delays, and administrative burdens. In addition to the veracity check provided, the outcomes of the check and matching exercise (i.e., FL-EM catch comparisons) provide a basis for sharing detailed feedback to skippers on where and how their FL data capture can be improved, and to a lesser extent, feedback to EM reviewers. The comparison of FL and EM estimates therefore not only serves as a check, but the outcome of the check provides the basis for guiding and incentivizing better recording, whether estimating catches, confirming time and location of fishing, and ensuring adherence to special catch handling protocols (e.g., halibut and sablefish measurement) and identifying areas of improvement. To avoid failure consequences, they are therefore incentivized to keep accurate records of all fishing events. This logic requires that failure consequences are sufficient to deter the fishers from taking a chance with poor recording practices.

TMP Administration

The groundfish trawl sector has tasked the Canadian Groundfish Research and Conservation Society (CGRCS), on its behalf, to negotiate and secure a contract(s) for the provision of the TMP. For the current fishing season (FY2025), the CGRCS has contracted with Vericatch and Archipelago Marine Research Ltd. (AMR) to deliver monitoring program components. The roles of these organizations are described below.

The TMP includes various oversight committees:

- The Trawl EM Technical Working Group – This committee is comprised of DFO management, DFO C&P (Conservation and Protection), CGRCS, AMR, and Vericatch (as required). The purpose of the committee has been to review program operations, troubleshoot problem areas, develop and modify audit rules and consequences, and assist in the development of new program components (e.g., biological sampling protocols).
- The Trawl EM Configuration Review Board and Trawl EM Configuration Review Committee – These two groups provide guidance on EM system installations on vessels. The Board was made up of industry representatives, DFO, C&P, and AMR, while the Committee included fishery experts (recommended by GTAC and approved by DFO) whose purpose was to tour vessels and provide recommendations on camera placements to best capture all catch handling activities. This Board existed in 2023/2024.
- The Trawl Audit Review Board – This Board is composed of DFO groundfish management whose purpose is to review audit reports and, in the case of failed audits, meet with vessel representatives to discuss the report, and to determine consequences before the vessel is permitted to resume fishing.
- Biosampling Technical Working Committee – This committee has representation from CGRCS, DFO, and AMR. The role is to oversee the collection and processing of biological samples collected by the vessel crew and delivered for processing on shore. While noted as an original biological objective of the TMP (see page 7) the

biological sampling element of the TMP has been omitted from this review, owing to its recent implementation.

The operational elements of the TMP are 100% funded by industry through user fees and an annual levy mechanism. The actual cost paid by industry and the basis for fees was not investigated; however, Gislason (2024) notes that monitoring costs in FY2021 were about 1.8% of vessel expenses, down from 5.4% over a decade earlier in FY2009, when at-sea monitoring was conducted with observers. The FY2021 figure of 1.8% likely does not include the capital cost of EM systems. It would be more reflective of the ongoing services associated with the provision of the combined program (i.e., EM, FL, DMP, and associated analyses and reports). DFO contributes staff time for their role in program oversight and the use of program information outputs.

Hail System

The Hail System serves multiple purposes of regulating access to fishing, notifying DFO and service providers of which vessels are fishing, and providing notification of offload intentions. Prior to a fishing trip commencement, all vessels must hail out to a call service to provide details of their intended trip and expected landing date and location. Upon landing, skipper hails provide notification of date, time and location of vessel offloading. The hail service is provided by Vericatch and is primarily delivered through an on-line platform to receive information and issue hail numbers. Note that in addition to its administrative role, the Hail System is critical to the QC of ensuring that all trawl fishing is monitored. It is very unlikely that fishing vessels would risk fishing without hailing out, and the vessel hails directly tie to the rest of the monitoring obligations (i.e., logbooks, dockside monitoring, and EM). In other words the hail system ensures that there is no illegal, unreported, and unregulated (IUU) fishing, since it is unlikely that a trip or even a trawl tow could be conducted secretly. This review found no evidence to the contrary.

Fisher Catch Reporting

Fisher Catch Estimation

A critical part of the skipper's reporting obligation is providing accurate estimates of retained and released catch. To some extent, skippers, crew, and EM reviewers all use similar strategies for catch estimation, employing a combination of volume estimates, species proportions, piece counts multiplied by weights, and deductions. The methods used are directly influenced by how fish are brought aboard and sorted. On the Wetfish vessels, all catch handling and sorting takes place on deck as the codend is emptied. Deck crew work through the pile of unsorted catch with the retained portion directed into fish holds, and the released portion separated out and directed overboard, except for halibut, sablefish and lingcod, which must be placed on a graduated board for camera-based measurements before releasing overboard. Measurements provide additional size data on released halibut and are the basis for distinguishing legal and undersize sablefish and lingcod. Under trawl regulations, all halibut and all undersized lingcod and sablefish must be released upon capture.

Based on conversations with skippers, catch estimates are a process of deduction, starting with a visual estimate of the total catch quantity, then estimating the portion and species composition of the retained portion. The released quantity is then determined by subtracting the retained quantity from the total catch. Species composition is estimated through both counts and visual estimation of the releases being directed overboard. Halibut, sablefish and large catch items (e.g., big skates, large sharks) are often estimated by applying a “standard” piece weight to counts of pieces. Vessels that split their catch aboard can make estimates from each split and combine the results to obtain a total catch estimate.

For RT vessels, the emptied codend is slowly spilled onto the trawl deck to separate halibut, sablefish, and lingcod for the measurement requirements, before the remaining catch is directed into a receiving tank accessible from the processing deck below. Below deck, unsorted catch travels on conveyors to an initial sorting area, where the released catch portion is removed before the remaining retained catch is directed onwards for processing. RT vessels are not permitted to mix catch from different tows in the receiving tank. Catch estimation for RT vessels differs from that for the Wetfish fleet. Total catch quantity is made by estimation of the codend volume as it is brought aboard, but species composition and retained/released portions are estimated at the factory level. The species composition and quantity of released catch are determined from counts made as catch is separated from the conveyor. The retained composition and quantity for the processed product are determined using a conversion factor to expand the product weight to the round weight. In the case of toted fish product, retained composition and quantity is estimated by tote counts, modified by ice allowance, and average product weight per tote. Similar to Wetfish boats, catch estimates of halibut, sablefish, lingcod, and large catch items are estimated separately.

Fisher Logbooks

Fisher logbooks are considered legal documents for skipper declarations of their fishing activities. Logbooks were 100% paper-based until around 2014, when an electronic log was introduced. Also, about this time a more comprehensive electronic logbook version was introduced for the ASOP. With the introduction of the current TMP, the expanded reporting requirements for skipper-generated catch data meant that skippers use the ASOP version of the electronic logbook. The term ‘Olog’ has been used for the application because of the observer origins, yet this application is the current Fisher Log, or FL.

The FL is recorded in a tablet-based application maintained and provided by Vericatch. Skippers are responsible for supplying an Android-based tablet to host the application. During the trip, skippers enter their catch data on the application, recording on a tow-by-tow basis, time and location of fishing, and quantities of retained and released catch by species or species group. There are some conveniences built into the application, such as drop-down lists for species and auto-population of current time and location (from built-in GPS). DFO licence conditions state that catch entries must be made by 8am the day following fishing activity and that reports must be filed within 48 hours after completion of the fishing trip. The application supports provision of a high level of catch detail, but skippers typically record just the most

common species occurrences as there is no minimum threshold specification. FL reports are filed by uploading the data to Vericatch's web-based platform. Once uploaded, the submitted data are copied to create an FL archive (legal copy) version, and an editable version, the 'Olog', which may be subsequently updated during the Audit process.

Dockside Monitoring Program

The DMP is a key part of the TMP, providing verified weights of landed catch by species or species group. Landed weights are validated by dockside Observers and are trusted estimates of the total retained catch to the level of species or species group as sorted at dockside, but only to the resolution of a complete trip (i.e., all tows combined). Note that trip catches are allocated, by proportion, to quota areas using FL location data.

The DMP has been operated by AMR since its inception. AMR is certified by Fisheries and Oceans Canada as a designated provider of DMP services. The DMP contract is administered by CGRCS, which represents trawl license holders' interests in ensuring that the DMP service is provided according to the needs of the fishery. For example, the program delivers data according to quality and timelines specified by DFO, while it offers services at geographic locations and call-out timelines specified by CGRCS.

Vessel offloading is not permitted unless the dockside Observer is present. At offloading, the dockside Observer ensures that the catch is sorted to the required species or species categories and that all weights are witnessed. When the offloading operations are completed, the dockside Observer confirms that vessel holds are empty and reports their tabulated catch data to AMR. The dockside Observer may also conduct other duties such as ensuring vessel FL reports are submitted, biosampling, and collecting hard drives from the vessel EM system.

Several quality assurance measures are carried out as part of the DMP service. All dockside monitors are certified by DFO after successfully meeting specified training standards. Dockside Observers often reside in the ports where the DMP services are provided, and their activities are regularly supervised by a senior dockside monitor (port supervisor). The data associated with each offload, including tote-by-tote weights, total offload tallies, and plant tallies are entered by the Observer into a web-based application where they are verified by AMR data managers before uploading into Vericatch's trawl data platform.

Electronic Monitoring

Several elements of the EM component enable the technology to be an effective tool for catch monitoring on fishing vessels. Much of the emphasis is typically directed towards the electronic equipment aboard the vessel (Figure 3), although the way it is installed, the vessel crew's obligations to support its effective use, and quality control measures within the program all contribute to a successful monitoring outcome. EM systems capture a vast quantity of data, and specialized analytical tools are needed to efficiently access and convert the video and sensor data into meaningful fisheries data. This section covers each element of the EM system and its

operational requirements; the details of the data analysis and audit are covered in the next section.

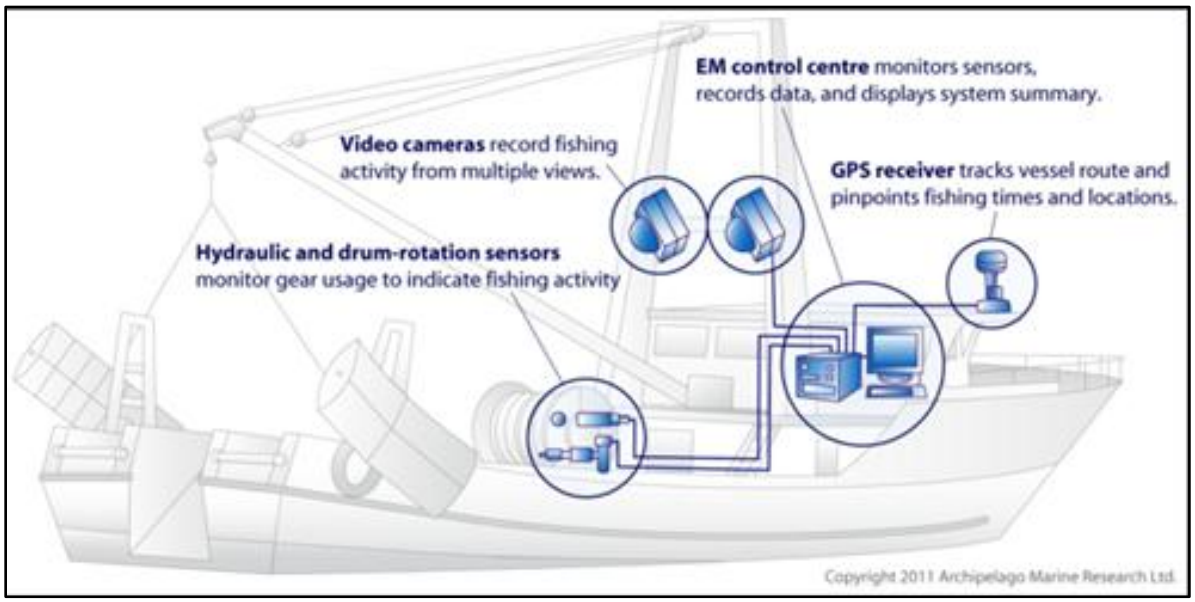


Figure 3. EM System Layout on a Typical Wetfish Trawl Vessel.

Vessel EM Requirements

The EM system operational requirements are the same for all Option A trawl vessels. EM systems must operate continuously from vessel departure until commencement of catch offloading. During this time, both sensors and cameras are operating, and the EM system is continuously recording data. Vessel masters are required to regularly monitor EM system performance. The display console provides continuous feedback on basic operation, data storage capacity, and camera views. The vessel master must regularly examine camera views presented on the display console to ensure cameras are operating (a view is present), the image quality is clear (camera domes are kept clean), and that cameras are aimed correctly for the intended field of view. During catch handling operations, there are also obligations to ensure that all activities occur within the field of view of cameras and that camera views are not obstructed in any way.

If a malfunction occurs, the vessel is required to contact AMR to attempt to resolve the issue. If the malfunction cannot be resolved, the vessel is required to contact DFO to determine whether they must cease fishing and return to port.

EM System Performance Requirements

The performance of the EM Systems is managed with organized response processes. Skipper feedback is the most important and timely method to correct system problems. AMR maintains a dedicated call line for reporting technical issues so that EM technician staff can respond quickly with remote support or provide scheduled service when the vessel returns to port. As well, EM System performance issues are identified during data analysis; technical staff,

skippers, and DFO are alerted if any issues are detected. Lastly, AMR maintains maintenance records for all deployed EM systems to track system age, repairs or upgrades performed, and other technical information.

The chain of custody of EM system data is essential to overall program quality. EM system data is encrypted using a secure technology that renders it unalterable without a software key to unlock the encryption. EM system data drives are usually collected from the vessel by AMR staff and forwarded to port or head-office-based locations. In some cases, skippers will swap drives while at sea, or remove drives to deliver to AMR, but in either case, data encryption ensures that the chain of custody is maintained.

The core element of the hardware on the vessel is the EM control centre. It supports multiple high-definition cameras and multiple sensors, including GPS, hydraulic pressure, and winch activity. Most vessels in the trawl fleet use model V5 EM systems, although a few have adopted the latest model, called FishView Vantage™, which incorporates 2-way satellite connectivity for real-time system checks and remote monitoring. All EM systems are designed to be tamper evident as opposed to being tamper-proof. In other words, the system continuously records operation and diagnostic information to enable detection of system tampering and failure. The lifespan of an EM system is 5-10 years, and replacement costs are about $CDN 10-15k for a complete system, depending on the components required. Since operational failure can interrupt fishing operations, it is incumbent upon fishers to maintain their EM systems and replace them when needed.

EM Data Processing

A typical EM data set from a fishing trip comprises hundreds of gigabytes or more, containing multiple high-resolution camera images and thousands of lines of sensor data recordings. It would be impossible to make sense of all this information without a specialized data visualization and analysis tool. AMR uses its product, FishView Interpret™, which provides a temporal display of the fishing trip aligned with a spatial display of the cruise track on a coastal chart (Figure 4). The coastal chart contains GIS shape files of fishery management areas and areas closed to trawl fishing. This part of the software enables the reviewer to view the entire fishing trip, noting fishing locations, camera activity, and other information. The accuracy of GPS position data is often less than 100m. Camera imagery and sensor data are time-referenced to enable navigation along the fishing trip timeline. The integration enables both quick spot-checking throughout the trip and/or detailed examination of a particular fishing event. In the latter case, the reviewer can activate relevant camera views, play through a catch retrieval event at a controllable speed, and simultaneously enter observations in a data entry table. In this way, simple observations can be reviewed many times real-time (fast forward) or by making multiple passes of the same event, examining different activities on camera with each pass. For example, a reviewer might focus on the composition of retained catch on one pass, quantities released on another, and measurements of halibut on a third pass.

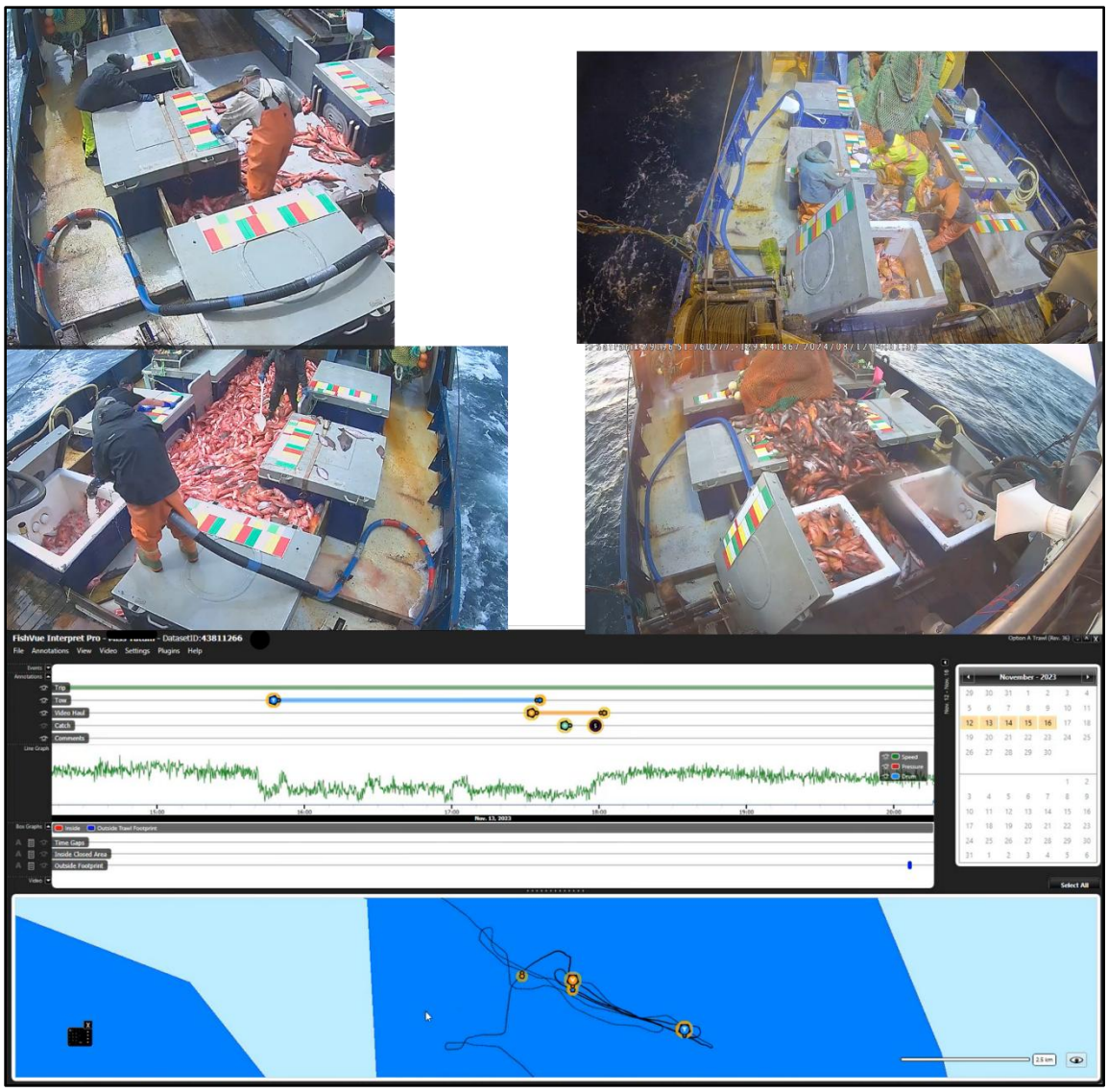


Figure 4. FishView InterpretTM Display Showing Temporal and Spatial Plots, Along with Synchronized Imagery.

The quality of EM data review is a critical component of the TMP. With respect to sensor data, there has seldom been an issue with data quality, as GPS positional accuracy is routinely monitored and other sensor data is considered accurate, when present. The quality of imagery and interpretations made from it are a potential key limitation of the program. For this reason, camera views and image quality are regularly monitored and corrected when necessary. In terms of interpretation, AMR follows several steps to ensure accurate determinations from the images. All viewers receive standardized training that emphasizes both species identification and the use and operation of FishView Interpret™ software. In the past, individuals with at-sea observer experience were often recruited for EM viewing work. With the decline in the number of active observers, AMR shifted to expanding standardized training to newly hired EM viewers, then assigning their work according to experience and performance, starting with less complex data sets (e.g., groundfish fixed gear trips, low diversity trawl trips), then moving to higher levels of catch estimation complexity. AMR also deploys EM viewers, when practical, to attend trawl

offload events and participate in trawl research charter trips for additional experience. EM viewer certification training includes testing with standardized reference imagery for assessment. EM viewer performance is also tested on an ongoing basis, with 10% of all reviewed trawl fishing events selected for re-review by all EM viewing staff. AMR reports that, under current (2025) activity levels, EM review work is carried out by 3-4 individuals with about 1 new hire per year.

The steps identified above enable continuous monitoring of EM viewer competency and serve to identify remedial training needs. However, it is important to note that there is a subjective element to image interpretation that will cause variability in resolution, particularly when image quality is suboptimal. It is also important to recognize that EM viewing capacity is not easily adaptable to fluctuations in the volume of work. This is true for annual fluctuations in activity as well as seasonal fluctuations, where analysis timelines are the same, regardless of the number of vessels landing at the same time.

The "Audit"

The core QA/QC element of the TMP is referred to as the "Audit", shown schematically in Figure 5. It is the overall process wherein all data streams are examined for completeness, tested for overall adequacy, and then feedback is issued through a 'Scorecard' (Appendix 3). Importantly, the EM review process occurs independently from, and without being informed by, information from the DMP or FL. In our interviews with stakeholders, it was evident that most had a general understanding of the process, but few understood the detailed testing and scoring methods. As an examination of the scoring methodology is central to this project, it seemed important to provide a simple explanation of the Audit process and how catch data are tested. The overall audit method is also described in detail by DFO (DFO, 2024).

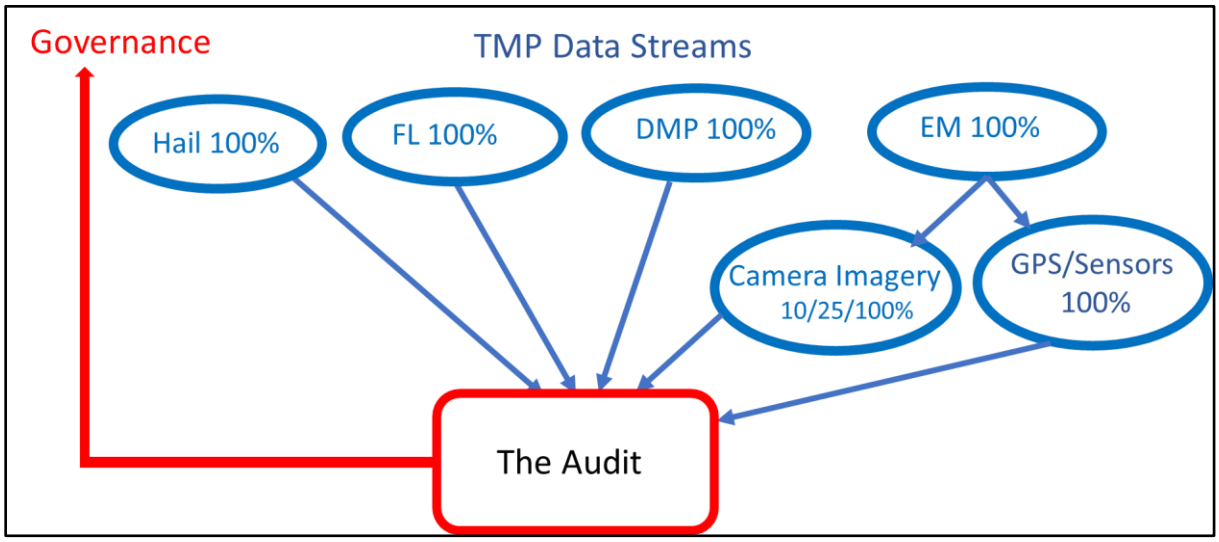


Figure 5. Conceptual Diagram of the TMP Audit Process. 10 and 25% Sampling Rate for Species Level Catch for Wetfish and RT Vessels, respectively. Note that 100% for Camera Imagery Reflects Total Catch Estimates Made for All Fishing Events.

Once a fishing trip and offload are completed, and the FL, EM (i.e., video, GPS and sensors), and DMP data have been provided, the Audit is performed to determine if the FL record meets data quality acceptability standards to be used as the official at-sea record of the trip. The Audit has two parts: data set conformity and catch record alignment between FL and EM (or DMP). These two elements come together to form an audit result (or outcome) to indicate whether the FL record meets acceptability standards.

In terms of data set conformity, a fundamental prerequisite is that FL and EM data records conform to one another in terms of timeline and locations as follows:

- Time gaps – the EM data record must be complete from the beginning to the end of the fishing trip with no breaks in the sensor or image data. Sensor data should be recorded at 10-second intervals, and image data should be recorded continuously from several cameras at rates of 5-10 frames per second. Small time gaps in the data set occur periodically (e.g., generator reset, EM system taken off-line for testing) but may not exceed 4 minutes in duration.
- Fishing event alignment – the FL and EM data record must both show the same number of fishing events, and all must align in terms of fishing area (groundfish management area), time, and location for the start position. Acceptability criteria for this test are an exact match for groundfish management area, within 60 minutes and within 2 nautical miles for time and location, respectively. AMR may automatically update vessel location information in cases where fishing events are aligned, but there is an obvious error in position reporting.

Once the data set is accepted in terms of basic level of conformity, described above, the audit proceeds to the next step: examining the imagery from all fishing events to estimate the total catch for each event. This information is compared with the FL estimate for alignment, and both EM and FL are totalled and compared against the landed catch. While this step is not used in data quality scores, it serves as an important check for basic alignment of EM and FL catch, as well as for anomalies in released catch. The total estimated catch from all fishing events should not exceed the landed catch by a large amount (5-8%, Appendix 1). In many cases, trips may fail to meet basic conformity requirements yet contain sufficient FL and EM data for further testing. In these cases, the available data are reviewed, but the trip will still be considered an automatic failure regardless of the test score results.

In assessing catch record alignment between EM and FL data sets, EM viewers follow prescribed protocols for randomly selecting fishing events from the pool of all fishing events on the trip. Wetfish boats are sampled at a target rate of 10%, wherein 10% of the fishing events are randomly selected then examined in detail. The RT fleet is sampled at a rate of 25%. In both cases, a minimum of one fishing event per trip is examined.

Each selected fishing event is examined in detail to identify and categorize catch composition. Catch sorting operations are viewed through different cameras to estimate the quantity and species composition of both retained and released catch. Releases of individual halibut, sablefish and lingcod must be placed against a graduated measurement board for size and

weight determination. Pre-determined measurements and area specifications (e.g., bin volumes) unique to different fishing vessels are used by EM viewers to enable weight estimates from volumes of fish observed on deck, within bins and nets, etc. EM viewers also monitor for factors that impair assessment (e.g., poor image quality, standing in front of cameras, etc.), as well as fishery violations, all of which are separate from the audit scoring but are incorporated into the audit report and thus the overall Audit.

Scored Test Areas

Catch comparisons are made between both the DMP and the FL records, as well as between the FL and the EM records for reviewed fishing events. These two comparison levels are incorporated into five separate test areas, each with prescribed species or species groups to be tested (see Table 1). The five test areas are as follows:

1. FL versus DMP (T1-DMP) - The FL retained catch totals for the trip are compared with landed catch totals as determined from DMP. There are 11 possible tests if all species categories occur in the landed catch. The scored comparison is not performed with landed catches of prohibited species (e.g., halibut, undersized sablefish and lingcod, and salmon), as well as other species not identified by category (e.g., invertebrates, etc.).
2. FL versus EM Retained Catch (T2-Ret) - The second test occurs at the fishing event level and compares FL and EM estimates of retained catch. Like T1, retained catch is examined and tested by species or species groups, but the groups differ. Skates and flatfishes are each grouped in a single test, with a possible total of eight tests if all retained catch categories are present. Like T1, scored testing is not performed with retained quantities of prohibited species, as well as other species not identified by a retained catch category (e.g., ratfish, invertebrates, etc.).
3. FL versus EM Released Catch (T3-Rel) - The third test is similar to T2, except the comparison is made for released catch. Released catch is tested using 13 possible species groups. There are two tests for lingcod that distinguish legal and sublegal size groups as verified by presentation on a graduated measurement board. Lingcod are considered legal-sized if they are not presented for measurement. Rockfishes are not included in the scored test as all should be retained, and similarly, halibut and sablefish are not included.
4. FL versus EM – Released Halibut (T4-Hal) – This test is specific to released halibut, where catch quantities are based on counting the number of pieces and applying a standard weight per piece (standardized by DFO as either 8 or 12 lbs, depending upon fishing area). In addition to estimating the number of pieces, all fish must be placed on a measurement grid to enable EM-based length assessment (to 5 cm); however, whether length presentation occurs or not does not form part of the test. A single test score is produced for each reviewed fishing event in which halibut are present.
5. FL versus EM – Released Sablefish (T5-Sab) – This test is specific for released sablefish, with legal and sublegal assessed separately. Similar to lingcod, released sablefish must be placed on a measurement grid to enable size differentiation. Those not presented for measurement are counted as legal size, except in cases where subsampling is carried out (those pieces not subsampled do not have to be placed over

the measurement grid). Piece counts are usually converted to weights using an estimated average piece weight, although there is no common standard used for piece weight. If the release quantities are high, a sample is generally used, although the sample collection method used is selected *ad hoc* at the moment, owing to the wide variety of contexts even within a trip, and applied to the release volume. Size presentation is not part of the test, but the EM reviewer classifies any sablefish as legal size if not presented for measurement determination. Legal and sublegal-sized sablefish (if present) are scored separately for each reviewed fishing event.

Table 1. Species Group Designations for FL and EM (or DMP) Catch Comparisons.

Species or Species Category	T1 - Landed Catch FL-DMP	T2 - Retained FL-EM	T3 - Rel Other Species FL-EM	T4 - Rel Halibut FL-EM	T5 - Rel Sablefish FL-EM
Rockfish	Tested	Tested	(prohibited)		
Hake	Tested	Tested	Tested		
Pollock	Tested	Tested	Tested		
Turbot	Tested		Tested		
Longnose Skate	Tested	Combined	Tested		
Big Skate	Tested	and Tested	Tested		
Other flatfish	Tested		Tested		
Pacific Cod	Tested	Tested	Tested		
Spiny Dogfish	Tested	Tested	Tested		
Lingcod (legal)	Tested	Combined and Tested	Tested		
Lingcod (sublegal)	(prohibited)	(not distinguished)	Tested		
Sablefish (legal)	Tested	Combined and Tested			Tested
Sablefish (sublegal)	(prohibited)	(not distinguished)			Tested
Halibut	(prohibited)	(prohibited)		Tested	
Salmon					
Other Species					

Comparison Test Scores

The five test areas described above use a similar method to assign a score to each individual species (or species group) comparison. In all cases, the catch estimates are grouped by the designated species categories, shown in Table 1, and the EM (or DMP) value is compared with the FL value. This comparison pair forms the basis for a single alignment test. Most often, this pair consists of two weight estimates, one by FL and the other by EM (or DMP). Less often, only one side records a weight and the other does not, usually because catch numbers are low and the species is difficult to detect consistently. In the former instance with two weight estimates, the absolute difference, or delta, between FL and EM (or DMP) is considered in relation to the size of the EM (or DMP) estimate, or 'reference value'. In the latter case, where the comparison pair is a weight and a null value, the delta is the difference between zero and the one estimate, with the reference value being either EM or FL, whichever recorded the weight. Throughout this paper, these two types of comparison pairs are distinguished because of their unique scoring effects. Hereafter, we refer to comparison pairs when both EM (or DMP) and FL detect and report a catch weight as 'matched pairs', while 'null comparisons' refer to instances when there is only one detection. The following notation is also used: EM+FL+ for matched pairs and either EM+FL- or EM-FL+ for null comparisons. In the case of T1 - DMP tests, the notation would be DMP+FL+, DMP+FL-, and DMP-FL+, respectively.

The delta and the reference values, as defined in the previous paragraph, are then used to determine the test score by referring to the tolerance definitions (Table 2). All comparison tests have a specified threshold value; when the EM estimate exceeds the threshold, the score is derived by the percent difference (delta divided by reference value multiplied by 100), whereas the absolute difference (delta) is used when the EM weight is below the threshold. For example, the score for retained Pacific cod is based on the delta if the EM estimate is less than 1,000 lbs, and on the percent difference when the EM estimate is above 1,000 lbs. This dual-scale approach allows higher tolerance at low catch amounts while maintaining low tolerances for higher catch quantities. Threshold values differ by test category. For example, DMP, retained, and released quantities have different weight ranges, as do species like pollock, hake, and arrowtooth flounder, which typically occur in larger quantities than species like lingcod and Pacific cod.

Table 2. Score Definitions by Test Area (see text for explanation).

Test	Score	Threshold	Score Assigned						
Area	Method	Range (lbs)	10	9	8	7	5	3	0
T1 - Landed Catch (FL-DMP)	Percent	> 1,000	0-2%	>2-10%	>10-20%	>20-30%	>30-40%	>40-50%	>50%
All Species Categories	Weight	< 1,000	0-150	151-250	251-400	401-550	551-700	701-850	851-1000
T2 - Retained Catch (FL-EM)	Percent	>500	0-2%	>2-10%	>10-20%	>20-30%	>30-40%	>40-50%	>50%
All Species Categories	Weight	<500	0-75	75-150	151-225	226-300	301-375	376-450	451-500
T3 - Released Catch (FL-EM)	Percent	>5,000	0-2%	>2-10%	>10-20%	>20-30%	>30-40%	>40-50%	>50%
Hake Only	Weight	<5,000	0-500	501-1000	1001-1500	1501-2000	2001-2500	2501-3000	3001-5000
T3 - Released Catch (FL-EM)	Percent	>2,000	0-2%	>2-10%	>10-20%	>20-30%	>30-40%	>40-50%	>50%
Turbot and Pollock Only	Weight	<2,000	0-200	201-400	401-600	601-800	801-1000	1001-1200	1201-2000
T3 - Released Catch (FL-EM)	Percent	>200	0-2%	>2-10%	>10-20%	>20-30%	>30-40%	>40-50%	>50%
All Other Species	Weight	<200	0-25	26-50	51-75	76-100	101-125	126-150	151-200
T4 & T5 - Released Catch (FL-EM)	Percent	>200	0-2%	>2-10%	>10-20%	>20-30%	>30-40%	>40-50%	>50%
Halibut and Sablefish	Weight	<200	0-25	26-50	51-75	76-100	101-125	126-150	151-200

Test Area Averages and Trip Scores

The next step in the scoring method is to combine individual comparison test scores to obtain test-area averages. Figure 6 is a simplified diagram of the T1, T2, and T3 test-averaging process (T4 and T5 are just one species), illustrating the elegance of the method in combining two dissimilar sets of information into a single score. In the diagram, the shape of the fish represents a species (or species group), while the size represents the size of the estimate. Note that both groups report four species, but only three are in common with both (species a, b, and d are matched pairs), and each group reports a species not reported by the other group (species c and e are null comparisons). The scoring method uses the delta and reference values to determine a score for each of the five species reported, with a gradient from green to red indicating high to low scores, respectively. The outcome on the right shows a range of high and low scores, with a test area average calculated from the values of the assigned scores.

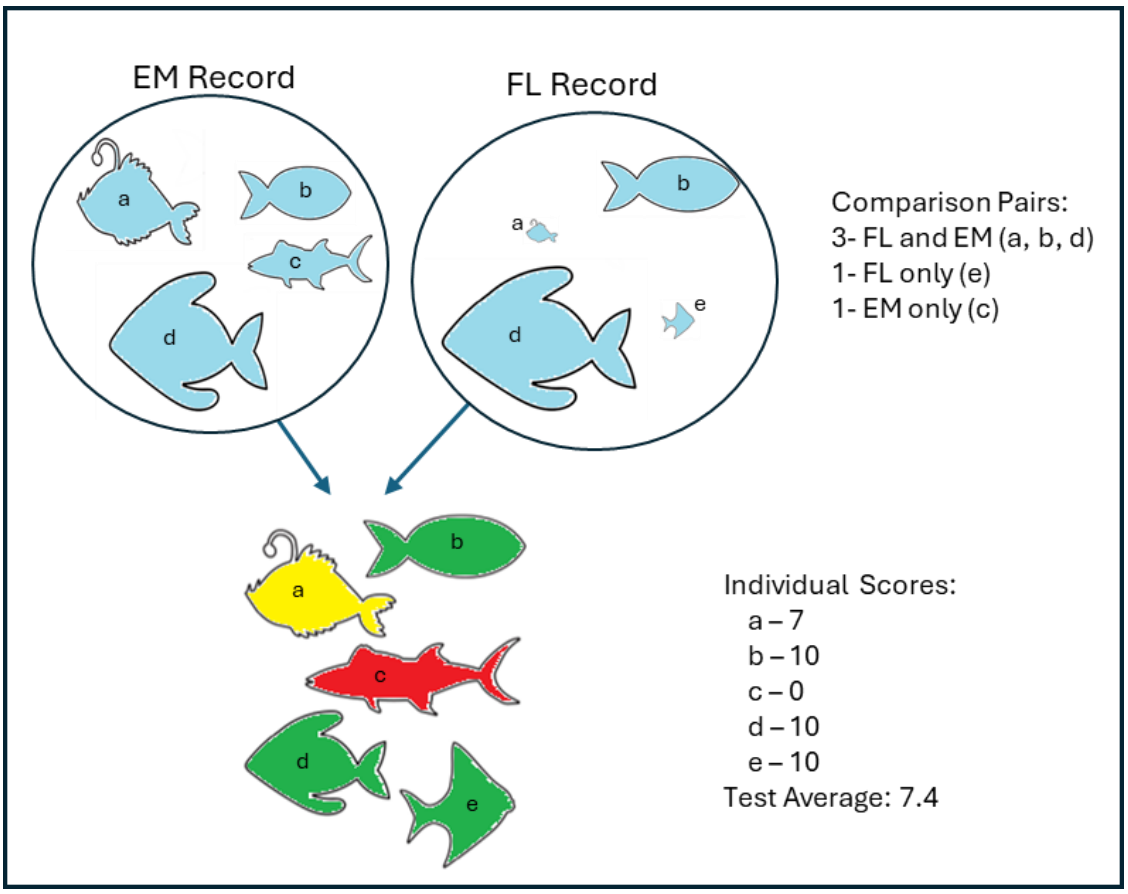


Figure 6. Simplified Example of Test Area Scoring Methods.

Once the five test area scores are calculated (assuming halibut and sablefish are present), they are averaged again to produce a trip score. Table 3 provides an example of the testing process for a typical fishing trip, starting with each test area, then the trip score. Included are landings data (FL-DMP comparisons) and FL-EM comparisons for one reviewed fishing event, with reported retained and released catch of various species. In total, 22 separate comparison tests are performed, with the number varying by test type. Since the tests are averaged first within each test group before averaging among groups, the relative weighting for each test varies. For example, the single halibut test accounts for 20% of the total trip score, while individual DMP tests account for just 4%.

Testing outcomes shown in Table 3 form the basis of the Audit Scorecard (Appendix 3), which compares the species across the five test areas, along with reported weights and test scores. This Scorecard is automatically generated as a PDF report and disseminated to skippers, DFO, and other specified parties. Also included is a second report called the Trip Summary Report (see Appendix 4). This report was a precursor to the Audit Scorecard and contains more reviewed tow details, total catch estimates by tow, and more granular species catch detail. The Trip Summary Report is not part of trip scoring but is useful for interpreting scoring patterns.

Table 3. Example of Catch Scoring for a Trawl Fishing Trip.

Test Area	Species or Sp Category	Estimated Weight (lbs) FL	EM/DMP	Test Scores Single	T1-T5	Trip
T1 - Landed Catch (FL-DMP)						
	Turbot	3,700	935	6		
	Big Skate	0	6	10		
	Lingcod	29,825	24,657	7		
	Longnose Skate	230	293	10		
	Other Flatfish	5,450	4,976	9		
	Pacific Cod	1,570	1,950	8		
	Rockfish	60,050	55,157	9		
	Sablefish	0	84	10		
	T1 Score				**8.4**	
T2 - Ret Catch (FL-EM)						
	Flatfish, Skate	50	0	10		
	Rockfish	19,700	20,896	9		
	Pacific Cod	200	6	8		
	Pollock	1,000	0	0		
	Sablefish	0	20	10		
	Lingcod	75	75	10		
	T2 Score				**7.8**	
T3- Rel Other Species (FL-EM)						
	Turbot	3,000	2,500	8		
	Flatfish	0	50	9		
	Pollock	8	8	10		
	Lingcod (legal)	40	24	10		
	Lingcod (sublegal)	10	45	9		
	T3 Score				**9.2**	
T4 - Rel Halibut (FL-EM)						
	All	60	72	10		
	T4 Score				**10**	
T5 - Rel Sablefish (FL-EM)						
	Legal	10	75	9		
	Sublegal	30	40	10		
	T5 Score				**9.5**	
						8.98

It should also be noted that the outcome of the audit process includes the generation of occurrence reports, incidents where evidence suggests that the vessel did not fully comply with licence conditions. These compliance incidents include releases of mandatory retention species, landings of undersized commercial species, or failure to comply with TMP requirements (obstructed views, camera failures, etc.). These reports are automatically generated and uploaded to the data portal for DFO. As of FY2026, occurrence reports are also noted on the Audit Scorecard, alerting the skipper to any issues.

While the test and scoring method is simple and useful for the purpose of compiling dissimilar catch information and providing a measure of overall FL data quality, the testing method converts absolute weights into rank scores. Rank scores are 'ordinal data' and have limitations in their use over the original weight data (also called 'continuous data'), creating scoring anomalies that are examined later in this document.

Audit Outcomes and Consequences

The intent of the scoring method is to provide a clear definition of FL data quality and a structured approach to the consequences that would apply when an audit produces a failure outcome. Audit outcomes and associated outcomes are specified by DFO (DFO, 2024), bringing together overall trip scores, test area scores, and data conformity measures (e.g., data gaps, fishing event alignment, video loss, etc.). The audit outcomes and category definitions are shown in Table 4. In practice, audit outcomes of a Pass and Category 1 are basically the same, as the FL data are accepted with no further action required. For Category 2-4 outcomes, the FL data are not automatically accepted, and the skipper is obliged to contact DFO within 14 days of when audit results are issued. Failure to do so will result in a hold being placed, preventing the vessel from hailing out to go fishing. For Category 2 and 3 outcomes, results are reviewed by DFO and discussed with the skipper. Based on the scoring results, DFO may require additional review of the fishing events. Decisions for additional review are made on a case-by-case basis, for example, T1-DMP comparison tests may contribute to the low score, and there would be little point in requiring additional review. Based on the Scorecard results and results from additional review, DFO may direct the replacement of FL estimates with EM catch estimates in the final data record, albeit done rarely. In the case of a Category 4 outcome, the data set would be examined by DFO and discussed with the skipper, the most common outcome being to ensure corrective measures are in place to minimize future occurrences. Hence, Category 2, 3, and 4 test results carry the consequence of the vessel potentially being delayed in fishing, having to communicate with DFO to review the audit results, having additional review of the data record, and having their data record revised with estimates derived from the EM record. The latter measures generally result in additional data processing fees being levied to the vessel.

Table 4. Schedule of Audit Outcomes, Outcome Definitions and Consequences.

Audit Outcome	Outcome Definitions	Consequences
Pass	Trip Score, 8 or greater	
	T1, T4, and T5 individual scores, 8 or greater	FL Data Accepted
	T2 and T3 average scores, 8 or greater	No Further Action Required
Category 1	Trip Score, greater than 7	
	One or more individual test scores (T1-T5) < 8	
Category 2	Trip Score, >4 and <=7	
	One or more individual test scores (T1-T5) < 8	FL Data Not Accepted
Category 3	Trip Score, <=4	Contact DFO
	One or more individual test scores (T1-T5) < 8	More Review Possible
Category 4	EM data set does not meet requirements	

Data Architecture

The overall data architecture spans from service providers to DFO, as shown in Figure 7. AMR hosts data processing associated with the DMP, EM, and Observer programs, while Vericatch hosts processes associated with the hail and FL programs. All FL (both legal unaltered copy and EM or Olog copy), Hail, and DMP are hosted on the Vericatch platform. AMR pulls data from the DFO FOS and Vericatch platforms to manage operations and compare data with EM to produce Audits. All EM data are housed within AMR's data management platform. Vericatch also maintains the interface for all finished data and manages synchronization with DFO servers.

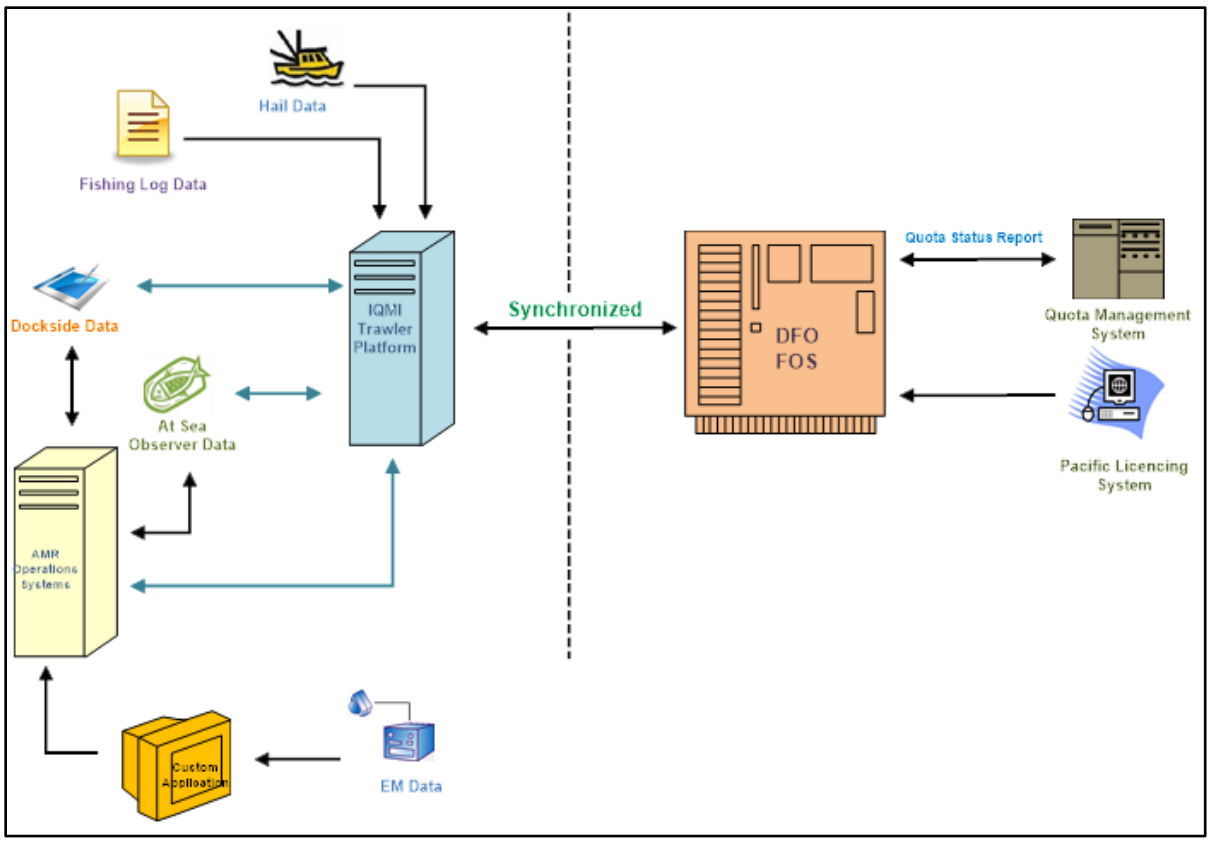


Figure 7. TMP Data System Architecture.

Materials and Methods

Overall Approach

The authors of this study acknowledge that they have both been retired for six or more years, were not involved in the development and implementation of the TMP and had little specific knowledge of the current status of the fishery. We began the study by becoming updated on the current trawl fishery, then by determining what information was available for an in-depth analysis of the TMP. Fortunately, there were numerous excellent resources available, including published materials and ready access to individuals directly involved with TMP design, operations, and TMP data systems. It was recognized early in the study that the trawl fishery had changed considerably over the past decade, particularly since the reduction of the hake fishery and the implementation of the current EM-based TMP. It became evident that historical program data would not be nearly as useful as current fishery year data; hence, the analysis focused on data sets that were as up to date as possible. This meant data sets would not include a full fishery year, and data sets pulled early in the study would differ from those taken later.

As mentioned in the Introduction, the Audit brings together different information sources for assessing the robustness of FL in five test areas. For the purposes of this study, the primary focus is on released catch (i.e., T3, T4, and T5), with less emphasis on retained catch. The logic for this decision was that the TMP is most dependent upon EM for estimating releases at sea, while DMP data are trusted as the source of landed catch data. Scoring inaccuracies with DMP (T1) or retained catch (T2) affect the spatial attribution of catch, whereas inaccuracies of released catch affect the estimations of quantities. Much of the scoring analysis is centred on released catch, although many of the observations apply to all test areas and are explained as such.

Background Materials

We began the investigation reviewing a large body of documents and reports which are referenced within this report, but fall within the following categories:

- BC Groundfish Fishery Management – These are fishery-specific documents by DFO, including management plans, license conditions, and TMP specifications documents.
- TMP Program Reports - These documents originate from DFO, CGRCS, or service providers, and provide much of the operational detail within the program. The largest part is AMR-generated Audit Scorecards and trip summary reports, created for each completed fishing trip. Also included are meeting presentations, reports describing TMP development, and specifications for certain program elements (e.g., halibut measurement).
- Fishery and Monitoring Program Literature – This includes a wide diversity of published primary and secondary literature describing the fishery and monitoring systems, with a

focus on EM technology. The material is mostly recent (in the last 15 years) and spans multiple geographies, including Canada, the US, Europe, and Australia.

Stakeholder Input

Early in the project, interviews were conducted with over 25 stakeholders and interested parties representing a wide range of perspectives, including DFO management, DFO science, DFO conservation and protection, CGRCS, trawl skippers, fleet managers, TMP service providers, an ENGO representative, representatives of other groundfish license sectors, and a lawyer with extensive BC fishery litigation experience.

The interviews captured various subject matter specific perspectives, but were consistent in exploring the following themes:

- Stakeholder background information;
- Understanding of the TMP;
- How well the program works and its strengths/weaknesses;
- Program and scoring system fairness;
- EM versus Observers for at-sea monitoring;
- Levels of trust in the data outputs in the program;
- Incentives/disincentives to comply with program requirements; and
- Opinions on specific issues (e.g., length presentation for EM measurement, electronic reporting, etc.).

Trawl Catch Data

To compare weight estimates between EM and FL for this study, AMR generated a data set using FL and EM catch observations by species (to the level of detail reported) for all reviewed fishing tows. We received all available data (as of mid-January 2026) for FY2025, as well as full data sets for the preceding three years (FY2022, FY2023, FY2024). Unlike Audit Scorecards and Trip Summary Reports, the fishery data set included FL and EM catch data, but no DMP catch data. The FY2025 data set included catch estimates from 715 reviewed tows on 523 completed trips by 21 vessels. Catch estimates were listed by species (when possible), which were collapsed into species or species categories corresponding to those used for comparison tests (Table 1). The data set included approximately 8,000 species-level weight estimates, which collapsed into about 5,000 comparison test pairs, about half of which were estimates of released catch. The detailed catch data set was accompanied by a trip-level data set showing pertinent details of the trip, including trip dates, data process schedule, audit outcome, and other information. These data sets were used in our analysis of data adequacy and scoring methodology.

Audit Scorecards and Trip Summary Reports

As mentioned, the Audit generates two reports, Audit Scorecards and Trip Summary Reports. These reports were examined based on the completed Audits for the FY2025 season, which included 461 completed fishing trips through mid-November 2025. These two reports are produced as PDF files and are the sources of information detailing the comparison tests and scoring results for completed fishing trips. While the queries could be rewritten to produce data instead of a PDF report, for the purpose of this study, we manually copied the data of interest into a spreadsheet. Due to time limitations, a random sample of 30 trips from each of Pass, Category 1, and Category 2+3 trip reports was transcribed for analysis. Category 4 trips were ignored, as the interest was only in trips with complete data sets. Because of the overwhelming dominance of Category 1 trips (i.e., 73% of 461 trips), sampling of this group was expanded, such that the sample rates were Pass (38%), Category 1 (21%), and Category 2+3 (73%).

The following information was extracted from the reports:

- Vessel details – vessel, skipper, hail #, and dates;
- Audit details – Category designation, T1-T5, and trip scores;
- Scoring details – number of tests by test area, with tests categorized by matched/null comparison, pass/fail, and % or weight method;
- Catch details – DMP/FL landed catch, EM/FL quantities for released and retained catch by reviewed tow; and
- Species of Concern – SOC notations made in FL/DMP or FL/EM portions of the report.

Statistical Analysis

This review relies on the most current data (FY2025 to January 2026) at the time of document preparation. While we suggest that the comparisons of EM and FL throughout the document provide insight into the veracity of the TMP, we emphasize that virtually all the paired comparisons are confounded by underlying influences in the data, which preclude exhaustive, classical statistical analyses. We emphasize that the work to follow is intended to capture qualitative signals in the data. We suggest that attempting classical statistics would be inappropriate, provide few benefits, and possibly lead to misleading conclusions.

Some of these confounding influences include, but are not limited to:

- the declining size and scale of the fishery from ASOP to EM, as well as a shrinking fishery within the EM period;
- the varying proportions of RT participation, varying from 6 vessels to one vessel and Wetfish vessels;
- the presence of a crew dedicated to catch estimation on some vessels;
- a mix of direct estimation of weight and conversion of numbers to weight or a fixed weight per piece logic (for salmon and halibut);
- too few observations for some species or category; and
- that alignment is conditioned by the choice of rolled-up species categories.

Additionally, statistical levels of significance are highly influenced by the sample size, or the "power" of each test. Testing the slope of one comparison with a sample size of 100 pairs and finding the slope (numerical bias) is not significantly different from another could result because the second sample is a comparison of 200 observations, even though both sets indicate the same slope and scatter. That is, the confidence limits of the slope will be smaller for the 200-pair comparison than that for the comparison with fewer observations. Finally, the impact of a few large data pairs is common through the comparisons. Overall, these data, like all fisheries data, are idiosyncratic and do not align well with the theoretical demands or assumptions of classical statistical tests

Slope characterizes how FL estimates increase relative to EM increases. A slope of 1.00 indicates congruence (e.g., both FL and EM estimates equal 100 lbs). In comparing FL to EM below, a slope >1.0 indicates a tendency (on average) for FL to be greater than EM; conversely, a slope <1.0 indicates a tendency for EM estimates to be higher. For example, a slope of 1.2 provides a point estimate that FL estimates, on average, a value 20% larger ("overestimate") than the matching EM estimate. We emphasize, as above, that these values are included to provide qualitative insight into slope and imprecision, not for rigorous hypothesis testing.

Simple, or Least Squares Regression (LSR), is a statistical method for finding the best-fitting linear model between *y* (FL) and *x* (EM). In brief, it finds the straight line and its slope that minimizes the distance between observed data points and the fitted line. However, the LSR model "assumes" that the *x* variable is measured without error. For example, this would be the case where *x* is preset, such as studying the effect of dosage (*x*) on recovery (*y*) , where dosages are precisely administered (e.g., 0mg, 10mg, 20mg) but recovery time (*y*) is more crudely measured.

Since the TMP and this review assumes that the individual EM estimates vary around the hypothetical "correct" value (or even around a theoretical average EM value if repeated measurements were collected), not only is LSR technically inappropriate for the comparisons below, but the impact is that the larger the imprecision in EM estimates, the more the fit in using simple LSR underestimates the slope, in turn, increasingly indicating that the FL underestimates the EM. This review used Type II regression, which is a more appropriate model when there is uncertainty in *x* (Ricker, 1973; Legendre and Legendre, 2024).

In addition to displaying the alignment of EM to FL estimates in simple scatterplots, we show the fit to a linear regression and estimates of slope (steepness or numerical bias) and precision (degree of scatter, R^2, or coefficient of determination). In Model II Major Axis (MA) regression, the R^2 is calculated the same way as in LSR. Specifically, it is computed as the squared Pearson correlation coefficient between the two variables *x* and *y*.

Note also, that in this analysis the regression treatment fits the lines by allowing the intercept to be freely estimated, as opposed to forcing the fit to pass through the origin (x=0 and y=0). There are strengths and weaknesses to both approaches. In theory, allowing the intercepts to be freely estimated can partially "absorb" disagreement allowing the slope to appear artificially closer to .

However, in most of the comparisons (see Figures 8-11 and Tables 9 and 12), the high proportion of low valued pairs (0,1; 1,0 etc.) either renders an intercept virtually equal to the origin or the intercept is very small relative to the average *x* and *y* values. In either case, the two approaches would render similar slopes. The exceptions were species for which there are few and highly imprecise data (see Figure 11 for Big skate released).

This review employs a range of terms to describe and evaluate the data produced by the TMP. While these terms are rigorously defined in a substantial body of theoretical literature for catch estimation (Allard and Benoît, 2019; Davies et al., 2023; DFO, 2020), the most important distinction for this document is between absolute and relative measures.

With respect to absolute values, *quality* refers specifically to **statistical data quality** in an absolute sense. That is, the degree to which an estimate (e.g., total catch) approximates the *true* value. This metric is characterized using standard statistical properties: **numerical bias** and **variability**, along with their corresponding converses or opposites, **accuracy** and **precision** (i.e., bias=1/accuracy; variability=1/precision). Thus, absolute measures are not contextual and, in themselves, provide no statement about the adequacy of whether they are good enough, or whether they meet the needs of the catch monitoring.

The term *bias* is used strictly in its statistical sense, referring to the expected difference between an estimate and the true value. It does not imply intentional misreporting. To reinforce this distinction, the term is presented as *numerical bias* throughout this document. An estimation process is considered **negatively** or **positively** numerically biased if it tends, on average, to under- or overestimate the true value of a statistic, respectively. Formally, numerical bias is defined as the expected value of the difference between estimated and true values, assuming the estimation process could be repeated many times.

Precision is a measure of the degree to which those repeated estimations are close in value (scatter). Precision does not have a sign (it is a positive number). If the same estimation process is repeated many times, the estimates may differ due to the randomness of the sampling process or other characteristics of the estimation process. Similarly, *numerical accuracy* is used here in its strict statistical sense, as distinct from its common, non-technical usage. It refers to a quantitative property derived from statistical analysis and is independent of context. As with bias, the qualifier *numerical* is used to distinguish this meaning from a different non-technical interpretation of “accuracy” that relates to how well an estimate meets a specific practical need.

Although numerical measures such as numerical bias and accuracy are essential for characterizing data quality, they do not, on their own, determine whether an estimate is adequate for its intended purpose. For this reason, the review also considers **relative adequacy**—that is, whether TMP-derived estimates are adequate, or “good enough”, to support decision-making. For example, one may ask whether estimates of total catch are adequate for effective quota management or conversely, whether the inaccuracy is sufficiently large to be consequential.

Importantly, judgments of relative adequacy are inherently:

- Dynamic: What is considered adequate at one moment may not remain so, as stock status or management objectives change, or regulatory requirements evolve.
- Context-dependent: The level of accuracy required may differ among species, fisheries, or management applications.
- Scale-dependent (granularity): The TMP may provide sufficient accuracy for aggregated measures, such as annual coastwide catch, while being less adequate for more detailed applications, such as the management of individual vessel quotas.

Results

TMP Operational Performance

The TMP has been operating continuously since its inception in April 2020 and appears to be fully deployed and operational. Table 5 provides a summary of performance outcomes for all fishing trips, listed by time period, since August 2021. Note that audit outcome designations were not introduced until February 2023, and the values shown for previous time periods were retrospectively constructed. While specific audit outcome results will be discussed later, the focal point here is the proportion of Category 4 outcomes – those fishing trips that failed to meet monitoring specifications. The proportion of Category 4 trips has been small for every year of the program (FY2021- FY2025), ranging from 2% to 7%. Interviews with skippers revealed one instance where a vessel had to return to port because the EM system was inoperable, but there were no other instances reported, nor were there comments concerning the reliability of EM systems. AMR reported that there have been no instances of aborted fishing trips related to EM equipment failure in the current fishing year (FY2025). Based on the information received, there were no instances when vessel fishing schedules were held up or delayed. A review of AMR's at-sea call logs indicated 19 calls in total from 9 vessels seeking assistance on a range of technical issues, of which about half were resolved remotely, with one resulting in a Category 4 outcome. This level of at-sea calls appears low considering there were over 500 seemingly satisfactory fishing trips completed.

Table 5. Summary of Audit Outcomes by Time Period.

Period	Pass	Cat1	Cat2	Cat3	Cat4	Total
Aug21-Feb22	30	143	63	4	9	249
Feb22-Jul22	37	162	56	1	4	260
Jul22-Feb23	37	299	75	0	25	436
Feb23-Feb24	84	395	83	0	43	605
Feb24-Feb25	68	439	56	0	27	590
Feb25-Nov25	74	334	40	1	11	460
Row Percentages						
Aug21-Feb22	12%	57%	25%	2%	4%	249
Feb22-Jul22	14%	62%	22%	0%	2%	260
Jul22-Feb23	8%	69%	17%	0%	6%	436
Feb23-Feb24	14%	65%	14%	0%	7%	605
Feb24-Feb25	12%	74%	9%	0%	5%	590
Feb25-Nov25	16%	73%	9%	0%	2%	460

Program performance assessment also concerns timelines. A review of FY2025 audit report tracking data indicated that 98% of the trip reviews were completed within 10 calendar days (the requirement is 10 business days), while about 40% were completed within five days. These

completion timelines are significant because they encompass all the necessary steps of data gathering, processing, and reporting following completion of a fishing trip.

Another dimension of overall program performance concerns fleet sample rates. As noted above, the Wetfish fleet is sampled with a target rate of 10% with a minimum of one tow per fishing trip, while the RT fleet is sampled at 25%. Applying these percentages to the number of tows on any given fishing trip yields different sample rates – that is, a single tow on an 8-tow trip yields 12.5%, while a single tow on a 12-tow trip yields an 8.3% sample rate. Across the fleet for all trips, the realized sample rates were higher than specified, with the Wetfish fleet averaging 12.9% and 13.4%, and the RT fleet 27.8% and 25.9% for the past two years, respectively.

It is helpful to look more closely at the Category 4 failures before turning to the successfully monitored audit outcomes (i.e., Pass, Category 1 to 3). Category 4 causes are summarized for time periods starting July 2022 (Table 6). The main cause of failure is a “data gap”, meaning some or all of the image and sensor data were not recorded. The reasons for a data gap vary, from a failure to turn on the system, a hard drive failure, or a system fault. The second cause of a Category 4 outcome relates to difficulties in aligning FL and EM data sets, each reporting different numbers of fishing events. Prior to February 2024, FL and EM defined the start and end of fishing events differently, such that FL reported a single fishing event when the vessel made multiple passes over the same spot with doors coming in and out (but with the net remaining in the water), while EM defined a single cycle of doors going out and in as a single event. Operational definitions used by FL and EM were aligned, and the issue has mostly disappeared. The third cause for a Category 4 designation is camera failures, when no image data are available. These instances are minor because vessels carry multiple cameras with overlapping views to guard against situations where the view from a single camera may be unusable (e.g., moisture, dirt, glare, etc.).

Table 6. Summary of Category 4 Failures by Time Period.

	Category 4 Issue			Totals	
Period	**Camera**	**Data Gap**	**# Tows**	**Cat 4**	**All Trips**
Jul22-Feb23	2	13	10	25	436
Feb23-Feb24	0	23	20	43	605
Feb24-Feb25	1	21	5	27	590
Feb25-Nov25	0	11	0	11	460
Row Percentages					
Jul22-Feb23	0.5%	3.0%	2.3%	5.7%	436
Feb23-Feb24	0.0%	3.8%	3.3%	7.1%	605
Feb24-Feb25	0.2%	3.6%	0.8%	4.6%	590
Feb25-Nov25	0.0%	2.4%	0.0%	2.4%	460

During the most recent fishing year, all eleven Category 4 failures were caused by data gaps. In two instances there was 100% data loss caused by a hard drive failure, and the remaining

nine were partial data losses ranging from a few minutes to several hours, the causes being both human error (forgetting to turn the system on) and technical error. Catch scoring was possible to some level for all eleven trips, ranging from T1-DMP only for the two trips with hard drive failure, to a full audit on the remaining nine trips, except that some fishing events were reselected when data gaps prevented viewing the selected fishing event (4 of 9 instances).

Turning towards successfully monitored outcomes, the success rate ranged from 93% to 98% of all fishing trips since August 2021 (Table 5, all trips other than Category 4). Among these outcomes, the majority were Pass and Category 1, outnumbering Category 2 outcomes by a factor of three for most years, and a factor of nine for recent years. Category 3 outcomes have always been very low (4 or fewer). Category 1 has consistently been the most common outcome, with the proportion increasing from 57% to 73% since inception, the increase due to the decline of Category 2, rather than changes in Pass outcomes. As mentioned, Pass and Category 1 have the same operational consequence of meeting program standards, despite Pass representing a higher standard of acceptance.

Vessel-specific audit outcomes from FY2025 are summarized (Table 7) with vessel names replaced with anonymous number values. Most of the vessels completed 20 or more fishing trips. Interestingly, most vessels had three or more different audit outcomes across trips completed during the year, with the Category 1 outcome dominating for all vessels. The proportion of Pass outcomes was high (27%-49%) for three vessels, but more normally, most vessels had only a few Pass outcomes. Two-thirds of the fleet had one or more Category 2 audit outcomes, while five vessels had five or more trips with Category 2 results. Yet, even among these vessels, Category 1 was the dominant audit outcome for trips completed.

Table 7. Number of Fishing Trips by Vessel, Summed by Audit Outcome (FY2025 data).

Vessel Ref	Pass	Cat 1	Cat 2	Cat 3	Cat 4	Total
1	18	23	1			42
2	19	17	3			39
3	10	22	2	1	1	36
4	4	29				33
5	2	29	1			32
6	5	23	4			32
7	7	22	2			31
8		23	5		1	29
9	2	25	1			28
10	4	23				27
11	1	20	3		1	25
12	2	22	1			25
13		18	4			22
14	2	19			1	22
15		18	1		3	22
16	4	15	1		1	21
17	1	11	5		2	19
18		11	8			19
19	4	12			1	17
20		1				1
21		1				1
22		1				1
Total Trips	**85**	**385**	**42**	**1**	**11**	**524**
Total Vessels	15	22	15	1	8	22
% Trips	**16.2%**	**73.5%**	**8.0%**	**0.2%**	**2.1%**	**100.0%**
% Vessels	68.2%	100.0%	68.2%	4.5%	36.4%	100.0%

Category 2 and 3 designations represent fishing trips with trip scores of 7 or lower (see Table 4). To gain a better perspective of follow-up actions taken for these trips, Table 8 provides a summary of outcomes resulting from Category 2 and 3 trips from the two most recent fishing years. Among these Category 2 and 3 trips, half or less were escalated to require additional review, and of those trips, most were at the 10% or 20% rate (i.e., one or two tows). Only a few trips (4 over two years) saw additional review rates of 30% and 100%. The FL record was updated with EM results in 16 trips in each of the two years, amounting to about one-third of all Category 2 and 3 trips. The updated data were estimates of released halibut and sablefish, and less frequently, lingcod. It is important to note that the updated data was made only on reviewed tows (i.e., cases where EM data were available).

Table 8. Summary of Trips With Additional Review by Time Period.

Period	# Trips	w/ ++Review	w/ FL Update	10%	20%	30%	100%
Feb 24 to Feb 25	56	23	16	14	7	2	0
Feb 25-Nov 25	41	25	16	17	8	1	1

TMP Data Adequacy

The most critical and problematic element of the TMP (and most fishery catch monitoring programs) is the generation of adequate estimates of released catch, notwithstanding the need for adequate DMP estimates of landed catch. To reiterate, TMP treats FL as the main and only complete source (census coverage) of releases by tows and species (or species group). However, the TMP includes a random review of tows by EM to compare individual catch estimates with matching FL estimates draw inference on the veracity of the FL. It is like the BCHLCMP fishery monitoring upon which elements of the TMP are based (Stanley et al. 2009, 2011, and 2015).

The EM estimates are therefore treated, with reservation as a quasi-standard in that there is no reason to assume the observations are deliberately altered, and there is more time for careful review and even re-review. But, it should be emphasized that the EM estimates are associated with their own degree of uncertainty owing to many factors, for example:

- Species identification;
- Estimation of quantities;
- Increasing experience by each reviewer (see also Grinnell 2010, for evolving skills of ASO's);
- Variation among reviewers;
- Differences in viewing and catch handling conditions on trawl vessels.

Alignment of Released Estimates of Major Commercial Species

The following section focuses on the alignment of EM and FL release estimates for FY2025. The alignment statistics are provided for selected species in Table 9, and Figures 8 and 9. The theoretical 1:1 perfect match line (dashed line) is indicated in the plots. All paired release estimates for FY2025 for all species are provided in Fig. 8A. It indicates a tendency for FL to overestimate by about 18% with imprecision (R^2=0.79).

Note that the following species-specific analyses, tables and figures, omit salmon. The tracking methods for salmon catches is still in a state of flux. Firstly, while salmon retention is technically prohibited the TMP currently requires retention and relinquishment at dock for better bycatch estimation and bycatch quota reconciliation against a coastwide cap. While the audit still has a release test, it doesn't test retention. Secondly, fisher-based retention estimation techniques vary from estimating weight directly to converting head counts to weight (fixed at 1 lb). Finally, there are too few data for conducting a meaningful comparison of FL to EM sources (see

recommendations and note the constantly evolving treatment of salmon monitoring methodology summarized in Appendix 2.

During FY2025, many of the skippers were reporting salmon as released even though they were retained owing to a misunderstanding of in the new protocol. As the results for FY2025 were collected during this state of flux in the program. This issue is expected to be clarified for FY 2026/2027.

Table 9. Regression Statistics for FL Releases vs. EM Releases by Species or Category. N represents individual comparison tests.

Species or Categories	Slope	R^2	N	Intercept
All Species Observations	1.18	0.79	3,142	-14.9
Arrowtooth flounder	0.86	0.78	251	9.5
Lingcod	1.11	0.74	259	1.4
Hake	1.69	0.75	145	143.9
Sablefish (lg pair removed)	0.86	0.78	251	9.5
Other Flatfish	1.34	0.64	201	-9.6
Rockfish	NA	NA	636	NA
Pollock	0.92	0.68	163	52.1
Pacific cod	1.57	0.57	57	-3.7
Longnose skate	1.05	0.42	198	-3.0
Big skate	58.48	0.00	16	1056.5
Haibut	1.11	0.93	317	-4.9
Spiny dogfish	1.00	0.91	377	2.7

The individual scatterplots of the alignment at the individual species level for lingcod, longnose skate, Big skate, and Pacific halibut are provided in Figure 8B-F. Figure 9A-B show the same plots for Spiny dogfish and sablefish releases. These show the variety among the comparisons. Two additional sablefish plots are included (Figure 9B-9C) to examine the relationship of releases in comparison to the sablefish Scorecard tolerances of 75 lbs with catches of <200 lbs and 20% for catches of >200 lbs.

Overall, FL tends to overestimate EM in 7 cases, slope >1.0. Arrowtooth flounder, sablefish and pollock, with reasonable sample sizes indicate FL underestimates. Big skate, and rockfish results are based on too small a sample size infer a pattern.

The FL tended to overestimate releases for lingcod (by 11%) with notable imprecision (R^2=0.74) but since only a small proportion of the total catch is released with most retained and recorded by DMP, overall catch is well estimated by the TMP (see Tables 10 and 11). Similarly, for hake (69%), other flatfish (34%), Pacific cod (54%), and other species, since the majority of the total catch is retained for these species, the TMP can be assumed to provide sufficiently adequate

estimates of total catch, even though the slopes in the release data indicate a significant numerical bias. In particular, the FL vs. EM release comparisons for these species' veracity check does not point to a concern of serious over-reporting, or, more importantly, under-reporting of releases by FL. This satisfies the objective of estimating total catch for quota management and stock assessment. For this not to be the case, the scatterplot would indicate small FL estimates paired with large EM estimates, or slope values well below 1.0. Furthermore, the slope estimate possibly provides a quantitative estimate of a potential correction factor to align FL and EM, if required. It is worth noting that the FL tends to overestimate more often, even though it might be assumed that there is an incentive to underreport to avoid IVQ constraints, even if modestly. It is noteworthy that during interviews, the skippers were concerned with achieving a good Scorecard result to avoid having to consult with DFO staff over the misalignment in the results and risk delays in processing catch settlements.

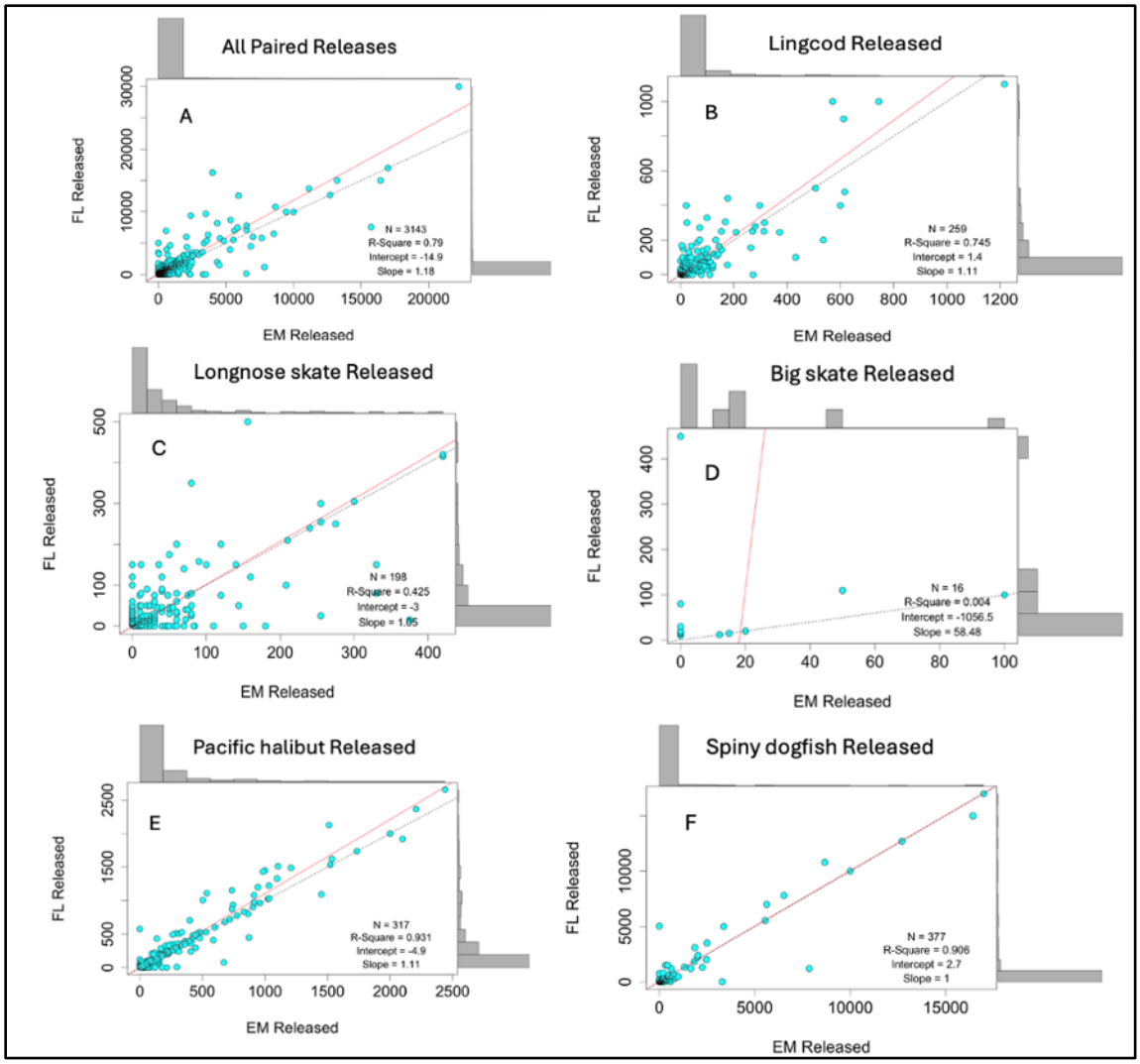


Figure 8. FL and EM Release Estimates All Species Combined (A), then Lingcod (B), Longnose skate (C), Big Skate (D), Halibut (E), and Spiny Dogfish (F). Sample Size, Regression Coefficient of Determinations, Intercept, and Slope are Shown within Each Plot.

Although there is little numerical bias (slope=1.04) indicated for release estimates of Longnose skate, the precision is poor (Figure 8C). The Big skate comparison is uninformative owing to the lack of data.

FL vs. EM halibut release estimates indicate reasonable precision (R^2=0.93) (Figure 8E). FL tends to overestimate EM by approximately 11%. As with other species, there is no evidence of large-scale unreported releases (i.e., very low FL versus EM).

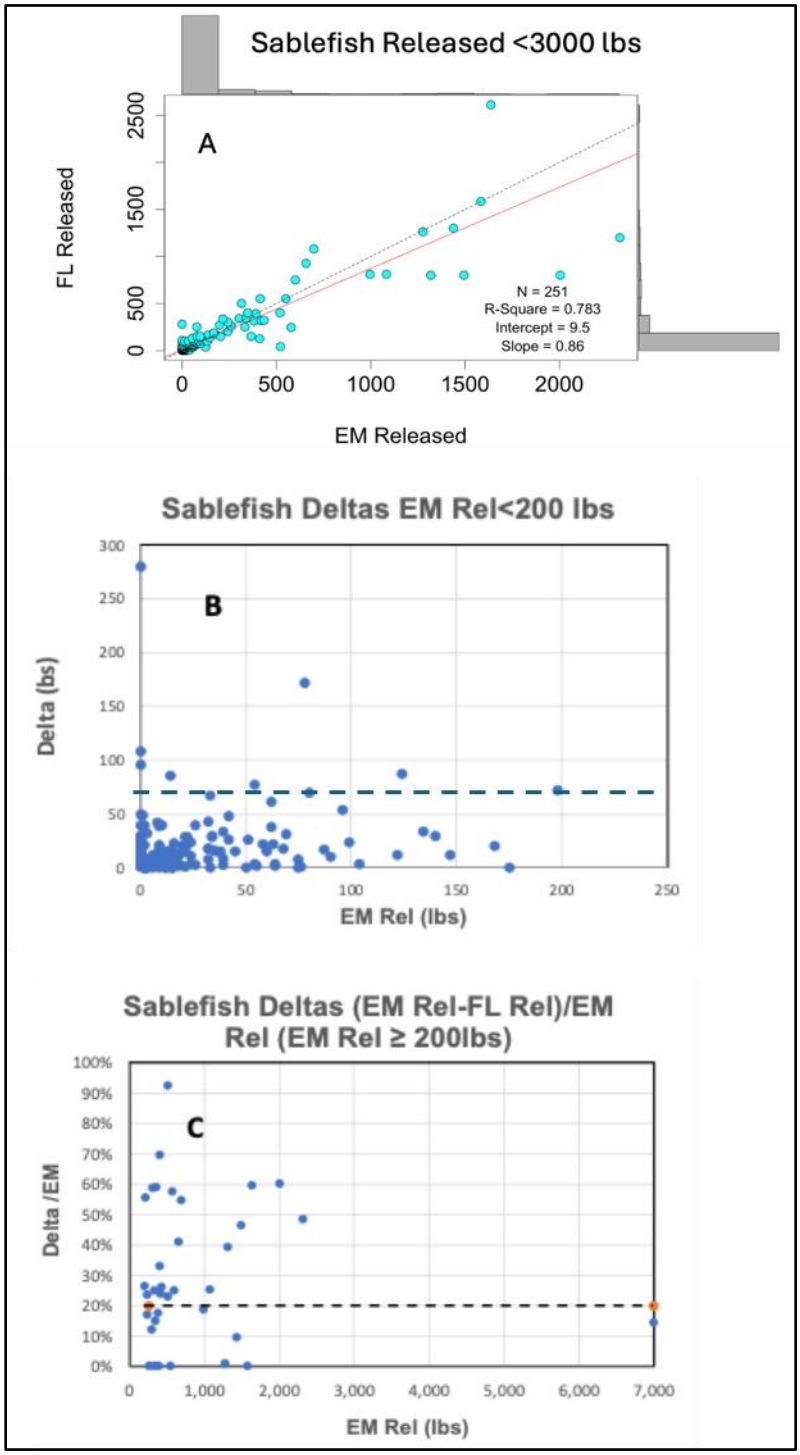


Figure 9. FL and EM Release Estimates for Spiny dogfish (A) and Sablefish (B). Plots C and D Show the Delta (% Difference) in Relation to the EM Estimate, with Catch Values and < 200 lbs and ≥ 200 lbs Plotted Separately.

Spiny dogfish (Figure 8E) releases appear better estimated than most other species (or species groups) even with larger catches. There is no numerical bias (Slope=1.0) and there is higher precision (R^2=0.091). While there is currently no fishery for Spiny dogfish, recent assessment work has noted a declining population trend (DFO 2025). Should this trend, Spiny dogfish may

be designated an SOC. This in turn will place a premium on adequate release estimates. However, like Pacific halibut, there is currently no evidence of larger-scale unreported releases.

The overall fit of sablefish releases (Figure 9A) would be influenced by one large pair (EM 7,000 lbs: FL 6,000 lbs). This document therefore provides a truncated version with pairs that only have EM estimates of <3,000 lbs (i.e., removing one pair). There is relatively poor alignment between FL and EM estimates of sablefish. It indicates a bias for FL to underestimate sablefish releases by about 14% (Slope=0.86) and weak precision (R^2= 0.78). The misalignment is notable.

Some of the imprecision in sablefish could result from the lack of a standard average weight used to determine sablefish releases, unlike halibut. EM reviewers will use an average weight for the pieces released. Skippers all use different methods to determine which weight to use.

The sum of the differences (deltas) between total FL releases and total EM releases (FL-EM) is divided by total EM to show how large the difference is in proportion to the total estimate (in EM) in Table 10. In other words, it characterizes the relative magnitude of the difference between FL and EM. The average across species (excluding rockfishes because of near 100% retention) varies between 30% and 20%, there is no distinct trend over the 4 years, although FY2025 indicates the lowest value. However, the averaging is highly influenced by certain species. This includes, for example, Pacific cod, which ranges from 37% to as high as 169% over the years. Like many of the quantitative data summaries in the review, these comparisons must be reviewed with caution.

This misalignment is important for species that are mostly released (e.g., Spiny dogfish) or for species-specific IVQ management among vessels, but arguably less important for those quota species, for which the major proportion of catch is verified at DMP (Tables 10 and 11). For the commercial species, the minimum requirement of the FL vs. EM veracity check is to provide reassurance that the fishery is not releasing relatively large quantities above the FL estimates, quantities that could jeopardize effective quota management and stock assessment.

The aggregated category of “other flatfish” precludes a species-specific examination of alignment between the FL and EM. Nevertheless, Tables 10 and 13 show that approximately 91% of this group is retained and therefore sorted to species at dockside. The Conditions of Licence (COL) for this fishery require a minimum cod-end mesh size of 152.6 mm (6.01 in.) on the north coast and 140 mm (5.51 in.) on the remainder of the coast. One north coast skipper who reported using a 165 mm (6.5 in.) cod-end stated that he releases virtually no undersized flatfish.

Overall, the magnitude of the total delta between FL and EM is small compared to the total catch (Table 11).

Table 10. Difference Between Total Releases (lbs) by EM and FL Relative to Total EM by Fishery Year.

			FY 2022					FY 2023		
Species	**Rel EM**	**Rel FL**	**Rel Delta (EM-FL)**	**Rel % Difference**	**Abs Rel% Difference**	**Rel EM**	**Rel FL**	**Rel Delta (EM-FL)**	**Rel % Difference**	**Abs Rel% Difference**
Spiny dogfish	224,748	218,993	5,755	3%	3%	178,349	213,667	-35,318	-20%	20%
Big skate	6,946	5,788	1,158	17%	17%	5,722	2,646	3,076	54%	54%
Longnose skate	27,578	18,621	8,957	32%	32%	28,772	21,578	7,194	25%	25%
Pacific cod	6,226	3,934	2,292	37%	37%	516	1,288	-772	-150%	150%
Pacific hake	294,251	504,707	-210,456	-72%	72%	376,602	413,149	-36,547	-10%	10%
Walleye pollock	434,658	571,800	-137,142	-32%	32%	1,061,794	973,832	87,962	8%	8%
Sablefish	181,311	226,329	-45,018	-25%	25%	233,298	232,410	888	0%	0%
Lingcod	16,353	18,809	-2,456	-15%	15%	26,241	35,948	-9,707	-37%	37%
Flatfish	93,388	90,867	2,521	3%	3%	106,442	82,006	24,436	23%	23%
Turbot	88,753	151,355	-62,602	-71%	71%	185,938	204,244	-18,306	-10%	10%
Pacific halibut	150,720	190,076	-39,356	-26%	26%	225,705	219,637	6,068	3%	3%
Average					30%					31%
			FY 2024					**FY 2025 (to Jan 13, 2026)**		
Species	**Rel EM**	**Rel FL**	**Rel Delta (EM-FL)**	**Rel % Difference**	**Abs Rel% Difference**	**Rel EM**	**Rel FL**	**Rel Delta (EM-FL)**	**Rel % Difference**	**Abs Rel% Difference**
Spiny dogfish	179,774	234,445	-54,671	-30%	30%	144,292	140,715	3,577	2%	2%
Big skate	3,101	2,962	139	4%	4%	741	307	434	59%	59%
Longnose skate	19,218	15,029	4,189	22%	22%	9,375	9,191	184	2%	2%
Pacific cod	1,431	3,849	-2,418	-169%	169%	1,673	2,418	-745	-45%	45%
Pacific hake	48,310	76,826	-28,516	-59%	59%	110,568	167,529	-56,961	-52%	52%
Walleye pollock	99,970	106,239	-6,269	-6%	6%	83,044	81,427	1,617	2%	2%
Sablefish	143,307	151,956	-8,649	-6%	6%	36,515	33,866	2,649	7%	7%
Lingcod	25,675	28,688	-3,013	-12%	12%	17,210	19,882	-2,672	-16%	16%
Flatfish	187,014	164,201	22,813	12%	12%	39,659	51,251	-11,592	-29%	29%
Turbot	141,283	281,231	-139,948	-99%	99%	69,423	70,888	-1,465	-2%	2%
Pacific halibut	129,524	123,538	5,986	5%	5%	69,052	74,951	-5,899	-9%	9%
Average					39%					20%

Table 11. Estimated EM-FL Released Catch Delta Compared Against Total Catch by Species for FY2025.

Species	Total catch (lbs)	Total EM-FL (lbs)	Delta Relative to Total Catch
Arrowtooth flounder	2,318,293	8,752	0.4%
Big skate	28,843	3,599	12.5%
Hake	9,340,726	-215,237	-2.3%
Halibut	491,608	-46,662	-9.5%
Lingcod	805,757	-19,843	-2.5%
Longnose skate	115,776	696	0.6%
Pacific cod	501,075	-5,351	-1.1%
Sablefish	626,923	23,637	3.8%
Spiny dogfish	778,119	6,148	0.8%
Walleye pollock	14,798,461	7,994	0.1%

Release Estimation by Tow

The previous discussion focused on the larger-scale picture of whether the TMP generally provides sufficient catch estimates for total catch; the uncertainty examined in this study does not appear to jeopardize quota management. This is supported by Tables 10 and 11, which show the relative sizes of the discrepancy between total FL and total EM released as a proportion of total EM. Misalignment of catches at the individual tow level by species is more important for IVQ management at a more granular level and for aligning catches to area, although this issue was not emphasized during stakeholder interviews.

In Figure 9B, sablefish deltas of EM Rel - FL Rel are plotted against EM weight for EM Rel estimates of <200 lbs. The horizontal bar indicates the Tolerance of 75 lbs in the Audit Scorecards. The points above the bar correspond to Scorecard failures. Of 216 paired observations where EM<200 lbs., the delta exceeded the tolerance limit 7 times (Scorecard fails=3%). All seven were cases of FL overestimating the releases.

In Figure 9C, sablefish deltas of EM Rel - FL Rel against EM weight for EM Rel estimates for larger catches of ≥ 200 lbs. The horizontal bar indicates the tolerance level of 20%, where the delta is expressed relative to EM estimate. The points above the bar indicate Scorecard fails. Of 35 paired observations where EM ≥200 lbs, the deltas exceeded the Tolerance of being within 20% for 26 of 35 times or 74% of the events. All were cases of FL underestimating releases in excess of the 20% Tolerance and represent fails.

Alignment of Retained Catch Estimates

One historical concern with the ASOP was that, because observers did not work around the clock and there were reports of coercion (Grinnell, 2010), retained catch may have been “dumped” prior to unloading. This could not be tested under the ASOP. Under the current TMP, however, it is tested because EM review records include reviewers’ estimates of total catch for every tow. Total catch estimates by fishing event for FL and EM can be compared, and totals for the trip can be compared with DMP. This was not examined in the present report because the information was only available in the PDF versions of the Trip Summary Report. However, this opportunity is illustrated in Figure 10 which plots total catch (Rel+Ret) for FL and EM from tows sampled in the FY2025 Audit Scorecards. The slope and R^2 are close to 1.0 indicating close alignment. More particularly, the estimate of FL total is not noticeably lower than EM total catch which, if present, would indicate dumping. This is based on 712 reviewed tows but the approach could be applied to 100% of the tows going forward. As well, the total catch estimate from EM could be summed by trip and compared with the DMP result.

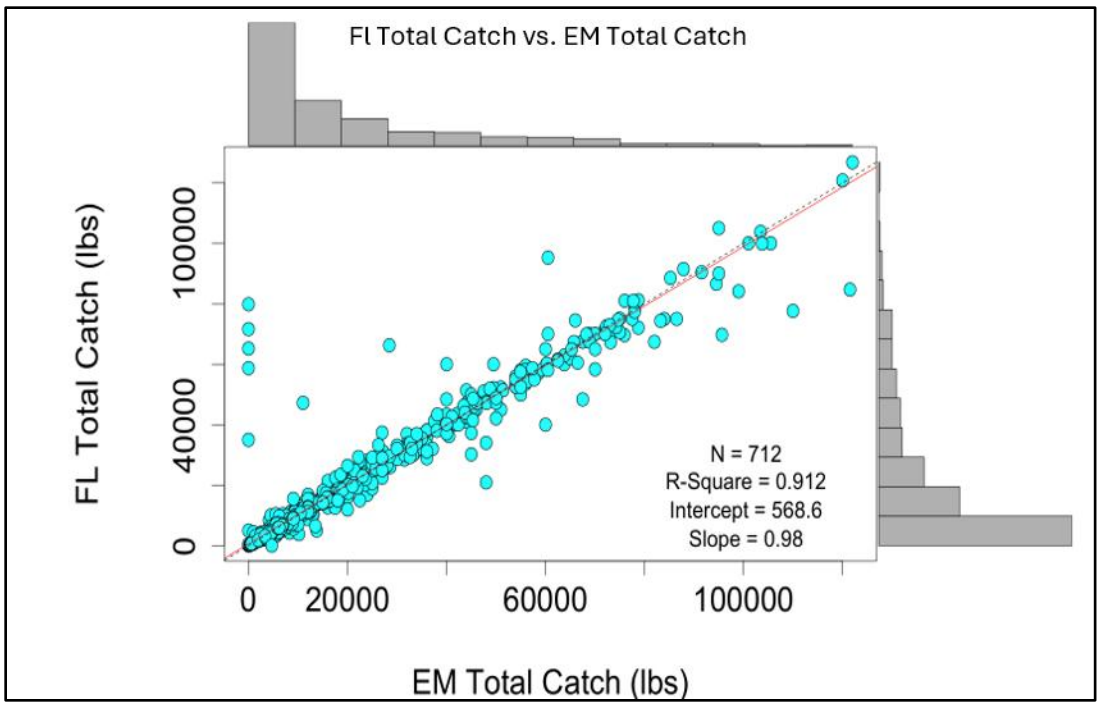


Figure 10. Total FL Catch Versus Total EM Catch for Sampled Tows from Audit Scorecard Sample (FY2025 partial).

The alignment of FL and EM retained estimates are presented in Table 12 and Figure 11. The alignment of FL total retained catch vs. DMP at the coarsest level of Trip total catch was examined for 130 randomly chosen trips taken from the Audit Scorecard sample (Figure 10A).

Table 12. Regression Statistics for FL Retained to EM Retained. Halibut and spiny dogfish omitted because of too few observations. Salmon omitted because retained "weights" standardized to be 1lb in RVs.

Test and Species	Slope	R^2	N	Intercept
FL vs. DMP (sampled trips)				
All Species	0.98	1.00	130	1,248.30
FL vs. EM (sampled tows)				
All Species Observations	0.99	0.94	2,928	87.60
Arrowtooth flounder	1.00	0.97	193	107.80
Lingcod	0.93	0.96	364	56.80
Hake	0.96	0.92	137	315.90
Sablefish	0.96	0.94	209	56.50
Other flatfish	0.93	0.89	399	111.10
Rockfishes	1.00	0.98	635	-74.80
Pollock	0.98	0.87	234	773.50
Pacific cod	1.30	0.85	273	39.40
Longnose skate	1.45	0.38	127	31.20
Big skate	1.66	0.24	89	-0.59

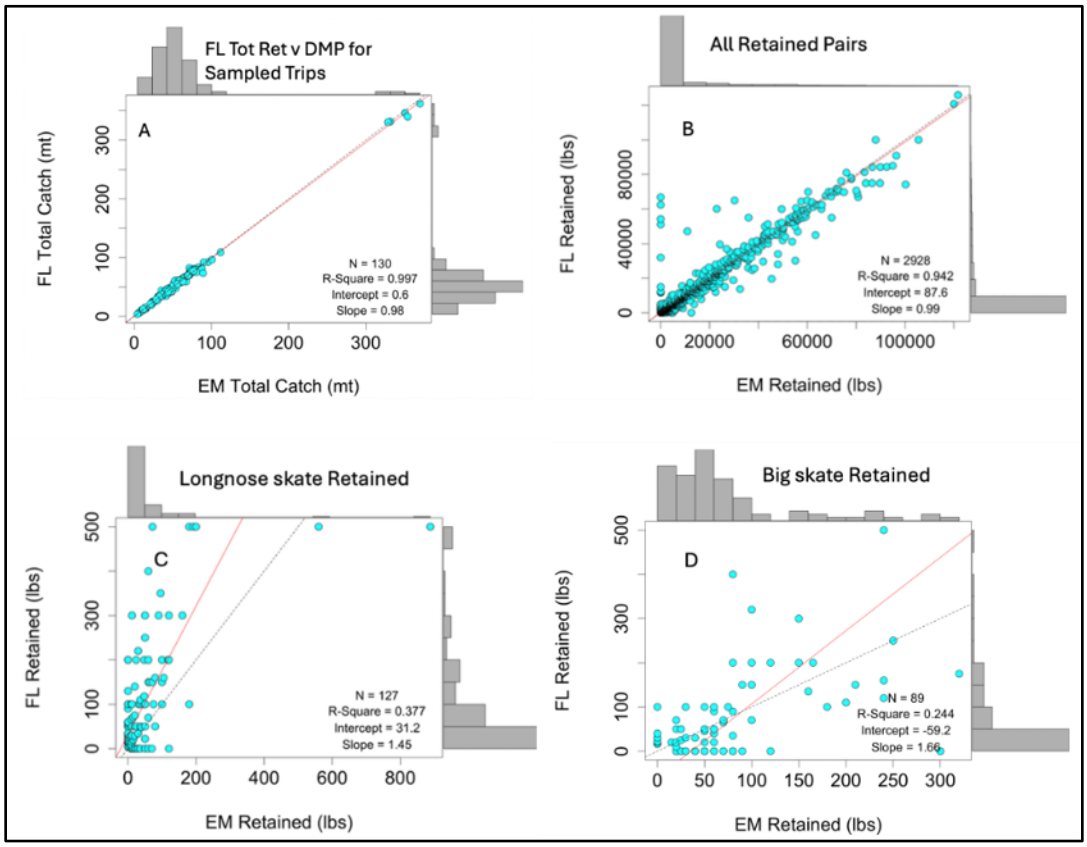


Figure 11. Retained Catch Comparisons, Showing All Species Combined (i.e., Trip Level DMP vs FL, 2A), All Retained Comparison Pairs (i.e., Set Level FL vs EM, 2B), Then for Longnose Skate (2C) and Big Skate (2D).

Figure 11A compares the estimates skippers recorded as being delivered to dockside versus what the DMP recorded for all species (or species group) and tows combined. The slope of 0.98 and R^2 of 1.0 indicate close alignment. The five large pairs influence the R^2, but the fit of pairs less than 300,000 lb is similar. The larger volumes (70,000-90,000 lbs) represent three trips by an RT vessel.

Since DMP provides what is treated as a verified estimate of retained catch by species, there is less concern about the alignment of FL and EM, only that they are aligned well enough that DMP total catch can be allocated to quota area proportionately by using FL catch by tow. This appears to be the case. Similarly, Figure 11B indicates a slope close to 1.0 and sufficient precision across all paired observations.

Most species catch estimates appeared to be reasonably aligned with respect to numerical bias and precision of FL and EM retained estimates (Table 12 and Figure 11C-D). These include:

- Arrowtooth flounder (slope=1.0, R^2=0.97);
- Lingcod (slope=0.93, R^2=0.96);
- Hake (slope=0.96, R^2=0.93);
- Sablefish (slope=0.96, R^2=0.94);

- Other flatfish (slope=0.93, R^2=0.90);
- Rockfish (slope=1.0, R^2=0.98);
- Pollock (slope=0.98, R^2=0.88).
  Pacific cod estimates of FL and EM retained indicate a bias of FL overestimating by approximately 30% (Slope=1.3). Precision is poor with an R^2 of 0.85.
- Big and longnose skate reflect poor alignment, with FL overestimating longnose skate (45%) and poor precision (R^2=0.38) and big skate (66%) with low precision (R^2=0.0.24 (Table 12, Figures 10C and 10D). There are numerous unmatched pairs with EM or FL estimates of 0.

As noted above, FL estimates are most important for allocating catch to tows and, therefore, the location of catch. The relative size of the overall discrepancy between retained FL and EM and total EM is provided in Table 13.

Table 13. Relative Difference for Total EM Ret and Total FL Ret (lbs.) by Species by Fishery Year. Omitted halibut, and dogfish) from yearly averages owing to influence of low values (Ret EM estimates of <5,000 lbs (FY 2025/2026) and changing recording practices.

	2022/2023					2023/2024				
Species	**Ret EM**	**Ret FL**	**Ret Delta (EM-FL)**	**Ret % Difference**	**Abs Value**	**Ret EM**	**Ret FL**	**Ret Delta (EM-FL)**	**Ret % Difference**	**Abs Value**
Spiny dogfish	12,721	6,252	6,469	51%	51%	8,722	11,237	-2,515	-29%	29%
Big skate	35,854	29,346	6,508	18%	18%	38,526	31,316	7,210	19%	19%
Longnose skate	7,623	12,275	-4,652	-61%	61%	6,513	9,740	-3,227	-50%	50%
Salmon	16,406	5,196	11,210	68%	68%	12,638	6,139	6,499	51%	51%
Pacific cod	230,993	260,145	-29,152	-13%	13%	88,650	110,721	-22,071	-25%	25%
Pacific hake	16,603,183	18,520,366	-1,917,183	-12%	12%	12,130,513	13,924,938	-1,794,425	-15%	15%
Walleye pollock	5,894,337	5,721,022	173,315	3%	3%	3,571,453	3,364,025	207,428	6%	6%
Rockfish	7,318,586	7,323,572	-4,986	0%	0%	8,550,417	8,268,488	281,929	3%	3%
Sablefish	57,404	62,274	-4,870	-8%	8%	70,953	82,054	-11,101	-16%	16%
Lingcod	163,691	199,200	-35,509	-22%	22%	142,213	206,122	-63,909	-45%	45%
Flatfish	927,128	890,501	36,627	4%	4%	824,886	732,674	92,212	11%	11%
Turbot	2,016,493	2,123,622	-107,129	-5%	5%	1,917,408	2,175,807	-258,399	-13%	13%
Pacific halibut	952	48	904	95%	95%	1,444	1,472	-28	-2%	2%
Average					**15%**					**20%**
	2024/2025					**2025/2026 (partial)**				
Species	**Ret EM**	**Ret FL**	**Ret Delta (EM-FL)**	**Ret % Difference**	**Abs Value**	**Ret EM**	**Ret FL**	**Ret Delta (EM-FL)**	**Ret % Difference**	**Abs Value**
Spiny dogfish	5,328	9,958	-4,630	-87%	87%	2,797	669	2,128	76%	76%
Big skate	16,719	15,867	852	5%	5%	6,815	6,038	777	11%	11%
Longnose skate	7,954	12,763	-4,809	-60%	60%	5,849	12,215	-6,366	-109%	109%
Salmon	2,768	4,405	-1,637	-59%	59%	2,855	3,600	-745	-26%	26%
Pacific cod	51,189	81,009	-29,820	-58%	58%	89,992	127,464	-37,472	-42%	42%
Pacific hake	2,996,963	3,051,927	-54,964	-2%	2%	2,364,857	2,306,542	58,315	2%	2%
Walleye pollock	5,686,486	5,457,032	229,454	4%	4%	3,912,600	3,747,998	164,602	4%	4%
Rockfish	7,463,682	7,169,581	294,101	4%	4%	6,691,194	6,588,690	102,504	2%	2%
Sablefish	89,814	88,692	1,122	1%	1%	87,533	95,977	-8,444	-10%	10%
Lingcod	219,214	261,653	-42,439	-19%	19%	165,326	174,854	-9,528	-6%	6%
Flatfish	735,355	712,069	23,286	3%	3%	563,366	569,965	-6,599	-1%	1%
Turbot	1,834,573	1,884,292	-49,719	-3%	3%	455,938	476,054	-20,116	-4%	4%
Pacific halibut	252	2,080	-1,828	-725%	725%	-	572	-572	-	-
Average					**16%**					**19%**

Trends in Catch Alignment (FY2020–FY2025)

Tables 10 and 13 indicate the relative difference of the discrepancy between the sum of EM minus the sum of FL, divided by total EM catches, total retained or released by year from FY2022 – FY2025. This ratio shows the difference between EM and Rel totals relative to total catch over time.

Estimation of Non-Commercial Releases and Species Richness

The frequencies of all observations for all species categories (N=142) are summarized in Table 14 for EM from the AMR database for FY2025. The species categories are as they were recorded ("nominal") with spelling and nominal species names intact as recorded. No attempts were made in this review particularly below to remove or alter unlikely or dubious nominal species observations.

The EM and FL estimates of coral and sponge released catch were not analyzed, as the data set contained only three records with a total estimated catch of less than 10 lbs.

Table 14. List of Nominal Species Categories Recorded by EM Review for FY2025 (partial).

Species	N	Weight (lbs)	Species	N	Weight (lbs)	Species	N	Weight (lbs)
Turbot	451	525,451	Rosethorn Rockfish	25	19	Cat Sharks	2	10
Lingcod	399	183,886	Splitnose Rockfish	24	2,705	Chionocetes Tanneri	2	-
Bocaccio	390	270,564	Chum Salmon	23	44	Curlfin Sole	2	-
Yellowtail Rockfish	385	653,636	Octopus	23	369	Generic Rockfish	2	4
Spiny Dogfish	383	147,089	Jack Mackeral	21	436	Giant Wrymouth	2	15
Ratfish	336	48,210	Pop / Yellowmouth	21	77,512	Salmon Shark	2	250
Canary Rockfish	320	543,068	Red Rockfish	21	3,710	Sockeye Salmon	2	-
Pacific Halibut	317	69,052	Skates	17	1,081	Soft Corals	2	2
Petrale Sole	315	194,196	Lampreys	11	1	Aurora Rockfish	1	1
Long Nose Skate	297	15,274	Sculpins	11	39	Barracudinas	1	-
Pacific Cod	297	91,665	Yellowmouth / Canary	11	58,967	Butter Sole	1	-
Sablefish	296	124,048	Shortbelly Rockfish	10	62	California Sea Lion	1	200
Silvergray Rockfish	293	324,765	Yellowfin Sole	10	-	Capelin	1	5
Pollock	279	3,995,644	Brown Cat Shark	9	19	Cephalopods	1	-
Yellowtail / Widow	254	2,250,432	Inanimate Objects	9	634	Clingfishes	1	-
Rex Sole	244	30,374	Morning Sun Starfish	9	-	Copper Rockfish	1	8
Dover Sole	233	206,062	Rougheye / Shortraker	9	2,873	Eelpout	1	1
Widow Rockfish	230	622,014	Yelloweye / Canary	9	194	Gastropoda	1	-
English Sole	207	92,426	Clupeidae	8	-	Giant Squid	1	-
Pacific Ocean Perch	196	992,319	Coho Salmon	8	5	Glass Shrimp	1	-
Hake	188	2,475,450	Pink Salmon	8	-	Grenadiers	1	5
Redstripe Rockfish	178	83,720	Giant Pacific Octopus	7	41	Herring Gull	1	3
Chinook Salmon	156	495	Unid. Org. Matter	7	4,445	Lanternfishes	1	-
Red Banded Rockfish	145	24,306	Deep Sea Skate	6	58	Longfin Dragonfish	1	-
Salmon (General)	108	2,465	Greenspotted Rockfish	6	-	Mola Mola	1	-
Shortspine Thorn.	106	3,601	Prawn	6	2	Oncorhynchus Spp.	1	-
Flatfish General	102	77,260	Crabs	5	50	Pacific Staghorn Sculpin	1	-
Big Skate	97	7,556	Quillback Rockfish	5	23	Pacific Viperfish	1	-
Yellowmouth	96	319,838	Alaska Skate	4	5	Pallid Slipskin	1	-
Boccacio / Silvergray	90	332,588	Cabezon	4	6	Purple Sea Urchin	1	-
Sharpchin Rockfish	83	13,263	Round Fish	4	36	Red Sea Urchin	1	100
Pacific Herring	75	26,962	Wrymouths	4	28	Shark Species	1	-
Sandpaper Skate	69	1,651	Broad Skate	3	62	Shrimp	1	54
Yelloweye Rockfish	69	688	Brown Rockfish	3	5,136	Skipjack Tuna	1	9
Greenstripe Rockfish	68	1,802	Chilipepper Rockfish	3	10	Sleeper Shark	1	200
Squids	58	7,544	Harlequin Rockfish	3	5	Smelts	1	1
Rougheye Rockfish	54	103,882	Humbolt Squid	3	-	Snipe Eels	1	-
Dungeness Crab	51	2,495	Longspine Thorn.	3	-	Sponges	1	3
Flathead Sole	51	2,090	Pacific Lamprey	3	-	Thornyheads	1	-
Darkblotch Rockfish	48	9,094	Prowfish	3	10	Tiger Rockfish	1	3
Starfish	43	1,381	Ragfish	3	151	Wolf Eel	1	10
American Shad	42	260	Sand Sole	3	50	Yellmouth / Rough	1	1,511
Eulachon	39	399	Sea Cucumbers	3	31			
Scyphozoa (Jellyfish)	39	64	Unknown Fish	3	31			
Shortraker Rockfish	33	1,739	Aleutian Skate	2	-			
Starry Flounder	32	3,894	Anemones	2	25			
Pacific Sanddab	28	3	Atka Mackerel	2	204			
Rock Sole	28	3,860	Bigmouth Sculpin	2	-			
Slender Sole	28	10	Black Rockfish	2	-			
Red Squid	25	-	Blackfin Sculpin	2	-			

A reasonable concern about the current TMP is whether the transition from ASOP to EM-based TMP resulted in recording fewer minor (non-quota) and, particularly, small species. The number of recorded species during the two periods FY2015-FY2020 and FY2021-FY2025 are summarised below (Figure 12 and Tables 15 and 16).

Note that there is a separate "marine mammal interactions" reporting tool within the DFO databases, which was not used in this review and may offer better information than the TMP FL and EM data. The audit report is intended to capture these rarer interactions (e.g., marine mammals, sharks other than dogfish) through the species-of-interest report, but it isn't part of the Scorecard.

Table 15. List of the Number of Species and Species Categories Recorded Observers or EM by Fishery Year.

Source	FY	Species	Sp_Groups	Total
ASO	2015/2016	144	30	174
ASO	2016/2017	143	31	174
ASO	2017/2018	146	28	174
ASO	2018/2019	131	31	162
ASO	2019/2020	129	32	161
EM	2020/2021	106	20	126
EM	2021/2022	115	25	140
EM	2022/2023	123	23	146
EM	2023/2024	120	26	146
EM	2024/2025	116	23	139
EM	2015/2026	122	16	138

As noted above, the review used the “nominal” observations in the overall databases without any attempt to delete questionable identifications. This treatment indicates an approximate 18% decline in the recorded species categories coincident with the transition to the EM period. The species resolution appears stable within the two distinct periods. As might be expected, the catches are relatively more concentrated among dominant species (or species group), such as hake, as shown in the cumulative frequency (Figure 12).

It should be noted that, in this treatment, we made little or no attempt to examine factors that may influence these observations. For example, the total catch and the number of tows declined from the ASO to the EM periods. Furthermore, the proportion of catch contributed by the RT vessels declined, but there is some basis for considering that these vessels could provide more species resolution by EM through the presence of more cameras. It was also reported in interviews that at least one RT vessel commonly dedicates a crew member to catch monitoring. It is also known that species identification capability varied widely among observers and over the course of the career of an individual observer (Grinnell, 2010). Other influences that could differ between periods include changes in codend mesh size and the area trawled, which would entrain changes in relative proportions of catch contributed by different habitats, depths, and seasonal effects. These changes also included the cessation of thornyhead-directed fishing in 2019, where many small deep water incidental species of fishes and invertebrates are encountered.

A more complex investigation would have to filter or attempt to correct across these many influential impacts on observed species composition. Notwithstanding these caveats, the treatment indicates that the number of species categories identified decreased during the transition but has appeared stable since.

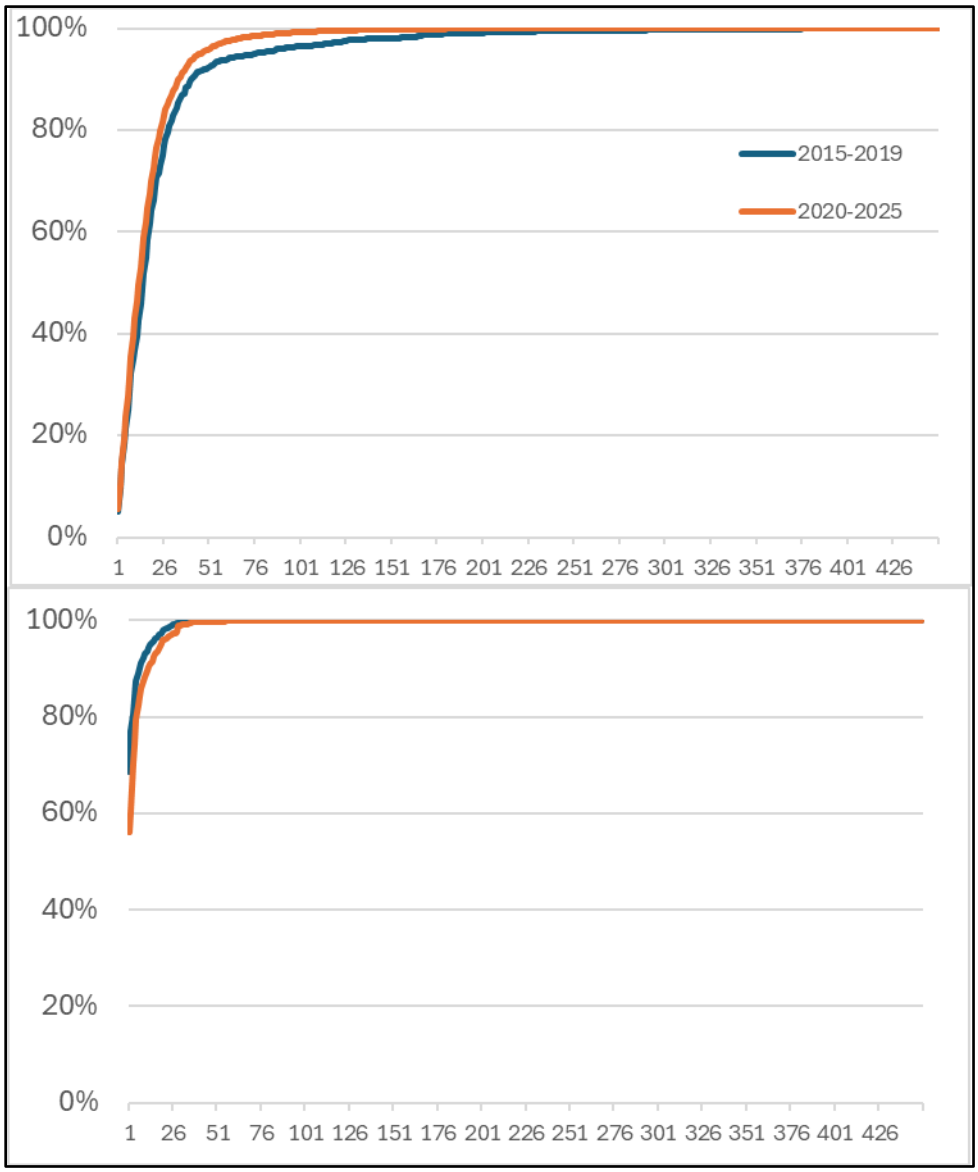


Figure 12. Cumulative Frequency of Catch Species for 2019-2020 (blue) Compared with 2021-2025 (red), Showing the Incidence of Occurrence by Species Reporting Events (upper graph) and by Species Weight (lower graph).

Species richness was also examined by examining which species (or species group) were present in only one of the periods (Table 16). This required a slightly more critical review of the available lists. In this case, a species (or species group) was omitted if one species of that group was present (i.e., *Oreo* sp. and *Oxeye oreo*, Rougheye/Blackspotted rockfish and Blackspotted rockfish). In total, 58 different categories were recorded during the ASOP period,

but not in the EM period. Conversely, 18 categories were recorded in the EM period but not in the preceding ASOP period.

Table 16. List of Species Categories Reported in either but not Both ASO and EM Periods (data from DFO FOS database).

ASO Only			EM Only
Black-Footed Albatross	Mosshead Sculpin	Slender Snipe Eel	Bearded Rattail
Blacktip Poacher	North Pacific Frostfish	Smooth Abyssal Grenadier	Black Bristlemouth
Blue Lanternfish	Northern Lampfish	Smooth Alligatorfish	Broad Skate
Buffalo Sculpin	Northern Rockfish	Snake Prickleback	Calico Surfperch
Closespine Snipe Eel	Northern Ronquil	Snakehead Eelpout	Capelin
Crested Bigscale	Northern Spearnose Poacher	Soft Eelpout	Clingfishes
Daggertooths	Opah	Soft Eelpout	Cow Sharks
Daubed Shanny	Oxeye Oreo	Sooty Shearwater	Crested Sculpin
Deacon Rockfish	Pacific Barracuda	Spiny Tapirfishes	Double-Crested Cormorant
Dwarf Wrymouth	Pacific Rainbow Smelt	Spotfin Sculpin	Herring Gull
Electric Rays	Pacific Spiny Lumpsucker	Sticklebacks	Pacific Softpout
Emarginate Snailfish	Rainbow Trout (aka Steelhead)	Thornback Sculpin	Pallid Eelpout
Giant Blobsculpin	Ribbed Sculpin	Threadfin Slickhead	Pygmy Poacher
Gray Whale	Ronquils	Thresher Sharks	River Lamprey
Great Sculpin	Sailfin Sculpin	Whales & Porpoises	Scaleless Black Dragonfishes
Grunt Sculpin	Scorpionfishes	White Sturgeon	Slipskin Snailfish
Lobefin Snailfish	Shortfin Eelpout	Whitebarred Prickleback	Surf Smelt
Longfin Sculpin	Shortfin Mako	Whitespotted Greenling	Tubeshoulders
Lumpfishes and Snailfishes	Silverspotted Sculpin		
Mackerel Sharks	Skilfish		

Readers are reminded as above, that the list of species in Table 16 should be considered as "nominal" and viewed with considerable caution. Some of the identifications appear unlikely (i.e., white sturgeon), and some could be incorrect because of grouping issues (i.e., cowsharks, lumpfishes, snailfishes, electric rays/Pacific electric ray, and seabirds).

This review did not attempt to explore in detail the TMP precision and numerical bias for non-commercial and rarely caught categories. The alignment approach used above would be informative for more commonly encountered species or species groups, such as ratfish, but the information content of such an approach attenuates quickly as the data become sparser.

At the other extreme, we briefly examined coral and sponge records as an example. In FY2025, the TMP EM-sampled dataset and the corresponding FL entries for the same tows (>10% coverage for Wetfish vessels and >25% for RT vessels) produced three observations from three different tows:

1. EM – Soft coral – 2 lbs (RT vessel)
2. FL – Soft coral – 1 lb (RT vessel)
3. EM – Sponges – 3 lbs (Wetfish vessels)

However, while the values may reasonably reflect the scale of the unintended catch, they should be interpreted as **minimum** estimates of the true magnitude. The inverse of the review

rates can expand the number of observations to prepare a nominal projection of total catch, but with so few observations, the resulting estimates lack credibility. It is noted elsewhere that increasing the low review rates may provide more data for less frequently encountered species, but this positive effect will attenuate rapidly for the bottom end of this frequency table where current frequencies are less than 10. Future analyses my develop meaningful inferences by combining years, as in Table 15.

Catch Scoring Analysis

As mentioned, scoring methods for assessing FL data quality are integral to the operation of the TMP, irrespective of the overall data adequacy achieved in the program. There needs to be a simple way to rapidly assess FL data quality, identify problem areas, direct appropriate feedback, and determine the utility of FL from the fishing trip. The purpose of this analysis is to determine how well these needs are satisfied by the scoring methods in place within the TMP.

The scoring method is an elegant way of bringing together the large number of FL-EM (or DMP) comparison pairs across five test areas to arrive at an audit outcome determination. Ultimately, these individual comparisons involve assessing the size of the delta relative to the reference value to arrive at a score. Before examining the scoring methods, it is useful to gain some perspective on the size and scale of the deltas within the data set. The FY2025 data set was summarized in two ways to show the magnitude of the deltas with FL-EM released catch estimates. The first case shows the FL-EM deltas for total released catch for reviewed fishing events (Table 17). The deltas are tabulated by size intervals and by audit outcome category, expressed as percentages to enable comparison between audit outcomes. The percentages in each column reflect the proportion of reviewed fishing events (bottom row) within each delta size category (rows). In the case of Pass, 49% of the reviewed fishing events reported zero releases, 2% where releases occurred but the delta was zero (exact match between FL and EM), and the remainder reported with non-zero deltas. Categories 1 and 2 were more similar to each other than either was to Pass, with both having about half of the reviewed fishing events showing deltas of more than 100 lbs. Importantly, fishing events with zero release were about five times greater with Pass than those of Categories 1 and 2. Categories 3 and 4 are included for completeness purposes, but they contribute little to this discussion.

Table 17. EM-FL Delta for Released Catch by Weight Interval and Audit Outcome (FL used when EM-FL+).

Delta (lbs)	Pass	1	2	3	4	Total
Zero Release	49%	10%	12%	0%	9%	15%
Exact Match	2%	1%	0%	0%	0%	1%
1-20	31%	18%	7%	0%	14%	18%
21-50	8%	14%	10%	0%	5%	12%
51-100	4%	11%	11%	0%	9%	10%
101-500	4%	30%	34%	100%	45%	28%
>500	1%	16%	27%	0%	18%	16%
Total	**90**	**498**	**92**	**1**	**22**	**703**

The second method for summarizing deltas was at the comparison pair level (Table 18). Shown are the frequency of released catch FL-EM deltas observed by species (or species group) taken from reviewed tows in the FY2025 data set (the same data set used for Table 17). The percentage values show the distribution of deltas by species. For example, with sablefish, 9% of the 252 comparison pairs were exact matches and 28% had deltas of less than 20 lbs. Overall, there were approximately 2,500 comparison pairs, with deltas ranging in size from zero (exact match, 9%) to over 500 lbs (4%). About half the deltas were 20 lbs or less. Higher deltas occur with pelagic species (hake, pollock), albeit with larger volumes, but the distribution was generally consistent across other species. Interestingly, 33% of halibut pairs were exact matches, showing that applying standard estimations (counts and standard piece weights) contributes to higher FL to EM alignment.

The lower portion of the table shows the frequency of deltas by audit outcome. As these results are based on the number of comparison pairs, as opposed to reviewed fishing events (Table 17), audit outcome patterns differ because the number of tests per reviewed fishing event differ. Pass outcomes were based on 104 tests from 90 reviewed events, while Category 2+3 included 305 tests from 93 reviewed events, differing by a factor of three. On a comparison-test basis, all audit categories were similar with well over half the tests having deltas of less than 20 lbs. In this pool of about 2,500 comparison tests, about a quarter had deltas of over 50 lbs, and 4% were over 500 lbs. The size and frequency of deltas are just part of the story influencing audit outcomes, as will be shown.

Table 18. Distribution of EM-FL Deltas by Species Group and Audit Outcome (FL Used if EM=0).

	Delta (lbs)						Total
	Zero	1-20	21-50	51-100	101-500	>500	Tests
By Species Group							
Arrowtooth Flounder	3%	43%	19%	12%	17%	5%	339
Big Skate	0%	75%	6%	13%	6%	0%	16
Hake	1%	34%	17%	9%	17%	22%	145
Halibut	33%	28%	19%	8%	11%	2%	317
Lingcod	7%	58%	16%	10%	10%	0%	260
Longnose Skate	8%	46%	29%	9%	8%	0%	198
OtherFlatfish	4%	46%	14%	10%	18%	7%	201
Pacific Cod	7%	56%	25%	9%	4%	0%	57
Rockfish	0%	68%	11%	4%	11%	7%	56
Sablefish	9%	61%	14%	6%	7%	2%	252
Spiny Dogfish	8%	43%	20%	12%	12%	5%	377
Walleye Pollock	2%	45%	16%	11%	17%	10%	163
All Species	9%	46%	18%	10%	12%	5%	2,381
By Audit Outcome							
Pass	12%	45%	19%	9%	12%	4%	104
Cat 1	6%	50%	18%	10%	12%	4%	1,923
Cat 2 + Cat 3	24%	43%	21%	12%	1%	0%	305
Cat 4	12%	45%	19%	9%	12%	4%	104
All Audit Categories	9%	48%	19%	10%	10%	4%	2,436

How the Number of Tests Affects Scores

Averaging individual comparison test scores replaces the diversity of source scores with a single value. The greater the number of tests, the greater the tolerance for variability among the source comparison tests. This simple concept is shown more clearly in Table 19, which presents different numbers of tests and score values. In the first instance, there are four examples with five tests, each with different individual score results but very similar average score values. The different examples show different score patterns, in particular, the number of high scores required to offset the low scores. Examples 5(3) and 5(4) show that scores of 0 or 3 can emerge with three scores of 10. When there are only 3 tests, the lowest possible score is 3; with 10 tests, there are 2 zero scores; and with 20 tests, there are 5 zero scores, but these must be offset by several scores of 10 to achieve an average score over 7.

Table 19. Distribution of EM-FL Deltas by Species Group and Audit Outcome (FL Used if EM=0).

Number	Test Number and Score Result																			Avg
of Tests	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Score
5 (1)	3	7	7	9	10																7.20
5 (2)	7	7	7	8	7																7.20
5 (3)	0	7	10	10	10																7.40
5 (4)	3	5	10	10	10																7.60
3	3	10	10																		7.67
10	0	0	5	10	10	10	10	10	10	10											7.50
20	0	0	0	0	0	7	10	10	10	10	10	10	10	10	10	10	10	10	10	10	7.35

The four examples of the five tests in Table 19 can also be used to illustrate test area averaging for a trip score. For example, 5(1) could reflect a score of 3 for T1, scores of 7 for T2 and T3, and scores of 9 and 10 for T4 and T5, respectively. All four examples have trip scores greater than 7 and would have a Category 1 designation, illustrating that scores as low as zero can occur at both the individual test and the test area levels, but multiple scores of 10 are needed for counterbalance to achieve a Category 1 outcome. By comparison, an audit outcome of Pass would require that all five cells in the 5 test examples be 8 or greater.

Shifting from the theoretical to actual fishery results, Table 20 provides a summary of the average number of tests performed within each of the five test areas, taken from the Audit Scorecard sample. The sampled trips are subdivided firstly by fleet type (Wetfish and RT vessels), then by trip target species, and lastly by audit outcome for Wetfish boats (the sample size for RT vessels was too small for this treatment). While the number of sampled trips varies by subdivision, a few patterns emerge. Firstly, trips by the RT fleet have many more tests per test area than trips by Wetfish boats, owing mostly to the higher sample rate (>25%) of the RT fleet. Among Wetfish fleet trips, those targeting flatfish (i.e., trips with flatfish as the principal landed species) average more tests than those trips targeting rockfish, followed by pollock. Wetfish fleet trips with a Pass averaged less than half the number of tests than either Category 1 or the Category 2+3 group. With a higher average number of tests moving from Pass to Category 1 to Category 2+3, the trend indicates more complex catch compositions and likely greater difficulty in estimating released catch.

Table 20. Average Number of Comparison Tests by Test Group and Trip Characteristic.

	Trips	T1-DMP	T2-Ret	T3-Rel	T4-Hal	T5-Sab	Total
Fleet Type							
RT	13	5.2	8.3	11.3	1.8	2.1	25.1
Wetfish	117	7.1	3.6	3.3	1.1	1.3	13.8
Target Species (Wetfish Fleet)							
Flatfish	18	9.0	5.4	3.2	1.2	1.2	19.3
Pollock	9	5.9	2.4	1.0			9.0
Rockfish	82	6.8	3.3	3.7	1.1	1.3	13.0
Audit Result (Wetfish Fleet)							
Pass	25	4.6	1.6	1.5	1.0		6.5
Cat 1	64	7.6	4.0	3.2	1.1	1.2	15.1
Cat2+3	28	8.4	4.4	3.7	1.2	1.5	17.3

An interesting feature of the results in Table 20 is the variation in the average number of tests across the five test areas. While the trip score calculation applies equal weighting to the five test areas, differences in the number of test areas result in differences in the relative weighting of a single comparison test. For example, T4 and T5 averaged slightly over one test per trip, whereas the average number of T1 tests was over five times greater, each T1 comparison test counting for less than a fifth of a T4 or T5 test. Similarly, a single T2 or T3 comparison test would count for about one-third of a halibut or sablefish test, but nearly twice that of a DMP test.

Paired and Null Comparison Tests

As mentioned above, for FL-EM catch estimation comparisons there are two types of test pairs, paired comparisons (FL+EM+) and null comparisons (EM+FL- or EM-FL+). These two types of comparison pairs are to be expected in the data set because a catch species will not always be seen in both sources. While null comparison pairs are statistically valid for the purposes of determining FL-EM alignment, they present problems for scoring.

Before considering scoring effects with paired and null comparison test pairs, it is helpful to examine their prevalence. Table 21 summarizes test results from the sample of Audit Scorecards, showing the number of comparison tests by test area and audit outcome. The results confirm the desired result that most comparison tests are based on matched pairs of observations, regardless of test area or audit outcome, with T5 sablefish tests being the lone exception. The highest proportion of matched pair tests across all audit outcomes was for T4 halibut. Null tests make up a very large component of comparison tests, with many test area-audit outcome cells showing proportions of 20-40%. T5 sablefish is particularly high when null tests account for about 60% of all comparison tests. The high proportion of null tests with sablefish is partly due to the way legal and sublegal size categories are scored, as will be explained below. Also showing a high null test proportion were T1-DMP comparison tests with

Pass audit outcomes (43%). The number of null tests based on catch estimates provided by EM (or DMP) outnumbers FL only by a large margin (i.e., FL-EM+ or FL-DMP+). Proportions of null tests versus paired tests among Category 1 and Category 2+3 audit outcomes appear more similar than Pass was to Category 1. It is important to note that the results in Table 21 are not adjusted for the entire fishery because they are based on samples from Audit Scorecards.

Table 21. Paired and Null Comparison Tests by Test Area and Audit Outcome.

Audit Outcome and Test Type	T1 - DMP	T2 - Ret	T3 - Rel	T4 - Hal	T5 - Sab	Total
Pass (28 trips sampled)						
Total Tests	132	54	14	3	0	203
Paired Tests (#)	75	42	9	3		129
Paired Tests (%)	57%	78%	64%	100%		64%
Null Tests (FL+)	2	4	0	0		6
Null Tests (EM/DMP+)	55	8	5	0		68
Total Null Tests (%)	43%	22%	36%	0%		36%
Cat 1 (72 trips sampled)						
Total Tests	527	328	238	46	47	1186
Paired Tests (#)	391	248	150	40	19	848
Paired Tests (%)	74%	76%	63%	87%	40%	72%
Null Tests (FL+)	9	46	34	4	14	107
Null Tests (EM/DMP+)	127	34	54	2	15	232
Total Null Tests (%)	26%	24%	37%	13%	62%	29%
Cat 2 and Cat 3 (30 trips sampled)						
Total Tests	244	147	120	25	16	552
Paired Tests (#)	188	106	70	21	7	392
Paired Tests (%)	77%	72%	58%	84%	44%	71%
Null Tests (FL+)	4	19	18	1	4	46
Null Tests (EM/DMP+)	52	22	31	3	5	113
Total Null Tests (%)	23%	28%	41%	16%	56%	29%

A more comprehensive examination of paired and null tests for released catch was made from the FY2025 data set (Table 22). Again, the proportion of paired and null comparison tests were summarized by audit outcome category, and the estimated catch quantities associated with these test types were also included. As before, there is a consistent pattern of paired comparison tests being more dominant than null tests, accounting for nearly 60% of all tests. As well, among null tests, EM more often reported catch items not reported in the FL record (26% vs 16% overall). The right side of Table 22 shows catch quantities associated with the comparison test types. Even though paired tests make up about 60% of all comparison tests, the catch quantity associated with these tests makes up 90% or more of the total released catch. Interestingly, among null tests, EM and FL contributed equally to the total catch, even though EM contributed 50% more null tests than FL. These results generally align with those

from the Audit Scorecards (Table 21) and confirm that null tests are mostly associated with small catch quantities.

Table 22. Species Comparison Test Symmetry in Relation to # Tests and Released Quantity.

Audit	Number of Tests				Released Quantity (lbs)			
Outcome	EM+ Fl+	EM+ FL-	EM- FL+	Total	EM+ Fl+	EM+ FL-	EM- FL+	Total
Pass	58%	35%	7%	138	97%	2%	1%	26,316
Cat 1	58%	25%	16%	2,352	88%	7%	5%	461,787
Cat 2 + Cat 3	56%	29%	15%	491	91%	4%	4%	169,748
Cat 4	59%	22%	19%	130	90%	8%	2%	42,023
n/a	6%	9%	84%	32	3%	2%	95%	12,780
Total	57%	26%	16%	3,143	88%	6%	6%	712,654

Using the same data set as Table 22, the relative contributions of null and paired comparison tests are examined for released catch by species category (Table 23). Note that the results include species groups included in the Audit as well as those not. As these data are based on the same data set as Table 22, the overall patterns are the same, namely the proportion of paired tests is high (57%), and these make up a high percentage (88%) of total released catch, although among tested species, these numbers are slightly higher (64% and 91%). At the individual species level, these patterns vary widely, with some species' results (both tests and catch weight) originating mostly from paired tests (e.g., halibut, sablefish). In contrast, null tests are a larger contribution to species results for others (e.g., Pacific cod, rockfish, and all skate categories). Rockfish stands out among all species categories, with a low proportion of matched pairs (3%), mostly reported by EM (63%), yet the FL catch estimate accounted for 94% of the total quantity reported. Species with a sizable proportion of null tests and catch is of concern because of the lower certainty associated with these tests, as well as the fact that EM estimates are used for FL data quality assessment and do not form part of the final catch record.

Table 23. Species Comparison Test Symmetry by Species Group.

Species	Number of Tests				Released Quantity (lbs)			
Category	EM+ FL+	EM+ FL-	EM- FL+	Total	EM+ FL+	EM+ FL-	EM- FL+	Total
Tested								
Turbot	63%	22%	15%	339	85%	6%	9%	76,709
Big Skate	13%	44%	44%	16	13%	66%	22%	948
Longnose Skate	52%	30%	18%	198	73%	16%	11%	10,554
OtherFlatfish	46%	34%	19%	201	82%	9%	9%	43,697
Hake	66%	27%	8%	145	98%	2%	0%	110,806
Pacific Cod	32%	37%	32%	57	56%	14%	30%	2,407
Walleye Pollock	51%	38%	11%	163	92%	3%	4%	86,956
Lingcod	73%	17%	10%	260	93%	5%	3%	17,654
Sablefish	69%	12%	19%	252	97%	1%	3%	37,493
Halibut	87%	6%	6%	317	97%	1%	1%	70,063
Spiny Dogfish	63%	21%	16%	377	91%	4%	5%	151,641
Subtotal	64%	22%	14%	2,325	91%	4%	5%	608,928
Not Tested								
Salmon	3%	28%	69%	29	5%	22%	73%	387
Rockfish	4%	63%	34%	56	1%	5%	94%	9,822
OtherFish	52%	32%	16%	449	81%	16%	4%	77,524
OtherSharks	31%	15%	54%	13	67%	2%	31%	695
OtherSkates	24%	62%	14%	87	29%	66%	4%	2,959
Invertebrates	35%	34%	31%	168	60%	9%	32%	7,067
Other	0%	100%	0%	16	0%	100%	0%	5,272
Subtotal	39%	38%	22%	818	66%	20%	14%	103,726
Total	57%	26%	16%	3,143	88%	6%	6%	712,654

Comparison Test Scoring

As shown in Tables 22 and 23, the contribution of null tests to overall estimates of released catch is small, both overall and for most species. Yet their effects on scoring results should be examined, as they comprise about 40% of the total comparison tests. Null and paired tests differ in ways that affect scoring, as will be shown. When null and paired tests are combined into a single test area average or trip score, the relative contribution of each test is assumed to be the same despite their underlying differences.

To gain a better understanding of score patterns associated with paired and null comparison tests, released catch (T3 to T5) data from the 2025/2026 fishery data set were analyzed to examine comparison test type (i.e., FL+EM-, FL-EM+, and FL+EM+), in relation to test scores. Each comparison pair in the data set was scored according to the species group tolerances (Table 2), with the score results clustered into groups of high (score 9-10), medium (score 5-8), and low (score 0-3). Note that there is no score of 4 (see Table 2). Paired comparison tests

based on weight difference were tabulated separately from those based on percent difference (see Table 3) to examine scoring differences. All null tests were reported as weight-based because the score is always zero when the reference weight exceeds the weight threshold (i.e., the percent difference between zero and the estimate is 100%). The methods used for this analysis followed the same procedures as those used in the Audi scoring, except that sablefish and lingcod were not distinguished (these data were not available).

Table 24 provides a summary of this analysis, including comparison test results for all species combined. The table shows the total number of tests, followed by percentages calculated vertically and horizontally. Among all 2,905 comparison tests shown, most (82%) were high-scoring, while low-scoring tests accounted for about 7% of the total. The pattern of high score dominance was similar between the null and paired tests, but the proportion of null tests was greater (89% versus 79%). Test score distribution was the same for EM+FL- as EM-FL+ based tests, but the contribution by EM+FL- was greater (i.e., 17% versus 11%). Among the paired tests, there were about eight times as many weight-based comparison tests as percent-based tests, with the scoring distributions opposite one another. In other words, weight-based paired tests were far more common, and scores were often high, whereas percent-based paired tests were much less common, and nearly half the scores were low. Based on these results, it appears that scores from percent-based comparison tests are much more likely to be low than comparison tests scored from weight-based tolerances. The results also indicate, unfortunately, that null tests contribute significantly to overall test scores, leading to higher scores and skewing overall scores upward.

Table 24. Summary of Test Score Patterns and Test Method.

Score	Null Tests (all wt)			Paired Tests			Grand
Category	EM+FL-	EM-FL+	Total	%	wt	Total	Total
By Numbers							
High (9-10)	448	276	724	39	1,625	1,664	**2,388**
Mid (5-8)	33	20	53	84	171	255	**308**
Low (0-3)	22	11	33	105	71	176	**209**
Total	**503**	**307**	**810**	**228**	**1,867**	**2,095**	**2,905**
Column Percentages							
High (9-10)	89%	90%	89%	17%	87%	79%	**82%**
Mid (5-8)	7%	7%	7%	37%	9%	12%	**11%**
Low (0-3)	4%	4%	4%	46%	4%	8%	**7%**
Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**
Row Percentages							
High (9-10)	19%	12%	30%	2%	68%	70%	**100%**
Mid (5-8)	11%	6%	17%	27%	56%	83%	**100%**
Low (0-3)	11%	5%	16%	50%	34%	84%	**100%**
Total	**17%**	**11%**	**28%**	**8%**	**64%**	**72%**	**100%**

To further explore potential bias resulting from one-sided comparison tests, Table 25 shows score patterns for paired tests separated from null and paired tests combined (shown in parentheses when different). The top portion of the table shows score results for T3, T4, and T5 by audit outcome category for paired comparison tests, and the bottom portion shows the data as column percentages. Comparing the results for paired tests alone with those including null tests (i.e., values in parentheses), inclusion of null tests consistently elevates high scores by 2-4%, and correspondingly, depresses mid-range and lower scores. The effect appears to be stronger with Category 3 than with Category 1 and Pass audit outcomes. Influences from null tests do not appear to greatly skew scores for Pass and Category 1 outcomes, but this is mostly because comparison test scores are already high. Among Category 2 outcomes, skewing influences of null tests are greater, elevating high scores by as much as 6% and depressing mid and low scores. While these magnitudes are not great, their differential effect across the audit outcomes poses an equity problem.

Table 25. Score Distributions by Audit Category Using Matched Tests (All Tests Shown in Brackets, when Different).

	Audit Outcome					
	Pass	**Cat 1**	**Cat 2**	**Cat 3**	**Cat 4**	**Total**
T-3 Species Groups						
High (9-10)	54	1050	166	0	60	1330 (1949)
Mid (5-8)	3	142	33	0	4	182 (228)
Low (0-3)	0	99	32	1	5	137 (168)
Total	**57**	**1,291**	**231**	**1**	**69**	1649 (2343)
T-4 Halibut						
High (9-10)	8	125	9	0	3	145 (175)
Mid (5-8)	0	41	3	0	2	46 (52)
Low (0-3)	0	6	13	0	1	20 (23)
Total	**8**	**172**	**25**	**0**	**6**	211 (250)
T-5 Sablefish						
High (9-10)	13	160	10	0	6	189 (284)
Mid (5-8)	0	23	4	0	0	27 (28)
Low (0-3)	0	14	5	0	0	19 (20)
Total	**13**	**197**	**19**	**0**	**6**	235 (312)
Total Tests	78 (105)	1660 (2,282)	275 (407)	1 (2)	81 (109)	2095 (2905)
Total Trips	84	384	42	1	9	520
Avg Tests/Trip	**0.9 (1.3)**	**4.3 (5.9)**	**6.5 (9.7)**	**1.0 (2.0)**	**9.0 (12.1)**	4.0 (5.6)
By Column Percentages						
T-3 Species Groups						
High (9-10)	94% (96%)	81% (84%)	72% (78%)	0%	87% (88%)	81% (83%)
Mid (5-8)	5% (4%)	11% (10%)	14% (12%)	0%	6% (5%)	11% (10%)
Low (0-3)	0%	8% (7%)	14% (10%)	100%	7% (6%)	8% (7%)
Total	**57 (78)**	**1,291 (1,813)**	**231 (355)**	**1 (2)**	**69 (95)**	100% (100%)
T-4 Halibut						
High (9-10)	100%	73% (74%)	36%		50% (57%)	69% (70%)
Mid (5-8)	0%	24% (23%)	12% (11%)		33% (29%)	22% (21%)
Low (0-3)	0%	3%	52% (54%)		17% (14%)	9% (9%)
Total	**8 (9)**	**172 (206)**	**25 (28)**	**0**	**6 (7)**	100% (100%)
T-5 Sablefish						
High (9-10)	100%	81% (85%)	53% (63%)		100%	80% (85%)
Mid (5-8)	0%	12% (9%)	21% (17%)		0%	11% (9%)
Low (0-3)	0%	7% (6%)	26% (21%)		0%	8% (6%)
Total	**13 (18)**	**197 (263)**	**19 (24)**	**0**	**6 (7)**	100% (100%)

As shown in previous tables, the outcome of null tests is more likely to produce a high than a low score, and the bias has a differential effect by audit outcome; yet the overall results do not appear to greatly alter score patterns at a consolidated fishery level. These results do not yet examine the potential for scoring bias at the fishing trip or Audit Scorecard level. Examples of null test effects were evident when reviewing Audit Scorecards. Figure 12 provides the T1-DMP test results taken from an Audit Scorecard. The figure shows both FL and DMP estimates for a

landing, with FL reporting one species not evident in the DMP tally, and DMP reporting two species not evident in the FL report. We assume the DMP is correct and the FL record of one species group was incorrect, yet all three null tests were scored as 10, essentially elevating the FL score for observations that did not take place. These three tests were sufficient to skew the T-1 DMP average from 5.75 (without) to 7.57 (with).

Species	Obs Log	DMP	Difference (Weight)	Difference (Percent)	Score
Arrowtooth Flounder	0	98	**98**	100%	10
Lingcod	150	506	**356**	70%	8
Other Flatfish	190	579	**389**	67%	8
Pacific Cod	370	1,084	714	**66%**	0
Pollock	50	0	**50**	100%	10
Rockfish	96,240	74,507	21,733	**29%**	7
Sablefish	0	18	**18**	100%	10

Figure 13. Example of T1-DMP Test Results Showing Null Comparison Tests with Score of 10 (red arrows).

As mentioned, scoring data in the Audit Scorecards were not accessible in a database format, preventing a comprehensive analysis of test area and trip score level bias effects, as shown in Figure 13. The analysis undertaken on the sample of Audit Scorecards yielded comparison test patterns in each of the five test areas. Instances where null tests yielded a score of 10 were tracked and tabulated within the sample of Scorecard reports. Table 26 enumerates total tests and instances of null comparisons yielding a score of 10 within each of the five test areas. Among T1 tests with a Pass, a third of the total T1-DMP tests were null tests scored as 10, whereas the percentage was 19% and 17% for Category 1 and Category 2+3, respectively. The results were similarly high (29% to 36%) for T3 tests among all audit outcomes. Null comparisons with 10 scores were lowest in the T-4 halibut tests, owing to fewer null tests in this group. Surprisingly, over half of the 47 Category 1 T-5 sablefish tests in the sample were null tests scored as 10. This result is comparable to sablefish released catch data (summarised in Table 23), taking into consideration that Audit scoring distinguishes legal and sublegal sablefish and would report more tests overall than the values shown in Table 23, which are not differentiated by size category.

Table 26. Null Comparison Tests with a Score of 10 by Test Area and Audit Outcome.

	T1 - DMP	T2 - Ret	T3 - Rel	T4 - Hal	T5 - Sab	Total
Pass (28 trips)						
# Tests	132	54	14	3	0	203
# Null	43	9	5	0		57
% Null	33%	17%	36%	0%		28%
Cat 1 (72 trips)						
# Tests	527	328	238	46	47	1186
# Null	99	69	77	5	25	275
% Null	19%	21%	32%	11%	53%	23%
Cat 2+3 (30 trips)						
# Tests	244	147	120	25	16	552
# Null	41	28	35	0	3	107
% Null	17%	19%	29%	0%	19%	19%

This issue with size-based sablefish scoring raises another avenue for skewing score results, an issue that is also relevant for releases of lingcod. The FL record reports released sablefish and lingcod weights separately for legal and sublegal sizes. Both species require the fish to be presented on a graduated surface to enable EM reviewer size estimation. When the fish are not placed against the measurement grid, the EM reviewer automatically classifies them as legal-sized. FL more often reports releases of sablefish and lingcod as sublegal, but both sizes occur. Differences in classifications between FL and EM, plus unmeasured fish counted as legal by EM, result in many more null tests than would occur if sized-based categories were not used. For example, FL might report 40 lbs of sublegal size and 0 lbs of legal, while EM viewer estimates 20 lbs of legal and 20 lbs of sublegal. Hence, what started as a score of 10 (40 vs 40) becomes two scores of 10 (40 vs 20, and 0 vs 20). With a high percentage of comparison tests already scoring 10 based on weight values for lingcod (68% of 260 tests) and sablefish (72% of 252 tests), the size-based scoring method provides an avenue for adding more high scores of 10 into the averaging pool.

Discussion

TMP Design

One of the central questions that led to this review was the appropriateness of the TMP, using EM, as compared to the previous ASOP, using observers. Almost universally among monitored commercial fisheries of the world, at-sea observers are considered the gold standard and much lauded for the collection of independent catch and other fishing information. EM is a relative newcomer for at-sea monitoring, particularly among trawl fisheries. One method for assessing the TMP would be to make a comparison of EM and observer program data, but as mentioned in the Introduction this approach would likely be problematic and inconclusive. The approach we chose was to base our assessment of the current TMP as a monitoring system. As one reviewer of this report stated, the TMP is a comprehensive quality assurance system which is a key element of any reputable catch monitoring program. Unlike a single method approach, using Observers or EM, the TMP is a comprehensive system (see Figure 2) with several avenues for evaluation, feedback and improvement. In our view, the TMP is a vast program design improvement over the previous ASOP, yet there are strengths and weaknesses of EM versus Observers in meeting the monitoring, as will be discussed in the next sections.

In addition to its advantages as a quality assurance system, the TMP provides other benefits. The direct involvement of the skipper and crew in catch monitoring, coupled with EM to incentivize accurate recording, serves to build engagement and strengthen trust in the data collected. This approach of shared involvement in building the data systems will strengthen trust by both harvesters and managers in the information used to manage the fishery. Secondly, EM has many benefits over human-based monitoring, including cost, deployment logistics, 24/7 operability, simultaneous vantage points, constant tracking multiple sensors, etc. These advantages reduce the costs and risks for the fishery, making it less dependent on observer availability such as occurred during the pandemic. Another EM related benefit is with data and data processing. As demonstrated in the TMP, EM systems can capture very complex data sets and reviewing software enables very efficient ways to convert sensor and image data sets into meaningful fisheries data. The extent to which Artificial Intelligence can further replace human-based image interpretation is debatable but there is no question it will have a role in further streamlining data collection and analysis processes.

TMP Data Adequacy

Perhaps the most basic question concerning the appropriateness of the TMP is in terms of the data outputs and their adequacy for fishery management purposes. This assessment was based on a partial FY2025/26 data set, as well as with full data sets from FY2024/25 and FY2023/24. Much of the analysis was in comparing EM and FL estimates from released catch and making determinations on their level of alignment. While there are limitations to this approach, it is important to note that the purpose of the testing of the alignment was not one of developing predictive equations with confidence limits, as one would use in sample-based

approach, but one of assessing whether EM estimates are sufficiently adequate to provide guardrails that incentivize accurate FL catch reporting. Admittedly, neither FL or EM catch estimation is a precise process, but the results of our analysis support their intended purpose for the TMP. That is, EM reliably detects catch species and catch quantities sufficiently to test FL reporting practices. The lack of follow up re-examination of imagery with large EM-FL variances is a key weakness of the TMP, as will be discussed later.

The alignment of FL of EM estimates could be examined with more comprehensive and rigorous statistical techniques. However, such techniques will always struggle with the lack of a" true" version and the many reasons why these data are very idiosyncratic. In our opinion, they are useful for qualitative inference. Effort spent attempting to "standardize" these data into rigorous sensibility could be better spent elsewhere. They and do not lend themselves to traditional statistic estimates of variance and confidence limits.

Released Catch

Fishing log release estimates appear to align sufficiently well with EM to provide reasonable confidence in the TMP's ability to generate adequate release estimates at broader scales, such as coastwide or area-specific totals especially for commonly caught species. Confidence in these estimates could be further improved by using numerical bias-correction factors derived as described above. The significance of error in release estimation is also reduced for catch categories in which most of the catch is retained and subsequently verified through DMP. For these major species, misalignment at the finer tow level will be more consequential for IVQ management, although this issue was not strongly emphasized during stakeholder interviews, with the exception of sablefish.

Confidence in catch estimates declines, however, as the number of observations per category decreases, particularly when precision is already limited. As a result, misalignment is more pronounced for categories such as skates. For categories with very few observations, such as coral and sponges, little quantitative credibility can be assumed. Although the program does record these observations, the inferences that can be drawn are generally limited to presence–absence and, in many cases, are only marginally useful for tracking catch incidence over time. This limitation could be reduced by increasing sampling rates, but for rarely observed species, meaningful improvement would likely require 100% sampling.

Retained Catch

Because DMP provides what is considered the "true" estimate of retained catch by species, there is much less concern about alignment between retained catch recorded in the FL and EM. The primary requirement is that these sources align well enough to allow DMP catches to be allocated accurately to quota area.

Interviews indicated that some skippers reallocate catch by tow after examining DMP results. Although this practice is not permitted under the TMP, its prevalence could be investigated

using time stamps in the FL data-entry system. Interviewees noted that, unlike in earlier periods of the fishery, there is now limited incentive to alter the reported capture area.

Species Composition

A comparison of ASOP and TMP catch data suggests that both programs are capable of yielding similar levels of species granularity. While ASOP has greater control over observer training to ensure identification accuracy, the ability of observers, and particularly new observers, can be overestimated (Grinnell 2010). The species list provided in the FL catch log application includes several hundred selections, depending on the grouping, specific species, and other general categories of taxonomic groups. The level of identification within the fleet is variable and there is no standard, other than an encouragement to identify catch to the lowest level possible. Catch records from some trips distinguished catch at a very basic level while catch from other trips showed very high species definition. To meet the general catch-reporting requirement, the species list could be greatly simplified to make it easier to use and direct attention to the species or categories that matter. It is recognized that there is an interest in highly granular catch reporting, but this should be carried out in a very targeted manner by qualified personnel. For example, many fisheries around the world use sentinel programs with specific vessels designated for higher-quality scientific data collection.

Some inference may be drawn by comparison with the trawl surveys. However, these analyses would be highly confounded by the many differences operationally between the two sources. To make only an abbreviated list, the impact of different trawl gear (rollers), codend and intermediate mesh size, spatial and depth differences, tow duration, seasonal and time of day differences. As noted above, these questions might be better addressed though use of strategic allocation of Observers.

TMP Performance

Operational Performance

Based on the audit outcomes records (Table 5), it appears that the Option A Trawl fleet responded quickly and effectively in terms of hosting EM systems aboard, producing complete EM data sets, and complying with FL reporting requirements. The success rate, measured as the percentage of trips with complete EM data sets (i.e., outcomes other than Category 4), has remained in the mid-90% range for most years since FY2021, with FY2025 at 98%. Given that the fleet operates year-round across a large coastal geography, recording continuous video from multiple cameras, and continuous GPS and other sensors (recorded every 10 seconds) from trip start to end, such high fleet coverage rates are a significant achievement. Further, the efforts of skippers and crew to put catch quantification and reporting into practice are also a noteworthy achievement, given that in the 23 years prior to the pandemic, these responsibilities largely fell to observers. Under the ASOP, skippers were required to keep fishing logs, but the official catch record came from observers. Under the TMP, skipper fishing logs became the

official record and received an unprecedented level of scrutiny. These accomplishments are significant and cannot be overstated.

Achievement of nearly 100% fleet monitoring from the start of the TMP relies on operationalizing processes to ensure the EM, FL and DMP data streams work reliably, and the resources are available to efficiently collate, evaluate, and report fishery data on a timely basis. These are vastly more complex processes than the conceptually similar GHLCMP. Examination of FY2025 data indicates that the Scorecard methodology is delivering nearly all (98%) within the required timeline (10 business days). Trends in audit outcomes over time (Table 5) demonstrate that there is a fleetwide response toward improved FL data quality as measured by Scorecards. The operationalizing of FL-EM catch comparisons has been in place from the start of the program, but the methodology is a continuously evolving process (see Appendix 2) and will be discussed in the next section. Despite this evolution in catch data comparison methodology, critical monitoring issues, including time and location of fishing locations and adherence to retention requirements, have been continuously in place from the start of the program.

Before getting into the details of the study findings, it is important to put the Option A Trawl Fishery into perspective with respect to other similar fisheries. Firstly, 100% at-sea monitoring by either observers or EM is extremely rare among trawl fisheries globally. The use of an integrated monitoring design, such as the TMP, is virtually non-existent in other similar fisheries as far as the authors are aware. The point is that, in our view, the TMP is a very cutting-edge program that will likely serve as a model for other fisheries to adopt. The pioneering approach taken with the TMP has put in on an iterative path of testing, implementing, evaluation and improvement.

Functional Performance

Audit Results

High rates of acceptable audit outcomes, coupled with a positive (with qualification) assessment of data adequacy, are strong signals that the TMP is working well and effective at driving FL data quality. However, this positive result should not imply that the TMP continue without a commitment to an ongoing process of improvement, correcting structural issues within the program and targeting areas with a reasonable expectation of further improvements in data quality. This commitment ensures that the TMP maintains a continued focus on FL data quality, and is responsive to temporal changes, such as catch composition, fleet and skipper makeup, and other issues. With this forward-looking focus, the next sections examine different aspects of the TMP to identify shortcomings and possible remedies.

Annual summary reports of audit outcomes demonstrated that the quality of FL catch reporting has improved over time; the vast majority (~90%) of fishing trips meet the acceptance levels with Pass or Category 1 audit outcomes, and improvements seen are from the decline in Category 2 and 3 outcomes. An examination of Audit Scorecard outcomes across the fleet

indicated that, during the year, many vessels had some of each audit outcome, with most trips being Category 1 (Table 7). This result, combined with the increased average number of comparison tests from Pass to Category 1 and 2 audit outcomes (Table 20), is mostly reflective of the greater estimation difficulty associated with more diverse catch assemblages. In reviewing Audit Scorecards there were many instances where Category 2 outcomes showed positive catch reporting efforts that were not evident in Pass or Category 1 outcomes (and the reverse). Pass audit outcomes also stood out, with a high proportion (49%) of reviewed fishing events with zero release, compared with Category 1 and 2 outcomes (10% and 12%, respectively; Table 17). The unique characteristics of trips with Pass outcomes (low released catch) explain why improvements over time have not resulted in a higher proportion of Pass outcomes, but more Category 1 outcomes.

The audit outcome patterns seem to indicate that the levels from Pass to Category 3 reflect a gradient of both catch composition complexity and FL catch estimation quality. Higher numbers of released catch comparison tests at these lower audit outcome levels make it difficult to use audit outcomes as a measure of skipper catch estimation performance. It would be helpful to examine scoring patterns associated with different types of bottom trawl fishing, as well as longitudinal studies of unique skipper/vessel results over time.

In our view, another effective means to search for the best way to examine different methods of comparison test scoring would be the development of a simulation model. Similar to the current Scorecard application used by AMR the model would generate an audit scorecard result, but with the ability to vary all the settings in the scorecard including species groupings, tolerance levels, test area weighting, etc., and generate fleetwide scoring patterns using existing EM-FL catch comparisons from all reviewed fishing events from the most recent year(s). We believe that this would be a cost-effective and efficient way to consider a variety of future catch scoring options. In short, the simulator would process all the trips for a given year with one version of the settings and produce a summary. The settings could be changed for a second run to compare outcomes, and so forth.

Skipper Catch Estimation

Before discussing the different elements of the EM-FL catch comparison and scoring, it is important to consider the impacts of catch-reporting requirements on skippers and crew. The challenges were repeatedly mentioned in all interviews with skippers. Wetfish vessels carry small crews of 3-4 individuals, and the net retrieval operation is perhaps the busiest and most dangerous time while fishing. Everyone is busy as the net and catch come aboard. Regardless of whether the codend comes aboard in one lift, or the catch is brought aboard in smaller splits, skipper and crew attention is fully occupied with multiple simultaneous activities: safely maneuvering the fishing gear, emptying catch from the net in a way that it can be sorted, then sorting out retained catch to be directed into different holds, while sorting out released and prohibited catch and directing it overboard. Catch accounting responsibilities add to this busy operation, including estimating retained catch quantities by species (or species categories) and released catch quantities by species, with special handling, enumeration, and length presentation requirements for halibut, lingcod, and sablefish.

On-deck operations with the RT fleet are more limited, with the majority of catch directed to the receiving tank and halibut, lingcod, and sablefish removed on deck (when possible) for enumeration, measurement, presentation, and release. Below deck, the factory conveyor systems make it more conducive, compared with the Wetfish fleet, for quantifying of released and retained catch, so long as systematic recording procedures are in place over the hours-long catch processing duration. RT vessel skippers indicated that they often carry a dedicated crew member some of whom have previous at-sea observer experience to oversee catch accounting responsibilities, but it is unlikely that they can fully attend all catch sorting operations. Hence, Wetfish and RT fleets have a different set of challenges with respect to catch accounting: Wetfish, having to cope with many simultaneous activities in a short amount of time; and RT having less hectic on deck catch accounting (as compared to Wetfish) but needing to maintain continuous conveyor-based catch accounting over an hours-long time period.

It was clear from skipper interviews that the achievement of catch accounting responsibilities is difficult and has taken time to put in place. The requirements are even harder to achieve with an inexperienced crew, poor weather, and fouled/broken gear. One Wetfish fishing enterprise has a policy of continually taking inexperienced crew to provide vocational training in their community. Skippers and crews do not have the advantages of EM with the ability to rewind and re-examine certain parts of different simultaneous activities from multiple viewing angles. It was clear from discussions with skippers that meeting catch accounting responsibilities involved a balancing act across multiple competing activities, and any improvements would require adjusting the balance, with increases offset by decreases. Skippers noted they missed the years with observer-based monitoring and decreased skipper reporting responsibilities but did not miss the extra costs and challenges with hosting an observer aboard. Given a choice, they would prefer the current EM-based monitoring over observers. If the government were to fund the observer program, some indicated they would prefer having an observer aboard to oversee catch reporting requirements.

Whereas the ASOP utilizes tools and structured procedures for observer catch estimation, the TMP does not provide this guidance other than some information videos provided on AMR's website. The quality of data improves with every step of the estimation process, including designating who carries out what responsibilities, fish handling patterns, systemized estimation methods, ancillary tools (measuring/counting devices, recording forms, etc.), practices in recording and logging data, and timelines for all steps. The estimation procedures should also give some guidance with respect to minimum thresholds of species and quantities that should be recorded. This is important given the scoring methodology currently used for FL data quality assessment. In our view, more attention should be given to catch estimation procedures and the automation tools to make this process as efficient as possible.

Electronic Fishing Logs

The electronic application used to record skipper catch information is a critical tool for the TMP. This study did not include an examination of this application, other than reviewing the documentation and discussing the use of the application with stakeholders. The application is

based on an earlier developed catch log used by observers and not tailored specifically for the TMP. Stakeholder feedback indicated that the application is old, overly basic and difficult to use. Efficiencies such as automatic recording of time and location information is only accurate if data are entered at the time of the activity, and there is a clear GPS signal. As most entries are made well after the fishing activity has concluded time and location entries are usually manually entered. As far as we are aware, data entries are not time stamped, nor does the application record update history, making it difficult to ensure entries are completed on a timely basis, as per the license condition requirements. Given the critical role of this application for the TMP, it should be a priority to ensure that it meets the TMP needs and is user-friendly for accurate and timely reporting of catch information. There should be stricter time requirements for the entry of catch information following the licence condition specifications. As one reviewer to this document noted, providing good tools and embedding clear catch reporting procedures leads to better quality data collection.

Fleet Sample Rates

The design and efficiency of the TMP depend on a sample-based approach in which feedback from the EM review is used to test FL data quality. In cases where the quality of an FL data set does not meet standards of acceptability (i.e., Category 2 or 3) additional review may be triggered. These two processes will be considered separately.

The sample rates for Wetfish and RT fleets follow precedent: a 10% sampling rate used in the GHLCMP and a 25% rate for the RT fleet, which was already in place as part of the ASOP. The realized rates are higher than specified rates owing to the set selection method. There is no statistical basis for the sample rates and setting separate sample rates for Wetfish and RT fleets seems illogical. Both fleets are guided by the same rate-setting logic of incentivising desirable behaviours (i.e., discouraging intentional violations and promoting accurate FL reporting) as opposed to meeting statistical requirements for catch estimation. Higher sampling rates would likely increase the detection of intentional violations, at least to a point where the probability of detection outweighs the benefits of avoiding detection. In the case of driving accurate FL data quality, there is no guarantee that higher fleet sampling rates will result in improvements, so long as the rates are adequate to perform a data quality assessment. In terms of setting optimal rates, there is no reason to believe that the Wetfish and RT fleets differ with respect to these rate-setting criteria.

In our view, fleet sample rates deserve further consideration. The sample rate for the Wetfish fleet is problematic in meeting its objectives. In terms of discouraging intentional violations, a 10% review seems hard to support, as there is a 90% chance a violation will not be detected on any given trip. It should be noted that the probability of detection is higher than 10% with repeated misreporting over multiple trips, but this raises the question of whether the TMP objective is to deter chronic or occasional offenders. In terms of the FL data quality objective, the number of sets selected for review (usually one per trip) results in a small number of EM-FL comparison tests generated (3.3 T3 tests, Table 20). In our view, elevating the minimum sets reviewed per trip for the Wetfish fleet would better meet TMP objectives and improve the optics

for compliance. In terms of the RT fleet, the fleet sample rates appear more than adequate but should adopt the same sample rates as the Wetfish fleet.

The classical trade-off in deterrence theory balances the certainty of punishment against the severity of punishment. But, without delving too deeply into the issue, apparently increasing the likelihood of being caught is far more effective at preventing wrongdoing than increasing the harshness of the penalties.

Additional viewing increases the sample rates for those fishing trips with audit outcomes of Category 2, 3, or 4. In the FY2025 fishing year, such audit outcomes were low (9% of all trips), and only about half were designated for additional review. Most commonly, additional review was performed at a sample rate of 20% or less (i.e., additional review doubled the sample rate from 10% to 20%). In most of these cases of additional review, FL data were substituted with EM estimates for a few selected species (mainly halibut and sablefish), but for only the fishing events reviewed (Table 8).

The purpose of additional review seems muddled. There do not appear to be any clear guidelines to inform additional review decisions as low audit scores originate from any of the five test areas, and additional review is most beneficial for low released catch scores (T3 to T5), and no benefit for low scores with retained catch comparisons to DMP (T1). The specified level of additional review is scaled to the trip score instead of the specific test areas or species test results. Further, scoring anomalies, as will be discussed in the next section, influence test area and trip scores. As a result, additional reviewing decisions may be incorrectly applied. The results from additional reviewing are provided in an updated Audit Scorecard, but this does not affect the original trip score. It is not clear what specific remedial actions are provided as a result of additional review, other than the skipper having to discuss audit outcome with DFO.

Data replacement is limited to a few species and circumstances, but the practice of applying corrections to only reviewed events is illogical when the results suggest the quality deficiencies also lie in more numerous non-reviewed fishing events. Given that there is only a handful of corrections made (16 in the FY2025 data set, Table 8), they provide negligible value in improving the overall catch database. The corrections would have much greater significance for the vessel quota holdings, but limiting their extent to reviewed sets greatly weakens the strength of this accountability measure. Separate from data quality and data replacement functions, additional review was also thought to provide a disincentive because of the additional monitoring costs of processing more EM imagery. However, the cost burden associated with additional review is small ($100-$200) and, as mentioned, is not applicable in all instances of low audit outcomes. None of the skippers interviewed in this study seemed concerned about cost burden associated with additional viewing.

In our view, the main purpose of additional review should be for data quality assessment purposes as the TMP depends upon applying corrective feedback to improve future data reporting. Decisions about the level and extent of additional review should be a nimble process that guarantees a clear assessment of FL data quality and identifies corrective measures, if

needed. This would start with a selection of fishing events from the trip, sufficient in number to form an initial assessment of data quality. The results of this initial review should guide further data quality assessment steps instead of producing an Audit Scorecard and waiting for a determination of next steps by DFO. For example, there may be too few comparisons to inform a proper assessment. The selected fishing event may not be representative of the tows on the trip in general. There may also be specific low scoring anomalies in reviewed sets that should be re-examined, instead of leaving them as unresolved. There may also be other anomalies such as poor practices observed (e.g., blocked cameras, improper catch handling protocols, etc.). In such cases, the review should be expanded, perhaps in a targeted manner, investigating fishing events for specific purposes, perhaps focusing on certain species or only released catch. There may be some logic to automatic review for sets that exceed thresholds of critical species. The most likely application of this measure would be for sablefish, halibut, and lingcod, but this may apply to other species. As well, there may be a place for wider application of this measure, such as when there are first signs of a large year class of an important (i.e., sablefish) species entering the fishery, particularly if it leads to large but relatively infrequent catch events. The result of a broader review process, outlined here, would yield a larger sample size and provide greater confidence in the assessment of data quality. The Audit Scorecard produced should provide results from the entire assessment, not just the initial sample of fishing events.

Data replacement resulting from additional review is clearly a necessary element of the TMP, but its application needs more consideration. In cases where the intent is to apply adjusted released catch values to the vessel quota holdings, the adjustments should be based on a review of all fishing events for the trip. In this case, the review would be targeted (only certain species) rather than a full review. There will also be cases where FL data quality for an entire trip is deemed unacceptably low, such that estimates from the EM review is a better option. Such cases are likely very rare and most likely apply just to released catch. These trips would require a full review of all fishing events to produce the replacement data.

The process of determining acceptance levels for FL estimates of certain species, or all catch, requires further consideration, and the EM data review procedures should accommodate data quality assessment and data replacement needs. Consideration should be given as to how real-time adaptive modifications to a fixed sample rate can be adopted without unnecessary loss of design rigour and degradation of subsequent analyses.

TMP Audit Methodology

Considerable attention has been given in this paper to describing audit methodology, examining of testing methods and scoring results. The audit process is complex and nearly all stakeholders interviewed indicated that did not fully understand it. Indeed, a common comment was that it should be greatly simplified. We do not agree. The audit is necessarily complex because it attempts to make a data quality assessment of a complex fish harvesting process. The audit methodology is the backbone of the QA/QC process of the TMP. Thus, it should be expected that the audit will be comprehensive and will continue to evolve over time, adapting to the data needs and program capacity to fulfill them.

The TMP would benefit from improved user-friendly documentation describing the overall program and audit procedures. This paper explained many areas of the program more simply than provided by DFO (2023) with that in mind. As well, there would be a benefit to restructuring the scorecards and trip summary reporting, consolidating key information in a single report, as will be discussed next.

Audit Scorecards

Audit Scorecards fail to deliver specific results that easily translate to an appropriate program response. This is a key shortcoming of Audit Scorecards that deserves attention. In this vein, the Audit Scorecard should include all relevant information linking the comparison tests to the corrective measures. Instead of separate feedback avenues the report should encompass all issues related to the catch reporting responsibility. For example, in addition to species comparison test results, other information could be included, such as species composition comparison, total catch by tow, repeat viewing results, occurrence reports, camera image quality, catch handling relative to cameras, compliance with fish presentation on measuring boards, FL data entry practices, etc. The EM review process includes an estimate of total catch for all tows, yet this information is not scored or highlighted in any manner. Consolidation of all relevant information into a single report would enable a more informed retrospective analysis. It would likely be more useful for this reporting to be web-based and ultimately reside in databases as opposed to the current practice of query-generated results issued in PDF reports.

Test Scoring

Scoring methodology is simple and provides a way to combine large numbers of different comparisons into a single score that serves as a proxy for overall FL data quality. This method is very useful and should be continued, but its limitations should be considered. The main limitation is that score values are rank or ordinal data; each score is only a category used to qualitatively capture the gradient from poor to good EM-FL alignment. There are seven score values: 0, 3, 5, 7, 8, 9, and 10. Importantly, the ranks carry no numerical relativity among scores, unlike continuous values. It is therefore incorrect to infer that an 8 is 14% superior to a 7. A score of 10 is not twice two scores of 5, nor are three scores of 3 equivalent to a score of 9. For this reason, mathematicians caution that while one can examine the frequency and medians of ranks, averaging is discouraged because it assumes the values have numerical relativity. In the Scorecard, averages serve a useful qualitative purpose by distinguishing levels of quality, but assuming equivalency between two different trips with the same score would be incorrect without reverting to the original comparison data. As was shown in Table 19, two instances with the same average score of 7 can have vastly different underlying source scores. This issue makes it problematic to compare different tests within a test area and to compare test-area averages or trip scores.

This issue, averaging rank scores, is further complicated by the number of tests used to calculate an average. Some hypothetical examples were presented (Table 19) to show how few tests are needed for low test scores to occur, so long as they are offset by high scores. Indeed, the average number of tests varies widely by test area (i.e., T1-T5), fleet type (Wetfish or RT

vessels), trip target species and audit outcome (Table 20). Hence, different types of fishing trips with similar trip scores may have vastly different source scores and data quality.

The key problems with the scoring method therefore arise from estimates being examined qualitatively and from individual values being lost in a test area or trip score averaging process. These deficiencies can be overcome with a secondary step of setting a threshold level of acceptability for each comparison test. The threshold could be determined using different criteria such as test score value, the size of the delta, the percent of released catch composition, species or species group, or a combination of these, with the intent of flagging anomalies. For example, instances with comparison test outcomes with a score of 5 or less could automatically trigger re-examination to confirm the delta and attribute the estimation error (EM or FL). The re-examination outcome, rather than being treated as an unresolved difference, would enable the TMP to more specifically apply corrective action, whether it be the EM program, the skipper, or both. The present practice of not consistently resolving differences in EM and FL score results is an easily resolved weakness in the program.

Null and Paired Value Comparison Tests

Another problem with the audit scoring method is that it requires the same scoring method for both null and paired-comparison tests. Null tests make up about 40% of the total comparison tests and were present in all five test areas (Table 21). From a fairness and equity perspective, applying the same scoring method for both test types is difficult to justify. Paired comparisons provide a true comparison between two independent catch estimates. Null comparison tests contain a non-zero estimate and a null value, the latter serving as a proxy for an absent record. The faulty logic becomes evident in the scoring, where the delta is always equal to the reference value, so the score value is simply based on the size of the reference value. Taking T3 released catch as an example, a score of 10 would require a null estimate of less than 25 lbs, whereas a paired test score of 10 includes any comparison with a difference of less than 25 lbs, so long as the EM estimate is less than 200 lbs. Since most (72%) null tests are with weights of 25 lbs or less, a score of 10 is a given, regardless of whether the estimate was provided by FL or EM (or DMP). In the example shown (Figure 13), the T1 DMP score was elevated by 3 scores of 10, even though 2 of these were not reported in the FL record. It is clearly incorrect that FL data quality assessment results are influenced by null tests arising from EM or DMP only observations. It is also hard to support the use of null tests based on FL-only observations because the results assume EM or DMP errors, even though this has not been confirmed. Most importantly, null tests mask the contribution of true comparisons by paired tests. While, on a whole-fishery level, the effects of null tests appear small, they can skew scores at the trip level enough to elevate audit outcomes. This influence is further complicated by EM contributing about 50% more null tests than FL (i.e., EM provides an estimate and FL does not) for released catch (26% vs 16%, Table 23), and DMP contributing 10-15 times more null tests than FL (Table 21). These occurrences weaken the overall scoring method because they inflate averages that would occur with true matched-pair comparison tests.

Null tests would seem to be an obvious way to knowingly manipulate Audit Scorecard results to achieve a higher test score. As shown, average scores can be elevated by adding more high-

scoring comparison tests. A simple way to achieve this result would be to include low-weight values in the FL record for catch species (or categories) known not to appear in the catch. Hence, these occurrences would not be recorded by EM (or DMP) and would show up as zero values in the catch record. The resulting null comparison test would be scored as a 10, the high score serving to mitigate low score effects and elevate the average score. There was no evidence of this observed, but it would be difficult to detect. Efforts to manipulate audit outcomes would be easier to achieve in the T1-DMP test than in those involving fishing events because the fishing events selected for review are unknown.

The other issue with null tests, aside from scores and averages, is the uncertainty of detection. In the case of paired tests, both confirm the presence of a catch species, whereas in null tests, detection is not confirmed without re-examination. In a catch estimation sense, the issue is trivial for the numerous small catch occurrences. In terms of test scoring, the magnitude of the delta matters; low deltas yield high scores and are often not questioned, while high deltas yield low scores which weakens the TMP as they likely should have been detected but are left unresolved. During stakeholder interviews, it was common to hear skipper comments acknowledging that they sometimes made estimation mistakes, but it was equally common to hear skipper comments that they were right, and the EM reviewer was wrong. AMR does offer an open-door policy with skippers to resolve differences, but this is likely insufficient to address the reputational damage to the TMP caused by unresolved differences. More importantly, unresolved differences deprive the TMP of the opportunity to learn and apply corrective measures.

Irrespective of scoring effects, null comparison pairs are a normal feature when comparing two independent sets of catch observations. FL or EM should be encouraged to use their best effort to identify and estimate all catch species, which will lead to different accounts of catch species and quantities (i.e., null comparison pairs). However, the scoring method should be changed to eliminate bias and the potential for gaming the scoring results. It would be incorrect to eliminate scoring altogether because gaming would not be eliminated (it could cause the opposite problem, forcing some tests not to occur), and also does not serve to include cases where catch observations were not made when they should have been.

Weight and Percent-Based Scoring

Moving to other aspects of scoring, analysis of the released catch data set indicated that score outcome was different for weight-based or percent-based methods (Table 24). It is unknown whether the intent of the tolerances was to more evenly balance scores across the range of expected distribution of values, but the proportion of percent-based tests was only 8%, whereas 92% were weight-based tests. Overall, weight-based tests skewed toward high scores and percent tests toward low scores. Examination of the specified tolerances (Table 2) partly explains this result, as the percent-based score intervals are much more restrictive. For example, with T3 other species (threshold value of 200), a score of 10 by weight must be within 25 lbs, while a score of 10 by percent must be less than 4 lbs when the EM weight is 200lbs. Even though the percent method scales to larger numbers, achieving a score of 10 with a weight of 2,000 lbs would require a delta of less than 40 lbs, a much harder threshold to achieve

than occurs with the weight-based scale. These issues may be trivial in relation to their overall intent of driving data quality performance, but there may be some benefit to setting tolerances more in line with more realistic estimation expectations.

Score ranges defined in the tolerance tables (Table 2) appear to result in high scores with weight-based testing and toward low scores with percent-based testing (Table 24), at least with released catch. It would be useful to examine scoring patterns more broadly to determine if changes should be made. If most percent-based released catch tests yield low scores, the ranges could be adjusted to reflect that high percent-based test scores are unachievable. Similarly, if most weight-based release catch scores are high, perhaps the tolerance ranges may be too generous. The goal here is to establish scores that align with normal patterns yet set realistic targets for achievement.

Lingcod and Sablefish Estimation

Another scoring-related anomaly concerns released catch of lingcod and sablefish. Both species must be presented over a graduated measurement grid to enable distinguishing legal and sublegal sizes. As the two size categories for each species are separately scored, there is no single score for released quantities of each species, a value which is important for data adequacy work. As shown, over two-thirds of tests for either species are for weights of less than 25 lbs. The allocation of these numbers to legal and sublegal bins further adds to high-score tests, as do practices of assigning fish as legal size when they are not presented for measurement. Another issue confounding the FL and EM comparison is that piece count estimates yield different weights because there are no piece weight standards. A better approach would be to score using a single species weight, as is the case for all other catch species, and to develop a separate avenue for reporting compliance with measurement presentation requirements.

Test Area Weighting

The Audit Scorecard currently applies equal weighting of the five test area scores in the trip score calculation. Under this weighting, retained catch (T1 and T2) testing accounts for 40% of the trip score while released catch (T3-T5) accounts for 60%. This weighting attribution between retained and released test scores seems misaligned with the relative value of the information. The greatest risk of a trip with poor retained catch estimates is more difficulty in attributing the trusted DMP landed catch to area of capture. While misallocation spatially of retained catch to area of capture can be is important if it can lead to less precise tracking of area-specific quotas, this was not flagged as a significant issue during interviews. In our view, the greatest value of the TMP for catch estimation, as with most monitoring programs, is with released catch (i.e., T3-5). Aside from the FL record there is no other source for this information, and there is no easy way to remedy poor FL released catch reporting. The quality of FL estimates of released catch other than halibut and sablefish (i.e., T3) is potentially at greatest risk because typical releases involve multiple species and the test score averaging method overshadows comparison tests with poor results. Further, the weighting given to T1 tests (DMP-FL) is also hard to justify, because the FL record does not have to be submitted until after the offload has been completed. During stakeholder interviews, it was evident that some

skippers use DMP values to adjust the FL retained catch record before submission. Future T1 tests should ensure independence between DMP and FL estimates requiring that FL estimates be logged in advance of the offloading operation.

The greater importance of the TMP for released estimates should be more strongly reflected in score weighting between the test areas. Given the high burden to skippers with existing catch reporting requirements, the emphasis should shift towards released catch, with documentation of retained catch kept to a minimum.

Incentive Structures

The TMP is built around the principle of incentivizing good FL recording practices in a way that promotes acceptable data records, while discouraging poor recording practices in a way that those providing substandard FL records can face greater data processing costs, time delays, and/or administrative burden. It is important to consider the different incentives in place, as they serve to enforce the cooperation required for the success of the TMP. One area of discussion during stakeholder interviews concerned incentive mechanisms.

There are a few measures used to encourage compliance, yet it is difficult to understand their relative importance. A commonly mentioned incentive is cost: structuring program fees so that those who comply with the program pay lower fees than those who do not. The primary tool for this is the additional viewing charges applied to trips with low audit outcomes yet, as mentioned, the cost burden is small ($100-$200) and is not applicable in all instances. None of the skippers mentioned cost burden as a factor motivating compliance. Another obvious program incentive is through penalties arising from enforcement measures. Additional viewing likely serves little value as a program incentive.

Another program incentive mentioned concerns penalties by DFO for fishery violations. DFO licence conditions specifically state a requirement to comply with TMP but enforcement actions are only applied to ‘fishery violations’ such as discarding rockfish, landing undersize fish, etc. It is unclear why DFO enforcement measures have not been used to address program-related issues within the TMP.

Based on skipper interviews, the most commonly mentioned incentive driving program compliance was receiving a ‘red report’ (failed Scorecards have a lot of red ink on the front page, while green is used on passing Scorecards), having to talk with DFO about the deficiencies, and having the settlement delayed by the longer audit processing time. Skippers also noted the pride they felt in getting a good test result, an acknowledgement that they were fulfilling their responsibilities correctly. Also mentioned was the desire to help make the TMP work, as the program was much less expensive and did not require hosting an observer. It is unknown if this incentive would persist once the TMP becomes permanent. During development of the GHLCMP, one outcome of poor FL reporting was (or remains) the punitive threat of requiring the vessel to take an observer, at vessel expense. There are a number of reasons why this would be impractical, other than as an empty threat.

While there are measures in place to motivate desired behaviour, the TMP lacks a clear incentive structure in place, as well as a clearly defined pathway for escalation. This may not have seemed to be necessary as the program is largely performing to expectation as defined by the high proportion (89%) of Pass and Category 1 audit outcomes (Table 5). However, as was mentioned, audit categories and Scorecard results do not report FL data quality in a way that would translate to effective escalation measures. Unresolved comparison test differences add to this uncertainty, causing administrative burden in determining appropriate program response.

The TMP should define incentive structures and escalation processes. This not only makes the program more responsive to problem areas but also sets clearer performance expectations and defends program integrity. Further attention should be given to this topic and the following are some dimensions for consideration in strengthening the incentive framework:

- Educational – providing directed feedback on the specific issue, escalating to more hands-on effort with follow-up monitoring.
- Compliance – fines levied by DFO for failure to meet program requirements (in the same way tickets are issued for landing undersized fish), escalating to higher penalties;
- Financial – increased charges for program services, or fines levied that are paid into a program pool, escalating to higher charges or posting performance bonds.
- Program Participation – forfeiture of FL reporting (i.e., 100% EM-based catch estimates), escalating to suspension or cancellation of services.

Recommendations

The TMP Design

Overall, the TMP appears to be operating effectively, providing high levels of fleet coverage, with data compilation, analysis and reporting delivering timely results to meet the operational needs of the fishery. The TMP, also called the 'Pilot EM Program' should be formally adopted, and acknowledging the previous ASOP has been abandoned. Delays in adoption will interfere with needed investments to improve the TMP.

Data Adequacy

Catch data produced by the TMP appear to be adequate to meet the general management needs of the fishery. As might be expected, it works best for species that are caught frequently in larger volumes. Efforts should be made to improve FL estimation accuracy, including measures such as re-examination on large EM-FL estimation deltas, improved scoring methods, more clearly defined catch reporting standards (ie., species granularity and quantity thresholds), and improved catch reporting tools.

With improvements to catch reporting specifications, the TMP may want to distinguish basic requirements that all vessels follow, and enhanced requirements by certain selected vessels to provide more granular catch data for certain segments of the fishery. This enhanced sampling could be accomplished with strategically deployed observers, or through the adoption of sentinel programs, where designated vessel crew provide enhanced data collection services. This would be most valuable for species that are caught in smaller volumes, less frequently and are more difficult to identify.

The dubious identifications and limited sample sizes render rigorous estimation of the many non-commercial species problematic, but we stress that the TMP continue to collect these observations as well as possible, as these observations inevitably become useful for examining some issues, at least qualitatively.

Audit Results

The basic methodology of providing an audit result in a timely fashion following completion of a fishing trip is important. This practice should continue to be directed at data quality. The TMP may want to consider other types of audit feedback, such as FL data quality across multiple trips, by the same vessel/skipper over years, or across different types of trawl fishing (e.g., flatfish, rockfish, midwater).

Skipper Catch Estimates

The catch reporting burden to skippers is recognized and efforts should be made to align obligations to the requirements. In our view, assuming finite resources, the alignment would put less emphasis on retained catch and more on released catch. Catch reporting responsibilities would benefit with clearer standards and improved tools, as previously mentioned.

Electronic Logbooks

The electronic FL should be redesigned to ensure it provides the greatest utility to meet program needs and be as user-friendly as possible. In our view, the application should auto-populate fields where possible, time stamp all entries, and track history for data entry corrections. Improved catch reporting standards should simplify the electronic logbook.

Fleet Sample Rates

In our view, the 10% sample rate used for the wetfish fleet yields too little data to form an adequate data quality assessment and meet the objective of deterring fishery violations. Increasing the sample rate could be accomplished in various ways but perhaps the simplest would be to maintain 10% for the Wetfish fleet, with a minimum sample of two fishing events, instead of one. The RT fleet sampling requirements should align with this approach as there is no logic to support separate rates.

We noted also that data adequacy specifications for certain species may require sample rates different than what may be necessary to meet basic program integrity needs. Hence, a sampling strategy recognizing both needs should be developed.

Additional Review

In our view, additional review should be a standard practice used when the initial sample fails to satisfy certain basic pre-defined data quality thresholds. It should be an automatic process, not one requiring pre-approval of DFO. The Audit Scorecard should be generated using all review information, including that from the additional review. Importantly, additional review should not be presented as a punitive measure, but a normal requirement of the TMP to gain a full understanding of FL data quality, although it follows that better quality FL and EM data sets will be easier to examine and likely require less additional review.

Data Replacement

EM data are occasionally used in place of FL data in certain instances of low scoring comparison tests. Data replacement is limited to fishing events reviewed when the results suggest the quality deficiencies also lie in non-reviewed fishing events. Some guidelines are provided (see discussion) and we recommend that guidelines be made to clearly specify when FL data are deemed unacceptable and EM replacement is necessary.

Audit Scorecards

The Audit Scorecard is a useful communication tool and should be continued as a routine post-trip reporting tool. The Audit Scorecard should not lose sight of the main objective: identifying significant EM to FL discrepancies within the catch data and implementing corrective measures to minimize them in the future. In this vein, the Audit Scorecard should include all relevant information linking the comparison tests to the corrective measures. Some ideas for this were mentioned in the discussion. It is unrealistic to distill data quality into a single trip numerical score when acceptance should be based on multiple factors, some of which are binary

(measurement protocols followed?) and others qualitatively ranked. It would likely be more useful for this reporting to be web-based as opposed to the current practice of issuing PDF reports. The audit outcome should tie to a specific escalation pathway, if needed. If not already available, application development should include building permanent and automatically populated databases to house all the results of these tests and outcomes.

Test Scoring

The overall scoring methodology is a very practical way of assimilating several different tests; however, changes are needed. Low scores with high deltas should not get lost in the averaging method, as their investigation provides important feedback for both EM and FL catch estimation activities. For this reason, it is recommended that all test scores of 5 or lower be re-examined by the EM viewer to confirm the delta and to reveal the causes. This should be an automatic process, not one that requires DFO approval.

Null and Paired Value Comparison Tests

A lot of attention was given to the comparison tests themselves, in particular null and paired value comparison tests. Given that null tests account for 40% of all released catch comparison pairs (Table 22), they overshadow the true comparisons (paired tests) and elevate scores; their effect needs to be mitigated in a way that prevents manipulating the audit outcome. In our view, the best approach would be to cap null comparison test scores to a value of 8, so that low scores still occur and bias effects of high null test scores are negated. A null test score of 8 would have much less potential to skew average scores, as was shown in the Figure 13 example.

Weight and Percent-Based Scoring

As noted, the score ranges in the tolerance tables appear to skew toward high scores with weight-based testing and toward low scores with percent-based testing, at least with released catch. It would be useful to examine scoring patterns more broadly to determine if changes should be made. If most percent-based released catch tests yield low scores, the ranges could be adjusted to reflect that high percent-based test scores are unachievable. Similarly, if most weight-based release catch scores are high, perhaps the tolerance ranges may be too generous. The goal should be to establish score rules that align with normal patterns yet set realistic targets for achievement.

Lingcod and Sablefish Estimation

The scoring analysis also investigated size-based scoring methods applied to lingcod and sablefish. With low release catch quantities for these species in two-thirds of the comparison pairs, testing legal and sublegal separately yields many more high scores that numerically bias the results, while not providing an assessment of species catch quantity itself. Assessment of compliance with measurement presentation requirements should be separately evaluated from catch estimates.

Salmon Catch Estimation

Salmon catch estimation protocols appear to have evolved continually since 2000 (see Appendix 2). While it is to the credit of the TMP to work to improve this element, these protocols should be stabilized and documented in order to support ongoing quality assurance monitoring.

Test Area Weighting

The value of the TMP for assessment of released over retained catch was mentioned several times. Given the high burden to skippers with existing catch reporting requirements, the emphasis should shift towards released catch, with documentation of retained catch kept to a minimum, assuming finite resources. The weighting of the T1 to T5 test areas should be revised to place greater emphasis on release estimates. T1 DMP and T2 Retained should carry a combined weighting of no more than 20% with the remainder split between T3 (40%), T4 (20%) and T5 (20%). Given the current practice of FL reports not being required until after the offload has been completed, T1 testing has little value as compared with the other test areas.

Ongoing Program Monitoring

Much of the analysis presented in this report should be incorporated into routine processes for ongoing monitoring. This would require some revisions to data bases (mentioned below) to enable access to certain data and data views. It will also require development of a dedicated quality assurance program with structured reporting. We recommend that a review of the TMP, similar to this one, be conducted approximately every five years. No monitoring program can be assumed to continue to be effective without ongoing reviews and revision.

Simulation Monitoring

One of the weaknesses of this study was that Audit Scorecard data are query generated for the report and not easily available in data base format for analysis. The scoring performed in this study was limited to the sample taken, and the summary data extracted from the sample. It would be helpful to examine scoring patterns in a broader fashion, such as with different types of bottom trawl fishing. In our view, the best way to examine different methods of comparison test scoring would be the development of a simulation model. Built out from the core of the current Audit Scorecard query developed and used by AMR the model would generate an audit scorecard result, but with the ability to vary all the settings in the scorecard including species groupings, tolerance levels, test area weighting, etc., and generate fleetwide scoring patterns using existing EM-FL catch comparisons from all reviewed fishing events from the most recent year(s). We believe that this would be a cost-effective and efficient way to consider a variety of future catch scoring options. In short, the simulator would process all the trips for a given year with one version of the settings and produce a summary. The settings could be changed for a second run to compare outcomes, and so forth.

TMP Resources

Uncertainty about the future of the TMP in its current form has likely contributed to a hesitancy to invest in areas beneficial to the program. It is hoped that the parties can quickly resolve this

and solidify the future of the TMP. The lack of investment in the program will soon become a limiting factor. Areas of investment include:

- Program Documentation – All stakeholders would benefit from simple, easy to understand multimedia material (documents, videos, etc.) to explain the overall program, and provide guidance on specific issues (catch estimation, FL tools, score methods, etc.).
- Audit Outcome Reporting – The Audit Scorecards and Trip Summary Reports should be revised into a common feedback document that is easier to read and provides more directed feedback. The report should include all feedback and not use multiple avenues for communicating post trip assessments.
- FL Estimation Methods and Tools – This area needs attention, as mentioned.
- Database Issues – As recommended, ongoing monitoring of the TMP should occur and databases developed to easily enable these processes.
- Technology Investments – The drawback of technology-based monitoring is that it requires an ongoing investment in EM systems and other data gathering instruments. Some areas of technology are evolving rapidly and there may be benefits to fleetwide adoption in certain areas. It is suggested that technology investments be considered proactively and planned for, rather than waiting until they become obsolete and require replacement in an ad hoc manner.

Incentives

Last, but not least, the TMP should further develop the set of incentive measures that both apply penalties to areas of deficiency and rewards to areas of exceptional performance. These measures solidify program compliance expectations and specify consequences to be applied as a routine operational procedure, not one that requires administrative process. In our view, incentive systems are essential, not optional, components of an effective and efficient monitoring program.

Acknowledgements

Several people and organizations contributed to this project. The authors would like to recognize the guidance and advice provided by the project steering committee, which included DFO staff (Management, Science, and C&P), Trawl Industry representatives and an Environmental Non-Government Organization. In addition, during the initial stages of the project, we interviewed several stakeholders including trawl skippers, fleet managers, crew, representatives from other groundfish license sectors, TMP project leads, a lawyer with a specialization in fishery issues, and service providers. These individuals were generous in their time and provided much essential information, but they shall remain anonymous as was agreed prior to the interviews. We would like to recognize Archipelago Marine Research Ltd. for their role in the project providing much of the data and data summaries used in the review, as well as providing staff time explaining many aspects of the project. Vericatch also contributed to the study with information about their role in the program. We particularly wish to acknowledge the support provided by Norm Olsen, in providing historical fishery data from the DFO GFFOS system, and assisting with statistical analysis of EM and FL catch data. This report was externally reviewed by five anonymous experts in the field, whose comments benefited the final version. We also wish to acknowledge the editorial expertise of Hannah McElderry who reviewed final versions of this report. The authors wish to credit ChatGPT which assisted with the development of the French translation of the Abstract and with the compilation of the Executive Summary.

References

Allard, J, and Benoît, H.P. 2019. Unified framework for the statistical assessment of fishery monitoring programs. DFO Can. Sci. Advis. Sec. Res. Doc. 2019/013. vi + 63 p.

Davies, T. K., Quinn, E. and Jardim, E. 2023. A novel framework to evaluate the accuracy of information used in MSC fisheries assessment: Development challenges and solutions. Marine Policy:158 (2023) 105869.

DFO, 2019. Fishery monitoring policy. dfo-mpo.gc.ca.

DFO, 2020. Assessment of the dependability of fishery catch monitoring programs. Can. Sci. Sec. Sci. Advis. Rep. 2020/022.

DFO, 2024. Electronic monitoring program for Option A trawl vessels. Version 7.3 – April 10, 2024.

DFO, 2025. Pacific Region Integrated Fisheries Management Plan, Groundfish, Effective February 21, 2025. Pacific Region Integrated Fisheries Management Plan. 341 p.

DFO, 2025. Pacific Spiny Dogfish (*Squalus suckleyi*) Population Modelling for Outside Waters of British Columbia in 2024. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2025/032.

Gale, K. S. P., Forrest, R., Rooper, C., Nephin, J., Wallace, S., Driscoll, J., and Rubidge, E. 2022. Evaluation of collaboratively developed management measures to reduce coral and sponge bycatch in a fully monitored multispecies trawl fishery. Can. J. Fish. Aquat. Sci. 79:9. 2022. https://doi.org/10.1139/cjfas-2021-0274.

Gislason, G. S. and Associates. 2024. 2021-2022 BC Groundfish Trawl Industry Profile, Development and Contribution to the British Columbia Economy First Report of a Triennial Series Prepared for: Canadian Groundfish Research and Conservation Society. Prepared by: GSGislason and Associates Ltd. Vancouver, B.C.

Gravelle, K. M. 2015. Validation of catch estimates for select species within the British Columbia groundfish hook-and-line catch monitoring program. Master of Science Thesis, Royal Roads University, Victoria, BC. Available from VIURRSpace: https://www.viurrspace.ca/items/0e5d0f47-b6a5-42fc-b431-ea8582b20062.

Grinnell, M. H. 2010. Evaluating the reliability and equitability of at-sea observer reports in the B.C. offshore groundfish trawl fishery. M.R.M. Simon Fraser. Report (Simon Fraser University, School of Resource and Environmental Management). W.A.C. Bennett Library (Burnaby). SH329.F56 G75 2010.

Legendre, P. 2024. Model II regression user's guide R Edition. (version 1.7.4).

McElderry, H. 2008. At-sea observing using video-based Electronic Monitoring. Electronic Monitoring Workshop July 2008. https://www.researchgate.net/publication/242517381_At-Sea_Observing_Using_VideoBased_Electronic_Monitoring

McElderry, H., J. Schrader, and J. Illingworth. 2003. The efficacy of video-based electronic monitoring for the halibut longline fishery. Can. Sci. Adv. Sec. Res. Doc. 2003/042.

McElderry, H., McCullogh, D., Schrader, J., and Illingworth, J. 2007. Pilot study to test the effectiveness of electronic monitoring in Canterbury fisheries. DOC Research & Development Series 264. Department of Conservation, Wellington. 27 pp. http://www.doc.gov.nz/documents/science-and-technical /drds264.pdf.

Michelin, M., Elliott, M., Bucher, M., Zimring, M., and Sweeney, M. 2018. Catalyzing the growth of Electronic Monitoring in fisheries: Building greater transparency and accountability at sea. California Environmental Associates and The Nature Conservancy https://www.ceaconsulting.com/wp-content/uploads/CEA-EM-Report-9-10-18-download.pdf

Ricker, W. E. 1973. Linear regressions in fishery research. Journal of the Fisheries Board of Canada, 30(3): 409-434.

Stanley, R.D., N. Olsen, and A. Fedoruk. 2009. The accuracy of yelloweye rockfish catch estimates from the British Columbia Groundfish Integration Project. Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science 1:354–362.

Stanley, R.D., McElderry, H., Mawani, T., and Koolman, J. 2011. The advantages of an audit over a census approach to the review of video imagery in fishery monitoring. ICES Journal of Marine Science. 2011: 68(8): 1621-1627. doi: 10.1093/icesjms/fsr058.

Stanley, R.D., Karim, T., Koolman, J., and McElderry, H. 2015. Design and implementation of electronic monitoring in the British Columbia groundfish hook and line fishery: a retrospective view of the ingredients of success. ICES Jour. Mar. Sci.; doi:10.1093/icesjms/fsu212.

Wallace, S., Turris, B., Driscoll, J., Bodtker, K., Mose, B., and Munro, G. 2015. Canada's pacific groundfish trawl habitat agreement: A global first in an ecosystem approach to bottom trawl impacts. Marine Policy 60 (2015) 240:248.

Appendices

Appendix 1 - Fishery History

Table A1 1. Summary of Option A Trawl Fishery Activity by Fishing Year (2-15-2025). Catch Represents Total Landings by Target Species. (to November 2025).*

	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025*
#RTV Vessels	7	7	7	7	7	8	8	8	9	5	2
#WetVessels	44	45	45	44	42	40	36	33	31	28	23
#Vessels	51	52	52	51	49	48	44	41	40	33	25
#Trips	1,050	1,267	1,360	1,401	1,301	1,141	906	847	783	773	501
#Hauls	11,718	12,472	11,998	11,867	11,637	10,021	9,213	9,001	7,815	6,336	4,267
Effort (hr)	1,988	2,544	2,907	2,868	2,977	2,572	1,986	1,791	1,492	1,451	986
Days	4,690	5,155	4,926	4,818	5,072	4,680	3,922	3,734	2,886	2,554	1,667
Catch (000's MT)											
Flatfish	15,091	14,972	15,709	13,670	12,068	4,383	8,464	8,351	6,308	7,710	2,997
Rockfish	18,131	20,349	16,720	15,432	13,889	11,380	14,791	16,408	16,702	16,320	14,475
Pollock	3,180	1,348	1,587	1,562	5,493	5,858	7,637	5,620	4,797	9,735	4,652
Other	2,195	844	1,057	561	958	748	1,419	2,318	1,975	1,822	1,233
Inside Hake	9,064	7,183	6,021	8,848	5,761	9,819	7,319	7,685	7,209	9,272	1,452
OutsideHake	39,436	70,675	83,434	97,569	98,836	97,704	61,195	38,076	29,227	3,355	4,358
All Trawl	87,097	115,371	124,528	137,641	137,005	129,893	100,824	78,458	66,217	48,214	29,168

*Table A1 1. Annual Totals for Option A Trawl Fishery. (Source DFO FOS, * - Data to Nov. 2025).*

Year	#Vessels	#Trips	#Hauls	Effort (hr)	Trip Days	Ttl Ret	Ttl Rel	Ttl Catch	% Release
2010	61	1,472	14,821	3,123	5,992	75,925	8,158	84,082	9.7%
2011	59	1,443	14,866	3,278	5,724	79,342	7,458	86,800	8.6%
2012	58	1,377	14,550	3,427	6,079	80,294	7,011	87,305	8.0%
2013	58	1,388	14,186	3,368	9,902	88,817	6,779	95,595	7.1%
2014	55	1,154	12,947	2,616	4,935	79,297	5,761	85,058	6.8%
2015	51	1,050	11,718	1,988	4,690	79,498	7,599	87,097	8.7%
2016	52	1,267	12,472	2,544	5,155	108,943	6,428	115,371	5.6%
2017	52	1,360	11,998	2,907	4,926	117,504	7,024	124,528	5.6%
2018	51	1,401	11,867	2,868	4,818	130,735	6,906	137,641	5.0%
2019	49	1,301	11,637	2,977	5,072	129,410	7,595	137,005	5.5%
2020	48	1,141	10,021	2,572	4,680	126,011	3,882	129,893	3.0%
2021	44	906	9,213	1,986	3,922	97,040	3,784	100,824	3.8%
2022	41	847	9,001	1,791	3,734	73,706	4,753	78,458	6.1%
2023	40	783	7,815	1,492	2,886	60,673	5,543	66,217	8.4%
2024	33	773	6,336	1,451	2,554	44,998	3,216	48,214	6.7%
2025*	25	501	4,267	986	1,667	27,410	1,758	29,168	6.0%

Appendix 2 - TMP Chronology of Major Changes to the TMP Since Implementation.

The chronology of the TMP evolution is summarized below, using information from the following unpublished documents:

AMR 2022, 2023, 2024, 2025. AMR Annual Program Reports to CGRCS for FY2022-FY2025.

DFO 2020. Chronology of Events – Evolution of events – Evolution of the Groundfish Option A EM Program. 2020-2025. Trawl EM Program., Dec 2, 2020.

DFO 2022. Final Groundfish Trawl Electronic Monitoring Program_v7.1_June 2022.pdf and versions 7.2 and 7.3.

June 2020

- Category 2 scoring design
- Appendix B to DFO program document adding report card template – May 2022 update
- Appendix C to DFO program document adding new procedure to change at-sea log to include unreported tows detected during the audit
- April 3, 2020, Fishery Management Order (FMO) 2020-01 formally suspending At-Sea Observer Programs (ASOP).
- April 14, 2020, Implementation of the “Emergency Electronic Monitoring Pilot Program”,
- April 2020 DFO begins receiving “tow comparison reports” for trips that do not meet initial EM standards. These reports are essentially the same as the “tow comparison” reports that are provided as supplementary information with the “audit Scorecard”, which was developed later subsequently.
- May 15, 2020, FMO replaced with FMO 2020-03, authorizing at-sea observers (ASOs) to board vessels where safe working procedures related to COVID are in place. No ASOs were deployed during this time period.
- Major health/safety concerns raised by GTAC members regarding continued risk of COVID-19 posed to fishing operations. Request to continue development of EM program and halt reinstatement of ASOP program.

September 2020

- Establishment of an ad hoc “Trawl EM Technical Working Group” address data/technical issues arising from EM program. Membership: CGRCS, DSTA, GMU, C&P, AMR, Vericatch (ad hoc).
- After October 29, 2020, vessel operators hailing out as “Option A quota Observed” have been required to complete Observer log and Skipper logbook. October 2019, Vericatch rolled out new version, so the Observer logbook populates Skipper logbook. Now one logbook.
- Vericatch worked on having GPS being added automatically be included.
- Delayed reinstatement of ASOP program pending further discussion with industry.

October 2020

- Updated Emergency EM Requirements that came into effect starting Oct 29, 2020.

November 2020

- Postponement of full reinstatement of ASOP and continued permission to deploy EM as an "alternative mitigating management measure". ASOs offered for those wishing to elect to take an observer, subject to observer availability and safe work procedures, etc. including:
  - mandatory completion of ASOP log (and FLOG) to fix ongoing logbook entry issues given reduced detail offered by FLOG only and restore trip type selection;
  - required use and installation of upgraded V5 systems across all Option A vessels;
  - measuring grid requirements (mainly for measuring legal/sublegal sablefish and lingcod discards).
- CGRCS letter to Skippers and Vessel Owners regarding new EM rules - Oct 20, 2020.pdf
- DFO worked to establish initial program standards, which were then reviewed with the Trawl EM TWG throughout winter/spring. These discussions refined the standards and shaped the initial audit program design.

2021

- August 15, 2021, New Emergency EM requirements (v5) came into effect starting August 15, 2021, including additional above and below deck camera requirements and implementation of the "audit score card". A grace period for Category 2 and 3 consequences was granted, which began at the start of the 2021/2022 fishing season.
- Developed the comprehensive Option A groundfish trawl EM audit program score summary report at the direction of the CGRCS, DSTA, and DFO to meet the improved audit reporting standards implemented August 15, 2021.
- Developed the comprehensive Option A groundfish trawl EM audit program score summary report at the direction of the CGRCS, DSTA, and DFO to meet the improved audit reporting standards implemented August 15, 2021.
- Made eight revisions to the format, content and data comparisons for the Option A groundfish trawl EM audit score summary report between November 2021 and May 2024 based on feedback from the DSTA, CGRCS, DFO, and skippers.
- Developed the comprehensive Option A groundfish trawl EM audit program summary report at the direction of the CGRCS, DSTA, and DFO to meet the improved audit reporting standards implemented August 15, 2021.

2022

- Category 1 and 4 consequences commenced. Earlier versions (i.e., v1-4) of this document were not released beyond the Trawl EM TWG.
- March 29, 2022, AMR hosted a catch estimation workshop for trawl vessel crews/skippers to improve catch estimation methods on Wetfish vessels and RTvessels. Training covered methods for estimating total catch and species composition, including visual estimates, volume calculations, and sampling techniques such as basket and timed interval samples. It also included guidance on completing logbooks using worksheets and Vericatch "Trawler", along with tips and prevention or common errors. A summary of the training was posted publicly on AMR's website for future use.
- DSTA, CGRCS and DFO worked to revise Pacific hake biological sample procedures for Wetfish vessel deliveries in April 2022 to target the number of samples required for stock assessment. Samples and data delivered to DFO for annual Pacific hake stock assessments in 2022, 2023, and 2024.
- Worked with the DSTA, CGRCS, and DFO to revise Pacific hake biological sampling.
- Collaborated with the DSTA, CGRCS, and DFO to develop a new Pacific hake biological sample program for vessels freezing catch at sea that was implemented July 13, 2022.

- The groundfish trawl biological sampling technical workgroup designed a shore-based biological sampling program for rockfish between November 2022 and March 2023.
- Implemented the groundfish trawl enhanced salmon monitoring and sampling program in September 2022. The sampling program has continued under a DFO contract with funding from DFO for the 2023/2024 to 2025/2026 groundfish seasons.
- Presented an overview of the Option A groundfish trawl EM program to the Council of the Haida Nation together with the CGRCS and DSTA in February 2022.
- Conducted groundfish trawl catch estimation and reporting workshops for groundfish trawl skippers and crew in March 2022 and provided a video of the training for future reference in April 2022.
- July 1, 2022, Updated v7.1 EM program requirements came into effect starting July 1, 2022, including implementation of compounding additional EM review requirements under Category 2 - adjusted release thresholds specifically for higher-volume species (arrowtooth, hake, and pollock).
- Also, formally removed “Emergency” from program title since program transitioned from emergency COVID response to a more formalized audit system.
- AMR maintained Canadian General Standards Board (CGSB) listing required for ASOP providers until the requirement was removed by DFO in 2019. CGSB listing and audits replaced by DFO annual reporting and National Verification Strategy (NVS). Archipelago audited by the DFO NVS team in October 2022.
- Implemented the groundfish trawl enhanced salmon monitoring and sampling program in August 2022. The program has continued under a DFO contract with funding from DFO for the 2023/2024 groundfish season.

2023

- January 2023 contract awarded to CGRCS to facilitate trawl EM system configuration reviews, including a standardized “EM system checklist”. Vessel configuration reviews were conducted either by a select group of trusted industry experts (e.g., Jon Eis, former ASO), or by C&P (Trevor Ruelle) who completed the RT vessel reviews due to conflict-of-interest concerns. Camera/EM system adjustments were made to active vessels over the next 6-12 months.
- AMR implemented a pilot rockfish biological sampling program in Ucluelet on March 20, 2023. The results of the program were reviewed by the technical working group in May and August and has been successful in collecting biological samples from rockfish species that are representative of the catch and in accordance with the sample design.
- AMR implemented the rockfish biological sampling program in Port Hardy and Prince Rupert on June 19, 2023.
- Developed Pacific halibut length sampling procedures for the option A groundfish trawl EM program in August 2023 at the direction of the CGRCS, DSTA and DFO. Using EM program halibut data, provided summaries to the groundfish trawl EM technical working group in August 2023, September 2023, April 2024, July 2024, December 2024, and June 2025 to inform the development and review of the Pacific halibut length sampling procedures.
- The groundfish trawl biological sampling technical working group worked on expanding the biological sampling program to include additional species in 2023, 2024, and 2025. Sample program design was tested for each new species using groundfish trawl catch data to simulate the number and distribution of samples that would be collected under different scenarios.
- Arrowtooth flounder biological sample collection commenced on freezer RT vessels in October 2023 and for Wetfish vessel landings in September 2024.
- AMR designation as an ASOP Corporation renewed for two years by DFO on April 1, 2023.

- Made seven revisions to the format, content and data comparisons for the Option A groundfish trawl EM audit score summary report between November 2021 and February 2023 based on feedback from the DSTA, CGRCS, DFO and skippers.
- The current version of the audit score summary report was implemented for the 2023/2024season on February 21, 2023. The most recent update included revisions to aggregate audit scores for retained arrowtooth flounder, flatfishes, big skate and longnose skate as well as updates to the report format and the addition of a relative trip score to the report.
- The groundfish trawl biological sampling technical working group worked on the design of a biological sampling program for arrowtooth flounder from May to September 2023. The design was \tested using groundfish trawl catch data to simulate the number and distribution of samples that would be collected. The design of the arrowtooth flounder biological sampling program was finalized in the fall of 2023.
- Updated above-deck camera requirements for vessels with receiving tanks
- Introduction of a combined retained species test for flounders, soles and skates
- Addition of eulachon and corals and sponges to species of interest section in audit an automated EM sensor and location data use process when standards not met.
- Addition of relative trip score to audit (Section 5.2)
- Eulachon and corals and sponges added to species of interest report (Section 6)

2024

- Pacific halibut length sampling pilot program was implemented in October 2023 followed by fleetwide implementation in May 2024. To date (2024) 5,358 halibut length measurements verified by EM have been delivered to DFO science staff.
- Developed sablefish and lingcod length measurement procedures for the option A groundfish trawl EM program between 2023 and 2024 in collaboration with the CGRCS, DSTA and DFO. Implemented in 2024.
- Made eight revisions to the format, content and data comparisons for the Option A groundfish trawl EM audit score summary report between November 2021 and May 2024 based on feedback from the DSTA, CGRCS, DFO and skippers.
- Implemented (May 15) 14-day audit flagging feature
- Automated additional review
- Pacific cod biological sample collection commenced at wet boat landing in September 2024.

2025

- Provided EM program sablefish release data summaries to the CGRCS and groundfish trawl EM technical working group in April, May of 2025 and July 2025 to inform the review of the sablefish and lingcod release measurement procedures.
- Met with the CGRCS and representatives from the Canadian Sablefish Association to review the EM program sablefish release data summaries in May of 2025.
- Participated in MSC certification site visits in with the DSTA and CGRCS in April 2025. Provided an overview of the Option A groundfish trawl EM, DMP and biological sampling program to staff from MRAG Americas responsible for the groundfish trawl MSC assessment.
- Gulf hake biological sampling procedures finalized, and sampling set to commence in October 2025
- Updated DMP offload monitoring procedures required for the implementation of individual chinook salmon bycatch limits for the 2024/2025 groundfish trawl fishery in February 2024.
- A groundfish trawl salmon excluder survey was completed by the Pacific Legacy No. 1 in April 2024. AMR completed net camera imagery review, data analysis and reporting

required for the CGRCS to evaluate the effectiveness of the groundfish trawl salmon excluder device.

- The technical report delivered to the CGRCS and DFO in July 2024 included survey results and recommendations for future work to test the effectiveness of salmon excluder devices.
- The current version of the audit score summary report was implemented for the 2024/2025 season on May 15, 2024.
- Flatfish biological sample collection commenced for Wetfish vessel landings in May 2025. Flatfish species targeted for biological sampling include Petrale Sole, English Sole, Dover Sole and Rock Sole.
- Updated DMP offload monitoring procedures required for the implementation of individual chinook salmon bycatch limits for the 2024/2025 groundfish trawl fishery in February 2024.

Appendix 3 – Example Audit Scorecard



Option A Trawl EM Audit Score Summary Report

Vessel	[redacted]	VRN	[redacted]		
Skipper	[redacted]	Fishery	Groundfish Trawl Trawl Option A - EM		
Offload Date	[redacted]	Validation Record #	[redacted]		
Hail In Number	[redacted]	FOS ID	[redacted]	AMR ID / Hail Out Number	[redacted]
Audit Date	[redacted]	DFO Contact Date	[redacted]		

1. Audit Result Summary

Trip Score	Audit Result
8.66	Category 1 - QSR automatically generated

Catch Comparison	Program Standard	Test Result	Score
Dmp to Observer Log Retained Test	All audit test scores must be 8 or better	6 of 8 tests passed	8.63
Video to Observer Log Retained by Species Test	The average score for all audit tests must be 8 or better	3 of 4 tests passed	8
Video to Observer Log Released by Species Test (Excluding Halibut and Sablefish)	The average score for all audit tests must be 8 or better	5 of 6 tests passed	8
Video to Observer Log Released Sablefish Test (Legal, Sublegal, Unmeasured and Undetermined)	All audit test scores must be 8 or better	2 of 2 tests passed	10

If you have questions about the data reported here please contact Archipelago at 250-383-4535 or at trawlersupport@archipelago.ca.
The DFO Review Board Contacts:
Lindsay Richardson-Deranger 604-345-4731 Lindsay.Richardson-Deranger@dfo-mpo.gc.ca

Audit Score Category Definitions
Pass: All audit tests meet program standards
Category 1: One or more audit tests do not meet standards. Trip score is between 7.01 and 10
Category 2: One or more audit tests do not meet standards. Trips score is between 3.01 and 7

Trip Score 6.01 to 7: Additional 10% of remaining tows are reviewed.
Trip Score 5.01 to 6: Additional 20% of remaining tows are reviewed.
Trip Score 4.01 to 5: Additional 30% of remaining tows are reviewed.
Trip Score 3.01 to 4: Additional 40% of remaning tows are reviewed.

Category 3: One or more audit tests do not meet standards. Trips score is between 0 and 3

May be subject to 100% review.

Category 4: Number of observer log tows does not match EM or there are EM sensor or video gaps

2. EM Sensor Data to Observer Log Comparison

Sensor Comparisons	Results (info only, does not affect scores)
Skipper Events to EM Events	Match
Skipper GMU Area to EM GMU Area	Skipper GMU area on 10 of 10 sets match EM GMU Areas
Start Position	Skipper position on 8 of 10 sets match EM within 2 Nm
Start Date and Time	Skipper date/time on 10 of 10 sets match EM within 60 minutes

3. Summary Of Test Results

3.1 Observer Log Retained Catch To Dockside Monitoring Record (DMP) (Weights in pounds)

For each species or group tested, the total retained weight for the trip (as recorded in the Observer Log) are compared to the DMP weight for that species or group. Species or group with a DMP weight greater than 1000 pounds are scored based on percent difference while those under 1000 pounds are scored based on the difference in weight. The difference scored is noted in bold text in the table below. All scores must be 8 or better to meet program standards.

Species	Obs Log	DMP	Difference (Weight)	Difference (Percent)	Score
Arrowtooth Flounder	250	885	**635**	72%	5
Lingcod	775	783	**8**	1%	10
Longnose Skate	0	25	**25**	100%	10
Other Flatfish	2,610	3,911	1,301	**33%**	5
Pacific Cod	450	349	**101**	29%	10
Pacific Halibut	0	19	**19**	100%	N/A
Pollock	0	20	**20**	100%	10
Rockfish	85,600	92,410	6,810	**7%**	9
Sablefish	24	49	**25**	51%	10
Salmon	10	8	**2**	25%	N/A

3.2 Observer Log Catch to EM Video Estimates of Retained and Released Catch (Weights in round pounds)

The Retained and Released weights by species or group reported in the observer log are compared to EM for 10% of tows on wet vessels and 25% of tows on receiving tank vessels. Retained weights for species or group with an EM weight greater than 500 pounds are scored based on percent difference while those under 500 pounds are scored based on the difference in weight. Released weights for species or group (other than pollock, Pacific hake and Arrowtooth flounder) with an EM weight greater than 200 pounds are scored based on percent difference while those under 200 pounds are scored based on the difference in weight. Released weights for pollock and Arrowtooth flounder with an EM weight greater than 2000 pounds are scored based on percent difference while those under 2000 pounds are scored based on the difference in weight. Released weights for Pacific hake with an EM weight greater than 5000 pounds are scored based on percent difference while those under 5000 pounds are scored based on the difference in weight.
The average score across all species and tests between observer log and EM must be 8 or better to meet the program standards. All Halibut and Sablefish scores must be 8 or better to meet the program standards. See section 4 for average scores.
Beginning February 21, 2023 Retained Observer Log to EM Video Estimates of Retained Arrowtooth Flounder, Flatfish, Longnose Skate and Big Skate are combined for the purpose of determining audit scores.

Arrowtooth Flounder

	Retained			Released		
Tow	Obs Log	Video	Score	Obs Log	Video	Score
1			N/A	125	60	10

Flatfish

	Retained			Released		
Tow	Obs Log	Video	Score	Obs Log	Video	Score
1			N/A	100	103	10

Longnose Skate

	Retained			Released		
Tow	Obs Log	Video	Score	Obs Log	Video	Score
1			N/A	75	48	9

Big Skate

	Retained			Released		
Tow	Obs Log	Video	Score	Obs Log	Video	Score

Arrowtooth Flounder, Flatfish, Longnose Skate and Big Skate Combined

	Retained		
Tow	Obs Log	Video	Score
1			

Rockfish

	Retained			Released		
Tow	Obs Log	Video	Score	Obs Log	Video	Score
1	1,435	2,603	3			N/A

Pacific Cod

	Retained			Released		
Tow	Obs Log	Video	Score	Obs Log	Video	Score
1	75	0	10			N/A

Pollock

	Retained			Released		
Tow	Obs Log	Video	Score	Obs Log	Video	Score
1			N/A	10	0	10

Pacific Hake

	Retained			Released		
Tow	Obs Log	Video	Score	Obs Log	Video	Score

Spiny Dogfish

	Retained			Released		
Tow	Obs Log	Video	Score	Obs Log	Video	Score
1			N/A	2,200	2,000	9

Pacific Halibut

	Retained			Released		
Tow	Obs Log	Video	Score	Obs Log	Video	Score

Sablefish Totals

	Retained			Released		
Tow	Obs Log	Video	Score	Obs Log	Video	Score
1	12	12	10	2	3	N/A

Sablefish Releases

	Released (Sublegal)			Released (Legal)		
Tow	Obs Log	Video	Score	Obs Log	Video	Score
1	2	0	10	0	3	10

Sablefish Size Unkown Considered Legal by EM

	Released Unmeasured	Released Size Not Determined
Tow	Video	Video
1	3	

Lingcod Totals

	Retained			Released		
Tow	Obs Log	Video	Score	Obs Log	Video	Score
1	200	72	9	328	99	N/A

Lingcod Releases

	Released (Sublegal)			Released (Legal)		
Tow	Obs Log	Video	Score	Obs Log	Video	Score
1	328	99	0	0		N/A

Lingcod Size Unkown Considered Legal by EM

	Released Unmeasured	Released Size Not Determined
Tow	Video	Video

Salmon

	Retained			Released		
Tow	Obs Log	Video	Score	Obs Log	Video	Score
1	10	0	N/A			N/A

4. Summary of Audit Scores

Catch Comparison	Average Score (Used to determine Trip Score)	Weighted Average Score (Not used for Trip Score at this time)
DMP to Observer Log Retained	8.63	8.82
Video to Observer Log Retained	8	3.38
Video to Observer Log Released	8	8.69
Video to Observer Log Sablefish Released	10	10

Average scores for each catch comparison are determined by summing all scores for all species and tests and dividing by the number of scores. Weighted average scores for each catch comparison are determined by multiplying each test score by the DMP or EM weight (or observer log weight when the DMP or EM weight is zero), summing the scores for each catch comparison and dividing by the total DMP or EM weight for the catch comparison.

For example, if the DMP weights for species A is 2000 lbs and for species B is 6000 lbs and the score for species A is 7 and for species B is 9 the weighted average score is 8.5 as follows:

((2000 X 7) + (6000 X 9)) / 8000 = 8.5

5. Species of Interest Summary

Does not affect audit results or scores but is provided to the DFO review board for consideration. The Species of Interest Catch Report will be included with every audit and contains test results for sea birds, sharks (other than dogfish), marine mammals, Eulachon, Corals and Sponges. This catch report is included to increase the accuracy of recording of these species in the fishing logbook with an end result of having independent data on these species.

Species of Interest Observer Log to DMP (Weights in round pounds)

Species	Obs Log	DMP	Difference (Weight)

Species of Interest Observer Log to EM Video Catch (Weights in round pounds)

Species	Tow	Obs Log	Video	Difference (Weight)

Appendix 4 – Example Trip Summary Report

ARCHIPELAGO
MARINE RESEARCH LTD.

Trawl Option A Auditing Report

Trip Number: [redacted] Vessel: [redacted] Fishing Area: 3D,5A
Trip Start Date: [redacted] Total Sets : (Em-Sensor / Flog) 10 / 10

Part 1. Fishing Log To EM Data Comparisons

*** Total catch is retained and released combined.**

Haul # 1	Date	Time Difference	Latitude	Longitude	Position Difference (Nm)
FLOG Haul Start	May 02,25 04:53	0 Mins	[redacted]		
EM Haul Start	May 02,25 05:02	9 Mins			0.15
FLOG Haul End	May 02,25 07:43	0 Mins			
EM Haul End	May 02,25 07:58	15 Mins			0.32

Total Catch for Event*:		Wgt. Difference:	Difference
FLOG	4,597		
EM	5,000	403 lbs	8.77 %

Haul # 2	Date	Time Difference	Latitude	Longitude	Position Difference (Nm)
FLOG Haul Start	May 02,25 16:35	0 Mins	[redacted]		
EM Haul Start	May 02,25 16:40	5 Mins			0.21
FLOG Haul End	May 02,25 20:39	0 Mins			
EM Haul End	May 02,25 20:46	7 Mins			0.08

Total Catch for Event*:		Wgt. Difference:	Difference
FLOG	18,866		
EM	20,000	1,134 lbs	6.01 %

Haul # 3	Date	Time Difference	Latitude	Longitude	Position Difference (Nm)
FLOG Haul Start	May 03,25 06:12	0 Mins	[redacted]		
EM Haul Start	May 03,25 06:16	4 Mins			**12.75**
FLOG Haul End	May 03,25 07:45	0 Mins			
EM Haul End	May 03,25 07:51	6 Mins			0.06

Total Catch for Event*:		Wgt. Difference:	Difference
FLOG	2,127		
EM	2,000	-127 lbs	-5.97 %

Haul # 4	Date	Time Difference	Latitude	Longitude	Position Difference (Nm)
FLOG Haul Start	May 03,25 08:31	0 Mins	[redacted]		
EM Haul Start	May 03,25 08:34	3 Mins			0.08
FLOG Haul End	May 03,25 10:29	0 Mins			
EM Haul End	May 03,25 10:37	8 Mins			0.09

Total Catch for Event*:		Wgt. Difference:	Difference
FLOG	12,175		
EM	12,000	-175 lbs	-1.44 %

Haul # 5	Date	Time Difference	Latitude	Longitude	Position Difference (Nm)

Printed: May 12, 2025 Page 1 of 7

FLOG Haul Start	May 03,25 11:09	0 Mins	[redacted]	
EM Haul Start	May 03,25 11:18	9 Mins		0.95
FLOG Haul End	May 03,25 13:23	0 Mins		
EM Haul End	May 03,25 13:16	6 Mins		**1.94**

Total Catch for Event*:		Wgt. Difference:	Difference
FLOG	6,202		
EM	6,000	-202 lbs	-3.26 %

Haul # 6	Date	Time Difference	Latitude	Longitude	Position Difference (Nm)
FLOG Haul Start	May 03,25 15:06	0 Mins	[redacted]		
EM Haul Start	May 03,25 15:10	4 Mins			0.17
FLOG Haul End	May 03,25 17:03	0 Mins			
EM Haul End	May 03,25 17:11	8 Mins			0.16

Total Catch for Event*:		Wgt. Difference:	Difference
FLOG	5,072		
EM	5,500	428 lbs	8.44 %

Haul # 7	Date	Time Difference	Latitude	Longitude	Position Difference (Nm)
FLOG Haul Start	May 03,25 18:24	0 Mins	[redacted]		
EM Haul Start	May 03,25 18:28	4 Mins			0.16
FLOG Haul End	May 03,25 21:23	0 Mins			
EM Haul End	May 03,25 21:30	7 Mins			0.04

Total Catch for Event*:		Wgt. Difference:	Difference
FLOG	28,385		
EM	28,000	-385 lbs	-1.36 %

Haul # 8	Date	Time Difference	Latitude	Longitude	Position Difference (Nm)
FLOG Haul Start	May 04,25 06:55	0 Mins	[redacted]		
EM Haul Start	May 04,25 06:58	3 Mins			0.14
FLOG Haul End	May 04,25 09:22	0 Mins			
EM Haul End	May 04,25 09:29	7 Mins			0.11

Total Catch for Event*:		Wgt. Difference:	Difference
FLOG	8,660		
EM	11,000	2,340 lbs	27.02 %

Haul # 9	Date	Time Difference	Latitude	Longitude	Position Difference (Nm)
FLOG Haul Start	May 04,25 10:23	0 Mins	[redacted]		
EM Haul Start	May 04,25 10:29	6 Mins			**6.38**
FLOG Haul End	May 04,25 11:47	0 Mins			
EM Haul End	May 04,25 11:55	8 Mins			0.04

Total Catch for Event*:		Wgt. Difference:	Difference
FLOG	4,087		
EM	5,000	913 lbs	22.34 %

Haul # 10	Date	Time Difference	Latitude	Longitude	Position Difference (Nm)
FLOG Haul Start	May 04,25 13:17	0 Mins	[redacted]		
EM Haul Start	May 04,25 13:21	4 Mins			0.15
FLOG Haul End	May 04,25 15:41	0 Mins			
EM Haul End	May 04,25 15:50	9 Mins			0.07

Total Catch for Event*:		Wgt. Difference:	Difference
FLOG	5,423		
EM	8,000	2,577 lbs	47.52 %

Part 2. Total Catch Summary

Calculated Released Weight:

Total Est. Wt. from Video Review	102,500
Actual Landed Wt. from DMP	98,945
CALCULATED TOTAL RELEASED =	3,555

Total Wt. From Fishing Log:

Released	5,875
Retained	89,719
TOTAL CATCH =	95,594

Difference between FLOG released and calculated released = 2,320

Difference between FLOG retained and DMP landed = -9,226 lbs. 9.32 %

Part 3. Species Comparisons

Retained	FLOG	DMP	Difference	Difference %
bocaccio	2,685	5,672	-2,987	**-52.66 %**
canary rockfish	47,775	46,550	1,225	2.63 %
chinook salmon	10	8	2	25.00 %
darkblotch rockfish	0	8	-8	-100.00 %
dover sole	1,060	2,139	-1,079	**-50.44 %**
english sole	100	144	-44	-30.56 %
greenspotted rockfish	55	0	55	100 %
greenstripe rockfish	0	447	-447	-100.00 %
lingcod	775	783	-8	-1.02 %
long nose skate	0	25	-25	-100.00 %
pacific cod	450	349	101	28.94 %
pacific halibut	0	19	-19	-100.00 %
pacific ocean perch	3,000	7,764	-4,764	**-61.36 %**
petrale sole	550	266	284	106.77 %
pollock	0	20	-20	-100.00 %
ratfishes	0	490	-490	-100.00 %
red banded rockfish	530	918	-388	-42.27 %
redstripe rockfish	750	1,758	-1,008	**-57.34 %**
rex sole	900	1,362	-462	-33.92 %
rosethorn rockfish	10	75	-65	-86.67 %
sablefish	24	49	-25	-51.02 %
sharpchin rockfish	3,600	4,067	-467	-11.48 %
shortspine thornyheads	45	415	-370	-89.16 %
silvergray rockfish	5,625	7,953	-2,328	-29.27 %
splitnose rockfish	0	42	-42	-100.00 %
turbot	250	885	-635	-71.75 %
widow rockfish	360	574	-214	-37.28 %
yelloweye rockfish	210	167	43	25.75 %
yellowmouth	11,200	7,905	3,295	41.68 %
yellowtail rockfish	9,755	8,095	1,660	20.51 %

Species Comparison for Event#: 1	FLOG	EM	Difference	Difference (%)
Bocaccio				
Retained	25	600	-575	-95.83 %
Canary Rockfish				
Retained	500	68	432	635.29 %
Chinook Salmon				
Retained	10	0	10	100 %
Greenspotted Rockfish				
Retained	50	0	50	100 %
Lingcod				
Released	328	0	328	100 %
Released - Sub-legal	0	99	-99	-100.00 %
Retained	200	72	128	177.78 %
Long Nose Skate				
Released	75	48	27	56.25 %
Pacific Cod				
Retained	75	0	75	100 %
Pacific Ocean Perch				
Retained	0	1,929	-1,929	-100.00 %
Petrale Sole				
Released	0	3	-3	-100.00 %
Pollock				
Released	10	0	10	100 %
Ratfish				
Released	25	0	25	100 %
Red Banded Rockfish				
Retained	25	3	22	733.33 %
Redstripe Rockfish				
Retained	500	0	500	100 %
Rex Sole				
Released	100	100	0	0.00 %
Rosethorn Rockfish				
Retained	10	0	10	100 %
Sablefish				
Released	2	0	2	100 %
Released - Unmeasured	0	3	-3	-100.00 %
Retained	12	12	0	0.00 %
Silvergray Rockfish				
Retained	250	0	250	100 %
Spiny Dogfish				
Released	2,200	2,000	200	10.00 %
Turbot				
Released	125	60	65	108.33 %
Yelloweye Rockfish				
Retained	50	0	50	100 %
Yellowtail Rockfish				
Retained	25	3	22	733.33 %

			ASOP	EM	Difference	Difference %
Arrowtooth Flounder	**Turbot**	Discard	125	60	65	108.33 %
	Totals:		**125**	**60**		
Lingcod	**Lingcod**	**Discard**	**328**	**99**	**229**	**231.31 %**
		Retained	**200**	**72**	**128**	**177.78 %**
	Totals:		**528**	**171**		
Longnose Skate	**Long Nose Skate**	Discard	75	48	27	56.25 %
	Totals:		**75**	**48**		
Pacific Cod	**Pacific Cod**	Retained	75	0	75	100 %
	Totals:		**75**	**0**		
Pollock	**Pollock**	Discard	10	0	10	100 %
	Totals:		**10**	**0**		
Sablefish	**Sablefish**	**Discard**	**2**	**3**	**-1**	**-33.33 %**
		Retained	**12**	**12**	**0**	**0.00 %**
	Totals:		**14**	**15**		
Spiny Dogfish	**Spiny Dogfish**	Discard	2,200	2,000	200	10.00 %
	Totals:		**2,200**	**2,000**		
	Ratfish	Discard	25	0	25	100 %
	Totals:		**25**	**0**		
Other Flatfish	**Petrale Sole**	Discard	0	3	-3	-100.00 %
	Rex Sole	Discard	100	100	0	0.00 %
	Totals:		**100**	**103**		
Rockfish	**Bocaccio**	Retained	25	600	-575	-95.83 %
	Canary Rockfish	Retained	500	68	432	635.29 %
	Greenspotted Rockfish	Retained	50	0	50	100 %
	Pacific Ocean Perch	Retained	0	1,929	-1,929	-100.00 %

Rockfish	**Red Banded Rockfish**	Retained	25	3	22	733.33 %
	Redstripe Rockfish	Retained	500	0	500	100 %
	Rosethorn Rockfish	Retained	10	0	10	100 %
	Silvergray Rockfish	Retained	250	0	250	100 %
	Yelloweye Rockfish	Retained	50	0	50	100 %
	Yellowtail Rockfish	Retained	25	3	22	733.33 %
	Totals:		**1,435**	**2,603**		
Salmon	**Chinook Salmon**	Retained	10	0	10	100 %
	Totals:		**10**	**0**		

All DMP weights are "Net Weight" (Product Wt. x conversion factor).