Draft 2024 Annual Deployment Plan and Partial Coverage Cost Efficiencies Analysis

September 2023

DRAFT PREPARED FOR THE PCFMAC

Fisheries Monitoring and Analysis Division, Alaska Fisheries Science Center
National Marine Fisheries Service
7600 Sand Point Way NE
Seattle, WA 98115

Alaska Regional Office
National Marine Fisheries Service
P.O. Box 21668
709 W. 9th Street
Juneau, Alaska 99802

Table of Contents

Executive Summary

This draft 2024 Annual Deployment Plan (ADP) documents how the National Marine Fisheries Service (NMFS) intends to assign at-sea and shoreside fishery observers and electronic monitoring (EM) to vessels and processing plants engaged in halibut and groundfish fishing operations in the North Pacific during the calendar year 2024.

Purpose

Observers and Electronic monitoring (EM) in the North Pacific are funded through industry funds. In the partial coverage fleet, funds are provided through an ex-vessel fee. At the October 2019 North Pacific Fishery Management Council (Council) meeting, the Council identified developing cost efficiencies in the partial coverage component of the Observer Program as one of its highest priorities moving forward. In 2022 NMFS recommended the development of an integrated analysis of the partial coverage category. In response to that request, this document presents a draft 2024 Annual Deployment Plan (ADP), in which alternative sampling designs are evaluated in an effort to improve the effectiveness and efficiency of partial coverage fisheries monitoring in the North Pacific. Unlike former draft ADPs, this analysis explores alternative ways that monitoring resources (observers and EM) could be deployed in a cost-effective manner while improving data quality and scientific utility for stock assessment, catch accounting, and other fishery management purposes.

This draft ADP incorporates all monitoring methods that are now part of the partial coverage Observer Program, including traditional at-sea observer coverage, EM for fixed-gear vessels, EM for pollock trawl catcher vessels, and shoreside observers. Additionally, it models the costs of supporting all observer and EM monitoring exclusively from fee revenues, whereas, in past ADPs, EM programs have been supported with fee revenue plus additional outside funding.

The overarching purpose of this draft ADP is to compare alternative scientifically robust, cost-effective sampling plans with the goal of choosing an appropriate sampling plan for deployment in 2024 and beyond. In addition, the analysis seeks to achieve the following Council goals:

- Efficiently distribute monitoring such that more monitoring is achieved for the available budget
- Increase monitoring on trawl-fisheries for PSC accounting
- Monitoring that has least impact on fishing operations
- A partial coverage program that isn't contentious

The ADP specifies the scientific sampling deployment design to be used in the partial coverage category and includes three elements: 1) the selection method to accomplish random sampling; 2) division of the population of partial coverage trips into selection pools, referred to as stratification; and 3) the distribution of monitoring resources to the strata, referred to as allocation.

Selection Method

In 2024, observers and EM will be deployed to vessels to monitor catch at-sea or to monitor shoreside deliveries using random selection methods. Trip-selection refers to the randomized method of selecting fishing trips which are the sampling unit. Under trip selection, vessel operators and owners log their trips into the Observer Deploy and Declare System (ODDS) and are notified if the trip is selected for coverage. In shoreside processing facilities, observers implement a systematic random sample design by selecting a random starting point and monitoring every n^{th} delivery thereafter; individual deliveries are the sample unit.

Stratification

This analysis evaluated four stratification definitions: the *Current* definition (2023) based on gear (hook-and-line, pot, trawl) and monitoring type (observer, EM); the *FMP* definition included FMP (BSAI, GOA); and two additional fixed-gear stratification definitions (*Fixed*) that combined fixed gears (HAL, POT) into a single strata to account for trips fishing HAL + POT on a single trip.

To increase the chances of monitoring trips in both the BSAI and the GOA in an efficient manner, we examined stratification definitions using FMP (BSAI, GOA). Stratifying by monitoring tool (observers, EM), gear type (hook-and-line, pot, trawl) and FMP resulted in strata with enough trips to provide a reasonable likelihood of being sampled. However, further splitting the BS and AI into separate strata produced some areas with very low total effort which were likely to go unmonitored.

Stratification by gear type (HAL, POT, TWL) has been used in past ADPs. However, fixed gear usage is changing: approximately 15% of observer-pool and 20% EM-pool fixed gear trips fished both hook-and-line and pot gear on a single trip in 2022. These trips cannot be unambiguously placed into strata defined by the use of a single gear type, and as a result, standard estimation methods could produce biased estimates. Alternative stratification definitions were evaluated to correct this issue. The stratification definition that performed best combined all trips that fish with HAL, POT, or both gear types on a single trip into a single stratum and included FMP (*Fixed FMP)*. The *Fixed FMP* stratification addressed the issue of assigning trips fishing with multiple gear types to strata without creating strata with low effort or high likelihood to be unmonitored. Analysts recommend use of the *Fixed FMP* stratification for 2024.

Allocation

Four alternatives for how fisheries monitoring assets are allocated among strata were explored. These included *Equal rates*, *Status quo*, *Cost-weighted boxes*, and *Proximity* allocation. *Equal rates* provides unbiased estimation from samples in the case where there is little to no prior information about the fishery. *Equal rates* are presented to provide a baseline from which to evaluate other designs since we can use information from the fishery to better inform our allocation strategy. *Equal rates* by default, is an allocation that is not affected by changes to stratification and therefore cannot differentially allocate samples to FMPs.

The *Status quo* allocation sets rates through a baseline + optimization algorithm for observers and by policy for EM. The budgets explored ($3.5, $4.5, and $5.25 million, assumed from fees only) were not sufficient to provide optimized *Status quo* allocation for observers resulting in little to no differences under the alternative stratification definitions because when monitoring rates are under the 15% minimum, the observer strata are allocated equally. The *Status quo* allocation results in large amounts of EM sampled trips which contributes to the large number of trips sampled overall which improves cost efficiency by reducing the variable cost of monitoring trips of different durations. EM cost efficiency improves as sample size increases more so than at-sea observer coverage. *Status quo* allocation also results in the lowest CV for chinook PSC. *Status quo* allocation results in little overlap (low interspersion) between observed trips and EM monitored fleets and between the observed and the unmonitored fleet. The lack of interspersion means that data from observers (age, length, maturity) would need to be used to account for missing data elements which are not collected by EM. However, this creates a problem because the lack of interspersion of the observer data is unlikely to be representative of the EM fishing activity. The *Status quo* allocation method resulted in the fewest observer samples collected at-sea, meaning less age, length, maturity, and stock of origin data will be available for use in stock assessments and stock of origin (genetics) analyses. The *Status quo* allocation resulted in the highest variability in PSC estimates of Pacific halibut from trawl gear and crab PSC. In addition, the review of imagery from fixed gear EM collected at-sea is too slow to be of any practical use for in-season management of quotas.

Alternative allocation methods to *Status quo* were developed to improve the cost-efficiency and the scientific merit of fisheries monitoring data. The *Proximity* and *Cost-weighted boxes* allocation methods employ algorithms to prioritize sampling strata that are expected to otherwise result in datagaps. Unlike *Status quo*, these allocation methods integrate EM into the deployment process, treating EM strata in the same way as observer strata. *Cost-weighted boxes* prioritizes the utilization of cheaper monitoring methods whereas *Proximity* prioritizes the sampling of smaller strata. These two allocation methods performed similarly. Relative to *Status quo,* both *Cost-weighted boxes* and *Proximity* had relatively good interspersion, ability to detect monitoring effects, improvements in data timeliness, and increased CVs of halibut and crab PSC, but both had a relatively poor CV of chinook PSC relative to *Status quo*.

The participation in the EM program is voluntary which causes a cost inefficiency because vessels that opt-in to EM but fish very little incur high fixed costs (EM equipment installation and maintenance). This results in the slightly better performance of *Proximity* over *Cost-weighted boxes* because the cost inefficiency is, by nature, present in *Cost-weighted boxes*. If the pool of EM vessels were pared down to include those that regularly fish, the cost efficiencies would be maximized and *Cost-weighted boxes* would likely outperform *Proximity* allocation. Analysts do not recommend the designs employing the *Status quo* allocation method or *Equal rates* allocation method due to low observer coverage and their inability to differentially allocate samples to FMPs across the range of budgets evaluated.

Summary

The current sampling design of the partial coverage monitoring in the North Pacific (*Current* stratification + *Status quo* allocation) balances the myriad objectives of the Observer Program including expansion of EM and large total sample size. However, in the past, the Observer Program has required additional funds beyond fee revenue to support the program. Because of the large uncertainty in obtaining additional funding for the program going forward, this analysis assumed budgets which relied exclusively on fees. At the fee-only budget levels examined here, the benefits of the current sampling design are not realized, resulting in poor quality data which would negatively impact fisheries management decisions. However, the analyses presented here demonstrate that the *Fixed FMP* stratification coupled with either the *Cost-weighted boxes* or *Proximity* allocations appear to have the ability to provide the most effective data from both EM and observers and collect it most efficiently at the variable budget levels.

1. Introduction

Authority

This draft 2024 Annual Deployment Plan (ADP) evaluates how the National Marine Fisheries Service (NMFS) intends to assign at-sea and shoreside fishery observers and electronic monitoring (EM) to vessels and processing plants engaged in halibut and groundfish fishing operations in the North Pacific. This plan is developed under the authority of the Magnuson-Stevens Fishery Conservation and Management Act (Magnuson-Stevens Act) (16 U.S.C. 1862), the Fishery Management Plan for Groundfish of the Bering Sea and Aleutian Islands Management Area (BSAI FMP), the Fishery Management Plan for Groundfish of the Gulf of Alaska (GOA FMP), and the Northern Pacific Halibut Act of 1982. The ADP outlines the science-driven method for deployment of observers and EM systems to support statistically reliable data collection. The ADP is a core element in implementation of section 313 of the Magnuson-Stevens Act, which authorizes the North Pacific Fishery Management Council (Council) to prepare a fishery research plan in consultation with NMFS.

The Council's role in the annual deployment plan process is described in the analysis that was developed to support the restructured observer program (NPFMC 2011) and in the preamble to the proposed rule to implement the restructured observer program (77 FR 23326). The preamble to the proposed rule notes that:

> *NMFS would consult with the Council each year on the deployment plan for the upcoming year. The Council would select a meeting for the annual report consultation that provides sufficient time for Council review and input to NMFS. The Council would likely need to schedule this review for its October meeting. The Council would not formally approve or disapprove the annual report, including the deployment plan, but NMFS would consult with the Council on the annual report to provide an opportunity for Council input. The final deployment plan would be developed per NMFS' discretion to meet data needs for conservation and management. (77 FR 23344 & 23345).*

The ADP follows the process envisioned by the Council and NMFS when the restructured observer program was developed and implemented. As a result, both the ADP development and the evaluation of data collected by observers and EM is an ongoing process. NMFS works with the Council throughout the annual review and deployment cycle to identify improved analytical methods and ensure Council and public input is considered.

More details on the legal authority of the ADP are found in the Final Rule for Amendment 86 to the BSAI FMP and Amendment 76 to the GOA FMP (77 FR 70062, November 21, 2012). Further details on the integration of EM deployment into the ADP process are found in the final rule to integrate EM into the Observer Program (82 FR 36991).

North Pacific Groundfish and Halibut Observer Program

NMFS implements the Council's fishery research plan through the North Pacific Groundfish and Halibut Observer Program (Observer Program). The Observer Program provides the regulatory framework and support infrastructure for stationing observers and EM systems to collect data necessary for the conservation, management, and scientific understanding of the commercial groundfish and Pacific halibut fisheries of the BSAI and GOA management areas. EM is broadly defined as technological tools which collect fishing data to support stock assessment and fishery management. In the North Pacific, EM is usually more specifically referencing video imagery and sensors to provide catch and discard information and compliance monitoring after video review.

The Observer Program is the largest observer program in the country and is responsible for monitoring a fleet of nearly a thousand vessels that fish a combination of hook-and-line, pot, and trawl gear across the Alaska Exclusive Economic Zone (EEZ) area of roughly 3.77 M km^2. The deployment of monitoring assets (observers and/or EM) is the first stage of a hierarchical sampling design (Cahalan and Faunce 2020). Since 2013, the trip has been the primary sampling unit. Fishing trips made by vessels are assigned to either full and partial coverage.

In full coverage, every trip is monitored by 1 or 2 observers if monitoring is completed at sea, or by an EM system at sea and an observer at the processing plant receiving the trip's catch. For full coverage trips, vessel and processing plant owners/operators are responsible for procuring observer and EM hardware services directly through NMFS-authorized companies. There are currently three NMFS-permitted observer service provider companies and two NMFS-approved EM hardware companies.

For partial coverage trips, vessel owners/operators declare each trip in a NMFS database and if the trip is selected for coverage, a NMFS-contracted observer provider company arranges for coverage. Funding for partial coverage is obtained from an ex-vessel fee on landings from the prior year and is used by NMFS to pay for observer and EM services. In the partial coverage component, the ADP specifies the scientific sampling design and the selection rate—the portion of trips that are sampled. NMFS and the Council recognized that selection rates in partial coverage, for any given year, would be dependent on available revenue generated from fees on groundfish and halibut landings. The annual apportionment of the budgets for observer deployment and EM system deployment is also reflected in the ADP process. The ADP process allows NMFS to adjust deployment in each year so that sampling can be achieved within financial constraints. While fisher participation in observer monitoring is automatic, if a vessel wishes to participate in at-sea EM they must volunteer, be approved by NMFS, and follow a vessel monitoring plan. Cost efficiency of an EM vessel may change over time, but hardware infrastructure cannot be easily or cheaply modified to respond to different fishing effort patterns. As a result of these different rules of participation, the selection rates for observer coverage change from one calendar year to the next to achieve efficiency, cost savings, and data collection goals while the selection rates for EM have been set by policy.

Observer Program Data Collection

Data collection through the Observer Program provides a reliable and verifiable method for NMFS to gain fishery discard and biological information on fish, and data concerning seabird and marine mammal interactions with fisheries. These data contribute to the best available scientific information used to manage the fisheries in the North Pacific. The design of the holistic monitoring program that meets mandates of the Magnuson-Stevens Act, Marine Mammal Protection Act (MMPA), and Endangered Species Act (ESA) ensures that multiple monitoring programs are not required on the fleet. Observers and EM systems provide fishery-dependent information that is used to estimate total catch and interactions with protected species. Managers use these data to manage groundfish and prohibited species catch within established limits and to document and quantify fishery interactions with protected species. Much of this information is expeditiously available (e.g., daily or at the end of a trip, depending on the type of vessel) to ensure effective management. Scientists also use fishery-dependent data to assess fish stocks, evaluate marine mammal and seabird interactions with fishing gear, characterize fishing impacts on habitat, and provide data for fisheries and ecosystem research and fishing fleet behavior. While both observers and EM systems provide fishery-dependent data, these monitoring methods provide different information on catch and interactions with protected species. Table 1-1 summarizes the broad suite of data collection through the different monitoring approaches under the Observer Program.

ADP Process

On an annual basis, NMFS develops an ADP to explain how observers and EM will be deployed for the upcoming calendar year, and prepares an Annual Report that evaluates the performance of the prior year's ADP implementation. NMFS and the Council created this ADP / Annual Report process to provide flexibility in the deployment of monitoring assets used to gather reliable data for estimation of catch in the groundfish and halibut fisheries off Alaska.

The Annual Report is presented to the Council in June each year and informs the Council and the public about how well various aspects of the program are working. The review highlights areas where improvements are recommended to 1) collect the data necessary to manage the groundfish and halibut fisheries, 2) maintain the scientific goal of unbiased data collection, and 3) accomplish the most effective and efficient use of the funds collected through the observer fees.

A draft ADP that outlines sampling for the upcoming year is prepared in October each year and a final ADP is completed in December. The ADP allows for partial coverage strata definitions, participation requirements, allocation methods, and selection rates to change each year. Strata help define how trips will be monitored (for example which vessels belong to observer or EM selection pools and the requirements necessary to participate in each) and may be based on factors such as gear type, vessel length, home or landing port, availability of EM systems, funding, and monitoring goals.

Since 2013, aspects of deployment have been adjusted through the ADP (Table 1-2). The modifications have included moving types of partial coverage trips between selection pools or strata, varying the selection unit from vessel to trip, and changes in selection rates used to deploy observers and EM in the partial coverage category.

The flexibility offered by the ADP allows NMFS and the Council to achieve transparency, accountability, and efficiency from the Observer Program to meet its myriad objectives. The ADP process ensures that the best available information is used to evaluate deployment, including scientific review and Council input, to annually determine deployment methods. The Observer Program is accountable to operate within annual financial constraints that are dependent on the amount of fee revenue collected from groundfish and halibut landings in the prior year and the anticipated future costs of monitoring and fishing effort.

Cost Efficiencies Analysis

At the October 2019 Council meeting, the Council recommended an increase in the observer fee percentage from 1.25 percent to 1.65 percent for the Partial Coverage Observer Program and dovetailed that recommendation with continued development of mechanisms to improve cost efficiency in the program as its highest priority moving forward. Specifically, the Council requested work to focus on:

- *Pelagic trawl EM combined with shoreside sampling;*
- *Integrated monitoring plan for fixed gear that combines EM, shoreside sampling, and at-sea observer coverage as needed (e.g., consider whether the 15% hurdle is still the appropriate baseline level for observer coverage in combination with EM coverage; develop average weight protocols to support the use of EM);*
- *Optimizing the size and composition of the fixed gear observed and EM fleets, taking into account both cost priorities and data needs for average weights and biological samples (including consideration of expansion of the zero-coverage pool to include vessels fishing from remote ports harvesting small amounts of fish).*

In January 2020, the Council's Partial Coverage Fishery Monitoring Advisory Committee (PCFMAC) reviewed a cost efficiencies work plan[1] that considered 6 potential options for improving cost efficiencies in the partial coverage program, including development of a pelagic trawl EM program.

Implementation of the pelagic trawl EM program was addressed through an Exempted Fishing Permit (EFP)[2] to evaluate the efficacy of EM systems and shoreside observers. The trawl EM program is designed to use EM for compliance monitoring, meaning that EM video data does not directly feed into catch accounting or stock assessments. Instead, catch accounting uses

[1] https://meetings.npfmc.org/Meeting/Details/1224

[2] The EFP application, permits, and reports can be found under the heading "Electronic Monitoring - Trawl Catcher Vessels" on the NMFS website: https://www.fisheries.noaa.gov/alaska/resources-fishing/exempted-fishing-permits-alaska

industry-reported data (verified through EM) and data collected by shoreside observers. Maximized retention ensures that unsorted catch will be delivered and available to be sampled by shoreside observers, allowing for non-biased data to be collected at the trip level by shoreside observers at the processing plant. The project was a collaborative process among project partners that included NMFS staff, EFP permit holders, EM service providers, video reviewers, and observer providers. The Council's Trawl EM Committee also met multiple times to review progress and provide recommendations to the Council. In October 2022, the Council took Final Action to implement EM on pelagic trawl pollock catcher vessels and tenders delivering to shoreside processors in the Bering Sea and Gulf of Alaska. The intended timeline is to continue the EFP in 2024 and implement the regulatory program in 2025 to ensure there is no gap between the EFP and the regulated program.

The purpose of this document is to address the Council's priorities for improving cost efficiencies in the partial coverage program and to outline the draft 2024 ADP that maintains a monitoring program that meets NMFS's data collection mandates. The overarching goal is to develop a fishery monitoring design that balances statistically rigorous data collection with minimizing the impacts on fishing operations while maximizing the amount of sampling conducted under a given budget. The total budget available for the partial coverage program is determined by the fee percentage and the resulting revenue from the fees that are collected. As such, this analysis focused on the cost per unit of monitoring as opposed to dynamic total annual cost of the program and the intent is to collect the best and most data for a given budget.

Chapters 3 and 4 provide an evaluation of the partial coverage category and evaluate several stratification methods (ways to divide the sample population of trips into groups, or strata) and allocation approaches (how much to sample in each stratum). This integrated evaluation of data collection methods (observers and EM) incorporates the goal of spending the limited, available funding more efficiently such that the most coverage (both EM and observers) is achieved for a range of budgets. The analysis evaluates the trade-offs between different monitoring designs, including:

- Relative per unit cost efficiency of each design
- Statistical efficiency of each design
- Relative impact on data quality (e.g. timeliness, ability detect rare events)
- Relative scalability of each design

Between April 2021 and May 2023, the Council's Fisheries Monitoring Advisory Committee (FMAC) and the PCFMAC have met multiple times and received updates on the cost efficiencies analysis. Through this process, there were additional cost efficiency ideas outside of deployment designs that did not involve stratification or allocation. These additional cost considerations are summarized in Chapter 5.

Chapter 6 will provide NMFS recommendations for the 2024 ADP. However, these are not yet presented in this version of the document. NMFS will consider input from the PCFMAC at their September meeting and finalize agency recommendations prior to the October Council meeting.

Table 1-1. Data types collected by at-sea observers, trawl EM with shoreside observers, and fixed gear EM. A green checkmark indicates that the data type is collected, a red x indicates that the data type is not collected, and blue arrows indicate that some aspects of the data type are collected.

Data Types Collected		At-sea Observers	Trawl EM + Shoreside Observers	Fixed Gear EM
Catch				
	Trip Characteristics (e.g. duration, total effort)	✔	✔	✔
	Haul Characteristics (e.g. location, effort, depth, gear performance)	✔	⇔	⇔
	Haul Level Species Composition - Counts	✔	✖	✔
	Haul Level Species Composition - Weights	✔	✖	✖
	Trip Level Species Composition - Counts	✔	✔	✔
	Trip Level Species Composition - Weights	✔	✔	✖
	Speciation of similar species (e.g. large red rockfish, king crabs)	✔	✔	✖
	Haul Specific Salmon PSC Enumeration	✔	✖	⇔
	Trip Specific Salmon PSC Enumeration	✔	✔	⇔
	USCG Marine Casualty Information	✔	⇔	⇔
Biologicals				
	Sex Length Data (fish and crab)	✔	✔	✖
	Pacific Halibut size and mortality assessment	✔	✔	✖
	Trip specific age structures (e.g. otoliths, scales, fin rays)	✔	✔	✖
	Trip Specific tissue for genetic analyses	✔	✔	✖
	Tagged organism information	✔	✔	✖
	Stomach samples (trophic interactions)	✔	⇔	✖
	Maturity information	✔	⇔	✖

Data Types Collected		At-sea Observers	Trawl EM + Shoreside Observers	Fixed Gear EM
Protected Species				
	Marine mammal injury and mortality	✔	⇔	⇔
	Marine mammal tissue (genetics, tropic information, contaminants)	✔	✖	✖
	Marine mammal interactions (non-lethal; non-injury)	✔	✖	⇔
	Marine mammal sightings	✔	✖	✖
	Verify use of seabird avoidance methods	✔	n/a	✔
	Seabird mortality (catch by gear)	✔	✔	✔
	Seabird mortality (vessel interactions)	✔	⇔	⇔
	ESA-listed seabird carcass	✔	⇔	✖

Table 1-2. Sampling strata and selection pools in the partial coverage category from 2013 to the present. The partial coverage selection rates set through the Annual Deployment Plan are noted and the realized coverage rates evaluated in each Annual Report are noted in parentheses. PreIm = Pre-implementation, prior to a fully regulated program; CP = catcher/processor vessel; CV = catcher vessel; GOA= Gulf of Alaska; BS = Bering Sea; H&L = hook-and-line gear; LOA = vessel length overall.

<table>
<tr><th>Year</th><th colspan="5">Observer Trip Selection
Trip-selection across all ports
Observer coverage required on all randomly selected trips</th><th>Port-based Trip Selection*</th><th colspan="2">Fixed-Gear EM trip selection pool
EM required on randomly selected</th><th>Trawl EM</th><th>Observer vessel selection pool</th><th colspan="2">No selection pool
Observer coverage not required</th></tr>
<tr><td>2023</td><td colspan="2">Trawl: 23%</td><td>H&L: 18%</td><td colspan="2">Pot: 17%</td><td rowspan="2"></td><td colspan="2" rowspan="8">Fixed gear
(H&L and Pot)
EM: 30%</td><td rowspan="7">100% at-sea EM; 33% shoreside monitoring in GOA and 100% shoreside monitoring in BS</td><td rowspan="9"></td><td rowspan="9">Vessels <40’ LOA and Jig gear</td><td></td></tr>
<tr><td>2022</td><td colspan="2">Trawl: 29.7% (29)</td><td>H&L: 19% (14.6)</td><td colspan="2">Pot: 17.5% (18.1)</td><td rowspan="8">EM Innovation Research 2-4 vessels</td></tr>
<tr><td rowspan="2">2021</td><td colspan="2">Sep. 1 - Dec. 31:
Trawl: 21% (28.2)</td><td>H&L:18% (17.2)</td><td colspan="2">Pot: 18% (20.5)</td><td>Deployment in all ports</td></tr>
<tr><td colspan="5">Jan. 1 - Aug. 31: Limited waivers due to COVID-19</td><td>Deployment in 13 ports</td></tr>
<tr><td rowspan="3">2020</td><td colspan="5">Jul. 1 – Dec. 31: Limited waivers due to COVID-19</td><td>Deployment in 13 ports</td></tr>
<tr><td colspan="5">Mar. 26 - Jun. 30: Waivers issued due to COVID-19</td><td>Deployment in Kodiak only</td></tr>
<tr><td colspan="2">Jan. 1 – Mar. 25:
Trawl: 20% (22.4)</td><td>H&L: 15% (13.4)</td><td colspan="2">Pot: 15% (15.5)</td><td>Deployment in all ports</td></tr>
<tr><td>2019</td><td>Trawl: 24% (25.2)</td><td>Trawl Tender: 27% (35.7)</td><td>H&L: 18% (17.6)</td><td>Pot: 15% (14.0)</td><td>Tender Pot: 16% (29.5)</td><td rowspan="2"></td><td rowspan="2"></td></tr>
<tr><td>2018</td><td>Trawl: 20% (20.3)</td><td>Trawl Tender: 17% (35.0)</td><td>H&L: 17% (15.5)</td><td>Pot: 16% (15.5)</td><td>Tender Pot: 17% (29.0)</td><td>H&L EM: 30%</td><td>Pot EM PreIm: 30% (not used in catch accounting)</td></tr>
</table>

<table>
<tr><th rowspan="2">Year</th><th colspan="6">Observer Trip Selection

Trip-selection across all ports
Observer coverage required on all randomly selected trips</th><th rowspan="2">Port-based Trip Selection*</th><th rowspan="2">Fixed-Gear EM trip selection pool
EM required on randomly selected</th><th rowspan="2">Trawl EM</th><th rowspan="2">Observer vessel selection pool</th><th rowspan="2" colspan="2">No selection pool
Observer coverage not required</th></tr>
<tr></tr>
<tr><td>2017</td><td>Trawl: 18% (20.7)</td><td>Trawl Tender: 14% (18.8)</td><td>H&L: 11% (12.0)</td><td>H&L Tender : 25% (0)</td><td>Pot: 4% (7.7)</td><td>Tender Pot: 4% (5.3)</td><td rowspan="5"></td><td rowspan="5"></td><td rowspan="5"></td><td rowspan="3"></td><td rowspan="4"></td><td>EM PreIm ~90 vessels</td></tr>
<tr><td>2016</td><td colspan="2">Trawl: 28% (28.0)</td><td colspan="2">H&L: 15% (15.0)</td><td colspan="2">Pot: 15% (14.7)</td><td>EM PreIm 60 vessels</td></tr>
<tr><td>2015</td><td colspan="3">Large Vessel: 24% (23.4)
Trawl CVs, Small CPs,
H&L/Pot CVs ≥ 57.5’</td><td colspan="3">Small Vessel: 12% (11.2)
H&L/Pot CVs >40’ and <57.5’</td><td>EM PreIm 12 vessels</td></tr>
<tr><td>2014</td><td colspan="6">All Trawl CVs and H&L/Pot vessels ≥ 57.5’ LOA: 16% (15.1)</td><td>H&L/Pot CVs >40’ and <57.5’: 12% (15.6)</td><td>Voluntary EM</td></tr>
<tr><td>2013</td><td colspan="6">All Trawl CVs and H&L/Pot vessels ≥ 57.5’ LOA: 14.5% (14.8)</td><td>H&L/Pot CVs >40’ and <57.5’: 11% (10.6)</td><td colspan="2">Vessels <40’ LOA and Jig gear</td></tr>
</table>

*Observer coverage on randomly selected trips in specific ports. This protocol was implemented in response to the COVID-19 pandemic when travel and lodging conditions in specific ports allowed observers to meet and maintain applicable health mandates for deployment into the commercial fisheries.

2. Budget and Cost Assumptions

Budget

Observer and EM coverage in the partial observer coverage category is funded through a system of fees[3] based on the ex-vessel value of groundfish and halibut landings. Landings accruing against an IFQ allocation or a Federal Total Allowable Catch (TAC) for groundfish from vessels in the partial coverage category are assessed a fee using standard ex-vessels prices multiplied by the landed catch weight of groundfish and halibut. Prior to 2021, the fee percentage was 1.25% of the ex-vessel value of landings and as of 2021, the fee percentage is 1.65%.

Table 2-1 presents both the total fees assessed under the fee percentage that was used in that year and the fees that would have been assessed had a 1.65% fee percentage been in place. For years since 2021, those two numbers will be the same. Between 2013 and 2022, the total fees that were either assessed or would have been assessed under a 1.65% landing fee averaged $4,424,474, with a minimum of $3,169,843 in 2021 and a maximum of $5,611,917 in 2013. The projected fee revenue for 2023 is $4.71M[4]. In recent years, FMA has also received approximately $750K in Congressionally allocated funds separate from the landing fee, but those are not guaranteed in future years.

Based on this range of funds available in past years, all sampling designs in this analysis were run with $3.5M, $4.5M, and $5.25M budgets. A $3.5M budget represents a scenario where fee revenue is low (e.g., low market prices) or expenses outpace revenues. A $4.5M budget represents a scenario with recent fee revenues. A “high” budget of $5.25M assumes that $750K in additional (e.g., federal) funds are available.

Cost Assumptions

The partial coverage monitoring program has three monitoring methods: at-sea observers, fixed-gear electronic monitoring, and at-sea compliance electronic monitoring with shoreside observers to sample offloads. To estimate the costs of monitoring, cost models were constructed for each monitoring method (Figure 2-1). Each model incorporates the best information available and different assumptions to reflect specific cost structures (fixed and variable costs) and known patterns of economy of scale. All of the cost models consider monitoring costs in how they pertain to the partial coverage monitoring program budget (e.g., costs for trawl EM trips in the Bering Sea are excluded because those trips are in full coverage) and assume that all future monitoring expenses will be supported by the fee revenue (including equipment replacement and maintenance costs).

[3] A summary of the observer fee is available at: https://media.fisheries.noaa.gov/2022-01/observerfees_1.pdf
[4]https://meetings.npfmc.org/CommentReview/DownloadFile?p=2934117d-9379-4080-ab99-4d5e6733a58f.pdf&fileName=2023%20Observer%20Fee%20update.pdf

At-sea Observers

Detailed monitoring expenses were compiled from internal reports for years 2017-2022 including expenses for sea days (where an observer is assigned to a vessel) and travel, and were then inflation-adjusted to 2024 dollars. Travel costs scaled linearly with the number of days purchased and therefore the cost per sea day purchased was assumed to vary similarly. The number of sea days purchased is dependent on the sample rates allocated to strata requiring at-sea observers.

To estimate the cost per sea day, a linear model was used to represent the relationship between the total sea day costs as a function of total sea days purchased using the existing partial coverage observer contract's sea day costs. The existing partial coverage observer contract is currently structured such that a minimum of 2,000 'base days' are purchased and additional 'option days' can be purchased at a lower rate. The cost-per-sea day therefore decreases as more days are purchased. However, the future partial coverage observer contract (to be enacted in August 2024) will allow NMFS to purchase fewer than 2,000 days. The cost-per-sea day linear model is therefore assumed to be valid in instances when fewer than 2,000 sea days are purchased (i.e., cost-per-sea day increases if fewer total sea days are purchased). None of the designs allocate more than 2,000 at-sea observer days under any of the budget scenarios in this analysis. NMFS does not have other information about the upcoming partial coverage observer contract to inform any other aspects of the cost model such as day costs or breakpoints.

The cost model can be written as:

$$\text{At-sea Observer costs} = (\text{sea days}) \times (\text{sea day rate} + \text{travel rate}),$$

where the sea day rate is modeled as a linear function of sea day costs of the current partial observer coverage contract:

$$\text{Sea day rate} = \text{intercept} + (\text{sea days}) \times (-\ \text{slope})$$

The intercept is positive and the slope is negative, leading to a lower sea day rate as more sea days are purchased. In other words, the slope represents the per-additional-day cost savings of a sea day. Exact values of the model's intercept and slope are not given to preserve confidential business information about the existing partial coverage observer contract.

At-sea Fixed-Gear Electronic Monitoring

Fixed-gear EM costs, vessel counts, and review days were compiled from Annual Reports (2015-2021; see NMFS 2022 for example). Costs were inflation-adjusted to 2024 dollars and separated into fixed costs (equipment install and maintenance) and variable costs (video review). Fixed costs scale with the number of vessels in the fixed-gear EM vessel pool. The fixed cost rate, or the average cost per vessel per year, was calculated as the total of the fixed costs divided by 172 vessels and 7 years, and was estimated as $5,679.90 per vessel per year. Variable costs scale with the number of sea days reviewed and is a function of the number of trips selected for monitoring (i.e., the sample rate). The cost per review day was calculated as the total review costs divided by the total review days between 2018 and 2021, a 4-year span where both review costs and number of days reviewed was reported, and was estimated as $150.32 per sea day. Review costs assume three EM video reviewers (the same number as in 2022 and 2023).

Although equipment installation was funded by grants in the past, the cost model assumes that future equipment installation, replacement, and maintenance costs will be funded by the fees from the partial coverage program.

The fixed-gear EM cost model is defined as:

Fixed-gear EM costs = (# vessels x $5,679.90) + (sea days x $150.32)

The cost model in this analysis assumed that the size of the fixed-gear EM vessel pool was 172 vessels, the number that participated in 2022. Note that 179 fixed gear vessels participated in 2023; therefore, fixed-gear EM cost estimates are slightly underestimated. The vessel count in 2024 will be finalized in November and updated in the Final ADP.

Trawl EM (at-sea compliance monitoring and shoreside Observers)

The costs of the trawl EM program were compiled from the trawl EM analysis (NMFS and NPFMC 2022), inflation-adjusted to 2024 dollars, and were separated into fixed costs (service provider fees and overhead, equipment installation and maintenance) and variable costs (compliance monitoring review and shoreside observer sampling).

The fixed cost rate, which is a per-vessel rate, was assumed as the per-vessel cost of equipment ($14,496) amortized over 5-years and inflation-adjusted to $4,100.71. The per-vessel equipment maintenance cost was estimated as $275,391 divided by the 68 vessels participating in 2022, and inflation-adjusted to $4,746.04. Therefore, the total per-vessel-per-year rate was estimated as $8,846.75. However, the fixed costs were assumed to be funded by the partial coverage observer fee only for GOA-only participant vessels. Therefore, the total fixed costs is estimated as the per-vessel cost multiplied by the number of GOA-only vessels expected to participate, which is 39 vessels for 2024.

Variable costs were estimated for both compliance video review and shoreside monitoring. The cost per compliance review day was estimated as $116,611 for 4,882 review days, inflation-adjusted to $27.99 per review day. This rate is applied to the total expected number of days fished by all Trawl EM vessels on trips fished in the GOA. The cost per shoreside observer day was estimated using the midpoint of a low and high estimate, $775, inflation-adjusted to $908.22 per day. The number of plant observer days required was assumed to be reliably predicted by the number of trips fished by Trawl EM vessels in the GOA, where in 2022, 432 trips required 548 observer plant days with a sampling rate of 33.33%, resulting in a conversion factor of 3.8059 days per trip.

The trawl EM cost model is defined as:

$$\text{Trawl-EM costs} = (\#\ \text{GOA-only vessels} \times \$8{,}846.75) + (\#\ \text{sea days} \times \$27.99) + (\#\ \text{trips} \times \text{sample rate} \times 3.8059 \times \$908.22)$$

Table 2-1. Partial coverage observer landing fees assessed between 2013 and 2022, and their equivalent had a 1.65% fee percentage been in place.

Year	Fee percentage	Fees assessed	1.65% fee equivalent
2013	1.25%	$4,251,452	$5,611,917
2014		$3,458,716	$4,565,505
2015		$3,775,956	$4,984,262
2016		$3,769,758	$4,976,081
2017		$3,821,263	$5,044,067
2018		$3,407,658	$4,498,109
2019		$2,895,378	$3,821,899
2020		$2,469,241	$3,259,398
2021	1.65%	$3,169,843	$3,169,843
2022		$4,313,661	$4,313,661

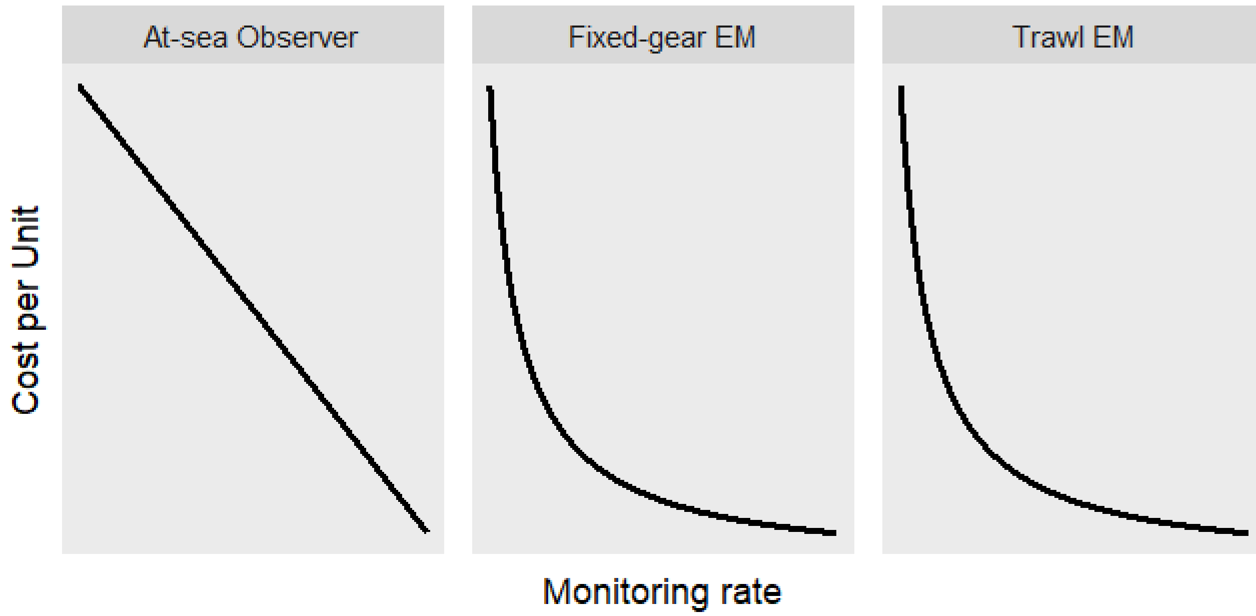

Figure 2-1. Per-unit monitoring costs (i.e., trips or offloads) as a function of monitoring rate for at-sea observers, fixed-gear EM, and trawl EM monitoring methods. The scales of both axes are intentionally masked and the y-axes are not aligned to discourage comparisons and preserve confidentiality of partial coverage observer contract costs. The cost of monitoring decreases for all monitoring methods the more the method is utilized. Both EM methods have high fixed costs (equipment purchases and maintenance) but relatively cheaper recurring costs (video review), hence per-unit costs decrease in a non-linear fashion with increasing monitoring rates.

3. Deployment Designs

Fundamental Elements to all designs

The effective design and components of successful fishery monitoring programs have been described previously (e.g., Cahalan and Faunce 2020, Cotter and Pilling 2007, Vølstad et al. 2014, ICES 2004). The key design elements common to these programs include randomized data collections over spatial and temporal scales (a probability sample), the collection of sufficient data, and the use of stratification and prespecification of sampling intensity to control precision of estimates, while also making efficient use of available funding (Cahalan and Faunce, 2020). To construct a monitoring design, several components need to be defined or identified, starting with the monitoring objectives. Because there are a large number of fisheries (i.e., managed species, stock assessments, monitored quotas), a single objective is hard to identify in the traditional sense of designing a program to minimize variance of catch estimates. However, we can develop a program that collects data required by a large number of data users as is outlined in the Observer Program data collection section of Chapter 1. NMFS' overarching goal is to design a deployment plan that can be used to monitor federally managed fisheries throughout the EEZ of Alaska in order to provide the best scientific information available (MSA National Standard 2) while adhering to the remaining National Standards.

Another fundamental element of a monitoring program's design is randomized sample collections, generally known as a probability sample. The basic elements of a probability sample are identification of the target population and sample frame, and a prescribed method for selecting sample units from which data are collected (e.g., Cochran 1977; Thompson 2012).

The target population is the population that we want to know about. For the ADP, the target population is all commercial groundfish and halibut fishing activity under federal jurisdiction in Alaska. Note that this includes trips taken by vessels in the zero selection fleet. The sample frame is the list of all discrete non-overlapping sampling units in the population. How sample units are defined will also define the sample frame. The sample frame should encompass the entirety of the target population and any discrepancies between the sample frame and target population will result in bias (undercoverage, the sample frame does not include portions of the population) or inefficient sampling (overcoverage, the sample frame includes sample units that are outside the target population). For the 2024 ADP, the sample frame is the list of all fishing trips that are available to be monitored and are included in the sampled strata; the zero selection stratum is not sampled (undercoverage). Because the sample frame does not include the entire target population, there is potential for bias in any estimates derived from data collected under the 2024 ADP.

Once the sample frame and any discrepancies between the sample frame and target population are identified, the population is divided into discrete strata where each sample unit must be assigned to one and only one stratum. Stratification is often used to decrease variability in the parameter being estimated, simplify logistics, or decrease costs. Sampling rates and methods cannot vary within a stratum but can differ between strata. The methods used to allocate sampling effort to each of the strata will reflect a particular sampling goal; for example, the allocation method used in 2023 sets a minimum baseline sample rate in each stratum (15%) to increase the probability of a monitored trip of a given gear type occurring in a NMFS area. In the 2023 ADP's allocation method, any additional sampling effort available is allocated to strata in order to decrease the overall between-trip variance in Chinook PSC, halibut PSC, and discards of groundfish and halibut. Probability methods are used to select samples; at-sea monitoring uses the ODDS system to randomly select fishing trips and observers use specified randomization methods to select deliveries to be monitored.

The sample design (sampling strata, sample frames and units, sample selection) links the sample to the underlying population, and estimators that use data collected under a particular design will incorporate that design into the estimation process to avoid bias. Hence, it is important to identify any operational barriers that may prevent a design from being executed as planned (e.g., if we can't reliably know if a vessel will use pot [POT] gear or hook-and-line [HAL] gear, stratifying by gear type may cause a mismatch between what is planned and realized). These differences between the sample design and the implementation of the design may increase the potential for biased estimates based on the data collected. Note that participation in EM is defined separately for fixed gear and trawl gear. The number of vessels participating in any future year will be determined through the ADP process.

Within a design, the ability to assess sub-populations of fishing activity will be higher when monitoring rates are higher due to the greater amount of data and increased probability of subpopulation data occurring in the sample. If data summaries are needed for smaller portions of the population (e.g., specific fishing areas or times), then sampling rates should be high enough to ensure that at least one (or more) fishing trips are monitored in that specific area. Hence, the goals of two novel allocation methods discussed below are to minimize gaps in monitoring coverage and increase the chances that data from all fisheries are collected.

Stratification

Stratification can be used to isolate portions of the fisheries (sub-populations) that are of particular interest, to focus sampling on portions of the population where minimal fishing occurs (e.g., Aleutian Islands), to simplify implementation (e.g., sampling supervision based in local field offices), to control costs by decreasing logistical constraints (e.g., travel times), and in some situations, to control variance (Cochran, 1977).

Stratification of the sample frame requires that each sample unit exists in one and only one stratum (e.g., Cochran 1977, Thompson 2012). Additionally, assignment of sample units to a stratum must be based on characteristics that are known before the fishing trip or delivery occurs. In the case of Alaska commercial fisheries, strata could be defined by gear type used, FMP where fishing will occur, and the monitoring method to be used (EM or observers). Stratification is most effective when there are either large differences between strata in sampling methods or the variable of interest, or there are large differences in monitoring costs (Cochran, 1977). Sampling methods and sampling rates need to be consistent within each stratum, but can vary between strata. Hence, the monitoring method is a necessary component of the stratification definition.

In this analysis, several stratification definitions were evaluated based on the underlying structure of the fisheries. In addition to the current stratification (generally defined by gear type and monitoring method), we will be evaluating strata definitions that include FMP and those that accommodate trips that fish with multiple gear types and therefore cannot be reliably placed in only one of the current strata.

2023 (Current) Stratification

The current stratification definition has seven partial coverage strata defined by gear type fished and monitoring method (Table 3-1). For all strata monitored at sea using either observers or EM systems, the sampling unit is defined as the fishing trip. For strata where monitoring occurs at shoreside processing plants, the sampling unit is defined as the shoreside delivery of catch.

The *2023 (Current)* stratification has been defined in the ADPs since 2020 with slight changes implemented as new monitoring programs are established (i.e., trawl EM). While the 2023 stratification is effective logistically, it has two notable drawbacks. First, trips occurring in the fixed gear strata often fish with both hook-and-line and pot gears, violating the strata definition specifying that each sample unit occupies only one gear-based stratum. Secondly, these strata are relatively large and at the moderate to low sample rates afforded in the past few years, important portions of the fisheries are at risk of not being sampled. One notable example is the Aleutian Island pot cod fishery where no fishing trips were monitored in 2021. Similarly, there were few monitored trips in the BSAI sablefish fishery in 2020. Because of these shortfalls, stratification definitions that subdivide the current strata by splitting them regionally by FMP and that address the mixed-trip stratification violation are evaluated below. The 2023 (current) stratification definition will also be included in the final suite of monitoring designs to be evaluated.

Stratification incorporating FMP

There are several fisheries where stock assessment data needs are not currently being met; AI pot Pacific cod being a notable example. One issue with the current stratification definition is that in order to provide reasonable opportunity to collect samples from this particular small subset of a stratum, the sample rate must be elevated for the entire stratum. In most partial coverage strata, the majority of fishing effort occurs in the GOA with a smaller proportion in the BSAI. However, if the same sample rate is applied in both regions, because the BSAI has fewer trips, it will not only have fewer trips monitored but will also be much more likely than the GOA to have no monitored trips. By incorporating FMP into the stratification definitions, we can increase sampling rates in FMP areas with low fishing effort to increase the probability of monitoring some trips in those areas. Targeted sampling in the AI and BSAI would increase data collections to support these stock assessments. Based on 2022 fishing effort, there are very few trips occurring in multiple FMP areas (Figure 3-1), hence the majority of trips could be unambiguously placed in a single stratum. Because vessel operators know which FMP area they will be fishing in, FMP stratum assignment could occur when trips are logged into the ODDS system.

We used the binomial distribution to estimate the probability of not having any trips monitored in either a stratum (or a subpopulation within the stratum defined by FMP region, for example, probability of zero monitored trips in the AI), and the probability that fewer than 3 will be monitored.

Table 3-2 shows the number of trips in each stratum in 2022, probabilities of no trips being monitored and probabilities of fewer than 3 trips being monitored for the proposed stratification definitions under a 15% sample rate. For this example, all strata are sampled at the same rate, hence we can use Table 3-2 to evaluate the probability that a subpopulation (e.g., FMP area - gear type combination) of the stratum will be sampled. Under the 2023 (current) stratification, AK-wide, all strata are relatively large and will likely have more than 3 trips monitored over the course of a year. If we consider how likely we are to have samples from the BSAI and GOA separately, we see that in the EM HAL stratum there is a 12% chance of fewer than 3 trips monitored in the BSAI. If we further subdivide the population, separating the AI from the BS, there is an 86% chance of fewer than 3 trips being monitored, and a 23% chance of no data being collected in the AI.

Stratification at finer scales allows for increased sample rates in smaller strata; however, strata with very few fishing trips will need to be sampled at high rates to ensure some trips are sampled. In spite of these higher rates, the overall sample size can be expected to be relatively small. Referring back to the example above where we want to ensure at least one monitored EM HAL trip in the AI, we can sample the EM HAL AI stratum (stratify by FMP as well as monitoring method and gear type) at 58%, to expect an average of 5 trips, and a high probability of at least one monitored trip. Alternatively, we can increase sample rates in the EM HAL BSAI stratum to 50% to achieve the same outcome (a high probability of monitoring at least one trip).

This highlights the benefit of using FMP area to define strata; if we stratify by monitoring tool (at-sea Observers, Fixed-gear EM and Trawl EM), gear type (HAL, POT, TRW), and FMP, we can increase sampling rates in FMP areas that have few fishing trips, increasing the chances that we have monitoring data from those areas. For example, if we didn't stratify by FMP but also we wanted to have at least one EM HAL trip monitored in the AI, we would need to sample the entire EM HAL stratum (722 trips) at a rate of approximately 58% (418 trips); to monitor 3 or more trips, we would need a sample rate of approximately 79% (570 trips). However, if we expect to monitor at least one trip, we could sample an EM_HAL BSAI (32 trips) strata at a rate of 50% (16 samples); to monitor at least one trip, and a rate of 72% (23 trips) to monitor 3 or more trips.

Noting that there are few trips in the AI for some gear types and monitoring methods, stratification that incorporates a BSAI and GOA component is recommended over separation of the AI into a separate stratum. *S*tratification definitions that include BSAI/GOA FMP components will be included in the suite of final designs that are evaluated. Stratification by BSAI and GOA within each gear type and monitoring method will decrease the chances of spatial-temporal gaps in monitoring coverage. Separating the BS from the AI will not be considered in this analysis because it creates strata with so few trips that such strata may not exist perennially.

Fixed Gear Stratification

Approximately 15% of observer-pool and 20% EM-pool fixed gear trips use both HAL and POT gear on a single trip in 2022 (Fig. 3-2). These trips cannot be unambiguously placed into strata defined by the use of a single gear type, and as a result, stratification assumptions are violated and standard statistical methods are prone to estimation errors. There are two options to correct this issue: 1) create a separate stratum for mixed-gear trips or 2) combine all trips that fish with either HAL, POT, or both into a single stratum.

To evaluate different strata definitions, we used the binomial distribution to estimate the probability of not having any trips monitored in either a stratum or a subpopulation within a stratum again defined by FMP (i.e., probability of zero monitored trips in a region), and the probability of 3 or more trips will be monitored. For each gear type, strata were constructed from 2022 fishing effort data, separating mixed fixed gear trips (both HAL and POT on the same trip) and mixed trawl gear trips (non-pelagic and pelagic) into separate strata. Trips where a single gear type was used (HAL, POT, non-pelagic, or pelagic trawl) were separated into gear-specific strata (Table 3-3).

Stratification that includes mixed-gear strata but does not incorporate FMP results in more than 100 trips in each stratum (AK-wide, Table 3-3). Incorporating FMP into the stratification definition creates several strata with 50 or fewer fishing trips and high probabilities of monitoring no trips or fewer than 3 trips (BSAI vs GOA, Table 3-3). Including AI as a separate stratum from BS exacerbates the low population size problem within some strata (BS vs AI vs GOA blue *N* values in Table 3-3). Stratification where mixed-gear trips are separate from single-gear trips does not appear to be compatible with stratification by FMP due to the small stratum sizes it creates. Moreover, due to annual changes in fishing effort, some strata are so small that they may not actually contain *any* trips. Such inconsistency between years would greatly affect processes that require the use of past fishing to predict the future, including fishing effort predictions and allocation.

However, by instead combining the HAL, POT, and mixed-gear trips within a monitoring method into a single fixed-gear stratum (Table 3-4), low stratum sizes are only evident in two AI strata (blue *N* values in Table 3-4). If data are necessary from EM fixed gear or Observed trawl trips in the AI to meet data users analytic needs (e.g., AI-specific stock assessments), then those two additional strata could be included in the final stratification definitions and monitored at higher sample rates. When strata are defined when all fixed-gear trips are combined but also split by BSAI and GOA, all strata are adequately large to provide a very high likelihood that at least three samples will be collected under a 15% selection rate.

Stratification definitions that include combining trips that fish with HAL, POT, or both gears into a single stratum will be included in final monitoring designs for evaluation. There is a high proportion of trips that fish both gear types and an expectation that this behavior will persist in the future. To maintain statistical integrity without creating strata with few fishing trips, we recommend creation of combined fixed gear strata for each monitoring method.

Stratifications Evaluated

The stratification definitions that will be included in the final designs are presented in Table 3-4. As noted above, creation of separate AI strata offers few advantages and is not pursued further. AI strata would contain few trips and by sampling a stratum defined with BSAI FMP at higher rates would ensure some AI data are collected without necessitating increased sampling on the larger AK-wide strata. Hence, *FMP* stratification definition includes the BSAI and GOA. Similarly, creation of separate strata for mixed HAL and POT trips is difficult to implement unambiguously, and if FMP is included in the stratification definition, creates strata with few fishing trips. For these reasons, we did not evaluate the *Mixed HAL-POT FMP* stratification definition further.

Allocation Methods

There are a variety of allocation methods that are used to distribute available sample units to individual strata. Each method is designed to achieve different sampling objectives ranging from minimization of variance of estimated parameters to minimizing costs of sampling to decreasing the potential for spatio-temporal gaps in data collections.

In many sampling situations, the goal is to minimize the variance of a single estimated parameter *(i.e., optimal allocation,* Cochran 1977) under the constraint that the sample sizes for each stratum sum to the total sample size afforded. These allocation methods can be expanded to include the stratum-specific costs of sampling and thereby minimize both variance of the estimate and the costs. However, monitoring programs rarely have a single objective and as a result, many estimated parameters are derived from the collected data. While there are tools that "optimize" allocation for multiple objectives (e.g., compromise allocation, Cochran 1977), these focus on setting monitoring levels to achieve a specified coefficient of variation on estimates of bycatch, either for a single species or a limited number of species. Simultaneously minimizing the variance of a suite of parameters with different underlying distributions (e.g., bycatch of different species) is difficult since allocation to minimize variance of one parameter (species) may have a deleterious effect on another parameter.

For large scale monitoring programs, there are often many multiple parameters being estimated (e.g., bycatch for different species) and data are used in a wide variety of analyses. In addition, when novel fisheries issues arise, data that had previously been low priority can become vitally important (e.g., deep sea corals). Monitoring programs need to employ allocation methods that ensure data utility is high regardless of specific analyses because how data are to be used by researchers and anticipating future data uses is not always known . One approach is to ensure that data are available on a relevant spatio-temporal scale and that all fishing activities are represented in the sample. Within a design, the ability to assess sub-populations of fishing activity (e.g. specific fisheries) will be higher at higher monitoring rates due to the increased probability of subpopulation data occurring in the sample and the larger number of monitored trips (more data). If data summaries are needed for smaller portions of the population (e.g., specific fishing areas or times), then sampling rates should be high enough to ensure that at least one (or more) fishing trips are monitored in that specific area.

To that end, two novel allocation methods are included in this evaluation; both of which are designed to minimize data gaps. The first method, the *Cost-weighted Box* allocation method, also prioritizes monitoring in low-cost strata while the second method, the *Proximity* allocation method, protects against low sample size (few monitored trips) within a stratum.

In addition to the allocation methods that are evaluated annually in the ADPs (NMFS 2022), we evaluate a generalized version of the current method, an equal rates allocation, and two novel methods.

Equal Rate Allocation

With this allocation method, each sampling stratum is sampled at the same rate, and as a result the distribution of samples is *proportional* to the size of the strata (*i.e.,* the number of monitored trips is proportional to the number of trips in the strata). While all trips have equal probability of being included in the sample, the number of trips selected in each stratum will vary depending on the total number of fishing trips.

This allocation method is frequently used when little is known about the population being sampled and strata definitions are based on cost and logistic concerns. For example, without data from previous monitoring, allocation methods that minimize variance or costs, or that decrease the potential for data gaps have no computational basis. This type of allocation is frequently used in pilot studies to collect data used to develop more complex sample designs for future implementation. Since all strata have the same sampling rate, strata can be combined in the analysis stage without the need to weight the data or use stratified statistical methods.

Under this approach, all strata including the EM strata, are sampled at the same maximal rate afforded under the given budget scenario. In the GOA trawl EM stratum, the number of participating vessels and the selection rate of deliveries is set as close to the at-sea trip selection rate for other strata as a systematic random sample will allow, and sampling occurs shoreside (random selection of deliveries are sampled).

The previous 3 years of fishing effort data are used to predict trip lengths, costs, and total trips that will occur in future years. These effort projections are used to calculate final rates once sample allocation is completed (e.g., target number monitored trips/total projected trips).

Equal Rates (*Proportional)* allocation methods are included in the final suite of designs evaluated.

Status Quo (Baseline 15% minimum plus optimization)

This objective of *Status Quo (Baseline 15% plus optimization)* allocation method is to monitor the at-sea observed strata at a 15% rate and then allocate any additional sampling effort to minimize the variance of the at-sea discards of groundfish, halibut PSC, and salmon PSC. The algorithm that allocates these additional samples only applies to at-sea observed strata; EM monitoring rates are set by policy.

In this design, sampling rates for partial coverage observer strata are able to differ from one another and rates are set according to methods most recently described in the 2023 ADP (NMFS 2023). Trawl EM is assumed to be a regulated program, and both trawl and fixed gear EM sampling rates are set at the rates currently determined by policy and most recently described in the 2023 ADP (NMFS 2022); fixed-gear EM monitoring is set at 30% of trips and trawl EM is set at 33.33% of shoreside deliveries, with the required funds carved off the monitoring program budget before allocating remaining funds to monitoring the at-sea observer strata.

Each at-sea observer stratum's minimum rate is set such that there is 95% confidence level of obtaining at least a 15% sampling rate, given random selection. If these minimum rates are afforded by the observer strata, then the remaining funds are used to allocate sampling effort according to Cochran's compromise allocation method using the variance of groundfish discards, halibut PSC, and Chinook PSC (optimized metrics). If funds are insufficient to monitor the at-sea observer strata at the rates required to achieve a 95% confidence level of obtaining a 15% sample rate, but can afford >15%, then monitoring effort is allocated to strata to maximize the confidence of obtaining 15% under the given budget. If these minimum rates cannot be afforded, then the proportional allocation method is applied to the at-sea observer strata.

The *Status Quo* allocation scheme has been shaped by input from policy-makers who specified the choice of optimized metrics (groundfish discards, PSC halibut, PSC chinook salmon) used to allocate samples above the 15% baseline. If *optimized* samples are afforded, this allocation method is designed to minimize the combined variance of the optimization metrics.

It should be noted that at the budget levels evaluated in this analysis, this allocation method did not afford enough samples to obtain the 15% minimum sample rate for observed strata and therefore reverted to equal (proportional) allocation to the at-sea observer strata.

The *Status Quo* (*Baseline 15% plus optimization*) allocation method is included in the final suite of designs evaluated.

EM Integrated Baseline 15% plus optimization

This allocation scheme is designed to integrate the EM strata into the *Status Quo* allocation method by discontinuing the policy that sets selection rates for trawl and fixed-gear EM strata and the associated required set-aside of funds. Sampling effort is allocated to *all* strata such that each stratum has a 95% probability of obtaining a 15% sample rate, with remaining sampling effort allocated to minimize the combined variance of groundfish discards, PSC halibut, and PSC chinook salmon. Note that the EM strata are allocated sampling effort under the same allocation algorithm as observed strata.

As above, each stratum's minimum rate is set such that there is 95% confidence level of obtaining at least a 15% sampling rate, given random selection. If these minimum rates are afforded for all strata, then the remaining funds are used to allocate sampling effort according to Cochran's compromise allocation method using the variance of groundfish discards, halibut PSC, and Chinook PSC (optimized metrics). If the strata cannot achieve the rates required to achieve a 95% confidence level of obtaining a 15% sample rate, but can afford >15%, then all strata are allocated to maximize the confidence level of obtaining a 15% monitoring rate. If the funds are insufficient to allocate sampling effort to strata at a 15% monitoring rate, then the *Equal rates* allocation method is applied to all strata.

It should be noted that at the budget levels evaluated in this analysis, this allocation method did not afford enough samples to obtain the 15% minimum sample rate and therefore reverted to *equal rates* allocation. Although this allocation scheme has merit in that sample rates are not set by policy, it was ultimately *excluded from this analysis for budgetary reasons*.

The *EM Integrated Baseline 15% plus optimization* allocation method is not included in the final suite of designs evaluated.

Cost-weighted Box Allocation

The objective of the *CWB* allocation method is to maximize the proportion of boxes monitored or near monitored boxes while penalizing strata with high monitoring costs. A weighting factor is used to distribute the available sample resources to strata. This allocation method applies to all sampled strata (i.e., does not apply to zero selection stratum); the EM strata are allocated sampling effort under the same allocation algorithm as observed strata.

The *Cost-weighted Box* (*CWB*) allocation method allocates greater sampling effort to strata that are more likely to have gaps in monitoring coverage that may result from random selection of trips that are widely dispersed within the stratum. This method distributes sampling effort to reduce the probability of spatiotemporal gaps in monitoring coverage due to randomization while prioritizing monitoring methods with lower per-sample unit (trip or delivery) costs.

Gaps occur when a predefined spatiotemporal unit of fishing effort is not expected to contain a monitored trip or have a neighboring monitored trip under a specified monitoring rate. These spatio-temporal units, or "boxes" are defined as spatial hexagonal cells 200 km across and temporal blocks 1-week in length (Fig. 3-3). Under this definition, trips are allowed to span multiple boxes in both time and space (based on the landing reports). Moreover, the "neighborhood" of a box includes the trips in immediately adjacent spatial or temporal boxes, hence, the overall extent of the neighborhood of a box is 600 km across and 3 consecutive weeks. The development of the box size definition is presented in Appendix A. For any given monitoring rate, boxes containing a greater number of fishing trips have a higher probability of being monitored, and the neighborhood of the box will also have a higher probability of containing a monitored trip.

The *CWB* allocation method allocates sampling effort to strata to minimize the proportion of boxes without data while prioritizing monitoring in strata with lower per-trip monitoring costs. For a given sample rate, strata with fishing trips that are distributed widely in space and time (e.g., many boxes and each with few trips) are more likely to have a greater proportion of boxes with unmonitored neighborhoods. In contrast, given the same sample rate, strata with highly concentrated fishing effort in time and space (e.g., few boxes, each with many trips) will have a lower proportion of boxes with unsampled neighborhoods.

Note that this allocation method does not allocate samples in order to get data from all boxes and hence, it is not allocating sample effort to ensure we have data from any particular box. It is also important to note that the proportion of boxes that are expected to have unmonitored neighborhoods is not minimized per se, but rather sample effort is distributed in a manner that reduces gaps most effectively.

The probability that we do not have any data in a neighborhood, $\widehat{A}_b$, is estimated using the binomial approximation of the hypergeometric distribution:

$$\hat{A}_b = (1 - r_h)^{t_{G_b}}$$

where box *b* is defined as the cell of interest. G_b defines the neighborhood adjacent to "box" *b* defined as the cell of interest and the adjacent 20 cells (6 spatial cells in the same week and 7 cells in the week prior and 7 in the week after = 21 cells total), t_G is the number of trips in a neighborhood, and r_h is the initial (assumed) sample rate used to estimate the probability that a hexagon is unmonitored.

Using this, we calculated the expected proportion of boxes that will not be monitored given a sampling rate, $\widehat{P}_h$, as the average across all cells of the probability of having no data:

$$\hat{P}_h = \frac{\sum_{b=1}^{B_h} \hat{A}_b}{B_h}$$

where again, b indexes the hexagons in stratum, and B_h is the total number of hexagons in the stratum. Figure Methods-cwb1 shows the decrease in the proportion of boxes without monitored trips in their neighborhoods with increasing sample rate. The rate of decrease varies across strata depending on the distribution of fishing effort within each stratum; strata with more dispersed fishing (e.g., EM Pot stratum) will initially decrease in P_h slower than those with more aggregated fishing (e.g., EM Trawl stratum) (Fig. 3-4).

In addition to reducing spatiotemporal data gaps, this method allocates higher sampling effort to strata with lower costs per trip. Each stratum's cost per trip depends on several factors: the monitoring method it uses, the monitoring rate used (the cost per trip decreases with higher rates, see Figure 2-1 in 'cost assumptions'), and the overall monitoring budget.

The final CWB index balances the average proportion of boxes without neighboring monitored trips by the stratum size, N_h and the inverse of stratum-specific costs, C_h. This index is also standardized to the sum across strata of all indexed so that W_h, is constrained between 0 and 1.

$$W_h = \frac{N_h \frac{\sum_{b=1}^{B_h} \hat{A}_b}{B_h C_h}}{\sum_{h=1}^{H} N_h \frac{\sum_{b=1}^{B_h} \hat{A}_b}{B_h C_h}} = \frac{N_h P_h C_h}{\sum_{h=1}^{H} N_h C_h}$$

This *CWB* index is multiplied by the total number of trips that can be monitored under the specified budget to generate the number of trips expected to be monitored for each stratum; the stratum-specific monitoring rate is this monitored number of trips divided by the total trips in the stratum. For a predetermined sample rate, different strata will have different sample rates, and the total cost is the product of number of trips in the stratum, N_h, the monitoring rate, r_h, and the stratum-specific cost per trip c_h, summed over the $h = 1, \ldots H$ strata:

$$Cost = \sum_{h=1}^{H} N_h * r_h * c_h$$

Because the expected proportion of boxes without sampled neighborhoods and cost per trip varies with allocated monitoring rates (algorithm output), the algorithm for computing the *CWB* index relies on an iterative process.

In the initial iteration, each stratum's P_h and cost per trip C_h, are calculated with a 15% monitoring rate (i.e., r_{0h} = 0.15) and used to estimate W_h . These estimated W_h are used to estimate updated monitoring rates which are compared to the previous iteration's rate (initially 0.15). If the updated rate and the rate from the previous iteration differ, the midpoints between the rates is used as the new assumed rate for each stratum in a second iteration. This process is repeated until the previous and current iteration rates are close or converge to the same rate, indicating that further iterations are unnecessary. This completes the allocation process.

The *Cost-weighted Box* allocation method is included in the final suite of designs evaluated.

Proximity Allocation

The objective of the Proximity allocation method is to maximize the proportion of trips that have monitored neighbors while controlling for low stratum-specific sample sizes. This method is designed to spread sampled trips throughout the fisheries to increase the proportion of trips that are sampled or near a sampled neighbor and to be consistent between strata within a specified budget, while also protecting against small sample sizes within a stratum. This allocation method applies to all sampled strata (i.e., does not apply to zero selection stratum); the EM strata are allocated sampling effort under the same allocation algorithm as observed strata. In this allocation method, we use the box and neighborhood definitions used in the *CWB* allocation, however, instead of minimizing the *proportion of boxes* with no monitored trips in their neighborhoods, the goal is to maximize the *number of trips* with neighboring monitored trips. The proximity index is based on the proportion of unmonitored trips that are expected to have a monitored neighbor; hence, as sampling rate increases, the proximity index also increases. The proximity index is a function of the available budget, each stratum's monitoring cost and size (total number of trips or offloads), sample rate, and spatiotemporal distribution of fishing effort. Strata with clustered fishing effort will achieve a specified proximity index at a lower sample rate than strata with more diffuse fishing effort; more samples are allocated towards strata with trips that are more spread out in space and time.

Proximity Index

As in the CWB method, we use the binomial approximation to the hypergeometric distribution to generate the probability that there are no monitored trips in the neighborhood of box *b*, $\widehat{A}_b$. The expected number of trips that have neighbors is the sum of the number of trips in the neighborhood, w_b, multiplied by the probability that one or more of those trips are sampled, $\left(1 - \widehat{A}_b\right)$.

As previously described, trips are allowed to span multiple boxes, and contribute equally to each box (e.g., a trip that crosses three boxes is counted as 0.33 trips in each box).

The proximity index, $\widehat{T}$, is the average of the expected proportion of trips with monitored neighbors averaged over the b = 1, …..B boxes in the stratum. For a given budget, we can maximize the amount of interspersion over all strata, essentially keeping the proximity index constant across strata while increasing sample rates until reaching the budget cap.

$$\hat{T} = \frac{\sum_{b=1}^{B} w_b * \left(1 - \widehat{A}_b\right)}{B}$$

The proximity index is useful for prioritizing the allocation of samples to highly spatiotemporally dispersed strata. However, strata with highly concentrated fishing effort and relatively small stratum sizes were allocated a small portion of the total sample amount. For these strata, a large proportion of unmonitored trips are located near monitored trips even at low sample rates, and allocation based solely on this index can result in small sample sizes for these strata. Since variance is a function of sample size, these small sample sizes can lead to catch estimates with high variability. In addition, estimated length and age composition data that drive some stock assessments will be sparse, leading to stock assessment harvest recommendations with higher uncertainty.

Small Sample Size Buffering

To buffer against low sample sizes (numbers of monitored trips) within a stratum that can occur when using the proximity index to allocate sampling effort, we incorporated the variance scaling used to estimate the variance of an estimated parameter (such as the sample mean). The variance of the sample mean is a function of the base (population) variance, the sample size, and the proportion of the population that is sampled.

All populations have a base variance, Eq (); the variability in a measured parameter (e.g., length) between all sample units (both in the sample and unsampled). Note that we are not summing only over those sample units that were sampled, but all samples in the population (i.e., i = 1 to N rather than i = 1 to n). For the ADP, the population variance is the between trip variance over all trips in a stratum and will be different for different species (years, gear types, *etc.*)

$$Var(x) = \frac{\sum_{i=1}^{N}(x_i - \overline{x})^2}{N - 1}$$

The estimated variance of the mean (or other parameter of interest) has two additional terms in addition to the population variance: the finite population correction factor (FPC, *(N-n)/N*) and the sample size (1/*n*)

$$V\hat{a}r(\overline{x}) = \frac{(N-n)}{N}\frac{1}{n}\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{(n-1)}$$

As the sample size increases, we know more about the population, and hence estimates will have less variance. The estimated variance will decrease with increasing sample rates until all sample units are included in the sample (sample rate = 100%), at which point we have a census of the population and there is no variance. Similarly, in addition to the variance savings we gain by increasing sample size, as we sample a larger and larger portion of the population, our uncertainty about the estimate decreases further.

We can see the impact of these two components, the sample size (Fig. 3-5, left panel) and the FPC (Fig. 3-5, right panel), on variance by plotting against the sample rate.

The FPC and sample rate combine to form a single variance scaling factor, *F*, (see also Fig.methods-proximity:2 below).

$$F = \frac{(N-n)}{N}\frac{1}{n}$$

In more complex sampling designs, this variance scaling factor is also more complex reflecting differences in sample size within trips, within post-strata, and within strata. However, for our purposes, we can use the simple version of the variance scaling factor to buffer the proximity index and decrease the potential for strata to have very low sample sizes.

Proximity Allocation Index

The final proximity allocation index used to allocate sampling effort to strata, D_h, is the product of $\widehat{T}$ (the average of the expected proportion of trips with monitored neighbors) and $(1 - F)$ (the variance scaling factor), where all terms are as previously defined:

$$\hat{D}_h = (1 - F_h) * \hat{T}_h$$

The full version highlights the estimation process, noting that the stratum-specific sample size, n_h, is an estimated parameter (product of stratum size, N_h, and stratum monitoring rate, r_h):

$$\widehat{D}_h = \left[1 - \frac{N_h}{(N_h - N_h r_h)} \frac{N_h r_h}{1}\right] \frac{\sum_{b=1}^{B_h} w_{bh} * (1 - \hat{A}_{bh})}{B_h} = \left[1 - \frac{N_h}{(N_h - n_h)} \frac{n_h}{1}\right] \frac{\sum_{b=1}^{B_h} w_{bh} * (1 - \hat{A}_{bh})}{B_h}$$

.

Similar to the CWB allocation method, this equation cannot be solved for the stratum sample sizes or monitoring rates: $\widehat{T}$, F, and costs are functions of sample size and hence numerical methods are used to determine the strata sample sizes that maximize the proximity allocation index while not exceeding the predetermined budget. As previously, the overall cost is the product of the number of trips in the stratum, the stratum-specific sampling rate, and the cost per trip for that stratum:

$$Cost = \sum_{h=1}^{H} N_h * r_h * c_h$$

where r_h is the stratum specific rate for the final proximity allocation index value, c_h is the cost per trip for stratum *h*, and N_h is the total number of trips for stratum *h*. Proximity allocation index values were calculated for each stratum over a range of sampling rates 0.0001 to 0.9950, and the associated monitoring costs. Based on these estimates, we identified the sample rates for each stratum associated with a maximal proximity allocation index value for the budget.

Using 2022 effort data and a budget of $4.5 million USD, we see that as the sample rate increases, the average proportion of trips with monitored neighbors increases rapidly reaching values close to 1 at monitoring rates close to 40% for all strata while at that same monitoring rate, the variance savings is more than 90% for all strata (variance scales to less than 10% of the population variance, Fig. 3-6).

The *Proximity* allocation method is included in the final suite of designs evaluated.

Allocation methods evaluated

In addition to the allocation methods currently used in the Draft ADP process to determine strata-specific sample rates (*Baseline 15% plus optimization* and *Equal Rates*), two novel methods were evaluated (*Cost-weighted Box* and *Proximity* allocation). These methods have substantively different allocation objectives which prioritize obtaining a representative sample of fishing activities over a range of budget scenarios. While a simple random sample (Equal Rates allocation) will achieve the same result on average, at lower sampling rates gaps in coverage may occur at lower sample sizes and some fishing activities (fisheries) might not be represented in the data. Both novel methods aim to increase the probability that unmonitored trips will have neighboring monitored trips and thus decrease the potential to have no monitored trips in a particular time of area (fishery).

In addition, a variation of the *Baseline 15% plus optimization* allocation that more fully integrates the EM strata into the allocation methods was evaluated. The *Baseline 15% plus optimization* allocation method dedicates funding for monitoring of 30% of fixed-gear trips (Fixed gear EM stratum) and 33% of trawl deliveries (Trawl EM stratum) before allocating any sampling resources to observed strata. This decreases monitoring resources available for use in the observed strata, resulting in lower monitoring rates and increased per-trip costs of observer coverage. Hence, we considered an allocation method that follows the same process but without the policy-set EM monitoring rates. By integrating the EM strata into the allocation methodology, we anticipated that this would allow higher observer coverage rates in the observed strata. Unfortunately, at the highest budget amount evaluated, this allocation method did not allocate more than 15% sampling across strata (reverted to equal rates) and was not pursued further.

A summary of the allocation methods presented above are presented in Table 3-6. The allocation methods that were incorporated into the final monitoring designs that were evaluated are the *Equal Rates*, *Baseline 15% plus optimization, Cost-weighted Box* and *Proximity* methods.

Final designs

Monitoring designs consist of both stratification definitions and a method to allocate sampling resources to those strata. In this evaluation, we investigated three novel stratification definitions and four novel allocation methods in addition to the stratification definition and allocation methods used in 2023. Of these, two novel stratification definitions and three novel allocation methods were combined into the final monitoring designs evaluated (Table 3-7).

The *Mixed HAL-POT* strata definition created strata with few trips and would be difficult to implement since vessel operators would be required to declare whether they would fish multiple gears in trip when logging their trips in ODDS in spite of those decisions often being made after fishing has begun. This stratification definition offered few advantages over the combined fixed-gear stratum where both HAL and POT gear trips are combined into one stratum, either observed or monitored with EM.

The *EM integrated Baseline 15% plus optimization* allocation method defaulted to the *Equal Rates* method under all stratification definitions and budget scenarios evaluated. Although this design has the advantage of integrating the EM strata into the overall monitoring design while otherwise maintaining the 2023 allocation objectives (15% baseline coverage and optimized allocation to decrease variance), those objectives were not achievable. This allocation method was not evaluated further and hence not included in any monitoring designs.

Table 3-1: Stratification definitions used in 2023, the current stratification definition. In the trawl EM strata, EM systems are used to monitor whether at-sea discards occurred and that unsorted catch is being delivered for shoreside sampling by the observer. Species count and fishing time and location data are collected by EM in the fixed-gear EM strata. Shaded cells = full coverage; OA = Open Access; IFQ = Individual Fishing Quotas; Exempted CPs are exempted from full coverage requirements.

Stratum	Fishing Activity	At-Sea		Shoreside	
		Monitoring	Frequency	Monitoring	Frequency
CP-plus	All trips taken by non-exempt CPs and CVs in non-IFQ catch-share programs	Observer	100%	None	0%
BSAI at-sea / shoreside observer	Trips taken by vessels participating under A91 (AFA fisheries of the BSAI)	Observer	100%	Observer	100%
BSAI trawl EM	Trips in the BSAI taken by trawl vessels that opted to carry compliance EM while pollock is open in the BSAI	EM	100%	Observer	100%
OA-IFQ Longline CVs	CV trips using longline gear not included in other strata and those taken on exempted CPs	Observer	Set in ADP	None	0%
OA-IFQ Pot CVs	CV pot trips not included in other strata and those taken on exempted CPs	Observer	Set in ADP	None	0%
OA-IFQ Trawl CVs	CV trawl trips not included in other strata	Observer	Set in ADP	None	0%
Fixed Gear HAL EM	Trips taken by vessels that opt-in and are approved by NMFS in each ADP and fish with HAL gear	EM	30%	None	0%
Fixed Gear POT EM	Trips taken by vessels that opt-in and are approved by NMFS in each ADP and fish with POT gear	EM	30%	None	0%
GOA at-sea / shoreside observer	Trips taken by vessels targeting pollock that are not part of the Trawl EM stratum	Observer	Set in ADP	Observers	Set in ADP
GOA trawl EM	Trips in the GOA taken by vessels that opted to carry compliance EM while pollock is open	EM	100%	Observers	33%
Zero	Catch taken on vessels less than 40 ft. LOA and on vessels fishing with jig gear.	None	0%	None	0%

Table 3-2. Number of trips (N), probability of selecting no trips at a 15% sample rate and 2022 effort data (not inclusive of PCTC or future changes in EM vessel participation) (P(0)), and probability of selecting fewer than 3 trips (P(<3)) under a 15% sampling rate under the current 2023 stratification and 2022 fishing effort where strata are defined by gear type and monitoring method (first column) with no stratification by FMP (AK-wide), for a stratification definitions that separate the BSAI from the GOA FMPs (BSAI vs GOA), and for a stratification definitions that separate each FMP (BS vs AI vs GOA). Highlighted blue text indicates proposed strata with few trips; red shading indicates proposed strata with high probability of obtaining no or few monitored trips in a 15% sample of trips. Strata are defined as fixed-gear EM HAL: EM_HAL; fixed-gear EM POT: EM_POT; trawl EM: EM_TRW; observed HAL: OB_HAL; observed POT: OB_POT; observed TRW: OB_TRW; zero coverage: ZERO. Grey shading delineates how population size changes when strata are subdivided by region.

15% sample	AK-wide			BSAI vs GOA				BS vs AI vs GOA			
STRATA	***N***	**P(0)**	**P(<3)**	**FMP**	***N***	**P(0)**	**P(<3)**	**FMP**	***N***	**P(0)**	**P(<3)**
EM_HAL	722	0.00	0.00	BSAI	32	0.00	0.12	AI	***9***	0.23	0.86
								BS	***23***	0.02	0.31
				GOA	690	0.00	0.00	GOA	690	0.00	0.00
EM_POT	353	0.00	0.00	BSAI	57	0.00	0.00	AI	***1***	0.85	1.00
								BS	56	0.00	0.00
				GOA	296	0.00	0.00	GOA	296	0.00	0.00
EM_TRW	620	0.00	0.00	GOA	620	0.00	0.00	GOA	620	0.00	0.00
OB_HAL	1352	0.00	0.00	BSAI	106	0.00	0.00	AI	***27***	0.01	0.20
								BS	78	0.00	0.00
				GOA	1246	0.00	0.00	GOA	1247	0.00	0.00
OB_POT	1086	0.00	0.00	BSAI	255	0.00	0.00	AI	***14***	0.10	0.65
								BS	241	0.00	0.00
				GOA	831	0.00	0.00	GOA	831	0.00	0.00
OB_TRW	631	0.00	0.00	BSAI	115	0.00	0.00	AI	***5***	0.44	0.97
								BS	110	0.00	0.00
				GOA	516	0.00	0.00	GOA	516	0.00	0.00
ZERO	1601	0.00	0.00	BSAI	134	0.00	0.00	BS	134	0.00	0.00
				GOA	1467	0.00	0.00	GOA	1467	0.00	0.00

Table 3-3: Number of trips (N), probability of selecting no trips at a 15% sample rate (P(0)), and probability of selecting fewer than 3 trips (P(<3)) under a 15% sampling rate and 2022 effort data (not inclusive of PCTC or future changes in EM vessel participation) where strata are defined by gear type and monitoring method (first column) with no stratification by FMP (AK-wide), for definitions that separate the BSAI from the GOA (BSAI vs GOA), and that separate FMPs (BS vs AI vs GOA). Highlighted blue text indicates proposed strata with few trips; red shading indicates proposed strata with high probability of obtaining no or few monitored trips in a 15% sample of trips. Strata are defined as fixed gear EM HAL: EM_FG_HAL; mixed HAL-POT EM: EM_FG_MIXED; fixed gear EM POT: EM_FG_POT; Pelagic trawl EM: EM_TRW_PTR; observed fixed gear HAL: OB_FG_HAL; observed mixed HAL-POT: OB_FG_MIXED; observed fixed gear POT: OB_FG_POT; observed pelagic trawl (PTR): OB_TRW_PTR; observed non-pelagic trawl (NPT): OB_TRW_NPT; observed mixed PTR-NPT: OB_TRW_MIXED. Grey shading delineates how population size changes when strata are subdivided by region.

15% sample	AK-wide			BSAI vs GOA				BS vs AI vs GOA			
STRATA	***N***	**P(0)**	**P(<3)**	**FMP**	***N***	**P(0)**	**P(<3)**	**FMP**	***N***	**P(0)**	**P(<3)**
EM_FG_HAL	668	0.00	0.00	BSAI	***28***	0.01	0.18	AI	***6***	0.38	0.95
								BS	22	0.03	0.34
				GOA	640	0.00	0.00	GOA	640	0.00	0.00
EM_FG_MIXED	100	0.00	0.00	BSAI	***6***	0.37	0.96	AI	***4***	0.52	0.99
								BS	***2***	0.72	1.00
				GOA	94	0.00	0.00	GOA	94	0.00	0.00
EM_FG_POT	307	0.00	0.00	BSAI	55	0.00	0.00	BS	55	0.00	0.00
				GOA	252	0.00	0.00	GOA	252	0.00	0.00
EM_TRW_PTR	620	0.00	0.00	GOA	620	0.00	0.00	GOA	620	0.00	0.00
OB_FG_HAL	1228	0.00	0.00	BSAI	101	0.00	0.00	AI	27	0.01	0.20
								BS	73	0.00	0.00
				GOA	1127	0.00	0.00	GOA	1128	0.00	0.00
OB_FG_MIXED	229	0.00	0.00	BSAI	***11***	0.16	0.79	AI	***1***	0.85	1.00
								BS	***10***	0.19	0.83
				GOA	218	0.00	0.00	GOA	218	0.00	0.00
OB_FG_POT	981	0.00	0.00	BSAI	249	0.00	0.00	AI	***13***	0.12	0.69
								BS	236	0.00	0.00
				GOA	732	0.00	0.00	GOA	732	0.00	0.00
OB_TRW_MIXED	160	0.00	0.00	GOA	160	0.00	0.00	GOA	160	0.00	0.00
OB_TRW_NPT	182	0.00	0.00	BSAI	114	0.00	0.00	AI	***5***	0.44	0.98
								BS	109	0.00	0.00
				GOA	68	0.00	0.00	GOA	68	0.00	0.00
OB_TRW_PTR	289	0.00	0.00	BSAI	***1***	0.85	1.00	BS	***1***	0.85	1.00
				GOA	288	0.00	0.00	GOA	288	0.00	0.00

Table 3-4: Number of trips (N), probability of selecting no trips at a 15% sample rate (P(0)), and probability of selecting fewer than 3 trips (P(<3)) under a 15% sampling rate and 2022 effort data (not inclusive of PCTC or future changes in EM vessel participation) where strata are defined by combined fixed gear and trawl gears, and by monitoring method (first column) with no stratification by FMP (AK-wide), for definitions that separate the BSAI from the GOA (BSAI vs GOA), and that separate FMPs (BS vs AI vs GOA). Highlighted blue text indicates proposed strata with few trips; red shading indicates proposed strata with high probability of obtaining no or few monitored trips in a 15% sample of trips. Strata are defined as fixed gear EM HAL: EM_FG_HAL; mixed HAL-POT EM: EM_FG_MIXED; combined fixed gear EM: EM_FIXED; trawl EM: EM_TRW; observed fixed gear: OB_FIXED; observed trawl: OB_TRW; zero coverage: ZERO. Grey shading delineates how population size changes when strata are subdivided by region.

15% sample	**AK-wide**			**BSAI vs GOA**				**BS vs AI vs GOA**			
STRATA	***N***	**P(0)**	**P(<3)**	**FMP**	***N***	**P(0)**	**P(<3)**	**FMP**	***N***	**P(0)**	**P(<3)**
EM_FIXED	1075	0.00	0.00	BSAI	89	0.00	0.00	AI	***10***	0.20	0.82
								BS	79	0.00	0.00
				GOA	986	0.00	0.00	GOA	986	0.00	0.00
EM_TRW	620	0.00	0.00	GOA	620	0.00	0.00	GOA	620	0.00	0.00
OB_FIXED	2438	0.00	0.00	BSAI	361	0.00	0.00	AI	41	0.00	0.04
								BS	319	0.00	0.00
				GOA	2077	0.00	0.00	GOA	2078	0.00	0.00
OB_TRW	631	0.00	0.00	BSAI	115	0.00	0.00	AI	***5***	0.44	0.97
								BS	110	0.00	0.00
				GOA	516	0.00	0.00	GOA	516	0.00	0.00
ZERO	1601	0.00	0.00	BSAI	134	0.00	0.00	BS	134	0.00	0.00
				GOA	1467	0.00	0.00	GOA	1467	0.00	0.00

Table 3-5: Stratification definitions to be included in final designs that will be evaluated.

Stratification	Number of Sampled Strata	Definition	Rationale
2023 (CURRENT)	6	Monitoring Method (Observer, EM Fixed Gear, EM Trawl) and Gear Type (HAL, POT, TRW)	Current stratification definition
FMP	11	Monitoring Method (Observer, EM Fixed Gear, EM Trawl) and Gear Type (HAL, POT, TRW) and FMP (BSAI, GOA)	Potential to reduce the likelihood of data gaps in the BSAI
Combined fixed gear - FMP (FIXED-FMP)	7	Monitoring Method (Observer, EM Fixed Gear, EM Trawl) and Gear Type (FIXED, TRW) and FMP (BSAI, GOA)	Maintains statistical integrity without creating small strata

Table 3-6: Objective, rational for, benefits and shortcomings of allocation methods considered for evaluation. At the highest budget evaluated, the *EM integrated Baseline 15% plus optimization* (shaded row) allocation method did not allocate above 15% for any stratum and hence reverted to equal rate allocation as was not evaluated further.

Allocation Method	Objective	Rational	Benefits	Shortcomings
Equal Rates	Sample each stratum proportionally to the size of the stratum	Simple allocation relies on few assumptions	Data can be combined across strata without use of stratified estimators	At low sample size, can be prone to data gaps
Baseline 15% plus optimization (status quo)	1 Monitor the at-sea observed strata at a 15% rate 2 Minimize combined variance of at-sea discards of groundfish, halibut PSC, and salmon PSC	Management of halibut PSC and salmon PSC relies on low variance estimates	Baseline rate to decrease data gaps	1 High EM monitoring rates result in low at-sea rates - reverts to equal rates 2 Policy based monitoring rate specification; 3 At low funding, at-sea baseline rates are not reached 4 Uses between-trip (not CAS) variance
EM integrated Baseline 15% plus optimization	1 Monitor all strata at a 15% rate; 2 Minimize combined variance of discards of groundfish, halibut PSC, and salmon PSC	Management of halibut PSC and salmon PSC relies on low variance estimates	1 Baseline rate to decrease data gaps	1 At low funding, at-sea baseline rates are not reached 2 Uses between-trip (not CAS) variance
Cost Weighted Boxes	1 Control prevalence of data gaps; 2 Minimize overall costs	Data users need representative data at varied spatio-temporal scales and limited funding resources need to be well-spent	1 High data utility 2 Fewer data gaps 3 Limits sampling in high-cost strata	1 Iterative process to set stratum weightings
Proximity	1 Control prevalence of data gaps; 2 Control prevalence of low sample size	Data users need representative data at varied spatio-temporal scales and sufficient sample size	1 High data utility 2 Fewer data gaps 3 Fewer low-sample strata (lower variance)	1 Iterative process to allocate sample effort

Table 3-7: Monitoring designs (combinations of stratification definitions and allocation methods) considered for evaluation. Shading indicated designs that were evaluated (green) and those that were removed from consideration (red). Dark green indicates the design used in 2023.

<table>
<tr><th>Allocation Method</th><th>2023 Stratification</th><th>FMP</th><th>Combined Fixed Gear and FMP</th><th>Separate Mixed-HAL & POT strata</th></tr>
<tr><td>Equal Rates</td><td>YES: integrated EM, baseline comparison</td><td>YES: integrated EM</td><td>YES: integrated EM</td><td rowspan="2">NO: this stratification increased logistical difficulties and resulted in strata with few trips when coupled with stratification by FMP.</td></tr>
<tr><td>Baseline 15% plus optimization (status quo)</td><td>YES: both the stratification definition and allocation method were used in 2023</td><td>YES</td><td>YES</td></tr>
<tr><td>EM integrated Baseline 15% plus optimization</td><td colspan="3">NO: under highest budget evaluated, this design defaulted to equal rates and was not pursued further</td><td>NO: Defaults to equal rate allocation and creates strata with few trips</td></tr>
<tr><td>Cost Weighted Boxes</td><td>YES: 2023 stratification definition and gap minimization Council requested cost efficiencies design</td><td>YES: integrated EM</td><td>YES: integrated EM</td><td rowspan="2">NO: this stratification increased logistical difficulties and resulted in strata with few trips when coupled with stratification by FMP.</td></tr>
<tr><td>Proximity</td><td>YES: 2023 stratification and gap minimization with sample size buffer</td><td>YES: integrated EM</td><td>YES: integrated EM</td></tr>
</table>

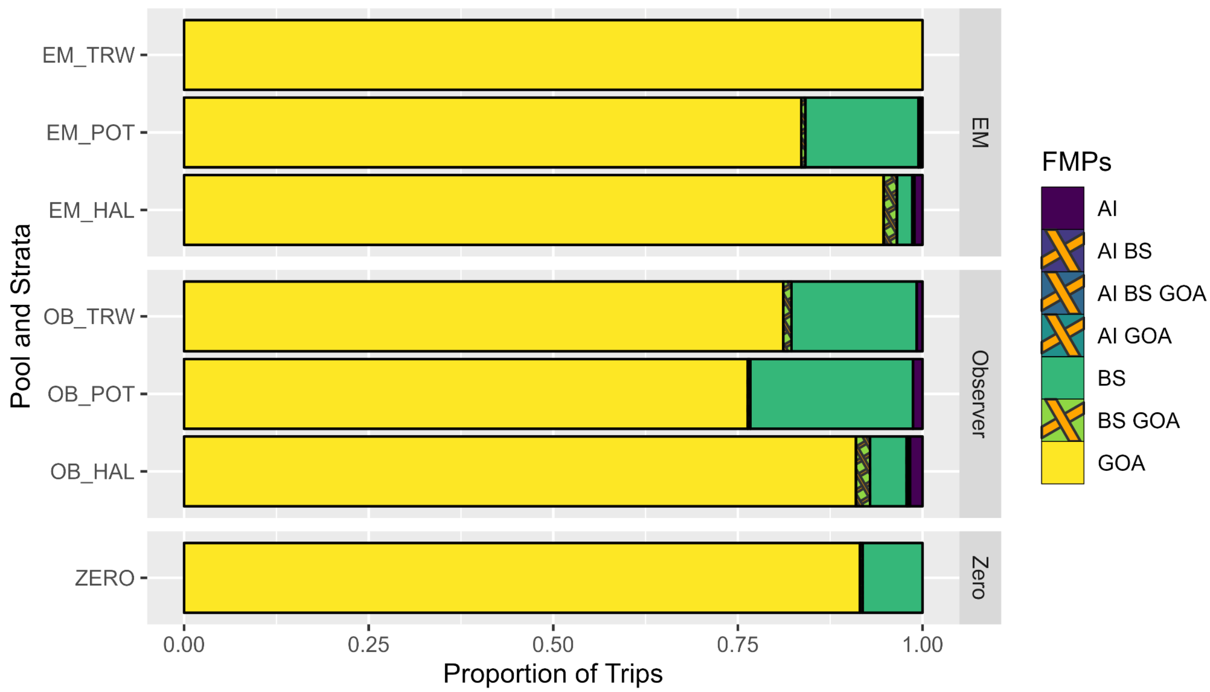

Figure 3-1. Proportions of trips by FMP area fished for each sampling stratum based on 2022 fishing effort data. Trips that fished in more than one FMP are represented by orange cross-hatching.

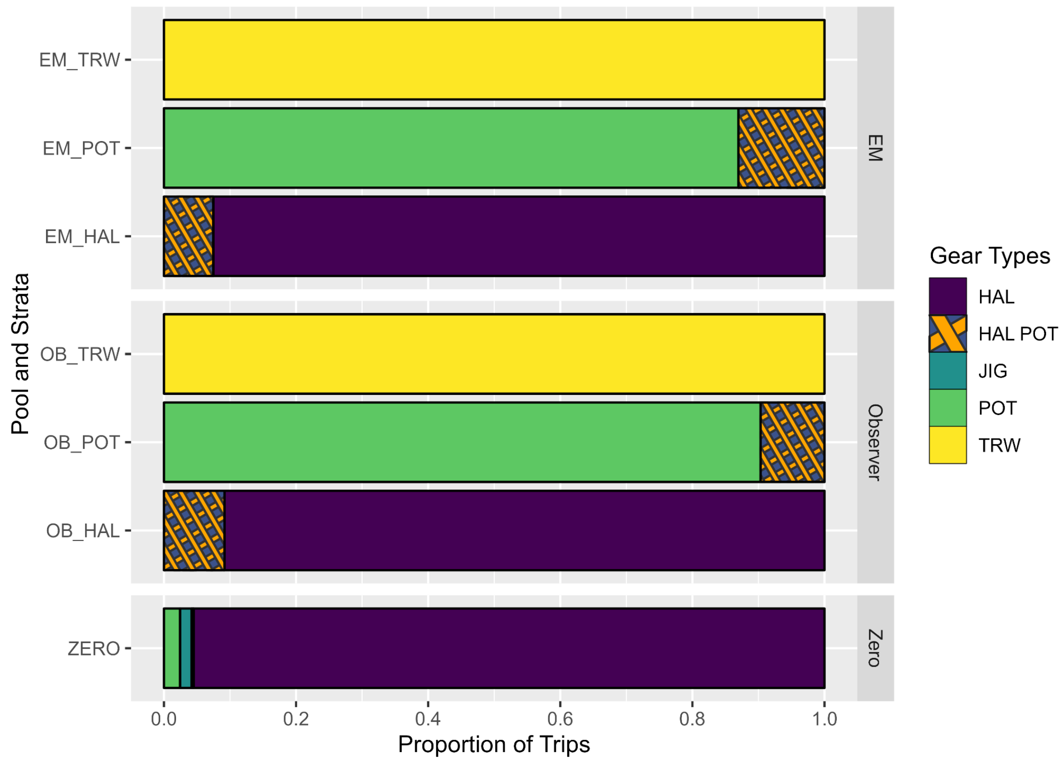

Figure 3-2. Proportions of trips that utilized each gear type for each sampling stratum based on 2022 fishing effort data. Trips that used more than one gear type are represented by orange cross-hatching.

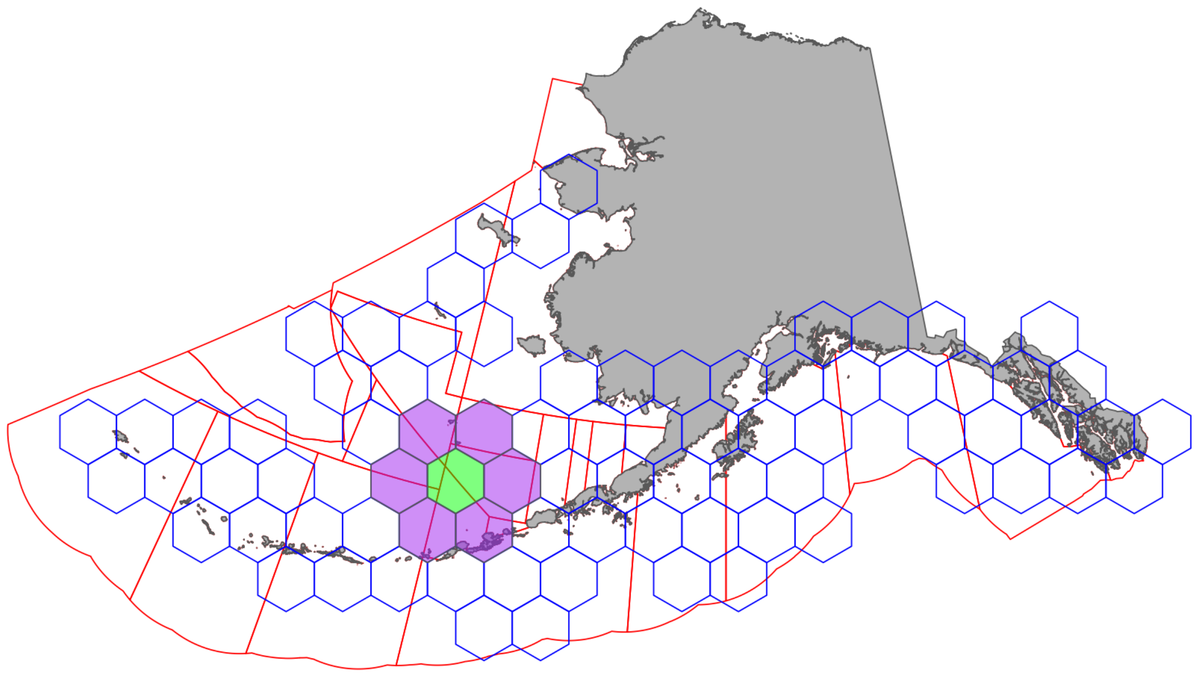

Figure 3-3. Grid of 200 km-wide hexagon cells (blue) used as the spatial components of the ‘box’ definition where 1-week bins define the temporal component; this definition is used to categorize fishing effort in space and time. For each gear type, boxes are also allowed to neighbor adjacent cells in both time and space (i.e., the ‘neighborhood) when determining the number of nearby trips that may be monitored. For example, the green cell and the trips that fish within it in a given week represents a single box, but the adjacent purple cells represent the spatial extent of its neighborhood (600-km wide) and the adjacent weeks (spanning 3-weeks) represent the temporal extent of the neighborhood. Boxes with more trips in their neighborhoods are more likely to be neighboring sampled trips. The cost-weighted boxes and proximity allocation methods and the interspersion evaluation metric use this same box definition. NMFS reporting areas (red) are shown for comparison.

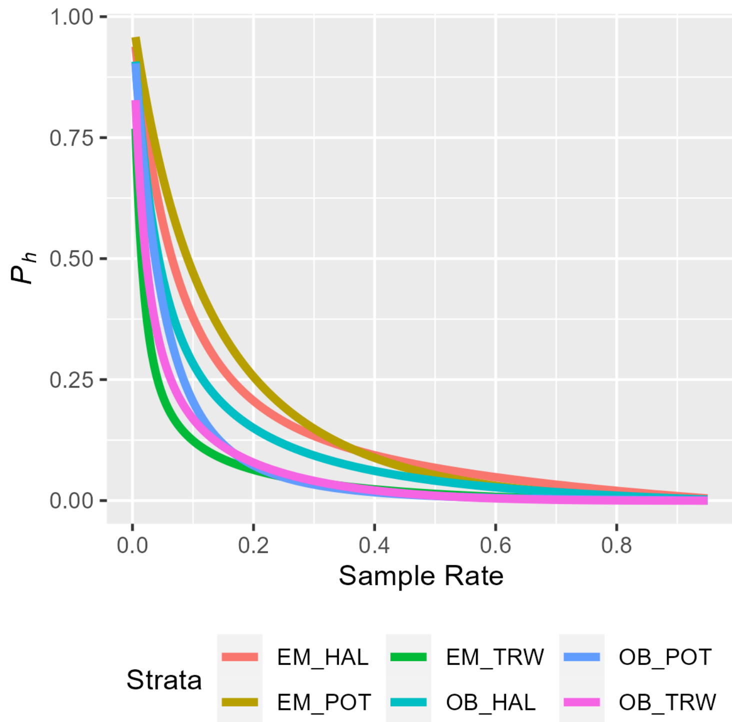

Figure 3-4. Average proportion of boxes without monitored trips in their neighborhood (P_h) as a function of sample rate for a range of 2022 strata.

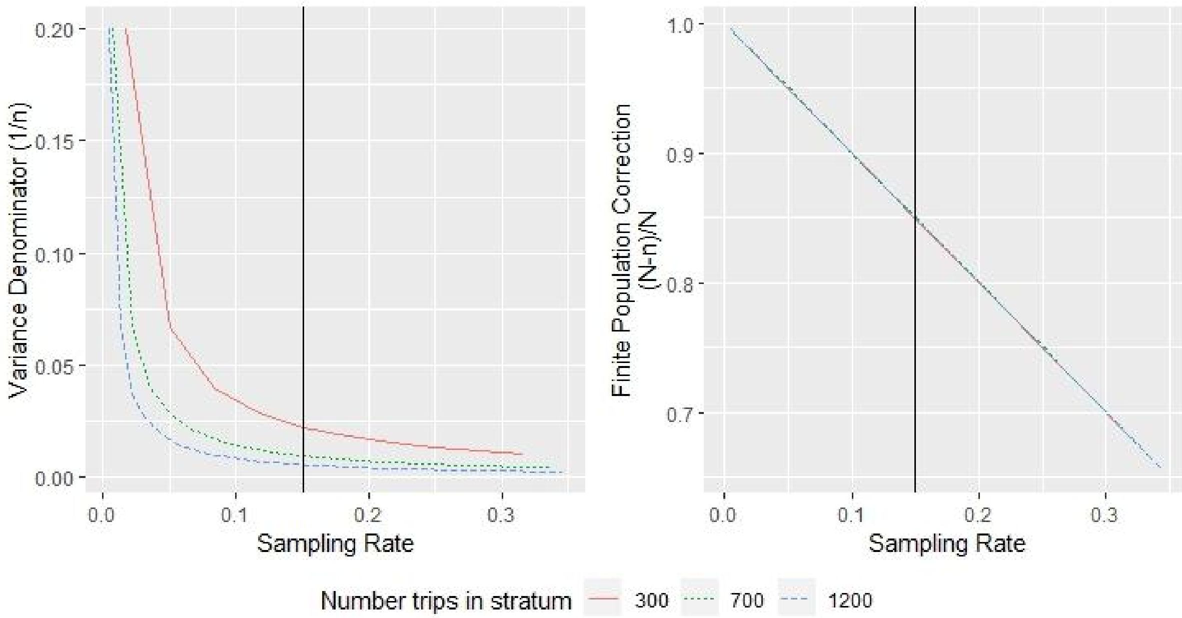

Figure 3-5. Components of the variance scaling factor (*F*), sample size in left panel and the finite population correction factor (FPC) in right panel, as a function of sample rate for three strata sizes.

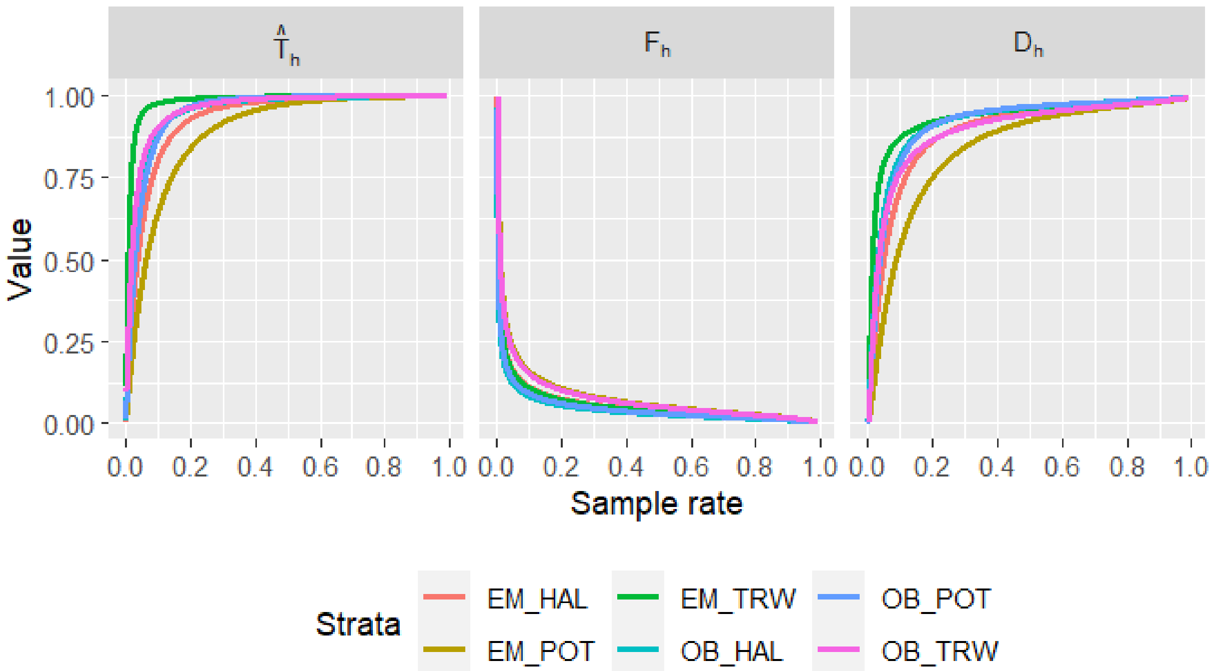

Figure 3-6. Proximity index ($\widehat{T}_h$), variance scaling factor (F_h), and the proximity allocation index (D_h) for each 2022 strata at a range of sampling rates under a $4.5 million USD budget.

4. Evaluating Design Performance

Fishery monitoring programs with less than complete coverage must be regularly evaluated, because sampling makes inference from the resulting data susceptible to higher levels of imprecision and potential bias compared to programs with complete (100%) coverage. In this section, various metrics are used to provide insight into how well potential future partial coverage monitoring programs can be expected to perform. These metrics address monitoring output, efficiency, and effectiveness. Metrics that address monitoring output include the number of trips sampled by observers and monitored by EM. Metrics that address monitoring efficiency include the variance in expenses, data timeliness, and trip-level variance. Metrics that address monitoring effectiveness include the power to detect rare events, the power to detect differences between monitored and unmonitored trips, and how far apart monitored trips are in space and time from unmonitored trips (interspersion).

Evaluation Metrics

Trips Sampled

The number of trips sampled by a design can be used as a measure of output, given identical budgets for monitoring. It is important to consider that the three monitoring programs collect different types of data and to therefore quantify those samples separately. Broadly, the counts of the expected number of samples (i.e., monitored trips) from each design can be categorized into two groups: those that collect biological data (e.g., sex/length/weight, tissues) and those that collect catch composition data. The expected number of samples collected by a monitoring method is calculated as the sum of the expected number of samples collected by each of its strata, or each stratum's sample rate multiplied by the total number of trips in the stratum. Although stratification methods may differ between designs, summaries can be grouped by monitoring method and gear type, as these do not differ between the proposed designs. These gear-specific sampled trip counts can then be summarized as the total within each monitoring method.

Biological data is collected by at-sea observers for all or a subset of hauls within a sampled trip. In addition, shoreside observers collect biological samples in the Trawl EM stratum, but those samples are collected at the offload rather than the haul level. Counts of the expected number of sampled trips by both at-sea observers and shoreside observers will be presented, as well as the combined total.

Catch composition data are collected by all three monitoring methods. At-sea observers and fixed-gear EM collect composition data from all or a subset of hauls. Additionally, because all Trawl EM trips utilize compliance monitoring to enforce maximized catch retention, landing reports inform catch composition of the entire trip. Counts of the expected number of sampled trips by at-sea observer, fixed-gear-EM, and all Trawl EM trips will be presented, as well as the combined total.

Variance in Expenses

The monitoring program will ideally collect as many samples as the budget will allow, and designs with lower variability in expenditures have a lower risk of going over budget. The actual expenditures incurred by the monitoring program will vary depending on many factors. In an evaluation of expenditure variance, these factors can be divided into those that are assumed to either vary or not vary between designs. Factors assumed to not vary between designs are differences in predicted versus realized fishing effort and predicted versus realized trip duration. Factors that affect expenditures that vary between designs are driven by random sampling, such as the stratification and allocation methods, as each stratum has a different population of trips, each with a different trip duration that affects the cost of sampling.

By applying the allocated rates from each design to our expectation of fishing effort and simulating sampling 10,000 times, we can build a distribution of realized monitoring expenditures and quantify the variance of the outcomes.

Power to Detect Rare Events

This is a novel analysis for an ADP that addresses mandates to monitor mammals and seabirds under various statutes. These organisms are rarely caught as bycatch. Therefore it is worth exploring how likely it would be for monitoring under the 2024 ADP and beyond to detect these rare events. The power to detect bycatch is a function of the bycatch (individuals) per unit effort (BPUE), the variance of the bycatch divided by its mean, the number of trips in the population, and the number of trips sampled. All of these values will differ between designs.

For this analysis, trip data from 2013-2022 were used to generate expected BPUE. Dead and injured mammals, as recorded by the North Pacific Groundfish and Observer Program, were used in the analysis. Three species of mammals were investigated: killer whales, humpback whales, and Steller sea lions. In a ten year period, no killer or humpback whales were recorded as dead bycatch in partial coverage trips, so they were dropped from the analysis.

Bird bycatch data were obtained from catch data that is used to generate catch and bycatch estimates by the NMFS Catch Accounting System (CAS). Species of interest were short-tailed albatross, Laysan Albatross, spectacled eiders, and Steller's eider. Although eiders are coded in CAS data as 'Other Birds', a query of the FMA data confirmed that no dead eiders were observed during the ten year period (2013-2022). Therefore, eiders were dropped from the analysis.

Mammal and seabird bycatch data from each trip were assigned to partial coverage sampling strata under the proposed stratification definitions being evaluated in this document, and estimates of the mean and variance of BPUE for each species were generated for each. This process was relatively straightforward for birds because of past efforts to align its data source with ADP analyses. However for mammals this process needed to be done manually. Since FMA data was the sole source of mammal bycatch information, it was not possible to generate estimates for Steller sea lions from strata that were expected to use EM in 2024.

For the stratification FIXED_FMP, mean and variance of BPUE were first aggregated to each FMP x gear (HAL, POT and TRW), and then these estimates were aggregated to each FMP x FIXED from

$$\bar{y} \quad = \frac{1}{N}\sum_{h=1}^{L} N_h \, \bar{y}_h$$

and

$$v\hat{a}r(y \quad) = \sum_{h=1}^{L} \left(\frac{N_h}{N}\right)^2 s_h^2$$

Where *h* is the stratum, and *N* is the number of trips. The result was the estimated mean and variance for each 2024 sampling stratum and species over a ten year period. These values were then paired with the expected number of trips for 2024 and the expected number of trips observed for 2024 that result from each design's allocation. These formed the basis for a second aggregation to generate design level estimates. The mean BPUE was calculated as before, however in this formulation, *h* denotes the 2024 sampling stratum and N and n denote expected values for 2024. Variance was calculated from

$$v\hat{a}r(y_{st}) = \sum_{h=1}^{L} \left(\frac{N_h}{N}\right)^2 \left(\frac{N_h - n_h}{N_h}\right)\frac{s_h^2}{n_h}$$

The power for future partial fishery monitoring programs to detect mammal and seabird rare event bycatch were generated using the R package *ObsCovgTools* described by Curtis and Carretta (2020)[5]. Estimates of bycatch from each design were generated by multiplying the BPUE by N. **Estimates of bycatch are for this analysis only and should not be confused with official estimates from NMFS.** They are underestimates because they do not include estimates from the zero coverage stratum.

Power to Detect Monitoring Effects

Monitoring effects occur when fishing events that are monitored have different properties than those that are not. Monitoring effects are important to identify because they have the potential to

[5] An online version of this test can be found at https://connect.fisheries.noaa.gov/content/bb44513d-4956-44dd-b0a6-673c9f2a3e3a/

cause bias and jeopardize the ability to make inferences about the entire fleet using data from monitored vessels and trips (Benôit and Allard, 2009). Monitoring effects have been shown to occur in the North Pacific prior to restructuring of the observer program in 2013 (Faunce and Barbeaux 2011) as well as in annual reviews of the restructured program (Faunce et al. 2017; Ganz et al. 2018, Ganz et al. 2019, Ganz et al. 2020).

For this analysis, 2020-2022 fishing trips were used to generate expected monitoring effects in retained weight (mt) of groundfish, the number of retained species, and the duration (days) of the fishing trip. These metrics are among those used to test for differences between monitored and unmonitored trips in the annual evaluation of observer deployment done by the National Marine Fisheries Service (NMFS; AFSC and AKRO, 2021). The *days fished* (or *duration*) metric is an indicator of fishing effort, the *number of species* is a measure of catch diversity, and *landed catch* is a measure of magnitude.

Similar to the power analysis for rare species, this analysis re-aggregates the trip-level data to calculate the mean and variance for each 2024 stratum, for each design, but also splits monitored and unmonitored trips. The metric of interest is not BPUE as before, but duration, species, or landed catch between monitoring status. For designs with the stratification FIXED_FMP, values for each monitoring status, each metric, were first aggregated to each FMP x gear, and then these estimates were aggregated to each FMP x FIXED using a stratified estimator as described in the power analysis for rare species. As in that analysis, the resulting values for each metric x stratum and monitoring status were then paired with the expected number of trips for 2024 and the expected number of trips observed for 2024 from each allocation method in designs.

A two-sample test of independence was conducted for each metric. This test treats the monitored and unmonitored conditions as different distributions. The test calculates the power to detect a difference between monitored and unmonitored trips given past magnitude (mean) and variance and the expected sample size in 2024. Power here is the probability to detect a difference with 95% confidence (α) given that one exists (i.e. the differences are real and not due to random error). High power is desirable, and power increases with the magnitude of the differences between monitored and unmonitored trips, lower variances in the distributions, and greater numbers of monitored trips. Tests were performed with the R package *pwrss* (Bulus, 2023)[6]. For presentations, the effect size was calculated as the mean of monitored trips subtracted from the mean of unmonitored trips (in this way negative values indicate that monitored trips were larger while positive values indicate that unmonitored trips were larger). The effect size was converted to a relative percentage of the monitored mean, or RPM, from the effect size divided by the mean of the monitored trips and the result multiplied by 100.

[6] An online version of this test can be found at https://pwrss.shinyapps.io/index/.

Data Timeliness
We defined data timeliness as the length of time between a trip's end date and the date at which data were available to the CAS. We chose this definition because it allowed us to measure data timeliness consistently across EM and observed strata. For trawl EM trips, we had to approximate this definition of data timeliness by adding one day to the difference between offload end date and the date that data were submitted to the AFSC. We added one day because that is the time it typically takes data that have been submitted to the AFSC to reach the CAS.

We used data from 2022 to calculate an average data timeliness for each stratum. These averages, along with the distributions that produced them, are shown in Figure 4-1. Data from 2022 were chosen in part because they represent the most recent full year of data that were collected during a year in which COVID-19 did not significantly impact deployment. Trips were excluded if there were no data associated with them.

For trips in the fixed-gear EM strata, we adjusted data timeliness for designs in which the expected number of monitored trips differed from the number of trips monitored in 2022. We did this because we expect review times to be impacted by the number of trips needing review. As an example, 285 of the fixed-gear EM trips that occurred in 2022 had been reviewed by the time data timeliness was calculated. This number differs from the number of monitored EM HAL trips in the 2022 Annual Report, since video review continued after the report was published. If a design results in 100 fixed-gear EM trips expected to be monitored, the data timeliness value assigned to those trips would be $217.97 \times \frac{100}{285} = 76.48$ days for EM HAL and $210.53 \times \frac{100}{285} = 73.87$ days for EM POT.

After adjusting for the number of fixed-gear EM trips expected to be monitored, we applied data timeliness values to each trip in a simulation, based on the stratum of the trip. We assigned the same data timeliness value to monitored and unmonitored trips within a stratum, since estimates for unmonitored trips will rely on data from monitored trips. As an example, if a year of fishing effort had an equal number of trips (monitored or unmonitored) in each of the current partial coverage strata, and if data timeliness for fixed-gear EM strata didn't require an adjustment, the average data timeliness would be 73.41 days (Table 4-1).

Trip-Level Variance
We estimated the trip-level variance associated with estimates of Chinook PSC, halibut PSC, groundfish discards, and crab PSC for each design and year of fishing effort. We chose Chinook PSC, halibut PSC, and groundfish discards because the status quo allocation strategy uses optimized days to reduce the variance associated with estimates of these quantities. We chose crab PSC given recent efforts to better understand crab stocks and because past ADPs have considered including this metric among the metrics used by status quo allocation.

By presenting trip-level variance as an evaluation metric, we show how the precision of these estimates can be expected to change with different designs. However, we are referring to trip-level (i.e., between-trip) variance, which is not equal to the actual variance estimates associated with these quantities. Actual variance estimates take into account variation at levels of the sampling hierarchy within the trip (e.g. haul-level and sample-level) and use different estimation processes (see the 2019 Annual Report: AFSC and AKRO 2021, Appendix C). In contrast, trip-level variance is a simplification that assumes catch is known at the trip-level without error, and calculates the between-trip variance using the traditional variance formula (the second equation in the Small Sample Size Buffering section of Chapter 3). Applied within strata, this equation produces the curves in Figure 4-2, which show the relationship between sample size and trip-level standard error (the square-root of variance).

Although trip-level standard error is driven by sample size, we typically present ADP sampling objectives in terms of rates. Figure 4-3 shows the same curves, with monitoring rate instead of sample size as the horizontal axis, and with a 15% monitoring rate shown by the dashed red lines.

If all strata were sampled at a 15% rate, the trip-level standard error for each species within a stratum would be equal to the quantity on the vertical axis that corresponds to the point at which the dashed red line intersects the curve. To get the total trip-level standard error for each species or species group, we added variances across strata and took the square root of that sum. We present results as CVs, which are equal to the standard error for each species or species group divided by the estimate of catch for that species or species group. The precision of estimates increases as CV decreases. As an example, Table 4-2 shows the CVs that result from the hypothetical design in which all partial coverage strata in 2022 were monitored at a rate of 15%.

Although the simplifying assumptions built into this evaluation metric result in CVs that are not equal to those we would expect out of the CAS, changes in the CVs produced here do show the relative impact that changes to selection rates have on uncertainty at the trip-level, which is the level at which observer and EM deployment is planned.

Interspersion

The interspersion metric measures the proportion of trips that neighbor monitored trips using a method that is similar to the computation of the proximity index, $\widehat{T}$, that is used in the *Proximity* allocation method. This metric is a measure of 1) the extent to which all of the proposed designs result in interspersing monitored and unmonitored trips thereby increasing the potential that all fishing activities are represented in the sample, 2) whether trips in strata where the full suite of data are not collected (EM and zero selection trips) are neighboring observed trips so that missing data elements (e.g., mean weights, species identification for similar species, protected species data) are available to complement EM data and as proxy values for zero selection pool fisheries, and 3) whether monitored fixed-gear EM trips are interspersed with unmonitored fixed-gear EM trips. For many analyses, EM data (trawl and fixed gear) are incomplete and missing data elements are obtained from the collected observed strata data. The importance of these data dependencies vary with the analysis being conducted; however for estimation of catch and bycatch, and for stock assessments, there is a strong reliance on observer collected data.

Although the interspersion metric shares many computational characteristics of the proximity index, it differs in one important way: the interspersion metric calculations are not specific to each stratum. Rather, the interspersion metric is the expected proportion of trips within a monitoring method that are neighboring monitored trips.

Interspersion is calculated for each data dependency combination and defines sampled neighboring trips as those that fished with the same gear type (hook-and-line, pot, or trawl) within the same neighborhood (3-week period and 7 hex-cells, spanning 21 boxes).

The interspersion data dependencies quantified for this evaluation are:

- Observer to Observer (biological and composition data to assess interspersion of monitored and unmonitored trips)
- Observer to Fixed-gear EM (biological data to support estimation based on EM strata data)
- Observer to Zero (biological and composition data to support estimation in the absence of independently collected data)
- Fixed-gear EM to Fixed-gear EM (composition data to assess interspersion of monitored and unmonitored trips)
- EM_TRW to EM_TRW (biological data collected shoreside to assess interspersion of monitored and unmonitored trips)

To compute the interspersion index, we use the binomial approximation to the hypergeometric distribution to generate the probability that there are no monitored trips in the neighborhood of box *b*, $\widehat{A}_b$. The difference is that the interspersion neighborhood, G_b is defined as all trips of the same gear type within each stratum, *h*, and then combined across strata.

$$\hat{A}_b = \sum_{h=1}^{H} (1 - r_h)^{t_{G_b}}$$

The expected number of trips that have neighbors is the sum of the number of trips in the neighborhood, w_b, multiplied by the probability that one or more of those trips are sampled, $\left(1 - \widehat{A}_b\right)$.

As in the proximity index, trips are allowed to span multiple boxes, and contribute equally to each box (e.g., a trip that crosses three boxes is counted as 0.33 trips in each box).

The interspersion metric, $\widehat{I}$, is the average of the expected proportion of trips with monitored neighbors averaged over the b = 1, …..B boxes in the stratum.

$$\hat{I} = \frac{\sum_{b=1}^{B} w_b * \left(1 - \widehat{\mathrm{A}}_b\right)}{B}$$

Although EM_TRW is included in these comparisons, since EM is used to identify deliveries of sorted catch (at-sea discards) that the observer should not sample (i.e., compliance monitoring), EM_TRW is not compared with OB_TRW.

It should be noted that these data dependencies highlight the importance of samples from at-sea observers on fixed-gear vessels as they are three-fold (i.e., other fixed-gear trips in the at-sea observer selection pool, fixed-gear EM selection pool, and zero selection pool.

Table 4-1. The data timeliness score that would result from a year of fishing effort that had an equal number of trips in each of the current partial coverage strata.

Stratum	Data timeliness (days)
EM HAL	217.97
EM POT	210.53
EM TRW	7.31
HAL	2.25
POT	1.19
TRW	1.19
Average	73.41

Table 4-2. Example trip-level CVs that result if all partial coverage strata in 2022 were monitored at a rate of 15%.

Metric	CV
Chinook PSC	0.15
Halibut PSC	0.17
Groundfish discards	0.06
Crab PSC	0.24

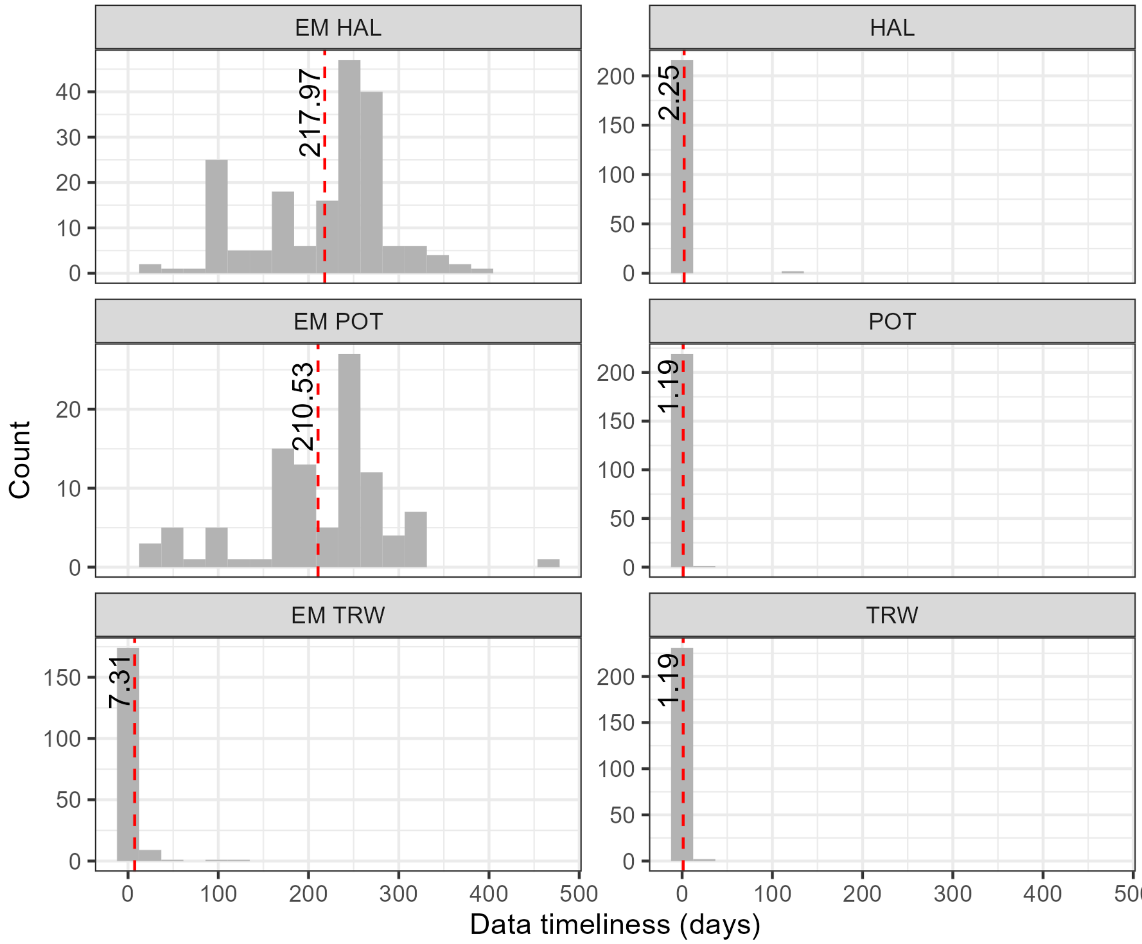

Figure 4-1. Distributions of data timeliness in days by stratum. Dashed red lines and annotations show mean data timeliness.

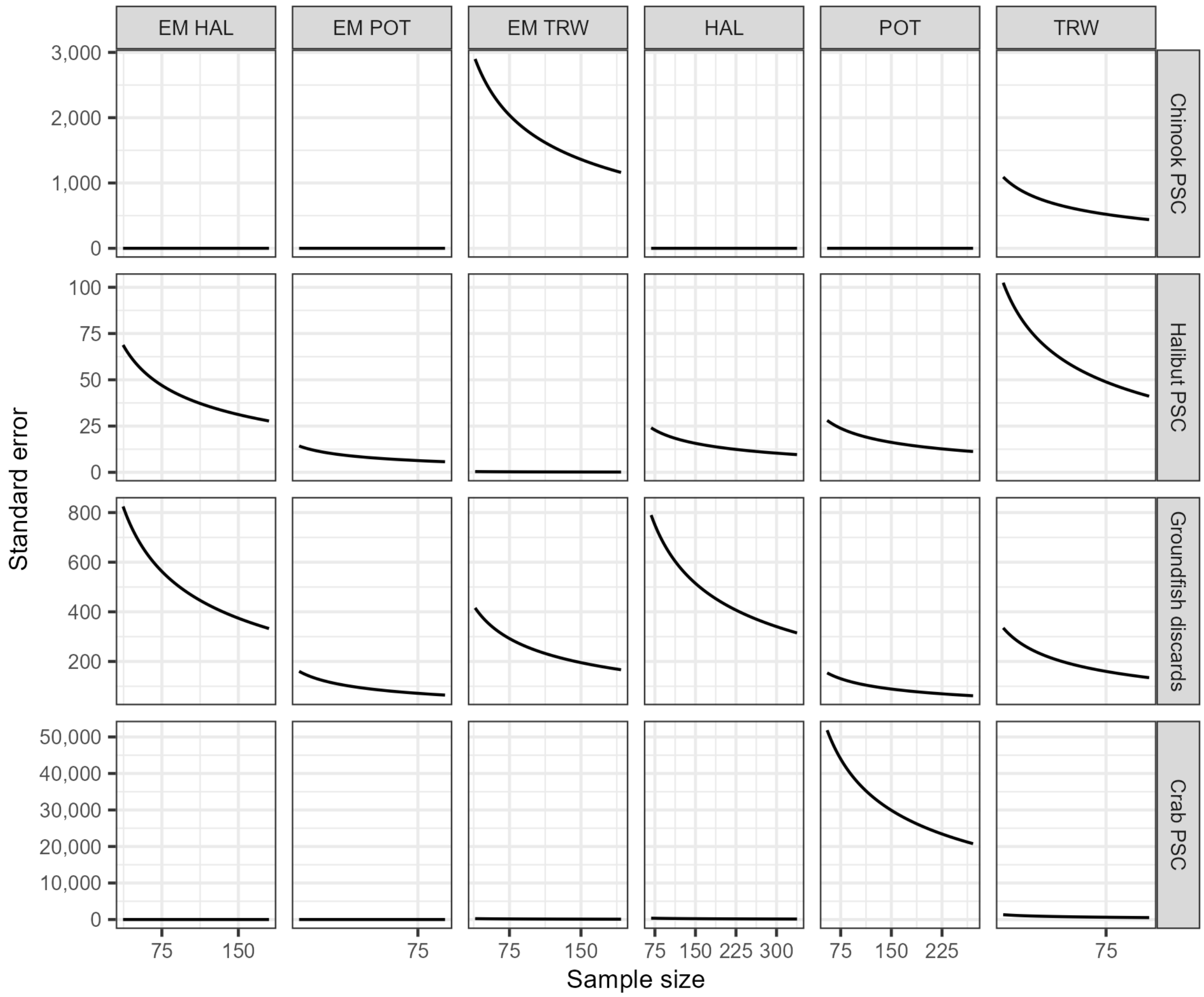

Figure 4-2. Trip-level standard error by sample size for each species or species group and partial coverage stratum, using data from 2022.

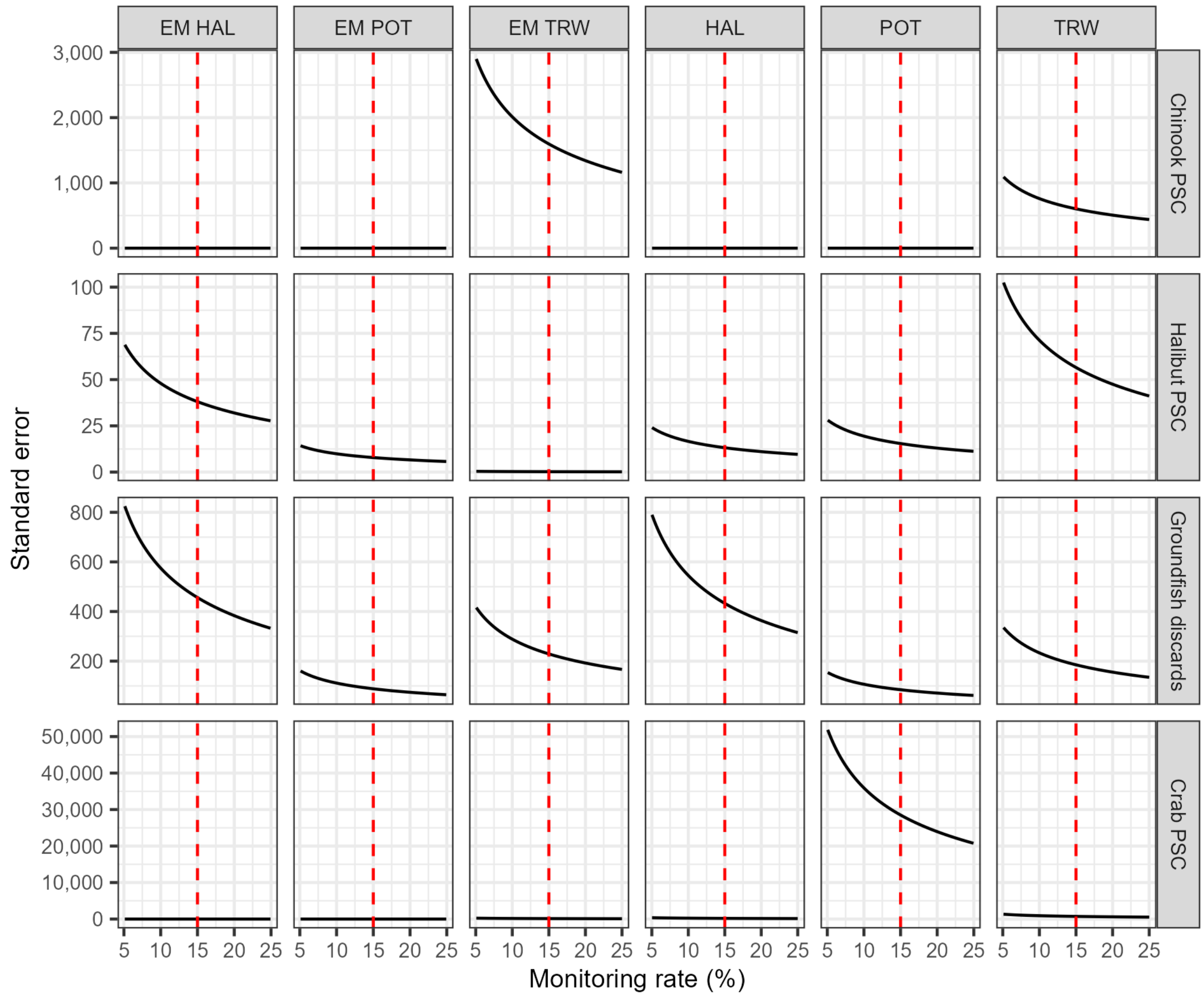

Figure 4-3. Trip-level standard error by monitoring rate for each species or species group and partial coverage stratum, using data from 2022. Dashed red lines indicate a 15% monitoring rate.

5. Results and Discussion

Design Summaries

Summaries of the number of trips, number of monitored trips, resulting coverage rates, and associated cost efficiencies are presented. The total and monitored efforts from each design are shown for different budget scenarios in Tables 5-1, 5-2, and 5-3. The *FMP* stratification resulted in the most sampling stratum, while the *Current* stratification has the least. Observer coverage rates are below 15% for all stratification and allocations at low budgets with the exception of some stratum with *Proximity* allocation. However, a design's performance or value should not be inferred from the sample rates allocated to each stratum; instead, the reader should understand how these allocations lead to the differences between designs in the evaluation metrics.

How different allocations and costs result in cost efficiency for three methods of monitoring is reported in Table 5-4. We can compare how each allocation differs from *Equal* allocation and how that impacts the cost per trip. For example, in the lowest budget, the amount of the total budget spent on monitoring with EMTRW is 16%, which is similar to the 13% of the budget spent on this monitoring method in the *Cost weighted boxes* and *Proximity* allocations. However, the *status quo* allocation is an outlier, spending over a third of the budget on EMTRW. The *status quo* allocation also puts over a third of the total budget into monitoring with EMFG (EM Fixed Gear). As a result, the *status quo* allocation has the greatest cost efficiency for EM monitoring, and the lowest cost efficiency for observers. Because EM costs drop so dramatically with an increase in sample size compared to observers (Figure 2-1), we would expect this allocation to result in the greatest number of monitored trips among allocations.

Evaluation Metrics

Trips Sampled

The number of trips sampled is a measure of monitoring output, and is an indirect measure of cost efficiency because all designs are evaluated under the same budget scenarios (*Samples:Biological* and *Samples:Composition* in Figures CS-1, CS-2 & CS-3). The largest differences between the proposed designs are between the *equal rates* and *status quo* allocation schemes. S*tatus quo* allocates more samples to the EM methods than the other designs, and *equal* allocates relatively fewer samples to the EM methods. *CWB* and *Proximity* are relatively similar in the number of samples allocated to each stratum. The current policy of 30% EM review on resulting observer coverage is evident by comparing observer coverage rates between the *status quo* and *equal* allocations. Observer coverage rates in *Status quo* allocation are halved from *equal* allocations at low budgets and reduced by over 25% at high budgets.

To allow more direct comparisons of the number of trips expected to be monitored in each design these values were summarized by monitoring method and gear type (Table 5-5). Note that within a budget level, while the number of expected samples collected under the *equal rates* and *status* *quo* allocation schemes do not vary under different stratifications. This is because the rates allocated to each monitoring method are identical and the stratifications do not affect which monitoring method a trip utilizes.

Variance in Expenses

The variance of expenditures (presented as coefficient of variation, or CV) were largely identical across designs with the exception of those utilizing the *status quo* allocation method at low budget levels that were less (Figure 5-1). This discrepancy is due to the fact that *status_quo* allocates a relatively larger portion of the budget to *EM_TRW*, which has a cost structure that is not dictated by the durations of sampled fishing trips because the monitoring is conducted at the offload. This cost efficiency of the *status quo* allocation diminishes as the budget increases from $3.5M to $5.25M. This is because the allocations to fixed-gear and trawl EM monitoring methods are static in the *status quo* allocation across the budget levels. As more money is available to allocate to observer monitoring, the variance in expenditures in the *status quo* allocation increases. Generally, as the budget increases, the budget coefficient of variation decreases for all designs (*Cost* in Figures CS-1, CS-2 & CS-3).

Power to Detect Rare Events

Figures 5-2, 5-3 and 5-4 illustrate the number of mammal and seabirds that could be expected to be killed in partial coverage strata in 2024 under each design and its power to detect this bycatch given that it occurs at rates seen in the past and the number of trips and monitored trips expected in 2024. These values aggregated for the design (all strata combined) are included in final summary tables.

Bycatch of Laysan albatross is much more common than for Steller sea lion or Short-tailed albatross. Laysan albatross can be expected to be caught as bycatch in EM HAL and Observer fixed gear in both FMPs. The average power to detect this species is 31.5% across designs at the low budget, 42.7% at the middle budget, and 49.6% at the high budget. However the number of high power test results (all for Laysan Albatross) is only 1 for the low budget, 5 for the middle budget and 7 for the high budget.

Only one short-tailed albatross would be expected to be encountered for the entire year. This species makes a good example of this analysis, because here power to detect is at its purest - what is the power to detect a single individual in the entire year given stratifications, sample allocations and fishing effort? The average power to detect a single individual of this species in the monitored partial coverage fleet is just 6% across designs at the low budget, 11.5% at the middle budget, and 15.3% at the high budget.

Bycatch of three Steller sea lions can be expected to occur in fixed gear and trawl gear in the Gulf of Alaska. Although estimates are only presented for observed stratum, we see no reason why bycatch would not also occur with similar gear on trips monitored with EM. This level of bycatch is quite low, and consequently the power to detect averaged only 7.5% across designs in the low budget, 12.5% in the middle budget, and 16.6% in the high budget.

In aggregate the power to detect Steller sea lions and Short-tailed albatross are so low and similar between designs they become uninformative (*Power to Detect* in Figures CS-1, CS-2 & CS-3). We can however look at power trends for Laysan albatross for clues about design choices. It appears that total bycatch differs between EM monitored stratum between FMPs, making stratifications that split FMPs more attractive than the *Current* stratification. The *FMP* stratification results in some stratum with a single bycatch event, which is very low, while the *Fixed FMP* stratification avoids this. Among allocations, the *Proximity* appears to perform best as evidenced by the appearance of high power tests as budgets increase.

Power to detect Monitoring Effects

Figures 5-5, 5-6 and 5-7 illustrate the effect size as a percentage of monitored trips and the resulting power of the test for three budgets. Across designs, the *FMP* and *Fixed FMP* stratifications highlight large monitoring effects in the OB_POT/FIXED-BSAI strata for trip duration and retained catch, and that these differences between FMPs are hidden or masked in the *Current* stratification (The relative effect size drops from 45-51% down to 20% because both FMPs are combined in the *Current* stratification).

The conventionally acceptable *minimum* power of 0.8 was selected to highlight (Krzywinski and Altman 2013). Power to detect results at this level of power were dismal at all funding levels with some improvements - 3 tests were above 0.8 at the lowest budget, 10 tests were above this level at the medium budget, and 13 tests were above this level at the high budget. All of these high power tests were in the metric *duration*, despite differences in landed catch relative to monitored trips of 45% in the FMP stratification (OB_POT-BSAI stratum). The *Fixed FMP* stratification with all but *Status quo* allocation appears attractive, because these designs isolate the monitoring effect and while large, allocate enough samples to adequately detect differences.

Data Timeliness

The main driver of results within the data timeliness metric is the amount of sampling expected within the fixed-gear EM strata. Because of the long review times associated with fixed-gear EM, and because those times scale with the number of trips expected to be reviewed, designs and budgets that allocate less sampling toward fixed-gear EM strata perform better in this metric. Within all stratifications and across all budget levels, *status quo* allocation performs the worst in this metric (*Data Timeliness* in Figures CS-1, CS-2 & CS-3), due to the fact that the fixed-gear EM sampling rate for that allocation strategy is set at 30%, which is higher than in any other design. *Equal rates* performs the best in this metric, and the *cost-weighted boxes* and *proximity* allocations perform similarly to each other, between *equal rates* and *status quo*.

Unlike most performance metrics, the results for data timeliness get worse as the budget increases. This is due to the fact that, with more funding, NMFS could afford more fixed-gear

EM sampling, thereby increasing the workload for video reviewers, and increasing the average time it takes for review. In this draft, we do not analyze a scenario in which funds are used to hire additional fixed-gear EM reviewers. In such a scenario, we would expect data timeliness scores to improve, and the scores of other metrics to worsen (relative to a scenario in which the number of fixed-gear EM video reviewers is kept at current levels), given that spending more on fixed-gear EM video review reduces the funding available for other sampling.

Trip-Level Variance

Results are mixed across the different metrics within the trip-level variance category (*Trip-Level Variance* in Figures CS-1, CS-2 & CS-3). Although *status quo* allocation is designed to minimize the type of variance that these metrics measure, none of the budget scenarios resulted in a situation in which observer days could be afforded above the 15% baseline that is a component of *status quo* allocation. Across all three budgets, *status quo* allocation reverts to equal rates for observer starata, while fixed-gear EM sampling is set at 30%, and shoreside sampling for trawl EM is set at 33%. If funding levels were sufficient to afford optimized days, we would expect *status quo* allocation to outperform other allocation strategies within this category of metrics. Across the budget levels we analyzed, no design appears to out-perform any other design across all trip-level variance metrics.

Interspersion

Gains and losses in interspersion for each data dependency group are presented in Figures CS-1, CS-2 & CS-3. As the budget increases, differences between the designs generally become less pronounced. When considering interspersion Alaska-wide (*Interspersion (AK)* in Figures CS-1, CS-2 & CS-3), within all three budget scenarios, the *equal rates* allocation method generally had lower EM_FIXED to EM_FIXED interspersion in comparison to the other interspersion metrics. In the low budget scenario, the *Status quo* allocation method had markedly higher interspersion in the OB to EM_TRW (0.996) and EM_FIXED to EM_FIXED (0.957) comparisons than in the rest of the comparisons (all others below 0.575). Interspersion scores from the *Cost-weighted boxes* and *Proximity* allocation methods generally responded similarly to changes in the budget and stratification, but *Cost-weighted boxes* had higher interspersion indices than *Proximity* in most cases. At low budgets, EM methods have very high costs, so *CWB* tends to allocate more to less expensive at-sea-observers, resulting in higher OB to EM and OB to EM_TRW indices. At higher budgets, at EM per-trip costs decrease, CWB tends to allocate more sample effort to EM strata and those patterns are reversed.

When looking at FMP-specific *Insterspersion* (*Interspersion (FMP)* in Figures CS-1, CS-2 & CS-3) under *Equal Rates* and *Status Quo* allocation methods, values are identical across stratification definitions. Since the minimum 15% coverage baseline is not achieved, *Status Quo* allocation reverts to *Equal Rates* allocation for both FMPs, resulting in no overall benefit. *Proximity* allocates more samples to the BSAI than the GOA compared to *CWB* due to smaller stratum size and high dispersion of fishing effort in space and time, resulting in greater interspersion of samples in the BSAI. This comes as a trade-off of interspersion in the GOA.

The gear-specific interspersion metrics that were summarized above to provide monitoring-level measures of interspersion are provided in Appendix B.

Table 5-1. Design summaries, including each stratum's allocated sample rate (as a percentage) and expected sample size (*n*), at the low budget level, $3.5M.

Budget: $3.5M			Allocation scheme							
			EQUAL		STATUS_QUO		CWB		PROX	
Stratification	Stratum	*N*	Rate	*n*	Rate	*n*	Rate	*n*	Rate	*n*
CURRENT	EM_HAL	722	5.85	42	30.00	217	6.77	49	8.77	63
	EM_POT	353	5.85	21	30.00	106	7.57	27	15.37	54
	EM_TRW	768	5.85	45	33.33	256	2.55	20	2.65	20
	OB_HAL	1,352	5.85	79	2.84	38	6.38	86	5.52	75
	OB_POT	1,086	5.85	64	2.84	31	5.51	60	6.44	70
	OB_TRW	389	5.85	23	2.84	11	7.95	31	6.42	25
FMP	EM_HAL-BSAI	32	5.85	2	30.00	10	6.34	2	38.88	12
	EM_HAL-GOA	690	5.85	40	30.00	207	6.56	45	6.33	44
	EM_POT-BSAI	57	5.85	3	30.00	17	8.77	5	18.81	11
	EM_POT-GOA	296	5.85	17	30.00	89	7.87	23	12.56	37
	EM_TRW-GOA	768	5.85	45	33.33	256	2.69	21	1.92	15
	OB_HAL-BSAI	106	5.85	6	2.84	3	7.45	8	23.87	25
	OB_HAL-GOA	1,246	5.85	73	2.84	35	5.71	71	3.83	48
	OB_POT-BSAI	255	5.85	15	2.84	7	5.94	15	7.01	18
	OB_POT-GOA	831	5.85	49	2.84	24	6.01	50	5.35	44
	OB_TRW-BSAI	21	5.85	1	2.84	1	12.42	3	31.01	7
	OB_TRW-GOA	368	5.85	22	2.84	10	7.59	28	4.34	16
FIXED_FMP	EM_FIXED-BSAI	89	5.85	5	30.00	27	9.00	8	29.58	26
	EM_FIXED-GOA	986	5.85	58	30.00	296	6.35	63	7.64	75
	EM_TRW-GOA	768	5.85	45	33.33	256	3.62	28	3.36	26
	OB_FIXED-BSAI	361	5.85	21	2.84	10	8.45	30	12.87	46
	OB_FIXED-GOA	2,077	5.85	122	2.84	59	5.16	107	3.96	82
	OB_TRW-BSAI	21	5.85	1	2.84	1	14.29	3	45.42	10
	OB_TRW-GOA	368	5.85	22	2.84	10	8.73	32	7.77	29

Table 5-2. Design summaries, including each stratum's allocated sample rate (as a percentage) and expected sample size (*n*), $4.5M.

Budget: $4.5M			Allocation scheme							
			EQUAL		STATUS_QUO		CWB		PROX	
Stratification	Stratum	N	Rate	n	Rate	n	Rate	n	Rate	n
CURRENT	EM_HAL	722	9.06	65	30.00	217	15.45	112	13.50	97
	EM_POT	353	9.06	32	30.00	106	17.34	61	23.14	82
	EM_TRW	768	9.06	70	33.33	256	5.44	42	4.67	36
	OB_HAL	1,352	9.06	122	6.07	82	9.78	132	8.60	116
	OB_POT	1,086	9.06	98	6.07	66	8.05	87	9.66	105
	OB_TRW	389	9.06	35	6.07	24	11.93	46	11.22	44
FMP	EM_HAL-BSAI	32	9.06	3	30.00	10	19.42	6	54.45	17
	EM_HAL-GOA	690	9.06	63	30.00	207	13.87	96	9.94	69
	EM_POT-BSAI	57	9.06	5	30.00	17	21.98	13	31.23	18
	EM_POT-GOA	296	9.06	27	30.00	89	18.05	53	19.52	58
	EM_TRW-GOA	768	9.06	70	33.33	256	5.82	45	3.22	25
	OB_HAL-BSAI	106	9.06	10	6.07	6	12.76	14	35.15	37
	OB_HAL-GOA	1,246	9.06	113	6.07	76	8.32	104	5.96	74
	OB_POT-BSAI	255	9.06	23	6.07	15	8.99	23	11.28	29
	OB_POT-GOA	831	9.06	75	6.07	50	9.01	75	8.35	69
	OB_TRW-BSAI	21	9.06	2	6.07	1	18.38	4	44.29	9
	OB_TRW-GOA	368	9.06	33	6.07	22	11.40	42	7.47	27
FIXED_FMP	EM_FIXED-BSAI	89	9.06	8	30.00	27	23.15	21	44.89	40
	EM_FIXED-GOA	986	9.06	89	30.00	296	12.47	123	11.30	111
	EM_TRW-GOA	768	9.06	70	33.33	256	7.65	59	6.00	46
	OB_FIXED-BSAI	361	9.06	33	6.07	22	13.89	50	20.45	74
	OB_FIXED-GOA	2,077	9.06	188	6.07	126	7.46	155	5.89	122
	OB_TRW-BSAI	21	9.06	2	6.07	1	20.93	4	61.66	13
	OB_TRW-GOA	368	9.06	33	6.07	22	13.21	49	13.20	49

Table 5-3. Design summaries, including each stratum's allocated sample rate (as a percentage) and expected sample size (n), \$5.25M.

Budget: \$5.25M			Allocation scheme							
			EQUAL		STATUS_QUO		CWB		PROX	
Stratification	Stratum	N	Rate	n	Rate	n	Rate	n	Rate	n
CURRENT	EM_HAL	722	11.75	85	30.00	217	22.56	163	17.25	125
	EM_POT	353	11.75	41	30.00	106	24.51	87	28.88	102
	EM_TRW	768	11.75	90	33.33	256	8.13	62	6.87	53
	OB_HAL	1,352	11.75	159	8.78	119	12.68	171	11.16	151
	OB_POT	1,086	11.75	128	8.78	95	10.07	109	12.15	132
	OB_TRW	389	11.75	46	8.78	34	15.20	59	15.65	61
FMP	EM_HAL-BSAI	32	11.75	4	30.00	10	31.02	10	64.45	21
	EM_HAL-GOA	690	11.75	81	30.00	207	19.32	133	12.98	90
	EM_POT-BSAI	57	11.75	7	30.00	17	33.07	19	42.01	24
	EM_POT-GOA	296	11.75	35	30.00	89	25.53	76	25.14	74
	EM_TRW-GOA	768	11.75	90	33.33	256	8.77	67	4.68	36
	OB_HAL-BSAI	106	11.75	12	8.78	9	17.38	18	43.65	46
	OB_HAL-GOA	1,246	11.75	146	8.78	109	10.44	130	7.76	97
	OB_POT-BSAI	255	11.75	30	8.78	22	11.53	29	15.14	39
	OB_POT-GOA	831	11.75	98	8.78	73	11.47	95	10.91	91
	OB_TRW-BSAI	21	11.75	2	8.78	2	22.79	5	54.72	11
	OB_TRW-GOA	368	11.75	43	8.78	32	14.61	54	10.62	39
FIXED_FMP	EM_FIXED-BSAI	89	11.75	10	30.00	27	35.55	32	55.21	49
	EM_FIXED-GOA	986	11.75	116	30.00	296	17.15	169	14.18	140
	EM_TRW-GOA	768	11.75	90	33.33	256	11.22	86	8.80	68
	OB_FIXED-BSAI	361	11.75	42	8.78	32	18.60	67	26.60	96
	OB_FIXED-GOA	2,077	11.75	244	8.78	182	9.30	193	7.43	154
	OB_TRW-BSAI	21	11.75	2	8.78	2	25.67	5	71.70	15
	OB_TRW-GOA	368	11.75	43	8.78	32	16.91	62	18.10	67

Table 5-4. Design summaries grouped by monitoring method, including the proportion of funds allocated to each monitoring method, cost per trip ($), and expected sample sizes (*n*).

			Allocation scheme											
			EQUAL			STATUS_QUO			CWB			PROX		
Budget	Stratification	Monitoring Method	% of Budget	Cost per trip	n	% of Budget	Cost per trip	n	% of Budget	Cost per trip	n	% of Budget	Cost per trip	n
$3.5M	CURRENT	EMFG	29.4	16,337	63	35.3	3,829	323	29.6	13,724	76	30.6	9,112	118
		EMTRW	16.0	12,471	45	36.9	5,039	256	13.5	24,136	20	13.6	23,356	20
		OB	54.6	11,548	165	27.9	12,147	80	56.9	11,256	177	55.8	11,517	170
	FMP	EMFG	29.4	16,373	63	35.3	3,834	322	29.6	13,721	76	30.5	10,228	104
		EMTRW	16.0	12,471	45	36.9	5,039	256	13.6	23,060	21	13.0	30,921	15
		OB	54.6	11,538	166	27.9	12,149	80	56.8	11,380	175	56.5	12,490	158
	FIXED_FMP	EMFG	29.4	16,316	63	35.3	3,830	322	29.5	14,638	71	30.4	10,452	102
		EMTRW	16.0	12,471	45	36.9	5,039	256	14.3	18,023	28	14.1	19,151	26
		OB	54.6	11,541	166	27.9	12,135	80	56.2	11,369	173	55.6	11,627	167
$4.5M	CURRENT	EMFG	23.4	10,831	97	27.4	3,829	323	24.8	6,455	173	24.9	6,256	179
		EMTRW	14.3	9,277	70	28.7	5,039	256	12.2	13,150	42	11.8	14,748	36
		OB	62.2	10,923	256	43.9	11,498	172	63.0	10,659	266	63.4	10,769	265
	FMP	EMFG	23.4	10,840	97	27.4	3,834	322	24.7	6,618	168	24.8	6,895	162
		EMTRW	14.3	9,277	70	28.7	5,039	256	12.4	12,517	45	10.9	19,833	25
		OB	62.2	10,935	256	43.9	11,494	172	62.8	10,855	261	64.2	11,732	246
	FIXED_FMP	EMFG	23.4	10,830	97	27.4	3,830	322	24.3	7,627	143	24.6	7,304	151
		EMTRW	14.3	9,277	70	28.7	5,039	256	13.5	10,350	59	12.5	12,245	46
		OB	62.2	10,930	256	43.9	11,525	171	62.2	10,839	258	63.0	10,990	258
$5.25M	CURRENT	EMFG	20.5	8,534	126	23.5	3,829	323	22.4	4,718	249	22.1	5,116	226
		EMTRW	13.7	7,944	90	24.6	5,039	256	11.8	9,943	62	11.2	11,132	53
		OB	65.8	10,392	332	51.9	10,977	248	65.8	10,160	340	66.7	10,185	344
	FMP	EMFG	20.5	8,549	126	23.5	3,834	322	22.3	4,925	238	22.0	5,551	208
		EMTRW	13.7	7,944	90	24.6	5,039	256	12.1	9,469	67	10.1	14,724	36
		OB	65.8	10,395	332	51.9	10,969	248	65.6	10,359	332	67.9	11,047	323
	FIXED_FMP	EMFG	20.5	8,525	126	23.5	3,830	322	21.7	5,690	201	21.7	6,016	189
		EMTRW	13.7	7,944	90	24.6	5,039	256	13.4	8,156	86	12.2	9,449	68
		OB	65.8	10,398	332	51.9	10,984	248	64.9	10,394	328	66.1	10,446	332

Table 5-5. Count of expected number of trips sampled, grouped by gear type, by each sample design (stratification and allocation combination) and budget.

			$3.5M				$4.5M				$5.25M			
Stratification	Monitoring Method	Gear	EQUAL	STATUS_QUO	CWB	PROX	EQUAL	STATUS_QUO	CWB	PROX	EQUAL	STATUS_QUO	CWB	PROX
CURRENT	OB	HAL	79	38	85	75	122	81	130	116	158	118	169	150
		POT	64	31	61	70	99	66	89	105	129	96	112	132
		TRW	23	11	31	25	35	24	46	44	46	34	59	61
	EM_FG	HAL	42	215	49	64	65	215	111	99	84	215	162	127
		POT	21	107	27	53	32	107	61	80	42	107	87	100
	EM_TRW	TRW	45	256	20	20	70	256	42	36	90	256	62	53
FMP	OB	HAL	79	38	79	73	122	81	117	112	158	118	148	143
		POT	64	31	66	62	99	66	98	98	129	96	125	129
		TRW	23	11	31	22	35	24	46	37	46	34	59	51
	EM_FG	HAL	42	215	47	57	65	215	102	87	84	215	144	112
		POT	21	107	28	47	32	107	66	75	42	107	94	97
	EM_TRW	TRW	45	256	21	15	70	256	45	25	90	256	67	36
FIXED_FMP	OB	HAL	79	38	73	63	122	81	107	95	158	118	135	120
		POT	64	31	65	66	99	66	98	102	129	96	126	130
		TRW	23	11	35	38	35	24	53	62	46	34	68	82
	EM_FG	HAL	42	215	46	62	65	215	93	92	84	215	129	115
		POT	21	107	24	40	32	107	51	60	42	107	72	74
	EM_TRW	TRW	45	256	28	26	70	256	59	46	90	256	86	68

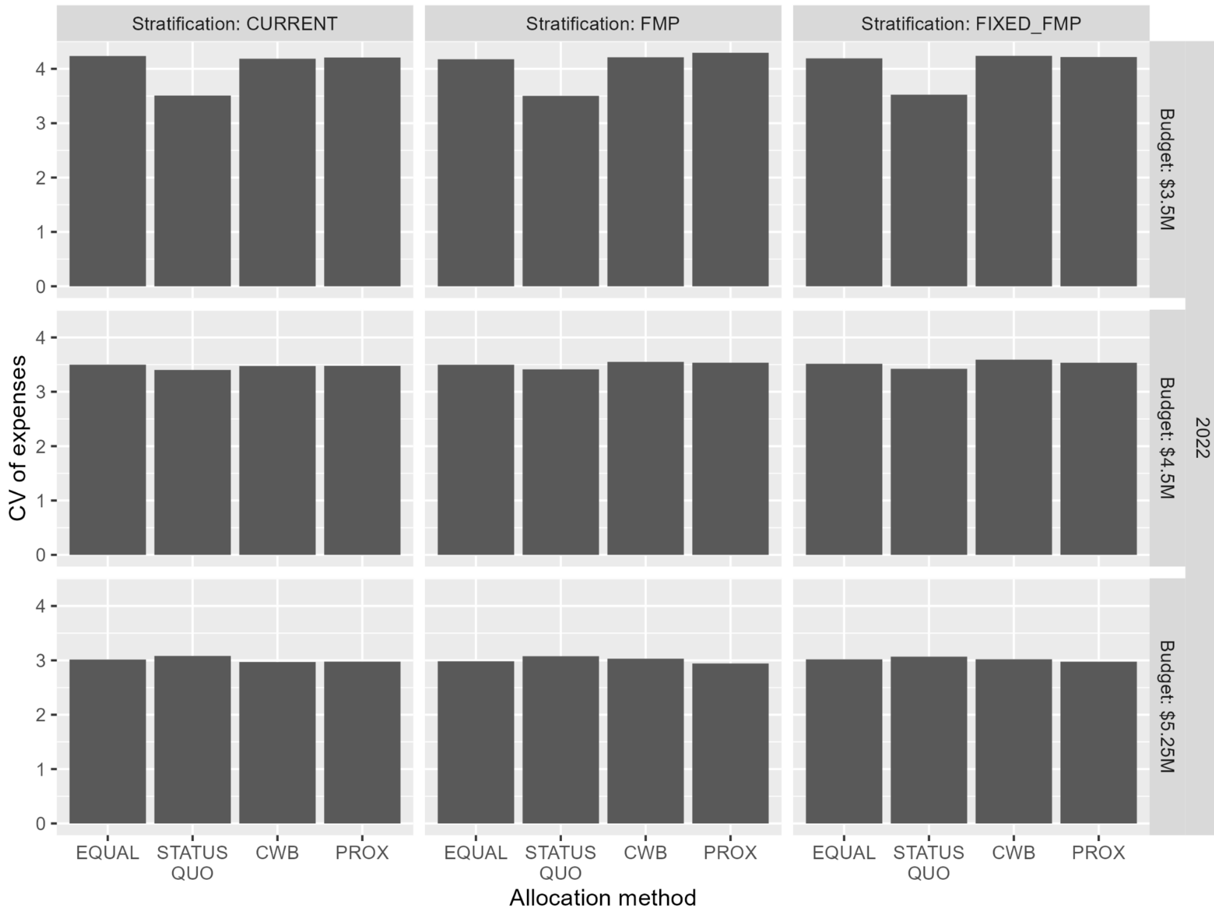

Figure 5-1. Coefficient of variation of monitoring expenditures from 10,000 simulated deployments of each of the monitoring designs.

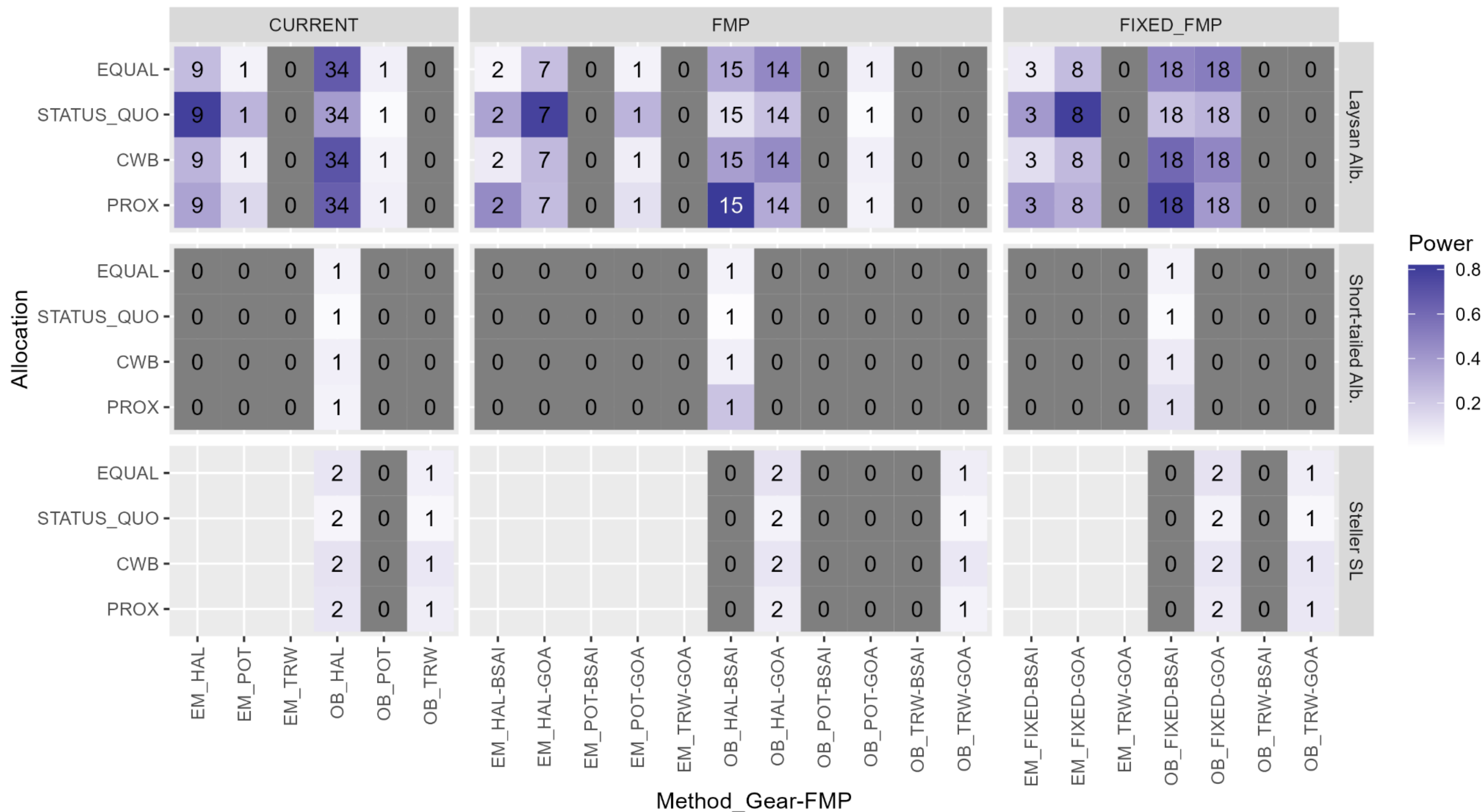

Figure 5-2. Results of power analysis for monitoring effects under a $3.5M budget. Values in each cell denote the bycatch for 2024, estimated from the mean bycatch per unit effort and the number of trips in each stratum. The resulting power to detect bycatch given it is present at rates in the past are shown as colors in each cell. Higher power is desirable. Cells with a power greater or equal to 0.8 are denoted with white text. Bycatch values from past data were not possible to calculate from EM stratum for Steller Sea Lion.

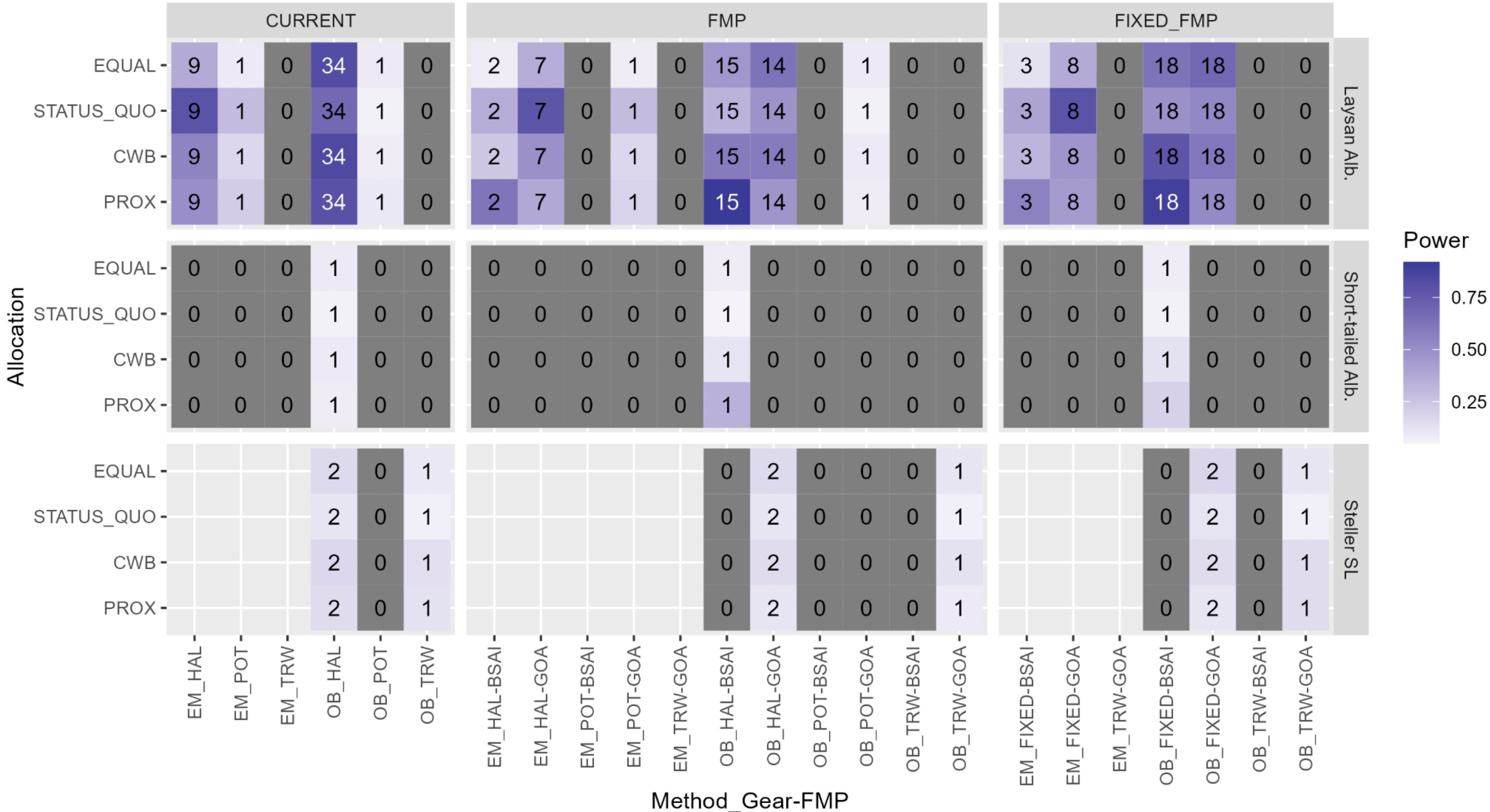

Figure 5-3. Results of power analysis for monitoring effects under a $4.5M budget. Values in each cell denote the bycatch for 2024, estimated from the mean bycatch per unit effort and the number of trips in each stratum. The resulting power to detect bycatch given it is present at rates in the past are shown as colors in each cell. Higher power is desirable. Cells with a power greater or equal to 0.8 are denoted with white text. Bycatch values from past data were not possible to calculate from EM stratum for Steller Sea Lion.

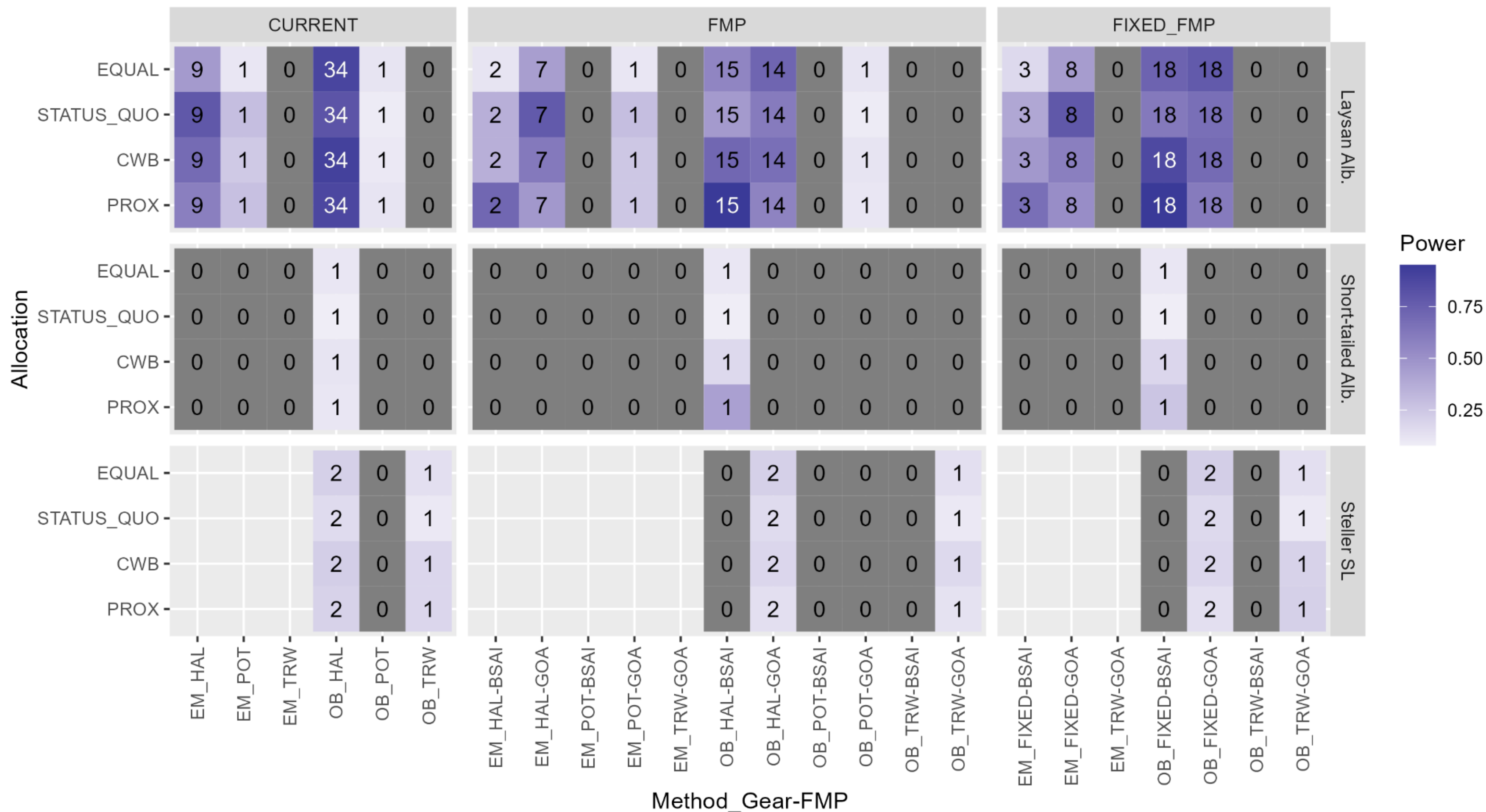

Figure 5-4. Results of power analysis for monitoring effects under a $5.25M budget. Values in each cell denote the bycatch for 2024, estimated from the mean bycatch per unit effort and the number of trips in each stratum. The resulting power to detect bycatch given it is present at rates in the past are shown as colors in each cell. Higher power is desirable. Cells with a power greater or equal to 0.8 are denoted with white text. Bycatch values from past data were not possible to calculate from EM stratum for Steller Sea Lion.

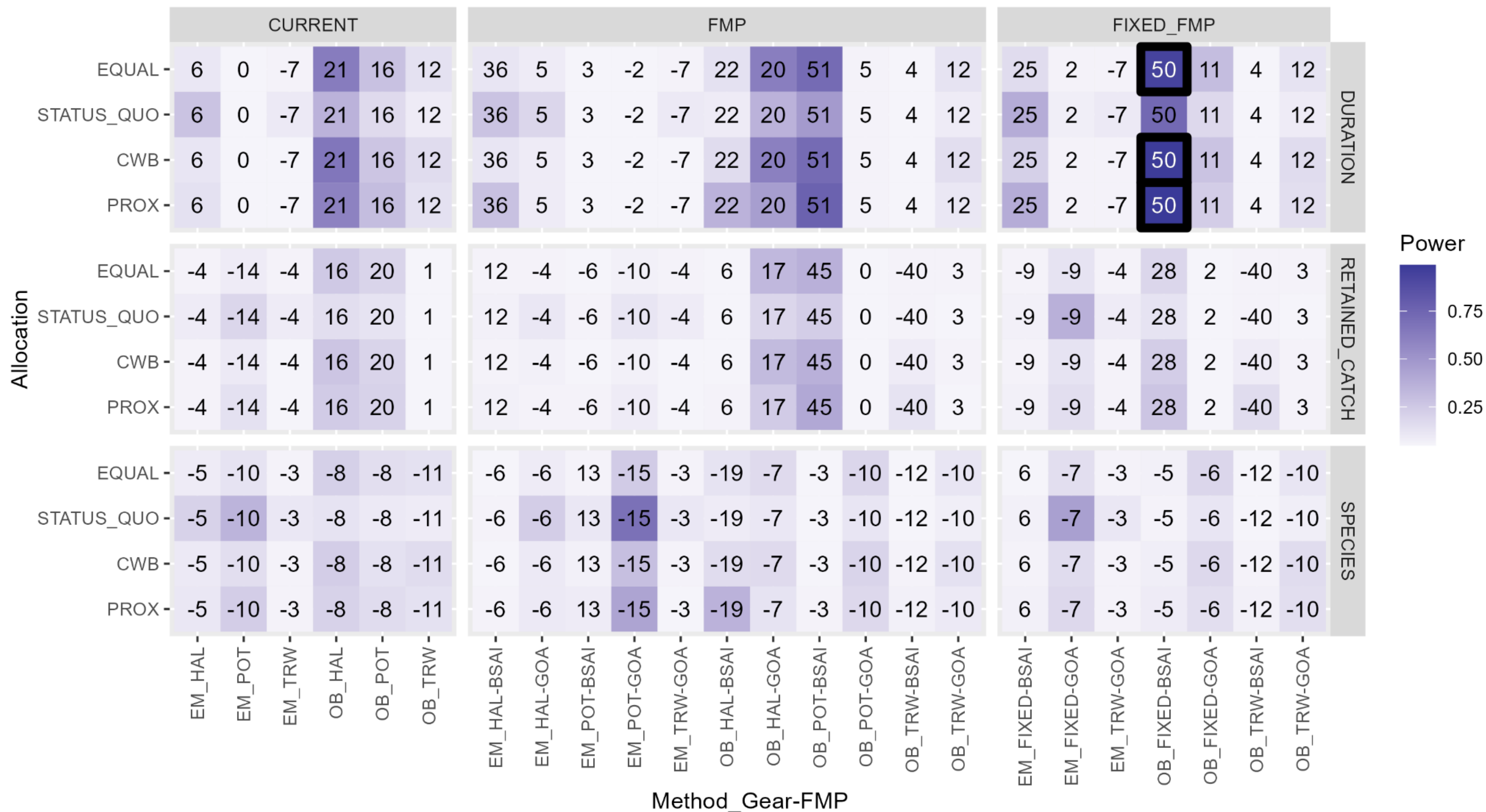

Figure 5-5. Results of power analysis for monitoring effects under a $3.5M budget. Values in each cell denote the effect size, or the mean from monitored trips subtracted from the mean of unmonitored trips (expressed as a percentage of the mean of monitored trips, where negative values indicate monitored trips were greater, while positive values indicate unmonitored trips were greater). The resulting power to detect differences of the effect size given the expected total number of trips and expected number of sampled trips from each design in 2024 are shown as colors in each cell. Higher power is desirable. Cells with a power greater or equal to 0.8 are highlighted.

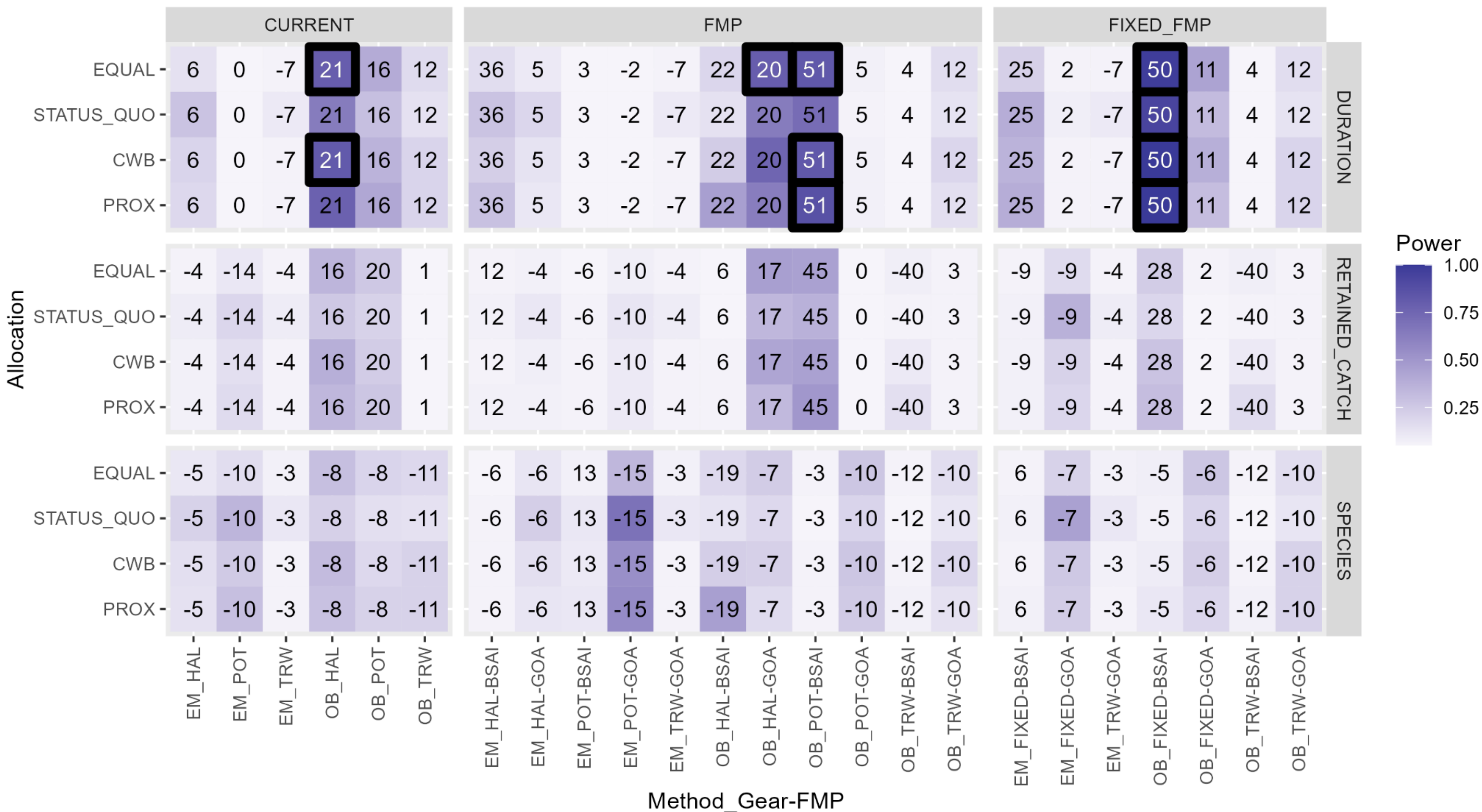

Figure 5-6. Results of power analysis for monitoring effects under a $4.5M budget. Values in each cell denote the effect size, or the mean from monitored trips subtracted from the mean of unmonitored trips (expressed as a percentage of the mean of monitored trips, where negative values indicate monitored trips were greater, while positive values indicate unmonitored trips were greater). The resulting power to detect differences of the effect size given the expected total number of trips and expected number of sampled trips from each design in 2024 are shown as colors in each cell. Higher power is desirable. Cells with a power greater or equal to 0.8 are highlighted.

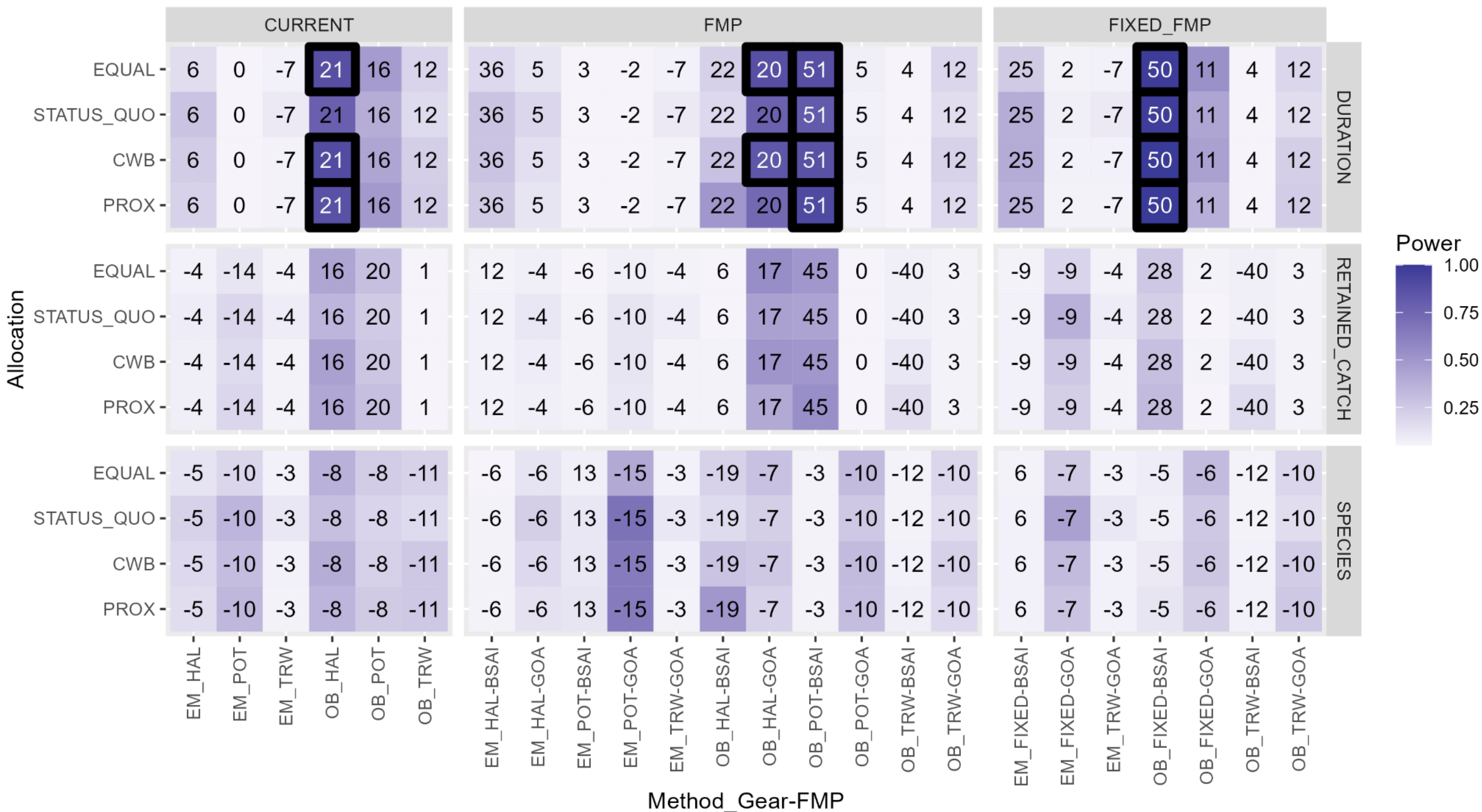

Figure 5-7. Results of power analysis for monitoring effects under a $5.25M budget. Values in each cell denote the effect size, or the mean from monitored trips subtracted from the mean of unmonitored trips (expressed as a percentage of the mean of monitored trips, where negative values indicate monitored trips were greater, while positive values indicate unmonitored trips were greater). The resulting power to detect differences of the effect size given the expected total number of trips and expected number of sampled trips from each design in 2024 are shown as colors in each cell. Higher power is desirable. Cells with a power greater or equal to 0.8 are highlighted.

Comparative Summary of Results

Stratification

This analysis evaluated four stratification definitions: the *Current* (2023) definition based on gear (hook-and-line, pot, trawl) and monitoring type (observer, EM), as well as stratification definitions that included FMP (BSAI, GOA), and two additional fixed-gear stratification definitions that combined fixed gears (HAL, POT) into a single strata to account for trips fishing in both gear types on a single trip.

To increase the chances of monitoring trips in both the BSAI and the GOA in a more efficient manner, we examined including FMP in alternative stratification definitions. Stratifying by monitoring tool (observers and EM), gear type (hook-and-line, pot, trawl) and FMP (where the BSAI was separate from the GOA) resulted in more strata, each with fewer trips but generally enough to provide a reasonable likelihood of being sampled. However, further splitting the BS and AI resulted in a larger number of strata, some of which had very low total effort and were likely to be entirely missed by fisheries monitoring (e.g., see Table Methods-FMP-strata:1).

The use of gear type-based stratification definitions did not adhere to statistical rules of stratification where each sampling unit (i.e., fishing trip) could only exist in one stratum. Although stratification by gear type has been used in past ADPs, fixed gear usage is changing; approximately 15% of observer-pool and 20% EM-pool fixed gear trips fished both hook-and-line and pot gear on a single trip in 2022. These trips cannot be unambiguously placed into strata defined by the use of a single gear type, and as a result, use of standard estimation methods may produce biased estimates. Two alternative strategies were evaluated to correct this issue: combine all trips that fish with hook-and-line, pot or both gear types on a single trip into a single 'fixed-gear' stratum, or isolate trips that fish with both gear types into separate stratum. Ultimately, a version of the former strategy, *Fixed FMP*, was chosen as a promising stratification definition because it addressed the issue of assigning trips fishing with multiple gear types to strata while also categorizing trips by monitoring tool and FMP without creating strata with low effort or high likelihood to be unmonitored. Analysts recommend use of the *Fixed FMP* stratification for 2024.

Allocation

Four alternatives for how fisheries monitoring assets are allocated among strata were explored. These included *Equal rates*, *Status quo*, *Cost-weighted boxes*, and *Proximity* allocation. *Equal rates* provides unbiased estimation from samples in the case where there is little to no prior information about the fishery. *Equal rates* are presented to provide a baseline from which to evaluate other designs since we can use information from the fishery to better inform our allocation strategy. *Equal rates* by default, is an allocation that is not affected by changes to stratification and therefore cannot differentially allocate samples to FMPs.

The *Status quo* allocation sets rates through a baseline + optimization algorithm for observers and by policy for EM. The budgets explored ($3.5, $4.5, and $5.25 million, assuming fee revenues only) were not sufficient to provide optimized *Status quo* allocation for observers resulting in many more EM sampled trips compared to trips with observers. This also resulted in little to no differences under the alternative stratification definitions because when monitoring rates are under the 15% minimum, the observer strata are allocated equally. Large EM allocation improves cost efficiency of the overall program by reducing the variable cost of monitoring trips of different durations and EM cost efficiency improves more with increased sample size than does at-sea observer coverage. The *Status quo* allocation results in large amounts of EM sampled trips which contributes to the large number of trips sampled overall. The *Status quo* allocation also results in the lowest CV for chinook PSC. The main problem with the *Status quo* allocation is that resulting observer coverage is low and the allocation results in little overlap between (or interspersion of) observed trips and the EM monitored fleets and between the observed and the unmonitored fleet. Because at-sea observers collect the full suite of data while other monitoring methods collect only a portion of the data elements, this lack of *interspersion* is means that and analyses using data collected by other monitoring methods rely of observer data for those missing elements (i.e. mean weights used to convert count data into catch weight). The *Status quo* allocation method resulted in the fewest observer samples collected at-sea, meaning less age, length, maturity, and stock of origin data will be available for use in stock assessments and stock of origin (genetics) analyses. The *Status quo* allocation resulted in the highest variability in PSC estimates of Pacific halibut from trawl gear and crab PSC. In addition, the review of imagery from fixed gear EM collected at-sea is too slow to be of any practical use for in-season management of quotas.

Alternative allocation methods to *Status quo* were developed to provide more robust data and improve the cost-efficiency and scientific merit of fisheries monitoring. Like the optimization portion of observer coverage under the *Status quo* allocation method, the *Proximity* and *Cost-weighted boxes* allocation methods employ algorithms to prioritize sampling strata that are expected to otherwise result in datagaps. However unlike *Status quo*, these allocation methods integrate EM into the deployment process, treating EM strata in the same way as observer strata. The allocation methods differ in that *Cost-weighted boxes* also prioritizes the utilization of cheaper monitoring methods whereas *Proximity* also prioritizes the sampling of smaller strata with higher risk of sample size issues. Their performance was similar with relatively good interspersion, ability to detect monitoring effects, improvements in data timeliness, and increased CVs of halibut and crab PSC, but had a relatively poor CV of chinook PSC. Analysts believe that the slightly better performance of *Proximity* over *Cost-weighted boxes* allocations seen in this document are caused by the fact that there is cost inefficiency built into the voluntary nature of EM participation. If the pool of EM vessels was pared down to include those that regularly fish, the high fixed costs of EM equipment installation and maintenance can be overcome and cost efficiencies can be maximized. Under these conditions analysts believe that the *Cost-weighted boxes* allocation has the potential to outperform *Proximity* allocation. Analysts do not recommend the designs employing the *Status quo* allocation method or *Equal rates* allocation method due to low observer coverage and their inability to differentially allocate samples to FMPs across the range of budgets evaluated.

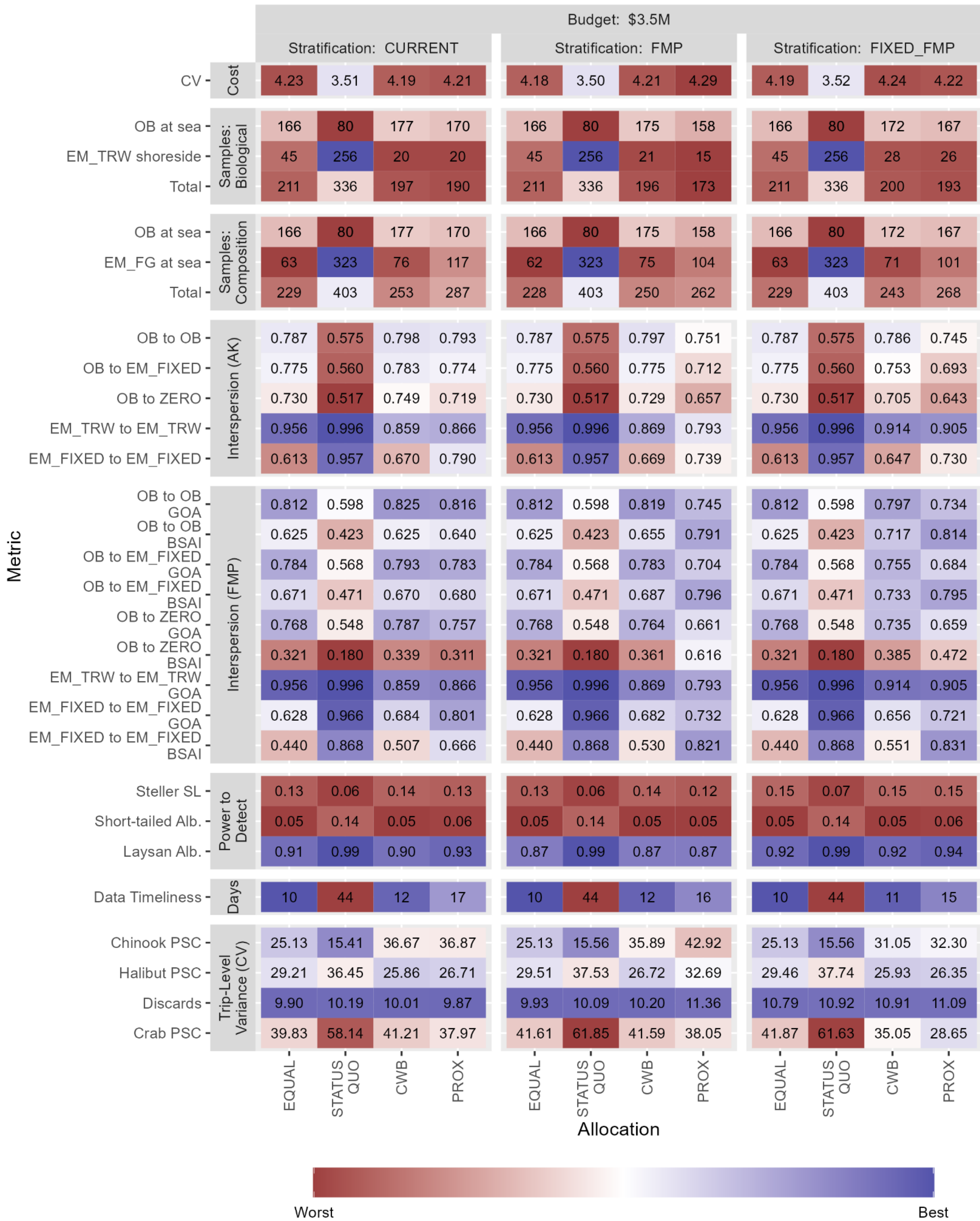

Figure CS-1. Summary of evaluation metrics for the low budget scenario. Colors are scaled within a suite of metrics (e.g. Cost, Samples Biological, Samples Composition, Interspersion) across budgets.

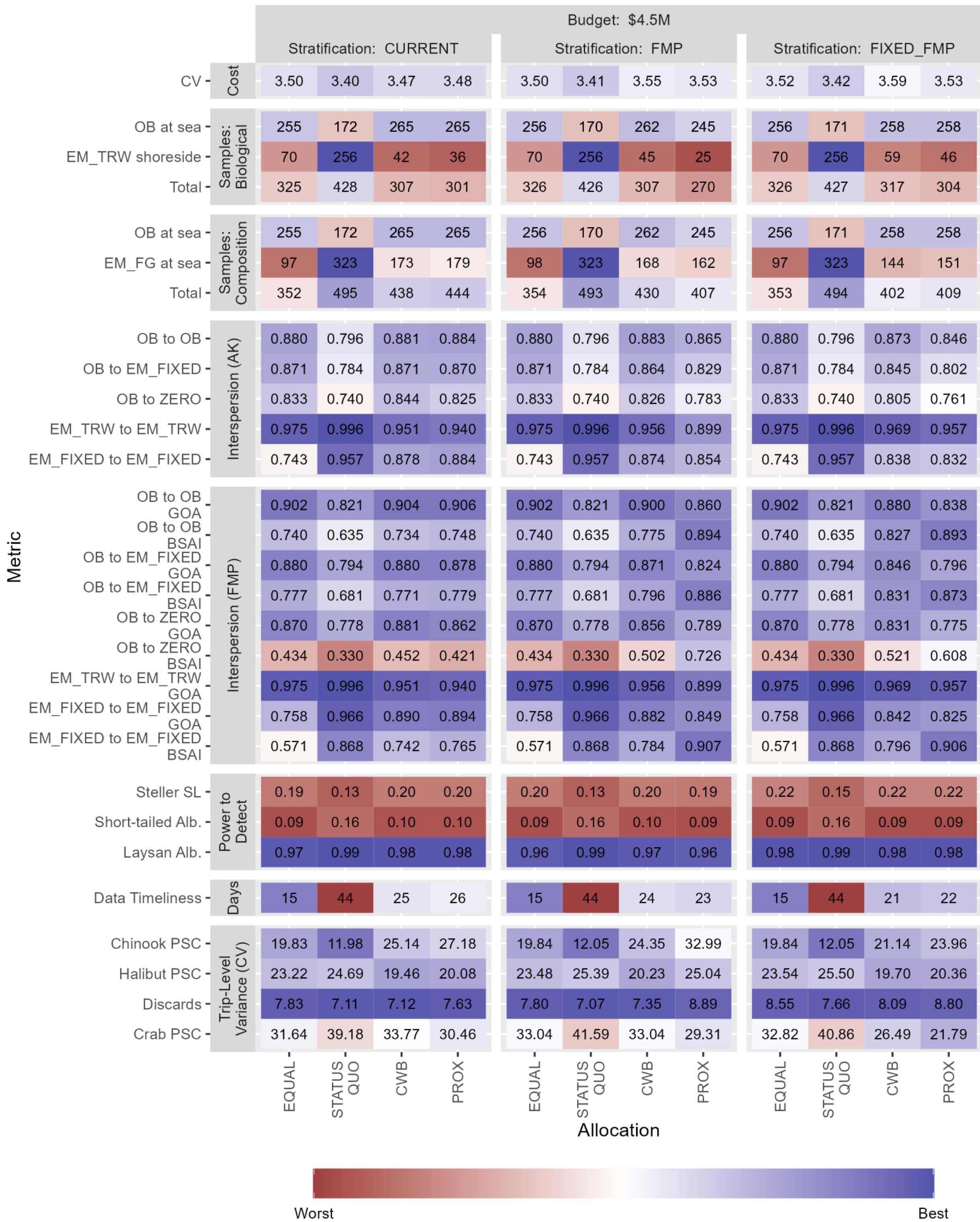

Figure CS-2. Summary of evaluation metrics for the moderate budget scenario. Colors are scaled within a suite of metrics (e.g. Cost, Samples Biological, Samples Composition, Interspersion) across budgets.

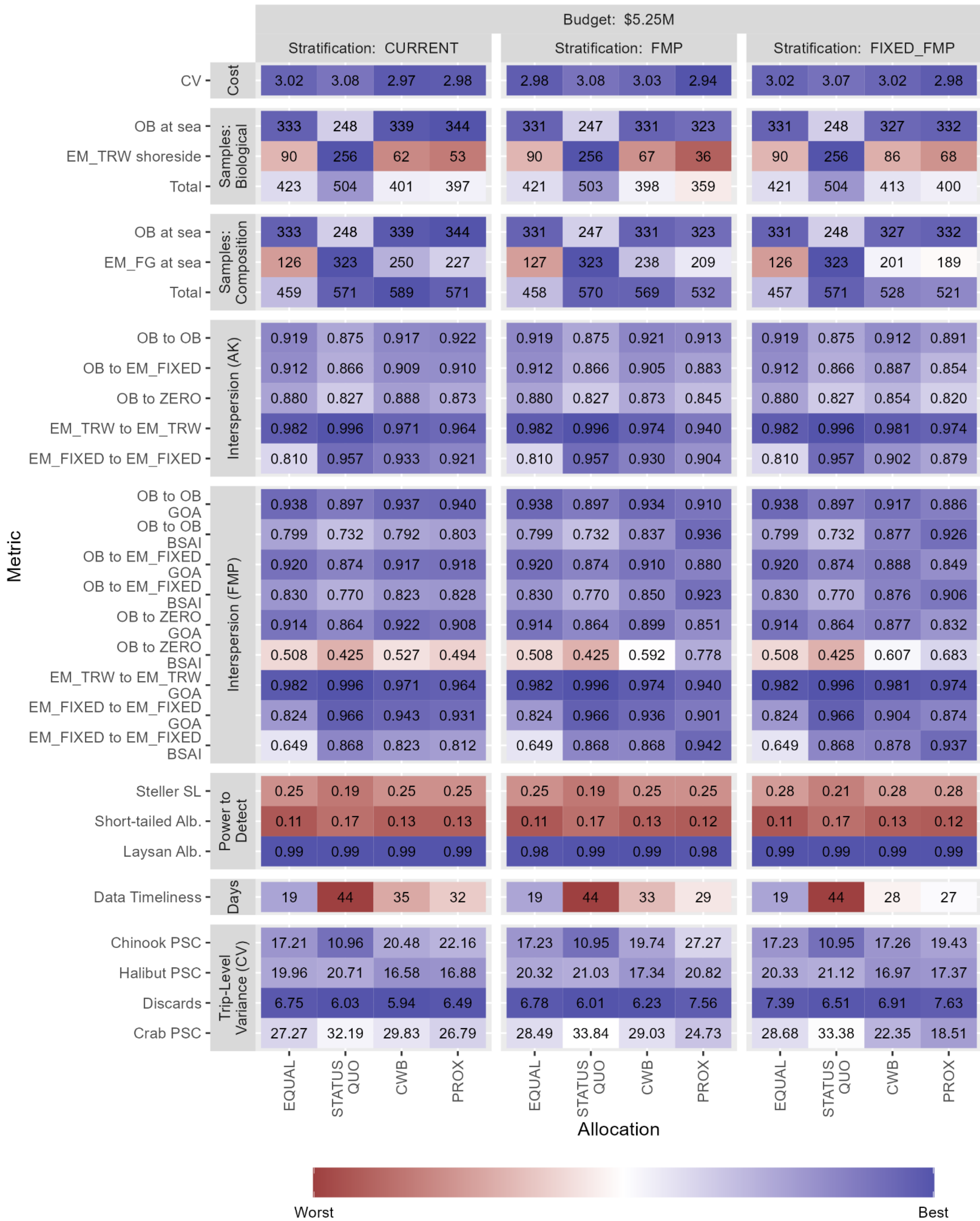

Figure CS-3. Summary of evaluation metrics for the high budget scenario. Colors are scaled within a suite of metrics (e.g. Cost, Samples Biological, Samples Composition, Interspersion) across budgets.

6. Additional Cost Efficiency Considerations and Ideas

Zero Selection

The Zero Selection pool is composed of vessels that will have no probability of carrying an observer on any trips. The definition of Zero Selection needs to use criteria that are identified ahead of time and are predictable from year to year. Currently, vessels are placed in Zero Selection based on the vessel's size and gear – hook and line and pot vessels under 40 ft and jig vessel regardless of length are placed in Zero Selection; hook and line and pot vessels 40 ft and over (and all trawl) are included in the sampling frame.

Increasing the number of vessels in Zero Selection would reduce the number of monitored vessels, thereby increasing the selection rate on the remaining vessels, but potentially not changing the total number of monitored trips. However, data quality would be reduced. This happens because as more vessels move into Zero Selection, the data being collected on monitored vessels is less representative of true fishing behavior. Removing vessels that take very few trips per year from the EM pool and adding them to Zero Selection could improve the efficiency of the EM program. However, the impact of these changes on observer deployment rates is unclear; a large number of vessels would need to be moved to Zero Selection to substantially increase monitoring rates in other strata. At less than 100% coverage, it is unknown what effect increases in deployment rates would have on the presence of monitoring effects.

Observer Procurement & Duties

In meetings with industry, a regular topic of conversation has been the potential cost efficiencies that might be realized by procuring observers in a way other than the current contract-based system, or modifying the structure of the current contract, or changing observer duties to take advantage of "down time" to increase cost efficiency. This section outlines those ideas and Table 5-1 summarizes their status.

Hiring Observers as Federal Employees

We worked with an economist to estimate the cost of hiring observers as federal employees. To do so, we first identified the salary that federal observers would start at, which is equivalent to a GS 5-1 at $44,649 annually. To account for benefits, we set total compensation equal to $72,331, based on an analysis by the Bureau of Labor Statistics, which found that 38% of the compensation to local and state government employees came in the form of benefits (BLS 2023). Dividing that amount by 52.25 weeks per year, a weekly rate of $2,064 would be paid to observers, assuming a 60-hour work week and a 50% increase in the hourly pay rate beyond 40 hours per week.

To estimate the costs for federal observers in low and high sampling effort scenarios, we multiplied the weekly salary by the 75th percentile and 90% of the maximum number of partial coverage observers that have been deployed by month, on average, from 2013 through 2022. We then summed these monthly costs across the year, including training costs. We calculated training costs at 40 hours per week, assuming a 70% retention rate. We assumed that new hires would require 3 weeks of training and returning hires would require 1 week of training. We then added the cost of per diem and travel within Alaska (airfare, baggage fees, and ground transportation), based on past invoices from the partial coverage observer provider. Finally, we added the estimated cost of airfare, including extra baggage fees, between Seattle and Alaska (assumed $1,500 per round trip) for the number of unique deployments expected from each end of the sampling effort range. This process is summarized in Table 5-2.

Under a model in which observers are hired as federal employees, observers would be paid a salary funded from the landing fee, and their supervisors would be paid using separate federal funds, as the landing fee must "not be used to pay any costs of administrative overhead or other costs not directly incurred in carrying out the [fisheries research] plan" (16 U.S.C. 1862[b][2][c]). Therefore, in order to hire observers as federal employees, federal funds separate from the landing fee would have to be identified for supervisors. Note also that the estimates provided here were based on 2 supervisors working a total of 20.5 months during the year at a ZP3-1 pay scale with 38% of total compensation coming from benefits (supervisors are estimated to cost $8,691.84/month of time). Observer provider input from the FMAC indicated that, from their experience, this number of supervisors was too low. If observers were hired as federal employees, the results presented in Table 6-2 suggest that the Observer Program may be able to realize a reduction in cost per observed sea day. The low sampling effort cost per day estimate ($1,260) presented in Table 5-2 is 9%, 10% and 16% and lower than the costs per day from the 2020, 2021, and 2022 Annual Reports ($1,381 , $1,393, and $1,492, respectively). The high sampling effort estimate ($1,237) is 10%, 11%, and 17% lower. Doubling the amount of supervision resulted in federal observer costs per day of $1,319 for the low sampling effort scenario and $1,276 for the high sampling effort scenario.

Multi-Provider / Voucher Program to Procure Observers

An idea that has been discussed to potentially create cost efficiencies would be to enable partial coverage vessels to procure observers directly from observer providers. NMFS would then use the observer fee to reimburse vessels for coverage at a set daily amount rather than using the fees to fund a federal contract with an observer provider company, as is currently the case. Under this "voucher" approach, a vessel owner would be responsible for securing an observer to monitor trips that were selected in ODDS. The observer providers would charge a market rate that encompasses the daily rate to cover that vessel's trip, as well as associated variable costs (travel and board). If the market rate exceeds the fixed daily rate dollar value ascribed to the voucher, the vessel owner selected for coverage would pay the difference directly to the provider.

In 2017, the Observer Advisory Committee reviewed a discussion paper (NPFMC 2017; section 3.5) that evaluated the multi-provider / voucher approach. The paper outlined legal issues, explained the complication of setting a voucher amount that is equitable, and discussed ways that it could introduce bias. The idea was discussed again in PCFMAC in March 2022[7] and at that time the committee did not want to divert NMFS staff resources for a new task and recommended that if a discussion paper was proposed and initiated by the Council, it be developed by Council staff and considered separately from the Cost Efficiencies Analysis/2024 ADP.

NMFS is not currently exploring observer cooperatives, voucher programs, or any type of multi-provider approaches. Changing how observers are procured would not change the sampling design (i.e., how data are collected), rather, the extent of resources (i.e., number of sea days, number of EM vessels that could be afforded, etc.) within the sampling design could potentially change. However, thus far, NMFS has not seen evidence that lower day rates would occur as a result of a multi-provider approach. The 2021 annual report (NMFS 2022) illustrates the relationship between the fully loaded cost per invoiced day for full observer coverage as a function of the number of days invoiced[8]. Compared to a partial coverage observer that may be deployed onto multiple vessels for one to five days at a time, an observer deployed onto the majority of full coverage vessels boards once and may stay on that vessel for a month or more. However, short-duration trips in full coverage (even with competition among full coverage observer providers) are much more expensive than the overall average daily observer rate for full coverage ($344 per day). In addition, depending on how such an approach was implemented, it could shift administrative overhead onto the FMA division, which currently has no infrastructure or administrative budget to oversee this type of program. If the approach was done through a federal observer contract, whether with one provider or several, it would continue to require a certain level of guaranteed work, so base rates for multiple providers would also be necessary.

Have observers review EM video

Another idea that has been proposed to increase cost efficiency is to have partial coverage observers review EM video during "down time" when they are in port. NMFS did a preliminary analysis of observer "down-time" and did not find evidence of observer free time that could be dedicated to video review so there was low potential of substantial cost savings. In addition, there are a variety of logistic difficulties, including field computers, video review software, and the observers needing to track hard-drives that make this approach complicated. As a result, NMFS has not pursued this idea further.

Structure of Partial Coverage Contract

NOAA's Acquisition and Grants Office (AGO) secures and administers the contract for the particle coverage observer provider for NMFS. FMA staff participate in contracting by initiating requirements documents, providing funding, and participating in the contract review and award process through formal source evaluation boards. The processes for federal contracts follow the

---

[7] PCFMAC March 2022 meeting minutes

[8] See Figure 2-3 in the 2021 Observer Program Annual Report

Federal Acquisition Regulations (FAR) and Commerce Acquisition Regulations (CAR). NMFS receives legal guidance on the FAR and CAR through NOAA contract attorneys and AGO staff.

Contracts for observer services in the partial coverage category are awarded through a competitive process, allowing any company that provides these services to bid. The partial observer coverage for the first 2 years (2013 and 2014) of the program was procured through a two-year contract awarded to AIS Inc. A second contract was awarded for the subsequent five years of the program to AIS, Inc. in April 2015. A third contract was completed and subsequently awarded for up to five years of the program to AIS, Inc. in July of 2019.

In 2024 a new partial coverage contract will be awarded. The structure of the new contract includes several components designed to improve efficiency and reduce costs. For example….

- Increase guaranteed days to the maximum realistic to get our max price per day as low as possible.
- The incorporation of plant days to support EM on trawl vessels, which reduces travel costs and may add flexibility for the provider to reduce lodging costs.
- Moving from half day to hourly billing.
- Comparative cost of observer deployment of recent past programs will be provided by the bidder.
- Contract is not solely evaluated on the cost of observer deployment.

Similar to the last contract, NMFS included the provision for observers to participate in NMFS fishery-independent surveys using funds made available through AFSC. This allows the provider to give additional work opportunities to their employees during the summer season when observing is more limited. This provides their employees continuity in employment, additional experience, and may help to reduce employee turnover, thereby increasing overall efficiency. NMFS benefits from trained observers with sea experience to help to conduct their survey fieldwork.

Biological Data Collection Modifications

Several ideas have been proposed to modify the biological data collection by observers. While these approaches do no create cost efficiencies directly, they do have the potential to reduce impact from loss of biological data from EM or potentially provide more data for stock assessments:

- **Use fishery-independent longline survey data for weights to inform fixed-gear EM**: Under the fixed-gear EM program catch is accounted for in numbers of fish and NMFS then uses average length / weight data collected by observers to convert numbers of fish to weight of catch. As the number of fixed-gear EM vessels increase, this could create data gaps and therefore the idea of using survey data, instead of observer data, to generate catch estimates. Stock assessment authors were consulted on this idea and they raised several concerns

  - This is problematic for the growing EM sablefish pot fishery because of gear selectivity differences. Current commercial pots are not standardized (e.g., escape rings will further change selectivity)
  - Average weights in fishery may be higher than survey because the fishery is targeting larger fish at ideal depths, rather than mirroring the survey
  - Weight data is only one component of observer data used in assessments
    - Loss of catch-at-age data will add more uncertainty to the assessment, especially for fisheries which are rapidly changing (e.g., sablefish)
    - Observer data is highly influential data source in the assessment to inform age class strength
    - Assessment is attempting to estimate contemporary selectivity differently from the historic, single gear (H&L) fishery
  - If full retention requirements for sablefish were to be removed, the assessment would have no data to understand discard information
- **Opportunistically deploy idle observers for focused collection of biological data**: Opportunistic deployments do not add value to a statistically rigorous sampling plan and do not result in the best data. In addition, predicting where and when observers will be 'idle' is challenging and the cost of at-sea observer data is more expensive than "idle" days so this has the potential to increase the number of at-sea observer days without increasing the value of the data. As such, NMFS is not planning to evaluate further.
- **Specify differing observer sampling protocols regionally or temporally based on data needs:** While this idea intuitively seems like it could be a way to reduce data gaps, we achieve the highest quality data from standardized sampling protocols and it is most efficient to have observers with skills that are interchangeable and it would be inefficient to have specialized observers and this could result in extra costs to get the "right" type of observer to a port. NMFS is not planning to evaluate further.

Reduce flexibility for fishery participants

There are a number of elements built into the partial coverage program that provide flexibility to fishery participants but, in general, these flexibilities are costly. Three ideas were proposed by NMFS that could reduce the cost per unit of monitoring, however due to the impact on fishery participants, the PCFMAC did not support moving any of these ideas forward for further evaluation (Table 5-1).

- **Require vessels to pick up observers in specific ports:** The current partial coverage program allows vessels to operate out of any port with a Federal Fishing Permitted processor. This flexibility allows vessels to operate as they usually would but increases costs for travel and observer down-time. There are potential programmatic cost savings by reducing the number of ports from which observers can deploy.

- **Trip selection compared to Vessel Selection.** There is an opportunity for cost efficiencies under the partial coverage program by re-evaluating trip selection as the sole method for assigning observers and EM. From 2013 to 2014, the partial coverage program used the vessel rather than the trip for vessels greater than or equal to 40 ft and less than 57.5 ft as the primary sampling unit from which to randomize observer deployment. Under this approach, selected vessels were required to carry observers for all trips during their selected 2-month period. Using vessel selection reduces the need for observer travel, and when combined with full monitoring for a period of time, generates representative data. This concept was abandoned in 2015 due to poor rates of observation of selected vessels (i.e., vessels disproportionately canceled their trips or did not fish in their selected time period). However, cost savings could be accomplished through a different form of vessel selection that increases the amount of time observers spend on a selected vessel to reduce travel cost and observer down-time
- **Extend notification before a trip:** The current partial coverage program requires a three-day notice for deploying at-sea observers. Utilizing a three-day window is expensive, as it gives both the agency and the observer provider a relatively short advance warning. This design was utilized to increase the level of flexibility afforded to fishermen to minimize the impact of their fishing trip (e.g., timing of the trip). Cost savings could potentially be incurred by extending the length of the notice for deploying at-sea observers, though this change would require buy-in from the industry by logging their fishing trips in the ODDS system further in advance from their departure date.

EM Improvement Projects

In addition to developing trawl EM, NMFS continues to work collaboratively with industry partners on the EM development, improvments, and cost efficiency projects. Table 5-1 provides information on the status of these projects as well as the status of ongoing work and projects in development.

Improving Fixed Gear EM Data Timeliness

It is possible to improve the timeliness with which fixed gear EM hard drives are reviewed. Figure 4-1 shows that, in 2022, it took an average of 218 days for EM HAL and 211 days for EM POT for data to be available to the CAS following the end of a fixed gear EM trip. As it typically takes only one day for data to become available to the CAS after they have been received by the AFSC, most of that time is attributable to video review. We worked with the fixed gear EM video review provider (the PSMFC) to estimate the cost of reducing video review times.

The PSMFC currently has three video reviewers who review fixed gear EM video. They estimate that a 28-day review time is achievable with the current number of reviewers if there are no hard drives from the previous year to review at the beginning of each year. Such backlogs have been common since the beginning of the fixed gear EM program. In order to achieve a seven-day review time for most of the year, the PSMFC estimates that the number of EM reviewers would need to be doubled to six, for a total additional annual cost of $300,000 ($100,000 per reviewer). The PSMFC estimates that the seven-day review time would still rely on there being no backlog of hard drives to review at the beginning of each year, and that there would be portions of the year when up to 10 reviewers would be needed to achieve a seven-day review time. However, the PSMFC also estimated that, for much of the year, three reviewers would be able to achieve a seven-day review time. Therefore, as stated, six reviewers would be able to achieve a seven-day review time for most of the year. The NMFS chose a seven-day review time, as that is the timeliness with which data are needed if they are to inform inseason management decisions. All scenarios in this analysis assume status quo review staffing (3 reviewers total).

Table 6-1. Summary of potential ways to reduce costs in the partial coverage component of the observer program that are not related to sampling design (i.e. separate from stratification and allocation of observers and EM) and the current status.

Approach	Description	Potential cost efficiency	Requires regulations change?	Status
Zero Coverage	Change the definition of zero selection	If vessels that take very few trips per year were added to Zero Selection and taken out of the EM, then it could improve the efficiency of the EM program	No	The definition of Zero Selection needs to use criteria that are identified ahead of time and are predictable from year to year. NMFS is not planning to change the definition at this time. However, the agency will move vessels that have not used their EM system into Zero Selection.
Observer Procurement and Duties	Hire observers (as federal employees and/or contractors) that would live in Alaska ports	Could reduce travel expenses if observers live in communities where fishing occurs	Maybe - would need to be further evaluated.	Preliminary review indicates that there might be cost savings for fee funds; however, FMA would incur additional expenses. This concept would need to be further analyzed before moving ahead.
	Voucher program to procure observers from multiple providers	Allow vessels in partial coverage, once selected in ODDS, to procure observers through current observer companies and then to be reimbursed by NMFS at the end of the season from the observer fees collected.	Yes	In 2017, the OAC reviewed a discussion paper (see section 3.5). No further work planned at this time.
	Have observers review EM video	Partial coverage observers would review EM video during "down time" when they are in port.	No	NMFS is not planning to evaluate due to the logistical complexity and the low potential of substantial cost savings.
	Partial coverage contract structure		No	NMFS has incorporated several cost efficiency measures into the structure of the new partial coverage contract.
Biological sample data collection	Using survey data for average weights and biological data	Potential method to reduce impact from loss of biological data from EM.	No	NMFS is not planning to evaluate further. Would have a negative impact on stock assessment.

Approach	Description	Potential cost efficiency	Requires regulations change?	Status
	Opportunistically deploy idle observers for focused collection of biological data	No cost efficiencies, but could potentially provide more data for stock assessments.	No	NMFS is not planning to evaluate further. Opportunistic deployments do not result in the best data.
	Specify differing observer sampling protocols regionally or temporally based on data needs	No cost efficiencies, but could potentially provide more data for stock assessments.	No	NMFS is not planning to evaluate further. It is inefficient to have specialized observers and this could result in extra costs to get the "right" type of observer to a port.
Reduce fishery flexibilities	Require vessels to pick up observers in particular ports	Potential cost savings by reducing the number of ports from which observers can deploy.	Yes -would need to be a regulation requiring vessels to pick up observers in, and return them to, one of the ports listed in the ADP.	In March 2022, PCFMAC did not support continued evaluation. NMFS is not planning to evaluate further.
	Instead of selecting one trip at a time for coverage, select multiple trips.	Potentially reduce travel costs for partial coverage observers.	No	In March 2022, PCFMAC raised concerns about negative impacts for industry and the potential to introduce bias. NMFS is not planning to evaluate further.
	Extending the length of the notice for deploying at-sea observers	Current regulations that specify vessels must register an anticipated trip in ODDS a minimum of 72 hours prior to embarking on each trip. The 72 hour window is expensive, as it gives both the agency and the observer provider a relatively short advance warning. Potential cost savings by requiring vessels to log fishing trips in ODDS further in advance from their departure date.	Yes	In Sep. 2021, PCMFAC noted the logistical challenges of this idea and did not support it. NMFS is not planning to evaluate further.
EM Improvement Projects	EM monitoring in GOA Rockfish fishery (maximized retention with compliance monitoring & shoreside observers in plant)	Would reduce costs for full coverage vessels. Could potentially result in cost savings of partial coverage monitoring in Kodiak by increasing the scale of shoreside monitoring.	Yes	Project in development.
	Improving and enhancing EM Data in Western GOA by testing EM configurations on vessels that fish using multiple	Reduce video review time and reduce catch handling burden for boats	No. Could be implemented through VMPs and definitions of EM selection	Ongoing projects by Aleutians East Borough funded through NFWF

Approach	Description	Potential cost efficiency	Requires regulations change?	Status
	gear types and evaluating catch handling and EM data review protocols for pot vessels		pools in the ADP	
	Real time electronic logbook data collection and reporting in groundfish and halibut fishery	Reduce data entry that is currently being done by the video reviewers.	No	Ongoing project. NMFS is working with a third party logbook company on the process for logbook data to be submitted to NMFS and the logbook to become "NMFS approved".
	Reduce time delay for EM data	Evaluate cost to get fixed-gear EM data in a timely fashion that is useful for inseason management. Could better leverage EM & reduce data gaps	No	In order to achieve a seven-day review time for most of the year, the PSMFC estimates that the number of EM reviewers would need to be doubled to six, for a total additional annual cost of $300,000
	Eligibility to be in the EM pool	Evaluate ways to optimize the fixed gear EM program for cost efficiencies by modifying ongoing eligibility for the fixed-gear EM program to ensure EM equipment is used cost effectively (for example, not installed on vessels not fishing or taking very few trips). Currently once NMFS approves vessels in the EM pool there isn't a mechanism to remove them.	Yes. While vessels can be removed for not following their VMP, they can't be removed for being cost inefficient	NMFS could consider as a longer term improvement which is more consistent with Trawl EM.
	Require fixed gear EM vessels to run EM systems on all trips & post-select trips to be submitted. Vessels could be told in ODDS in advance to run their cameras on all trips, and then be told to mail hard drives only for trips that were selected	Could better enable space-based strata by determining which strata the boat was in based on what they did on the trip, rather than what they think they are going to do. This approach would eliminate any monitoring effect.	No.	Proposed by NMFS but not supported by PCFMAC nor Council. NMFS would consider if annual report analysis shows evidence of monitoring effect and after evaluating catch handling protocols on pot vessels.

Table 6-2. Components of the cost estimate for hiring observers as federal employees. Due to rounding, totals will differ from the sum or multiplication of individual components.

Cost Component	Low sampling effort	High sampling effort
GS 5-1 Base annual salary	$44,649	$44,649
GS 5-1 Annual salary with benefits	$72,331	$72,331
Weekly rate (60/hours per week)	$2,423	$2,423
Observer weeks	756	1,148
Labor (weekly rate ✕ observer weeks)	$1,832,026	$2,781,501
Training	$65,064	$89,982
Travel to and from Alaska	$60,000	$90,000
Per diem and lodging	$1,115,973	$1,689,268
Travel within Alaska	$560,728	$851,229
Supervision	$178,183	$178,183
Total	**$3,811,973**	**$5,680,163**
Sea days (4 per week per observer)	3,025	4,593
Cost per observed sea day	$1,260	$1,237
Cost per observed sea day (2 x supervision)	$1,319	$1,276

7. References

AFSC (Alaska Fisheries Science Center) and AKRO (Alaska Regional Office). 2021. North Pacific Observer Program 2019 Annual Report. AFSC Processed Rep. 2021-05, 205 p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

Benoît, H. P. and J. Allard. 2009. Can the data from at-sea observer surveys be used to make general inferences about catch composition and discards? Can. J. Fish. Aquat. Sci. 66, 2025–2039. https://doi.org/10.1139/F09-116.

BLS (Bureau of Labor Statistics). 2023. Employer Costs for Employee Compensation – March 2023. U. S. Department of Labor. https://www.bls.gov/news.release/pdf/ecec.pdf

Bulus, M. 2023. *pwrss*: Statistical Power and Sample Size Calculation Tools. R package version 0.3.1. https://CRAN.R-project.org/package=pwrss.

Cahalan, J. and C. Faunce. 2020. Development and implementation of a fully randomized sampling design for a fishery monitoring program. Fish. Bull. 118(1): 87-99. http://dx.doi.org/10.7755/FB.118.1.8

Cochran, W.G. 1977. Sampling Techniques, 3rd edition. John Wiley and Sons, Inc. New York. 428 p.

Condie, H.M., A. Grant, and T.L. Catchpole. 2014. Incentivising selective fishing under a policy to ban discards; lessons from European and global fisheries. Mar. Policy 45, 287-292. https://doi.org/10.1016/j.marpol.2013.09.001

Cotter, A. J. R., and G. M. Pilling.2007. Landings, logbooks and observer surveys: improving the protocols for sampling commercial fisheries. Fish Fish. 8:123–152.

Figus E, K. R. Criddle. 2019. Characterizing preferences of fishermen to inform decision-making: A case study of the Pacific halibut (*Hippoglossus stenolepis*) fishery off Alaska. PLoS ONE 14(3): e0212537. https://doi.org/10.1371/journal.pone.0212537.

Curtis K. A., Carretta J. V. 2020. ObsCovgTools: Assessing observer coverage needed to document and estimate rare event bycatch. Fish. Res. 225: 105493. https://doi.org/10.1016/j.fishres.2020.105493.

Faunce, C., J. Sullivan, S. Barbeaux, J. Cahalan, J. Gasper, S. Lowe, and R. Webster. 2017. Deployment performance review of the 2016 North Pacific Groundfish and Halibut Observer Program. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-358, 75 p. http://doi.org/10.7289/V5/TM-AFSC-358.

Faunce, C. H. and S. Barbeaux. 2011. The frequency and quantity of Alaskan groundfish catcher-vessel landings made with and without an observer. ICES J. M. Sci. 68, 1757–1763. https://doi.org/10.1093/icesjms/fsr090

Ganz, P., C. Faunce, G. Mayhew, S. Barbeaux, J. Cahalan, J. Gasper, S. Lowe, and R. Webster. 2020. Deployment performance review of the 2019 North Pacific Observer Program. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-411, 87 p. https://doi.org/10.25923/8xpb-rg73.

Ganz, P., S. Barbeaux, J. Cahalan, J. Gasper, S. Lowe, R. Webster, and C. Faunce. 2019. Deployment performance review of the 2018 North Pacific Observer Program. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-397, 64 p. https://doi.org/10.25923/cvrp-je11.

Ganz, P., S. Barbeaux, J. Cahalan, J. Gasper, S. Lowe, R. Webster, and C. Faunce. 2018. Deployment performance review of the 2017 North Pacific Observer Program. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-379, 68 p. https://doi.org/10.7289/V5/TM-AFSC-379.

ICES (International Council for the Exploration of the Sea). 2014. Report of the third workshop on practical implementation of statistical sound catch sampling programmes (WKPICS3); Copenhagen, Denmark, 19–22 November 2013. ICES CM 2013/ACOM:54, 109 p. [Available from website.]

Krzywinski, M. and N. Altman. 2013. Power and sample size. Nat. Meth. 10, 1139–1140. https://doi.org/10.1038/nmeth.2738.

NMFS (National Marine Fisheries Service). 2022. 2023 Annual Deployment Plan for Observers and Electronic Monitoring in the Groundfish and Halibut Fisheries off Alaska. National Oceanic and Atmospheric Administration, 709 West 9th Street. Juneau, Alaska 99802. Available at: https://www.fisheries.noaa.gov/resource/document/2023-annual-deployment-plan-observers-and-electronic-monitoring-groundfish-and

NMFS and NPFMC (North Pacific Fishery Management Council). 2022. Draft for Final Review - Environmental Assessment/Regulatory Impact Review for Proposed Amendments to the Fishery Management Plans for Groundfish of the Bering Sea / Aleutian Islands Management Area and for Groundfish of the Gulf of Alaska: Integrating Electronic Monitoring on Pollock Catcher Vessels using Pelagic Trawl Gear and Tender Vessels in the North Pacific Observer Program. National Oceanic and Atmospheric Administration, 709 West 9th Street. Juneau, Alaska 99802. Available online at: https://meetings.npfmc.org/CommentReview/DownloadFile?p=e31b9c56-d3a4-4d1e-b621-b0b4bd892b5a.pdf&fileName=C3%20Trawl%20EM%20Analysis.pdf

NPFMC. 2017. Options to address low selection rates in partial coverage. North Pacific Fishery Management Council, 1007 West Third Ave., Suite 400, L92 Building, 4th floor, Anchorage, Alaska 99501. Available at: https://meetings.npfmc.org/CommentReview/DownloadFile?p=afd96563-e9d1-4986-94bf-13b870573f9c.pdf&fileName=C5%20OAC%20DP%20on%20low%20sampling%20917.pdf

Thompson, S. K. 2012. Sampling, 3rd ed., 472 p. John Wiley and Sons Inc., Hoboken, NJ.

Vølstad, J. H., P. S. Afonso, A. P. Baloi, N. de Premegi, J. Meisfjord, and M. Cardinale. 2014. Probability-based survey to monitor catch and effort in coastal small-scale fisheries. Fish. Res. 151:39–46.

8. Appendices

Appendix A. Box size definition

A method for spatiotemporally categorizing fishing effort is needed by the 2024 ADP analysis for allocating monitoring effort and evaluating the extent to which the proposed sample designs mitigate data gaps. At any given sample rate, strata with more dispersed distributions of fishing effort are more likely to have gaps than strata with fishing effort that is highly clumped in space and time.

Two of the proposed allocation methods, *Cost-weighted Boxes* and *Proximity*, rely on defining an appropriate spatio-temporal resolution to group fishing effort and define “neighboring” fishing trips. Dividing fishing effort into equally-sized units of time and space is an important step in understanding where and when fishing occurs within each stratum, and therefore where and when gaps in monitoring may occur. By applying the same method of binning fishing effort into discrete units of time and space across proposed monitoring designs and evaluations, we can quantify the likelihood that a stratum may have gaps in coverage and devise methods to decrease the potential for those data gaps.

Methods

In this analysis, *boxes* are defined as discrete bins of time and space into which all fishing trips are categorized. Once boxes are defined, a grid of spatial cells can be superimposed on a map of the Alaska EEZ and the calendar year can be divided on the selected temporal scale. Time can be defined on any scale (days, weeks, months). Spatial units in this analysis will be defined by the width of hexagonal cells; the geometric arrangement of an iso-area hexagon grid is easy to apply and the boundaries of the cells are unlikely to match those of geography or management areas. Fishing trips can be assigned to the hexagonal boxes and temporal periods based on the ADFG Statistical Area and fishing dates recorded on landing reports from each trip. ADFG Statistical Areas for each trip are used. Trips may span more than one box either spatially, temporally, or both, and such trips will be placed equally in (split equally between) the boxes that it spans. For example, a fishing trip reported to have fished in three ADFG Statistical Areas will contribute 0.33 trips to each area.

To define boxes, the extent of both the temporal and spatial components need to be carefully specified because box size will affect its utility in defining spatio-temporal resolution. The spatio-temporal extent of each box must be defined such that the distribution of fishing effort can be captured and we can identify where gaps in monitoring coverage may occur. Using too broad of a box definition (e.g., 2000-km wide hex cells and 3-month bins) results in fewer boxes, almost all of which contain so many trips that it is virtually impossible for them to go unsampled. Additionally, large boxes have less resolution and may not adequately represent spatio-temporal patterns in fishing activity. Such a broad box definition is also more prone to edge effects because the location of the box boundaries (where the grid is centered) will influence how trips are categorized into each box. For instance, shifting a 2,000km hex grid only 10km to the east or west may greatly impact the total number of boxes, the number of trips contained in each box, and the number of boxes that contain large proportions of land or regions where fishing does not occur. Conversely, using too narrow of a box definition will categorize trips into many boxes, each with few trips and a high likelihood of being unsampled. In considering box definitions, it is important to balance the extent of both the spatial and temporal units; pairing a narrow spatial unit (e.g., 100 km-wide hexagon cells) with a very broad temporal unit (e.g., 3-month temporal bins) will result in overemphasizing the spatial distribution of fishing effort; with 100 km wide hexagon cells and 3-month temporal bins, fishing effort will be categorized into over 100 separate hexagon cells and only 4-5 temporal units. This disparity makes identifying potential monitoring gaps challenging.

One final consideration when developing box definitions is avoiding box-size or grid placement artifacts that influence our evaluation of coverage gap propensity. For example, consider a box definition with two adjacent boxes, one with only one trip and the other with 25 trips. Because the first box has only 1 trip, any sample is less likely to contain a monitored trip from that box than from the second box which has 25 trips; a potential data gap. A shift in the box definition, either by shifting the leading edge of the grid or by changing the box size, may result in a single box of 26 trips which could change the interpretation of how many boxes are likely to contain no monitored trips during sampling.

To mitigate this issue, we define *neighborhoods* for each box that allow each box to rely on adjacent boxes in time and/or space when determining the likelihood of a data gap; the data gap is defined by the box's neighborhood. Therefore, defining a neighborhood for each box shifts the importance away from how many trips occur in individual boxes and instead focuses the importance towards the number of sampled trips occurring in close proximity to a box. Gaps can therefore be identified as isolated boxes containing few trips where no trips are expected to be monitored in the box or its neighboring boxes. Extending the previous example, if the box definition allows boxes to neighbor adjacent boxes, the one-trip box would add the 25 trips from the adjacent box in its neighborhood and be interpreted as being very likely to be either sampled OR neighboring a box with monitored trip; thus removing the impact of the artifacts.

Allowing boxes to neighbor adjacent boxes (i.e., have neighborhoods) also allows us to define the boxes with a higher spatio-temporal resolution. For instance, if fishing effort was binned into boxes defined by 200 km-wide hexagon spatial cells and 1-week temporal bins but without the ability to seek adjacent cells for monitored trips, most strata would be composed of boxes with very few trips (Table A-1). In most strata, 75% of boxes would have 4 or fewer trips, and if a 15% sample rate were used, these boxes would be more likely than not (52%) to be unsampled. However, if we allow these boxes to neighbor adjacent boxes, the number of trips in the neighborhoods of each box is greatly increased, and all strata have only 25% of boxes containing 4 or fewer trips. Smaller boxes are able to capture the spatiotemporal distribution of fishing effort at a finer resolution, and by defining neighborhoods, gaps are not identified by the few trips that occur in smaller boxes.

A large number of box definitions were evaluated, with spatial units ranging from 100 km to 750 km-wide hexagonal cells, temporal units ranging from 1 week to 2 months, and neighborhoods defined by the number of of adjacent boxes in both space in time varying from zero (no neighboring boxes, *i.e.*, the neighborhood is the size of the box), one (*i.e.*, immediately adjacent boxes), or two (*i.e.*, 2 rings of spatial hexagon cells and within ± 2 units of time). These box definitions were applied to partial coverage fishing effort between 2018 and 2022 using the current (2023) stratum definitions. For each box definition, the total number of boxes and the distribution of the number of the trips contained in the boxes and their neighborhoods were quantified. The purpose of this exercise was to identify a box definition that supports our ability to determine the likelihood of spatiotemporal monitoring gaps in each stratum under a 15% trip selection rate that balances the relative spatial and temporal resolution of the box definition..

Results and Discussion

A majority of box definitions were removed from consideration because extremes in the spatial or temporal extents of the boxes resulted in an impaired ability to differentiate boxes with a low or high probability of sampling (*e.g.*, created extremely large boxes where no boxes are data gaps or extremely small boxes where every box becomes a data gap). Additional designs were not considered because their neighborhood definitions were too extreme (no-neighboring with narrow spatio-temporal extent leads to small boxes, or neighborhoods of 2 layers of adjacent boxes coupled with broad spatiotemporal extents leads to large boxes). A 'goldilocks' region of box definitions was identified where boxes and their associated neighborhoods were neither large or small and the distribution of trips within boxes allowed for identification of potential gaps without focusing on individual boxes.

Ultimately, a box defined spatially by hexagon cells 200 km-wide, temporally by 1-week bins, and neighboring 1-box deep (adjacent neighbors) was chosen as the final box definition to employ in the 2024 ADP analysis (Figure 3-3). Box definitions varying slightly from the chosen definition are likely just as valid choices, but this particular definition was chosen for several reasons. Firstly, simplicity; because it uses round numbers, an easy-to-interpret time scale, and neighboring boxes are easily identified (they are directly adjacent). Secondly, the spatial extent of the 200 km-wide cell and its neighborhood is consistent with the spatial size of NMFS areas (Figure A-1). Finally, the total number of unique spatial and temporal units among the boxes was relatively similar across strata (Table A-2), thus the importance of the temporal and spatial distribution of fishing effort is relatively equal between strata.

Table A-1. Quantiles of the number of trips per box defined by 200 km-wide hexagon spatial cells and 1-week temporal bins when either counting only trips in the box (no neighbors) or including trips in adjacent boxes (with neighbors). These counts are from fishing effort in 2022 with the current stratification definition.

	No neighbors				With neighbors			
	Quantiles (# of trips per box)				Quantiles (# of trips per box)			
Strata	10%	25%	50%	75%	10%	25%	50%	75%
OB_HAL	1	1	2	4	3	7	17	30
OB_POT	1	1	2	4	7	11	19	27
OB_TRW	1	1	3	8	4	8	20	43
EM_HAL	1	1	1	3	2	6	11	18
EM_POT	1	1	1	2	3	4	7	11
EM_TRW	1	2	5	15	8	19	67	106

Table A-2. Total number of spatial and temporal units populated by each stratum where boxes are defined by 200 km-wide hexagon cells and 1-week bins for 2022 fishing effort and current stratification definitions.

Strata	Total hexagon cells	Total week bins
OB_HAL	54	49
OB_POT	40	47
OB_TRW	14	33
EM_HAL	44	48
EM_POT	18	47
EM_TRW	14	24

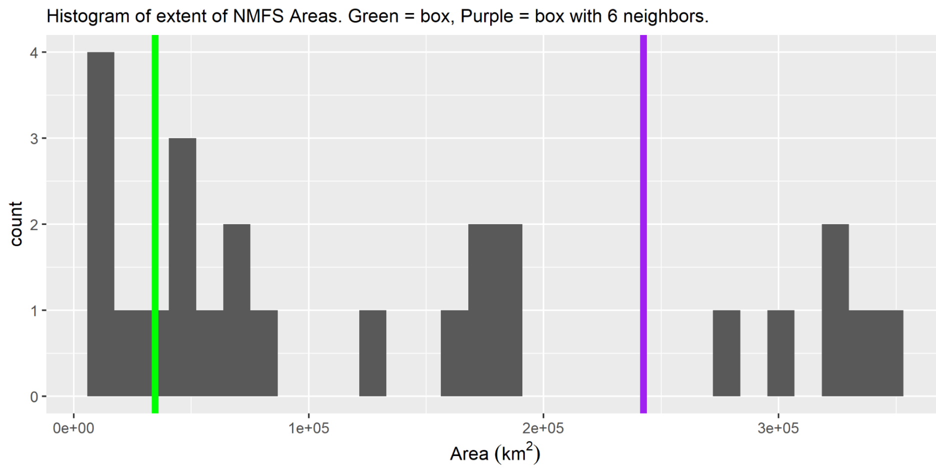

Figure A-1. Histogram of the spatial extents (in km^2) of NMFS Areas, the 200 km-wide box (green line), and its neighborhood (purple line).

Appendix B. Gear-specific Interspersion summaries

The figures included in this appendix are meant to accompany the Interspersion evaluation metrics provided in Chapter 5: Results and Discussion. *Interspersion* is a summary of the expected proportion of trips within a monitoring method (at-sea observer, at-sea fixed-gear EM, and shoreside observers with trawl EM) that are neighboring sampled trips using the same gear type, in the same time and space (neighborhood). The final *Interspersion* metric results presented in Chapter 5 is a summary across gear types within a monitoring method of the values below (Tables B1 through B3). The following figures show the gear-specific interspersion values achieved by all proposed designs at the three budget levels evaluated. Note that in this analysis, the *Equal Rates* and *Status Quo* allocation methods did not differ between stratification definitions, so those results were shown under the *Current* stratification definition but omitted due to redundancy with the *FMP* and *FIXED-FMP* stratum definitions.

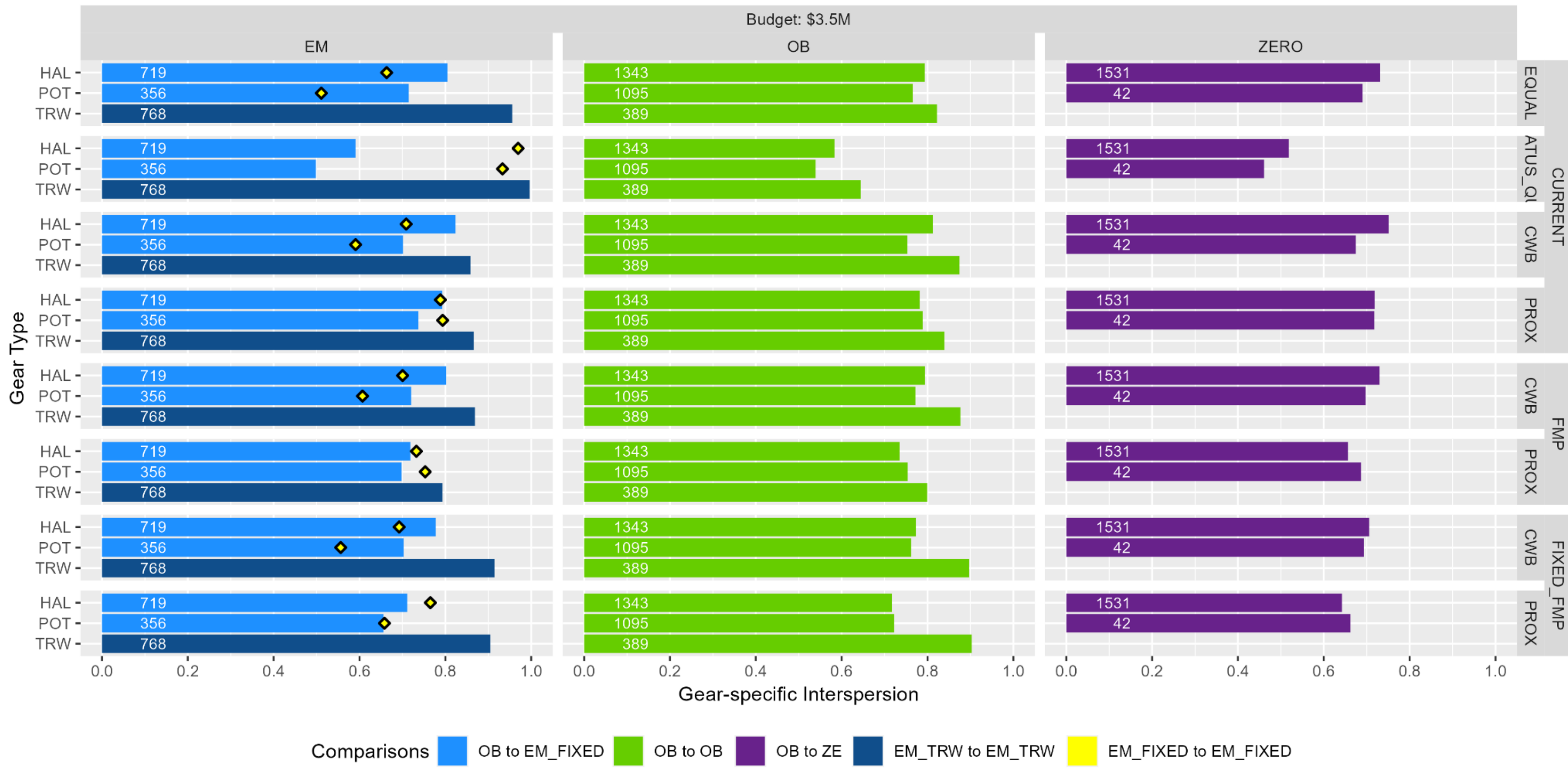

Figure B-1. Alaska-wide gear-specific interspersion of all designs with a $3.5M budget for each design. Bars depict the interspersion value of OB - EM (left, blue bars), OB - OB (center, green bars), and OB - ZERO (right, purple bars) and the number of trips in each gear group is included on the bars. EM - EM interspersion values are included as yellow diamonds in the left panels. Monitoring design is on the right axis with 3 stratification definitions: CURRENT (2023), FMP (includes separate strata for BSAI and GOA), and FIXED_FMP (combined fixed gear: HAL, POT, and mixed gear trips, and FMP: BSAI and GOA) and each allocation method: EQUAL *Equal rates*, STATUS QUO *Status quo*, CWB *Cost-weighted boxes*, and PROX *Proximity*.

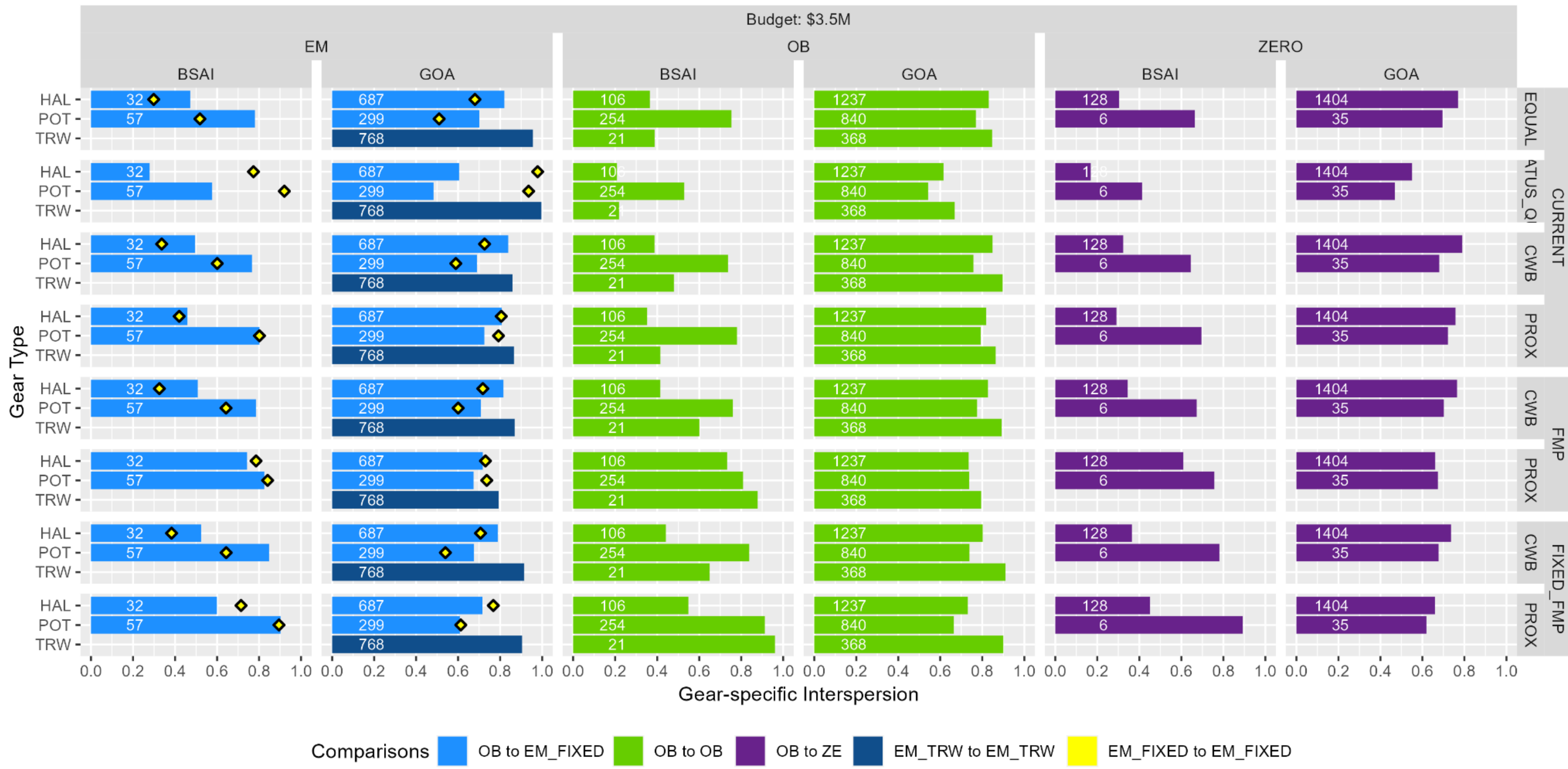

Figure B-2. Gear-specific interspersion for each FMP (BSAI, GOA) and all designs with a $3.5M budget for each design. Bars depict the interspersion value of OB - EM (left, blue bars), OB - OB (center, green bars), and OB - ZERO (right, purple bars) and the number of trips in each gear group is included on the bars. EM - EM interspersion values are included as yellow diamonds in the left panels. Monitoring design is on the right axis with 3 stratification definitions: CURRENT (2023), FMP (includes separate strata for BSAI and GOA), and FIXED_FMP (combined fixed gear: HAL, POT, and mixed gear trips, and FMP: BSAI and GOA) and each allocation method: EQUAL *Equal Rates*, STATUS QUO *Status Quo*, CWB *Cost-weighted boxes*, and PROX *Proximity*.

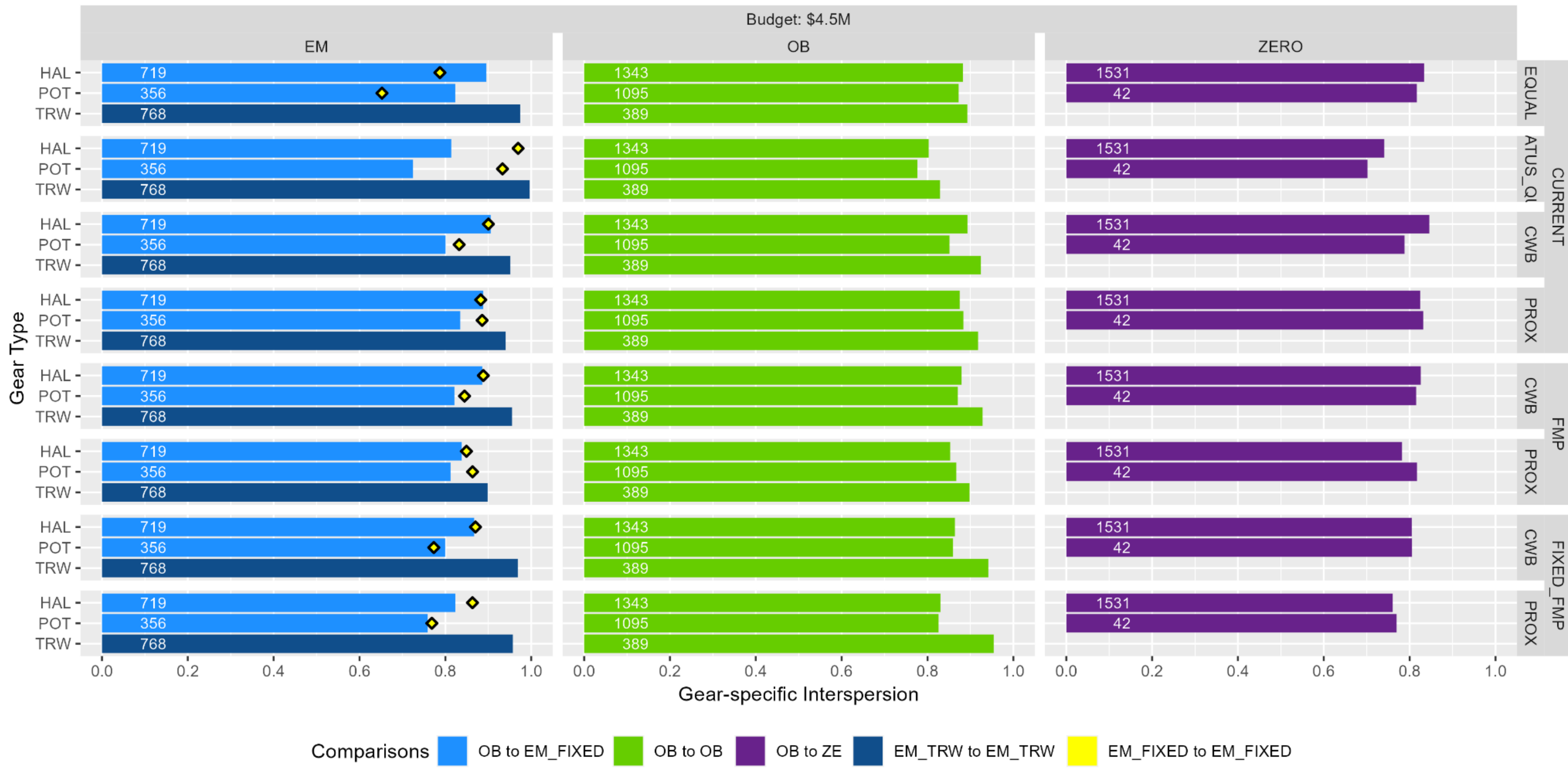

Figure B-3. Gear-specific interspersion of all designs with a $4.5M budget for each design. Bars depict the interspersion value of OB - EM (left, blue bars), OB - OB (center, green bars), and OB - ZERO (right, purple bars) and the number of trips in each gear group is included on the bars. EM - EM interspersion values are included as yellow diamonds in the left panels. Monitoring design is on the right axis with 3 stratification definitions: CURRENT (2023), FMP (includes separate strata for BSAI and GOA), and FIXED_FMP (combined fixed gear: HAL, POT, and mixed gear trips, and FMP: BSAI and GOA) and each allocation method: EQUAL *Equal Rates*, STATUS QUO *Status Quo*, CWB *Cost-weighted boxes*, and PROX *Proximity*.

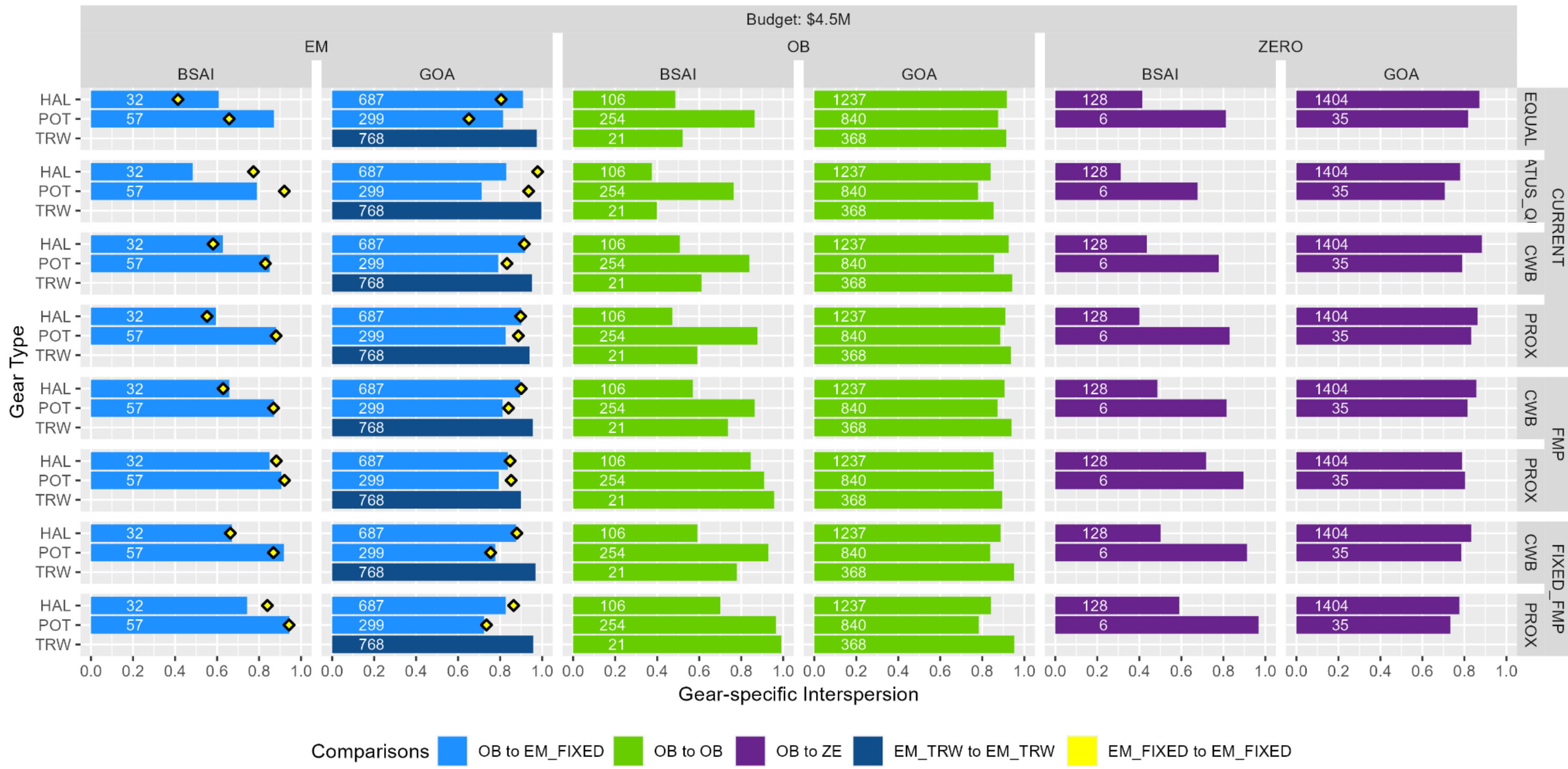

Figure B-4. Gear-specific interspersion for each FMP (BSAI, GOA) and all designs with a $4.5M budget for each design. Bars depict the interspersion value of OB - EM (left, blue bars), OB - OB (center, green bars), and OB - ZERO (right, purple bars) and the number of trips in each gear group is included on the bars. EM - EM interspersion values are included as yellow diamonds in the left panels. Monitoring design is on the right axis with 3 stratification definitions: CURRENT (2023), FMP (includes separate strata for BSAI and GOA), and FIXED_FMP (combined fixed gear: HAL, POT, and mixed gear trips, and FMP: BSAI and GOA) and each allocation method: EQUAL *Equal Rates*, STATUS QUO *Status Quo*, CWB *Cost-weighted boxes*, and PROX *Proximity*.

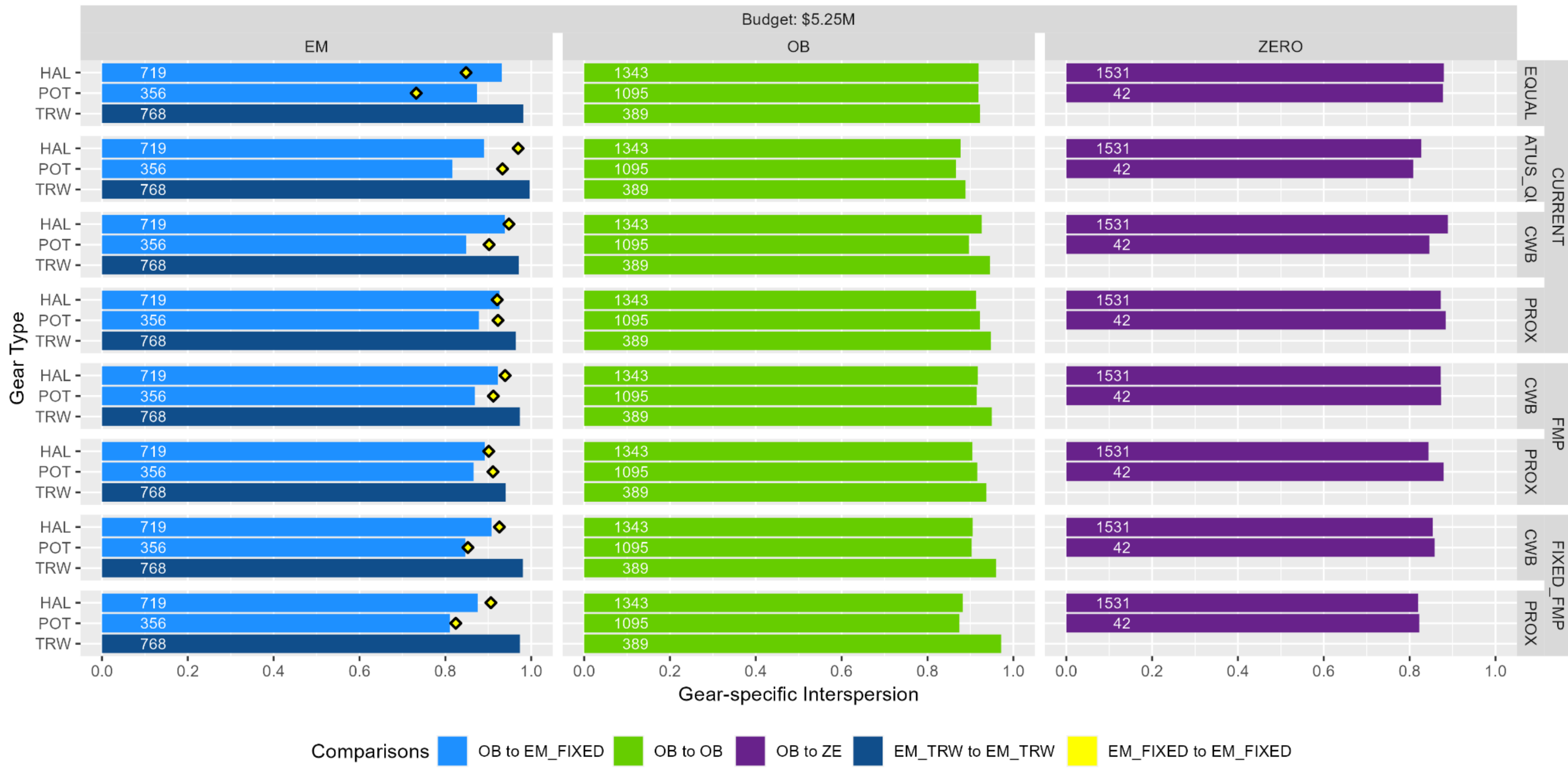

Figure B-5. Gear-specific interspersion of all designs with a $5.25M budget for each design. Bars depict the interspersion value of OB - EM (left, blue bars), OB - OB (center, green bars), and OB - ZERO (right, purple bars) and the number of trips in each gear group is included on the bars. EM - EM interspersion values are included as yellow diamonds in the left panels. Monitoring design is on the right axis with 3 stratification definitions: CURRENT (2023), FMP (includes separate strata for BSAI and GOA), and FIXED_FMP (combined fixed gear: HAL, POT, and mixed gear trips, and FMP: BSAI and GOA) and each allocation method: EQUAL *Equal Rates*, STATUS QUO *Status Quo*, CWB *Cost-weighted boxes*, and PROX *Proximity*.

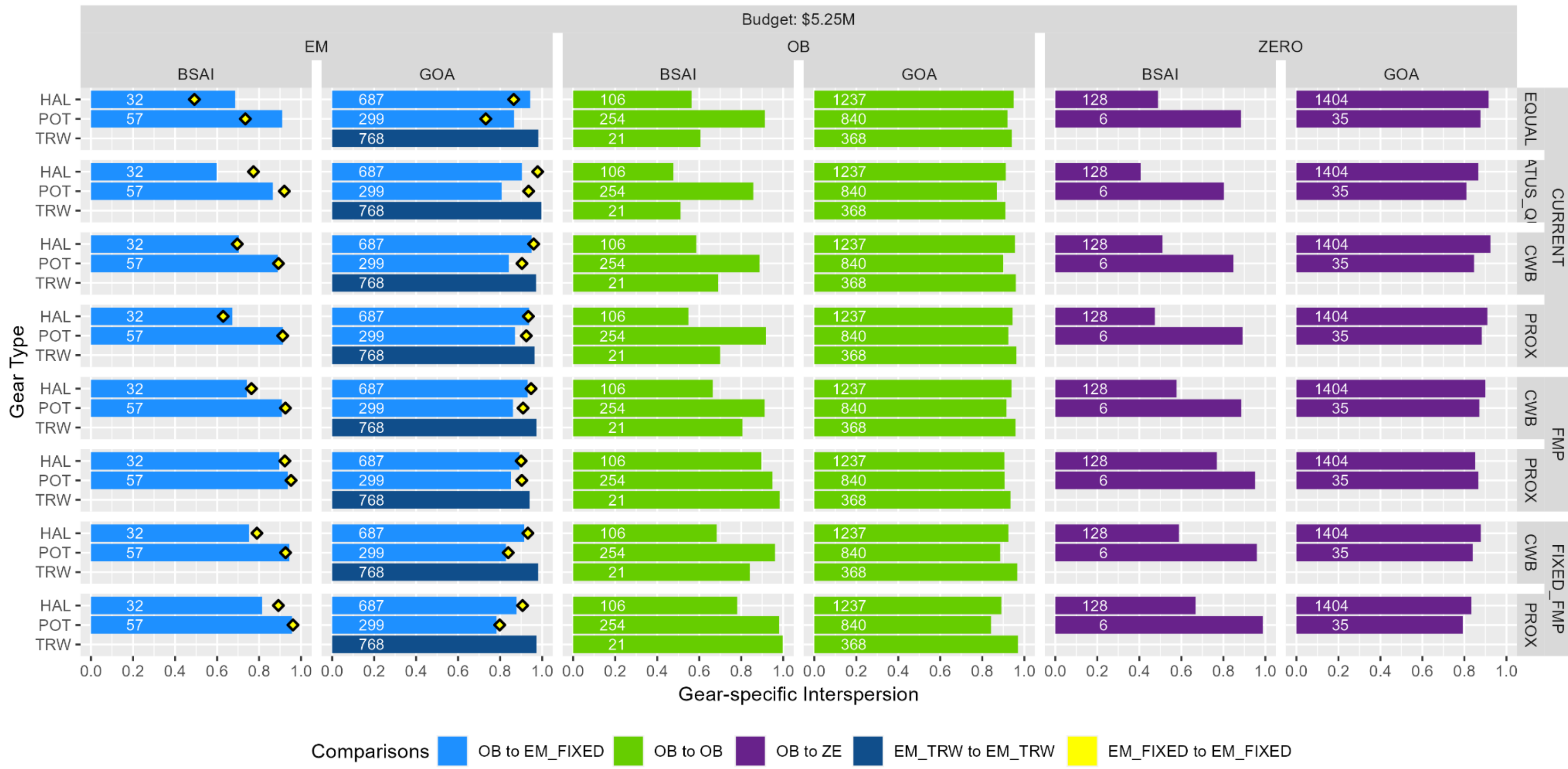

Figure B-6. Gear-specific interspersion for each FMP (BSAI, GOA) and all designs with a $5.25M budget for each design. Bars depict the interspersion value of OB - EM (left, blue bars), OB - OB (center, green bars), and OB - ZERO (right, purple bars) and the number of trips in each gear group is included on the bars. EM - EM interspersion values are included as yellow diamonds in the left panels. Monitoring design is on the right axis with 3 stratification definitions: CURRENT (2023), FMP (includes separate strata for BSAI and GOA), and FIXED_FMP (combined fixed gear: HAL, POT, and mixed gear trips, and FMP: BSAI and GOA) and each allocation method: EQUAL *Equal Rates*, STATUS QUO *Status Quo*, CWB *Cost-weighted boxes*, and PROX *Proximity*.