NOAA Technical Memorandum NMFS-AFSC-302
doi:10.7289/V55M63NW

Deployment Performance Review of the 2014 North Pacific Groundfish and Halibut Observer Program

C. Faunce, J. Gasper, J. Cahalan, S. Lowe, R. Webster,
and T. A'mar

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Alaska Fisheries Science Center

June 2015

NOAA Technical Memorandum NMFS

The National Marine Fisheries Service's Alaska Fisheries Science Center uses the NOAA Technical Memorandum series to issue informal scientific and technical publications when complete formal review and editorial processing are not appropriate or feasible. Documents within this series reflect sound professional work and may be referenced in the formal scientific and technical literature.

The NMFS-AFSC Technical Memorandum series of the Alaska Fisheries Science Center continues the NMFS-F/NWC series established in 1970 by the Northwest Fisheries Center. The NMFS-NWFSC series is currently used by the Northwest Fisheries Science Center.

This document should be cited as follows:

Faunce, C., J. Gasper, J. Cahalan, S. Lowe, R. Webster, and T. A'mar. 2015. Deployment performance review of the 2014 North Pacific Groundfish and Halibut Observer Program. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-302, 55 p. doi:10.7289/V55M63NW.

Document available: http://www.afsc.noaa.gov/Publications/AFSC TM/NOAA-TM-AFSC-302.pdf.

NOAA Technical Memorandum NMFS-AFSC-302
doi:10.7289/V55M63NW

Deployment Performance Review of the 2014 North Pacific Groundfish and Halibut Observer Program

by
C. Faunce[1], J. Gasper[2], J. Cahalan[3], S. Lowe[4],
R. Webster[5], and T. A'mar[4]

[1]Fisheries Monitoring and Analysis Division
Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
7600 Sand Point Way NE
Seattle, WA 98115

[2]Sustainable Fisheries Division
Alaska Regional Office
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
709 West 9th Street, Juneau, AK, 99801

[3]Pacific States Marine Fisheries Commission
7600 Sand Point Way NE
Seattle, WA 98115

[4]Resource Ecology and Fisheries
Management Division
Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
7600 Sand Point Way NE
Seattle, WA 98115

[5]International Pacific Halibut Commission
2320 West Commodore Way, Suite 300
Seattle, WA 98199

www.afsc.noaa.gov

U.S. DEPARTMENT OF COMMERCE
Penny S. Pritzker, Secretary
National Oceanic and Atmospheric Administration
Kathryn D. Sullivan, Under Secretary and Administrator
National Marine Fisheries Service
Eileen Sobeck, Assistant Administrator for Fisheries

June 2015

This document is available to the public through:

National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161

www.ntis.gov

ABSTRACT

This report contains the analyses and findings of the Alaska Fisheries Science Center's Fisheries Monitoring and Analysis Division's Observer Science Committee on the efficiency and effectiveness of observer deployment in 2014. This review is performed as part of the annual review process implemented by the North Pacific Fishery Management Council for the North Pacific Groundfish and Halibut Observer Program. This report is intended to inform the Council, the National Marine Fisheries Service (NMFS), and the public how well various aspects of the observer program are working, and it contains recommendations for improvement.

Several performance metrics were evaluated. These included metrics that assess whether targeted levels of observer coverage were achieved, metrics that assess whether data from the observed fleet was representative of the unobserved fleet, and metrics that evaluated the probability of obtaining data throughout the fleet given the current sample rate (deployment rate). Observer deployments are representative if they are not spatially or temporally biased and if the characteristics of observed trips are similar to those from unobserved trips. Non-representative spatial or temporal deployments are evidence of a deployment effect and differences between the observed and unobserved fleet fishing characteristics are evidence of an observer effect.

This evaluation of observer deployment was conducted at the stratum level as defined in the 2014 Annual Deployment Plan (ADP; released in December of 2013), since each stratum is defined by its own sampling unit (i.e., vessels or trips) and/or rate of sampling. Representativeness of trip selection was examined at a finer scale to help clarify initial findings.

In 2014, observers were deployed and collected data for 4,367 days. This amount is within 9% of the observer days budgeted at the start of the year in the 2014 ADP. In total, 417 vessels and 5,883 trips (43%) were observed in the fleet from 11 deployment strata. Both full coverage strata were completely observed, the coverage rate in the trip selection stratum was within its expected range, the coverage rate within the six vessel selection strata were within expected ranges (4 strata) or greater than expected amounts of coverage (2 strata), and there was no coverage in the no selection strata (2 strata).

In trip selection, vessels log trips into the Observer Declare and Deploy System (ODDS) where logged trips are randomly selected for observer coverage based on pre-determined selection probabilities. Exact binomial tests confirmed that the ODDS was selecting trips at the correct rate (16%). However trips selected to be observed were cancelled at nearly four (3.7) times the rate of unselected trips. In order to prevent cancellation of observed trips from decreasing the sample size, the ODDS was constructed to automatically assign observer coverage to the user's next logged trip; however, the order and timing of observed trips can be changed by the user. The lag in the number of observed trips relative to the number of total trips in this stratum was evidence of temporal bias in trip selection. There was no evidence of spatial bias in this stratum. We recommend that the ODDS protocols of 1) allowing selected trips to be cancelled and 2) allowing multiple trips to be logged prior to sailing be re-evaluated.

There was evidence of an observer effect in trip selection. Initial tests found differences in the vessel length and landed catch of observed trips compared to unobserved trips. When differences were examined by gear, it was found that hook-and-line vessels landed 14.4% less catch that had 9.1% more landed species when observed than when unobserved. Trawl vessels were found to have fished in 4.4% fewer areas and fished on trips that were 8.4% shorter in

duration when observed compared to when unobserved. We recommend Fisheries Management Plan (FMP) area and gear type be included as potential strata definitions in the 2016 ADP.

In vessel selection, NMFS randomly selects some vessels for 100% coverage during six separate 2-month selection periods (strata). Sampling frames for each strata are constructed based on past fishing activity. This method has three drawbacks. First, vessels that fished in the past are not necessarily those that will fish in the future. Second, vessels that did not fish in the past but do fish in the selection period have no chance of being selected. The error in terms of the percentage of the selection frame represented by these two factors averaged 40% among selection periods. Third, vessel selections need to be inflated since some selected vessels will request and be granted a release from observer coverage by NMFS. The number of selected vessels was on average 8 times higher than the targeted number of vessels in 2014.
Releases from coverage in the vessel selection strata occurred in each selection period in 2014. Between 37% and 67% of vessels that were selected and fished in a given time period were granted a release from observer coverage by NMFS. There was no evidence of spatial bias caused by releasing vessels in those areas where more than 10 trips occurred; however, four of the areas with fewer than 10 trips were not observed due to releases and there were more trips released from coverage than were observed. However, vessel selection coverage rates remained within or above expected values (23% of vessel selection vessels were observed during 2014). There was strong evidence of an observer effect in vessel selection; the number of permutations tests with p-values < 0.05 was 9 times the expected amount indicating fishing characteristics differed between observed and unobserved trips.

Based in part on the recommendations contained within the 2013 version of this document, the vessel selection method of observer deployment was discontinued in 2015. Vessels previously in vessel selection now participate in trip selection. To increase sampling efficiency in future years, we recommend that a list of vessels that cannot carry a human observer be generated and updated each year. This list could then be used to define new deployment strata for non-human monitoring in future ADPs.

Observers monitor the deliveries of trawl pollock at shoreside processing plants and collect genetic tissue samples from salmon bycatch. For vessels fishing pollock in the full-coverage stratum, this task is performed by designated plant observers, whereas for vessels fishing pollock in partial-coverage strata, only those trips that are observed at sea are also monitored at the plant. Although all pollock deliveries were monitored in the full coverage pollock fishery, this was not the case for pollock fishery trips in the partial coverage stratum where the proportion of deliveries monitored for salmon bycatch was lower than the target sample rate. Observer coverage rates were especially low in ports with high tendering activity. When tendering activity was removed from the analysis, the proportion of pollock fishery partial-coverage deliveries that were monitored fell within the expected range, indicating salmon sampling goals were not achieved in the partial coverage pollock fleet as a result of tendering activities.
Tendering activity is problematic for the observer program since observers rely on the captain to define the fishery, there are insufficient resources in most partial-coverage plants to reliably detect salmon bycatch if they are rare, and observers cannot be stationed on tendering vessels. Vessels that delivered to a tender were 11.5% shorter in length, fished 29.1% longer in duration, and had catch that was 1.3% less “pure” than vessels that did not deliver to a tender. Because of these differences in trip characteristics when a vessel was delivering to a tender versus delivering shoreside, an examination of potential observer effects was conducted within trips delivering to tender vessels. Trips delivering to tenders occurred on vessels that were 8.8% shorter in length

and delivered catch that was 6% less “pure” when observed relative to unobserved trips, providing evidence of an observer effect within the fleet of vessels delivering to tenders. The benefits to monitoring salmon bycatch in the presence of deployment and observer effects caused by tendering activity should be addressed in future ADPs.

We examined the probability of having no observer coverage as a function of how many trips were in a NMFS area for each partial coverage stratum since each stratum has its own coverage rate. The probability of having no data from a NMFS area decreased as total effort and sampling rate increased. While this result is not surprising, it helps to highlight that the smaller the population of interest (e.g., fishing activities in a NMFS area), the higher the overall coverage rate needs to be in order to meet observer coverage goals and to obtain data that are representative of fishing activities.

Catcher processors (CPs) in the partial coverage stratum pose a sampling problem to the observer program. Since there are few trips made by these vessels, the probability that they are observed under the current selection rates is low. In addition, by nature of being CPs, fishing trip characteristics tend to be different from the other trips in partial coverage. With this in mind, we recommend that catcher processor vessels in partial coverage be treated as a new stratum with a potentially different selection rate.

We repeat our 2013 recommendation that NMFS should develop a trip identifier to better link databases containing logged trips that are used in the logistics of observer deployment (ODDS) to the actual landings records of fishing trips (*e*Landings).

CONTENTS

INTRODUCTION

Background of the North Pacific Groundfish and Halibut Observer Program

Fisheries observers are people who collect independent information that is used to determine the effect of fishing to natural resources. The National Marine Fisheries Service (NMFS) uses its observer program in Alaska to enable the use of tools such as catch quotas to manage against the over or under-harvest of fishes. The data collection by observers is currently the only reliable and verifiable method for fishery discard information that facilitates estimation of total catch, as well as seabird and marine mammal interactions with fisheries. In addition, observers also collect biological information such as length, sex, weight, ageing structures (e.g., otoliths, spines, scales and vertebrae), and stomachs to support ecosystem studies and stock assessments.
The observer program in the North Pacific has a long history. Observers were first deployed onto fishing vessels in the Bering Sea in 1973 and into the remainder of the North Pacific in 1975 (Nelson et al. 1981, Wall et al. 1981). Fisheries in the North Pacific were initially prosecuted exclusively by foreign and later by "joint venture" operations where a developing domestic fleet of catcher vessels delivered to foreign-owned processing vessels. During the foreign and joint venture operations, foreign vessels carried fisheries observers at their expense, while domestic vessels were exempted from this observer coverage. As foreign vessels' rights to fish in the U.S. Exclusive Economic Zone (EEZ) were reduced over time and the domestic fishery grew, it became obvious that observer coverage would be necessary for the emerging domestic fleet. At the onset of fully domestic fishery operations in 1990, the North Pacific Groundfish Observer Program was established as an interim observer program with rules governing observer coverage codified in regulations. This interim program would be extended four times over the next 20 years by the North Pacific Fishery Management Council (Council) -- the last without a sunset date.

The regulations established in 1990 required vessels 60-125 feet in length (overall) and all vessels fishing pot gear to carry observers at their own cost for 30% of their fishing days in a calendar quarter plus at least one trip in each fishery they participate in (termed the "30% fleet"), and vessels greater than 125 feet in length to carry an observer for 100% of their fishing days at their expense. Some vessels were not required to carry observers. These included vessels less than 60 feet, vessels fishing jig gear or vessels fishing with trawl gear that deliver unsorted codends to processing vessels (termed "catcher processors" or CPs if the vessel also has catching ability and "mothership" or M if the vessel does not) and catcher vessels that fished for Pacific halibut (*Hippoglossus stenolepis*). For shoreside processors, the rules governing observer coverage were based on the estimated tonnage processed in a calendar month: plants that processed less than 500 metric tons (t) a month were exempted from coverage, those that processed between 500 t and 1,000 t a month were required to be observed for 30% of the calendar days, and those that processed more than 1,000 t a month were required to be observed for each day in the month.

Soon after the establishment of the domestic observer program, concerns over the ability and incentive for fishers to bias observer data through the ability to manipulate observer coverage prompted efforts by NMFS and the Council to provide a mechanism for NMFS to gain control over where and when observers were deployed (Faunce and Barbeaux 2011). From 1992 to 2008, several attempts to "restructure" the program were made. In 2010, the Council unanimously decided to move forward with the restructured observer program. In 2012, the Final Rule 77 FR 70062 was published to implement Amendment 86 to the Fishery Management Plan for Groundfish of the BSAI Management Area and Amendment 76 to the Fishery Management Plan for Groundfish of the GOA. Amendments 86/76 added a funding and deployment system for observer coverage to the existing North Pacific Groundfish Observer Program and amended existing observer coverage requirements for vessels and processing plants.

Methods of Observer Deployment

Since 2013, vessels and processors in the groundfish and halibut fisheries off Alaska belong to either full- or partial-coverage categories. Vessels and processors in the full-coverage category obtain observers by contracting directly with observer providers. Vessels and processors in the partial coverage category obtain observers through a NMFS-contracted provider(s), and pay a 1.25% fee on all groundfish and halibut landings.

Decisions as to the type of vessel operation that would be contained within the full- and partial-coverage category were included as part of the supporting analysis in the Public Review Draft of Observer Amendments 86/76 (NPFMC 2011).

The full-coverage category in 2014 included the following:

- Catcher/processors (CPs)
- Motherships
- Catcher-vessels while participating in American Fisheries Act (AFA) or Community Development Quota (CDQ) pollock fisheries
- Catcher-vessels while participating in CDQ groundfish fisheries (except: when fishing longline sablefish and halibut, or when fishing pot gear)
- Catcher-vessels while participating in the Central Gulf of Alaska Rockfish Program (RP)
- Inshore processors when receiving or processing Bering Sea pollock.

Vessels and processors in the 2014 partial coverage category included the following:

- Catcher vessels designated on a Federal Fisheries Permit (FFP) when directed fishing for groundfish in Federally managed or parallel fisheries, except those in the full coverage category
- Catcher-vessels when fishing for halibut IFQ or CDQ

- Catcher-vessels when fishing for sablefish IFQ or fixed gear sablefish CDQ

- Shoreside or stationary floating processors, except those in the full coverage category.

In the partial coverage category, three deployment strata were implemented: trip selection, vessel selection, and zero selection. In the trip selection stratum, vessel owners and/or operators are provided with a username and initial password to the Observer Declare and Deploy System (ODDS) -- a web-based application used to log intended trips (odds.afsc.noaa.gov). Each trip is assigned a random number, and if the number is below a pre-determined threshold, the trip is selected for observer coverage. Users "close" (indicating to NMFS that the trip has been completed) their unselected trips by either selecting one or more landing reports made by their vessel that are provided to them within the logged date range for the trip, or can manually enter the port, processor, and date of landing. Catcher vessels ≥ 57.5 feet in length are placed in the trip-selection stratum.

Catcher-vessels 40-57.5 feet were placed in the vessel selection stratum. In this stratum, the observer program uses past activity in the year prior to create a list of vessels eligible to carry observers, and a sample of vessels from this list are selected for observer coverage. Selected vessels are subject to observation on all of their fishing trips during a 2-month period. Vessels less than 40 feet in length and those fishing jig or troll gear were not subject to observer coverage in 2014 and constitute the "zero selection" stratum.

Small changes have been made to the original design of the observer program since 2013 that were retained in 2014 (NMFS 2013, NMFS 2014a). These changes include:

- Vessels while participating in Electronic Monitoring Pilot Studies were excluded from the obligations of carrying human observers.

THE ANNUAL DEPLOYMENT PLAN AND REVIEW

Analysis and evaluation of the data collected by observers is an on-going process. Decisions as to the amount of coverage (i.e., selection probabilities that are assigned to each partial-coverage category) are based on available funding, the cost of observer coverage, anticipated effort, and the inclusive, cooperative decision-making process of the Council. Each June, NMFS provides the Council with a comprehensive evaluation of past years' observer activities, costs, sampling levels, and implementation issues as well as recommended changes for the coming year. The June report aims to identify areas where improvements are needed to 1) collect the data necessary to manage the groundfish and halibut fisheries, 2) maintain the scientific goals of unbiased data collection, and 3) accomplish the most effective and efficient use of the funds collected through the observer fee. It is intended that this review will inform the Council and the public of how well various aspects of the program are working, and consequently lead to recommendations for improvement. This June report is termed the Annual Report. The NMFS releases a draft Annual Deployment Plan (ADP) by 1 September of each year to allow review by the Groundfish and Crab Plan Teams, as well as the Science and Scientific and Statistical Committee (SSC) and the Council. Based on input from its advisory bodies and the public, the

Council may choose to clarify objectives and provide recommendations. Upon analysis of the Council recommendations, NMFS will make any necessary adjustments to finalize the ADP and release it to the public; ideally the ADP is released to the public prior to the December Council meeting. The final ADP contains the deployment rates that will be programmed into ODDS and used in the following year (NMFS 2014a).

The Observer Science Committee

Each year the Alaska Fisheries Science Center's (AFSC) Fisheries Monitoring and Analysis (FMA) Division establishes an *ad hoc* Observer Science Committee (OSC) for the North Pacific Groundfish and Halibut Observer Program (Observer Program). The OSC is intended to provide scientific advice in the areas of regulatory management, natural science, mathematics, and statistics as they relate to observer deployment and sampling in the groundfish and halibut fisheries of the Bering Sea and Aleutian Islands (BSAI) and the Gulf of Alaska (GOA). OSC members must have practical, analytical, and scientific expertise relating to the observer sampling of groundfish and halibut fisheries of the BSAI and GOA and/or the use of the resulting data. If possible, the OSC is represented by at least one member of the AFSC/FMA (Observer Program) Division, one member of the AFSC/Stock Assessment and Multispecies Assessments Program, one member of the Alaska Regional Office/ Sustainable Fisheries Division (SF), and one member of the International Pacific Halibut Commission (IPHC).

Deployment Performance Review

The following sections contain the OSC review of the deployment of observers in 2014 relative to the intended sampling plan and goals of restructured observer program. This report identifies where possible biases exist and provides recommendations for further evaluation, including potential improvements to the observer deployment process that should be considered during the development of the 2016 Annual Deployment Plan (ADP).

The goal of sampling under the restructured program is to randomize the deployment of observers into fisheries to collect representative data used to estimate catch and bycatch, assess stock status, and determine biological parameters used in ecosystem modeling efforts and salmon stock-of-origin analyses. Therefore, this evaluation focuses on the randomization of observer deployments (primary sampling units) under the restructured Observer Program, and how departures from a random sample affect data quality. It does not evaluate the catch estimation process that is evaluated and summarized in separate documents (Cahalan et al. 2014).

Observer Deployment Performance Metrics

Performance metrics have been developed to assess the efficiency and effectiveness of observer deployment into the partial coverage strata. These metrics reflect four mechanisms that can impact the quality of the data: sample frame discrepancies, non-response, trip differences, and sample size.

Sample frame discrepancies (under- and over-coverage of the sample frame) are used to quantify the differences between the sampled population and the population for which estimates (inferences) are made, as well as to identify possible mechanisms of bias. Non-response assessments are made to quantify the differences between the selected sample (selected trips or vessels expected to be observed) and the actual observed sample (observed trips or vessels after non-response drivers such as releases) that may lead to bias in the resulting data. Other measures that address potential observer deployment effects (sensu the "observer effects" of Benoit and Allard [2009]) are focused on the representativeness of the sample; for example whether observed trips have similar characteristics to unobserved trips such as areas fished, numbers of species landed, and trip duration. Adequacy of sample size is evaluated by assessing whether sample sizes were large enough to ensure data were captured for all types of fisheries. Specifically, the probability of selecting a sample and observing no trips in a specified area is used to evaluate the adequacy of the sample rates used in 2014.

This chapter is an evaluation of whether the deployment of observers is representative of fishing effort. This focus on observer deployment is important because it represents the first tier of the observer hierarchical sampling design from which all haul, species composition, length, age, sex, maturity, and genetic data collections depend on.

It has been argued that variance of the resulting catch estimate be used as a performance metric to determine adequate sample size for observer programs (NMFS 2004). However, given the multiple sources of variance that results from the complex nature of the sampling and estimation routines used in the North Pacific, final variance and catch estimates are neither the only metric nor necessarily the best metric for evaluating stratification and randomization of sampling of primary sample units (trips, vessels). For example, an analytical focus on variance does not evaluate the overall *quality* of the underlying data collection process.

The performance measures listed below are meant to assess the representativeness of the data collected by the observer program through the implementation of the 2014 Annual Deployment Plan.

Description of Performance Metrics Used in this Evaluation

1. Deployment rates for each stratum: This is the basic level of evaluation comparing sampling rates targeted and achieved. Implementation challenges can be identified in this step, such as: sample frame inadequacy, selection biases, and issues with sample unit definitions (e.g., tender trips). Specifically, this section assesses the following:

    a. Sample rates (partial selection strata) and number of samples (vessel selection strata) relative to intended values.
    b. Quantification of under- and over-coverage rates (sample frame discrepancies). Over-coverage of a population occurs when the sample frame includes elements (trips or vessels) that are not part of the target population. When these elements are included in the random sample, effort (time, cost) is expended needlessly. Under-coverage results from having a sample frame that does not include a portion of the target population which can lead to biased data if that portion of the population differs from the population included in the sample frame.

c. Non-response rates. Non-response occurs when randomly selected elements (trips or vessels) are not actually sampled. If these trips or vessels have different fishing behavior (e.g., catch, areas fished) than the rest of the population, the data collected will not represent the entire fleet (non-response bias).

2. Representativeness of the sample: Randomized sampling is a method used to ensure that the results of sampling reflect the underlying population. Departures from randomization can lead to non-representative data and hence potential bias in estimators of parameters of interest. A randomized sample design is expected to achieve a rate of observed events (relative to the trip or vessel strata) that is similar across both space and time. The hypergeometric distribution is used to construct several of these metrics. This distribution describes the probability of selecting sample units (e.g., trips) with specific characteristics (e.g., NMFS Reporting Area) based on a sample taken from a population with known characteristics (e.g., trips that occurred in a NMFS Reporting Area). Representativeness of the sample was divided into three separate components:

   a. Temporal representativeness
      i. Effort plots: plots of expected and actual observed effort over time. Areas where these two lines deviate from each other are indicative of periods with differential realized sample rates (and potential temporal bias).
   b. Spatial representativeness
      i. Maps: Maps provide a visual depiction of the spatial distribution of observer coverage relative to effort in each partial coverage stratum, as well as where low or high coverage rates occurred.
      ii. Probability of selecting a sample and observing a fewer or greater number of trips within an area than would be expected given the implemented sample rates. This probability of observing as many or a more extreme number of trips for each NMFS Area and deployment stratum is determined using the hypergeometric distribution.
   c. Representativeness of trip characteristics
      i. Consistency of trip characteristics for observed and unobserved portions of the stratum. Attributes include:
         – Trip duration
         – Vessel size
         – The number of NMFS Areas visited during the trip
         – The amount of landed catch
         – The number of species in the landed catch (also known as species richness)
         – The proportion of the total landed catch that was due to the most prevalent species (a measure of species diversity).

3. Adequacy of sample size: A well-designed sampling program will have a sample large enough to reasonably ensure that the entire target population is sampled (represented in the data). This determination was made through an examination of the probability of selecting a

sample and having cells (e.g., defined by NMFS Reporting Area and strata) with no observer coverage as determined using the hypergeometric distribution.

EVALUATION OF OBSERVER DEPLOYMENT IN 2014

The deployment of observers into the 2014 Federal fisheries in Alaska needs to be evaluated at the level of the deployment stratum because each stratum is defined by a different sampling unit (trips, vessels) and sampling rate (i.e., time period). In the 2014 ADP, simulated sampling of 2012 fishing effort was conducted to achieve a set of selection rates that was anticipated to result in about a 1 in 10 chance of going over budget. Following a Council request that coverage rates in trip selection be higher than those in vessel selection, vessel selection rates were selected to be less than trip selection by the same relative amount as in the 2013 ADP (vessel selection rates = 0.74 * trip selection rates).

In the 2014 ADP, initial rates were selected assuming a NMFS budget of $4.8 million. However, actual budgets were not known at the time of the 2014 December Council Meeting (NMFS 2014a). Increases in the available budget changed the tolerance for risk for NMFS for 2014. In addition, NMFS noted that the effort in 2012 was unrealistically high for 2014. Therefore, simulations were re-run after the final ADP and before the start of 2014 with a rate set such that expected expenditures were equal to the budget. This is equivalent to the point estimate and is a rate such that the likelihood of deploying over budget is equal to that of deploying under budget. NMFS programmed the ODDS to select 16% of trips and vessel selection draws were conducted to achieve the specified number of vessels in the 2014 ADP (NMFS 2014a).

Tracking Costs

The selection rates translate into costs through fishing effort. Therefore, how close anticipated costs are to actual costs is a function of how well NMFS predicts effort and how well NMFS achieves its sampling rate.

To inform the observer program of costs throughout the year, two sources of information were used. The first was the range of observer days expected to be observed from the ADP simulations. The second was the amount of observer days for which the program had data for (equivalent to payable days). Based on simulations of 2012 fishing data made a year in advance of deployment, the FMA expected observed fishing effort to be 4,718 days at the end of 2014. In 2014, the FMA deployed observers for 4,367 days, or 92.6% of anticipated budget.

On the whole, the results above imply that the FMA was very good at anticipating fishing effort and achieving its desired selection rate. However, upon closer inspection this appears to be the result of lower than expected observer days in trip selection and higher than anticipated observer days in vessel selection (Fig. 1). The reasons for these discrepancies will be explored in greater detail within each stratum separately.

Performance of the Observer Declare and Deploy System in Trip Selection

Random selection of trips in the trip selection stratum is facilitated by the ODDS. The ODDS generates a random number according to pre-determined rates and assigns each logged trip to either "selected to be observed" (selected) or "not selected to be observed" (not selected) categories. The NMFS observer provider has access to all selected trip information necessary to schedule observer logistics. Users of the system are given flexibility to accommodate their fishing operations; up to three trips may be logged in advance of fishing.

Logged trips have different dispositions. They may be closed by a vessel operator after fishing (the desired outcome), or cancelled prior to fishing. Trips can be cancelled by the user or the observer provider. In the former case the trip is recorded as selected and cancelled while in the latter case the logged trip is recorded as a trip waiver. Any remaining trips that have not been closed at the end of the calendar year are automatically cancelled by the ODDS. The number of trips logged in the ODDS in 2014 and their dispositions is summarized in Table 1. Of 4,687 trips logged, a total of 570 trips were cancelled (12.2%), and 15 trips were waived (0.3%).
The ratio of the number of trips cancelled by users that had been selected and those that had not been selected for coverage is useful to determine the amount of potential manipulation of trips. If users were trying to avoid observer coverage, then we would expect the cancellation rate (%) to be higher for selected trips compared to not-selected trips. We found that 5% of non-selected trips logged had been cancelled compared to 18.5% of selected trips logged that had been cancelled. In 2014, ODDS users disproportionately cancelled trips that had been selected for observer coverage compared to trips that had not been selected for coverage.

The flexibility offered by the ODDS means that the outcome of random selection is known to the vessel operator for up to three logged trips. In the case where ODDS users disproportionately cancel selected trips, observer coverage is expected to be less than programmed selection rates. To reduce this potential bias, ODDS is programmed to automatically select the vessel's next logged trip if a previously selected trip was cancelled by the user. Although these "inherited" trips preserve the *number* of selected trips in the year, they cannot prevent the *delay* of selected trips during the year. Therefore the potential for temporal bias is still present.

The extent to which trip selections are altered can be determined by comparing the rate of trip observation expected from 1) random selection of all logged trips (initial selection rate) and 2) random selection of remaining trips after they have had dates changed and are closed or cancelled (final selection rate). In either case, the proportion of trips selected to be observed should fall within what would be expected given the binomial distribution (since each trip is either selected or not selected). The rate obtained in the initial selection process was 15.5% and was within the range of values expected from a binomial distribution (exact binomial test p-value = 0.342). This means that the ODDS was selecting trips according to the programmed rate. The final selection rate after trips were closed and cancelled was 16.6%. The fact that the final selection rate is greater than the initial selection rate results from the fact that cancelled trips that were originally selected for coverage are preserved through the inherit process, while cancelled trips that were not originally selected for coverage are not. These rates and the potential impact of trip selection waivers is presented in Table 2.

Differences in the initial and final selection rates are evident throughout the year. While the original selection rate rapidly rises from zero to approach the programmed rate within a month after the start of the year, that of the final selection rate lags that of the initial rate and does not approach the programmed selection rate until several months later (Fig. 2). Near mid-year, the final selection rate eclipses that of the initial selection rate and remains higher through the remainder of the year. These patterns are consistent with the hypothesis that the disproportionate cancellation of selected trips results in a greater number of selected trips later in the year as the result of the inherit process. Had vessel operators not disproportionately cancelled their initially selected trips, the final selection rate would have been lower.

It is important to remember that ODDS only provides the expectation as to what levels of observer coverage levels should be from actual fishing events. While the 2014 ODDS provided users with a list of Report IDs from *e*Landings from which to close their logged trips, there is no way to know that such linkages between logged and realized trips are accurate. At a minimum, all trips logged should be closed or cancelled by the end of the year. In order to prevent 2014 ODDS trips from bleeding into 2015, trips that were not closed by the end of the year were automatically closed (cancelled) by ODDS. The number of these auto closed trips provides a minimum estimate of the potential mismatch between ODDS and *e*Landings. A total of 259 trip selection trips were auto closed at the end of 2014 by the NMFS (5.9%).

Evaluation of Deployment Rates

This section compares the coverage rate achieved against the expected coverage rates. Unlike the earlier evaluation of the ODDS, data for this evaluation derive from a special database generated for this purpose that utilizes data within the Catch Accounting System (managed by the AKRO), the observer program database NORPAC (managed by the AFSC), and *e*Landings (under joint management by ADF&G and NMFS). Separate rate evaluations are conducted depending on whether the unit of observer deployment was at-sea fishing trips or dockside deliveries of pollock.

At-Sea Deployments

Observers were deployed onto at-sea fishing trips by vessels designated as belonging to full or partial selection categories. There are two deployment strata to evaluate in full coverage; trips belonging to vessels defined in regulation (e.g., AFA, termed regulatory full coverage), and those made by vessels that volunteered to carry full observer coverage when fishing in the BSAI (termed voluntary full coverage). Deployment strata in the partial coverage category include: trips by vessels in trip selection during the year, trips made by vessels in vessel selection during six 2-month selection periods, and trips made by vessels in the no selection category. This last category includes two strata: those vessels designated as belonging to the no selection category in the 2014 ADP, and those that were removed from vessel selection because they had agreed to carry electronic monitoring technology.

Rate evaluations are based on trips for the year with the exception of the vessel selection stratum that is evaluated in terms of vessels in a 2-month time period. Evaluations for the full coverage category and the no selection category are straightforward - either the coverage achieved was equal to 100% or 0%, respectively, or it was not. For trip and vessel selection strata, observed rates were expected to fall between upper and lower bounds on the expected value that were generated from the 0.025 and 0.975 quantiles of a binomial distribution (aka a 95% "confidence bound") for each time period. Coverage levels were considered to have met expectation goals if the actual value was equal to one of the upper or lower confidence bounds, or fell within them. For the trip selection stratum, the expected coverage rate was the rate programmed into ODDS. For the vessel selection strata, the expected number of vessels observed was taken from the 2014 ADP, and the expected bounds for the binomial distribution were determined from Vr where V is the total number of vessels that fished in the stratum each time period and r is the expected rate of coverage from the 2013 ADP (12%).

The 2014 Observer Program had 11 different deployment strata to be evaluated (Table 3). The program met expected rates of coverage for the full-coverage regulatory and full-coverage voluntary strata, the trip selection stratum, four of six time-periods within vessel selection, and the partial coverage no selection. Observer coverage was higher than expected from a 12% selection rate in two of the six time periods within vessel selection. Among all fishing in Federal fisheries of Alaska, 5,883 trips (43%) and 417 vessels (32.8%) were observed.

Coverage Rates in Vessel Selection

Two factors that impact the ability to achieve a target number of vessels to be observed in vessel selection difficult include 1) the lack of a complete sampling frame, and 2) policies that grant releases from observer coverage based on certain conditions. A sampling frame should include all the elements of the population of interest. Hence, a sampling frame for vessel selection would consist of a list of vessels that actually fish in each 2-month deployment period. This list was not available for the vessel selection strata prior to each selection period of 2014. In trip selection, only vessels that intend to fish log trips into ODDS. Consequently, the trip selection sampling frame for the observer program is equal to the target population. However in vessel selection, without a similar notification system informing NMFS of their intent to fish, the sample frame is based on past fishing behavior (specifically whether the vessel landed catch in the same 2-month period the year prior). NMFS used 2012 data to plan for coverage given anticipated budgets for the 2014 ADP, but used data from 2013 to generate lists of vessels to select from for 2014. Obviously the list of vessels that fished 2 years ago or last year may not be the same as the list of vessels that fish in the current year. This introduces two potential sources of error. The first is the selection of vessels that fished prior to 2014 but did not fish during 2014. This is called "over-coverage" and results in sampling inefficiency (this term over-coverage derives from survey research methods and should not be confused with having too much observer coverage). To meet the target sample size (number of vessels), additional vessels are selected to carry observers. The amount of this "over-draw" was based on the expected proportion of vessels in the selection frame that will not fish in 2014 *plus* the proportion of vessels that are selected and will fish, but are expected to be granted a release from observer coverage. The greater this combined proportion, the greater the inefficiency of the sampling process and the greater the amount of over-draw in the selection. For vessel selections 3-6, data from the current year, but from two time periods earlier to accommodate a 60-day advance notice of selection, were also used to construct the sampling frame (e.g., the first time-period selection results could not inform future selections until the third time period selection, the fourth time period selection was informed by the first and second selection results).

The second source of error introduced by an incomplete sampling frame is that a portion of the population has no chance of being selected for observer coverage (no way to select "new" vessels). A new vessel in this case is one that did not fish during a time period in 2013 but will fish in the same time period in 2014; these are not included in the selection frame. These "new" vessels then have no chance to be selected for observer coverage. This is called "under-coverage" and is of particular concern because it represents a potential bias (the term under-coverage derives from survey research methods and should not be confused with having too little observer coverage). Bias would result if these new vessels in 2014 fish differently than vessels that fished in 2013 and were in the selection frame. These combined effects make vessel selection imprecise and inefficient for the NMFS.

Vessels in the vessel selection strata can be classified in numerous ways depending on their fishing, selection, and observation status. Table 4 presents these values for each time period. The number of vessels that fished in 2014 was lower than the number of vessels anticipated to fish in the ADP in all but the second time period (row 6 vs. row 1 in Table 4). Values of the relative amount of overdraw, (expressed as the number of selected vessels divided by the target number of vessels to be observed) ranged between 7.3 and 9.9 (average = 8.6) among time periods. Between 10 and 71 vessels were selected and actually fished among time periods (Table 4, line 10). Between 5 and 35 vessels were selected, fished, and actually observed among time periods (Table 4, line 15).

The number of vessels that would be expected to carry observers after considering release policies is difficult to determine because a release may be granted that is only for a part of the coverage period, or for only some activities. For example, if a vessel is granted a conditional release based on a life raft with insufficient capacity, then we would expect all fishing to be released from coverage. However, if a release was granted for only those trips during which an IFQ holder is on board, the vessel would carry an observer when fishing without an IFQ holder; that is, outside of IFQ fisheries. In this example the vessel has received a release based on certain criteria; in some situations there is an observer on board, whereas on other trips there is not. The data summaries pertaining to the expected number of observed vessels are presented in a generalized level in Table 4 on lines 12-20.

To measure the performance of the vessel selection process, data in Table 4 were expressed as relative percentages (Table 5). Over- and under-coverage rates in the vessel selection sampling frame are not additive since the former is a percentage of the sampling frame, and the latter is a percent difference from the true frame (i.e., the list of vessels that actually fished). Values in these metrics were greatest in the last selection period (Table 5, rows 1 and 2). If being selected for coverage has no effect on the likelihood that a vessel fishes in Federal waters, we would expect that the percentage of vessels that were in the selection frames and did not fish to be approximately equal to the percentage of vessels that were in the selection frame and were selected for coverage and did not fish. Comparing the first and third lines of Table 5 shows that this was the case in the latter four time periods. Only in the second time period did a greater percentage of selected vessels not fish compared to the percentage of vessels that were not selected. With the exception of the second time period, it appears that the act of being selected for coverage did not greatly increase the percentage of vessels that chose not to fish in Federal waters.

The loss of information on trips that should be observed is also presented in Table 5. This type of non-response is represented by the number of vessels that were selected, fished, but were not observed, divided by the number of vessels that fished. It can be caused by conditional release, loss of observer data due to poor quality or failure to follow protocols, or non-compliance. The rate of non-response for "expected to be observed" vessels ranged between 36.8 and 66.2% and gradually increased from the start of the year to a peak in the fourth selection period before decreasing until the end of the year (Table 5, line 4). As expected, a similar pattern was evident in the percentage of vessels released from coverage (36.8 %, line 7).

By dividing the number of desired vessels to be observed from the 2014 ADP by the number of vessels that actually fished in 2014, the expected proportion of vessels to be observed is obtained (Table 5, line 8). Dividing the number of observed vessels by the number of vessels that actually fished in 2014 gives the actual proportion of vessels observed (Table 5, line 9). The achieved coverage rate in vessel selection was close to that expected given the number of vessels that fished, and was greater than expected in the first and third selection periods.

Types of Non-response in Vessel Selection

There were two types of releases granted in 2014: Temporary exemptions and conditional releases. Temporary exemptions were granted when a vessel had more bunk space than life raft capacity. Conditional releases were granted when all available bunks were planned to be occupied by either crew or crew and IFQ holders. Table 6 summarizes the number of vessels that received each type of release in vessel selection.

Spatial Patterns of Non-response in Vessel Selection

The effect of non-response (expected to be observed but were not) on the spatial distribution of observer coverage was evaluated (Table 7). In total, 54% of the vessels and 55% of the trips resulting from these vessels, were in the non-response category (expected to be observed but were not). Non-response percentages by NMFS Area must be interpreted with caution when only a few vessels are present within each category (consider the extreme case where only one vessel fishes-- the only possible percentages are either zero or 100%). With this caveat in mind, where there were more than 10 trips in a NMFS Area, the non-response percentages were similar between areas. No observer data was obtained from four NMFS Areas as a result of conditional releases (Table 7).

Cost Trajectories Revisited

The results of the trip and vessel selection rate evaluations allow us to re-evaluate the results of the cost trajectories in Figure 1. It appears that for trip selection the difference between the expected days observed and actual days observed was due to changes in fishing effort between 2012 and 2014. This conclusion is supported by the fact that random selection in ODDS was according to programmed rates and the rate of observed trips conformed with expectations. For vessel selection the difference between the expected days observed and actual days observed was due to the inability to construct an adequate sampling frame. Supporting evidence comes from the fact that under- and over- coverage among time periods averaged 40.3%.

Coverage Rates for Dockside Monitoring

Observers were assigned to monitor deliveries of walleye pollock (*Gadus chalcogrammus*). The objective of this monitoring is to obtain a count of the number of salmon caught as bycatch and to obtain genetic samples from these fish in each observed pollock delivery. There have been many iterations of the sampling design used to obtain genetic samples from salmon bycatch for the purposes of stock of origin (Faunce 2015). For 2014, the level of dockside monitoring of walleye pollock should be 100% in the full coverage category, and within acceptable tolerance of expected values for a deployment rate of 16% in the partial coverage category. This is because the observer program gains substantial logistical efficiency by having observers that participate in at-sea coverage also monitor corresponding offloads, and all deliveries of this species are expected to occur with trawl gear that is restricted to trip selection.

One issue that arises with this observer program objective is how pollock deliveries are defined. The problem facing the observer is that his or her sampling protocols are dictated by the answer given by the captain as to whether or not this trip will be a pollock trip. Asking the captain for the expected fishery is necessary, since catch is not known before a trip begins. However, the fact that the captain told the observer this was a pollock trip is not recorded in landings records or the observer data. The assignment of a pollock delivery is necessarily made once the fish have been delivered and a landing report has been generated. One approach is to label any delivery where the predominant species is pollock as a pollock delivery (i.e., trip target = pollock) while another is to use a minimum threshold of the landed catch that is comprised of pollock. The first method is referred to as the target definition, while the latter is the (minimum) ratio definition. A minimum percentage in the delivery of 20% was used here to define the ratio method since that is the definition of directed pollock fishing used by the CAS to assign a trip to a management program. Since there are different ways that a delivery can be assigned to the pollock fishery that are not known to the observer prior to monitoring the delivery, there is the potential for the observer to monitor a delivery that is not a pollock delivery, and to not monitor a delivery that is a pollock delivery.

The number of deliveries identified as belonging to the pollock fishery using both definitions is presented in Table 8. There was very good agreement (99%) between definitions across all ports. Among ports, two deliveries at Kodiak in full coverage were not identified by the target definition and King Cove had a relatively low rate of agreement in partial coverage (84.4%). From these results, we defined pollock deliveries using the minimum ratio definition and evaluated observer coverage accordingly.

In partial coverage, unbiased estimates of salmon stock of origin should arise from samples of individual fish obtained from samples of pollock deliveries given randomization protocols. However a random sample of pollock deliveries is not possible because of tendering activity. This activity occurs when a vessel delivers caught fish to a tender and that tender vessel then delivers the fish to a shoreside processing plant. Since tender vessels can provide fuel and food, it is possible that a catcher vessel can remain at sea on a single trip for the entire season. If that trip were logged into ODDS and not selected, the vessels' entire season activity would not be observed. Furthermore, the tender vessel does not log their own trips since they are not fishing and cannot be observed.

The relative impact of tendering activity can be illustrated by comparing the observer coverage rates by port for all pollock deliveries to those without tender deliveries. While very few pollock deliveries were unobserved in full coverage (0.31%), the chance that the coverage rate in partial coverage resulted from 16% random deployment was extremely small (exact binomial test p-value = 0.001; Table 9). However, when deliveries of pollock from tender trips were removed, this likelihood is dramatically increased by two orders of magnitude (p-value = 0.1). The majority of pollock deliveries in the ports of Akutan and King Cove from the partial coverage category were tender deliveries (Table 9).

REPRESENTATIVENESS OF THE SAMPLE

Temporal Patterns in Trip Selection

An examination of temporal patterns in trip selection is warranted since ODDS data demonstrated that observed trips were disproportionately cancelled and coverage levels after trips were logged lagged that of originally logged trips. Under the hypothesis that there is no temporal bias in the observation of trips during the year, the number of observed trips should be close to the expected value of 16%. The cumulative number of trip selection trips was multiplied by 0.16 to obtain the expected number of observed trips, and acceptable bounds of the number of observed trips were obtained from the 0.025 and 0.975 quantiles from the normal approximation of the binomial distribution (the 95% "confidence bounds").

The number of observed trips achieved was outside of their expected values during start of the year (Fig. 3). We would expect that 5% of our observed values would fall outside of our upper and lower expected bounds, and the value was 15.3%. At the end of the year, the likelihood that the number of trips observed resulted from random selection at 16% (exact binomial test p-value) was 0.1. These results mean that while coverage rates were lower than expected at the beginning of the year, the final coverage rate was within expected ranges.

Spatial Representativeness

Under a strictly random selection of trips and with a large enough sample size, the spatial distribution of selected trips should reflect the spatial distribution of all trips. However, the interpretation of results when the number of observed trips deviates from expected values is not straightforward. The hypergeometric distribution was used to calculate the probability of having a given number of items with a certain characteristics (e.g., trip selection trips in NMFS Area 610) in a sample taken from a population (all trips in a stratum) where the number of items with that same characteristic is known (the number of trips in a NMFS Reporting Area based on landings data). The expected number of trips based on this distribution is the number of trips selected divided by the total number of trips (= sample rate) multiplied by the number of trips that fished in an area.

Using this method, we compared the expected number of sample units (trips in trip selection and vessels in vessel selection) with the observed number of sample units in each NMFS Reporting Area and stratum combination (Fig. 4). The size of the data points in Figure 4 represent the probability of observing that number of sample units or a number of sample units farther from the expected number (more extreme). Small data points indicate an observed number of trips or vessels that is unlikely ($p < 0.05$) given randomized observer deployment. Given that there were

17 NMFS Areas fished in trip selection, we would expect there to be 0.05 * 17 = 1 small data points for this stratum. There was indeed 1.

Observations deviated more from expected in vessel selection than in trip selection. Given that there were 69 NMFS Area time period combinations fished in vessel selection, we would expect there to be 0.05 * 69 = 4 small data points. There were 14 small data points. All but one of these combinations had greater number of observed vessels than expected under random deployment. There was a near even distribution between trips taken in the BSAI (8) and the GOA (6). These results should be interpreted with caution however. It is not known which of these outcomes is real and which four are by chance. In addition, vessels may fish in more NMFS Areas when observed than when unobserved, and counts of vessels among NMFS Areas within a 2-month time period are not independent. Not accounting for the clustering of sampling units would result in an inflated number of cells with extreme outcomes than actually exist, although the use of vessel as the unit of measure in this analysis should help alleviate this effect.

The same data in the above analyses can also be presented in maps. Trip selection coverage rates among NMFS Areas ranged from 11.1% to 33.3% (median = 14.5; Fig. 5). The likelihood of this amount of coverage in trip selection is depicted in Figure 6. Vessel selection coverage rates among NMFS Areas were more variable, and ranged from 0% to 100 (median = 16; Fig. 7 and Fig. 8). The likelihood of this amount of coverage in vessel selection is depicted in Figure 9 and Figure 10.

Taken together, the spatial distribution of observer coverage in trip selection is what would be expected under a random sample of trips. However there was a greater number of observed vessels in vessel selection strata than would be expected under random deployment. These results highlight the difficulty in obtaining an adequate sampling frame in vessel selection.

Trip Metrics

This section is focused on answering three questions related to the deployment of observers: 1) are tendered trips identical to non-tendered trips, 2) are observed tendered trips identical to unobserved tendered trips?, and 3) are observed non-tender trips identical to unobserved non-tender trips?

Six trip metrics were examined in each question. These metrics are as follows: the number of NMFS Areas visited in a trip, trip duration (days), the weight of the landed catch (t), the vessel length (m), the number of species in the landed catch, and the proportion (0 to 1) of the landed catch that was due to the most predominant species (pMax). Total catch is comprised of retained and discarded portions. While it may be desirable to compare discarded catch or total catch between groups, there is a problem with this logic since discarded catch from catcher boats is not available from unobserved trips. Therefore retained catch represents the only "apples to apples" comparison available.

The metric vessel length was not included in the 2013 Annual Report. If observers are deployed randomly into the fleet, then the distribution of vessel lengths on observed trips should be equal to that of unobserved trips. Since fishing power is positively correlated to vessel length, this metric is used to help interpret the results from landed weight of catch. For example, differences

between landed catch weight on observed and unobserved trips have different meaning if there is also a difference in vessel length between observed and unobserved trips. Differences in weight *and* length are interpreted as a failure to achieve a random sample of vessels of different sizes, whereas differences in weight only lend more evidence that there is an observer effect.
The number of species within the landed portion of the catch is a measure of species richness. Our pMax metric follows the concepts behind Hill's diversity number N1 that depicts the number of abundant species (Hill 1973) and is a measure of how "pure" catch is since a value of 1 would indicate that only the predominant (and presumed desirable) species was landed.

In the 2013 version of this report comparisons of trips were conducted using simple histograms and visual inspection (Faunce et al. 2014). Here we employ permutation tests (a.k.a randomization tests) to answer the question "How likely is the difference we found given these two groups have the same distribution (in the metric we are comparing)?" Permutation tests compare the actual difference found between two groups to the distribution of many differences derived by randomizing the labels defining the two groups (e.g., observed and unobserved). By randomizing group assignment, the combined distribution of randomized differences represents the sampling distribution under the null hypothesis that the two groups are equal. In this report 10,000 randomized trials are run for each test. The p-value from the test is calculated as the number of randomized trials with greater absolute differences than the actual difference divided by the number of randomized trials. Similar to the other statistical tests used in this report, low p-values indicate rare events and provide evidence to reject the null hypothesis of equality. In an attempt to improve clarity, although five values are calculated in each test; 1) the difference between groups, 2) the mean difference between groups from randomized trials, 3) 1 expressed as a percentage of the mean value of the metric being tested, 4) 2 expressed as a percentage of the mean value of the metric being tests, and 5) the p-value of the test, only Nos. 1, 3 and 5 are presented in relevant tables.

Are Tender Trips Identical to Non-tender Trips?

This comparison is the basis for examining if there is a tendering effect (i.e., differential trip characteristics when vessels use tenders compared to when they do not). Under the null hypothesis tendered and non-tendered trips are the same. Permutation tests examine whether the difference in trip metrics found between these two groups could have arisen from random differences under the null hypothesis. Low p-values (< 0.05) indicate that there is reason to reject the null hypothesis and conclude that there is a tendering effect. In these comparisons differences were calculated by subtracting non-tender trip values from tendered trip values. Of the six metrics compared, three had low p-values. Vessel that delivered to tenders were 11.5% shorter and fished 29.1% longer than non-tendered trips (Table 10). The catch of tender trips was 1.3% less “pure” (more diverse) than non-tendered trips. Although some of these results are small, the likelihood that tendered and non-tendered trips were the same in 2014 is very small.

Are Observed Tendered Trips Identical to Unobserved Tendered Trips?

The finding that tendered trips are different from non-tendered trips necessitates separate examination of an observer effect within tendered and non-tendered trips. This comparison is the basis for examining if there is an observer effect (i.e., differential behavior when observed compared to when not observed) within tendered trips. Under the null hypothesis observed and unobserved tendered trips are the same. Permutation tests examine whether the difference in trip metrics found between these two groups could have arisen from random differences under the null hypothesis. Low p-values (<0.05) indicate that there is reason to reject the null hypothesis and conclude that there is an observer effect. In these comparisons differences were calculated by subtracting unobserved trip values from observed trip values. Of the six metrics compared, 2 had low p-values. Observed vessels that delivered to tenders were 8.8% shorter and catch was 6% less "pure" (more diverse) than unobserved tendered trips (Table 11). There is evidence that observed tender trips in 2014 were different than unobserved tendered trips.

Are Observed Non-tendered Trips Identical to Unobserved Non-tendered Trips?

This comparison is the basis for examining if there is an observer effect (i.e., differential behavior when observed compared to when not observed) within non-tendered trips. Under the null hypothesis observed and unobserved non-tendered trips are the same. Permutation tests examine whether the difference in trip metrics found between these two groups could have arisen from random differences under the null hypothesis. Low p-values (< 0.05) indicate that there is reason to reject the null hypothesis and conclude that there is evidence for an observer effect. In these comparisons differences were calculated by subtracting unobserved trip values from observed trip values. Separate comparisons are made for each partial coverage observer deployment stratum.

Trip Selection

The results of permutation tests for this question are presented in Table 12 (TS Rows). Vessels that carried observers were 2.6% shorter and landed 9.1% less catch than unobserved vessels (Fig. 11).

An additional analysis was carried out following these results to repeat the permutation tests which include the variable of gear (Table 13). The results for trip selection are presented in Figure 12. Since we have six metrics and three gear types, we have 18 tests of which we expect 1 to have low p-values. Instead there were 4. These results demonstrate that the effect of vessel length in trip selection was confounded by gear type. Hook-and-line vessels in trip selection that were observed landed 14.4% less catch and 9.1% more species than unobserved vessels (Table 13, row 13). Trawl vessels in trip selection that were observed fished in 4.2% fewer areas on trips that were 8.4% shorter in duration than unobserved vessels (Table 13, row 15). There were no low p-value tests for trip selection vessels that fished pot gear (Fig. 12). Taken together, there is evidence of an observer effect in trip selection hook and line and trawl gear.

Vessel Selection

Unlike trip selection that has only one time period and six trip metrics, vessel selection has six time periods and six metrics. This means that even without considering gear, there are 36 permutation tests. Under the assumption that observed and unobserved trips are the same, the distribution of resulting p-values from many tests should be uniform (we expect that only 5% of p-values to be below a value of 0.05, only 10% of the values to be below 0.1, etc.). Hence, rather than placing undue emphasis on a particular test result, here a strong deviation from the expected frequency of *all* of the resulting p-values was used as criteria to broadly reject the null hypothesis (Murdoch et al. 2008). If tests of interest are those that have p-values less than 0.05, we would expect there to be two tests of interest, and instead there were 18 (Table 12; Fig. 13). Evidence of an observer effect was found in vessel selection.

This finding of several potential observer effects can result in incorrect conclusions being drawn about the population based on the sample data. These differences may result from true behavioral changes by the vessel (e.g., fishing shorter trips to decrease costs associated with additional persons onboard or fishing to avoid bycatch of limiting species when observed) or from sparse observed trip data resulting in non-representative information from which to draw inferences. The finding of observer effects in 2014 does not guarantee that they will be found with the same level of coverage in future years. Conversely, the absence of an observer effect in these comparisons does not necessarily imply there is no observer effect; other metrics and types of behavior (e.g., precise fishing location, differences in technique, time of day) that were not examined here may have revealed differences between observed and unobserved trips. However, observer effects do not exist in full coverage fleets and there is little justification for lowering coverage rates below those that were used when the observer effect was found.

ADEQUACY OF THE SAMPLE SIZE

In a well-designed sampling program, the observer coverage rate should be large enough to reasonably ensure that the range of fishing activities and characteristics are represented in the sample data. The Catch Accounting System post-stratifies data into groups of fishing activities with similar characteristics (gear, NMFS Area, trip targets) within weekly periods. At low numbers of trips and low sampling rates, the probability of no observer data within a particular post-stratum is increased and may result in expansions of bycatch rates from one type of fishing activity against landings for a different type of fishing activity. For this reason it is important to have a large enough sample (observed trips and vessels) to have reasonable expectation of observing all types of fishing.

Over the course of an entire year, some NMFS Areas have low fishing effort and as a result have a relatively high probability of being missed by the simple random sampling represented by observer deployments. The fishing effort data for each stratum (trip and vessel selection for each time period) and the sample size (number of observed trips in trip selection and vessels in vessel selection) over the course of 2014 was used to illustrate their combined effect on the probability of a NMFS Area containing observer data using the hypergeometric distribution (Fig. 14). From this figure it can be seen how 1) the likelihood of at least one observation is increased with sampling units (trips or vessels fishing) and 2) is also increased with an increase in the selection rate. The results in Figure 14 should be interpreted as an optimistic simplification since including additional factors such as week, gear, and target will decrease cell size and increase the probabilities of obtaining no observer data in the random sample. Sample size requirements to ensure data are present in all cells of interest will be evaluated during the planning process for 2016 and are the focus of other analyses conducted by the NMFS (NMFS 2015).

RECOMMENDATIONS TO IMPROVE DATA QUALITY

The Observer Science Committee made the following recommendations in its 2013 review of observer deployment to be considered in developing the 2014 ADP (Faunce et al. 2014, NMFS 2014b). Following each recommendation is the outcome of that recommendation for 2014 in italics.

Recommendations from 2013

- The sampling frame in vessel selection would be improved through a check-in system whereby vessels would notify the Observer Program of their intent to fish and would in return be notified of whether the vessel would require an observer and the duration of the observation period. This type of check-in system is identical to the procedure currently used in trip selection. Use of such a system would greatly reduce errors due to oversampling and improve the efficiency of the selection process.

  Such a check-in system was not implemented in 2014. However, noting the problems with vessel selection, this method of observer deployment was discontinued in 2015 (NMFS 2014c).

- The conditional release policy imparts bias into the observer data. If such releases are continued, then they should apply to all fishing activities within the sampling unit (all trips made by a vessel during the time period, and not only during certain fishing activities).

  This recommendation was not adopted in 2014. Consequently it was difficult to identify the trips within vessel selection time periods that were required to carry observers but did not.

- The selection rate in ODDS should remain constant throughout the year. Changing the selection rate creates temporal strata. Rather than reduce the selection rate in ODDS to reduce the risk of cost overages, we recommend that NMFS use budget buffers if possible to mitigate for the rare event of overage.

  This recommendation was adopted in the 2014 ADP.

- Data analyses continue to be hampered by the lack of a trip identifier. We recommend that the linkage between ODDS and *e*Landings be strengthened.

  A trip identifier has not been implemented to date.

Recommendations from 2014

Below are the Observer Science Committees recommendations to improve the 2016 ADP:

- Providing vessel operators the flexibility in ODDS to log 3 trips also provides vessels with the ability to delay observer coverage and potentially bias observer data. The current protocols of 1) allowing selected trips to be cancelled in ODDS and 2) allowing multiple trips to be logged prior to sailing should be re-evaluated. Changing these protocols should reduce the time lag in observer coverage and temporal bias exhibited in trip-selection during 2013 and in 2014.

- The ability of a Catcher Processor to retain product for more than several days without spoilage means that trip durations and landed catch per trip are likely to be larger from catcher vessels that cannot freeze their catch. An expansion of the number of Catcher

Processors in the partial coverage class would necessitate their treatment as a separate stratum with a potentially different selection rate in ODDS.

- The use of metrics known before a trip begins is necessary for the designation of deployment strata. Each trip must be assigned to one and only one deployment strata at the time it is logged. The merits of deploying observers by gear and FMP should be explored in future ADPs. There are FMPs and gear types for example that have low effort and are highly likely to be missed in random selection procedures without high selection rates.

- The assumption used in the ADP that fishing effort in the following year will be equal to the fishing effort two years prior should be improved upon. The NMFS should develop better tools such as models to predict fishing effort.

- The practice of granting releases whereby vessels are sometimes subject to human observer coverage and sometimes not subject to human observer coverage should be discontinued. We recommend that a list of vessels that cannot carry an observer be generated. The list should be updated each calendar year. This list defines a new strata to be observed with alternatives to human monitoring, and should be included in the annual deployment plan and annual review.

- We repeat our 2013 recommendation that the linkage between ODDS and *e*Landings be strengthened through the use of a trip identifier.

Tender vessel activities are problematic for the observer program for several reasons. First, the regulatory definition of a trip means that an operator of a vessel in partial coverage can use an unselected logged trip to deliver to a tender for an extended duration of time unobserved. In the extreme, the vessel could take a single trip that encompasses the entire fishing effort by the vessel. Second, vessels that act as tenders are not covered under the safety requirement of the MSA, meaning that they cannot be used to deploy or house observers. Third, the catch that is delivered to a tender is not accessible to an observer. Finally, the tender vessel, by its very nature, mixes catch from multiple deliveries, meaning that salmon bycatch if identified by an observer dockside could not be attributed to a catcher vessel trip.

- The ability of the observer program to obtain a representative sample of salmon bycatch from the GOA pollock fishery for genetic stock composition analysis is compromised by three factors. In increasing magnitude these factors are: 1) the fact that observers are dependent on the response of the captain on whether or not the trip is a pollock trip, 2) insufficient resources to ensure perfect detection of salmon in the delivery at the processing facility, and 3) the inability to be deployed to or monitor tender deliveries. We do not see an easy solution to 1; deployment into fishery is problematic since catch that determines fishery has not yet occurred at the time of deployment. The GOA Chinook stock compositions have been remarkably stable between the years of 2010-2015 (Guyon et al. 2015, slide 12). Alternatives to the *status quo* monitoring of pollock deliveries include: 1) the collection of genetic tissues by citizen or third party other than the observer program or 2) providing additional funds to institute a more rigorous dockside monitoring by the observer program. Of these, the former is cost-effective to the observer program while the latter is more expensive. Costs to the observer program to obtain genetic bycatch material reduces the available revenue for at-sea observer coverage; it is this at-sea observer coverage which should be the primary deployment objective of the observer program since observers are the only source of discard at-sea information for NMFS to use in fisheries management.

ACKNOWLEDGMENTS

The 2014 version of ODDS was maintained by Glenn Campbell, Paul Packer, Martin Park, Julie Blair, and Craig Faunce.

CITATIONS

Benoit, 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.

Cahalan, J., Mondragon, J., and J. Gasper. 2014. Catch sampling and estimation in the Federal groundfish fisheries off Alaska: 2015 Edition. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-286, 46 p.

Faunce, C. H. 2015. Evolution of observer methods to obtain genetic material from Chinook salmon bycatch in the Alaska pollock fishery. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-288, 28 p.

Faunce, C., Cahalan, J., Gasper, J., A'mar, T., Lowe, S., Wallace, F. and Webster, R. 2014. Deployment performance review of the 2013 North Pacific Groundfish and Halibut Observer Program. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-281. 74 p.

Faunce, C. H., and Barbeaux, S. J. 2011. The frequency and quantity of Alaskan groundfish catcher-vessel landings made with and without an observer. ICES J. Mar. Sci. 68:1757-1763.

Guyon, J. R., C. M. Guthrie III, A. R. Munro, J. Jasper, and W. D. Templin. 2015. Genetic stock composition analysis of the Chinook salmon bycatch in the Gulf of Alaska walleye pollock (*Gadus chalcogrammus*) trawl fisheries. 26 p. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-291, presentation made for the April 2015 meeting of the North Pacific Fisheries Management Council. Presentation available online at http://legistar2.granicus.com/npfmc/meetings/2015/4/923_A_North_Pacific_Council_15-04-06_Meeting_Agenda.pdf.

Hill, M.O. 1973. Diversity and evenness: A unifying notation and its consequences. Ecology 61: 225-236.

Murdoch, D. J., Y. Tsai, and J. Adcock. 2008. P-values are random variables. Am. Stat. 62:242-245.

Nelson Jr., R., R. French, R. and J. Wall. 1981. Sampling by U.S. observers on foreign fishing vessels in the eastern Bering Sea and Aleutian Island region, 1977-78. Mar. Fish. Rev. 43:1-19.

NMFS. 2015. Draft Supplement to the Environmental Assessment for Restructuring the Program for Observer Procurement and Deployment in the North Pacific. NMFS, Alaska Regional Office, Juneau. May 2015.

NMFS. 2014a. 2014 Annual Deployment Plan for Observers in the Groundfish and Halibut Fisheries off Alaska. December 2013. 21 p. plus appendices. Available online at http://alaskafisheries.noaa.gov/sustainablefisheries/observers/adp2014.pdf.

NMFS. 2014b. North Pacific Groundfish and Halibut Observer Program 2013 Annual Report. National Oceanic and Atmospheric Administration, 709 West 9th Street. Juneau, Alaska 99802. 106 p. plus appendices. Available online at http://alaskafisheries.noaa.gov/sustainablefisheries/observers/annualrpt2013.pdf.

NMFS. 2014c. 2015 Annual Deployment Plan for Observers in the Groundfish and Halibut Fisheries off Alaska. December 2014. 34 p. Available online at http://alaskafisheries.noaa.gov/sustainablefisheries/observers/final2015adp.pdf.

NMFS. 2013. 2013 Annual Deployment Plan for Observers in the Groundfish and Halibut Fisheries off Alaska. 39 p. plus appendices. National Oceanic and Atmospheric Administration, 709 West 9th Street. Juneau, Alaska 99802. Available online at http://alaskafisheries.noaa.gov/sustainablefisheries/observers/ADP_Final_2013.pdf.
NMFS. 2004. Evaluating bycatch: a national approach to standardized bycatch monitoring programs. U.S. Dep. Commer., NOAA Tech. Memo. NMFSF/SPO-66, 108 p. Available online at http://www.nmfs.noaa.gov/by_catch/SPO_final_rev_12204.pdf.
NPFMC (North Pacific Fishery Management Council). 2011. Environmental Assessment/Regulatory Impact Review/Initial Regulatory Flexibility Analysis for Proposed Amendment 86 to the Fishery Management Plan for Groundfish of the Bering sea/Aleutian Islands Management Area and Amendment 76 to the Fishery Management Plan for Groundfish of the Gulf of Alaska: Restructuring the Program for Observer Procurement and Deployment in the North Pacific. March 2011. 239 pages plus appendices. Available online at http://alaskafisheries.noaa.gov/analyses/observer/amd86_amd76_earirirfa0311.pdf.
Wall, J., R. French, and R. Nelson Jr. 1981. Foreign fisheries in the Gulf of Alaska, 1977-78. Mar. Fish. Rev. 43:20-35.

TABLES AND FIGURES

Table 1. -- Disposition of trips in the ODSS for 2014. TS: Trip selection. “Paper” indicates trips that were logged when the ODDS was not available.

Strata	Random Selection	Logged	Cancelled by System	Cancelled by User	Waived	Paper
Trip-Selection	Not Selected	3692	258	183	0	0
Trip-Selection	Selected	675	0	125	0	0
Trip-Selection	Not Assigned	16	1	0	15	1
Voluntary 100%	Not Assigned	304	0	3	0	0
Total		4687	259	311	15	1

Table 2. -- Number of logged trip selection trips that were selected using the initial random number generator (Random Selection Only) and those that remained after user manipulation (Final Expected). The relative impact of waivers in trip selection is shown in the last column.

Variable	Random Selection Only	Final Expected	Final Expected if No Waivers
Selected	675	635	650
Total	4367	3816	3816
Selection %	15.5	16.6	17.0

Table 3. -- Number of total vessels (V), sampled vessels (v), total trips (N), sampled trips (n) for each observer deployment stratum in 2014. Totals are unique vessels. Expected coverages are in percent for trip selection and number of vessels for vessel selection. TS: Trip selection, VS: Vessel selection, ZS: Zero selection, EM: Electronic Monitoring, WIE= within expected.

Coverage	Strata	Time Period	V	v	N	n	% Trips Observed	% Observed (Deployment Type)	Expected Coverage	Expected Coverage (min)	Expected Coverage (max)	Meets Expectations?
Full	Regulatory	Year	166	166	4,588	4,587	100.0	100				
Full	Voluntary	Year	30	30	310	310	100.0	100				
Full	Total	Year	171	171	4,898	4,897	100.0	100				
Partial	TS Total	Year	293	199	4,390	662	15.1	15.1	16	14	16.2	Yes
Partial	VS	Jan. - Feb.	50	12	293	69	23.5	24	9	2	11	No*
Partial	VS	Mar. - Apr.	160	17	471	46	9.8	10.6	16	12	28	Yes
Partial	VS	May - Jun.	173	35	434	79	18.2	20.2	24	13	29	No*
Partial	VS	Jul - Aug.	135	24	289	62	21.5	17.8	18	9	24	Yes
Partial	VS	Sep. - Oct.	168	19	476	49	10.3	11.3	20	12	29	Yes
Partial	VS	Nov. - Dec.	32	5	116	19	16.4	15.6	5	1	8	Yes
Partial	VS Total	Year	375	86	2,079	324	15.6	22.9				
Partial	ZS	Year	484	0	2,305	0	0.0	0				
Partial	ZS (EM)	Sep. - Oct.	5	0	15	0	0.0	0				
Partial	ZS Total	Year	489	0	2320	0	0.0	0				
Total Fleet	Total	Year	1,27	417	13,687	5,883	43.0					

**Observed > Expected*

Table 4. -- The number of vessels that fall under specific criteria within the vessel selection strata.

Row	Metric	Jan.-Feb.	Mar.-Apr.	May-Jun.	Jul.-Aug.	Sep.-Oct.	Nov.-Dec.
1	Anticipated to fish (2014 ADP)	85	154	233	177	200	48
2	In selection frame (2013 data); F	66	158	215	150	159	46
3	In frame and fished; fY	36	116	129	76	96	16
4	In frame and did not fish; fN	30	42	86	74	63	30
5	Not in frame and fished (potential bias); f0	14	44	44	59	72	16
6	Active (fished = true frame); f*=f0 + fY	50	160	173	135	168	32
7	Desired to be observed; vT	9	16	24	18	20	5
8	Selected for coverage; vS	27	43	117	141	90	27
9	Selected by did not fish; vN	8	16	51	70	43	17
10	Selected and fished; vF	19	27	66	71	47	10
11	Selected, fished, and never released	12	15	33	23	18	6
12	Selected, fished, and had released trips; vR	7	12	33	48	29	4
13	Selected, fished, released for the entire period	6	12	32	43	26	4
14	Selected, fished, released part of the period	1	0	1	5	3	0
15	Selected and obseved total, v	12	17	35	24	19	5
16	Selected with at least one non-released trip (Expected Observed)	13	15	34	28	21	6
17	Selected, not released, all data present	11	15	29	19	19	5
18	Selected, not released, some data missing	1	0	1	1	0	0
19	Selected, not released, all data missing (potential violation)	1	0	4	8	2	1
20	Selected, released, but observer data; v?	0	2	5	4	0	0

Table 5. -- Vessel-selection rates expressed as percentages (all rate formulations multiplied by 100). Abbreviations follow Table 4.

Row	Metric	Jan.-Feb.	Mar.-Apr.	May-Jun.	Jul.-Aug.	Sep.-Oct.	Nov.-Dec.
1	Error in sampling frame due to over-coverage (% of sample frame); fN/F	45.5	26.6	40.0	49.3	39.6	65.2
2	Error in sampling Frame due to under-coverage (% of True Frame); f0/f*	28.0	27.5	25.4	43.7	42.9	50.0
3	Error due to non-response: selected and did not fish; vN/vS	29.6	37.2	43.6	49.6	47.8	63.0
4	Error due to non-response: Selected, fished, and not observed (vF-v)/vF	36.8	37.0	47.0	66.2	59.6	50.0
5	Chance of selection if in frame and fished; vF/fY	28.8	17.1	30.7	47.3	29.6	21.7
6	Chance of selection if not in frame	0.0	0.0	0.0	0.0	0.0	0.0
7	Percent selected boats that fished and given some sort of release; vR/vF	36.8	44.4	50.0	67.6	61.7	40.0
8	Percent coverage desired; vT/f*	18.0	10.0	13.9	13.3	11.9	15.6
9	Percent coverage acheived; v/f*	24.0	10.6	20.2	17.8	11.3	15.6

Table 6. -- Number of vessels that received two types of releases from observer coverage in each time period of 2014 vessel selection.

Time Period	Temporary Exemption	Conditional Release	Total Vessels
Jan. - Feb.	1	6	7
Mar. - Apr.	1	11	12
May - Jun.	2	31	33
Jul - Aug.	7	41	48
Sep. - Oct.	3	26	29
Nov. - Dec.	0	4	4
Year	10	91	101

Table 7. -- The total number of trips and vessels in the vessel selection strata that were either observed or conditionally released. The number of vessels and trips are not unique among individual cells of this table (trips and vessels can cross NMFS Reporting areas), so totals should be interpreted with caution. NMFS Reporting Areas > = 600 are located in the Gulf of Alaska, Areas 541-543 are located in the Aleutian Islands, and other areas are located in the Bering Sea.

NMFS Reporting Area	Observed trips	Released trips	Released trips (%)	Observed vessels	Released vessels	Released vessels (%)
513	0	1	100	0	1	100
514	0	5	100	0	1	100
517	0	1	100	0	1	100
518	6	4	40	4	2	33
519	13	18	58	4	2	33
521	0	1	100	0	1	100
523	1	0	0	1	0	0
541	17	10	37	4	4	50
542	5	4	44	2	3	60
543	1	1	50	1	1	50
610	47	46	49	12	14	54
620	42	24	36	18	17	49
630	107	156	59	33	40	55
640	16	21	57	9	12	57
649	7	12	63	4	6	60
650	59	82	58	26	31	54
659	32	29	48	23	26	53
Total	324	390	55	86	101	54

Table 8. -- Comparison of the number of pollock deliveries during 2014 as defined by predominant species (Target Definition) and at least 20% pollock (Ratio Definition) by port, coverage category and Fishery Management Plan.

FMP	Coverage	Port	Ratio Definition	Target Definition	Agreement %
Bering Sea	Full	Akutan	737	736	99.9
Bering Sea	Full	Dutch Harbor	783	782	99.9
Bering Sea	Full	Inshore Floating Processor	310	310	100.0
Bering Sea	Full	King Cove	83	83	100.0
Gulf of Alaska	Full	Kodiak	2	0	0.0
Gulf of Alaska	Partial	Akutan	20	20	100.0
Gulf of Alaska	Partial	Inshore Floating Processor	13	13	100.0
Gulf of Alaska	Partial	King Cove	135	114	84.4
Gulf of Alaska	Partial	Kodiak	1,196	1,197	100.1
Gulf of Alaska	Partial	Sand Point	238	228	95.8
Gulf of Alaska	Partial	Seward	3	3	100.0
Total			3,520	3,486	99.0

Table 9. -- The number of pollock deliveries by observation status and tendering status. BSAI: Bering Sea and Aleutian Islands, GOA: Gulf of Alaska, IFP: Inshore Floating Processor, H: Harbor.

FMP	Coverage	Port	Trips Total	Trips Observed	% Observed	p-value Trips Observed	% Tender Trips	% Observed without Tenders	p-value Trips Observed without Tenders
BSAI	Full	Akutan	737	735	100		0	100	
BSAI	Full	Dutch H.	783	782	100		0	100	
BSAI	Full	IFP	310	309	100		0	100	
BSAI	Full	King Cove	83	83	100		0	100	
GOA	Full	Kodiak	2	0	0		0	0	
GOA	Partial	Akutan	20	0	0		80	0	
GOA	Partial	IFP	13	1	8		15	9	
GOA	Partial	King Cove	135	3	2		92	27	
GOA	Partial	Kodiak	1,196	167	14		0	14	
GOA	Partial	Sand Point	238	38	16		3	16	
GOA	Partial	Seward	3	1	33		0	33	
Total	Full		1,915	1,909	100		0	100	
Total	Partial		1,605	210	13	0.001	9	14	0.1

Table 10. -- Results of permutation tests between tendered and non-tendered trips in 2014. OD: Observed Difference. Differences are calculated from tendered minus non-tendered trips.

NMFS Areas	Days Fished	Landed Catch	pMax Species	Species Landed	Vessel Length	Metric
0.008	0.884	4.826	-0.012	0.169	-7.883	Observed Difference
0.748	29.104	9.525	-1.277	4.869	-11.538	OD (%)
0.364	**0.000**	0.092	**0.006**	0.196	**0.000**	p-value

Table 11. -- Results of permutation tests between observed and unobserved tendered trips in 2014. OD: Observed difference. Differences are calculated from observed minus unobserved.

NMFS Areas	Days Fished	Landed Catch	pMax Species	Species Landed	Vessel Length	Metric
0.007	-0.536	-23.384	-0.057	0.433	-5.444	Observed Difference
0.642	-14.177	-42.727	-6.032	11.989	-8.824	OD (%)
1.000	0.355	0.068	**0.001**	0.317	**0.025**	p-value

Table 12. -- Results of permutation tests between observed and unobserved non tendered trips in 2014. OD: Observed difference. Differences are calculated from observed minus unobserved. TS: Trip selection, VS: Vessel selection.

Row	Strata	Time Period	NMFS Areas	Days Fished	Landed Catch	pMax Species	Species Landed	Vessel Length	Metric
1	TS	Jan. - Dec.	-0.015	0.160	-4.970	0.002	-0.119	-1.967	OD
2	VS	Jan. - Feb.	0.038	0.232	1.750	-0.013	0.756	2.531	OD
3	VS	Mar. - Apr.	0.078	0.893	1.096	-0.070	1.088	3.558	OD
4	VS	May - Jun.	0.020	0.329	-0.858	0.013	-0.077	3.008	OD
5	VS	Jul - Aug.	0.090	1.806	2.203	0.035	-0.336	3.332	OD
6	VS	Sep. - Oct.	0.034	1.941	0.669	0.012	-0.135	4.399	OD
7	VS	Nov. - Dec.	-0.021	0.939	18.862	0.018	-0.672	8.787	OD
8	TS	Jan. - Dec.	-1.319	4.299	-9.142	0.166	-2.762	-2.555	OD (%)
9	VS	Jan. - Feb.	3.730	7.420	15.164	-1.409	25.650	5.568	OD (%)
10	VS	Mar. - Apr.	7.516	23.093	14.787	-7.749	29.112	7.342	OD (%)
11	VS	May - Jun.	1.850	7.576	-18.618	1.489	-2.056	6.123	OD (%)
12	VS	Jul - Aug.	8.237	39.948	47.956	3.926	-10.958	6.877	OD (%)
13	VS	Sep. - Oct.	3.213	50.276	12.119	1.343	-4.387	9.082	OD (%)
14	VS	Nov. - Dec.	-2.048	28.256	139.685	1.977	-21.756	19.443	OD (%)
15	TS	Jan. - Dec.	0.320	0.128	**0.015**	0.817	0.370	**0.028**	p-value
16	VS	Jan. - Feb.	0.127	0.130	0.095	0.165	**0.002**	**0.000**	p-value
17	VS	Mar. - Apr.	**0.022**	**0.001**	0.208	**0.000**	**0.000**	**0.000**	p-value
18	VS	May - Jun.	0.617	0.178	0.066	0.460	0.775	**0.000**	p-value
19	VS	Jul - Aug.	**0.046**	**0.000**	**0.000**	0.083	0.267	**0.000**	p-value
20	VS	Sep. - Oct.	0.504	**0.000**	0.382	0.532	0.647	**0.000**	p-value
21	VS	Nov. - Dec.	1.000	**0.002**	**0.000**	0.305	**0.029**	**0.000**	p-value

Table 13. -- Table of observed differences (OD), percent observed differences (% OD), and the permutation test p-values. HAL: Hook and Line gear. TS: Trip selection. VS: Vessel selection. NaN: Not a number, NA: Not Available. These codes arise because all selected vessels that made Pot trips in the September – October selection period (2) were released and there were no observed trips.

Row	Strata	Time Period	Gear	NMFS Areas	Days Fished	Landed Catch	pMax Species	Species Landed	Vessel Length	Metric
1	TS	Jan. - Dec.	HAL	0.019	0.434	-1.673	-0.018	0.332	-1.392	OD
2	TS	Jan. - Dec.	Pots	0.001	0.093	-0.230	-0.006	0.003	-2.101	OD
3	TS	Jan. - Dec.	Trawl	-0.047	-0.213	-3.793	0.016	-0.211	-0.862	OD
4	VS	Jan. - Feb.	HAL	0.058	0.619	-0.026	-0.030	0.918	-0.110	OD
5	VS	Jan. - Feb.	Pots	0.000	-0.586	7.574	-0.002	1.043	6.599	OD
6	VS	Mar. - Apr.	HAL	0.078	0.893	1.102	-0.163	1.090	3.558	OD
7	VS	May - Jun.	HAL	0.020	0.329	-0.858	0.013	-0.077	3.008	OD
8	VS	Jul - Aug.	HAL	0.090	1.806	2.203	0.035	-0.336	3.332	OD
9	VS	Sep. - Oct.	HAL	0.024	1.674	1.275	0.024	-0.345	5.289	OD
10	VS	Sep. - Oct.	Pots	NaN	NaN	NaN	NaN	NaN	NaN	OD
11	VS	Nov. - Dec.	HAL	-0.022	1.378	1.826	-0.042	-0.189	6.222	OD
12	VS	Nov. - Dec.	Pots	0.000	1.222	42.281	0.055	0.056	3.833	OD
13	TS	Jan. - Dec.	HAL	1.555	7.371	-14.380	-2.112	9.070	-2.171	OD (%)
14	TS	Jan. - Dec.	Pots	0.116	2.535	-0.644	-0.563	0.151	-2.776	OD (%)
15	TS	Jan. - Dec.	Trawl	-4.243	-8.359	-4.348	1.833	-3.597	-1.016	OD (%)
16	VS	Jan. - Feb.	HAL	5.671	19.236	-0.222	-3.216	28.603	-0.250	OD (%)
17	VS	Jan. - Feb.	Pots	0.000	-21.403	72.540	-0.236	54.883	12.970	OD (%)
18	VS	Mar. - Apr.	HAL	7.516	23.093	14.879	-16.585	29.192	7.342	OD (%)
19	VS	May - Jun.	HAL	1.850	7.576	-18.618	1.489	-2.056	6.123	OD (%)
20	VS	Jul - Aug.	HAL	8.237	39.948	47.956	3.926	-10.958	6.877	OD (%)
21	VS	Sep. - Oct.	HAL	2.296	40.552	25.554	2.699	-10.567	11.085	OD (%)
22	VS	Sep. - Oct.	Pots	NaN	NaN	NaN	NaN	NaN	NaN	OD (%)
23	VS	Nov. - Dec.	HAL	-2.179	41.005	16.762	-4.674	-5.776	14.206	OD (%)
24	VS	Nov. - Dec.	Pots	0.000	39.855	136.808	5.584	2.976	7.038	OD (%)
25	TS	Jan. - Dec.	HAL	0.653	0.104	**0.015**	0.104	**0.032**	0.099	p-value
26	TS	Jan. - Dec.	Pots	1.000	0.529	0.924	0.124	1.000	0.346	p-value
27	TS	Jan. - Dec.	Trawl	**0.025**	**0.000**	0.160	0.120	0.340	0.451	p-value
28	VS	Jan. - Feb.	HAL	0.089	**0.001**	0.980	**0.006**	**0.003**	0.871	p-value
29	VS	Jan. - Feb.	Pots	1.000	**0.029**	**0.002**	0.727	**0.000**	**0.000**	p-value
30	VS	Mar. - Apr.	HAL	0.025	**0.001**	0.203	0.103	**0.000**	**0.000**	p-value
31	VS	May - Jun.	HAL	0.613	0.170	0.071	0.456	0.781	**0.000**	p-value
32	VS	Jul - Aug.	HAL	**0.042**	**0.000**	**0.000**	0.090	0.275	**0.000**	p-value
33	VS	Sep. - Oct.	HAL	0.754	**0.000**	0.077	0.239	0.248	**0.000**	p-value
34	VS	Sep. - Oct.	Pots	NA	NA	NA	NA	NA	NA	p-value
35	VS	Nov. - Dec.	HAL	1.000	**0.001**	0.340	0.061	0.673	**0.000**	p-value
36	VS	Nov. - Dec.	Pots	1.000	**0.014**	**0.000**	**0.035**	1.000	0.140	p-value

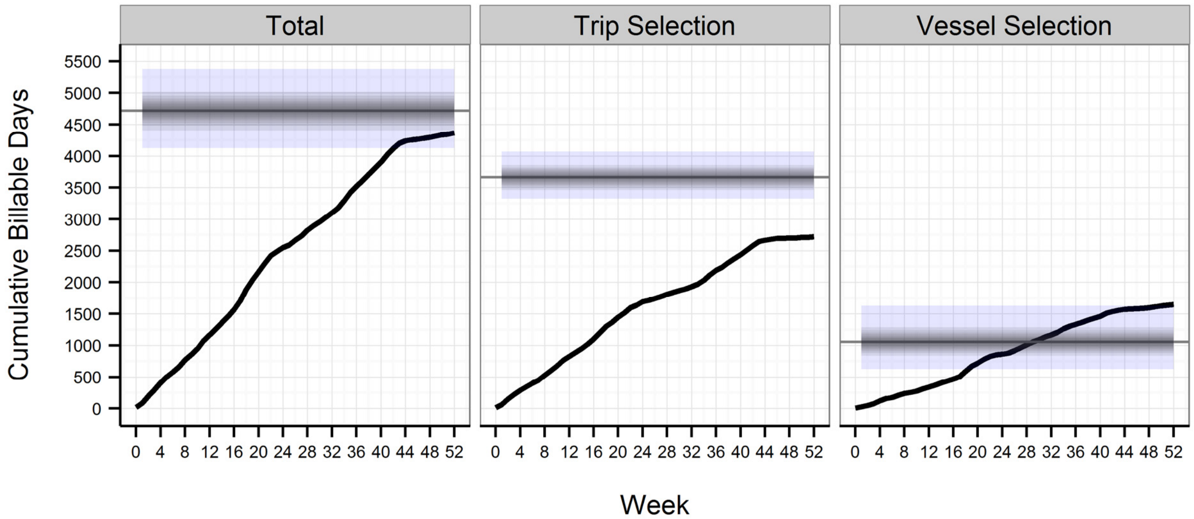

Figure 1. -- Cumulative plots of the number of billable days expected from observer data in 2014. Horizontal bands denote the range of potential billable days that were estimated in December 2013. Shading is proportional to the expected likelihood from 2014 ADP simulations.

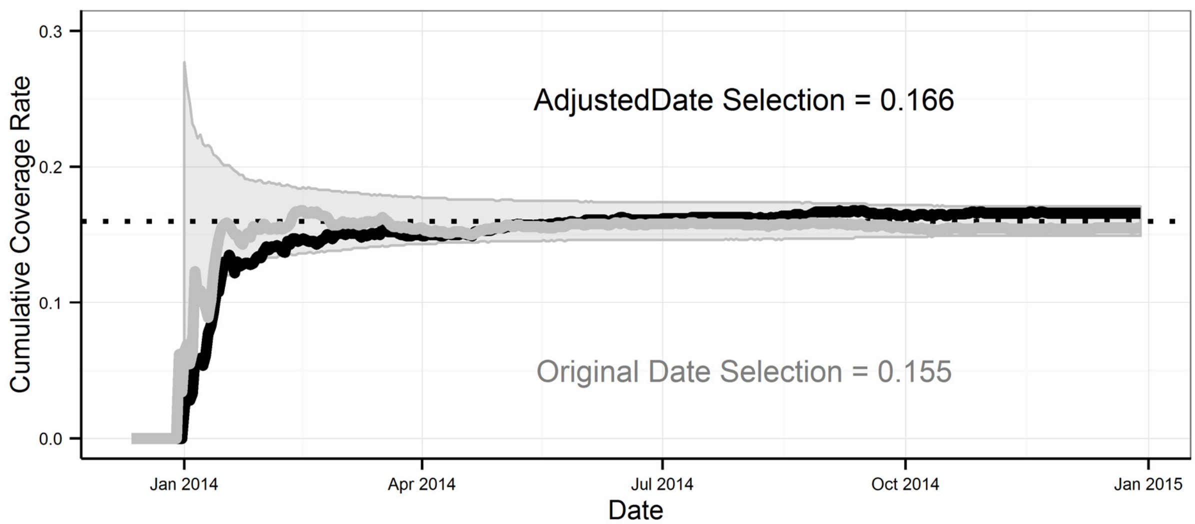

Figure 2. -- Rate of selected trips logged into ODDS organized by original date entered for all trips (grey line and grey text), and final date considering only non-cancelled trips (black line and black text). The programmed selection rate is depicted as the dotted line. Grey shaded areas denote the range of coverage rates that correspond to the 95% 'confidence intervals' expected from the binomial distribution. The final coverage rate was higher than if trip dates had not been altered and trips not cancelled.

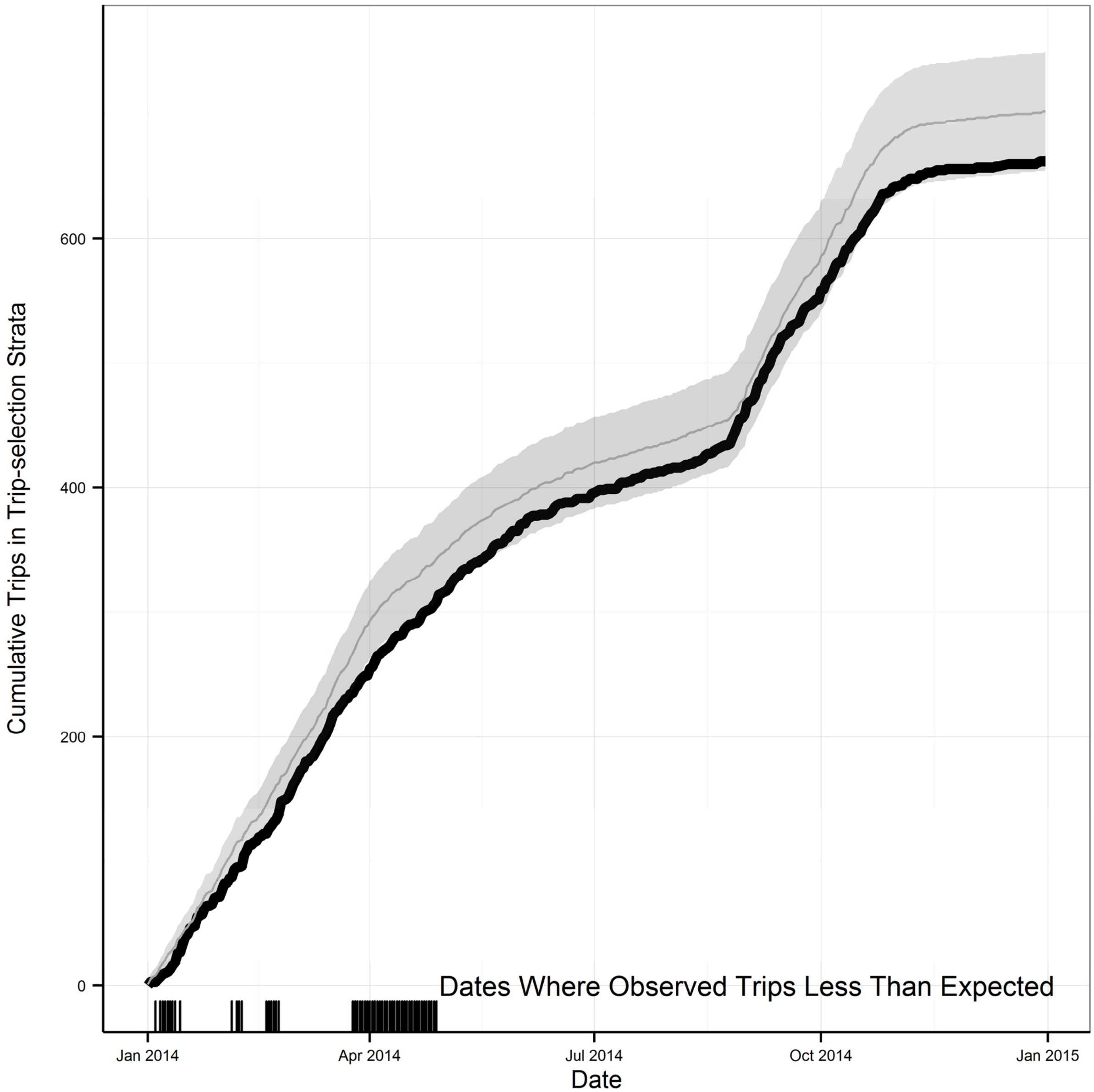

Figure 3. -- Cumulative number of trips observed during 2014 (black line) compared to the expected cumulative number of trips from an observation rate of 16%. Grey shaded areas denote the range of coverage rates that correspond to the 95% 'confidence intervals' expected from the binomial distribution. Dates where the observed number of trips is less or more than the range of expected values are depicted as tick marks on the x-axis.

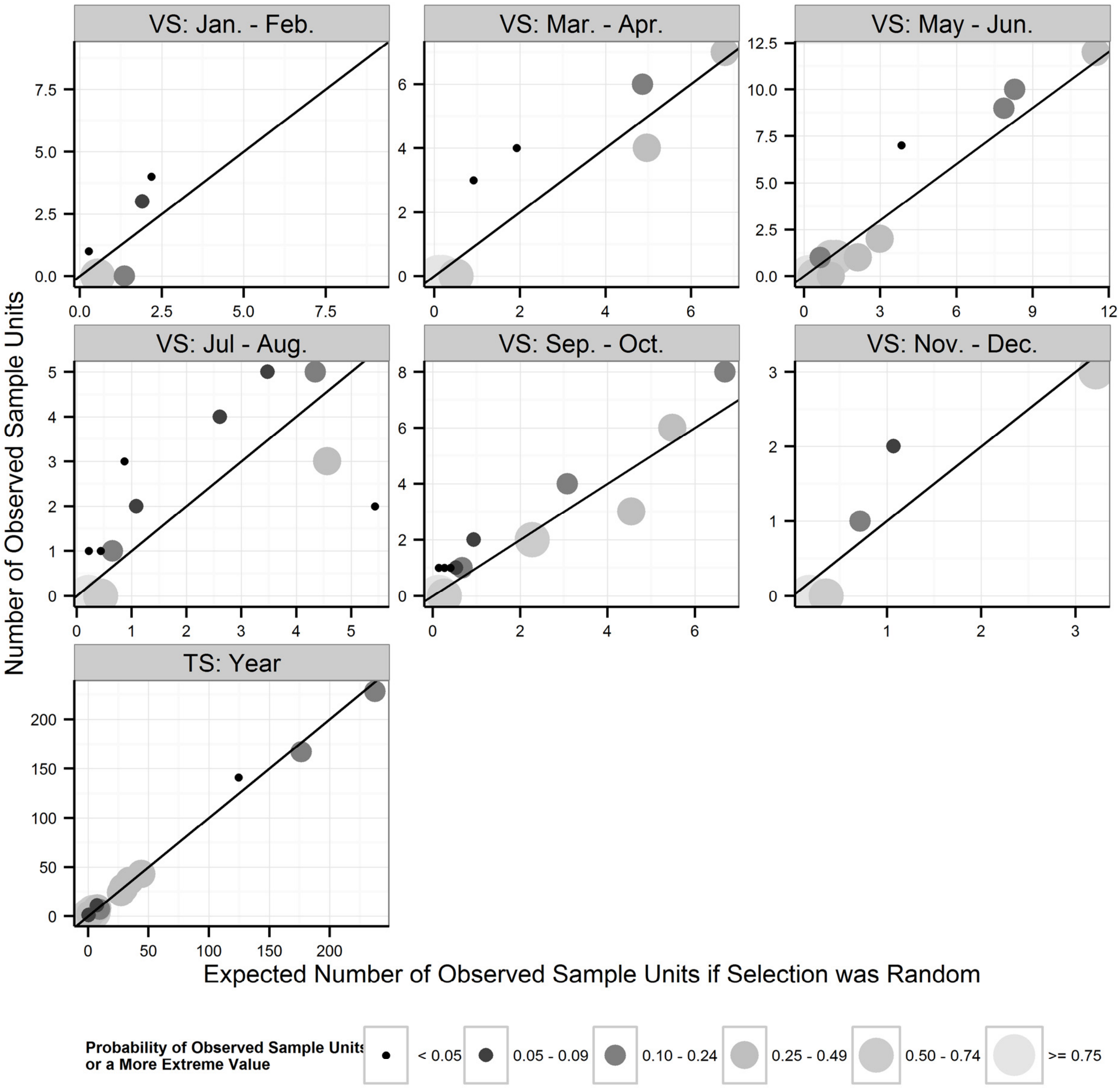

Figure 4. -- Comparison plots depicting the number of observed sample units (trips for trip selection = TS; vessel for vessel selection = VS) compared to the number of expected observed sample units from the hypergeometric distribution. Each point on a plot represents a NMFS Area. The size of the plot is proportional to the probability of the observed number of sample units or a more extreme outcome (more if above the solid 1:1 line, less if below it).

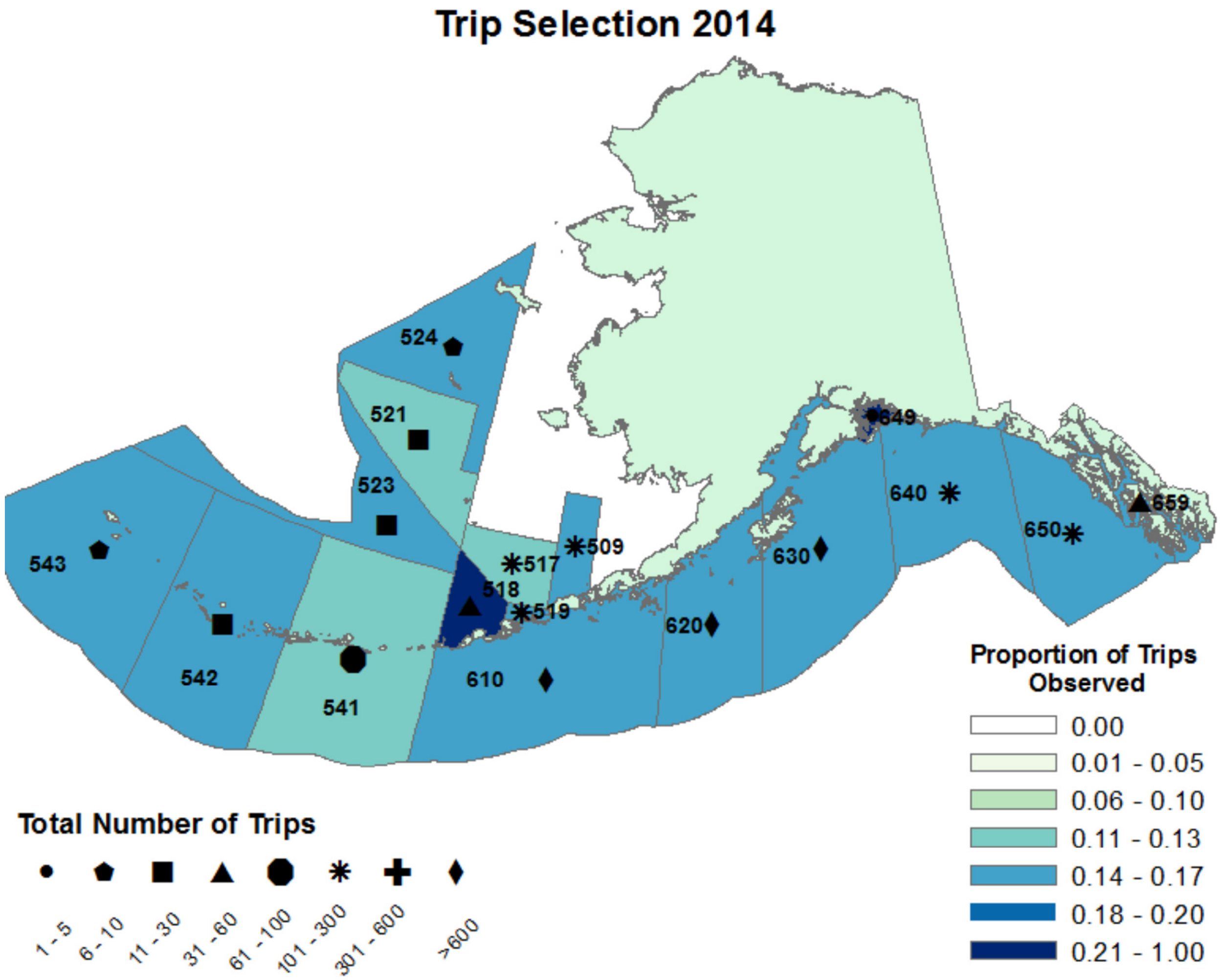

Figure 5. -- Proportion of trips observed in each NMFS Reporting Area in the trip selection strata. The color of the Reporting Area reflects the proportion of trips that were observed while the symbol indicates the total number of fishing trips that occurred in that area.

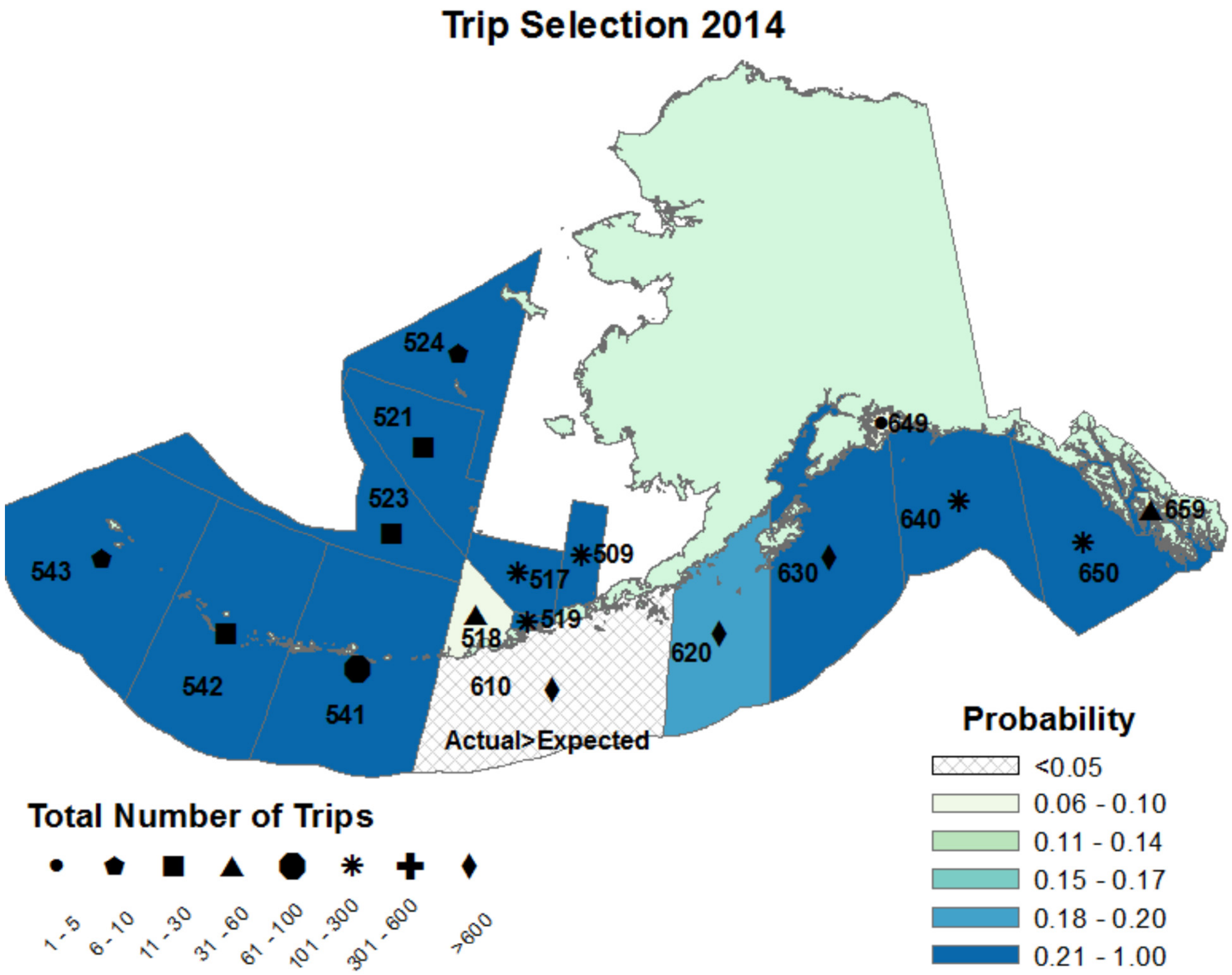

Figure 6. -- The probability of observing a number of trips in trip selection as far or farther from expected values (probability of observing a more extreme value). The symbol indicates the total number of fishing trips that occurred in that area.

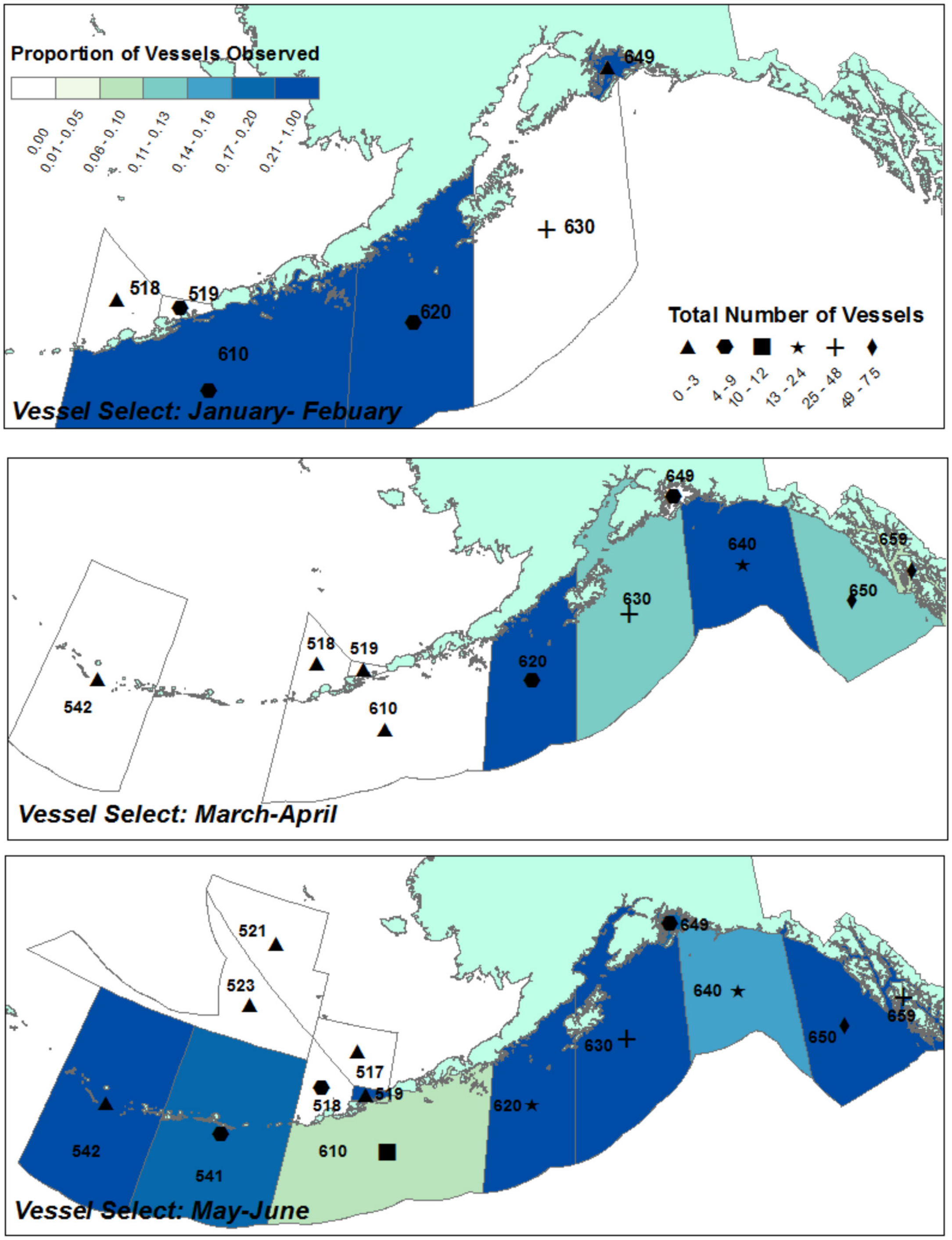

Figure 7. -- Proportion of vessels observed in each NMFS Reporting Area in the vessel selection strata during the first half of 2014. The color of the Reporting Area reflects the proportion of vessels that were observed while the symbol indicates the total number of fishing vessels that occurred in that area.

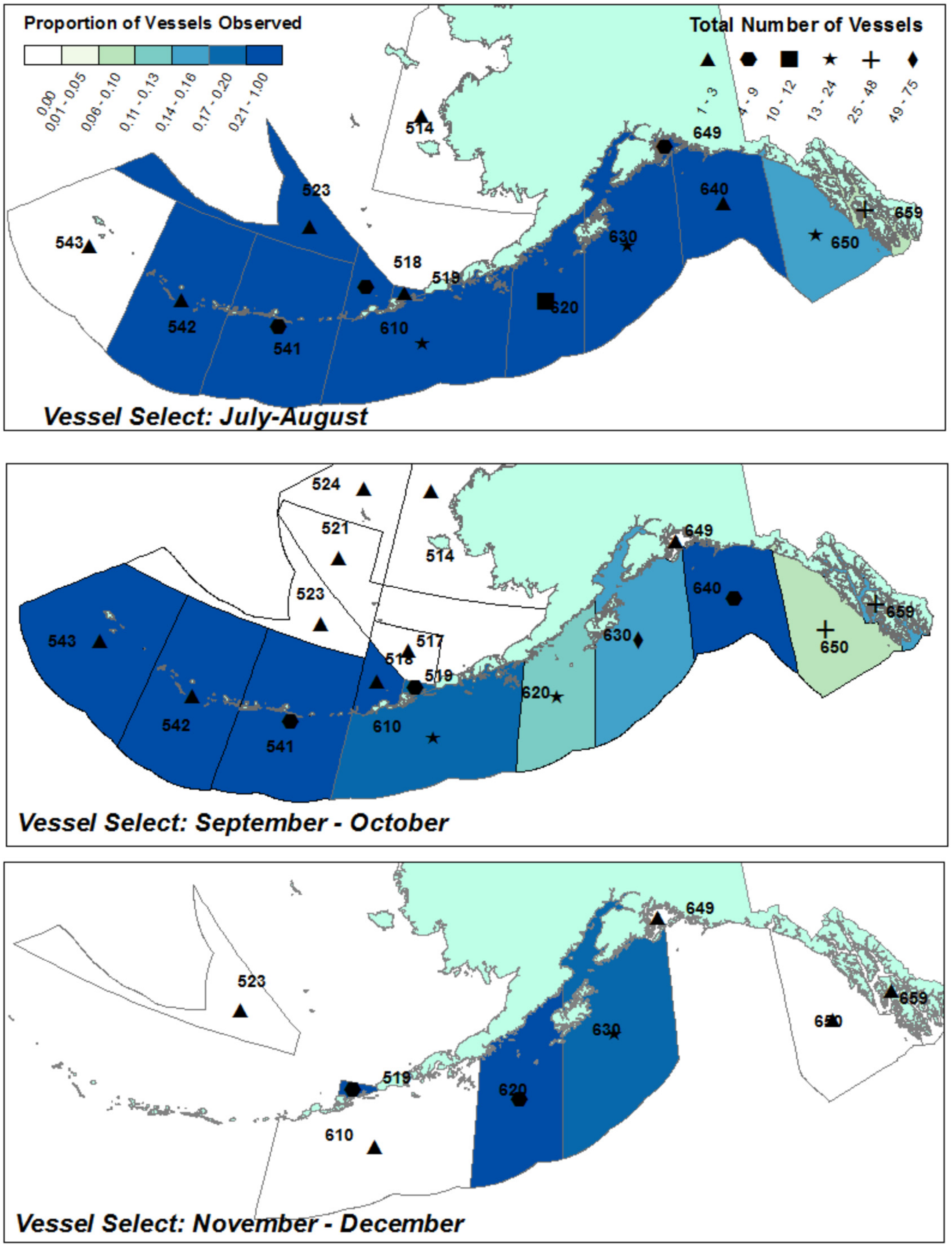

Figure 8. -- Proportion of vessels observed in each NMFS Reporting Area in the vessel selection strata during the second half of 2014. The color of the Reporting Area reflects the proportion of vessels that were observed while the symbol indicates the total number of fishing vessels that occurred in that area.

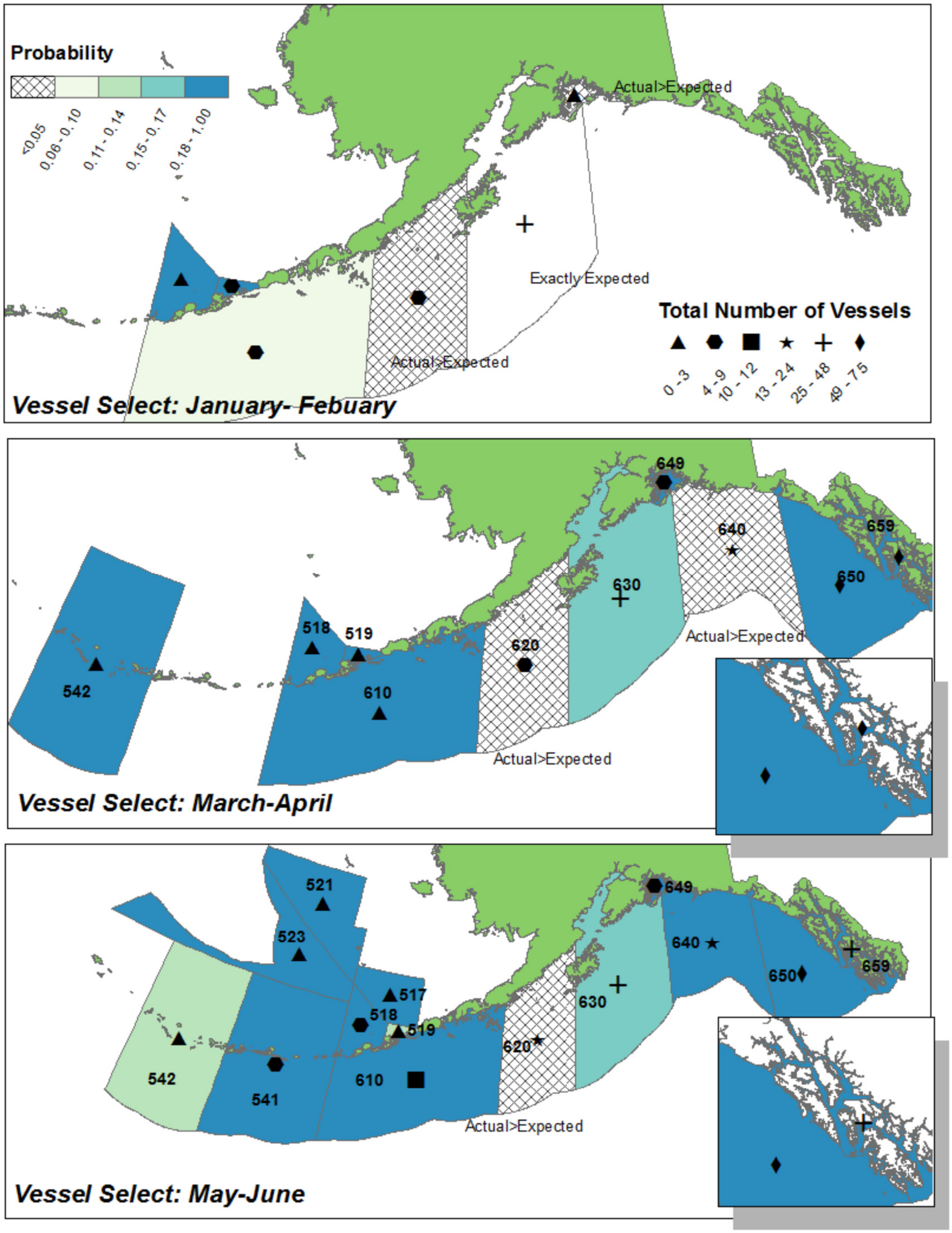

Figure 9. -- The probability of observing a number of trips in vessel selection as far or farther from expected values (probability of observing a more extreme value) during the first half of 2014. The symbol indicates the total number of fishing trips that occurred in that area.

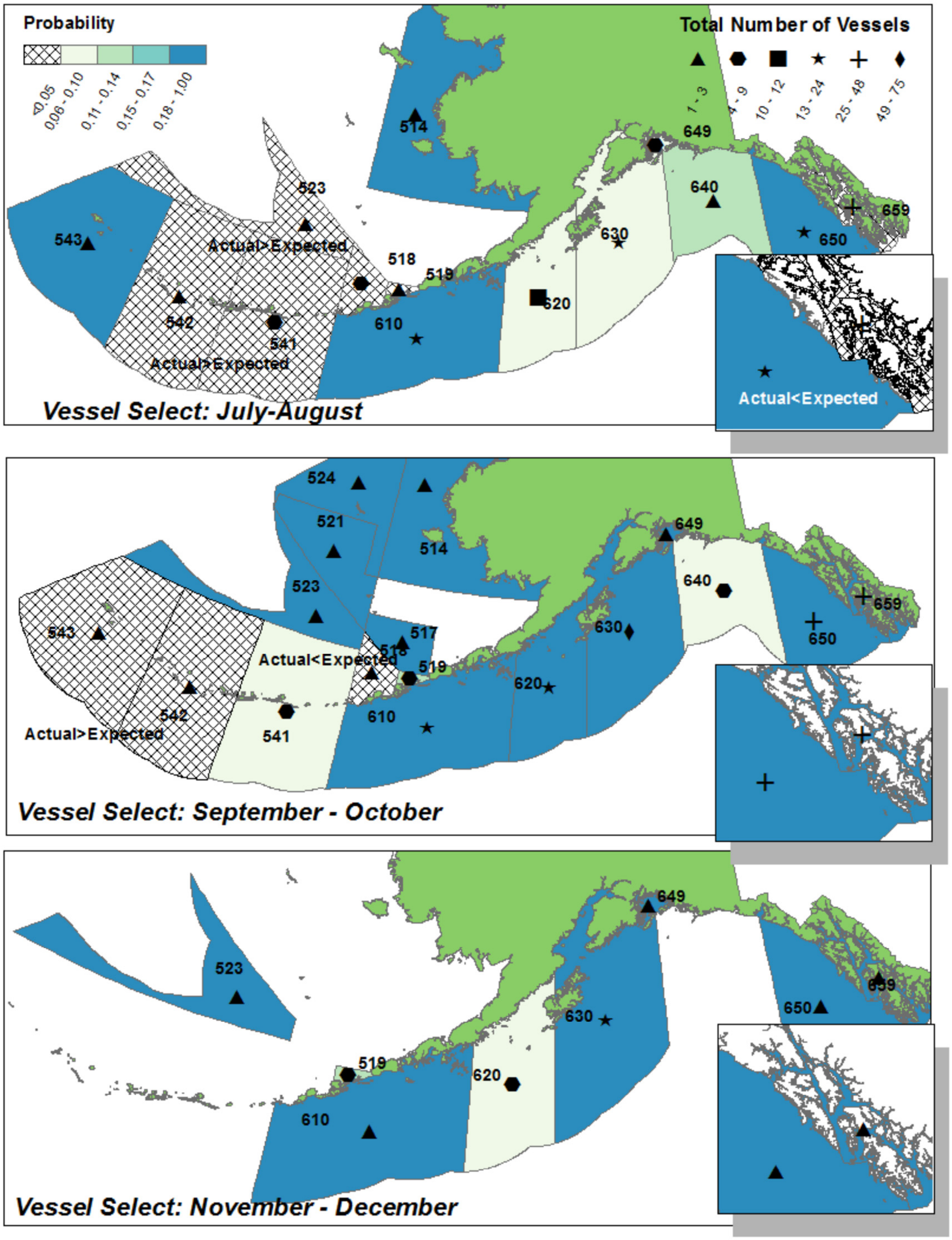

Figure 10. -- The probability of observing a number of trips in vessel selection as far or farther from expected values (probability of observing a more extreme value) during the second half of 2014. The symbol indicates the total number of fishing trips that occurred in that area.

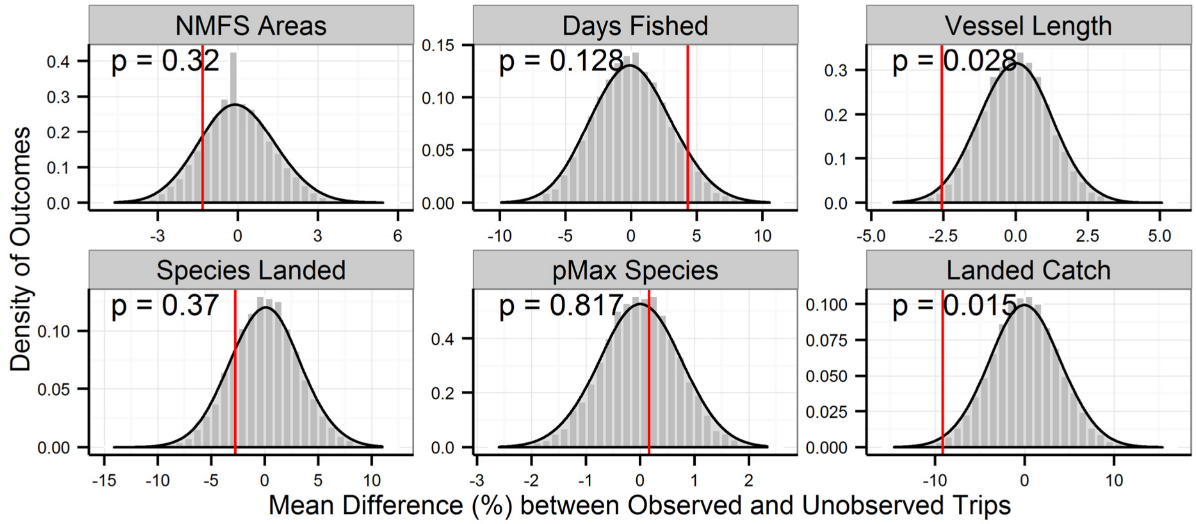

Figure 11. -- Results of permutation tests for each trip metric in trip selection. In each panel, the grey bars depict the distribution of differences between observed and unobserved trips where the assignment of observation status had been randomized (this represents the sampling distribution under the null hypothesis that observed and unobserved trips are the same). The vertical line denotes the actual difference between observed and unobserved trips. Values on the x-axis have been scaled to reflect the relative (%) differences in each metric. The corresponding p-value for each test is denoted in the upper left corner. Low p-values are reason to reject the null hypothesis and conclude that there is an observer effect. The finding that vessel length and landed catch are lower for observed trips is cause for further investigation into the potential drivers of this observer effect.

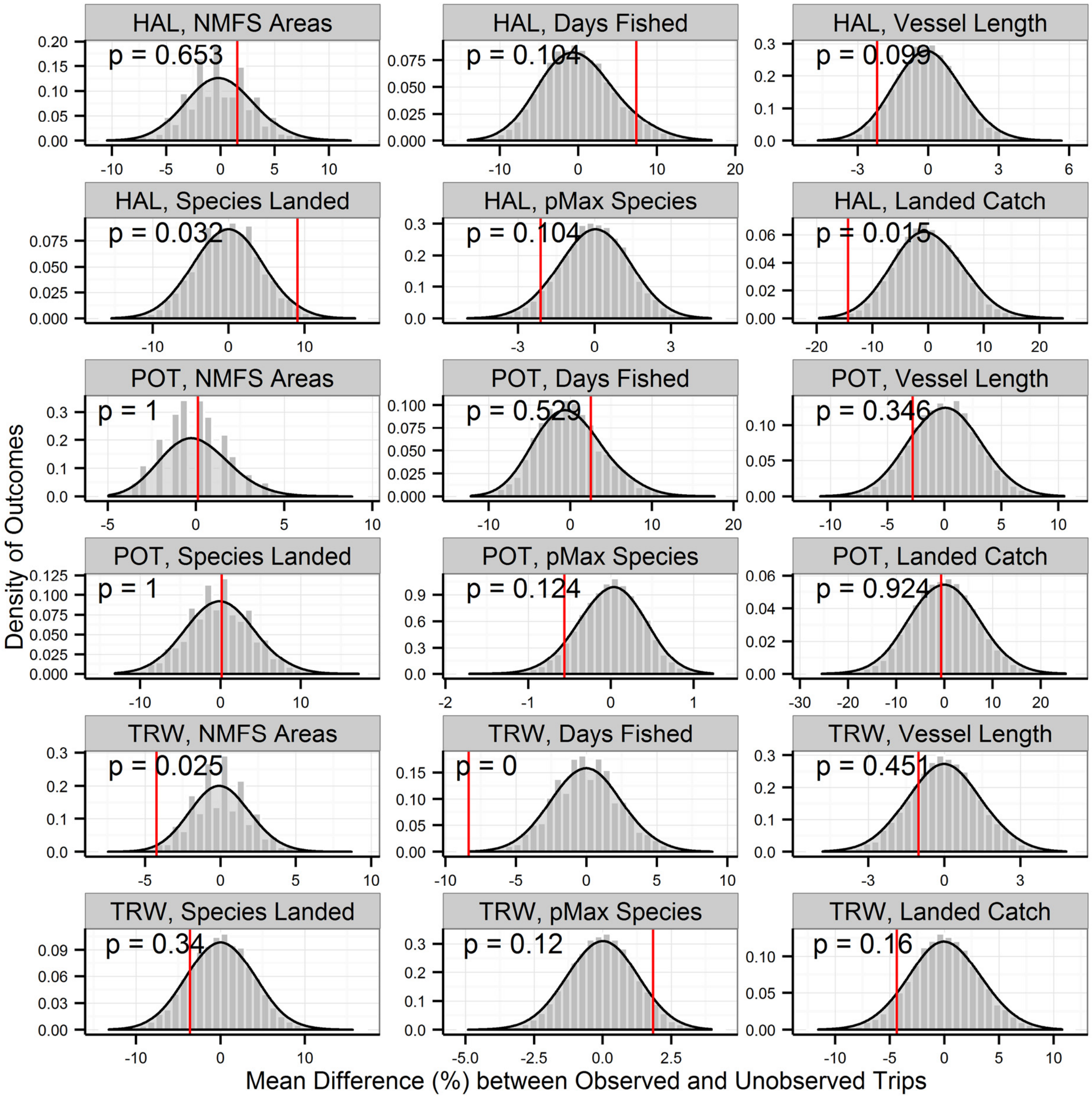

Figure 12. -- Results of permutation tests for each trip metric in trip selection separated by gear type. In each panel, the grey bars depict the distribution of differences between observed and unobserved trips where the assignment of observation status had been randomized (this represents the sampling distribution under the null hypothesis that observed and unobserved trips are the same). The vertical line denotes the actual difference between observed and unobserved trips. Values on the x-axis have been scaled to reflect the relative (%) differences in each metric. The corresponding p-value for each test is denoted in the upper left corner. Low p-values are reason to reject the null hypothesis and conclude that there is an observer effect. Evidence of an observer effect is present in hook and line and trawl gear.

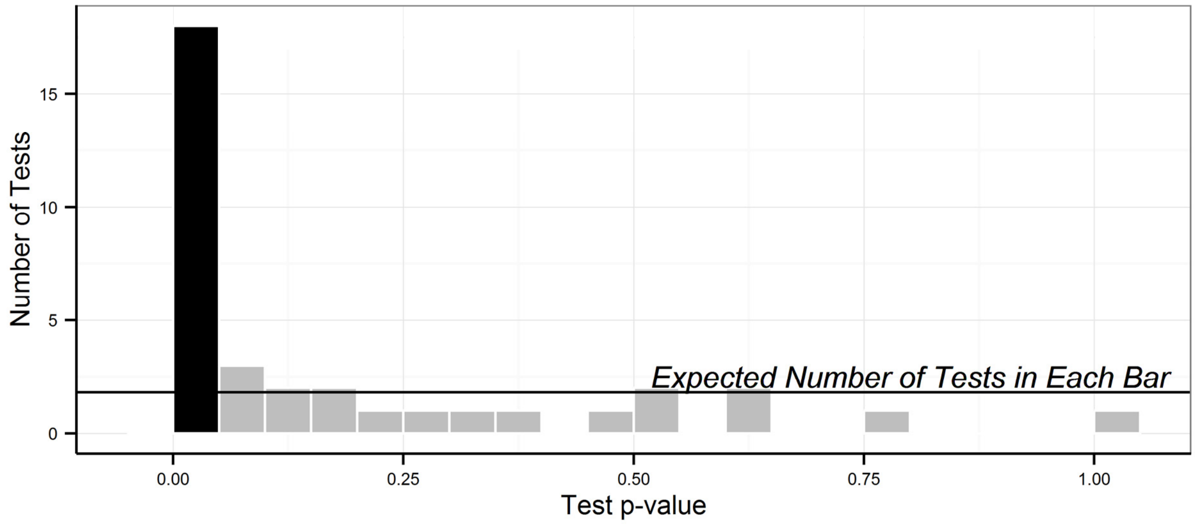

Figure 13. -- Histogram of the p-values from permutation tests on six trip metrics from within six time periods of vessel selection. Under the null hypothesis that observed and unobserved trips are the same, we would expect a distribution of p-values to roughly follow the horizontal solid line. The preponderance of low p-value test results denoted in black is reason to conclude that an observer effect was present in vessel selection.

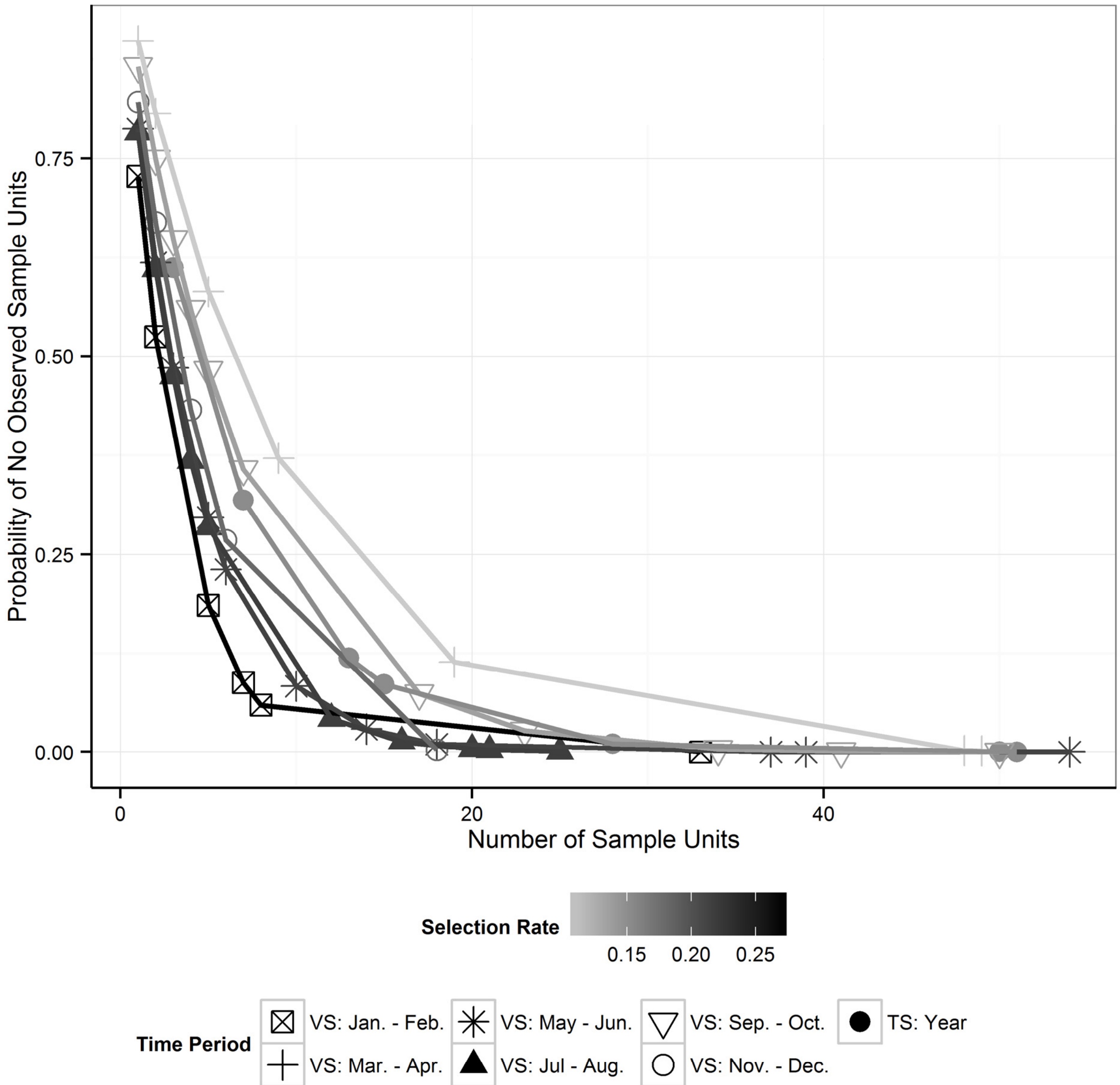

Figure 14. -- Probability of selecting a sample and observing no sample units (trips in trip selection and vessels in vessel selection) as a function of the number of sample units and selection rate that occurred in a NMFS Area, time period, and stratum. The x-axis has been truncated to increase resolution at smaller numbers of sampling units. The likelihood of having no observer data decreases with increasing total fishing effort and selection rate.

RECENT TECHNICAL MEMORANDUMS

Copies of this and other NOAA Technical Memorandums are available from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22167 (web site: *www.ntis.gov)*. Paper and electronic (.pdf) copies vary in price.

AFSC-

301 ALLEN, B. M., and R. P. ANGLISS. 2015. Alaska marine mammal stock assessments, 2014, 304 p. NTIS number pending.

300 HELKER, V. T., B. M. ALLEN, and L. A. JEMISON. 2015. Human-caused injury and mortality of NMFS-managed Alaska marine mammal stocks, 2009-2013, 94 p. NTIS number pending.

299 LEW, D. K., G. SAMPSON, A. HIMES-CORNELL, J. LEE, and B. GARBER-YONTS. 2015. Costs, earnings, and employment in the Alaska saltwater sport fishing charter sector, 2011-2013, 134 p. NTIS No. PB2015-104937.

298 PRESCOTT, M. M., and M. ZIMMERMANN. 2015. Smooth sheet bathymetry of Norton Sound, 23 p. NTIS No. PB2015-104936.

297 VON SZALAY, P. G. 2015. The impact of sample size reduction on the precision of biomass estimates in the Gulf of Alaska, 17 p. NTIS number pending.

296 STONE, R. L., D. STEVENSON, and S. BROOKE. 2015. Assessment of a pilot study to collect coral bycatch data from the Alaska commercial fishing fleet, 45 p. NTIS No. PB2015-104935.

295 DALY, B. J., C. E. ARMISTEAD, and R. J. FOY. 2015. The 2013 eastern Bering Sea continental shelf bottom trawl survey: Results for commercial crab species, 165 p. NTIS No. PB2015-104934.

294 FOWLER, C. W., and L. K. JOHNSON. 2015. Reality-based marine protected areas for the eastern Bering Sea, 109 p. NTIS No. PB2015-104933.

293 JOHNSON, S. W., A. D. NEFF, and M. R. LINDEBERG. 2015. A handy field guide to the nearshore marine fishes of Alaska, 211 p. NTIS No. PB2015-104932.

292 WHITTLE, J. A., S. C. VULSTEK, C. M. KONDZELA, and J. R. GUYON. 2015. Genetic stock composition analysis of chum salmon bycatch from the 2013 Bering Sea walleye pollock trawl fishery, 50 p. NTIS No. PB2015-104931.

291 GUYON, J. R., C. M. GUTHRIE III, A. R. MUNRO, J. JASPER, and W. D. TEMPLIN. 2015. Genetic stock composition analysis of the Chinook salmon bycatch in the Gulf of Alaska walleye pollock (*Gadus chalcogrammus*) trawl fisheries, 26 p. NTIS number pending.

290 GUTHRIE, C. M. III, HV. T. NGUYEN, and J. R. GUYON. 2015. Genetic stock composition analysis of the Chinook salmon bycatch from the 2013 Bering Sea walleye pollock (*Gadus chalcogrammus*) trawl fishery, 21 p. NTIS number pending.

289 GUYON, J. R., HV. T. NGUYEN, C. M. GUTHRIE III, J. BONNEY, K. MCGAULEY, K. HANSEN, and J. GAUVIN. 2015. Genetic stock composition analysis of Chinook salmon bycatch samples from the rockfish and arrowtooth flounder 2013 Gulf of Alaska trawl fisheries and the Gulf of Alaska salmon excluder device test, 19 p. NTIS number pending.

288 FAUNCE, C. H. 2015. Evolution of observer methods to obtain genetic material from Chinook salmon bycatch in the Alaska pollock fishery, 28 p. NTIS No. PB2015-103731.

287 ZIMMERMANN, M., and M. M. PRESCOTT. 2015. Smooth sheet bathymetry of the central Gulf of Alaska, 54p. NTIS No. PB2015-103730.

286 CAHALAN, J., J. GASPER, and J. MONDRAGON. 2014. Catch sampling and estimation in the federal groundfish fisheries off Alaska, 2005 edition, 46 p. NTIS No. PB2015102748.