NOAA Technical Memorandum NMFS-AFSC-322
doi:10.7289/V5/TM-AFSC-322

Deployment Performance Review of the 2015 North Pacific Groundfish and Halibut Observer Program

C. Faunce, J. Gasper, J. Cahalan, S. Lowe, S. Barbeaux, and R. Webster

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

June 2016

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, S. Barbeaux, and R. Webster. 2016. Deployment performance review of the 2015 North Pacific Groundfish and Halibut Observer Program. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-322, 54 p. doi:10.7289/V5/TM-AFSC-322.

Document available: http://www.afsc.noaa.gov/Publications/AFSC-TM/NOAA-TM-AFSC-322.pdf.

NOAA Technical Memorandum NMFS-AFSC-322
doi:10.7289/V5/TM-AFSC-322

Deployment Performance Review of the 2015 North Pacific Groundfish and Halibut Observer Program

C. Faunce[1], J. Gasper[2], J. Cahalan[3], S. Lowe[4], S. Barbeaux[4], and R. Webster[5]

[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 2016

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 following the 2015 Annual Deployment Plan (ADP). Responses to comments by the North Pacific Fishery Management Council's Science and Statistical Committee from the 2014 version of this report are also included.

In 2015, observers were deployed and collected data for 5,318 days - 3.6% fewer days than were budgeted for and predicted in the 2015 ADP. In total, 4,859 trips (39.1%) were observed on 498 vessels belonging to five strata (two trip-selection strata - a "*t*" stratum comprised of fixed gear catcher vessels 40-57.5 feet in length and a "*T*" stratum defined by larger catcher vessels, an EM voluntary stratum, a full-coverage stratum, and a zero-coverage stratum). All partial coverage strata had coverage rates within expected ranges with the exception of the new EM Voluntary (EM) stratum that had two observed trips although the selection rate was zero.

Overall selection rates in the Observer Declare and Deploy System (ODDS) were as expected for each trip-selection stratum. The proportion of trips selected computed using the initial observer-status (selected or not selected for coverage) of logged trips differed from the proportion selected using the final observer-status after trips are cancelled or dates have been changed. Trips that were selected to be observed were cancelled at a higher rate than unselected trips. Despite this behavior, a temporal bias was not evident in the trips belonging to either trip-selection strata. Spatial bias was evident in the *t* stratum with coverage rates lower than expected in four NMFS areas. There was some evidence of an observer effect in the partial coverage category. In particular, observed trips were 8 – 14% shorter in both trip-selection strata and landed catch was 1% less diverse in the *T* stratum than unobserved trips.

Observers monitor deliveries of trawl-caught walleye pollock (*Gadus chalcogrammus*) at shoreside processing plants and collect genetic tissue samples from salmon bycatch. Although all pollock deliveries were monitored in the full coverage pollock fishery, coverage rates were below expected in the partial coverage category due to tendering activity at King Cove. Tendered trips were found to have different trip characteristics than non-tendered trips. There was some evidence of an observer effect within partial coverage tendered trips. Tendered trips landed 33-101% fewer fish, by weight, when observed and landed catch with 25% less species in the *t* stratum than unobserved tendered trips. In the *T* stratum, observed tendered trips were also 51% shorter than unobserved tendered trips. Similarly, for non-tendered trips, observed trips were 3-5% shorter and fishing occurred in 3% fewer areas in the *t* stratum and were 1% less diverse in the *T* stratum.

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. 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 highlights 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 all fishing activities.

Based on these combined results, we support the continued implementation and expansion of *t*Landings to document unobserved tendered deliveries, and make four recommendations to improve the 2017 ADP. The recommendations are: 1) tendered vessels should be evaluated as separate strata, 2) the treatment of partial coverage catcher processors should be evaluated as separate strata, 3) sampling rates should be high enough that the probability of achieving three observed trips in each NMFS Area is low, and 4) NMFS should work with its partial coverage contractor to explore the possibility of eliminating trip cancellations in ODDS.

CONTENTS

INTRODUCTION

Background of the North Pacific Groundfish and Halibut Observer Program

Fisheries observers collect independent information that is used to determine the effects of fishing on 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 Bering Sea and Aleutian Islands (BSAI) Management Area and Amendment 76 to the Fishery Management Plan for Groundfish of the Gulf of Alaska (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. The ‘restructured’ North Pacific Groundfish and Halibut Observer Program (hereafter termed ‘Observer Program’) began in 2013 with the randomization of deployments among trips and vessels. Since 2013, the Observer Program has employed a hierarchical

sampling design with randomization at all levels (trips > haul > species composition > length, age, sex, maturity and tissues for genetic analysis).

THE ANNUAL DEPLOYMENT PLAN AND REVIEW

Analysis and evaluation of the data collected by observers is an ongoing process. NMFS considers Council input in making decisions as to the amount of coverage (i.e., selection probabilities that are assigned to each partial-coverage category). These decisions are based on available funding, the cost of observer coverage, and anticipated effort. The restructure of the Observer Program established new annual reporting processes. 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 Annual 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 Annual Report will inform the Council and the public of how well various aspects of the program are working, and consequently lead to recommendations for improvement. The NMFS also releases a draft and final Annual Deployment Plan (ADP). The ADP defines deployment strata and establishes selection rates given available budgets and anticipated fishing effort. A draft ADP is released by 1 September of each year to allow review by the Council's Groundfish Plan Teams, as well as the 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 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 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 BSAI and the 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).

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 percentage-based (currently 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 first included as part of the supporting analysis in the Public Review Draft of Observer Amendments 86/76 (NPFMC 2011) and adjusted by subsequent Council actions.

The full-coverage category in 2015 included the following (NMFS 2014a):

- 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 2015 partial coverage category included the following (NMFS 2014a):

- 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 2015, the partial coverage category was divided into two deployment pools: trip-selection and no selection. Trip-selection participants were 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 logged trip was instantly assigned a random number, and if the number is below a pre-determined threshold, the trip was selected for observer coverage. Users "close" (indicating to NMFS that the trip has been completed) their 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.

In 2015, the trip-selection pool was further divided into two sampling strata (NMFS 2014a). Catcher vessels that fished hook-and-line or pot gear on vessels greater than or equal to 40 feet LOA, but less than 57.5 feet LOA were in the small vessel trip-selection stratum (t). This stratum is identical in composition to the 2014 ADP 'vessel-selection' stratum (NMFS 2013). The large-vessel trip-selection stratum (T) was constituted of three classes of vessels: 1) all catcher vessels fishing trawl gear, 2) catcher vessels fishing hook-and-line or pot gear that are also greater than or equal to 57.5 feet LOA, and 3)

catcher-processor vessels exempted from full coverage requirements (50 CFR 679.51(a)(2)(iv)). This stratum is identical in composition to the 2014 ADP 'trip-selection' stratum (NMFS 2013).

In 2015 the no-selection pool was expanded to include an additional sampling strata (NMFS 2014a). As in 2013-2014, catcher vessels less than 40 feet LOA, or vessels fishing with jig gear (which includes handline, jig, troll, and dinglebar troll gear) constituted the 'zero selection' stratum and were not subject to observer coverage. New in 2015, vessels that volunteered, and were selected to participate in Electronic Monitoring Cooperative Research constituted the 'EM voluntary' stratum and were also not subject to observer coverage. The rules governing the composition of, participation in, sampling rates and methods of this stratum are dictated by the Council's EM Workgroup (See EM Workgroup at http://www.npfmc.org/observer-program/). The EM Workgroup works independently from the ADP process. Rules governing the EM stratum were defined prior to the ADP in 2015.

DEPLOYMENT PERFORMANCE REVIEW

The following sections contain the OSC review of the deployment of observers in 2015 relative to the intended sampling plan and goals of the 2015 ADP (NMFS 2014a). 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 2017 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 population and ecosystem modeling efforts in addition to salmon bycatch stock-of-origin determinations. 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.

Observer Deployment Performance Metrics

Performance metrics have been developed to assess whether the trip-selection process (through the implementation of the 2015 ADP) provides a representative sample of the catch in the North Pacific in 2015. These metrics reflect four mechanisms that can impact the quality of the data: sample frame discrepancies, non-response, trip differences, and sample size. In cases where the vessel is the sampling unit, sample frame discrepancies (under- and over-coverage of the sample frame) were 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.

The performance metrics used in this evaluation are as follows:

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 (vessel-selection only), selection biases, and issues with sample unit definitions (e.g., tendered trips). Specifically, this section assesses the following:
    a. Sample rates and number of samples relative to intended values.
    b. (Vessel-selection strata only) 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. (Vessel-selection strata only) 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 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. These data are used to identify departures from anticipated sampling rates.

   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 (t)
- 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 (pMax, an inverse measure of species diversity – in increase in pMax indicates a decline in 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.

CHANGES TO THIS REPORT FROM LAST YEAR

Deployment Strata

- Vessels could volunteer to carry Electronic Monitoring (EM) in 2015. These vessels were placed into the zero-selection category during 2015 for as long as they continued to volunteer.

- The vessel-selection method of deployment that uses vessel:time periods as a sample unit to deploy observers within the partial coverage category was suspended in 2014 and was not in use during 2015 (Note however, that this method was reintroduced for EM Voluntary vessels starting in 2016).

Methodological changes

- The methodology used to define full-coverage trips changed between past Annual Reports and this version. In the past, full-coverage trips were defined in the data from vessel:week combinations. In this report, the field TRIP_SEQ from observer data were used to define trips in this coverage category, since all trips are observed. Trip counts and durations in this coverage category now reflect actual fishing trips rather than units used in quota monitoring. Hence comparisons between the number of trips in this coverage category from this report are not comparable to past reports.

- In this report, considerable effort was made to make use of VMS using geo-fencing algorithms to define trip durations for unobserved partial coverage trips that delivered to a tender vessel.

- Tables with the number of trips for each condition used in permutation tests (i.e., sample size) are now included.

- A permutation test to examine whether observed trips were different from unobserved trips for all partial coverage activity (tendered and non-tendered trips combined) by stratum is now included.

- Permutation tests to examine whether partial coverage trips delivering to a tender are different from those that did not deliver to a tender were performed using 2015 and 2016 stratum definitions (i.e., gear type; NMFS 2015a). In 2014 these tests were conducted on all partial coverage strata combined.

- Permutation tests to examine whether observed trips were different from unobserved trips within tendered and non-tendered trips are now conducted for each partial coverage stratum. In 2014 tests for an observer effect within tendered trips were conducted on all partial coverage strata combined, and tests for an observer effect within non-tendered trips were conducted by strata:gear combinations.

EVALUATION OF OBSERVER DEPLOYMENT IN 2015

The deployment of observers into the 2015 Federal fisheries in Alaska needs to be evaluated at the level of the deployment stratum because each stratum is defined by a different sampling rate. Following the June 5th 2014 Council Motion that the *t* stratum (2014 vessel-selection stratum) have a lower selection rate than the *T* stratum (2014 trip-selection stratum), a suite of possible combinations of rates for these two stratum were evaluated in the 2015 ADP. Possible rates were determined by iteratively increasing the selection rate for the *t* stratum from 12% to 19% and determining the corresponding rate for *T* that maximized the expenditure of available funds. This translates to a set of selection rates that was anticipated to result in about a 1 in 2 chance of going over budget.

Following the 2015 ADP, the ODDS was programmed to randomly select observer coverage for 12% of logged trips in the *t* stratum and 24% of logged trips in the *T* stratum. These rates were the expected rates of observer coverage in these strata.

Evaluating Effort Predictions

Each year the NMFS sets an annual budget in terms of observer days. Therefore how close anticipated observed effort is to actual invoiced effort in each ADP is a function of how well the NMFS predicts effort and how well the NMFS achieves its sampling rate. The observer day budget for 2015 was set at 5,518 days for the 2015 ADP (NMFS 2014a). Based on simulations of 2013 fishing data made a year in advance of deployment, the FMA predicted it would observe 5,517 fishing days at the end of 2015. In 2015, the FMA paid for 5,318 observer days-- 936.5 days were paid in the *t* stratum and 4,381.5 days were paid in the *T* stratum. This total observer day budget was 3.6% lower than predicted (Fig. 1). For comparison, in 2014 the expended budget was 7.4% less than predicted in the 2014 ADP (NMFS 2015b).

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. Industry 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 or cancelled prior to fishing. Trips can be cancelled by the user or the observer provider. A NMFS waiver is issued in cases where the observer provider cannot provide an observer for a selected trip in time. Any remaining trips that have not been closed at the end of the calendar year are automatically cancelled by the ODDS to prevent 2015 ODDS trips from affecting the deployment rates set for the 2016 ADP. The number of trips logged in the ODDS in 2015 and their dispositions is summarized in Table 1. Of 7,046 trips logged, a total of 931 trips were cancelled (13.2%-- 552, or 7.8% by ODDS) and 12 trips were waived (0.2%). The cancellation rate (calculated from the number of trips cancelled by the user divided by the number of trips not cancelled by the ODDS) ranged from 2.8% to 3.8% for non-selected trips, and 13% to 23% for selected trips from the *t* and *T* strata, respectively.

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 changed from the time they are entered 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 11.6% for the *t* stratum and 23.8% for the *T* stratum. These values were well within the range of values expected from a binomial distribution (exact binomial test p-values = 0.595 and 0.770 for *t* and *T,* respectively). This means that the ODDS was selecting trips according to the programmed rate. The final selection rate after trips were closed and cancelled was 12.6% for the *t* stratum and 25.4% for the *T* stratum. The fact that the final selection rates were greater than the initial selection rates 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 were evident throughout 2015. Whereas the original selection rate approached the programmed rate within partial coverage strata after only a month, the final selection rate lagged that of the initial rate and did not approach the programmed selection rate until several months later (Fig. 2). After several months, the final selection rate eclipsed that of the initial selection rate and remained the higher rate through the remainder of the year. These patterns are consistent with the hypothesis that trips selected for coverage are being delayed, and cancellation of selected trips results in a greater number of selected trips later in the year as the result of the inherit process. It is important to remember that ODDS only provides the *expectation* as to what levels of observer coverage levels should be resulting from actual fishing events. While the 2015 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.

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 Alaska Regional Office), the Observer Program database NORPAC (managed by the AFSC), and *e*Landings (under joint management by Alaska Department of Fish and Game -- ADF&G and the 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 undertaken by vessels designated as belonging to full or partial coverage 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 the *t* and *T* strata of the trip-selection pool, and the zero-selection pool. Although the EM voluntary vessels are contained within the zero-selection pool, they are separated in this Annual Report due to the pre-implementation status of EM in the North Pacific.

Rate evaluations are based on trips for the year (in 2016, the EM Voluntary stratum will contain four time periods). Evaluations for the full coverage category and the no selection pool are straightforward - either the coverage achieved was equal to 100% or 0%, respectively, or it was not. For partial coverage 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"). 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. The expected

coverage rate for partial coverage category strata in 2015 was the rate programmed into ODDS for each stratum.

The 2015 Observer Program had 5 different deployment strata to be evaluated. With one possible exception, the program met expected rates of coverage for all of these strata (Table 3). Observer coverage was higher than expected from within the EM Voluntary stratum. This was because within this stratum one vessel agreed to simultaneously carry EM and an observer on two trips so that resulting data from the two methods could be compared. Among all fishing in Federal fisheries of Alaska, 4,859 trips (39.1%) and 498 vessels (42.1%) were observed.

Coverage rates for dockside monitoring

Observers were assigned to monitor deliveries of walleye pollock (*Gadus chalcogrammus*). The objective of this monitoring was 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 2015a). The sampling design used for this objective in 2015 remained unchanged from that used since 2014; all deliveries of walleye pollock that are observed at sea were also observed dockside. While all Bering Sea pollock trips and deliveries are observed, this is not the case in the Gulf of Alaska (NMFS 2015c).

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 to analyzing the data 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 NMFS Office of Law Enforcement (OLE)[1]. 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.

Given this design, the level of dockside monitoring of walleye pollock should be 100% in the full coverage category, and within acceptable tolerance of the deployment rate of 24% in the partial coverage category (since all trawl catcher vessels in partial coverage participating in this fishery are within the *T* stratum). 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 always possible from the partial coverage fleet 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 (it is also possible the vessels' entire season activity is 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. Very few pollock deliveries were unobserved in full coverage (1%). In contrast, the chance that the coverage rate in partial coverage resulted from a random deployment at the expected rate was extremely small (exact binomial test p-value = 0.016; Table 4). However, when deliveries of pollock from tender trips were removed, this likelihood

[1] Very few differences were found between the two methods in 2014 (NMFS 2015b).

was dramatically increased (p-value = 0.796). The majority of pollock deliveries in the port of King Cove from the partial coverage category were tender deliveries and very few of these were observed (Table 4).

SAMPLE PROPERTIES

Temporal Patterns in Trip Selection

The cumulative number of fishing trips in each stratum was multiplied by their selection rate 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"). Under the assumption that there is no temporal bias in observer coverage, 5% or less of values should fall outside of upper and lower expected bounds. The number of observed trips achieved was outside of their expected values on 2 days (0.6%) very early in the year and within only one stratum (Fig. 3). For comparison, the number of observed trips achieved was outside of their expected values on 15.3% of the year (NMFS 2015b). Tests that the observed rate at the end of the year derived from a binomial distribution sampled at the expected selection rates for each stratum were 0.273 for *t* and 0.338 for *T* (Fig. 3). Based on these combined results, no evidence of temporal bias was found in 2015.

Spatial Patterns in Trip Selection

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., *t* strata 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. This evaluation does not test whether the resulting coverage rate in a NMFS Area for a stratum is equal to the stratum selection rate, but instead tests whether the resulting coverage rate in a NMFS Area for a stratum is unexpected compared to the stratum-wide actual realized observation rate.

Using this method, the expected number of trips with the observed number of trips in each NMFS Reporting Area and stratum combination were compared (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.

The t *stratum*

Given that there were 18 NMFS Areas fished in *t*, we would expect there to be $0.05 \times 18 = 1$ small data points for this stratum. Instead, there were five small data points for this stratum. Coverage rates among NMFS Areas for this stratum ranged from 0% to 50% (median = 10.8). The likelihood of this amount of coverage in this stratum is depicted in Figure 5. The probability of these coverage rates in each NMFS area are depicted in Figure 6.

Although there were no observed trips in the *t* stratum in NMFS Areas 542 and 543 (Western and Central Aleutian Islands), Areas 513 and 514 (Western Alaska) and Area 518 (Bogoslof Islands), this outcome was not unexpected given the low amount of fishing effort in all of these areas but Area 518. Low fishing effort can result in more than expected observer coverage as well, for example NMFS Area 524 in the Bering Sea. The fact that there were only three observed trips in the Aleutian Islands FMP (Areas 541:543) and no observed trips in Western Alaska (Areas 513:514) is a consequence of a low selection rate and low fishing effort in these stratum:location combinations. While not entirely unexpected, these results mean that there was little information on at-sea discards or biological tissues

collected to support in-season quota management or stock assessment from trips in these areas undertaken on vessels within the *t* stratum.

Within the *t* stratum there were four NMFS Areas that had less than expected observer coverage and did *not* have low fishing effort. These results are more powerful than those already discussed associated with low fishing effort. In the Bering Sea, NMFS Area 518 was also under-represented in Observer Coverage during September - October vessel-selection period of 2014 (NMFS 2015b). This NMFS Area is exhibiting an annual pattern of lower observer coverage than expected. In addition, Under-sampling of trips was also evident in the Western and Central Gulf of Alaska (NMFS Areas 610-620), while higher than expected coverage resulted in the Eastern Central Gulf Area (630).

From these results, it is concluded that spatial bias occurred in the partial coverage *t* stratum during 2015.

The T *stratum*

Given that there were 19 NMFS Areas fished in *T*, we would expect there to be $0.05 \times 19 = 1$ small data points for this stratum. Instead, there were two small data points for this stratum. Coverage rates among NMFS Areas in this stratum ranged from 0% to 40% (median = 24.1. The likelihood of this amount of coverage in this stratum is depicted in Figure 7. The probability of these coverage rates in each NMFS area is depicted in Figure 8.

Although no trips in this stratum were observed within NMFS Areas 649 (Prince William Sound) and the Western Aleutians (Area 543), this outcome was not surprising given that fewer than five trips occurred within these areas.

In general, there was a trend towards outcomes with lower probabilities in Western Alaska in the Bering Sea and Southeast Alaska in the Gulf of Alaska. However, only NMFS Areas 519 in the Bering Sea and Area 542 in the Aleutian Islands had substantially less observer coverage than expected from within the *T* stratum. There were no consistent patterns between 2014 and 2015.

Taken together, there is no clear evidence of spatial bias within the *T* stratum.

Trip Metrics

This section is focused on answering four questions related to the deployment of observers: 1) are observed trips identical to unobserved trips, 2) are tendered trips identical to non-tendered trips, 3) are observed tendered trips identical to unobserved tendered trips, and 4) are observed non-tender trips identical to unobserved non-tender trips.

Permutation tests (a.k.a., randomization tests) were used to answer each question. Each test evaluates 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). Difference values in all permutation tests were calculated by subtracting the mean metric value for the "No" condition from the mean metric value for the "Yes" condition. For example, the difference between vessel lengths in a permutation test for a tendering effect would be the mean value for non-tendered trips subtracted from the mean value for tendered trips. By randomizing group assignments, the combined distribution of randomized differences represents the sampling distribution under the null hypothesis that the two groups are equal. In this report 1,000 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) No. 1 expressed as a percentage of the mean value of the metric being tested; 4) No. 2 expressed as a percentage of the mean value of the metric being tests; and 5) the p-value of the test, only No.’s 1, 3 and 5 are presented in relevant tables.

Six trip metrics were examined in each permutation test. These metrics include: 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). The metric vessel length is used to help interpret the results from landed weight of catch, since fishing power positively correlated to vessel length. Specifically, 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. 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.

Since there are six metrics within each permutation test, and each is evaluated to be unusual if the p-value is < 0.05, we would expect by random chance to have $0.05 \times 6 = 0.3$ tests to have low p-values.

Are observed trips identical to unobserved 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 all partial coverage trips. Sample sizes for this test are presented in Table 5.

Of the six metrics compared in the *t* stratum, 1 had low p-values. Observed trips in this stratum were 13.6% shorter in duration than unobserved trips (Table 6).

Of the six metrics compared in the *T* stratum, 2 had low p-values. Observed trips in this stratum were 8.4% shorter in duration and landed catch that was 1% less diverse than unobserved trips (Table 6).

A visual depiction of individual results of this permutation test is given in Figure 9 for illustration purposes. In both strata, observed trips were shorter than unobserved trips.

Taken together, there was some evidence of an observer effect within the *t* stratum and evidence of an observer effect within the *T* stratum.

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. Two separate evaluations were conducted. In the first, the 2015 stratum definitions were used. In the second, the 2016 stratum definitions were used.

Using 2015 stratum definitions

Sample sizes for this test are presented in Table 7.

Of the six metrics compared in the *t* stratum, four had low p-values. Trips in this stratum that delivered to tenders were 9.3% longer in length, landed catch with 20.1% fewer species and was 8.8% less diverse, and weighed 389.9% more than trips that did not deliver to a tender (Table 8).

Of the six metrics compared in the *T* stratum, six had low p-values. Trips in this stratum that delivered to tenders occurred in 8.3% fewer areas, lasted 22.2% longer, occurred on vessels 17.4% shorter in length, landed catch that contained 16.3% fewer species and was 6% less diverse, and weighed 27.6% more than trips that did not deliver to a tender (Table 8).

Tendered and non-tendered trips were not the same in 2015.

Using 2016 stratum definitions

Trips from 2015 were re-coded according to the stratum definitions used in the 2016 ADP. These new codes denote three gear types: Hook and Line (*HAL*), Pot (*POT*), and Trawl (*TRW*). Sample sizes for this test are presented in Table 9.

Of the six metrics compared in the *HAL* 2016 stratum definition, one had low p-values. Trips in this stratum that delivered to tenders landed 149.4% more catch than trips that did not deliver to a tender (Table 10).

Of the six metrics compared in the *POT* 2016 stratum, four had low p-values. Trips in this stratum that delivered to tenders lasted 37.7% longer, occurred on vessels 9.5% shorter in length, landed catch that contained 22.8% more species and weighed 43.9% more than trips that did not deliver to a tender (Table 10).

Of the six metrics compared in the *TRW* 2016 stratum, five had low p-values. Trips in this stratum that delivered to tenders fished in 9.2% fewer areas, lasted 52.6% longer, occurred on vessels 30% shorter in length, landed catch that contained 12.9% less species and was 5.1% less diverse than trips that did not deliver to a tender (Table 10).

Taken together, a tendering effect was evident within all gear types during 2015. In *HAL*, only one metric had a low p-value, however this could be due to a low number of tendered trips (n) in this gear type (n = 12, N = 3,145). Nonetheless, the magnitude of the effect size (OD%) for landed catch cannot be ignored.

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. Sample sizes for this test are presented in Table 11.

Of the six metrics compared in the *t* stratum, one had low p-values. Observed trips in this stratum that delivered to tenders landed catch with 24.7% fewer species than unobserved trips that delivered to tenders. We also note that the magnitude of the difference in landed catch (101% less on observed trips) was large, and therefore, despite the fact that the associated p-value was greater than 0.05 (Table 12), such a difference should not be ignored.

Of the six metrics compared in the *T* stratum, one had low p-values. Observed trips in this stratum that delivered to tenders lasted 50.8% shorter than unobserved trips that delivered to tenders. We also note that the magnitude of the difference in landed catch (33% less on observed trips) was large, and therefore,

despite the fact that the associated p-value was greater than 0.05 (Table 12), such a difference should not be ignored.

From the above results, we conclude that there is some evidence of an observer effect within trips that delivered to tenders in 2015.

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. Sample sizes for this test are presented in Table 13.

Of the six metrics compared in the *t* stratum, two had low p-values. Observed non-tendered trips in this stratum fished in 3.4% fewer NMFS Areas and for 13.1% fewer days than unobserved non-tendered trips (Table 14).

Of the six metrics compared in the *T* stratum, two had low p-values. Observed trips in this stratum that did not deliver to tenders lasted 5.2% shorter and was 1.2 less diverse than unobserved trips that did not deliver to tenders (Table 14).

The fact that both strata fished for shorter durations but had similar catches is evidence of an observer effect within non-tendered trips in 2015.

Gear, tender, and observed status combinations

One of the first analyses presented in the 2013 Annual Report was a comparison of trip durations for combinations of observed and tendered status by stratum (NMFS 2014b). The rationale for this plot and focus on this metric was because of the concern that tendered trips were longer than non-tendered trips and therefore were being avoided for observer coverage. Frequency distributions showed that tendered trips had a long right tail compared to non-tendered trips, and that there were few observed trips in that long right tail (NMFS 2014b; Fig. 14). The OSC concluded that there were no major differences between observed and unobserved tendered trips based on the fact that there were observed trips

(however few) in those long duration tendered trips. Since 2013, permutation tests have replaced these frequency plots. However, these permutation tests do not visually map the data for observed and tendered states together. To accomplish this, a plot of the trip durations for these states is included as Figure 10. While tendered trips can be as long as a month, there appears to be a lack of observed tendered trips with Pot and Trawl Gear longer than 10 days. Whether this is due to an observer effect through intentional manipulation of trips (facilitated by the flexibility in ODDS and the current trip definitions), is due to the structure of the data (observed trips and trips with VMS are shortened since all unobserved non-VMS deliveries to a tender are lumped into the same trip), or simply low sample size is unknown.

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 and the number of observed trips over the course of 2015 was used to illustrate their combined effect on the probability of a NMFS Area containing observer data using the hypergeometric distribution (Fig. 11). From this figure it can be seen how 1) the likelihood of at least one observation is increased with fishing effort and 2) is also increased with an increase in the selection rate. Given our sampling rates in the two partial coverage trip selection stratum, there is a low

expectation that any trips will be observed unless more than 20 trips occur in a NMFS area for the *t* stratum and over 40 trips within a NMFS Area for the *T* stratum. 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 2017.

RESPONSE TO SSC COMMENTS

The SSC has requested that a specific section with responses to SSC comments be provided in the written report, as is done for SAFE documents. This section address comments made by the SSC (in italics) in response to the presentation of the 2014 Annual Report made at the June 2015 Council meeting.

- *The SSC recommends developing the necessary procedures for calculating the variances associated with point estimates. There is a critical need to calculate the variances associated with the point estimates (e.g., target catch, bycatch, PSC) to aid with optimization of the observer deployment sampling design and to assess uncertainty in estimates of catch. Consider, as a first-step, the calculation of variance using standard multi-stage cluster sampling (Thompson 2012), wherein the stage-specific variance is calculated along with the mean.*

The purpose of this chapter is to assess whether or not the goals of the Observer Program to obtain a representative sample was achieved at the deployment level of the design. The catch estimation process is evaluated and summarized in separate documents (Cahalan et al. 2014). Estimation of variances is outside the scope of this report, and is scheduled to be presented at the June 2016 Council Meeting (SSC only) as a separate agenda item.

- *In addition to sample size needs for spatial and temporal coverage, develop accuracy and precision objectives for catch, PSC, and bycatch.*

The purpose of this chapter is to assess whether or not the goals of the Observer Program to obtain a representative sample was achieved at the deployment level of the design. Setting of precision objectives for catch, PSC, and bycatch are outside the scope of this report.

- *The SSC recommends that an appropriate level of stratification for sampling beyond, or as a replacement for, vessel length be investigated.*

With the exception of the strata within the zero coverage pool of vessels, the 2016 ADP defines sampling strata based on gear and not vessel length (NMFS 2015a). This decision was reached after examining 12 alternative designs that defined strata based on alternative factors than vessel length (however holding to the requirement that stratum definitions be based on factors known before fishing begins; Faunce 2015b).

- *The SSC recommends that sampling issues with tendered trips be addressed. There is a critical need that regulatory action be considered. Our primary concerns are with the potential for bias, caused by trips that are tendered versus those that are not, and the inability to collect a representative sample of salmon PSC from tendered trips.*

We found evidence that trips delivering to tenders were different from trips that did not deliver to tenders, and these differences were present in both 2015 and 2016 strata definitions. However, there was only some evidence of an observer effect within tendered trips when considering retained catch of groundfish. This means that with respect to groundfish, there is not overwhelming evidence that observer data was biased because of tendering activity.

Tendering activity has however affected the ability of the observer program to achieve a random sample of trips to obtain samples of salmon caught as bycatch within the partial coverage portion of the trawl pollock fishery. Evidence of a substantial impact on the observer program sampling of genetics within the pollock fishery was only found at King Cove during 2015. The creation of a new sampling strata may allow flexibility in observer sampling methodology to accommodate tendering activity. However, tendering activity would need to be known prior to deployment and sampling methods would still need to be developed. This approach does not require regulatory action and does not unnecessarily expose human lives to increased safety risk.

- *The SSC recommends that the policy of allowing trip cancellation and logging multiple trips prior to sailing be reevaluated.*

For the October 2015 Council meeting, the NMFS proposed reducing the number of 'open' trips (i.e., the number of trips in which the disposition of the trip had not yet been recorded by the vessel operator) from three to two as part of the draft 2016 ADP, citing temporal bias and that trips selected to

be observed were cancelled at a much higher rate than trips that had not been selected to be observed. The Council heard public testimony that the differential cancelation rate was a function of how the ODDS operates- selected trips must be cancelled or dates changed whereas non-selected trips can remain open and closed at any time. Based on this testimony, the Council did not support the request to change the number of open trips in ODDS for 2016. We again point out that disproportionate cancellation of selected trips compared to non-selected trips occurred in 2015, and offer that the ability to cancel trips in ODDS be re-evaluated (see recommendations section below).

- *Evaluate performance relative to the success of observer deployments. For example, report on those statistic associated with numbers of successfully completed trips versus total observed trips, and differences in trip metrics associated with trips where there were observer complaints versus those without complaints.*

Numerous tables in this chapter contain the information on the number of total vessels and trips and the number of observed vessels and trips. The observer statement database is not readily useful for analytics, and essentially stores data in a single open text field. Multiple potential violations of a given nature are grouped into a single statement organized by an observer cruise, which includes multiple trips that are not linked to those definitions used in this Chapter. In addition, some categories used for statements are not intuitive (e.g., AFA, A80, Miscellaneous Violations). Considerable work is required to conduct the analyses suggested in the latter part of this comment, and it may be supplemental to the contents of this chapter, the purpose of which is to assess whether or not the goals of the Observer Program to obtain a representative sample was achieved at the deployment level of the design.

- *Examine the potential association of prohibited species catch (PSC) with trip attributes on observed vessels. If associations are found, PSC rates in shoreside offloads from unobserved vessels could be compared for evidence of bias.*

The purpose of this chapter is to assess whether or not the goals of the observer program to obtain a representative sample was achieved at the deployment level of the design for vessels in the partial coverage fleet. The comparison of PSC rates in shoreside offloads from unobserved vessels is outside the scope of this report. Dockside monitoring by observers only occurs within the trawl pollock fishery. In the Bering Sea, all trips are observed, hence, there are no comparisons between observed and unobserved

deliveries. Shoreside sampling for PSC in the partial coverage fleet only occurs in the Gulf of Alaska trawl pollock fishery and only for salmon. Further, the SSC comment indicates that a comparison of PSC rates between observed and unobserved vessels should be undertaken; however, if the intent was to compare CAS rates, we note that rates from observed vessels are applied to unobserved vessels in CAS, therefore, this would not be a valid comparison.

RECOMMENDATIONS TO IMPROVE DATA QUALITY

The Observer Science Committee made the following recommendations in its 2014 review of observer deployment to be considered in developing the 2016 ADP (NMFS 2015b, Faunce et al. 2015). Following each italicized recommendation is the outcome 'to-date' of that recommendation.

Recommendations from the 2014 Annual Deployment Review

- *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 Council did not support the recommendation that the number of ODDS trips be reduced from three to two.

- *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.*

  See our recommendation on this issue.

- *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 2016 ADP removed the vessel length definitions from the partial coverage strata with the exception of those strata within the zero selection pool.

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

The 2016 ADP utilized relationship models between the number of invoiced days and fishing days throughout the year to project the expected number of fishing days until the end of the year, and then used the relationship between past fishing activity as a proportion of the quota over the past three years to account for expected fishing effort in one stratum (NMFS 2015a).

- *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.*

A list of vessels unable to accommodate an observer has not been made. To accommodate the ephemeral nature of vessels participation in the EM voluntary pool, strata are defined by vessel: date combinations and can be changed upon notification by the vessel of their intent to no longer participate in the program (wherein they are returned to the partial coverage stratum specified in the 2016 ADP).

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

In 2016 a voluntary field was added to *e*Landings to accommodate the entry of the ODDS trip number.

- *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.

The Council initiated a discussion paper on this subject at its February 2016 meeting (NPFMC 2016). In addition to evaluating the impacts of Council motion alternatives, this discussion paper focused on a where, when, and how many tendering trips occur in the fleet and their trends over time. These data should not be confused with the results in this review that examine the differences between tendered and non-tendered; observed and unobserved trip states within a single year. For example, an increase or decrease in the proportion of the catch landed to tenders in a given NMFS Area or FMP over time does not equate to a change in the observer effect or bias. In this report, we found only some evidence for an observer effect within tendered trips. Notwithstanding, we did find substantial differences between tendered and non-tendered trips.

OSC Recommendations to Improve Data Quality and Guide the 2017 ADP

We appreciate the efforts made by the NMFS and the Council to address all of the recommendations made in the 2013 version of this report (ODDS rates should remain constant in time, revisions to conditional release policies, and improvements to vessel-selection; NMFS 2014b, Faunce et al. 2014). The Observer Science Committees recommendations to improve the 2017 ADP are as follows:

- The OSC recommends that tendered vessels be addressed differently in future ADPs. In any proposed solution to this issue, particular attention must be paid to ensure the safety of observers.
  - Tendered trips should be evaluated as separate strata in future ADPs -.
  - There is not a way to identify the duration of fishing trips made by catcher boats delivering to tenders without an observer or VMS on-board. The OSC recommends that NMFS and

Council address this data gap. The OSC supports the continued expansion and implementation of *t*Landings.

- The OSC reiterates our 2014 recommendation that the expansion of the pool of partial coverage catcher processors warrants their treatment as separate strata in future ADPs.

- Three observed trips are needed to calculate variance. The OSC recommends that sampling rates in future ADPs be high enough in each stratum to maximize the probability of achieving three observed trips in each NMFS Area. In simulated sampling evaluations of 2014 data, most observer data gaps disappeared or were severely minimized at deployment rates greater than or equal to 15% (relative to a 50% probability of a post-strata being empty; NMFS 2015c, p.98). In 2015, selection rates in the *t* stratum were 12%, and an actual observation rate of 11.2% was achieved. At this level of coverage numerous NMFS Areas without any observer coverage resulted. The temporal bias present in the *T* stratum in 2014 when selection rate was 15% was no longer present in 2015 when selection rates were set at 24%.

- The OSC recommends that NMFS should work with its partial coverage contractor and the OAC to explore the possibility of eliminating the ability to cancel a trip in ODDS, since the ability to change dates is already facilitated.

CITATIONS

AFSC (Alaska Fisheries Science Center). 2015. 2015 Observer Sampling Manual. Fisheries Monitoring and Analysis Division, North Pacific Groundfish Observer Program. AFSC, 7600 Sand Point Way N.E., Seattle, Washington, 98115. Available online at http://www.afsc.noaa.gov/FMA/Manual_pages/MANUAL_pdfs/manual2015.pdf.

Cahalan, J., J. Mondragon, 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. 2015a. 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. H. 2015b. An initial analysis of alternative sample designs for the deployment of observers in Alaska, U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-307, 33 p.

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.

Faunce, C., J. Cahalan, J. Gasper, T. A'mar, S. Lowe, F. Wallace, and R. Webster. 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. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-291, 26 p.

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

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 (National Marine Fisheries Service). 2015a. 2016 Annual Deployment Plan for Observers in the Groundfish and Halibut Fisheries off Alaska. National Oceanic and Atmospheric Administration, 709 West 9th Street. Juneau, Alaska 99802. Available online at: https://alaskafisheries.noaa.gov/sites/default/files/final2016adp.pdf.

NMFS. 2015b. North Pacific Groundfish and Halibut Observer Program 2014 Annual Report. National Oceanic and Atmospheric Administration, 709 West 9th Street. Juneau, Alaska 99802. 101 p. plus appendices. Available online at https://alaskafisheries.noaa.gov/sites/default/files/annualrpt2014.pdf.

NMFS. 2015c. 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. Available online at https://alaskafisheries.noaa.gov/sites/default/files/analyses/finalea_restructuring0915.pdf.

NMFS 2014a. 2015 Annual Deployment Plan for Observers in the Groundfish and Halibut Fisheries off Alaska. National Oceanic and Atmospheric Administration, 709 West 9th Street. Juneau, Alaska 99802. Available online at https://alaskafisheries.noaa.gov/sites/default/files/final2015adp.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. Available online at https://alaskafisheries.noaa.gov/sites/default/files/annualrpt2013.pdf.

NMFS 2013. 2014 Annual Deployment Plan for Observers in the Groundfish and Halibut Fisheries off Alaska. National Oceanic and Atmospheric Administration, 709 West 9th Street. Juneau, Alaska 99802. Available online at https://alaskafisheries.noaa.gov/sites/default/files/adp2014.pdf.

NPFMC (North Pacific Fishery Management Council). 2016. Deployment of Observers on Catcher Vessels Delivering to Tender Vessels Discussion Paper. 40 p. Available online under agenda item C-9 at http://legistar2.granicus.com/npfmc/meetings/2016/2/934_A_North_Pacific_Council_16-02-01_Meeting_Agenda.pdf.

NPFMC. 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_earirfa0311.pdf.

Thompson, S. K. 2012. Sampling, 3rd Edition. John Wiley & Sons Inc., Hoboken, NJ. 472 pp.

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

Table 1. -- Disposition of trips in the ODDS for 2015. “Paper” indicates trips that were logged when the ODDS was not available.

Strata	Random Selection	Logged	Cancelled by System	Cancelled by User	Waived	Paper
T	Not Selected	3534	248	125	1	0
T	Selected	1104	0	146	3	0
Vol. EM	Not Assigned	83	4	3	0	0
Vol. 100%	Not Assigned	178	2	0	0	0
t	Not Selected	1898	298	46	1	0
t	Selected	249	0	59	7	0
Total		7046	552	379	12	0

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 also shown (No Waivers).

Variable	*t*	*T*
Random Selection Only: Selected	249	1104
Random Selection Only: Total	2147	4638
Random Selection Only: Selection Rate	11.6	23.8
Final Expected: Selected	220	1047
Final Expected: Total	1744	4119
Final Expected: Selection Rate	12.6	25.4
Final Expected No Waivers: Selected	228	1051
Final Expected No Waivers: Total	1744	4119
Final Expected No Waivers: Selection Rate	13.1	25.5

Table 3. -- Number of total vessels (*V*), sampled vessels (*v*), total trips (*N*), sampled trips (*n*) for each observer deployment stratum in 2015. The number of vessels are not additive – total vessels are unique. Expected coverages are in percent for trip selection. EM: Electronic Monitoring.

Coverage	Strata	*V*	*v*	*N*	*n*	% Trips Observed	Expected Coverage	Expected Coverage (min)	Expected Coverage (max)	Meets Expectations?
Full	Full	170	170	3524	3522	99.9				
Partial	EM Voluntary	13	1	92	2	2.2				
Partial	*t*	354	138	2148	241	11.2	12	9.9	12.6	Yes
Partial	*T*	289	234	4676	1094	23.4	24	22.2	24.6	Yes
Partial	Zero Selection	415	0	2001	0	0.0				
Total Fleet	Total	1184	498	12441	4859	39.1				

Table 4. -- The number of pollock deliveries by observation and tendering status. IFP: Inshore Floating Processor, Hbr: Harbor.

FMP	Coverage Category	Port	Total Trips (*N*)	Observed Trips (*n*)	% Observed	p-value Trips Observed	% Tender Trips	% Observed no Tenders	p-value Trips Observed no Tenders
Bering Sea	Full	Akutan	723	722	100		0	100	
Bering Sea	Full	Dutch Hbr.	830	830	100		0	100	
Bering Sea	Full	IFP	299	298	100		0	100	
Bering Sea	Full	King Cove	69	69	100		0	100	
Bering Sea	Full	Sand Point	15	15	100		0	100	
Bering Sea	Partial	IFP	1	0	0		0	0	
Gulf of Alaska	Full	Kodiak	8	0	0		0	0	
Gulf of Alaska	Partial	IFP	5	1	20		0	20	
Gulf of Alaska	Partial	King Cove	204	6	3		92	35	
Gulf of Alaska	Partial	Kodiak	1294	306	24		0	24	
Gulf of Alaska	Partial	Sand Point	360	90	25		4	26	
Total	Full		1944	1934	99		0	99	
Total	Partial		1864	403	22	0.016	11	24	0.796

Table 5. -- Number of trips by observation status in the 2015 trip-selection strata.

Strata	Observed	Unobserved
t	241	1907
T	1094	3582

Table 6. -- Results of permutation tests between observed and unobserved trips in the 2015 trip-selection strata. OD: Observed Difference.

Strata	NMFS Areas	Days Fished	Vessel Length	Species Landed	pMax Species	Landed Catch	Metric
t	-0.033	-0.561	0.379	-0.009	0.013	-0.994	OD
T	0.000	-0.311	1.050	-0.189	0.009	-2.156	OD
t	-3.1	-13.6	0.8	-0.2	1.4	-14.1	OD (%)
T	-0.0	-8.4	1.4	-4.5	1.0	-4.0	OD (%)
t	0.053	0.000	0.306	0.964	0.124	0.196	p-value
T	1.000	0.000	0.135	0.064	0.031	0.213	p-value

Table 7. -- Number of trips by tendered status in the 2015 trip-selection strata.

Strata	Tendered	Non-tendered
t	64	2084
T	313	4363

Table 8. -- Results of permutation tests between tendered and non-tendered trips in the 2015 trip-selection strata. OD: Observed Difference.

Strata	NMFS Areas	Days Fished	Vessel Length	Species Landed	pMax Species	Landed Catch	Metric
t	-0.044	0.234	4.419	-0.719	0.079	27.417	OD
T	-0.092	0.819	-13.217	-0.685	0.056	15.039	OD
t	-4.1	5.7	9.3	-20.1	8.8	389.9	OD (%)
T	-8.3	22.2	-17.4	-16.3	6.0	27.6	OD (%)
t	0.182	0.393	0.000	0.006	0.000	0.000	p-value
T	0.000	0.000	0.000	0.000	0.000	0.000	p-value

Table 9. -- Number of trips by tendered status in the 2015 trip-selection strata re-coded as 2016 trip-selection strata.

2016 Strata	Tendered	Non-tendered
HAL	12	3133
POT	185	1027
TRW	180	2287

Table 10. -- Results of permutation tests between tendered and non-tendered trips in the 2015 trip-selection strata re-coded as 2016 trip-selection strata. OD: Observed Difference.

2016 Strata	NMFS Areas	Days Fished	Vessel Length	Species Landed	pMax Species	Landed Catch	Metric
HAL	0.056	1.142	0.121	-0.794	0.027	11.600	OD
POT	0.003	1.344	-6.635	0.411	0.000	14.240	OD
TRW	-0.101	1.432	-24.891	-0.701	0.048	0.538	OD
HAL	5.0	23.9	0.2	-21.4	3.1	149.4	OD (%)
POT	0.3	37.7	-9.5	22.8	0.0	43.9	OD (%)
TRW	-9.2	52.6	-30.0	-12.9	5.1	0.6	OD (%)
HAL	0.651	0.104	0.980	0.205	0.465	0.001	p-value
POT	1.000	0.000	0.000	0.000	0.905	0.000	p-value
TRW	0.000	0.000	0.000	0.002	0.000	0.901	p-value

Table 11. -- Number of tendered trips by observation status in the 2015 trip-selection strata.

Strata	Observed	Unobserved
t	17	47
T	45	268

Table 12. -- Results of permutation tests between observed and unobserved tendered trips in the 2015 trip-selection strata. OD: Observed Difference.

Strata	NMFS Areas	Days Fished	Vessel Length	Species Landed	pMax Species	Landed Catch	Metric
t	0.059	-1.189	1.618	-0.711	-0.007	-34.103	OD
T	-0.026	-2.261	3.498	-0.516	0.007	-22.792	OD
t	5.8	-27.4	3.1	-24.7	-0.7	-101.4	OD (%)
T	-2.6	-50.8	5.5	-14.5	0.7	-33.3	OD (%)
t	0.263	0.264	0.066	0.036	0.824	0.081	p-value
T	0.507	0.005	0.097	0.216	0.123	0.092	p-value

Table 13. -- Number of non-tendered trips by observation status in the 2015 trip-selection strata.

Strata	Observed	Unobserved
t	224	1860
T	1049	3314

Table 14. -- Results of permutation tests between observed and unobserved non-tendered trips in the 2015 trip-selection strata. OD: Observed Difference.

Strata	NMFS Areas	Days Fished	Vessel Length	Species Landed	pMax Species	Landed Catch	Metric
t	-0.037	-0.538	0.087	0.070	0.010	-0.281	OD
T	-0.002	-0.189	0.466	-0.198	0.011	-0.632	OD
t	-3.5	-13.1	0.2	2.0	1.2	-4.5	OD (%)
T	-0.2	-5.2	0.6	-4.7	1.2	-1.2	OD (%)
t	0.046	0.001	0.826	0.674	0.219	0.480	p-value
T	0.869	0.008	0.510	0.063	0.010	0.698	p-value

FIGURES

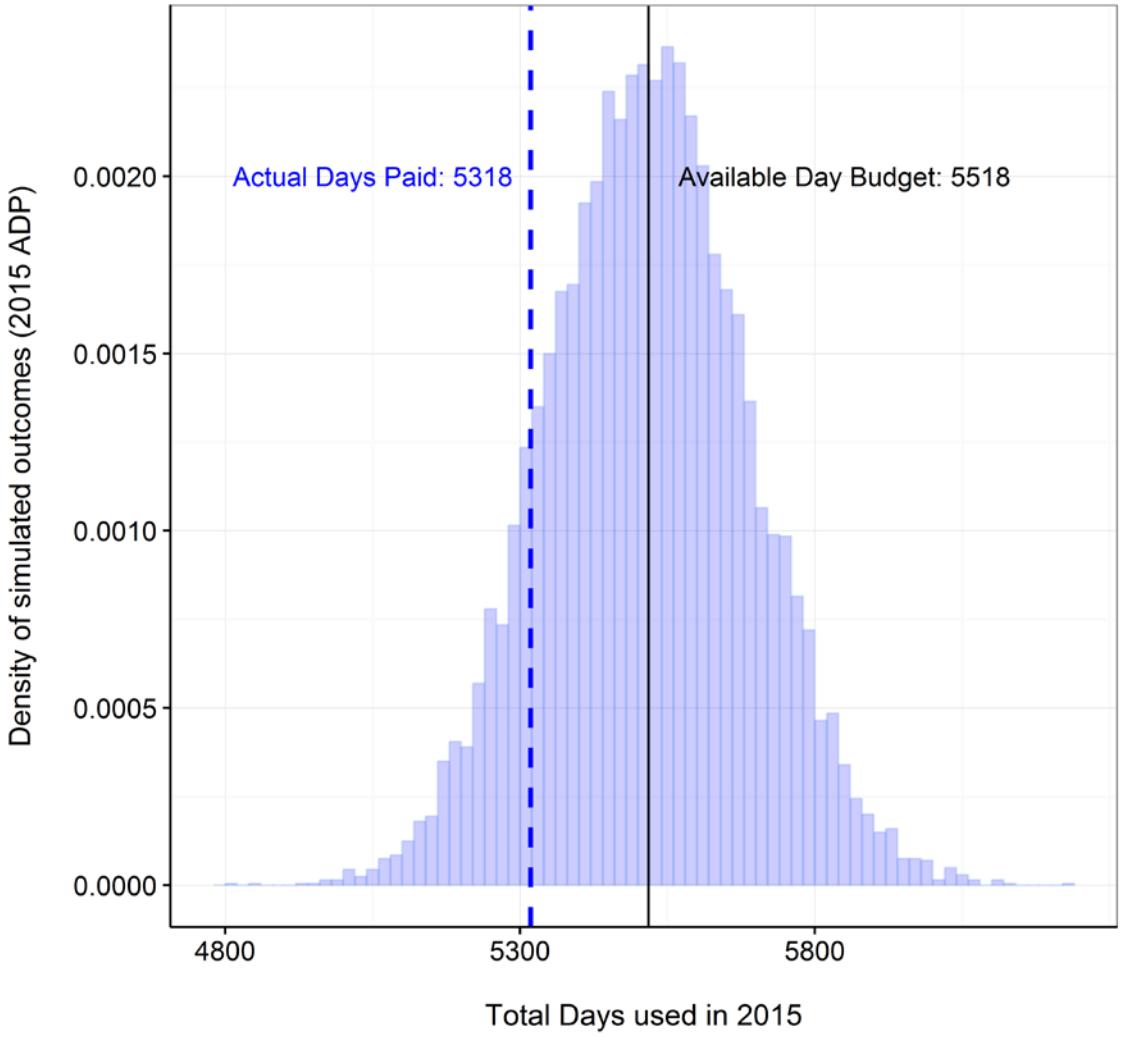

Figure 1. -- Actual paid sea-days in 2015 (dotted line) in relation to the range of potential budgetary outcomes estimated in December 2014 for the Final 2015 Annual Deployment Plan (vertical bars).

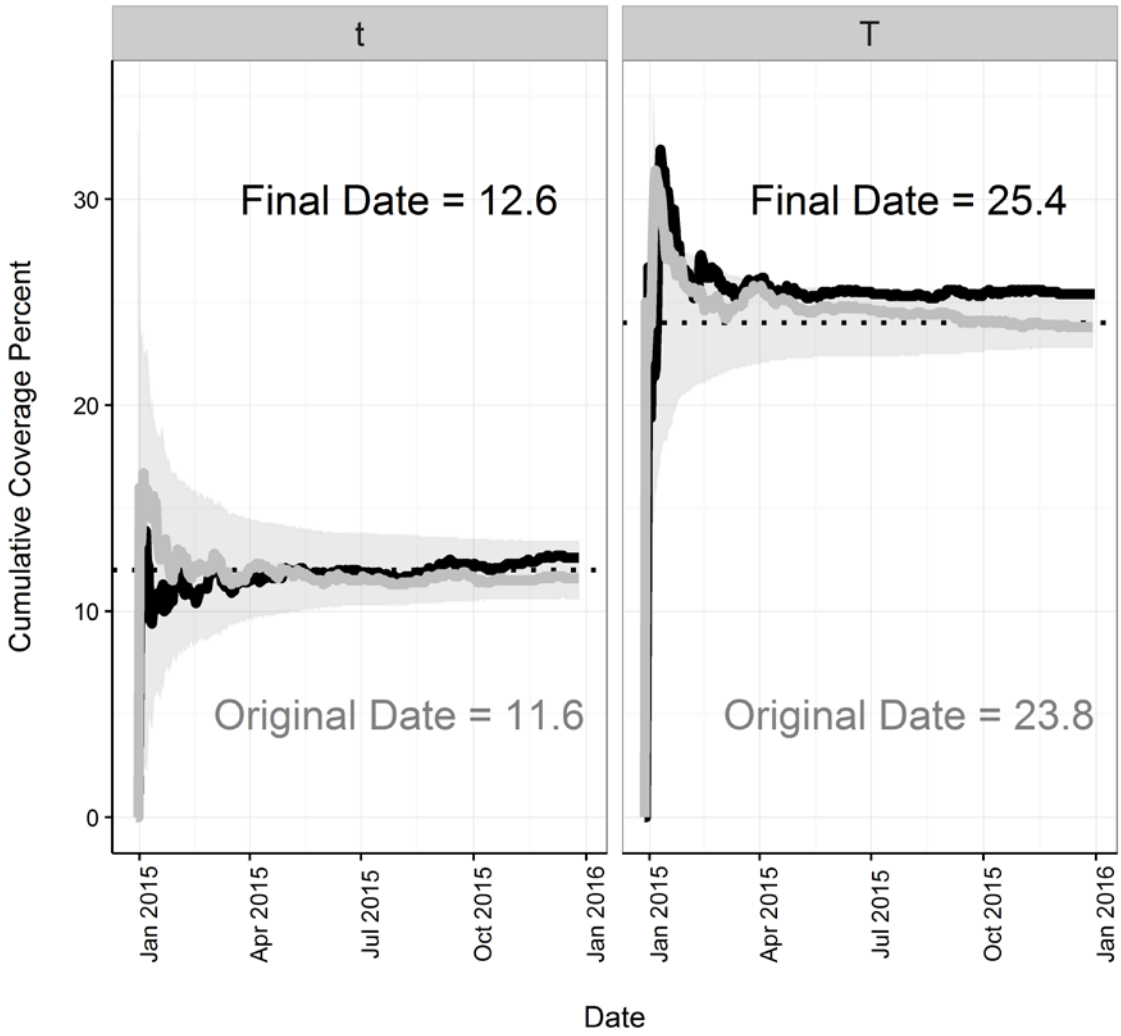

Figure 2. -- Rate of selected trips logged into ODDS organized by original date entered for all trips (grey line and grey text), and final date considering only non-cancelled trips (black line and black text). The programmed selection rate is depicted as the dotted line. Grey shaded areas denote the range of coverage rates corresponding 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 / or cancelled.

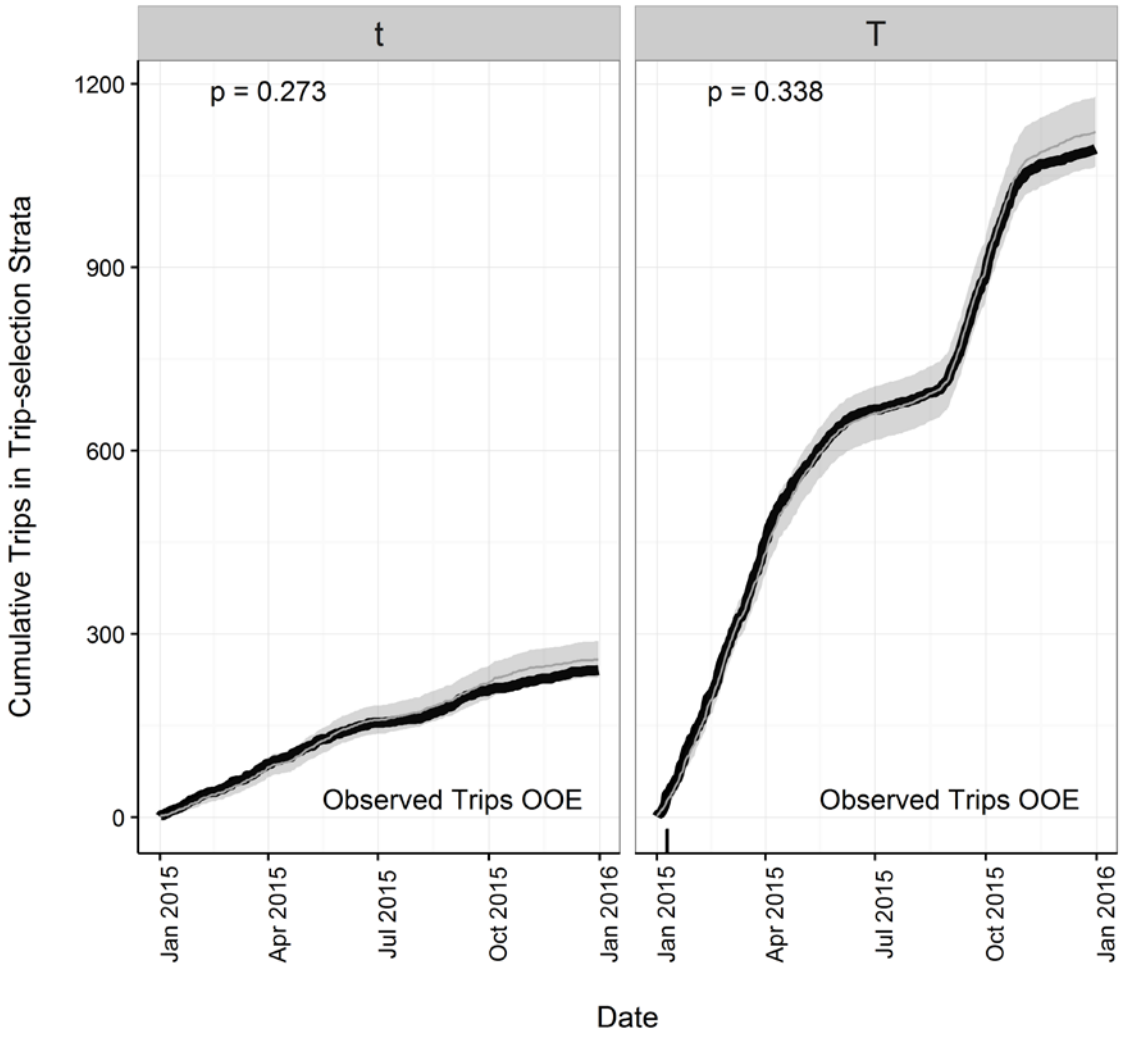

Figure 3. -- Cumulative number of trips observed during 2015 (black line) compared to the expected range of observed trips (shaded area) given fishing effort and sampling rates. Dates where the observed number of trips is outside of expected (less or more than the range; OOE) are depicted as tick marks on the horizontal x-axis. The results of tests that the observed rate derived from a binomial distribution sampled at the selection rate are denoted as p-values.

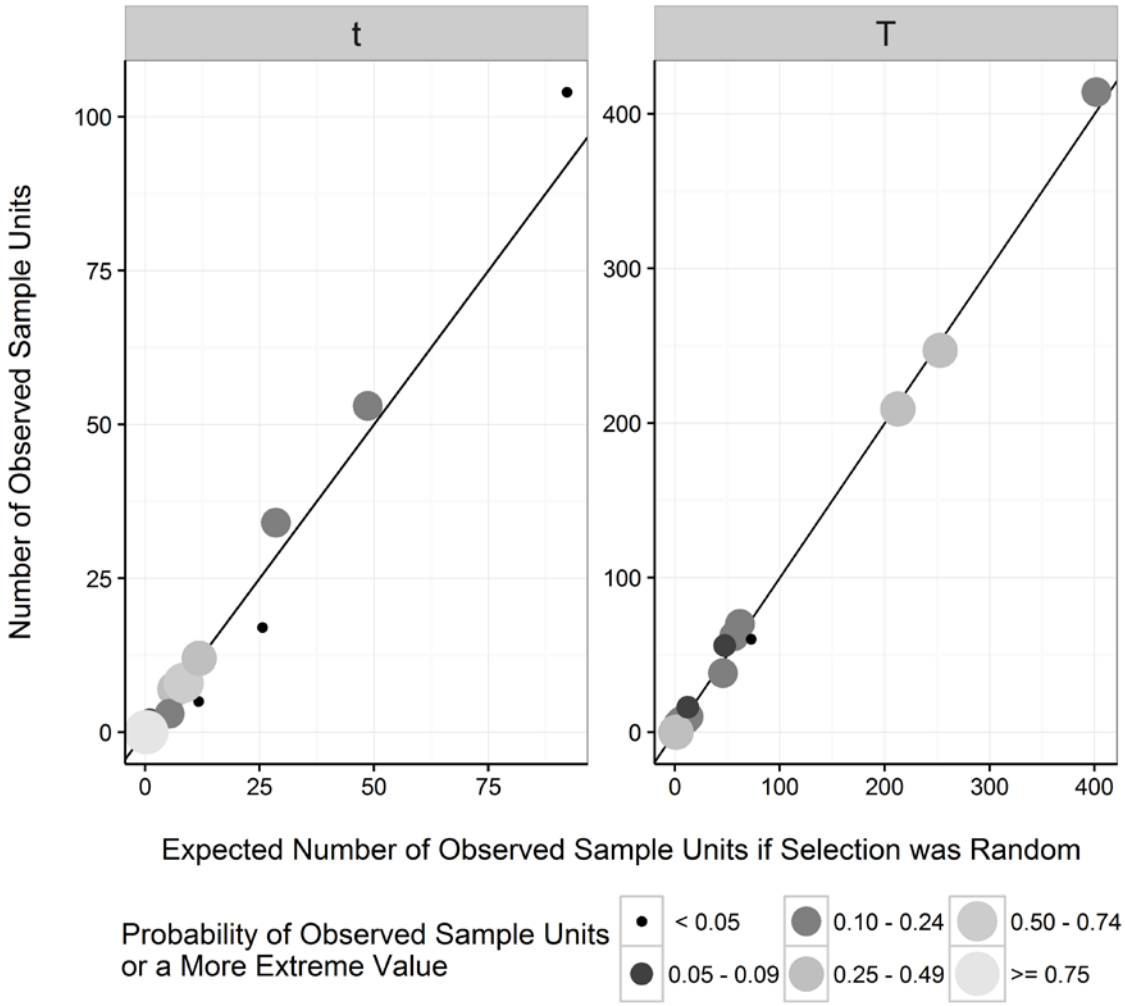

Figure 4. -- Comparison plots depicting the number of observed sample units compared to the number of expected observed sample units for each partial coverage stratum. Each point on a plot represents a NMFS Area. The smaller the point, the more unusual the result.

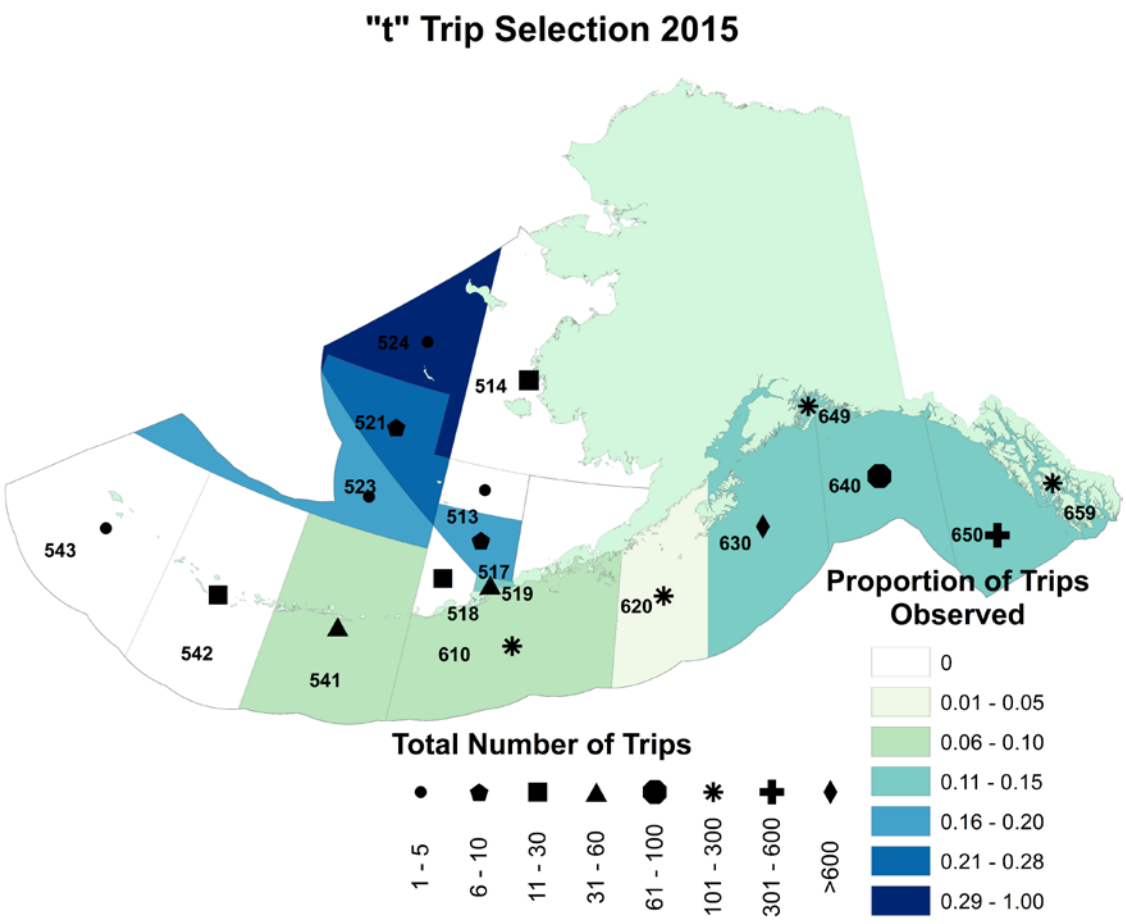

Figure 5. -- Proportion of trips observed in each NMFS Reporting Area in the 'small t' stratum. The color of the Reporting Area reflects the proportion of trips that were observed while the symbol indicates the total number of fishing trips that occurred in that area.

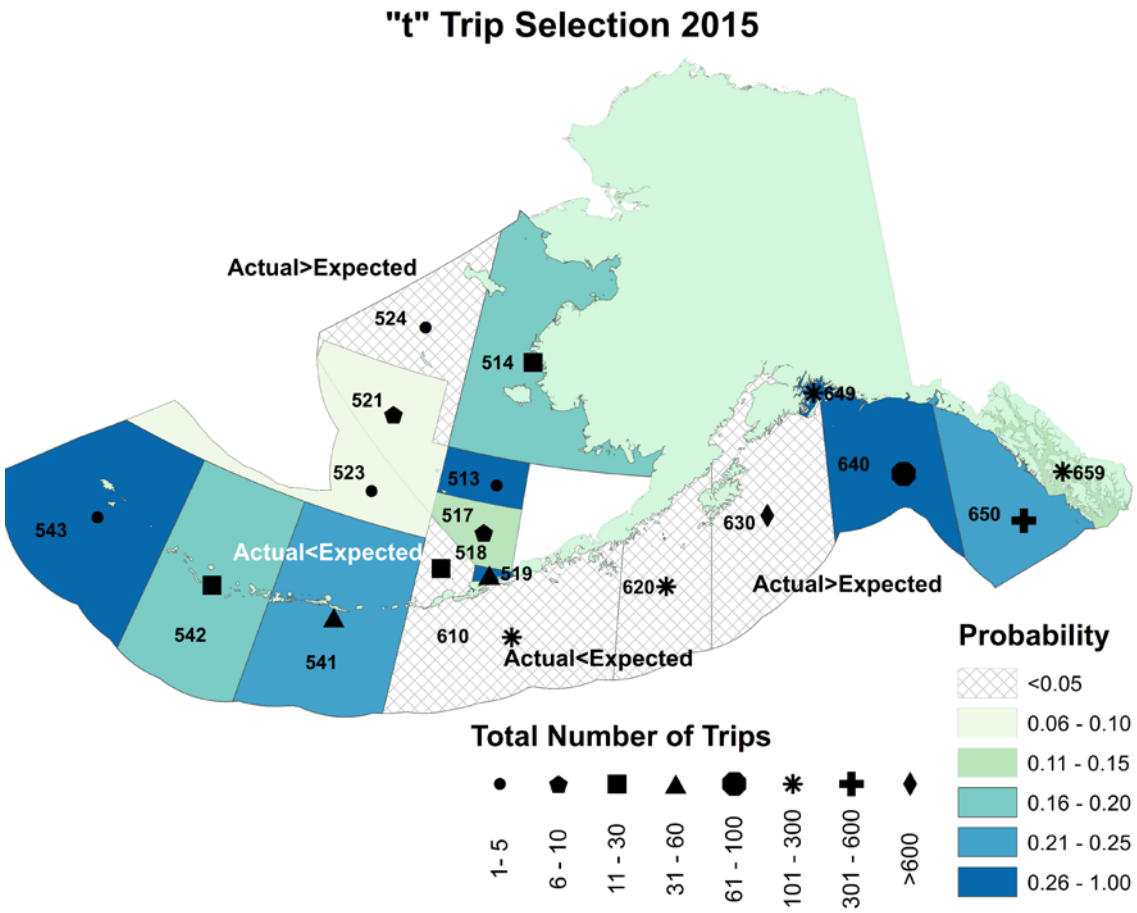

Figure 6. -- The probability of observing a number of trips in the 'small t' stratum as far or farther from expected values (probability of observing a more extreme value). The symbol indicates the total number of fishing trips that occurred in that area.

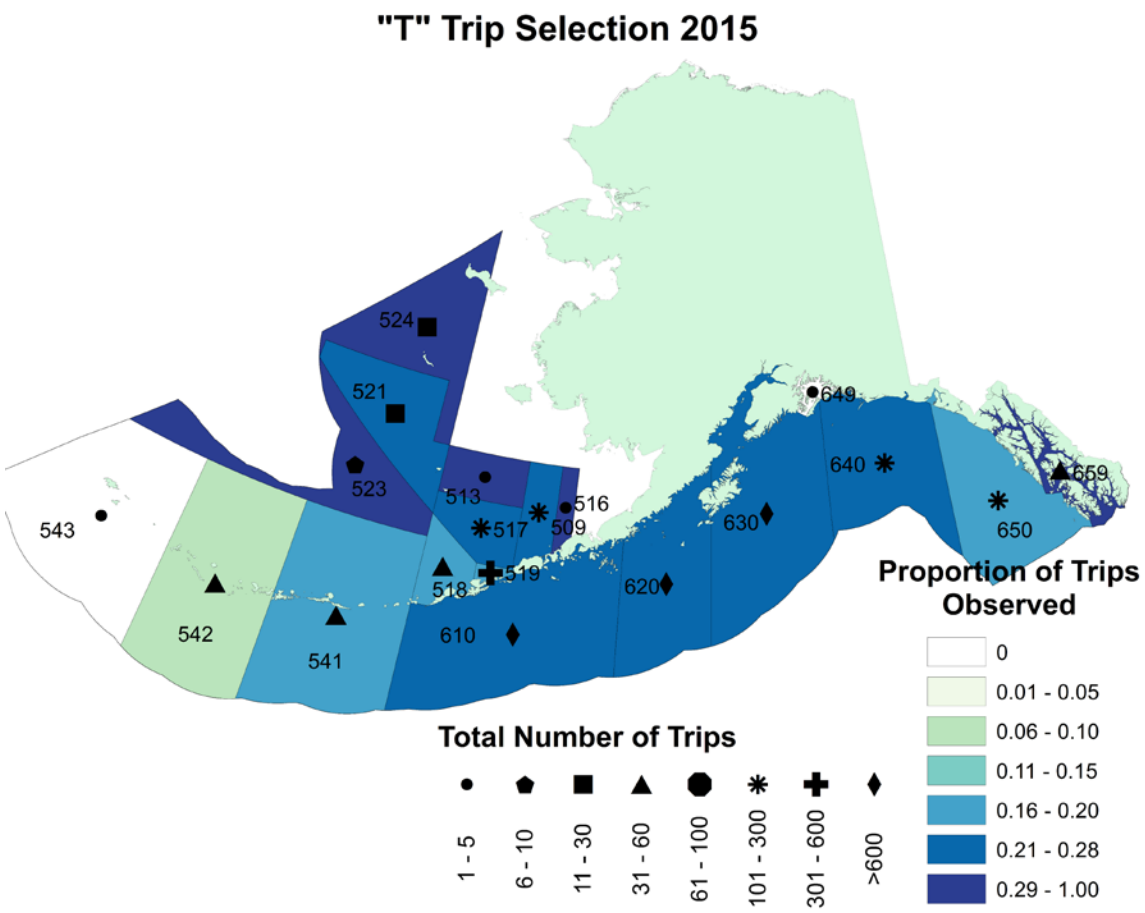

Figure 7. -- Proportion of vessels observed in each NMFS Reporting Area in the 'large T' stratum. The color of the Reporting Area reflects the proportion of trips that were observed while the symbol indicates the total number of trips that occurred in that area.

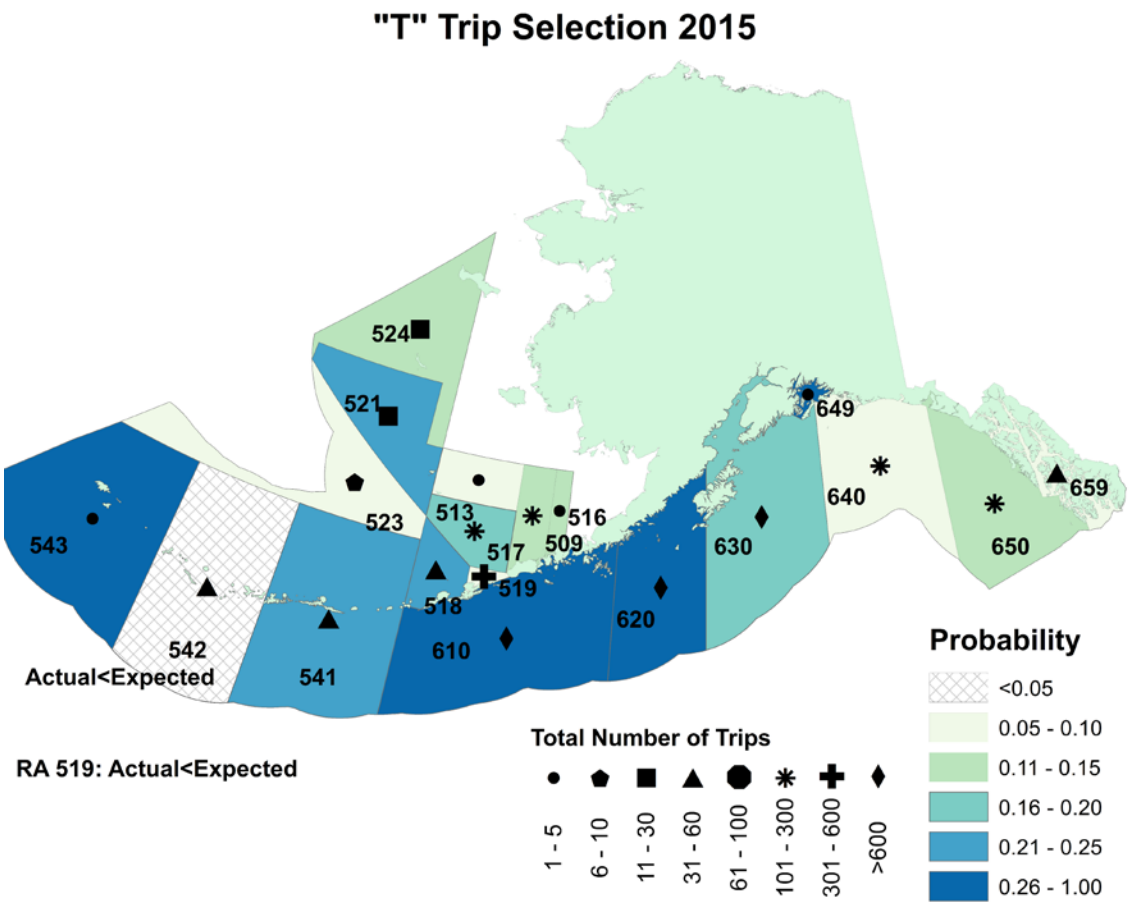

Figure 8. -- The probability of observing a number of trips in the 'Large T' stratum as far or farther from expected values (probability of observing a more extreme value). The symbol indicates the total number of fishing trips that occurred in that area.

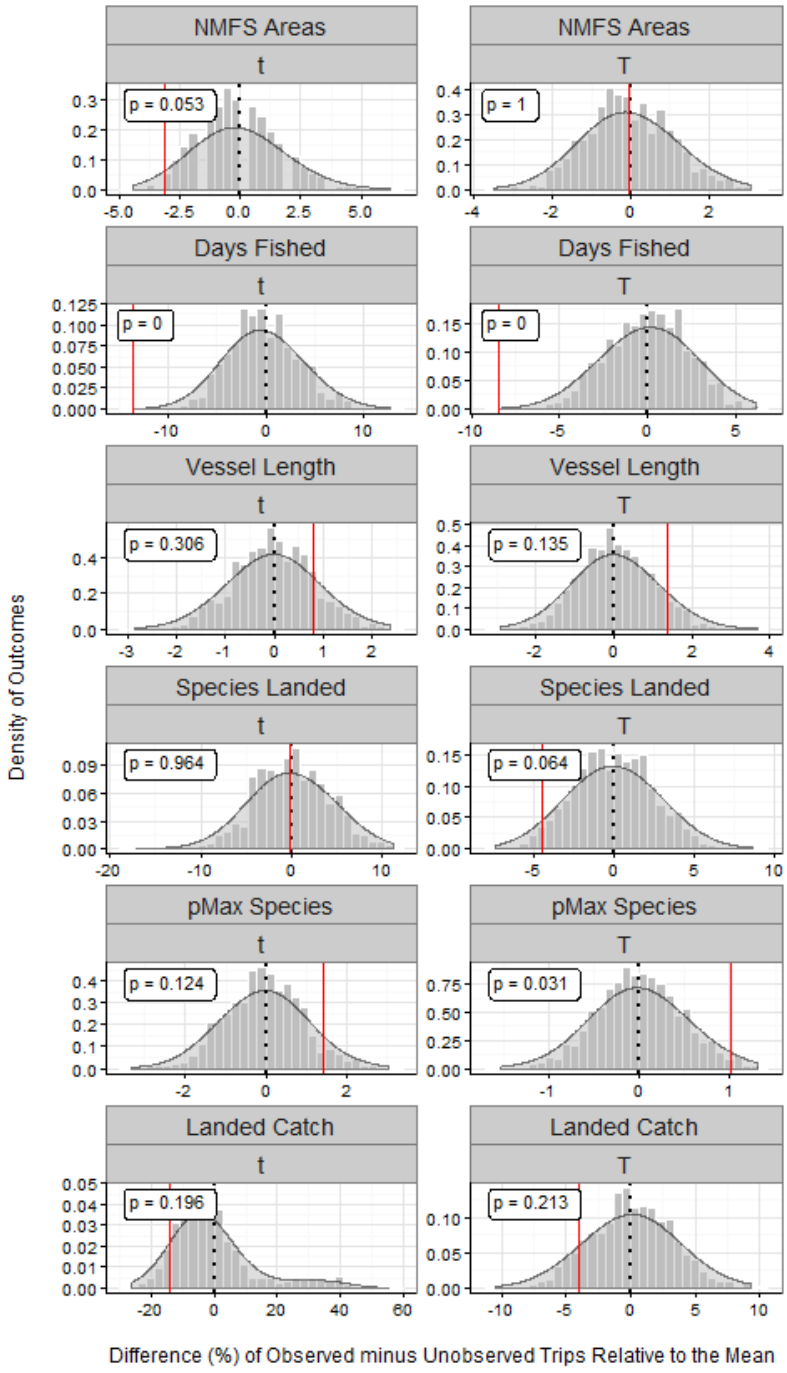

Figure 9. -- Results of permutation tests between observed and unobserved trips for each strata in the partial coverage category of the 2015 ADP. In each panel, the grey bars depict the distribution of differences between observed and unobserved trips where the assignment of observed status had been 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.

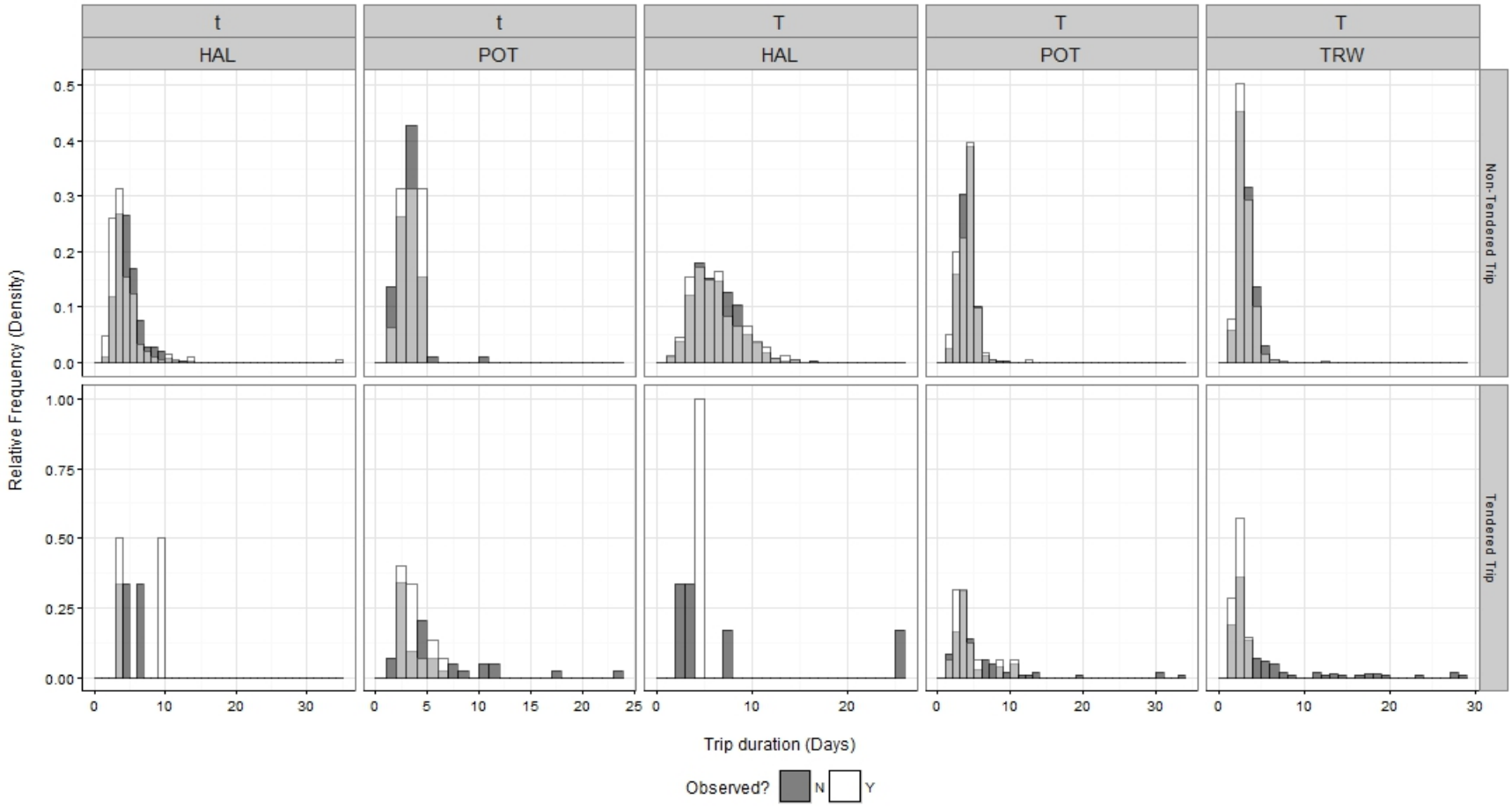

Figure 10. -- Distribution of trip duration for vessels in the partial coverage category by stratum, gear, and observation status. Observed trips are depicted as translucent white bars overtop of solid black bars for unobserved trips. Trip durations where both observed and unobserved status exist are depicted as gray (This is not the same as 'stacked bars', in which the height of the bar would reflect observed and unobserved on top of one another- this plot is has each observation status in front of the other).

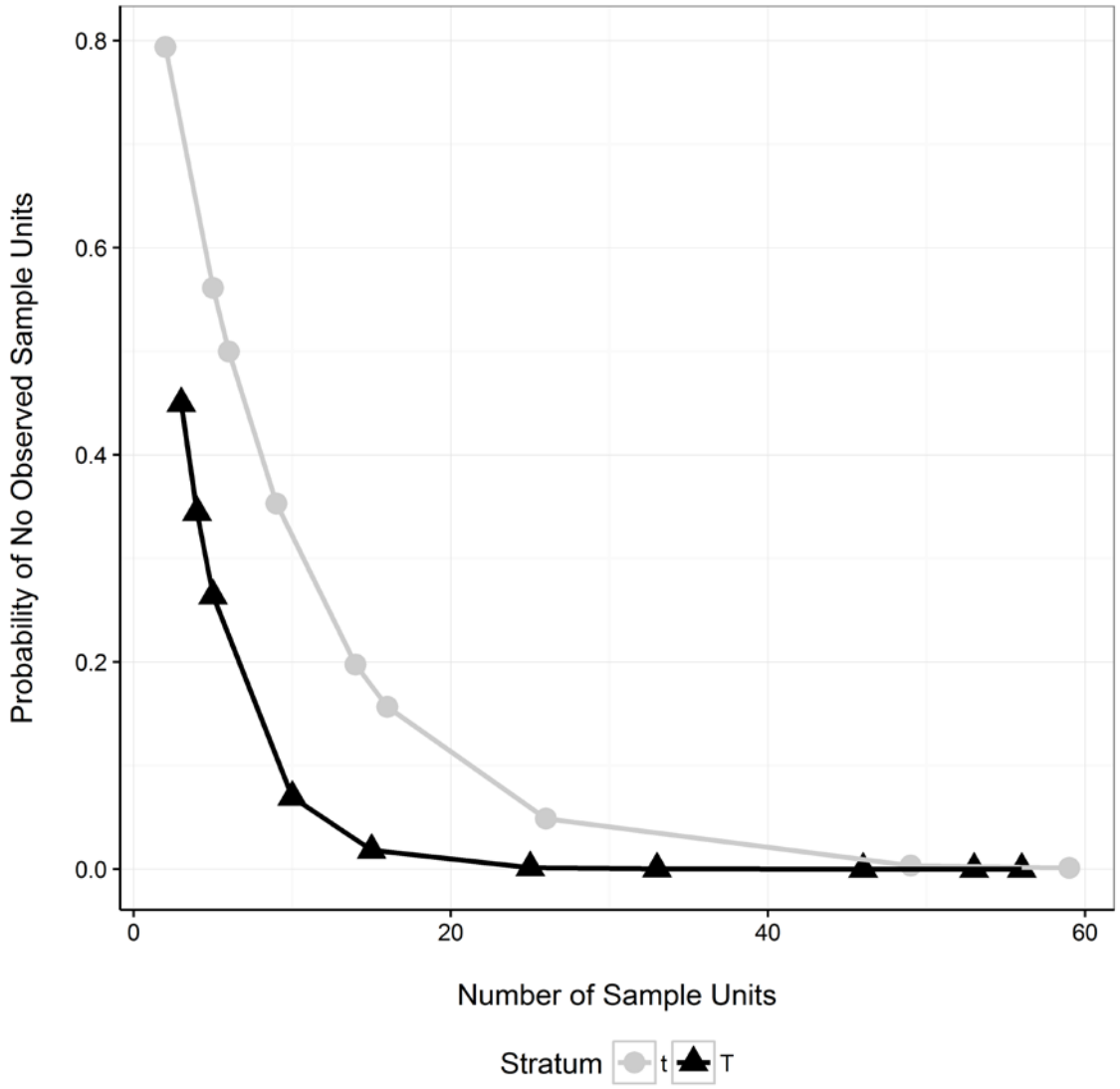

Figure 11. -- Probability of selecting a sample and observing no trips as a function of the number of sample units and selection rate that occurred in a NMFS Area, time period, and stratum. The x-axis has been truncated to increase resolution at smaller numbers of sampling units. The likelihood of having no observer data decreases with increasing total fishing effort and selection rate. The *t* stratum has a selection rate of 12%. The *T* stratum had a selection rate of 24%.

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-

321 FRITZ, L., K. SWEENEY, R. TOWELL, and T. GELATT. 2016. Aerial and ship-based surveys of Steller sea lions (*Eumetopias jubatus*) conducted in Alaska in June-July 2013 through 2015, and an update on the status and trend of the western distinct population segment in Alaska, 72 p. NTIS number pending.

320 ECHAVE, K. B. 2016. Feasibility of tagging sablefish, *Anoplopoma fimbria*, with pop-off satellite tags in the northeast Pacific Ocean, 38 p. NTIS number pending.

319 BARBEAUX, S. J., S. ROMAIN, E. LOGERWELL, and D. FRASER. 2016. Aleutian Islands Cooperative Acoustic Survey Study, 2007, 35 p. NTIS number pending.

318 WHITEHOUSE, G. W., and K. Y. AYDIN. 2016. Trophic structure of the eastern Chukchi Sea: An updated mass balance food web model, 175 p. NTIS number pending.

317 SMITH, T. A., and R. A. MCCONNAUGHEY. 2016. The applicability of sonars for habitat mapping: a bibliography, 129 p. NTIS number pending.

316 TESTA, J. W. (editor). 2016. Fur seal investigations, 2013-2014, 126 p. NTIS number pending.

315 HELKER, V. T., M. M. MUTO, and L. A. JEMISON. 2016. Human-caused injury and mortality of NMFS-managed Alaska marine mammal stocks, 2010-2014, 89 p. NTIS number pending.

314 KONDZELA, C. M., J. A. WHITTLE, D. YATES, S. C. VULSTEK, H. T. NGUYEN, and J. R. GUYON. 2016. Genetic stock composition analysis of chum salmon from the prohibited species catch of the 2014 Bering Sea walleye pollock trawl fishery and Gulf of Alaska groundfish fisheries, 49 p. NTIS No. PB2016-102777.

313 GODDARD, P., R. WILBORN, C. ROOPER, K. WILLIAMS, R. TOWLER, M. SIGLER, and P. MALECHA. 2016. Results of the 2014 underwater camera survey of the eastern Bering Slope and Outer Shelf, 304 p. NTIS number pending.

312 LAAKE, J. L., S. R. MELIN, A. J. ORR, D. J. GREIG, K. C. PRAGER, R. L. DeLONG, and J. D. HARRIS. 2016. California sea lion sex- and age-specific morphometry, 21 p. NTIS No. PB2016-102562.

311 GUTHRIE, C. M. III, HV. T. NGUYEN, and J. R. GUYON. 2016. Genetic stock composition analysis of the Chinook salmon bycatch samples from the 2014 Gulf of Alaska trawl fisheries, 31 p. NTIS No. PB2016-101398.

310 GUTHRIE, C. M. III, HV. T. NGUYEN, and J. R. GUYON. 2016. Genetic stock composition analysis of the Chinook salmon bycatch from the 2014 Bering Sea walleye pollock (Gadus chalcogrammus) trawl fishery, 25 p. NTIS No. PB2016-101397.

309 GUTHRIE, C. M. III, Hv. T. NGUYEN, and J. R. GUYON. 2016. Genetic stock composition analysis of the Chinook salmon bycatch from the 2014 Bering Sea walleye pollock (*Gadus chalcogrammus*) trawl fishery, 25 p. NTIS No. PB2016-101396.

308 DALY, B. J., C. E. ARMISTEAD, and R. J. FOY. 2015. The 2015 eastern Bering Sea continental shelf bottom trawl survey: Results for commercial crab, 167 p. NTIS No. PB2016-100681.

307 FAUNCE, C. H. 2015. An initial analysis of alternative sample designs for the deployment of observers in Alaska, 33 p. NTIS No. PB206-100705.