Deployment Performance Review of the 2015 North Pacific Observer Program

Observer Science Committee

Thu Mar 09 13:48:55 2017

# Introduction

# Deployment Performance 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.

## Observer Science Committee

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

This chapter contains the OSC review of the deployment of observers in 2016 relative to the intended sampling plan and goals of the 2016 ADP (**CITATION** - NMFS 2014a). This review 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 2018 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 for comparing targeted and achieved sampling rates. 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., tender trips). Specifically, this section assesses the following:
   1. Sample rates and number of samples relative to intended values.
   2. (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.
   3. (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:
   1. Temporal representativeness
      * 1. 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).
   2. Spatial representativeness
      * 1. 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.
        2. 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.
   3. Representativeness of trip characteristics
      * 1. 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 a measure of species diversity where an increase in pMax indicates a decline in diversity).

1. 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 having cells (e.g., defined by NMFS Reporting Area and strata) with no observer coverage.

# Changes to this report from last year

*Deployment Strata*

* Strata definitions
* EM
* Trip versus vessel selection

*Methodological changes*

* New analyses (and subsequent new sections, tables, or figures)
* Changes in definitions that may impact old analyses
* Changes to CAS, Valhalla, or ODDS

# Evaluation of Observer Deployment in 2016

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 2016 ADP, the ODDS was programmed to randomly select logged trips at a rate of 28% in the *TRW* stratum, 15% in the *HAL* stratum, and 15% in the *POT* 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 2016 was set at 5107 days for the 2016 ADP (NMFS 2014a). Based on simulations of 2014 fishing data made a year in advance of deployment, the FMA predicted it would observe 4900 fishing days at the end of 2016. In 2016, the FMA paid for 5107 observer days, which was 8.4 % lower than predicted (Figure 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 2016 ODDS trips from affecting the deployment rates set for the 2017 ADP. The number of trips logged in the ODDS in 2016 and their dispositions is summarized in Table 1. Of the trips logged, a total of 1052 trips were cancelled, including by ODDS (7.2%) and 541 by users (7.6%). A total of 28 trips were waived (0.40% of trips remaining after cancellations). 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 4.00% for Declared Gear - Trawl to 8.60% for Declared Gear - Pot gear for selected trips, and 0.00% for BSAI Cod 100% Voluntary Coverage to 25.30% for Declared Gear - Pot gear for non-selected trips.

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 percentages of selected trips from either inherits or waivers is found in Table 2. **INTERPRETATION NEEDED**.

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 (random selection only) and 2) random selection of remaining trips after they have had dates changed and are closed or cancelled (final selection rate). In either case, the proportion of trips selected to be observed should fall within what would be expected given the binomial distribution (since each trip is either selected or not selected). The rate obtained in the initial selection process was 15.88% for the *HAL* stratum, 14.27% for the *POT* stratum, and 28.39% for the *TRW* stratum. These values were well within the range of values expected from a binomial distribution (exact binomial test p-values = 0.483, 0.341, and 0.933 for *HAL*, *POT*, and *TRW* respectively). This means that the ODDS was selecting trips according to the programmed rate. The final selection rate after trips were closed and cancelled or waived was 17.72% for the *HAL* stratum, 14.42% for the *POT* stratum, and 29.55% for the *TRW* 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 (exact binomial test p-values = 0.003, 0.462, and 0.170 for *HAL*, *POT*, and *TRW* respectively). These rates and the potential impact of trip selection waivers is presented in Table 3.

Differences in the initial and final selection rates were evident throughout 2016. 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 (Figure 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 2016 ODDS provided users with a list of Report IDs from eLandings 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 AKRO), the Observer Program database NORPAC (managed by the AFSC), and *e*Landings (under joint management by Alaska Department of Fish and Game -- ADF&G; the Internationshal Pacific Halibut Commission -- IPHC; 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 *TRW*, *HAL*, and *POT* strata in the trip-selection pool, *EM Voluntary 30%* and *EM Voluntary 100%* strata in the vessel selection pool,, *EM Voluntary 30%* and *EM Voluntary 100%* strata in the vessel selection pool,, *EM Voluntary 30%* and *EM Voluntary 100%* strata in the vessel selection pool,, *EM Voluntary 30%* and *EM Voluntary 100%* strata in the vessel selection pool,, *EM Voluntary 30%* and *EM Voluntary 100%* strata in the vessel selection pool,, *EM Voluntary 30%* and *EM Voluntary 100%* strata in the vessel selection pool,and the zero-selection pool which included 3vessels participating inEM research.

**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 2016 was the rate programmed into ODDS for each stratum. Although the EM voluntary vessels are considered part the partial coverage category, we do not evaluate whether they met expectations based on the pre-implementation status of EM in the North Pacific.

The 2016 Observer Program had 15 different deployment strata to be evaluated. **Interpretation needed** the program ?met expected rates of coverage for all of these strata (Table 4). Among all fishing in Federal fisheries of Alaska, 6140 trips (46.5%) and 513 vessels (45.5%) were observed.

### Coverage Rates in Vessel Selection (Voluntary EM)

Vessel selection in the EM Voluntary category was based on fishing acticity in 4 selection time periods: Jan-Feb, Mar-Jun, Jul-Oct, and Nov-Dec. The ability to achieve a target number of observed vessels in vessel selection within each time period is hindered by the difficulty identifying a complete sampling frame, which should include all the elements of the population of interest. A complete sampling frame for vessel selection would consist of a list of vessels that actually fish in each time period. In trip selection, only vessels that intend to fish log trips into ODDS. Consequently, the trip selection sampling frame for the observer program is equal to the target population. However in vessel selection, without a similar notification system informing NMFS of their intent to fish, the sample frame is based on past fishing behavior (specifically whether the vessel landed catch in the same time period the year prior). Several lists were compiled by NMFS in an attempt to achieve a complete sampling frame. First, the EM Working Group and NMPFMC published a list of vessels anticipated to fish in the EM Voluntary category in the Final 2016 Electronic Monitoring Pre-Implementation Plan (EMWG/NPFMC 2015; Table 5, Row 1). Second, NMFS compiled a list of eligible vessels at the beginning of 2016 based on vessel owners who contacted NMFS and indicated they were interested in participating in EM. Vessel owners could contact NMFS at any time throughout the year to be removed from this list. These vessels are not yet considered in the sampling frame, because we don't know when they will fish. In order to get this information, vessel operators contacted NMFS at least 30 days prior to fishing activity within in a selection time period. These vessels, which we categorize as the *EM Voluntary 30%* stratum, were considered "within the sampling frame" for that time period and were selected for EM coverage at a 30% rate (Table 5, Rows 2 and 8). Vessels who missed the 30 day cutoff prior to the start of the time period but still wanted to volunteer for EM were selected at a 100% rate if equipment was available. They are categorized as the *EM Voluntary 100%* stratum and are not considered within the sampling frame (Table 5, Row 9) Finally, we compared the list of selected vessels within these two strata to the list of vessels from Pacific States Marine Fisheries Commission that had EM data within a given time period (these vessels were considered observed) (Table 5, Row 13).

**INTERPRETATION** The number of vessels that fished in 2016 was lower than the number of vessels anticipated to fish in the Final 2016 Electronic Monitoring Pre-Implementation Plan in all but the Jan-Feb time period (Table 5, Row 6 vs. 1). **Values of the relative amount of overdraw, (expressed as the number of selected vessels divided by the target number of vessels to be observed) ranged between 7.3 and 9.9 (average = 8.6) among time periods.** Between 0 and 14 vessels were selected and actually fished among time periods (Table 5, Row 12). Between 0 and 14 vessels had video data reviewed among time periods (Table 5, Row 13).

To measure the performance of the vessel selection process, data in Table 5 were expressed as relative percentages (Table 6). Over- and under-coverage rates in the vessel selection sampling frame are not additive since the former is a percentage of the sampling frame, and the latter is a percent difference from the true frame (i.e., the list of vessels that actually fished). **INTERPRETATION Over-coverage rates were 0 for all time periods, and under-coverage rates were greatest in the Nov-Dec time period** (Table 6, Rows 1 and 2). **INTERPRETATION :** If being selected for coverage has no effect on the likelihood that a vessel fishes in Federal waters, we would expect that the percentage of vessels that were in the selection frames and did not fish to be approximately equal to the percentage of vessels that were in the selection frame and were selected for coverage and did not fish. A comparison of Rows 1 and 3 of Table 6 shows that this was the case for the voluntary EM in 2016.

### 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 2016 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). For the analysis, pollock trips are defined as any delivery where the predominant species is pollock (i.e. trip target = pollock).

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 **NEEDS TO BE CHANGED** **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 (~0%). 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; Table 7). However, when deliveries of pollock from tender trips were removed, this likelihood was dramatically increased (p-value =0.003). 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 7).

# Sample quality

## 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 48 days (0.06%) very early in the year and within only one stratum (Figure 3). For comparison, the number of observed trips achieved was outside their expected values for (LAST YEAR ADP) of the year (LAST YR ADP REF). Results from the exact binomial test suggest that observed rates at the end of the year were within the expectation for all strata (*HAL* stratum expected rate=0.15, realized rate=0.152, p-value=0.743; *POT* stratum expected rate=0.15, realized rate=0.149, p-value=0.937; *TRW* stratum expected rate=0.28, realized rate=0.28, p-value=0.983; Figure 3). Based on these combined results, **INTERPRETATION NEEDED** evidence of temporal bias was found in 2016.

## 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., *TRW* 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 (Figure 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 HAL stratum

Given that there were 17 NMFS Areas fished in *HAL*, we would expect there to be 0.05 x 17 = 1 small data points for this stratum. There were 2. The percent of trips observed among NMFS Areas in this stratum ranged from 0% to 33.3% (median = 15.6%, Figure 5). The probability of these coverage rates or rates that deviated further from expected values is depicted in Figure 6.

### The POT stratum

Given that there were 9 NMFS Areas fished in *POT*, we would expect there to be 0.05 x 9 = 0 small data points for this stratum. There were 1. The percent of trips observed among NMFS Areas in this stratum ranged from 0% to 100% (median = 15.3%, Figure 7). The probability of these coverage rates or rates that deviated further from expected values is depicted in Figure 8.

### The TRW stratum

Given that there were 6 NMFS Areas fished in *TRW*, we would expect there to be 0.05 x 6 = 0 small data points for this stratum. There were 1. The percent of trips observed among NMFS Areas in this stratum ranged from 22.807% to 29.8% (median = 25.8%, Figure 9). The probability of these coverage rates or rates that deviated further from expected values is depicted in Figure 10.

## 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) #1 expressed as a percentage of the mean value of the metric being tested, 4) #2 expressed as a percentage of the mean value of the metric being tests, and 5) the p-value of the test, only #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 is 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 x 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 8.

Of the six metrics compared in the *HAL* stratum, 4 had low p-values. Observed trips in this stratum were 5.6% shorter in duration, occurred on vessels 3.9% longer in length, landed 6.3% more species, and landed catch that weighed 9.7% less than unobserved trips (Table 9).

Of the six metrics compared in the *POT* stratum, 1 had low p-values. Observed trips in this stratum landed 8.8% fewer species than unobserved trips (Table 9).

Of the six metrics compared in the *TRW* stratum, 5 had low p-values. Observed trips in this stratum occurred in 2.9% fewer areas, were 12.8% shorter in duration, landed 15.5% fewer species, landed catch that was 2.6% less diverse, and landed catch that weighed 10.1% less than unobserved trips (Table 9).

A visual depiction of individual results of this permutation test is given in Figure 11 for illustration purposes. **INTERPRETATION NEEDED**

### 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. Sample sizes for this test are presented in Table 10. *FLAG Should any strata omitted due to sample sizes?*

Of the six metrics compared in the *HAL* stratum, 1 had low p-values. Trips in this stratum that delivered to a tender landed catch that weighed 311.7% more than trips that did not deliver to a tender (Table 11).

Of the six metrics compared in the *POT* stratum, 4 had low p-values. Trips in this stratum that delivered to a tender were 28.5% longer in duration, occurred on vessels 16.2% shorter in length, landed 33.5% more species, and landed catch that weighed 33.2% more than trips that did not deliver to a tender (Table 11).

Of the six metrics compared in the *TRW* stratum, 5 had low p-values. Trips in this stratum that delivered to a tender occurred in 8.2% fewer areas, were 27.9% longer in duration, occurred on vessels 30.3% shorter in length, landed 17% fewer species, and landed catch that was 7.1% less diverse than trips that did not deliver to a tender (Table 11).

**INTERPRETATION NEEDED** Taken together, a tendering effect was evident within all gear types during 2015.

### 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 12. *FLAG Should any strata omitted due to sample sizes?*

Of the six metrics compared in the *POT* stratum, 1 had low p-values. Observed trips in this stratum that delivered to a tender landed 27.5% fewer species than unobserved trips that delivered to a tender (Table 13).

Of the six metrics compared in the *TRW* stratum, 3 had low p-values. Observed trips in this stratum that delivered to a tender were 87.9% shorter in duration, landed 15.9% fewer species, and landed catch that weighed 69.5% less than unobserved trips that delivered to a tender (Table 13).

From the above results, we conclude that there is **INTERPRETATION** evidence of an observer effect within trips that delivered to tenders in2015.

### 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 14.

Of the six metrics compared in the *HAL* stratum, 3 had low p-values. Observed non-tendered trips in this stratum were 5.6% shorter in duration, occurred on vessels 3.9% longer in length, and landed 6.2% more species than unobserved non-tendered trips (Table 15).

Of the six metrics compared in the *POT* stratum, there were no metrics with low p-values (Table 15).

Of the six metrics compared in the *TRW* stratum, 3 had low p-values. Observed non-tendered trips in this stratum occurred on vessels 3.2% longer in length, landed 14.2% fewer species, and landed catch that was 2.3% less diverse than unobserved non-tendered trips (Table 15).

**INTERPRETATION NEEDED** 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 2013 AR). 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 2013 AR; 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 12. 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 ten days. Whether this is due to an observer effect through intentional manipulation of trips (facillitated by the flexibility in ODDS and the current trip definitions), 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 (**FLAG: Maybe cite Fig. 2 from Gasper, Cahalan, and Mondragon 2014, Tech Memo 286**). 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 2016 was used to illustrate their combined effect on the probability of a NMFS Area containing observer data using the hypergeometric distribution (Figure 13). 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 3 partial coverage trip selection strata, the probability of having no observed trips in a NMFS Reporting Areas increases quickly above 0.05 when there are fewer than 18 trips in the *HAL* stratum, 18 trips in the *POT* stratum, 19 trips in the *POT* stratum, and 9 trips in the *TRW* stratum in a given area. 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.

# 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 2015 Annual Report made at the June 2016 Council meeting.

# Recommendations to Improve Data Quality

The Observer Science Committee made the following recommendations in its 2015 review of observer deployment to be considered in developing the 2017 ADP (Faunce et al. 2015, **citation**). Following each italicized recommendation is the outcome of that recommendation.

## Recommendations from the 2015 Annual Deployment Review

## OSC Recommendations to improve data quality and guide the 2018 ADP

# Acknowledgements

The 2016 Observer Science Committee was comprised of Craig Faunce (AFSC/FMA), Jason Gasper (AKRO/SF), Jennifer Cahalan (PSMFC), Steve Barbeaux (AFSC/REFM), Sandra Lowe (AFSC/REFM), and Ray Webster (IPHC).

# 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., Mondragon, J., and J. Gasper. 2014. Catch Sampling and Estimation in the Federal Groundfish Fisheries off Alaska: 2015 Edition. 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., Gasper, J., Cahalan, J., Lowe, S., Webster, R., and A'mar, T. 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., Cahalan, J., Gasper, J., A'mar, T., Lowe, S., Wallace, F. and Webster, R. 2014. Deployment performance review of the 2013 North Pacific Groundfish and Halibut Observer Program. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-281. 74 p.

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

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

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

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_earirirfa0311.pdf>.

Wall, J., R. French, and R. Nelson Jr. 1981. Foreign fisheries in the Gulf of Alaska, 1977-78. Mar. Fish. Rev. 43:20-35.

# TABLES

Table 1. Disposition of trips in the ODDS for 2016. "Waived and Not Selected" indicates that a trip inherited a selection and was later waived. "Paper" indicates that a trip was logged when the ODDS was not available.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Strata | Disposition | Logged | Cancelled by System | Trips remaining | Cancelled by User | Waived | Paper | Cancellation (%) | Waiver (%) |
| BSAI Cod 100% Voluntary Coverage | Selected | 141 | 0 | 141 | 0 | 0 | 0 | 0.00 | 0.00 |
| Declared Gear - Longline | Not Selected | 2394 | 377 | 2017 | 86 | 5 | 0 | 4.30 | 0.20 |
| Declared Gear - Longline | Selected | 452 | 1 | 451 | 108 | 14 | 0 | 23.90 | 3.10 |
| Declared Gear - Pot gear | Not Selected | 1141 | 29 | 1112 | 96 | 3 | 0 | 8.60 | 0.30 |
| Declared Gear - Pot gear | Selected | 190 | 0 | 190 | 48 | 6 | 0 | 25.30 | 3.20 |
| Declared Gear - Trawl | Not Selected | 2023 | 104 | 1919 | 76 | 0 | 0 | 4.00 | 0.00 |
| Declared Gear - Trawl | Selected | 802 | 0 | 802 | 127 | 0 | 0 | 15.80 | 0.00 |
| Total |  | 7143 | 511 | 6632 | 541 | 28 | 0 | 8.20 | 0.40 |

Table 2. Number of logged trips in each trip-selection strata (*TRW*, *HAL*, and *POT*) 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).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Strata | Total Trips | Randomly Selected (*r*) | Selected from Inherits (*i*) | Selected but Waived (*w*) | Total Final Selected (*T*=*r*+*i*-*w*) | % Selected from Inherits ((*i*/*T*)\*100) | % Waivers from Total Selected (*w*/(*T*+*w*)\*100) |
| Declared Gear - Longline | 2274 | 343 | 73 | 13 | 403 | 18.10 | 3.10 |
| Declared Gear - Pot gear | 1158 | 142 | 31 | 6 | 167 | 18.60 | 3.50 |
| Declared Gear - Trawl | 2518 | 675 | 69 | 0 | 744 | 9.30 | 0.00 |

Table 3. Number of logged trips in each trip-selection strata (*TRW*, *HAL*, and *POT*) 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).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Strata | Trip Disposition | Selected trips | Total Trips | Actual Selection (%) | Programmed Selection (%) | p-value (H0: Actual = Programmed) |
| Declared Gear - Longline | Initial Random Selection (*a*) | 452 | 2846 | 15.88 | 15.41 | 0.483 |
| Declared Gear - Longline | After Cancellations (*a*-*b*) | 343 | 2274 | 15.08 | 15.41 | 0.684 |
| Declared Gear - Longline | With Inherits (*a*-*b*+*c*) | 416 | 2274 | 18.29 | 15.41 | 0.000 |
| Declared Gear - Longline | With Inherits and Waivers (*a*-*b*+*c*-*d*) | 403 | 2274 | 17.72 | 15.41 | 0.003 |
| Declared Gear - Pot gear | Initial Random Selection (*a*) | 190 | 1331 | 14.27 | 15.24 | 0.341 |
| Declared Gear - Pot gear | After Cancellations (*a*-*b*) | 142 | 1158 | 12.26 | 15.24 | 0.004 |
| Declared Gear - Pot gear | With Inherits (*a*-*b*+*c*) | 173 | 1158 | 14.94 | 15.24 | 0.806 |
| Declared Gear - Pot gear | With Inherits and Waivers (*a*-*b*+*c*-*d*) | 167 | 1158 | 14.42 | 15.24 | 0.462 |
| Declared Gear - Trawl | Initial Random Selection (*a*) | 802 | 2825 | 28.39 | 28.31 | 0.933 |
| Declared Gear - Trawl | After Cancellations (*a*-*b*) | 675 | 2518 | 26.81 | 28.31 | 0.097 |
| Declared Gear - Trawl | With Inherits (*a*-*b*+*c*) | 744 | 2518 | 29.55 | 28.31 | 0.170 |
| Declared Gear - Trawl | With Inherits and Waivers (*a*-*b*+*c*-*d*) | 744 | 2518 | 29.55 | 28.31 | 0.170 |

Table 4. Number of total vessels (*V*), sampled vessels (*v*), total trips (*N*), sampled trips (*n*) for each observer deployment stratum in 2016. The number of vessels are not additive - total vessels are unique. EM: Electronic Monitoring.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Coverage | Strata | Deployment Method | Time Period | *V* | *v* | *N* | *n* | % Observed by Deployment Type | Expected Coverage (%) | Minimum Expected Coverage (%) | Maximum Expected Coverage (%) | Meets Expectations? |
| Full | Regulatory | Trip Selection | Year | 170 | 170 | 4573 | 4573 | 100 |  |  |  |  |
| Full | Voluntary | Trip Selection | Year | 23 | 23 | 137 | 136 | 99.3 |  |  |  |  |
| Full Coverage Total | NA | Trip Selection | Year | 172 | 172 | 4710 | 4709 | 100 |  |  |  |  |
| Partial | HAL | Trip Selection | Year | 465 | 244 | 2619 | 399 | 15.2 | 15 | 13.9 | 16.7 | Yes |
| Partial | POT | Trip Selection | Year | 113 | 73 | 1271 | 189 | 14.9 | 15 | 13 | 16.9 | Yes |
| Partial | TRW | Trip Selection | Year | 85 | 82 | 2738 | 767 | 28 | 28 | 26.3 | 29.7 | Yes |
| Gear-based Total | NA | Trip Selection | Year | 594 | 365 | 6628 | 1355 | 20.4 |  |  |  |  |
| NA | EM Voluntary 100% | Vessel Selection | Jan-Feb | 0 | 0 | 0 | 0 |  |  |  |  |  |
| Partial | EM Voluntary 100% | Vessel Selection | Mar-Jun | 7 | 4 | 26 | 16 | 57.1 |  |  |  |  |
| Partial | EM Voluntary 100% | Vessel Selection | Jul-Oct | 9 | 5 | 25 | 6 | 55.6 |  |  |  |  |
| Partial | EM Voluntary 100% | Vessel Selection | Nov-Dec | 1 | 0 | 10 | 0 | 0 |  |  |  |  |
| Partial | EM Voluntary 30% | Vessel Selection | Jan-Feb | 4 | 2 | 22 | 11 | 50 |  |  |  |  |
| Partial | EM Voluntary 30% | Vessel Selection | Mar-Jun | 28 | 10 | 94 | 33 | 35.7 |  |  |  |  |
| Partial | EM Voluntary 30% | Vessel Selection | Jul-Oct | 16 | 6 | 51 | 10 | 37.5 |  |  |  |  |
| NA | EM Voluntary 30% | Vessel Selection | Nov-Dec | 0 | 0 | 0 | 0 |  |  |  |  |  |
| EM Voluntary Total | NA | Vessel Selection | Year | 41 | 24 | 228 | 76 | 58.5 |  |  |  |  |
| Partial | Zero Coverage | Trip Selection | Year | 376 | 0 | 1596 | 0 | 0 |  |  |  |  |
| Partial | Zero Coverage EM Research | Trip Selection | Year | 3 | 0 | 30 | 0 | 0 |  |  |  |  |
| Zero Coverage Total | NA | Trip Selection | Year | 379 | 0 | 1626 | 0 | 0 |  |  |  |  |
| Total Fleet | Total |  |  | 1128 | 513 | 13192 | 6140 | 46.5% Trips; 45.5% Vessels |  |  |  |  |

Table 5. The number of vessels that fall under specific criteria within the EM Voluntary vessel selection strata.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Row | Metric | Jan-Feb | Mar-Jun | Jul-Oct | Nov-Dec |
| 1 | Anticipated to fish (Final 2016 EM Pre-Implementation Plan) | 3 | 38 | 36 | 2 |
| 2 | In selection frame (Notified will fish); F | 4 | 28 | 16 | 0 |
| 3 | In frame and fished; fY | 4 | 28 | 16 | 0 |
| 4 | In frame and did not fish; fN | 0 | 0 | 0 | 0 |
| 5 | Not in frame and fished (potential bias); f0 | 0 | 7 | 9 | 1 |
| 6 | Active (fished = true frame); f\* = f0 + fY | 4 | 35 | 25 | 1 |
| 7 | **Keep language?** Desired to be observed; vT | 2 | 11 | 8 | 1 |
| 8 | **NEW** Selected for coverage randomly (30%); vS\_30 | 2 | 10 | 6 | 0 |
| 9 | **NEW** Selected for coverage (100%); vS\_100 | 0 | 4 | 4 | 0 |
| 10 | Selected for coverage (Total); vS = vS\_30 + vS\_100 | 2 | 14 | 10 | 0 |
| 11 | Selected but did not fish; vN | 0 | 0 | 0 | 0 |
| 12 | Selected and fished; vF | 2 | 14 | 10 | 0 |
| 13 | Video data reviewed; v | 2 | 14 | 11 | 0 |

Table 6. Vessel-selection rates in EM Voluntary strata expressed as percetages (all rate formulations multiplied by 100). Abbreviations follow Table 5.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Row | Metric | Jan-Feb | Mar-Jun | Jul-Oct | Nov-Dec |
| 1 | Error in sampling frame due to over-coverage (% of sample frame); fN/F | 0.0 | 0.0 | 0.0 | 0.0 |
| 2 | Error in sampling frame due to under-coverage (% of true frame); f0/f\* | 0.0 | 20.0 | 36.0 | 100.0 |
| 3 | Error due to non-response: selected and did not fish; vN/vS | 0.0 | 0.0 | 0.0 | 0.0 |
| 4 | Error due to non-response: Selected, fished, and not observed (vF-v)/vF | 0.0 | 0.0 | -10.0 | 0.0 |
| 5 | **NEW** Chance of random selection if in frame and fished; vS\_30/fY | 50.0 | 35.7 | 37.5 | 0.0 |
| 6 | **NEW** Chance of selection if not in frame and fished; vS\_100/f0 | 0.0 | 57.1 | 44.4 | 0.0 |
| 7 | **Keep language?**Percent coverage desired; vT/f\* | 50.0 | 31.4 | 32.0 | 100.0 |
| 8 | Percent coverage acheived; v/f\* | 50.0 | 40.0 | 44.0 | 0.0 |

Table 7. 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 Non-tendered | p-value Trips Observed Non-tendered |
| Bering Sea | Full | Akutan | 751 | 751 | 100 | NA | 0 | 100 | NA |
| Bering Sea | Full | Dutch Hbr. | 806 | 806 | 100 | NA | 0 | 100 | NA |
| Bering Sea | Full | IFP | 339 | 339 | 100 | NA | 0 | 100 | NA |
| Bering Sea | Full | King Cove | 79 | 79 | 100 | NA | 0 | 100 | NA |
| Bering Sea | Full | Sand Point | 5 | 5 | 100 | NA | 0 | 100 | NA |
| Gulf of Alaska | Partial | Akutan | 158 | 47 | 30 | NA | 2 | 30 | NA |
| Gulf of Alaska | Partial | Dutch Hbr. | 7 | 4 | 57 | NA | 0 | 57 | NA |
| Gulf of Alaska | Partial | IFP | 29 | 2 | 7 | NA | 0 | 7 | NA |
| Gulf of Alaska | Partial | King Cove | 115 | 0 | 0 | NA | 97 | 0 | NA |
| Gulf of Alaska | Partial | Kodiak | 1097 | 315 | 29 | NA | 0 | 29 | NA |
| Gulf of Alaska | Partial | Sand Point | 540 | 58 | 11 | NA | 21 | 13 | NA |
| Total | Full |  | 1980 | 1980 | 100 | NA | 0 | 100 | NA |
| Total | Partial |  | 1945 | 426 | 22 | 0 | 12 | 25 | 0.003 |

Table 8. Number of trips by observation status in the 2016 trip-selection strata.

|  |  |  |
| --- | --- | --- |
| Strata | Observed | Unobserved |
| HAL | 399 | 2220 |
| POT | 189 | 1082 |
| TRW | 767 | 1971 |

Table 9. Results of permutation tests between observed and unobserved trips in the 2016 trip-selection strata. OD: Observed Difference.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Strata | Metric | NMFS Areas | Days Fished | Vessel Length | Species Landed | pMax Species | Landed Catch |
| HAL | Observed Difference | 0.019 | -0.289 | 2.141 | 0.233 | -0.002 | -0.688 |
| HAL | OD (%) | 1.699 | -5.617 | 3.888 | 6.255 | -0.249 | -9.697 |
| HAL | p-value | 0.371 | 0.043 | 0.000 | 0.049 | 0.758 | 0.027 |
| POT | Observed Difference | 0.016 | -0.108 | 0.183 | -0.168 | 0.002 | -2.455 |
| POT | OD (%) | 1.526 | -2.969 | 0.261 | -8.788 | 0.212 | -7.784 |
| POT | p-value | 0.410 | 0.452 | 0.919 | 0.022 | 0.204 | 0.353 |
| TRW | Observed Difference | -0.032 | -0.344 | 0.316 | -0.844 | 0.024 | -8.383 |
| TRW | OD (%) | -2.875 | -12.816 | 0.388 | -15.491 | 2.566 | -10.119 |
| TRW | p-value | 0.020 | 0.000 | 0.691 | 0.000 | 0.000 | 0.001 |

Table 10. Number of trips by tendered status in the 2016 trip-selection strata.

|  |  |  |
| --- | --- | --- |
| Strata | Tendered | Non-tendered |
| HAL | 3 | 2616 |
| POT | 142 | 1129 |
| TRW | 272 | 2466 |

Table 11. Results of permutation tests between tendered and non-tendered trips in the 2016 trip-selection strata. OD: Observed Difference.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Strata | Metric | NMFS Areas | Days Fished | Vessel Length | Species Landed | pMax Species | Landed Catch |
| HAL | Observed Difference | -0.139 | 0.864 | 2.936 | -1.067 | 0.060 | 22.126 |
| HAL | OD (%) | -12.216 | 16.826 | 5.332 | -28.578 | 6.864 | 311.694 |
| HAL | p-value | 1.000 | 0.530 | 0.655 | 0.432 | 0.457 | 0.005 |
| POT | Observed Difference | -0.007 | 1.036 | -11.365 | 0.642 | -0.001 | 10.467 |
| POT | OD (%) | -0.617 | 28.458 | -16.161 | 33.500 | -0.102 | 33.179 |
| POT | p-value | 0.850 | 0.000 | 0.000 | 0.000 | 0.592 | 0.000 |
| TRW | Observed Difference | -0.090 | 0.748 | -24.651 | -0.927 | 0.067 | 2.340 |
| TRW | OD (%) | -8.206 | 27.864 | -30.315 | -17.031 | 7.124 | 2.824 |
| TRW | p-value | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.490 |

Table 12. Number of tendered trips by observation status in the 2016 trip-selection strata.

|  |  |  |
| --- | --- | --- |
| Strata | Observed | Unobserved |
| HAL | 0 | 3 |
| POT | 18 | 124 |
| TRW | 122 | 150 |

Table 13. Results of permutation tests between observed and unobserved tendered trips in the 2016 trip-selection strata. OD: Observed Difference.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Strata | Metric | NMFS Areas | Days Fished | Vessel Length | Species Landed | pMax Species | Landed Catch |
| HAL | Observed Difference | NaN | NaN | NaN | NaN | NaN | NaN |
| HAL | OD (%) | NaN | NaN | NaN | NaN | NaN | NaN |
| HAL | p-value | NA | NA | NA | NA | NA | NA |
| POT | Observed Difference | 0.071 | -1.026 | -4.584 | -0.684 | 0.008 | -25.033 |
| POT | OD (%) | 6.746 | -22.485 | -7.612 | -27.503 | 0.768 | -61.291 |
| POT | p-value | 0.209 | 0.253 | 0.182 | 0.040 | 0.133 | 0.057 |
| TRW | Observed Difference | 0.003 | -2.949 | 0.492 | -0.735 | 0.000 | -59.025 |
| TRW | OD (%) | 0.302 | -87.884 | 0.832 | -15.944 | 0.021 | -69.482 |
| TRW | p-value | 1.000 | 0.000 | 0.531 | 0.000 | 0.861 | 0.000 |

Table 14. Number of non-tendered trips by observation status in the 2016 trip-selection strata.

|  |  |  |
| --- | --- | --- |
| Strata | Observed | Unobserved |
| HAL | 399 | 2217 |
| POT | 171 | 958 |
| TRW | 645 | 1821 |

Table 15. Results of permutation tests between observed and unobserved non-tendered trips in the 2016 trip-selection strata. OD: Observed Difference.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Strata | Metric | NMFS Areas | Days Fished | Vessel Length | Species Landed | pMax Species | Landed Catch |
| HAL | Observed Difference | 0.019 | -0.287 | 2.146 | 0.232 | -0.002 | -0.659 |
| HAL | OD (%) | 1.683 | -5.596 | 3.897 | 6.215 | -0.240 | -9.311 |
| HAL | p-value | 0.387 | 0.045 | 0.001 | 0.043 | 0.773 | 0.054 |
| POT | Observed Difference | 0.010 | 0.013 | 0.456 | -0.099 | 0.001 | 0.214 |
| POT | OD (%) | 0.952 | 0.382 | 0.637 | -5.360 | 0.150 | 0.704 |
| POT | p-value | 0.706 | 0.905 | 0.798 | 0.186 | 0.417 | 0.930 |
| TRW | Observed Difference | -0.028 | -0.054 | 2.667 | -0.785 | 0.021 | -1.605 |
| TRW | OD (%) | -2.562 | -2.064 | 3.184 | -14.173 | 2.302 | -1.943 |
| TRW | p-value | 0.055 | 0.438 | 0.002 | 0.000 | 0.001 | 0.446 |

# FIGURES

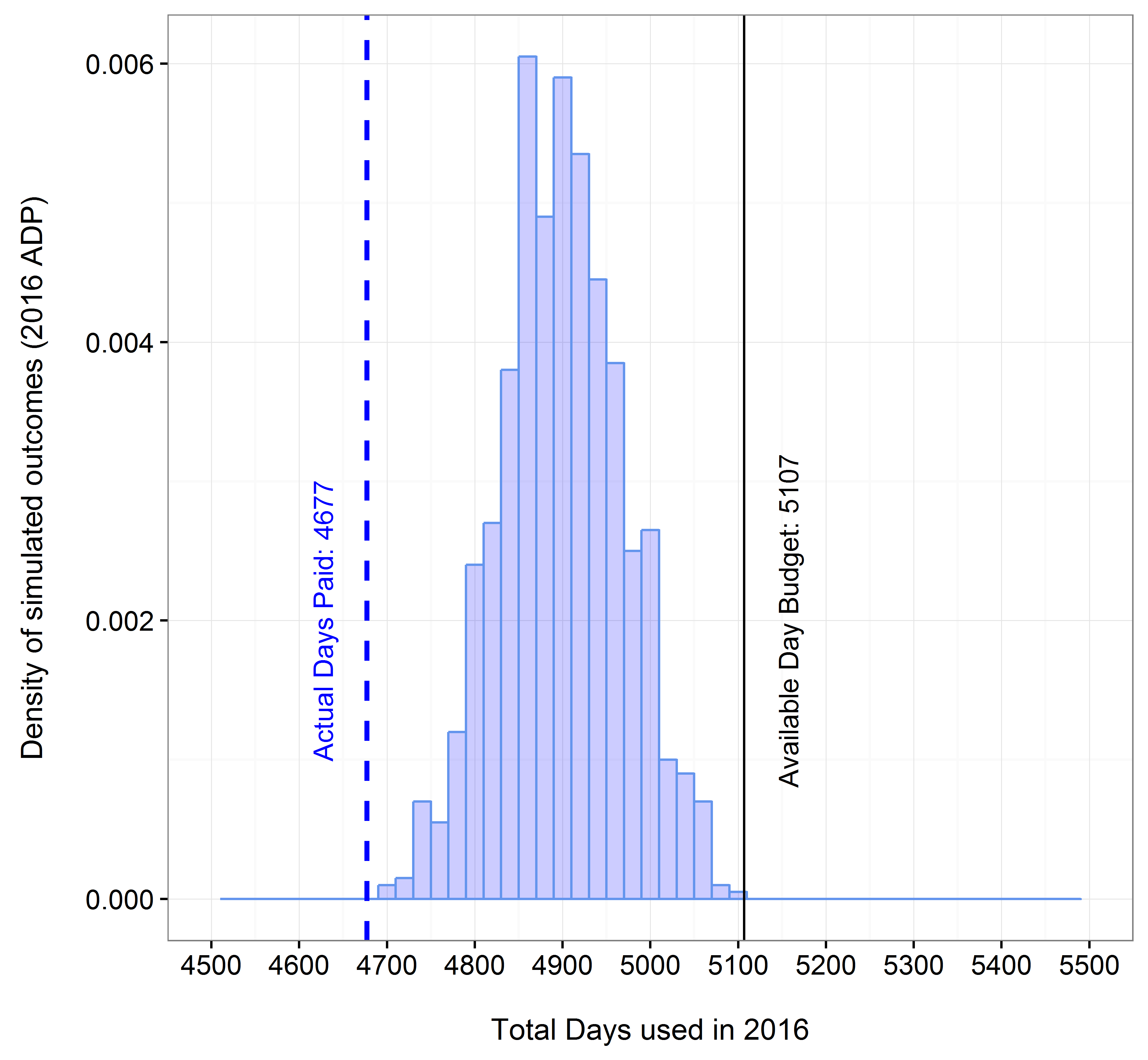


Figure 1. Actual paid sea-days in 2016 (dotted line) in relation to the range of potential budgetary outcomes estimated in December 2015 for the Final 2016 Annual Deployment Plan (vertical bars).

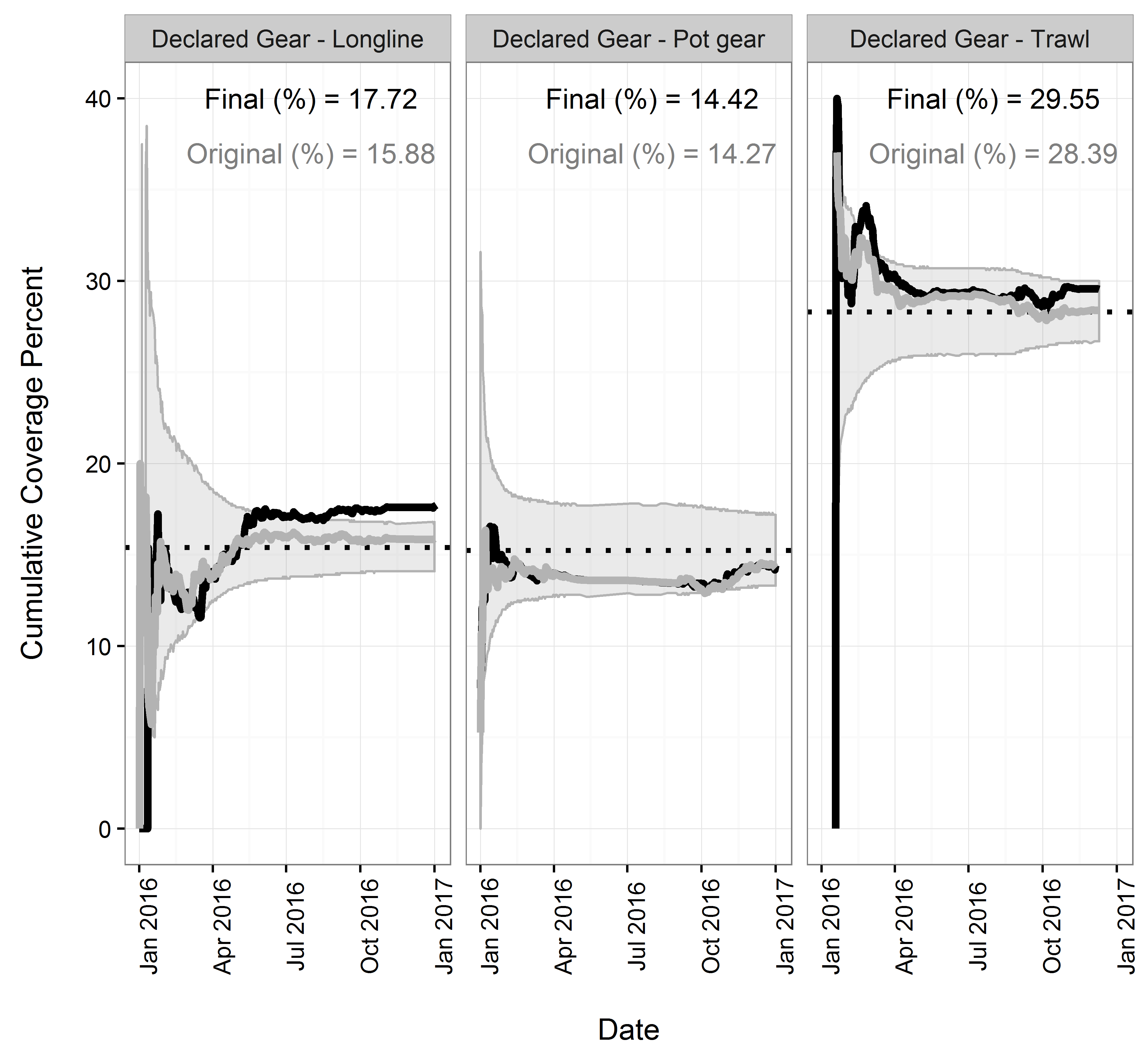


Figure 2. Rate of selected trips logged into ODDS organized by original date entered for all trips (grey line and grey text), and final date considering only non-cancelled trips (black line and black text). The programmed selection rate is depicted as the dotted line. Grey shaded areas denote the range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

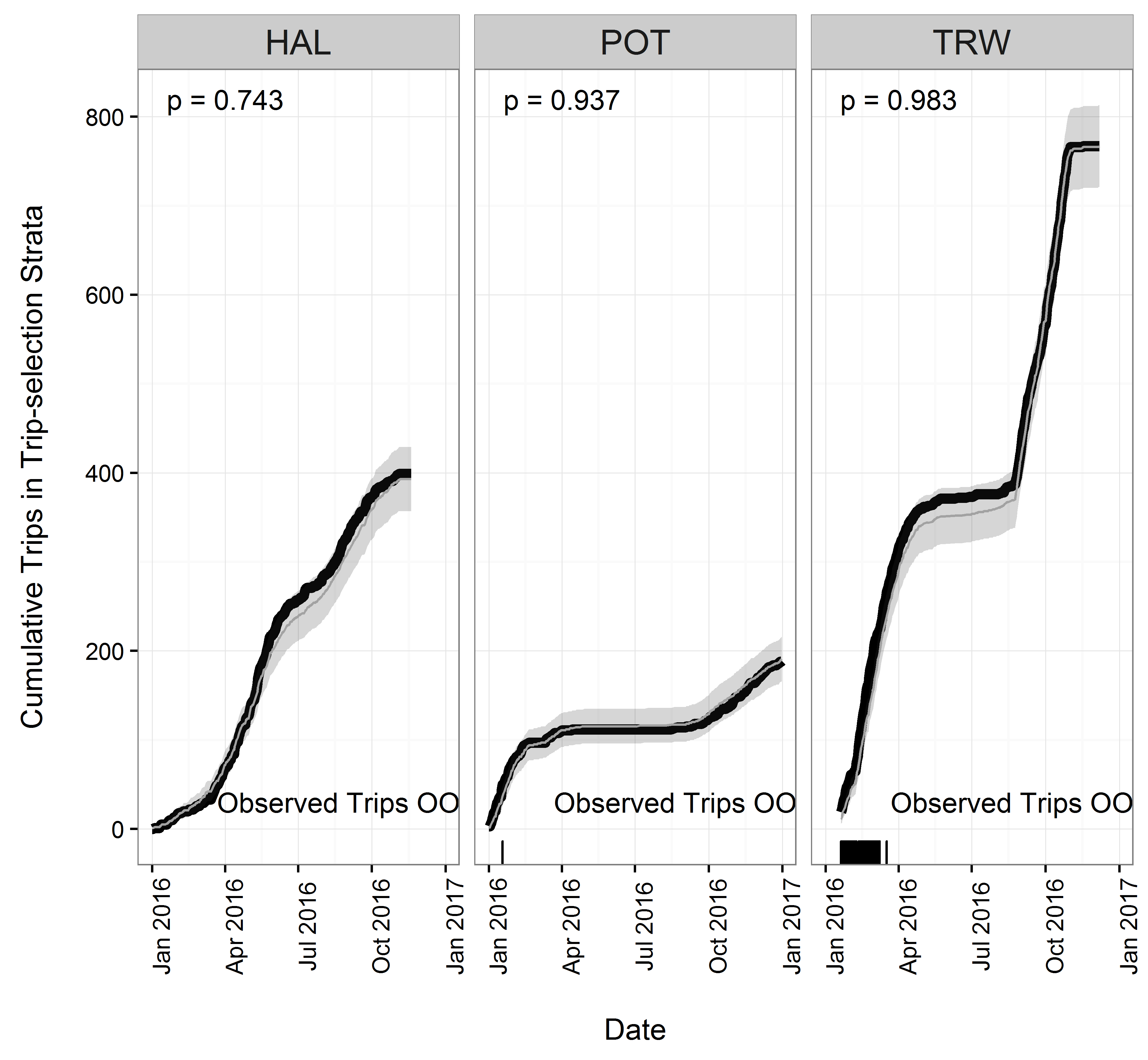


Figure 3. Cumulative number of trips observed during 2016 (black line) compared to the expected range of observed trips (shaded area) given fishing effort and sampling rates. Dates where the observed number of trips is outside of expected (less or more than the range; OOE) are depicted as tick marks 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.

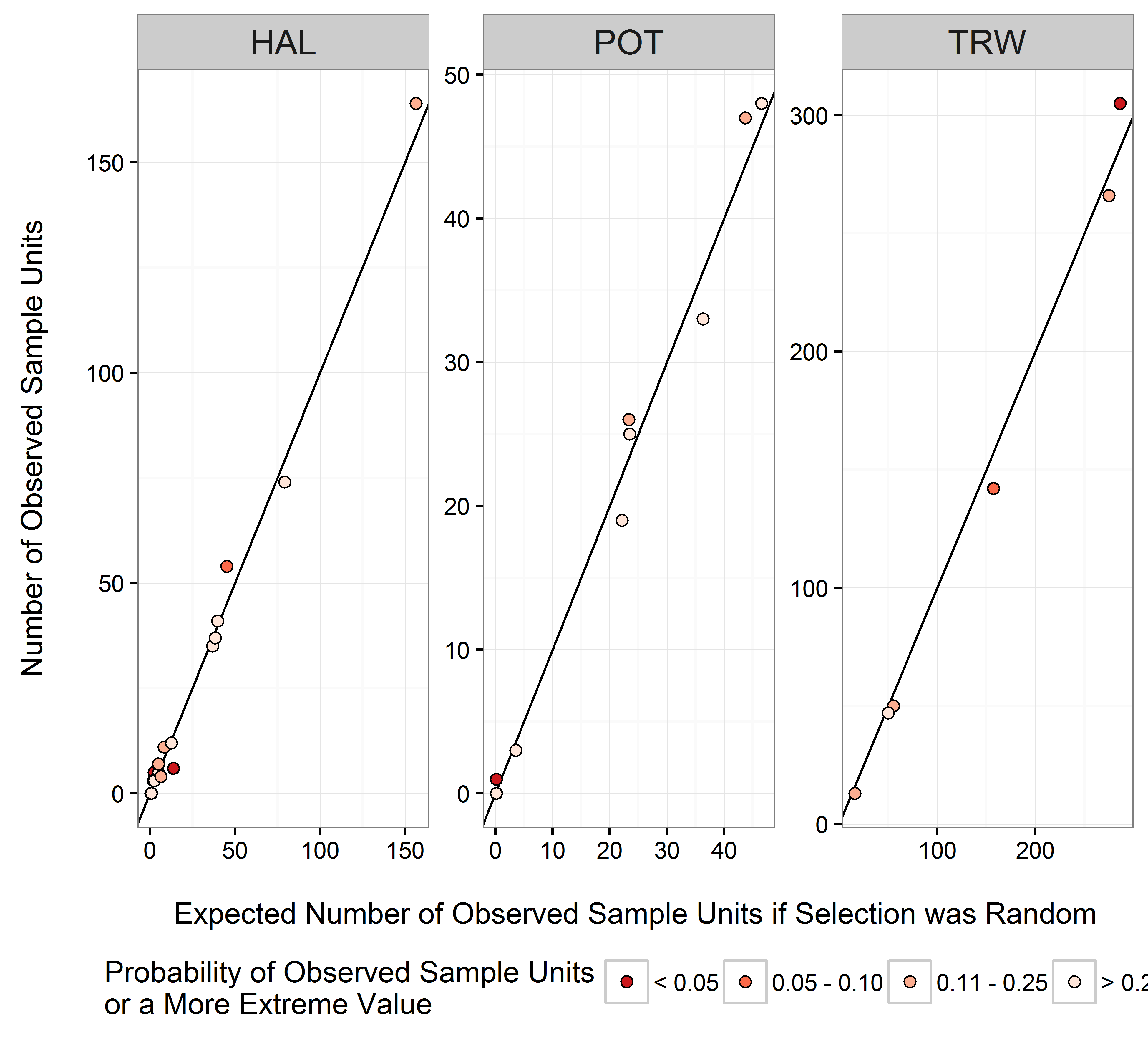


Figure 4. Comparison plots depicting the number of observed sample units compared to the number of expected observed sample units for each partial coverage stratum. Each point on a plot represents a NMFS Area. The smaller the point, the more unusual the result.

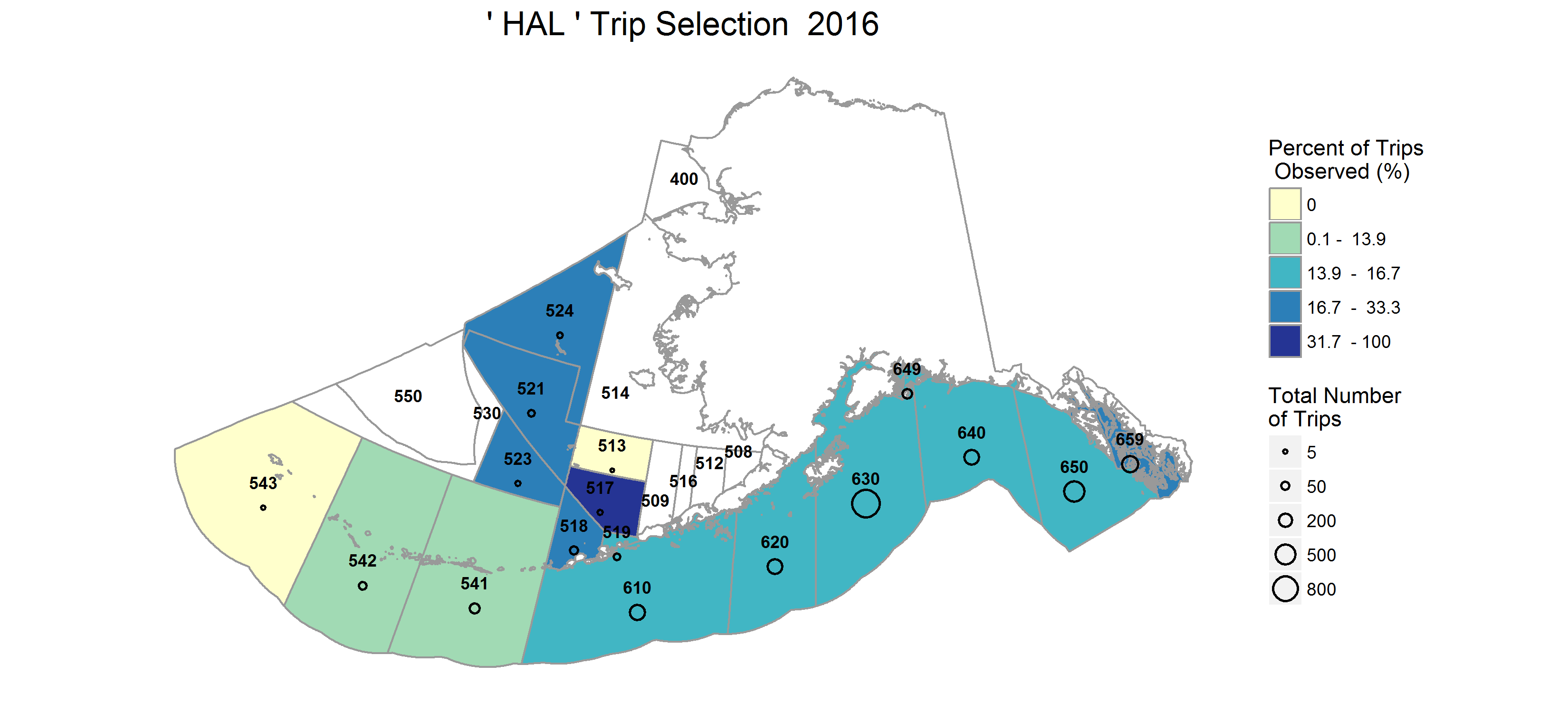


Figure 5. Percent of trips observed by NMFS Reporting Area in the *HAL* stratum.

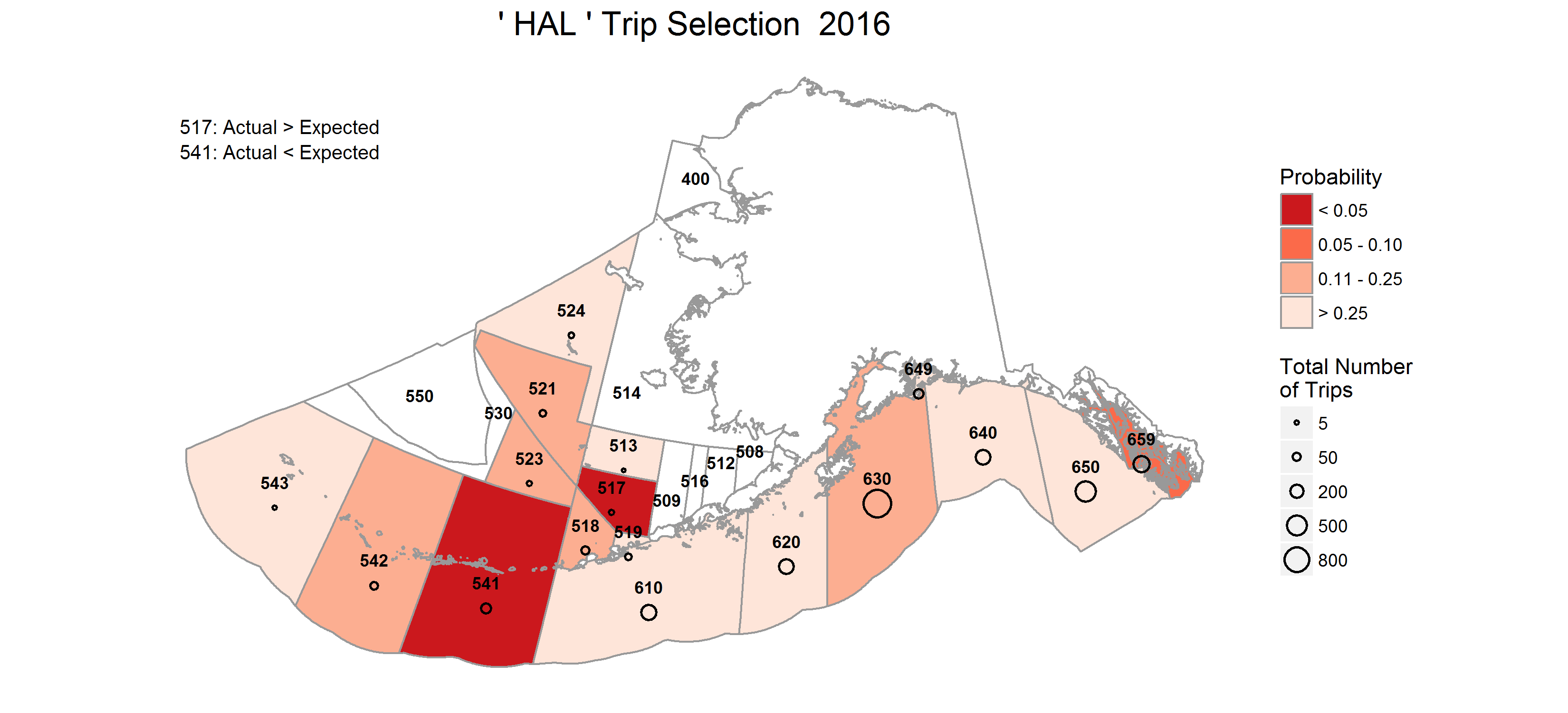


Figure 6. Probability of observing the realized or more extreme outcome (coverage rate) in a NMFS Reporting Area in the *HAL* stratum. Reporting Areas where unlikely outcomes occurred are shaded in darker colors.

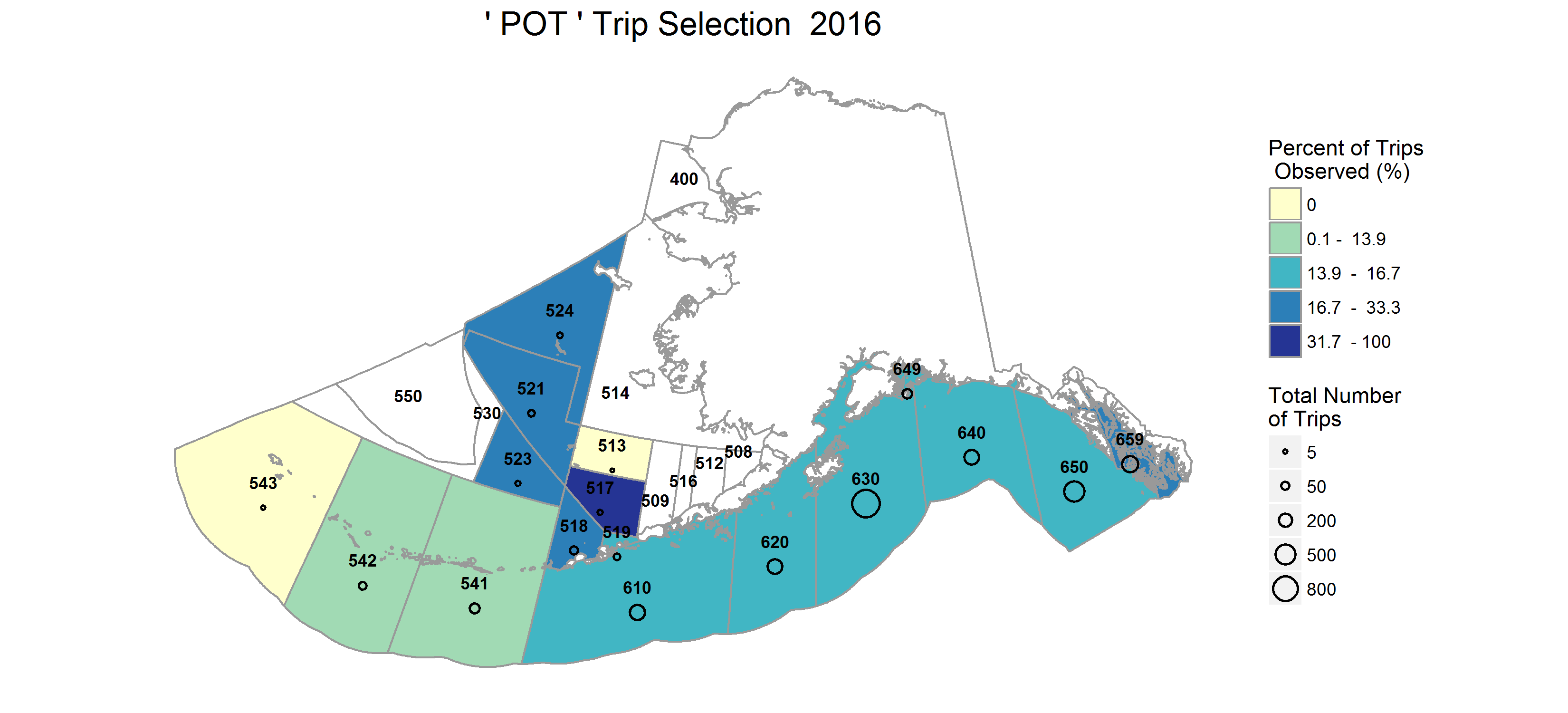


Figure 7. Percent of trips observed by NMFS Reporting Area in the *POT* stratum.

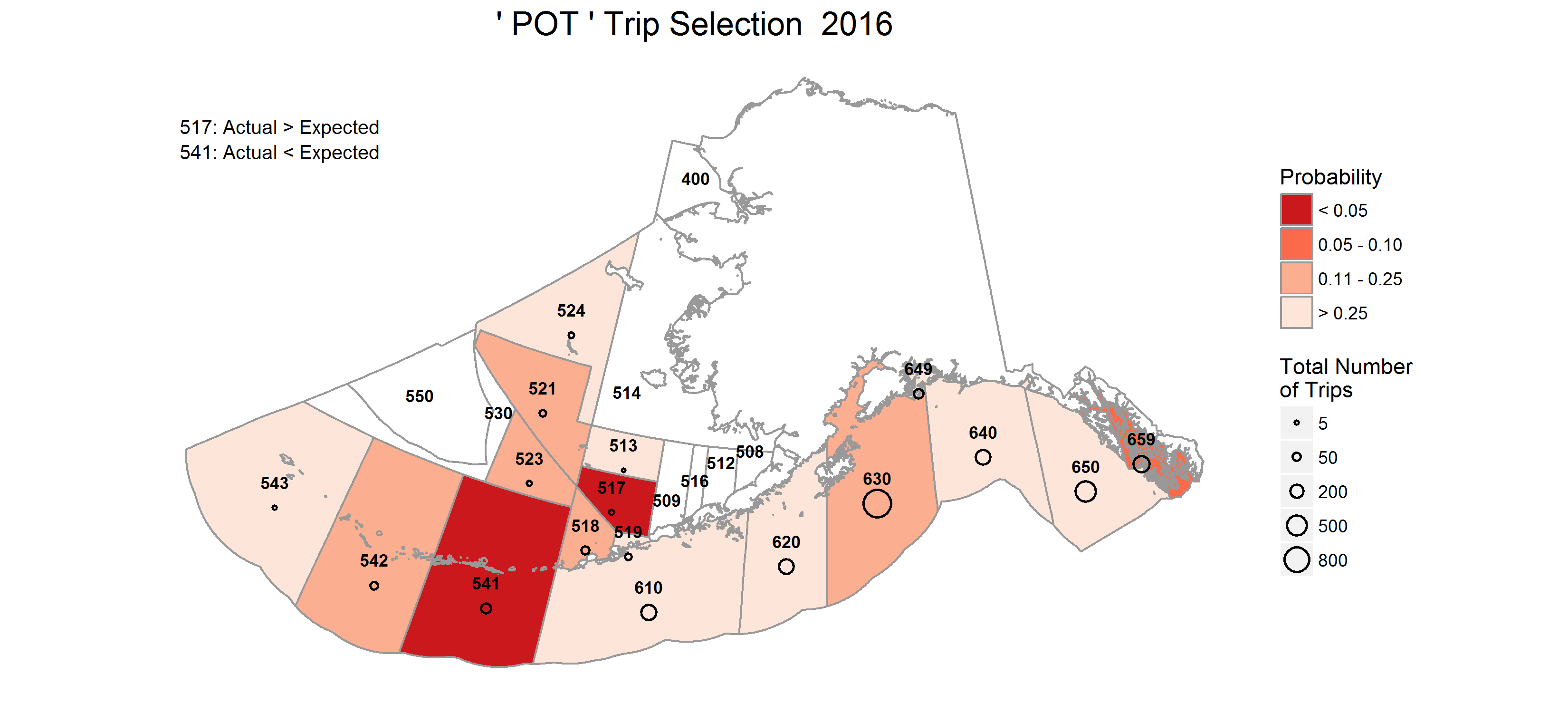


Figure 8. Probability of observing the realized or more extreme outcome (coverage rate) in a NMFS Reporting Area in the *POT* stratum. Reporting Areas where unlikely outcomes occurred are shaded in darker colors.

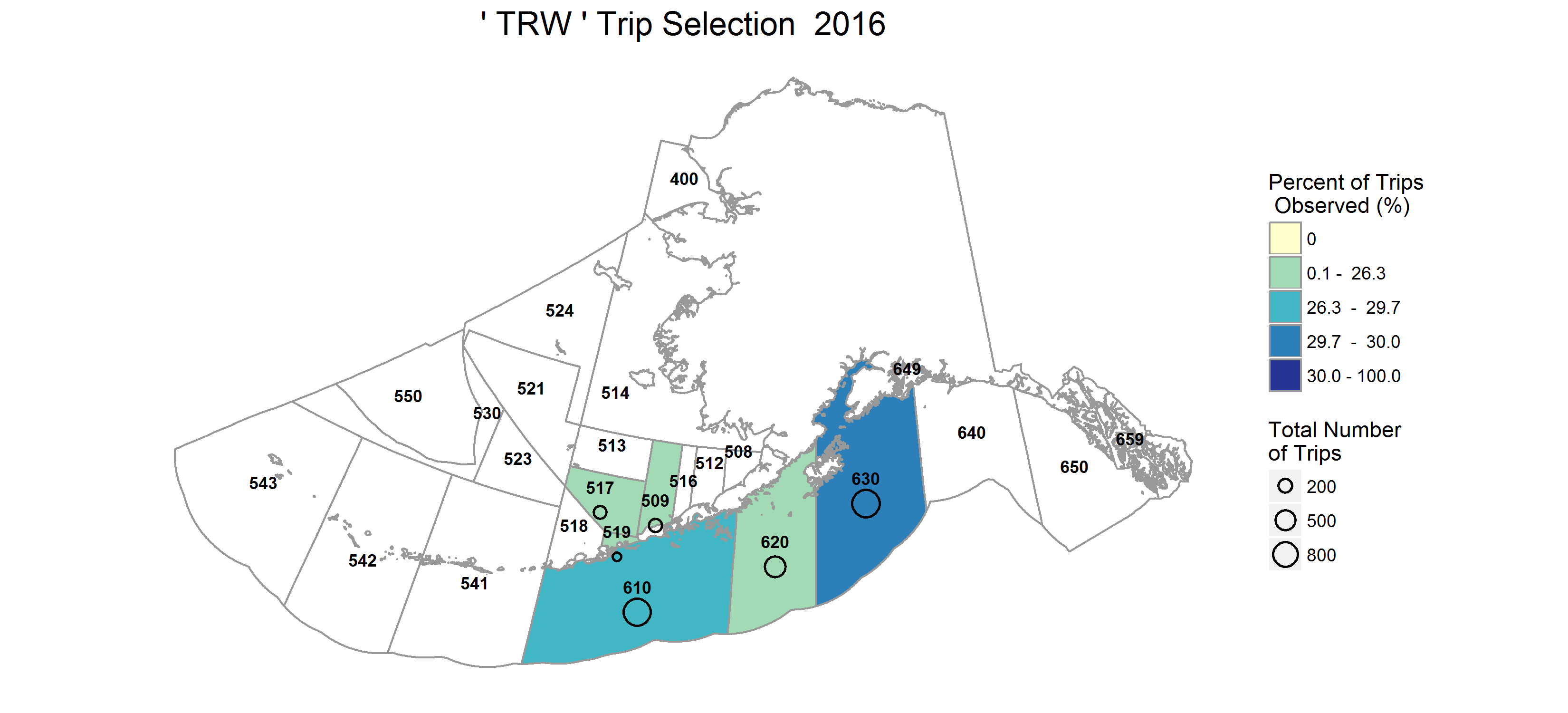


Figure 9. Percent of trips observed by NMFS Reporting Area in the *TRW* stratum.

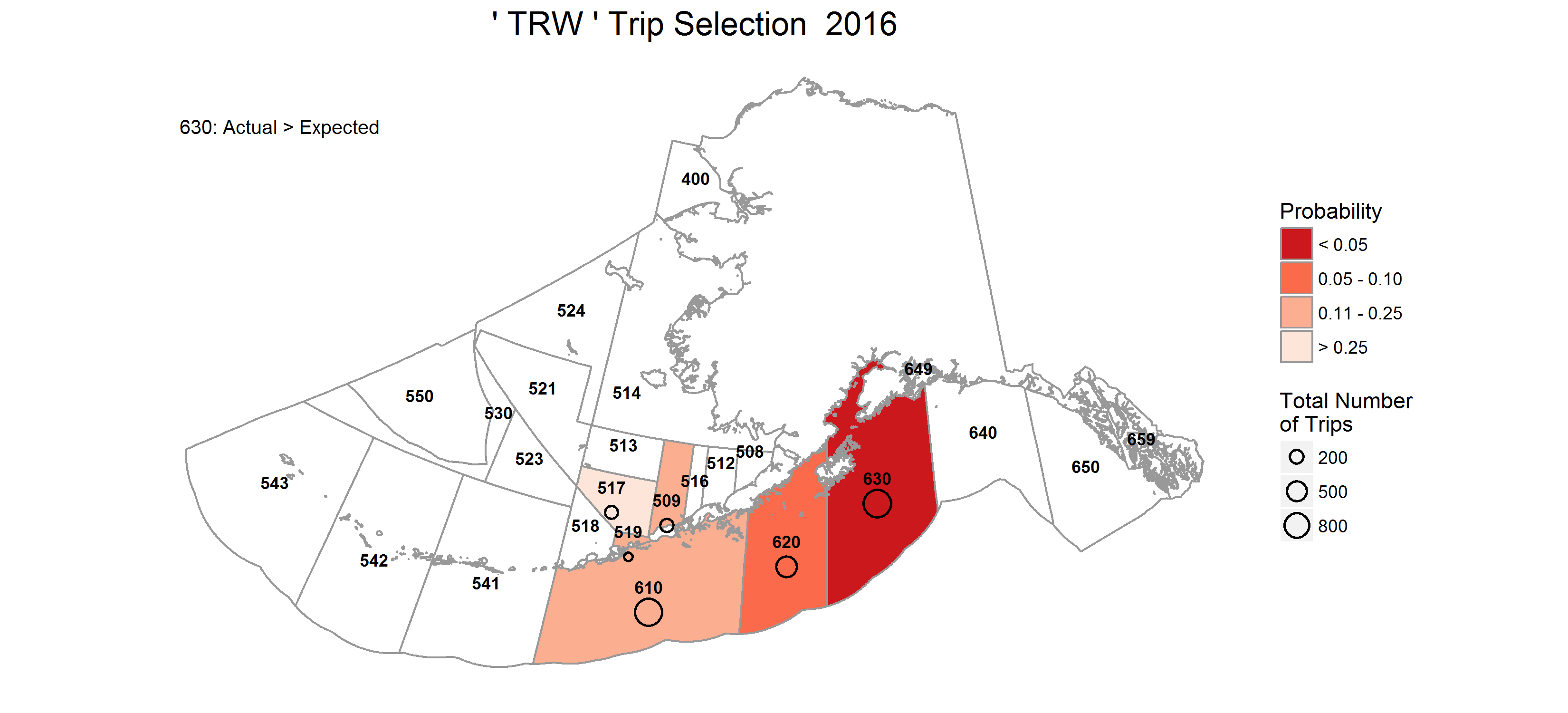


Figure 10. Probability of observing the realized or more extreme outcome (coverage rate) in a NMFS Reporting Area in the *TRW* stratum. Reporting Areas where unlikely outcomes occurred are shaded in darker colors.

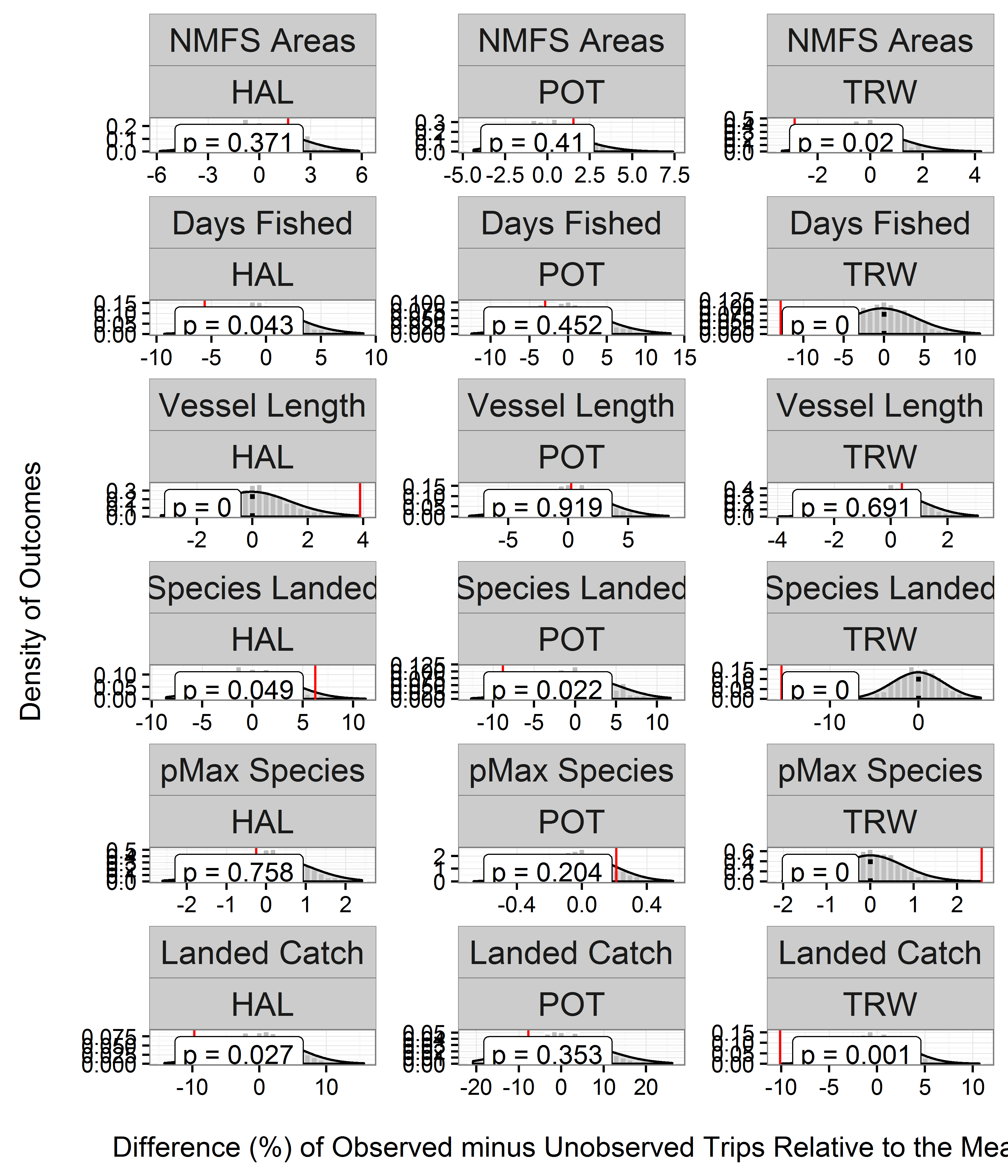


Figure 11. Results of permutaion tests between observed and unobserved trips for each strata in the partial coverage category of the 2016 ADP. In each panel, the grey bars depict the distribution of differences between observed and unobserved trips where the assignment of observed status has been randomized (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.

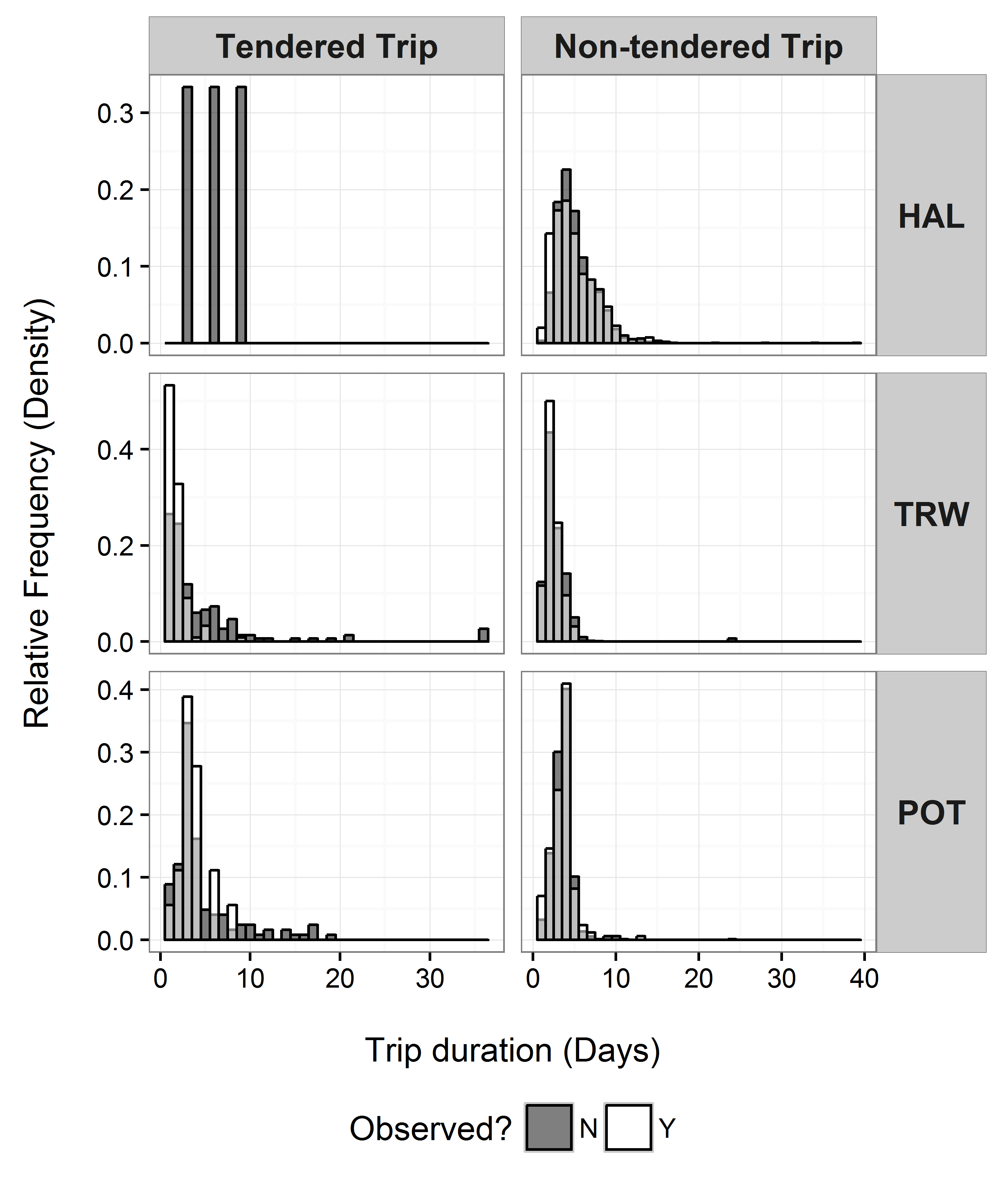


Figure 12. Distribution of trip durations for vessels in the partial coverage category by gear and observation status. Observed trips are depicted as transparent white bars overtop of solid black bars for unobserved trips. Trip durations where both observed and unobserved status exist are depicted in gray (This is not the same as a stacked bar chart, in which the height of the bar would reflect observed and unobserved on top of one another- this plot has each observation status in front of the other).

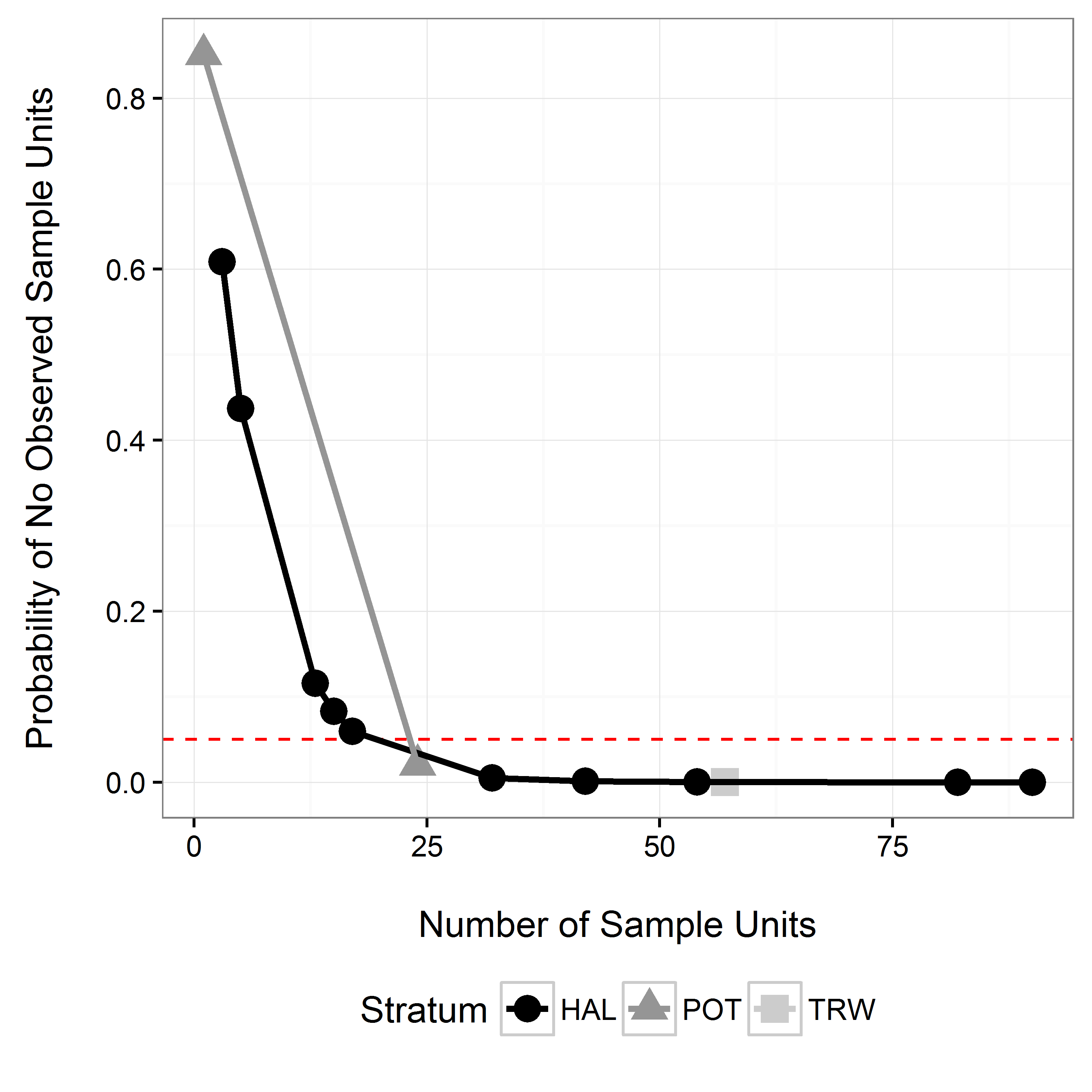


Figure 13. Probability of selecting a sample and observing no trips as a function of the number of sample units and selection rate that occurred in a NMFS Area, time period, and stratum. The x-axis has been truncated to increase resolution at smaller numbers of sampling units. The likelihood of having no observer data decreases with increasing total fishing effort and selection rate. The selection rate is 28% in the *TRW* stratum, 15% in the *HAL* stratum, and 15% in the *POT* stratum.