NOAA Technical Memorandum NMFS-AFSC-307
doi:10.7289/V5/TM-AFSC-307

An Initial Analysis of Alternative Sample Designs for the Deployment of Observers in Alaska

C. H. Faunce

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

November 2015

NOAA Technical Memorandum NMFS

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

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

This document should be cited as follows:

Faunce, C. H. 2015. 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. doi:10.7289/V5/TM-AFSC-307.

Document available: http://www.afsc.noaa.gov/Publications/AFSC TM/NOAA TM AFSC 307.pdf

NOAA Technical Memorandum NMFS-AFSC-307
doi:10.7289/V5/TM-AFSC-307

An Initial Analysis of Alternative Sample Designs for the Deployment of Observers in Alaska

by
C. H. Faunce

[1]Alaska Fisheries Science Center
Fisheries Monitoring and Analysis Division
7600 Sand Point Way N.E.
Seattle, WA 98115

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

November 2014

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

Changes in regulation enacted in 2013 have enabled the Alaska Fisheries Science Center's Fishery Monitoring and Analysis Division (FMA) and Alaska Regional Office's Sustainable Fisheries Division to work collaboratively on an Annual Deployment Plan (ADP). Each ADP documents how the National Marine Fisheries Service (NMFS) plans to deploy observers into fishing activities for the coming year under the limits of available funding. Draft ADPs are presented to the North Pacific Fishery Management Council (Council) during September - October and are finalized in December. The sampling design for observer deployment has two elements: how the population is subdivided (i.e., stratification schemes) and how available samples are allocated (i.e., allocation strategies).

Here the relative performance of 12 alternative sampling designs for the deployment of observers into at-sea operations are compared in support of the 2016 draft ADP. These designs were defined by combinations of six potential stratifications and whether allocations were proportional to effort (defined by trips) or to reduce the weighted variance of total groundfish retained and discarded ("optimal") allocation. Performance metrics included a gap analyses as well as single-stage estimates of precision and accuracy. These three metrics were combined into a single score for design comparison. Gap analysis scores were used in a "hurdle-model" approach for final design evaluations. The four designs with above-average gap analyses and final scores were proposed for consideration in the draft 2016 ADP; two of these stratified by gear (Trawl, Hook and Line, and Pot). Preliminary anticipated coverage rates were generated for the four candidate designs. For the draft 2016 ADP, the NMFS proposed that the sample design for the observer deployment be defined in units of trips stratified by three gear types with sample sizes allocated according to a blended optimal allocation strategy. At their October 2016 meeting, the Council supported this design for the final 2016 ADP.

CONTENTS

INTRODUCTION

The North Pacific Groundfish and Halibut Observer Program (observer program) uses a hierarchical sampling design with randomization at all levels to achieve unbiased data from fishing operations in the region. The fishing trip represents the primary sampling unit of this design. Since 2013, fleet operations in Alaska have been divided into two portions; vessels and shore-based industry operations that are subject to complete observation at the level of the trip or delivery are termed "full-coverage" while the remainder are termed "partial-coverage". Definitions of full- and partial coverage are set in Federal Regulations.

Observer deployment hereafter refers to how trips and deliveries are selected for observer coverage in the partial-coverage category of the Alaska fishing industry. All fishing trips subject to partial observer coverage constitute the target population for observer deployment. A sampling frame for the deployment of observers is constructed though the use of a mandatory log-in system known as the Observer Declare and Deploy System (ODDS)[1].

Since 2013, the observer program has been required to provide an Annual Report and an Annual Deployment Plan (ADP) to the North Pacific Fishery Management Council (Council). The Annual Report is presented in June and contains information on how well various aspects of the observer program are performing in addition to recommendations for future ADPs. The draft and final ADP are presented in September and December, respectively, and describe the observer deployment for the coming year. Three separate advisory bodies provide their comments and perspectives to the Council at each meeting. These include the Observer Advisory Committee, the Advisory Panel, and the Science and Statistical Committee (SSC). Members on the Observer Advisory Committee and the Advisory Panel represent major segments of the fishing industry as well as observers, consumers, environmental/conservation, and sport fishermen.

---

[1] http://odds.afsc.noaa.gov

Science and Statistical Committee members are scientists with expertise in biology, economics, statistics, and social science.

Partial coverage observers are trained prior and debriefed after their respective deployments by the observer program. Observers are employees of an observer provider company who is responsible for the logistical aspects of deployment. Funds to deploy observers in partial coverage are obtained by NMFS through a landings fee, and these funds are contracted to the observer provider company. The Council has the authority to change the fee up to a maximum of 2% of landed value. The fee currently stands at 1.25% and is scheduled to be re-assessed in 2018.

Concerns over the costs of the observer program and resulting data quality has led to scrutiny, even legal challenge of observer deployment[2]. The ADP process provides a mechanism for NMFS and the Council to re-evaluate deployment and improve efficiency in the sampling design. In the most recent Annual Report (NMFS 2015a), the NMFS recommended that future ADPs explore alternative ways to subdivide the population of partial coverage trips. The corresponding SSC report added that such an endeavor will require estimates of uncertainty and likely involve tradeoffs in quality among the multiple measures produced by the observer program (NPFMC 2015a).

What follows is a comparison of alternative stratum definitions and sample allocations for the deployment of observers into the fleet of vessels in partial coverage. These analyses are performed in support of the 2016 ADP following the findings and recommendations contained in the 2014 Annual Report and the SSC response to those findings.

---

[2] *The Boat Company and the Fixed Gear Alliance v. P. Pritzker*. 2014. U.S. District Court for the District of Alaska-Cross-Motions for Summary Judgment. Case No. 3:12-cv-0250-HRH.

METHODS

The sampling design for observer deployment involves two elements; how the population of partial coverage trips are subdivided, and what proportion of the total observer deployments are to occur within these subdivisions. The first of these is termed *stratification*, while the second is termed *allocation*.

Stratification Schemes

Stratification is the division of sample units in the population into subpopulations. The subpopulations are individually called stratum (strata if plural). Stratified random sampling is the act of obtaining independently random samples from within each stratum in the population. For this reason, strata need to be defined based on criteria known prior to the draw of the sample. This means that elements of fishing trips known prior to departure are valuable in defining deployment strata, whereas catch is not.

There are numerous reasons for creating strata. These include the following: when a separate estimate for a subpopulation is desired, when administrative convenience (field logistics) permits it, and to increase the precision of sample-based estimates of the total. Increased precision is accomplished through the division of a heterogeneous population into homogenous subpopulations since the variance in the population total is dependent on the variances of the individual stratum means (Cochran 1977).

The collection of strata that together subdivide the population of trips in partial coverage constitutes a *stratification scheme*. In this study six stratification schemes were considered. These stratification schemes (with numbered lists of the individual strata) are as follows:

1. STRATA 2010

This stratification scheme was influenced by the classification and regression tree (CART) model performed on total retained groundfish that was conducted prior to restructure of the observer program

(NPFMC et al. 2011). Strata are as follows:

1) Trawl gear.

2) Fixed gear (Hook and Line and Pot) ≥ 57.5' LOA.

3) Fixed gear 40 - 57.5' LOA.

2. STRATA 1315

This stratification scheme has been employed by the observer program during 2013-2015. Strata are as follows:

1) Fixed and trawl gear ≥ 57.5' LOA (a.k.a. the “T” stratum of 2015).

2) Fixed and trawl gear 40 - 57.5' LOA (a.k.a the “t” stratum of 2015).

3. STRATA 16

This stratification is identical to STRATA 1315 with the modification that proposed changes to the full and partial coverage category of the fleet occur in 2016 following NPFMC (2015b). These proposed changes include the following: 1) small catcher processors (CPs) that were in full coverage be placed into the partial coverage category, 2) some “AFA” trawl catcher vessels when fishing in the Bering Sea and Aleutian Islands voluntarily choose to belong to the full coverage category, and 3) vessels selected by the EM Workgroup of the Council for 2016 are removed from the partial coverage category. How data were prepared to accommodate these changes are provided later in the subheading “data preparation.”

4. STRATA Gear

This stratification uses the partial coverage definitions in STRATA 2016 but divides the trips into three strata:

1) Hook and Line ≥ 40' LOA.

2) Pot ≥ 40' LOA.

3) Trawl.

5. STRATA GFMP2

This stratification uses the partial coverage definitions in STRATA 2016 and builds on the stratifications in STRATA Gear. Gear-based stratifications are additionally divided by whether they occur in the Fishery Management Plan (FMP) of the Gulf of Alaska (GOA) or the Bering Sea and Aleutian Islands (BSAI). Strata are as follows:

1) Hook and Line BSAI ≥ 40' LOA.

2) Hook and Line GOA ≥ 40' LOA.

3) Pot BSAI ≥ 40' LOA.

4) Pot GOA ≥ 40' LOA.

5) Trawl BSAI.

6) Trawl GOA.

An alternative to this stratification that further subdivided the BSAI into separate Bering Sea and Aleutian Islands portions was also investigated. However, this nine strata scheme was abandoned since two strata had less than 20 trips in several strata meaning that there was very little chance of obtaining even a single observed trip through random sampling.

6. STRATA HALFYR

This stratification uses the partial coverage definitions in STRATA 2016 but is structured to allow maximum flexibility in setting observer coverage rates by the observer program. The observer partial coverage contract period is offset from the calendar year by 6 months. Days purchased for one 12-month contract option period cannot easily be transferred to another 12 month contract option period. However, fishing effort is related to available quota which is set by the calendar year. Given that 1) the fishing effort and available contract days can differ for each 6-month period of the calendar year, and 2) the prosecution of the fisheries in fixed gear is largely divided into months 1-3 and 9-11, this stratification scheme has five strata:

1) Hook and Line ≥ 40' LOA first half of the year (First).

2) Hook and Line ≥ 40' LOA second half of the year (Second).

3) Pot ≥ 40' LOA First.

4) Pot ≥ 40' LOA Second.

5) Trawl gear.

The stratification schemes 1-6 can be thought of as a continuum. The first scheme represents a simplified view of the major divisions in the fleet while the second scheme represents what was first implemented by NMFS and the Council with Amendment 86 to the BSAI FMP and Amendment 76 to the GOA FMP. Comparisons between the second and the third stratification schemes should represent the relative "impact" of anticipated changes to the partial coverage and full coverage categories scheduled for the 2016 ADP as the result of actions scheduled by the Council. The relative "impact" of alternative definitions of the partial coverage category of the fleet for 2016 can be measured by independently comparing the third stratification scheme to the remaining stratification schemes.

Sample Allocation

Sample allocation is the term for how available observer deployments are apportioned to strata. "Optimal" allocation is that which achieves the most precision for the least cost (*c*). If *n* is the number of observed trips afforded for the year among all partial coverage fishing trips (*N*) that occur within *H* strata, and the estimate of catch from these trips has S^2 variance, the number of samples that is considered optimum for each stratum (n_h) is denoted by equation 1,

$$n_h = n * \frac{N_h S_h / \sqrt{c_h}}{\sum_{h=1}^{H} \left(N_h S_h / \sqrt{c_h} \right)} \qquad \text{(Cochran 1977).} \qquad \text{Eq. (1)}$$

The partial coverage contract of the observer program pays for observer days according to the intersection of two variables: fixed costs for each deployment day, and variable costs in terms of transportation. While the fixed cost component of observer days are known and equal between deployments of observers, variable costs are not. A matrix of flight costs between each port in Alaska as a function of time before departure and booking date would need to be obtained for such costs to be considered in sample allocation formulae. For this reason, fully loaded rates that incorporate the fixed cost per day and the averaged travel cost per day are normally used by the observer program to present costs. This fully loaded rate (cost per day) is equal for each deployment. Assuming equal trip duration, the costs of each trip (in terms of a daily rate) are assumed to be equal between strata.

Neyman (1934), proposed a special case of optimum allocation that is arguably the most widely used and known concept of stratification and optimal allocation of the sample. Under the constraint that costs of obtaining each sample unit in a stratum is equal across strata, the optimal allocation of samples within each stratum is proportional to its relative weighting of the total number of units in the stratum and the square root of the variance ($W_{h\ opt}$, equation 2),

$$n_{h\ opt} = n * W_{h\ opt}, \quad \text{where } W_{h\ opt} = \frac{N_h S_h}{\sum_{h=1}^{H} N_h S_h}. \qquad \text{Eq. (2)}$$

"Neyman allocation" has important implications on how strata are defined in a sampling program. If strata are defined such that they comprise groups of similar values of the target metric, then overall variance will be reduced through stratification. If, however, strata poorly discriminate between similar sample units according to the target metric, overall variance will not be substantially reduced. In the special case where variance is unknown or considered equal among strata, then n_h is set proportional to the weight of the strata ($W_{h\ prs}$) and the resulting allocation is known as *proportional allocation* (PRS, equation 3),

$$n_{h\ prs} = n * W_{h\ prs}, \quad where\ W_{h\ prs} = \frac{N_h}{\sum_{h=1}^{H} N_h}. \qquad \text{Eq. (3)}$$

There are three problems that arise with Neyman allocation. First, it is possible that formulae may result in $n_h > N_h$. In this case, Cochran (1977) recommends setting the n_h with the largest $n_h : N_h$ ratio equal to N_h (100% coverage) and then re-calculating Neyman allocation with the remaining strata. The second problem is that resulting n_h are not integers. Rounding offers a simple solution; however, it is possible to end up with the situation where the sum of $n_h > n$. The third challenge is how to allocate when there is more than one target metric. In these cases, Cochran (1977) shows that the *compromised optimal allocation* (m_h; OPT) is derived from the average number of optimal sample sizes measured across L metrics (equation 4),

$$m_h = n * \overline{W}_{h\,opt,} \text{ where } \overline{W}_{h\,opt} = \frac{\sum_{l=1}^{L} W_{lh}}{L}. \qquad \text{Eq. (4)}$$

It is worth noting that unless n_h among all metrics are positively correlated, the resulting compromise allocations may be substantially different from n_h for any individual target metric.

Data Preparation

A database containing species-specific catch amounts, dates, locations, and disposition, observation status, and associated ADP strata for 2013 and 2014 was enhanced with additional information from the Alaska Regional Office and FMA to assign past fishing trips to stratification schemes 3-6. First, past fishing activity by nine CPs in the second half of the year were relabeled as belonging to the partial coverage category and not the full coverage category. Second, past partial coverage activities by AFA catcher vessels that volunteered for full coverage during the 2014 BSAI Pacific cod fishery were relabeled as belonging to the full coverage category. This decision was made as a compromise between the larger list of vessels that volunteered in 2013 and the corresponding smaller list from 2015. Third, past partial coverage fishing activities by 56 vessels identified by the Electronic

Monitoring workgroup of the Council[3] as the “2016 EM Selection Pool” were removed from these stratification schemes. Following the STRATA 1315, all trips corresponding to Hook and Line and Pot gear < 40 ' LOA in addition to all trips with jig gear were removed from the analyses since these trips have no selection probability.

Evaluation of Alternative Designs

The evaluation of alternative designs involves three major steps: simulation of observer deployments, gap analysis, and distance rankings (Fig. 1). The following sections describe these steps in greater detail.

Simulation of Observer Deployments

Two trip metrics were used in this analysis: total retained weight of groundfish and total discarded weight of groundfish. This first metric is identical to that used in NPFMC et al. (2011) to generate the STRATA 2010 stratification scheme. The second metric is the product of observer discard rates applied to total retained groundfish on each catcher vessel trip combined with some of the “prohibited species catch” (PSC) algorithms of the Catch Accounting System (Cahalan et al. 2014). Total groundfish discarded in this study includes PSC of Pacific halibut (*Hippoglossus stenolepis*), but not crab or salmon species that are managed in numbers of fish.

The population of partial coverage trips from 2013 and 2014 corresponding to each stratification scheme was used to generate Neyman optimal allocations for each metric (Eq. 2) that were adjusted following Cochran (1977) and rounded if necessary. If the sum of $n_h \neq n$, the stratum with the greatest n_h value was reduced or increased by the difference. These values were then subsequently used to generate OPT allocations (Eq. 4) for each stratification scheme and PRS allocations were also generated (Eq. 3)

---

[3] List of vessels modified after the EM Working Group’s “opt-in” criteria. Preliminary report of these vessels can be found online at: http://www.npfmc.org/wp-content/PDFdocuments/conservation_issues/Observer/EM/EM%20Selection%20Pool%20Opt-In%20Characteristics.pdf.

using a sample size of 2,000. This initial choice of *n* was used as a rough approximation of the combined number of trips sampled in 2013 and 2014 and is of relatively minor consequence since it is used only to convert allocation weightings into sample sizes for each stratum in simulations which will be described in the following paragraph.

The six stratification schemes described previously, combined with PRS and OPT, generated 12 alternative sampling designs (6 stratification schemes x 2 allocation strategies). For each design, stratified random sampling without replacement was performed on the population of partial coverage trips for 10,000 iterations. In each iteration, Horvitz-Thompson estimates (Horvitz and Thompson 1952) with corresponding standard errors (SE = the square root of the variance of the estimate) of each metric were obtained[4]. For comparison, each estimate ($\hat{X}$) divided by the known true value (X; a measure of accuracy) was plotted against its corresponding SE (a measure of precision) with ellipses corresponding to the 95% region assuming a multivariate normal distribution.

Although the independent values of accuracy and precision from each iteration provide meaningful ways to explore the data from each simulation, summaries provide a much easier way to compare sampling designs. The absolute percent error for each simulation was used as the basis for comparative metrics for each sampling design (Eq. 5),

$$|\%\ error| = \left|\left[\left(\frac{X - \hat{X}}{X}\right)\right] * 100\right| \ . \qquad \text{Eq. (5)}$$

Specifically, the mean and variance of the absolute percent error among all simulations were generated for each sampling design (MPE and VAR, respectively). Relative measures of accuracy and precision were then generated from Eq. 6,

[4] Note that this estimator reduces to a simple design-based estimator in this case where sampling probabilities for all trips in a stratum are equal.

$$I_{rel} = \left(\frac{I}{I_{min}}\right) * 100 \ , \qquad \text{Eq. (6)}$$

where *I* denotes the index (MPE or VAR here), *rel* denotes relative, and *min* denotes the minimum value among sampling designs. These relative indices range from 100 to infinity, are without units, and are created so that smaller is better.

Gap Analysis

Previous evaluations of observer deployments by the observer program have placed a high value on the results of gap analysis (Faunce et al. 2014, 2015). This is because of the invaluable service observers provide in the generation of total catch estimates; if there is no observer data in a given domain of interest, then data must be borrowed from similar or adjacent sampling units and incorrect inference about the total catch can result. This has implications for the in-season quota management used in Alaska.

Unlike the simulations described previously, in gap analysis the interest is in predicting the performance of each sampling plan using the most recent data. For this reason gap analyses and all subsequent analyses were performed on the 2014 subset of the source data (Fig. 1). The number of partial coverage trips corresponding to each stratification scheme was summed into domains defined by Gear, NMFS Area, and Target combinations that are roughly equal to those used by the Catch Accounting System for catcher vessels delivering shoreside (Cahalan et al. 2014). Gear was defined as three types following the STRATA Gear stratification scheme and NMFS Area was combined for the Bering Sea into one area.

The number of budgeted observer days (*D*) was converted into budgeted observer deployments (i.e., observed trips) by dividing it by the average trip duration (d_i) during 2013 and 2014 within the partial coverage category of the STRATA16 stratification scheme. Using the previously identified weighted sample allocations this revised value for *n* was then used to calculate OPT and PRS sample sizes for each strata in each stratification scheme (Eq. 2 and 3).

The hypergeometric distribution was used to calculate the probability of observing at least three trips within a domain for each sampling design. These probabilities were made Boolean based on whether or not they exceeded 50%. This value was chosen as the minimum acceptable value since it represents equal chance of meeting the needs of variance calculation within a domain. The proportion of domains that did pass this criteria were plotted for comparison (larger is better). The proportion of domains that did *not* pass this criteria represented a G score (*G*). This G score for each sampling plan was divided by its minimum among sampling designs and multiplied by 100 to provide a relative metric for the gap analysis (Eq. 6). This relative G score ranges from 100 to infinity, is without units, and is created such that smaller is better.

Distance Rankings

The relative indices of VAR, MPE, and G were used to generate a single Euclidean distance (*E*) to compare each sampling design (equation 7),

$$E = \sqrt{(MPE_{rel} - 100)^2 + (VAR_{rel} - 100)^2 + (G_{rel} - 100)^2}. \qquad \text{Eq. (7)}$$

The choice to use relative metrics provides equal weighting of the input values on the resulting distance metric. Euclidean distances are widely used in multivariate statistics such Principal Components Analysis (PCA) and have been used in fisheries to provide a single metric in Productivity and Susceptibility Analyses (PSA; Patrick et al. 2010, Ormseth and Spencer 2011). This value for *E* is counterintuitive, since lower values are better. To ease in interpretation, *E* was rescaled by subtracting the value by the maximum among all sampling designs (Eq. 8),

$$rescaled\ E = E_{max} - E. \qquad \text{Eq. (8)}$$

This rescaled *E* distance for each sampling design is interpreted as larger is better. These values were plotted for visual comparison.

Calculation of Preliminary Coverage Rates

The calculation of observer coverage rates is desired by the public, Council, and required by ODDS for 2016. Potential coverage rates were calculated only for the sampling plans with above-average gap analyses and above-average distance metrics (hence, gap analyses results have been used twice in the final evaluation). Similar to the gap analyses, the most recent available data (2014) were used in determining preliminary coverage rates under the necessary assumption that these best represent future fishing effort. The number of expected observed trips in each stratum from gap analyses divided by the number of trips in the stratum yielded the expected coverage rate. These calculations were repeated for the Neyman optimal allocations for both catch metrics for comparison with the compromised optimal allocation.

RESULTS AND DISCUSSION

The PRS designs outperformed all but one OPT design in gap analyses (Fig. 2). This is because PRS allocates observer deployments proportional to fishing effort and thereby ensures that observer coverage is allocated at the same rate to all fishing activities (akin to Gear:Area:Target used in gap analyses), whereas OPT strategies instead allocate proportional to the product of effort and variance in the target metric(s). For example, if many vessels fishing with the same gear have highly varying catches of retained and or discarded groundfish of the same principal species, then OPT strategies will tend to allocate many observed trips to that area whereas PRS would not. In contrast, if a few vessels fished in many different areas with similar catch characteristics, then PRS would (with sufficient sample size) ensure that at least some of those trips were observed, whereas OPT allocation would not suggest allocating observer deployments to those trips and they would be missed.

Several trends are evident from plots of catch estimates and associated precision (Fig. 3). Regardless of the sampling design, the mean estimate always reflected the true value. This is because of the “law of large numbers”, which states that as sample size grows, the mean of the sample will get closer and closer to the population mean. This applies only if a random and unbiased sample is achieved, which is the case in the perfectly executed deployments simulated here. As designed, the OPT allocations resulted in greater precision in groundfish retained estimates than PRS allocations. The current stratification (STRATA 1315) had the least precision, whereas the Gear and FMP stratification (STRATA GFMP2) had the greatest precision among stratification schemes (Fig. 3). However, OPT allocation did not always result in greater precision in the total discards of groundfish. The total discard estimates from the STRATA 2010, STRATA 1315, and STRATA 16 stratification schemes were nearly identical when OPT and PRS allocations were compared (Fig. 3). However, the remaining three stratification schemes did exhibit a lower range of SE values for discarded catch for OPT allocation than for PRS allocation. This is largely due to how a few trips with extremely high total discards were handled in each allocation. The PRS allocation of all stratification schemes in some iterations captured these high discard trips in their estimates and sometimes did not. Consequently on some iterations the resulting total discard estimate had low precision, whereas sometimes it had rather high precision. The PRS “clouds” of all stratification schemes in Fig. 3 contain a distinct patch of estimates for total discards that is different from the remaining points. These are the estimates resulting from the inclusion of those high discard trips. In contrast, OPT allocation puts more observer samples in the strata with the greater variability in the total discards. The stratification schemes that separate Gear by itself (STRATA Gear) or as a function of something other than vessel length (FMP in the case of STRATA GFMP2 and Half Year in the case of STRATA HALFYR) were able to adequately reduce the “impact” of high discard trips on the overall estimates of total discards of groundfish.

The improved performance of OPT allocation over PRS allocation is readily apparent when total groundfish and discard estimates of precision and accuracy are averaged across all iterations for each sampling design (Fig. 4). However, it should be noted that the best performing designs in terms of catch estimates performed here are not those that performed well in gap analyses. The STRATA GEAR stratification scheme with was the only one examined with OPT allocation that also had above-average gap analysis scores. The STRATA GEAR.OPT sampling design had the greatest overall distance score among all of the sampling designs considered, and this stratification scheme with PRS allocation was among the four designs with above-average distance scores (Fig. 5).

The four sampling designs with above-average distance scores were considered as possible candidates for consideration in the 2016 ADP. Details on the relative allocations for these designs and how the expected number of budgeted trips translates into anticipated trip-coverage rates is provided in Table 1. Table 2 is provided as an example of how allocations differ depending on the target metrics chosen for the STRATA GEAR stratification scheme.

Caveats and Potential Improvements

There are a number of assumptions that were made that affect the utility of the results of this effort and need to be discussed. Herein a simple approach towards optimal allocation with multiple objectives was employed. The rounding methods used here to adjust Neyman allocations can be improved upon by using the methods proposed by Wright (2012, 2014). The assumption that all trips were of equal duration permitted Neyman allocations and simplified calculations of expected sample sizes and coverage rates for 2016. This assumption is largely true for trips in partial coverage by catcher vessels that deliver their catch to land-based processors. This is because these vessels lack freezing capacity and trip duration is limited by the risk of product spoilage. However, catcher processors and vessels delivering to at-sea tenders are not limited in this way and can have much longer trips. An improvement on the methods used here would be to use actual trip duration to estimate average trip cost for each stratum, and use these values in Eq. 1. Alternatively, stratum definitions that include tender trips or those made by catcher processors could also be included in future iterations of this work.

The decision to conduct simulated samples from the population of partial coverage trips from both 2013 and 2014 combined was to incorporate between-year variance in the data. Consequently, the results of these simulations should represent the "optimal of the average". This is a desirable feature since the results can be interpreted as general predictions about how a given sampling design will perform on a new population of trip data (2016 partial coverage fishing), compared to the alternative of being an excellent design for prior trip data and a poor performer on new data.

Simulations were performed under the simple assumption that deployment is executed perfectly (e.g., there are no "deployment effects" or "observer effects", *sensu* Benôit and Allard 2009). This is likely to be untrue in reality since observer effects have been demonstrated in the observer program over multiple years (e.g., Faunce and Barbeaux 2011, Faunce et al. 2015). However, it is beyond the scope of this analysis to incorporate potential observer effects into simulated deployments.

The catch on each sampled trip was assumed to be known without variance. Obviously this was an oversimplification. The simulations and catch estimates produced in this effort are single-stage and should not be confused with the estimates and associated variance that will arise from the five-stage sampling design of the observer program (Cahalan et al. 2014). Previous studies have demonstrated that although the vessel was a significant factor in estimating total discards, the first stage of nested sampling designs (vessel or trip) is often, but not always, the stage with the *least* amount of variance (Allen et al. 2002, Borges et al. 2004). An examination of the variance components of the hierarchical design of the observer program is warranted.

It is important to recognize that the result of the simulations performed here change as a function of the target metrics chosen, how gap analyses are performed, the choice of evaluation metrics and how they are weighted for final comparisons. The choice of target and evaluation metrics as well as their equal weighting lies with the author. Different choices will yield different results. It is possible that future iterations of this work can be interactive and facilitate custom user inputs[5].

Finally, for all of the reasons already listed in this section, the resulting coverage rates presented in this study should only be considered preliminary estimates that are likely high relative to what will be presented in the final ADP or realized in 2016. Once a stratification design for the final ADP is established, more robust procedures that take true trip duration into account will be used to estimate expected coverage rates following the final 2015 ADP (NMFS 2014).

SUMMARY AND CONCLUSIONS

The analyses performed here, while far from perfect, represent a necessary and important first step towards providing comparisons of alternative sampling designs for observer deployment for consideration by NMFS, the Council, and the public. The results presented here demonstrate that 1) Neyman allocations derived from multiple target metrics can be compared to a compromised optimal allocation, 2) compromised allocation largely (but not always) results in greater precision in resulting single-stage estimates than are obtained from proportional allocation, and 3) proportional allocation outperformed compromised allocations in gap analyses. Consequently, this endeavor supports the 2016 ADP following the findings and recommendations contained in the 2014 Annual Report and the SSC response to those findings. All but one of the sampling plans with above-average gap analyses scores and above-average total distance scores included proportional allocation, which may be more robust than compromised optimal allocation to new data. It is cautioned here that what is “optimal” in the past may not be so in the future. The stratification scheme STRATA GEAR, which stratifies partial coverage by three gear types was included as two of the four best performing sampling plans in this study.

[5]Quantitative staff of the AFSC and AKRO use the R programming language. For examples of how R can be used interactively, see http://shiny.rstudio.com/.

Based on the results of this study, the NMFS recommended gear-based deployment with optimal allocation for the draft 2016 ADP (NMFS 2015b). At their October meeting, the Council supported this design for use in the final 2016 ADP.

ACKNOWLEDGMENTS

An earlier version of this work was originally prepared as Appendix B in NMFS (2015b). I am thankful for the support from J. Gasper, Alaska Regional Office, Sustainable Fisheries Division, in providing the necessary data and for improving my functions and code. The R language for statistical computing (version 3.1.1) was used to perform analyses and generate associated figures (R Core Team 2014). Stratified sampling without replacement was conducted using the R package ‘sampling’ (Tille and Matei 2015). Horvitz-Thompson estimates and corresponding SE were calculated using the R package ‘survey’ (Lumley 2004, 2010). Figures were generated using the package ‘ggplot2’ Wickham (2009).

CITATIONS

Allen, M., D. Kilpatrick, M. Armstrong, R. Briggs, G. Course, and N. Perez. 2002. Multistage cluster sampling design and optimal sample sizes for estimation of fish discards from commercial trawlers. Fish. Res. 55: 11-24.

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

Borges, L., A. F. Zuur, E. Rogan, and R. Officer. 2004. Optimum sampling levels in discard sampling programs. Can. J. Fish. Aquat. Sci., 61: 1918-1928.

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

Cochran, W. G. 1977. Sampling Techniques (Third Edition), New York, NY: John Wiley & Sons.

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

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. Gasper, J. Cahalan, S. Lowe, T. A'mar, 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.

Horvitz, D. G., and D. J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. J. Am. Stat. Assoc. 47: 663-685.

Lumley, T. 2004. Analysis of complex survey samples. J. Stat. Software 9: 1-19.

Lumley, T. 2010. Complex Surveys: a Guide to Analysis Using R. Wiley, Hoboken, NJ. 276 p.

Neyman, J. 1934. On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. J. Royal Stat. Soc. 97: 558-606.

NPFMC (North Pacific Fishery Management Council), NMFS (National Marine Fisheries Service), and Alaska Fisheries Science Center. 2011. Appendix 9: Identification of stratum for the deployment of fishery observers onboard North Pacific groundfish catcher vessels, p. A-45:49. *In* Restructuring the Program for Observer Procurement and Deployment in the North Pacific: 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. March 2011. Available online at: http://alaskafisheries.noaa.gov/analyses/observer/amd86_amd76_earirirfa0311.pdf.

NMFS. 2015a. North Pacific Groundfish and Halibut Observer Program 2014 Annual Report. Published May 2015. 101 p. Available online at: https://alaskafisheries.noaa.gov/sustainablefisheries/observers/annualrpt2014.pdf.

NMFS. 2015b. Draft 2016 Annual Deployment Plan for observers in the groundfish and halibut fisheries off Alaska. Available online at: http://alaskafisheries.noaa.gov/sustainablefisheries/observers/draft2016adp.pdf.

NPFMC. 2015a. Report of the Science and Statistical Committee to the North Pacific Fishery Management Council, June 1-3, 2015. Sitka, Alaska. 27 p. Available online at: http://www.npfmc.org/meeting-minutes/.

NPFMC. 2015b. Documents and Council Motions. Accessed 8/31/2015 at http://www.npfmc.org/observer-program/.

NMFS. 2014. Appendix C: Calculation of the Selection Rate for the 2015 Annual Deployment Plan. *In* 2015 Annual Deployment Plan for Observers in the Groundfish and Halibut Fisheries off Alaska. December 2014. National Oceanic and Atmospheric Administration, 709 West 9th Street. Juneau, Alaska 99802.

Ormseth, O. A., and P. D. Spencer. 2011. An assessment of vulnerability in Alaska groundfish. Fish. Res. 112:127-133.

Patrick, W. S., P. Spencer, J. Link, J. Cope, J. Field, D. Kobayashi, P. Lawson, T. Gedamke, E. Cortes, O. Ormseth, K. Bigelow, and W. Overholtz. 2010. Using productivity and susceptibility indices to assess the vulnerability of United States fish stocks to overfishing. Fish. Bull. 108:305-322.

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at: http://www.R-project.org/.

Tille, Y., and A. Matei. 2015. Sampling: Survey Sampling. R package version 2.7. Available online at: http://CRAN.R-project.org/package=sampling.

Wickham, H. 2009. ggplot2: elegant graphics for data analysis. Springer New York.

Wright, T. 2012. The equivalence of Neyman optimum allocation for sampling & equal proportions for apportioning the U.S. House of Representatives. Am. Stat. 66: 217-224.

Wright, T. 2014. A simple method of exact optimal sample allocation under stratification with any mixed constraint patterns. Research Report Series (Statistics #2014-07), U.S. Census Bureau, Washington, D.C. Available online at: https://www.census.gov/srd/papers/pdf/rrs2014-07.pdf.

TABLES AND FIGURES

Table 1.-- Sampling designs with above-average gap analysis results and above-average distance values recommended for consideration in the 2016 ADP. Sampling designs are defined by their stratification schemes and sampling allocations (OPT = optimal, PRS = proportional). Gear stratum abbreviations are HAL = Hook and Line, POT = Pot, and TRW = Trawl. FMP stratum abbreviations are: BSAI = Bering Sea and Aleutian Islands, GOA = Gulf of Alaska. The total number of trips in each stratum, their relative proportion (Proportion N), and relative allocation under compromised optimal allocation (Relative m_h) are also provided for comparison. The number of samples afforded in each stratum (n_h) is the product of the number of samples afforded total (*n*) and either the PRS weighted allocation (W_h) for proportional allocation or the OPT weighted allocation (m_h) for compromised optimal allocation. The weighted allocation used in each rate calculation is depicted in bold. The anticipated preliminary coverage rate (Rate) is n_h divided by N_h.

Sampling design (Strata Scheme. Allocation)	Stratum (h)	Trips (N_h)	PRS weighted allocation (W_h)	OPT weighted allocation (m_h)	n_h	Rate*
GEAR.OPT	HAL	2775	0.522	**0.339**	419	0.151
GEAR.OPT	POT	1253	0.190	**0.152**	187	0.149
GEAR.OPT	TRW	1992	0.288	**0.510**	630	0.316
GEAR.PRS	HAL	2775	**0.522**	0.339	646	0.233
GEAR.PRS	POT	1253	**0.190**	0.152	235	0.188
GEAR.PRS	TRW	1992	**0.288**	0.510	357	0.179
FMP.PRS	HAL_BSAI	323	**0.067**	0.032	83	0.257
FMP.PRS	HAL_GOA	2452	**0.454**	0.311	562	0.229
FMP.PRS	POT_BSAI	546	**0.082**	0.089	101	0.185
FMP.PRS	POT_GOA	707	**0.108**	0.052	134	0.190
FMP.PRS	TRW_BSAI	119	**0.021**	0.025	26	0.218
FMP.PRS	TRW_GOA	1873	**0.267**	0.491	331	0.177
HALFYR.PRS	HAL_First	1665	**0.302**	0.183	373	0.224
HALFYR.PRS	HAL_Second	1110	**0.220**	0.154	272	0.245
HALFYR.PRS	POT_First	650	**0.106**	0.099	131	0.202
HALFYR.PRS	POT_Second	603	**0.084**	0.049	104	0.172
HALFYR.PRS	TRW	1992	**0.288**	0.515	357	0.179

*NOTE: RATES PROVIDED HERE ARE FOR COMPARISON PURPOSES ONLY AND ARE MADE UNDER THE ASSUMPTION THAT EACH TRIP IS IDENTICAL IN LENGTH, THAT OBSERVER DEPLOYMENTS ARE PERFECTLY EXECUTED, AND FISHING EFFORT IN 2014 IS EQUIVALENT TO FISHING EFFORT IN 2016.

Table 2.-- Comparison of observer coverage rates* for the STRATA GEAR stratification scheme that result from proportional allocation and compromised optimal allocation (Relative m_h; OPT). Also depicted is how the OPT coverage rates differ from those that would have resulted from either the Neyman allocation based on total groundfish discarded (Discarded) or total groundfish retained (Retained). The sampling design GEAR.OPT was the only design with OPT allocation with above-average gap analysis scores and above-average distance scores.

Stratification scheme	Stratum (h)	Proportional (PRS)	Relative m_h (OPT)	Neyman allocation (Discarded)	Neyman allocation (Retained)
GEAR	HAL	0.233	0.151	0.231	0.071
GEAR	POT	0.188	0.149	0.049	0.251
GEAR	TRW	0.179	0.316	0.269	0.363

*NOTE: RATES PROVIDED HERE ARE FOR COMPARISON PURPOSES ONLY AND ARE MADE UNDER THE ASSUMPTION THAT EACH TRIP IS IDENTICAL IN LENGTH, THAT OBSERVER DEPLOYMENTS ARE PERFECTLY EXECUTED, AND FISHING EFFORT IN 2014 IS EQUIVALENT TO FISHING EFFORT IN 2016.

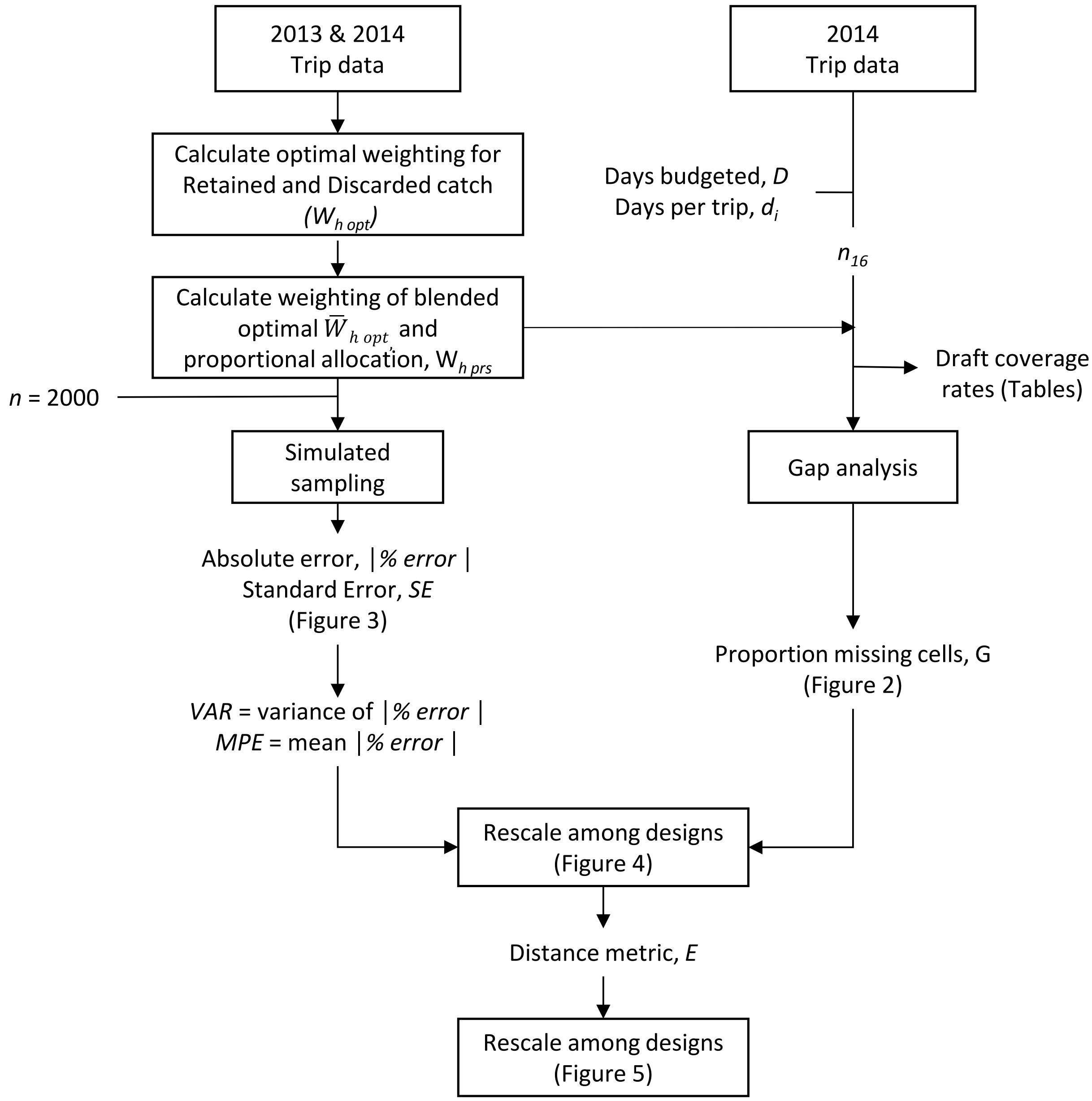

Figure 1.-- Flow chart depicting methods used in this analysis. In the text the left branch, right branch, and lower levels are broadly referred to as simulation of observer deployments, gap analysis, and distance rankings, respectively.

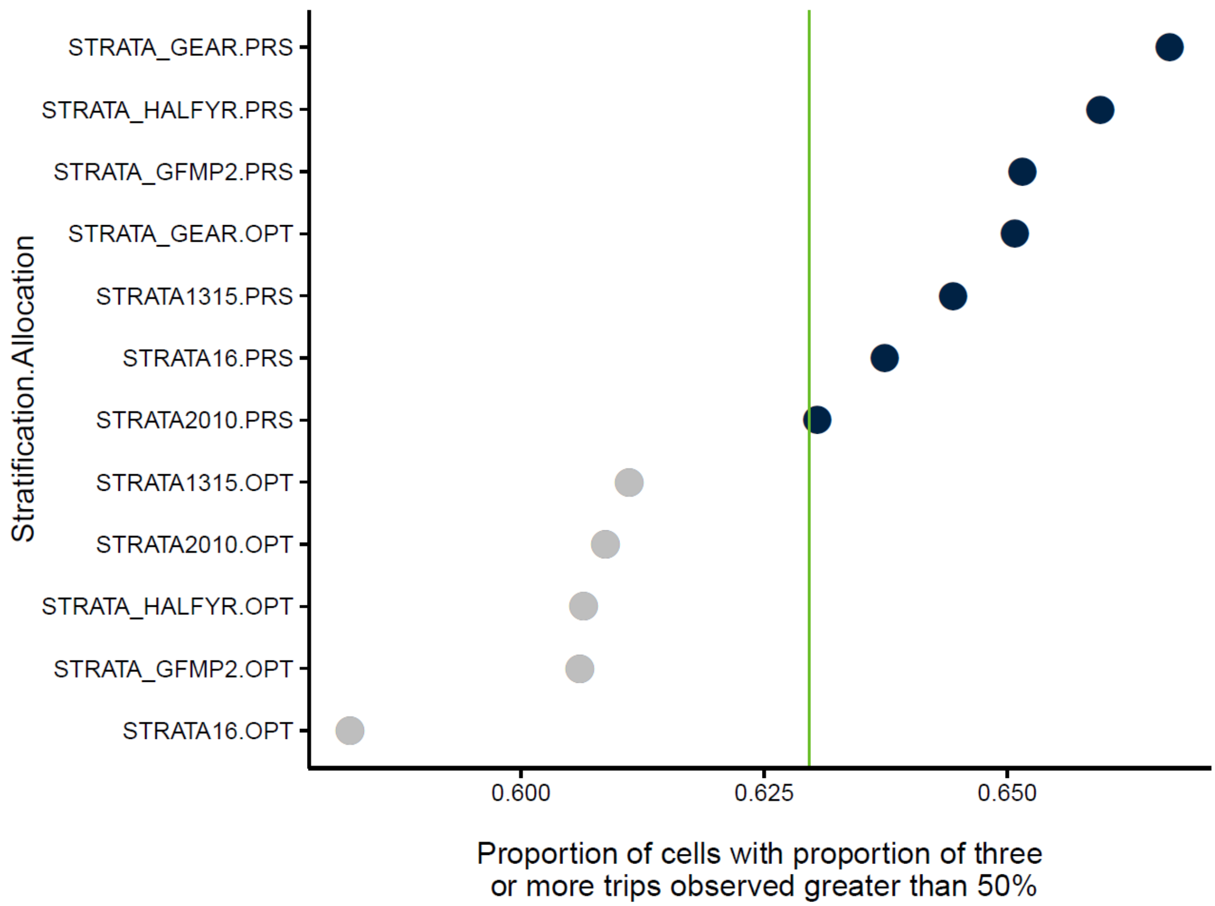

Figure 2.-- Comparison of gap analysis results for the 12 sampling designs under consideration for the 2016 ADP. See text for details on strata definitions and allocation strategy definitions. Green vertical line denotes mean among sampling designs.

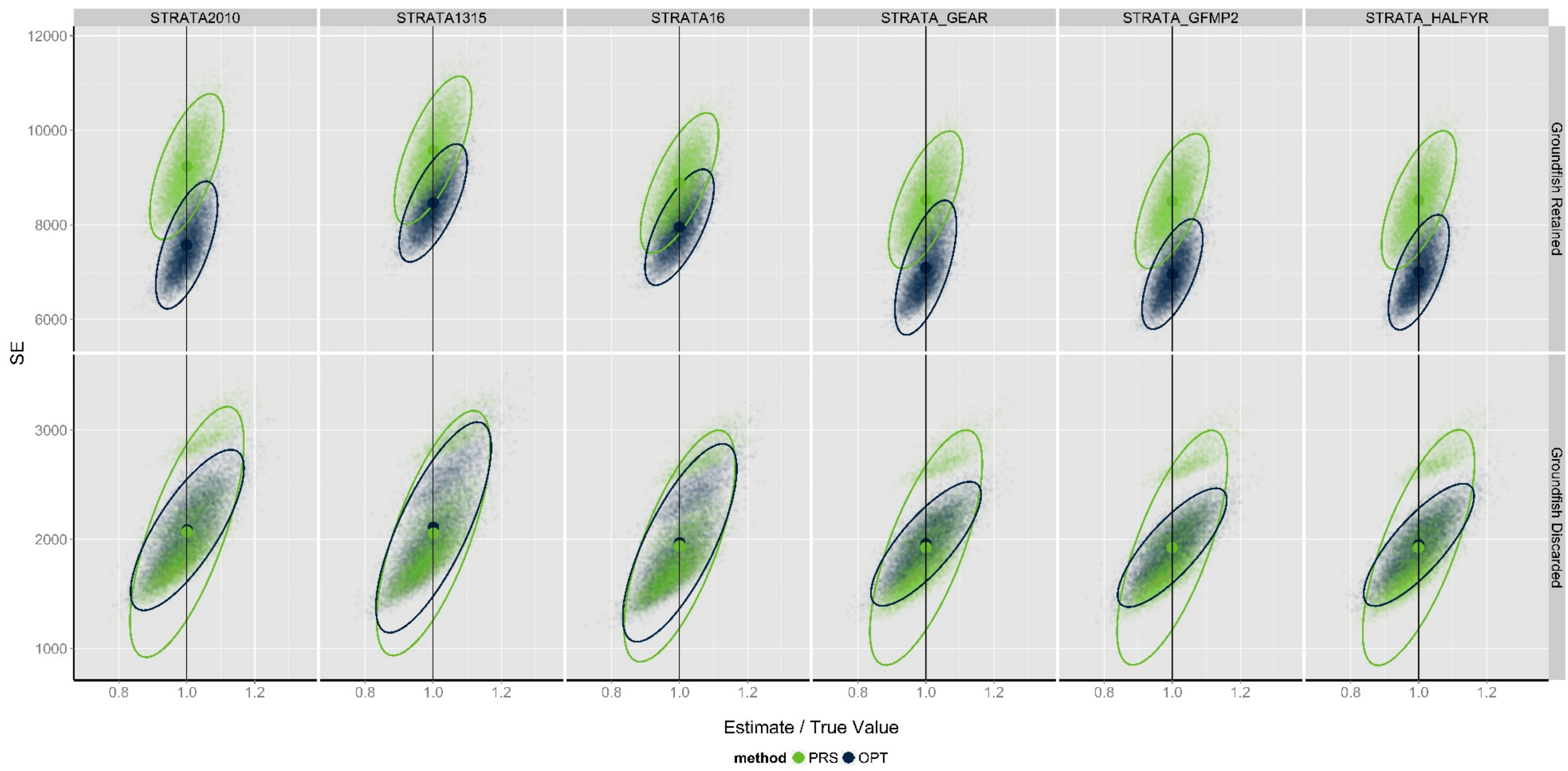

Figure 3.-- Comparison of the relative accuracy (horizontal-axis) and relative precision (vertical-axis) in the single-stage catch estimates for total retained groundfish (top panels) and total discarded groundfish (bottom panels) estimated from stratified random sampling according to six stratification schemes (columns) and two allocation strategies (colors; PRS = proportional, OPT = compromised optimal). The vertical line at 1.0 denotes the true value. As expected from the law of large numbers, distribution means from each design approximates the true value.

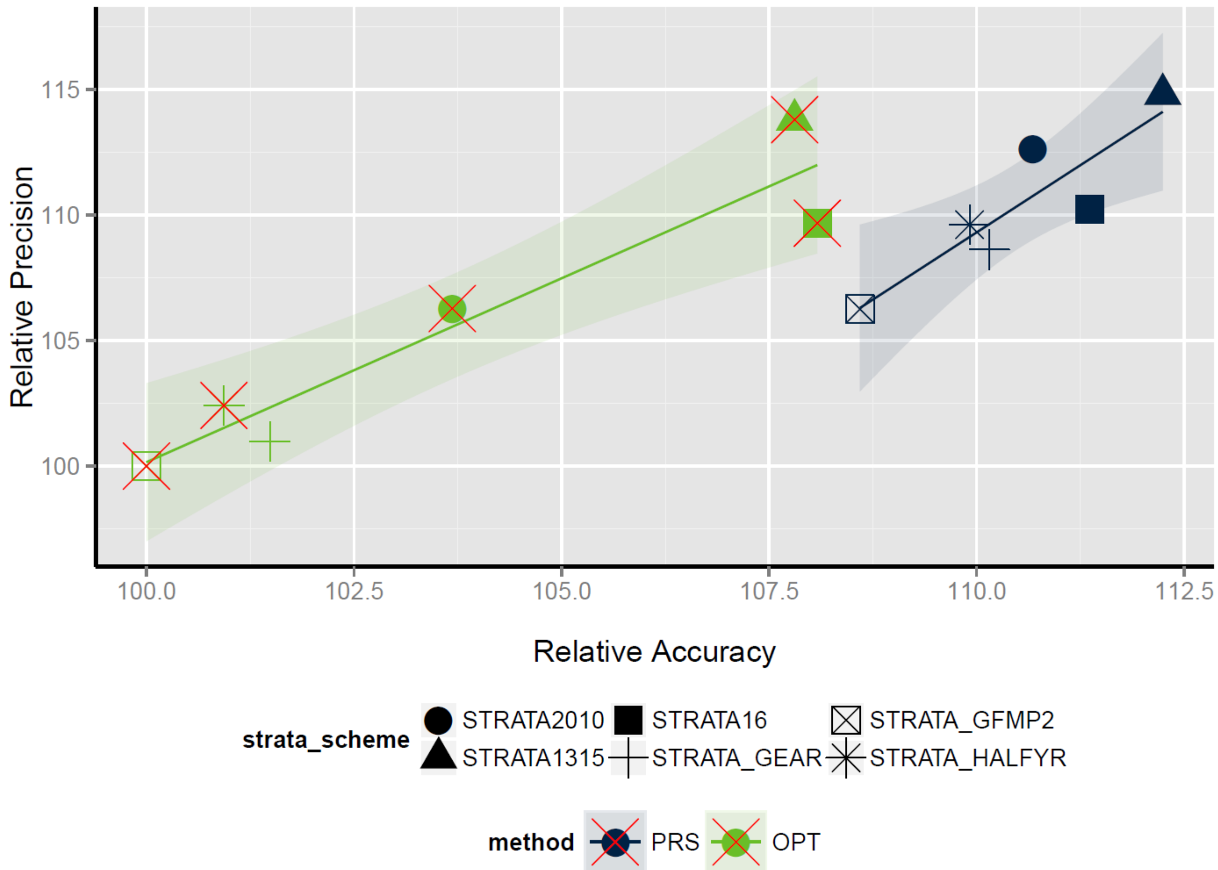

Figure 4.-- Relative accuracy and precision of total groundfish retained and discarded as measured by the comparison of means from each sampling design (strata_scheme = stratification scheme, PRS = proportional allocation, OPT = compromised optimal allocation). Sampling designs with below-average coverage in gap analyses are denoted with a red “x”.

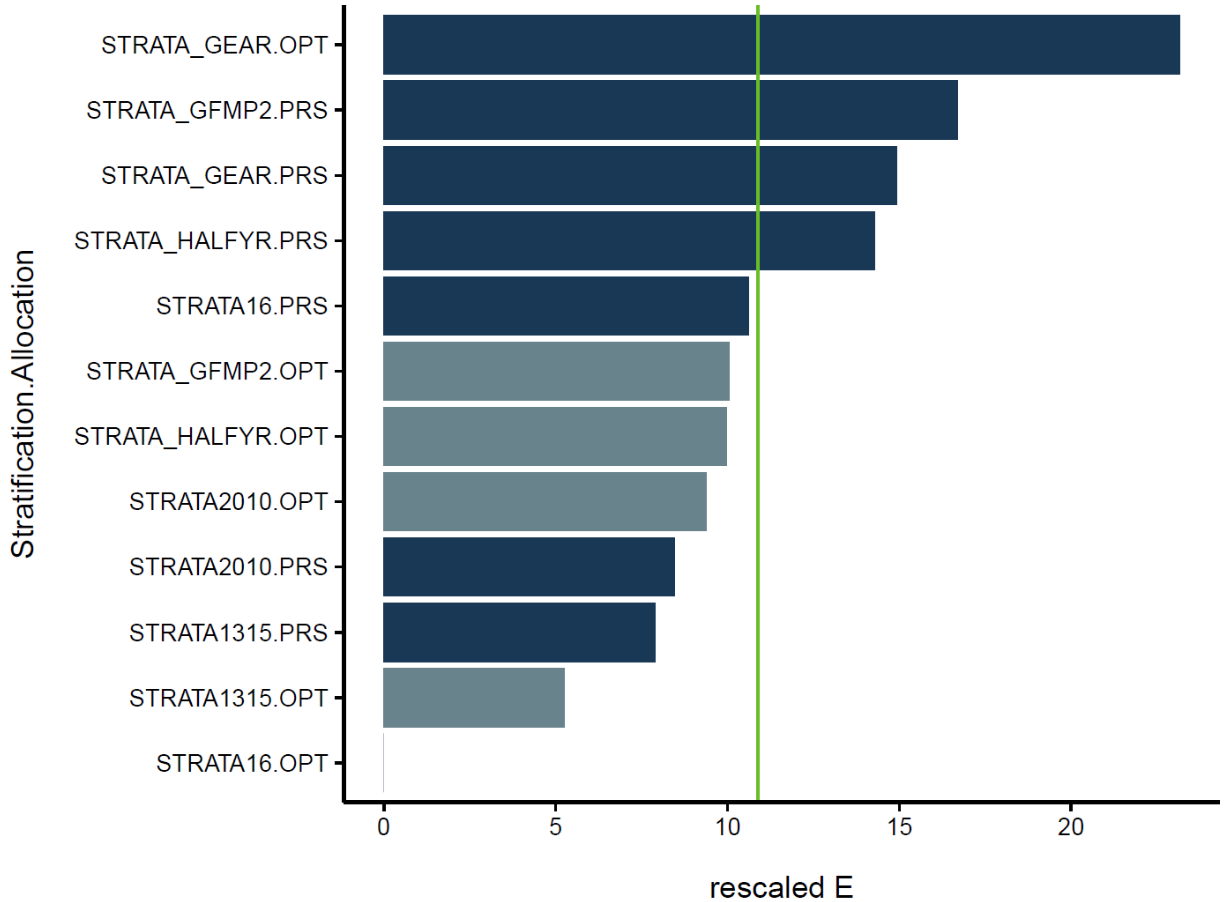

Figure 5.-- Relative distance scores of the twelve sampling designs examined in this study. Vertical line denotes mean across all designs. Those designs with below-average scores in gap analyses are colored light grey. Only those four designs with above-average gap analyses scores and above-average distance scores are examined for preliminary coverage rates.

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.

<u>AFSC-</u>

306 HIMES-CORNELL, A., S. KASPERSKI, K. KENT, C. MAGUIRE, M. DOWNS, S. WEIDLICH, and S. RUSSELL. 2015. Social baseline of the Gulf of Alaska groundfish trawl fishery: Results of the 2014 social survey, 98 p. plus Appendices. NTIS number pending.

305 FISSEL, B. E. 2015. Methods for the Alaska groundfish first-wholesale price projections: Section 6 of the Economic Status of the Groundfish Fisheries off Alaska, 39 p. NTIS No. PB2016-100044 .

304 HIMES-CORNELL, A., and K. KENT. 2015. Industry perceptions of measures to affect access to quota shares, active participation, and lease rates in the Bering Sea and Aleutian Islands crab fisheries. U.S. Dep. Commer., 67 p. NTIS No. PB2016-100043.

303 GRAY, A. K., C. J. RODGVELLER, and C. R. LUNSFORD. 2015. Evidence of multiple paternity in quillback rockfish (*Sebastes maliger*), 25 p. NTIS No. PB2016-100036.

302 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, 55 p. NTIS PB2015-105670.

301 ALLEN, B. M., and R. P. ANGLISS. 2015. Alaska marine mammal stock assessments, 2014, 304 p. NTIS No. PB2015-105669.

300 HELKER, V. T., B. M. ALLEN, and L. A. JEMISON. 2015. Human-caused injury and mortality of NMFS-managed Alaska marine mammal stocks, 2009-2013, 94 p. NTIS No. PB2015-105160.

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

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

297 VON SZALAY, P. G. 2015. The impact of sample size reduction on the precision of biomass estimates in the Gulf of Alaska, 17 p. NTIS No. PB2016-100042.

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

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

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

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

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

291 GUYON, J. R., C. M. GUTHRIE III, A. R. MUNRO, J. JASPER, and W. D. TEMPLIN. 2015. Genetic stock composition analysis of the Chinook salmon bycatch in the Gulf of Alaska walleye pollock (*Gadus chalcogrammus*) trawl fisheries, 26 p. NTIS No. PB2016-100041.