NOAA Technical Memorandum NMFS-AFSC-402

Examination of the Statistical Performance of the Mean Basket Weight Protocol in Survey Catch Processing

L. J. Rugolo, B. J. Turnock, and P. G. von Szalay

March 2020

U.S. DEPARTMENT OF COMMERCE

National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Alaska Fisheries Science Center

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:

Rugolo, L. J., B. J. Turnock, and P. G. von Szalay. 2020. Examination of the statistical performance of the mean basket weight protocol in survey catch processing. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-402, 77 p.

This document is available online at:

Document available: https://repository.library.noaa.gov/

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

NOAA FISHERIES

Examination of the Statistical Performance of the Mean Basket Weight Protocol in Survey Catch Processing

L. J. Rugolo, B. J. Turnock, and P. G. von Szalay

Alaska Fisheries Science Center
Resource Assessment and Conservation Engineering Division
7600 Sand Point Way
Seattle, WA 98115

U.S. DEPARTMENT OF COMMERCE

National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Alaska Fisheries Science Center

NOAA Techncial Memorandum NOAA-TM-AFSC-402

March 2020

ABSTRACT

The Alaska Fisheries Science Center's biennial bottom trawl surveys in the Gulf of Alaska (GOA) and Aleutian Islands (AI) are essential for assessing groundfish stock status and for formulating best management practices. When the catch of a species is too large to completely sample, one of several catch estimation procedures is employed. The 'counted basket' method is common in which the catch is processed using weighed and unweighed (counted) baskets, and total catch weight is estimated using the estimated mean basket weight and total basket count. Thirty consistently-filled basket weights are recommended to estimate the mean. No analytical foundation exists to demonstrate the accuracy of this approach in estimating large volume catches. We assess the error associated with this estimation procedure, and examine its statistical performance in terms of the representativeness and robustness of the catch estimates. If shown to be robust, then its use would address ergonomic and safety concerns of whole-haul processing large catches.

Three data sets were the subject of our analysis. Two were large volume (> 7.0 metric tons [t]) catches from the 2016 AI survey (Atka mackerel [*Pleurogrammus monopterygius*]) and the 2017 GOA survey (Pacific ocean perch [*Sebastes alutus*]). While they were completely sampled, the baskets were not consistently filled and, consequently, there was greater variability in these data than if the protocol had been applied. The third data set was a medium volume (~ 3.0+ t) catch of Pacific ocean perch collected from the 2018 AI survey, completely sampled and processed consistent with design specifications. The performance of the counted basket method was simulated via bootstrap analysis over a range of population sizes (~ 1.0 t to 7.0+ t) and sample sizes (n = 1 to 50 baskets). For the two non-standard data sets, the point estimates of catch were accurate and relatively precise over the range of catch volumes and sample sizes. At the target of n = 30 baskets, the CVs of the estimates were ≈ 1.5% for all catch volumes, and there was almost negligible gain in the accuracy of the estimates for a large volume catch (7.0+ t) relative to a medium volume (~3.0 t) catch. The 95% confidence bounds on the total catch estimate deviated by an average 2.4% from the observed at n = 30, and averaged 2.6% for

large volume catches and 2.3% for medium volume catches. We surmised that the catch estimates would not deviate by more than 2.5% from the observed, and undoubtedly less, if the method is applied consistent with design requirements. Based on these non-standard data, we found the counted basket approach to be statistically robust in terms of performance measures, and that it produced reliable estimates of mean basket weight and non-subsample total weight.

Using the 2018 data set, the catch estimates were highly accurate and highly precise over the range of catch volumes and sample sizes. The CV on the estimates of catch was ~ 1.0% at $n = 15$, and $< 0.7\%$ at $n = 30$ baskets. There was negligible gain in the precision of the estimates at $n > 20$ to 25 weights. For the non-standard data, the CVs exceeded 1.0% even at $n = 50$, whereas they were $< 1.0\%$ at $n \geq 17$ and equaled 0.68% at $n = 30$ for the 2018 data set. At $n = 30$, the mean basket weight estimate was indistinguishable from the observed, and the difference was less than the level of precision (0.01 kg) of the weight scales. The difference between the non-subsample total weight estimate and the observed was similarly negligible. The 95% confidence bounds deviated by $\leq 1.1\%$ from the observed at $n = 30$, which supported our supposition that the bounds would not deviate by more than 2.5% from the mean for data collected consistent with the design. In general, the CVs on the point estimates for the 2018 data set were less than one-half the magnitude of those for non-standard data for all catch volumes and sample sizes, while the deviations in the 95% confidence bounds for the non-standard data were greater by a multiple of 2.5 than for the 2018 data set.

Results of this analysis reinforce the validity of the counted basket approach in processing large-volume catches on the GOA and AI surveys, particularly when the haul data are collected in accordance with design prescriptions. The approach results in highly precise and highly accurate estimates of the catch relative to whole-haul processing, and we conclude that whole-hauling large volume catches is not statistically necessary to achieve reliable estimates of the catch.

CONTENTS

INTRODUCTION

The Alaska Fisheries Science Center (AFSC) conducts a biennial bottom trawl survey in the Gulf of Alaska (GOA) and Aleutian Islands (AI) to assess trends in the abundance and distribution of groundfish fishery resources under the aegis of the North Pacific Fishery Management Council. The AI survey has been conducted since 1980 and that in the GOA since 1984. The objective of these surveys is to collect data essential to the formulation of best management practices consistent with mandates of the Magnuson-Stevens Fishery Conservation and Management Act. Survey data serve as the underpinning of 'best available science' used to assess stock status, specify overfishing definitions, and to establish fishery controls.

Survey sampling protocols for the AFSC's bottom trawl surveys are described in Stauffer (2004). Upon completion of each successful haul, catch processing includes identifying, enumerating and weighing all taxa, and collecting biological data on select commercially and ecologically important fish species. Key biological metrics collected for these species are length- and sex-composition, stomach contents for feeding habit analysis, and otoliths for ageing analysis. These data are derived commonly from a sample of the catch that depends on the total size of the catch and the target collection goals for each metric.

The total catch can range in weight from a few 10s of kilograms to in excess of 10 metric tons (t) depending on the abundance of fish at a station. Customarily, catches are of a size (< 5 t) enabling the complete or whole-haul processing of the catch in which a total weight is measured for each species. When the total catch for any species exceeds target sample goals for biological metrics, it's assigned to two categories: the 'subsample' – the weight of the catch component set aside for length and sex data collections, and from which otoliths are taken; and, the 'non-subsample' – the weight of the component not needed for biological collections which is discarded. When the catch for an individual species is less than target collection goals, it's all assigned to the subsample category. On both the GOA and AI surveys, with emphasis on the latter, the total catch is occasionally too large (> 6 t) to efficiently whole-haul process given time constraints, or the catch of a dominant species is similarly too large to completely sample. In such instances, one of several specified catch estimation procedures is employed as specified in the annual GOA or AI Scientific Operations Plans. These large catch sampling procedures are derivative of either crane-mounted load cell weights or volumetric sampling approaches in which a portion of the catch is weighed and the remaining portion estimated volumetrically, then weight imputed on estimated volume.

Among the large volume catch processing methods, the 'counted basket variant' (aka, 'mean basket protocol') is often employed. In this procedure for the case of a dominant species in the catch, the catch is emptied onto the sorting table in successive splits and the catch processed using a combination of weighed and unweighed (counted) baskets. Baskets comprising the subsample are weighed and set aside for processing. For the non-subsample, the procedure prescribes that baskets are consistently filled (hence, the volumetric element), at least 30 baskets weighed, and a count taken of all remaining baskets in excess of the 30 weights. The estimate of the total catch weight is the sum of the subsample weight, the non-subsample baskets for which weights were taken, and the product of the count of the discarded baskets and the mean basket weight derived using both subsample and non-subsample weights. A spreadsheet is used to tabulate these data and make the product expansion of the basket count and mean basket weight to estimate the total weight of discarded baskets. The Deck Lead has the option to exclude from the product expansion any basket weight judged not comparable with baskets consistently filled (e.g., partial baskets at the end of a split). Although subjective, designating such weights as 'outliers' allows more robustness in the mean basket weight estimate and, therefore, in the estimate of total catch weight.

The mean basket protocol has been used on the GOA and AI surveys *ad hoc* for many years, perhaps decades. It was formalized as an alternative catch processing method, and described in the 2015 GOA Scientific Operation Plan (Anon. 2015). Despite that, there is no requirement to employ it for large volume catches of a dominant species, and neither are criteria defined that would trigger its implementation. As a result, this protocol has been used non-uniformly over GOA and AI survey operations. Notwithstanding any logistical considerations which may limit its use for a particular haul or vessel, the inconsistency of its deployment among staff resides, at least partially, in uncertainty concerning the validity of the method for total weight estimation. The resulting weight estimate includes process error additional to that of a complete sample weight observation presumably with only measurement error. No analytical foundation has been provided to date to demonstrate the accuracy of this method in processing large volume catches.

The goal of this study was to assess the validity of and error associated with the mean basket weight protocol in catch processing. We examine its statistical performance in terms of the representativeness and robustness of the estimate of the mean basket weight, and its extrapolation to total non-subsample weight. Our aim is to provide the technical basis to evaluate this method of catch processing and to advance its understanding and use as a catch sampling procedure. Our hope is that the analysis might 1) address concerns about the counted basket variant approach in large-volume catch processing, 2) engender discussion on control rules for its implementation, and 3) enable increased efficiency of catch processing without loss of precision. If the analysis shows that this method is robust in terms of catch estimation relative to whole-haul processing, then it can be applied (without significant loss of precision or accuracy) to address ergonomic and safety issues associated with complete processing of large catches. That is, is whole-hauling large volume catches defensible in light of findings that may demonstrate the scientific validity of this approach in catch estimation?

MATERIALS AND METHODS

The Data

We utilized two data sets from the GOA and AI surveys to conduct the study. The first (Data Set-1) was a haul from the 2017 GOA survey that consisted of 7,590 kg of Pacific ocean perch (*Sebastes alutus*) split into 247 non-subsample baskets totaling 7,438 kg, and 5 subsample baskets totaling 151 kg. The second (Data Set-2) was a haul from the 2016 AI survey consisting of 7,207 kg of Atka mackerel (*Pleurogrammus monopterygius*) split into 213 non-subsample baskets (7,136 kg), and 2 subsample baskets (71 kg). Neither of these hauls was processed using the counted basket variant approach. They were completely sampled (i.e., all baskets were individually weighed and not level-filled prior to weighing). Consequently, there was greater variability in the basket weights than expected had the counted basket variant been employed. Results of the analysis of a third, medium volume (~ 3.0+ t) catch of Pacific ocean perch collected from the 2018 AI survey, completely sampled and processed consistent with design specifications, are presented in the Appendix.

Analytical Approach

Our analysis consisted of a two phases. The first was simple exploratory data analyses to understand the data characteristics and identify outlier basket weights in the data. Specifically, to 1) inspect the data and identify outlier basket weights, 2) test whether the respective non-subsample and subsample weights were consistent with having been drawn from the same population, and 3) evaluate the distributions of basket weights relative to assumptions about normality. In the second phase, we conducted simulations using bootstrap analysis to 1) evaluate performance statistics of the distribution of the mean basket weight, 2) estimate non-parametric confidence limits on the distribution of mean weight, 3) examine performance of the estimated total non-subsample catch weight to the true catch weight, and 4) simulate outcomes over a range of sample sizes and range of non-subsample population sizes. All analyses were implemented using R Development Software (R Core Team 2013).

Exploratory Data Analysis

A suite of first-order statistics (mean, standard deviation, coefficient of variation (CV), upper and lower 95% confidence limits) were estimated for each data set to provide insight into the distribution and variability of basket weights in the data and to identify outlier weights. Since both hauls were completely sampled, outlier basket weights were included in the data. In processing a haul using the counted basket approach, the designation of a basket weight as an outlier is the judgment of the Deck Lead, and there are no rules for such designations in the Scientific Operations Plan. Most often, outlier weights represent partially- or incompletely-filled baskets for which weights were taken. Outlier weights are excluded from the estimation of mean basket weight but included in the sum of the non-subsample total weight after the product expansion of counted baskets and mean basket weight. Since the hauls analyzed in this study were not processed using the counted basket approach but whole-hauled, outlier weights in these data were expected.

Initial data inspection was performed by way of graphical analysis to illustrate the distribution of basket weights in the non-subsample and subsample, inspect data ranges, including measures of central tendency and quantiles of the distributions, and test departures from normality.

For the 2017 GOA Pacific ocean perch data, 3 of the 247 non-subsample basket weights were assumed to be partially-filled baskets, designated as outliers and removed from the analysis. The resultant Data Set-1 subject to analysis consisted of 244 non-subsample basket weights totaling 7,389 kg, and the original 5 subsample baskets totaling 151 kg.

For the 2016 AI Atka mackerel data, 5 of the 213 non-subsample basket weights were assumed to be partially-filled baskets, designated as outliers and removed from the analysis. The resultant Data Set-2 subject to analysis consisted of 208 non-subsample basket weights totaling 7,043 kg, and the original two subsample baskets totaling 71 kg.

Randomization test-- To assess whether the non-subsample and subsample basket weights in each data set were consistent with having been drawn from the same population, a non-parametric Randomization Test of the difference in the respective means [H_o: $\mu1 - \mu2 = 0$] was conducted. This test is freed from constraints of random sampling from a known error distribution with equal variances, and the null hypothesis is based on the sampling distribution of the test statistic generated by the data. Given the small number of subsample weights in both sets, the goal of this analysis was to test for unusually large departures in the means of the respective categories. It was not known whether the subsample baskets for either haul were filled to meet target biological data collection goals and, therefore, may have differed substantially from observed non-subsample basket weights. Under the mean basket weight protocol, the non-subsample and subsample basket weights are combined to estimate the mean basket weight used in the product expansion of counted baskets.

Under the null hypothesis, the observed basket weights comprising the non-subsample and subsample populations is a random realization generated by chance, and the observed weights are exchangeable between groups. The observed test statistic, T_{OBS}, was the difference in mean weight between non-subsample and subsample groups. The distribution of the test statistic under H_o was obtained by resampling the basket weights without replacement and calculating 10,000 values of the test statistic under rearrangement of the labels on the data points. The p-value of the test represented the probability that the test statistic was at least as extreme as T_{OBS} at $\alpha = 0.05$ given that the null hypothesis was true (i.e., $P[|\mu1- \mu2|] > T_{OBS}$).

Simulation Analyses

For each data set, we simulated implementation of the mean basket weight protocol via bootstrap analysis. The performance of the mean basket weight estimate was assessed in terms of its distribution, coefficient of variation, and 95% non-parametric confidence bounds relative to observed mean weight. For the estimated non-subsample total weight, we examined the bootstrapped estimates of total weight, its coefficient of variation, 95% confidence bounds relative to the observed total weight, and the range of and the 95% confidence bounds as a proportion of the estimated total weight.

Results of the counted basket variant approach applied to these data were evaluated for varying catch volumes versus varying sample sizes (n) used to estimate the mean basket weight. For the large catch volume simulations, we bootstrapped the mean non-subsample basket weight from the 244 observed basket weights (Data Set-1) and the 208 observed basket weights (Data Set-2) by resampling 10,000 times without replacement for sample sizes n = 1 to 50. Both non-subsample population sizes exceeded 7.0 t – a large volume catch by our definition. For the simulations involving lesser catch volumes, subsets of the non-subsample populations of sizes 50-200 baskets were derived as random draws without replacement from the observed non-subsample populations of 244 basket weights (Data Set-1) and 208 basket weights (Data Set-2). For each level of sample size n = 1 to 50 and each level of population size 50-200, we generated 10,000 non-subsample mean basket weights by resampling the non-subsample population subset without replacement. This analysis generated the 7,550 sample size-catch volume combinations for each data set.

We selected four simulation scenarios as being informative in terms of the performance of the mean basket protocol to survey catch processing.

Scenario-1: Large Volume Catch and Fixed Sample Size. For a fixed sample size of n = 30 basket weights, we bootstrapped the mean non-subsample basket weights from the 244 observed weights for Data Set-1 and the 208 observed weights for Data Set-2. This design was most similar to the recommended prescriptions of the approach in the Scientific Operations Plan and to customary large volume catches experienced on the GOA and AI surveys.

Scenario-2: Medium Volume Catch and Fixed Sample Size. We examined the performance of the mean basket protocol for a lesser catch volume using these data - on the order of 3.0 t versus 7.0+ t. For each data set, we generated a subset of 100 non-subsample basket weights from the population of 244 weights (Data Set-1) and 208 weights (Data Set-2), then bootstrapped the mean non-subsample basket weight using a fixed n = 30 basket weights. This design provided insight into the gain in robustness achieved when estimating a moderate-sized non-subsample population using the prescribed sample size of n = 30 baskets.

Scenario-3: Large Volume Catch and Variable Sample Size. As an extension to Scenario-1, we examined results of processing a large volume catch using a range of basket weights to estimate the mean. For each level of sample size n = 1 to 50, we bootstrapped the mean non-subsample basket weights from the 244 observed weights for Data Set-1 and the 208 observed weights for Data Set-2. This design provided a basis to evaluate the return on investment in terms of model diagnostics of weighing more or less than 30 baskets as prescribed in the Scientific Operations Plan.

Scenario-4: Variable Volume Catch and Variable Sample Size. We formulated results of the counted basket variant approach for a range of catch volumes and range of sample sizes to estimate the mean. For each data set, subsets of the non-subsample population of sizes 50-200 baskets (~ 1.5 t to ~6.0+ t) were derived from the observed non-subsample populations of 244 basket weights (Data Set-1) and 208 basket weights (Data Set-2). For each level of population size, results were examined for sample sizes n = 1 to 50. Since this simulation generated a set of model diagnostics for 7,550 sample size-catch volume combinations, select diagnostics relevant to assessing the performance of this sampling approach were examined.

RESULTS

Exploratory Data Analyses

<u>Data Set-1</u>: The original haul data consisted of n = 252 baskets (7,589.88 kg) split into a non-subsample (n = 247; 7,438.42 kg) and a subsample (n = 5; 151.46 kg). The mean of the non-subsample was 30.12 kg, with range = [13.93, 38.46], CV = 10.74%, and 95% bounds = [29.71, 30.52] (Table 1). The mean of the subsample was 30.29 kg with range = [27.84 32.32], CV = 6.46%, and 95% bounds = [28.58, 32.01] (Table 1). The distribution of non-subsample weights

was variable, left-skewed, and departed from normality (Fig. 1a). Three non-subsample weights below 20 kg were 3.6-5.0 standard deviations from the mean. They were designated as outliers and removed.

The resulting non-subsample was n = 244 baskets totaling 7,388.83 kg, with mean = 30.28 kg, range = [21.43, 38.46], CV = 9.47%, and 95% bounds = [29.92, 30.64] (Table 1). While the distribution of basket weights were less variable, the approximate 1.8× difference in the range of weights was still higher than expected had they been level-filled. Outlier removal resolved departures form normality in the non-subsample distribution (Fig. 1b). The original subsample consisting of n = 5 baskets (151.46 kg) was not changed. The revised data set was considered more suitable to examining the mean basket protocol in catch estimation despite the greater variation in the data had the protocol been applied to this haul.

<u>Data Set-2</u>: The original haul data consisted of n = 215 basket weights (7,207.12 kg) split into a non-subsample (n = 213; 7,136.36 kg) and subsample (n = 2; 70.76 kg). The mean of the non-subsample was 33.50 kg with range = [13.24, 40.88], CV = 11.31%, and 95% bounds = [33.00, 34.01] (Table 1). The mean of the subsample was 35.38 kg with range = [32.64, 38.12], CV = 10.95%, and 95% bounds = [30.01, 40.75] (Table 1). The distribution of non-subsample weights were highly variable, strongly left-skewed, and departed from normality for both tails of the distribution (Fig. 2a). Five non-subsample weights below 23 kg ranged from 2.8 to 5.4 standard deviations from the mean. They were designated as outliers and removed.

The resulting non-subsample consisted of 208 baskets totaling 7,042.64 kg, with mean = 33.86 kg, range = [25.29, 40.88], CV = 8.58%, and 95% bounds = [33.45, 34.27] (Table 1). Although variability in the distribution of weights decreased, an approximate 1.6× difference in the range of weights was still higher than expected if the baskets had been level-filled. The new distribution was slightly left-skewed, and it still departed from normality at both tails (Fig. 2b). The original subsample of two baskets (70.76 kg) was not changed. The revised data set was more suitable to examining the mean basket protocol, but recognized as inherently more variable than had the protocol been implemented for this haul.

Randomization Tests of Mean Differences

The distribution of differences in non-subsample versus subsample mean basket weights for Data Set-1 is shown in Figure 3. For this test, T_{OBS} = 0.0099 kg and p = 0.9945. The mean differences were shown to be strongly normally distributed (Fig. 4). For Data Set-2, the distribution of differences in non-subsample vs subsample mean basket weights is shown in Figure 5. For this test, T_{OBS} = 0.0099 kg and p = 0.4671. The mean differences were shown to depart slightly from normality at both tails (Fig. 6) resulting from the greater variability in these data. We could not reject the null hypothesis for either data set that the non-subsample and subsample basket weights were consistent with having been drawn from the same population.

Simulation Analyses

Scenario-1: Large Volume Catch and Fixed Sample Size. These results are for a 7.0+ t catch processed using a fixed sample size n = 30.

Data Set-1: The estimate of the mean basket weight and its 95% CI were 30.279 kg (SE = 0.49) and [29.46, 31.10] kg, respectively. This compared to the observed mean basket weight of 30.282 kg (Fig. 7). The non-subsample total weight was estimated at 7,388.11 kg compared to the observed total weight of 7,388.83 kg. The 95% confidence bounds on the total weight estimate was [7,188.89, 7,587.18] kg with a range of 398.29 kg, and these bounds represented deviations of -2.70% and 2.69%, respectively, from the mean estimate.

Data Set-2: The estimate of the mean basket weight and its 95% CI were 33.856 kg (SE = 0.50) and [33.02, 34.68] kg, respectively. This compared to the observed non-subsample mean weight of 33.859 kg (Fig. 8). The non-subsample total weight was estimated at 7,042.13 kg compared to the observed total weight of 7,042.64 kg. The 95% confidence bounds on the total weight estimate was [6,868.37, 7,213.65] kg with a range of 345.28 kg, and these bounds represented deviations of -2.47% and 2.43%, respectively, from the mean estimate.

Scenario-2: Medium Volume Catch and Fixed Sample Size. These results are for a ~3.0 t catch processed using a fixed sample size n = 30.

Data Set-1 Subset: The estimate of the mean basket weight and its 95% CI were 30.436 kg (SE = 0.42) and [29.74, 31.13] kg, respectively. This compared to the observed non-subsample mean weight of 30.441 kg (Fig. 9). The non-subsample total weight was estimated at 3,043.58 kg compared to the observed total weight of 3,044.13 kg. The 95% confidence bounds on the total weight estimate was [2,974.20, 3,112.77] kg with a range of 138.57 kg, and these bounds represented deviations of -2.28% and 2.27%, respectively, from the mean estimate.

Data Set-2 Subset: The estimate of the mean basket weight and its 95% CI were 34.282 kg (SE = 0.47) and [33.51, 35.05] kg, respectively. This compared to the observed non-subsample mean weight of 34.275 kg (Fig. 10). The non-subsample total weight was estimated at 3,428.19 kg compared to the observed total weight of 3,427.52 kg. The 95% confidence bounds on the total weight estimate was [3,350.83, 3,505.47] kg with a range of 154.63 kg, and these bounds represented deviations of -2.26% and 2.25%, respectively, from the mean estimate.

Scenario-3: Large Volume Catch and Variable Sample Size. These results are for a 7.0+ t catch processed using a range of sample sizes from 1 to 50.

Data Set-1: The set of performance measures for estimates of mean basket weight and total non-subsample weight versus sample size are shown in Table 2. For the mean basket weight, they are the estimate of the mean and its standard deviation, and the lower and upper 95% confidence bounds. For the non-subsample total weight, they are the total weight estimate and its standard error, the lower and upper 95% confidence bounds, the range of the 95% CI, the lower- and upper-bound deviations from the mean, and the lower and upper bound deviations as a percent of the mean. For reference, the observed mean non-subsample basket weight was 30.28 kg; the observed total weight was 7,388.83 kg.

The lower and upper 95% confidence bound deviations from the mean are shown as a proportion of the estimated non-subsample total weight versus sample size (Fig. 11). The proportion of both deviations declined rapidly over the range of sample sizes. At n = 30, the lower and upper bound deviations were 2.64% and 2.74%, respectively.

Figure 12 shows the lower and upper 95% confidence bounds on the non-subsample total weight estimate versus sample size. At n = 30, the lower-bound estimate was 7,194.67 kg and the upper bound estimated at 7,591.98 kg. These compared to the observed non-subsample total of 7,388.8 kg and equated to deviations of -194.2 kg and 203.1 kg, respectively, from the observed total weight (range = 393.31 kg).

The change in the coefficient of variation of the non-subsample total weight estimate versus sample size is presented in Figure 13. At n = 30, the CV was 1.63%. The CV declined rapidly with increasing sample size, and it fell to less than 2.0% at n ≥ 20. Figure 14 presents a diagnostic result of the bootstrap formulation. The mean of the bootstrapped mean non-subsample weight estimates varied within a narrow range of the observed total weight over the range of n as would be expected if correctly formulated.

The 95% confidence bounds on the non-subsample mean basket weight estimate versus sample size is presented in Figure 15. At n = 30, the lower-bound estimate was 29.49 kg and the upper bound 31.11 kg. These compared to the observed non-subsample mean basket weight of 30.28 kg and equated to deviations of -0.79 kg (-2.61%) and 0.83 kg (2.74%), respectively, from the observed mean.

Data Set-2: Performance measures for this simulation are presented in Table 3. For reference, the observed mean non-subsample basket weight was 33.86 kg; the observed total weight was 7,042.64 kg.

The lower and upper 95% confidence bound deviations from the mean are shown as a proportion of the estimated non-subsample total weight versus sample size in Figure 16. The proportion of both deviations declined rapidly over the range of sample sizes. At n = 30, the lower- and upper-bound deviations were 2.45% and 2.52%, respectively.

Figure 17 presents the lower and upper 95% confidence bounds on the non-subsample total weight estimate versus sample size. At n = 30, the lower bound estimate was 6,866.8 kg and the upper bound estimated at 7,216.9 kg. These compared to the observed non-subsample total of 7,042.64 kg and equated to deviations of -175.84 kg and 174.26 kg, respectively, from the observed total weight (range = 350.06 kg).

The change in the coefficient of variation of the non-subsample total weight estimate versus sample size is shown in Figure 18. At n = 30, the CV was 1.51%. The CV declined rapidly with increasing sample size, and it fell to less than 2.0% at n ≥ 20. Figure 19 presents a diagnostic result of the bootstrap formulation. The mean of the bootstrapped mean non-subsample weight estimates varied within a narrow range of the observed total weight over the range of n as would be expected if correctly formulated.

The 95% confidence bounds on the non-subsample mean basket weight estimate versus sample size is presented in Figure 20. At n = 30, the lower bound estimate was 33.01 kg and the upper bound 34.70 kg. These compared to the observed non-subsample mean basket weight of 33.86 kg and equated to deviations of -0.85 kg (-2.51%) and 0.84 kg (2.48%), respectively, from the observed mean.

Scenario-4: Variable Volume Catch and Variable Sample Size. These results are for a 151 levels of non-subsample population size (50 to 200 baskets; ~1.5 t to ~6.0+ t) processed using a range of sample sizes from 1 to 50.

<u>Data Set-1</u>: The coefficient of variation of the estimated mean basket weight versus sample size and catch volume is shown in Figure 21. This plot would be identical if shown for estimated non-subsample total weight. Despite slight irregularities in the response surface, the CV declined rapidly from small to large n, and was fairly low and consistent for n > 10 to 15.

Figure 22 presents slices of the response surface of Figure 21 for seven catch volumes (50 to 200, by 25). For all catch volumes, the CV declined rapidly from low to high n. For population = 50 baskets, the CV went to zero at n = 50 since all baskets were weighed. The patterns of CV for the other six catch volumes were similar and marginally larger with increasing population size. At n = 23, the CVs for all catch volumes were less than 2.0%. At n = 16 baskets, all CVs were < 2.5%, and they were all < 3.0% at n = 12. At the recommended n = 30 baskets, the CVs for all catch volumes were < 1.7%.

<u>Data Set-2</u>: Figure 23 presents the CV of the estimated mean basket weight versus sample size and catch volume. Despite slight irregularities in the response surface, the CV declined rapidly from small to large n, and were fairly low and consistent for n > 10 to 15.

Slices of the response surface of Figure 23 for seven catch volume levels is presented in Figure 24. For all catch volumes, the CV declined rapidly from low to high n. For the population = 50 baskets, the CV went to zero at n = 50 since all baskets were weighed. The patterns of CV for the other six catch volumes were similar and marginally larger with increasing population size. At n = 20, the CVs for all catch volumes were less than 2.0%. At n = 14 baskets, all CVs were < 2.5%, and they were all < 3.0% at n = 11. At the recommended n=30 baskets, the CVs for all catch volumes were < 1.5%.

DISCUSSION

This study illustrates the statistical robustness of the counted basket variant approach in survey catch processing. The principal finding is that this approach is robust in terms of estimation of non-subsample mean basket weight and total non-subsample catch weight. Of particular note in considering these findings is the fact that these data were not collected in a manner consistent with prescriptions of this catch processing method. That is, the baskets were not level-filled prior to weighing and, consequently, there was greater variability in these data than would be expected had the method been applied. While we cannot assess the magnitude of this effect on the estimated metrics, we proffer that the results would be even more robust if the counted basket variant method was applied to these hauls.

We conservatively processed these data *post hoc* to remove conspicuous outlier basket weights while retaining variability in the data (Table 1; Figs. 1 and 2). The Atka mackerel data were relatively more variable than those for Pacific ocean perch, and the distribution of Atka mackerel basket weights still departed from normality (Fig. 2) after outlier removal. The ranges of basket weights in both data sets (Table 1) were broad and variable, and this was desirable in terms of evaluating of the performance of this approach.

The randomization test of mean differences found that the respective non-subsample and subsample basket weights were consistent with having been drawn from the same population (Figs. 3 through 6). Although a minor aspect of the analysis, this was pertinent since, under the mean basket weight protocol, the non-subsample and subsample basket weights are combined to estimate the mean basket weight used to calculate total non-subsample weight.

Simulation Analyses

Simulation results suitably bracket the range of potential real-life scenarios in catch processing on the GOA and AI surveys. The first two scenarios (i.e., fixed n = 30 vs. large and medium catch volumes, respectively) are most similar to customary applications of the counted basket approach in survey operations. The last two scenarios (i.e., variable n = 1 to 50 vs. either large or variable catch volumes, respectively) provide scope to the results and a basis to evaluate the performance of the approach over a broad range of potential combinations of sample size and catch volume.

Scenario-1: Large Volume Catch and Fixed Sample Size

Processing a large catch volume (7.0+ t) using a fixed number (n = 30) of basket weights demonstrates the typical performance of this estimation procedure. A principal conclusion of this simulation is that estimates of non-subsample total weight and mean basket weight are exceedingly statistically robust.

For Pacific ocean perch (Data Set-1), the estimated mean basket weight (30.279 kg) compares favorably to the observed mean weight (30.282 kg) (Fig. 7), as does the estimated mean total non-subsample weight (7,388.11 kg) to the observed (7,388.83 kg). There's high confidence that estimated total non-subsample weight (or mean basket weight) doesn't deviate more than ± 2.70% from observed.

For Atka mackerel (Data Set-2), the estimated mean basket weight (30.856 kg) also compares favorably to the observed mean weight (30.859 kg) (Fig. 8), as does the mean total non-subsample weight estimate (7, 042.13 kg) to the observed (7, 042.64 kg). For these data, there's high confidence that estimated total non-subsample weight (or mean basket weight) doesn't deviate more than ± 2.50% from observed.

Scenario-2: Medium Volume Catch and Fixed Sample Size

Processing a moderate catch volume (~3.0 t) using n = 30 basket weights reveals the gain in robustness of the estimates versus those for a large volume catch. The same conclusion of statistical robustness of the catch estimates is made for results of this simulation.

For the Pacific ocean perch (Data Set-1), the estimated mean basket weight (30.436 kg) is close to the observed mean weight of (30.441 kg) (Fig. 9), as is the estimated mean total non-subsample weight (3,043.58 kg) to the observed (3,044.13 kg). For these estimates, there's high confidence that estimated total non-subsample weight (or mean basket weight) doesn't deviate by more than 2.70% from the observed.

For Atka mackerel (Data Set-2), the estimated mean basket weight (34.282 kg) compares favorably to the observed mean weight (34.275 kg) (Fig. 10), as does the mean total non-subsample weight estimate (3,428.19 kg) to the observed (3,427.52 kg). There's high confidence that estimated total non-subsample weight (or mean basket weight) doesn't deviate more than 2.30% from the observed.

In comparing results of the two previous scenarios, we found that there is negligible gain in the performance measures or accuracy of catch estimates between processing a large volume (~7.0+ t) versus medium volume (~3.0 t) catch using a fixed n = 30. The percent deviations of the lower and upper 95% confidence bounds from the mean are nearly identical between respective catch volumes for the same data, and this bolsters the conclusion that the counted basket approach for catch processing is robust over a range of catch volumes.

Scenario-3: Large Volume Catch and Variable Sample Size

Processing a large volume catch (7.0+ t) using variable sample sizes (1 to 50) illustrates the influence on catch estimation of the choice of sample size. Model diagnostics indicative of statistical robustness generally stabilize at $n > 15$ or 20, and result in catch estimates consistent with those at $n = 30$. We see that the precision of the estimates increases with increasing n, and they stabilize at low levels at $n \geq 20$. The gain in accuracy of the point estimates, or decrease in their variance, are relatively negligible beyond n = 25 to 30. These findings support the conclusion that the counted basket approach provides reliable catch estimates in terms of both accuracy and precision.

The measures of dispersion and CVs of the point estimates of mean basket weight and total non-subsample weight decrease with increasing n for Data Set-1 (Table 2). At $n > 10$, the percent deviation between point estimates and the lower or upper 95% confidence bound is < 5.0%, and < 3.0% for $n \geq 30$ baskets (Fig. 11). At $n = 30$, the percent deviation of the lower and upper 95% confidence bounds is 2.64% and 2.74%, respectively. This is consistent with the < 2.75% deviation between the observed total non-subsample weight and its 95% confidence bounds (Fig. 12). Over the range of n, the CVs on the catch estimates are < 2.0% at $n \geq 20$ and equal 1.63% at $n = 30$ (Fig. 13). Results for the estimate of mean basket weight are equivalent: the 95% confidence bounds are within ± < 1.0 kg of the observed mean at $n \geq 20$, and < 2.75% of the mean at $n = 30$ (Fig. 15).

The measures of dispersion and CVs of the point estimates of mean basket weight and total non-subsample weight decrease with increasing n for Data Set-2 (Table 3). At $n > 10$, the percent deviation between point estimates and the lower or upper 95% confidence bound is < 5.0%, and < 3.0% for $n \geq 30$ baskets (Fig. 16). At $n = 30$, the percent deviation of the lower and upper 95% confidence bounds is 2.64% and 2.74%, respectively. This is consistent with the < 2.75% deviation between the observed total non-subsample weight and its 95% confidence bounds (Fig. 17). Over the range of n, the CVs on the catch estimates are < 2.0% at $n \geq 20$, and equal 1.51% at $n = 30$ (Fig. 18). Results for the estimate of mean basket weight are equivalent: the 95% confidence bounds are within ± < 1.0 kg of the observed mean at $n \geq 20$, and < 2.75% of the mean at $n = 30$ (Fig. 20).

Scenario-4: Variable Volume Catch and Variable Sample Size

This simulation integrates processing variable catch volumes (50 to 200 baskets; ~1.5 t to ~6.0+ t) using variable sample sizes (1 to 50). We've previously seen that that the point estimates of mean basket weight and total non-subsample weight are relatively precise and accurate over the range of sample sizes for large volume catches (Tables 2 and 3). We've also seen a negligible gain in the accuracy of the estimates in processing a large volume (~7.0 t) catch versus a medium volume (~3.0 t) catch using a fixed $n = 30$ (Figs. 7 through 10, and attendant discussion). Given that the catch estimates are conserved over the range of n, we focus on the CV of the estimates to assess the interaction between catch volume and sample size.

The CVs of the catch estimates are seen to decline to low and consistent levels at $n \geq 10$ to 15 (Figs. 21 and 23). For Data Set-1, CV < 2.0% at n = 23, CV < 2.5% at n = 16, CV < 3.0% at n = 12, and CV < 1.7% at the target n = 30 baskets (Fig. 22). For Data Set-2, CV < 2.0% at n = 20, CV < 2.5% at n = 14, CV < 3.0% at n = 11, and CV < 1.5% at the target n = 30 baskets (Fig. 24). The principal conclusion is that, above a moderate threshold sample size ($n \geq 10$ to 15), the coefficients of variation of the catch estimates are acceptably low (< 2.5%) for all catch volumes and, at the target n = 30, the CVs are approximately 1.5% for all data. For the range of potential real-life scenarios of catch volumes on the GOA and AI surveys, selecting 25-30 basket weights to estimate the mean leads to statistically robust catch estimates.

CONCLUSION

The principal finding of the study is that the mean basket weight protocol is exceedingly robust in terms of the estimates of mean basket weight and total non-subsample weight when applied according to the prescriptions in the GOA and AI Scientific Operations Plans. We evaluated the statistical performance of the mean basket weight protocol in catch processing of two dominant species catch hauls on the GOA and AI surveys. Our evaluation focused on the representativeness and robustness of the estimates of mean basket weight and total non-subsample weight relative to observed measures. We simulated a range of catch volumes and sample sizes that encompass real-life scenarios experienced on these surveys. The findings are illustrative of the expected performance of this method of catch processing, and informative in terms of decisions concerning its implementation.

The point estimates of the catch were both accurate and relatively precise over a broad range of catch volumes (~1.5 t to 7.0+ t) and sample sizes (1 to 50 baskets). When the target sample of 30 baskets is used to estimate the mean, we found an almost negligible gain in the accuracy of the estimates for a large volume catch (7.0+ t) relative to a medium volume (~3.0 t) catch. Over the range of catch volumes, the coefficients of variation of the catch estimates are acceptably low (< 2.0%) for n ≈ 20 baskets. At the target sample size of n = 30, the CVs were all ≈ 1.5% for all levels of catch volume generated using these data.

A particular interest was evaluating how well this method performed in estimating non-subsample total catch weight. The first two simulations were most similar to standard applications of the method on the survey. We found that both the lower and upper 95% confidence bounds on the non-subsample total weight estimate deviated by an average 2.42% from the observed total weight and ranged from 2.25% to 2.70%. The average percent deviations were 2.57% for the large volume catch, and 2.26% for the medium volume catch. Comparable results were found for the latter two simulations at n = 30. Considering that these data were not collected consistent with design requirements in the operation plan, these results are likely overestimates of the expected percent deviations of the 95% confidence bounds from the observed. While we made no attempt to assess the magnitude of the overestimation, we conclude that the results would be even more robust if the mean basket weight protocol had been applied to these hauls, perhaps < 2.0% deviations.

A central conclusion of this work is that there's high confidence that the estimates of non-subsample mean basket weight and total catch weight will not deviate, on average, by more than 2.5% from observed weights, and undoubtedly less if the method is applied consistent with design specifications. For example, 2.5% deviations for the 95% confidence bounds equates to ± 25 kg for a 1.0 t catch, ± 50 kg for a 2.0 t catch, ± 125 kg for a 5.0 t catch, and ± 175 kg for a 7.0 t catch.

It's useful to view these results in a broader survey sampling context. The ultimate interest to the stock assessment process is the uncertainty of a stratum population estimate in which a large volume catch of a dominant species occurred. The additional uncertainty introduced to the estimate by applying this catch processing method will be small compared to the additional uncertainty from a single large volume catch. Depending on the number of tows in a stratum, one large volume catch can increase the CV of the population estimate much more than the ≈ 1.5% deviation on the catch estimated in this study. The large increase in CV resulting from one large volume catch will overwhelm the additional uncertainty from not completely sampling the catch. In addition, if the time expended in whole-hauling large volume catches results in fewer tows completed in a stratum, the uncertainty in population estimation may be larger than any uncertainty added by applying this catch estimation method.

Our goal was to address the validity of the mean basket weight protocol in catch processing. We evaluated its statistical performance in terms of the representativeness and robustness of the estimates of mean basket weight, and its extrapolation to total non-subsample weight. We found that this method is exceedingly robust in terms of catch estimation relative to whole-haul processing. We conclude that whole-hauling large volume catches is not statistically necessary for precise estimates of total catch weight. Further, application of the mean basket weight method can be used to mitigate safety and ergonomic issues associated with total catch processing of very large catches.

CITATIONS

Anon. 2015. Scientific Operations Plan: 2015 Gulf of Alaska Bottom Trawl Survey Cruise 2015-01. Resource Assessment and Conservation Engineering Division, Alaska Fisheries Science Center, NOAA, National Marine Fisheries Service, Seattle, WA 98115. 89 p.

R Core Team (2013). **R**: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Stauffer, G. 2004. NOAA protocols for groundfish bottom trawl surveys of the Nation's fisheries resources. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-F/SPO-65, 205 p.

Table 1. --Summary statistics for experimental data sets. Shown are respective number of baskets (n), total basket weight (kg), mean basket weight (kg), coefficient of variation of the mean, range of basket weights (kg), and the upper and lower 95% confidence bounds on the mean basket weight (kg).

Statistic:	**Data set summary statistics**					
	Data Set-1: Pacific ocean perch			**Data Set-2: Atka mackerel**		
	Non-subsample (Raw)	**Non-subsample (-Outliers)**	**Subsample**	**Non-subsample (Raw)**	**Non-subsample (-Outliers)**	**Subsample**
n:	247	244	5	213	208	2
Total Weight:	7,438.42	7,388.83	151.46	7,136.36	7,042.64	70.76
Mean:	30.12	30.28	30.29	33.50	33.86	35.38
CV (%):	0.68	0.61	2.89	0.77	0.62	7.74
Range [min,max]:	[13.93, 38.46]	[21.43, 38.46]	[27.84, 32.32]	[13.24, 40.88]	[25.29, 40.88]	[32.64, 38.12]
95% CI [lb,ub]:	[29.71, 30.52]	[29.92, 30.64]	[28.58, 32.01]	[33.00, 34.01]	[33.45, 34.27]	[30.01, 40.75]

Table 2. -- Performance measures and model output of mean basket weight and total non-subsample weight estimates versus sample sizes n = 1 to 50 for Data Set-1. Mean basket weight measures are: the estimate of the mean (Est.) and its standard deviation (SD), and the lower (LB) and upper (UB) 95% confidence bounds. Total non-subsample weight measures are: the total weight estimate (Est.) and its standard error (SE), the lower (LB) and upper (UB) 95% confidence bounds, the range of the 95% CI (CI Range), the lower (Δ LB) and upper (Δ UB) bound deviations from the mean, and the lower (% LB) and upper (% UB) bound deviations as a percent of the mean.

Pacific ocean perch – Model output for sample sizes [1-50] from 244 Non-subsample basket weights													
	Mean basket weight (kg)				Total Non-subsample weight (kg)								
n	Est.	SD	LB	UB	Est.	SE	LB	UB	CI Range	Δ LB	Δ UB	% LB	% UB
1	30.28	2.89	25.84	35.14	7387.39	705.17	6304.96	8574.16	2269.20	1082.43	1186.77	14.65	16.06
2	30.31	2.02	26.99	33.64	7394.44	492.44	6585.56	8208.16	1622.60	808.88	813.72	10.94	11.00
3	30.28	1.62	27.61	32.93	7389.01	395.81	6736.03	8034.11	1298.08	652.98	645.10	8.84	8.73
4	30.28	1.42	27.91	32.61	7389.44	346.79	6810.65	7955.62	1144.97	578.79	566.18	7.83	7.66
5	30.28	1.28	28.19	32.37	7387.37	312.58	6878.85	7897.79	1018.94	508.52	510.43	6.88	6.91
6	30.26	1.16	28.38	32.15	7382.70	282.17	6923.91	7845.41	921.51	458.79	462.72	6.21	6.27
7	30.27	1.06	28.54	32.04	7386.75	258.73	6963.06	7817.76	854.70	423.68	431.01	5.74	5.83
8	30.29	0.99	28.66	31.92	7389.65	242.57	6991.82	7789.09	797.27	397.83	399.44	5.38	5.41
9	30.29	0.94	28.76	31.83	7389.79	228.70	7016.36	7767.60	751.25	373.44	377.81	5.05	5.11
10	30.27	0.89	28.82	31.73	7384.96	216.02	7032.08	7741.63	709.55	352.88	356.67	4.78	4.83
11	30.28	0.84	28.88	31.66	7388.74	205.35	7047.61	7725.48	677.88	341.13	336.75	4.62	4.56
12	30.28	0.80	28.97	31.60	7387.81	195.78	7068.48	7710.40	641.92	319.34	322.59	4.32	4.37
13	30.30	0.77	29.03	31.55	7392.15	187.74	7084.26	7697.64	613.38	307.89	305.48	4.17	4.13
14	30.27	0.75	29.02	31.50	7385.01	181.94	7081.40	7687.05	605.64	303.61	302.04	4.11	4.09
15	30.28	0.72	29.10	31.47	7388.04	175.23	7099.75	7678.84	579.09	288.29	290.80	3.90	3.94
16	30.29	0.69	29.15	31.44	7390.50	169.45	7113.21	7670.45	557.24	277.29	279.95	3.75	3.79
17	30.28	0.68	29.16	31.38	7389.31	164.72	7115.90	7657.87	541.97	273.41	268.56	3.70	3.63
18	30.28	0.65	29.22	31.33	7388.63	158.56	7129.41	7645.47	516.06	259.22	256.84	3.51	3.48
19	30.28	0.64	29.24	31.34	7389.40	155.38	7133.66	7646.70	513.04	255.74	257.30	3.46	3.48
20	30.29	0.62	29.29	31.31	7390.69	150.24	7146.15	7640.37	494.22	244.54	249.68	3.31	3.38
21	30.28	0.60	29.28	31.28	7388.95	146.77	7145.48	7633.02	487.54	243.47	244.07	3.30	3.30
22	30.28	0.58	29.32	31.24	7388.83	141.93	7154.30	7621.45	467.15	234.53	232.62	3.17	3.15
23	30.27	0.57	29.34	31.22	7385.94	139.18	7159.49	7618.53	459.04	226.45	232.59	3.07	3.15
24	30.29	0.56	29.36	31.20	7389.71	135.99	7164.04	7613.51	449.47	225.66	223.81	3.05	3.03
25	30.29	0.54	29.39	31.19	7390.30	132.74	7170.67	7611.24	440.57	219.63	220.94	2.97	2.99
26	30.28	0.53	29.40	31.16	7388.47	130.40	7174.63	7602.76	428.13	213.83	214.29	2.89	2.90
27	30.27	0.53	29.41	31.15	7387.05	128.28	7176.31	7599.43	423.11	210.73	212.38	2.85	2.88
28	30.27	0.50	29.44	31.10	7385.16	123.16	7183.10	7589.27	406.17	202.06	204.11	2.74	2.76
29	30.28	0.50	29.45	31.12	7389.11	122.97	7184.71	7593.45	408.74	204.40	204.34	2.77	2.77
30	30.29	0.49	29.49	31.11	7389.86	120.42	7194.67	7591.98	397.31	195.20	202.12	2.64	2.74

Table 2. -- Continued.

Pacific ocean perch - Model output for sample sizes [1-50] from 244 Non-subsample basket weights													
	Mean basket weight (kg)				Total Non-subsample weight (kg)								
n	Est.	SD	LB	UB	Est.	SE	LB	UB	CI Range	Δ LB	Δ UB	% LB	% UB
31	30.28	0.48	29.49	31.09	7389.16	117.65	7194.93	7585.33	390.40	194.23	196.17	2.63	2.65
32	30.28	0.47	29.51	31.04	7387.72	114.66	7200.59	7574.37	373.78	187.13	186.65	2.53	2.53
33	30.28	0.47	29.51	31.05	7388.20	114.47	7199.63	7576.27	376.65	188.58	188.07	2.55	2.55
34	30.28	0.46	29.54	31.03	7388.41	111.13	7206.68	7570.39	363.70	181.72	181.98	2.46	2.46
35	30.28	0.45	29.56	31.02	7389.16	108.85	7211.45	7568.88	357.43	177.70	179.72	2.40	2.43
36	30.29	0.44	29.57	31.02	7390.84	107.35	7213.93	7567.86	353.94	176.91	177.03	2.39	2.40
37	30.28	0.43	29.56	31.00	7389.06	105.83	7213.83	7563.60	349.78	175.24	174.54	2.37	2.36
38	30.29	0.43	29.59	30.99	7389.88	104.34	7219.45	7562.01	342.56	170.43	172.13	2.31	2.33
39	30.28	0.42	29.60	30.98	7388.46	101.67	7223.21	7558.06	334.84	165.25	169.60	2.24	2.30
40	30.29	0.41	29.61	30.97	7389.62	101.17	7224.41	7557.17	332.76	165.21	167.55	2.24	2.27
41	30.28	0.40	29.62	30.94	7388.02	98.16	7227.22	7549.00	321.78	160.80	160.99	2.18	2.18
42	30.28	0.40	29.63	30.95	7388.94	98.02	7229.02	7551.16	322.14	159.92	162.22	2.16	2.20
43	30.28	0.40	29.62	30.94	7389.20	97.67	7227.79	7550.27	322.48	161.41	161.07	2.18	2.18
44	30.28	0.40	29.63	30.93	7388.62	96.57	7230.83	7547.14	316.31	157.79	158.52	2.14	2.15
45	30.28	0.39	29.65	30.92	7389.02	94.51	7234.22	7545.13	310.91	154.80	156.11	2.10	2.11
46	30.29	0.38	29.66	30.91	7389.86	93.20	7236.09	7542.52	306.43	153.77	152.66	2.08	2.07
47	30.28	0.38	29.66	30.90	7389.07	91.93	7237.61	7540.12	302.51	151.46	151.05	2.05	2.04
48	30.28	0.37	29.67	30.87	7388.29	90.19	7239.28	7532.89	293.61	149.01	144.60	2.02	1.96
49	30.28	0.37	29.69	30.89	7388.25	89.33	7243.16	7536.01	292.85	145.08	147.77	1.96	2.00
50	30.28	0.36	29.69	30.87	7388.77	87.60	7244.60	7533.01	288.41	144.17	144.24	1.95	1.95

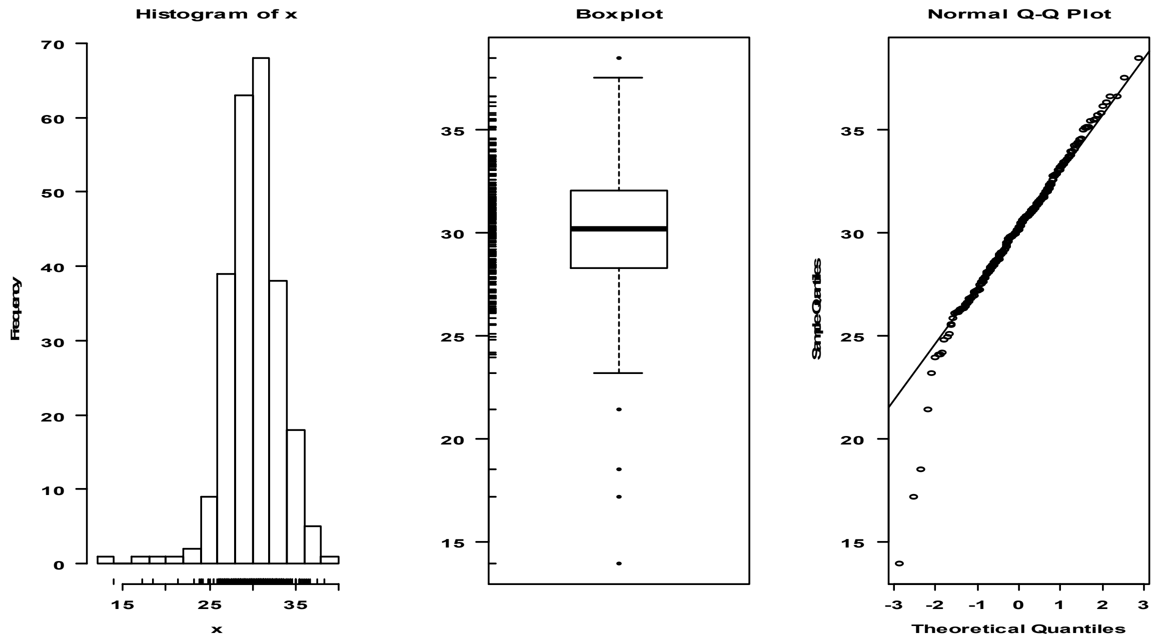
Table 3. --Performance measures and model output of mean basket weight and total non-subsample weight estimates versus sample sizes n = 1 to 50 for Data Set-2. Mean basket weight measures are: the estimate of the mean (Est.) and its standard deviation (SD), and the lower (LB) and upper (UB) 95% confidence bounds. Total non-subsample weight measures are: the total weight estimate (Est.) and its standard error (SE), the lower (LB) and upper (UB) 95% confidence bounds, the range of the 95% CI (CI Range), the lower (Δ LB) and upper (Δ UB) bound deviations from the mean, and the lower (% LB) and upper (% UB) bound deviations as a percent of the mean.

Atka mackerel - Model output for sample sizes [1-50] from 208 Non-subsample basket weights														
	Mean basket weight (kg)				Total Non-subsample weight (kg)									
n	Est.	SD	LB	UB	Est.	SE	LB	UB	CI Range	Δ LB	Δ UB	% LB	% UB
1	33.83	3.01	28.11	38.36	7037.42	625.09	5846.88	7978.88	2132.00	1190.54	941.46	16.92	13.38
2	33.89	2.11	30.24	37.20	7049.05	438.70	6288.88	7737.60	1448.72	760.17	688.55	10.78	9.77
3	33.86	1.73	30.87	36.64	7042.26	359.65	6420.27	7621.12	1200.85	621.99	578.86	8.83	8.22
4	33.88	1.49	31.37	36.26	7047.79	309.04	6523.92	7542.60	1018.68	523.87	494.81	7.43	7.02
5	33.86	1.32	31.62	35.97	7042.56	273.89	6576.13	7481.34	905.22	466.43	438.79	6.62	6.23
6	33.86	1.22	31.82	35.85	7043.07	252.93	6619.25	7456.45	837.20	423.82	413.38	6.02	5.87
7	33.85	1.11	31.99	35.64	7041.83	231.32	6653.92	7412.53	758.61	387.91	370.70	5.51	5.26
8	33.87	1.04	32.12	35.54	7044.75	216.80	6680.18	7393.10	712.92	364.57	348.35	5.18	4.94
9	33.86	0.98	32.23	35.46	7042.99	203.23	6704.07	7375.22	671.15	338.92	332.23	4.81	4.72
10	33.85	0.92	32.29	35.33	7040.19	191.54	6716.94	7349.47	632.53	323.24	309.29	4.59	4.39
11	33.86	0.89	32.39	35.29	7042.97	185.18	6736.74	7339.56	602.82	306.23	296.59	4.35	4.21
12	33.87	0.85	32.44	35.23	7044.51	176.56	6747.17	7327.49	580.32	297.34	282.98	4.22	4.02
13	33.86	0.80	32.54	35.16	7042.75	165.88	6767.84	7312.96	545.12	274.91	270.21	3.90	3.84
14	33.86	0.78	32.56	35.12	7042.62	162.44	6772.18	7304.37	532.18	270.43	261.75	3.84	3.72
15	33.85	0.74	32.61	35.04	7041.02	154.65	6782.88	7287.90	505.02	258.14	246.89	3.67	3.51
16	33.85	0.73	32.64	35.03	7041.28	150.98	6790.03	7286.37	496.34	251.25	245.09	3.57	3.48
17	33.85	0.69	32.71	34.97	7041.82	142.55	6803.92	7273.88	469.96	237.90	232.06	3.38	3.30
18	33.85	0.68	32.72	34.96	7041.74	141.22	6806.22	7271.10	464.88	235.52	229.36	3.34	3.26
19	33.86	0.65	32.79	34.94	7043.51	136.07	6819.33	7266.53	447.20	224.18	223.02	3.18	3.17
20	33.87	0.64	32.81	34.91	7044.08	132.42	6824.17	7261.80	437.63	219.92	217.72	3.12	3.09
21	33.86	0.63	32.84	34.89	7043.69	130.07	6830.03	7256.92	426.90	213.67	213.23	3.03	3.03
22	33.85	0.61	32.85	34.84	7041.07	126.30	6832.04	7247.67	415.62	209.03	206.60	2.97	2.93
23	33.86	0.60	32.86	34.83	7042.33	124.52	6835.51	7245.45	409.94	206.81	203.13	2.94	2.88
24	33.85	0.58	32.89	34.78	7040.22	119.84	6840.95	7234.59	393.64	199.27	194.37	2.83	2.76
25	33.87	0.56	32.93	34.79	7044.80	116.80	6850.36	7236.74	386.38	194.44	191.94	2.76	2.72
26	33.86	0.55	32.96	34.76	7043.37	114.60	6855.28	7230.48	375.20	188.09	187.11	2.67	2.66
27	33.85	0.54	32.95	34.73	7039.84	112.81	6853.14	7223.30	370.16	186.70	183.47	2.65	2.61
28	33.86	0.52	32.99	34.71	7042.84	108.95	6862.44	7219.68	357.24	180.40	176.84	2.56	2.51
29	33.86	0.52	33.00	34.71	7041.86	108.47	6863.43	7219.18	355.75	178.43	177.32	2.53	2.52
30	33.86	0.51	33.01	34.70	7042.37	106.05	6866.29	7217.53	351.24	176.08	175.16	2.50	2.49

Table 3. -- Continued.

Atka Mackerel - Model Output for Sample Sizes [1-50] from 208 Non-Subsample Basket Weights													
	Mean Basket Weight (kg)				Total Non-Subsample Weight (kg)								
n	Est.	SD	LB	UB	Est.	SE	LB	UB	CI Range	Δ LB	Δ UB	% LB	% UB
31	33.86	0.50	33.05	34.68	7043.59	104.32	6873.73	7212.77	339.04	169.86	169.18	2.41	2.40
32	33.86	0.49	33.05	34.66	7042.99	102.32	6874.08	7210.00	335.92	168.91	167.01	2.40	2.37
33	33.86	0.48	33.05	34.63	7042.08	99.29	6875.28	7203.04	327.76	166.80	160.96	2.37	2.29
34	33.86	0.47	33.09	34.64	7042.14	97.76	6882.48	7204.08	321.60	159.66	161.94	2.27	2.30
35	33.86	0.47	33.09	34.62	7043.85	97.02	6882.30	7201.20	318.89	161.55	157.35	2.29	2.23
36	33.86	0.46	33.12	34.62	7043.73	95.06	6889.08	7201.31	312.23	154.66	157.58	2.20	2.24
37	33.86	0.44	33.13	34.59	7042.72	91.84	6890.08	7194.61	304.52	152.64	151.89	2.17	2.16
38	33.86	0.44	33.13	34.57	7042.32	91.08	6891.86	7190.07	298.21	150.46	147.75	2.14	2.10
39	33.86	0.44	33.13	34.57	7042.27	91.46	6890.93	7190.88	299.95	151.34	148.61	2.15	2.11
40	33.86	0.42	33.15	34.55	7042.04	88.31	6895.56	7186.24	290.68	146.48	144.20	2.08	2.05
41	33.86	0.42	33.15	34.56	7041.86	88.33	6896.01	7187.52	291.50	145.85	145.65	2.07	2.07
42	33.86	0.42	33.15	34.54	7042.13	87.21	6896.09	7184.42	288.33	146.04	142.29	2.07	2.02
43	33.86	0.41	33.17	34.53	7042.45	85.44	6899.12	7182.43	283.32	143.33	139.98	2.04	1.99
44	33.85	0.41	33.18	34.52	7041.79	84.76	6900.64	7179.59	278.96	141.16	137.80	2.00	1.96
45	33.86	0.39	33.21	34.50	7042.08	81.94	6907.13	7176.18	269.06	134.95	134.11	1.92	1.90
46	33.86	0.39	33.21	34.50	7042.12	81.54	6906.96	7175.28	268.32	135.17	133.15	1.92	1.89
47	33.85	0.39	33.21	34.49	7041.61	81.28	6908.34	7174.19	265.84	133.27	132.58	1.89	1.88
48	33.86	0.38	33.22	34.49	7042.66	79.79	6909.59	7172.92	263.34	133.07	130.26	1.89	1.85
49	33.86	0.37	33.23	34.47	7042.97	77.54	6912.82	7169.97	257.16	130.16	127.00	1.85	1.80
50	33.86	0.37	33.23	34.47	7043.02	77.97	6912.30	7168.80	256.51	130.72	125.79	1.86	1.79

(a)



(b)

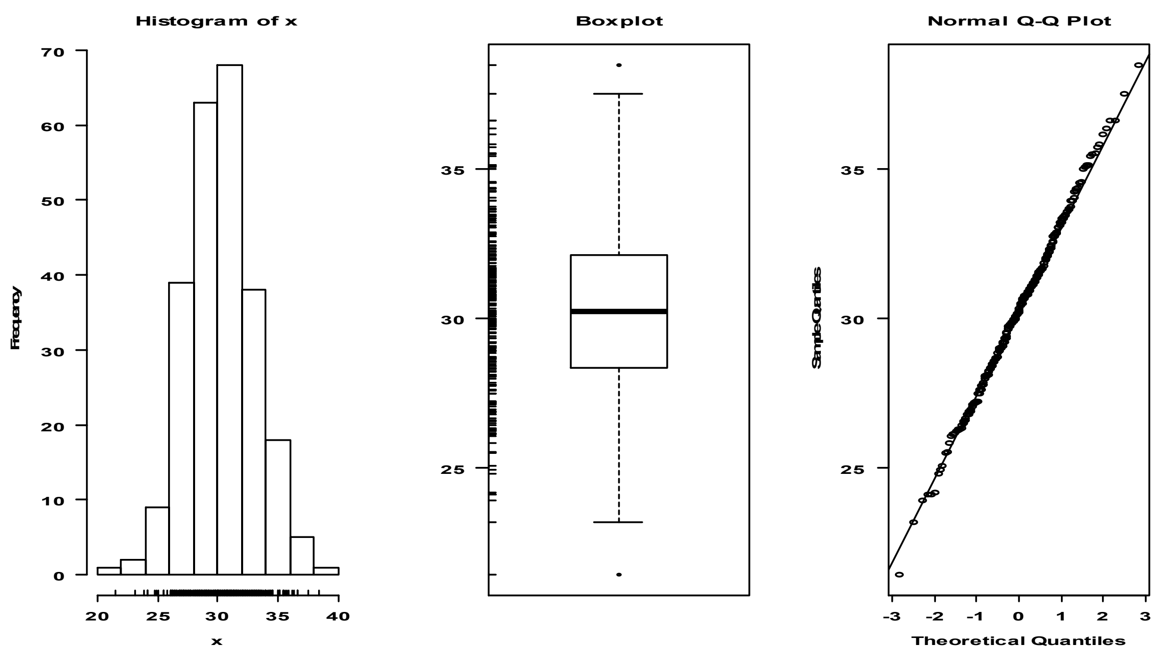


Figure 1. -- Histogram, boxplot and normal Q-Q plot of non-subsample basket weights for Data Set-1 (Pacific ocean perch). Panel (a) is raw n = 247 basket weights, and (b) is n = 244 after removal of three outliers.

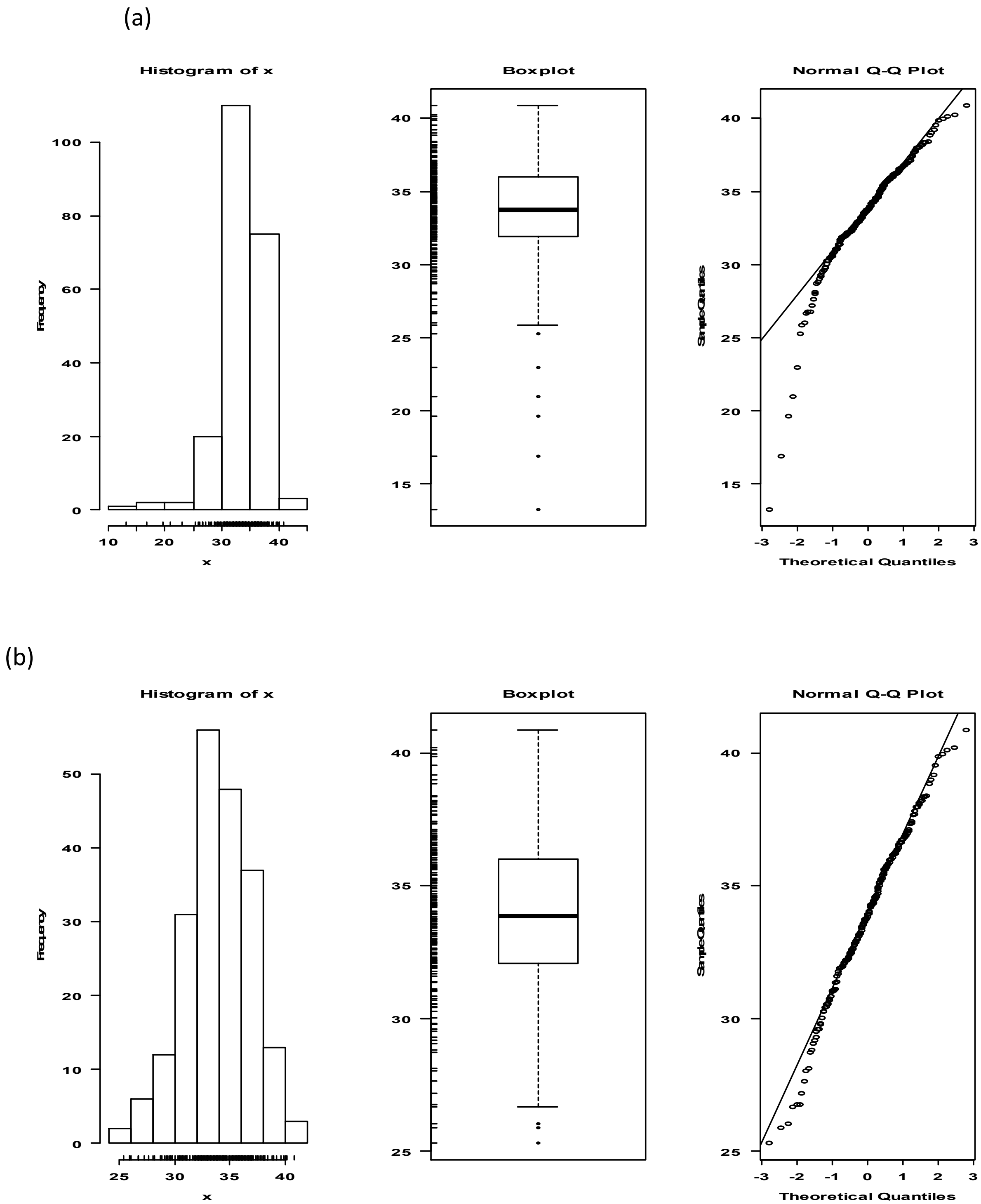


Figure 2. -- Histogram, boxplot and normal Q-Q plot of non-subsample basket weights for Data Set-2 (Atka mackerel). Panel (a) is raw n = 213 basket weights, and (b) is n = 208 after removal of five outliers.

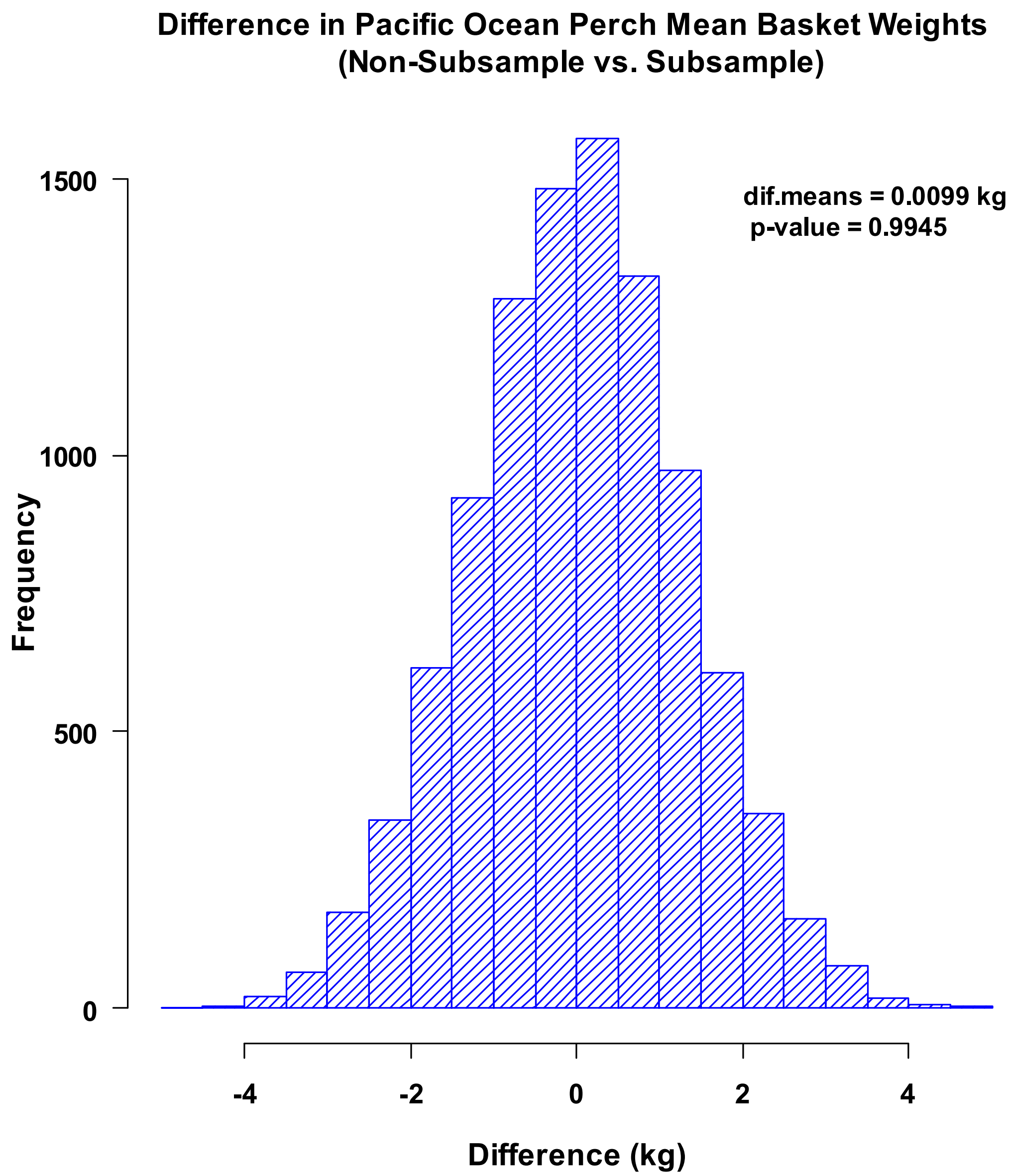


Figure 3. -- The distribution of 10,000 mean differences in non-subsample versus subsample mean basket weights for Data Set-1. T_{OBS} = 0.0099 kg and p = 0.9945.

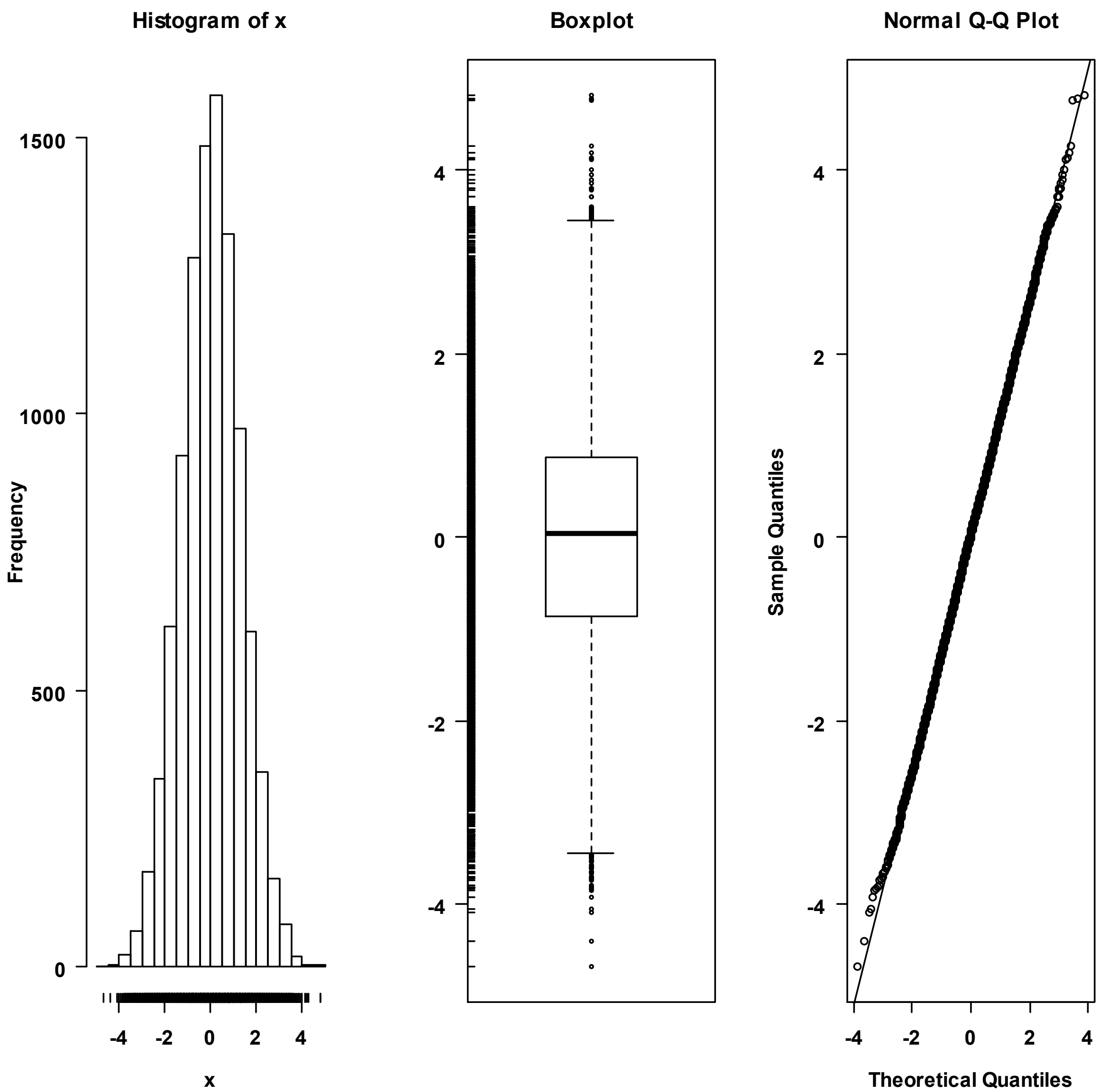


Figure 4. -- Histogram, boxplot and normal Q-Q plots of results of the randomization test of mean differences in non-subsample versus subsample mean basket weights for Data Set-1 (Pacific ocean perch). The data are 10,000 bootstrapped mean differences where, T_{OBS} = 0.0099 kg and p = 0.9945.

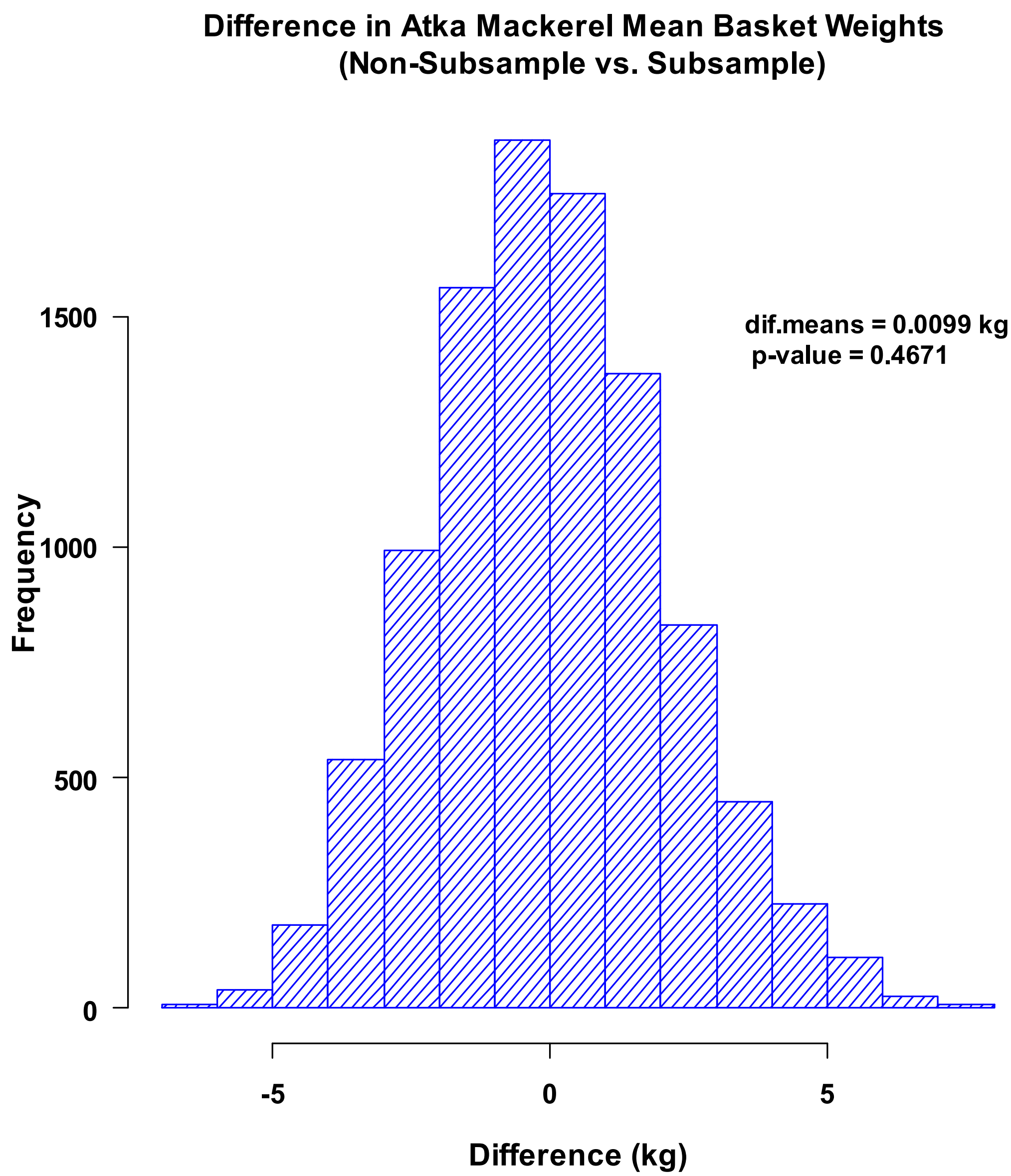


Figure 5. -- The distribution of 10,000 mean differences in non-subsample versus subsample mean basket weights for Data Set-2. T_{OBS} = 0.0099 kg and p = 0.4671.

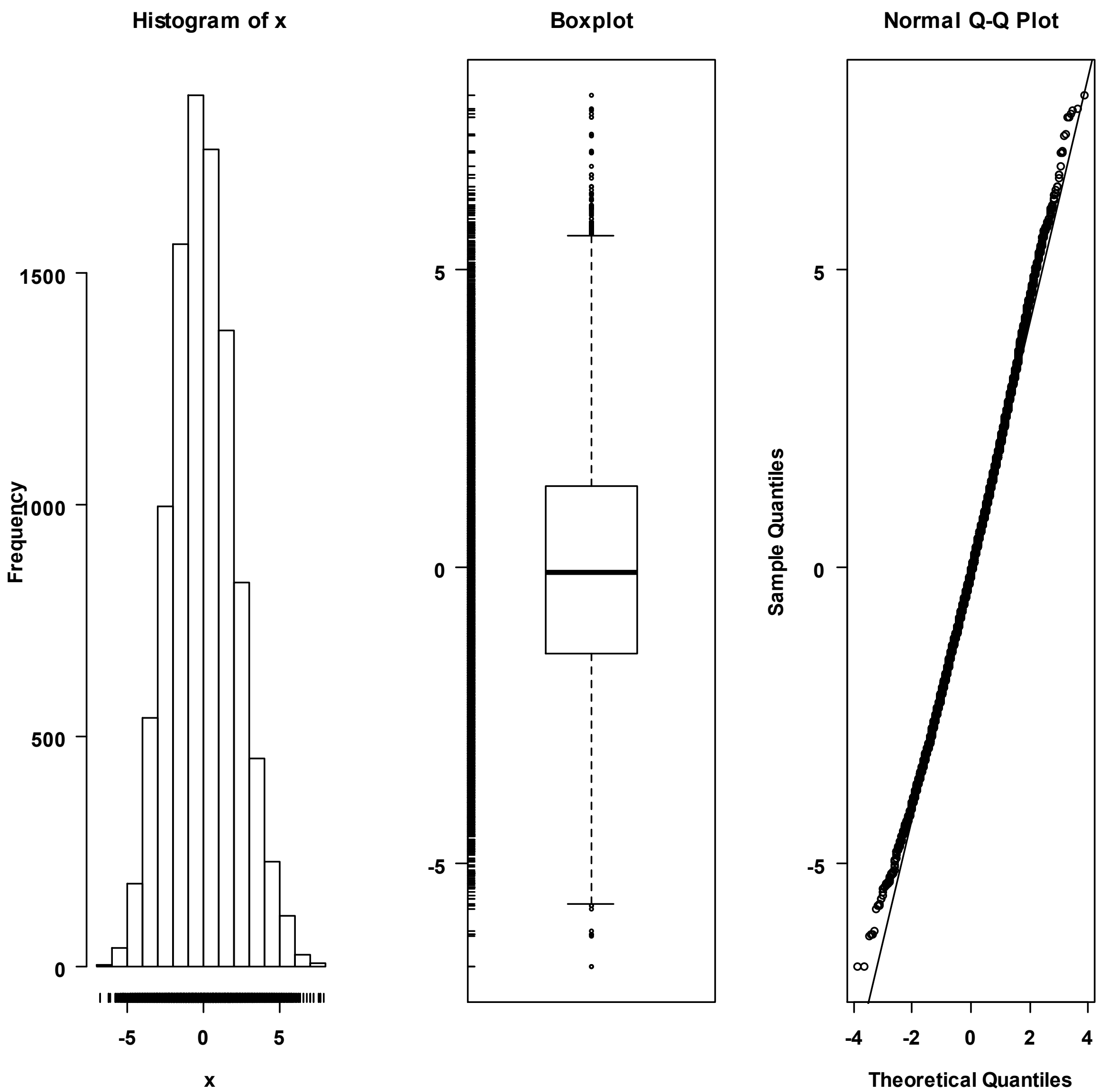


Figure 6. -- Histogram, boxplot and normal Q-Q plots of results of the randomization test of mean differences in non-subsample versus subsample mean basket weights for Data Set-2 (Atka mackerel). The data are 10,000 bootstrapped mean differences where, T_{OBS} = 0.0099 kg and p = 0.4671.

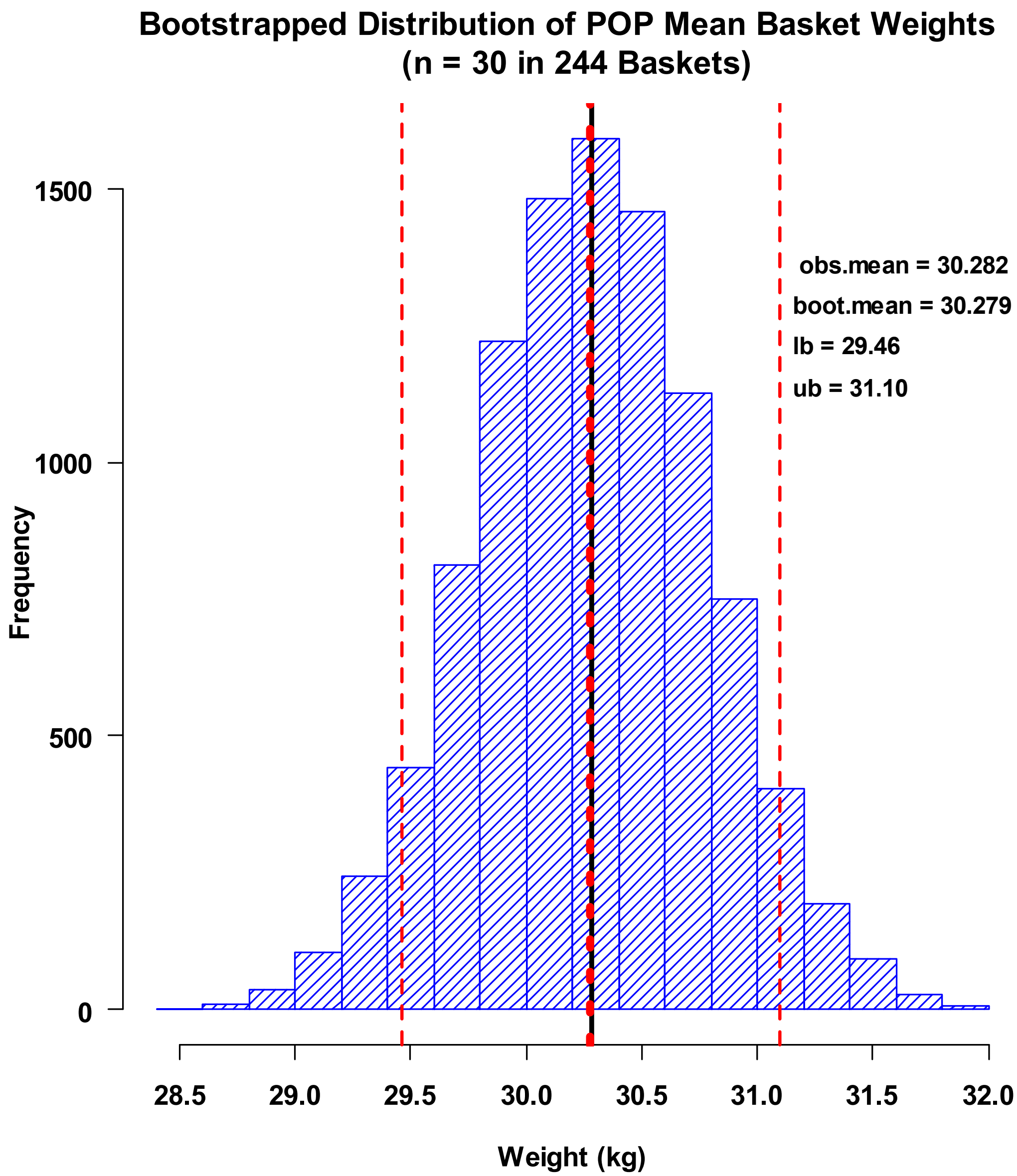


Figure 7. -- The distribution of the 10,000 mean non-subsample mean basket weights for Data Set-1. Shown are the mean of the bootstrapped means (bold dotted red), the upper and lower 95% confidence bounds (dashed red) of the bootstrapped mean, and the observed mean of the 244 non-subsample basket weights (solid black).

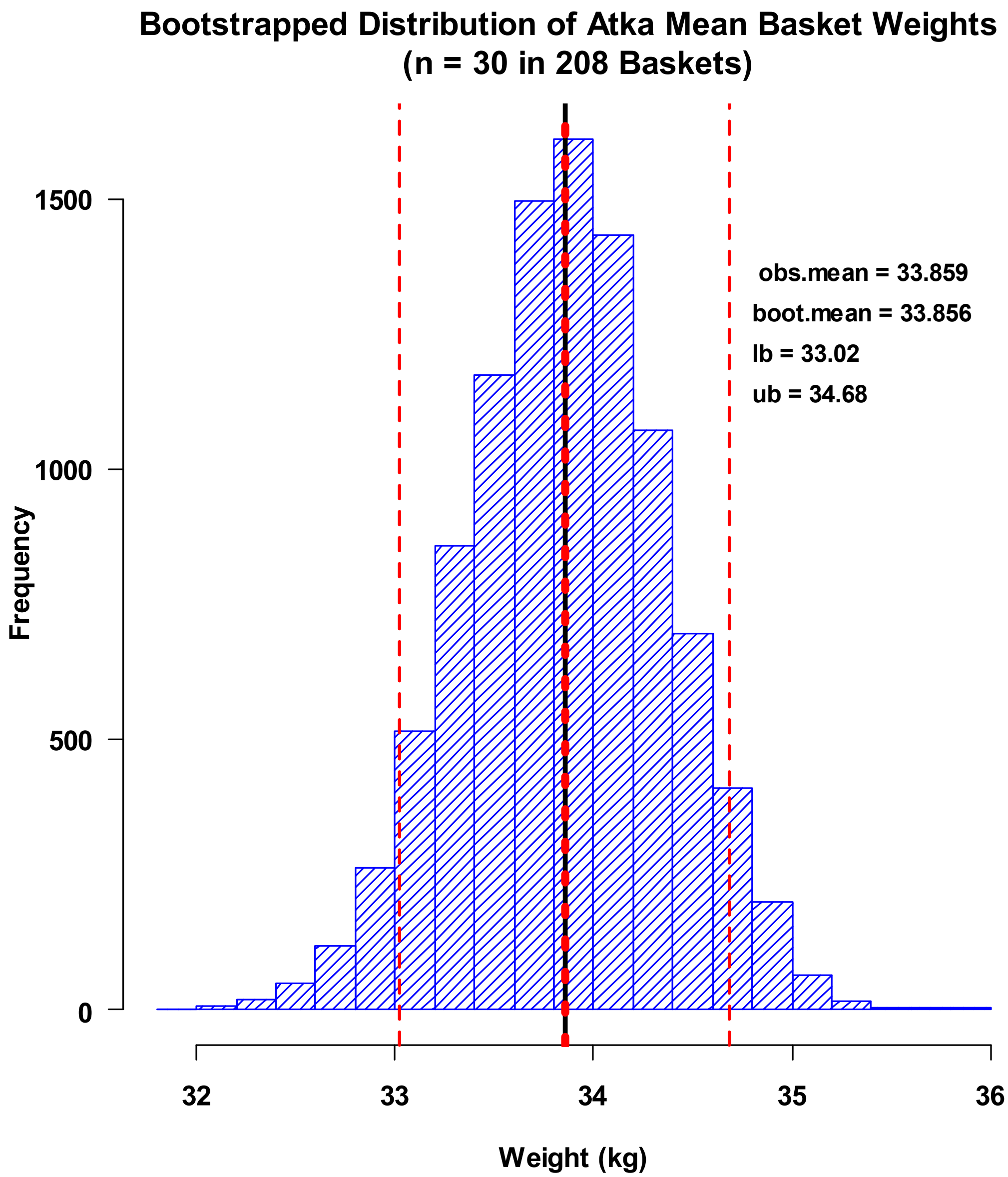


Figure 8. -- The distribution of the 10,000 mean non-subsample mean basket weights for Data Set-2. Shown are the mean of the bootstrapped means (bold dotted red), the upper and lower 95% confidence bounds (dashed red) of the bootstrapped mean, and the observed mean of the 208 non-subsample basket weights (solid black).

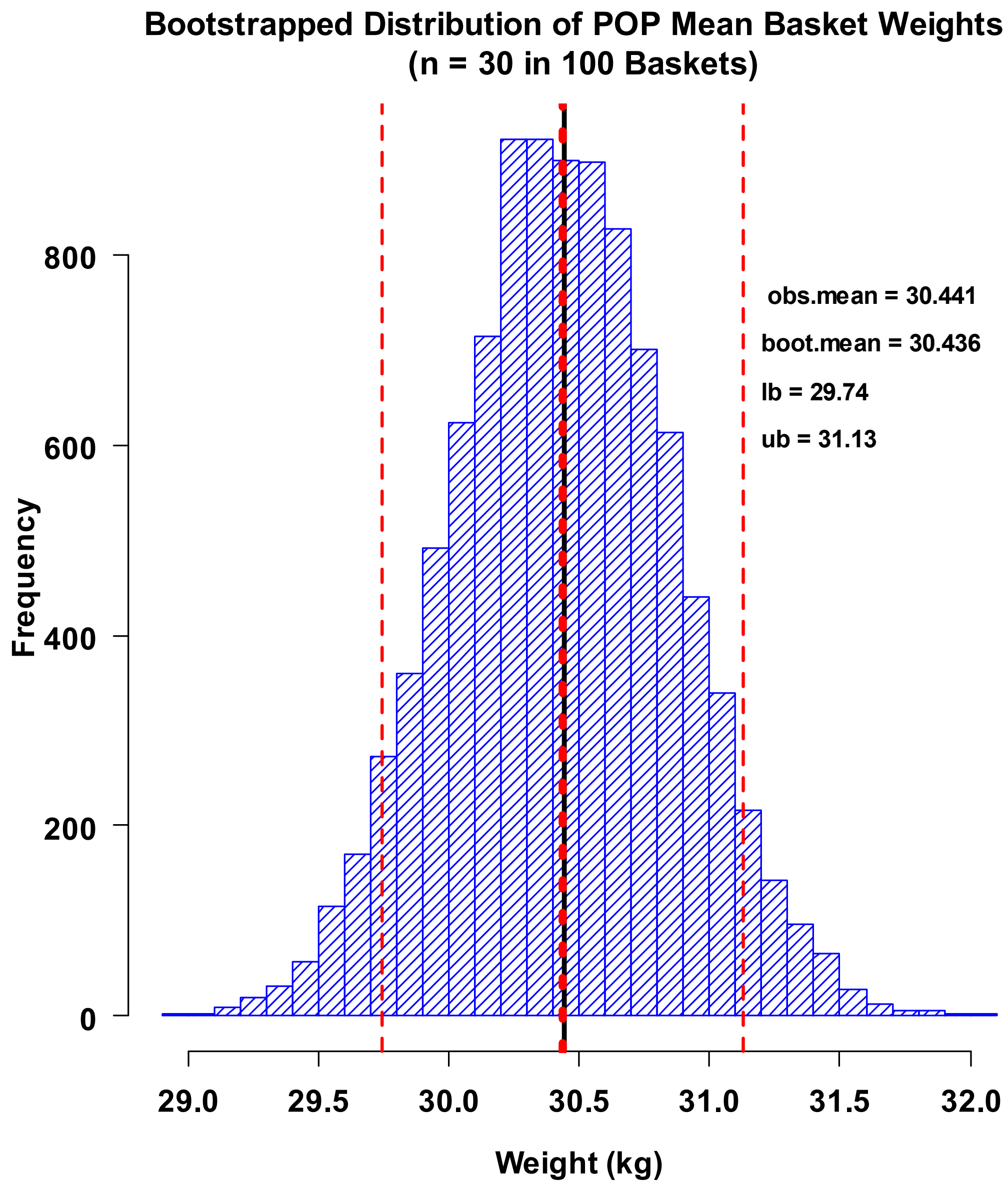


Figure 9. -- The distribution of the 10,000 mean non-subsample mean basket weights for Data Set-1. Shown are the mean of the bootstrapped means (bold dotted red), the upper and lower 95% confidence bounds (dashed red) of the bootstrapped mean, and the observed mean of the 100 non-subsample basket weights (solid black).

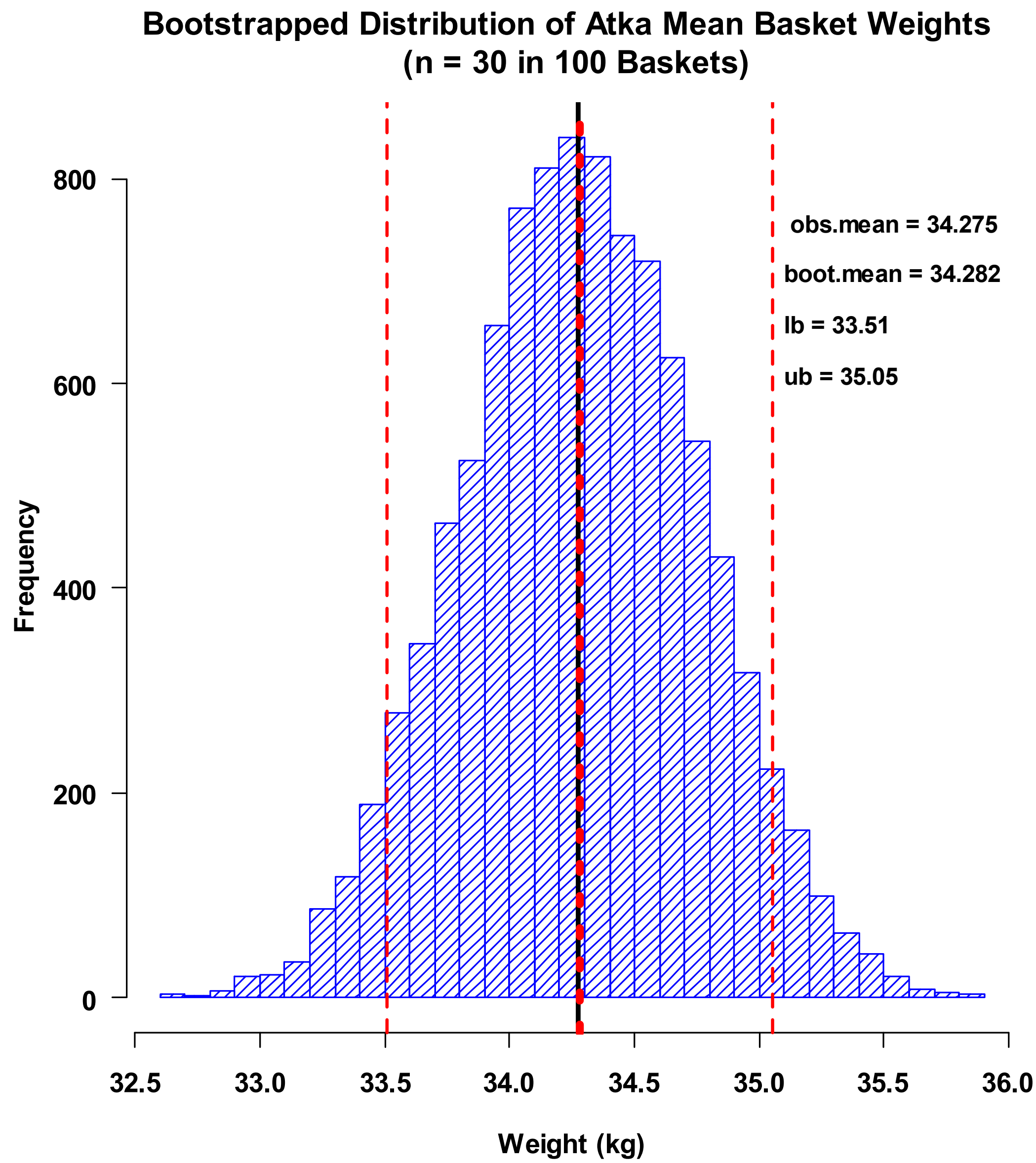


Figure 10. -- The distribution of the 10,000 mean non-subsample mean basket weights for Data Set-2. Shown are the mean of the bootstrapped means (bold dotted red), the upper and lower 95% confidence bounds (dashed red) of the bootstrapped mean, and the observed mean of the 100 non-subsample basket weights (solid black).

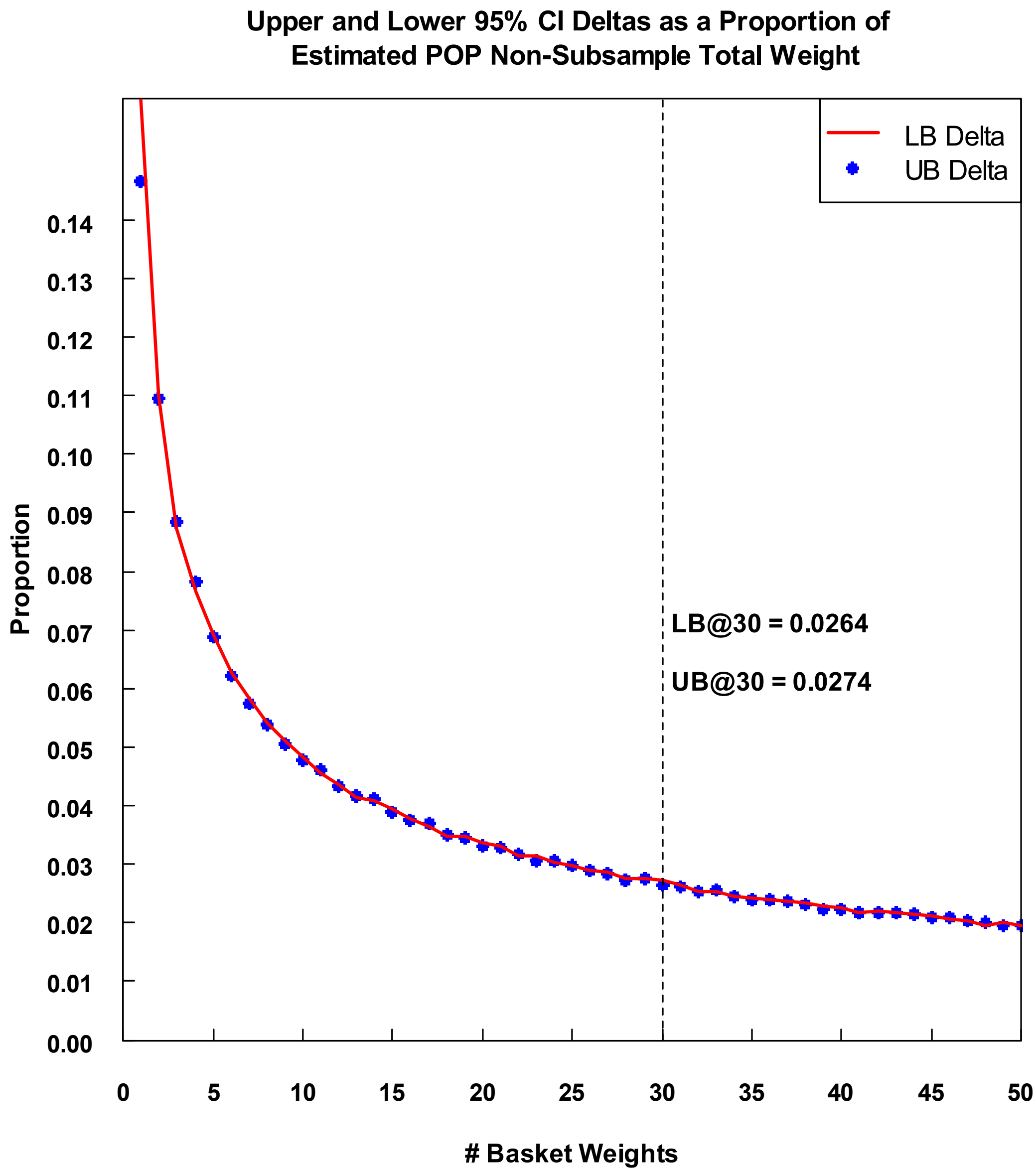


Figure 11. -- Lower and upper 95% confidence bounds deltas (deviations) as a proportion of the estimated non-subsample total weight versus sample size for Data Set-1. Indicated are the lower and upper proportions for sample size n = 30 (dashed line).

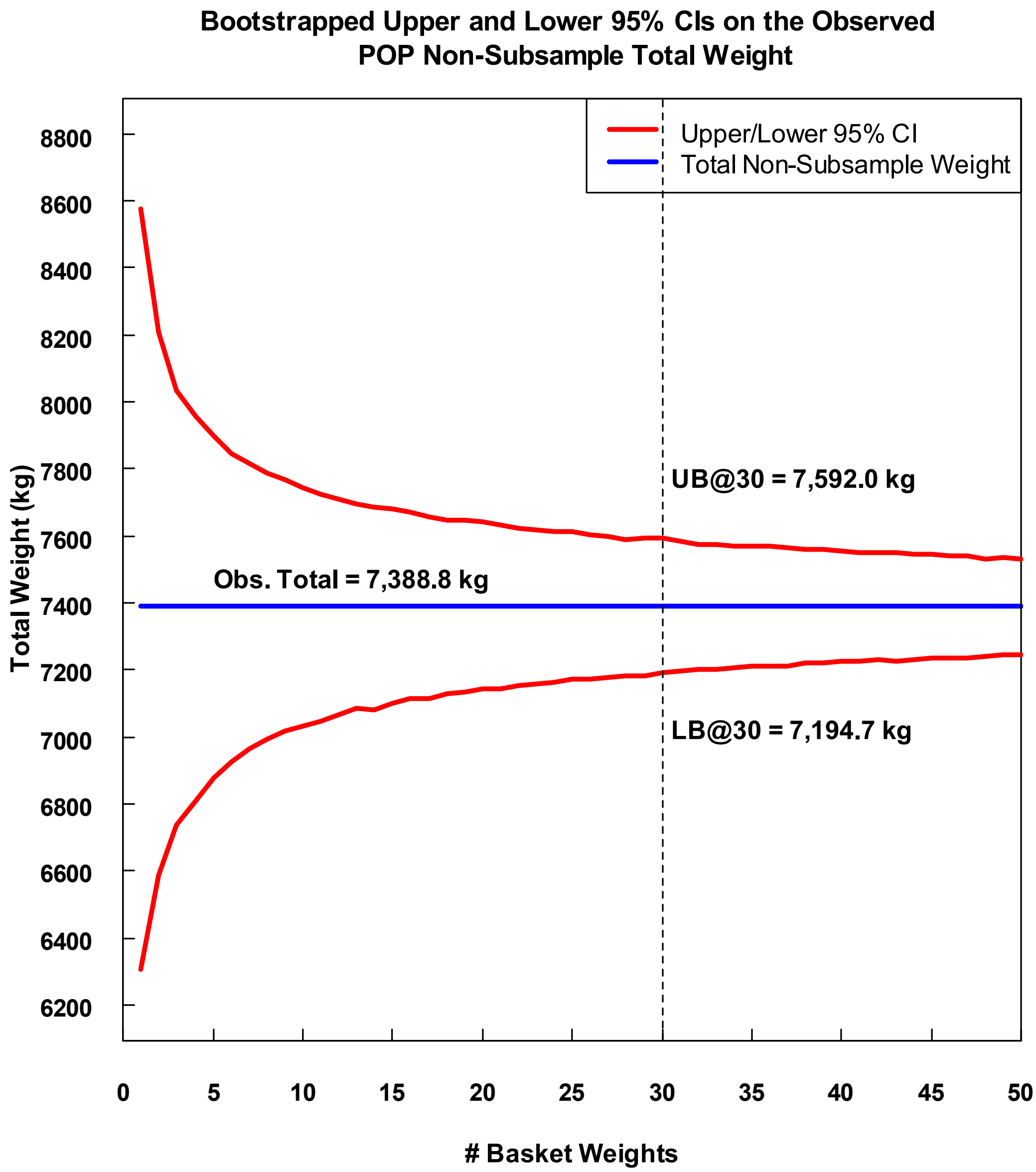


Figure 12. -- Lower and upper 95% confidence bounds on the non-subsample total weight estimate versus sample size for Data Set-1. Also shown are the observed total non-subsample weight (blue), and the lower and upper bounds for sample size n = 30 (dashed line).

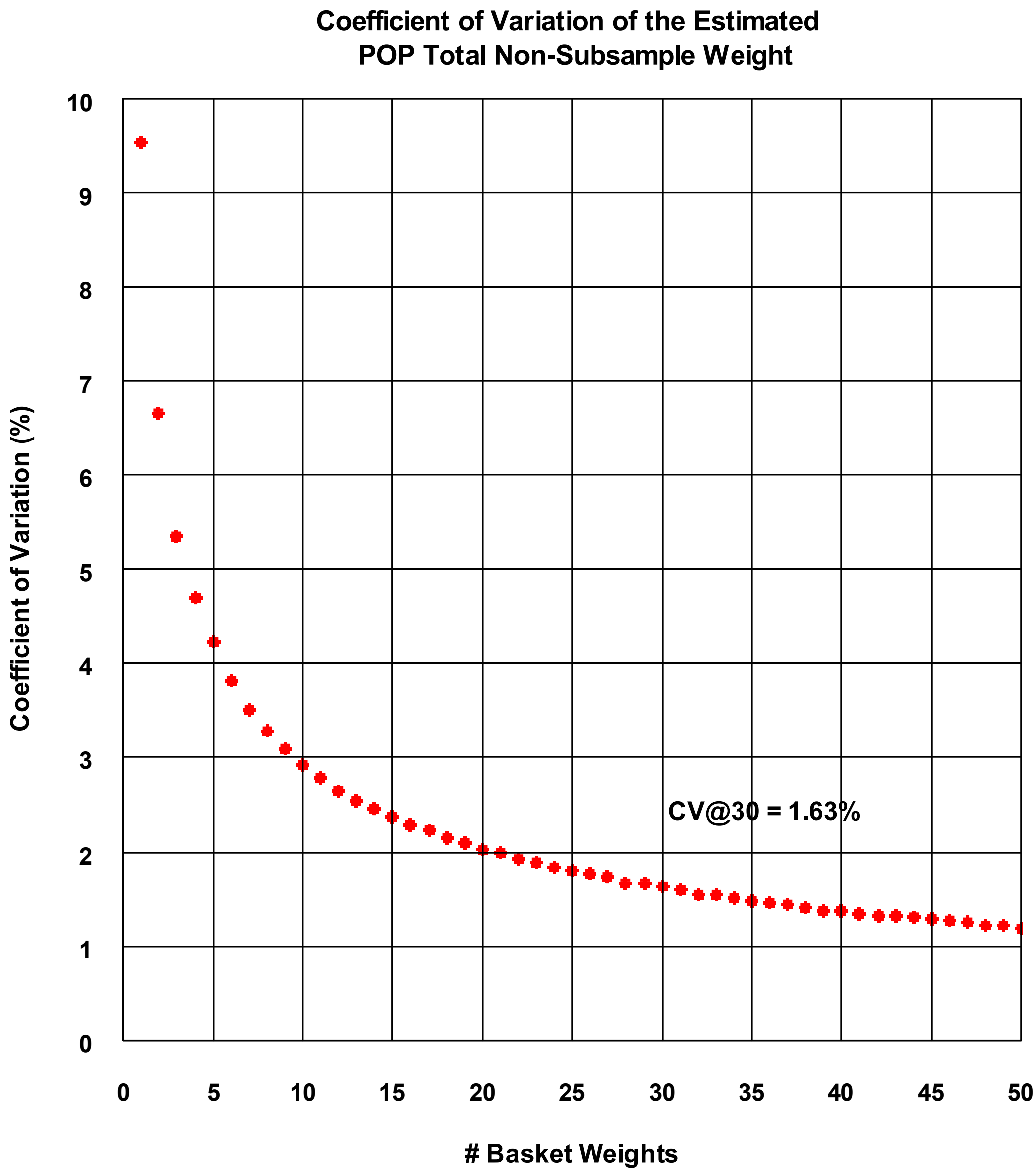


Figure 13. -- The coefficient of variation (CV) of the estimated non-subsample total weight versus sample size for Data Set-1. Indicated is the CV of the estimate at sample size n = 30.

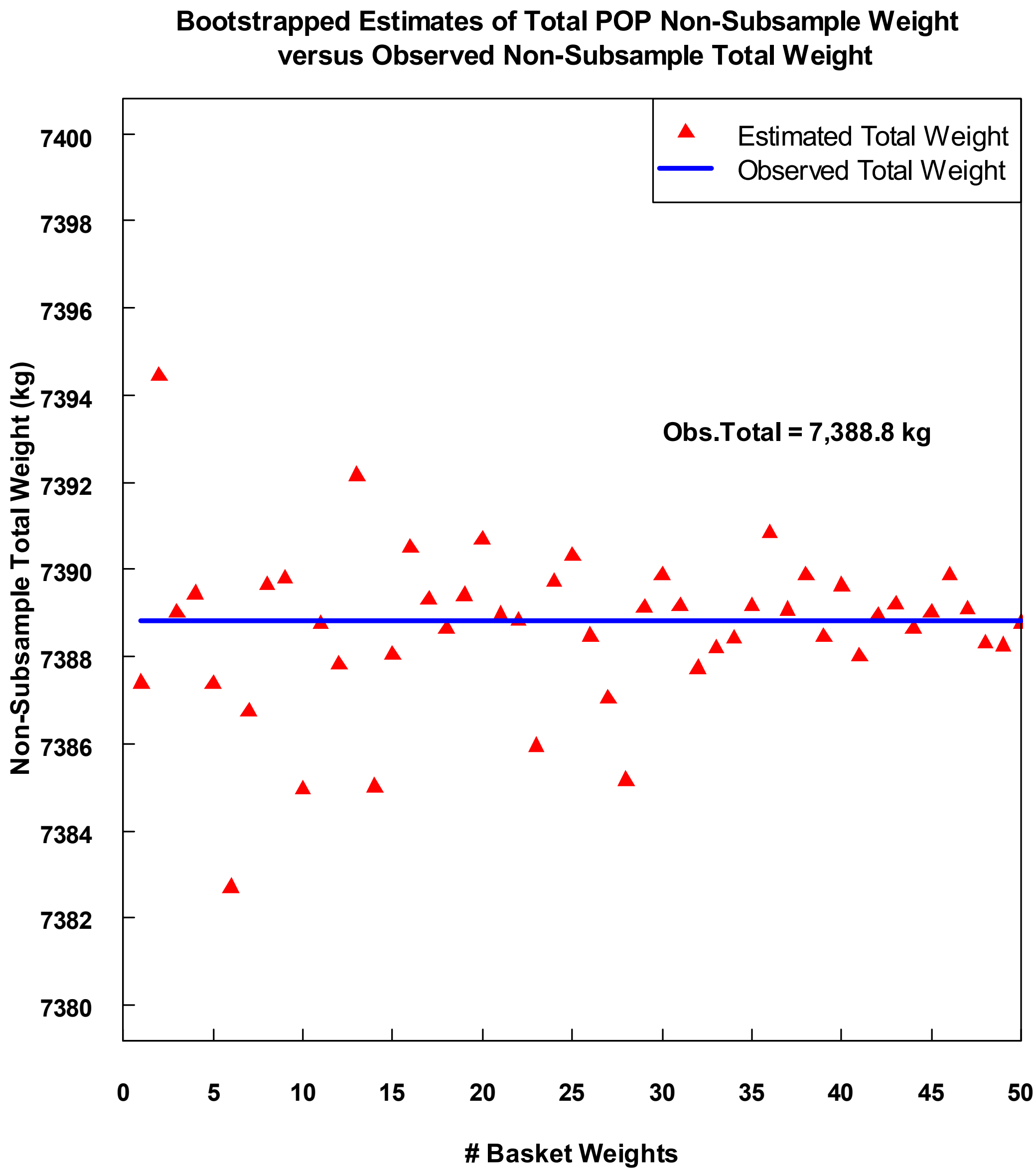


Figure 14. -- Estimates of the mean non-subsample total weight versus sample size for Data Set-1. Indicated is the observed total non-subsample weight (blue line).

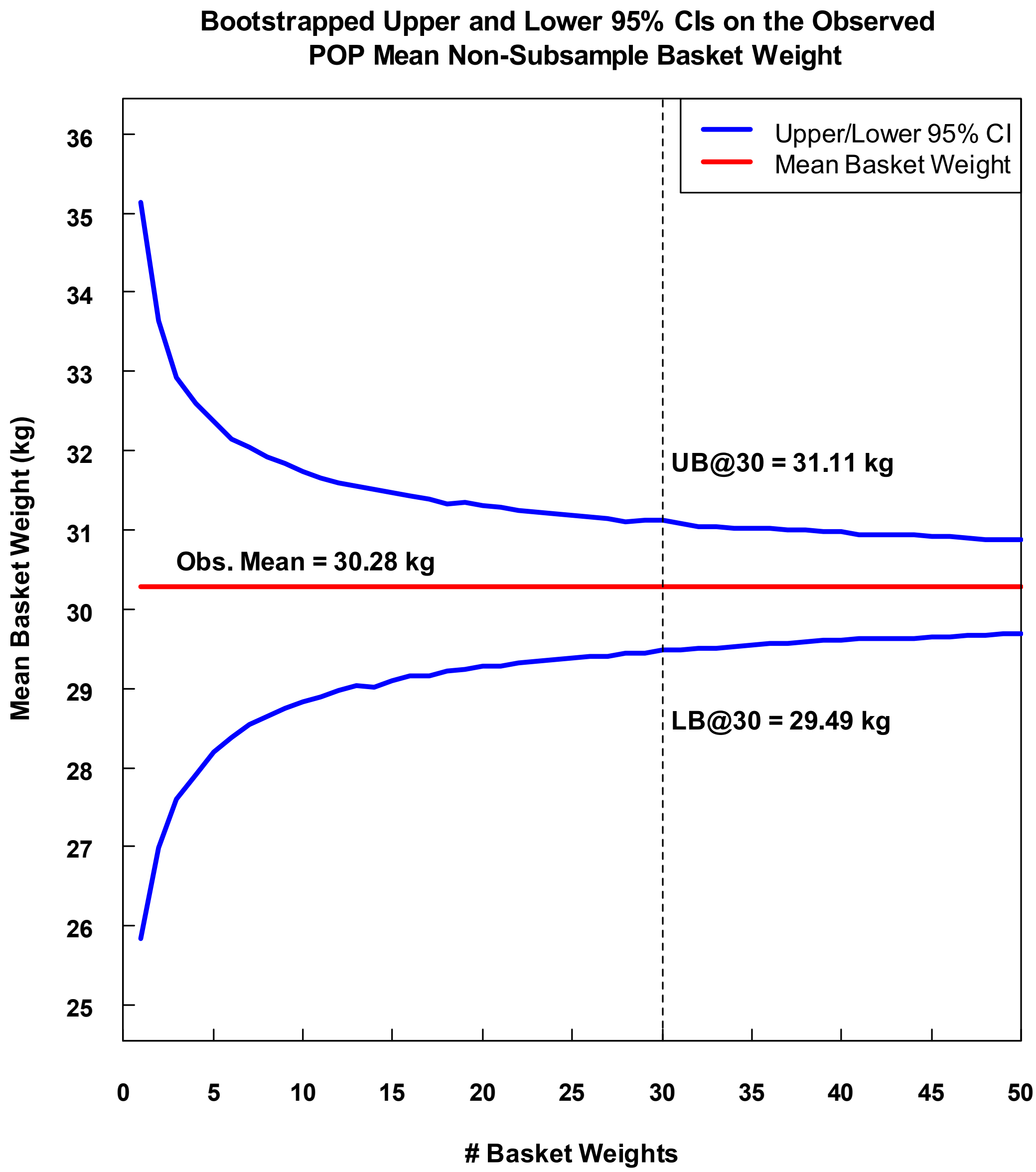


Figure 15. -- Lower and upper 95% confidence bounds on the non-subsample mean basket weight versus sample size for Data Set-1. Also shown are the observed non-subsample mean weight (red), and the lower and upper bounds for sample size n = 30 (dashed line).

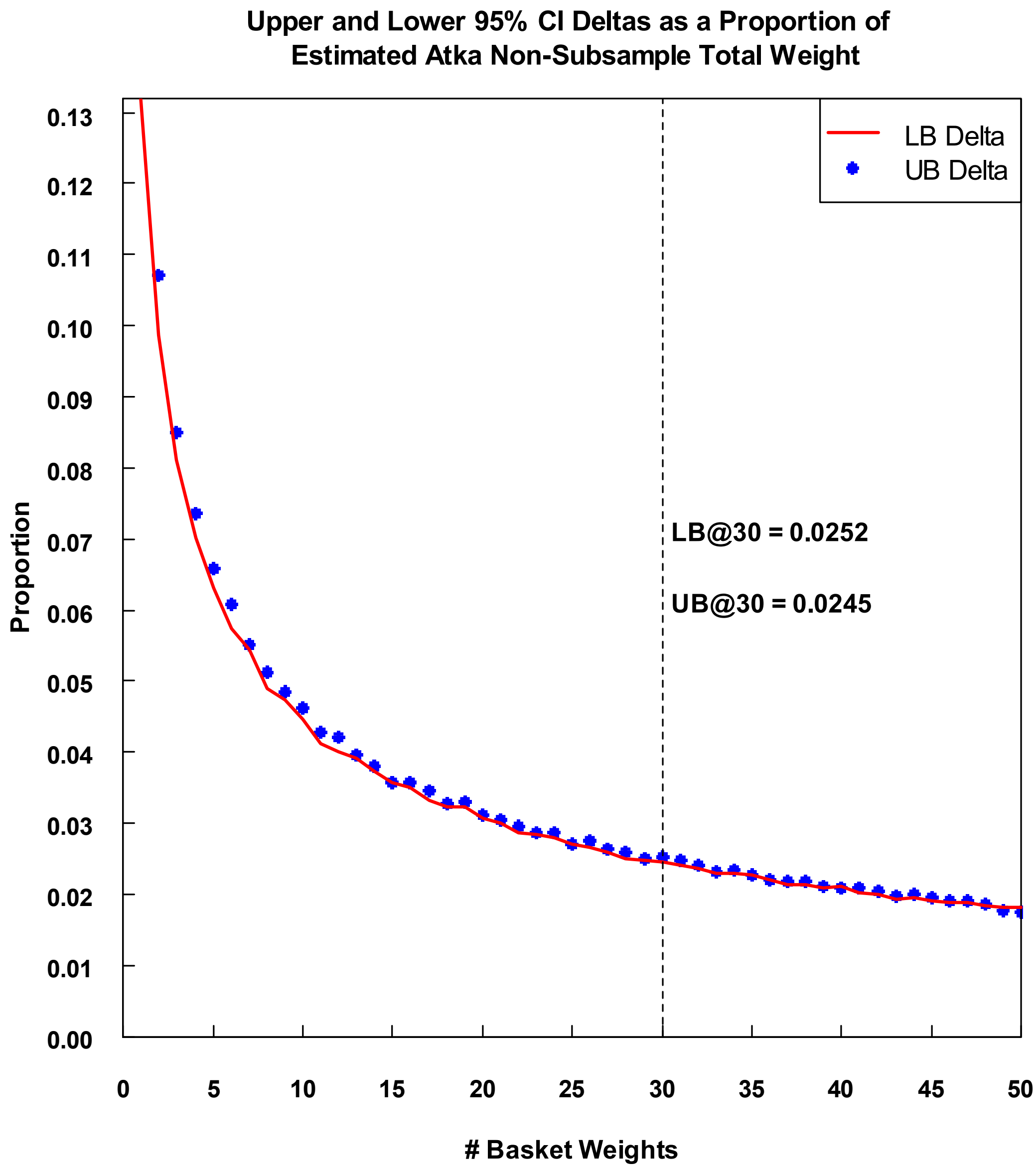


Figure 16. -- Lower and upper 95% confidence bounds deltas (deviations) as a proportion of the estimated non-subsample total weight versus sample size for Data Set-2. Indicated are the lower and upper proportions for sample size n = 30 (dashed line).

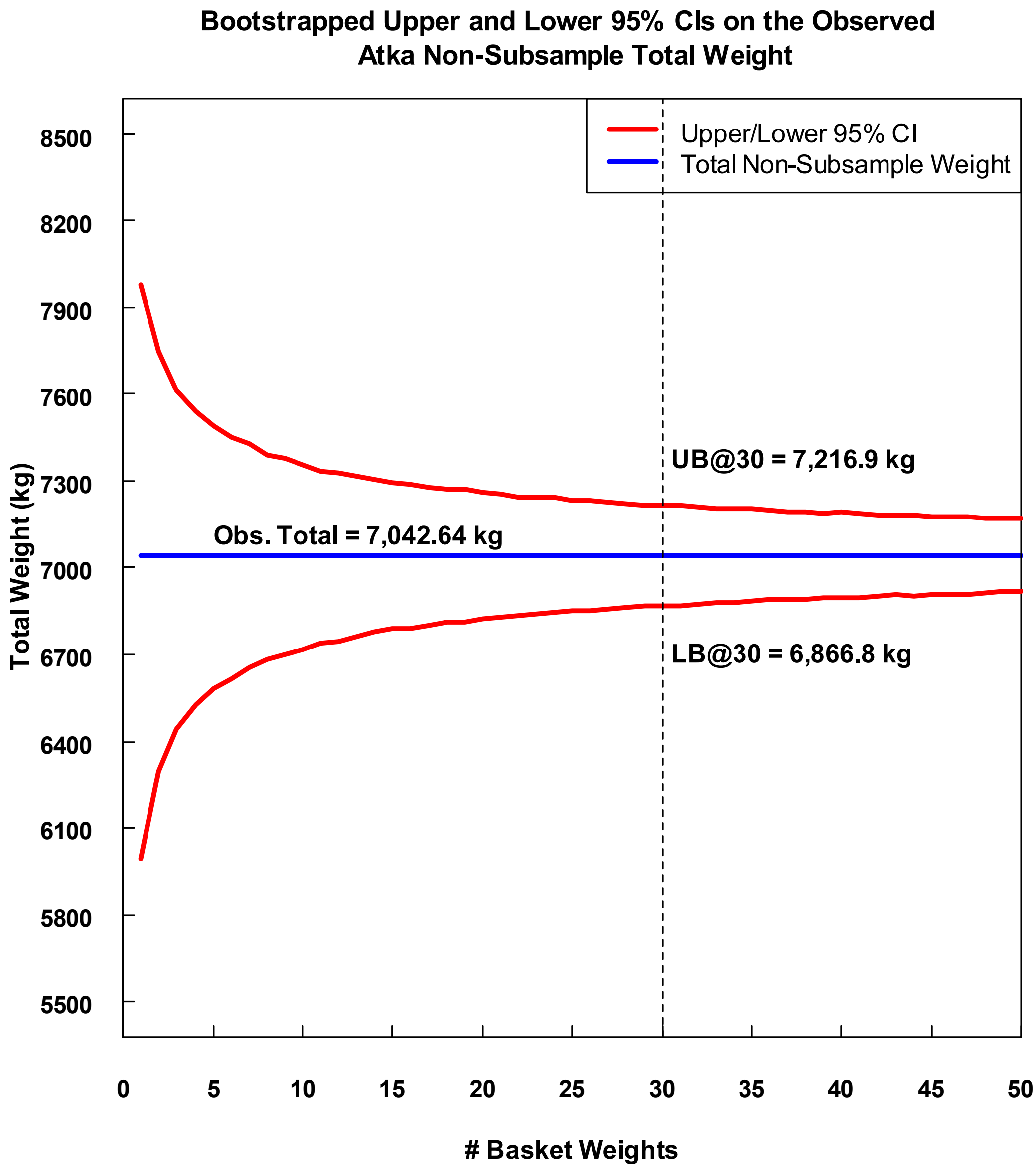


Figure 17. -- Lower and upper 95% confidence bounds on the non-subsample total weight estimate versus sample size for Data Set-2. Also shown are the observed total non-subsample weight (blue), and the lower and upper bounds for sample size n = 30 (dashed line).

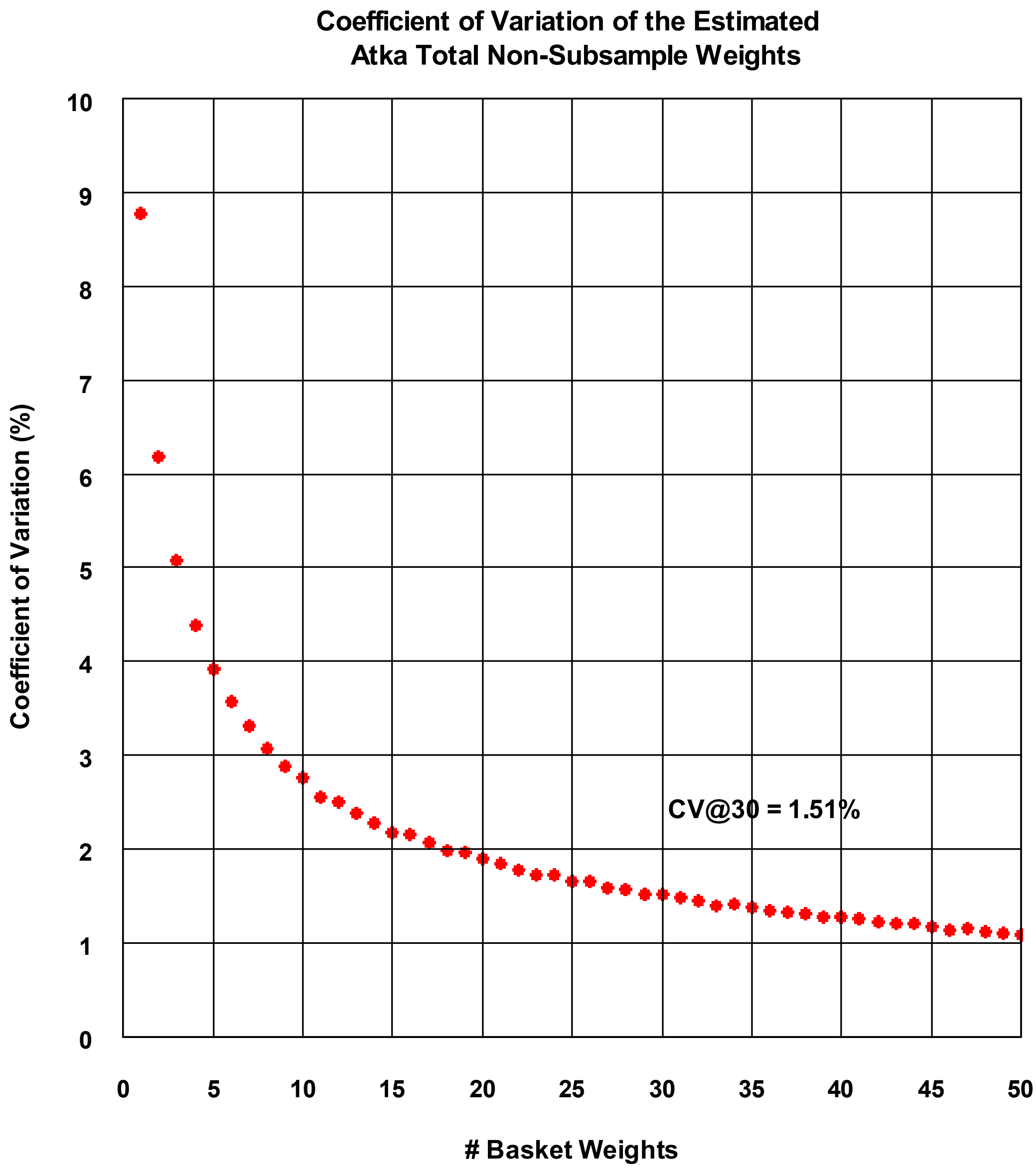


Figure 18. -- The coefficient of variation (CV) of the estimated non-subsample total weight versus sample size for Data Set-2. Indicated is the CV of the estimate at sample size n = 30.

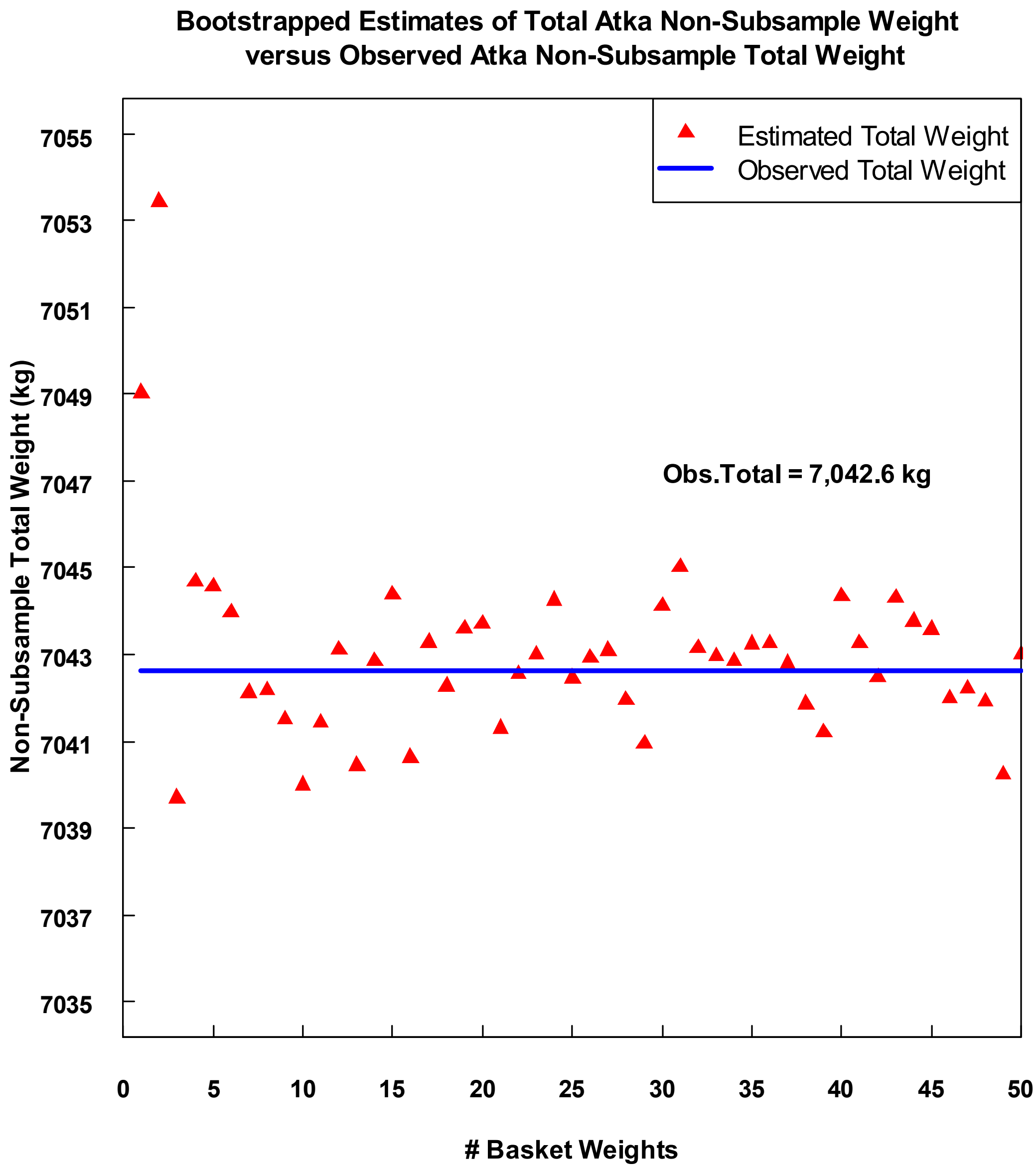


Figure 19. -- Estimates of the mean non-subsample total weight versus sample size for Data Set-2. Indicated is the observed total non-subsample weight (blue line).

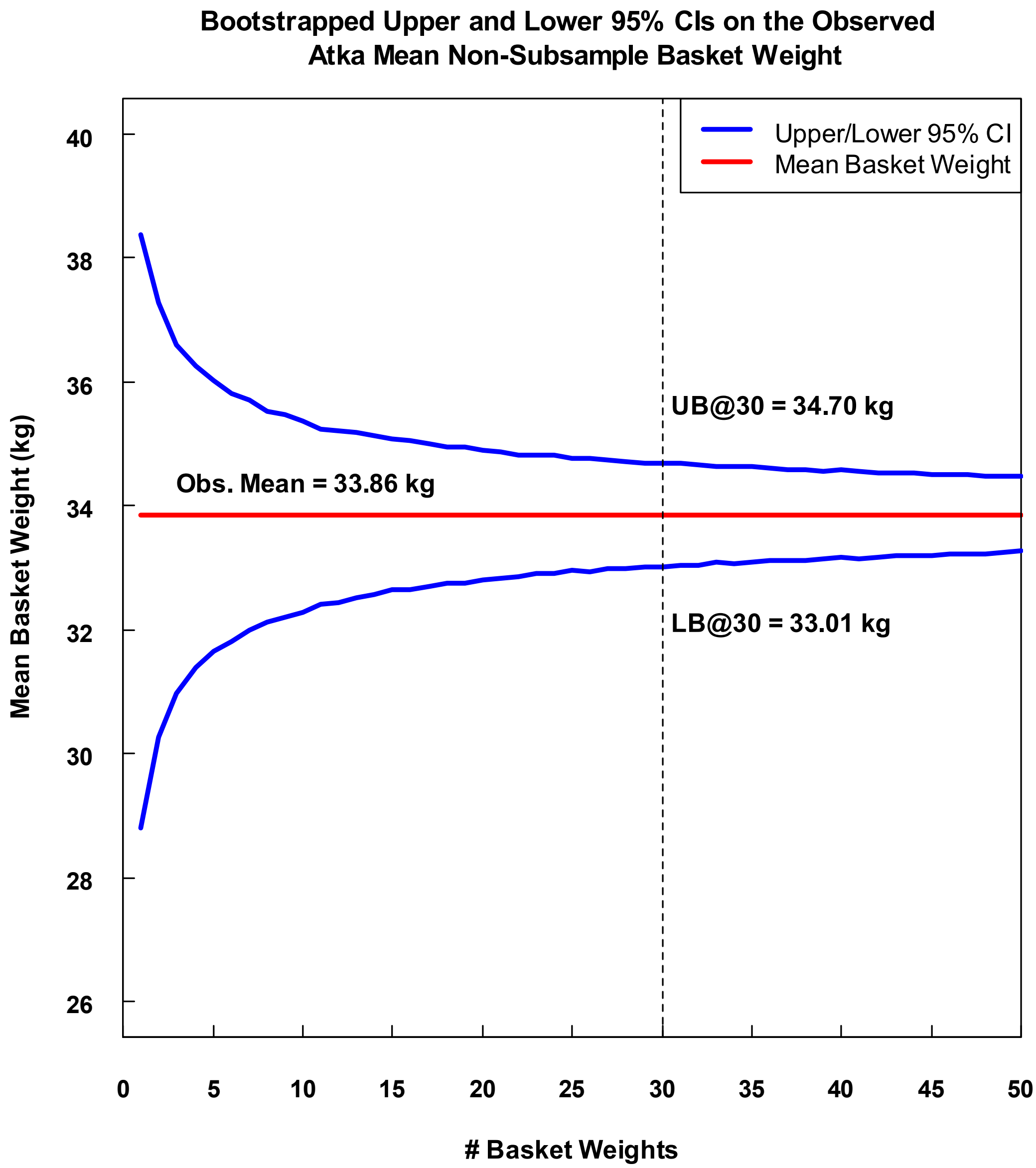


Figure 20. -- Lower and upper 95% confidence bounds on the non-subsample mean basket weight vs sample size for Data Set-2. Also shown are the observed non-subsample mean weight (red), and the lower and upper bounds for sample size n = 30 (dashed line).

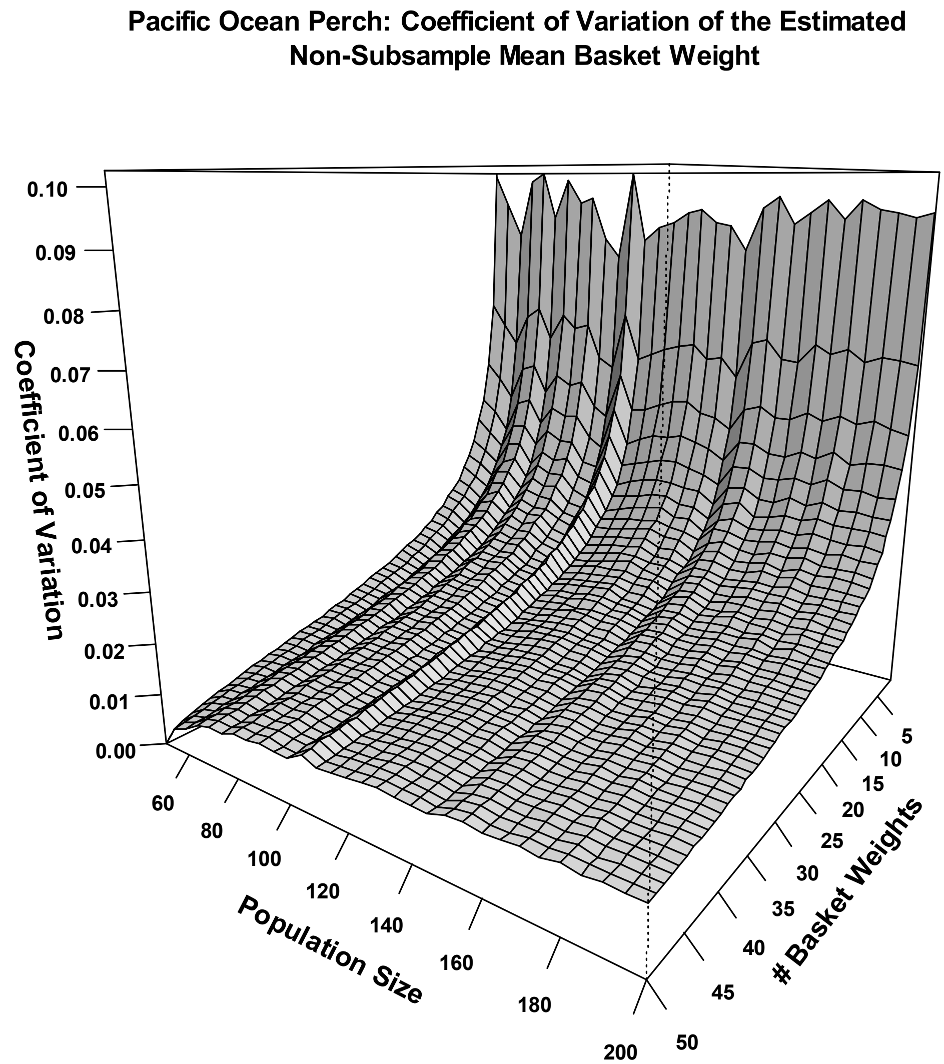


Figure 21. -- The coefficient of variation of the estimated mean basket weight vs sample size (n = 1 to 50) and catch volume (population size) from 50 to 200 baskets for Data Set-1. (The calculated coefficients of variation would be identical if shown for estimated non-subsample total weight.)

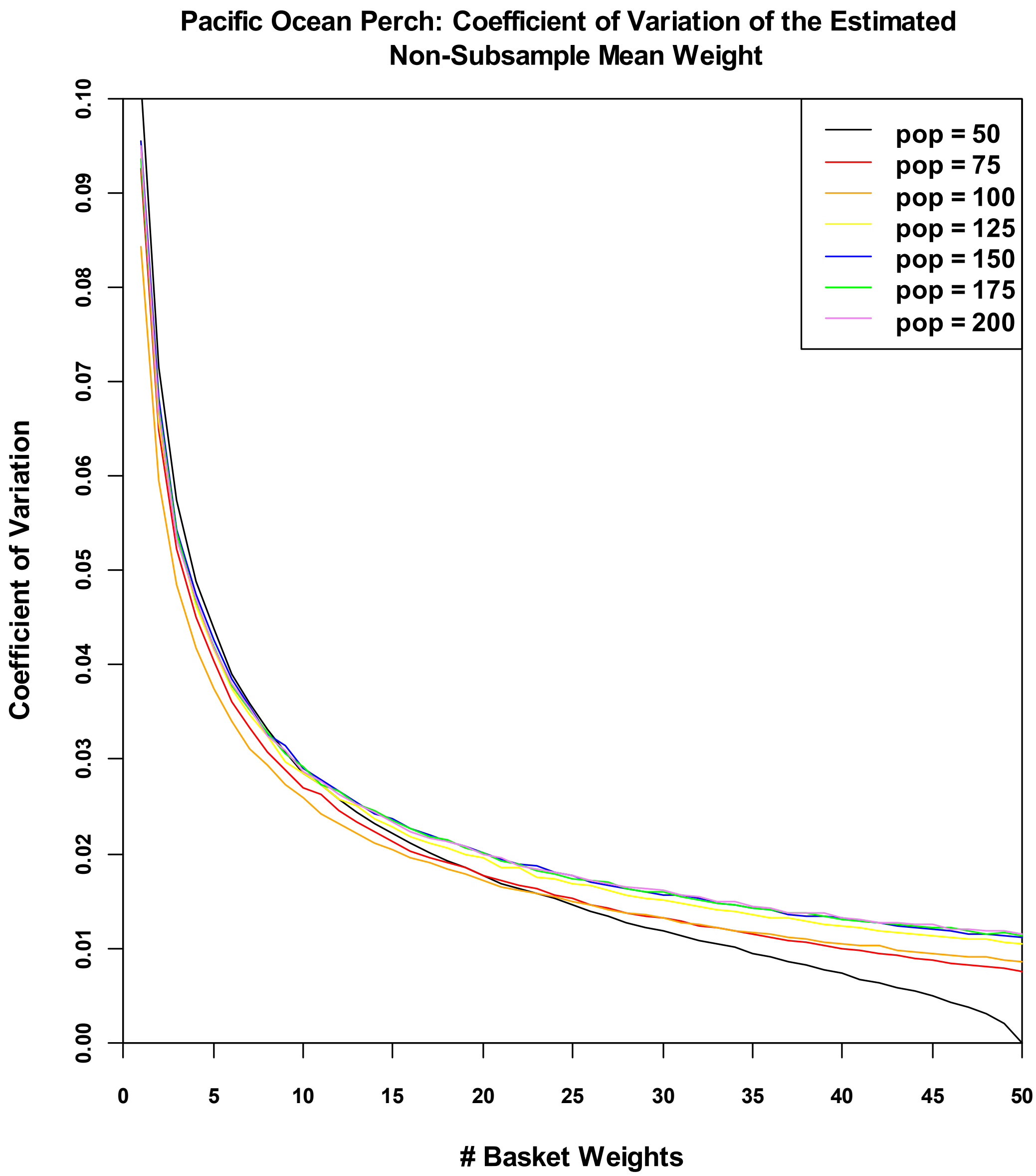


Figure 22. -- The coefficient of variation of the estimated mean basket weight versus sample size (n = 1 to 50) and catch volume (population size) from 50 to 200 (by 25) baskets for Data Set-1. (The calculated coefficients of variation would be identical if shown for estimated non-subsample total weight.)

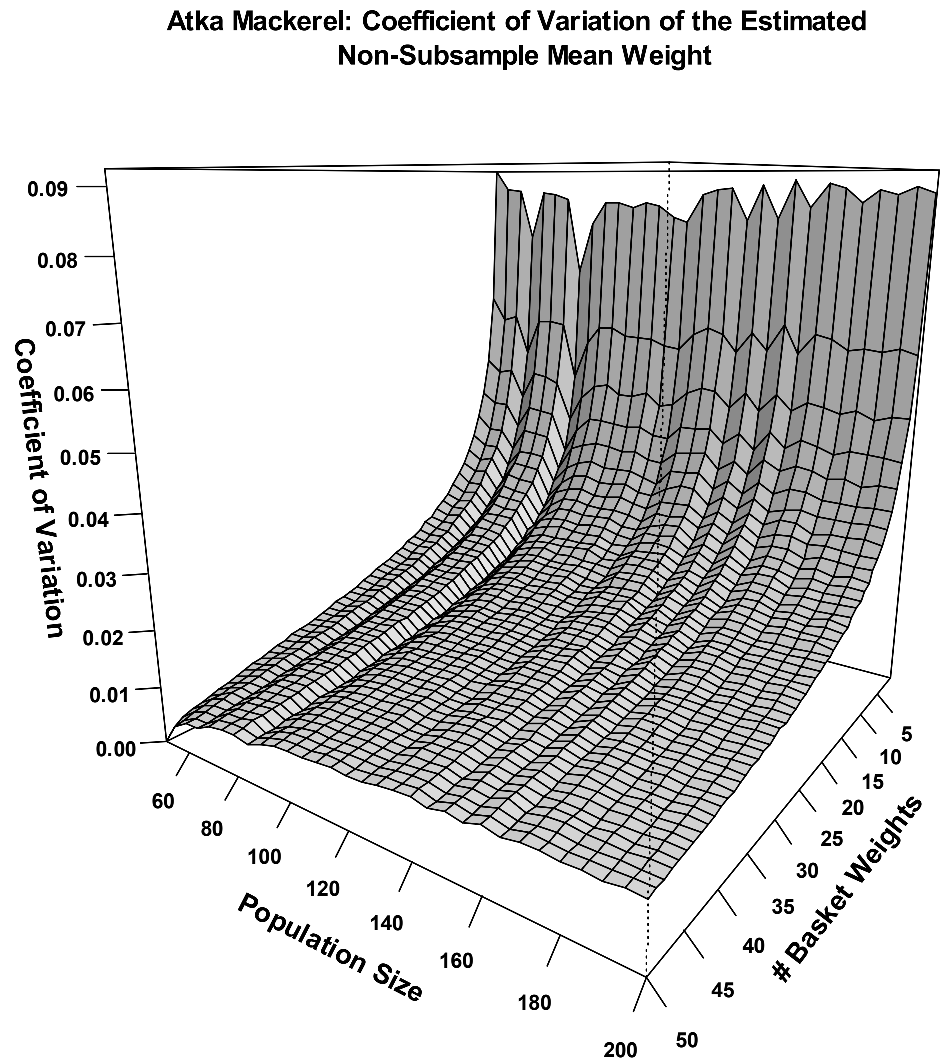


Figure 23. -- The coefficient of variation of the estimated mean basket weight versus sample size (n = 1 to 50) and catch volume (population size) from 50 to 200 baskets for Data Set-2. (The calculated coefficients of variation would be identical if shown for estimated non-subsample total weight.)

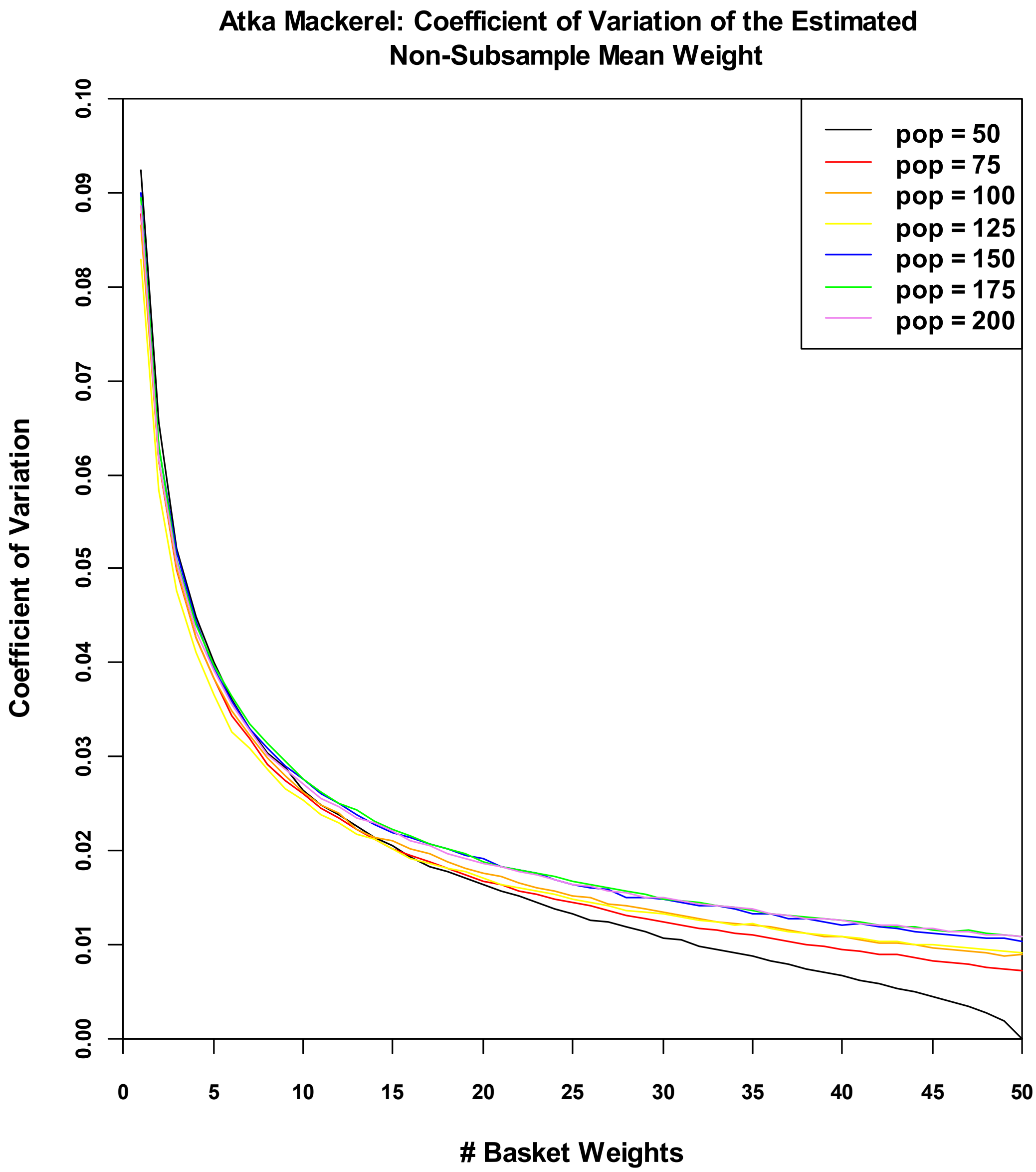


Figure 24. -- The coefficient of variation of the estimated mean basket weight versus sample size (n = 1 to 50) and catch volume (population size) from 50 to 200 (by 25) baskets for Data Set-2. (The calculated coefficients of variation would be identical if shown for estimated non-subsample total weight.)

ADDENDUM

Introduction

We conducted a special study during the 2018 AI survey aimed at a more informed assessment of the statistical performance of the counted basket approach in catch processing. For a single successful haul of Atka mackerel and Pacific ocean perch, the goal was to process the catch according to design requirements of the method with one exception. Instead of counting baskets and dumping that portion of the non-subsample that would not be weighed, weights on all baskets were taken. The target catch size was at least 100 non-subsample basket weights, but not substantially greater than 100 to mitigate safety and ergonomic concerns, or catches in the range of 3.5 t to 5.0 t. Basket weights taken in this manner would have the expected level of variability for baskets level-filled and dumped if the approach was applied. This contrasts the variability in basket weights in Data Set-1 (Pacific ocean perch) and Data Set-2 (Atka mackerel) analyzed previously that did not follow prescribed collection methods.

A principal finding of our analysis of Data Set-1 and Data Set-2 was that the counted basket approach is statistically robust in terms of performance measures, and that it produced reliable estimates of mean basket weight and total non-subsample weight relative to the observed. The point estimates of catch were accurate and relatively precise over a range of catch volumes. At the target sample size of n = 30 baskets, the CVs of the catch estimates were ≈ 1.5% for all catch volumes (~1.5 t to 7.0+ t) generated from these data. We also found that the 95% confidence bounds on the total non-subsample weight estimate deviated by an average 2.4% from the observed total weight; the average percent deviations were 2.6% for large volume catches, and 2.3% for medium volume catches. We surmised that these were likely overestimates of the expected percent deviations, and concluded that the estimates of non-subsample mean basket weight and total weight would not deviate, on average, by more than 2.5% from the observed, and undoubtedly less (perhaps < 2.0%) if the method is applied consistent with design specifications.

On the 2018 AI survey, we successfully processed a haul of Pacific ocean perch under the special study, but not for Atka mackerel due to logistical considerations. In this Addendum, we evaluate this 2018 Pacific ocean perch data set (hereafter, the 2018 data set) using the methods applied to Data Set-1 and Data Set-2, and compare the performance measures of the estimates of mean basket weight and non-subsample total weight to those for Data Set-1 and Data Set-2. Results of this analysis will provide greater insight into the expected performance of the counted basket approach in catch processing, and address our previous supposition that the 95% confidence bounds on the point estimates will not deviate by more than 2.0% from the observed if the method is applied as designed.

Materials and Methods

The 2018 data set consisted of 3,368 kg of Pacific ocean perch split into 111 non-subsample baskets (3,253 kg), and 4 subsample baskets (115 kg). The catch was completely sampled and processed consistent with design requirements (i.e., all baskets were level-filled and individually weighed). The mean of the non-subsample was 29.30 kg, with range = [5.33, 31.96], CV = 8.95%, and 95% bounds = [28.82, 29.79]. The mean of the subsample was 28.78 kg with range = [28.20, 29.60], CV = 2.06%, and 95% bounds = [28.20, 29.36]. The non-subsample weights were widely distributed, left-skewed, and departed from normality (Addendum Fig. 1a).

The one partially-filled non-subsample basket weight of 5.33 kg was designated as an outlier and removed. The resulting non-subsample was n = 110 baskets totaling 3,247 kg, with mean = 29.52 kg, range = [26.46, 31.96], CV = 4.32%, and 95% bounds = [29.28, 29.76]. The new distribution was slightly left-skewed, and it still departed from normality at both tails (Addendum Fig. 1b). The original subsample of four baskets was not changed.

We analyzed the 2018 data set using the methods applied to Data Set-1 and Data Set-2. We refer the reader to Materials and Methods section for detailed descriptions of those methods.

Results

We simulated the implementation of the mean basket weight protocol via bootstrap analysis using the 2018 data set. The results correspond to the scenarios in Simulation Analyses with the exception that the 2018 haul was a medium volume catch by our definition.

Simulation Analyses

In the analysis of Data Set-1 and Data Set-2, we simulated a medium volume catch as a population of 100 baskets derived as a random draw from the observed non-subsample populations of 244 basket weights (Data Set-1) and 208 basket weights (Data Set-2). Since the 2018 data set consisted of 110 non-subsample weights, the medium volume population here included all 110 weights.

Scenario-1: Medium Volume Catch and Fixed Sample Size

For a fixed sample size of n = 30, the mean basket weight estimate and its 95% CI were 29.519 kg (SE = 0.20) and [29.19, 29.74] kg, respectively. This compared to the observed non-subsample mean weight of 29.522 kg (Addendum Fig. 2). The non-subsample total weight was estimated at 3,247.10 kg compared to the observed total weight of 3,247.38 kg representing a 0.009% difference. The 95% confidence bounds on the total weight estimate was [3,211.12, 3,282.33] kg with a range of 71.21 kg, and these bounds represented deviations of -1.11% and 1.08%, respectively, from the mean.

Scenario-2: Medium Volume Catch and Variable Sample Size

The performance measures for estimates of mean basket weight and total non-subsample weight versus sample size (n = 1 to 50) are shown in Addendum Table 1. The lower and upper 95% confidence bound deviations from the mean are shown as a proportion of the estimated non-subsample total weight versus sample size (Addendum Fig. 3). Both deviations declined rapidly over the range of sample sizes. At n = 30, the lower and upper bound deviations were 1.11% and 1.08%, respectively.

Addendum Figure 4 shows the lower and upper 95% confidence bounds on the non-subsample total weight estimate versus sample size. At n = 30, the lower-bound estimate was 3,211.19 kg, and the upper bound was 3,282.91 kg. These compared to the observed non-subsample total weight of 3,247.38 kg, and equated to deviations of -36.19 kg and 35.53 kg, respectively, from the observed (range = 71.72 kg).

The change in the coefficient of variation of the non-subsample total weight estimate versus sample size is presented in Addendum Figure 5. At n = 30, the CV was 0.68%. The CV declined rapidly with increasing sample size, and it fell to less than 1.0% at n > 15. Addendum Figure 6 presents a diagnostic result of the bootstrap formulation. The mean of the bootstrapped mean non-subsample total weight estimates varied within a narrow range (95% were ± 1.0 kg) of the observed total weight over the range of n as would be expected if correctly formulated.

The 95% confidence bounds on the non-subsample mean basket weight estimate versus sample size is presented in Addendum Figure 7. At n = 30, the lower-bound estimate was 29.19 kg and the upper bound was 29.85 kg. These compared to the observed non-subsample mean basket weight of 29.52 kg, and equated to deviations of -0.33 kg (-1.11%) and 0.32 kg (1.09%), respectively, from the observed mean.

Scenario-3: Variable Volume Catch and Variable Sample Size

These results are for 51 levels of non-subsample population size (50 to 100 baskets; ~1.5 t to ~3.0 + t) processed using sample sizes from 1 to 50. The coefficient of variation of the estimated mean basket weight versus sample size and catch volume is shown in Addendum Figure 8. This plot would be identical if shown for the estimated non-subsample total weight. Despite slight irregularities in the response surface, the CV declined rapidly from small to large n, and was fairly low and consistent for n > 10.

Addendum Figure 9 presents slices of the response surface in Addendum Figure 8 for six catch volumes (50 to 100, by 10). For all catch volumes, the CV declined rapidly from low to high n.

For the population of 50 baskets, the CV went to zero at n = 50 since all baskets were weighed. The patterns of CV for the other five catch volumes were similar and only marginally larger with increasing population size. At n = 6, the CVs for all catch volumes were less than 2.0%. At n = 9, all CVs were all < 1.5%, and they were all < 1.0% at n = 17. At the recommended n = 30 baskets, the CVs for all catch volumes were < 0.7%.

Discussion

These results illustrate a marked increase in statistical robustness of the counted basket approach when it's applied to data that meet design requirements versus Data Set-1 and Data Set-2. Basket weights in Data Set-1 and Data Set-2 were more variable than in the 2018 data set even after *post hoc* removal of conspicuous outliers. We can consider Data Set-1 and Data Set-2 as non-standard data as they weren't collected consistent with the prescriptions of the approach in the operations plan.

Simulation Analyses

Simulation results cover real-life scenarios processing catch on the GOA and AI surveys. The first scenario (fixed n = 30) is the typical application of the approach in survey operations. The last two scenarios (i.e., variable n = 1 to 50 vs. either medium or variable catch volumes, respectively) provide scope to the results and a basis to evaluate its performance over a range of combinations of sample size and catch volume.

Scenario-1: Medium Volume Catch and Fixed Sample Size

The principal conclusion of this simulation is that the estimates of non-subsample total weight and mean basket weight are exceedingly precise and accurate. The estimated mean basket weight (29.519 kg) is virtually indistinguishable from the observed mean weight (29.522 kg) (Addendum Fig. 2), as is the estimated non-subsample total weight (3,247.10 kg) from the observed (3,247.38 kg). There is high confidence that the estimated non-subsample total weight (or mean weight) deviates by only ≈ 1.1% from the mean. This contrasts the results for Data Set-1 and Data Set-2 which found that the estimates deviated by ± 2.7% and ± 2.5%, respectively, from the observed. The deviations for these non-standard data are greater by a multiple of 2.5 and 2.3, respectively, than the deviations found using the 2018 data set.

Scenario-2: Medium Volume Catch and Variable Sample Size

Model diagnostics indicative of statistical robustness stabilize at n > 10, and result in catch estimates consistent with those at n = 30 (Addendum Table 1). The precision of the estimates increases with increasing n, and they stabilize at n ≥ 15. The gain in accuracy of the point estimates, or decrease in their variance, is relatively negligible beyond n = 10. These findings support the conclusion that the counted basket approach provides reliable catch estimates in terms of both accuracy and precision.

The measures of dispersion and CVs of the point estimates of mean basket weight and non-subsample total weight decrease with increasing n (Addendum Table 1). At n > 1, the percent deviation between the point estimates and the lower or upper 95% confidence bound is < 5.0%, it's < 3.0% for n ≥ 4 baskets, and ≤ 1.1% n ≥ 30 baskets (Addendum Fig. 3). This contrasts the results for Data Set-1 where the percent deviations were < 5.0% at n > 10, < 3.0% at n ≥ 25, and ≤ 2.7% at n ≥ 30 baskets (Table 2 and Fig. 11). The results for Data Set-2 were < 5.0% at n > 9, < 3.0% for n ≥ 22, and ≤ 2.5% at n ≥ 30 (Table 3 and Fig. 16)

The coefficient of variation of the non-subsample total weight estimate for the 2018 data set (Addendum Fig. 5) was found to be approximately one-half of that for Data Set-1 and Data Set-2 over the range of n = 1 to 50 (Figs. 13 and Fig. 18, respectively). For the 2018 data set, the CV was < 5.0% at n = 1, < 2.0% at n ≥ 5, < 1.0% at n ≥ 15, and it equaled 0.68% at n = 30. This contrasts results for Data Set-1 where the CVs were < 2.0% at n ≥ 20 and equaled 1.63% at n = 30 (Fig. 13), and those for Data Set-2 where the CVs were < 2.0% at n ≥ 20, and equaled 1.51% at n = 30 (Fig. 18). To achieve a CV on the non-subsample total catch estimate < 2.0%, for example, required a sample of at least n = 5 baskets for the 2018 data set versus at least n = 20 for both Data Set-1 and Data Set-2. For both the non-standard data sets, the CV on the estimate of non-subsample total catch never fell below 1.0% even at n = 50 baskets.

For the 2018 data set, the 95% confidence bounds on the non-subsample mean basket weight estimate varied within ± < 0.33 kg (≤ 1.1%) from the observed at n ≥ 30 (Addendum Fig. 7), whereas, it varied within ± < 1.0 kg (≤ 2.8%) for Data Set-1 and Data Set-2 (Figs. 15 and 17, respectively). Thus, the range on the point estimate for the non-standard data sets was approximately three times greater than for the 2018 data set.

Scenario-3: Variable Volume Catch and Variable Sample Size

This simulation integrates processing variable catch volumes (50 to 100 baskets; ~1.5 t to ~3.0 + t) using sample sizes from 1 to 50. In the analysis of Data Set-1 and Data Set-2, we saw that the estimates of mean basket weight and non-subsample total weight were relatively precise and accurate over the range of n, and that there was negligible gain in the accuracy of the estimates in processing large volume (~7.0 t) versus medium volume (~3.0 t) catches at n = 30 (Figs. 7 through 10). Since catch estimates are relatively conserved over n, we focus on the CV of the estimates to assess the interaction between catch volume and sample size for data taken consistent with design requirements.

Addendum Figure 8 and Addendum Figure 9 show that the CV of the estimates declined rapidly from low to high n, and were fairly low and consistent for n > 10 for all catch volumes. At n = 6, all CVs were < 2.0%, < 1.5% at n = 9, and < 1.0% at n = 17 baskets. At n = 30, the CVs for all catch volumes were < 0.7%. This contrasts results for Data Set-1 where the CVs were < 4.3% at n = 6, < 3.4% at n = 9, < 2.4% at n = 17, and < 1.7% at n = 30 (Fig. 22). For Data Set-2, the CVs were < 3.9% at n = 6, < 3.2% at n = 9, < 2.2% at n = 17, and < 1.6% at n = 30 (Fig. 24).

For the 2018 data set, the key finding is that, above a small threshold sample size (n ≥ 4), the coefficients of variation of the catch estimates are acceptably low (< 2.5%), they are exceedingly low (< 1.0%) for a moderate sample size (n = 17), and at the target n = 30, the CVs are < 0.7%. For the range of potential real-life scenarios of catch volumes on the GOA and AI surveys, selecting 25-30 basket weights to estimate the mean leads to highly statistically robust catch estimates.

Conclusion

The analysis of the 2018 data set bolsters our main conclusion from the earlier analysis that the counted basket variant approach is exceedingly statistically robust in estimating mean basket weight and non-subsample total weight, particularly so when applied to data taken according to the prescriptions in the operations plan. In the earlier analysis, we evaluated the performance of the approach using two non-standard dominant species hauls on the GOA and AI surveys. The special study in 2018 provided a data set that would be more illustrative of the expected performance of the approach, and more informative in terms of implementation decisions; for example, basket weight sample size goals.

The point estimates of the catch (i.e., mean basket weight and non-subsample total weight) were both highly accurate and highly precise over the range of catch volumes (~1.5 t to ~ 3.0 + t) and sample sizes (1 to 50 baskets). Even using a moderate number of basket weights to estimate the mean ($n \approx 15$), we can expect the CV on the catch estimates to be approximately 1.0%, and at the recommended $n = 30$ baskets, the CV will be < 0.7%. There is negligible gain in the precision of the point estimates by taking more than 20 to 25 basket weights; $n = 30$ weights is more than adequate to produce highly precise point estimates of the catch.

A general observation concerning the 2018 data set analysis is that the CVs on the point estimates are less than one-half the magnitude of those for the non-standard data sets (Data Set-1 and Data Set-2) over the range of $n = 1$ to 50 (Addendum Fig. 5 vs. Figs. 13 and 18, respectively). It will require substantially fewer basket weights to achieve the equivalent CV from data collected consistent with design requirements compared to non-standard data. For example, to achieve a CV on the non-subsample total catch estimate < 2.0% required as few as $n = 5$ basket weights for the 2018 data set, and at least $n = 20$ for the non-standard data sets.

Furthermore, for both Data Set-1 and Data Set-2, the CVs never fell below 1.0% even at $n = 50$ baskets, whereas they were < 1.0% at $n \geq 17$ and equaled 0.68% at $n = 30$ for the 2018 data set.

A specific goal in analyzing these 2018 data was to refine our expectation of the accuracy of this method in estimating non-subsample total weight, and to assess our earlier supposition that the 95% confidence bounds on point estimate will not vary, on average, by more than 2.0% from the mean. The first simulation (fixed $n = 30$) was the typical application of the approach on the survey. The estimated mean basket weight (29.519 kg) was indistinguishable from the observed mean weight (29.522 kg), and this difference is subsumed by the level of precision (0.01 kg) of the scales used to take basket weights. The difference between the estimate of non-subsample total weight (3,247.10 kg) and the observed total weight (3,247.38 kg) was similarly negligible considering practical considerations of taking accurate baskets weights on board a vessel platform. There is high (95%) confidence that the estimated non-subsample total weight deviates by an average of only ≤ 1.1% from the observed. We can expect, by extension, that upper and lower 95% confidence bound deviations will be ≤ ± 33 kg for a 3.0 t catch, ≤ ± 55 kg for a 5.0 t catch, and ≤ ± 77 kg for a 7.0 t catch.

The 2018 special study data set provided a basis to assess the degree of overestimation of the percent error in the mean basket weight and non-subsample total weight estimates using the non-standard Data Set-1 and Data Set-2. Results for the 2018 data set were dramatically more robust than even those for the non-standard data sets. In general, the coefficients of variation on the point estimates for the 2018 data set were less than one-half the magnitude of those for Data Set-1 and Data Set-2 for all catch volumes and sample sizes, while the deviations in the 95% confidence bounds for the non-standard data were greater by a multiple of 2.5 than for the 2018 data set.

Results of this analysis reinforce the validity of the counted basket variant approach in processing large-volume catches on the GOA and AI surveys, particularly when the haul data are collected consistent with design protocols. We found the approach to be even more statistically robust than our earlier findings, and that it results in highly precise and highly accurate estimates of mean basket weight and non-subsample total weight relative to whole-haul processing. We conclude that whole-hauling large volume catches is not statistically necessary to achieve precise and accurate estimates of total catch weight.

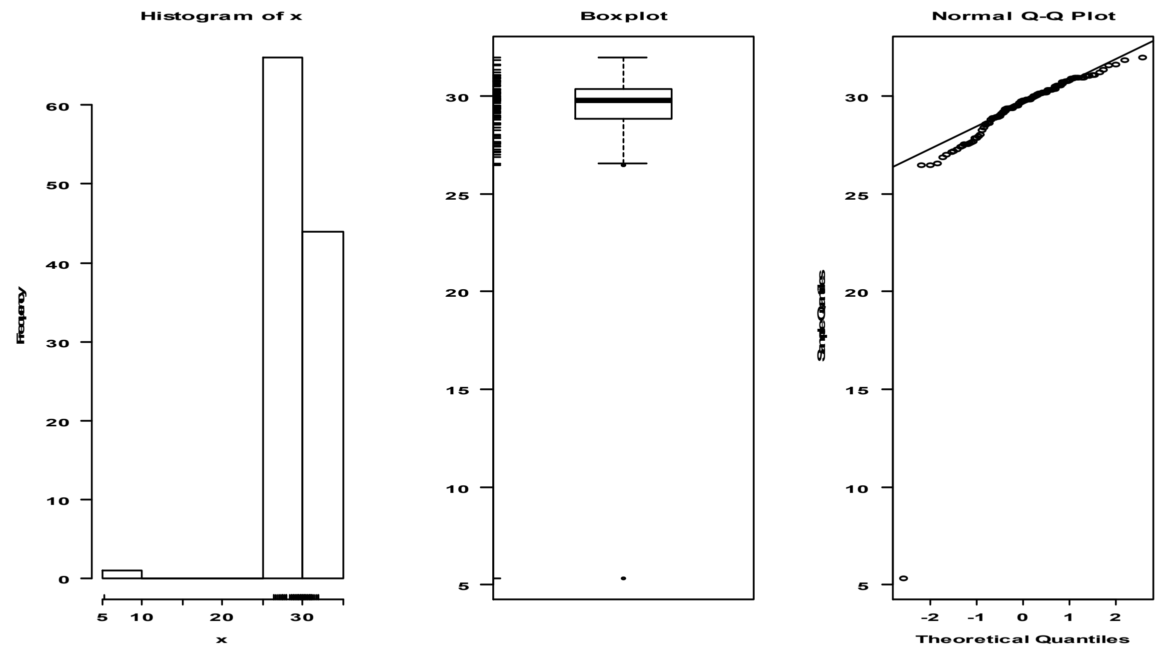
Addendum Table 1. -- Performance measures and model output of mean basket weight and total non-subsample weight estimates versus sample sizes n = 1 to 50. Mean basket weight measures are: the estimate of the mean (Est.) and its standard deviation (SD), and the lower (LB) and upper (UB) 95% confidence bounds. Total non-subsample weight measures are: the total weight estimate (Est.) and its standard error (SE), the lower (LB) and upper (UB) 95% confidence bounds, the range of the 95% CI (CI Range), the lower (Δ LB) and upper (Δ UB) bound deviations from the mean, and the lower (% LB) and upper (% UB) bound deviations as a percent of the mean.

Pacific ocean perch – Model output for sample sizes [1-50] from 110 Non-subsample basket weights													
	Mean basket weight (kg)				**Total Non-subsample weight (kg)**								
n	**Est.**	**SD**	**LB**	**UB**	**Est.**	**SE**	**LB**	**UB**	**CI Range**	**Δ LB**	**Δ UB**	**% LB**	**% UB**
1	29.51	1.27	27.14	31.20	3246.38	139.28	2985.40	3432.00	446.60	260.98	185.62	8.04	5.72
2	29.54	0.88	27.98	30.84	3249.72	96.85	3077.80	3392.40	314.60	171.92	142.68	5.29	4.39
3	29.51	0.73	28.23	30.63	3246.44	80.07	3105.67	3368.93	263.27	140.78	122.49	4.34	3.77
4	29.52	0.62	28.45	30.49	3247.28	68.25	3129.50	3353.35	223.85	117.78	106.07	3.63	3.27
5	29.53	0.55	28.58	30.39	3247.94	60.89	3143.36	3343.12	199.76	104.58	95.18	3.22	2.93
6	29.52	0.50	28.64	30.32	3247.08	55.50	3150.77	3335.20	184.43	96.32	88.12	2.97	2.71
7	29.52	0.47	28.72	30.26	3247.14	51.25	3158.89	3328.60	169.71	88.25	81.46	2.72	2.51
8	29.52	0.44	28.79	30.22	3247.59	48.27	3166.35	3324.20	157.85	81.24	76.61	2.50	2.36
9	29.53	0.40	28.85	30.17	3247.85	44.50	3173.13	3318.58	145.44	74.71	70.73	2.30	2.18
10	29.52	0.39	28.87	30.14	3247.39	42.54	3175.48	3315.40	139.92	71.91	68.01	2.21	2.09
11	29.52	0.37	28.90	30.12	3247.27	40.26	3179.40	3313.40	134.00	67.87	66.13	2.09	2.04
12	29.53	0.35	28.95	30.08	3248.40	37.96	3184.13	3308.62	124.48	64.26	60.22	1.98	1.85
13	29.52	0.33	28.95	30.06	3247.36	36.57	3184.75	3306.09	121.34	62.60	58.73	1.93	1.81
14	29.52	0.32	28.98	30.03	3246.69	35.09	3187.64	3303.14	115.50	59.05	56.45	1.82	1.74
15	29.53	0.31	29.02	30.02	3248.31	33.83	3192.05	3302.49	110.44	56.26	54.18	1.73	1.67
16	29.53	0.29	29.04	30.01	3248.04	32.39	3194.26	3300.69	106.43	53.78	52.64	1.66	1.62
17	29.52	0.28	29.05	29.98	3247.40	31.10	3195.56	3297.54	101.98	51.84	50.14	1.60	1.54
18	29.52	0.28	29.06	29.97	3247.14	30.33	3196.48	3297.07	100.59	50.66	49.92	1.56	1.54
19	29.53	0.26	29.08	29.95	3248.00	29.05	3199.15	3294.79	95.64	48.85	46.79	1.50	1.44
20	29.53	0.26	29.09	29.95	3247.94	28.85	3200.01	3294.72	94.71	47.93	46.78	1.48	1.44
21	29.52	0.25	29.09	29.93	3246.81	27.61	3200.27	3291.83	91.56	46.55	45.01	1.43	1.39
22	29.52	0.25	29.12	29.92	3247.64	26.96	3202.80	3291.20	88.40	44.84	43.56	1.38	1.34
23	29.52	0.24	29.13	29.91	3247.37	26.21	3203.77	3290.15	86.37	43.59	42.78	1.34	1.32
24	29.52	0.23	29.15	29.89	3247.34	25.04	3206.41	3288.27	81.86	40.93	40.93	1.26	1.26
25	29.52	0.22	29.15	29.89	3247.40	24.52	3206.10	3287.50	81.40	41.29	40.11	1.27	1.24
26	29.52	0.22	29.16	29.86	3247.23	23.66	3207.60	3285.11	77.51	39.63	37.88	1.22	1.17
27	29.52	0.21	29.17	29.87	3247.31	23.60	3208.41	3286.07	77.65	38.89	38.76	1.20	1.19
28	29.52	0.21	29.17	29.87	3247.47	23.19	3208.54	3285.31	76.76	38.92	37.84	1.20	1.17
29	29.52	0.21	29.18	29.85	3247.13	22.58	3209.34	3283.69	74.34	37.79	36.56	1.16	1.13
30	29.53	0.20	29.19	29.84	3247.78	22.00	3211.19	3282.91	71.72	36.58	35.14	1.13	1.08

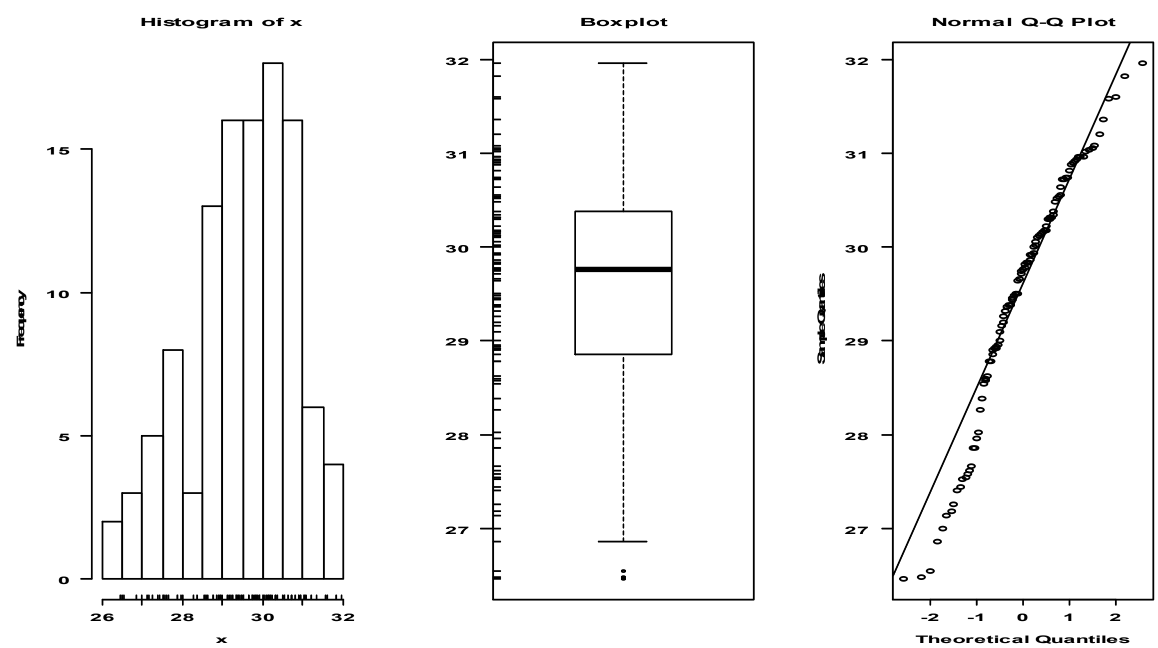
Addendum Table 1. -- Continued.

Pacific ocean perch - Model output for sample sizes [1-50] from 110 Non-subsample basket weights													
	Mean basket weight (kg)				Total Non-subsample weight (kg)								
n	Est.	SD	LB	UB	Est.	SE	LB	UB	CI Range	Δ LB	Δ UB	% LB	% UB
31	29.52	0.20	29.19	29.83	3247.17	21.46	3211.01	3281.76	70.75	36.16	34.59	1.11	1.07
32	29.52	0.19	29.21	29.83	3247.50	20.87	3212.89	3281.58	68.68	34.61	34.07	1.07	1.05
33	29.52	0.19	29.21	29.83	3247.30	20.46	3213.13	3280.80	67.67	34.17	33.50	1.05	1.03
34	29.52	0.18	29.22	29.82	3247.35	19.83	3214.39	3279.88	65.48	32.95	32.53	1.01	1.00
35	29.52	0.18	29.23	29.81	3247.48	19.55	3215.14	3279.26	64.11	32.33	31.78	1.00	0.98
36	29.52	0.17	29.24	29.81	3247.25	19.17	3216.03	3279.16	63.13	31.21	31.91	0.96	0.98
37	29.52	0.17	29.23	29.80	3247.47	18.97	3215.81	3278.54	62.73	31.66	31.07	0.98	0.96
38	29.52	0.17	29.24	29.80	3247.48	18.70	3216.63	3278.06	61.43	30.85	30.58	0.95	0.94
39	29.52	0.16	29.25	29.78	3247.38	17.92	3217.58	3276.08	58.50	29.79	28.70	0.92	0.88
40	29.52	0.16	29.25	29.79	3247.48	17.67	3218.00	3276.35	58.36	29.48	28.87	0.91	0.89
41	29.52	0.16	29.26	29.79	3247.50	17.54	3218.71	3276.39	57.68	28.79	28.89	0.89	0.89
42	29.52	0.15	29.27	29.78	3247.30	17.02	3219.44	3275.70	56.26	27.87	28.39	0.86	0.87
43	29.52	0.15	29.26	29.77	3247.18	17.00	3218.86	3275.03	56.18	28.32	27.85	0.87	0.86
44	29.52	0.15	29.28	29.76	3247.39	16.26	3220.50	3273.80	53.30	26.89	26.41	0.83	0.81
45	29.52	0.15	29.28	29.76	3247.34	16.27	3220.46	3273.89	53.44	26.88	26.56	0.83	0.82
46	29.52	0.14	29.29	29.76	3247.62	15.74	3221.57	3273.41	51.84	26.06	25.79	0.80	0.79
47	29.52	0.14	29.29	29.75	3247.21	15.55	3221.69	3272.34	50.65	25.52	25.12	0.79	0.77
48	29.52	0.14	29.29	29.75	3247.44	15.34	3222.04	3272.59	50.55	25.40	25.15	0.78	0.77
49	29.52	0.13	29.29	29.74	3247.08	14.84	3222.37	3271.13	48.76	24.70	24.05	0.76	0.74
50	29.52	0.13	29.31	29.74	3247.45	14.57	3223.57	3271.53	47.96	23.88	24.08	0.74	0.74

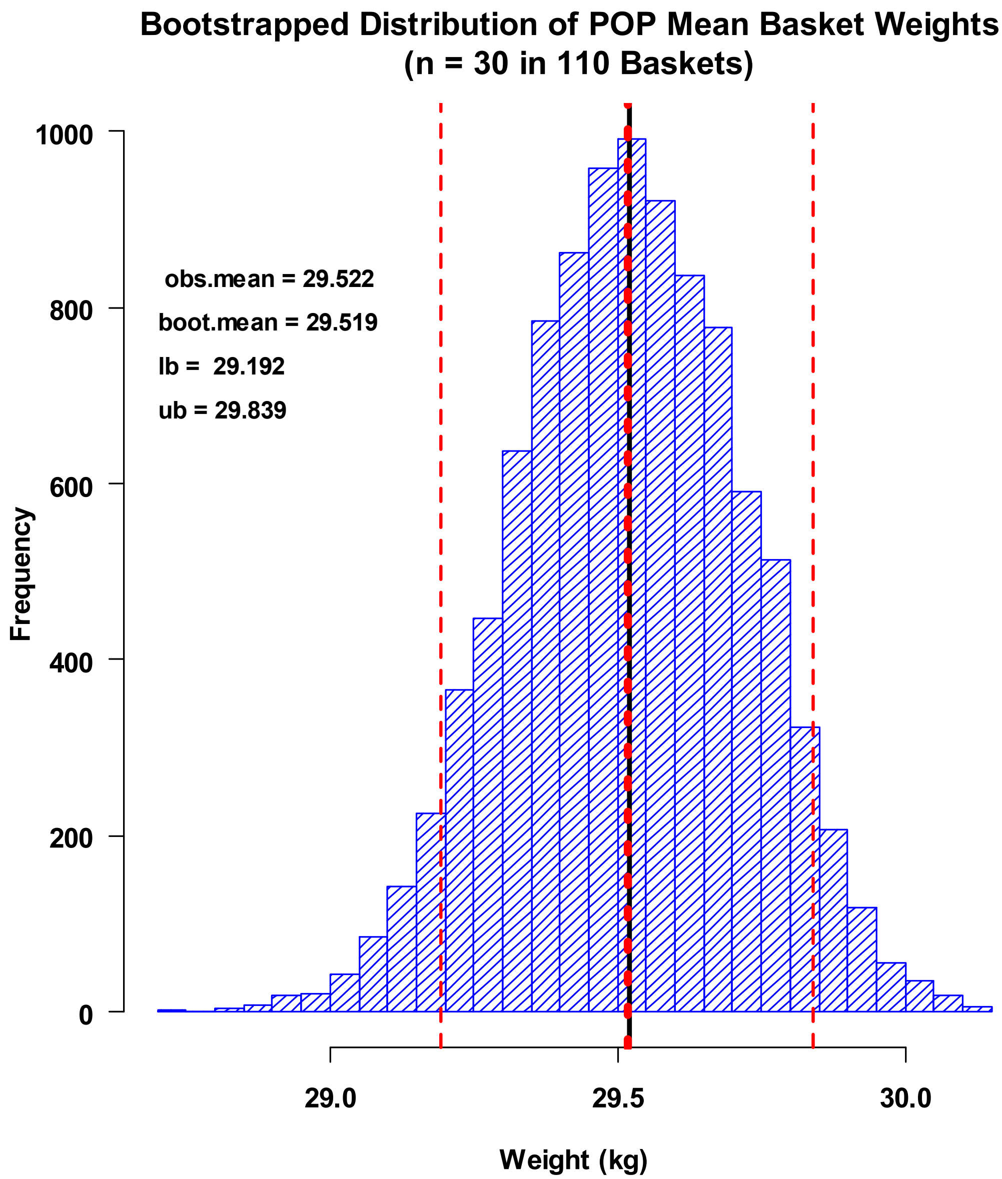
(a)



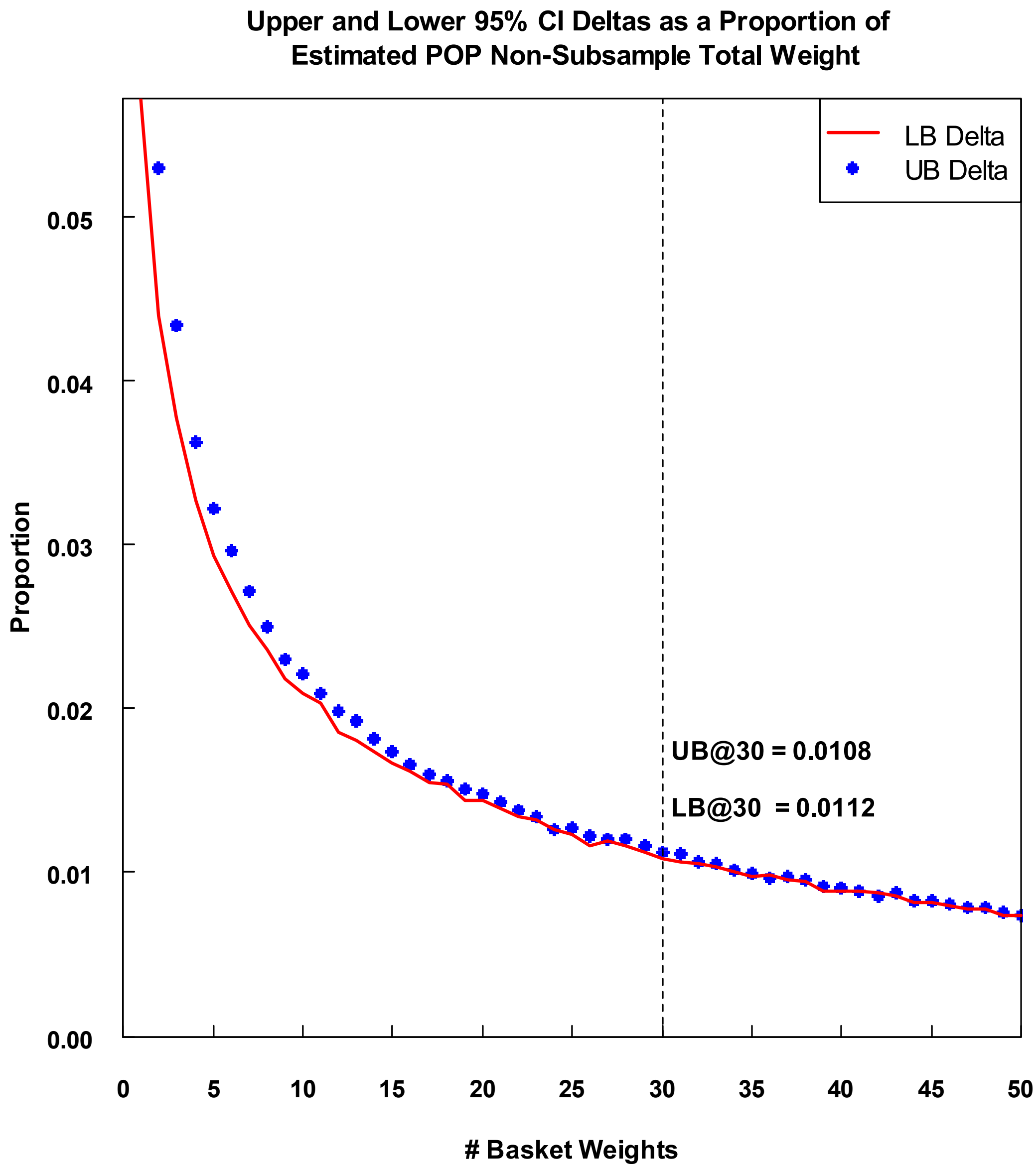
(b)



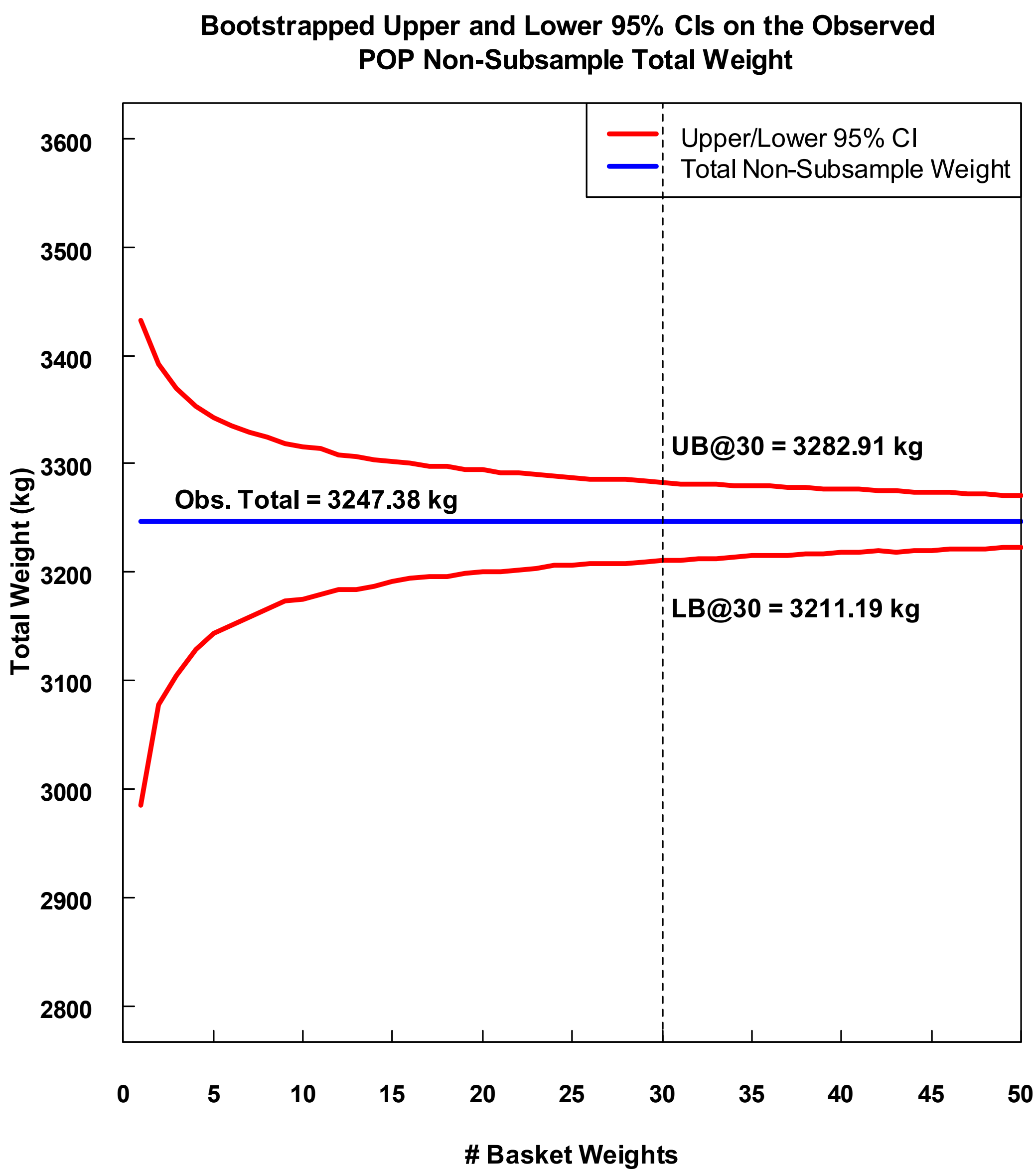
Addendum Figure 1. -- Histogram, boxplot and normal Q-Q plot of non-subsample basket weights for the 2018 data set. Panel (a) is raw n = 111 basket weights, and (b) is n = 110 after removal of one outlier.



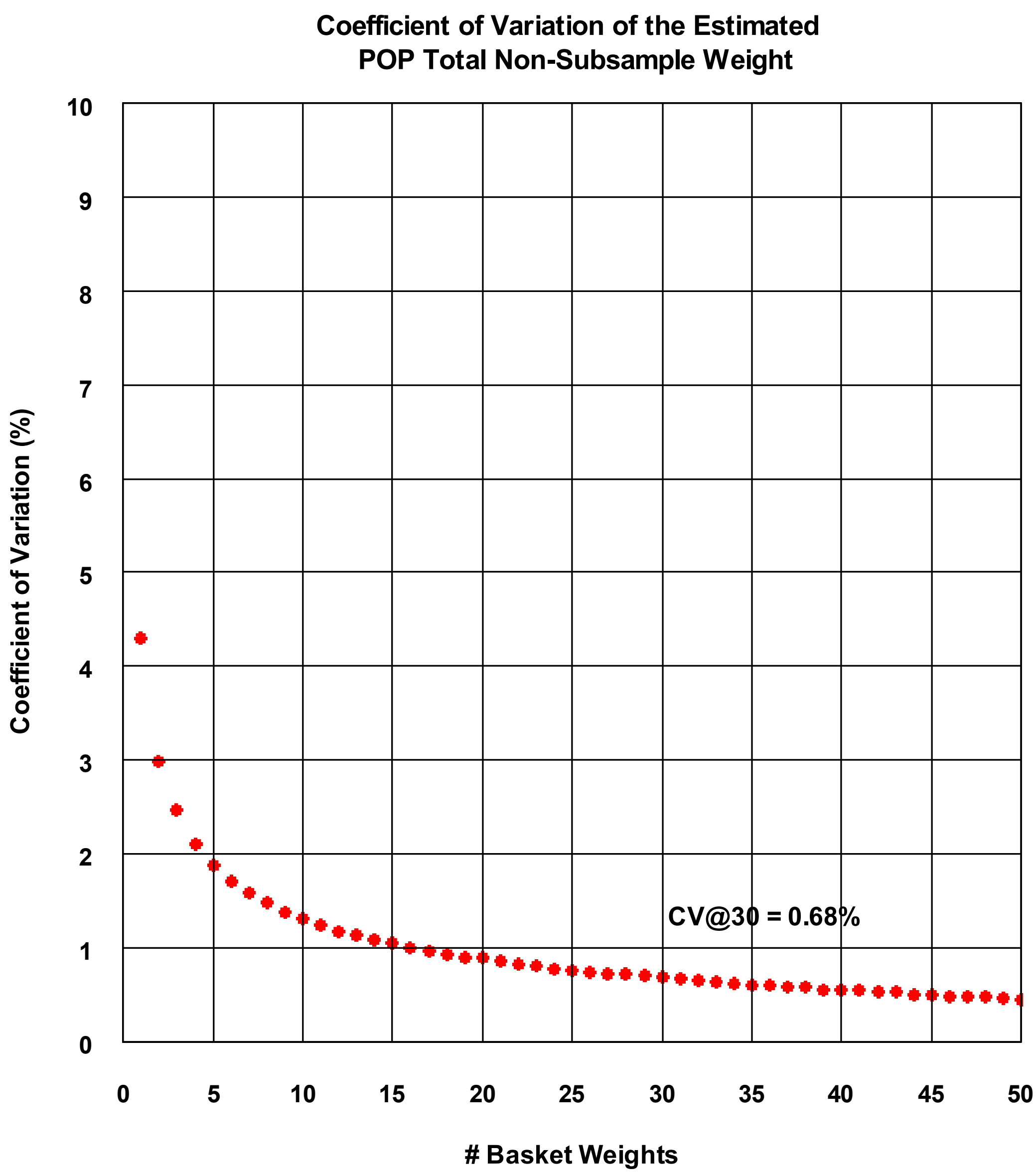
Addendum Figure 2. -- The distribution of the 10,000 mean non-subsample mean basket weights for the 2018 data set. Shown are the mean of the bootstrapped means (bold dotted red), the upper and lower 95% confidence bounds (dashed red) of the bootstrapped mean, and the observed mean of the 100 non-subsample basket weights (solid black).



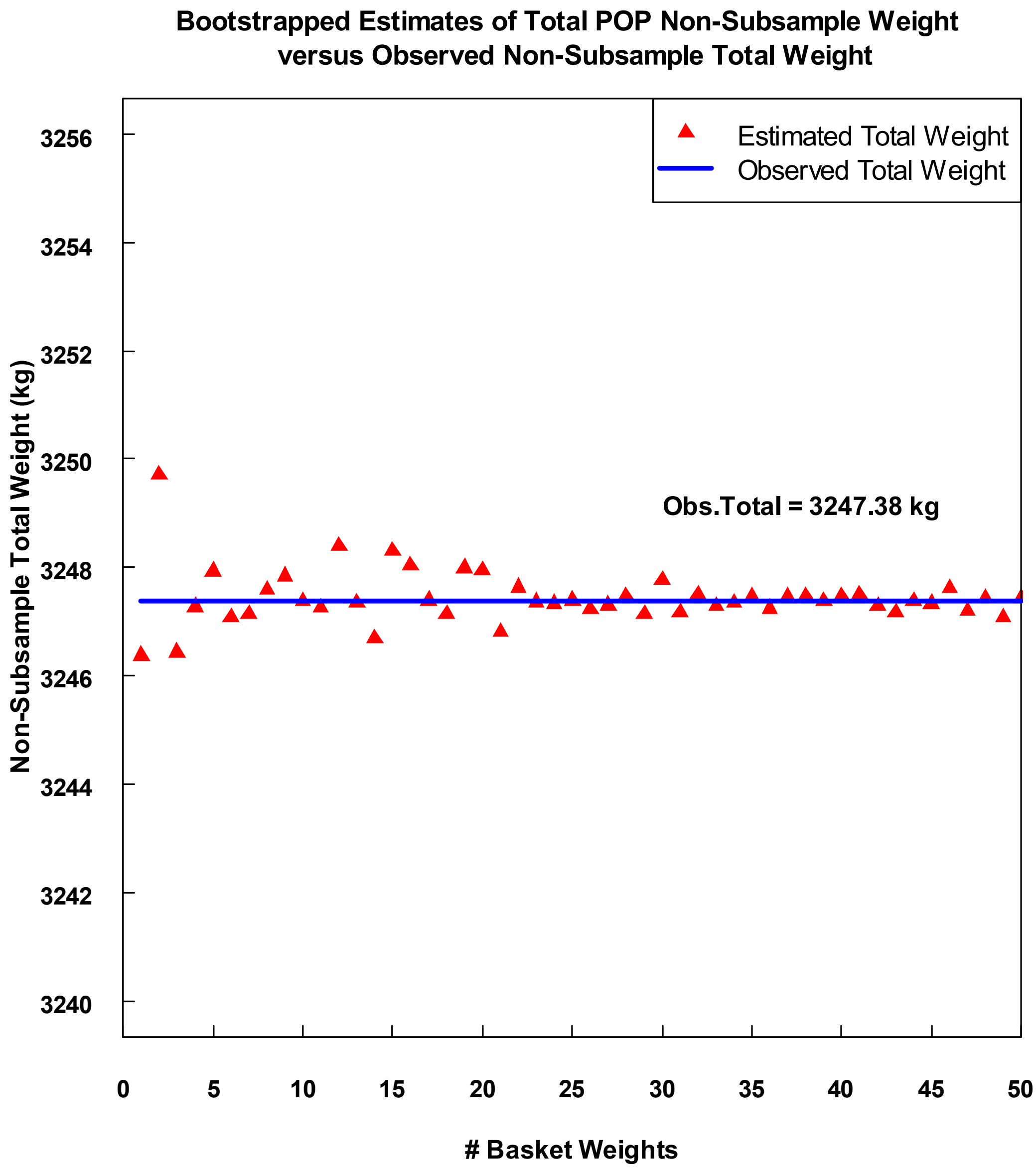
Addendum Figure 3. -- Lower and upper 95% confidence bounds deltas (deviations) as a proportion of the estimated non-subsample total weight versus sample size for the 2018 data set. Indicated are the lower and upper proportions for sample size n = 30 (dashed line).



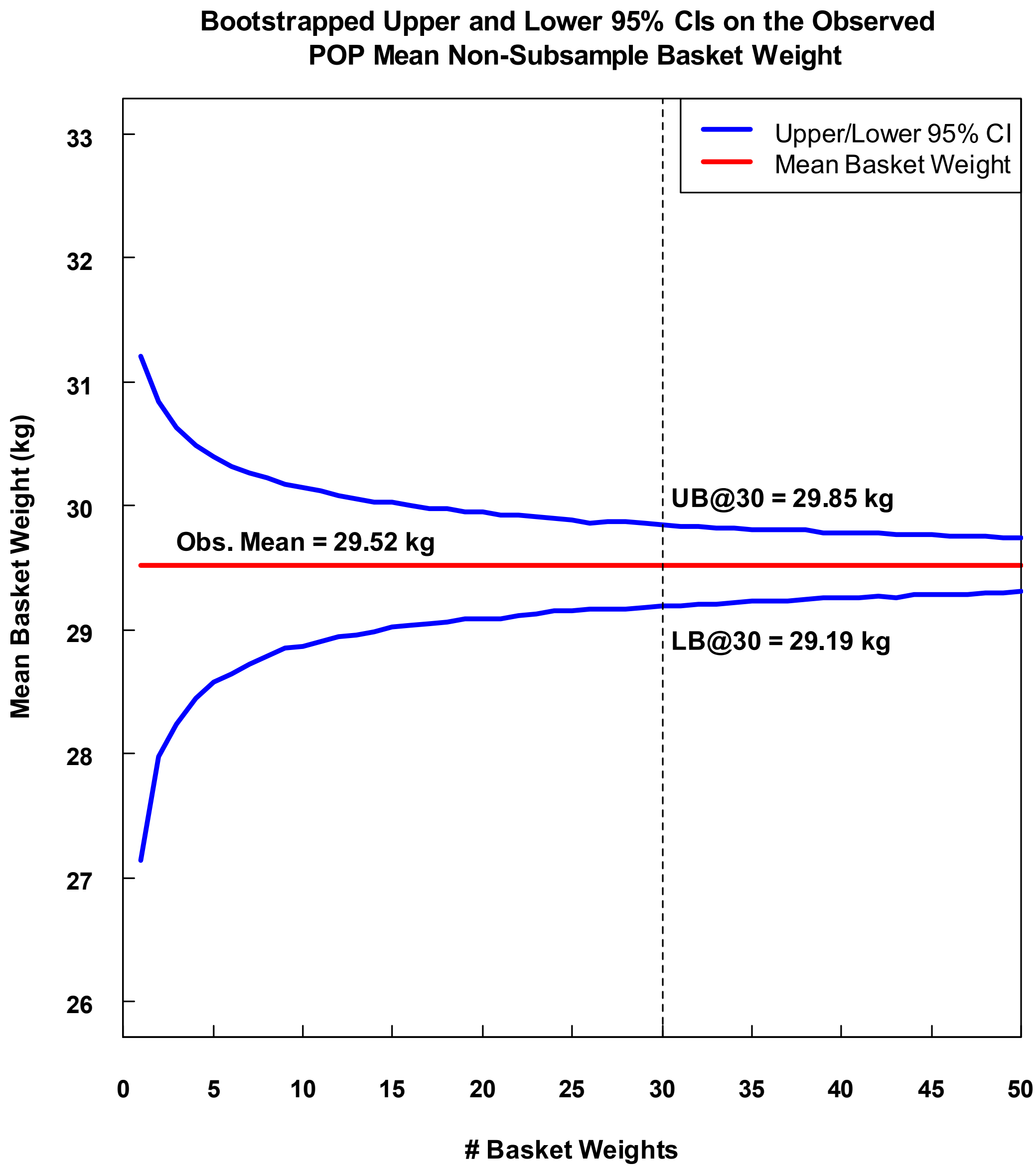
Addendum Figure 4. -- Lower and upper 95% confidence bounds on the non-subsample total weight estimate versus sample size for the 2018 data set. Also shown are the observed total non-subsample weight (blue), and the lower and upper bounds for sample size n = 30 (dashed line).



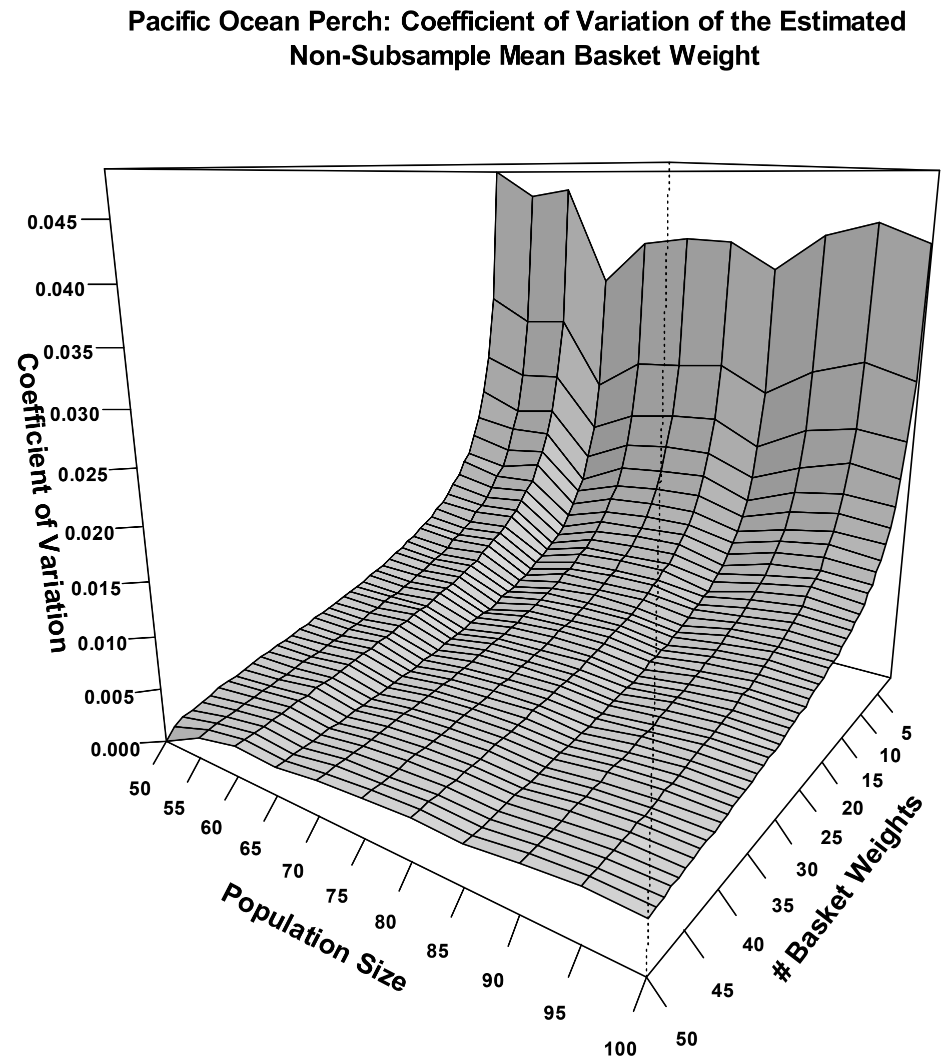
Addendum Figure 5. -- The coefficient of variation (CV) of the estimated non-subsample total weight versus sample size for the 2018 data set. Indicated is the CV of the estimate at sample size n = 30.



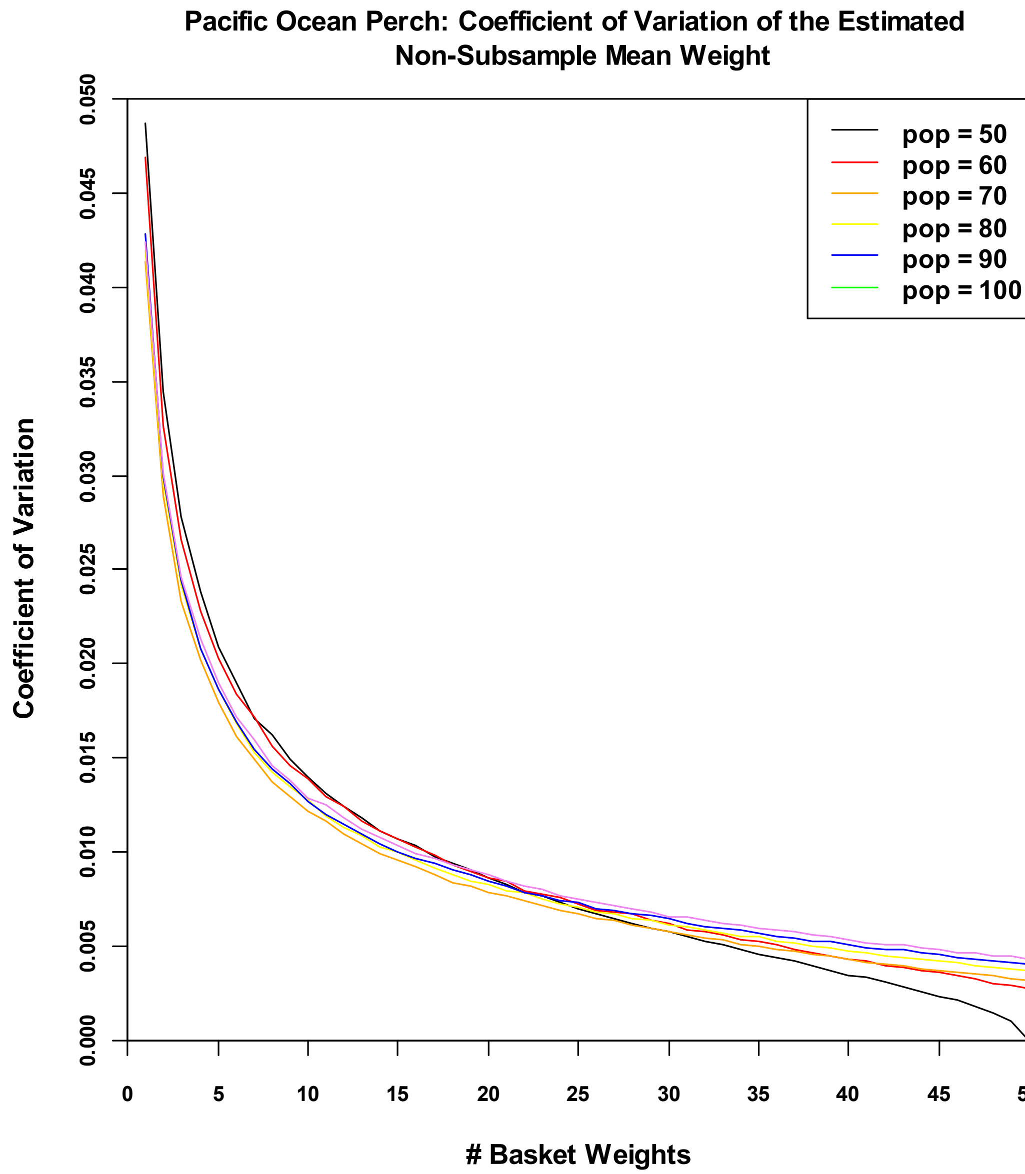
Addendum Figure 6. -- Estimates of the mean non-subsample total weight versus sample size for the 2018 data set. Indicated is the observed total non-subsample weight (blue line).



Addendum Figure 7. -- Lower and upper 95% confidence bounds on the non-subsample mean basket weight versus sample size for the 2018 data set. Also shown are the observed non-subsample mean weight (red), and the lower and upper bounds for sample size n = 30 (dashed line).



Addendum Figure 8. -- The coefficient of variation of the estimated mean basket weight vs sample size (n = 1 to 50) and catch volume (population size) from 50 to 100 baskets for the 2018 data set. (The calculated coefficients of variation would be identical if shown for estimated non-subsample total weight.)



Addendum Figure 9. -- The coefficient of variation of the estimated mean basket weight versus sample size (n = 1 to 50) and catch volume (population size) from 50 to 100 (by 10) baskets for the 2018 data set. (The calculated coefficients of variation would be identical if shown for estimated non-subsample total weight.)

U.S. Secretary of Commerce
Wilbur L. Ross, Jr.

Acting Under Secretary of Commerce for Oceans and Atmosphere
Dr. Neil Jacobs

Assistant Administrator for Fisheries
Chris Oliver

March 2020

www.nmfs.noaa.gov

OFFICIAL BUSINESS

National Marine Fisheries Service
Alaska Fisheries Science Center
7600 Sand Point Way N.E.
Seattle, WA 98115-6349