NOAA Technical Memorandum NMFS-AFSC-464

Subsampling Catches to Determine Sex-specific Length Frequency in Alaska Fisheries Science Center Bottom Trawl Surveys

P-J. F. Hulson, B. C. Williams, M. D. Bryan, J. Conner,
M. R. Siskey, W. T. Stockhausen, S. F. McDermott, and
W. C. Long

February 2023

U.S. DEPARTMENT OF COMMERCE

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

This document should be cited as follows:

Hulson, P-J. F., B. C. Williams, M. D. Bryan, J. Conner, M. R. Siskey, W. T. Stockhausen, S. F. McDermott, and W. C. Long. 2023. Subsampling catches to determine sex-specific length frequency in Alaska Fisheries Science Center bottom trawl surveys. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-464, 38 p.

Subsampling Catches to Determine Sex-specific Length Frequency in Alaska Fisheries Science Center Bottom Trawl Surveys

P-J. F. Hulson[1], B. C. Williams[1], M. D. Bryan[2], J. Conner[3], M. R. Siskey[4], W. T. Stockhausen[2], S. F. McDermott[3], and W. C. Long[3]

[1]Auke Bay Laboratories
Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
17109 Pt. Lena Loop Road
Juneau, AK 99801

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

[3]Resource Assessment and Conservation
Engineering Division
Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
7600 Sand Point Way NE
Seattle, WA 98115

[4]School of Aquatic and Fishery Sciences
University of Washington
1122 NE Boat St, Box 355020
Seattle, WA 98195-5020

U.S. DEPARTMENT OF COMMERCE

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

NOAA Technical Memorandum NOAA-TM-AFSC-464

February 2023

ABSTRACT

Length frequency sub-samples are collected to provide sex-specific length composition data for groundfish stock assessments during the Alaska Fisheries Science Center bottom trawl surveys. The overall workload on board of survey vessels, which included determining sex for each individual in the length frequency sub-samples (e.g., upwards of 200 per haul per species), has likely led to unacceptable numbers of acute and repetitive stress injuries of survey staff. In December 2021, a working group was formed to evaluate the effect of reducing length frequency sub-sample sizes for which sex is determined on length composition uncertainty. Over the spring of 2022, the working group developed a bootstrap-simulation approach that determines changes in uncertainty for reduced sub-sample sizes. For the majority of species limiting sub-samples to 100 fish per haul reduces length composition input sample sizes by ≤ 25% from status quo length composition input sample sizes. Therefore, upon meeting with stock assessment authors the joint recommendation between the working group and assessment authors to the survey staff was to limit length frequency sub-sampling for sex determination to 100 subsampled fish per haul.

CONTENTS

INTRODUCTION

The National Oceanic and Atmospheric Administration's (NOAA's) Alaska Fisheries Science Center (AFSC) has conducted effort-standardized bottom trawl surveys in the eastern Bering Sea (EBS), Gulf of Alaska (GOA) and Aleutian Islands (AI) since 1982 (Stauffer 2004). These resource surveys comprehensively catalog the biota encountered at each sampling station. Observations of sex-specific length distributions for designated species encountered within a catch sample are collected by 1) sorting the trawl sample by species, 2) weighing each species in aggregate, 3) obtaining a random sub-sample of target sample size, 4) sorting the sub-sample by sex (each fish is cut with a scalpel, gonads are identified and placed in a sex-specific receptacle), and 5) measuring and recording each fish length (each fish is placed on a measuring board, length is identified and the species, sex and measurement are recorded on a computer).

This length measurement process is repeated over 100,000 times in a given region's survey each year, representing a daunting amount of work for six field scientists per research vessel. Each year, this work flow contributes to acute and repetitive stress injuries, some requiring medical interventions and claims to the U.S. Office of Workers' Compensation Programs. Another consequence of the intensity of this work is unrecoverable errors in the observed data; as fatigue or injuries accumulate during the course of field work, so do data collection errors (e.g. failure to properly encode length measurements with the correct species or sex), despite extensive *in situ* quality assurance protocols.

In December of 2021, an AFSC working group was formed (the 'Survey Workload Optimization' Working Group, SWO) to evaluate the potential impacts of reductions to sex-sampling within the length frequency data collection process on survey data products. From the stock assessment perspective, the primary concern when reducing the sex-sampling within the length frequency data collections was the impact on the uncertainty in subsequent sex-specific length composition data. This is of particular importance to stock assessments that are sex-specific, such as the flatfish assessments conducted by AFSC. To that end, we evaluated the impact of reducing sex-sampling within the length frequency data collection to address two objectives: 1) the impact of reducing sampling for sex on the uncertainty in the sex-specific length composition, and, 2) identifying an acceptable level of increase in uncertainty in sex-specific length composition data due to sub-sampling for sex.

MATERIALS AND METHODS

Survey Data

Data collection for each AFSC groundfish survey is described in respective NOAA Technical Memorandums (EBS: Lauth et al. 2019, AI: von Szalay et al. 2017, GOA: von Szalay and Raring 2018). Fundamental methods of length sample collection are generally synchronized between these surveys, with species-specific exceptions for minimum length sub-sample size. Survey protocols include sampling bimodal length distributions separately (e.g., many juveniles and fewer adults within one species), which are then extrapolated to the total estimated number of individuals in each length mode category for that haul sample and aggregated. Some species' gonads are difficult to distinguish in the field when they are small, so a length threshold is determined (e.g., approximately 10-15 cm for walleye pollock) below which sex is coded as "unsexed".

Survey length sub-sampling protocols have been altered at least once during the time series. Historically, AFSC trawl survey standard operating procedures directed scientists to collect a minimum of 200 length specimens per species where more than 200 were in a trawl sample. However, analysis of length frequency sample size found that 10 random length specimens should be sampled for each length bin range of a sample (Gerritsen and McGrath 2007). This change in methodology has been applied to some species sampled in the EBS survey, although with varying consistency.

The Resource Assessment and Conservation Engineering (RACE) Division of AFSC hosts data collected from bottom trawl surveys on an internal SQL-based server ("RACEBASE"). For the analysis, RACEBASE was queried using the R package sumfish (https://github.com/afsc-gap-products/sumfish). Raw length sample data were summarized to determine the average number of sex-specific and total length frequency samples over the most recent three years by survey and species (EBS: 2017-2019; AI: 2014, 2016, 2018; GOA: 2017, 2019, 2021). Additionally we summarized the number of sex-specific lengths and total number of lengths collected per haul over time. The most frequently sampled species within each survey were then selected to be included in this analysis; these included primarily stocks that are assessed at AFSC with age-structured assessment models, but it also include stocks assessed solely with trawl survey biomass estimates.

Expanding Length Frequency to Population Abundance at Length

Length frequency samples collected by the AFSC bottom trawl surveys are expanded by area-swept type catch-per-unit-effort (CPUE) and stratum area to obtain estimates of population abundance-at-length (i.e., design-based expansion). In a design-based expansion process, this is often referred to as the 'first stage expansion' and is a common method to obtain population estimates at length from area-swept survey data (Miller and Skalski 2006, Ailloud and Hoenig 2019). Population abundance-at-length are computed for three sex categories (males, females, and unsexed) at the stratum level, which are then summed across strata to obtain the population abundance-at-length for the management-scale region (i.e., EBS, AI, or GOA). These can also be summed to any sub-region level.

In the first step of this process, we compute the overall population numbers in year-y for stratum-*st* ($\widehat{N}_{st,y}$) with

$$\widehat{N}_{st,y} = \overline{CPUE}_{st,y} \cdot A_{st} \quad , \tag{1}$$

where A_{st} is the area of stratum-*st* (in km^2), and $\overline{CPUE}_{st,y}$ is the average catch per unit effort (in numbers per km^2) of fish captured across the hauls within a strata, given by

$$\overline{CPUE}_{st,y} = \frac{1}{H_{st,y}} \sum_{h=1}^{H_{st,y}} C\,PUE_{h,st,y} = \frac{1}{H_{st,y}} \sum_{h=1}^{H_{st,y}} \frac{n_{h,st,y}}{E_{h,st,y}} \quad , \tag{2}$$

where $H_{st,y}$ is the number of hauls, $CPUE_{h,st,y}$ is the catch (in numbers) per unit effort for haul-h, $n_{h,st,y}$ is the catch (in numbers) in haul-h, and $E_{h,st,y}$ is the effort in haul-h. $E_{h,st,y}$ is computed as the net width multiplied by the distance fished of the tow, or, the area swept by the haul (in km^2). Next, the ratio of catch per unit effort among hauls per stratum ($\hat{p}_{CPUE,h,st,y}$) is computed by

$$\hat{p}_{CPUE,h,st,y} = \frac{CPUE_{h,st,y}}{\sum_{h=1}^{H_{st,y}} C\,PUE_{h,st,y}} \quad , \tag{3}$$

where $CPUE_{h,st,y}$ is the catch (in numbers) per unit effort within a haul-h. We then compute the sex-specific ratio of the total number of lengths sampled within a haul by length ($\hat{p}_{sx,l,h,st,y}$) with

$$\hat{p}_{sx,l,h,st,y} = \begin{cases} (1)\, \dfrac{N_{sx,l,h,st,y}}{N_{h,st,y}} \\ (2)\, \dfrac{\sum_{h=1}^{H_{st,y}} \left[N_{sx,l,h,st,y}/N_{h,st,y}\right]}{\sum_{sx=1}^{3} \sum_{l=1}^{L} \sum_{h=1}^{H_{st,y}} \left[N_{sx,l,h,st,y}/N_{h,st,y}\right]} \end{cases} \quad , \tag{4}$$

where $N_{sx,l,h,st,y}$ is the count of sampled fish, in numbers, by sex-*sx* and length-*l*. In some cases there are hauls that have catch for a species but without corresponding length frequency data, which is addressed differently between the EBS survey and the GOA and AI surveys. If no length frequency samples are missing from the data, case (1) is applied. Missing length frequency data in the EBS is handled by assigning the estimated proportion of the stratum population to a substitute length code, so that the abundance is not extrapolated to observed length categories. For missing length frequency data in the GOA and AI surveys, case (2) of equation (4) is applied in order to account for the unknown length frequency in these hauls. Finally, we estimate the sex-specific population abundance at length within strata-*st* with

$$\widehat{N}_{sx,l,st,y} = \widehat{N}_{st,y} \cdot \hat{p}_{CPUE,h,st,y} \cdot \hat{p}_{sx,l,h,st,y} \tag{5}$$

and to obtain the sex-specific estimates of population abundance at length in a management area one would simply sum $\widehat{N}_{sx,l,st,y}$ across strata.

Bootstrap-simulation Framework

To evaluate the effect of sub-sampling length frequency collections on uncertainty in calculated length compositions, we developed a bootstrap-simulation framework that 1) allows for reductions in the number of sexed length frequencies collected historically, and 2) conducts the first-stage expansion process for each bootstrap replicate of length frequency information to generate length compositions. In simple terms, the simulation framework that we developed selects a pre-determined number of fish from the length frequency collections that are then sexed, the remaining length frequency data (regardless of whether sex was actually determined in the historical data) are classified as ‘unsexed’.

The bootstrap-simulation framework is composed of a suite of nested resampling processes. Bootstrap resampling was performed either with replacement (wr) or without replacement (wor) depending on the needs of a particular protocol. Functions to conduct the sampling protocols were developed in a compartmentalized manner to provide for substantial flexibility in exploring desired resampling scenarios. The order of operations (Fig. 1) has the following steps, with steps 1-3 being optional switches:

1. Resample hauls (wr) from the set of hauls with associated catch-per-unit-effort (in numbers)
2. Within the resampled hauls from step 1, resample the observed lengths (wr)
3. From the resampled lengths in step 2, subset the lengths (wor) with observed sex (either male or female) and sample these sex-length pairs at the sub-sampling level determined in step 2; equation (6)
4. Calculate sex-specific population abundance at length, using equations (1) - (5)

The core of the bootstrap-simulation function (step 3 above) is designed to explore reductions in the sample size of lengths that are then sexed on a per haul basis. In this bootstrap-simulation, the number of lengths in a given haul must be less than or equal to the desired sample size x determined in step 2. In step 3, when the number of resampled lengths from step 2 in a haul n_l is less than x then n_l is used directly, if $n_l > x$ then a random draw of lengths is taken without replacement in step 3, such that

$$l_x = \begin{cases} n_l & n_l \leq x \\ n_{lx} & n_l > x,\ wor \end{cases} . \tag{6}$$

In the case where $n_l > x$, the random draw of lengths are taken without replacement because the lengths within the haul have already been resampled in step 2, and the remainder of the lengths that were not sampled are then moved into the ‘unsexed’ category. The bootstrap-simulation then repeats steps 1-3 iteratively for each sex sub-sample size determined in step 1, providing iterated sex-specific population abundance at length that was influenced by sex-subsampling.

We applied the bootstrap-simulation to species that were most commonly captured in the EBS, AI, and GOA bottom trawl surveys (Tables 1 - 3). The sub-sample levels that we evaluated for subsequent sex determination from the length frequency collections were 50, 75, 100, 125, 150, and 175 samples. We also ran the bootstrap-simulation for the historical number of sexed length frequency collections without subsetting in order to compare the base level uncertainty to the increase in uncertainty gained through sub-sampling. We ran the bootstrap-simulation for 500 iterations, a level at which the variability in population abundance-at-length results had stabilized (that is, the results did not appreciably change from one iteration to the next). We applied the bootstrap-simulation to the most recent 3 years of the respective bottom trawl surveys, which were the most indicative of the current sampling levels. The bootstrap-simulation was developed in R (R Core Team 2022) and is available via GitHub as an R package (https://github.com/BenWilliams-NOAA/swo).

Computing Effective and Input Sample Size

We used two performance metrics to quantify potential changes in uncertainty of length compositions (i.e., population abundance-at-length) owing to reduced sampling of sex-specific length frequency information. First, effective sample size was used to compare each bootstrap replicate of length composition to the original length composition calculated from the complete historical dataset. Second, input sample size was used to summarize uncertainty across bootstrap replicates as the harmonic mean of effective sample size replicates.

Effective sample size (ESS), as introduced by McAllister and Ianelli (1997), is a statistic that can evaluate the level of intra-haul correlation in composition samples that are collected on a survey (whether from age or length frequency collections). It is also a statistic that can evaluate the amount of uncertainty in an estimated composition compared to an observed composition. Effective sample size is usually calculated from stock assessment output, comparing the composition predicted by the model to the observed composition fed to the model. Effective sample size is given by

$$ESS = \frac{\sum_{c=1}^{C} E_c\,(1 - E_c)}{\sum_{c=1}^{C}(E_c - O_c)^2} \quad , \tag{7}$$

where E_c is the estimated proportion for category-c (which can be length, age, or any other arbitrary category across which proportions are computed) and O_c is the observed proportion. Therefore, higher ESS indicates less uncertainty in the size composition estimate.

In this bootstrap-simulation, we used effective sample size to calculate uncertainty in length compositions for each simulation replicate. Length composition derived from the historical bottom trawl surveys was treated as the observed proportions O_c in equation (7). For each iteration of the simulation for a determined sex sub-sample size, we computed a sex-specific estimated proportion (E_c) that was then compared to the historical sex-specific length composition (the effective sample size for the total length composition, as the sum of population abundance at length across sex categories, was also computed). Thus, for each iteration of the simulation, we computed an effective sample size that quantifies the amount of uncertainty that resulted from each iteration of sub-sampling sexed length frequency data.

Input sample size (ISS) is defined as a metric of uncertainty used in data-weighting procedures for stock assessment models. An input sample size is usually assigned to annual length compositions in the model fitting process, but there are a variety of methods for calculation – many of which are closely related to the information content of the data product in question. To summarize uncertainty across bootstrap replicates of ESS, we calculated ISS as the harmonic mean of effective sample size across bootstrap simulations

$$ISS = \left(\frac{\sum_{i=1}^{I} E\,SS_i}{I}\right)^{-1} \quad , \tag{8}$$

where ISS is the input sample size, ESS_i is the effective sample size for iteration-i, and I is the total number of iterations for which the bootstrap procedure is run. The harmonic mean has been shown to reduce bias in recovering the true sample size in simulations for a multinomial distribution. Due to this reduction in bias the harmonic mean has also been recommended to determine the 'input sample size' that is used in stock assessment models to fit compositional data (Stewart and Hamel 2014). Herein, when we use the term 'effective sample size or ESS' we are referring to the effective sample sizes that were computed for each iteration of the bootstrap-simulation from equation (7), when we use the term 'input sample size or ISS' we are referring to the harmonic mean of the iterated effective sample sizes from equation (8).

RESULTS

The species with the greatest average total number of lengths sampled varied by the regional-specific surveys (Tables 1 - 3). The top five species for the eastern Bering Sea shelf bottom trawl survey were walleye pollock, yellowfin sole, northern rock sole, flathead sole, and arrowtooth flounder; Pacific ocean perch, northern rockfish, walleye pollock, arrowtooth flounder, and Atka mackerel for the Aleutian Islands bottom trawl survey; and arrowtooth flounder, Pacific ocean perch, walleye pollock, flathead sole, and rex sole for the Gulf of Alaska bottom trawl survey. The species with the greatest number of lengths was similar among years as well. The species most frequently sampled for lengths across all the surveys was walleye pollock in the eastern Bering Sea survey (Table 2), with nearly 50% more samples in some years as the next most frequently sampled species (arrowtooth flounder in the Gulf of Alaska, Table 3). The number of sexed sampled lengths per haul for species of interest are summarized in (Figs. 2- 4) for the most recent three years of data. The species with the greatest number of hauls exceeding the sampling minimum of 200 samples are walleye pollock, yellowfin sole, and northern rock sole in the eastern Bering Sea; Pacific ocean perch, northern rockfish, and walleye pollock in the Aleutian Islands; and Pacific ocean perch, arrowtooth flounder, and walleye pollock in the Gulf of Alaska. The largest frequency of hauls that exceeded the sampling target of 200 length samples compared to all other species in each survey was for eastern Bering Sea walleye pollock (Fig. 3).

In general, the mean of the regional, replicated sex-specific ESS for length compositions increased as the number of collected lengths for which sex was determined increased (Figs. 5 - 7). For a number of species, the sub-sampling case of 175 sexed lengths was nearly equivalent to the base case (in which no sub-sampling occurred), indicating that for these species around 175 or less lengths have been subsequently sexed historically. There were notable exceptions to the sub-sampling case of 175 being similar to the base case, indicating that in the historical surveys there were a number of hauls in which more than 175 sex-specific lengths were measured. These exceptions include arrowtooth flounder in the eastern Bering Sea (Fig. 6) and Gulf of Alaska (Fig. 7), and walleye pollock in the eastern Bering Sea (Fig. 6). For all the species and regions investigated the mean of the replicated ESS for the total length composition (the sum of the female, male, and unsexed population at length estimates) were insensitive to the sexed sub-sampling cases.

Results of the sex-specific ISS for length composition mimicked the results of the replicated ESS for each region; as the sexed sub-sampling level increased the ISS increased (Figs. 8 - 10). It was usually the case that the magnitude of ISS was sensitive to the species and the survey year. In general, larger ISS were associated with species with larger sample sizes. While, for any given species, the effect of the survey year on the ISS is associated with the sample size for that year, it is also related to the structure of the length composition (i.e., uni- or multi-modal length compositions). Sex-specific ISS for length frequency ranged from around 50 to 1,200 for the Aleutian Islands bottom trawl survey (Fig. 8), 500 to 3,000 for the eastern Bering Sea bottom trawl survey (Fig. 9), and 50 to 3,000 for the Gulf of Alaska bottom trawl survey (Fig. 10). Similar to the results of the replicated ESS, the ISS for the total length composition remained stable across the sexed sub-sampling cases investigated. Total ISS for length were species-specific, with values around 50 to 500 for less frequently sampled species and 1,000 to 4,500 for more frequently sampled species in the Aleutian Islands, eastern Bering Sea, and Gulf of Alaska.

Compared to the base ISS (the case using the full historical data set), the proportion of the base ISS for each sexed sub-sampling case was similar for females and males, but they were different across species and survey years (Figs. 11 - 13). The sex-specific proportion of base ISS increased as the sexed sub-sample size increased, and for the smallest sexed sub-sampling case the sex-specific ISS was less than 0.5 of the base ISS for some species. Again, the proportion of the base ISS for the total length composition was relatively unchanged across the sexed sub-sampling cases. For a number of species and survey years the sex-specific proportion for the sexed sub-sampling cases were similar to the base ISS, particularly for sexed sub-sampling sizes greater than 100. For these species, the proportion of the sexed sub-sample size of 100 was greater than 0.75 of the base ISS. The most notable exception was for eastern Bering Sea walleye pollock in each year of the survey, whose proportion of the base ISS for each of the sexed sub-sampling cases investigated ranged from 0.25 to 0.5 (Fig. 12). The second most notable exception was for arrowtooth flounder in the Gulf of Alaska (Fig. 13). Other species for which ISS for sexed sub-sample sizes greater than 100 were less than 0.75 of the base ISS included eastern Bering Sea Pacific cod in 2019, eastern Bering Sea and Aleutian Islands flathead sole in 2018, eastern Bering Sea arrowtooth flounder in 2018 and 2019, and Gulf of Alaska flathead sole in 2021.

For some species the average amount of length frequency samples that would not have been subsequently sexed (for the sex sub-sampling cases of 100 - 175) ranged in the thousands of fish (Table 4). The largest reduction of sexed length frequency collections by region was for eastern Bering Sea walleye pollock, Aleutian Islands Pacific ocean perch, and Gulf of Alaska arrowtooth flounder. For each region, across the species investigated here, the range of sex samples that would not have been taken within a given survey ranged on average from thousands to tens of thousands of fish.

DISCUSSION

Two of the primary results from this analysis were 1) with an increasing level of sub-sampling (i.e., fewer samples) of sex-specific length frequency the uncertainty in the resulting sex-specific length composition increases, and 2) the uncertainty in the total length composition (which is the result of the sum of the sex-specific length population estimates) is unaffected by sub-sampling of sex from the length frequency data. Across the majority of stocks the resulting ISS for sex-specific length composition was reduced by less than 25% after decreasing the number of length frequency collections that are subsequently sexed to 100 - 175 fish when compared to the ISS that resulted from the full historical datasets without any sex sub-sampling. There were notable exceptions, however, primarily for walleye pollock in the eastern Bering Sea, for which the ISS for the sex sub-sampling cases was greatly reduced when compared to the base case. This reduction in the ISS can be explained by the large reduction in sexed sample sizes compared to the base case (nearly 15,000 less sexed samples on average for the sex sub-sample size case of 100). The amount of sampling for walleye pollock in the eastern Bering Sea survey is much larger than for any of the other species in any of the surveys, and this result points to a possible threshold in the reduction of ISS as it relates to the magnitude of sex-specific length composition sample size in any given haul during a survey. Although, while the ISS for sex-specific length composition resulted in a larger decrease compared to other species, the eastern Bering Sea walleye pollock stock assessment is a single-sex assessment and uses the total length composition (Ianelli et al. 2021). Therefore, the ISS for the assessment model would remained unchanged regardless of the sub-sampling level for sex determination of the length frequency collections. This same single-sex modeling approach is used for the eastern Bering Sea Pacific cod stock assessment (Thompson et al. 2021), so while the sex-specific ISS was reduced by a large amount compared to the base case in some years (e.g., 2019), the ISS for the total length composition was similar to the base case.

Another notable exception to the ISS being reduced by less than 25% for sex sub-sampling cases of 100 - 175 fish, though not as extreme as eastern Bering Sea walleye pollock, was for some of the flatfish stocks in the eastern Bering Sea and Gulf of Alaska. All of the flatfish assessments performed by the AFSC are sex-specific (e.g., Spies et al. 2021) due to clear sexual dimorphism in growth. Further, some of these flatfish assessments use the sex-specific length frequency samples in addition to the age specimen collections (which include information on sex, weight and age for individual fish) to estimate growth internally and fit age composition data in the assessment model through conditional age-at-length methods (e.g., Turnock et al. 2017). Additionally sex-specific, length-based selectivity is often estimated in these assessments to account for differences in availability to the survey. Therefore estimates of spawning stock biomass are highly dependent on understanding the trends in female length frequency. In addition, recent research has indicated that growth of certain flatfish stocks may be regionally-specific and thus spatially-explicit stock assessments have been developed to account for different growth rates in different regions (McGilliard and Palsson 2017, Bryan and Palsson 2021). Thus, the reduction in the sex-specific ISS for length composition based on the sexed sub-sampling level may have a greater influence on the results of these assessments. The influence of sub-sampling for sex determination of length frequency data and how that relates to assessment output sensitivity to reduction in ISS, conditional age-at-length results, and spatially-explicit assessments remains unknown. We recommend that additional flatfish-specific analysis would be performed to determine the influence of sex sub-sampling and that caution would be used when reducing flatfish sex sampling on the AFSC bottom trawl surveys in some cases.

On April 8, 2022 the SWO working group met with a subset of AFSC assessment authors as well as representatives from the Groundfish Assessment Program to discuss the results of this work and the need for changes to be made on the AFSC bottom trawl surveys as it pertains to sex determination and the physical health of survey staff. The group of assessment authors noted that a reduction of 25% or less in the sex-specific length composition ISS was acceptable in order to reduce injuries of survey staff, particularly if the amount of age specimen data collected was unchanged and remained at historical levels. It was agreed that in future AFSC bottom trawl surveys the sampling limit for sex determination of length frequency collections would be set at 100, but that if authors of certain stock assessments (e.g., flatfish assessments) were concerned with that level that the survey would be flexible and consider stock-specific limits of sex determination that were larger than 100. This joint recommendation will be applied starting with the 2022 eastern Bering Sea and Aleutian Islands bottom trawl surveys, and the 2023 Gulf of Alaska bottom trawl survey.

ACKNOWLEDGMENTS

We thank representatives from the Groundfish Assessment Program (Duane Stevenson, Ned Leyman, and Stan Kotwicki), the Marine Ecology and Stock Assessment program, and the Status of Stocks and Multispecies Assessments program for their participation and collaboration in this work and providing recommendations to the AFSC bottom trawl surveys. We thank Stan Kotwicki and Carey McGilliard for their helpful reviews of this manuscript.

CITATIONS

Ailloud, L.E., and Hoenig, J.M. 2019. A general theory of age-length keys: Combining the forward and inverse keys to estimate age composition from incomplete data. ICES Journal of Marine Science 76(6): 1515–1523.

Bryan, M.D., and Palsson, W. 2021. Assessment of the northern and southern rock sole (*Lepidopsetta polyxstra* and *bilineata*) stocks in the Gulf of Alaska. In: Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. North Pacific Fishery Management Council, 1007 West 3rd Ave., Suite 400, L92 Building, 4th floor, Anchorage, AK 99501.

Gerritsen, H.D., and McGrath, D. 2007. Precision estimates and suggested sample sizes for length-frequency data. Fishery Bulletin, U.S. 105(1): 116–120.

Ianelli, J.I., Fissel, B., Stienessen, S., Honkalehto, T., Siddon, E., and Allen-Akselrud, C. 2021. Assessment of the walleye pollock stock in the eastern Bering Sea. In: Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea and Aleutian Islands. North Pacific Fishery Management Council, 1007 West 3rd Ave., Suite 400, L92 Building, 4th floor, Anchorage, AK 99501.

Lauth, R.R., J, D.E., and J, C. 2019. Results of the 2017 eastern and northern Bering Sea continental shelf bottom trawl survey of groundfish and invertebrate fauna. U.S. Department of Commerce. NOAA Technical Memorandum NMFS-AFSC-396, 260 p.

McAllister, M.K., and Ianelli, J.N. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Canadian Journal of Fisheries and Aquatic Sciences 54(2): 284–300.

McGilliard, C.R., and Palsson, W. 2017. Assessment of the rex sole stock in the Gulf of Alaska. In: Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. North Pacific Fishery Management Council, 1007 West 3rd Ave., Suite 400, L92 Building, 4th floor, Anchorage, AK 99501.

Miller, T.J., and Skalski, J.R. 2006. Integrating design- and model-based inference to estimate length and age composition in North Pacific longline catches. Canadian Journal of Fisheries and Aquatic Sciences 63(5): 1092–1114.

R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.R-project.org/.

Spies, I., Haehn, R., Siddon, E., Conner, J., Markowitz, E., Yeung, C., and Ianelli, J. 2021. Assessment of the yellowfin sole stock in the eastern Bering Sea and Aleutian Isalnds. In: Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea and Aleutian Isalands. North Pacific Fishery Management Council, 1007 West 3rd Ave., Suite 400, L92 Building, 4th floor, Anchorage, AK 99501.

Stauffer, G. 2004. NOAA Protocols for Groundfish Bottom Trawl Surveys of the Nation’s Fishery Resources. U.S. Department of Commerce. NOAA Technical Memorandum NMFS-F/SPO-65, 205 p.

Stewart, I.J., and Hamel, O.S. 2014. Bootstrapping of sample sizes for length-or age-composition data used in stock assessments. Canadian Journal of Fisheries and Aquatic Sciences 71(4): 581–588.

Thompson, G.G., Barbeaux, S., Connor, J., Fissel, B., Hurst, T., Laurel, B., O'Leary, C.A., Rogers, L., Shotwell, S.K., Siddon, E., Spies, I., Thorson, J.T., and Tyrell, A. 2021. Assessment of the Pacific cod stock in the eastern Bering Sea. In: Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea and Aleutian Isalands. North Pacific Fishery Management Council, 1007 West 3rd Ave., Suite 400, L92 Building, 4th floor, Anchorage, AK 99501.

Turnock, B.J., McGilliard, C.R., and Palsson, W. 2017. Assessment of the flathead sole stock in the Gulf of Alaska. In: Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. North Pacific Fishery Management Council, 1007 West 3rd Ave., Suite 400, L92 Building, 4th floor, Anchorage, AK 99501.

von Szalay, P.G., and Raring, N.W. 2018. Data Report: 2017 Gulf of Alaska bottom trawl survey. U.S. Department of Commerce. NOAA Technical Memorandum NMFS-AFSC-374: 260 p.

von Szalay, P.G., Raring, N.W., Rooper, C.N., and A, L.E. 2017. Data Report: 2016 Aleutian Islands bottom trawl survey. U.S. Department of Commerce. NOAA Technical Memorandum NMFS-AFSC-349: 161 p.

Table 1. -- Total length frequency samples from the most recent three Aleutian Islands surveys for the species evaluated in the bootstrap-simulation for reduction in sexed length-frequency collections.

Common name	Scientific name	2014	2016	2018	Average
arrowtooth flounder	*Atheresthes stomias*	10,212	9,166	10,225	9,868
Atka mackerel	*Pleurogrammus monopterygius*	10,162	6,080	7,423	7,888
flathead sole	*Hippoglossoides elassodon*	4,962	3,904	4,941	4,602
Kamchatka flounder	*Atheresthes evermanni*	3,188	3,768	3,623	3,526
northern rock sole	*Lepidopsetta polyxystra*	10,477	10,018	11,767	10,754
northern rockfish	*Sebastes polyspinis*	14,884	15,116	14,784	14,928
Pacific cod	*Gadus macrocephalus*	4,556	6,506	6,107	5,723
Pacific ocean perch	*Sebastes alutus*	30,205	36,277	30,992	32,491
walleye pollock	*Gadus chalcogrammus*	10,305	9,142	12,972	10,806

Table 2. -- Total length frequency samples from the most recent three eastern Bering Sea surveys for the species evaluated in the bootstrap-simulation for reduction in sexed length-frequency collections.

Common name	Scientific name	2017	2018	2019	Average
Alaska plaice	*Pleuronectes quadrituberculatus*	8,347	14,030	9,080	10,486
arrowtooth flounder	*Atheresthes stomias*	11,679	18,756	14,349	14,928
flathead sole	*Hippoglossoides elassodon*	19,238	26,837	20,993	22,356
Greenland turbot	*Reinhardtius hippoglossoides*	373	259	166	266
Kamchatka flounder	*Atheresthes evermanni*	2,654	2,743	1,966	2,454
northern rock sole	*Lepidopsetta polyxystra*	22,466	28,397	23,231	24,698
Pacific cod	*Gadus macrocephalus*	11,827	11,317	11,288	11,477
walleye pollock	*Gadus chalcogrammus*	49,037	56,799	42,795	49,544
yellowfin sole	*Limanda aspera*	26,973	31,584	25,768	28,108

Table 3. -- Total length frequency samples from the most recent three Gulf of Alaska surveys for the species evaluated in the bootstrap-simulation for reduction in sexed length-frequency collections.

Common name	Scientific name	2017	2019	2021	Average
arrowtooth flounder	*Atheresthes stomias*	35,219	38,862	36,444	36,842
flathead sole	*Hippoglossoides elassodon*	16,086	14,292	15,321	15,233
northern rock sole	*Lepidopsetta polyxystra*	4,776	3,318	2,960	3,685
northern rockfish	*Sebastes polyspinis*	2,570	2,237	2,088	2,298
Pacific cod	*Gadus macrocephalus*	2,601	3,755	3,999	3,452
Pacific ocean perch	*Sebastes alutus*	20,710	26,445	22,802	23,319
rex sole	*Glyptocephalus zachirus*	10,374	13,205	15,054	12,878
sablefish	*Anoplopoma fimbria*	4,599	8,847	9,173	7,540
shortspine thornyhead	*Sebastolobus alascanus*	10,376	10,242	10,269	10,296
southern rock sole	*Lepidopsetta bilineata*	5,986	6,924	8,661	7,190
walleye pollock	*Gadus chalcogrammus*	12,790	15,100	22,427	16,772

Table 4. -- Range of average reduction in sexed length frequency samples for the sex subsampling cases of 100 - 175 fish sampled for sex determination from the most recent three surveys.

Common name	AI	EBS	GOA
Alaska plaice	–	101 - 708	–
arrowtooth flounder	189 - 1,237	361 - 2,048	1,661 - 8,325
Atka mackerel	149 - 1,701	–	–
flathead sole	72 - 612	523 - 3,541	219 - 2,211
Kamchatka flounder	–	0 - 27	–
northern rock sole	–	792 - 4,399	65 - 352
northern rockfish	719 - 4,765	–	20 - 480
Pacific cod	122 - 480	121 - 818	19 - 140
Pacific ocean perch	2,042 - 11,801	–	888 - 7,339
rex sole	–	–	39 - 786
sablefish	–	–	56 - 779
shortspine thornyhead	0 - 221	–	23 - 119
southern rock sole	–	–	14 - 417
walleye pollock	573 - 2,474	2,882 - 14,479	1,885 - 4,659
yellowfin sole	–	1,024 - 6,337	–
Total	3,866 - 23,291	5,804 - 32,357	4,889 - 25,694

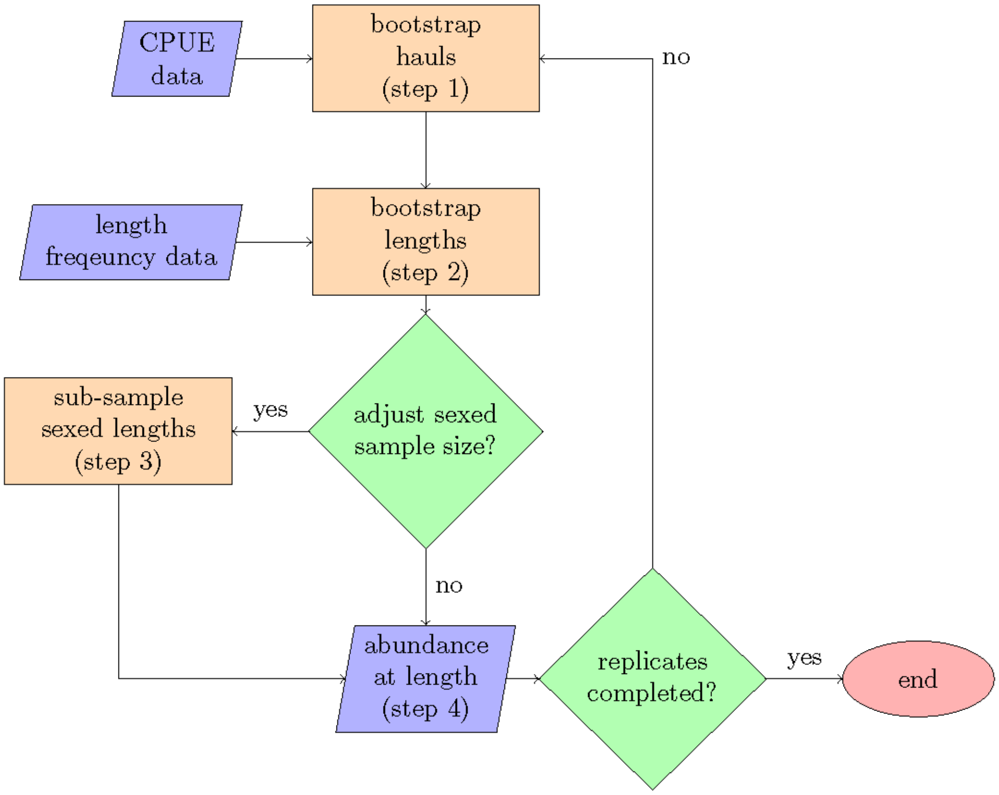

Figure 1. -- Bootstrap-simulation flow chart, the steps refer to the order of operations.

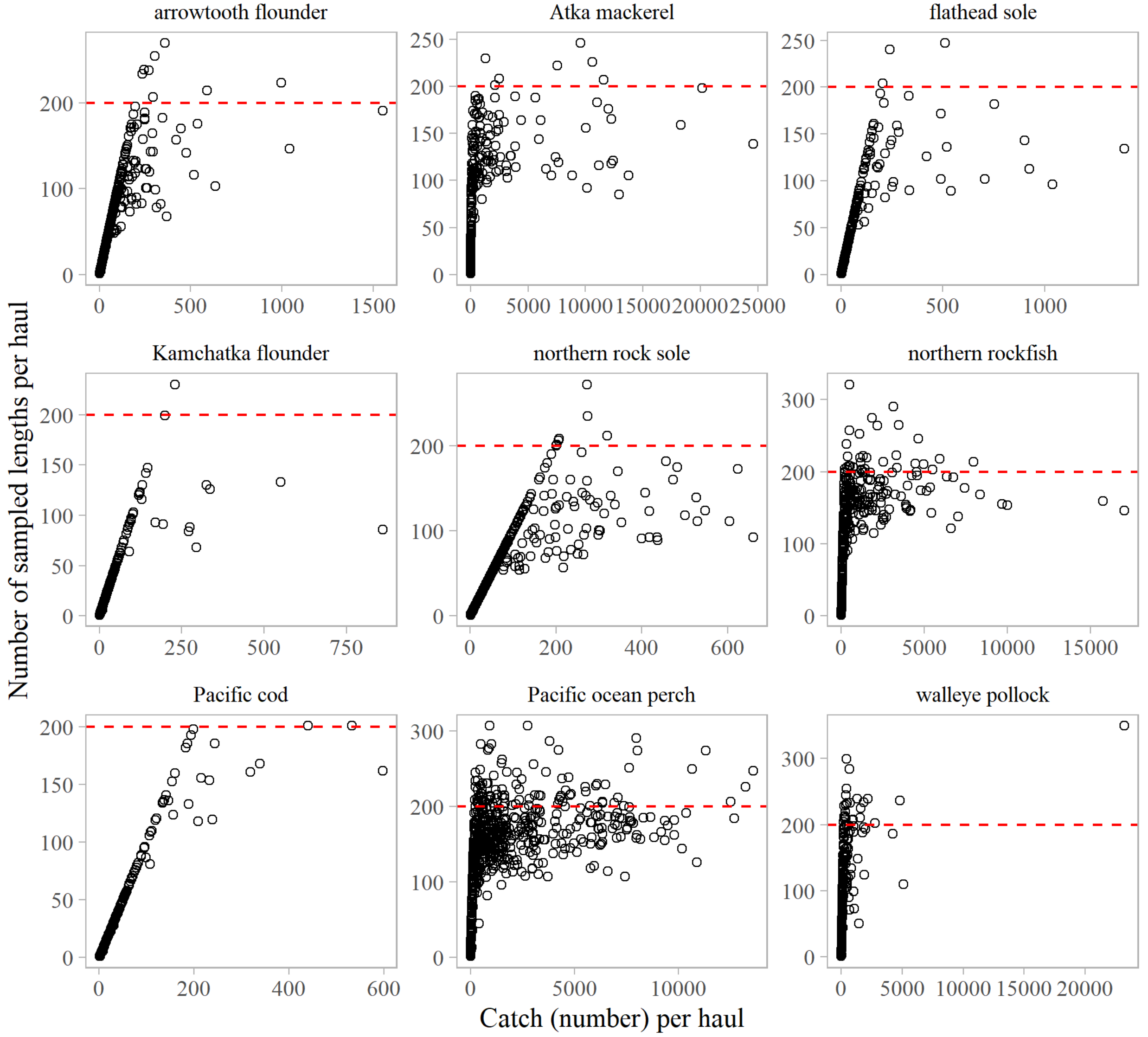

Figure 2. -- Number of sexed length samples collected per haul compared to catch per haul for the most recent three years of the Aleutian Islands bottom trawl survey. The red dashed-line represents the historical minimum sample size of 200.

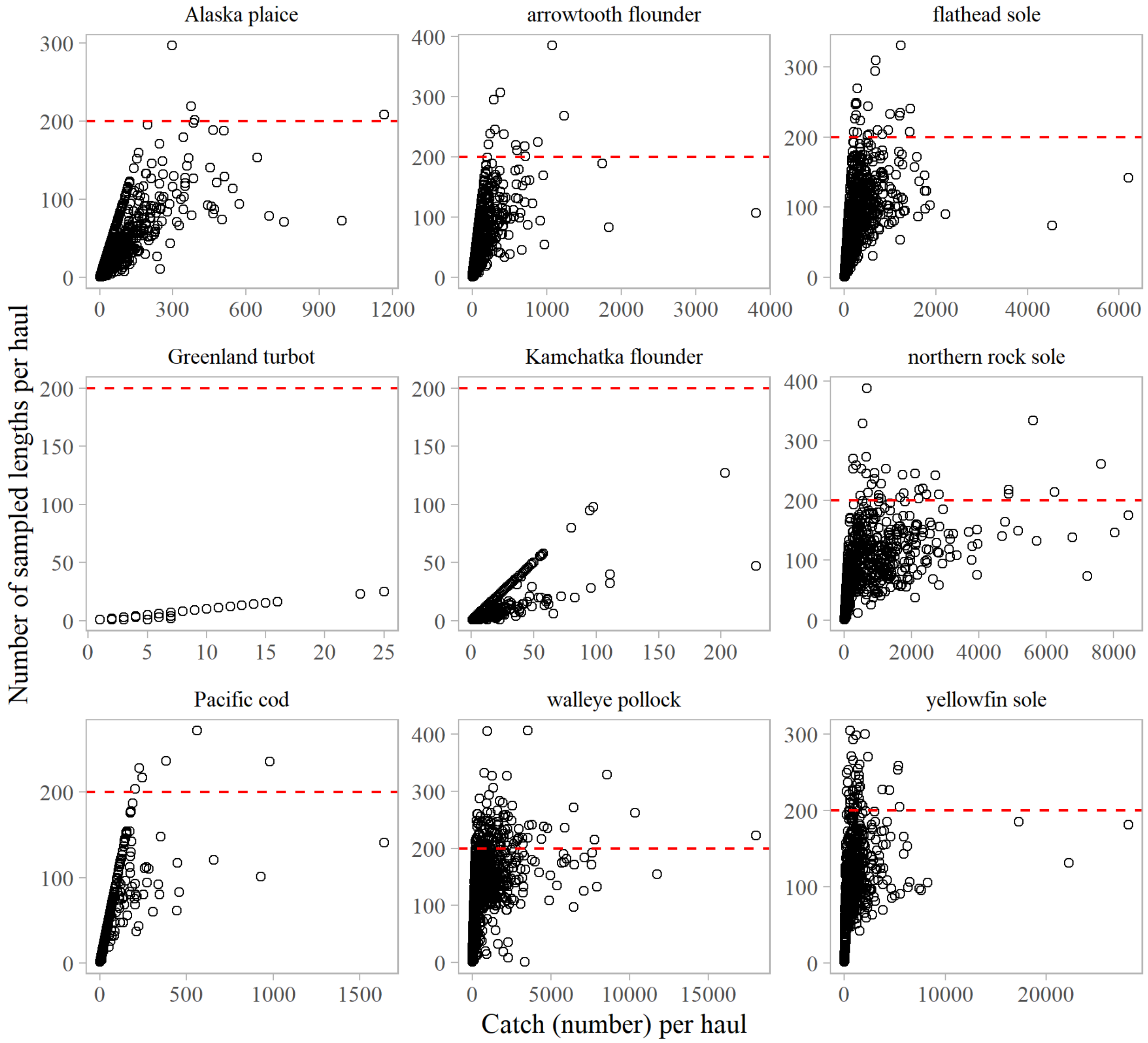

Figure 3. -- Number of sexed length samples collected per haul compared to catch per haul for the most recent three years of the eastern Bering Sea bottom trawl survey. The red dashed-line represents the historical minimum sample size of 200.

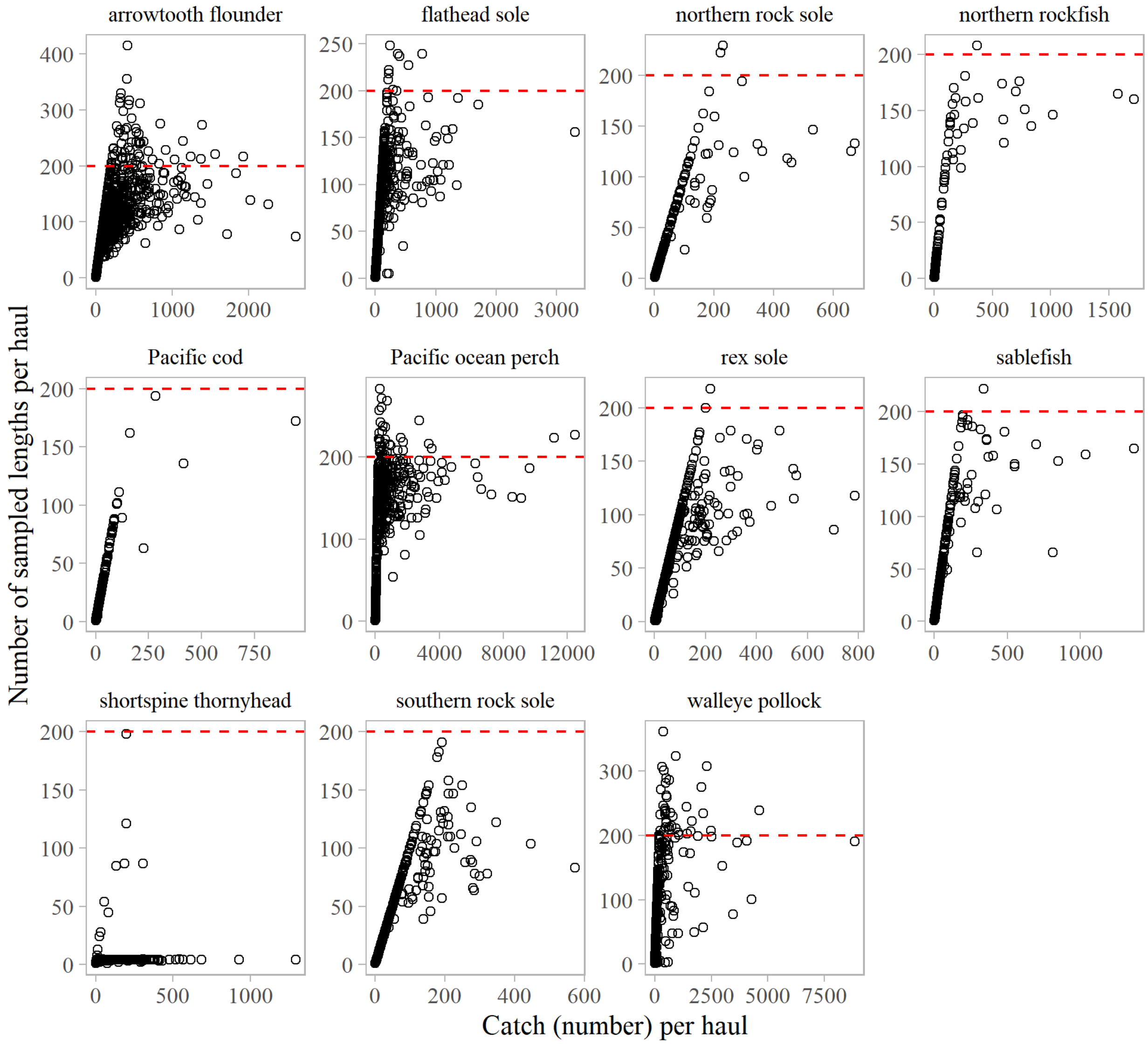

Figure 4. -- Number of sexed length samples collected per haul compared to catch per haul for the most recent three years of the Gulf of Alaska bottom trawl survey. The red dashed-line represents the historical minimum sample size of 200.

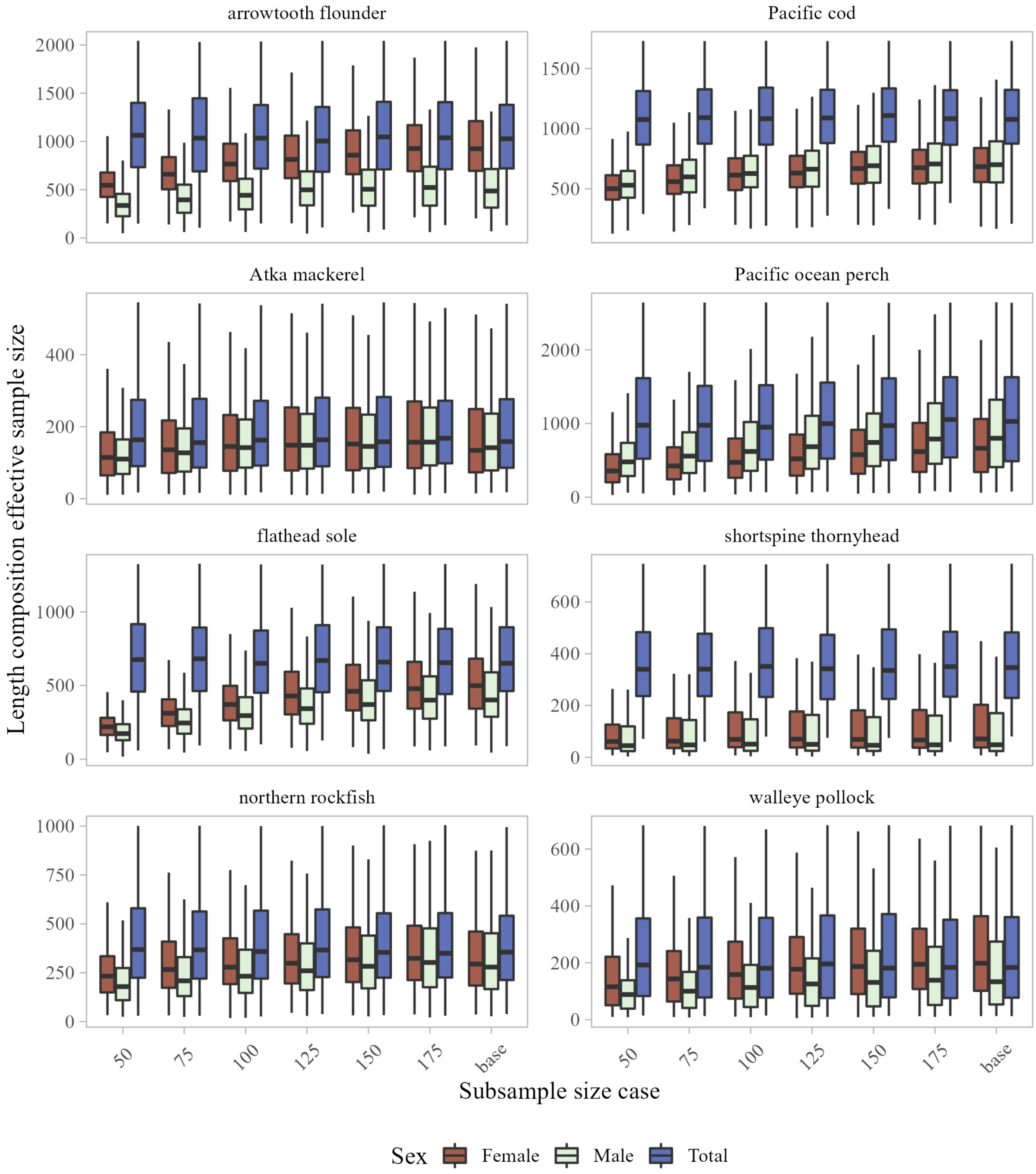

Figure 5. -- Replicated length composition effective sample size for females, males, and total (sum of females, males, and unsexed) for selected species in the Aleutian Islands for each sexed sub-sample size case (including the base case where no sub-sampling for sexing occurred).

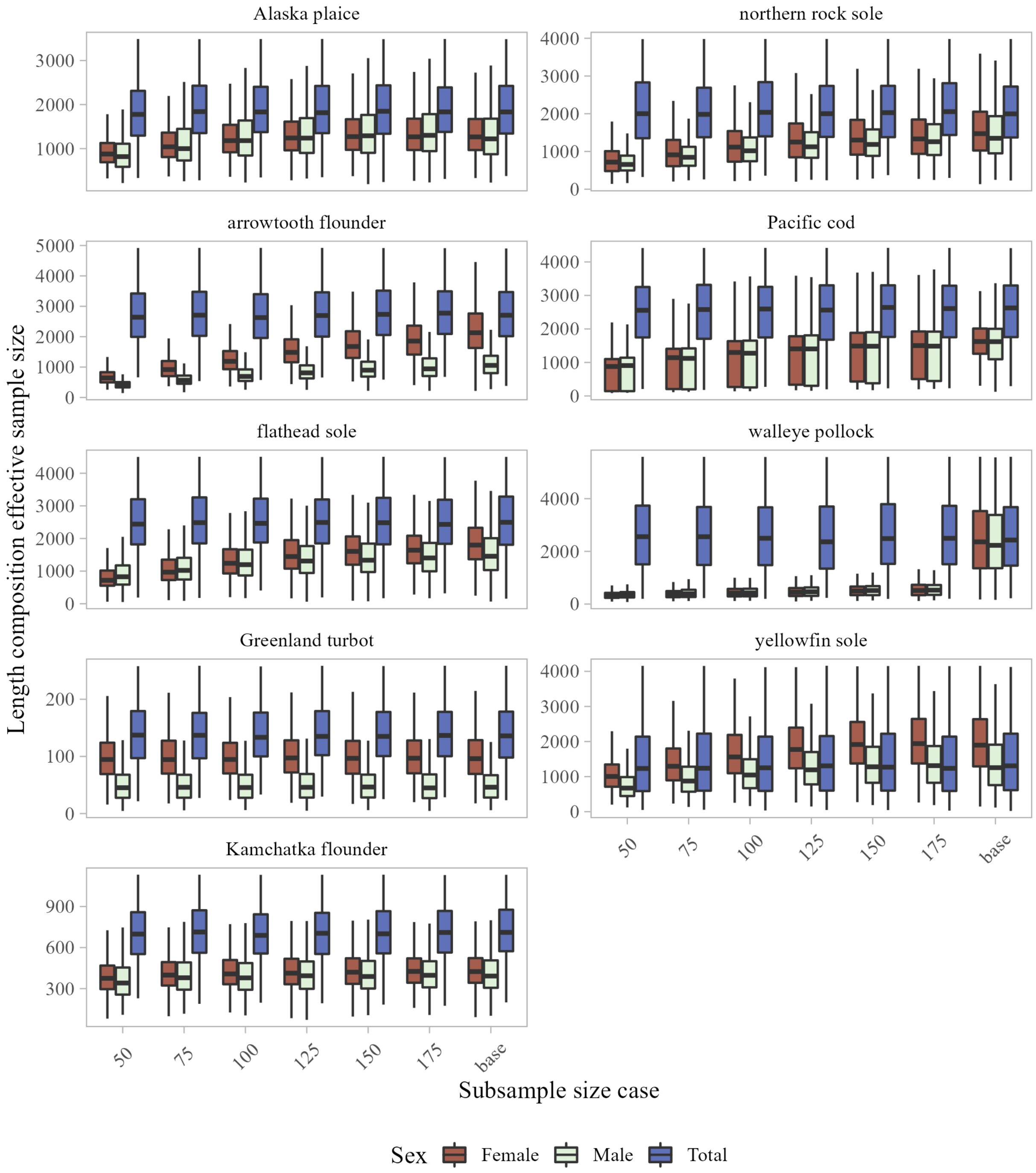

Figure 6. -- Replicated length composition effective sample size for females, males, and total (sum of females, males, and unsexed) for selected species in the eastern Bering Sea for each sexed sub-sample size case (including the base case where no sub-sampling for sexing occurred).

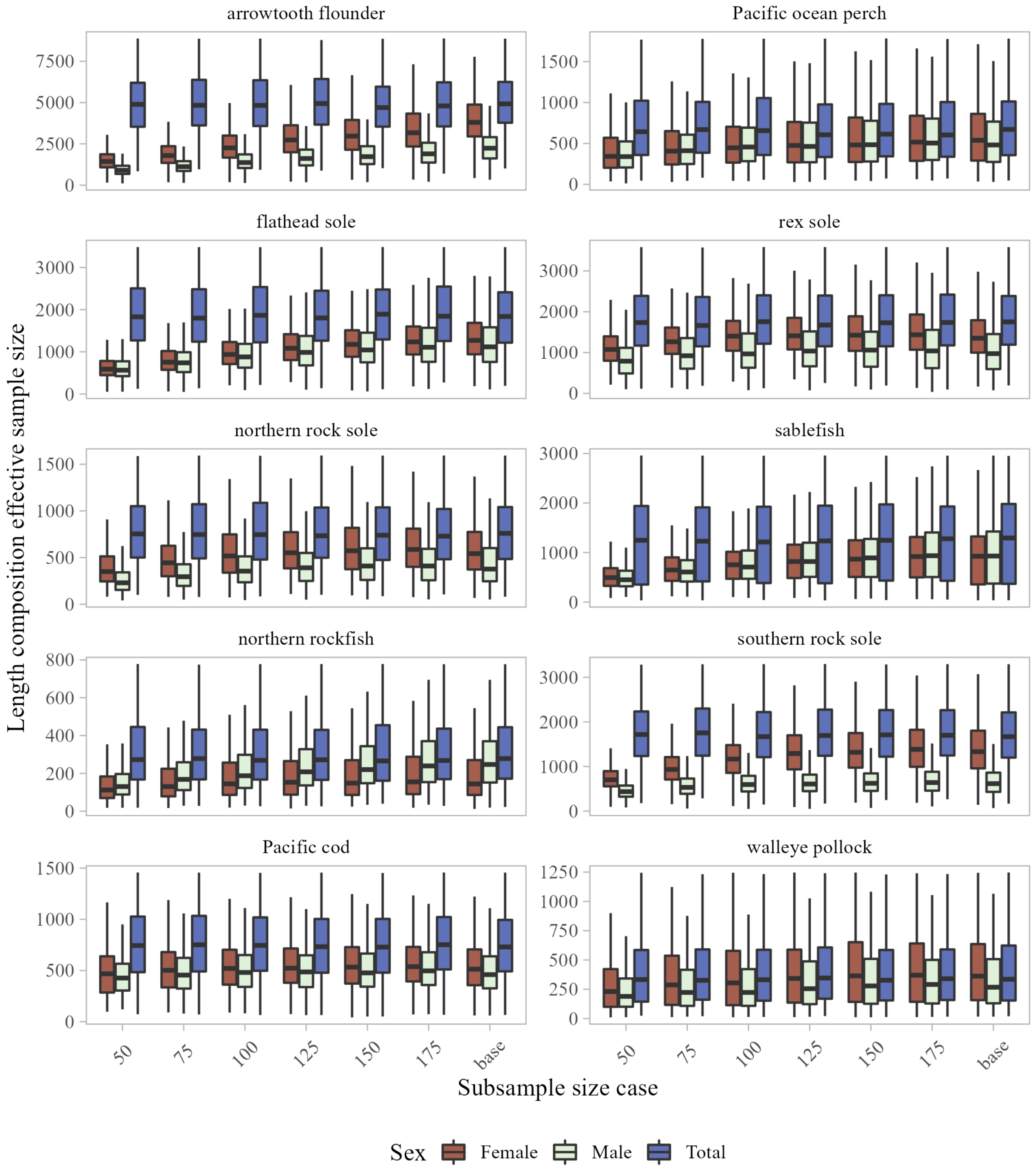

Figure 7. -- Replicated length composition effective sample size for females, males, and total (sum of females, males, and unsexed) for selected species in the Gulf of Alaska for each sexed sub-sample size case (including the base case where no sub-sampling for sexing occurred).

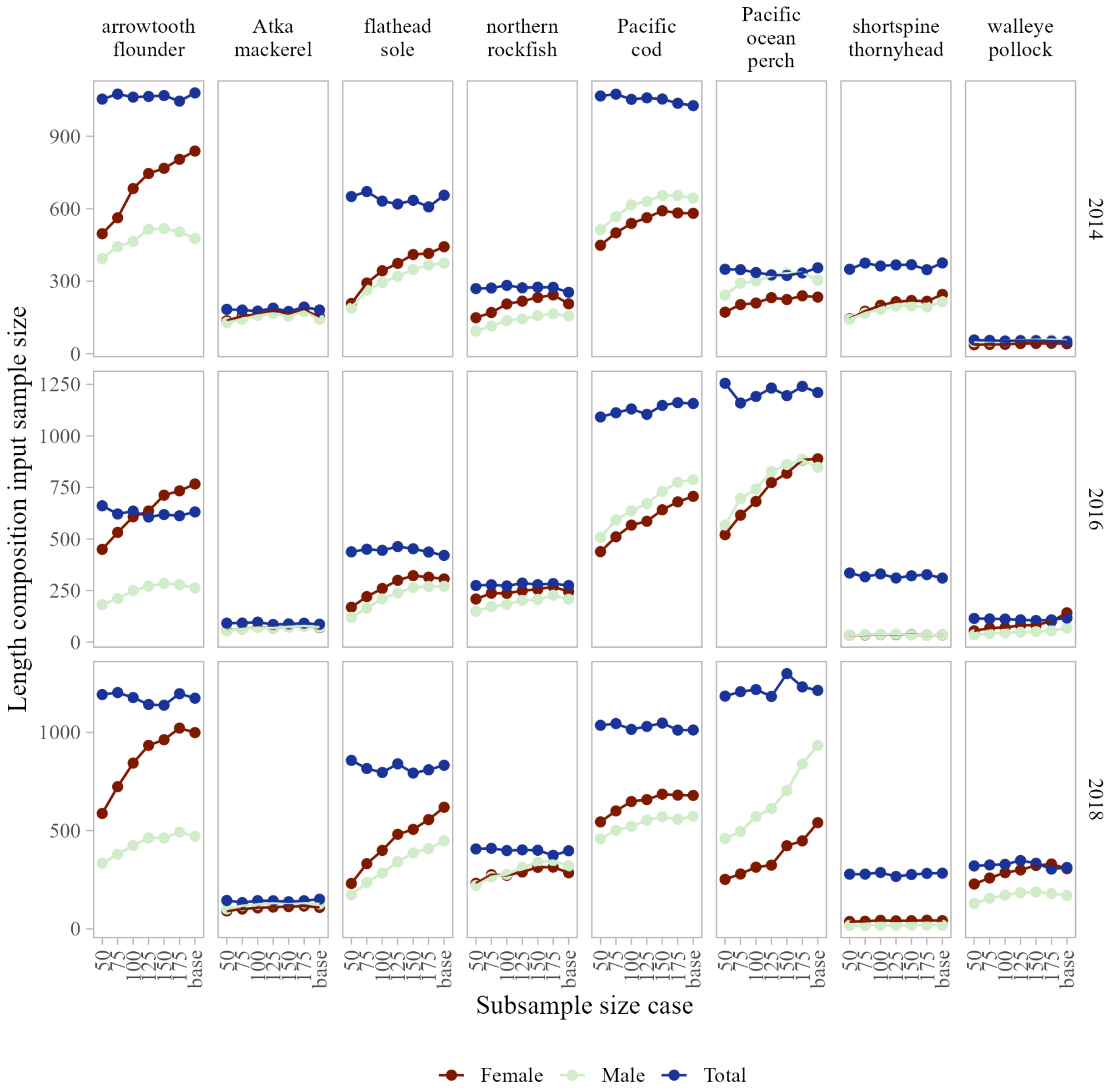

Figure 8. -- Length composition input sample size for females, males, and total (sum of female, male, and unsexed) of selected species in the Aleutian Islands for each sexed sub-sample size case (including the base case where no sub-sampling for sexing occurred).

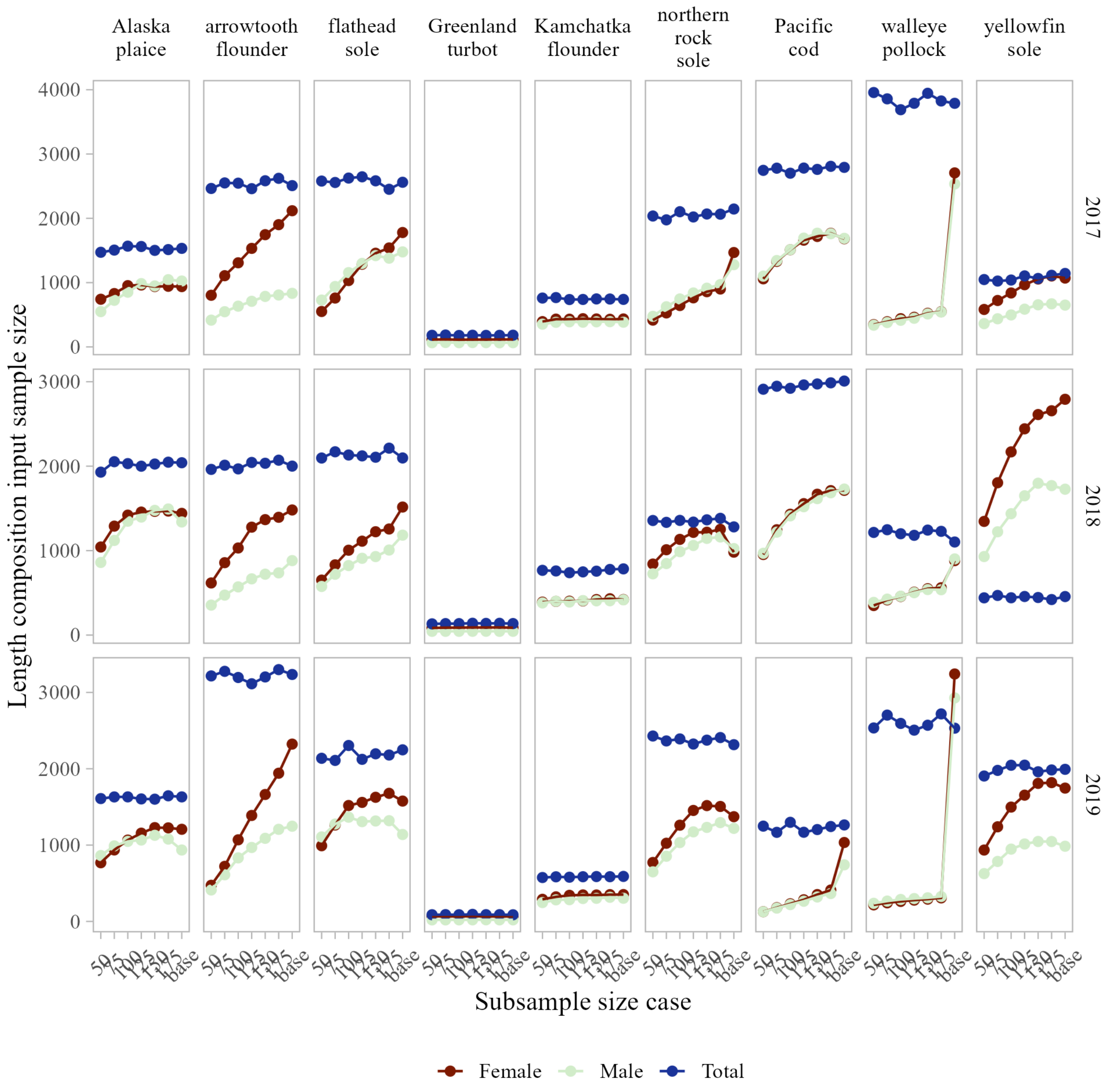

Figure 9. -- Length composition input sample size for females, males, and total (sum of female, male, and unsexed) of selected species in the eastern Bering Sea for each sexed sub-sample size case (including the base case where no sub-sampling for sexing occurred).

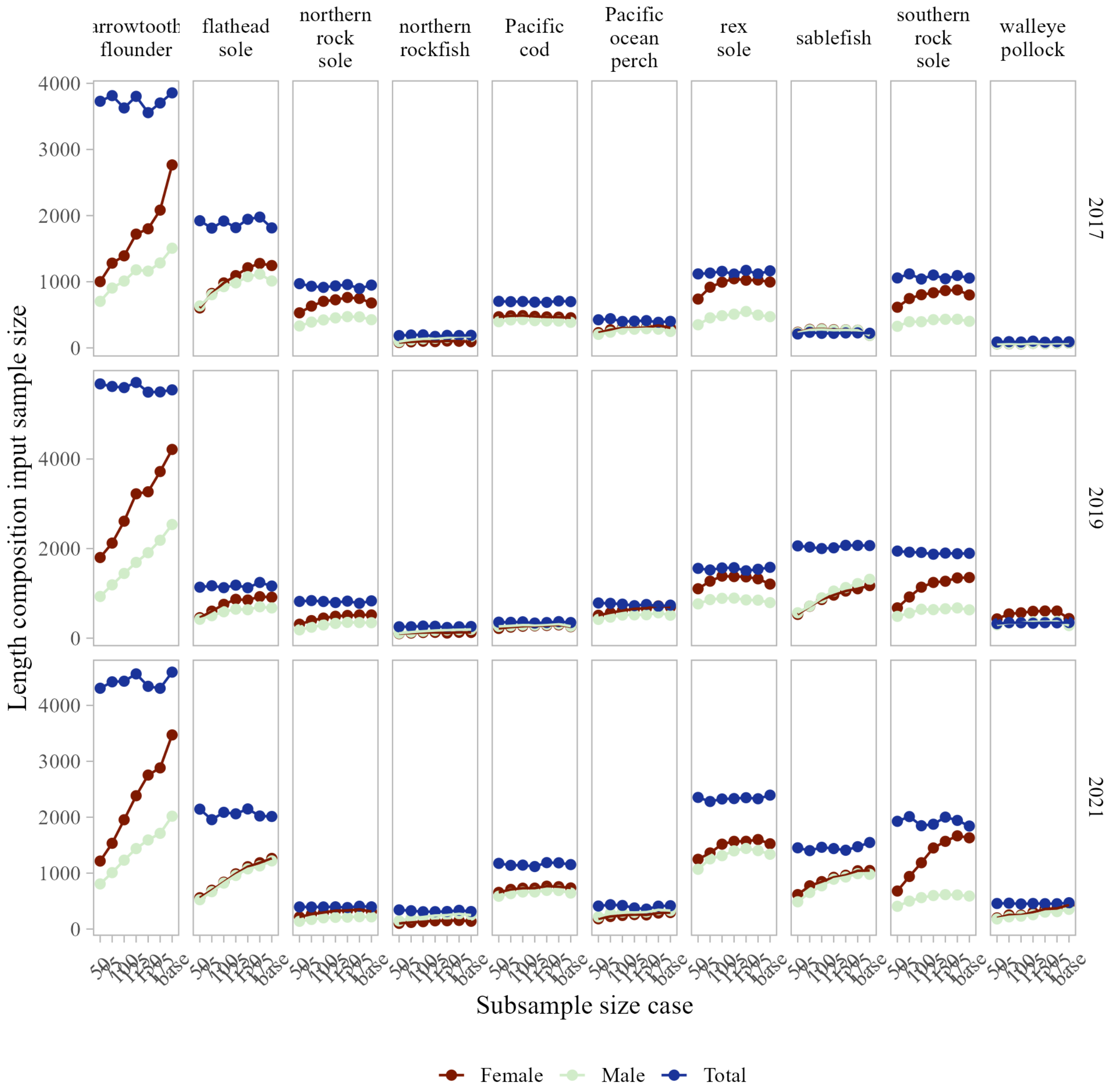

Figure 10. -- Length composition input sample size for females, males, and total (sum of female, male, and unsexed) of selected species in the Gulf of Alaska for each sexed sub-sample size case (including the base case where no sub-sampling for sexing occurred).

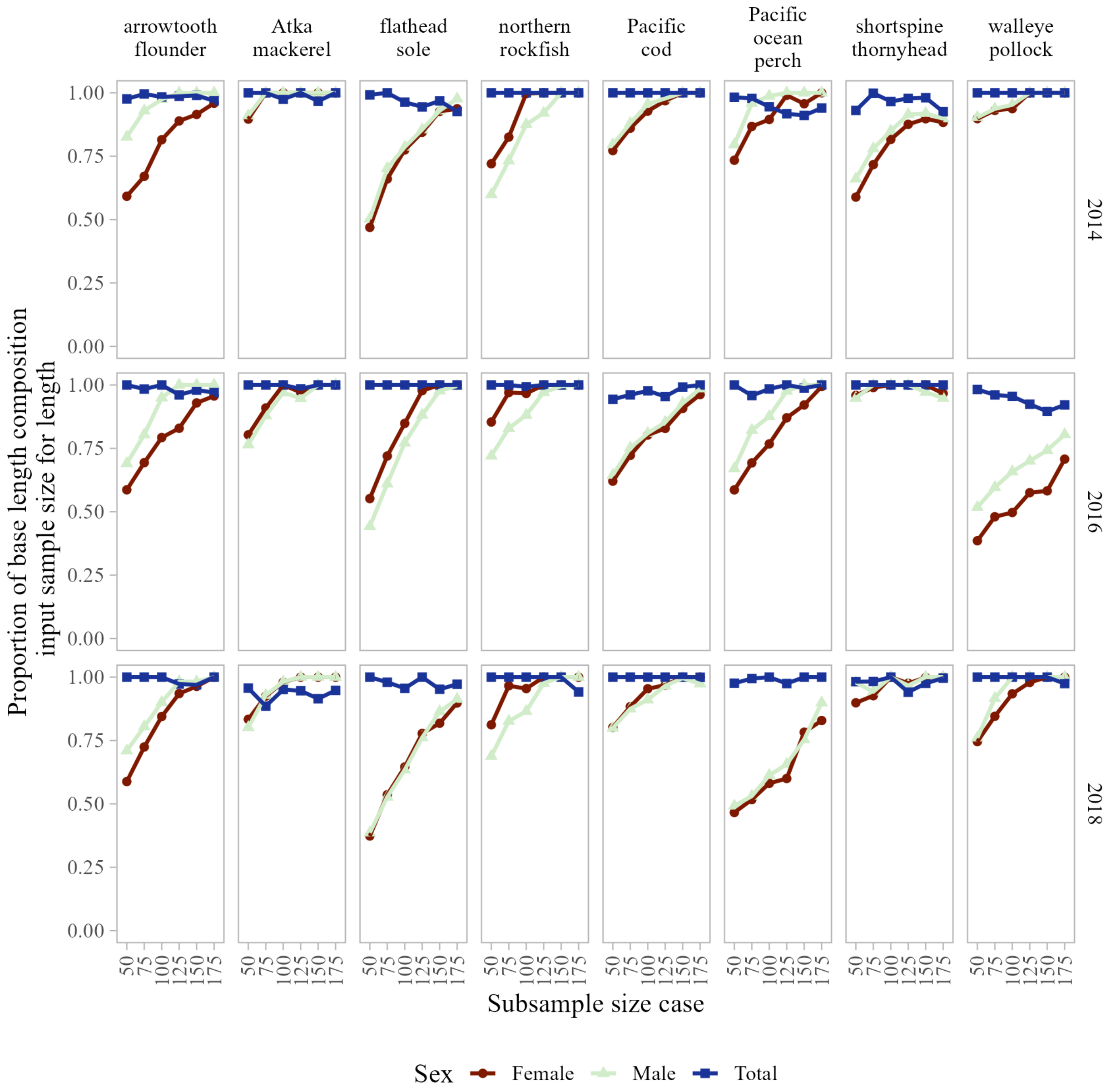

Figure 11. -- Proportion of base length composition input sample size for each of the sexed subsample size cases investigated for selected species in the Aleutian Islands.

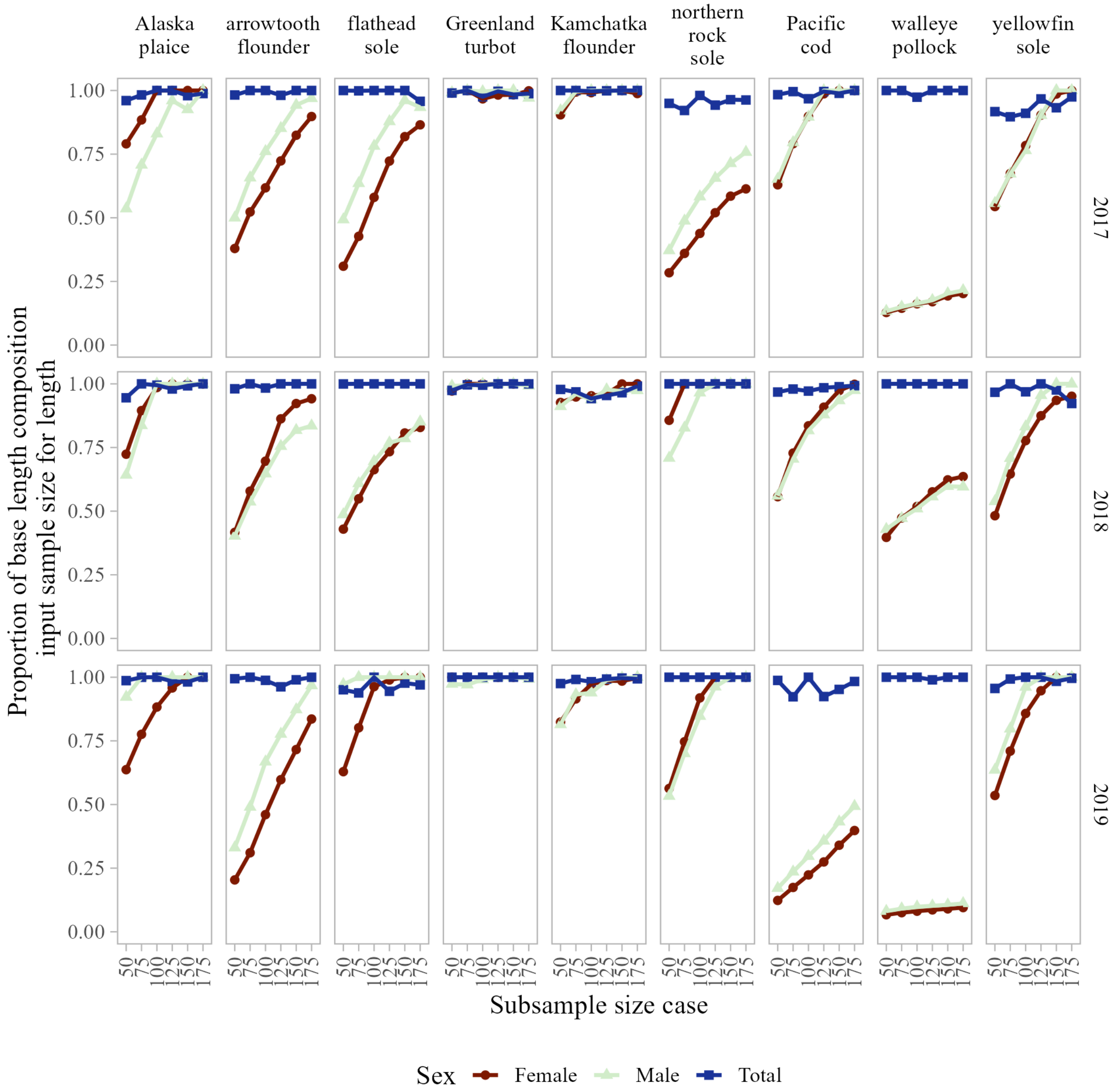

Figure 12. -- Proportion of base length composition input sample size for each of the sexed sub-sample size cases investigated for selected species in the eastern Bering Sea.

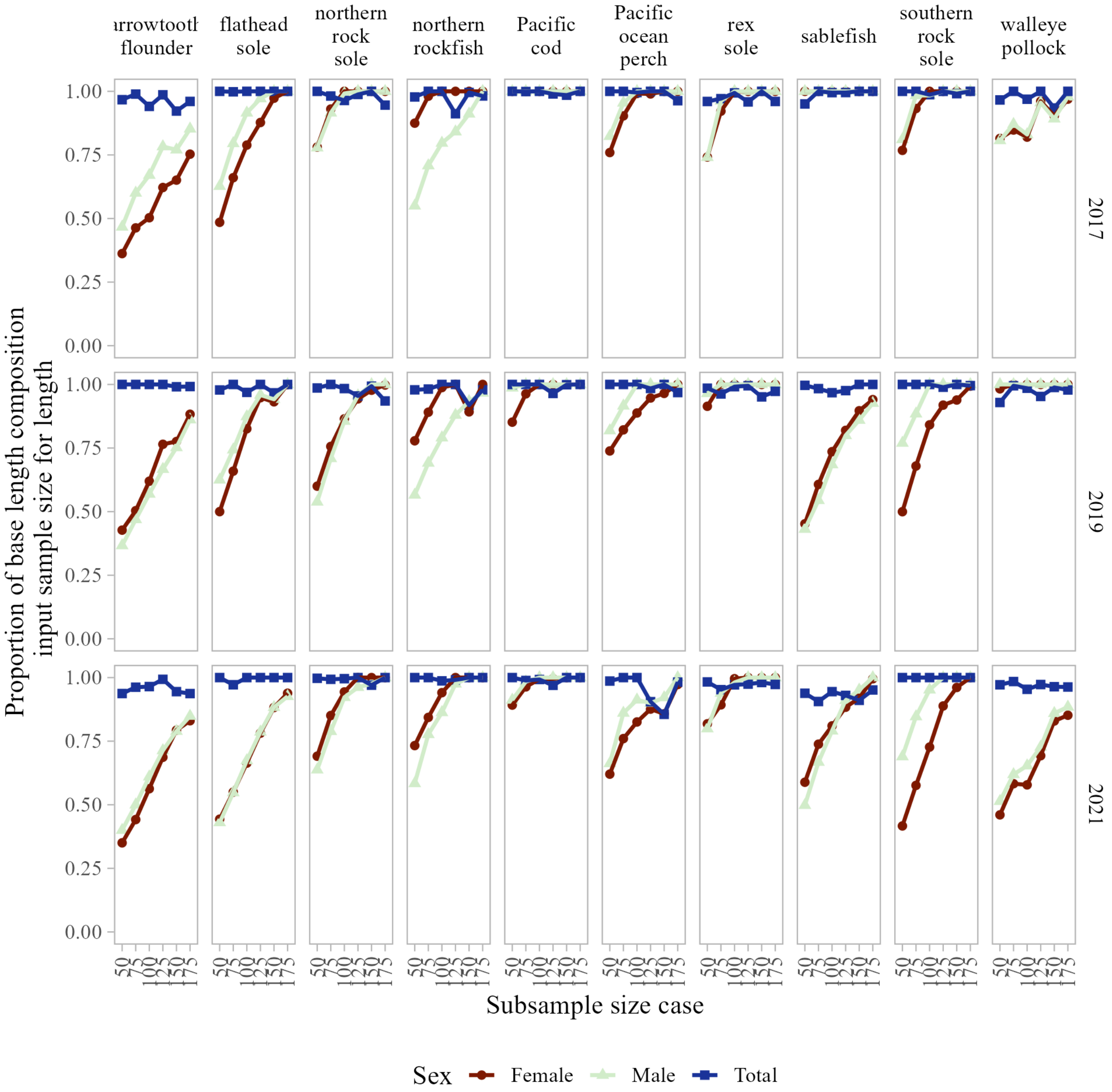

Figure 13. -- Proportion of base length composition input sample size for each of the sexed sub-sample size cases investigated for selected species in the Gulf of Alaska.

U.S. Secretary of Commerce
Gina M. Raimondo

Under Secretary of Commerce for
Oceans and Atmosphere
Dr. Richard W. Spinrad

Assistant Administrator, National Marine
Fisheries Service.
Janet Coit

February 2023

www.nmfs.noaa.gov

OFFICIAL BUSINESS

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