Fisheries Research 270 (2024) 106894

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Inclusion of ageing error and growth variability using a bootstrap estimation of age composition and conditional age-at-length input sample size for fisheries stock assessment models

Peter-John F. Hulson [a,*], Benjamin C. Williams [a]

[a] Auke Bay Laboratories, Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 17109 Point Lena Loop Rd., Juneau, AK 99801, USA

A R T I C L E I N F O

Handled by: A.E. Punt

Keywords:
Stock assessment
Aging error
Growth variability
Maximum likelihood
Uncertainty
Input sample size

A B S T R A C T

Statistical catch-at-age assessment models used for fisheries management integrate multiple sources of information that are statistically weighted in a joint likelihood framework; the relative statistical weighting between these sources of information is an important, yet often subjective aspect of stock assessment. Input sample size (ISS) is a quantity that is used to statistically weight composition data in these types of models. Both design-based bootstrap and model-based estimators have been proposed, however, these methods to determine ISS do not explicitly account for uncertainty from ageing error and growth variability that are inherent to expanded age composition and conditional age-at-length data. In this study, we evaluate the impact of including ageing error and growth variability within bootstrap methods that estimate age composition and conditional age-at-length ISS. We find that for all the stocks evaluated the ISS determined from bootstrap methods decreased as these addition sources of uncertainty were included. The decrease in ISS was species type specific, but generally decreased up to 40% when ageing error was introduced, up to 50% when growth variability was included, and up to 60% when both sources of uncertainty were included. These results indicate that there is more variability within age composition or conditional age-at-length data than would be accounted for with ISS estimates that do not include these sources of uncertainty. The method and results provided here allow for assessment scientists to statistically weight age composition and conditional age-at-length with ISS that takes into account ageing error and growth variability from either fishery-independent or fishery-dependent sources. Including these sources of uncertainty improves bootstrap estimates of ISS to capture all the sources of variability in age composition and conditional age-at-length and will subsequently improve stock assessment model quality.

1. Introduction

Compositional information on age and length are critical data products used in statistical catch-at-age assessment (SCAA) models as they facilitate the tracking of year classes and size-structure over time to facilitate our understanding of a fish stock's population dynamics (Quinn and Deriso, 1999), including the size and age based mortality processes through the selectivity of the fisheries. The two primary sources for age and length composition data used in SCAA models are fishery-independent and fishery-dependent. Fishery-independent sources typically include randomized and standardized collection of samples from hauls placed across space in a non-targeted framework. Fishery-dependent sources are based upon collection of age and length samples, randomized at some level, but obtained from hauls or trips targeting a specific species or species group. A common challenge in using compositional information in SCAA models to estimate population processes is the statistical weighting in the joint likelihood, as the statistical weighting affects the performance of the model. Due to the strong influence that compositional data can have in SCAA models, the statistical weight assigned to these data products are important for providing accurate advice for management (e.g., Hulson et al., 2012, Xu et al., 2020).

Regardless of the source of composition data (whether fishery-independent for fishery-dependent), it is commonly accepted that overdispersion of the data is inherent due to intra-haul correlation (e.g., Pennington and Volstad, 1994, Pennington et al., 2000). The concept of

* Corresponding author.
E-mail address: pete.hulson@noaa.gov (P.-J.F. Hulson).

https://doi.org/10.1016/j.fishres.2023.106894
Received 20 June 2023; Received in revised form 20 October 2023; Accepted 21 October 2023
Available online 3 November 2023
0165-7836/Published by Elsevier B.V.

effective sample size (ESS; introduced by McAllister and Ianelli, 1997), a reduced sample size from the actual number of fish measured or aged to account for this overdispersion, can be implemented within the likelihood function to weight the age or length composition data. The statistical weight assigned to annual composition data can follow a myriad of methods (e.g., fixed values as in Monnahan et al., 2021, number of samples or tows sampled upon as in Hulson et al., 2021 or Spencer and Ianelli., 2022, bootstrapping compositions as in Stewart and Hamel., 2014). The primary consideration when assigning a statistical weight to composition data is to account for the potential variability and correlation in the sampling process that result in overdispersion.

Throughout the development and implementation of the ESS concept in SCAA models a variety of terms have been used, often having multiple meanings for the same term. Typically, ESS is a term that has been used to denote the sample size used in statistical weighting of age or length composition data (e.g., Hulson et al., 2012, Punt et al., 2021), it has also be used to denote the performance of a SCAA estimates of composition data compared to the observed data (e.g., Thorson and Haltuch, 2019). Input sample size (ISS) has also been used as a term to denote the sample size used in statistical weighting of age or length composition data (e.g., Thorson and Haltuch, 2019, Thorson et al., 2023). In addition, realized sample size is a term introduced when using bootstrap methodologies (Stewart and Hamel, 2014). In order to provide consistency in the literature we propose the following usage of terms as it relates to this issue:

- Nominal sample size: the actual sample size obtained for age or length composition data from fishery-independent or fishery-dependent sources.
- Input sample size: the reduced sample size that accounts for overdispersion of age or length composition data used to statistically weight the composition data in SCAA models.
- Effective sample size: the statistic used to measure the difference in fit between SCAA model estimates of age or length composition data and the observed composition data.
- Realized sample size: the sample size that measures the difference between bootstrap estimates of age or length composition and the observed composition for a given bootstrap iteration.

Much of this terminology follows from Thorson et al. (2023) and we reiterate and expand upon it here in an attempt to convince researchers to adhere to a uniform set of terms across the fisheries literature when studying age and length composition data used in SCAA models.

When age is capable of being determined from otoliths, there is further variability in age composition data due to the ageing of the otolith, often called ‘ageing error’ (e.g., Punt et al., 2008). There are a number of factors that can influence the magnitude of ageing error, for example, the number of age classes or the sample size (Nesslage et al., 2022), but inherent to obtaining ages from otoliths is variability in the age readings across the laboratory age readers. To account for this source of variability, ageing laboratories regularly evaluate precision through obtaining multiple readings of the same otolith across different age readers (Morison et al., 2005). Several methods have been developed to account for ageing error in SCAA models when fitting age composition by integrating an ageing error matrix as an additional input data source for the model (Punt et al., 2008; Candy et al., 2012). The ageing error matrix is used to ‘correct’ the numbers-at-age estimated by the assessment model by assigning a certain proportion of fish in a given age class to adjacent age-classes based on the magnitude of the ageing error within the specific age-class. Since the development and implementation of ageing error matrices studies have been devoted to quantifying the effects of ageing error on assessment model estimates (e.g., Liao et al., 2013). Within each of these studies, and in each application of an ageing error matrix within a SCAA model, the age composition data will be statistically weighted by an ISS. As described previously, the ISS selected to statistically weight the age composition data should reflect the variability in the sampling process, thus, it should also include the variability in the age readings themselves.

In the process of obtaining an observed age composition, it may be the case that an age-length key (ALK) is employed to expand the estimated population numbers-at-length to population numbers-at-age (Quinn and Deriso, 1999; Ailloud and Hoenig, 2019). For either fishery-dependent or fishery-independent sources of age composition, if an expansion process is used to obtain an observed age composition it is generally the case that length frequency is expanded to some geographic area by weighting haul-level length frequency by haul-level catch-per-unit-effort (in numbers), this provides an estimated population numbers-at-length. Then, an ALK is constructed with age-length paired data and multiplied by the estimated population numbers-at-length to obtain estimated population numbers-at-age, often referred to as expanded age composition data (Siskey et al., 2023). An intrinsic component to the ALK is the variability in length for a given age. This variability in growth, or the range in lengths that are observed for a given age, is directly linked to the variability in the expanded age composition, and thus, should be reflected in the ISS selected to statistically weight the age composition data within a SCAA model.

To date, no method has been developed to integrate the variability in the ageing process when reading otoliths and in the growth process upon which age-length keys are based within the estimation of ISS used to statistically weight age composition data in SCAA models. The bootstrap method developed by Stewart and Hamel (2014) allows for resampling techniques to be employed at each level of the sampling design, and provides an objective avenue to estimate ISS that is based on the observation variability contained within the sampling process. In this study, we extend the methods of Stewart and Hamel (2014) to estimate age composition ISS that includes both ageing error and growth variability in the estimation process. We show, in a step-wise process, the added variability in age composition sample size from including ageing error and growth variability across a number of species that reflect differing life histories and levels of ageing difficulty.

2. Material and methods

2.1. Data

We used historical data collected from bottom trawl surveys conducted by the Alaska Fisheries Science Center (AFSC) in the Eastern Bering Sea (EBS: Lauth et al., 2019), Aleutian Islands (AI: von Szalay et al., 2017), and Gulf of Alaska (GOA: von Szalay and Raring, 2018). Within the AFSC bottom trawl surveys both length frequency data and age specimen data are collected, in addition to other survey data (e.g., catch, effort, location). Generally, a subsample of fish from each haul were processed at sea to collect their sex, length, and weight. A subsample of these fish have their sagittal otoliths collected; these otoliths were sent (with haul and specimen data) to the AFSC Age and Growth laboratory for age determination. Periodically, a subset of aged otoliths are selected for reader-tester agreement tests. These tests are used to evaluate the reproducibility of an age reading when two different readers age the same fish without knowledge of the other reader’s age determination of the otolith (Kimura and Lyons, 1991). The average annual bottom trawl survey age sample sizes by region, and the total number of otoliths used for reader-tester agreement tests are shown in Table 1 for the species evaluated. For the EBS the survey years included in this analysis were 1982 – 2022, for the GOA were 1990 – 2021, and for the AI were 1991 – 2018. As a point of clarification, we use the term ‘stock’ to identify a certain species within a distinct region, for example, the EBS walleye pollock stock. The stocks selected for this analysis all have greater than 5000 reader-tester paired otolith readings (Table 2) and are all assessed using integrated SCAA models that require input sample sizes for the age composition data.

Table 1
Average annual age samples from the AFSC bottom trawl surveys by region (rounded to the nearest 10), and total reader-tester age pairs (rounded to the nearest 100) for the stocks evaluated in the bootstrap-simulation.

Stock (species type)	Scientific name	AI	EBS	GOA	R-T
arrowtooth flounder (flatfish)	*Atheresthes stomias*	450	480	850	6100
flathead sole (flatfish)	*Hippoglossoides elassodon*	–	560	520	9400
northern rock sole (flatfish)	*Lepidopsetta polyxystra*	–	460	450	8900
northern rockfish (rockfish)	*Sebastes polyspinis*	570	–	450	6400
Pacific cod (gadid)	*Gadus macrocephalus*	800	1070	650	21,200
Pacific ocean perch (rockfish)	*Sebastes alutus*	940	–	1030	13,500
walleye pollock (gadid)	*Gadus chalcogrammus*	790	1500	1300	84,400
yellowfin sole (flatfish)	*Limanda aspera*	–	750	–	10,300

Table 2
Description and notation for bootstrap-simulation evaluations.

Uncertainty scenarios	
Base	Standard bootstrap-simulation (omitting steps 5 and 6 that include ageing error and growth variability in the Bootstrap-Simulation framework)
AE	Bootstrap-simulation including ageing error only
GV	Bootstrap-simulation including growth variability only
AE & GV	Bootstrap-simulation including both ageing error and growth variability
Treatments	
Growth data treatment	Resample lengths for a given age after pooling age-length data across survey years ('Pooled') or using annual age-length data ('Annual')
Length bin treatment	Implement 1 cm, 2 cm, and 5 cm length bins in the length data
Aggregation treatment	Aggregate length and age data before ('Pre-expansion') or after ('Post-expansion') length and age expansion

2.2. Length and age composition expansion

Details of how the length frequency and age collections are expanded to population abundance-at-length and -age then subsequently used as compositional data in stock assessment models at AFSC are provided in Hulson et al. (2023). Here we generalize these methods to provide the reader with a broad understanding of how length and age composition are expanded in the AFSC bottom trawl surveys.

Length frequency samples collected by the AFSC bottom trawl surveys are expanded by catch and stratum area to obtain estimates of population abundance-at-length. This is often referred to as the 'first stage expansion' and is a common method to obtain population abundance estimates at length from area-swept survey data (e.g., Miller and Skalski, 2006, Ailloud and Hoenig, 2019). To expand the species-specific length frequency samples to population-at-length we first compute the overall population numbers within a stratum by multiplying the average catch per unit effort within the strata (i.e., the number of fish per square kilometer averaged across the hauls performed within the strata) by the area of the strata (in square kilometers). The overall population numbers in year-y within stratum-s ($\widehat{N}_{s,y}$) is computed with

$$\widehat{N}_{s,y} = \overline{CPUE}_{s,y} \cdot A_s \tag{1}$$

where A_s is the area of stratum-s (in km^2), and $\overline{CPUE}_{s,y}$ is the species-specific average catch per unit effort of numbers captured across the hauls within a strata in year-y We then compute the relative catch per unit effort for each haul performed within the strata and the sex-specific relative length composition for each haul. The relative catch per unit effort for each haul ($\widehat{c}_{h,s,y}$) is computed by

$$\widehat{c}_{h,s,y} = \frac{CPUE_{h,s,y}}{\sum_{h=1}^{H_{s,y}} CPUE_{h,s,y}} \tag{2}$$

where $CPUE_{h,s,y}$ is the catch per unit effort of numbers caught within a haul-h for stratum-s in year-y. The sex-specific relative length composition for each haul ($\widehat{p}_{x,l,h,s,y}$) is computed with

$$\widehat{p}_{x,l,h,s,y} = \frac{\sum_{h=1}^{H_{s,y}} \left[N_{x,l,h,s,y}/N_{h,s,y}\right]}{\sum_{x=1}^{3}\sum_{l=1}^{L}\sum_{h=1}^{H_{s,y}} \left[N_{x,l,h,s,y}/N_{h,s,y}\right]} \tag{3}$$

where $N_{x,l,h,s,y}$ is the length frequency sampled, in numbers, by sex-x and length-l (in cm) within a haul-h for stratum-s in year-y. Note that when expanding length frequencies at AFSC the length bins are set at 1 cm (that span the size range for each species), as this is how the length bin structure is set in many of the stock assessment models employed at AFSC, however, these formulae can be used for other bin sizes (for example, 2 cm or larger). Finally, the expanded population abundance-at-length is obtained by multiplying the overall population numbers within the strata (Eq. (1)), the relative catch per unit effort of each haul (Eq. (2)), and the sex-specific relative length composition (Eq. (3)) with

$$\widehat{N}_{x,l,s,y} = \widehat{N}_{s,y} \cdot \sum_{h=1}^{H_{s,y}} \left[\widehat{p}_{C,h,s,y} \cdot \widehat{p}_{x,l,h,s,y}\right] \tag{4}$$

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). Strata are defined as regions with similar bathymetric characteristics (e.g., depth ranges), and population abundance-at-length within strata can also be summed to any sub-region level. We note that this formulation is equivalent to the design-based length composition expansion used in Stewart and Hamel (2014). The only difference is found in multiplying the sex-specific relative length composition by the relative catch-per-unit-effort for each haul in Eq. (4) here, where Stewart and Hamel (2014) multiply the sex-specific relative length composition by the predicted number of fish in a haul.

Age-length-keys (ALKs) generated from the age-length paired observations within a survey are then applied to estimated abundance-at-length to provide an estimate of abundance-at-age (e.g., Quinn and Deriso, 1999), referred to as the 'second stage expansion'. In the second stage expansion, the sex-specific estimates of population abundance-at-length (from Eq. (4)) are used to estimate sex-specific population abundance-at-age. The annual specimen data collected during the survey, which include observations of age-at-length, are first populated into sex-specific numbers at age and length ($N_{x,a,l,y}$). Next, the sex-specific numbers-at-age and length are converted to sex-specific proportions of age-at-length (i.e., age-length key) with

$$\widehat{p}_{x,a,l,y} = \frac{N_{x,a,l,y}}{\sum_{a=1}^{A} N_{x,a,l,y}} \tag{5}$$

The proportions of age-at-length are then expanded to population abundance-at-age with

$$\widehat{N}_{x,a,y} = \sum_{l=1}^{L} \widehat{p}_{x,a,l,y} \cdot \widehat{N}_{x,l,y} \tag{6}$$

where $\widehat{N}_{x,l,y}$ is the population abundance-at-length from Eq. (4) summed across strata.

For both the expanded population numbers-at-length and -age the formulae presented here perform the expansions for sex-specific data. Thus, population numbers-at-length and age for male, female, and unsexed categories are computed, and the total population numbers-at-length and -age are computed by summing across these sex categories. While these formulae are presented for specific sex categories, the methods developed in this study are also flexible to combining data across the sex categories (males, females, and unsexed) prior to the first and second stage expansions, thus, estimating a total (or combined sex) length and age composition without the need for summation after the first and second stage expansions.

2.3. Simulation-bootstrap framework

To evaluate the effect of the inclusion of ageing error and growth variability on uncertainty in age composition datasets we modified a bootstrap-simulation framework (Hulson et al., 2023) to include these additional sources of error. In simple terms, the simulation framework is a two-stage bootstrap that first resamples hauls, then resamples lengths and ages collected within the resampled hauls following from the methods in Stewart and Hamel (2014). The simulation framework was modified to account for ageing error by resampling from tester ages associated with a given reader age. Growth variability was incorporated by resampling from lengths associated with a given age and sex. We developed these simulations so that growth variability can be incorporated by either pooling the age-length data across all survey years and resampling the lengths for a given age, or using the annual age-length data and only resampling the lengths for a given age that were observed within the specific survey year.

The order of operations (Fig. 1) has the following schedule:

1. Resample hauls 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.
3. With the resampled length frequency data from step 2, calculate population abundance-at-length (Eqs. (1) - (4)).
4. Within the resampled hauls from step 1, resample the observed ages from the specimen data.
5. For the resampled ages in step 4, resample a length from the set of lengths observed for the given age.
6. For the resampled ages in step 4, resample an age from the set of tester ages for the given age.
7. With the resampled age data in steps 4–6 and the population abundance-at-length in step 3, calculate the population abundance-at-age (Eqs. (5) - (6)).

We also include functions that compute conditional age-at-length (CAAL) in addition to the expansion methods described above. To compute CAAL we perform step 1, then steps 4 – 6, and in step 7 we compute the ALK (Eq. (5)) without the abundance-at-age expansion. Steps 5 and 6 were designed to explore inclusion of growth variability and ageing error. The bootstrap-simulation repeats these steps providing iterated population abundance-at-age and CAAL for comparison to the historical (the full sample without any resampling of data) population abundance-at-age and CAAL.

2.4. Computing input sample size

A useful statistic that can quantify the variability in age composition is realized sample size, introduced by McAllister and Ianelli (1997); using the terminology of Stewart and Hamel (2014). This statistic

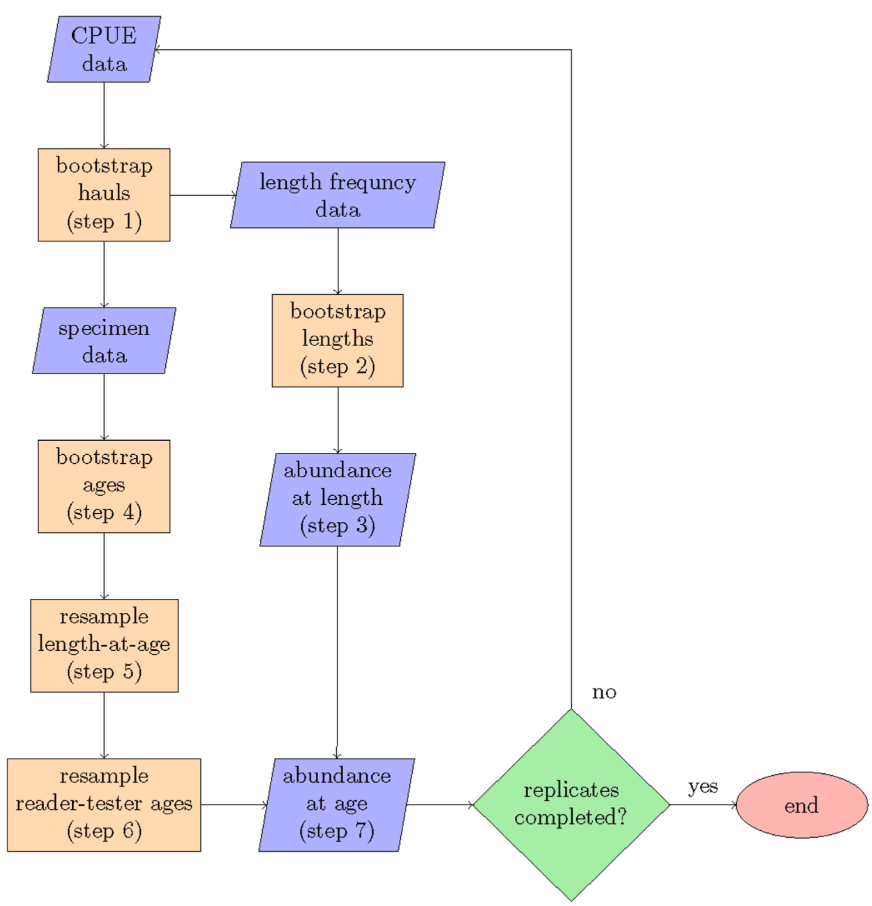

Fig. 1. Bootstrap-simulation flow chart, the steps refer to the order of operations.

evaluates the amount of uncertainty in an estimated composition compared to an observed composition and is given by:

$$R_{i,y} = \frac{\sum_{c=1}^{C} O_{c,y}\left(1 - O_{c,y}\right)}{\sum_{c=1}^{C} \left(E_{c,i,y} - O_{c,y}\right)^2} \tag{7}$$

where $E_{c,i,y}$ is the estimated proportion for category-c (which can be age or any other arbitrary category across which proportions are computed) for iteration-i in year-y and $O_{c,y}$ is the observed proportion. We note, that for the realized sample size of CAAL, there would be an additional subscript introduced in Eq. (7) for length bin-l (e.g., $O_{c,l,y}$), where category-c would be age, thus providing a realized sample size for each length bin within a given year's CAAL data. Here, the underlying age composition and CAAL derived from the historical bottom trawl surveys with the full and unsampled data was treated as the observed proportions in Eq. (7). For each iteration-i of the bootstrap-simulation we computed an estimated proportion ($E_{c,i,y}$) that was then compared to the observed age composition ($O_{c,y}$) to determine the realized sample size ($R_{i,y}$) of the resampled age composition or CAAL. Thus, across each iteration of the bootstrap-simulation we computed a realized sample size that indicated the amount of uncertainty in the resampled age composition or CAAL.

To summarize realized sample size across iterations we used the harmonic mean. This has been shown to reduce bias in recovering the true sample size in simulations for a multinomial distribution and has also been recommended for determining the ISS that is used in stock assessment models to fit compositional data (Stewart and Hamel, 2014). Thus, for the expanded age composition data we present the annual ISS computed from the harmonic mean of the annual iterated realized sample sizes. For CAAL the ISS for each length bin within the annual CAAL data was computed as the harmonic mean of the bin-specific realized sample size across the iterations. Then, to summarize the effect of additional uncertainty, we compute the arithmetic mean of the ISS across the length bins (rather than show the ISS for each year and length bin). While we present the results of the annual ISS for each stock evaluated when incorporating ageing error, growth variability, or both, we also compute the proportion of 'base' ISS in order to present the relative decrease in ISS when incorporating these sources of additional uncertainty. The 'relative ISS' is computed by dividing the ISS as determined after incorporating ageing error, growth variability, or both, by the base ISS without these sources of uncertainty.

2.5. Bootstrap-simulation scenarios and treatments

We applied the bootstrap-simulation in a step-wise manner to evaluate the consequences of adding each source of additional error to the age composition estimates across what we term 'uncertainty scenarios' (Table 2). First, we ran the standard bootstrap-simulation omitting steps 5 and 6 above ('Base' scenario). Next, we added ageing error ('AE' scenario) and growth variability ('GV' scenario) separately, thus, omitting either step 5 or 6 depending on the source of uncertainty desired. Finally, we added both ageing error and growth variability ('AE & GV' scenario) to the bootstrap-simulation framework. To increase reader-tester sample sizes for each species, we pooled reader-tester data across the three regions (we note that age readings for all three regions are produced in the same age reading laboratory at AFSC). To generalize the presentation of results we aggregate across regions and species types, thus, annual ISS and relative ISS results are shown for flatfish, gadids, and rockfish across the stocks and regions included in this analysis (Table 1). In the presentation of CAAL results we selected example stocks for each of the species types; GOA arrowtooth flounder (*Atheresthes stomias*) as an example for flatfish, GOA Pacific cod (*Gadus macrocephalus*) as an example for gadids, and GOA Pacific ocean perch (*Sebastes alutus*) as an example for rockfish.

We applied three bootstrap-simulation treatments across the uncertainty scenarios in order to evaluate the consistency of the results after incorporating each additional error source (Table 2). In the first treatment we evaluated the impact of pooling age-length data across all years ('Pooled') versus using the annual age-length data ('Annual') when resampling lengths for a given age to incorporate growth variability; we term this treatment the 'Growth data treatment'. In the second treatment we evaluate the impact of different length bins for the length frequency data by including 2 cm and 5 cm length bins in addition to the base bin of 1 cm for comparison; we term this treatment the 'Length bin treatment'. In the third treatment we show an example of aggregating length and age data prior to length and age expansion ('Pre-expansion') or after length and age expansion ('Post-expansion'); we term this treatment the 'Aggregation treatment'. For this treatment we selected two stocks to show as an example: GOA Pacific cod and GOA Pacific ocean perch. We selected these stocks because they do not exhibit differences in growth between females and males, which is the primary consideration for aggregating data either before or after length and age expansion.

The bootstrap-simulations were run for 500 iterations, a level at which the variability in population abundance-at-age results had stabilized. 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/surveyISS).

2.6. Evaluating sampling and life-history relationships to consequences of added uncertainty

For the three species types in this analysis (flatfish, gadids, and rockfish) we evaluated relationships between sampling rates and indicators of life-history traits across the uncertainty scenarios considered. To evaluate the relationship with sampling rates, and the consequence of added uncertainty in ISS, we present the relationship between the average ISS per age sampled and the number of ages collected which are comparable the type of results presented in Stewart and Hamel (2014). To evaluate the relationship between ISS and life-history and the consequences of additional sources of uncertainty we used two indicators. First, we compare relative ISS after incorporating ageing error with the age range of the stocks to assess the relationship with longevity. Second, we compare relative ISS after incorporating growth variability with the length range of the stocks to determine if the impacts of growth variability are related to the size of the species type considered. Finally, we rank the relative ISS after incorporating both ageing error and growth variability across stocks (and highlighting species types) to illustrate any species type impacts on ISS after incorporating these sources of variability.

3. Results

While the magnitude of age composition ISS was stock and region specific, there was a consistent reduction in age composition ISS for each species types with the introduction of additional sources of uncertainty in the bootstrap procedure (top panels of Fig. 2, shown by region and stock in Fig. S1). This reduction in age composition ISS resulted for both sex-specific (female and male) and total (combined sex) age composition ISS. The magnitude of adding ageing error compared to growth variability was species type specific (top panels Fig. 2) and stock and region specific (Fig. S1). For example, adding ageing error to rockfish age data resulted in smaller ISS on average than adding growth variability, but a larger ISS for flatfish and gadids. For all the species types, age composition ISS was the smallest when both ageing error and growth variability were included in the bootstrap-simulation procedure.

The relative age composition ISS across uncertainty scenarios revealed patterns among species types, where flatfish and gadids had similar reductions in age composition ISS, and greater reductions than rockfish age composition ISS, when the additional sources of uncertainty were included in the bootstrap-simulation procedure (bottom panels of

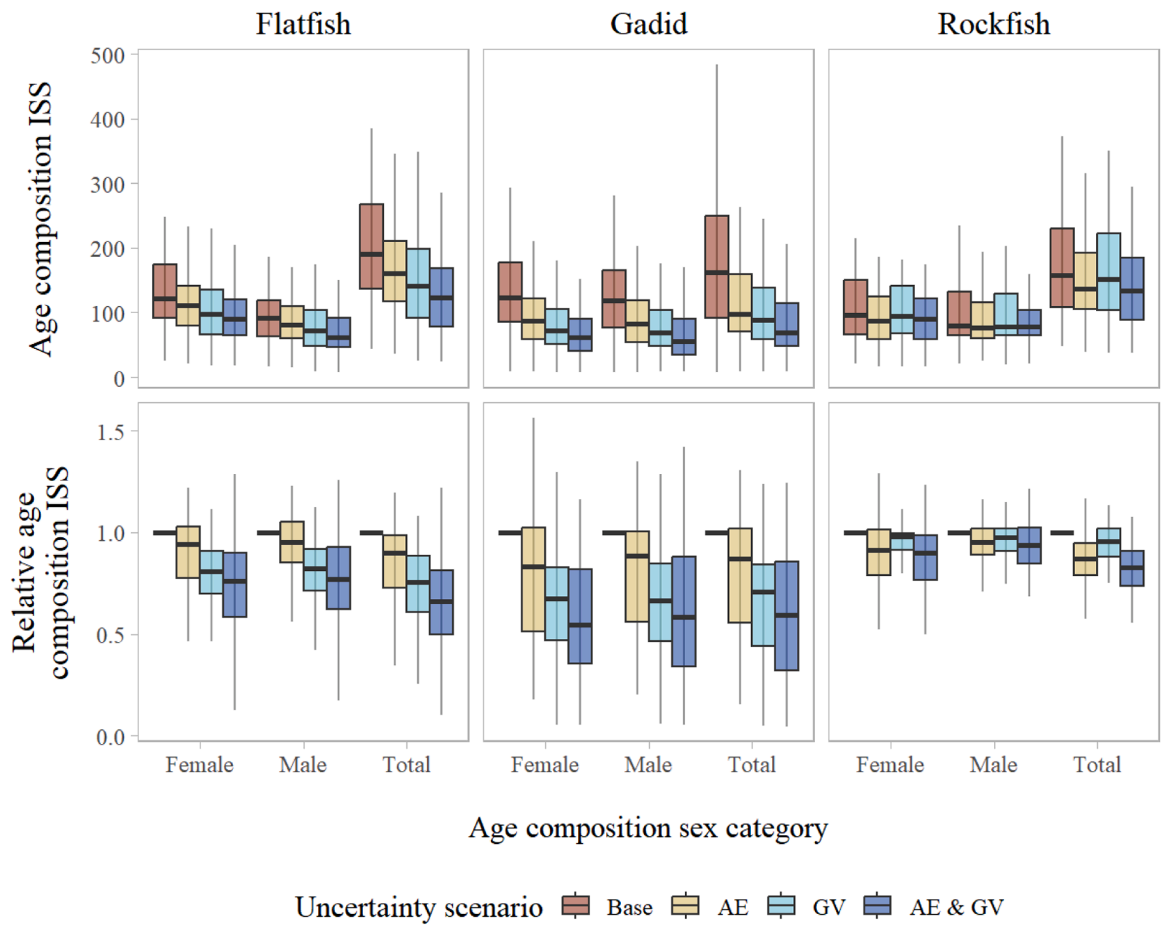

Fig. 2. Boxplots of annual age composition input sample size (top row) and relative age composition input sample size (bottom row) aggregated by species type across uncertainty scenarios and sex category for 1 cm length bins and pooled growth data. Uncertainty scenarios are described in Table 1. The boxplots shows the median (solid line), 25–75% percentile range (box limits, also called the inter-quartile range), and 1.5 times the inter-quartile range (whiskers).

Fig. 2). Including ageing error for rockfish had a larger proportional reduction in relative ISS than including growth variability. However, when pooling the growth data across years, growth variability had a larger proportional reduction for flatfish and gadid relative ISS than adding ageing error. Gadids exhibited the greatest variability in the proportional reduction in relative ISS and extended to the largest reduction in relative ISS, while rockfish had the smallest reduction in relative ISS across the uncertainty scenarios, in general. Overall, when adding both ageing error and growth variability the median decrease in age composition ISS was 72% for flatfish, 56% for gadids, and 88% for rockfish when compared to age composition ISS that does not include these sources of uncertainty. We also note an interesting result where there are a few instances that the relative ISS is greater than 1.0. This indicates that when implementing ageing error and growth variability there is a random chance that the age composition ISS could increase compared to the base ISS. However, this occurred for less than 25% of the stock-year age composition ISS.

Similar reductions in age composition ISS were observed in the growth data treatment whether using pooled or annual growth data for flatfish or rockfish, however, annual ISS decreased less for gadids (top of Fig. 3, shown for individual stocks in Figs. S1 and S2). In general, for all the species types, the relative ISS was smaller when using pooled growth data compared to annual growth data, indicating an increase in uncertainty when using pooled growth data as compared to annual growth data. Further, the variability in the relative ISS was reduced when using annual growth data as compared to pooled growth data in the growth variability uncertainty scenario (GV), for all species types, though this was particularly true for flatfish and gadids.

For the remaining treatments (length bin and aggregation treatments) we show results using the annual growth data when implementing growth variability (but note that the trend of results was consistent regardless of growth data treatment). Slight differences in age composition ISS were observed for the length bin treatment and aggregation treatment, and the relative ISS results were consistent across the species types (Figs. 4 and 5). As length bin size increased the age composition ISS increased across each of the uncertainty scenarios (top panels of Fig. 4, shown for individual stocks in Figs. S2 – S4). The increase in age composition ISS ranged from 4% to 13% for 2 cm bins and 10 – 23% for 5 cm bins across the species types and uncertainty scenarios compared to the age composition ISS using 1 cm length bins. In the length bin treatment the relative ISS, and variability in relative ISS, resulted in similar values for each of the uncertainty scenarios regardless of the size of the length bin for each of the species types (bottom panels, Fig. 4). An increase in age composition ISS resulted when aggregating combined sex data pre-expansion as compared to post-expansion in the aggregation treatment for the example stocks presented (top panels of Fig. 5). This increase in age composition ISS ranged from 4% to 11% for these stocks across the uncertainty scenarios when aggregating combined sex data pre-expansion as compared to post-expansion. Similar to the length bin treatment, the relative ISS remained largely unchanged whether the combined sex data were aggregated pre-expansion or post-expansion (bottom panels of Fig. 5).

Similar to results for expanded age composition, the magnitude of conditional age-at-length ISS (presented as the mean across length bins) was stock specific and decreased across the uncertainty scenarios as ageing error and growth variability was introduced (Fig. 6). The

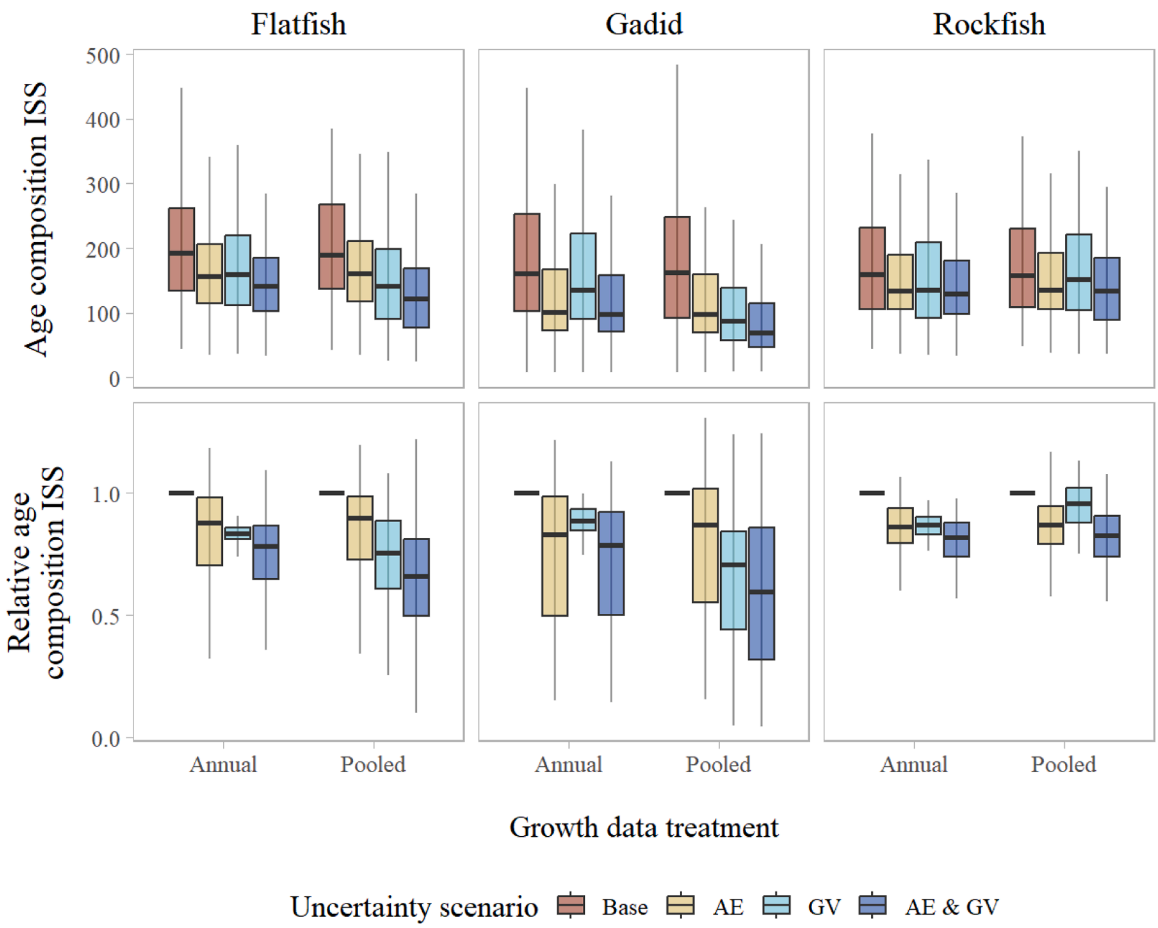

Fig. 3. Boxplot of annual age composition input sample size (top row) and relative age composition input sample size (bottom row) aggregated by species type across uncertainty scenarios within each growth data treatment (shown for total age composition expanded with 1 cm length bins). Uncertainty scenarios are described in Table 1. The boxplots shows the median (solid line), 25–75% percentile range (box limits, also called the inter-quartile range), and 1.5 times the inter-quartile range (whiskers).

magnitude of age composition ISS differed compared to conditional age-at-length ISS for the stock examples. For example, arrowtooth flounder age composition ISS was generally larger than either walleye pollock or Pacific ocean perch age composition ISS, though Pacific ocean perch conditional age-at-length ISS was generally larger than arrowtooth flounder and walleye pollock conditional age-at-length ISS (top two rows of Fig. 6). While the relative age composition and conditional age-at-length ISS decreased across the uncertainty scenarios, the magnitude of decrease within the uncertainty scenarios was different when comparing between age composition and conditional age-at-length (bottom two rows of Fig. 6). For example, the decrease in relative conditional age-at-length ISS when implementing growth variability was larger for arrowtooth flounder and Pacific ocean perch than the decrease in relative age composition ISS. Additionally, the decrease in relative conditional age-at-length ISS was larger for walleye pollock for each of the uncertainty scenarios as compared to relative age composition ISS.

A decreasing relationship was observed for each species type between the age composition ISS per age sample and the total number of age samples collected (top panels of Fig. 7). This decreasing relationship resulted for each of the uncertainty scenarios, but was not a significant relationship with R^2 values less than 0.23 for each linear model fit, where the majority were below 0.1 (shown in text in the top panels of Fig. 7). The median age composition ISS per age sample ranged from 0.3 to 0.37 for flatfish, 0.14 – 0.23 for gadids, and 0.25 – 0.29 for rockfish across the uncertainty scenarios (bottom panels of Fig. 7). We note that the uncertainty in these median values of age composition ISS per age sample was large, with coefficients of variation upwards of 28% for flatfish, 61% for gadids, and 33% for rockfish.

An increasing relationship resulted for each species type between the age composition ISS per sampled haul and the number of age samples per sampled haul (top panels of Fig. 8, we note these panels are analogous to Fig. 4 in Stewart and Hamel, 2014). The strongest relationship resulted for flatfish (with R^2 values of 0.8 – 0.93), was intermediate for rockfish (with R^2 values of 0.59 – 0.64), and was the weakest for gadids (with R^2 values of 0.39 – 0.56). The linear relationship also degraded as additional uncertainty was incorporated across the uncertainty scenarios for all the species types. The median age composition ISS per sampled haul ranged from 2.5 to 3.1 for flatfish, 0.9 – 1.5 for gadids, and 1.5 – 1.7 for rockfish across the uncertainty scenarios (bottom panels of Fig. 8, we note these panels are analogous to Fig. 3 in Stewart and Hamel, 2014). The uncertainty in these median values of age composition ISS per sampled haul was large, with coefficients of variation upwards of 74% for flatfish, 86% for gadids, and 57% for rockfish.

Comparing between statistics for longevity (as indicated by age range) and growth (as indicated by length range) resulted in generally similar trends in the relative age composition ISS by species types when adding either ageing error or growth variability (top panels of Fig. 9). The relative age composition ISS when adding ageing error had a decreasing trend when compared to longevity for each of the species types, indicating that the longer lived the stock the larger the effect of ageing error had on age composition ISS (top left panel of Fig. 9). A similar decreasing trend resulted for each species type when comparing relative age composition ISS after adding growth variability with the length range of the stocks. This indicates that the larger size range that a

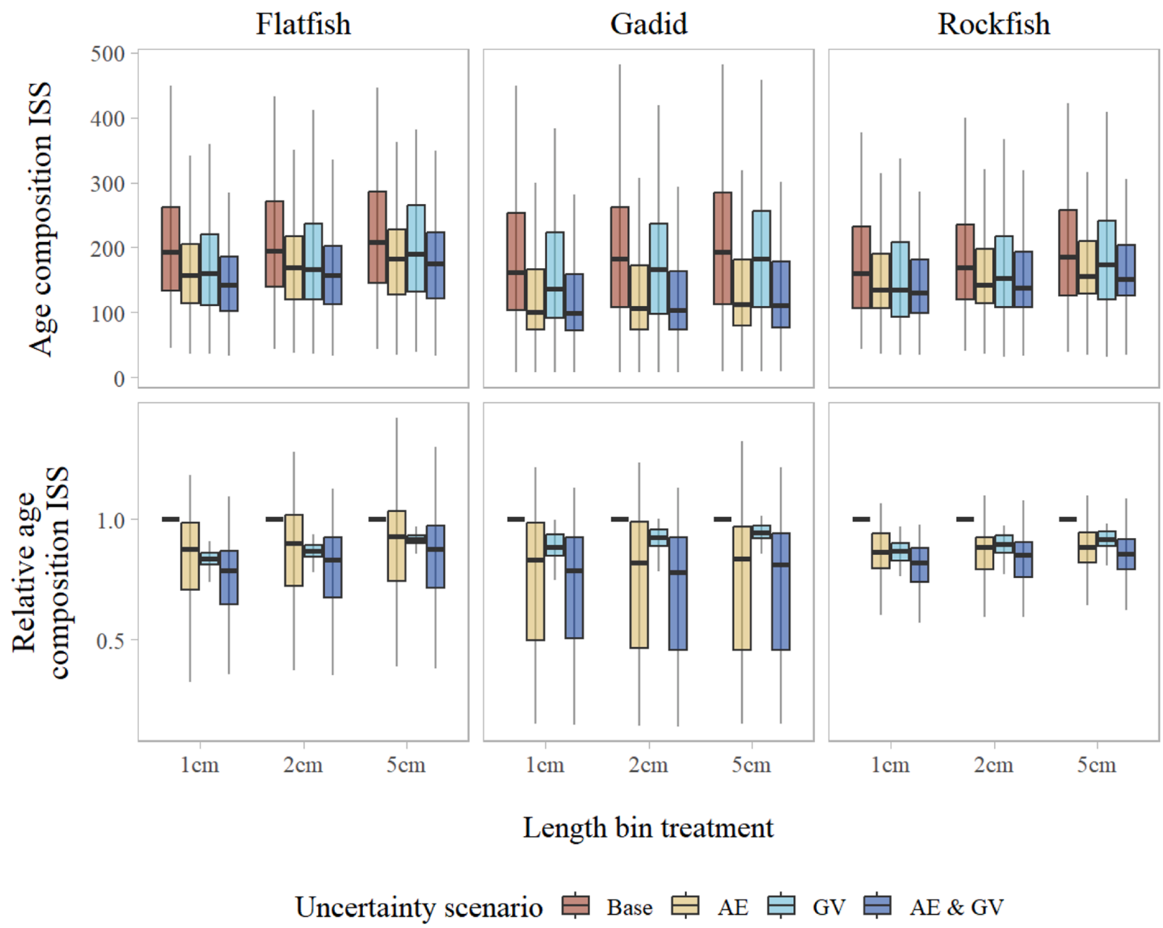

Fig. 4. Boxplot of annual age composition input sample size (top row) and relative age composition input sample size (bottom row) aggregated by species type across uncertainty scenarios within each length bin treatment (shown for total age composition expanded using annual growth data). Uncertainty scenarios are described in Table 1. The boxplots shows the median (solid line), 25–75% percentile range (box limits, also called the inter-quartile range), and 1.5 times the inter-quartile range (whiskers).

stock grows the more of an effect growth variability has on decreasing the age composition ISS (top right panel of Fig. 9). On average, the relative ISS when adding both ageing error and growth variability was largest for rockfish (85% of the base age composition ISS), intermediate for flatfish (78% of the base age composition ISS), and smallest for gadids (70% of the base age composition ISS, bottom panel of Fig. 9). The same trend resulted when evaluating the range in the relative age composition ISS when both ageing error and growth variability were added, where the range was smallest for rockfish, intermediate for flatfish, and largest for gadids.

4. Discussion

In this study we find that accounting for ageing error and growth variability using bootstrap procedures decreased age composition ISS for all stocks examined. This result was consistent across all the treatments that we applied, which included pooling growth data (annual data or across time), different length bins (1 cm, 2 cm, or 5 cm bins), and differences in aggregating data for total age composition (before or after length and age expansion). We also show that this result was consistent for conditional age-at-length ISS. The impact of the sources of uncertainty on resulting ISS was species type specific, with ageing error being more influential for rockfish than growth variability, and growth variability more influential than ageing error for flatfish and gadids. However, the influence of growth variability for gadids and flatfish was sensitive to the pooling of growth data; age composition ISS was smaller when growth variability was applied to pooled data as compared to annual data for gadids, and the range in the decrease in age composition ISS was smaller for annual data compared to pooled data for flatfish. We propose that these results are due to larger inter-annual growth variability observed in gadids and flatfish compared to rockfish. Further, the effects of ageing error are not unexpected for rockfish, as they are so long-lived. When considering both ageing error and growth uncertainty the largest reduction in ISS magnitude was for gadids, followed by flatfish, with the least effect observed for rockfish, though results varies by stock and region.

When applying the bootstrap procedure we developed in this study to estimate age composition ISS there are several considerations that should be made that are specific to the stock that is being analyzed. These considerations include: (1) the size of length bins used for length data, (2) whether to aggregate length and age data prior to or after length and age expansion, and (3) whether to pool growth data across time or use annual data.

In this study we found that age composition ISS can increase as the size of the length bin increases. For example, age composition ISS for 5 cm length bins was larger than when using 1 cm length bins, albeit, the increase was not significant. In Hulson et al. (in press) this result was also presented for length composition ISS where increasing the size of length bins increased the ISS. This result suggests that increasing the size of the length bin reduces the amount of uncertainty in the length frequency collections. In this case we recommend that length bins larger than 1 cm be considered in order to increase the age composition ISS.

We also show that aggregating combined sex, length and age data prior to length and age expansion can increase age composition ISS as

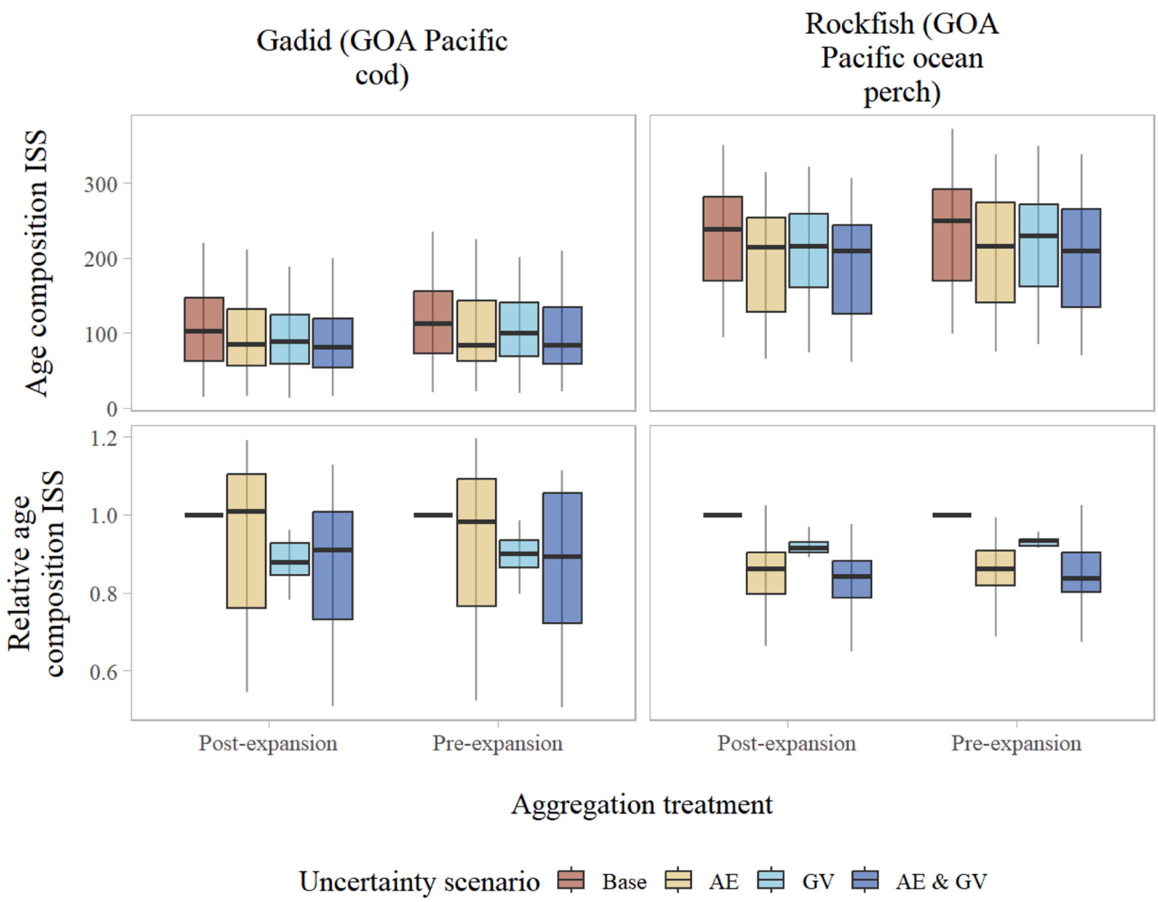

Fig. 5. Boxplot of annual age composition input sample size (top row) and relative age composition input sample size (bottom row) for the selected example species type stocks across uncertainty scenarios within each aggregation treatment (shown for total age composition expanded using annual growth data and 1 cm length bins). Uncertainty scenarios are described in Table 1. The boxplots shows the median (solid line), 25–75% percentile range (box limits, also called the inter-quartile range), and 1.5 times the inter-quartile range (whiskers).

compared to summing sex-specific length and age composition after expansion. While the increase was small (and not significant), this is an important consideration to be made, particularly if the stock that is being assessed does not display sex-specific differences in growth. For example, the rockfish and gadid stocks presented here do not display sex-specific differences in growth, and we recommend that pre-expansion aggregation of age and length data be considered. However, the flatfish stocks included in this analysis do exhibit sex-specific differences in growth, and thus, we recommend that sex-specific length and age compositions be constructed.

When implementing growth variability into the bootstrap-simulation procedure the resulting magnitude of age composition ISS was sensitive to how the growth data was pooled. This was particularly true for the gadid stocks evaluated in this study, where, for example, using pooled growth data resulted in smaller age composition ISS than when using annual growth data. Alternatively, for the rockfish stocks evaluated, there was not a large difference in age composition ISS magnitude whether using pooled growth data or annual growth data. We recommend that for stocks that exhibit inter-annual variability in growth, particularly inter-annual variability that may result in a trend in growth across time (i.e., increasing or decreasing size over time), that annual growth data be used in order to avoid over-estimating the effect of growth variability on age composition ISS. For stocks that don't exhibit inter-annual variability in growth, we recommend using pooled growth data in order to more adequately incorporate the potential growth variability by leveraging the larger sample size in pooled data compared to annual data.

It is well known that misspecification of ISS when fitting compositional data can lead to biased results in assessment model predictions (e.g., Stewart and Monnahan, 2017, Xu et al., 2020). Here, we show for gadids that the ISS for some stocks and years when adding additional uncertainty as compared to the base case could be as small as 21% of the base case ISS, as small as 41% of the base case for flatfish, and as small as 61% of the base case for rockfish. Without these additional sources of uncertainty taken into account, using the bootstrap procedure would inflate the ISS. While we did not investigate implications to specific SCAA model outcomes, we can infer that reductions of ISS of this scale will have downstream effects on model predictions and associated uncertainty. We note that these sources of uncertainty would not be contained only to fishery-independent sources, like evaluated here, but would also be inherent to age collections from fishery-dependent sources as well. The functions to bootstrap age composition data to determine ISS for fishery-dependent sources are currently under development. Future investigations into the impacts of adding ageing error and growth variability into ISS estimation on SCAA model results should also include fishery-dependent ISS implications. These investigations should also include data weighting methods that allow for adjustment to weighting between the data sources integrated, for example using the methods presented in Francis (2011) and Thorson et al. (2017). We note that while these methods can adjust the relative influence of composition data on SCAA model results, the initial ISS matters. In theory these re-weighting methods can overcome misspecification of ISS, however, in practice SCAA model results can be sensitive to the starting values of ISS. This implication points to the importance of using length and age composition ISS that adequately include the sources of uncertainty common to age and length

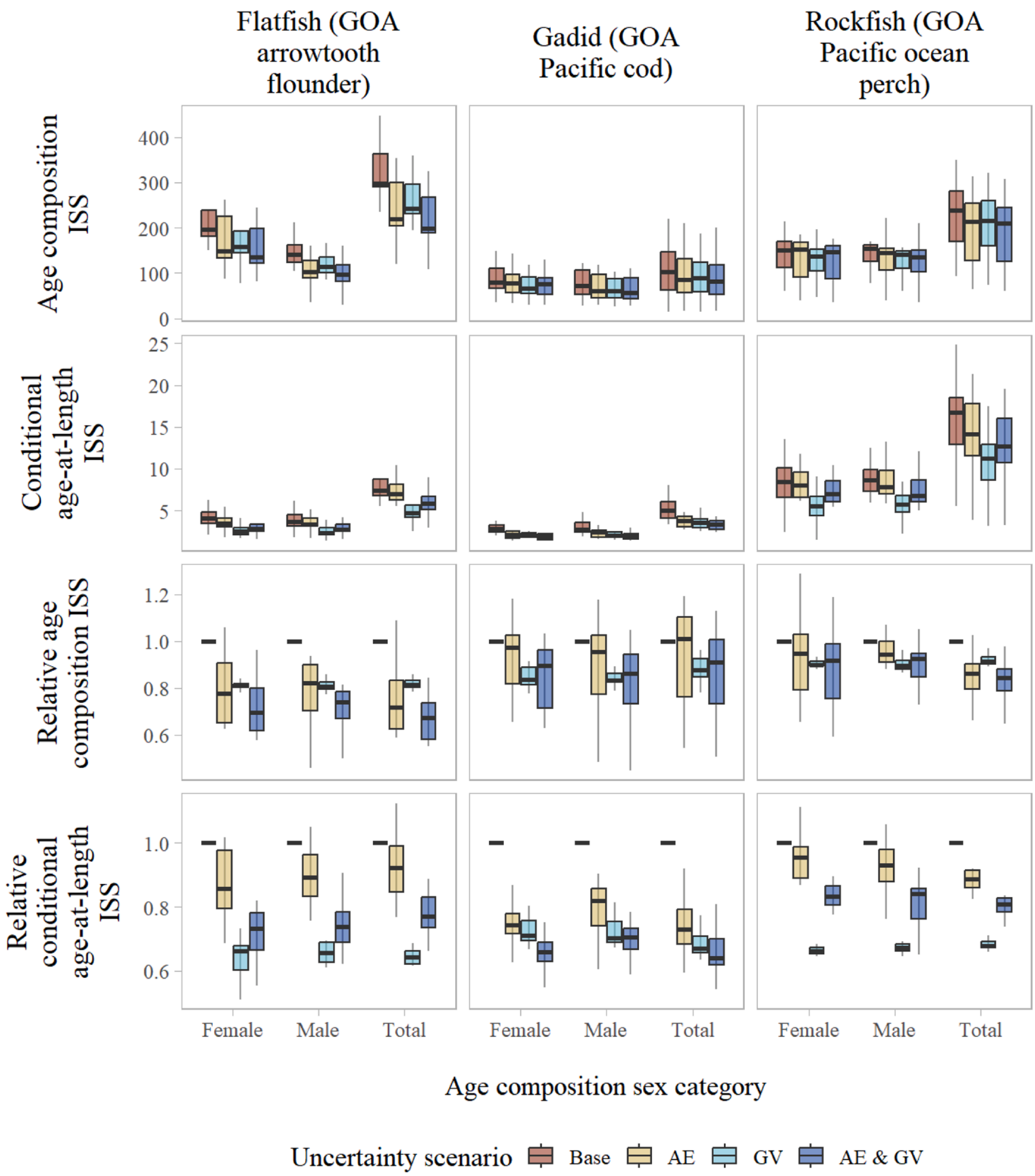

Fig. 6. Boxplot of annual age composition and conditional age-at-length input sample size (top two rows) and relative age composition and conditional age-at-length input sample size (bottom two rows) for the example species type stocks across uncertainty scenarios by sex category (using annual growth data and 1 cm length bins). Uncertainty scenarios are described in Table 1. The boxplots shows the median (solid line), 25–75% percentile range (box limits, also called the inter-quartile range), and 1.5 times the inter-quartile range (whiskers).

composition data.

A number of operational assessment models use a proxy in some form based on the number of sampled hauls or the nominal sample size when setting ISS for age and length composition data (e.g., Hulson et al., 2021, Barbeaux et al., 2022). Considering the proxy using hauls, (derived from Pennington et al., 2000 who investigated length frequency sampling and determined, on average, the ISS was one fish per haul) we note that the conclusion made in Pennington et al. (2000) was not that the number of hauls should be used as a proxy for ISS in the assessment model data fitting procedure, but, rather, that in order to potentially increase the ISS that samples should be taken from an increased number of hauls. Alternatively, considering the proxy using nominal sample size, the results of Stewart and Hamel (2014) have been used to scale nominal sample size to length composition ISS based on the average ISS per sample reported in that study. Here we find that the relationship between age composition ISS and either nominal sample size or sampled hauls to be highly variable and statistically insignificant. We recommend that length and age composition ISS be determined from a bootstrap procedure rather than scaling hauls or nominal sample size based on the relationship with bootstrap ISS results. We note that the bootstrap procedure presented here is not computationally burdensome and can be applied to specific stocks in a matter of minutes.

Previous work has investigated both the inclusion of ageing error (Punt et al., 2008; Liao et al., 2013) and growth (Taylor and Methot, 2013) within stock assessment models, however, none have applied these additional sources of uncertainty in the context of estimating ISS.

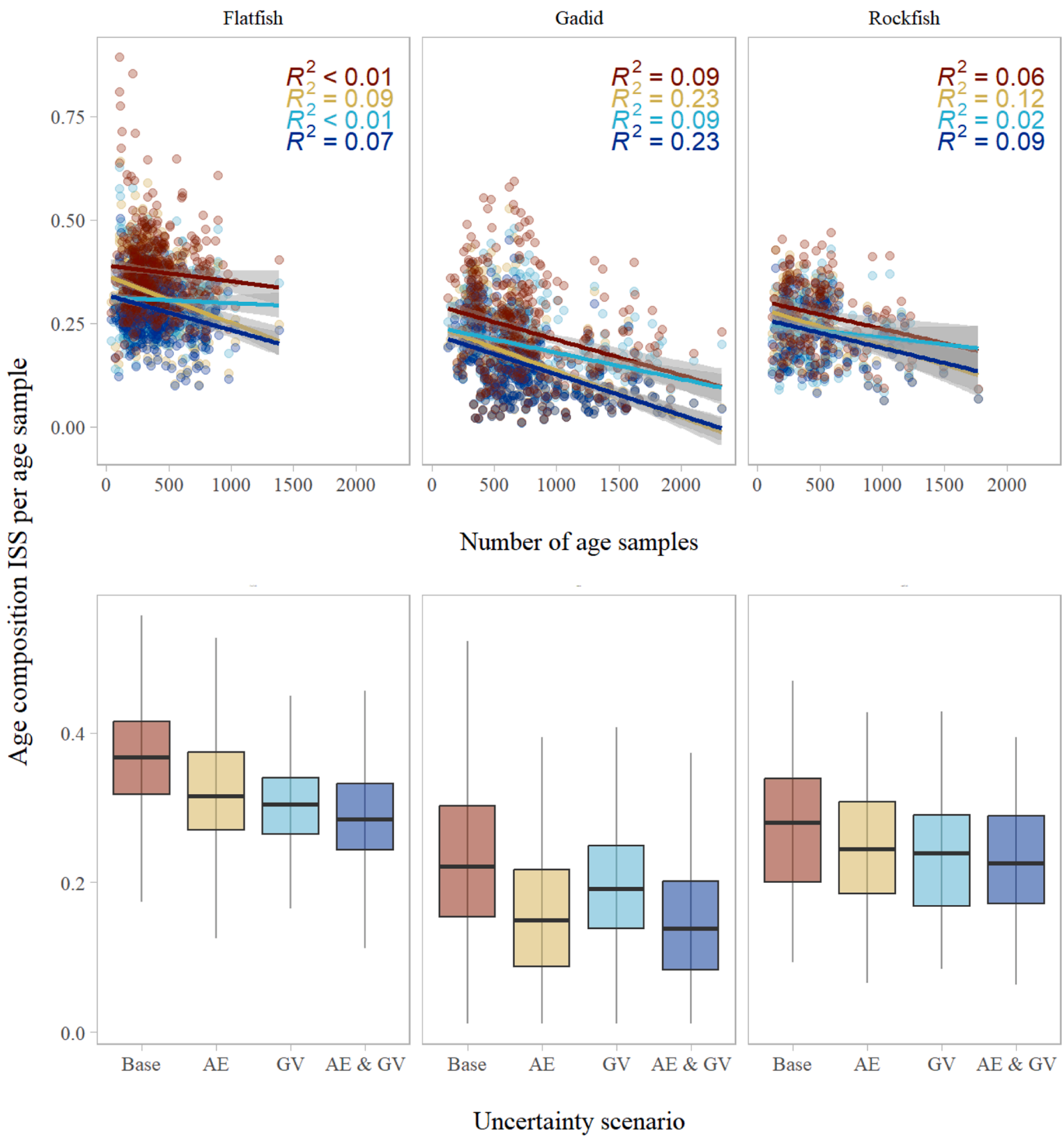

Fig. 7. Age composition input sample size per age sample compared to the number of ages sampled (top panels, points represent species-year combinations) and across uncertainty scenarios (bottom panels) aggregated by species types. Uncertainty scenarios are described in Table 1. Linear relationships are shown in the top panels, along with the R^2 values, for each uncertainty scenario (color). The boxplots in the bottom panels shows the median (solid line), 25–75% percentile range (box limits, also called the inter-quartile range), and 1.5 times the inter-quartile range (whiskers).

As it pertains to ageing error, in many current assessments an ageing error matrix is implemented either to be applied to marginal age composition data (e.g., Williams et al., 2022) or applied to CAAL data (e.g., Hulson et al., 2022) in order to account for ageing error that is inherent to the age composition and CAAL data. Use of an ageing error matrix effectively 'smudges' assessment model estimates of population-at-age into adjacent age classes prior to fitting the observed proportions in the age composition input data from either fishery-independent or fishery-dependent sources. Then, in the model fitting step, an ISS is used to fit the model's 'smudged' estimates of age composition to the observed age composition. Here, we suggest that unless ageing error is accounted for in the age composition or CAAL ISS used to fit these data then we are likely 'over-fitting' the model estimates to the 'observed' age composition and that the use of only an ageing error matrix only partially accounts for this source of uncertainty in the assessment modeling process. The results from this study suggest that in some cases the addition of ageing error to the bootstrap method decreases the age composition ISS to such an extent that an ISS that does not take into account this source of uncertainty can be 125–165% too large.

It is commonly the case that age composition is produced through a two-stage expansion process (Quinn and Deriso, 1999), in which length frequency data is expanded to population-at-length in the first stage, and an ALK is used to expand population-at-length to population-at-age in the second stage (Ailloud and Hoenig, 2019). In the second-stage of this expansion process, the ALK is produced through the use of age-length

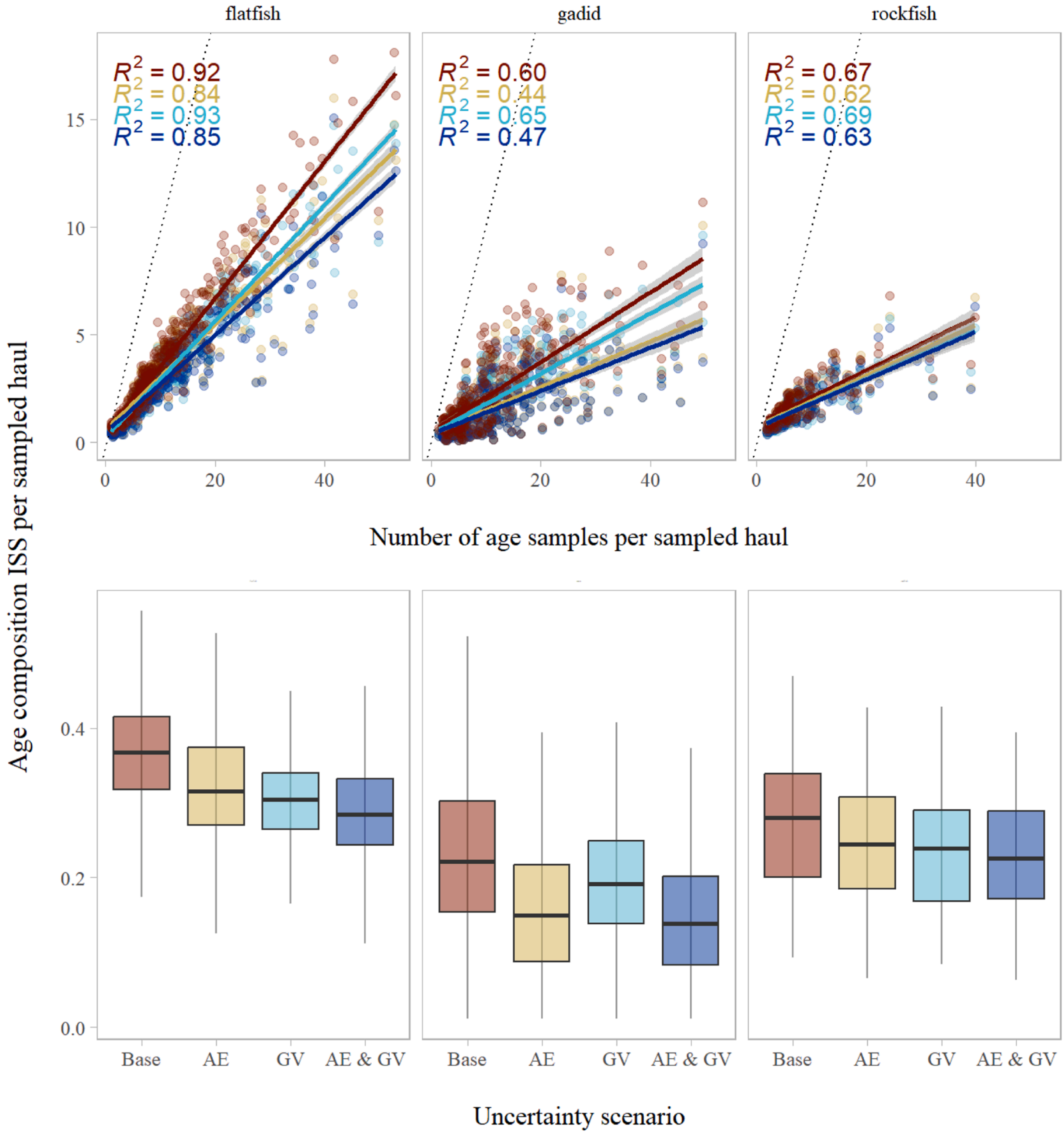

Fig. 8. Age composition input sample size per sampled haul compared to the number of age samples per sampled haul (top panels, points represent species-year combinations) and across uncertainty scenarios (bottom panels) aggregated by species types. Uncertainty scenarios are described in Table 1. Linear relationships are shown in the top panels, along with the R^2 values, for each uncertainty scenario (color; and a dashed 1–1 line is shown for reference). The boxplots in the bottom panels show the median (solid line), 25–75% percentile range (box limits, also called the inter-quartile range), and 1.5 times the inter-quartile range (whiskers).

paired data that are obtained in the age sampling collection, and within the ALK the variability in length-at-age is implicitly accounted for. In other cases CAAL data is used in conjunction with expanded length composition data so that growth can be estimated internally in an SCAA model (e.g., Hulson et al., 2022). The CAAL data is computed in a similar method as expanded age composition, in that the ALK that is produced is used directly in the model, rather than expanding by length composition to produce marginal age composition externally to the SCAA model. However, there have been no previous attempts to include this source of variability when considering the ISS that is used to fit the expanded age composition or CAAL data. We find that the magnitude of effect on age composition and CAAL ISS is species type dependent, and can have different implications when either using age composition or CAAL data. When implementing growth variability there was a 10–50% decrease in the magnitude of bootstrapped age composition and CAAL ISS, thus, if not taking this source of uncertainty into account the ISS for either source of information can be up to 200% too large. We also note that in some cases age composition data is not expanded by an ALK, but rather, the age specimens are used directly to compute the age composition (for example, in cases when age specimen data are collected randomly rather than stratified). We suggest that future analyses be conducted to quantify the relative benefit in ISS of using an ALK versus expanding randomly collected ages directly without benefitting from the expanded sample size of lengths.

An additional consideration is understanding effects of survey reduction effort, the focus of a number of recent studies (ICES, 2020,

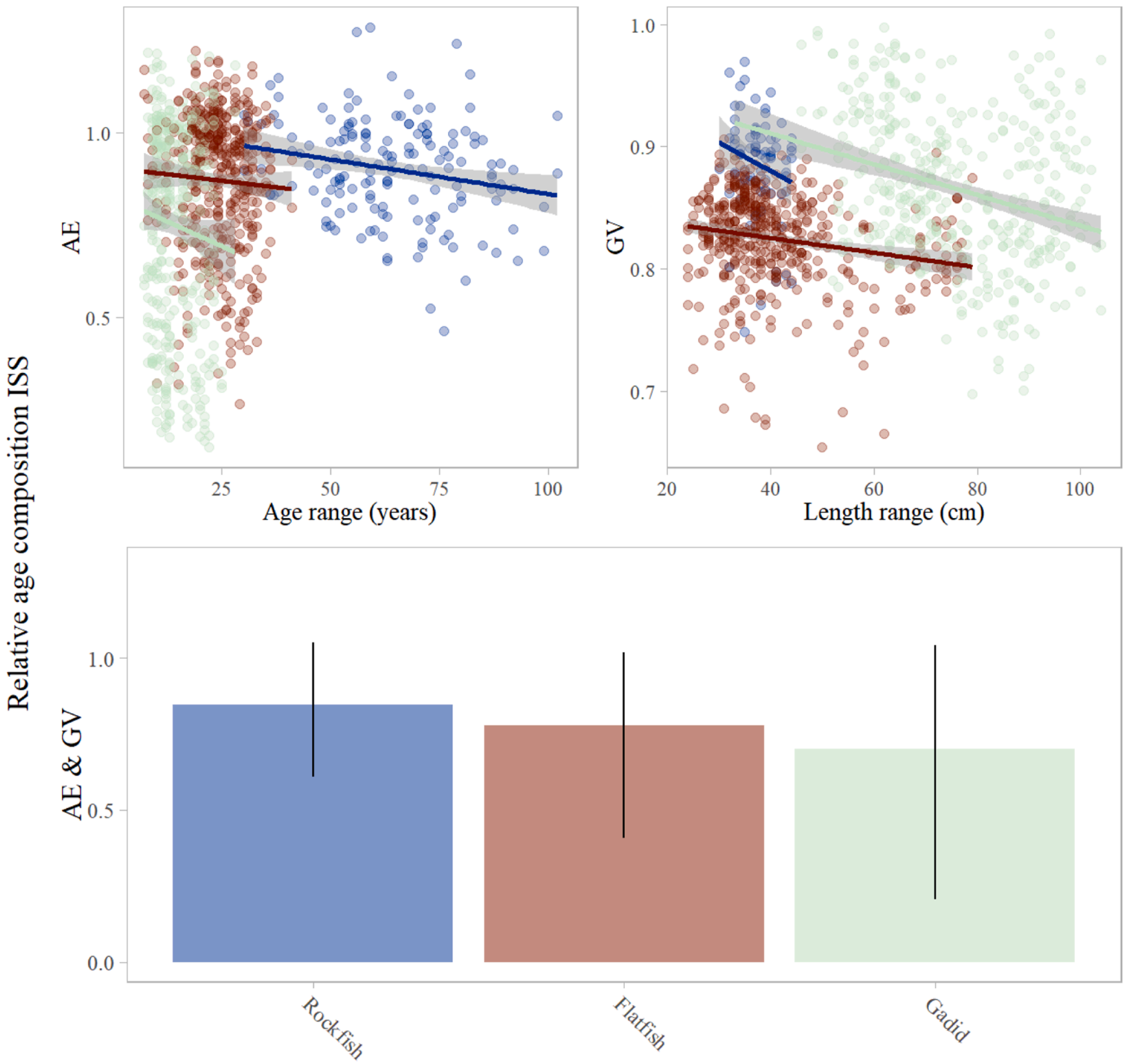

Fig. 9. Relative age composition input sample size when including ageing error ('AE') or growth variability ('GV') compared to longevity (as indicated by age range) and growth (as indicated by length range, top panels, points represent species-year combinations) and when including both ageing error and growth variability ('AE & GV', bottom panel, with the whiskers indicating the 95% confidence intervals) across the species types (colors) evaluated. For illustration of trends, linear relationships for each species type are shown in the top panels.

2023), as survey reductions may be inevitable in many regions due to declining budgets. A recent study investigated the reductions in length frequency and age collection effort, using AFSC bottom trawl survey (Hulson et al. in press) found that reduction in age collections had a larger effect on age composition uncertainty for flatfish and rockfish as compared to gadids. Here we find that including additional sources of uncertainty has a greater effect on gadids, and less impact on rockfish. It is potentially the case that the effect of decreases in sampling effort for gadids and flatfish would be smaller given the magnitude of effect by these sources of uncertainty as compared to rockfish. However, we acknowledge that this should be evaluated in future studies to understand the specific effects on stocks and when comparing among species types when including these additional sources of uncertainty in estimating age composition ISS.

5. Conclusions

Overall, we find that expanding upon the method introduced by Stewart and Hamel (2014) by including ageing error and growth variability into estimation of age composition ISS can have a large effect in reducing the magnitude of ISS. We provide two primary recommendations from this work. First, we recommend that stock assessment scientists consider the use of bootstrap methods to set age composition ISS, length composition ISS, and CAAL ISS. With modern computing power, for a single species using the R package we built (https://github.com/BenWilliams-NOAA/surveyISS), it takes less than an hour to obtain both age and length composition bootstrap ISS for a historical survey time series (longer than 40 years in some cases); for a single year it takes a matter of minutes. Second, for all estimates of age composition ISS and CAAL ISS we recommend implementing ageing error and growth variability to more explicitly and thoroughly take these sources of uncertainty into account in stock assessment models. We note that while we used fishery-independent data here as an example, these sources of uncertainty are also inherent to fishery-dependent data.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We thank Dan Goethel, Cole Monnahan, and two anonymous reviewers for their helpful reviews of this manuscript. We also thank all the AFSC survey staff who collected the data over the last 40 years used in this analysis. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contribution

Conceptualization: P-JFH Data curation: P-JFH Formal analysis: P-JFH Funding acquisition: N/A Investigation: P-JFH, BCW Methodology: P-JFH, BCW Project administration: P-JFH Resources: N/A Software: BCW, P-JFH Supervision: P-JFH Validation: P-JFH, BCW Visualization: P-JFH, BCW Writing – original draft: P-JFH, BCW Writing – review & editing: P-JFH, BCW.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.fishres.2023.106894.

References

Ailloud, L.E., 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 J. Mar. Sci. 76 (6), 1515–1523. https://doi.org/10.1093/icesjms/fsz072.

Barbeaux, S.J., Barnett, L., Connor, J., Nielson, J., Shotwell, S.K., Siddon, E., and Spies, I. 2022. 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 Islands. North Pacific Fishery Management Council, 1007 West 3rd Ave., Suite 400, L92 Building, 4th floor, Anchorage, AK 99501.

Candy, S.G., Nowara, G.B., Welsford, D., McKinlay, J.P., 2012. Estimating an ageing error matrix for Patagonian toothfish (*Dissostichus eleginoides*) otoliths using between-reader integer errors, readability scores, and continuation ratio models. Fish. Res. 115, 14–23. https://doi.org/10.1016/j.fishres.2011.11.005.

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/.

Francis, R.I.C., 2011. Data weighting in statistical fisheries stock assessment models. Estimating an ageing error matrix for Patagonian toothfish (Dissostichus eleginoides) otoliths using between-reader integer errors, readability scores, and continuation ratio models. Can. J. Fish. Aquat. Sci. 68, 1124–1138. https://doi.org/10.1139/F2011-025.

Hulson, P.-J.F., Williams, B., Bryan, M., Conner, J., and Siskey, M. in press. Reductions in sampling effort for fishery-independent age and length composition: balancing sampling efficiency, data uncertainty, and workforce health. Canadian Journal of Fisheries and Aquatic Sciences.

Hulson, P.-J.F., Hanselman, D.H., Quinn II, T.J., 2012. Determining effective sample size in integrated age-structured assessment models. ICES J. Mar. Sci. 69, 281–292. https://doi.org/10.1093/icesjms/fsr189.

Hulson, P.-J.F., Williams, B.C., Fissel, B.E., Ferriss, B.E., Hall, M., Yasumiishi, E.M., and Jones, D.T. 2021. Assessment of the Pacific ocean perch 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.

Hulson, P.-J.F., S.J. Barbeaux, B. Ferriss, S. McDermott, and I. Spies, 2022. Assessment of the Pacific cod stock in the Gulf of Alaska. In: Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. Anchorage, AK, North Pacific Fishery Management Council, 1007 West 3rd Ave., Suite 400, L92 Building, 4th floor, Anchorage, AK 99501.

Hulson, P.-J.F., Williams, B., Siskey, M., Bryan, M., Conner, J., 2023. Bottom trawl survey age and length composition input sample sizes for stocks assessed with statistical catch-at-age assessment models at the Alaska Fisheries Science Center. U. S. Department of Commerce. NOAA Tech. Memo. NMFS-AFSC-470 38.

ICES, 2020. Workshop on unavoidable survey effort reduction (WKUSER). ICES Sci. Rep. https://doi.org/10.17895/ices.pub.7453.

ICES, 2023. Workshop on unavoidable survey effort reduction 2 (WKUSER). ICES Sci. Rep. https://doi.org/10.17895/ices.pub.22086845.v1.

Kimura, D.K., Lyons, J.J., 1991. Between-reader bias and variability in the age-determination process. Fish. Bull., U. S 89, 53–60.

Lauth, R.R., Dawson, E.J., Conner, J., 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 Tech. Memo. NMFS-AFSC-396 260.

Liao, H., Sharov, A.F., Jones, C.M., Nelson, G.A., 2013. Quantifying the effects of aging bias in Atlantic striped bass stock assessment. Trans. Am. Fish. Soc. 142 (1), 193–207. https://doi.org/10.1080/00028487.2012.705255.

McAllister, M.K., Ianelli, J.N., 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Can. J. Fish. Aquat. Sci. 54 (2), 284–300. https://doi.org/10.1139/f96-285.

Miller, T.J., Skalski, J.R., 2006. Integrating design- and model-based inference to estimate length and age composition in North Pacific longline catches. Can. J. Fish. Aquat. Sci. 63 (5), 1092–1114. https://doi.org/10.1139/f06-022.

Monnahan, C.C., Dorn, M.W., Deary, A.L., Ferriss, B.E., Fissel, B.E., Honkalehto, T., Jones, D.T., Levine, M., Rogers, L., Shotwell, S.K., Tyrell, A., and Zador, S. 2021. Assessment of the walleye pollock 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.

Morison, A., Burnett, J., McCurdy, W., Moksness, E., 2005. Quality issues in the use of otoliths for fish age estimation. Mar. Freshw. Res. 56. https://doi.org/10.1071/MF04217.

Nesslage, G., Schueller, A.M., Rezek, A.R., Mroch III, R.M., 2022. Influence of sample size and number of age classes on characterization of ageing error in paired-age comparisons. Fish. Res. 249, 106255 https://doi.org/10.1016/j.fishres.2022.106255.

Pennington, M., Volstad, J.H., 1994. Assessing the effect of intra-haul correlation and variable density on estimates of population characteristics from marine surveys. Biometrics 50 (3), 725–732. https://doi.org/10.2307/2532786.

Pennington, M., Burmeister, L.M., Hjellvik, V., 2000. Assessing the precision of frequency distributions estimated from trawl-survey samples. Fish. Bull., U. S 100 (1), 74–80.

Punt, A.E., Smith, D.C., KrusicGolub, K., Robertson, S., 2008. Quantifying age-reading error for use in fisheries stock assessments, with application to species in Australia's southern and eastern scalefish and shark fishery. Can. J. Fish. Aquat. Sci. 65 (9), 1991–2005. https://doi.org/10.1139/F08-111.

Punt, A.E., Tuck, G.N., Day, J., Burch, P., Thomson, R.B., Bessell-Browne, P., 2021. The impact of alternative age-length sampling schemes on the performance of stock assessment methods. Fish. Res. 238. https://doi.org/10.1016/j.fishres.2021.105904.

Quinn, T., Deriso, R., 1999. Quantitative Fish Dynamics. Oxford University Press, New York, NY.

Siskey, M.R., Punt, A.E., Hulson, P.-J.F., Bryan, M.D., Ianelli, J.N., Thorson, J.T., 2023. The estimated impact of changes to otolith field-sampling and ageing effort on stock assessment inputs, outputs, and catch advice. Can. J. Fish. Aquat. Sci. 80 (1), 115–131. https://doi.org/10.1139/cjfas-2022-0050.

Spencer, P.D., and Ianelli, J.I. 2022. Assessment of the Pacific ocean perch stock in the Bering Sea/Aleutian Islands. 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.

Stewart, I.J., Hamel, O.S., 2014. Bootstrapping of sample sizes for length-or age-composition data used in stock assessments. Can. J. Fish. Aquat. Sci. 71 (4), 581–588. https://doi.org/10.1139/cjfas-2013-0289.

Stewart, I.J., Monnahan, C.C., 2017. Implications of process error in selectivity for approaches to weighting compositional data in fisheries stock assessments. Fish. Res. 192, 126–134. https://doi.org/10.1016/j.fishres.2016.06.018.

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.

Taylor, I.G., Methot, R.D., 2013. Hiding or dead? A computationally efficient model of selective fisheries mortality. Fish. Res. 142, 75–85. https://doi.org/10.1016/j.fishres.2012.08.021.

Thorson, J.T., Haltuch, M.A., 2019. Spatiotemporal analysis of compositional data: increased precision and improved workflow using model-based inputs to stock assessment. Can. J. Fish. Aquat. Sci. 76 (3), 401–414. https://doi.org/10.1139/cjfas-2018-0015.

Thorson, J.T., Johnson, K.F., Methot, R. D, K.F., Taylor, I. G, I.G., 2017. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. Fish. Res. 192, 84–93. https://doi.org/10.1016/j.fishres.2016.06.005.

Thorson, J.T., Monnahan, C.C., Hulson, P.-J.F., 2023. Data weighting: an iterative process linking surveys, data synthesis, and population models to evaluate mis-specification. Fish. Res.

Williams, B.C., Hulson, P.-J.F., Lunsford, C.R., and Ferriss, B. 2022. Assessment of the northern rockfish 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.

Xu, H., Thorson, J.T., Methot, R.D., 2020. Comparing the performance of three data-weighting methods when allowing for time-varying selectivity. Can. J. Fish. Aquat. Sci. 77 (2), 247–263. https://doi.org/10.1139/cjfas-2019-0107.