North American Journal of Fisheries Management 40:1007–1022, 2020

ISSN: 0275-5947 print / 1548-8675 online
DOI: 10.1002/nafm.10468

MANAGEMENT BRIEF

Assessing Accuracy and Bias of Protocols to Estimate Age of Pacific Salmon Using Scales

Julianne E. Harris*
U.S. Fish and Wildlife Service, Columbia River Fish and Wildlife Conservation Office, 1211 Southeast Cardinal Court, Suite 100, Vancouver, Washington 98683, USA

Abstract
Accurate estimates of age composition are essential for fisheries management, but evaluation of protocols to estimate age data are rarely conducted. Herein, known-age Pacific salmon *Oncorhynchus* spp. from six hatchery stocks were used to examine accuracy of multiple scale-aging protocols and to assess how errors in age assignment may affect age composition estimated from scale data. Accuracy reached 1.0 for one stock of Coho Salmon *Oncorhynchus kisutch*, over 0.9 for three stocks of spring Chinook Salmon *Oncorhynchus tshawytscha*, and was no higher than 0.8 for two stocks of fall Chinook Salmon. For all hatchery stocks, accuracy increased and bias decreased when readers had auxiliary information about a fish's fork length and sex. Results suggest that a protocol that collects scales from above the lateral line, provides auxiliary information on fork length and sex, allows readers to exclude individuals for which age appears unclear, and uses an arbitrator to assign age when two initial readers disagree may produce the most accurate and unbiased age data from scales. This study highlights the importance of assessing patterns in age errors both qualitatively and quantitatively. A general pattern of overestimating age for younger age-classes and underestimating age for older age-classes was evident. This pattern in age errors could affect age composition estimates for adult Pacific salmon and other species with few (e.g., two to six) age-classes. In addition, individual variation in demographic rates may affect individual patterns identified on scales; thus, for stocks with more extensive individual demographic variability, like fall Chinook Salmon stocks, accuracy may be lower and bias higher, resulting in less accurate estimates of age composition.

Identifying the age composition of a fish population or fishery is essential for proper stock assessment and management (Hilborn and Walters 1992; Campana 2001). One way to identify age composition is by examining growth increments on calcified structures, such as otoliths, fin rays, or scales, from individuals in the population (Campana 2001). Despite frequent use and application of age data from calcified structures in assessment and management, the probability that age is assigned correctly, the patterns in age errors, and the bias in estimated biological reference points are rarely assessed (Beamish and McFarlane 1983; Campana 2001; Zabel and Levin 2002). Age errors can arise when growth increments do not match the expected time intervals (i.e., process error) or growth increments are not interpreted correctly (i.e., observation error) due to the aging protocol or skill level of the biologist (i.e., reader) examining the structure for age (Campana 2001; Buckmeier 2002; Morison et al. 2005). Thus, it is essential to identify, evaluate, and adopt an aging protocol that is accurate, unbiased, and practical and to train biologists to use that protocol. Examining patterns in age errors and assessing how those errors may affect age composition estimates can improve the understanding of species and stock biology and the evaluation of demographic rates produced from the age data.

Scales have been used to produce age data for Pacific salmon *Oncorhynchus* spp. since the early 1900s (Gilbert 1912). Scale data are often used to estimate age composition for Pacific salmon because scales can be collected easily and nonlethally and there are protocols to interpret the growth patterns on scales to assign age; however, accuracy can vary among stocks and readers (McNicol and MacLellan 2010). Age errors often skew estimates of age composition towards dominant age-classes (Worthington et al. 1995); however, effects of age errors on age composition could vary based on the patterns in age errors and true age composition. The effects of patterns in

*Corresponding author: julianne_harris@fws.gov
Received October 4, 2019; accepted May 11, 2020

age errors may be particularly problematic for Pacific salmon collected as adults, since few age-classes are collected (often less than four) and one age-class can dominate a spawning run. Age composition data can contribute to estimates of exploitation, spawning potential, and future run size, as well as help evaluate effects of environmental factors on population persistence (Zabel and Levin 2002; Copeland et al. 2007). Thus, assessing and understanding the accuracy and bias of scale-aging techniques are vital for the use of scale data in management of Pacific salmon.

Since length and sex provide evidence about age for Pacific salmon, an aging protocol that allows readers to consider an individual's length and sex might reduce errors, improve accuracy, and reduce bias in age composition data. For Pacific salmon, demographics like age at maturity and growth are a function of both genetic and environmental factors (Hankin et al. 1993; Berejikian et al. 2011; Siegel et al. 2017). There is also individual variability in growth and observed stock-level changes in length at age and age at maturity over time (Ricker 1981; Morita et al. 2005; Lewis et al. 2015; Ohlberger et al. 2018). Thus, having auxiliary information about length and sex may also reduce the potential to identify changes in demographic rates or to link those changes to environmental factors or fishing pressure. Since demographic changes may be harder to identify, auxiliary information is rarely provided to readers as part of the scale-aging protocol. An inability to identify stock-level changes would be problematic for long-term management of Pacific salmon and other harvested species in a changing environment.

Pacific salmon are widely caught and harvested for economic, recreational, and cultural purposes throughout the Pacific Northwest and in the Columbia River basin specifically (Meengs and Lackey 2005; Bottom et al. 2009; Jordan et al. 2012). Historically, the Columbia River basin supported over 200 Pacific salmon stocks, but that number has declined dramatically (Johnson et al. 2019). In addition, many extant Pacific salmon stocks have declined in abundance from historic levels due to multiple complex and interconnected environmental and anthropogenic impacts, but they continue to support substantial fisheries, serve a vital ecological role in both marine and freshwater ecosystems, and are a focus of research and management efforts (Bottom et al 2009; Jordan et al. 2012). One response to declining populations has been for hatcheries to produce and release millions of juvenile Pacific salmon in the Columbia River basin each year for managed harvest as adults in future years. Some of these young Pacific salmon are tagged with either a passive integrated transponder (PIT) tag or a coded wire tag (CWT) for stock identification, harvest management, and research purposes and are thus of a known age when they return to the hatchery in the Columbia River basin to spawn as adults.

The primary purpose of this study was to use returning known-age Pacific salmon to assess multiple scale-aging protocols for six U.S. Fish and Wildlife Service (USFWS) National Fish Hatchery (hatchery) stocks in the Columbia River basin: three spring Chinook Salmon *Oncorhynchus tshawytscha*, two fall Chinook Salmon, and one Coho Salmon *O. kisutch*. The objectives of this study were to (1) assess the accuracy, precision, and bias of aging protocols, (2) examine the probability of accuracy as a function of multiple factors, and (3) identify and examine how errors in age-class assignment may affect age composition estimates using data from 1993–2016 as examples of variability in age composition. Factors we examined included fish stock, location of the scale on the fish's body, provision of auxiliary information on the fish's sex and fork length (mm), the fish's age, reader, and use of an arbitrator. Estimating age from scales is time consuming, and low accuracy and substantial bias can cause inaccurate stock assessment and poor management decisions. Scale age data from hatcheries are used with other data to forecast future run sizes of Pacific salmon for planning purposes and harvest allocations in the Columbia River basin and elsewhere (PFMC 2019). Evaluation techniques used herein and results of this study could inform standard operating procedures for scale aging of multiple species, particularly species of Pacific salmon, especially with regard to the provision of auxiliary information on fish length and sex. The results of this study also inform the use and value of scale age data in Pacific salmon stock assessment and management.

METHODS

Fish collection.— Scales were collected in 2016 from six hatchery stocks of known-age adult Pacific salmon: Eagle Creek–Coho Salmon (EC-CO), Carson–spring Chinook Salmon (C-SC), Warm Springs–spring Chinook Salmon (WS-SC), Little White Salmon–spring Chinook Salmon (LWS-SC), Spring Creek–fall Chinook Salmon (SC-FC), and Little White Salmon–fall Chinook Salmon (LWS-FC). Each fish was tagged with an individually numbered CWT prior to release as a juvenile from the hatchery, so the year of release and age at return was known; thus, all included fish were of known age. All returning adult hatchery fish were maintained at the hatchery alive in holding ponds during the spawning run before being collected from holding ponds, euthanized, and spawned (or surplused). Then all fish with a detected CWT were examined for fork length, sex, and any other marks, and then scales and snouts were collected. For this study, a random number generator was used to select 100 fish from the total number of tagged fish sampled from each hatchery stock.

Scale collection and preparation.— For each sampled fish, 20 scales were collected: 10 from each side of the fish, 5 above and 5 below the lateral line. Scales were collected using forceps from the area between the posterior margin of the dorsal fin and the anterior insertion of the anal fin, within the first three scale rows above or below the lateral line ("preferred area"). In the rare case in which the preferred area was damaged (i.e., scales not present in that area), scales were taken directly in front of or behind the preferred area. Dirt, blood, mucus or any other foreign matter was removed from the sampling area prior to removing scales. In the laboratory, scales were pressed onto acetate.

Readers.— Six readers examined scales for age estimation. The six readers (R1–R6) ranged in the number of years of experience reading scales of adult Pacific salmon from the stocks included in this study and other stocks. McNicol and MacLellan (2010) described readers with over 10 years of experience aging as having "deep experience" and those with less than 3 years as having "shallow experience." The readers R1 and R2 had deep experience (each >15 years) and were considered "experts," while R3 and R4 had "moderate experience" with 6–7 years of experience and R5 and R6 had shallow experience (1–3 years). All readers had experience assigning age to scales from multiple stocks and species of Pacific salmon and thus would be considered to have "broad experience," as described by McNicol and MacLellan (2010).

Age estimation protocols.— To evaluate different scale-aging protocols, each fish was assigned an age four different times (i.e., four protocols) by all six readers using an ST ViewScan III microfilm scanner with ST ViewScan III premium software (ST Imaging, Northbrook, Illinois). Readers knew the hatchery stock for each fish while examining its scales during all four protocols. Readers were provided each fish's sex and fork length as "auxiliary information" for two protocols. Between protocols, a non-reader used a random number generator to assign a new identification number to each fish to reduce the likelihood that readers would remember any particular fish's scales. The four protocols were as follows: (1) scales above the lateral line examined with no auxiliary information, (2) scales below the lateral line examined with no auxiliary information, (3) scales above the lateral line examined with auxiliary information, and (4) scales below the lateral line examined with auxiliary information. Readers examined scales without auxiliary information before examining with auxiliary information. Each reader assigned an age to each fish with at least one useable scale (i.e., clear enough scale impression and not regenerated) and recorded how confident the reader felt about the age assigned (1 = very confident, 2 = somewhat confident, 3 = not confident). The "confidence" metric was used to examine if provision of auxiliary information or correct age assignment influenced confidence (i.e., did readers feel more confident when they had more information and/or did they recognize individuals with scales that were more difficult to read, etc.). Readers did not communicate about results until all were finished examining all scales by all protocols.

Calculations and analyses.— Accuracy, precision, and bias were assessed for protocols, hatchery stocks, and readers (objective 1). Accuracy (i.e., "exactness") was calculated as the number of correctly aged fish divided by the total number of aged fish for that protocol, hatchery stock, and reader. A change in accuracy with an arbitrator was also examined. For this examination, pairs of readers and an arbitrator ($n = 3$ pairs) were selected using a random number generator (for each hatchery stock and protocol). When the two selected readers agreed on the fish's age, that age was assigned. Alternatively, when the two readers disagreed on age for an individual, the age assigned independently by the arbitrator was taken into account; if the arbitrator agreed with one reader, the age assigned by the arbitrator was assigned to the fish. If the arbitrator did not agree with either reader, the fish was removed from the estimate of accuracy. Arbitrators were "expert" readers (R1 and R2). Average accuracy of final ages, given an arbitrator, were calculated and compared to average accuracy of individual readers for each hatchery and protocol. Minimum "acceptable" accuracy was set at 0.9 (Buckmeier 2002).

To assess precision (i.e., "reproducibility") among readers for each hatchery stock and protocol, an index of average percent error (APE) and mean coefficient of variation (MCV) were calculated (Campana 2001). Reproducibility among readers can be important for consistency in large-scale age programs since different readers may examine scales in different years. Average percent error for each fish (APE_j) was calculated as follows:

$$APE_j = 100\% \cdot \frac{1}{R}\sum_{i=1}^{R}\frac{\left|X_{ij} - X_j\right|}{X_j},$$

where X_{ij} was the age assigned by reader i for fish j, X_j was the average age assigned for fish j, and R is the total number of times fish j was assigned an age. Average APE_j included all fish (i.e., APE) and was considered an index of precision for each hatchery stock and protocol. Coefficient of variation (CV_j) was calculated as follows:

$$CV_j = 100\% \cdot \frac{\sqrt{\sum_{i=1}^{R}\frac{\left(X_{ij}-X_j\right)^2}{R-1}}}{X_j}.$$

Again, the average of CV_j across all fish (i.e., MCV) for each hatchery stock and protocol was an assessment of precision. Precision was considered "acceptable" when APE was less than 5 and MCV was less than 7.6 (Campana 2001; McBride 2015).

Bias (i.e., systematic pattern in age errors) was assessed quantitatively and qualitatively for all hatchery stocks and protocols, including data from all readers combined. To assess bias quantitatively Bowker's unpooled test of symmetry was completed (Bowker 1948; Hoenig et al. 1995; McBride 2015). There are multiple tests of symmetry that can be used to examine bias in an age data set and they are not equivalent (McBride 2015). Bowker's unpooled model was chosen for this study because for most examined hatchery stocks, there were few age-classes of returning adults (two to four) and a pattern of overestimating age for younger age-classes and underestimating age for older age-classes was observed (see Results). This pattern in errors could result in no systematic underaging or overaging in the data set; thus, pooled tests of symmetry may suggest minimal bias, even when the estimated age distribution is biased. Bowker's unpooled test of symmetry was completed and the *P*-value was reported for each hatchery stock and protocol. In addition, Bowker's analysis was evaluated using a Bayesian approach. For the Bayesian approach, specific cell values (i.e., sums of incorrect ages) were first drawn from a Poisson distribution and then Bowker's test value was calculated, as described by Bowker (1948). Median values and 90% intervals for calculated Bowker test values were examined in relation to median and 90% values for a chi-square distribution with the number of degrees of freedom specified by Bowker (1948). For this bias assessment, 90% intervals were examined, as opposed to 95% intervals, since Bowker's test of symmetry is often examined as a one-sided test with an a priori alpha value of 0.05 and this allowed for comparison of the Bayesian approach to the standard approach. To examine patterns in bias qualitatively, Bland–Altman bubble plots (Bland and Altman 1986) were produced, as described by McBride (2015) for use with fish age data. Bland–Altman bubble plots simultaneously illustrate over- or underestimation in the full data set, as well as for all age-classes individually, in the same plot. Because sample size varied extensively by age-class, circle sizes in the Bland–Altman bubble plots were the proportion of fish in an age-class, as opposed to the actual number in that age-class; thus, patterns in errors were visible for age-classes that composed a smaller portion of the hatchery stock.

Probabilities of assigning the correct age (p_j) to a fish (j) were estimated as a function of multiple factors using binomial models (objective 2). First, logistic regression was conducted to estimate the probability of assigning the correct age to a fish (p_j) as a function of hatchery stock, scale location, auxiliary information about fork length and sex, and reader:

$$C_j \sim \text{binomial}(N_j, p_j),$$

where C_j was the number correctly assigned age and N_j was the total number assigned age. The logit link was used to examine p_j as a function of four categorical covariates: hatchery stock (h_j), scale location (i.e., above or below) in relation to the lateral line (l_j), auxiliary information about fork length and sex (I_j), and reader (r_j):

$$\text{Logit}(p_j) = \alpha + b_1(h_j) + b_2(l_j) + b_3(I_j) + b_4(r_j),$$

where $b_1 - b_4$ are slopes for the four categorical covariates and α represents the intercept. In the model, the probability of assigning the correct age was calculated (i.e., derived value) for the average reader by hatchery stock, protocol, and location relative to the lateral line.

To examine the effects of providing auxiliary information (i.e., fork length and sex) on the probability of assigning the correct age to a fish from its scales, two additional binomial analyses were conducted. First, variability among readers in the probability of assigning the correct age and in the effect of auxiliary information on the probability of assigning the correct age was examined. For this analysis, binomial models for each reader were developed with the logit link used to assess the probability of assigning the correct age as a function of auxiliary information. Probabilities with and without auxiliary information, and the change with auxiliary information (i.e., difference in the two values of p_j), were calculated for each reader. Second, the probability of assigning the correct age and effects of auxiliary information on that probability by age-class were assessed by hatchery stock. For each hatchery stock, the logit link was used to examine the probability of assigning the correct age as a function of age-class for protocols that included auxiliary information and for those that did not; the probability of correctly assigning age by age-class, with and without auxiliary information, was calculated for each hatchery stock.

All models in this study were analyzed using Bayesian methods with JAGS software (Plummer 2003) called from Program R (R Core Team 2013). Package jagsUI with function autojags was used (Kellner 2018) with three chains, an adaption of 1,000, a burn-in of 5,000, iteration interval of 5,000, and enough iterations saved to meet convergence, as assessed by all Rhat scores less than 1.1 for all estimated parameters (Gelman and Hill 2007; Kéry and Schaub 2012). The median of the posterior distribution of each estimated and calculated parameter was considered the expected value for the parameter. The 95% credible interval was used to describe variability, and substantial differences among parameters were defined by nonoverlapping 95% credible intervals (except bias analysis, see Methods). Prior distributions were selected to be uninformative: for binomial models, priors for intercept and slope values (e.g., α and $b_1 - b_2$) were normal with mean = 0 and standard deviation = 1,000; for Bowker analyses, priors for expected cell values were from a uniform distribution ranging from 0 to 200.

To examine if accuracy (i.e., exactness) and provision of auxiliary information on the fish's fork length and sex affected confidence in assigned age, confidence when age assignment was correct versus incorrect and with auxiliary information versus without auxiliary information was graphed. This qualitative assessment was conducted for each hatchery stock separately. To assess how auxiliary information affected confidence for each reader, average confidence for fish aged with auxiliary information was subtracted from the average value without auxiliary information. The change in the probability of assigning the correct age with auxiliary information (estimated by a binomial model, above) was graphically examined in relation to the average change in confidence with the inclusion of auxiliary information, for each reader.

Finally, errors by age-class assignment (due to scale reading errors) identified in this study were calculated to produce "correction proportions" and to examine how those errors may have affected the estimated age composition for four of six of the examined hatchery stocks from 1993 to 2016 (objective 3). The effects of age errors on estimated age composition are a function of the patterns in age errors and true age composition, which can vary among years. From 1993 to 2016, hatcheries used methods similar to protocol (3) to age scales to estimate the number of returning adults assigned to each age-class in a stock; thus, this analysis includes data from protocol (3) only.

First, by hatchery stock, age errors from this study were used to produce correction proportions for each assigned age-class. For each assigned age-class, correction proportions were calculated as proportions of individuals assigned to an age-class that were actually that age (i.e., assigned age = true age), as well as actually from any other true age in the hatchery (i.e., incorrectly assigned). For example, in this study, some SC-FC assigned age 3 were true age 3 but others were true age 2 and still others were true age 4. Correction proportions for SC-FC assigned age 3 were calculated as the proportions of fish assigned age 3 that were true age 2, true age 3, and true age 4 in this study.

Second, by hatchery stock, numbers of returning adults assigned to each age-class (1993–2016) and calculated correction proportions for each assigned age were used to estimate the numbers that belonged in each true age-class to produce a "corrected" age composition for each examined year. Correction proportions were used to estimate corrected age composition for each year for four of six hatchery stocks (C-SC, LWS-SC, SC-FC, and LWS-FC). Age composition was not corrected for EC-CO because there were no errors in age assignment using protocol (3). Age composition was also not corrected for WS-SC, since age composition there is estimated exclusively from returns of CWTs (i.e., no scales are used; thus, age composition is free from scale age errors).

After correcting age compositions, percent relative error (PRE_k; Vandergoot and Brendan 2015) by age-class (k) for each hatchery stock and year was calculated as follows:

$$\mathrm{PRE}_k = 100 \cdot [(Y_k - Z_k)/Z_k],$$

where Z_k is the corrected number of individuals of that age for that year and Y_k is the number assigned that age for that year using protocol (3). Within each hatchery stock, median PRE_k among years was considered the expectation of age-class bias and interquartile range of PRE_k among years was an assessment of variability in that bias among years.

This exercise to estimate and examine corrected age composition assumes the correction proportions calculated in this study for each hatchery stock (e.g., the proportion of fish assigned age 3 that were actually age 4) were appropriate for all years from 1993 to 2016, regardless of differences in readers, true age composition, demographic variability, environmental stochasticity, or any other factors that may change the probability of errors by age-class. Because assumptions are strict, results of this analysis should be evaluated with caution. The purpose of this analysis was not to identify age composition bias for any specific year but rather to generally examine how patterns in age errors may affect estimated age composition among years, over a realistic range in age compositions for the hatchery stock.

RESULTS

Scales from 100 individuals from each hatchery stock were collected. Due to a data recording error, one individual LWS-FC did not have a known age and was removed from the data set. Each reader assigned age (i.e., at least one of the five scales was usable) to at least 94 fish in each hatchery stock using each protocol.

Accuracy (i.e., exactness) was highest for EC-CO, followed by the three spring Chinook Salmon stocks, and finally, by the two fall Chinook Salmon stocks (Table 1). Average accuracy and both precision measures were considered "acceptable" (i.e., accuracy ≥ 0.90, APE < 5.0, and MCV < 7.6) for EC-CO and C-SC by all protocols and for WS-SC and LWS-SC when auxiliary information was available on each individual's sex and fork length. Average accuracy and estimates of precision were not considered adequate for either fall Chinook Salmon stock, regardless of protocol (Table 1). Average accuracy with use of an arbitrator when the pair of randomly selected readers disagreed on age was equivalent or higher to the average accuracy of one reader alone for all hatchery stocks and protocols (Table 1).

Results from analyses of Bowker's (1948) test of symmetry suggest differences in bias among hatchery stocks and protocols (Table 2). Calculated P-values were below

0.05 when median estimates from the Bayesian analysis did not fall within the 90% credible interval for the chi-square test (Table 2). For all four protocols, bias was minimal for EC-CO and substantial for LWS-FC (Table 2). For all hatchery stocks, bias was lower when auxiliary information on fork length and sex were provided to readers as compared with protocols without provision of auxiliary information (Table 2). Bias was lowest (expect for LWS-SC for which it was similar to the lowest) and not substantial according to Bayesian 90% credible intervals (except for LWS-FC) when auxiliary information was provided and scales were collected from above the lateral line. Bland–Altman bubble plots suggest minimal bias for EC-CO regardless of protocol (Figure 1). For all other hatchery stocks, mean over- or underestimation in the full data set was minimal, suggesting no systematic overestimation or underestimation of age; however, a pattern of overestimating age for younger age-classes and underestimating age for older age-classes was indicated (Figure 1). The pattern in age errors appeared less dramatic for protocols in which auxiliary information on fork length and sex was provided as compared with those without it (Figure 1).

Initial binomial model results suggest that the probability of assigning the correct age varied by hatchery stock, reader, and protocol (Figure 2). Within protocols, 95% credible intervals for the probability of assigning age

TABLE 1. Calculations of accuracy and precision by protocol, hatchery stock, and reader. Accuracy is shown by reader (R1–R6, ordered from most to least experienced), and also given is average reader accuracy (Ave) and average accuracy with an arbitrator (Arbitrator). Precision is measured as average percent error (APE) and mean coefficient of variation (MCV). Protocols are as follows: (1) scales above the lateral line without auxiliary information, (2) scales below the lateral line without auxiliary information, (3) scales above the lateral line with auxiliary information, and (4) scales below the lateral line with auxiliary information. Hatchery stocks are as follows: Eagle Creek–Coho Salmon (EC-CO), Carson–spring Chinook Salmon (C-SC), Warm Springs–spring Chinook Salmon (WS-SC), Little White Salmon River–Spring Chinook Salmon (LWS-SC), Spring Creek–fall Chinook Salmon (SC-FC), and Little White Salmon River–fall Chinook Salmon (LWS-FC).

	Accuracy								Precision	
Protocol and hatchery stock	R1	R2	R3	R4	R5	R6	Ave	Arbitrator	APE (%)	MCV (%)
Protocol 1										
EC-CO	1.00	1.00	1.00	0.98	1.00	1.00	1.00	1.00	0.20	0.29
C-SC	0.99	0.92	0.90	0.89	0.94	0.93	0.93	0.99	2.65	3.70
WS-SC	0.93	0.83	0.96	0.79	0.92	0.90	0.89	0.92	2.75	3.84
LWS-SC	0.95	0.86	0.85	0.88	0.89	0.73	0.86	0.92	3.79	5.13
SC-FC	0.80	0.73	0.71	0.48	0.68	0.78	0.70	0.80	9.49	12.13
LWS-FC	0.76	0.68	0.67	0.53	0.67	0.62	0.65	0.72	7.37	9.98
Protocol 2										
EC-CO	1.00	1.00	0.99	0.88	1.00	0.98	0.98	1.00	1.53	2.18
C-SC	0.99	0.93	0.88	0.85	0.89	0.96	0.92	0.98	2.79	3.66
WS-SC	0.95	0.82	0.97	0.82	0.94	0.92	0.90	0.94	2.32	3.31
LWS-SC	0.74	0.71	0.76	0.79	0.74	0.60	0.72	0.74	2.96	3.93
SC-FC	0.61	0.42	0.70	0.66	0.67	0.82	0.65	0.69	10.98	13.90
LWS-FC	0.64	0.71	0.70	0.32	0.65	0.59	0.60	0.74	9.18	12.70
Protocol 3										
EC-CO	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00
C-SC	0.99	0.96	0.98	0.97	0.91	0.98	0.96	0.98	1.15	1.54
WS-SC	0.94	0.96	0.98	0.98	0.97	0.95	0.96	0.96	0.64	0.88
LWS-SC	0.95	0.93	0.94	0.92	0.92	0.89	0.92	0.94	1.34	1.78
SC-FC	0.74	0.79	0.70	0.86	0.81	0.85	0.79	0.86	7.14	9.19
LWS-FC	0.66	0.82	0.80	0.82	0.66	0.63	0.73	0.79	5.44	6.83
Protocol 4										
EC-CO	1.00	1.00	1.00	0.98	1.00	0.99	0.99	1.00	0.35	0.52
C-SC	0.99	1.00	0.99	0.95	0.83	0.98	0.96	0.99	1.44	1.97
WS-SC	0.94	0.94	0.95	0.97	0.95	0.92	0.95	0.96	0.96	1.32
LWS-SC	0.95	0.93	0.94	0.94	0.88	0.92	0.93	0.95	1.32	1.84
SC-FC	0.82	0.75	0.73	0.89	0.75	0.85	0.80	0.87	6.30	8.37
LWS-FC	0.61	0.80	0.86	0.82	0.69	0.60	0.73	0.78	5.15	6.73

correctly overlapped only for C-SC and WS-SC, suggesting differences among stocks. The probability of correctly assigning age was lower when scales were collected below the lateral line as compared with above (b_2 [logit scale]: –0.21; 95% credible interval: –0.31 to –0.11) and when auxiliary information was not provided as compared with when it was provided (b_3 [logit scale]: –0.71; 95% credible interval: –0.82 to –0.60). The probability of assigning the correct age varied among readers (i.e., 95% credible intervals on reader slopes did not overlap for all readers), although there was no obvious pattern to suggest the probability increased with reader experience (see details below).

Auxiliary information on fork length and sex improved the probability of assigning the correct age, but the extent of improvement was not equivalent among readers (Figure 3). The expected probability was higher with auxiliary information for all readers; however, for R1 (expert) and R6 (shallow experience), 95% credible intervals overlapped suggesting that the improvement with auxiliary information was not substantial for those two readers (Figure 3A). The biggest improvements in the probability of assigning the correct age with auxiliary information occurred for R4 (moderate experience) and R2 (expert); those two readers had the lowest probabilities of assigning the correct age without auxiliary information and thus the most opportunity for improvement (Figure 3).

TABLE 2. Results of bias analyses by protocol and hatchery stock, including the median and 90% credible interval (in parentheses) for a chi-square distribution (χ^2) with degrees of freedom suggested for Bowker's test of symmetry, the median and 90% credible interval (in parentheses) for a Bayesian application of Bowker's test of symmetry (Bowker's χ^2), and *P*-values for the standard Bowker's test of symmetry. See Table 1 for protocol descriptions and hatchery stock abbreviations. An asterisk indicates a statistically significant value.

Protocol and hatchery stock	χ^2	Bowker's χ^2	*P*-value
Protocol 1			
EC-CO	0.5 (0.0–3.8)	0.5 (0.0–3.6)	0.16
C-SC	1.4 (0.1–5.9)	21.5 (10.9–34.0)	<0.01*
WS-SC	1.4 (0.1–6.0)	15.0 (7.3–25.4)	<0.01*
LWS-SC	1.4 (0.1–5.8)	21.1 (9.7–35.9)	<0.01*
SC-FC	3.4 (0.7–9.6)	19.2 (9.5–31.6)	<0.01*
LWS-FC	5.3 (1.7–12.5)	73.7 (55.6–93.2)	<0.01*
Protocol 2			
EC-CO	0.5 (0.0–3.9)	1.9 (0.0–8.0)	0.07
C-SC	1.4 (0.1–6.1)	23.2 (12.0–35.9)	<0.01*
WS-SC	1.4 (0.1–6.1)	10.7 (4.2–18.9)	<0.01*
LWS-SC	2.4 (0.4–7.8)	9.6 (2.7–19.9)	0.03*
SC-FC	4.4 (1.1–11.1)	57.7 (41.0–76.8)	<0.01*
LWS-FC	7.3 (2.7–15.4)	99.5 (79.7–120.4)	<0.01*
Protocol 3			
EC-CO			
C-SC	0.5 (0.0–4.0)	3.5 (0.1–11.2)	0.05
WS-SC	1.4 (0.1–6.0)	2.7 (0.3–8.3)	0.25
LWS-SC	2.3 (0.3–7.7)	2.5 (0.4–7.4)	0.77
SC-FC	2.4 (0.4–7.8)	11.5 (4.6–21.3)	0.01*
LWS-FC	4.4 (1.1–11.0)	53.0 (38.7–69.2)	<0.01*
Protocol 4			
EC-CO	0.4 (0.0–3.8)	0.5 (0.0–3.6)	0.56
C-SC	0.4 (0.0–3.8)	19.6 (10.4–29.8)	<0.01*
WS-SC	1.4 (0.1–6.0)	5.6 (1.1–13.4)	0.06
LWS-SC	1.4 (0.1–6.0)	2.0 (0.2–7.8)	0.64
SC-FC	2.3 (0.3–7.7)	13.0 (5.3–23.7)	<0.01*
LWS-FC	4.4 (1.2– 10.9)	55.1 (40.4–70.7)	<0.01*

The probability of assigning age correctly was not equivalent among age-classes, and patterns were not similar among all hatchery stocks (Figure 4). However, the use of auxiliary information increased the probability of assigning the correct age for all age-classes for all hatchery stocks, although not always substantially (Figure 4). For example, for LWS-SC the probability of assigning the correct age was generally lowest for the youngest age-class and highest for the middle age-class, which is the dominate age-class in the stock. In contrast, for LWS-FC the probability of assigning the correct age was generally higher for younger age-classes as compared with older age-classes. For EC-CO, both age-2 and age-3 fish were assigned age with a high probability of accuracy, regardless of whether auxiliary information was provided or not (Figure 4).

Confidence in assigned age increased with both accuracy and auxiliary information on sex and fork length. More specifically, confidence was higher in hatchery stocks that could be aged with higher overall accuracy, during protocols when auxiliary information was available, and when the reader assigned the correct age to the individual (Figure 5). The improvement in accuracy associated with having auxiliary information appeared to be generally positively associated with an improvement in confidence across readers (Figure 3B).

The number and proportion of individuals assigned to each age-class that were from that age-class and from any other potential age-classes (i.e., correction proportions) varied among examined hatchery stocks (Table 3). For all hatchery stocks, there was annual variability in "corrected" age composition (Figure 6). In general, median PRE_k was closer to zero (i.e., less biased overall) and the interquartile range of PRE_k was smaller (i.e., lower variability in bias among years) for age-classes that composed a higher proportion of a stock as compared with age-classes that were relatively rare (Figure 6). However, there were substantial differences among hatchery stocks and age-classes in terms of direction and magnitude of the bias (i.e., extent of overestimation or underestimation for that age-class), as well as annual variability in bias (Figure 6). Notably, 3-year-old fish from C-SC did not make up a large proportion of the hatchery stock but were all aged correctly, illustrating no bias or imprecision in the

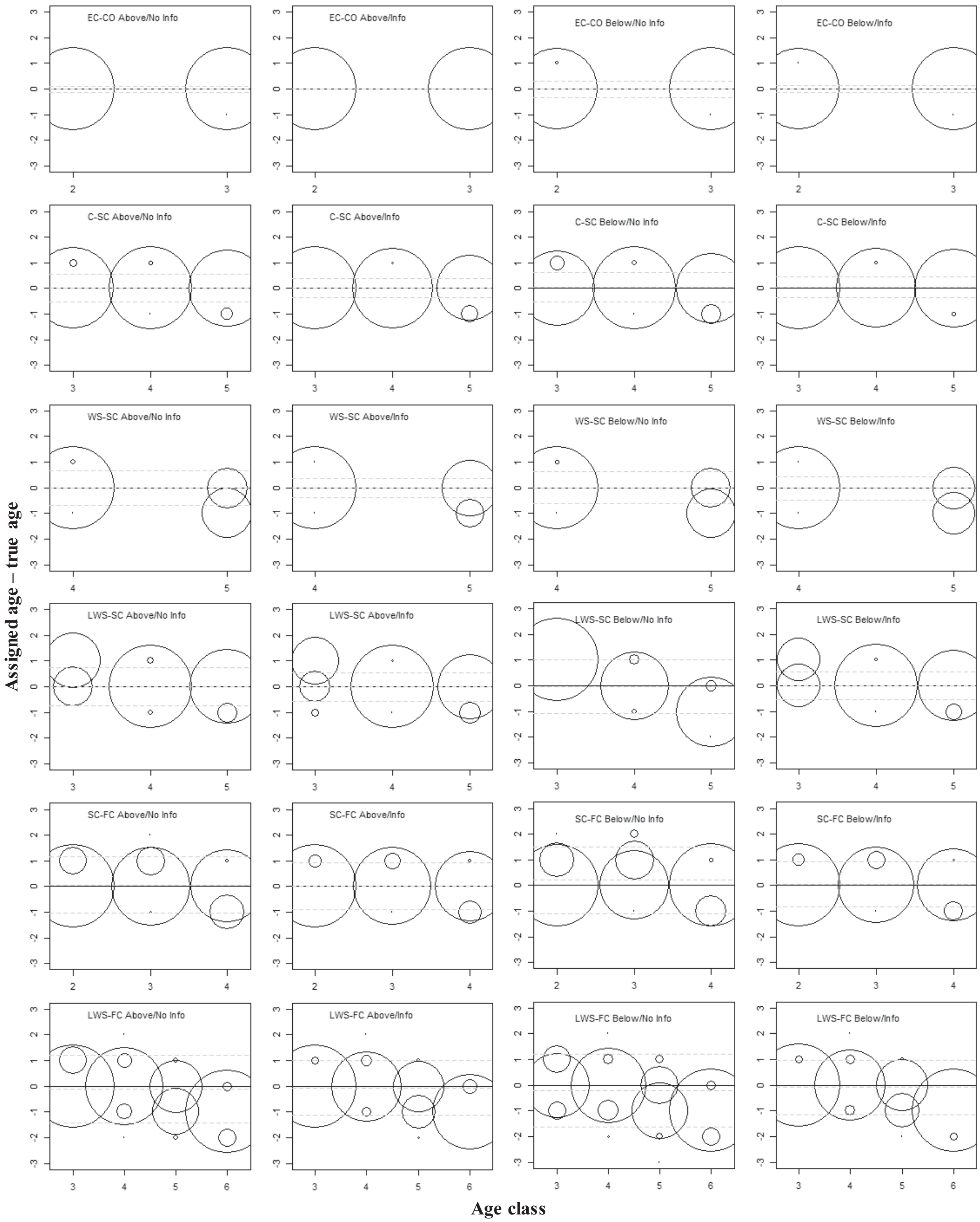

FIGURE 1. Bland–Altman bubble plots by hatchery stock and protocol (all readers combined). For each panel, the straight solid black line indicates perfect agreement between assigned age and true age. The middle dashed gray line indicates mean disagreement (flanked by dashed gray lines indicating two standard deviations from the mean) between assigned age and true age. These straight lines illustrate systematic over- or underestimation in the full data set. Circles indicate proportions of fish in each age-class with a specific assigned age minus true age calculation. Circle size illustrates any systematic over- or underestimation in age, by age-class. See Table 1 for hatchery stock abbreviations and protocol descriptions.

estimated number of fish for that age-class; however, the number of age-5 fish was generally overestimated (Figure 6). For all LWS-SC age-classes, median PRE_k was close to zero ($\leq|7.3|$), illustrating minimal bias overall, but interquartile range was large for age-3 and age-5 fish, illustrating that overestimating or underestimating the number of fish in these two age-classes was not uncommon (Figure 6). For the two fall Chinook Salmon stocks, especially for LWS-FC, expected bias and annual variability in bias were substantial for all age-classes, suggesting systematic overestimation of some age-classes and underestimation of others but also extensive variability in the extent of that overestimation or underestimation among years (Figure 6).

DISCUSSION

For many fish species, age data from scales are used in stock assessment and fisheries management because scales can be collected easily and nonlethally and can provide an accurate understanding of age composition (DeVries and Frie 1996; Campana 2001). The release of large numbers of coded-wire-tagged juvenile Pacific salmon from hatcheries provided an opportunity to assess, evaluate, and potentially improve the accuracy of age composition data produced from scales because at recapture the returning adults were of known age, something not usually available for wild fish. Results of this study suggest the highest accuracy and lowest bias for a protocol that (1) provides readers with auxiliary information on fork length and sex, (2) uses scales from the standard location above the lateral line, (3) removes individuals for which age appears unclear, and (4) incorporates an independent arbitrator when two initial readers disagreed on age. Inclusion of an arbitrator improved accuracy in two ways. First, if one reader was correct, the arbitrator tended (although not always) to agree with that reader. Second, when errors were of more than 1 year, the arbitrator would sometimes not agree with either reader or would suggest that there were no usable scales, thus removing that fish from assessment of accuracy. For species in which errors are often by more than 1 year, the arbitrator may have a higher probability of not agreeing with either reader and thus removing individuals that are more difficult to age correctly. Results from this study could help biologists improve protocols for aging scales; however, it is important to consider the biology of the species and stock, evaluate how potential demographic changes in the stock overtime may be detected, and examine how errors in age assignment may affect age composition data used in assessment and management.

The accuracy of scale aging and patterns in age errors varied among hatchery stocks, with accuracy generally lower and bias generally higher for species and stocks with more variability in age at maturity. All readers assigned age to Coho Salmon from the Eagle Creek hatchery with minimal errors. At Eagle Creek and elsewhere in United States, the vast majority of returning female Coho Salmon are age 3 and males are age 2 (i.e., jacks) or age 3 (Sandercock 1991; Kodama et al. 2012; Smith et al. 2015). Readers suggested that distinguishing among the two age-classes was easy with at least one usable scale, and confidence in assigned age was high. Sethi et al. (2017) found that juvenile Coho Salmon could be accurately aged with data on fork length alone. Fork length distributions of age-2 and age-3 Coho Salmon at Eagle Creek overlap only minimally, suggesting the potential to use fork length

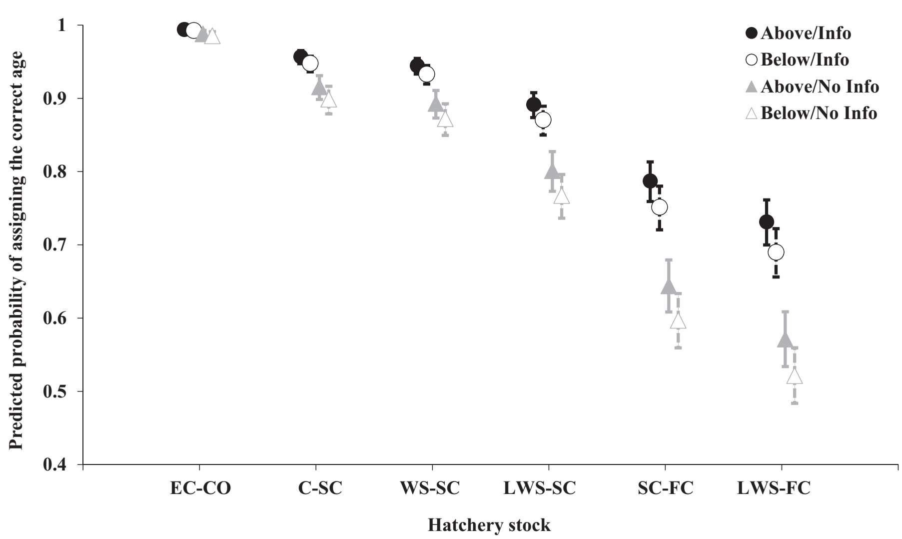

FIGURE 2. Predicted probabilities of assigning the correct age by scale location (i.e., “above” or “below” the lateral line) and by auxiliary information on fork length and sex of the fish (i.e., “info” or “no info”). Predicted probabilities were for the average reader. The error bars show 95% credible intervals. See Table 1 for hatchery stock abbreviations.

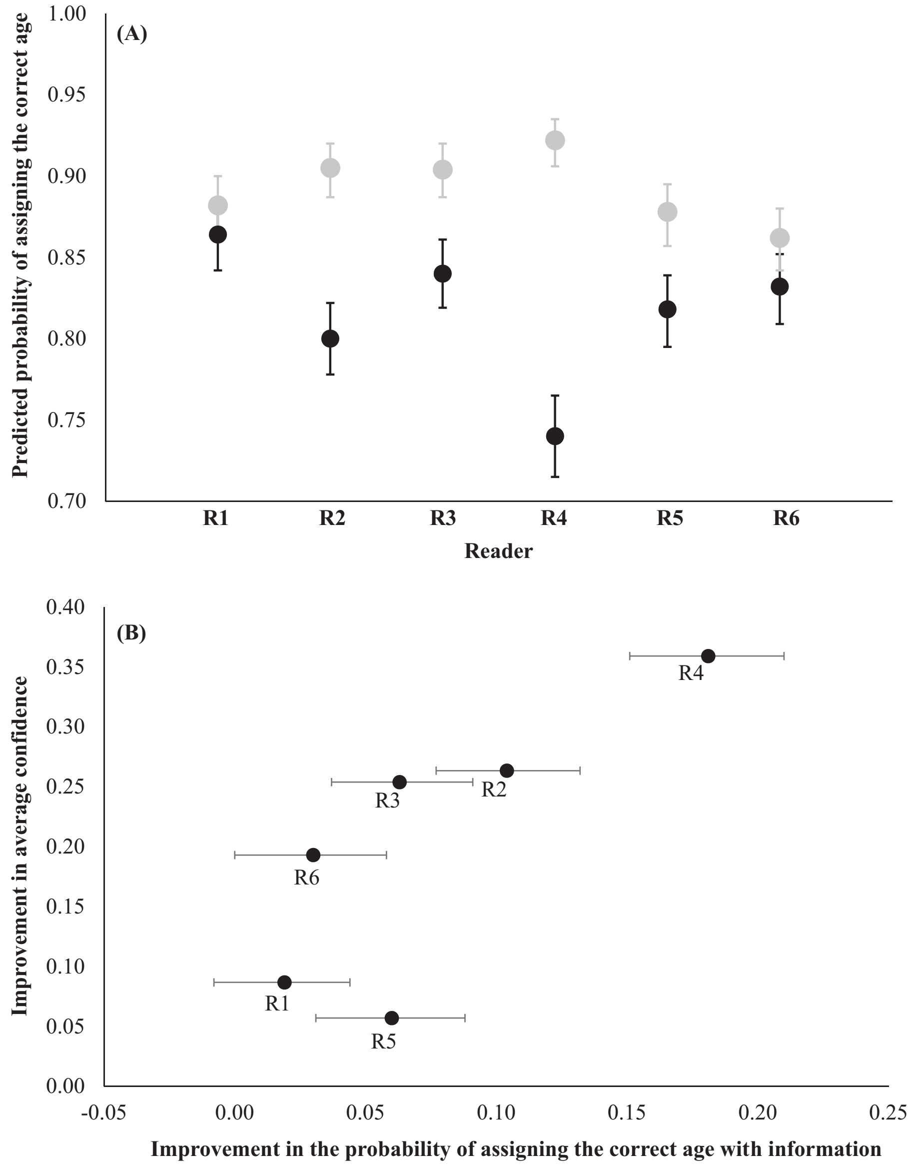

FIGURE 3. Panel **(A)** shows the predicted probabilities of assigning the correct age without (black) and with (gray) information, by reader. Estimates include all hatchery stocks examined. Panel **(B)** shows the improvement in the probability of assigning the correct age with provision of auxiliary information on fish fork length and sex, related to average improvement in confidence with provision of auxiliary information, by reader. The error bars show 95% credible intervals. Readers ("R") are ordered (i.e., R1–R6) from most (>15 years) to the least (<2 years) experienced.

to age adult Coho Salmon. If a protocol to produce age composition data for adult Coho Salmon using fork length could be developed and verified, it might improve sampling efficiency. In contrast, accuracy was the lowest and bias the highest for hatchery stocks of fall Chinook Salmon. Variability in age at maturity was highest for fall Chinook Salmon stocks, and fork length distributions among returning fish of different ages overlapped. The lowest accuracy observed was for fall Chinook Salmon from the Little White Salmon River. This stock had the most observed adult age-classes and the oldest fish (i.e., some age 6). Consideration of biological differences among species and stocks may be essential when comparing results among aging studies since stocks with high variability in demographic rates may have scales that are more difficult to correctly age.

Differences in the accuracy of scale aging among stocks of Chinook Salmon may be due to differences in the environmental conditions experienced by the fish. Branigan et al. (2019) observed variability in relative accuracy of scales, otoliths, and fin rays to age kokanee *Oncorhynchus nerka* (lacustrine Sockeye Salmon) of different stocks in

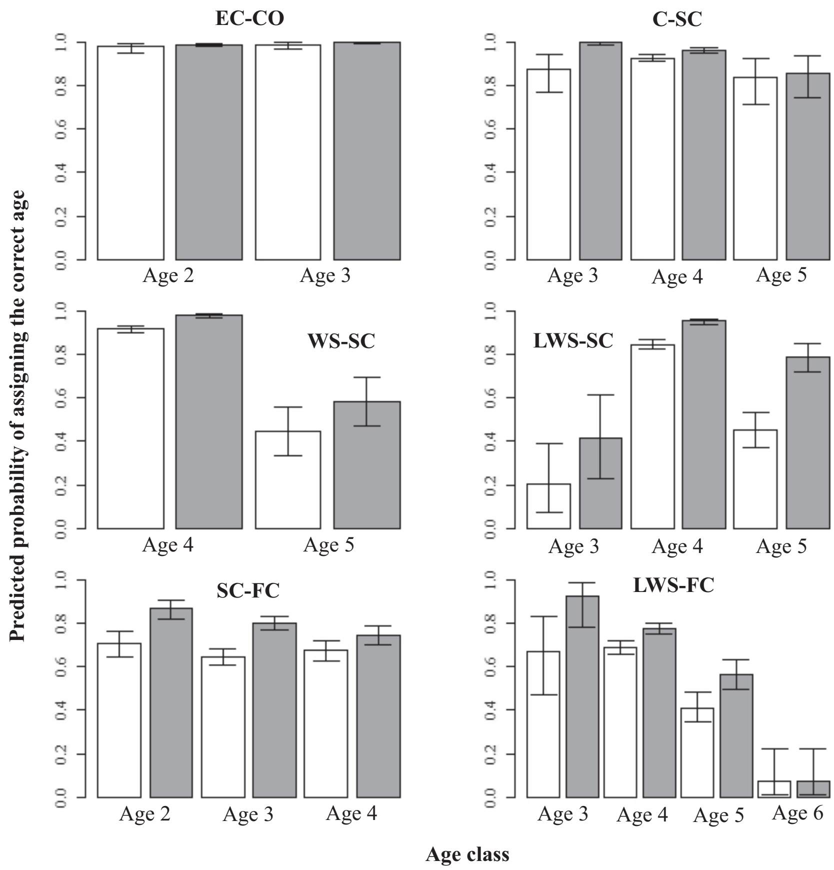

FIGURE 4. Probability of accurately assigning age (in years) by true age for each hatchery stock without (white bars) and with (gray bars) auxiliary information on fish fork length and sex. See Table 1 for hatchery stock abbreviations.

Idaho. They suggested that local environmental conditions, such as water chemistry and temperature, affected accuracy by altering the production of growth increments or the ability to evaluate them appropriately on the different aging structures (Branigan et al. 2019). In this study, accuracy of scales differed among hatchery stocks, potentially due in part to differences in local environmental conditions and differences in fish migratory behaviors. Variation in accuracy between spring and fall Chinook Salmon stocks may be due in part to differences in growth and migration that affect scale features, such as the timing of emigration to the ocean as juveniles (Koo and Isarankura 1967). Spring Chinook Salmon generally emigrate as yearlings (i.e., first annulus formed in freshwater), whereas fall Chinook Salmon generally emigrate as subyearlings (i.e., first annulus formed in marine water; Gilbert 1912; Healey 1991), and this difference can be detected by features on the scales (Koo and Isarankura 1967). However, due to the multiple dams in the Columbia River basin, some Snake River fall Chinook Salmon now overwinter in reservoirs and emigrate as yearlings, and this migration timing difference can be observed by patterns on scales (Conner et al. 2005). In addition, spring Chinook Salmon return as adults to the Columbia River basin in April–May and spawn in August, whereas fall Chinook Salmon enter in August–September and spawn September–October, and this seasonal difference in migration may affect growth increments on scales or the actual condition of scales. Because spatial and temporal differences in migration patterns and timing (freshwater and marine) can affect demographic rates and chemical deposition on scales, these differences among stocks could also affect the accuracy of ages assigned by scales. Thus, understanding species and stock biology, as well as the timing of scale collection (e.g., at spawning, at river entry, etc.), may be important.

Providing auxiliary information on fork length and sex increased accuracy, confidence, and reproducibility among readers, as well as reduced bias in age estimation. There is a strong, positive association between age and length for Pacific salmon and other species; thus, an individual's fork length can provide some indication of its age, especially for younger fish. An individual's sex can also be helpful for aging Pacific salmon since males exhibit higher

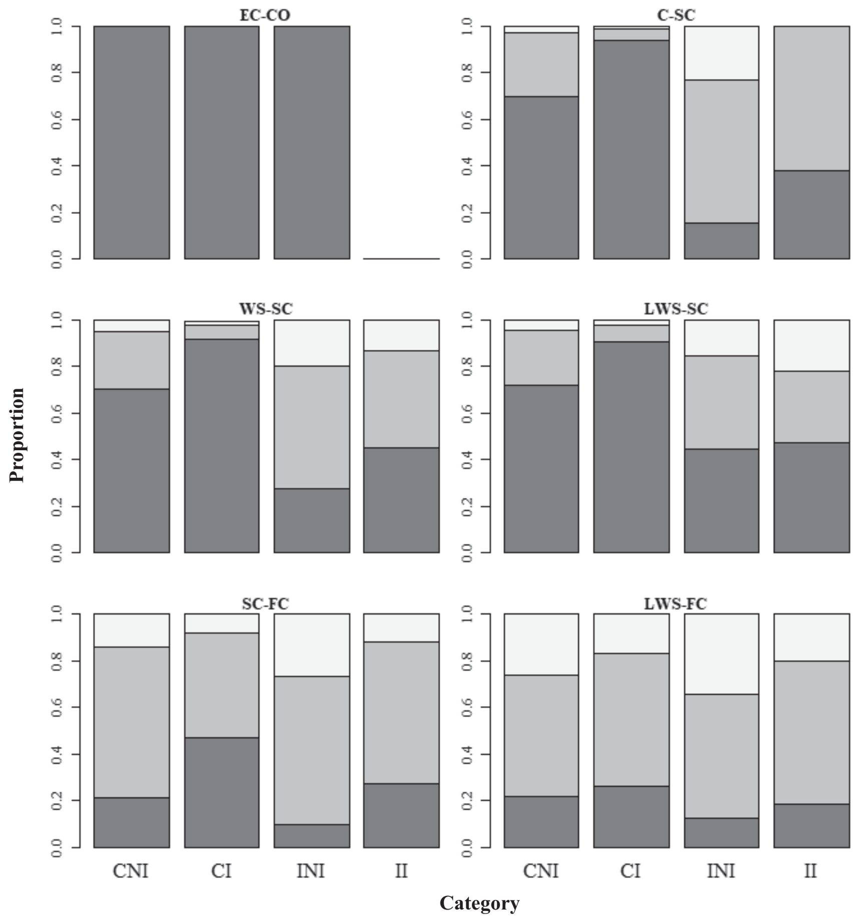

FIGURE 5. Reader confidence in the age assigned by accuracy and auxiliary information on fish sex and fork length for each hatchery stock. Confidence regarding each age was scored by each reader as "very confident" (dark gray), "moderately confident" (medium gray), and "not confident" (light gray). Confidence proportions are averages of all readers. Categories are abbreviated as follows: CNI = correct age without auxiliary information; CI = correct age with auxiliary information, INI = incorrect age without auxiliary information, and II = incorrect age with auxiliary information. See Table 1 for hatchery stock abbreviations. Note that no fish were incorrectly assigned age with auxiliary information for the EC-CO stock (i.e., no occurrences of II).

diversity in reproductive phenotype (i.e., jacks and hooknose; Gross 1985; Berejikian et al. 2011), which can be identified by features on scales (Johnson et al. 2012). However, changes in demographic rates, such as age at maturity and length at age, occur for Pacific salmon stocks (Ricker 1981; Lewis et al. 2015; Ohlberger et al. 2018). Identifying changes in size and age composition can help assess potential impacts of multiple stressors, such as climate change, fishing pressure, and hatcheries, on Pacific salmon stocks (Ohlberger et al. 2018; Cline et al. 2019). Results from this study suggest that auxiliary information on fork length and sex can improve accuracy and reproducibility and reduce bias but can also increase reader confidence, even when the reader is incorrect. Thus, inclusion of auxiliary information could result in reduced diligence in assessment of errors over time and reduce the potential to differentiate changes in length at age and age at maturity over time, both of which are important for stock assessment and management. Thus, if auxiliary information on fork length and sex are provided to readers, periodic validation studies (like this one) over time would likely be beneficial to examine for shifts in demographic rates. Changes in accuracy (i.e., Figure 3) and discussion with readers suggested that some readers relied on auxiliary information more than others did. Training to reduce reliance on auxiliary information and focus on scale features could simultaneously improve accuracy and reproducibility (e.g., in case readers vary among years), reduce bias, and increase the potential to detect demographic changes.

This study suggests the importance of examining bias (i.e., systematic pattern in age errors). In this study, both a qualitative and a quantitative assessment of bias were completed and they complemented each other. The

TABLE 3. Assigned age (total number assigned each age in parentheses) by true age for all readers combined and calculated "correction proportions" by hatchery stock (see Table 1 for hatchery stock abbreviations). Correction proportions are proportions of an assigned age (i.e., age assigned by a reader) that are in any true age (i.e., age known from coded wire tag). Correction proportions for each assigned age sum to 1.00 (i.e., include all individuals assigned that age).

Hatchery stock and assigned age	True age				
	2	3	4	5	6
C-SC					
3 ($n = 24$)		1 ($n = 24$)	0	0	
4 ($n = 530$)		0	0.99 ($n = 524$)	0.01 ($n = 6$)	
5 ($n = 38$)		0	0.40 ($n = 15$)	0.60 ($n = 23$)	
LWS-SC					
2 ($n = 1$)	0	1 ($n = 1$)	0	0	
3 ($n = 10$)	0	0.40 ($n = 4$)	0.60 ($n = 6$)	0	
4 ($n = 517$)	0	0.01 ($n = 6$)	0.95 ($n = 494$)	0.03 ($n = 17$)	
5 ($n = 66$)	0	0	0.23 ($n = 15$)	0.77 ($n = 51$)	
SC-FC					
2 ($n = 98$)	0.95 ($n = 93$)	0.05 ($n = 5$)	0	0	
3 ($n = 288$)	0.05 ($n = 15$)	0.78 ($n = 224$)	0.17 ($n = 49$)	0	
4 ($n = 194$)	0	0.23 ($n = 47$)	0.77 ($n = 147$)	0	
5 ($n = 7$)	0	0	1 ($n = 7$)	0	
LWS-FC					
3 ($n = 61$)		0.18 ($n = 11$)	0.77 ($n = 47$)	0.05 ($n = 3$)	0
4 ($n = 401$)		<0.01 ($n = 1$)	0.90 ($n = 362$)	0.09 ($n = 38$)	0
5 ($n = 124$)		0	0.46 ($n = 57$)	0.46 ($n = 57$)	0.08 ($n = 10$)
6 ($n = 7$)		0	0.14 ($n = 1$)	0.57 ($n = 4$)	0.29 ($n = 2$)

qualitative assessment of bias through a Bland–Altman bubble plot (Bland and Altman 1986) illustrated minimal overestimation or underestimation of age but showed patterns in age errors for individual age-classes for multiple hatchery stocks. Examination of specific patterns in the Bland–Altman bubble plots resulted in selection of Bowker's (1948) test of symmetry because examination of errors suggested overestimation of younger age-classes and underestimation of older age-classes, which could cause bias in age composition but not be detected by pooled tests of symmetry, which are recommended more generally (McBride 2015). When used, quantitative tests of symmetry often determine if bias is minimal enough to meet a standard a priori cutoff, such as an alpha level of 0.05. A Bayesian approach like that illustrated herein could allow for a better evaluation of bias, as opposed to just a set cutoff; however, further evaluation of how to appropriately incorporate error would be useful. Although used in this study to compare assigned age with known age, these bias analyses can be used to directly compare among structures, readers, or protocols, making them useful for studies with or without known-age fish.

Patterns in age errors, notably overestimating age for young fish and underestimating age for older fish, may have nonnegligible effects on age composition for some hatchery stocks in some years. For some stocks, age composition from scales often with other data is used to assess cohort strength and forecast future run sizes of specific cohorts. For example, for some spring Chinook Salmon stocks, numbers of returning age-3 and age-4 fish in a year are modeled to predict the numbers of age-4 and age-5 fish, respectively, that will return the next year. These estimates are vital for setting harvest allocations and for anticipating the potential to meet, fail to meet, or exceed brood stock collection needs for the hatchery in future years. Estimated age composition is a function of both true age composition and any pattern in age errors. Results of this study suggest that even when measures of accuracy and precision are considered high (i.e., accuracy > 0.9; APE and MCV < 2; Campana 2001; Buckmeier 2002; McBride 2015) and quantitative assessment of bias considered insubstantial, patterns in age errors may affect estimated age composition and thus predictions of future cohort sizes. Analysis used in this study assumes that patterns in age errors are consistent among years, which may not be true; thus, specific results must be evaluated with caution. However, the pattern of overestimating the age of younger age-classes and underestimating the age of older age-classes was consistent among hatchery stocks. This

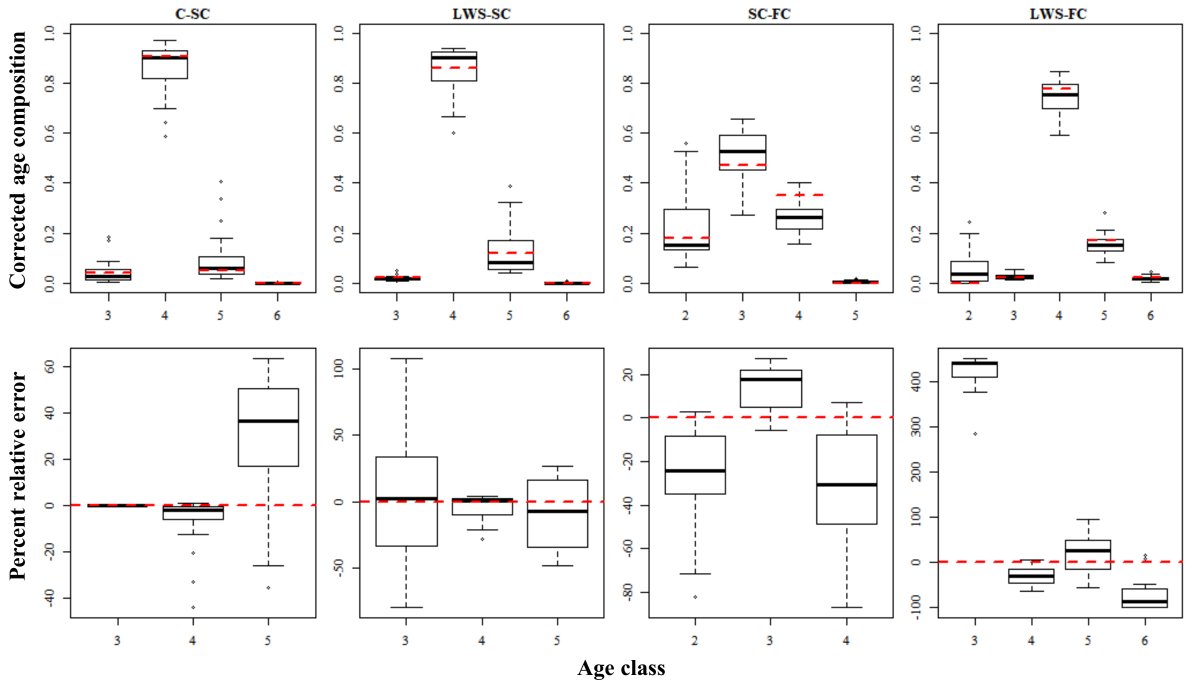

FIGURE 6. Box plots of corrected age composition proportions (upper row) and percentage relative error in assigned age (lower row) by age-class for four hatchery stocks for 1993–2016. Note the differences in the *y*-axis scales among the lower panels. For the box plots the horizontal line in each box indicates the median, the box dimensions represent the 25th to 75th percentile ranges, the whiskers show the 10th to 90th percentile ranges, and the dots are outliers. For the upper row, dashed red lines indicate proportions in this study, and for the lower row, dashed red lines indicate a percentage relative error of zero (i.e., no bias). See Table 1 for hatchery stock abbreviations. The differences in *x*-axis scales between the upper and lower panels were a result of rare age-classes that were not encountered sampling. [Color figure can viewed at afsjournals.org.]

pattern is likely at least partially due to the small number of age-classes and reader knowledge of species maturity patterns; thus, the general pattern may be somewhat consistent among years. Future studies to examine annual patterns in age errors would help better evaluate how age errors may affect estimates of age composition among years.

In conclusion, this study suggests that species and stock biology should be considered when applying scale age data in stock assessment and management. Although many hatchery fish are individually tagged, ages from scales may be the main or only source of demographic data available to assess a wild stock, especially for small stocks of conservation concern. Wild-origin and hatchery-raised Pacific salmon of the same species from the same river can differ in growth, survival, age structure, and migration patterns and have been differentiated by patterns on their scales (Conner et al. 2005); thus, not just species-specific, but also stock-specific, assessments of scale age data are optimal. However, producing a sample of known-age adults from wild populations (e.g., from tagging or genetics) may be impractical or impossible in some (or even most) cases. Therefore, results from this study suggest that with appropriate protocols, scale age data from stocks with low variation in age at maturity and observable patterns in length at age, like those of Coho Salmon and spring Chinook Salmon, would likely aid stock assessment and management. For stocks with extensive individual variation in behavior and demographics, such as fall Chinook Salmon stocks, the use of CWTs, PIT tags, or genetic techniques (Seamons et al. 2009; Beacham et al. 2019) may be necessary. Being cognizant of potential differences in protocols and biology are important when applying estimates of accuracy from one stock to another. For many species, integrated approaches that combine data from scales, tags, and/or genetic techniques could be beneficial for improving accuracy and efficiency. Continued work to improve accuracy and better understand the potential sources and implications of bias using known-age fish from hatcheries may be an effective way to develop and improve protocols to provide demographic data from scales for assessment and management of wild stocks.

ACKNOWLEDGMENTS

I thank James Archibald, Steven Lazzini, Ian MacDonald, Courtney Newlon, Jesse Rivera, and Howard Schaller for helpful discussions about study design and sampling protocols and for organization of data collection. I thank the USFWS Columbia River Fish and Wildlife Conservation Office marking crew for collecting and reading scales, especially Kyle Beard, Chuck Fuller, Geoff Gribble, Pat Kemper, Steven Lazzini, and Ian MacDonald. I appreciate Tim Copeland, Steven Haeseker, Richard McBride, William Simpson, Jeremy Voeltz, and two anonymous reviewers for providing helpful comments on previous drafts of this manuscript. Funding was provided by the USFWS Columbia River Fish and Wildlife Conservation Office. References to trade names do not imply endorsement by the U.S. Government. The findings and conclusions in this manuscript are those of the author and do not necessarily represent the views of the USFWS. There is no conflict of interest declared in this article.

REFERENCES

Beacham, T. D., C. Wallace, K. Jonsen, B. McIntosh, J. R. Candy, D. Willis, C. Lynch, J.-S. Moore, L. Bernatchez, and R. E. Withler. 2019. Comparison of coded-wire tagging with parentage-based tagging and genetic stock identification in a large-scale Coho Salmon fisheries application in British Columbia, Canada. Evolutionary Applications 12:230–254.

Beamish, R. J., and G. A. McFarlane. 1983. The forgotten requirement for age validation in fisheries biology. Transactions of the American Fisheries Society 112:735–743.

Berejikian, B. A., D. M. Van Doornik, and J. J. Atkins. 2011. Alternative male reproductive phenotypes affect offspring growth rates in Chinook Salmon. Transactions of the American Fisheries Society 140:1206–1212.

Bland, J. M., and D. G. Altman. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1:307–310.

Bottom, D. L., K. K. Jones, C. A. Simenstad, and C. L. Smith. 2009. Reconnecting social and ecological resilience in salmon fisheries. Ecology and Society 14:5.

Bowker, A. H. 1948. A test for symmetry in contingency tables. Journal of the American Statistical Association 43:572–574.

Branigan, P. R., K. A. Meyer, N. C. Whal, M. P. Corsi, and A. M. Dux. 2019. Accuracy and precision of age estimates obtained from three calcified structures for known-age kokanee. North American Journal of Fisheries Management 39:498–508.

Buckmeier, D. L. 2002. Assessment of reader accuracy and recommendations to reduce subjectivity in age estimation. Fisheries 27(11): 10–14.

Campana, S. E. 2001. Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. Fish Biology 59:197–242.

Cline, T. J., J. Ohlberger, and D. E. Schindler. 2019. Effects of warming climate and competition in the ocean for life histories of Pacific salmon. Nature Ecology and Evolution 3:935–942.

Conner, W. P., J. G. Sneva, K. F. Tiffan, R. K. Steinhorst, and D. Ross. 2005. Two alternative juvenile life history types for fall Chinook Salmon in the Snake River basin. Transactions of the American Fisheries Society 134:291–304.

Copeland, T., M. W. Hyatt, and J. Johnson. 2007. Comparison of methods used to age spring–summer Chinook Salmon in Idaho: validation and simulated effects on estimated age composition. North American Journal of Fisheries Management 27:1393–1401.

DeVries, D. R., and R. V. Frie. 1996. Determination of age and growth. Pages 483–512 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Gelman, A., and J. Hill. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York.

Gilbert, C. H. 1912. Age at maturity of the Pacific coast salmon of the genus *Oncorhynchus*. Bulletin of the U.S. Bureau of Fisheries 32:3–22.

Gross, M. R. 1985. Disruptive selection for alternative life histories in salmon. Nature 313:47–48.

Hankin, D. G., J. W. Nicholas, and T. W. Downey. 1993. Evidence for inheritance of age at maturity in Chinook Salmon (*Oncorhynchus tshawytscha*). Canadian Journal of Fisheries and Aquatic Science 50:347–358.

Healey, M. C. 1991. Life history of Chinook Salmon (*Oncorhynchus tshawytscha*). Pages 311–394 *in* C. Groot and L. Margolis, editors. Pacific salmon life histories. UBC Press, Vancouver.

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York.

Hoenig, J. M., M. J. Morgan, and C. A. Brown. 1995. Analyzing differences between two age determination methods by tests of symmetry. Canadian Journal of Fisheries and Aquatic Sciences 52:364–368.

Johnson, B. M., M. S. Johnson, and G. H. Thorgaard. 2019. Salmon genetics and management in the Columbia River basin. Northwest Science 92:346–363.

Johnson, J., T. Johnson, and T. Copeland. 2012. Defining life histories of precocious male parr, minijack, and jack Chinook Salmon using scale patterns. Transactions of the American Fisheries Society 141:1545–1556.

Jordan, S. J., T. O'Higgins, and J. A. Dittmar. 2012. Ecosystem services of coastal habitats and fisheries: multiscale ecological and economic models in support of ecosystem-based management. Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science 4:573–586.

Kellner, K. 2018. "jagsUI." A wrapper around "rjags" to streamline "JAGS" analyses. Available: https://cran.r-project.org/web/packages/jagsUI/jagsUI.pdf. (July 2019).

Kéry, M., and M. Schaub. 2012. Bayesian population analysis using WinBUGS: a hierarchical perspective. Academic Press, Waltham, Massachusetts.

Kodama, M., J. J. Hard, and K. A. Naish. 2012. Temporal variation in selection on body length and date of return in a wild population of Coho Salmon, *Oncorhynchus kisutch*. BMC Evolutionary Biology [online serial] 12:116.

Koo, T. S. Y., and A. Isarankura. 1967. Objective studies of scales of Columbia River Chinook Salmon, *Oncorhynchus tshawytscha* (Walbaum). Fishery Bulletin 66:165–180.

Lewis, B., W. S. Grant, R. E. Brenner, and T. Hamazaki. 2015. Changes in size and age of Chinook Salmon *Oncorhynchus tshawytscha* returning to Alaska. PLoS ONE [online serial] 10:e0130184.

McBride, R. S. 2015. Diagnosis of paired age agreement: a simulation of accuracy and precision effects. ICES Journal of Marine Science 72:2149–2167.

McNicol, R. E., and S. E. MacLellan. 2010. Accuracy of using scales to age mixed-stock Chinook Salmon of hatchery origin. Transactions of the American Fisheries Society 139:727–734.

Meengs, C. C., and R. T. Lackey. 2005. Estimating the size of historical Oregon salmon runs. Reviews in Fisheries Science 13:51–66.

Morison, A. K., J. Burnett, W. J. McCurdy, and E. Moksness. 2005. Quality issues in the use of otoliths for fish age estimation. Marine and Freshwater Research 56:773–782.

Morita, K. S. H., M. Fukuwaka Morita, and H. Matsuda. 2005. Rule of age and size at maturity of Chum Salmon (*Oncorhynchus keta*): implications of recent trends among *Oncorhynchus* spp. Canadian Journal of Fisheries and Aquatic Sciences 62:2752–2759.

Ohlberger, J. E., J. Ward, D. E. Schindler, and B. Lewis. 2018. Demographic changes in Chinook Salmon across the northwest Pacific Ocean. Fish and Fisheries 19:533–546.

PFMC (Pacific Fishery Management Council). 2019. Preseason report I: stock abundance analysis and environmental assessment part 1 for 2019 ocean salmon fishery regulations. PFMC, Portland, Oregon.

R Core Team. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available: https://www.r-project.org/. (July 2019).

Ricker, W. E. 1981. Changes in the average size and average age of Pacific salmon. Canadian Journal of Fisheries and Aquatic Sciences 38:1636–1656.

Sandercock, F. K. 1991. Life history of Coho Salmon (*Oncorhynchus kisutch*). Pages 395–445 *in* C. Groot and L. Margolis, editors. Pacific salmon life histories. UBC Press, Vancouver.

Seamons, T. R., M. B. Dauer, J. Sneva, and T. P. Quinn. 2009. Use of parentage assignment and DNA genotyping to validate scale analysis for estimating steelhead age and spawning history. North American Journal of Fisheries Management 29:396–403.

Sethi, S. A., J. Gerken, and J. Ashline. 2017. Accurate aging of juvenile salmonids using fork lengths. Fisheries Research 185:161–168.

Siegel, J. E., M. V. McPhee, and M. D. Adkison. 2017. Evidence that marine temperatures influence growth and maturation of western Alaskan Chinook Salmon. Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science [online serial] 9:441–456.

Smith, C. T., J. Baumsteiger, W. R. Areden, and D. K. Hawkins. 2015. Eliminating variation in age at spawning leads to genetic divergence within a single salmon population. Journal of Fish and Wildlife Management 6:4–18.

Vandergoot, C. S., and T. O. Brendan. 2015. Accuracy and precision of fishery and demographic estimates from a spatial tag-recovery model when inter-regional movements are treated as fixed. Fisheries Research 164:8–25.

Worthington, D. G., A. J. Fowler, and P. J. Doherty. 1995. Determining the most efficient method of age determination for estimating the age structure of a fish population. Canadian Journal of Fisheries and Aquatic Sciences 52:2320–2326.

Zabel, R. W., and P. S. Levin. 2002. Simple assumptions on age composition lead to erroneous conclusions on the nature of density dependence in age-structured populations. Oecologia 133:349–355.