esa

ECOSPHERE

COASTAL AND MARINE ECOLOGY

Aging Pacific cod (*Gadus macrocephalus*) from otoliths using Fourier-transformed near-infrared spectroscopy

JORDAN HEALY,[1] THOMAS E. HELSER,[2,]† IRINA M. BENSON,[2] AND LUKE TORNABENE[1]

[1]*School of Aquatic and Fisheries Sciences (SAFS), Univerity of Washington, 1122 NE Boat Street, Seattle, Washington 98195 USA*
[2]*NOAA Fisheries, Resource Ecology and Fisheries Management Division, Alaska Fisheries Science Center (ASFC), 7600 Sand Point Way NE, Seattle, Washington 98115 USA*

Citation: Healy, J., T. E. Helser, I. M. Benson, and L. Tornabene. 2021. Aging Pacific cod (*Gadus macrocephalus*) from otoliths using Fourier-transformed near-infrared spectroscopy. Ecosphere 12(8):e03697. 10.1002/ecs2.3697

Abstract. For decades, age-structured stock assessments have been a key component to managing fishery resources worldwide. Fisheries management systems have been under increasing demand to generate a greater volume and quality of age estimates. Traditional aging techniques, which require physical preparation followed by microscopic examination of fish otoliths, are labor-intensive, expensive, and inherently subjective among individual analysts, making repeatability and precision a challenge. Here we investigated an innovative approach to aging fish from their otoliths using Fourier-transformed near-infrared spectroscopy and partial least squares regression models. Models were fit to and validated on spectra and used to microscopically estimate ages of Pacific cod from three years of fishery-independent otolith data out of the Bering sea. Calibrated and validated models for each year, as well as on an ensemble of the three years, yielded high precision for the multiyear model (R^2 = 0.869, RMSE = 0.614, PA = 63%, CV = 7.412), and independent year models (R^2 = 0.844–0.891, RMSE = 0.555–0.615, PA = 65%, CV = 6.313–6.775). These metrics of model performance were highly comparable to precision from the traditional microscopic aging approach (R^2 = 0.763–0.869, RMSE = 0.639–0.737, PA = 63%–70%, CV = 5.671–6.698). In all cases, a two-sided Kolmogorov–Smirnov test showed no significant difference between reference and model estimated age distributions. Our results illustrate how Fourier-transformed near-infrared spectroscopy can be utilized on otoliths to predict age estimates with substantially greater efficiency, good precision, high repeatability, and no loss in data integrity compared to the traditional microscopic method for aging Pacific cod.

Key words: age estimation; chemometrics; fisheries management; near-infrared spectroscopy; Pacific cod; partial least squares regression; stock assessment.

Received 9 December 2020; revised 8 March 2021; accepted 12 March 2021. Corresponding Editor: Hunter S. Lenihan.

† **E-mail:** thomas.helser@noaa.gov

INTRODUCTION

In recent decades, marine fish stocks have been managed by a variety of integrated population modeling approaches that are able to utilize diverse types of data, such as fish lengths, ages, and indices of abundance (Maunder and Punt 2013). In the United States, the National Marine Fisheries Services, National Oceanographic, and Atmospheric Administration (NOAA fisheries) currently manages 474 stocks across 46 fisheries plans (https://www.fisheries.noaa.gov/find-species). In the Alaskan region, the eastern Bering Sea (EBS) Pacific cod (*Gadus macrocephalus*) fishery is

one of the largest fisheries stocks managed, bringing in a yearly average of 185,896 metric tons for years 1991 through 2018 (Thompson 2018). Pacific cod stocks are primarily managed with an age-structured, length-at-age-structure, or stage-structured stock assessment model (Thompson 2018). In most management cases, precise estimates of population age composition are a crucial piece of data, improving the precision of population projections (Ricker 1975, Punt and Hilborn 1997, Maunder and Punt 2013, Ono et al. 2015). Population growth rates of Pacific cod are estimated using age-structured stock assessment models in the Stock Synthesis model framework (Methot and Wetzel 2013, Thompson 2018, Methot et al. 2019). Thus, the acquisition and reliability of age data from Pacific cod populations is imperative to the management of the species.

Teleost fishes have calcium carbonate ($CaCO_3$) structures, called otoliths, in the inner ear that are important for spatial orientation and hearing (Campana 1999, Popper et al. 2005, Thomas and Swearer 2019). Otoliths are composed primarily of a $CaCO_3$ structure and a protein matrix, though the relative proportions of these organic constituents vary among species (Degans et al. 1969, Campana 1999, Zorica et al. 2010). These structures grow throughout the lifetime of the fish. As they grow, otoliths incorporate temporal and geospatial variability in chemical structure and morphology (Gillanders 2002, Elsdon and Gillanders 2004, Schaffler and Winkelman 2008, Matta and Kimura 2012, Matta et al. 2019). This variability provides valuable information to scientists studying the dynamics of fish life history, most notably the annual growth zones that allow age estimation. Temperature has been shown to have an effect on otolith development driving variability in the formation of opaque and translucent growth bands (Morales-Nin 2000, Otterlei et al. 2002, Neat et al. 2008, Hurst et al. 2010, Matta et al. 2010, Matta et al. 2018). In addition, patterns of temperature variability in marine ecosystems have been demonstrated to drive shifts in fish distribution and food availability for Pollock (*Gadus chalcogrammus*) and other *gadidae* species throughout the Bering Sea and Alaskan Gulf (Willie-Echeverrie and Wooster 1998, Kotwicki et al. 2005, Stabeno et al. 2012, Thorson et al. 2017).

Traditional age estimation typically involves some sort of physical preparation, such as sawing and toasting the otolith, followed by microscopic examination of annual growth patterns. For instance, EBS Pacific cod otoliths are roasted, sectioned on a saw, and examined under a dissecting microscope (Matta and Kimura 2012). This entire process can take between 6 and 10 min to derive a single age (T. E. Helser, *personal observation*). Trained analysts (sometimes referred to as age readers) examine and count the opaque and translucent bands (annuli) that correspond to organic-rich summer growth and winter growth, respectively, formed annually to derive an age estimate. Some species have growth patterns that are inherently difficult to read, making age estimation difficult and leading to low statistical precision. Moreover, the interaction between environment and otolith development may lead to additional difficulty interpreting annual banding in otoliths, which translates to less repeatability and more subjectivity in age estimates (Matta and Kimura 2012). These challenges can lead to increased time and effort by analysts, reduced precision, and increased relative bias and error in estimating fish ages (Høie et al. 2009, Morrongiello and Thresher 2015). The complexity encountered in otolith annual growth patterns can make distinguishing a true annulus from a false annulus (sometimes referred to as a check) a tenuous task, resulting in difficulty producing precise age estimates for Pacific cod (Beamish and McFarlane 1995). Despite Pacific cod being rather short-lived, this challenge is particularly apparent for fish under 6 yr of age (Roberson et al. 2005).

The goal to increase both the precision and cost efficiency of aging fish has led researchers to explore alternative age estimation techniques. In recent years, near-infrared spectroscopy (NIRs) has been explored as an alternative to microscopic age estimation to reduce cost, increase precision, and improve repeatability in estimating fish and elasmobranch ages (Wedding et al. 2014, Rigby et al. 2015, Helser et al. 2019, Passerotti et al. 2020, Wright et al. 2021). Applications of this technology are emerging in other ecological disciples including non-invasive gender identification in anurans (Vance et al. 2014) and Chinook salmon (Hampton et al. 2002) in addition to determination of dietary contents in fecal samples (Jean et al. 2014). The application of spectroscopy spans various studies with the

substantial benefit of a parsimonious increase in repeatability, precision, and speed of estimation (Foley et al. 1998).

Near-infrared spectroscopy functions by exciting covalent bonds in organic structures with near-infrared electromagnetic energy. The near-infrared wavenumber range runs from about 4000 to 12,500 cm^{-1}. The light interaction with the sample results in measurable vibrational frequencies from the organic covalent bonds (O-H, C=O, C-H, C-N, and N-H) of those molecules (Siesler et al. 2002, Conzen 2014). These vibrations are recorded resulting in a matrix of absorbance values for all samples measured along the range of wavenumbers (cm^{-1}) applied by the near-infrared light source. Spectral signatures for each sample are unique to each compound measured by the distinctive combination of organic structures that make up that compound. In otoliths, these spectral signatures are a good proxy for fish age. Using multivariate calibration methods, such as partial least squares regression (PLSr), they can be related to the reference ages, which are estimated by the traditional microscopic approach (Wedding et al. 2014, Rigby et al. 2015, Helser et al. 2019, Passerotti et al. 2020, Wright et al. 2021). The model development process is complete when a predictive calibration model is evaluated against reference ages and tested by predicting a completely new set of representative unknown samples (external validation).

Only two studies predicting fish age from otoliths using Fourier-transformed near-infrared spectroscopy (FT-NIRs) have been published in the northern hemisphere and, as such, additional studies are needed to verify its applicability to species in other ecosystems and stability over time. In this study, we investigate the application of FT-NIRs to estimate Pacific cod ages from otolith spectra, using the PLSr model algorithm with the purpose of examining its relative precision and addressing temporal model performance. Our primary objectives were specifically to (1) develop FT-NIRs predictive models and evaluate the model skill for estimating ages for Pacific cod from otoliths, (2) investigate whether model skill and precision varies across different independent sample collection years, and (3) compare whether a multiyear model ensemble provides better individual year prediction.

Materials and Methods

Otolith collection

Otoliths for Pacific cod are collected annually from bottom trawl surveys conducted in the eastern Bering Sea by the AFSC. These surveys are conducted using a 20 nautical mile fixed grid design at 0–50 m (inner shelf), 50–100 m (middle shelf), and 100–200 m (outer shelf) depths. The survey area averages 493,000 km^2 over approximately 300–400 hauls (Thompson 2018). Otoliths are collected in a random sample from each haul and transported to the Age & Growth Laboratory at the AFSC for age and growth studies. From there, they are ultimately archived at the University of Washington Fish Collection where they are available for use by other researchers (https://www.burkemuseum.org/collections-and research/biology/ichthyology/otolith-database/search.php). From this final otolith repository, 2010, 2016, and 2017 otolith collections were obtained for our study. Temperature has been shown to play an important role in otolith growth and development (Morales-Nin 2000, Otterlei et al. 2002, Neat et al. 2008, Hurst et al. 2010, Matta et al. 2010, Matta et al. 2018). These three years represented largely different temperature regimes in the Bearing sea. Therefore, they were picked to examine how well models would perform when analyzed across multiple years with different temperature regimes in the eastern Bering Sea. Broken, crystalized, and otherwise anomalous specimens were removed from analysis. This resulted in 1049 samples collected in 2010, 1525 samples collected in 2016, and 1267 samples collected in 2017 for a total of 3841 samples analyzed across the three years (Fig. 1).

Reference age determination

Reference ages for all samples ($n = 3841$) were determined by microscopically counting the annual growth rings of otoliths following the aging guidelines and procedures of the AFSC's Age and Growth Manual (Matta and Kimura 2012). This was done for only one of the paired sagittal otoliths belonging to a given fish specimen, leaving the other otolith for spectroscopic analysis. Including all physical otolith preparation, such as cutting and baking, estimating age this way takes approximately 6–10-min on average for Pacific cod (T. E. Helser, *personal*

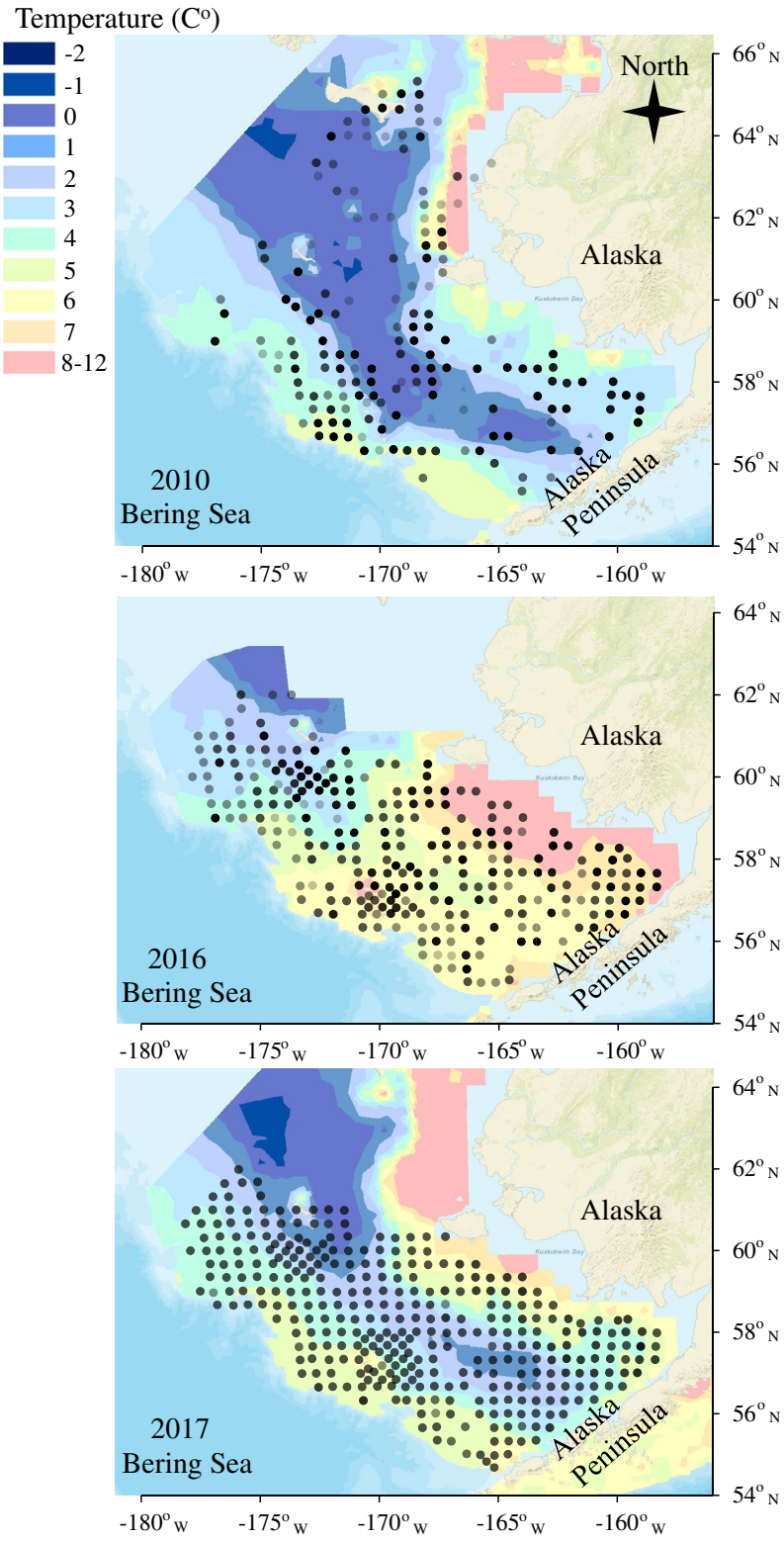

Fig. 1. Map of survey collected Pacific cod otoliths across the eastern Bering Sea shelf in the years 2010, 2016,

(Fig. 1. *Continued*)
and 2017, along with temperature data for those same years. Catch biomass denoted by point density. Base map provided by the AFSC public access ground fish database: https://apps-afsc.fisheries.noaa.gov/RACE/groundfish/survey_data/default.htm.

observation). Approximately 20–30% of all fish aged were randomly selected and aged again by a second reader. Hereafter, we will refer to the primary age reader as Reader1 and the secondary reader as Reader2. This re-aging method is used to estimate precision metrics such as percent agreement (PA) and coefficient of variation (CV) between two readers across all (n) samples (Beamish and Fournier 1981, Chang 1982, Chilton and Beamish 1982, Kimura and Lyons 1991, Campana et al. 1995, Campana 2001). This method of assessing agreement, and the metrics therefrom calculated, has been used for decades to provide stock assessment scientists with an estimate of aging precision for Pacific cod (Matta and Kimura 2012, Thompson 2018).

Aging precision and bias

In order to determine how precise an aging method is, two or more independent age estimates for the same sample need to be generated. In our study, bias between methods or readers was calculated as the difference in integer age between them, wherein A_1 is the age provided by the primary age reader or method and A_2 is the age provided by a secondary reader or method:

$$\text{bias} = (A_1 - A_2).$$

This metric of bias was used to calculate other metrics such as PA, based on the number of agreed upon ages (n_a) ± integer age. In this case, n_a occurs when bias is equal to zero for a given set of ages. This method of assessing agreement has been outlined in Kimura and Lyons (1991) and provides the agers with an estimate of absolute agreement:

$$\text{PA} = \left(\frac{n_a}{n}\right) \times 100.$$

PA shows the proportion of ages for which there is absolute agreement among readers (i.e., where bias = 0). However, this calculation was extended to situations wherein bias was not equal to zero, making it useful in assessing both precision and bias at once. In this case, the equation for PA was used to calculate the relative proportion of situations for which bias was ± some integer age. With samples for which bias between readers was not zero, n_a was exchanged for the number of ages estimated at a given integer value of bias between readers (n_b). Here we called this Percent Bias (PB):

$$\text{PB} = \left(\frac{n_b}{n}\right) \times 100.$$

In our study, PB was calculated for each absolute value of bias between readers in the reference data as well as between reference ages and model estimated ages. From there, values of PB were plotted over different values of bias between two readers or methods. In this way, bias was proportionally assessed in conjunction with PA, simultaneously providing an appraisal of the method's bias and precision. Another way in which bias was estimated, for better visual analysis, was by averaging it with the following equation:

$$\text{average bias} = \frac{\sum_{i=1}^{n} \text{bias}_i}{n}.$$

This was calculated with and without respect to age class to assess overall bias as well as bias at age. In this study, we used this equation for average bias to calculate bias in our model and our reference age estimation. As such, the term "average bias" was used in the context of the overall bias of an age estimating method, be it PLSr model or traditional microscopic age estimation. Finally, bias was visualized with a plot of a primary set of age estimates (Reader1 or reference ages) against a secondary set (Reader2 or model age estimates). In this case, average bias between two readers or methods was assessed as a shift away from the one-to-one line.

Ultimately, CV, PA, PB, and average bias were useful statistics for summarizing model performance in terms of precision and bias. To that end, these statistics were used, along with root mean squared error (RMSE):

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(A_{1i} - A_{2i})^2}{n}}$$

and coefficient of determination (R^2):

$$R^2 = \left(\frac{n\left[\sum_{i=1}^{n}(A_{1i}A_{2i})\right] - \left[\sum_{i=1}^{n}(A_{1i})\right]\left[\sum_{i=1}^{n}(A_{2i})\right]}{\sqrt{\left[n\left(\sum_{i=1}^{n}A_{1i}^2\right) - \left(\sum_{i=1}^{n}A_{1i}^2\right)^2\right]\left[n\left(\sum_{i=1}^{n}A_{2i}^2\right) - \left(\sum_{i=1}^{n}A_{2i}\right)^2\right]}} \right)^2,$$

which were used to determine the level of precision between one set of age estimates (A_1) and second set of age estimates (A_2) for the same samples.

Spectral data collection

Only one of the paired sagittal otoliths, for a given fish, was aged via microscopic examination, as described in the *Reference age determination* section of this study. This left an unaltered sister otolith for each fish. These unaltered whole otoliths were removed from their vials and gently blotted dry with Kimwipes to remove excess glycerol-thymol before placement on the aperture window of a Bruker Tango FT-NIR spectrometer. This spectrometer is a single channel instrument that measures diffuse reflectance using an integrated sphere. Metadata related to these samples (species name, vessel identification, and specimen number) were entered into the OPUS sample description window to be appended to the spectral scan file for later identification. These otoliths were then scanned by the spectrometer, collecting spectra in 16 cm^{-1} resolution at 8 wavenumber intervals with 64 co-added scans, which resulted in approximately 937 wavenumber covariates in an $n \times p$ dimensional x-matrix of spectral data, which was plotted as absorbance values from 12,500 to 4000 cm^{-1} (Fig. 2a). In keeping with protocol used by Helser et al. (2019), wavenumber covariates above 8000 cm^{-1} were omitted from the analysis as they yield little relevant information

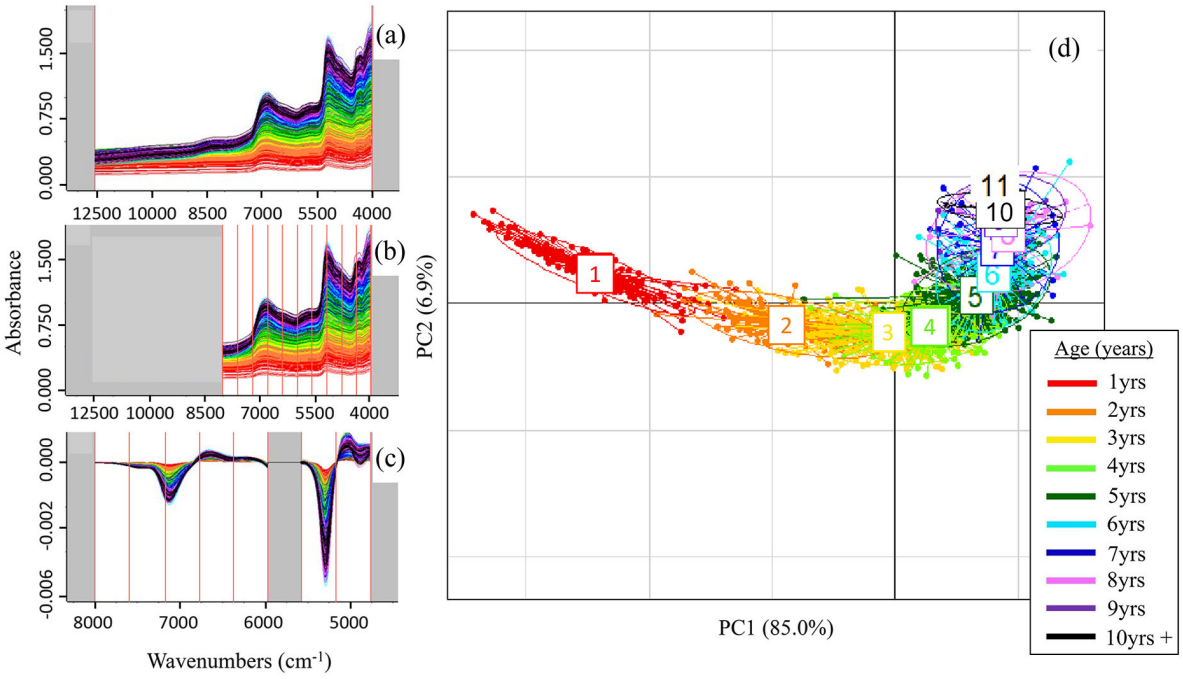

Fig. 2. Pacific cod raw otolith spectra for all three years combined (a). Wavenumber region from 12,000 to 8000 cm^{-1} removed and remaining spectra split into equal width intervals for iPLSr wavenumber selection (b). The resultant Savitzky-Golay first-derivative-transformed and second-order polynomial smoothed spectral signatures, with intervals selected by iPLSr (c). Ordination of spectra with PCA to show discrimination among ages (d). Data was color-coded by age for visualization.

(Fig. 2b). The entire process of removing, drying, positioning, and scanning otoliths, as well as entering scan file metadata, took around 1.5 min per otolith on average. This time estimate does not include total time spend building and validating PLSr models.

Model optimization

Spectra were pre-processed and analyzed in the chemometric software package OPUS (version 8.1, Bruker Optics, Billerica, Massachusetts, USA) and R statistical programing software (R Core Team, 2019) with the R packages prospectr (Stevens and Ramirez-Lopez 2020) and mdatools (Kucheryavskiy 2020). OPUS software was primarily used for all quantitative PLSr analyses of spectral data. Interval partial least squares regression (iPLSr) was used to select for the optimal regions of spectral data in conjunction with data transformations. The iPLSr method is a good option for spectral data as a technique for selecting the optimal combination of model wavenumber regions (x-matrix covariates) and data transformations, due to its relatively low computational demands and ability to handle the high dimensionality of a spectral data matrix (Norgaard et al. 2000, Bras et al. 2008, Mehmood et al. 2012, Wang et al. 2018). In this study, the iPLSr algorithm was used to fit and cross-validate models under combinations of vector normalization and Savitzky-Golay first derivative transformations, in tandem with all possible combinations of 10 equal width spectral intervals (Fig. 2b). The iPLSr algorithm built into the OPUS software included mean centering of the data during the model calibration. All possible combinations of intervals and transformations were analyzed in tandem using RMSE values from leave-one-out (LOO) cross-validation to find the optimal model in terms of selected wavenumber covariates (Table 1, Fig. 2c). Principal component analysis (PCA) was also performed on the iPLSr optimized x-matrices of spectra, as a tool to visualize separation among age in the ordinated spectra (Fig. 2d). Eigenvector loadings of the first 2–3 principal components (where >90% of the variance is explained) were also used to assess for heavily loaded wavenumber covariates which were important descriptors of the variation in the spectra (Fig. 3). In all of the models, iPLSr selected Savisky-Golay first derivative transformation over vector normalization. This was ultimately paired with a second-order polynomial smoothing function within a 17-point smoothing window to reduce noise amplification by the derivative.

Table 1. Wavenumbers selected for individual 2010, 2016, 2017, and multiyear iPLSr models.

Year	Wavenumber regions	No. covariates (p)	Model rank	RMSE
2010	7997-3949	506	10	0.594
2016	7672-7296, 6560-6176, 5440-3952	281	8	0.592
2017	7464-6752, 6408-5000, 4656-3952	353	10	0.588
Multiyear	8000-5968, 5573-4757	356	10	0.623

Models selection and validation

Spectral data and reference ages were fit using PLSr, a linear modeling technique that works well on FT-NIRs spectral data as it reduces the high dimensionality of an x-matrix of spectra down to its principal components (Wold et al. 1984, Geladi and Kowalski 1986, Norgaard et al. 2000). To address our study objectives, we first divided the data from each year into calibration and validation subsets, via a 50-50 random sample spit. We then developed separate PLSr models for each year, wherein models were independently calibrated on the requisite reference ages and spectral data in the calibration subset for that year. These were then validated against the validation subset for that same year. Under this scenario, annual variability within each year was independent of and, therefore, not accounted for by the other two years. We then developed a multiyear model consisting of all three years of data combined, again using a 50-50 random split of the data from each year into calibration and validation data subsets. In this case, the calibration subsets for each year were combined into a single multiyear calibration dataset, used to calibrate a multiyear PLSr model. This model was then validated two ways: (1) on the remaining validation subsets for each year separately and (2) with these same validation subsets combined. In this way, a multiyear

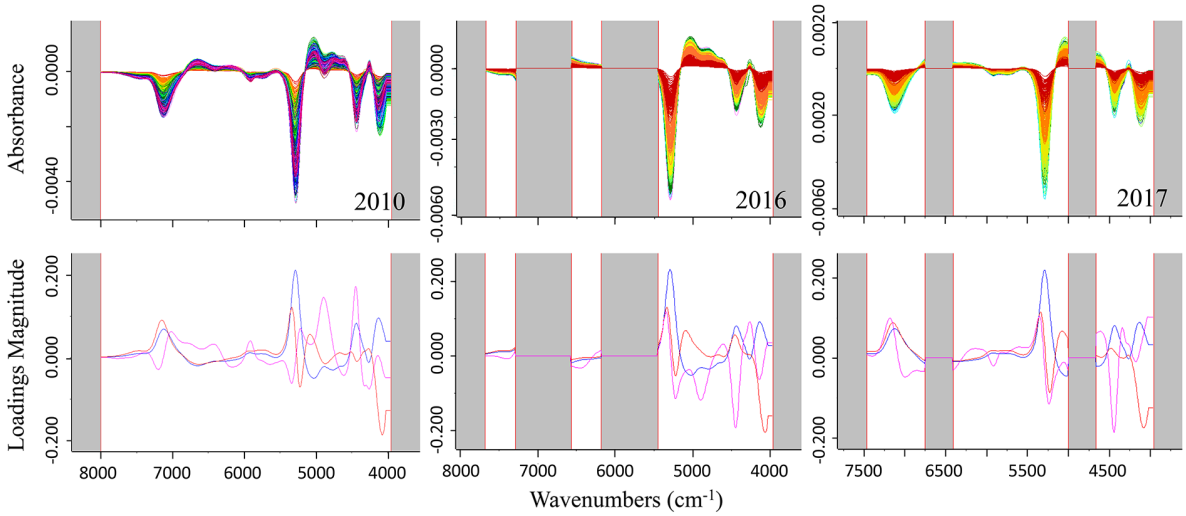

Fig. 3. Savitzky-Golay first-derivative-transformed second-order polynomial smoothed iPLSr selected wavenumber regions (top) and the magnitude of the eigenvector loadings across all wavenumber variables (bottom) for 2010, 2016, and 2017 models.

Table 2. Model performance statistics for the 2010, 2016, 2017, and multiyear iPLSr models.

	Individual year validations			Implicit-year validation	Explicit-year validations		
Statistics	2010	2016	2017	(2010 + 2016 + 2017)	2010	2016	2017
Model R^2	0.871	0.891	0.844	0.869	0.849	0.885	0.859
Reference R^2	0.869	0.858	0.763	0.842	0.869	0.858	0.763
Model PA (%)	65	65	65	63	60	61	65
Reference PA (%)	70	67	63	66	70	67	63
Model CV	6.775	6.378	6.313	7.412	8.657	7.430	6.331
Reference CV	5.671	6.626	6.698	6.093	5.671	6.626	6.698
Model RMSE	0.584	0.555	0.625	0.614	0.690	0.600	0.619
Reference RMSE	0.737	0.639	0.734	0.703	0.737	0.639	0.734
Model Bias	−0.0640	−0.0485	−0.0704	0.0018	0.0552	−0.0625	0.0072
Reference Bias	−0.0333	−0.0696	−0.1004	−0.0669	−0.0333	−0.0696	−0.1004
Model RPD	2.80	3.04	2.55	2.76	2.58	2.95	2.65

Notes: Three individual models were fitted and validated independently on each of the three years of data. The multiyear model was validated by implicit-year validation and explicit-year validation.

model was fitted to the reference age and spectra variability among the three years and validated to account for year directly (explicit-year validation) and indirectly (implicit-year validation). Overall, there were three independent year models (2010, 2016, and 2017) and a multiyear model that was validated implicitly and explicitly with respect to year (Table 2). The estimated ages from both the FT-NIRs and traditional aging methods were aggregated by integer age and compared using a Kolmogorov-Smirnov (K-S) test. The K-S test was applied at a significance level of $\alpha = 0.05$ to test the hypothesis that empirical distributions from reference and model estimated ages come from the same probability distribution.

Results

Reference age determination

Out of the total 3841 Pacific cod otoliths analyzed in this study, two independent age estimates were obtained by microscopic examination from a randomly selected subset ($n = 1372$) of

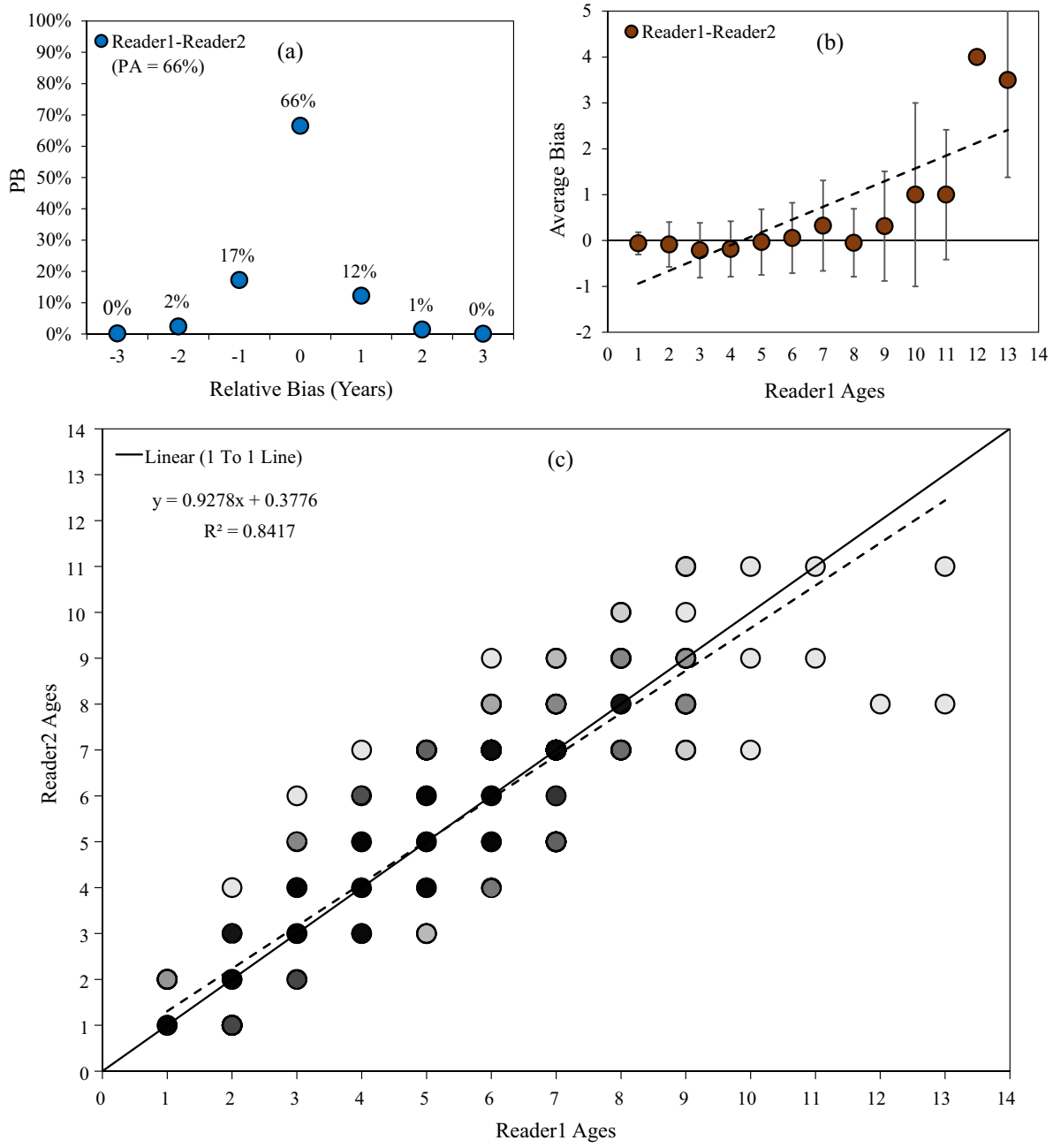

Fig. 4. Comparison of Reader1 and Reader2 estimates of Pacific cod ages, across all three years of our reference data, for visual analysis of between reader agreement (a). Average bias between two age readers for all age groups (b). Between-reader plot with 90% transparency in data points to illustrate point density (c). PB is percent bias for different ± values of bias. Error bars represent one standard deviation.

these otoliths (2010, 541; 2016, 333; and 2017, 498). This method for estimating microscopic aging precision is standard operating procedure for Bering Sea Pacific cod (Matta and Kimura 2012). Reader agreement plots suggested good agreement (PA = 66%) and little overall bias between the two independent microscopic estimates of age for the multiyear dataset (Fig. 4a).

Average bias at age was also minimal, except for samples >8 yr old (Fig. 4.b). There was also high correlation ($R^2 = 0.842$) between the age estimates of the two age readers (Fig. 4.c). Microscopic age estimation precision for multiyear ($R^2 = 0.842$, RMSE = 0.703, PA = 66%, CV = 6.093) and for individual year datasets (2010: $R^2 = 0.869$, RMSE = 0.737, PA = 70%, CV = 5.671, 2016: $R^2 = 0.858$, RMSE = 0.639, PA = 67%, CV = 6.626, 2017: $R^2 = 0.763$, RMSE = 0.734, PA = 63%, CV = 6.698) were comparable to the Pacific cod long term average for microscopic aging since 1990 (Matta and Kimura 2012: PA = 65%, CV = 7.74). However, notable variability existed among the three years in terms of between reader estimates of reference age precision. This variability was apparent across all precision statistics (Table 2).

Model optimization

After the optimal pre-processing and wavenumber selection was chosen for each model, using iPLSr, the resultant spectra showed reasonable separation among ages when ordinated with PCA, though the degree of this separation decreased with age (Fig. 2d). Model wavenumber selection with iPLSr yielded different wavenumber regions between the three survey years (Table 1, Fig. 3). The model fitted to 2010 data kept all 10 intervals of spectra from 7997 to 3949 cm^{-1}, whereas the 2016 and 2017 models were each split into three regions (6 and 8 intervals, respectively) and the multiyear model kept two regions (Table 1, Fig. 3). Prediction error from the iPLSr cross-validations of the multiyear model was higher (RMSE = 0.623) than for any of the individual year models (2017: RMSE = 0.588, 2016: RMSE = 0.592, 2010: RMSE = 0.594) (Table 1). The eigenvector loadings associated with wavenumber covariates showed variation in selected regions, with some of the selected regions having low amplitude loadings in the first three principal components (Fig. 3).

Model performance

Models fitted independently to each year showed good performance, with results being relatively indistinguishable among years (2010: $R^2 = 0.871$, RMSE = 0.584, PA = 65%, CV = 6.775; 2016: $R^2 = 0.891$, RMSE = 0.555, PA = 65%, CV = 6.378; 2017: $R^2 = 0.844$, RMSE = 0.615, PA = 65%, CV = 6.313;), especially in terms of R^2, PA, and CV (Table 2, Fig. 5). When all of the data were combined across years into a multiyear model and validated, using the intrinsic-year validation, precision was slightly lower than for models fitted to each individual year ($R^2 = 0.869$, RMSE = 0.614, PA = 63%, CV = 7.412) especially in terms of PA and CV, but less so in terms of R^2 and RMSE (Table 2, Fig. 6a). Model predictions had slight negative average bias in older fish (after about 6 yr of age), along with an increase in RMSE of prediction as age increased (Figs. 5, 6a). This trend in model average bias in older fish was consistent across all model validations. Though overall model average bias alternated between positive and negative, it remained relatively low in all cases (model average bias <0.07) (Table 2).

Explicit-year validation of the multiyear model yielded relatively high precision in estimating ages for each year when compared against reference ages (2010: $R^2 = 0.850$, RMSE = 0.690, PA = 60%, CV = 8.657; 2016: $R^2 = 0.885$, RMSE = 0.600, PA = 61%, CV = 7.430; 2017: $R^2 = 0.859$, RMSE = 0.606, PA = 65%, CV = 6.331) (Figs. 6, 8). In terms of PA and CV, models performed slightly poorer at estimating ages from earlier collections (Figs. 7, 8). This coincided with validation against the 2017 dataset performing better ($R^2 = 0.859$, RMSE = 0.619, PA = 65%, CV = 6.331) than the between reader estimates of reference age precision ($R^2 = 0.763$, RMSE = 0.734, PA = 63%, CV = 6.698) (Table 2). As was also the case with the individual model validations, RMSE universally increased as ages increased for all model validations and model bias trended slightly negative for fish over 6 years old (Fig. 6b).

All individual year PLSr models performed better than expected when compared to traditional precision metrics for the Pacific cod estimated by Matta and Kimura (2012), in terms of CV < 7.74 and PA > 65% (Table 2 and Fig. 8). CV and PA statistics from individual year models did not stray far from estimates of precision between two readers in the reference ages (Table 2, Fig. 8). When CV and PA were high between two readers in the reference data, they were relatively low between reference age and FT-NIRs age estimates for each of the four models (2010, 2016, 2017, and multiyear) (Table 2, Fig. 8). In terms of PA and CV, all individual and

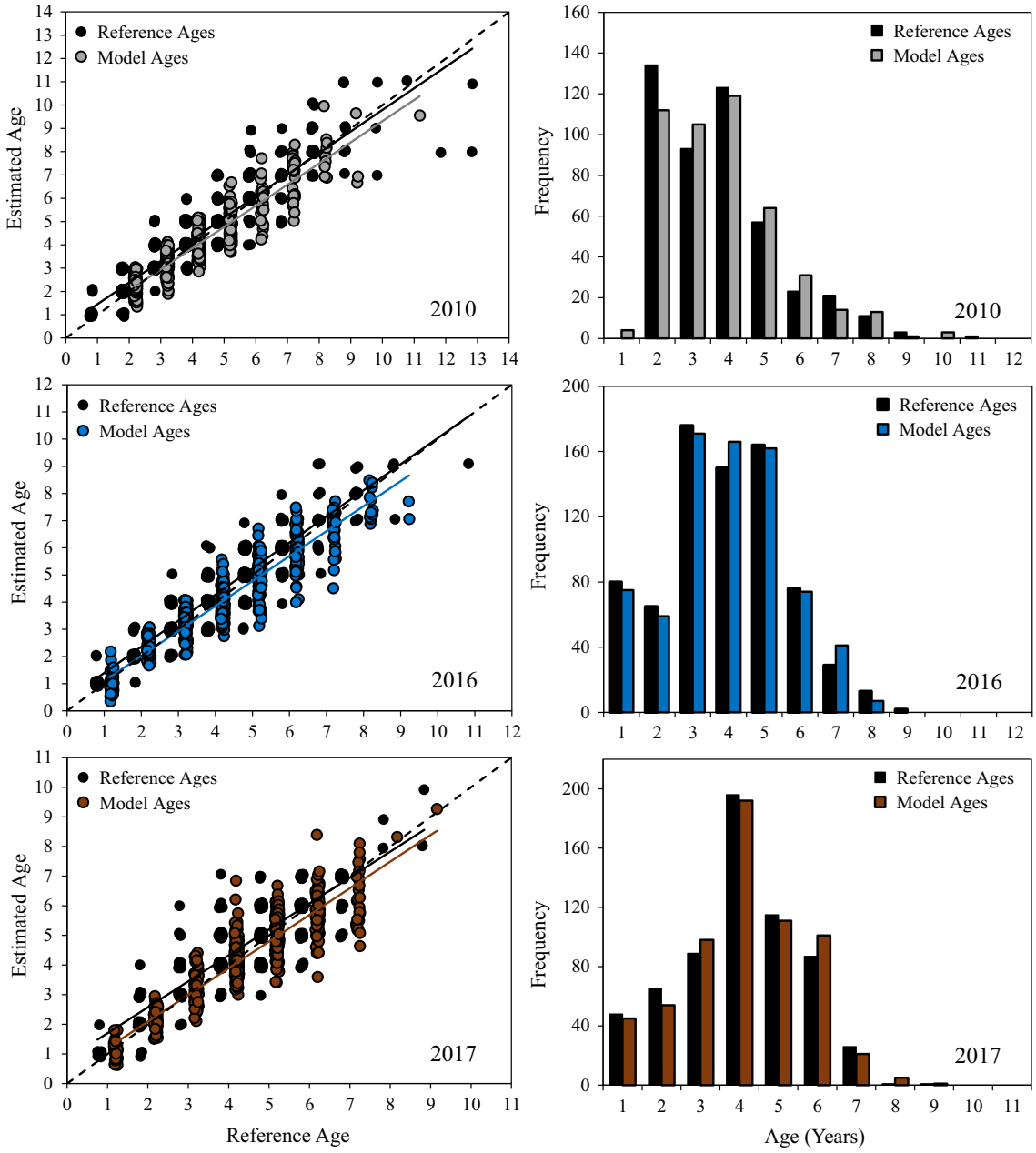

Fig. 5. Results of external validation of 2010, 2016, and 2017 individual year models. Panels on the left show the model predicted ages against the reference microscopic ages (colored) compared to Reader1 vs. Reader2 estimates of reference ages (black). Panels on the right show comparison of age distributions generated by models (colored) and by reference microscopic aging (black).

multiyear model validations performed better than the precision metrics by Matta and Kimura (2012), with the exception of the explicit-year validations of the multiyear model against 2010 and 2016 datasets (Fig. 8). The implicit-year validation performed well (R^2 = 0.869, RMSE = 0.614, PA = 63%, CV = 7.412), especially in terms of PA and CV (Table 2, Fig. 8).

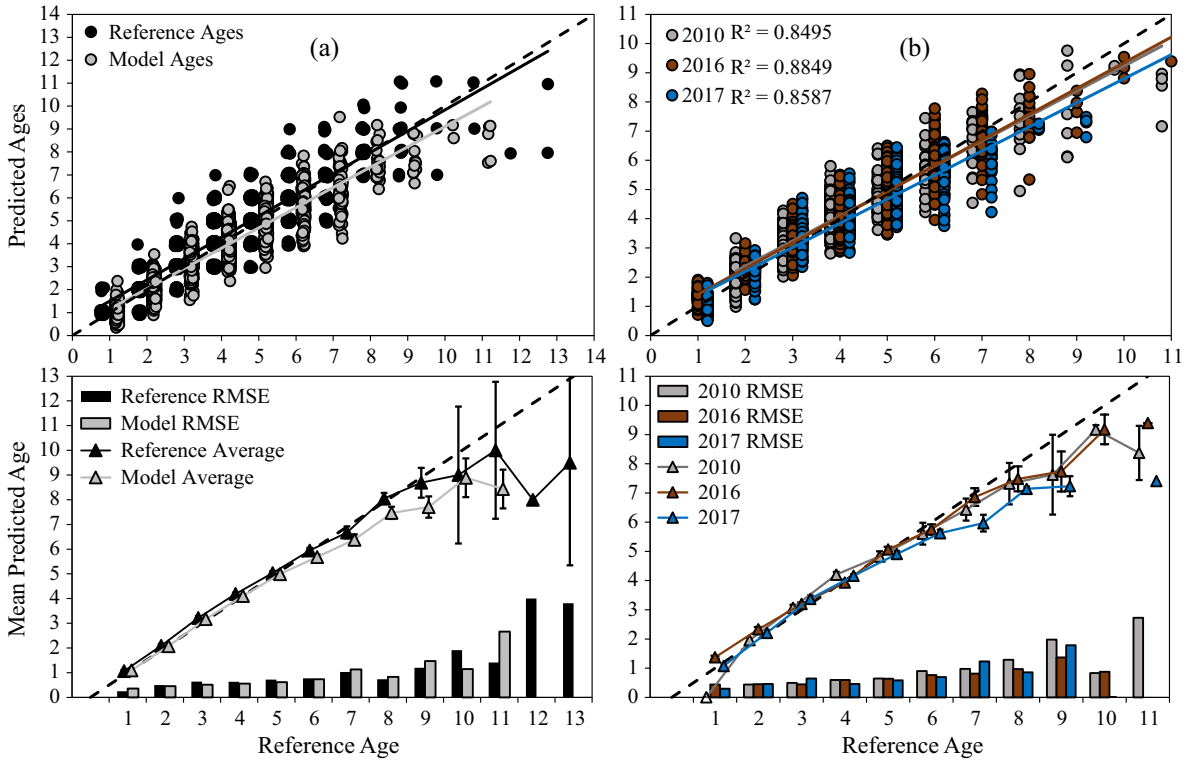

Fig. 6. Results of multiyear model validations. Comparison of Reader1 vs. Reader2 age estimates (black) against model vs. reference age estimates (gray) for the implicit-year validation of the multiyear model (a). Comparison of reference age estimates against multiyear model age estimates for each year of the explicit-year validation (b). For both panels (a) and (b), model average predictions are shown with 95% confidence intervals and RMSE estimates in the reference and model generated ages (bottom).

Overall, the age distributions generated by model validations and microscopic aging did not differ significantly according to two-sample K-S tests ($D = 0.021$ to 0.047, $P = 0.998$ to 0.683). Results of the two-sided K-S test showed that age distributions generated from individual model validations did not differ significantly from traditional reference age distributions for 2010 ($D = 0.039$, $P = 0.878$), 2016 ($D = 0.021$, $P = 0.996$), or 2017 ($D = 0.022$, $P = 0.998$) (Fig. 5). The same was true when comparing the reference age distributions to the model estimated age distributions from the implicit-year validation ($D = 0.020$, $P = 0.835$) and also for 2010 ($D = 0.047$, $P = 0.683$), 2016 ($D = 0.026$, $P = 0.956$), and 2017 ($D = 0.035$, $P = 0.839$) explicit-year validations of the multiyear model.

Discussion

Our study demonstrates that Pacific cod ages can be estimated from FT-NIRs spectra of otoliths, using PLSr models, with a high degree of precision that is comparable to, if not slightly better than, traditional aging methods. PA ranged from 60 to 65% and CV ranged from 6.31 to 8.66. These results were comparable with the historic 30-yr average precision for microscopic aging of Bearing sea Pacific cod at the AFSC, which reported a PA of 65% (±0 yr) and CV of 7.74% (Matta and Kimura 2012). Our models' R^2 (0.84 to 0.89) and RMSE (0.56 to 0.69) were comparable to previous studies exploring the application of FT-NIRs to aging fish. Analysis of walleye pollock (*Gadus chalcogrammus*) otoliths resulted in relatively good precision compared

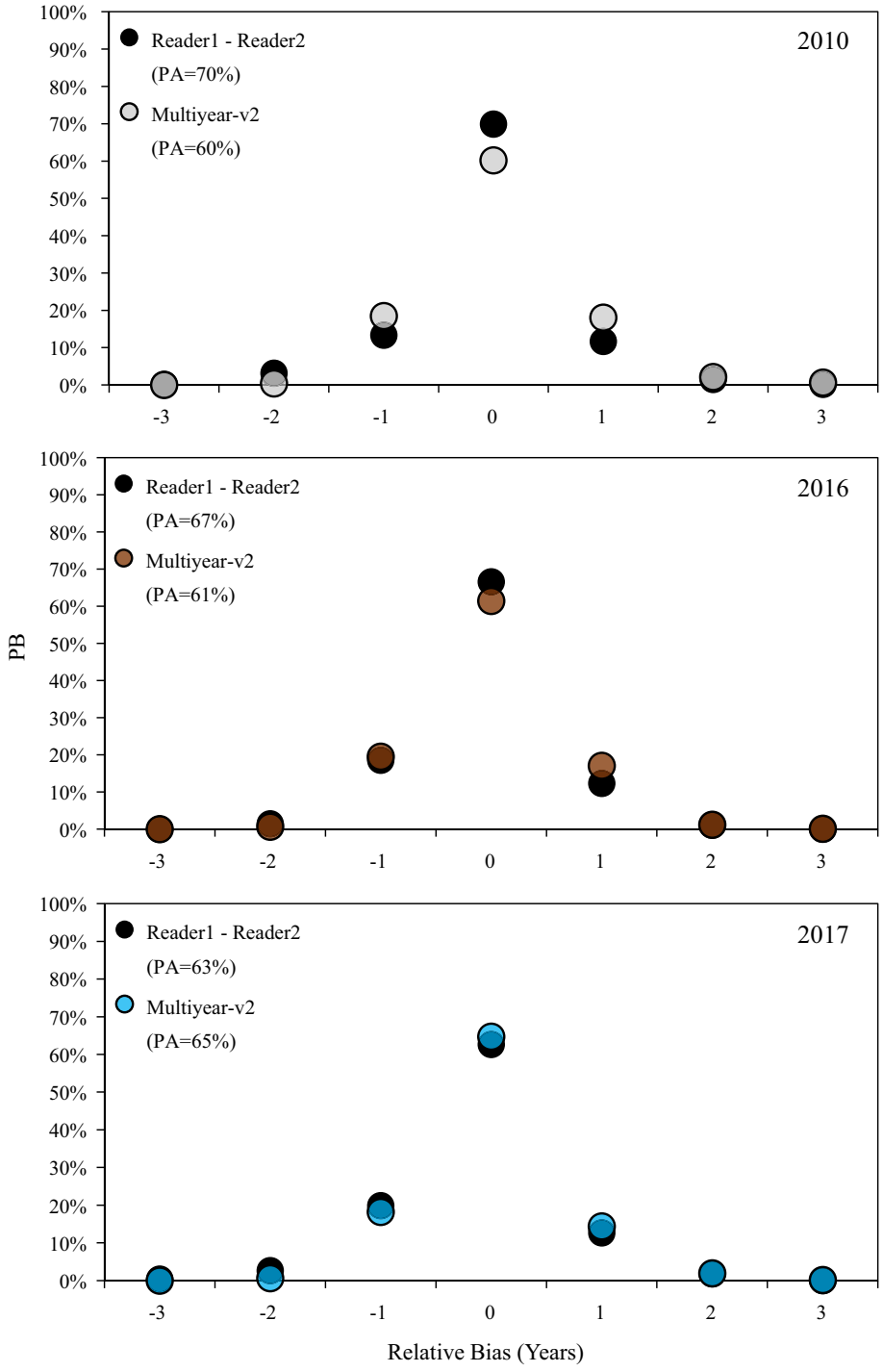

Fig. 7. Comparison of agreement between two readers in the reference ages (black) against agreement between reference and multiyear model estimated ages (colored) for 2010, 2016, and 2017 explicit-year validations of the multiyear model.

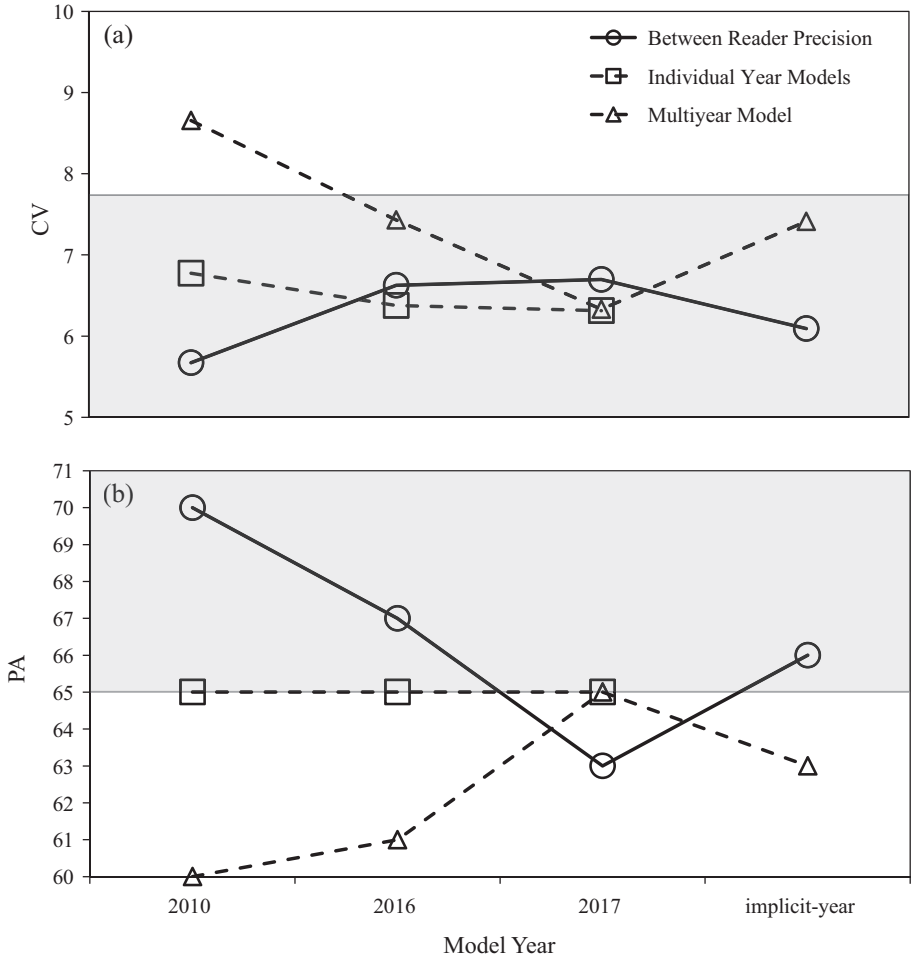

Fig. 8. CV (a) and PA (b) between two readers "Between Reader Precision," between reference ages and model estimated ages for each individual year models "Individual Year Models," and between reference ages and validations of multiyear model "Multiyear Model." Shaded regions indicate where the precision results are better than the 30-yr average for microscopic aging (PA = 65% and CV = 7.74), as set by Matta and Kimura (2012).

against ages generated by the traditional microscopic aging method (R^2 = 0.90, RMSE = 0.96, PA = 79%, age range = 1–18 yr) (Helser et al. 2019). Similar results have been found for saddletail snapper (*Lutjanus malabaricus*; R^2 = 0.94, RMSE = 1.54, age range = 1–23 yr) using otoliths (Wedding et al. 2014). Using mark–recapture ages as a model reference, FT-NIRs also yielded good results for estimating hammerhead (*Sphyrna mokarran*: R^2 = 0.89, RMSE = 0.87, age range = 1–10 yr) and spot-tail (*Carcharhinus sorrah:* R^2 = 0.84, RMSE = 0.88, age range = 1–10 yr) sharks ages from vertebrae (Rigby et al. 2015).

The differences in model performance among studies of FT-NIRs aging are likely driven by uncertainty in the reference age distributions, which in turn reduces model precision. This means that performance of FT-NIRs and PLSr models varies among studies and species mostly on the basis of species-specific reference age uncertainty. Therefore, in each study, an internal comparison of models against their reference age precision, as estimated between two readers, is the most appropriate approach for assessing model performance. Pacific cod is considered to be a highly difficult species to age microscopically (Matta and Kimura 2012), which is notable

in light of its short lifespan. The subjectivity among analysts means that traditional microscopic aging has bias and uncertainty, which differs among collections and analysts (Chilton and Beamish 1982, Kimura and Lyons 1991, Kimura et al. 1992, Beamish and McFarlane 1995). Subjectivity of age estimation may be circumvented by the more automated FT-NIRs based modeling approach. Furthermore, the precision with which FT-NIRs were used to estimate Pacific cod ages was highly promising in terms of repeatability and efficiency gains. Given, on average, a 1.5-min FT-NIRs age estimation time for Pacific cod, which includes physically processing and scanning otoliths, compared to an approximately 6–10-min microscopic aging time (T. E. Helser, *personal observation*), substantial efficiency gains can be realized.

The variation in individual model performance across years indicates a need to integrate a yearly effect into the modeling framework. Temperature regime shifts between years in the Bering Sea drive shifts in primary production, which permeate throughout the food web (Benson and Trites 2002, Polovina 2005, Overland et al. 2008, Stabeno et al. 2012). Environmental fluctuations and primary productivity can alter the ratio of calcium carbonate to protein in the otolith matrix (Beamish and McFarlane 1995). These otolith biochemical alterations are likely to affect their spectral signatures and otolith growth intervals, confounding microscopic aging precision (Morales-Nin 2000, Otterlei et al. 2002, Neat et al. 2008, Matta et al. 2010, Matta et al. 2018). Spectral variation and increased uncertainty in reference ages may also affect model performance, as the linear combinations between the spectral data and the reference ages will have increased error. Furthermore, Pacific cod are a mobile species and there is recent evidence to suggest that their distribution throughout the baring sea is affected by changing temperature regimes (Stabeno et al. 2012, Thompson 2018), and this relationship between temperature and distribution is ontogenetic (Barbeaux and Hollowed 2017). Additionally, there is evidence of a recent northward expansion in Pacific cod (Spies et al. 2019). This variation in Pacific cod distribution exists while the sampling grids for a given stock remain fixed. Therefore, we cannot be certain that we are getting precisely commensurate samples of the stock population from one survey season to the next. This is likely to drive substantial variation in stock composition among survey years along many dimensions, one of which may be the chemical structure of the otolith, therein driving variation in the spectral signatures. All these effects on the baring sea Pacific cod stocks may additively explain the discrepancy in iPLSr wavenumber selection among the three years we analyzed in this study. It is important to pursue future studies that utilize a multiyear model updating approach to estimating age from spectra, perhaps also spatially integrating temperature and other environmental effects. Furthermore, it is paramount to identify what environmental factors are most influential on driving variation in otolith spectra among years. The efficiency benefits of having a dynamic multiyear model framework, rather than generating and validating models on a yearly basis, make it a more desirable approach for future application.

While developing FT-NIRs predictive models, we wanted to investigate whether model skill and precision might vary with ecosystem changes across three different independent sample collection years and evaluate whether a multiyear model ensemble might provide better individual year prediction. The implicit-year validation of the multiyear model showed slightly lower precision, in terms of PA and CV, than any of the individual year validations (Table 2, Fig. 8). This precision was also slightly lower than between reader precision in the reference ages. Even So, absolute model average bias and RMSE for the implicit-year validation (model bias = 0.0018, RMSE = 0.614) were lower than they were between readers (reference bias = −0.0669, RMSE = 0.703) and lower than seen in independent year validations (Table 2). This lower precision and reduced bias makes sense as the variability among years, in terms of both spectra and reference ages, should lead to a larger model error, therein reducing precision slightly while still retaining the central tendency among years in the data, therein reducing model bias. This result further illustrates a need to combine multiple years of data into our models, in order to capture the greatest amount of variance in the system and reduce the bias. The explicit-year validation of the multiyear model showed a decline in precision when validating the model

against older years (Table 2, Figs. 7, 8). That the explicit-year validation of the multiyear model performed so differently from models validated independently for each year is an interesting result (Fig. 8). Given the large sample size in our study, this result does not seem likely to be merely an artifact of biased sampling of calibration and validation datasets for each model. One possible explanation can be found in the selection of wavenumber variables during iPLSr model optimization. Each of the four models were optimized to a suite of wavenumber predictors unique to each model. The multiyear model also differed from single year models, in terms of the suite of predictor wavenumbers that best defined it, therein resulting in a multiyear model less optimized for the spectral information found in each independent year that it was validated against (Table 1). Also, temperature has been shown to influence otolith development (Morales-Nin 2000, Otterlei et al. 2002, Neat et al. 2008, Hurst et al. 2010, Matta et al. 2010, Matta et al. 2018), in turn making precise microscopic age estimation difficult and effecting otolith chemistry (Morales-Nin 2000, Otterlei et al. 2002, Neat et al. 2008, Matta et al. 2010, Matta et al. 2018). Because of this, the relationship between spectra and ages may vary more between years than within them. This would result in models that estimate age better within a given year. Overall, the multiyear model provided good, but not necessarily improved, precision over the individual models. However, the fact that model validations varied in performance among years, whether individual or multiyear, further solidifies the need to incorporate yearly variation into a larger multiyear model framework.

Future research and considerations

As the results of this study have illustrated, estimating ages of fish from otoliths with FT-NIRs and PLSr models shows promise in terms of precision, repeatability, and efficiency. Additional research to consider may include (1) investigation of mechanisms which relate NIR light absorption to organic structures in the otoliths and how this relates to age, (2) analysis of the degree to which error and bias in the reference ages effects model performance, and (3) evaluation of performance of stock assessment models to FT-NIRs age estimates.

While the precise molecular constituents of fish otoliths which respond to NIR energy have yet to be fully elucidated, the answer may lay in the amount and relative contribution of the organic and carbonate fractions. Otoliths are composed primarily of a $CaCO_3$ structure and a protein matrix. These comprise roughly 96% and 3% of the otolith by mass, respectively (Campana 1999), though the relative proportions of these organic constitutes vary among species (Degans et al. 1969, Campana 1999, Zorica et al. 2010). Within the protein matrix of the otolith, 28 proteins have currently been described, about half of which are water-soluble (Hüssy et al. 2004, Thomas and Swearer 2019), and approximately 380 have been detected, most of which are not yet identified (Thomas et al. 2019). As fish age, they feed and grow and accrete proteins and $CaCO_3$ in a seasonally cyclical manner (Campana 1999), a phenomenon largely driven by temperature differences among these seasons that alters the concentration of soluble and insoluble proteins in the organic matrix (Dannevig 1956, Hüssy et al. 2004, Neat et al. 2008). Therefore, though these organics may be a good proxy for age, more work is needed to identify the functional relationship between age and organic development.

Model error in linear models summarizes the variance and covariance among explanatory variables and values of the response variable. In our case, this would be the variance and covariance between the spectral variables and the measured values in the response vector of ages. Spectral signatures from otoliths may have unexplained variance due to variation in their organic constituents between years, across biological gradients, among regions, and among individuals. This variation, like other sources of morphological and chemical variation, varies among samples irrespective of age. However, spectra are measured with near-perfect precision, rather than estimated (Foley et al. 1998). Conversely, reference ages are imprecisely estimated microscopically, rather than measured, therein introducing a second source of error that occurs at the level of the sample. The effect of this error source was somewhat illustrated by Rigby et al. (2015), wherein a PLSr model was fitted and validated with FT-NIRs spectra for "known age" specimens (mark–recapture data) and for specimens

that had microscopically estimated ages (vertebral band counts) in two shark species (*Sphyrna mokarran* and *Carcharhinus sorrah*). The PLSr model estimated ages from "known age" specimens with far greater precision than from band count specimens, illustrating that variability in microscopic age estimation skews FT-NIRs model results. Conducting similar studies with "known age" specimens across multiple species could help us estimate these two sources of error, which may be unique to each species. Generating model estimates under different simulated levels of error and bias for the ages of "known age" specimens may provide a robust sensitivity analysis to determining the degree to which aging error affects model performance.

As age data are primarily used in stock assessment for making population projections on stock biomass, it is worth considering how the FT-NIRs age data type might affect such models. The accuracy and precision with which age data are estimated are important for reliable age-structured stock assessment estimates, which is the leading stock assessment format used for Pacific cod (Thompson 2018). Therefore, the new data type needs to be used to estimate population parameters alongside microscopically estimated ages, to see how the error therein propagates through stock biomass estimates. This could be done with a sort of sensitivity analysis, by varying age estimate precision, to see how sensitive estimates of growth parameters and stock biomass are to values of error and bias in age data. This would involve a stock assessment simulation study, in which reasonable levels of error and bias in age structure distributions from both FT-NIRs and traditional aging methods are simulated through stock assessments with comparative operating and experimental models. Furthermore, if the observed model bias is in fact a result of uncertainty and the under-aging of older fish, it might be a good idea to consider a reduced age for the stock assessment model plus group. The current plus group cutoff age for Pacific cod is 12 yr. Perhaps this should be reduced to 7 or 8 yr, due to aging uncertainty. With questions such as these addressed, FT-NIRs could prove to be a promising innovation to the future of age estimation in fisheries programs around the world.

Acknowledgments

The authors would like to thank Delsa Anderl and her team of age readers, at the AFSC's Age and Growth Laboratories, for collaboration and assistance with scanning and microscopic aging of Pacific cod otoliths for this project. Thank you to Sandra Neidetcher and Craig Kastelle at NOAA Fisheries for providing valuable input on this manuscript during the writing process. We also extend a special thank you to Jason Erickson for hours of help and instruction in the use of OPUS® chemometric software for data analysis. Finally, thank you to all of the anonymous reviewers for their insightful feedback, which has been invaluable in shaping the final product of this paper. All funding was provided by the National Pacific Research Board (NPRB, Project #1819). The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect those of NOAA or the Department of Commerce. Reference to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA.

Literature Cited

Barbeaux, S. J., and A. B. Hollowed. 2017. Ontogeny matters: climate variability and effects on fish distribution in the eastern Bering sea. Fisheries Oceanography 27:1–15.

Beamish, R. J., and D. A. Fournier. 1981. A method for comparing the precision of a set of age determinations. Canadian Journal of Fisheries and Aquatic Sciences 38:882–983.

Beamish, R. J., and G. A. McFarlane. 1995. A discussion of the importance of aging errors, and an application to walleye pollock: the world's largest fishery. Pages 545–565 *in* D. H. Secor, J. M. Dean, and S. E. Campana, editors. Recent developments in fish otolith research. University of South Carolina Press, Columbia, South Carolina, USA.

Benson, A. J., and A. W. Trites. 2002. Ecological effects of regime shifts in the Bering Sea and eastern North Pacific Ocean. Fish and Fisheries 3:95–113.

Bras, L. P., M. Lopes, A. P. Ferreira, and J. C. Menezes. 2008. A bootstrap-based strategy for spectral interval selection in PLS regression. Journal of Chemometrics 22:695–700.

Campana, S. E. 1999. Chemistry and composition of fish otoliths: pathways, mechanisms and applications. Marine Ecology Progress Series 188:263–297.

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

Campana, S. E., M. C. Annand, and J. I. McMillan. 1995. Graphical and statistical methods for determining the consistency of age determinations. American Fisheries Society 124:131–138.

Chang, W. Y. B. 1982. A Statistical method for evaluating the reproducibility of age determination. Canadian Journal of Fisheries and Aquatic Sciences 39:1208–1210.

Chilton, D. E., and R. J. Beamish. 1982. Age determination methods for fishes studied by the groundfish program at the pacific biological station. Canadian Special Publication of Fisheries and Aquatic Sciences 60:1–102.

Conzen, J. 2014. Multivariate calibration. Third English edition. Bruker Optik GmbH, Drachten, The Netherlands.

Dannevig, E. H. 1956. Chemical composition of the zones in cod otoliths. Journal of Marine Science 21:156–159.

Degans, E. T., W. G. Dueser, and R. L. Haedrich. 1969. Molecular structure and composition of otoliths. Marine Biology 2:105–113.

Elsdon, T. S., and B. M. Gillanders. 2004. Fish otolith chemistry influenced by exposure to multiple environmental variables. Journal of Experimental Marine Biology and Ecology 313:269–284.

Foley, W. J., A. McIlwee, I. Lawler, L. Aragones, A. P. Woolnough, and N. Berding. 1998. Ecological applications of near infrared reflectance spectroscopy – a tool for rapid, cost-effective prediction of the composition of plant and animal tissues and aspects of animal performance. Oecologia 116:293–305.

Geladi, P., and B. R. Kowalski. 1986. Partial least squares regression: a tutorial. Analytica Chimica Acta 185:1–17.

Gillanders, M. B. 2002. Temporal and spatial variability in elemental composition of otoliths: implications for determining stock identity and connectivity of populations. Canadian Journal of Fisheries and Aquatic Sciences 59:669–679.

Hampton, K. A., A. G. Cavinato, D. M. Mayes, J. Steve, and T. L. Hoffnagle. 2002. Near infrared spectroscopic classification of gender and maturity in Chinook salmon (*Oncorhynchus tshawytscha*). Eastern Oregon Science Journal 18:31–36.

Helser, T. E., I. Benson, J. Erickson, J. Healy, C. Kastelle, and J. A. Short. 2019. A transformative approach to ageing fish otoliths using Fourier transform near-infrared spectroscopy: a case study of eastern Bering Sea walleye pollock (*Gadus chalcogrammus*). Canadian Journal of Fisheries and Aquatic Sciences 76:780–789.

Høie, H., R. S. Millner, S. McCully, K. H. Nedreaas, G. M. Pilling, and J. Skadal. 2009. Latitudinal differences in the timing of otolith growth: a comparison between the Barents Sea and southern North Sea. Fisheries Research 96:319–322.

Hurst, T. P., B. J. Laurel, and L. Ciannelli. 2010. Ontogenetic patterns and temperature-dependent growth rates in early life stages of Pacific cod (*Gadus macrocephalus*). Fishery Bulletin 108:382–392.

Hüssy, K., H. Mosegaard, and F. Jessen. 2004. Effect of age and temperature on amino acid composition and the content of different protein types of juvenile Atlantic cod (*Gadus morhua*) otoliths. Canadian Journal of Fisheries and Aquatic Sciences 61:1012–1020.

Jean, P. O., R. L. Bradley, M. A. Giroux, J. P. Tremblay, and S. D. Cote. 2014. Near infrared spectroscopy and fecal chemistry as predictors of the diet composition of white-tailed deer. Rangeland Ecology and Management 67:154–159.

Kimura, D. K., and J. J. Lyons. 1991. Between-reader bias and variability in the age determination process. Fishery Bulletin 89:53–60.

Kimura, D. K., J. J. Lyons, S. E. Maclellan, and B. J. Goetz. 1992. Effects of year-class strength on age-determination. Marine and Freshwater Research 43:1221–1228.

Kotwicki, S., T. W. Buckley, T. Honkalehto, and G. Walters. 2005. Variation in the distribution of walleye pollock (*Theragra chalcogramma*) with temperature and implications for seasonal migration. Fishery Bulletin 103:574–587.

Kucheryavskiy, S. 2020. mdatools: R package for chemometrics. Chemometrics and Intelligent Laboratory Systems 198:1–10.

Matta, M. E., B. A. Black, and T. K. Wilderbuer. 2010. Climate-driven synchrony in otolith growth-increment chronologies for three Bering Sea flatfish species. Marine Ecology Progress Series 413:137–145.

Matta, M. E., T. E. Helser, and B. A. Black. 2018. Intrinsic and environmental drivers of growth in an Alaskan rockfish: an otolith biochronology approach. Environmental Biology of Fishes 101:1571–1587.

Matta, M. E., and D. Kimura. 2012. Age determination manual of the Alaska Fisheries Science Center Age and Growth Program. NOAA Professional Paper NMFS 13. NOAA, National Marine Fisheries Services, Seattle, Washington, USA.

Matta, M. E., J. A. Miller, J. A. Short, T. E. Helser, T. P. Hurst, K. M. Rand, and O. A. Ormseth. 2019.

Spatial and temporal variation in otolith elemental signatures of age-0 Pacific cod (*Gadus macrocephalus*) in the Gulf of Alaska. Deep Sea Research Part II: Topical Studies in Oceanography 165:268–279.

Maunder, M. N., and A. A. Punt. 2013. A review of integrated analysis in fisheries stock assessment. Fisheries Research 142:61–74.

Mehmood, T., K. H. Liland, L. Snipen, and S. Saebø. 2012. A review of variable selection methods in partial least squares regression. Chemometrics and Intelligent Laboratory Systems 118:62–69.

Methot, R. D., and C. R. Wetzel. 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142:86–99.

Methot, R. D., C. R. Wetzel, I. G. Tayler, and K. Doering. 2019. Stock synthesis user manual. Version 3.30.14. NOAA Fisheries, Seattle, Washington, USA.

Morales-Nin, B. 2000. Review of the growth regulation processes of otolith daily increment formation. Fisheries Research 46:53–67.

Morrongiello, J. R., and R. E. Thresher. 2015. A statistical framework to explore ontogenetic growth variation among individuals and populations: a marine fish example. Ecological Society of America 85:93–115.

Neat, F. C., P. J. Wright, and R. J. Fryer. 2008. Temperature effects on otolith pattern formation in Atlantic cod *Gadus morhua*. Journal of Fish Biology 73:2527–2541.

Norgaard, L., A. Saudland, J. Wagner, J. P. Nielsen, L. Munck, and S. B. Engelsen. 2000. Interval partial least-squares regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy. Applied Spectroscopy 54:413–419.

Ono, K., et al. 2015. The importance of length and age composition data in statistical age-structured models for marine species. ICES Journal of Marines Science 72:31–43.

Otterlei, E., A. Folkvord, and G. Nyhammer. 2002. Temperature dependent otolith growth of larval and early juvenile Atlantic cod (*Gadus morhua*). ICES Journal of Marines Science 59:401–410.

Overland, J., S. Rodionov, S. Minobe, and N. Bond. 2008. North Pacific regime shifts: definitions, issues and recent transitions. Progress in Oceanography 77:92–102.

Passerotti, M. S., T. E. Helser, I. M. Benson, B. K. Barnett, J. C. Ballenger, W. J. Bubley, M. J. M. Reichert, and J. M. Quattro. 2020. Age estimation of red snapper (*Lutjanus campechanus*) using FT-NIR spectroscopy: feasibility of application to production ageing for management. ICES Journal of Marines Science 77:2144–2156.

Polovina, J. J. 2005. Climate variation, regime shifts, and implications for sustainable fisheries. Bulletin of Marine Science 76:233–244.

Popper, A. N., J. Ramcharitar, and S. E. Campana. 2005. Why Otoliths? Insights form inner ear physiology and fisheries biology. Marine and Freshwater Research 56:497–504.

Punt, A. E., and R. Hilborn. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries 7:35–63.

R Development Core Team. 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin of the Fisheries Research Board of Canada 191:1–367.

Rigby, C. L., B. B. Wedding, S. Grauf, and C. A. Simpfendorfer. 2015. Novel method for shark age estimation using near infrared spectroscopy. Marine and Freshwater Research 67:537–545.

Roberson, N. E., D. K. Kimura, D. R. Gunderso, and A. M. Shimada. 2005. Indirect validation of the age-reading method for Pacific cod (*Gadus macrocephalus*) using otoliths from marked and recaptured fish. Fishery Bulletin 103:153–160.

Schaffler, J. J., and D. L. Winkelman. 2008. Temporal and spatial variability in otolith trace-element signatures of juvenile striped bass from spawning locations in Lake Texoma, Oklahoma-Texas. Transactions of the American Fisheries Society 137:818–829.

Siesler, H. W., Y. Ozaki, S. Kawata, and H. M. Heise. 2002. Near-infrared spectroscopy. Principles, instruments, applications. First edition. WILEY-VCH Verlag GmbH, Weinheim, Germany.

Spies, I., K. M. Gruenthal, D. P. Drinan, A. B. Hollowed, D. E. Stevenson, C. M. Tarpey, and L. Hauser. 2019. Genetic evidence of a northward range expansion in the eastern Bering Sea stock of Pacific cod. Evolutionary Applications 13:362–375.

Stabeno, P. J., N. B. Kachel, S. E. Moore, J. M. Napp, M. Sigler, A. Yamaguchi, and A. N. Zerbini. 2012. Comparison of warm and cold years on the southeastern Bering Sea shelf and some implications for the ecosystem. Deep Sea Research Part II: Topical Studies in Oceanography 65–70:31–45.

Stevens, A., and L. Ramirez-Lopez. 2020. Miscellaneous functions for processing and sample selection of spectroscopic data. Vignette R package, Version 0.2.0. https://cran.r-project.org/web/packages/prospectr/vignettes/prospectr.html

Thomas, O. R. B., and S. E. Swearer. 2019. Otolith biochemistry: a review. Reviews in Fisheries Science and Aquaculture 27:458–489.

Thomas, O. R. B., S. E. Swearer, E. A. Kapp, P. Peng, G. Q. Tonkin-Hill, A. Papenfuss, A. Roberts, P. Bernard, and B. R. Roberts. 2019. The inner ear proteome of fish. FEBS Journal 286:66–81.

Thompson, G. G. 2018. Assessment of the Pacific cod stock in the eastern Bering Sea. NPFMC Bering Sea and Aleutian Islands SAFE. NOAA, National Marine Fisheries Services, Seattle, Washington, USA. https://www.fisheries.noaa.gov/resource/data/2018-assessment-pacific-cod-stock-eastern-bering-sea

Thorson, J. T., J. N. Ianelli, and S. Kotwicki. 2017. The relative influence of temperature and size-structure on fish distribution shifts: a case-study on Walleye pollock in the Bering Sea. Fish and Fisheries 18:1073–1084.

Vance, C. K., A. J. Kouba, and S. T. Willard. 2014. Near infrared spectroscopy applications in amphibian ecology and conservation: gender and species identification. NIR News 25:5–10.

Wang, L., Y. Lin, X. Wang, N. Xiao, Y. Xu, H. Li, and Q. Xu. 2018. A selective review and comparison for interval variable selection in spectroscopic modeling. Chemometrics and Intelligent Laboratory Systems 172:229–240.

Wedding, B. B., A. J. Forrest, C. Wright, S. Grauf, P. Exley, and S. E. Poole. 2014. A novel method for the age estimation of saddletail snapper (*Lutjanus malabaricus*) using Fourier transform-near infrared (FT-NIR) spectroscopy. Marine and Freshwater Research 65:894–900.

Willie-Echeverrie, T., and W. S. Wooster. 1998. Year-to-year variations in Bering Sea ice cover and some consequences for fish distributions. Fisheries Oceanography 7:159–170.

Wold, S., A. Ruhe, H. Wold, and W. J. Dunn. 1984. The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. SIAM Journal on Scientific and Statistical Computing 5:735–743.

Wright, C., B. B. Wedding, S. Grauf, and O. J. Whybird. 2021. Age estimation of barramundi (*Lates calcarifer*) over multiple seasons from the southern Gulf of Carpentaria using FT-NIR spectroscopy. Marine and Freshwater Research. https://doi.org/10.1071/MF20300

Zorica, B., G. Sinovcic, and V. C. Kec. 2010. Preliminary data on the study of otolith morphology of five pelagic fish species from the Adriatic Sea (Croatia). Acta Adriatica 51:89–96.