Environ Biol Fish (2013) 96:799–810
DOI 10.1007/s10641-012-0072-8

Statistical analysis on otolith data of anadromous fishes

Yongwen Gao · Robert Conrad · David Bean · David L. G. Noakes

Received: 14 March 2011 / Accepted: 18 July 2012 / Published online: 4 August 2012

Abstract Stable oxygen and carbon isotope ($\delta^{18}O$ and $\delta^{13}C$) analyses of otoliths are becoming increasingly common in fisheries science and management. However, little is known about the statistical properties of isotopic data and few attempts have been made to explore appropriate statistical methods that could be used for otolith data analysis. In this paper, we present a pilot study on $\delta^{18}O$ and $\delta^{13}C$ data from otoliths of two anadromous fish species, Atlantic salmon (*Salmo salar*) and Pacific sockeye salmon (*Oncorhynchus nerka*). The results indicated that the salmon otolith data were not normally distributed, so that linear discriminant function analysis and commonly-used statistical tests such as ANOVA and the *t*-test may not be appropriate. Using non-parametric *k*-sample nearest neighbor discriminant analysis, we were able to discriminate with high accuracy among five hatcheries for Atlantic salmon and the origins of wild and hatchery sockeye salmon. Analyses also indicated that the sample sizes required to estimate $\delta^{18}O$ and $\delta^{13}C$ means based on the different sources of variability (between group or within group) and precision levels ($\leq \pm 5.0$ %) were not large. These results and conclusions not only address the statistical considerations of isotopic data from otoliths, but also have practical importance for fisheries management as well.

Keywords Stable isotope analysis · $\delta^{18}O$ and $\delta^{13}C$ data · Statistical tests · Atlantic salmon · Sockeye salmon

Y. Gao (✉)
Makah Fisheries Management,
P.O. Box 115, Neah Bay, WA 98357, USA
e-mail: gaoy@olypen.com

Y. Gao
College of Fisheries, Huazhong Agricultural University,
Wuhan, Hubei 430070, China

R. Conrad
Northwest Indian Fisheries Commission,
6730 Martin Way East,
Olympia, WA 98516, USA

D. Bean
National Marine Fisheries Service,
17 Godfrey Drive,
Orono, ME 04473, USA

D. L. G. Noakes
Department of Fisheries and Wildlife,
Oregon State University,
Corvallis, OR 97331, USA

Introduction

Stable oxygen and carbon isotope ratio ($^{18}O/^{16}O$ or $\delta^{18}O$, and $^{13}C/^{12}C$ or $\delta^{13}C$) analyses of otoliths have become a powerful tool in fisheries science and management, particularly in stock identification (e.g., Edmonds and Fletcher 1997; Stephenson et al. 2001; Gao et al. 2001, 2010), fish behavior and migration (Nelson et al. 1989; Meyer-Richow et al. 1992; Gao and Beamish 2003a), and decadal-scale climate

regime shift investigations (Gao 2002; Gao and Beamish 2003b). As more and more isotopic data are accumulated, a series of questions have been raised in recent years. For instance, what are the proper methods of analysis for isotopic data and how to extract the most valuable information for management from coexisting fish species? Previously Gao and Beamish (1999) demonstrated that $\delta^{18}O$ and $\delta^{13}C$ of otoliths from sockeye salmon (*Oncorhynchus nerka*) provided information on freshwater and marine life history of the individual fish and might be useful in discriminating stocks of wild-origin salmon that had originated from different geographic locations. However, there has been no further examination of whether isotopic analysis may be useful for differentiating sockeye salmon from different hatcheries. Coghlan et al. (2007) used otolith chemistry and linear discriminant function analysis (LDFA) to discriminate between hatchery-reared and naturally-spawned brook trout (*Salvelinus fontinalis*) and rainbow trout (*Oncorhynchus mykiss*). Gibson et al. (2010) used LDFA to separate wild and hatchery red snapper (*Lutjanus campechanus*) in the Gulf of Mexico, and found that the most important feature for discriminating between wild and hatchery red snapper was $\delta^{13}C$. To our knowledge, there is a lack of understanding on statistical properties of otolith $\delta^{18}O$ and $\delta^{13}C$ data and few attempts have been made to explore what kind of statistical methods are appropriate.

Previous studies generally treated the isotopic data of otoliths as traditional biological data and used common parametric statistical analyses (e.g., ANOVA, *t*-test, LDFA) for data interpretation and analysis. Compared with traditional biological data in fisheries, the isotopic data of otoliths have the following distinctions: (1) the $\delta^{18}O$ and $\delta^{13}C$ data have much higher quality in terms of accuracy and precision, because all stable isotope laboratories over the world use the same international standard VPDB (Vienna Peedee belemnite) to calibrate the isotopic enrichment and the precision is generally better than 0.1×10^{-3} for both $\delta^{18}O$ and $\delta^{13}C$ values; (2) using well-preserved samples and the same otolith pair, investigators are able to re-run samples and check questionable $\delta^{18}O$ and $\delta^{13}C$ values; and (3) the errors of the $^{18}O/^{16}O$ and $^{13}C/^{12}C$ ratio measurements can be calculated and monitored under laboratory conditions. In contrast, the traditional biological data in fisheries are often of poor quality due to highly dynamic nature of the animals, and most often the samples are only available for a single-time measurement. Therefore, analysts should consider the differences between biological data and isotopic data of otoliths, and adopt statistical methods specific to their features.

From a statistical point of view, the essential properties of the $\delta^{18}O$ and $\delta^{13}C$ data from otoliths are determined by distribution patterns (Sokal and Rohlf 1981). If the isotopic data are normally distributed, for instance, traditional parametric statistical analyses can be used to compare the mean ($\overline{X}$) and the standard deviation (*SD*) of different fish populations. Previous research has not critically examined the sample size needed for otolith studies. Instead, some authors have reported their stable isotopic analysis based on several fish otoliths; others have presented their results with dozens to hundreds of otoliths. These critical questions still remain in fisheries science and need to be addressed: how many samples should be collected for an otolith project and what kind of information could be gathered by statistical analysis of isotopic data?

In this paper, we analyze the $\delta^{18}O$ and $\delta^{13}C$ data of otoliths from our previous investigations on Atlantic salmon (*Salmo salar*) and Pacific sockeye salmon (*Oncorhynchus nerka*). These anadromous species were primarily from hatchery sources such that all analytical results can be explained by hatchery observations and practices. The objectives of this manuscript are three-fold. First, using analyses on Atlantic salmon and sockeye salmon as examples, we examined whether the $\delta^{18}O$ and $\delta^{13}C$ data of otoliths were normally distributed; as this would help to determine the important statistical properties of anadromous fish otoliths. Second, we also evaluated the minimum sample size requirements for otolith isotopic data. Last, we investigated a practical management need for anadromous fishes by using otolith isotopic data for identification of wild and hatchery salmon from spawners or carcass samples, and discrimination of salmon smolts from different hatcheries in a watershed or broad region. Because the isotopic data of otoliths contain useful information on fish population and growth, we believe that the methods discussed in the manuscript should be able to reflect the statistical power of the chemical data and may extend to other fish species for otolith studies.

Materials and methods

Otolith sample collection and analyses

Otolith samples of Atlantic salmon smolts were collected from five hatcheries located in the State of Maine (Fig. 1): three private hatcheries (Oquossoc, Bingham, and Gardner Lake) that rear and transfer salmon smolts to ocean pens for commercial growth; and two federal hatcheries (the Craig Brook National Fish Hatchery (CBFH) and the Green Lake National Fish Hatchery (GLFH)) which provide smolts for salmon recovery (i.e., out-planting in rivers and streams). During the 2005–2006 sampling season, 40–50 otolith pairs of smolts were collected from each hatchery and preserved for stable isotope analyses (Gao and Bean 2008). Otolith samples of sockeye salmon were collected from Lake Ozette near the western Washington coast (cf. Fig. 1), from either salmon spawners or carcasses during field surveys in 2000 and 2002, respectively (MFM 2000). The 2000 otoliths (67 in total) were sampled with known origin (either wild or hatchery, based primarily on the sampling location and fin-clip markers), and were used to establish an isotopic standard for the mixed

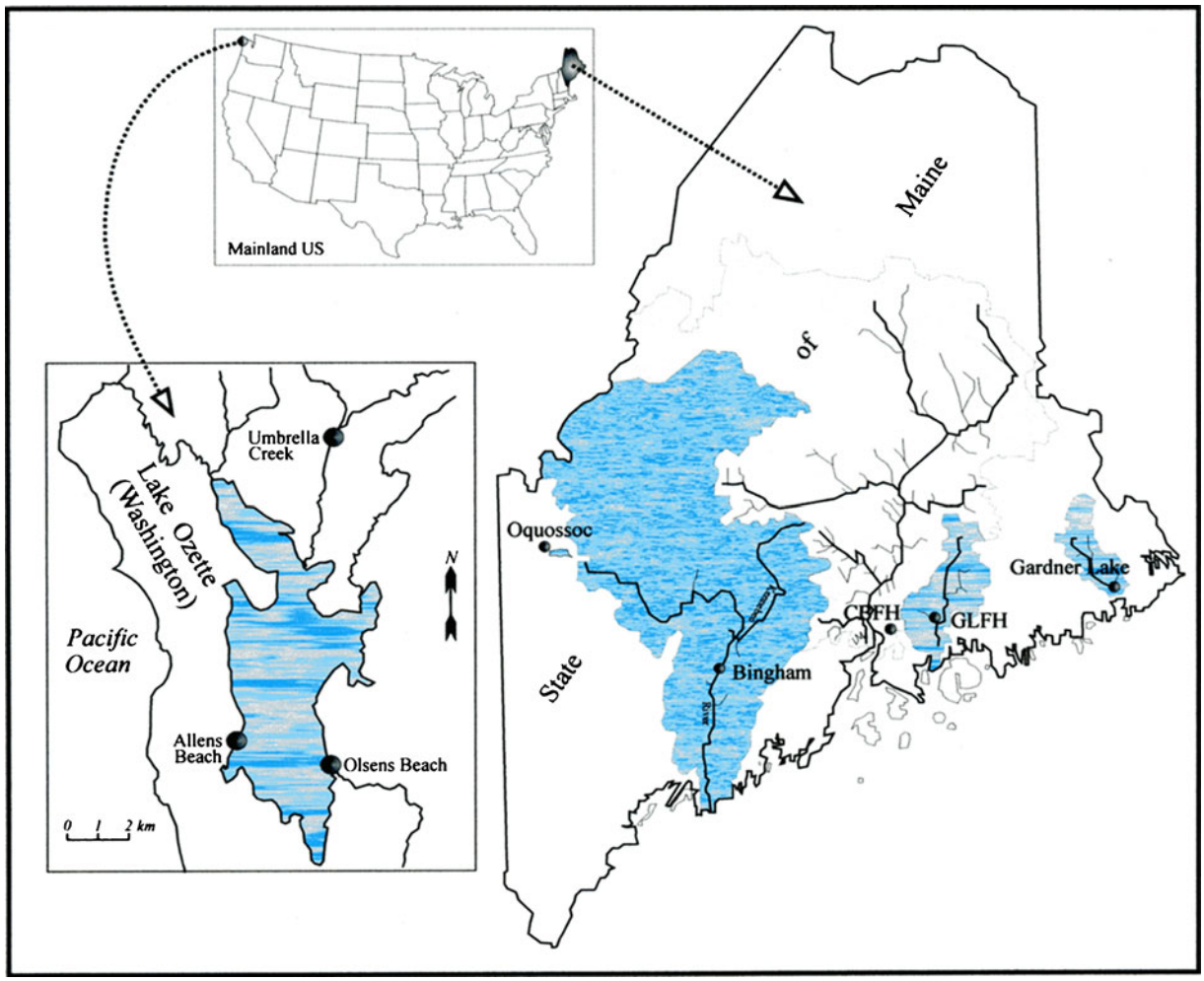

Fig. 1 Location map showing the otolith collections for Atlantic salmon (*Salmo salar*) in the State of Maine and sockeye salmon (*Oncorhynchus nerka*) in Lake Ozette at Washington State, respectively. There are three or four watersheds for the five hatcheries. The three private hatcheries (Oquossoc, Bingham, and Gardner Lake) belong to Cooke Aquaculture, whereas the Craig Brook and Green Lake National Fish Hatchery (CBFH and GLFH) are operated by the United States Fish and Wildlife Service (cf. Gao and Bean 2008). The Lake Ozette sockeye salmon are in a closely-related water system, and the lake beaches and tributary creeks are only about 6 km apart and directly connected in water flow (cf. MFM 2000)

origin samples. The 2002 otoliths (92 of mixed origin) were used for a blind test to evaluate the statistical method used for assignment.

Otolith samples were prepared in the Otolith Laboratory of Makah Fisheries Management at Neah Bay, Washington. The cleaned Atlantic salmon smolt otoliths were put into a jade mortar and ground with a pestle; and then we carefully collected the powder using a thin flat blade. When a sample was finished, the sampling tools (including mortar, pestle, and blade) were cleaned with a 10 % hydrochloric acid solution and rinsed with distilled water and wiped with Kimberly-Clark tissue-paper. Microsampling for sockeye salmon otoliths was conducted by using the Dremel method (Gao 1999), and one aragonite powder sample was taken from the nucleus of each otolith section. At least 50 μg of powder material was extracted from the otolith surface. Once a sample was taken, the powder was carefully tapped onto aluminum foil and placed into a metal cup. The otolith section and the sampling bit were subsequently cleaned using an Aero-Duster gas.

Analysis of otolith powder samples was performed in the Stable Isotope Laboratory at the University of Michigan Ann Arbor, using a Finnigan MAT Kiel preparation device that is coupled directly to the inlet of a Finnigan MAT 251 triple-collector gas-ratio mass spectrometer. All the isotope ratio measurements were reported in the standard δ notation (‰): $\delta^{18}O = \{[(^{18}O/^{16}O)_A / (^{18}O/^{16}O)_S] - 1\} \times 1000$, for instance, where A is otolith aragonite sample and S is an international standard (VPDB, Vienna Peedee belemnite). Calibration of isotopic enrichments to VPDB standard is based on daily analysis of NBS-19 (National Bureau of Standards) powdered carbonate and the analytical precision is better than 0.08‰ for both $\delta^{18}O$ and $\delta^{13}C$ values.

Statistical properties of the isotopic data

To explore the statistical properties of the isotopic data, a naïve examination was first performed on a set of 244 Atlantic salmon smolt otoliths without prior knowledge of the origin of the samples. Simple descriptive methods were used to characterize these data with frequency histograms and scatter plots. The sample mean ($\overline{X}$), standard deviation (SD), range, and the coefficient of variation ($CV=[SD/\overline{X}] \times 100$ %) were used to describe the data and its variability. The same analysis was also used for the set of 159 sockeye salmon otoliths.

During the naïve otolith data analyses, we were concerned about both the distribution of the Atlantic salmon and sockeye salmon data, because a multimodal distribution would provide evidence of the presence of more than one stock (hatchery), and the normality of the distributions as many methods of statistical analysis assume multivariate normality. The Shapiro-Wilk W statistic (Zar 1974) and Kolmogorov-Smirnov test (Conover 1980) were used to compare the distributions of the data to the normal distribution. Standard measures of skewness and kurtosis (Zar 1974), and their standard errors, were calculated for the distributions of $\delta^{18}O$ and $\delta^{13}C$ for all data and for each hatchery. In addition, methods described in Cochran (1977) were used to estimate sample sizes for continuous data based on the variability of the data relative to $\overline{X}$ and desired levels of precision ($\leq \pm 5.0$ %) and accuracy.

Discriminant analyses

Based on the distribution patterns of the isotopic data, we explored two methods to evaluate whether the $\delta^{18}O$ and $\delta^{13}C$ data could be used to accurately identify the different hatcheries for Atlantic salmon or separate origins (wild or hatchery) for sockeye salmon. Linear discriminant function analysis (LDFA, Manly 1994) was evaluated as a method for estimating a predictive model for group membership for both Atlantic salmon and sockeye salmon. The LDFA model is composed of a set of discriminant functions based on linear combinations of the predictor variables that provide the best discrimination between the groups (Dillon and Goldstein 1984). Cases are assumed to be independent. Predictor variables are assumed to have a multi-variate normal distribution, and group membership is assumed to be mutually exclusive (no case belongs to more than one group) and collectively exhaustive (all cases are members of a group). LDFA is a parametric method that assumes (among other things) a multi-variate normality of the data and equal variance-covariance matrices among groups. Box's M test (Box 1949) was used to test for the equality of group covariance matrices.

If the otolith isotopic data were not normally distributed and variance-covariance matrices were not equal among groups, LDFA was not appropriate without some transformation of the data to address heterogeneity of the group variance-covariance matrices. However, we had concerns about the multi-variate

normality of the isotopic data so we used *k*-sample nearest neighbor discriminant analysis (*k*NNA). The method was used by Bickford and Hannigan (2005) in identifying two hatchery stocks and a possible wild stock of walleye (*Sander vitreus*) with otolith chemistry data. The *k*NNA is a non-parametric method that does not require multi-variate normality of the data or the assumption of equal group variance-covariance matrices. We assumed equal prior probabilities in the group assignment during each analysis. The performance of models based on the 3, 4, and 5 nearest neighbors was evaluated. The Euclidean distance was used to measure the similarity of cases. For the sockeye salmon data, approximately half the samples of known origin were used to train the *k*NNA model and the remaining observations were withheld and used to evaluate the assignment accuracy of the model. A final model was trained using all of the data of known origin and used to estimate the samples of mixed origin. All analyses were conducted using the PASW statistical package (SPSS 2010).

Results

Atlantic salmon otoliths

The $\delta^{18}O$ values for the Atlantic salmon smolt otoliths ranged from −8.9 to −4.9‰, while the $\delta^{13}C$ of the same otoliths ranged from −12.1 to −9.0‰. Among five hatcheries, CBFH had the highest $\overline{X}$ in both $\delta^{18}O$ and $\delta^{13}C$ values (Table 1). Both Oquossoc and Bingham hatcheries had the lowest $\overline{X}$ in $\delta^{18}O$ and $\delta^{13}C$. The hatchery-specific analysis provided very precise estimates of the means and within-hatchery variation was relatively small. All *CV*s were less than 5 % and most were less than 2 % for both $\delta^{18}O$ and $\delta^{13}C$. Frequency histograms for all hatchery data combined indicated that the distributions appeared to be multimodal: the $\delta^{18}O$ distribution appeared tri-modal and the $\delta^{13}C$ distribution showed bi-modality (Fig. 2). For individual hatcheries, frequency histograms of $\delta^{18}O$ showed a distinct separation from Bingham to CBFH (Fig. 3a), whereas the $\delta^{13}C$ values indicated two overlapped groups relative to the normal distribution (Fig. 3b). Not surprisingly, the Kolmogorov-Smirnov (K-S) and Shapiro-Wilk (S-K) tests for normality were both significant ($P<0.001$) for each isotope, indicating that the isotopic data of Atlantic salmon smolt otoliths were not normally distributed (Table 2). Even though the data appear to be multimodal, the variability relative to the mean as measured by the *CV* is not extreme. The relative precision (= [half-width of confidence interval /$\overline{X}$] x 100 %) of the 95 % confidence interval for the isotope means is ±1.95‰ and ±0.97‰ for $\delta^{18}O$ and $\delta^{13}C$, respectively. Due to this relatively low variability, sample size requirements to estimate means with ±5.0 % relative precision or better were not large (Table 3). Because of greater between-group variability, $\delta^{18}O$ had larger sample size requirements than $\delta^{13}C$ for the combined hatchery data. However, because of smaller within-

Table 1 Summary statistics for isotopes of Atlantic salmon (*Salmo salar*) otoliths by hatchery

Hatchery	N	$\overline{X}$ (‰)	*SD*	*CV* (%)	Range (‰)
$\delta^{18}O$:					
Oquossoc	52	−8.33	0.128	1.5	−8.58 to −8.10
Bingham	41	−8.63	0.147	1.7	−8.88 to −8.17
CBFH	53	−5.51	0.177	3.2	−5.84 to −4.93
GLFH	51	−7.98	0.144	1.8	−8.23 to −7.69
Gardner	47	−7.07	0.157	2.2	−7.31 to −6.70
$\delta^{13}C$:					
Oquossoc	52	−11.06	0.214	1.9	−11.69 to −10.60
Bingham	41	−11.15	0.489	4.4	−12.12 to −10.26
CBFH	53	−9.33	0.146	1.6	−9.64 to −9.01
GLFH	51	−10.70	0.191	1.8	−11.11 to −10.14
Gardner	47	−9.65	0.239	2.5	−10.16 to −9.26

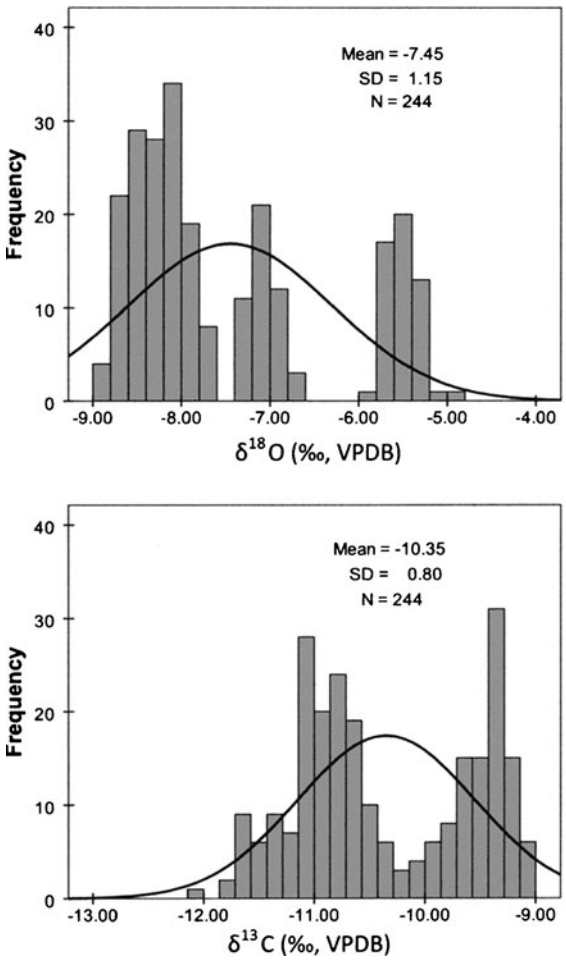

Fig. 2 Frequency histograms of $\delta^{18}O$ and $\delta^{13}C$ values for Atlantic salmon (*Salmo salar*) smolt otoliths relative to the normal distribution

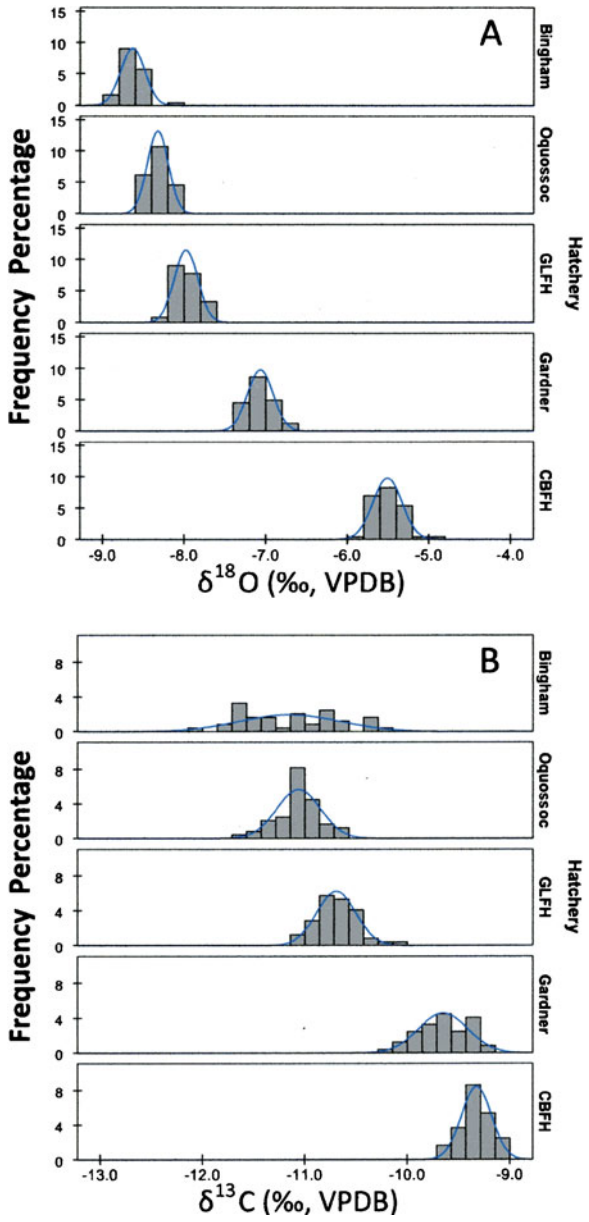

Fig. 3 Frequency histograms of otolith data relative to the normal distribution for the five Atlantic salmon (*Salmo salar*) hatcheries. **a** $\delta^{18}O$ values; and **b** $\delta^{13}C$ values

group variability, $\delta^{18}O$ generally had smaller sample size requirements than $\delta^{13}C$ when the data are considered separately by hatchery.

As reported previously (cf. Gao and Beamish 1999), values of $\delta^{18}O$ and $\delta^{13}C$ were highly correlated and there was generally a good separation of four of the five hatcheries (Fig. 4). The linear correlation coefficient for the combined data was 0.877. Although there is a significant ($P<0.001$) correlation between $\delta^{18}O$ and $\delta^{13}C$, it appears that it is not a strictly linear relationship as the cluster of data points associated with Gardner Lake Hatchery deviates from the line.

Using $\delta^{18}O$ and $\delta^{13}C$ as the discriminating variables, Box's M test for the equality of group covariance matrices was significant ($P<0.005$). Thus the LDFA was not appropriate and the method of *k*NNA was chosen for analysis of Atlantic salmon otoliths. The *k*NNA based on the four nearest neighbors provided the highest mean classification accuracy. The classification accuracies for the withheld samples for the *k*NNA model were relatively high for a five-group

Table 2 Significance (*p*) of the Kolmogorov-Smirnov (K-S) and Shapiro-Wilk (S-K) tests comparing the distribution of $\delta^{18}O$ and $\delta^{13}C$ data to the normal distribution and measures of skewness and kurtosis for Atlantic salmon (*Salmo salar*)

Hatchery Group	N	K-S test (*p*)	S-W test (*p*)	Skewness	Std. Error of Skewness	Kurtosis	Std. Error of Kurtosis
$\delta^{18}O$:							
All data	244	<0.001	<0.001	0.769	0.156	−0.846	0.310
Oquossoc & Bingham	93	0.169	0.024	−0.117	0.250	−0.983	0.495
CBFH & GLFH	104	<0.001	<0.001	−0.027	0.237	−1.966	0.469
Gardner	47	0.200	0.069	0.570	0.347	−0.334	0.681
$\delta^{13}C$**:**							
All data	244	<0.001	<0.001	0.094	0.156	−1.339	0.310
Private	140	<0.001	<0.001	0.342	0.205	−1.196	0.407
Federal	104	<0.001	<0.001	−0.075	0.237	−1.819	0.469

problem (Table 4). Mean unweighted classification accuracy for the five hatcheries was 89 %. Not surprisingly given Fig. 3, classification accuracies for CBFH and Gardner were 100 %. Oquossoc had the lowest classification accuracy (77 %) and was misclassified as either Bingham (8 %) or GLFH (15 %).

Sockeye salmon otoliths

The $\delta^{18}O$ values for the 159 sockeye salmon otoliths ranged from −6.6 to −3.1‰, whereas the related $\delta^{13}C$ values ranged from −14.1 to −9.5‰. The mean and variability for $\delta^{18}O$ were similar to those reported by Gao and Beamish (1999) for the freshwater portion of the sockeye salmon otoliths from the four stocks they analyzed. The frequency histograms for the hatchery and wild sockeye salmon showed that the distributions of isotopic data appeared uni-modal for both $\delta^{18}O$ and $\delta^{13}C$ (Fig. 5). Frequency histograms of $\delta^{13}C$ showed a larger overlap between the peaks of the two superimposed normal curves for the groups (hatchery and wild, from −11.0‰ to −12.0‰) than that of $\delta^{18}O$ (from −4.25‰ to −4.75‰), indicating a possible bimodal distribution for $\delta^{13}C$ (Fig. 6). However, the correlation of $\delta^{18}O$ and $\delta^{13}C$ for sockeye salmon otoliths showed two different relationships: the linear correlation for the hatchery fish was 0.506 compared to the linear correlation for the wild fish of 0.907 (Fig. 7). Overall, the distribution of the bivariate data from the 92 otoliths of mixed origin was between the two groups. Because the Box's M test for the equality of group covariance matrices was significant ($P<0.005$), the LDFA was also not appropriate.

For analysis of sockeye salmon otolith nuclei, the *k*NNA based on the three nearest neighbors provided the highest mean classification accuracy. Mean unweighted classification accuracy for the wild and hatchery groups was 92 % (Table 5). This model was then used to classify the 92 otolith samples of mixed origin. The results of this classification estimated that 45 of the samples were of wild origin and 47 of the samples were of hatchery origin, which is very close to the determination of origin based on location of sampling and fin-clip identifications (MFM 2000).

Table 3 Sample sizes required to estimate $\delta^{18}O$ and $\delta^{13}C$ means of Atlantic salmon (*Salmo salar*) otoliths based on two kinds of variability for three levels of relative precision

95 % Confidence Interval Half-width Size	$\delta^{18}O$	$\delta^{13}C$
a) Between-hatcheries:		
± 1.0 %	925	231
± 2.5 %	150	39
± 5.0 %	39	12
b) Within-hatchery:		
± 1.0 %	42	76
± 2.5 %	9	14
± 5.0 %	4	8

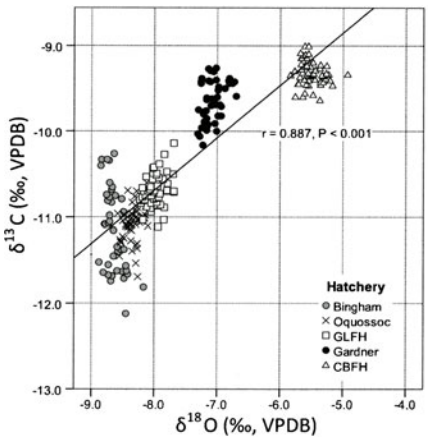

Fig. 4 Relationship between $\delta^{18}O$ and $\delta^{13}C$ for Atlantic salmon (*Salmo salar*) smolt otoliths with hatchery sources indicated. The linear relationship for the correlation of $\delta^{18}O$ and $\delta^{13}C$ is also shown

Discussion

The evaluation of the statistical properties of isotopic data and the extraction of unlocked information from otoliths are of paramount importance to fisheries management and otolith geochemistry. Otoliths are laminated calcium carbonate structures ($CaCO_3$) located in the inner ears of teleost fish (Carlstrom 1963) and the early fundamental work can be traced back to the late-1960s (Devereux 1967; Degens et al. 1969). The theory and practice of using otoliths as proxies are directed by Urey's (1947) hypothesis and the later experiments that calcium carbonates are precipitated in oxygen isotopic equilibrium with the surrounding waters in which the organism (including fish) lived, and thus retain the isotopic records of life history of the animal (e.g., Epstein et al. 1953; Friedman and O'Neil 1977; Grossman and Ku 1986; Patterson et al. 1993). Stable oxygen isotope ratios ($^{18}O/^{16}O$) of otoliths can provide unique information about water temperature and habitat conditions; while carbon isotope ratios ($^{13}C/^{12}C$) of otoliths reflect metabolic sources and dietary shifts (DeNiro and Epstein 1978; Mulcahy et al. 1979; Schwarcz et al. 1998). Thus the isotopic data of otoliths and correlation of $\delta^{18}O$ and $\delta^{13}C$ can depict the integration of ambient water and food conditions to which an individual fish was exposed (Keith and Weber 1965; Gao and Beamish 1999; Gao et al. 2010).

Our analyses based on otoliths of Atlantic salmon smolts and sockeye salmon showed that the statistical properties of isotopic data did contain important information on their populations. The $\delta^{18}O$ of Atlantic salmon showed tri-modality relative to the normal distribution, whereas the $\delta^{13}C$ of the same otolith samples showed a bi-modal distribution (cf. Fig. 2). These distribution patterns agree well with the principle of carbonate geochemistry and actual hatchery practices. From the theory of otolith studies (e.g., Kalish 1991; Schwarcz et al. 1998; Gao 2002), we know that the $\delta^{18}O$ of otoliths represents water conditions (e.g., temperature, salinity, and chemical composition) while the $\delta^{13}C$ of otoliths represents a fish's food conditions (food sources, dietary shift, and trophic-level shifts). The five Atlantic salmon hatcheries were located within three or four watersheds (depending on scales) in the region (cf. Fig. 1), and had two more different diets to feed the salmon juveniles or smolts. The three private hatcheries used the same feed but combined different manufacturers for early grow out diet and smolt diet, which have very

Table 4 Classification accuracies of withheld observations for each hatchery group in the 4-nearest neighbor analyses (*k*NNA) of Atlantic salmon (*Salmo salar*)

Hatchery	Classified Hatchery of Origin				
	Bingham	Oquossoc	GLFH	Gardner	CBFH
Bingham	80.0 %	20.0 %	0.0 %	0.0 %	0.0 %
Oquossoc	7.7 %	76.9 %	15.4 %	0.0 %	0.0 %
GLFH	0.0 %	12.0 %	88.0 %	0.0 %	0.0 %
Gardner	0.0 %	0.0 %	0.0 %	100.0 %	0.0 %
CBFH	0.0 %	0.0 %	0.0 %	0.0 %	100.0 %

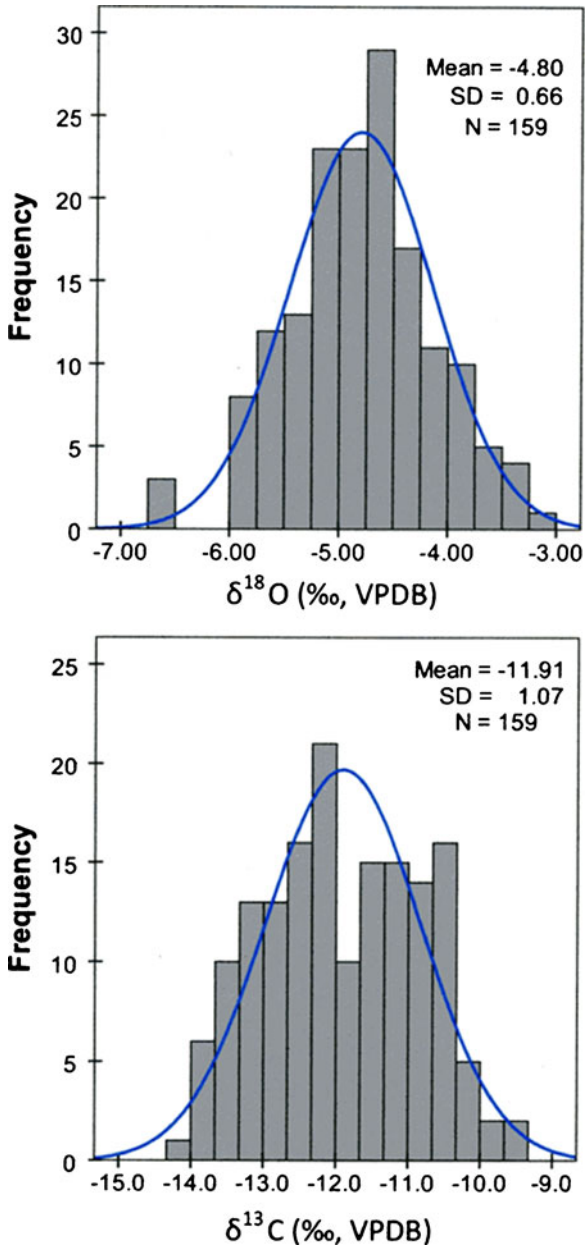

Fig. 5 Frequency histograms of $\delta^{18}O$ and $\delta^{13}C$ values for sockeye salmon (*Oncorhynchus nerka*) otoliths (wild, hatchery, and mixed-origin samples combined) relative to the normal distribution

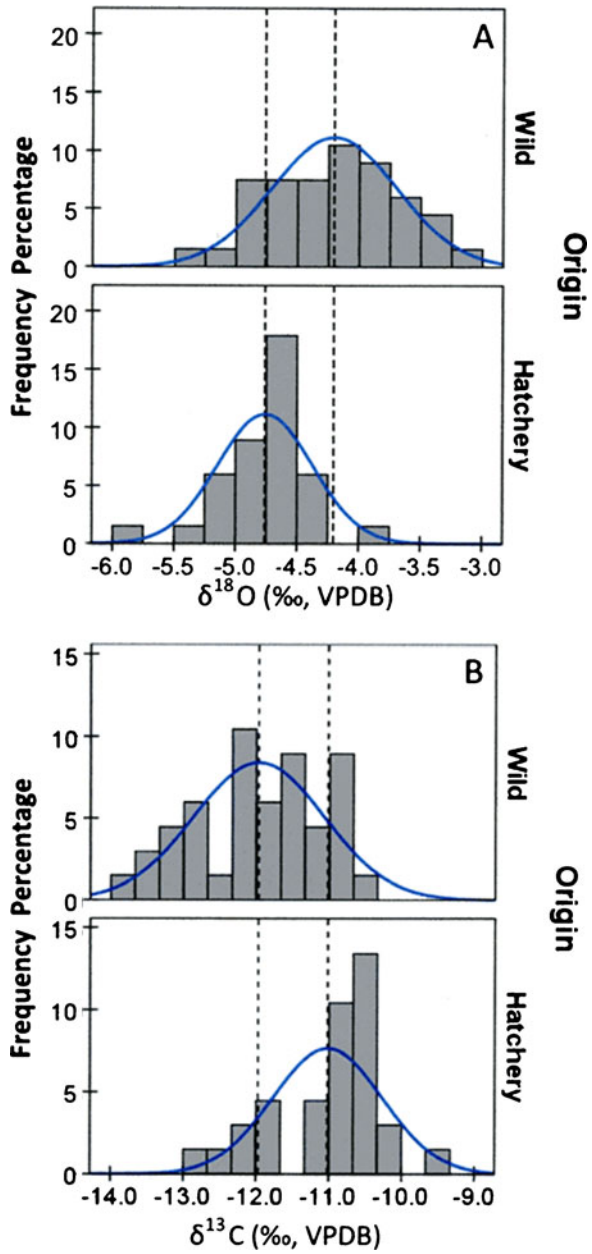

Fig. 6 Frequency histograms of otolith data (from both 2000 and 2002 samples) relative to the normal distribution for the wild and hatchery sockeye salmon (*Oncorhynchus nerka*). **a** $\delta^{18}O$ values; and **b** $\delta^{13}C$ values

similar protein, oil, and other ingredients (Mr. Greg Lambert, Manager for Cooke Aquaculture, pers. comm.). The two federal hatcheries purchased Nutra Fry from the same manufacturer Corey Feed Mills of Canada that have the same sources of ingredients and have similar fishmeal and oil concentrations (Mr. Fred Trasko, Assistant Manager at Green Lake Hatchery, pers. comm.). Because of mixing for the watersheds in the region and diet supply from the hatchery streams, there were some overlaps for the $\delta^{18}O$ and $\delta^{13}C$ compositions. Nevertheless, the tri-modal distribution of $\delta^{18}O$ of Atlantic salmon otoliths is supported by three discriminating watersheds in the Gulf of Maine, while

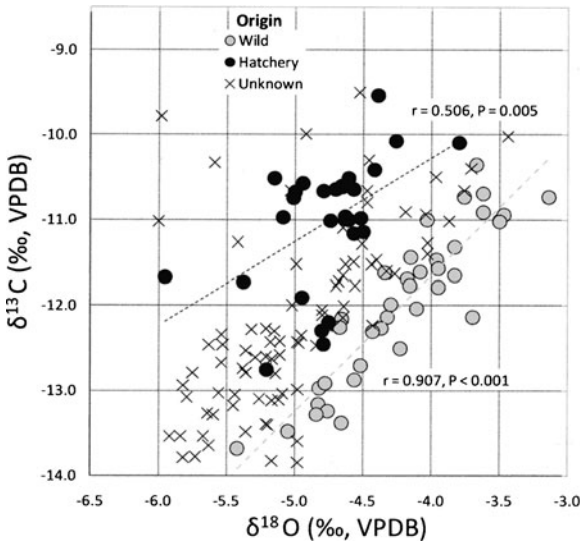

Fig. 7 Relationship between $\delta^{18}O$ and $\delta^{13}C$ for sockeye salmon (*Oncorhynchus nerka*) otoliths of wild, hatchery, and mixed origin. The linear relationships between $\delta^{18}O$ and $\delta^{13}C$ are presented for the hatchery and wild sockeye salmon group

the bi-modal distribution of $\delta^{13}C$ is supported by the fact that the five hatcheries had two discriminating foods (either two different feeds from commercial sources or feeds and diet from water sources). Slight differences may be explained by the hatchery practice that some of the hatcheries used different feeds and formulations during different growing stages (juvenile or smolt). For the isotopic data of sockeye salmon otoliths, the uni-modal distribution is supported by the fact that no $\delta^{18}O$ differences had been found between tributary and lake-beach waters (Gao, unpubli. data) because the water flow of tributary and lake beaches is directly connected (cf. Fig. 1).

Table 5 Classification accuracies of withheld observations for each group in the 3-nearest neighbor analyses (*k*NNA) of sockeye salmon (*Oncorhynchus nerka*)

Group of Origin	Classified Group of Origin	
	Natural	Hatchery
Natural	94.7 %	5.3 %
Hatchery	10.3 %	89.7 %
Unknown	48.9 %	51.1 %

Nevertheless, there are distinct differences in food sources between the hatchery and wild smolts, so that the $\delta^{13}C$ indicated a somewhat bi-modal distribution in the frequency histograms (cf. Fig. 6). These examples may only represent anadromous fishes, however, statistical analyses indicated that the otolith data of both Atlantic salmon and sockeye salmon were not normally distributed. The conclusion was also supported by Shapiro-Wilk and Kolmogorov-Smirnov tests for normality ($P<0.001$), with Box's M test used for the equality of group covariance matrices ($P<0.005$).

If the above results are applicable to other fish species, previous otolith chemistry studies, especially treating otolith isotopic data as traditional biological data and using parametric methods (e.g., ANOVA, *t*-test, and LDFA) for statistical analyses, should be re-examined; as the parametric methods must meet the requirement of the normal distribution. If one did not test isotopic data for normality, non-parametric analysis may be acceptable as an alternative as it has a distribution free assumption. In our study, we used *k*NNA to discriminate among five Atlantic salmon hatcheries and between hatchery and wild sockeye salmon. Another concern is sample size requirements because it is always a question for managers when making decisions for a project and developing budget proposals. Kalish (1991) mentioned, albeit briefly, that the appropriate sample size for oxygen isotope analysis was based on the mass spectrometer measurement error and the variation of the otolith $\delta^{18}O$ means. How many samples do we need for otolith analysis? We have demonstrated that for Atlantic salmon raised in the hatchery environment, relatively small samples are required to estimate hatchery means with high relative precision ($\leq \pm 5.0$ %). In particular, because of greater variability between hatcheries, the $\delta^{18}O$ had larger sample size requirements than $\delta^{13}C$ for the combined hatchery data (cf. Table 3). However, because of smaller variability within each hatchery, the $\delta^{18}O$ generally had smaller sample size requirements than $\delta^{13}C$ when the data were considered separately by hatchery. Therefore, previous isotopic studies based on several otoliths may be adequate for investigations in fish behavior, but not enough for stock or population studies in management, even for hatchery-reared fish.

Two critical issues have been a long-standing challenge to Pacific salmon hatchery practice and

management: (1) how to identify the origin of hatchery and wild salmon without biological markers (e.g., fin-clips, cold-wire tags), and (2) how to evaluate the different hatcheries within the same river-system or watershed? We have demonstrated that stable isotopic data from otoliths and appropriate statistical analyses can provide answers to both these questions. Using the *k*NNA method, we identified the origin of hatchery and wild sockeye salmon with an accuracy of 92 %. Using this analysis on a set of samples of mixed origin, there was a good agreement with the origin of fish determined from the location of field sampling and fin-clip markers. The accuracy of identification could be improved because the spawning areas of sockeye salmon in Lake Ozette are limited to only 6 km between lake beaches and tributary creeks (MFM 2000). For the more geographically distant Atlantic salmon hatcheries, statistical analyses of otolith data separated the private and federal hatcheries using *k*NNA with a mean classification accuracy of 89 %. Therefore, as reported for other fish species (e.g., Gao et al. 2001; Gao and Beamish 2003a; Bickford and Hannigan 2005), otolith isotopic data and statistical analyses appear to constitute a natural marker useful for identifying the sources of anadromous fishes, either from different origins (hatchery or wild) or from different hatcheries within a watershed.

One by-product from statistical properties of isotopic data in this study, fortunately, is the non-linear relationship for the correlation of $\delta^{18}O$ and $\delta^{13}C$ in Atlantic salmon smolt otoliths (cf. Fig. 4). A linear correlation had been documented by McConnaughey (1989a, b) as an indicator of kinetic isotopic fractionation or metabolic effect, particularly in rapidly growing carbonate skeletons. Gao et al. (2010) argued that the differences between mollusks and fishes may not be only due to growth rate, and that the correlation of $\delta^{18}O$ and $\delta^{13}C$ is not dominated by a "kinetic" effect like corals. For sockeye salmon otoliths, in fact, the hatchery and wild fish showed two significantly different linear relationships (cf. Fig. 7).

In summary, we have demonstrated that using *k*NNA methods we can discriminate among the different Atlantic salmon hatcheries in the State of Maine and between the hatchery and wild sockeye salmon in Lake Ozette with high accuracy. None of the Atlantic salmon smolt and sockeye salmon otolith data were normally distributed, so that the LDFA and the commonly-used parametric statistical tests may not be appropriate. The statistical properties of isotopic data appear consistent with the theory of carbonate geochemistry and the inherent relationship of $\delta^{18}O$ and $\delta^{13}C$ in otoliths, and are related to the sample size required to estimate the $\delta^{18}O$ and $\delta^{13}C$ means. Thus we conclude that the features of otoliths through the statistical power of the $\delta^{18}O$ and $\delta^{13}C$ data can be used to discriminate different hatcheries in a watershed and the different origins of hatchery and wild anadromous fishes.

Acknowledgements We are indebted to Cooke Aquaculture: Oren Kephart from Gardner Lake hatchery, Greg Lambert from Oquossoc hatchery, and Brian Wheeler from Bingham hatchery; and the United States Fish and Wildlife Service: Fred Trasko from GLNFH and Tom King from CBNFH, for providing Atlantic salmon otoliths or smolts in this study. We thank former Makah Fisheries Management staff Mike Haggerty, Andy Ritchie, and Caroline Peterschmidt for providing sockeye salmon otoliths. Many colleagues and scientists made comments and suggestions that greatly improved the quality of the manuscript. We are grateful to Lora Wingate at the Stable Isotope Laboratory, University of Michigan Ann Arbor, for her assistance with processing otolith samples in this study.

References

Bickford N, Hannigan R (2005) Stock identification of walleye via otolith chemistry in the Eleven Point River Arkansas. N Am J Fish Manag 25:1542–1549

Box GEP (1949) A general distribution theory for a class of likelihood criteria. Biometrika 36:317–346

Carlstrom D (1963) A crystallographic study of vertebrate otoliths. Biol Bull 124:441–463

Cochran WG (1977) Sampling techniques, 3rd edn. John Wiley and Sons, New York

Coghlan SM, Lyerly MS, Bly TR, Williams JS, Bowman D, Hannigan R (2007) Otolith chemistry discriminates among hatchery-reared and tributary-spawned salmonines in a tailwater system. N Am J Fish Manag 27:531–541

Conover WJ (1980) Practical nonparametric statistics, 2nd edn. John Wiley and Sons, New York

Degens ET, Deuser WG, Haedrick RL (1969) Molecular structure and composition of fish otoliths. Mar Biol 2:105–113

DeNiro MJ, Epstein S (1978) Influence of diet on the distribution of carbon isotopes in animals. Geochim Cosmochim Acta 42:495–506

Devereux I (1967) Temperature measurements from oxygen isotope ratios of fish otoliths. Science 155:1684–1685

Dillon WR, Goldstein M (1984) Multivariate analysis -methods and applications. John Wiley and Sons, New York

Edmonds JS, Fletcher WJ (1997) Stock discrimination of pilchards *Sardinops sagax* by stable isotope ratio analysis of otolith carbonate. Mar Ecol Prog Ser 152:241–247

Epstein S, Buchsbaum R, Lowenstam HA, Urey HC (1953) Revised carbonate-water isotopic temperature scale. Geol Soc Am Bull 64:1315–1326

Friedman I, O'Neil JR (1977) Compilation of stable isotope fractionation factors of geochemical interest. In: Fleischer M (ed) Data of Geochemistry. US Geol Surv Prof Paper 440-KK

Gao YW (1999) Microsampling of fish otoliths: a comparison between DM 2800 and Dremel in stable isotope analysis. Environ Biol Fish 55:443–448

Gao YW (2002) Regime shift signatures from stable oxygen isotopic records of otoliths of Atlantic cod (*Gadus morhua*). Isotopes Environ Health Stud 38:251–263

Gao YW, Beamish RJ (1999) Isotopic composition of otoliths as a chemical tracer in population identification of sockeye salmon (*Oncorhynchus nerka*). Can J Fish Aquat Sci 56:2062–2068

Gao YW, Beamish RJ (2003a) Stable isotopic composition of otoliths from tagged Pacific halibut, *Hippoglossus stenolepis*. Environ Biol Fish 67:253–261

Gao YW, Beamish RJ (2003b) Stable isotope variations in otoliths of Pacific halibut (*Hippoglossus stenolepis*) and indications of the possible 1990 regime shift. Fish Res 60:393–404

Gao YW, Bean D (2008) Stable isotopic analyses of otoliths in identification of hatchery origin of Atlantic salmon (*Salmo salar*) in Maine. Environ Biol Fish 83:429–437

Gao YW, Joner SH, Bargmann GG (2001) Stable isotopic composition of otoliths in identification of spawning stocks of Pacific herring (*Clupea pallasi*) in Puget Sound. Can J Fish Aquat Sci 58:2113–2120

Gao YW, Dettman DL, Piner KR, Wallace FR (2010) Isotopic correlation ($\delta^{18}O$ versus $\delta^{13}C$) of otoliths in identification of groundfish stocks. Trans Am Fish Soc 139:491–501

Gibson SM, Patterson WF, Phelps RP, Patterson WP, Chen Z (2010) Distinguishing wild from hatchery-produced juvenile red snapper with otolith chemical signatures. N Am J Fish Manag 30:1176–1186

Grossman EL, Ku T-L (1986) Oxygen and carbon isotope fractionation in biogenic aragonite: temperature effects. Chem Geol 59:59–74

Kalish JM (1991) Oxygen and carbon stable isotopes in the otoliths of wild and laboratory-reared Australian salmon (*Arripis trutta*). Mar Biol 110:37–47

Keith ML, Weber JN (1965) Systematic relationships between carbon and oxygen isotopes in carbonates deposited by modern corals and algae. Science 150:498–501

Manly BFJ (1994) Multivariate statistical methods—a primer, 2nd edn. Chapman and Hall/CRC, New York

McConnaughey T (1989a) ^{13}C and ^{18}O isotopic disequilibrium in biological carbonates: I Patterns. Geochim Cosmochim Acta 53:151–162

McConnaughey T (1989b) ^{13}C and ^{18}O isotopic disequilibrium in biological carbonates: II. In vitro simulation of kinetic isotope effects. Geochim Cosmochim Acta 53:163–171

Meyer-Richow VB, Cook I, Hendy CH (1992) How to obtain clues from otoliths of an adult fish about the aquatic environment it has been in as a larva. Comp Biochem Physiol 103A:333–335

MFM (Makah Fisheries Management) (2000) Lake Ozette sockeye hatchery and genetic management plan. Biological Assessment Section 7

Mulcahy SA, Killingley JS, Phleger CF, Berger WH (1979) Isotopic composition of otoliths from a bentho-pelagic fish, *Coryphaenoides acrolepis*, Macrouridae: Gadiformes. Oceanol Acta 2:423–427

Nelson CS, Northcote TG, Hendy CH (1989) Potential use of oxygen and carbon isotopic composition of otoliths to identify migratory and non-migratory stocks of the New Zealand common smelt: a pilot study. NZ J Mar Freshw Res 23:337–344

Patterson WP, Smith GR, Lohmann KC (1993) Continental paleothermometry and seasonality using the isotopic composition of aragonitic otoliths of freshwater fishes. In: Swart PK, Lohmann KC, McKenzie J, Savin S (eds) Climate change in continental isotopic records. Geophysical Monograph 78. pp. 191–202

Schwarcz HP, Gao YW, Campana S, Browne D, Knyf M, Brand U (1998) Stable carbon isotope variations in otoliths of cod (*Gadus morhua*). Can J Fish Aquat Sci 55:1798–1806

Sokal RR, Rohlf FJ (1981) Biometry. The principles and practice of statistics in biological research, 2nd edn. Freeman and Company, San Francisco

SPSS (2010) PASW Statistics 18. SPSS, Inc, Chicago

Stephenson PC, Edmonds JS, Moran MJ, Caputi N (2001) Analysis of stable isotope ratios to investigate stock structure of red emperor and Rankin cod in northern Western Australia. J Fish Biol 58:126–144

Urey HC (1947) The thermodynamic properties of isotopic substances. J Chem Soc 1947:562–581

Zar JH (1974) Biostatistical analysis. Prentice-Hall, Englewood Cliffs