Age validation of Dover sole (*Microstomus pacificus*) by means of bomb radiocarbon

Craig R. Kastelle (contact author)
Delsa M. Anderl
Daniel K. Kimura
Chris G. Johnston
Email address for C. R. Kastelle: Craig.Kastelle@noaa.gov
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Alaska Fisheries Science Center
7600 Sand Point Way NE.
Seattle, Washington 98115-6349

Abstract—We used bomb radiocarbon (^{14}C) in this age validation study of Dover sole (*Microstomus pacificus*). The otoliths of Dover sole, a commercially important fish in the North Pacific, are difficult to age and ages derived from the current break-and-burn method were not previously validated. The otoliths used in this study were chosen on the basis of estimated birth year and for the ease of interpreting growth zone patterns. Otolith cores, material representing years 0 through 3, were isolated and analyzed for ^{14}C. Additionally, a small number of otoliths with difficult-to-interpret growth patterns were analyzed for ^{14}C to help determine age interpretation. The measured Dover sole ^{14}C values in easier-to-interpret otoliths were compared with a ^{14}C reference chronology for Pacific halibut (*Hippoglossus stenolepis*) in the North Pacific. We used an objective statistical analysis where sums of squared residuals between otolith ^{14}C values of Dover sole and the reference chronology were examined. Our statistical analysis also included a procedure where the Dover sole ^{14}C values were standardized to the reference chronology. These procedures allowed an evaluation of aging error. The ^{14}C results indicated that the Dover sole age estimates from the easier-to-interpret otoliths with the break-and-burn method are accurate. This study validated Dover sole ages from 8 to 47 years.

Manuscript submitted 24 March 2008.
Manuscript accepted 13 June 2008.
Fish. Bull. 106:375–385 (2008).

The otoliths of Dover sole (*Microstomus pacificus*), a commercially important fish in the North Pacific, are difficult to interpret. Ages derived with the current otolith break-and-burn method have not been previously validated. The age data are important for population modeling and setting the total allowable catch (Stockhausen et al., 2005). The necessity of age validation studies is widely recognized (Beamish and McFarlane, 1983; Campana, 2001) and age validation has become the focus in an expanding number of studies at the Alaska Fisheries Science Center (AFSC) (e.g., Kastelle and Kimura, 2006; Kimura et al., 2006; Hutchinson et al., 2007; Kastelle et al., 2008).

Two methods of Dover sole otolith preparation are used by various agencies. At the AFSC; the Groundfish Program of the Pacific Biological Station, Nanaimo, B.C.; and the Northwest Fisheries Science Center, Dover sole are aged by the break-and-burn method (Chilton and Beamish, 1982) and are estimated to have a maximum age of 54 years. An alternative aging method employed at the Southwest Fisheries Science Center uses transverse thin sections of otoliths (Hunter et al., 1990). The maximum age estimated with thin sections is 58 years (Hunter et al., 1990). The two methods appear to produce similar results but any similarity has not been tested quantitatively, and neither method has been validated.

Workshops on the interlaboratory calibration of methods and on otolith interpretation for determining the age of Dover sole have been held periodically among agencies responsible for the management of this species. However, age validation has not been a focus of these workshops. Because the otoliths of this fish species are small and the species has a relatively long life expectancy, the precision of multiple readings on a sample of otoliths can be poor. The precision measured by percentage coefficient of variation (CV) between two age readers is reported to be 9.64% which is higher than that for most species aged at the AFSC (Kimura and Anderl, 2005). For many North Pacific Pleuronectiformes the CV is under 4% (Kimura and Anderl, 2005). The poor precision (high CV) is an indication of the difficulty in reading otoliths from Dover sole, and indicates that validating the accuracy of the ages is necessary.

The goal of this study was to use bomb-produced radiocarbon (^{14}C) to validate the ages of Gulf of Alaska (GOA) Dover sole determined by the otolith break-and-burn method. Otoliths were selected on the basis of estimated birth year and two descriptive categories: 1) otoliths with uniform growth zones that were easy to enumerate, and 2) otoliths with growth zones that were difficult to enumerate. The first category was used to validate general aging criteria and the second was used to help determine otolith interpretation. Otolith material corresponding to a time near birth was isolated by extracting the core up to the first 3 years and measured for its ^{14}C content. We analyzed the Δ^{14}C from the first category of otolith cores using statistical methods first reported by Kastelle et al. (2008).

Bomb radiocarbon age validation has been used on an increasing number of species and is considered one of the best methods to confirm the accuracy of fish ages (Campana, 2001). Recent uses in the North Pacific include that on the white shark (*Carcharodon carcharias*) (Kerr et al., 2006), quillback rockfish (*Sebastes malinger*) (Kerr et al., 2005), canary rockfish (*S. pinniger*) (Piner et al., 2005; Andrews et al., 2007), bocaccio rockfish (*S. paucispinis*) (Andrews et al., 2005; Piner et al., 2006), Pacific halibut (*Hippoglossus stenolepis*) (Piner and Wischniowski, 2004), and Pacific ocean perch (*S. alutus*) (Kastelle et al., 2008).

Radiocarbon fish age validation relies on a time reference provided by production of ^{14}C from atomic bomb testing. The above-ground testing of atomic bombs that introduced ^{14}C into the atmosphere and marine environment began in the 1950s and continued into the 1960s (Kalish, 1993; Nydal, 1993). This caused a rapid increase in marine ^{14}C lasting through about 1970—an increase that is recorded in calcified marine organisms and otoliths and provides a necessary time reference. To validate ages from a "validation species" (in this case Dover sole), a ^{14}C "reference chronology" is used, where the exact time frame of the ^{14}C increase is considered known. Two reference chronologies have been developed for the North Pacific Ocean: one from Pacific halibut (Piner and Wischniowski, 2004) and one from yelloweye rockfish (*S. ruberrimus*) (Kerr et al., 2004). The posited birth years for the validation species are calculated from ages estimated by otolith growth zone counts and date of collection. Specimens representing the validation species are chosen such that the range of posited birth years spans the period of rapid marine ^{14}C increase. Otolith core material deposited in the first one or two years of life from the validation species is analyzed and each core provides one ^{14}C data point. To evaluate the ages, the ^{14}C from the cores of the validation species is plotted with respect to the posited birth years and compared to the known ^{14}C values in the reference chronology. If there is a timing difference between the ^{14}C increase in the validation species and the reference chronology, then the estimated ages of the validation species are often assumed to be in error. Alternatively, if a timing difference is not present, the ages from the validation species are considered accurate. In a recent bomb radiocarbon age validation study of Pacific ocean perch, a series of new procedures was used to compare the ^{14}C measurements in the validation samples to the reference chronology (Kastelle et al., 2008). We used the same methods here—purposely biasing the ages to be validated by ±0, 1, 2, and 4 years; standardizing the validation sample ^{14}C values to the reference chronology; and evaluating the residuals between the validation samples and the reference chronology to see if inaccuracies in the age estimates were present.

There are two important assumptions when validating fish ages with the bomb radiocarbon method (Campana and Jones, 1998; Piner and Wischniowski, 2004; Piner et al., 2005; Kastelle et al., 2008). Assumption 1 is that the validation species must be biologically and environmentally similar to the species in the reference chronology during the first years of life. If both species are receiving their ^{14}C from the same sources, then the magnitude and timing of the ^{14}C increase should be similar (Andrews et al., 2007). A reference chronology based on the same species as that being investigated is best, and occasionally available (Campana, 1997; Campana et al., 2002; Piner and Wischniowski, 2004). Assumption 2 requires that the otolith core used for each ^{14}C analysis be uncontaminated and that it constitute a closed system. Therefore, an accurate extraction of the core without contamination from other carbon sources or different years is necessary. Dover sole otoliths presented a unique challenge in this regard because of their small size. For further information regarding radiocarbon age validation studies, one can consult the earlier mentioned studies from the North Pacific along with Kalish (1993, 1995) and Campana (1997).

Materials and methods

Otolith selection and coring procedures

The Gulf of Alaska (GOA) Dover sole otoliths used in this study were collected either during AFSC survey cruises or by AFSC fishery observers aboard commercial vessels. The survey cruises took place in 1984 and 2005; the otoliths were removed from the fish at sea, stored in a glycerin and thymol mixture, and archived for future age determination. The specimens collected from commercial harvests were caught in 1998 and treated similarly, except they were first stored dry for up to 3 months before storage in a glycerin and thymol mixture. The glycerin and thymol mixture is not expected to be a contaminant in otolith ^{14}C measurements (Campana et al., 2003).

After the archival period, the otoliths were aged at the AFSC for stock assessments. The initial ages were determined by the break-and-burn method (Chilton and Beamish, 1982) with the blind-side otolith. Assumed annual growth zones were counted by enumerating the translucent zones. The otolith growth over the course of one year is assumed to consist of an opaque zone and a translucent zone. After growth zones were read, the otoliths were archived again for varying durations up to 14 years. Samples where the initial age estimate placed the birth year near the era of marine ^{14}C increase were re-examined by age readers experienced in the interpretation Dover sole otolith growth zones, and considered for possible ^{14}C measurement. In the re-examination process otoliths were re-aged to assign a "final age" and were placed into two subjective categories based on the ease of interpretation of the growth zones:

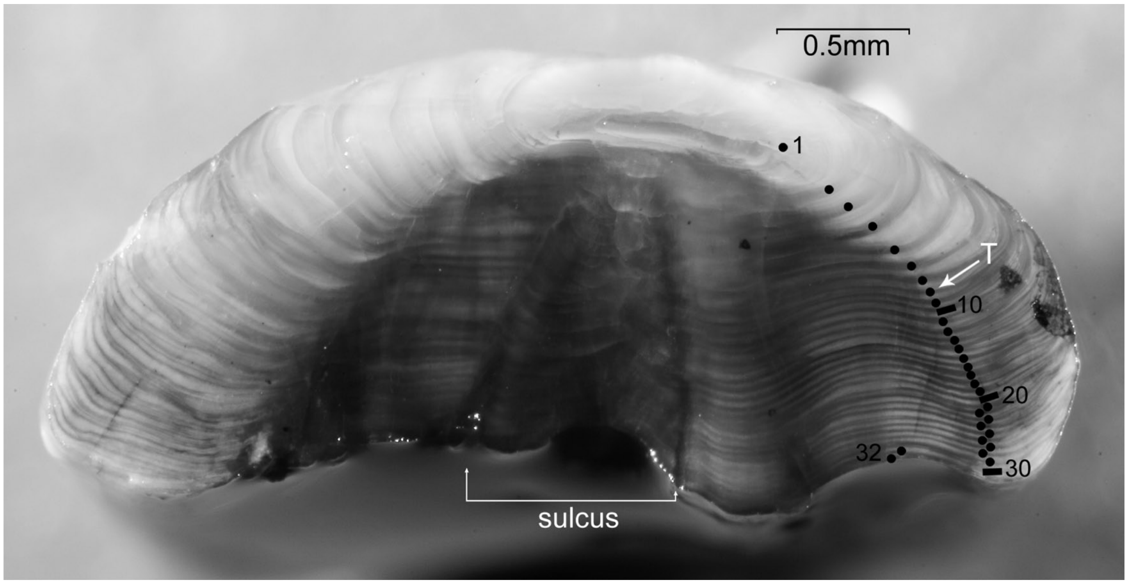

Figure 1

Category-1 otolith (specimen number 82) from Dover sole (*Microstomus pacificus*) treated by the break-and-burn method. The series of dots indicate the translucent zones that were counted to estimate a final age of 32 years. The "T" arrow points to the transition region from fast early growth to slower later growth. The sulcus region of the otolith is indicated as a reference point.

1 Category-1 otoliths had clear growth zones that were easy to interpret, in that most of the translucent zones appeared with minimal or no splitting in the proximal growth axes in at least one region (dorsal or ventral) on the break-and-burn cross section (Fig. 1). Splitting is defined as the branching of a single translucent zone into two or more translucent zones. Typically the translucent zones were spaced evenly, but with decreasing intervals, as the fish became older. Some samples in this category may have presented interpretative options, and different reading axes in the cross section could be chosen, but typically only small differences in age (of 1 or 2 years) would result.

2 Category-2 otoliths had growth zones that were difficult to interpret and that made these fish difficult to age. Many translucent zones had obvious splits and uneven spacing (Fig. 2). Widely different ages could be generated depending upon the interpretation chosen and which reading axis was used.

A sample selection process for ^{14}C measurement occurred after re-examination of the ages, and this selection process relied on two factors. First, in category 1 our intention was that the initial age and final age should agree within 3 years. For otoliths in category 2, which are more common for Dover sole than are category-1 otoliths, we did not use age agreement as a selection factor. Second, the estimated birth years, which were based on the final ages, had to be evenly distributed from about 1951 to 1977, to bracket the era of marine ^{14}C increase. For several specimens in category 1, the first factor was relaxed to evenly populate the years of marine ^{14}C increase. This process provided 43 specimens for Δ^{14}C analysis: 38 otoliths in category 1 and 5 in category 2 (Table 1). The range in catch years from 1984 to 2005 generated a large range in Dover sole ages for potential validation. In the remainder of this article, unless specified differently, the ages referred to are these final ages.

For the selected Dover sole specimens, the otolith core was extracted from the remaining whole eyed-side otolith for ^{14}C analysis. The core represented material deposited only in the first three years of life. Previous studies have often used a 1-year core (e.g., Campana, 1997), but that was not possible with Dover sole because their otoliths are very small, therefore, the minimum mass required for ^{14}C analysis mandated a 3-year core. For the otolith coring procedure we used a Buehler® EcoMet® (Buehler Ltd., Lake Bluff, IL) grinder with 320 grit wet or dry sandpaper to first remove otolith material on the proximal surface (the main growth axis in older otoliths). On some otoliths a small amount of material was also removed on the distal surface. Next, the grinder was used to remove material on the perimeter, in the anterior-posterior and dorsal-ventral axes, beyond the third year. After the exterior layers were removed with this procedure, the location of the third year's growth zone in each otolith became easier to see, and its location served as a primary guide in the coring

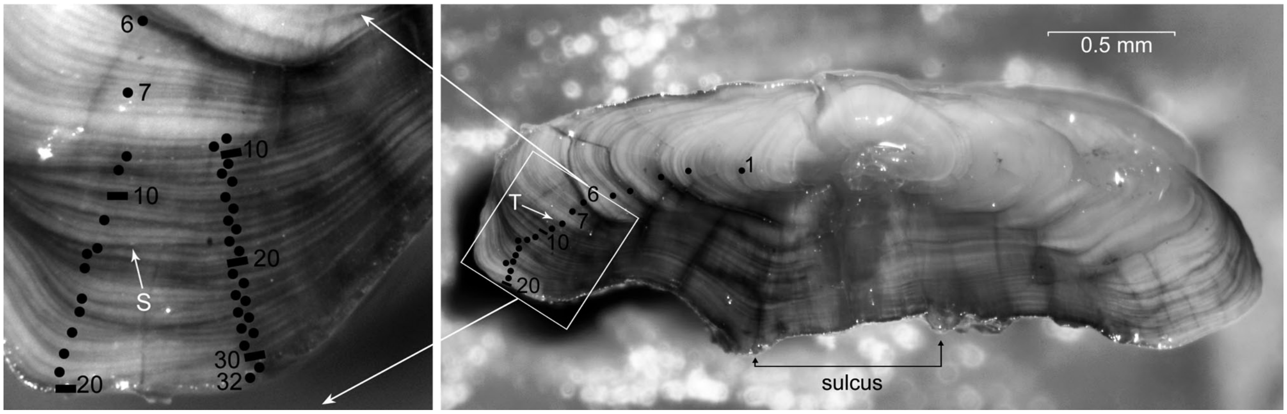

Figure 2

Category-2 otolith (specimen 225) from Dover sole (*Microstomus pacificus*) treated by the break-and-burn method. The "T" arrow points to the transition region from fast early growth to slower later growth. In the enlarged section, the two series of dots represent different options of interpreting the translucent zones; the final age of 20 years and the maximum age of 32 years are shown. The "S" points to a translucent zone that splits to form two translucent zones. The sulcus region of the otolith is indicated as a reference point.

process. Each core was tailored to the size and shape of the apparent third year. A secondary guide for the coring process was the weight and size of otoliths from 3-year-old fish in the 2005 survey collection, and an additional survey collection (not used for ^{14}C measurements) from 2003. These otoliths from 3-year-old fish had an average weight of 9.5 mg (±0.4 mg standard error) and an average size of $2.41 \times 3.84 \times 0.68$ mm (n=21). The core would occasionally break into several pieces as material was removed, but all salvageable pieces were retrieved and used. The core size was recorded for all intact cores and compared to otoliths from young fish. Finally, the cores were cleaned in an ultrasonic cleaner, dried, weighed, and then stored in acid-washed vials before ^{14}C analysis.

^{14}C analysis

The ^{14}C and ^{13}C of Dover sole otolith cores were measured at the National Ocean Sciences Accelerator Mass Spectrometry Facility at the Woods Hole Oceanographic Institution, Woods Hole, MA. Samples were treated at the Woods Hole Oceanographic Institution with a routine acid hydrolysis procedure to produce a graphite target, and analyzed with accelerator mass spectrometry. We report results as Δ^{14}C, which is defined in Stuiver and Polach (1977) as the relative difference between an international standard (base year 1950) and sample activity. The Δ^{14}C is normalized to 1950 and corrected for isotopic fractionation with the δ^{13}C measurement and normalized to a $\delta^{13}C_{VPDB}$ value of –25 ‰.

To evaluate the Dover sole ages, we compared the otolith Δ^{14}C results from the otolith cores with the Pacific halibut reference chronology using several procedures. Initially, the Δ^{14}C from the Dover sole otolith cores in both categories was plotted along with a loess (locally weighted least squares) smoothed curve of the Δ^{14}C in Pacific halibut. All loess-smoothed curves in this study were fitted by using Splus (Insightful Corp., Seattle, WA) (Chambers and Hastie, 1992) with a span of 2/3 and a degree of 2. To further analyze the Δ^{14}C results in comparison with the reference chronology, we used three procedures first introduced by Kastelle et al. (2008). The first procedure was to purposely bias the category-1 ages by 0, ±1, ±2, and ±4 years, generating seven sets of ages. For each of these seven sets of biased ages, posited birth years were calculated and a sum of squared residuals (*SSR*) between the Δ^{14}C in Dover sole otoliths and the loess smoothed data in the Pacific halibut reference Δ^{14}C chronology was calculated. The smallest of the seven *SSR*s indicated which purposely biased set of ages represented the best fit, and thereby indicated if an overall aging error in the Dover sole ages existed. We call this procedure the "unstandardized" analysis.

In the second procedure used to investigate the accuracy of Dover sole ages in category 1 we performed a "standardization" of the measured Δ^{14}C values (Kastelle et al., 2008). This is a linear transformation of the Δ^{14}C values in the Dover sole otolith cores which removes any difference in scale between the validation sample Δ^{14}C measurements and the reference chronology. It does not change the timing of the validation sample Δ^{14}C values or their relative magnitude. For this standardization, let $\{v_{y[j]}\}$ be the series of $j=1, \ldots, n$ validation observations of Δ^{14}C, where $y[j]$ refers to the year core j was formed. We defined the standardized series for $\{v_{y[j]}\}$ as $\{\hat{v}_{y[j]}=(v_{y[j]}+\mu)/\sigma\}$, where μ and σ can be estimated by a least squares fit (i.e., by minimizing the *SSR*) to the loess-smoothed curve of the reference chronology data set $\{l_{y[j]}\}$:

Table 1

Age estimates (in years) and radiocarbon measurements from Dover sole (*Microstomus pacificus*) otoliths collected from the Gulf of Alaska. Category 1 represents easy-to-age specimens and category 2 represents difficult-to-age specimens. Estimated birth year was calculated from final age and known capture date. The final ages are re-evaluated ages undertaken in the current study after initial aging for stock assessments. Postmeasurement min.–max. ages were generated after the radiocarbon values were known. Results are reported as Δ^{14}C (which is defined in Stuiver and Polach [1977] as the relative difference between an international standard [base year 1950] and sample activity), Δ^{14}C 95% confidence intervals (CI) were derived from accelerator mass spectrometry error, and the δ^{13}C measurements were used to correct for natural effects of isotopic fractionation.

			Age estimates (years)				
Specimen no.	Category	Estimated birth year	Final	Postmeasurement min.–max.	Δ^{14}C (‰)	Δ^{14}C 95% CI	δ^{13}C (‰)
3	1	1972	12		80.7	7.2	–1.34
11	1	1965	33		43.8	6.8	–1.86
17	1	1968	30		80.9	8.4	–1.99
18	1	1963	35		27.2	7.8	–1.6
20	1	1974	24		57.6	7.0	–1.5
21	1	1957	41		–86.3	6.2	–0.97
24	1	1958	47		–68.0	6.0	–1.65
34[1]	1	1953	45		–108.1	5.6	–0.97
36	1	1974	10		80.3	7.2	–1.77
48	1	1959	39		–64.9	6.6	–1.12
49	1	1969	29		81.5	7.4	–1.06
54	1	1957	27		–76.2	6.6	–0.98
66	1	1960	24		–65.2	8.0	–1.11
73	1	1964	20		49.2	10.2	–2.26
82	1	1966	32		68.0	7.2	–2.03
84	1	1977	21		36.2	8.0	–1.88
85	1	1966	32		86.4	8.2	–1.26
89	1	1959	39		–84.7	7.0	–1.14
102	1	1967	31		73.7	8.2	–1.12
130	1	1964	34		38.3	7.8	–0.76
138	1	1969	15		72.7	6.6	–1.4
143	1	1953	31		–108.2	5.6	–0.7
157[1]	1	1970	28		57.5	6.6	–0.95
159	1	1976	8		72.6	10.0	–1.88
160	1	1962	36		21.4	7.6	–2.07
210	1	1951	33		–109.5	5.4	–0.57
273	1	1963	21		9.5	8.6	–2.06
302	1	1975	9		64.9	9.8	–1.53
33	1	1959	39		–52.7	6.2	–1.92
34[1]	1	1963	35		41.5	9.0	–1.6
35	1	1970	28		54.3	6.4	–1.17
52	1	1969	29		83.6	8.4	–1.66
71	1	1955	43		–86.7	5.6	–0.71
157[1]	1	1974	10		62.2	6.8	–1.06
201	1	1955	29		–104.3	6.4	–1.32
257	1	1966	18		38.0	6.8	–1.34
271	1	1972	12		67.3	8.2	–1.41
310	1	1967	17		52.3	7.2	–2.78
95	2	1957	48	37–50	–120.1	5.8	–1.59
146	2	1958	26	16–26	75.7	10.2	–0.84
188	2	1953	52	37–59	–105.1	6.8	–0.92
225	2	1964	20	17–32	56.9	7.2	–1.33
317	2	1961	23	19–23	54.7	10.2	–1.41

[1] Specimens with repeated identification numbers do not represent repeat measurements but are different specimens with the same identification numbers from different collection years.

$$SSR = \sum_j (\hat{v}_{y[j]} - l_{y[j]})^2$$

$$= \sum_j [((v_{y[j]} + \mu)/\sigma) - l_{y[j]}]^2.$$

An iterative process to estimate μ and σ can be found in Kastelle et al. (2008). Again, to evaluate for aging error, we purposely biased the ages by 0, ±1, ±2, and ±4. For each of these seven sets of ages, we estimated μ and σ, standardized the validation sample Δ^{14}C observations with $\{\hat{v}_{y[j]}\}$, and calculated the *SSR*. As before, the smallest *SSR* indicated the best fit to the reference chronology and indicated if any overall aging bias (error) existed. It should be noted that if μ and σ are not estimated, but instead are set to $\mu=0$ and $\sigma=1$, this process becomes the unstandardized procedure described earlier.

The third and final method we employed was estimation of confidence intervals around the loess-smoothed reference chronology (Kastelle et al., 2008). For the estimated confidence intervals, simultaneous Bonferonni statistical inference was used (Miller, 1966) that calculates simultaneous (α=0.01) confidence intervals whose width is dependent on the number of distinct years at which comparisons are made between the reference chronology and validation samples, and on the variability in the reference chronology. As an aid in our comparison between the standardized Dover sole otolith Δ^{14}C values and the Pacific halibut reference chronology, not only were the *SSR*s evaluated, the standardized Δ^{14}C values were also viewed graphically when plotted against the posited birth years. This comparison generated seven plots, one for each purposeful bias, and included the loess-smoothed Pacific halibut reference chronology with confidence intervals.

The category-2 otoliths were re-examined again after the Δ^{14}C results were known. The goal of this re-aging process was to use the Δ^{14}C results as a guide for refining current otolith aging criteria. To facilitate this process, the Δ^{14}C in the category-2 otoliths was plotted with a loess-smoothed curve of category-1 results. This process allowed the age reader to learn how difficult-to-interpret otoliths are best aged. In this last re-aging process a minimum and maximum age were estimated for the category-2 otoliths.

The 3-year core in Dover sole otoliths needed to be taken into consideration when analyzing the results. We assumed a mid-point of approximately 1.5 years for core deposition because the core represents material from the first 3 years of life. This means that we assumed linear otolith growth during the first three years of life. Linear otolith growth may not be accurate, but was assumed for simplicity and because any error from this assumption is trivial. The Pacific halibut reference chronology is based on material from only the first year of life; therefore, we assumed a mid-point of 0.5 years for core deposition. These assumptions mandated an approximate allowance of a (1.5–0.5 =) 1.0-year shift in the comparison of the Dover sole results with the Pacific halibut reference chronology.

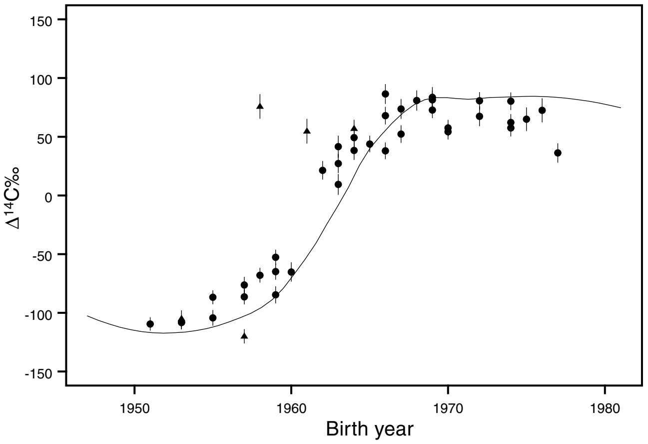

Figure 3

Δ^{14}C ‰ in otolith cores plotted against birth year of Dover sole (*Microstomus pacificus*), by category 1 (●) and category 2 (▲), with a 95% confidence interval shown for each point. Two category-1 specimens are overlapping at –108 Δ^{14}C and birth year 1953. The line (——) is a loess-smoothed curve of the Pacific halibut (*Hippoglossus stenolepis*) reference chronology (Piner and Wischniowski, 2004).

Results

Otolith selection and coring procedures

The selection process generated specimens for which final ages agreed with initial ages and was followed by successful coring. In all but three of the specimens in category 1, the agreement between the initial age and final age was within 3 years. In this category, the maximum final age was 47 years, and the minimum age was 8 years (Table 1). In category-2 specimens, four out of five had discrepancies of over 3 years between the initial age and final age. The percentage CV (Kimura and Anderl, 2005) between initial ages and final ages in both categories was 4.21%. The average core weight across all categories was 5.4 mg (±0.2 mg standard error) and the average size was 1.93 × 2.97 × 0.49 mm, which was smaller than the guide provided by the 3-year-olds described earlier.

^{14}C analysis

The Δ^{14}C in Dover sole otolith cores from category 1 followed the expected general pattern of initial low ^{14}C before atmospheric testing followed by an increase synchronous with testing. It displayed a rise in about 1955 (from below –100‰) and peaked in 1966 at over 85‰ (Table 1, Fig. 3). This trend in Dover sole radiocarbon

Table 2

Sum of squared residuals (*SSR*) between Δ^{14}C in category-1 Gulf of Alaska Dover sole (*Microstomus pacificus*) otoliths and loess-smoothed curve of the Pacific halibut (*Hippoglossus stenolepis*) reference chronology (Piner and Wischniowski, 2004). Results are tabled by unstandardized (No) or standardized (Yes) Δ^{14}C values. The μ and σ are coefficients in the linear standardization defined for the series $\{v_{y[j]}\}$ as $\{\hat{v}_{y[j]}=(v_{y[j]}+\mu)/\sigma\}$, where μ and σ were estimated by a least squares fit to the loess-smoothed curve of the reference chronology data set, or set to 0 and 1 respectively when unstandardized. When standardized, *SSR* was minimized with respect to μ and σ. Age bias was applied to each final age estimate such that –4 represents younger ages and +4 represents older ages.

Δ^{14}C Standardized to reference	Parameter	Age bias (years) –4	–2	–1	0	+1	+2	+4
No	μ	0	0	0	0	0	0	0
	σ	1	1	1	1	1	1	1
	SSR	40,429	13,179	11,651	20,140	38,281	66,696	153,202
Yes	μ	28.47	9.59	1.63	–6.17	–13.97	–22.80	–45.09
	σ	1.10	0.94	0.91	0.90	0.92	0.95	1.11
	SSR	15,544	7969	9533	16,834	29,191	45,706	84,612

values parallels those seen in the Pacific halibut reference chronology but the values appear to be shifted earlier in time by 1 or 2 years. Also, from 1969 on, the Dover sole values are mostly lower than those in the reference chronology (Fig. 3).

In the analysis of the category-1 ages, a purposeful bias of –1 or –2 years provided the lowest *SSR*s. In the unstandardized procedure, where the ages were purposely biased –1 years, the *SSR* was smallest at 11,651 (Table 2). In the standardized procedure when the ages were purposely biased –2 years, the *SSR* was smallest at 7969, where $\mu=9.59$ and $\sigma=0.94$ (Table 2). For the standardized analysis, a series of plots is presented, one plot for each of the seven sets of purposely biased ages (Fig. 4). Both the overall time shift and the low bias after 1969 were removed by the standardization procedure when the purposeful age bias was –2 years (Fig. 4B). Also, at this purposeful age bias and standardization, the Dover sole validation specimens were all within the 99% confidence intervals around the loess-smoothed reference chronology.

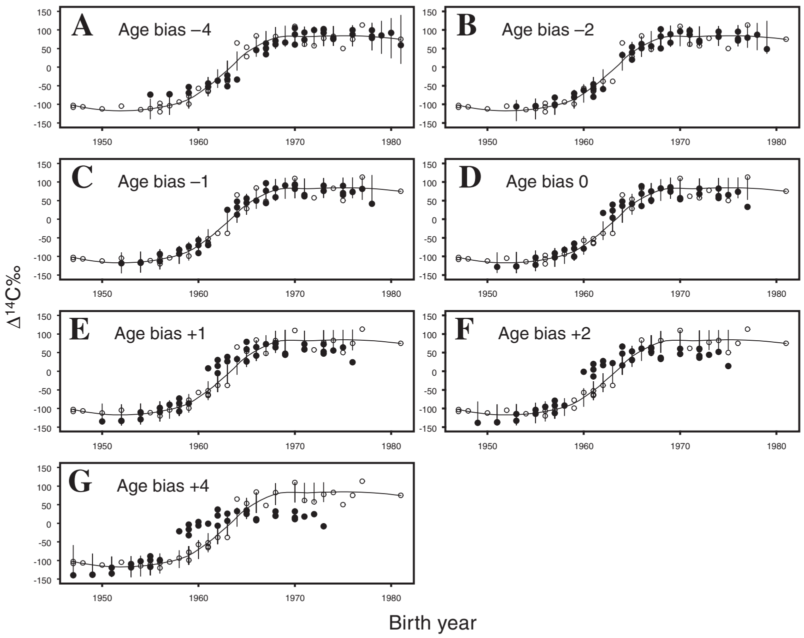

Figure 4

Series of seven plots for Δ^{14}C ‰ in otolith cores of Dover sole (*Microstomus pacificus*) category-1 (●) where the Δ^{14}C is standardized and each plot, (**A**) through (**G**), corresponds to a purposeful age bias of –4 to +4 years, respectively, with a loess-smoothed curve (——) of the Pacific halibut (*Hippoglossus stenolepis*) reference chronology (○) (Piner and Wischniowski, 2004) and 99% simultaneous confidence intervals around the mean smooth.

The general difficulty in aging category-2 specimens is evident in the Δ^{14}C results that are not synchronous with those from the category-1 specimens. The range of possible ages for category-2 specimens (Table 1) is shown by horizontal bars in Figure 5. Two of the three younger category-2 specimens were likely over-aged because only the right end of the horizontal bar is close to the values for the other specimens or the values of the loess-smoothed curve of category-1 specimens. Therefore, for these three specimens the choice of a younger age is more accurate. In the two older category-2 specimens, the results are less definitive because the Δ^{14}C values fall in the stable pre-bomb era. Figure 2 is an example of how a category-2 otolith may be aged given the different possible interpretive options.

Discussion

Age validation

This study is the first published age validation for GOA Dover sole. Ages in the range of 8 to 47 years were validated. Accurate ages were indicated by both the unstandardized and standardized Δ^{14}C results for category 1 when a purposeful age bias of -1 or -2 years produced the lowest *SSR*s. The shift of approximately 1 year (due to core size, as explained earlier) was expected, and when considered with the purposeful age bias and lowest *SSR*s, indicated that ages estimated by the break-and-burn method were accurate. All category-1 specimens were within the 99% simultaneous confidence intervals on the reference chronology (Fig. 4B) and this result also provides strong evidence for accurate ages.

Figure 5

Δ^{14}C ‰ in otolith cores of Dover sole (*Microstomus pacificus*) plotted against birth year, by category 1 (●) and category 2 (▲). Horizontal lines on category-2 specimens representing the minimum and maximum postmeasurement age, and a loess smoothed curve (——) has been fitted to category-1 data points.

The difference between the purposeful age bias of –1 and –2 was not resolved. The resolution of this bomb ^{14}C age validation study was limited by the approximate nature of the expected 1-year shift, variability in ^{14}C due to geographic location, and variability in ^{14}C measurement, but the general accuracy of the ages was validated. In using bomb Δ^{14}C for age validation, assumption 1 is the largest contributor to any concerns about resolution due to variability in Δ^{14}C because of depth or geographic area (Nydal, 1993; Kalish, 1995; Andrews et al., 2007).

Our method of separating the specimens into two categories addressed specific objectives. Use of only the clearest specimens to validate the age estimates for category-1 specimens is common practice (Piner and Wischniowski, 2004; Kerr et al., 2005; Kastelle et al., 2008). In a long-lived species such as Dover sole, age determination is difficult (Chilton and Beamish, 1982; Kimura and Anderl, 2005) and subjective interpretations of growth patterns (Fig. 2) must be made. When determining the age of a species on a routine basis, a set of species-specific interpretive rules or "aging criteria" are applied to all specimens, i.e., to specimens that are easy to interpret and to those that are difficult to interpret (Kastelle et al., 2008). By choosing validation samples like category 1 where there is little ambiguity, the basic methods and aging criteria can be validated and then applied to all samples, including ones like category 2. If the specimens were chosen randomly, without consideration for the difficulty in estimating age or interpreting otoliths, the spread of validation sample points around the reference chronology would increase. This spread would provide less informative results regarding the basic aging criteria, as shown by the spread in the category-2 specimens. The types of specimens represented by category 2 are more common in Dover sole; hence this sample design was selected to provide the most informative results. The exact percentage represented by the two categories was not determined in this study. A further reason for using different categories was to provide a tool to develop aging criteria for the hard-to-age specimens. Following the validation, the less common category-1 specimens were used to provide the loess fit in Figure 5 to which the category-2 specimens were compared.

The Δ^{14}C results displayed a gap in specimens at a birth year of 1961. Similar gaps have been seen in other studies (Kerr et al., 2004; Piner and Wischniowski, 2004; Piner et al., 2005; Kastelle et al., 2008) and are likely due to two reasons. First, in the early- to mid-1960s, marine ^{14}C was likely increasing so quickly that even if otoliths were accurately aged, a gap would likely be present because of the limited time range when these mid-range values of Δ^{14}C existed. Second, in our sample selection process, otoliths that could be categorized unambiguously as category 1 and aged such that they represented the 1961 birth year were not present.

The utility of Δ^{14}C standardization is apparent in this study. Without the standardization, most of the validation sample Δ^{14}C points before about 1968 are above the reference chronology, or to its left, and below the reference chronology after 1969. With the Δ^{14}C standardization and the purposeful bias of –2 years applied to the category-1 ages, it is clear that the Dover sole Δ^{14}C points and Pacific halibut reference chronology are in synchrony and that ages are accurate. As previously explained, the shift to the left is due to the core size. The consistently low Dover sole Δ^{14}C points after 1969 are likely due to different environmental regimes or biological differences of the two species (Kalish, 1993, 1995; Nydal, 1993; Andrews et al., 2007). The standardization procedure is ideal for correcting this type of bias, where a difference in range of Δ^{14}C exists. Previously, this procedure was used by Kastelle et al. (2008) to re-analyze validation data for black drum (*Pogonias cromis*), originally presented by Campana and Jones (1998). For the re-analysis, the black drum Δ^{14}C values were standardized to a Northern Hemisphere atmospheric Δ^{14}C reference chronology in a comparison where they were dramatically different in scale but similar in timing (Kastelle et al., 2008). If little difference in range exists when the standardization is applied, the estimated values of μ and σ will be close to 0 and 1, respectively, provided the overall fit is good. This situation would indicate that the standardization had little effect and that the correct evaluation of any aging error will still be made by considering the SSRs. This was the case for Pacific ocean perch analyzed previously with this method (Kastelle et al., 2008). Therefore, we feel this standardization method can be applied generally.

We chose the Pacific halibut reference chronology for several reasons. First, this reference chronology is based mostly on juvenile fish (Piner and Wischniowski, 2004). It also represents a wide geographic area in the GOA, similar to that for Dover sole. Finally, although the early life history of Dover sole in the GOA is not well understood, the pelagic larvae are found in the upper 30 m of the water column, and immature fish are known to concentrate in nearshore areas and shallow waters over the continental shelf (Abookire et al., 2001; Abookire and Bailey, 2007). Juvenile Pacific halibut are typically found in shallow nearshore areas (Norcross et al., 1995; Abookire et al., 2001); therefore comparisons with Dover sole for this age validation were reasonable.

The otolith cores from the validation samples were smaller than the measured guide otoliths from 3-year-olds. Some of this difference may be explained by the presence of newly deposited opaque material beyond the third translucent zone in the measured 3-year-olds. In the cored otoliths, this same material was often ground away to expose the third translucent zone, thereby producing a size difference. Also, a few of the cores may have been incorrectly centered during the grinding process, and therefore may have incorporated a little material from beyond the third year. Conversely, too much material could have been removed, down to the second year's growth zone. The latter is more likely the case as evidenced by the small core weights. On the proximal side, the coring process may have inadvertently removed some material belonging to the third year in an effort to remove all material from later years in the region of the sulcus groove. We considered the average age of the material represented by the cores to be approximately 1.5 years. However, if the cores were too small and some material inside the third translucent zone was removed, than the average age of the material may have been closer to 1 year, indicating less of a required shift. The difference was only 0.5 years; hence we considered any error from this consideration to be negligible. Less of a required shift was also indicated if the otoliths were accumulating more mass during their first year than in subsequent years (see *Materials and methods* where linear otolith growth is assumed). As mentioned previously, this age validation method can not resolve either type of error when potentially very small. However, it is probably not coincidental that the purposeful age bias of –1 year for the unstandardized Δ^{14}C results was the best fit.

Dover sole otoliths often display a transition in growth rate typically seen as a pattern of decreased spacing between presumed annual growth zones. This occurs when the fish is estimated to be about 6- to 8-years-old, with growth zones deposited prior to the transition representing younger and faster growth and post-transition zones representing slower growth (Figs. 1 and 2). An association between the transition timing and maturity has not been documented in Dover sole, but Abookire and Macewicz (2003) reported that 50% maturity occurs at 6.7 years which roughly coincides with the observed transition. They used specimens aged by the same experienced age readers as in this study; hence some level of circularity exists. However, our studies' results lead us to believe that the timing of the transition pattern is likely associated with the onset of maturation. In other species such as orange roughy (*Hoplostethus atlanticus*) a decrease in the annual growth zone width is documented to correspond to the onset of maturity (Francis and Horn, 1997).

Considerations concerning category-2 results

Category-2 Δ^{14}C results confirm that Dover sole are often a difficult-to-age species. This difficulty was exemplified by the range in possible ages of even the youngest category-2 specimens and was especially evident in the two older category-2 specimens. The CV of 4.21% for these hand-picked specimens, where the majority (38 out of 43) of otoliths were deemed to be clear, although better than the typical CV of 9.64% for Dover sole, was high in comparison to that for many other flatfish species aged at the AFSC (Kimura and Anderl, 2005). The category-2 otoliths are typical of many Dover sole samples aged at the AFSC where subjective decisions are made by necessity in the age reading process.

Results from the five difficult-to-age specimens in category 2 did not indicate consistent over-aging or under-aging. Correct decisions in how to interpret the growth zone patterns were made for specimens 225 and 188, and a reasonably good choice was made in specimen 95 especially when its high age was considered. It is clear that the choice of a mid-range or older age was the most accurate for specimens 95 and 188 when compared to the loess-smoothed curve of category-1 specimens. However, specimens 146 and 317 were demonstrated to be over-aged in the comparison to category-1 specimens. The underlying difficulty in aging and interpreting Dover sole otoliths lies in the framework of splits that can make up a single translucent zone (Fig. 2). In specimens from category 1, the translucent zones were usually compact and well defined; hence any apparent splits could easily be interpreted and decisions could be made as to how they should be enumerated. In category-2 specimens, the potential annual zones were not well defined, often due to splits that created interpretative options. Splits in a potential annual zone that occur before the transition to slow growth are especially problematic. In hindsight, it was reasonable to conclude that the two specimens (numbers 146 and 317) that diverged from the loess-smoothed category-1 data were over-aged probably because of broad splits. The specimen with the largest discrepancy, number 146, was one where the pattern in the broken-and-burnt cross section could be aged at 16 years (more in line with the loess-smoothed category-1 data), or on a second reading axis could be interpreted as 26 years. Similarly, in Figure 2 an age of 32 years was estimated along an axis closer to the sulcus, but a more correct age estimate of 20 years was chosen from an adjacent reading axis. Some of this discrepancy could be the result of splitting translucent zones. In reality, the five specimens in category 2 are not enough to draw firm conclusions on how these difficult otoliths should be aged. A further study of category-2 type specimens may help to refine the aging criteria for Dover sole otoliths.

The region in the otolith cross section before the transition to slow growth can be an area of splitting or diffuse translucent zones. This is a situation that can lead to over-aging; therefore the reader must exercise care to count only prominent translucent zones. Splitting and diffuse translucent zones are a frequent problem in age reading many species (e.g., Francis and Horn, 1997; Gregg et al., 2006; Hutchinson et al., 2007). The fish in category 1 were correctly aged by counting only the prominent translucent zones preceding the transition to slow growth.

Conclusions

The ages estimated for the GOA Dover sole were validated as accurate based on the easy-to-age otoliths in category 1. When the age bias of –1 or –2 years was applied, the Δ^{14}C in the validation samples had the same timing as the Δ^{14}C in the Pacific halibut reference chronology and was consistent with the expected 1-year core-size shift. An age-structured stock assessment model is used for management of the GOA Dover sole commercial fisheries and hence the age data validated here are important for population modeling and setting the total allowable catch (Stockhausen et al., 2005). In the future, analysis of additional difficult-to-age category-2 otoliths may help to further answer questions regarding aging criteria for these specimens. In reality, the lower-than-average between-reader precision for growth zone counts will likely persist, but now we have a high degree of confidence in the accuracy of ages estimated from specimens with clear growth zones.

Acknowledgments

We thank the staff at the Age and Growth Program of the Alaska Fisheries Science Center for support during this study. We also wish to thank T. Wilderbuer and W. Stockhausen of the Alaska Fisheries Science Center for helpful reviews and comments on early versions of this manuscript. K. Mckinney provided photographic support for which we are grateful. We thank A. Andrews of Moss Landing Marine Laboratories and two anonymous reviewers for insightful comments on the manuscript. S. Handwork and K. Elder of the National Ocean Sciences Accelerator Mass Spectrometry Facility at the Woods Hole Oceanographic Institution provided support and technical advice regarding the Δ^{14}C measurements for which we are grateful.

Literature cited

Abookire, A. A., and K. M. Bailey.
2007. The distribution of life cycle stages of two deep-water pleuronectids, Dover sole (*Microstomus pacificus*) and rex sole (*Glyptocephalus zachirus*), at the northern extent of their range in the Gulf of Alaska. J. Sea Res. 57:198–208.

Abookire, A. A., and B. J. Macewicz.
2003. Latitudinal variation in reproductive biology and growth of female Dover sole (*Microstomus pacificus*) in the North Pacific, with emphasis on the Gulf of Alaska stock. J. Sea Res. 50:187–197.

Abookire, A. A., J. F. Piatt, and B. L. Norcross.
2001. Juvenile groundfish habitat in Kachemak Bay, Alaska, during late summer. Alaska Fish. Res. Bull. 8(1):45–56.

Andrews, A. H., E. J. Burton, L. A. Kerr, G. M. Cailliet, K. H. Coale, C. C. Lundstrom, and T. A. Brown.
2005. Bomb radiocarbon and lead-radium disequilibria in otoliths of bocaccio rockfish (*Sebastes paucispinis*): a determination of age and longevity for a difficult-to-age fish. Mar. Freshw. Res. 56:517–528.

Andrews, A. H., L. A. Kerr, G. M. Cailliet, T. A. Brown, C. C. Lundstrom, and R. D. Stanley.
2007. Age validation of canary rockfish (*Sebastes pinniger*) using two independent otolith techniques: lead-radium and bomb radiocarbon dating. Mar. Freshw. Res. 58:531–541.

Beamish, R. J., and G. A. McFarlane.
1983. The forgotten requirement for age validation in fisheries biology. Trans. Am. Fish. Soc. 112(6):735–743.

Campana, S. E.
1997. Use of radiocarbon from nuclear fallout as a dated marker in the otoliths of haddock *Melanogrammus aeglefinus*. Mar. Ecol. Prog. Ser. 150:49–56.
2001. Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. J. Fish Biol. 59:197–242.

Campana, S. E., and C. M. Jones.
1998. Radiocarbon from nuclear testing applied to age validation of black drum, *Pogonias cromis*. Fish. Bull. 96(2):185–192.

Campana, S. E., L. J. Natanson, and S. Myklevoll.
2002. Bomb dating and age determination of large pelagic sharks. Can. J. Fish. Aquat. Sci. 59:450–455.

Campana, S. E., R. D. Stanley, and S. Wischniowski.
2003. Suitability of glycerin-preserved otoliths for age validation using bomb radiocarbon. J. Fish Biol. 63:848–854.

Chambers, J. M., and T. J. Hastie.
1992. Statistical Models in S, 608 p. Wadsworth and Brooks/Cole Advanced Books and Software, Pacific Grove, CA.

Chilton, D. E., and R. J. Beamish.
1982. Age determination methods for fishes studied by the groundfish program at the Pacific Biological Station. Can. Spec. Publ. Fish. Aquat. Sci. 60, 102 p.

Francis, R. I. C. C., and P. L. Horn.
1997. Transition zone in otoliths of orange roughy (*Hoplostethus atlanticus*) and its relationship to the onset of maturity. Mar. Biol. 129:681–687.

Gregg, J. L., D. M. Anderl, and D. K. Kimura.
2006. Improving the precision of otolith-based age estimates for Greenland halibut (*Reinhardtius hippoglossoides*) with preparation methods adapted for fragile sagittae. Fish. Bull. 104:643–648.

Hunter, J. R., J. L. Butler, C. Kimbrell, and E. A. Lynn.
1990. Bathymetric patterns in size, age, sexual maturity, water content, and caloric density of Dover sole, *Microstomus pacificus*. Calif. Coop. Oceanic Fish. Invest. Rep. 31:132–144.

Hutchinson, C. E., C. R. Kastelle, D. K. Kimura, and D. R. Gunderson.
2007. Using radiometric ages to develop conventional ageing methods for shortraker rockfish (*Sebastes borealis*). *In* Biology, assessment, and management of North Pacific rockfishes (J. Heifetz, J. DiCosimo, A. J. Gharrett, M. S. Love, V. M. O'Connell, and R. D. Stanley, eds.), p. 237–249. Alaska Sea Grant College Program, AK-SG-07-01, Univ. Alaska Fairbanks, Fairbanks, AK.

Kalish, J. M.
1993. Pre- and post-bomb radiocarbon in fish otoliths. Earth Planet. Sci. Lett. 114:549—554.
1995. Radiocarbon and fish biology. *In* Recent development in fish otolith research (D. H. Secor, J. M. Dean, and S. E. Campana, eds), p. 637–653. Belle W. Baruch Library in Marine Science Number 19. Univ. South Carolina Press, Columbia, S.C.

Kastelle, C. R., and D. K. Kimura.
2006. Age validation of walleye pollock (*Theragra chalcogramma*) from the Gulf of Alaska using the disequilibrium of Pb-210 and Ra-226. ICES J. Mar. Sci. 63:1520–1529.

Kastelle, C. R., D. K. Kimura, and B. J. Goetz.
2008. Bomb radiocarbon age validation of Pacific ocean perch (*Sebastes alutus*) using new statistical methods. Can. J. Fish. Aquat. Sci. 65:1101–1112.

Kerr, L. A., A. H. Andrews, G. M. Cailliet, T. A. Brown, and K. H. Coale.
2006. Investigations of $\Delta^{14}C$, $\delta^{13}C$, and $\delta^{15}N$ in vertebrae of white shark (*Carcharodon carcharias*) from the eastern North Pacific Ocean. Environ. Biol. Fishes 77:337–353.

Kerr, L. A., A. H. Andrews, B. R. Frantz, K. H. Coale, T. A. Brown, and G. M. Cailliet.
2004. Radiocarbon in otoliths of yelloweye rockfish (*Sebastes ruberrimus*): a reference time series for the coastal waters of southeast Alaska. Can. J. Fish. Aquat. Sci. 61:443–451.

Kerr, L. A., A. H. Andrews, K. Munk, K. H. Coale, B. R. Frantz, G. M. Cailliet, and T. A. Brown.
2005. Age validation of quillback rockfish (*Sebastes malinger*) using bomb radiocarbon. Fish. Bull. 103:97–107.

Kimura, D. K., and D. M. Anderl.
2005. Quality control of age data at the Alaska Fisheries Science Center. Mar. Freshw. Res. 56:783–789.

Kimura, D. K., C. R. Kastelle, B. J. Goetz, C. M. Gburski, and A. V. Buslov.
2006. Corroborating the ages of walleye pollock (*Theragra chalcogramma*). Mar. Freshw. Res. 57:323–332.

Miller, R. G.
1966. Simultaneous statistical inference, 272 p. McGraw-Hill, New York, NY.

Norcross, B. L., B. A. Holladay, and F. J. Muter.
1995. Nursery area characteristics of Pleuronectids in coastal Alaska, USA. Neth. J. Sea Res. 34:161–175.

Nydal, R.
1993. Application of bomb ^{14}C as a tracer in the global carbon cycle. Trends Geophys. Res. 2:355–364.

Piner, K. R., O. S. Hamel, J. L. Menkel, J. R. Wallace, and C. E. Hutchinson.
2005. Age validation of canary rockfish (*Sebastes pinniger*) from off the Oregon coast (USA) using the bomb radiocarbon method. Can. J. Fish. Aquat. Sci. 62:1060–1066.

Piner, K. R., J. R. Wallace, O. S. Hamel, and R. Mikus.
2006. Evaluation of ageing accuracy of bocaccio (*Sebastes paucispinis*) rockfish using bomb radiocarbon. Fish. Res. 77:200–206.

Piner, K. R., and S. G. Wischniowski.
2004. Pacific halibut chronology of bomb radiocarbon in otolliths from 1944 to 1981 and a validation of ageing methods. J. Fish Biol. 64:1060–1071.

Stockhausen, W. T., B. J. Turnock, Z. T. A'mar, M. E. Wilkins, and M. H. Martin.
2005. Gulf of Alaska Dover sole. *In* Appendix B, Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska, p. 351–397. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage, AK 99510.

Stuiver, M., and H. A. Polach.
1977. Discussion: reporting of ^{14}C data. Radiocarbon 19(3):355–363.