Abstract.– At the Alaska Fisheries Science Center, one in five age readings produced for routine stock assessments are re-aged independently by a second age-reader. The Center now has a large database of repeated age readings that covers a variety of groundfish species and years. The purpose of this paper is to point out the problems and utility of interpreting such a database. The main problem of interpretation is fundamental, and relates to the fact that the true age of a fish is seldom known. Nevertheless, from a pragmatic point of view, these data can still provide useful insights into the age-determination process. Data from six marine fish species are used to show the overall levels of between-reader bias, agreement, and variability that have occurred on production age readings. Other uses for these data include objectively ranking the relative difficulty in ageing different species, maintaining quality control, examining between-reader differences in ageing criteria, and evaluating the possible importance of between-reader bias and variability in later analysis and modeling applications. Assuming reader bias is negligible, modeling results presented here indicate that estimated percentage agreements are consistent with the hypothesis that age determinations are normally distributed with a constant coefficient of variation over relatively wide age ranges. This result supports use of the coefficient of variation for measuring variability in age precision studies.

Manuscript accepted 6 August 1990.
Fishery Bulletin, U.S. 89:53–60 (1991).

Between-Reader Bias and Variability in the Age-Determination Process

Daniel K. Kimura
Julaine J. Lyons
Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA
7600 Sand Point Way NE, Seattle, Washington 98115-0070

In this paper we evaluate a unique database developed for all marine fish species being routinely aged at the Alaska Fisheries Science Center. Here, large subsamples (one in five age readings produced for routine stock assessments) have been re-aged independently by a second experienced age-reader, mainly for the purpose of maintaining quality control. However, it became apparent that this database could be used to provide additional insights into the age-determination process.

Most everyone familiar with the ageing of fish knows this process is fraught with difficulties. At the very least, there must be random variability about some true age. Most likely there is also bias in the ageing methodology at some ages, as well as between-reader differences. Because reader bias is probably affected by the individual reader, the true age of the fish being aged, and perhaps even individual fish, the analysis of repeated age readings made by different readers does not easily fall under the purview of classical statistical theory.

The types of analysis that can be performed on age-determination data are dependent on the kind of data collected and the assumptions the data analyst is willing to make. For example, if replicated readings are made by each reader, it is possible to perform a variance components analysis, assuming that reader effects are random and unbiased (Kimura et al. 1979). Comparative calibration is the area of statistical analysis that compares different methods of measurement (e.g., different readers) where all methods of measurement are assumed to contain error, and perhaps bias (Theobald and Mallinson 1978). Recently Kimura (unpubl.) examined the limits of possible inference for the functional comparative calibration model.

In the analyses presented here we examine between-reader bias and variability based on subsamples aged independently by two age-readers. For these types of data, we define between-reader bias as the average difference ($\bar{a}_1 - \bar{a}_2$) in ages assigned by these readers when ageing the same specimens of the same nominal age. Thus between-reader bias presumably arises from the two readers using different ageing criteria. If the average difference between age-readers is negligible at some nominal age, then between-reader bias at that age is defined to be negligible, regardless of what the unknown absolute bias of the readers might be.

Estimates of between-reader ageing variability from these types of data can be computed by averaging the sample variances calculated from the two age readings (df = 1) from each age structure over some nominal age. These sample variances (between-reader variances) probably overestimate measurement error, because they include a component of variability that might be thought of as between-reader bias.

Age determination is a statistical process that has a characteristic level of variability. This variability is species-dependent, and provides a basis for comparing the ageing of different species. For example, a species that can be aged with a larger percentage agreement (percentage of specimens aged the same on two occasions by the same or different reader), or smaller coefficient of variation, provides a statistical confirmation of the statement that species "x" is easier to age than species "y."

Between-reader bias provides a measure of the adequacy of criteria for distinguishing ages in a particular species of some nominal age. Presumably, if there is no between-reader bias, ageing criteria are being applied similarly by both readers, and the data only contain random measurement error. If between-reader bias and measurement error are independent, at this point between-reader variance would also be minimized. Significant between-reader bias may indicate a lack of resolving power in the criteria, insufficient training, or even peculiarities in the structures being aged. Between-reader variance is generally an indicator of overall "ageability," but is not as effective as between-reader bias measurements for pointing out between-reader differences in criteria.

Species often have a characteristic age above which between-reader biases become larger. This age may be interpreted as a line distinguishing which ages are more reliable. For age-readers themselves, between-reader bias is usually of more interest than variability. This is because while measurement error is inherent in the age-determination process, between-reader bias can be controlled to a greater extent.

In age-determination studies the term "precision" is used to describe "agreement," or variability between readings of the same specimen by the same or different age-reader. The term "accuracy" is reserved to describe a comparison of ages generated by readers with the "true" age for specimens of known age.

By emphasizing the importance of between-reader bias and variability, we do not mean to denigrate the obvious importance accuracy and age validation play in the age-determination process (Beamish and McFarlane 1983, 1987). Validation (the comparison of ages determined by counting rings on hard parts with known ages) can be carried out in a variety of ways, all of which are difficult. These include combining an external tag with an oxytetracycline (OTC) injection that labels calcium rings with a mark visible under ultraviolet light, following unusually strong year-classes through time, ageing young fish of known ages, and, most recently, measuring the activity of naturally occurring radioisotopes. Scientists at the Pacific Biological Station (Beamish et al. 1983, Cass and Beamish 1983, Leaman and Nagtegaal 1987, McFarlane and Beamish 1987) have made wide use of the OTC mark. And, recently, two studies appear to have succeeded in validating longevity in rockfish using radioisotopes (Bennett et al. 1982, Campana et al. 1990). Typically, validation can be carried out on only a very few fish. Often, doubts remain concerning criteria for certain age groups, or structures that look different. Nevertheless, the validation process is a critical one, and age-readers must constantly strive to improve the accuracy of their age determinations.

Because we seldom knew the true age of a fish, absolute bias and total mean-square error in the ageing process were not known. Therefore, our discussions here will be limited to between-reader bias and variability. These quantities are defined by the between-reader bias and coefficient-of-variation formulas described in the following section.

Materials and methods

The Ageing Unit at the Alaska Fisheries Science Center has the broad responsibility of ageing commercially important fish species and fish stocks in U.S. waters from California to the eastern Bering Sea. Historically, data have been accumulated from three principal sources: scientific surveys using various fishing gear, and foreign and domestic vessels fishing in U.S. waters. The present data consist of ages read from the otoliths (ear bones) of various groundfish species collected using assorted gear.

Since 1981, the preferred method of reading ages from these structures has been to either break or saw the otolith cross-wise, burn the exposed surface, and read the cross-section under a microscope (Chilton and Beamish 1982). Only young or unusually clear specimens of select species can be read from the intact surface.

In 1983 a quality-control program was initiated wherein 20% of all routine age readings would be independently re-aged by an age-reader (i.e., the tester) particularly experienced in a species. Statistics were calculated on these reader/tester data (one reading per otolith from each age-reader) in the following manner:

1 mean ($\bar{x}$) = (tester + reader)/2.
2 standard deviation (SD) = $\sqrt{[(\text{tester} - \bar{x})^2 + (\text{reader} - \bar{x})^2]}$
3 nominal age (age): $\bar{x}$ (rounded), or tester age
4 n (count): sample size (number of specimens aged)
5 percentage agreement: (n agree/n) × 100
6 coefficient of variation (CV) = (SD/$\bar{x}$) × 100
7 between-reader bias: reader age − tester age
8 percentage bias: [(reader age − tester age)/$\bar{x}$] × 100

Elements 5–8 were averaged over the "n" specimens of the same nominal age, and over all ages (weighted by n) for overall statistics.

Table 1
Statistics comparing reader/tester data for Pacific whiting in 1986. Bias is between-reader bias; nominal age is mean.

Age (yr)	Count	Percentage agreement	CV (%)	Bias (yr)	Percentage bias
2	258	100.0	0.0	0.00	0.0
3	58	50.0	15.1	-0.14	-5.3
4	8	62.5	7.6	-0.38	-10.7
5	13	92.3	3.6	0.23	5.1
6	651	95.5	0.6	-0.00	-0.2
7	118	20.3	9.9	0.23	3.4
8	66	37.9	11.9	-0.23	-2.9
9	153	77.8	2.4	-0.02	-0.2
10	28	10.7	12.1	0.25	2.4
11	23	26.1	12.0	0.00	-0.2
12	13	0.0	12.0	-0.92	-7.9
13	29	75.9	1.4	0.03	0.3
14	11	0.0	6.1	0.45	3.3
15	4	25.0	13.1	0.25	1.5
16	2	100.0	0.0	0.00	0.0
Average		78.6	3.2		

Table 2
Statistics comparing reader/tester data for yellowfin sole in 1986. Bias is between-reader bias; nominal age is mean.

Age (yr)	Count	Percentage agreement	CV (%)	Bias (yr)	Percentage bias
3	1	100.0	0.0	0.00	0.0
4	4	50.0	10.1	0.50	14.3
5	11	45.5	9.7	0.64	13.7
6	15	86.7	1.7	0.13	2.4
7	64	79.7	2.2	0.14	2.2
8	32	68.8	2.9	0.06	0.8
9	23	78.3	1.8	-0.13	-1.5
10	39	66.7	2.7	0.21	2.1
11	19	57.9	3.2	-0.16	-1.5
12	32	62.5	3.0	0.25	2.1
13	19	52.6	3.3	-0.05	-0.4
14	14	42.9	6.9	0.79	5.7
15	22	68.2	2.2	0.45	3.1
16	26	34.6	3.6	0.04	0.2
17	12	25.0	5.3	0.08	0.5
18	16	43.8	3.0	0.38	2.2
19	12	58.3	1.9	0.33	1.8
20	5	0.0	7.1	0.00	0.0
23	3	0.0	5.1	-1.00	-4.3
25	1	0.0	2.9	1.00	4.1
26	1	0.0	2.8	1.00	3.9
Average		60.9	3.2		

When the main purpose of an analysis is to compare criteria of the age readers, the nominal age for classification should be the tester age. When the overall characteristics of ageing a species is of interest, perhaps $\bar{x}$ is the better nominal age. The estimated statistics by age look different depending on which nominal age is being used. For example, if $\bar{x}$ is used, the percentage agreement for younger ages will appear larger.

Both the average percent error (Beamish and Fournier 1981) and the coefficient of variation have been proposed as an "age independent" method of estimating precision for age-determination studies. Assuming normality, Chang (1982) favored the coefficient of variation on the basis of efficiency, and we favor the coefficient of variation on the basis of common usage. Under differing distributional assumptions, the average percent error may actually be superior.

If age determinations are independently and normally distributed about some true age, then the percentage agreement at each age can be predicted from the area under the normal curve. Suppose the age of an "a"-year-old fish can be determined with a certain coefficient of variation. The difference between two independent age determinations (b and c, say), would be distributed as

$$z = b - c \sim N[0, 2CV^2a^2].$$

The predicted percentage agreement (ppa) at age "a" is then

$$ppa = [\Phi(z_2) - \Phi(z_1)] \times 100,$$

where Φ = the cumulative distribution function of the unit normal distribution,

$z_1 = -0.5/[(1.4142)(CV)(a)]$ and,

$z_2 = +0.5/[(1.4142)(CV)(a)]$.

Results and discussion

The above statistics were calculated for several species sampled in 1986 and subsequently aged (Tables 1–6). These are overall statistics calculated using data from all readers and testers, and therefore represent group rather than individual performance. A summary of percentage agreements and coefficients of variation averaged over all ages is given below:

Species	Percentage agreement	CV
Pacific whiting *Merluccius productus*	78.6	0.032
yellowfin sole *Limanda aspera*	60.9	0.032
Pacific ocean perch *Sebastes alutus*	40.8	0.049
walleye pollock *Theragra chalcogramma*	63.8	0.050
Atka mackerel *Pleurogrammus monopterygius*	66.8	0.068
sablefish *Anoplopoma fimbria*	43.7	0.129

Table 3
Statistics comparing reader/tester data for Pacific ocean perch in 1986. Bias is between-reader bias; nominal age is mean.

Age (yr)	Count	Percentage agreement	CV (%)	Bias (yr)	Percentage bias	Age (yr)	Count	Percentage agreement	CV (%)	Bias (yr)	Percentage bias
2	1	100.0	0.0	0.00	0.0	34	2	0.0	12.7	1.00	3.0
3	7	71.4	8.1	0.29	11.4	36	1	0.0	3.9	-2.00	-5.6
4	1	0.0	20.2	1.00	28.6	37	2	0.0	1.9	1.00	2.7
5	34	73.5	4.2	-0.03	-0.7	38	3	66.7	0.6	0.33	0.9
6	29	51.7	6.2	-0.28	-5.0	39	5	0.0	8.0	0.80	2.0
7	9	33.3	10.7	-0.78	-11.7	40	2	0.0	2.7	-1.50	-3.8
8	62	71.0	2.9	0.02	0.2	41	8	12.5	2.2	-0.50	-1.2
9	42	45.2	4.7	0.05	0.5	42	6	50.0	1.7	0.00	0.0
10	41	36.6	5.8	0.20	2.1	43	4	25.0	3.7	0.75	1.7
11	12	25.0	7.7	0.33	3.1	44	1	0.0	4.9	3.00	6.9
12	21	38.1	5.2	0.19	1.6	45	3	66.7	1.0	0.67	1.5
13	6	16.7	7.3	-0.67	-5.2	46	3	0.0	2.1	-0.67	-1.4
14	8	0.0	6.5	-1.00	-7.3	47	1	0.0	3.0	2.00	4.3
15	9	33.3	5.3	0.44	3.0	48	1	100.0	0.0	0.00	0.0
16	14	21.4	4.8	0.50	3.2	49	1	0.0	1.5	1.00	2.1
17	5	20.0	5.0	0.40	2.4	50	2	0.0	1.4	0.00	0.0
18	7	14.3	6.8	0.00	-0.0	52	5	0.0	7.4	4.20	8.1
19	8	37.5	7.0	1.13	5.9	54	2	0.0	2.6	-1.00	-1.9
20	7	28.6	4.6	-0.71	-3.7	55	2	50.0	3.2	-2.50	-4.6
21	1	0.0	10.3	3.00	14.6	56	1	0.0	3.8	3.00	5.4
22	1	100.0	0.0	0.00	0.0	57	2	0.0	7.4	-6.00	-10.5
23	2	0.0	6.1	0.00	0.0	59	3	0.0	6.0	-3.00	-5.1
24	2	50.0	1.5	-0.50	-2.1	60	1	0.0	2.4	-2.00	-3.3
25	4	25.0	4.3	-1.50	-6.1	61	5	0.0	2.8	-1.60	-2.6
26	2	0.0	2.8	0.00	0.0	65	2	50.0	1.6	-1.50	-2.3
27	3	33.3	3.6	1.33	5.0	66	1	0.0	1.1	-1.00	-1.5
28	2	100.0	0.0	0.00	0.0	68	1	0.0	3.1	3.00	4.4
29	2	0.0	17.1	3.00	10.3	72	1	0.0	1.0	1.00	1.4
30	1	0.0	4.7	2.00	6.7	73	1	0.0	8.8	9.00	12.4
31	2	0.0	4.6	0.00	0.0	75	1	100.0	0.0	0.00	0.0
32	2	0.0	4.4	0.00	0.0	78	1	0.0	3.6	-4.00	-5.1
33	1	100.0	0.0	0.00	0.0	Average		40.8	4.9		

Table 4
Statistics comparing reader/tester data for pollock in 1986. Bias is between-reader bias; nominal age is mean.

Age (yr)	Count	Percentage agreement	CV (%)	Bias (yr)	Percentage bias
1	18	100.0	0.0	0.00	0.0
2	64	93.8	2.9	0.03	2.1
3	132	92.4	2.1	0.00	0.0
4	159	74.8	5.3	0.04	1.0
5	136	66.2	5.6	0.07	1.5
6	119	64.7	5.2	0.06	1.1
7	113	49.6	6.5	-0.19	-2.7
8	181	56.9	4.7	-0.19	-2.5
9	85	21.2	7.1	0.11	1.2
10	26	15.4	10.7	-0.08	-0.8
11	17	41.2	7.4	0.41	3.8
12	8	25.0	5.3	-0.38	-3.2
13	2	0.0	8.3	-0.50	-3.7
Average		63.8	5.0		

Table 5
Statistics comparing reader/tester data for Atka mackerel in 1986. Bias is between-reader bias; nominal age is mean.

Age (yr)	Count	Percentage agreement	CV (%)	Bias (yr)	Percentage bias
1	15	100.0	0.0	0.00	0.0
2	76	97.4	1.2	0.00	0.0
3	47	55.3	14.2	0.02	0.9
4	28	64.3	10.3	-0.11	-2.8
5	29	48.3	10.1	0.03	1.1
6	28	50.0	6.8	0.11	1.8
7	20	65.0	5.4	-0.40	-6.0
8	35	40.0	7.2	-0.26	-3.4
9	4	25.0	8.1	0.00	0.3
10	4	50.0	3.7	-0.50	-5.3
Average		66.8	6.8		

Table 6

Statistics comparing reader/tester data for sablefish in 1986. Bias is between-reader bias; nominal age is mean.

Age (yr)	Count	Percentage agreement	CV (%)	Bias (yr)	Percentage bias
1	13	100.0	0.0	0.00	0.0
2	43	88.4	5.5	0.02	1.6
3	43	58.1	13.4	-0.21	-7.8
4	24	58.3	10.1	-0.25	-7.1
5	49	53.1	10.2	0.51	10.8
6	50	34.0	12.7	0.82	14.4
7	29	37.9	12.1	0.79	11.7
8	17	11.8	15.0	1.18	15.2
9	21	0.0	20.6	2.33	26.9
10	11	0.0	24.2	3.00	30.3
11	11	9.1	17.4	1.73	16.0
12	12	8.3	24.1	3.83	32.7
13	5	0.0	29.9	5.40	42.3
14	4	0.0	36.7	4.50	33.3
15	1	0.0	4.9	1.00	6.9
16	1	0.0	31.9	7.00	45.2
18	2	0.0	8.1	1.00	5.7
19	2	0.0	35.5	9.50	50.2
23	1	0.0	15.7	-5.00	-22.2
29	2	50.0	1.2	0.50	1.8
Average		43.7	12.9		

Table 7

Statistics comparing reader/tester data by individual readers for pollock in 1986. Reader A is less experienced than reader B. Bias is between-reader bias; nominal age is tester age.

Age (yr)	Count	Percentage agreement	CV (%)	Bias (yr)	Percentage bias
Results for Reader A					
1	21	85.7	6.7	0.14	9.5
2	46	87.0	4.1	0.09	2.9
3	95	80.0	4.2	0.16	4.3
4	94	72.3	4.9	0.02	-0.3
5	76	69.7	4.2	0.09	1.2
6	59	62.7	4.8	0.07	0.5
7	63	38.1	7.5	-0.17	-3.5
8	122	44.3	6.0	-0.18	-3.0
9	30	26.7	10.0	-0.30	-5.0
10	5	0.0	7.4	-1.00	-10.5
11	3	0.0	6.7	-1.00	-9.5
12	5	20.0	12.5	-1.80	-17.6
13	1	0.0	11.8	-2.00	-16.7
14	1	0.0	10.9	-2.00	-15.4
Average		61.0	5.5		
Results for Reader B					
2	11	100.0	0.0	0.00	0.0
3	30	86.7	3.2	0.00	-0.8
4	57	80.7	3.3	0.05	0.7
5	34	76.5	4.0	0.06	0.4
6	44	79.5	2.8	0.00	-0.4
7	35	80.0	2.0	0.09	1.0
8	54	74.1	2.5	-0.02	-0.5
9	29	24.1	7.1	-0.28	-3.8
10	8	50.0	5.3	0.00	-0.6
11	14	50.0	6.0	-0.57	-6.0
12	3	33.3	3.9	0.00	-0.2
Average		72.4	3.4		

Because percentage agreement decreases with the age of fish (Tables 1–6), and age distributions vary greatly among different species and among samples of the same species, percentage agreement lends itself only to age-specific comparisons. This is illustrated above by Pacific ocean perch and sablefish which show a similar percentage agreement; however, Pacific ocean perch is much more "ageable" than sablefish, as reflected in the corresponding coefficients of variation.

Although percentage agreement and coefficient of variation both reflect the relative difficulty of ageing each species, only the coefficient of variation adjusts for the absolute age of the fish. Therefore, one might conclude that the easiest species to age are Pacific whiting (Table 1) and yellowfin sole (Table 2); and the medium–difficult species are Pacific ocean perch (Table 3) and walleye pollock (Table 4).

The most difficult species to age—species with unresolved criteria, or species for which readers needed further training—were Atka mackerel (Table 5) and sablefish (Table 6). In fact, in this study the age-reader for Atka mackerel was inexperienced with the species, and there were unresolved criteria for sablefish.

The most important usage of reader/tester data is in maintaining quality control. Unlike the data presented in Tables 1–6, for quality-control purposes we need to compare only one tester with one reader, with the nominal age being the tester age. For pollock (Table 7) there were significant between-reader biases at older ages in the case of inexperienced reader A. There were no such between-reader biases for the experienced reader B. To ensure data quality, these types of between-reader biases are constantly reviewed and the samples partially re-aged, before the data are released for use.

Sablefish is an especially difficult species to age (Table 8). There were so many problems with age determination that we suspended ageing, reviewed criteria with other ageing labs, and re-aged several large samples. It is probable that between-reader bias for this species can be reduced, but it is doubtful that the coefficient of variation for this species can be substantially reduced. Nevertheless, the availability of reader/tester data was useful in revealing problems. Also, data users deserve a quantitative presentation of variability in age determinations and may be able to use these

Table 8

Statistics comparing reader/tester data for individual Reader A ageing sablefish in 1986. Bias is between-reader bias; nominal age is tester age.

Age (yr)	Count	Percentage agreement	CV (%)	Bias (yr)	Percentage bias
1	16	81.3	8.8	0.19	12.5
2	47	80.9	7.3	0.19	4.6
3	40	62.5	10.9	-0.08	-6.6
4	41	34.1	14.7	0.68	12.5
5	62	41.9	14.1	0.82	10.0
6	40	42.5	12.2	1.05	12.7
7	29	37.9	11.9	1.28	14.6
8	23	8.7	21.6	2.83	24.2
9	15	0.0	21.4	2.87	23.7
10	6	0.0	14.4	1.67	13.3
11	8	12.5	15.6	2.88	17.2
12	6	16.7	13.3	1.17	6.4
14	1	0.0	4.9	1.00	6.9
16	2	0.0	19.2	-1.00	-9.9
17	1	0.0	11.5	3.00	16.2
18	1	0.0	4.0	-1.00	-5.7
25	1	0.0	15.7	-5.00	-22.2
28	1	0.0	2.5	1.00	3.5
29	1	100.0	0.0	0.00	0.0
Average		43.7	12.9		

data for making decisions on aspects of their data analysis and modeling.

We examined the question of consistency between the percentage agreement and coefficient of variation measures of variability when analyzing between-reader data. Earlier we showed that assuming a constant coefficient of variation, and the normal error model, the percentage agreement can be predicted for all nominal ages. By comparing these theoretical curves with estimated percentage agreements calculated from data (Tables 1–6), some confidence in the consistency of the two measures can be derived. However, this comparison is crude due to the probable existence of between-reader biases that are not factored into the analysis.

Four different values for the coefficients of variation were used to calculate theoretical percentage agreement curves (Fig. 1A). These curves were then compared with estimated percentage agreement values for yellowfin sole (CV = 0.032, Fig. 1B), walleye pollock (CV = 0.050, Fig. 1C), and sablefish (CV = 0.129, Fig. 1D).

The percentage agreements for all three species appear consistent with the hypothesis that the coefficient of variation is constant over a wide age range, although the percentage agreements for pollock are biased low. These results support averaging the coefficient of variation across age ranges, and generally support using the coefficient of variation for interpreting precision data from age-determination studies. However, there is considerable variation in these data which makes our results somewhat tentative.

An important factor that also affects the ageing process is the presence of a strong year-class. For example, if two adjacent year-classes have absolute strengths of 10 and 100 fish, a 10% imprecision of ±1 year will add 5 fish from the strong cohort to the weak one (a 50% change) but only one-half a fish from the weak year-class to the strong (a 0.5% change). The data on Pacific whiting (Table 1) show how this phenomenon can lead to poor percentage agreements for weaker year-classes.

Users of age data are often concerned that after some age, say 9 years, for example, the dominant year-classes become spread over several ages. Since percentage agreement by such an age has often decreased to 50% or less, it is expected that age distributions will be smoothed. The only reason the ages would not be smoothed is if the dominant year-class is being anticipated by the age-reader. For example, if samples are 90% 10-year-olds, it would be difficult for the age-reader not to anticipate that age and between-reader agreement would be high. However, if say 9-, 10-, and 11-year- olds occur in equal numbers, the agreement would not be nearly as good.

Two possible ways of handling this problem are evident. A controversial method would be to assure that all age-readers are reasonably coached as to the probable occurrence of a strong year-class; the other is to group the older ages in any model analyzing these data (e.g., Deriso et al. 1989). Both approaches avoid asking the age-reader to perform the impossible.

Finally, interpretation and analysis of repeated readings given here assume that the repeated readings were statistically independent. In the present context, this simply means that each reader did not have information regarding the other reader's results. When repeated readings are not made on an independent basis, or are of inadequate sample size, the data will be difficult or impossible to interpret statistically. From such a database, it is impossible to make assertions regarding precision.

Acknowledgments

Mr. George Hirschhorn had the foresight to initiate the reader/tester quality control system. We thank the age readers of the Ageing Unit at the Alaska Fisheries Science Center who provided the data for this study. We also thank the Scientific Editor and three anonymous referees whose comments contributed significantly to this paper.

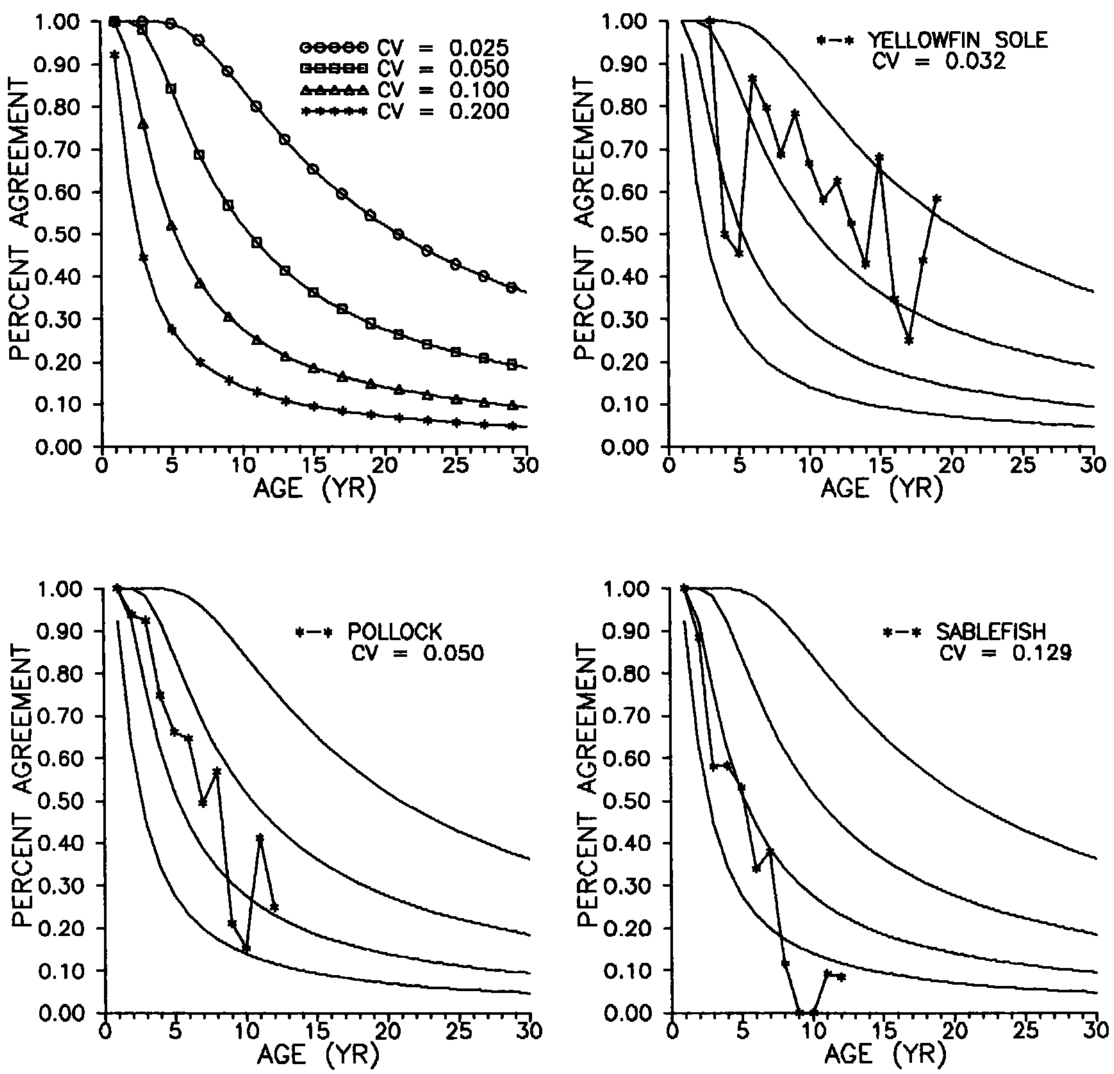

Figure 1

Predicted percentage agreement curves assuming different coefficients of variation and normally distributed error (upper left); overlayed with estimated percentage agreements for yellowfin sole (upper right); overlayed with estimated percentage agreements for pollock (lower left); and overlayed with estimated percentage agreements for sablefish (lower right).

Citations

Beamish, R.J., and D.A. Fournier

1981 A method for comparing the precision of a set of age determinations. Can. J. Fish. Aquat. Sci. 38:982–983.

Beamish, R.J., and G.A. McFarlane

1983 The forgotten requirement for age validation in fisheries biology. Trans. Am. Fish. Soc. 112:735–743.

1987 Current trends in age determination methodology. *In* Summerfelt, R.C., and G.E. Hall (eds.), Age and growth of fish, p. 15–42. Iowa State Univ. Press, Ames.

Beamish, R.J., G.A. McFarlane, and D.E. Chilton

1983 Use of oxytetracycline and other methods to validate a method of age determination for sablefish (*Anoplopoma fimbria*). *In* Proceedings of the international sablefish symposium, p. 95–116. Alaska Sea Grant Rep. 83-8, Univ. Alaska, Fairbanks.

Bennett, J.T., G.W. Boehlert, and K.K. Turekian

1982 Confirmation of longevity in *Sebastes diploproa* (Pisces: Scorpaenidae) from Ra-226/Pb-210 measurements in otoliths. Mar. Biol. (Berl.) 71:209–215.

Campana, S.E., K.C.T. Zwanenburg, and J.N. Smith

1990 ^{210}Pb/^{226}Ra determination of longevity in redfish. Can. J. Fish. Aquat. Sci. 47:163–165.

Cass, A.J., and R.J. Beamish

1983 First evidence of validity of the fin-ray method of age

determination for marine fishes. N. Am. J. Fish. Manage. 3:182–188.
Chang, W.Y.B.
1982 A statistical method for evaluating the reproducibility of age determination. Can. J. Fish. Aquat. Sci. 39:1208–1210.
Chilton, D.E., and R.J. Beamish
1982 Age determination methods for fishes studied by the groundfish program at the Pacific Biological Station. Can. Spec. Publ. Fish. Aquat. Sci. 60, 102 p.
Deriso, R.B., T.J. Quinn II, and P.R. Neal
1989 Further aspects of catch-age analysis with auxiliary information. *In* Beamish, R.J., and G.A. McFarlane (eds.), Effects of ocean variability on recruitment and an evaluation of parameters used in stock assessment, p. 127–135. Can. Spec. Publ. Fish. Aquat. Sci. 108.
Kimura, D.K., R.R. Mandapat, and S.L. Oxford
1979 Method, validity, and variability in the age determination of yellowtail rockfish (*Sebastes flavidus*), using otoliths. J. Fish. Res. Board Can. 36:377–383.
Leaman, B.M., and D.A. Nagtegaal
1987 Age validation and revised natural mortality rate for yellowtail rockfish. Trans. Am. Fish. Soc. 116:171–175.
McFarlane, G.A., and R.J. Beamish
1987 Validation of the dorsal spine method of age determination for spiny dogfish. *In* Summerfelt, R.C., and G.E. Hall (eds.), Age and growth of fish, p. 287–300. Iowa State Univ. Press, Ames.
Theobald, C.M., and J.R. Mallinson
1978 Comparative calibration, linear structural relationships and congeneric measurements. Biometrics 34:39–45.

National Oceanic and Atmospheric Administration

ERRATA NOTICE

One or more conditions of the original document may affect the quality of the image, such as:

Discolored pages
Faded or light ink
Binding intrudes into the text

This has been a co-operative project between the NOAA Central Library and the Climate Database Modernization Program, National Climate Data Center (NCDC). To view the original document, please contact the NOAA Central Library in Silver Spring, MD at (301) 713-2607 x124 or www.reference@nodc.noaa.gov.

LASON
Imaging Contractor
12200 Kiln Court
Beltsville, MD 20704-1387
March 21, 2005