PRECISION OF AGE DETERMINATION AND THE EFFECT ON ESTIMATES OF ABUNDANCE AND MORTALITY AMONG PACIFIC HERRING

By

Linda K. Brannian

Regional Information Report[1] No. 2A88-11

Alaska Department of Fish and Game
Division of Commercial Fisheries
333 Raspberry Road
Anchorage, Alaska

November 1988

[1] Contribution 88-11 from the Lower Cook Inlet area. The Regional Information Report Series was established in 1988 to provide an information access system for all unpublished divisional reports. These reports frequently serve diverse ad hoc informational purposes or archive basic uninterpreted data. To accommodate needs for up-to-data information, reports in this series may contain preliminary data.

AUTHORS

Linda K. Brannian is Region II Biometrician for the Alaska Department of Fish and Game, Division of Commercial Fisheries, 333 Raspberry Rd, Anchorage, AK 99518.

ACKNOWLEDGEMENTS

The author would like to thank Charles Meacham for his initial assignment and his encouragement for completion of this project. My thanks also go to the participants of this study, without whose experience in ageing scales this project would not have been possible. Their enthusiasm and quick response time on ageing scales was greatly appreciated. I would also like to thank Fritz Funk, Henry Yuen, Kathy Rowell, and Chuck Meacham for review of the manuscript.

TABLE OF CONTENTS

	Page
LIST OF TABLES	iv
LIST OF FIGURES	v
LIST OF APPENDICES	vi
ABSTRACT	vii
INTRODUCTION	1
METHODS	3
RESULTS AND DISCUSSION	6
LITERATURE CITED	10
APPENDICES	32

LIST OF TABLES

Table		Page
1	Comparison between the age composition estimates of reader repetitions. The critical value of the chi-square statistic was 20.203 for a CER of 0.0167	12
2	Chi-square statistics resulting from the comparison of the age frequencies of the first repetition of two readers. The critical value of the chi-square statistic was 22.78 with 9 degrees of freedom and a CER of 0.0067	13
3	Mean age in number of years as determined from herring scales for each area-reader-repetition combination. Unreadable scales were not included	14
4	Percent agreement and the average percent error (APE) between repetitions of ageing Pacific herring scales	15
5	Number of scales designated unreadable for each area-reader-repetition combination and the mean age of those scales as read by other readers	16
6	Comparison between the number of scales deemed unreadable between repetitions of readers. The critical value of the chi-square statistic was 5.731 for a CER of 0.0167	17

LIST OF FIGURES

Figure		Page
1	Major stocks of Pacific herring in Central Region of the State of Alaska Division of Commercial Fisheries	18
2	Proportion by age class as determined through repetitive ageings of Pacific herring scales by reader 1	19
3	Proportion by age class as determined through repetitive ageings of Pacific herring scales by reader 2	20
4	Proportion by age class as determined through repetitive ageings of Pacific herring scales by reader 3	21
5	Proportion by age class as determined through repetitive ageings of Pacific herring scales by reader 4	22
6	Proportion by age class as determined through repetitive ageings of Pacific herring scales by reader 5	23
7	Proportion by age class as determined through repetitive ageings of Pacific herring scales by reader 6	24
8	Average absolute difference between individual ages and their mean age from Pacific herring scales	25
9	Minimum and maximum observed for each mean age from Pacific herring scales	26
10	Standard deviation of age determinations by mean age for Pacific herring	27
11	Probabilities by age for correctly ageing a fish, ageing the fish 1 year too young, or 1 year too old. Probabilities used to model ageing error as a multinomial distribution with three outcomes	28
12	Abundance of herring by age for strong (top) and weak (bottom) cohorts. Observed lines as they differ from actual are the result of ageing imprecision	29
13	Average percent error of observed from actual for strong and weak cohorts	30
14	Abundance of herring by age for strong and weak cohorts. Observed lines as they differ from actual are the result of ageing imprecision	31

LIST OF APPENDICES

APPENDIX A Page

A.1 Date and time spent ageing 240 herring scales for each reader and repetition.......... 32

A.2 Percent by age class of 240 herring scales as read by six readers for two repetitions. Only ageable scales were used.... 33

A.3 Average percent error, average absolute difference and the minimum and maximum designated age for Pacific herring scales by age.......... 34

ABSTRACT

Scales have been used to determine the age structure of Pacific herring (*Clupea harengus pallasi*) populations in coastal waters of Alaska. Neither an age validation or method-structure comparison has been conducted for Alaskan stocks. Furthermore, age determination was assumed precise up to age 8 with older aged fish being pooled. Now that the time series of catch-at-age observations from sac roe fisheries is extensive enough for cohort analysis, a re-analysis of data for older aged fish was required. Catch-at-age analysis gives equal weight to all year classes and ages. Therefore, an understanding of the effect ageing imprecision would have on age determination of Pacific herring scales as it represented the historic data base was necessary.

A test set of 240 scales comprising 80 each from 3 major South Central Alaska herring stocks (Prince William Sound, Lower Cook Inlet, and Bristol Bay) were randomly chosen from those collected in 1983-87. Six persons were asked to age the test set twice. A significant difference was detected in the percent classified unreadable among reader's first repetition but no significant difference in the resulting age composition among readers was detected. Percent agreement between reader repetitions ranged from 70% to 85.7%. Precision was greatest for scales from Prince William Sound herring and for all areas decreased with increasing mean age of fish. Age-specific error was modeled as a multinomial random variable for simulating imprecise age determination. Using this error structure, adjacent strong and weak cohorts of herring were aged through their life span. The resulting patterns of decline over time suggested a potential bias in the estimation of abundance and mortality.

KEY WORDS: Ageing error, Precision, Pacific herring, *Clupea harengus*, Natural mortality

INTRODUCTION

Estimates of age composition are fundamental to the understanding of the biology and dynamics of fish populations but are often difficult to obtain. Difficulties arise in obtaining an unbiased random sample of the target population and validating a structure and method for accurate and precise age determination. Beamish and McFarlane (1983) stress the importance of validating all age groups of a fish for each species and population of interest. Ageing studies can be grouped into three categories (1) validation of age determination by various structures, (2) comparison of ageing structures and methods, and (3) evaluation of precision, repeatability, and agreement among readers of a single structure. Most studies involve the later two issues either proposing a choice of structure for ageing or evaluating the within and between reader precision or repeatability of age determinations. Following the definitions of Chilton and Beamish (1982) precision refers to the degree of repeatability and thus relates to the variability between readers or readings. Accuracy relates to the degree of closeness to the true age of the fish. Age may be validated using fish of known age through aquaculture or from mark and recapture data which can only validate for the years between capture. Even lacking validation, an accurate but imprecise estimate is as worthless as an inaccurate but precise one.

Great importance is placed on the accurate and precise determination of ages. Barlow (1984) discussed the consequences of not including all ages in the Chapman-Robson mortality estimate when age estimation is imprecise. Most cohort analysis assumes age is known without error. Fournier and Archibald (1982) discussed the consequences of ageing error and the possible need to pool across the oldest aged fish if accuracy is very poor. Deriso et al. (1985) describe a modification of catch-at-age analysis where observed catches are affected by random errors in the ageing of fish which they modelled by sampling from a multinomial distribution. The influence of measurement error on return/spawner relationships was investigated by Walters and Ludwig (1981) and Ludwig and Walters (1981) and involved measures of stock size which could include imprecise allocation to year class through ageing error.

Long lived fishes with highly variable recruitment often result in strong year classes which must support fisheries for a number of years. When catch-at-age analysis treats all cohorts equally even small ageing errors can greatly affect our perception of a cohort's decay through time. Large recruitment will appear smaller and vice-a-versa. Weak year classes neighboring strong ones will appear to have low fishing mortality at young ages and abnormally high mortalities at the oldest ages (Fournier and Archibald 1982).

Scales are the structure used to determine the age of Pacific herring (*Clupea harengus pallasi*) in British Columbia since the turn of the century and are the generally accepted method used by management agencies along the east Pacific coast. Prior to the use of otolith cross sections Messieh and Tibbo (1970) found otolith ages to be lower than scales and concluded that scales were more accurate for Atlantic herring (*Clupea harengus harengus*). Chilton and Stocker (1987) compared age determination using the break and burn method of processing otoliths with that of reading scales for Pacific herring. They found no significant difference in the age composition resulting from each structure, though otolith reading detected older aged fish (13-16 years). They concluded for British Columbia herring that scales were the most suitable structure for age determination due to ease of collection, preparation, and the generally younger mean age of their stocks. Otoliths have also been used for age determination in California (Spratt 1981). Finally, Chilton and Beamish (1982) recommend that scales not be used to age fishes suspected to be long lived (age 10 years or greater).

Neither an age validation or method-structure comparison has been conducted for Alaskan stocks of Pacific herring. Following the lead of East Pacific fisheries management agencies, scales have been collected to be read for age determination. Department archives contain data back to 1978 for Bristol Bay and 1973 for Prince William Sound, and Lower Cook Inlet. This corresponds to the development of sac roe fisheries on the three major stocks of herring in South Central Alaska (Figure 1). Furthermore for Prince William Sound and Bristol Bay age determination was assumed to be precise only up to age 8 and all fish age 9 or older were pooled into a category as age 9^+ (McCurdy 1986, McBride et al. 1981). Only recently has the time series of catch-at-age observations from the sac roe fisheries on these stocks been long enough to allow catch-at-age or "cohort" analysis (Pope 1972). Preliminary cohort analysis was conducted on the Bristol Bay stock (Wespestad 1982) and recently a more exhaustive approach was taken for Prince William Sound (Funk and Sandone *In press*). In addition age frequency distributions have been used to estimate total and natural mortality rates in Prince William Sound (Funk and Sandone *In press*), Bristol Bay (Brannian and Rowell *In press*) and Lower Cook Inlet (Yuen *In press*). Re-analysis and summarization of data for older age groups (up to 13 years) was performed in order to follow a cohort through to its apparent depletion. This was undertaken for Bristol Bay (Sandone *In press*) and Prince William Sound (Sandone 1988a and 1988b).

Given the interest in modeling the population dynamics of Alaskan herring stocks coupled with the addition of new scale readers to the staff, a review of our ageing was appropriate. Most Alaskan stocks have few strong cohorts supporting their sac roe fisheries and therefore are the case most affected by ageing error. The Bristol Bay herring stock has had successfully strong year classes on the average once every 5 years (1959-83 spawning events) over the past 25 years (Funk 1988). The 1978 year class was forecast to represent 32% of the population in 1988 as 10-year-old fish. No significant recruitment has been seen since this year class recruited to the fishery as age 4 in 1982. For Prince William Sound, only one strong year class (1976) has been observed since 1973 with the next largest (1980) being half that size (Funk and Sandone *In press*).

The objectives of this study were:

1. To assess the precision and repeatability of age determinations of Pacific herring scales collected from Prince William Sound, Lower Cook Inlet, and Bristol Bay.

2. To determine whether precision varies with mean age of fish, stock, or reader.

3. To determine if the percent of unreadable scales varies among readers or stocks and if so the effect of differences on age composition.

4. To investigate the consequences of ageing error on the estimation of abundance for strong and weak year classes and the resulting estimates of mortality.

METHODS

A random sample of 80 herring scales was selected from each previous 5 year (1983-87) collection for Prince William Sound (PWS), Lower Cook Inlet (LCI), and Bristol Bay (BB). Current year samples were excluded as they had been recently aged by at least one member participating in the study. Scales were not screened for readability nor selected to represent a particular age composition. Study results were to be representative of the age composition data base of these three stocks.

Scales had been collected from the preferred area of Pacific herring above the left pectoral fin and lateral line approximately 3-4 scales posterior to the fin origin. They were cleaned, dipped in a 10% mucilage solution, and positioned unsculptured side down on a labeled glass slide. Slides contained up to 10 scales each (or 20 for LCI) and were stored with a cover slide.

The label on each scale slide containing the location, date, and capture gear was obscured from the reader. The scales on each slide were read for age determination in the same order that they were mounted on the glass, with the label on the left, the first scale on the upper left and the tenth or 20^{th} (for LCI) positioned at the lower right. The existence of slides of 2 different dimensions resulted in a reader knowing whether or not scales were from Lower Cook Inlet. The order in which scale slides were read was randomly chosen by the author for each reader and replicate. They were asked to record the time expended, equipment used, and date on which herring scales were last aged. At least 2 weeks and more commonly a month lapsed between repetitive ageing of the test sample. Readers did not know the length of the fish or year of capture. It was thought that local knowledge of an area's dominate age class present in any one year might prejudice ageing for that time-location combination and were therefore concealed. Scale images were projected on a microfiche screen and read in a darkened room.

The average percent error (APE) suggested by Beamish and Fournier (1981) and Chilton and Beamish (1982) was used to compare the precision or repeatability of readers. A sample of N fish were aged, M number of times by each reader. The age in years was X_{ij} for reading i of fish j. The APE for a particular reader was defined as:

$$APE = 100\ N^{-1} \sum_{j=1}^{N} \left(M^{-1} \sum_{i=1}^{M} \left| \frac{X_{ij} - X_j}{X_j} \right| \right) \qquad [1]$$

where X_j is the average across i readings of fish j. The overall APE was calculated across all readers and readings of a scale or when M=12.

The mean age (X_j) across all readings of fish j was used as its true age in order to evaluate how precision of ageing varied with age of fish. Mean ages were grouped according to their nearest integer age in order to estimate the average associated error in age determination for each age k. Thus, the number of age k fish (N) became only those fish, j, whose mean age (X_j) fell into the range of k-0.5, k+0.499. Average absolute difference (AAD) was used to measure precision as it was unweighted by mean age. Thus the AAD for age k was equation 1 not divided by the mean age (X_j) across the 12 readings (M=12).

Chi-square tests were used to test the hypothesis that similar criteria for age determination were used by a reader for each ageing repetition and a common criteria among different readers. Age frequencies for readers and repetitions consisted of age classes 4-11 with a separate category for old (≥12 years) and young (≤3 years) fish. The Bonferroni inequality (SAS 1985) was used to set a comparisonwise error rate (CER) given a type one error rate (α) where c is the number of simultaneous comparisons being made. Therefore

$$CER = \alpha/c$$

forms the critical value of the test or in a sense the new "α" for this multiple comparison.

When average differences between readings, readers, or areas are less than 1 year these differences would not be evident when such ages are subsequently reported as integers. Therefore statements concerning significant differences resulting from statistical analysis would not be meaningful. Tests for significance were not conducted when results differed by less than 1 year.

Ultimately, age compositions are used to describe various year class strengths, and to facilitate forecasting and catch-at-age analysis for the population of interest. It was of interest to simulate the effect ageing error would have on the estimation of total or natural mortality and our perception of a cohort's decline in abundance through time. The decline of a year class through time was modeled as:

$$N_0 = RS^0$$

where S is e^{-Z}, the rate of survival, and total mortality Z can be further broken into the component due to fishing (F) and natural mortality (M) such that

$$Z = F + M.$$

Size of the cohort is R and represents abundance as it recruits into the fishery which for the purposes of this simulation was set at age 2. The resulting population size (N_k) can be estimated annually as it declines through time as:

$$N_1 = RS^1$$
$$N_2 = RS^2$$
$$\vdots$$
$$N_k = RS^k \quad \text{or } N_k = Re^{-Zk} \qquad [2]$$

A linear form of this relationship is used to estimate mortality:

$$\ln(N_k) = \ln(R) - Zk \qquad [3]$$

If catch data were the only population information available and fishing mortality or rate of exploitation (μ) had been constant the following relationship can be used to estimate mortality:

$$C_k = \mu RS^k \qquad \text{or linearized as} \quad \ln(C_k) = \ln(\mu R) - Zk.$$

Ageing error entered into this model when allocating catch or total return to year class y of age k through scale ageing. Ageing error was defined as the imprecision in determining age as seen among readers and repetitions in this study. It was assumed to follow a multinomial distribution with three outcomes correctly aged, over or under aged by one year. If fish X_{jk} were really of age k, the ages that would be determined from multiple ageings of its scales would be multinomially distributed where the probabilities of the three age outcomes of k-1, k, k+1 were estimated from this study. Therefore for cohort $N_{k,y}$ only a proportion, p_{ky}, of year class y at age k would be properly aged as k and some proportion would be aged older (k+1) and younger (k-1). Proportions were estimated from calculating the area under the standard normal curve having standardized ages $((X_j-k)/\sigma_k)$. The following probabilities of being aged younger (k-1), correctly (k), or older (k+1) were defined:

$P(-1.5 \le x \le -0.5) = p_{k-1,y}$ Proportion of year class y aged as k-1 years

$P(-0.5 \le x \le 0.5) = p_{k,y}$ Proportion of year class y aged correctly as k

$P(0.5 \le x \le 1.5) = p_{k+1,y}$ Proportion of year class y aged as k+1 years

The number of age k fish for year class y was estimated as:

$$\hat{N}_{k,y} = p_{k,y}R_yS^k + p_{k+1,y+1}R_{y+1}S^{k'} + p_{k-1,y-1}R_{y-1}S^{k''} \qquad [4]$$

where:

$k' = k-1$
$k'' = k+1$

For example, due to ageing error the observed number of 8-year-old fish in 1988 from the 1980 year class would be:

$$\hat{N}_{8,80} = p_{8,80}R_{80}S^8 + p_{7+1,81}R_{81}S^7 + p_{9-1,79}R_{79}S^9$$

Thus the observed size of a cohort consists of a portion of the 1980 year class ($N_{8,80}$) correctly aged and portions of age 7, 1981 year class and age 9, 1979 year class incorrectly age as age 8 due to ageing error.

RESULTS AND DISCUSSION

Six readers participated in the study. Each was a permanent staff member either experienced in ageing herring scales or with current herring ageing responsibilities. Experience in ageing scales ranged from 1 to 10 years (Appendix A.1). Each participant was most familiar ageing herring scales from their own area and no reader had aged scales from all areas. The average length of time spent ageing the test set of 240 scales was 1 h 42 min or 2.4 scales per minute. The time needed varied from 2 h 44 min to 52 min (Appendix A.1). All participants had either aged scales within the last 2 months, the period of the sac roe fishery, or re-aged a test set of their own 1988 scales to re-familiarize themselves before examining the test set.

Histograms depicting the age composition as estimated by each reader and repetition show similar contributions (Figures 2-7, Appendix A.2) for ages 5-8 with a small showing of fish up to 14 years of age and as young as age 2. In order to contain the type 1 error rate (α=0.05) a simultaneous comparison between repetitions was made for the 6 readers. No significant difference was detected (Table 1) between the age frequencies estimated repetitively by any of the 6 readers and the p-values were all 0.822 or larger. This indicates that readers are using a consistent criteria to determine age and the resulting age compositions are reproducible.

It was also of interest to investigate whether the readers use similar criteria and can produce a common age composition. No significant difference (χ^2=36.3, df=45, p-value= 0.819) was detected among the age frequencies of each reader's first repetition. In other words, though Figures 2-7 may seem slightly different, statistically, they could all have been produced using the same ageing criteria. The largest contribution to the chi-square statistic was reader 5 and the number in the oldest age category reported by reader 1. The non-significant result was somewhat surprising in viewing the histograms, especially when histograms of reader 3 and 5 are compared. When all 15 pairwise comparisons were made (Table 2) the largest discrepancy was between reader 3 and 5. Reader 5 was overall the most unique judging by the magnitude of the chi-square statistic in each of its pairwise comparisons. Yet a simultaneous comparison needed to keep the type 1 error rate to an acceptable limit found all pairs to be non-significant (if α=0.2 the CER = 0.2/15 = 0.013). In other words, when the extreme histograms are chosen they appear different but are not statistically different in the context of the 4 other readers.

Mean ages as determined by each reader and repetition varied by less than 0.4 years (Table 3). The mean age of scales from Bristol Bay were oldest (7.3-7.9 years) and Prince William Sound the youngest (5.5-5.8 years). The range in mean ages was 0.6 years for Bristol Bay, 0.4 years for Lower Cook Inlet, and 0.3 years for Prince William Sound. There was also a slightly greater difference between readers than readings. No statistical analysis was conducted for even if differences were significant they would not be meaningful when age is reported as an integer.

Percent agreement for each reader, the number of times the second age determination differed from the first, ranged from 70% to 85.7% (Table 4).

Percent agreement was highly correlated (r=0.8) with experience. Overall agreement among the total 12 estimates of age for each scale was 28.2%. When comparing only the first reading by each reader (6 total) there was 34% agreement. Unlike the APE, percent agreement treats all differences between repetitive ageings equally whether they are of one year or more. In contrast, the APE weights the difference by the mean age. Therefore both the magnitude of the difference and the associated mean age are important. For example, a one year discrepancy contributes a greater amount to the APE if the fish was young versus old. The greatest penalty to precision was due to discrepancies in determining if a scale was regenerated or unreadable which were given values of 18 and 19, respectively. Even so, the same relative pattern in precision is demonstrated by the readers. Reader 4 was the most precise using either statistic (greatest percent agreement and smallest APE). In general readers were most precise when ageing Prince William Sound herring which could be a function of their younger mean age. The next most precise area varied by reader and appeared to be the area or stock where they were most experienced in ageing.

Precision in age determination also varied with age of fish. The average absolute difference (AAD) increased steadily for ages 3-13 (Figure 8). The lack of precision was expected for older aged fish but was surprising for the younger ages. Figure 9 presents the minimum and maximum age designation for each age category. Given the trend in Figure 8 one would expect the two lines to form a fan opening toward the older ages. This picture, instead, depicts the worse case in age designations (minimum and maximum) and shows lines that are fairly parallel for mean age 5-11 years. The difference between the minimum and maximum age for these ages was 3 to 5 years (Appendix A.3). From Figure 9 it is evident that the magnitude of the discrepancy is the same over a wide range of ages (5-11 years) although Figure 8 shows the frequency of occurrence to be less for younger aged fish. Figure 10 presents the observed increase in the standard deviation in age determination by age class. This was further smoothed by fitting a line between ages 3 and 12 (Appendix A.3). Even at the oldest ages (≥11 years) if ageing imprecision is normally distributed over 90% of the fish being aged are within ±1 year of their true age.

Readers may affect the estimation of age composition and mean age through designating scales un-ageable due to regeneration or otherwise. This would occur, for example, if there was a greater likelihood of designating older scales unreadable. Percent unreadable ranged from 15% to 2.9% (Table 5) between readers and repetitions indicating dissimilar criteria for designating a scale ageable. Only 4 of the 240 scales were deemed unreadable across all 12 reader repetitions. Fifty-seven scales were deemed unreadable at least once. There was a significant difference (α=0.05) between the frequency of unreadable scales (χ^2=34.7, df=5, p-value <0.001) among the readers first repetition. The frequency of unreadable scales did not vary significantly (α=0.05) among stocks (χ^2=1.1, df=2, p-value=0.57) indicating there was no propensity for regeneration or difficulty in reading in any of the three area's scale collections. In general this difference between readers did not carry over to between reader repetitions. Only reader 3 demonstrated a significantly different (Table 6) criteria for determining readability between ageing repetitions. All other readers were able to reproduce their percent readable.

The effect of differing criteria for designating scales unreadable was evaluated by comparing the mean age of readable scales from Table 3 with the mean age, as designated by others for all unreadable scales for that reader and repetition (Table 5). The difference between mean age of readable and unreadable scales was less than 1 year for all readers except reader 4 which was based on 3 fish and differed by up to 2.2 years. Where the number of unreadable scales was largest (>22 for reader 2 and 3), the mean ages of readable and unreadable scales were very similar. There does not seem to be a bias towards rejecting older aged fish when determining readability. Yet the significantly different unreadable rates should be taken into account when developing sample size goals for a particular stock and reader combination.

It was of interest to assess what affect ageing imprecision, demonstrated in this study, would have on estimates of abundance. A total mortality rate of 0.4 (Z) was used to model the exponential decline of a herring cohort through time (equation 2). How observed abundance (equation 4) differed from actual abundance (equation 2) at age would be due to ageing imprecision. Ageing imprecision was modeled as a random error in the ageing of fish and was simulated by sampling from a multinomial distribution with three outcomes. The multinomial probabilities of being correctly aged (p_k of equation 4) was found to decrease with increasing age ranging from 100% through age 4 to 44% correctly aged at 13 years. The probability of underageing one year was set equal to the probability of over ageing one year (p_{k-1}, p_{k+1}) and increased to 28% by age 13 (Figure 11).

Ageing error has the smallest affect on the estimation of year class strength when recruitment is constant over time, and generally affects only the youngest and oldest observed age classes. In this case the number incorrectly aged a year too old and too young are nearly equaled by the number of adjacent year classes mis-apportioned. It would be equal if the cohorts decay through time were linear. The result using equation 2 for actual abundance and equation 4 to calculate what is observed due to ageing error is a very small overestimation of abundance in all but the youngest age class.

The case of interest is when year class strength varies and especially when strong and weak year classes are adjacent. The two scenarios modeled were (1) the abundance of a strong year class from a recruitment pattern of weak followed by strong followed by weak recruitment and (2) a weak year class from the center of a recruitment pattern of weak-weak-strong. The strong cohorts were 10 times in abundance to the weak. Using the multinomial probabilities from Figure 11 and a total mortality of 0.4 strong and weak cohorts were aged through time with and without ageing error (Figure 12). Ageing error does not seem to affect our perception of abundance through age 5 of a strong year class and age 3 of a weak one. For older ages represented in these two models imprecise age determination and variable year class strength interact. Abundance of a strong year class is underestimated and a weak cohorts is overestimated (Figure 12). The overestimation of a weak cohort increased with increasing age to a 95% error by age 12 (Figure 13). The percent error in the underestimation of a strong year class increased in magnitude to 46% by age 12.

To understand how ageing imprecision might affect the estimate of mortality, data of Figure 12 were transformed as in equation 3 (Figure 14). The slope of the lines are estimates of total mortality. The resulting estimate of total mortality using ages 3-14 years was 0.47 for a strong year class and 0.33 for a weak cohort, a 18% error. When these two cohorts were standardized to a common scale (say using proportion by age) and a common line fit, the resulting estimate of mortality was 0.396.

From this study it appears that ageing error is not a problem for younger aged fish. In addition the abundance we observe by estimating age composition with error can be a problem for weak cohorts neighboring strong cohorts. The problem is less severe for strong year classes. Mortality should be estimated across year classes of various strengths. A validation or structure comparison is recommended to study the accuracy in ageing older herring.

LITERATURE CITED

Barlow, J. 1984. Mortality estimation: biased results from unbiased ages. Canadian Journal of Fisheries and Aquatic Sciences 41:1843-1847.

Beamish, R.J. and D.A. Fournier. 1981. A method for comparing the precision of a set of age determinations. Canadian Journal of Fisheries and Aquatic Sciences 38:982-983

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

Brannian, L.K. and K.A. Rowell. *In Press*. Forecast of the Pacific herring biomass in Togiak District, Bristol Bay, 1989. Alaska Department of Fish and Game, Division of Commercial Fisheries, Region II, Regional Information Report 2A88-#, Anchorage.

Chilton, D.E. and R.J. Beamish. 1982. Age determination methods for fishes studied by the groundfish program at the Pacific biological station. Department of Fisheries and Oceans, Canadian Special Publication of Fisheries and Aquatic Sciences 60, Ottawa.

Chilton, D.E. and M. Stocker. 1987. A comparison of otolith and scale methods for aging Pacific herring. North American Journal of Fisheries Management 7:202-206.

Deriso, R.B., T.J.Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Canadian Journal of Fisheries and Aquatic Sciences 42:815-824.

Fournier, D. and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Canadian Journal of Fisheries and Aquatic Sciences 39:1195-1207.

Funk, F.C. 1988. Status of Bristol Bay herring stocks in 1988. Alaska Department of Fish and Game, Division of Commercial Fisheries, Regional Information Report 5J88-02, Juneau.

Funk, F.C. and G.J. Sandone. *In Press*. Stock assessment of Prince William Sound Herring 1973-1987, using cohort analysis. Alaska Department of Fish and Game, Division of Commercial Fisheries, Fishery Research Bulletin, Juneau.

Ludwig, D. and C.J. Walters. 1981. Measurement errors and uncertainty in parameter estimates for stock and recruitment. Canadian Journal of Fisheries and Aquatic Sciences 38:711-720.

McBride, D., C. Whitmore, and D. Bergstrom. 1981. Age, sex, and size composition of Pacific herring, *Clupea harengus pallasi* (Vallenciennes), from selected coastal spawning sites along the eastern Bering Sea, 1979-1980. Alaska Department of Fish and Game, Division of Commercial Fisheries, Technical Data Report 61, Juneau.

McCurdy, M.L. 1986. Prince William Sound herring commercial fishery age, weight, length and sex data 1973 through 1983. Alaska Department of Fish and Game, Division of Commercial Fisheries, Prince William Sound Data Report 86-3, Anchorage.

Messieh, S.N. and S.N. Tibbo. 1970. A critique on the use of otoliths for ageing Gulf of St. Lawrence herring (*Clupea harengus L.*). Journal Du Conseil, Conseil International pour l'Exploration de la Mer 33(2):181-191.

Pope, J.G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. Research Bulletin of the International Northwest Atlantic Fisheries Commission 9:65-74.

SAS 1985. SAS user's guide, statistics, version 5 edition. SAS Institute Inc., Cary, North Carolina.

Sandone, G.J. 1988a. Age, sex, and size composition of Pacific herring sampled from the Prince William Sound Management Area, 1973-1983. Alaska Department of Fish and Game, Division of Commercial Fisheries, Region II, Regional Information Report 2A88-06, Anchorage.

Sandone, G.J. 1988b. Age, sex, and size composition of Pacific herring sampled from the Prince William Sound Management Area, 1984-1987. Alaska Department of Fish and Game, Division of Commercial Fisheries, Region II, Regional Information Report 2A88-08, Anchorage.

Spratt, J.D. 1981. Status of the Pacific herring, *Clupea harengus pallasi*, resource in California 1972 to 1980. California Department of Fish and Game, Fish Bulletin 171.

Walters, C.J. and D. Ludwig. 1981. Effects of measurement errors on the assessment of stock-recruitment relationships. Canadian Journal of Fisheries and Aquatic Sciences 38:704-710.

Wespestad, V.G. 1982. Cohort analysis of catch data on Pacific herring in the eastern Bering Sea, 1959-1981. National Marine Fisheries Service, Northwest and Alaska Fisheries Center, NOAA Technical Memorandum F/NWC-24, Seattle.

Table 1. Comparison between the age composition estimates of reader repetitions. The critical value of the chi-square statistic was 20.203 for a CER of 0.0167.[a]

	Reader					
	1	2	3	4	5	6
Chi-square (χ^2)	0.514	0.582	5.137	1.676	3.081	0.692
Probability ($x>\chi^2$)[b]	>0.999	>0.999	0.822	0.996	0.961	>0.999
Degrees of Freedom	9	9	9	9	9	9

[a] The Bonferroni inequality (SAS 1985) was used to set a comparisonwise error rate (CER) where $\alpha = 0.1$ and c is the number of simultaneous comparisons being made. CER = α/c.

[b] P-value for the above chi-square statistic.

Table 2. Chi-square statistics resulting from the comparison of the age frequencies of the first repetition of two readers. The critical value of the chi-square statistic was 22.78 with 9 degrees of freedom and a CER of 0.013.[a]

	Chi-square Statistic (P-value)				
Reader	2	3	4	5	6
1	3.626 (0.934)	7.673 (0.567)	5.978 (0.742)	10.052 (0.346)	6.115 (0.728)
2		4.565 (0.870)	2.493 (0.981)	10.079 (0.344)	3.962 (0.914)
3			3.154 (0.958)	14.102 (0.119)	3.285 (0.952)
4				7.850 (0.549)	1.972 (0.992)
5					6.996 (0.637)

[a] The Bonferroni inequality (SAS 1985) was used to set a comparisonwise error rate (CER) where $\alpha = 0.2$ and c is the number of simultaneous comparisons being made. CER = α/c.

Table 3. Mean age in number of years as determined from herring scales for each area-reader-repetition combination. Unreadable scales were not included.

Reader	Repetition	Mean Age by Area [a]			
		BB	LCI	PWS	Total
1	1	7.6	6.7	5.6	6.7
	2	7.7	6.5	5.7	6.6
2	1	7.4	6.6	5.7	6.6
	2	7.5	6.8	5.6	6.6
3	1	7.3	6.4	5.5	6.4
	2	7.6	6.8	5.6	6.7
4	1	7.4	6.7	5.5	6.6
	2	7.5	6.7	5.6	6.6
5	1	7.9	6.6	5.8	6.8
	2	7.7	6.5	5.7	6.6
6	1	7.4	6.4	5.6	6.5
	2	7.4	6.4	5.6	6.5

[a] Area designations are Bristol Bay (BB), Lower Cook Inlet (LCI), and Prince William Sound (PWS).

Table 4. Percent agreement and the average percent error (APE) between repetitions of ageing Pacific herring scales.

Reader	Percent Agreement	APE	APE by Area BB	LCI	PWS
1	82.9	3.0	1.9	3.8	3.3
2	80.4	3.3	3.7	3.8	2.5
3	70.0	5.1	7.2	6.0	2.3
4	85.7	1.4	1.8	1.5	1.0
5	70.4	3.9	4.8	4.6	2.3
6	80.0	2.7	2.7	3.1	2.2

Table 5. Number of scales designated unreadable for each area-reader -repetition combination and the mean age of those scales as read by other readers.

Reader	Repetition	Number of Scales BB	LCI	PWS	Total	Mean[a] Age	Percent Unreadable
1	1	3	5	6	14	6.6	5.8
	2	3	2	3	8	6.4	3.3
2	1	11	15	10	36	6.6	15.0
	2	9	10	8	27	6.4	11.3
3	1	11	9	3	23	7.2	9.6
	2	3	4	2	9	5.9	3.8
4	1	2	3	2	7	4.4	2.9
	2	2	3	2	7	5.5	2.9[b]
5	1	4	3	6	13	5.8	5.4
	2	7	5	5	17	6.9	7.1
6	1	4	4	3	11	6.6	4.6
	2	4	3	5	12	6.9	5.0

[a] Four scales were found to be un-readable by all readers and therefore the mean age is based on 4 less than the total un-aged.
[b] Two scales were damaged in shipping before reader 4 finished the second repetition. Percent unreadable was based on 238 scales.

Table 6. Comparison between the number of scales deemed unreadable between repetitions of readers. The critical value of the chi-square statistic was 5.731 for a CER of 0.0167.[a]

	Reader					
	1	2	3	4	5	6
Chi-square (χ^2)	1.715	1.480	6.562	<0.001	0.569	0.046
Probability ($x>\chi^2$)	0.190	0.224	0.010	>0.999	0.451	0.830
Degrees of Freedom	1	1	1	1	1	1

[a] The Bonferroni inequality (SAS 1985) was used to set a comparisonwise error rate (CER) where $\alpha = 0.1$ and c is the number of simultaneous comparisons being made. CER = α/c.

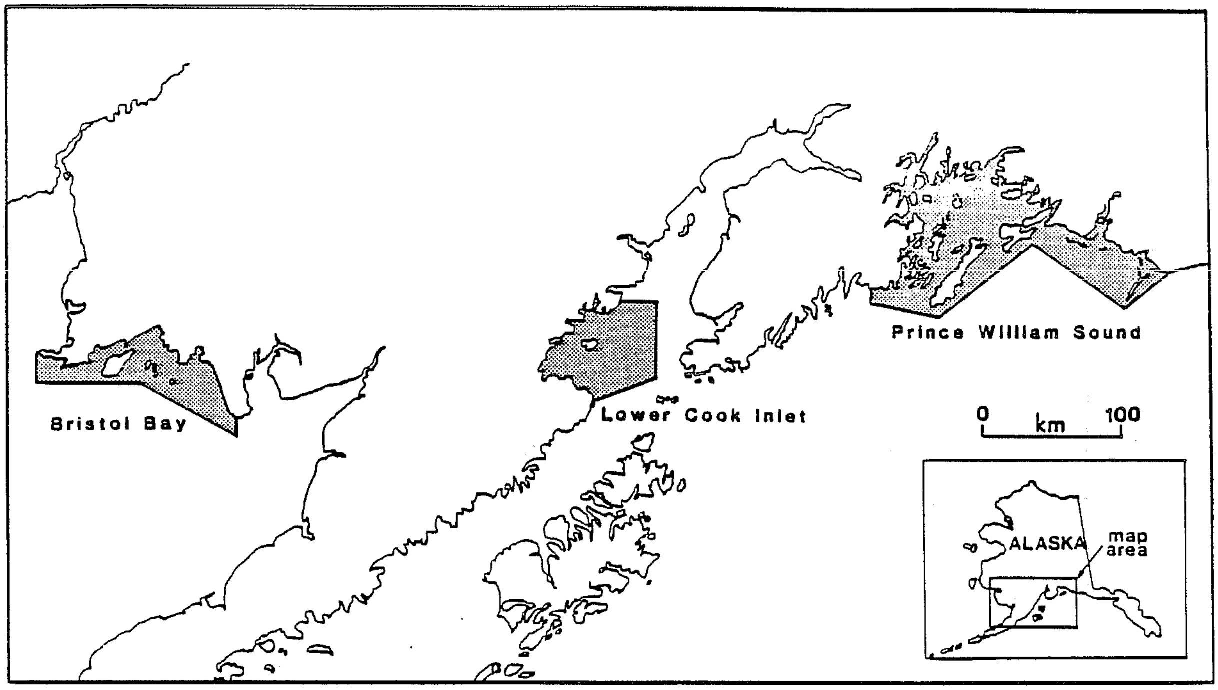

Figure 1. Major stocks of Pacific herring in Central Region of the State of Alaska Division of Commercial Fisheries.

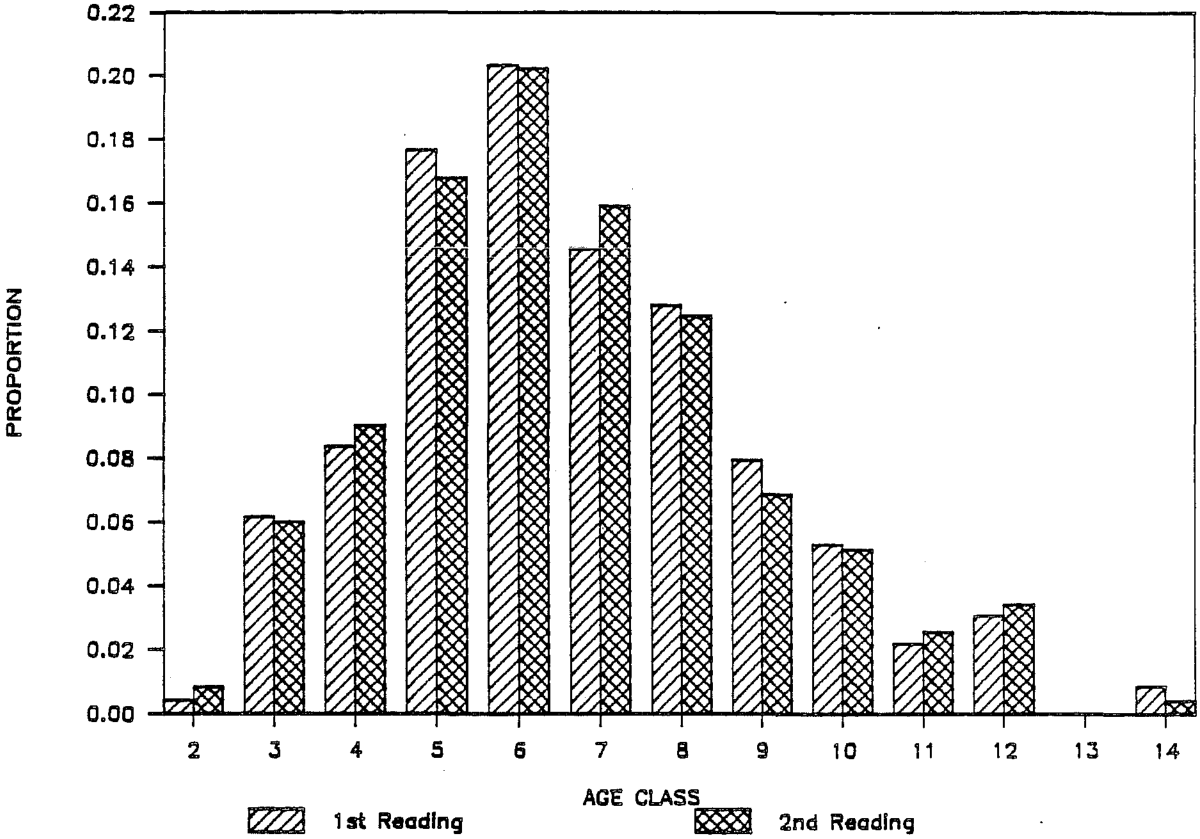

Figure 2. Proportion by age class as determined through repetitive ageings of Pacific herring scales by reader 1.

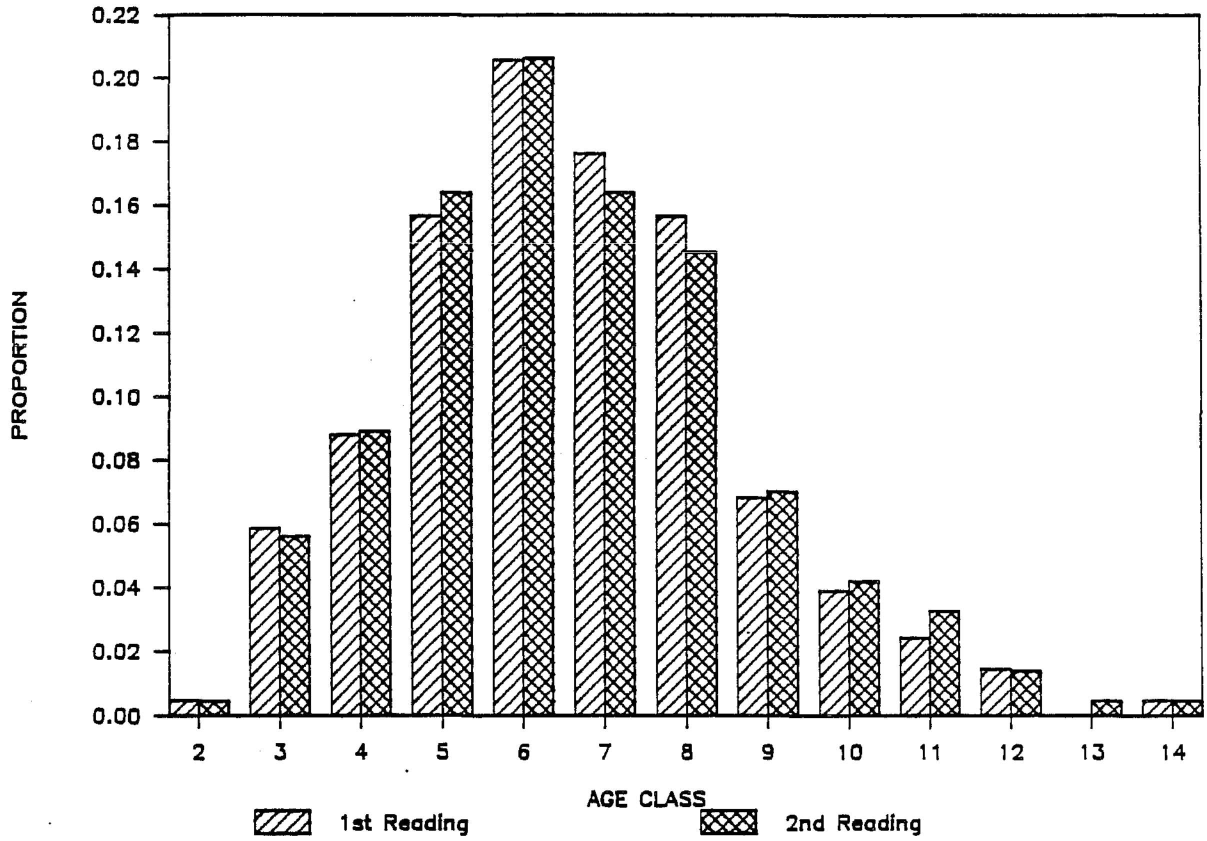

Figure 3. Proportion by age class as determined through repetitive ageings of Pacific herring scales by reader 2.

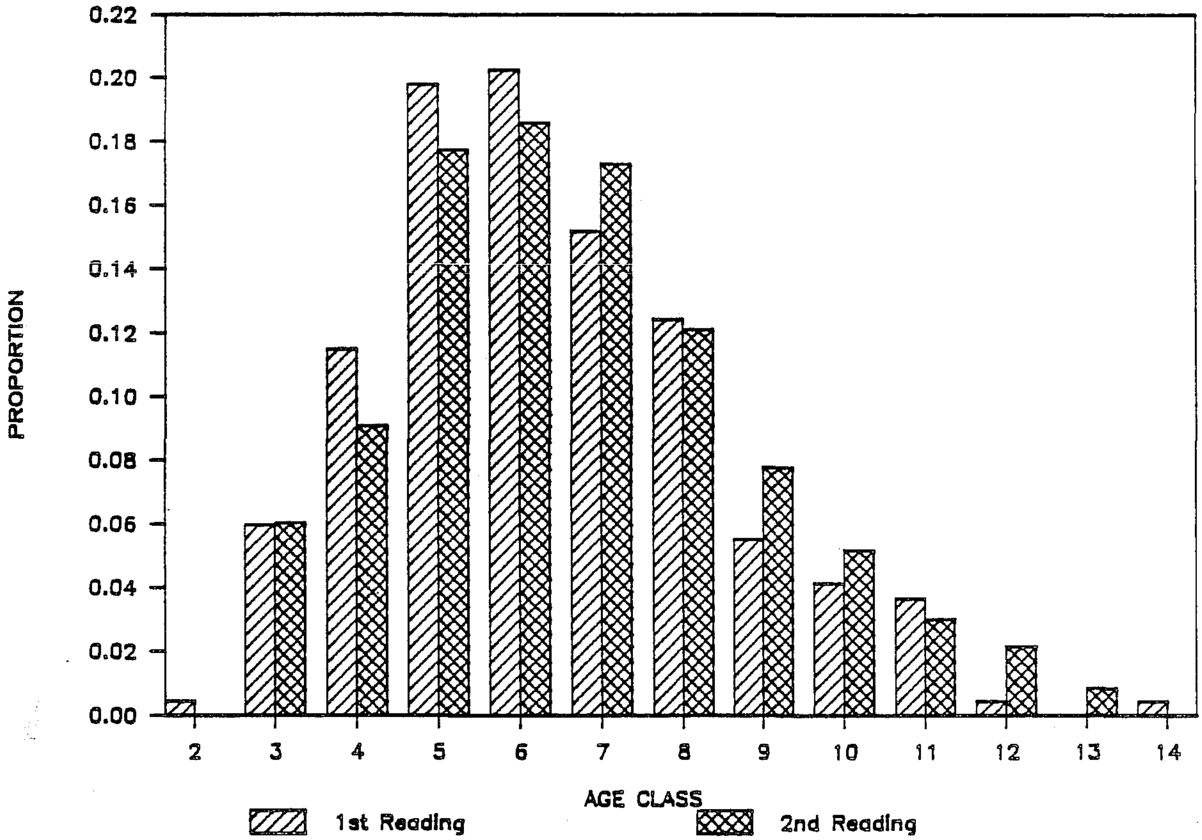

Figure 4. Proportion by age class as determined through repetitive ageings of Pacific herring scales by reader 3.

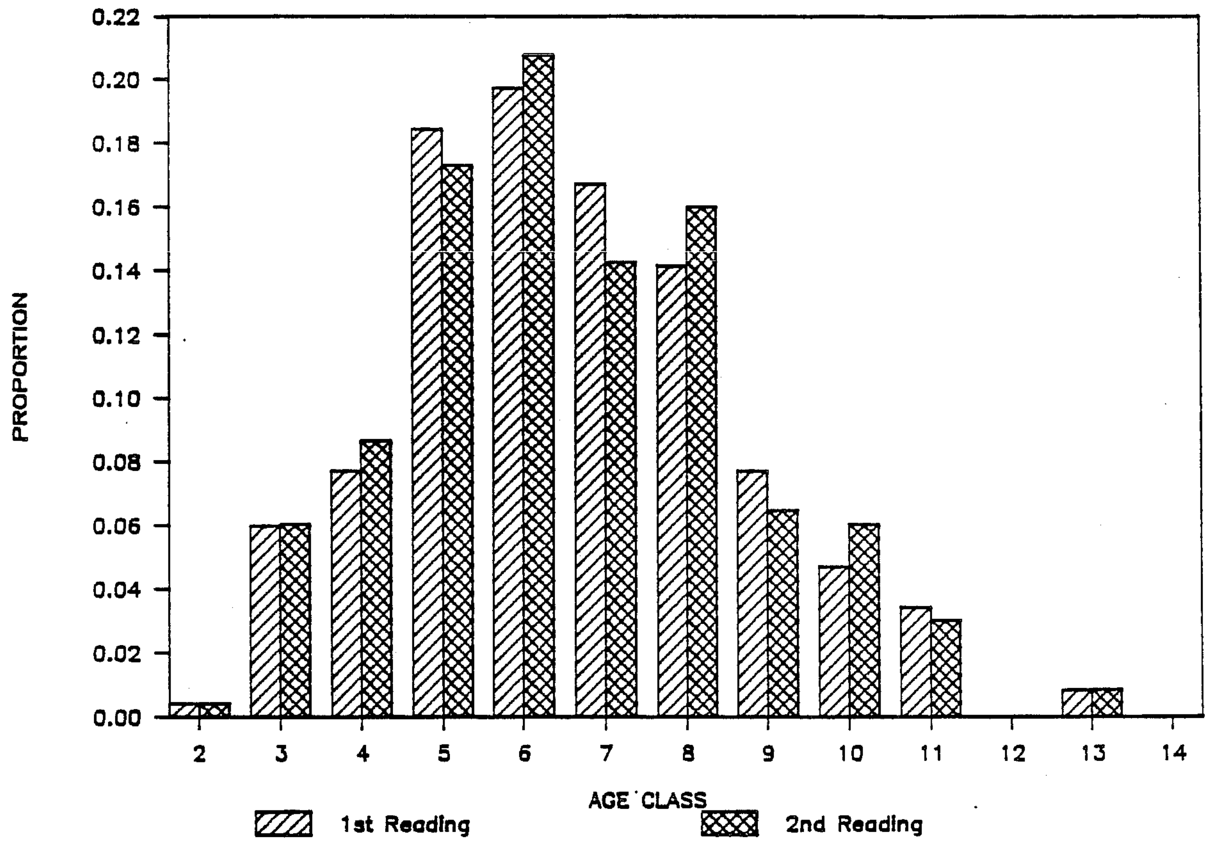

Figure 5. Proportion by age class as determined through repetitive ageings of Pacific herring scales by reader 4.

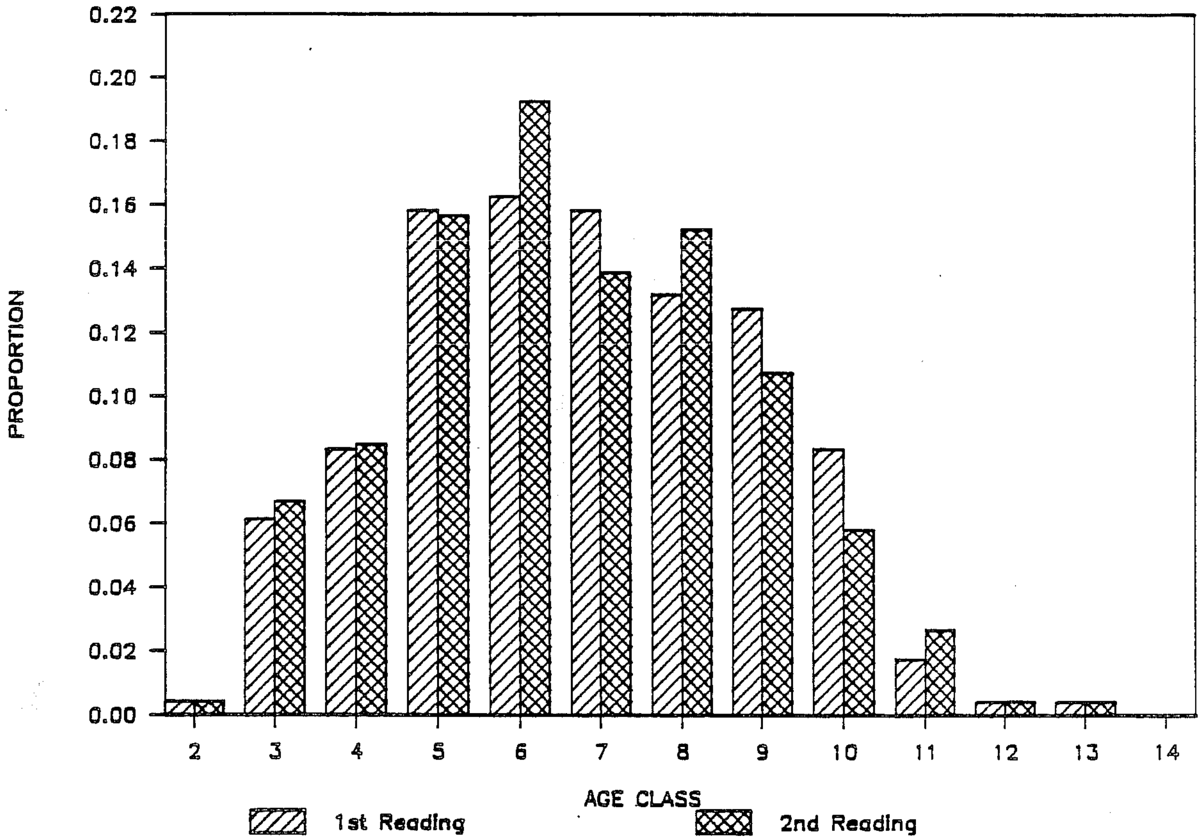

Figure 6. Proportion by age class as determined through repetitive ageings of Pacific herring scales by reader 5.

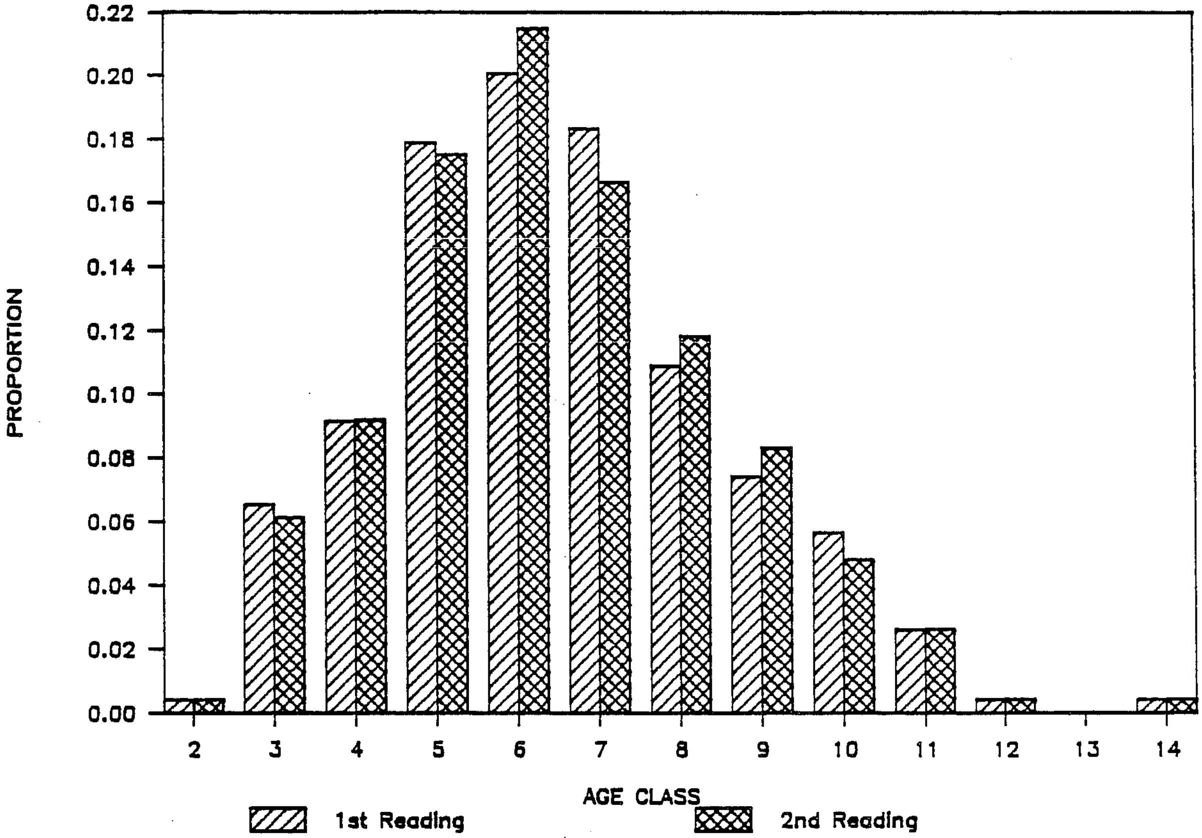

Figure 7. Proportion by age class as determined through repetitive ageings of Pacific herring scales by reader 6.

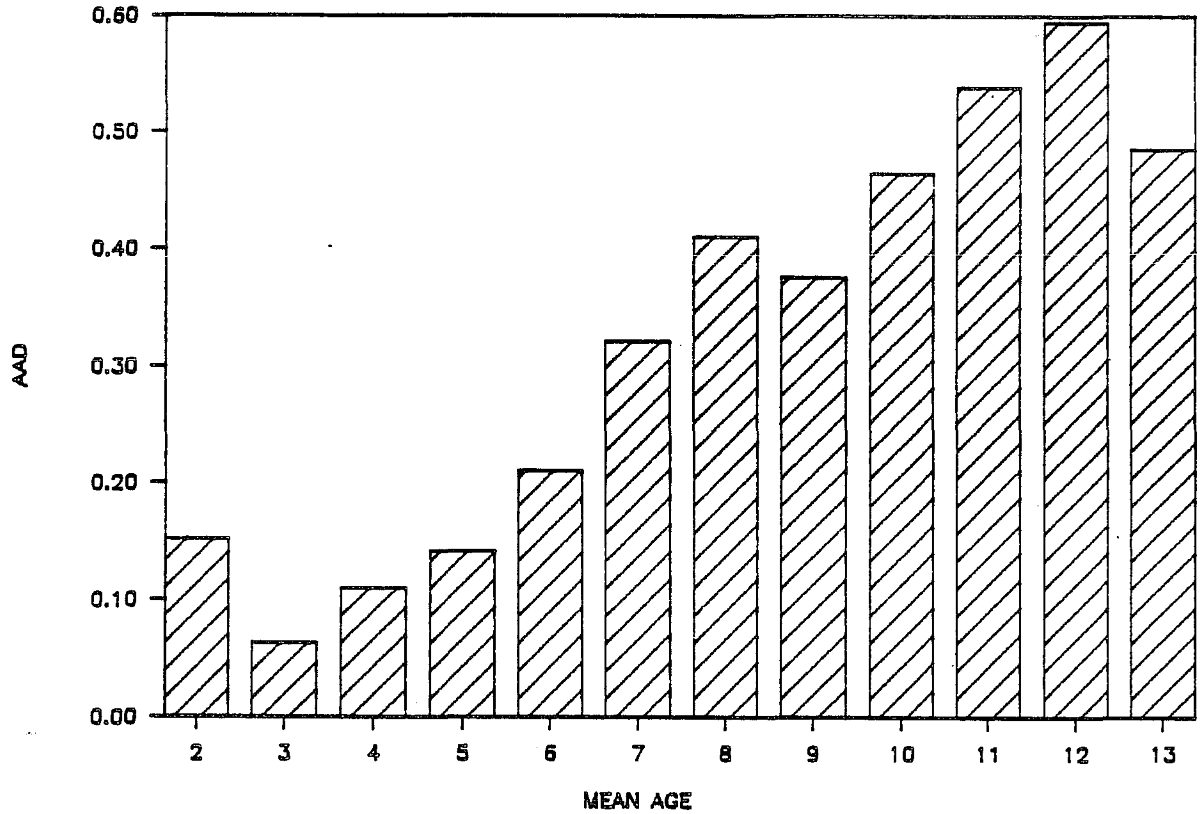

Figure 8. Average absolute difference between individual ages and their mean age from Pacific herring scales.

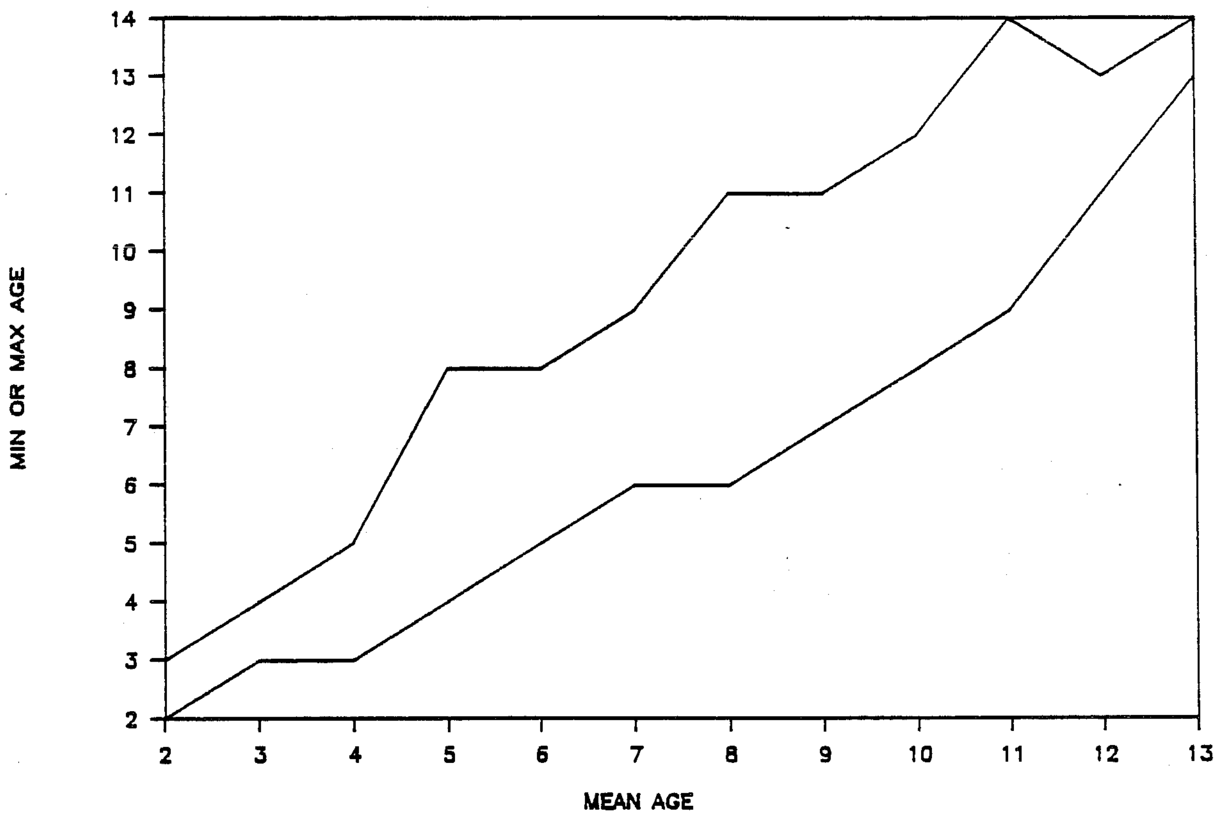

Figure 9. Minimum and maximum observed for each mean age from Pacific herring scales.

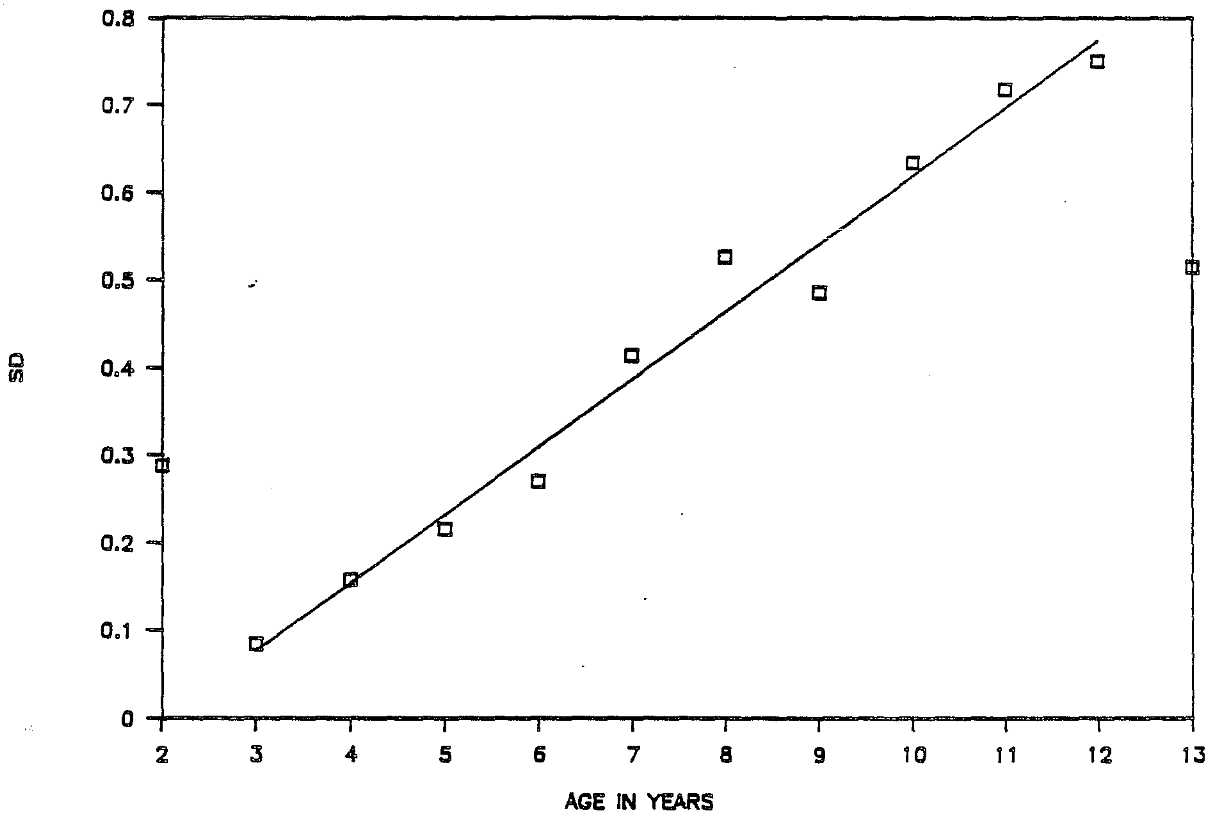

Figure 10. Standard deviation of age determinations by mean age for Pacific herring.

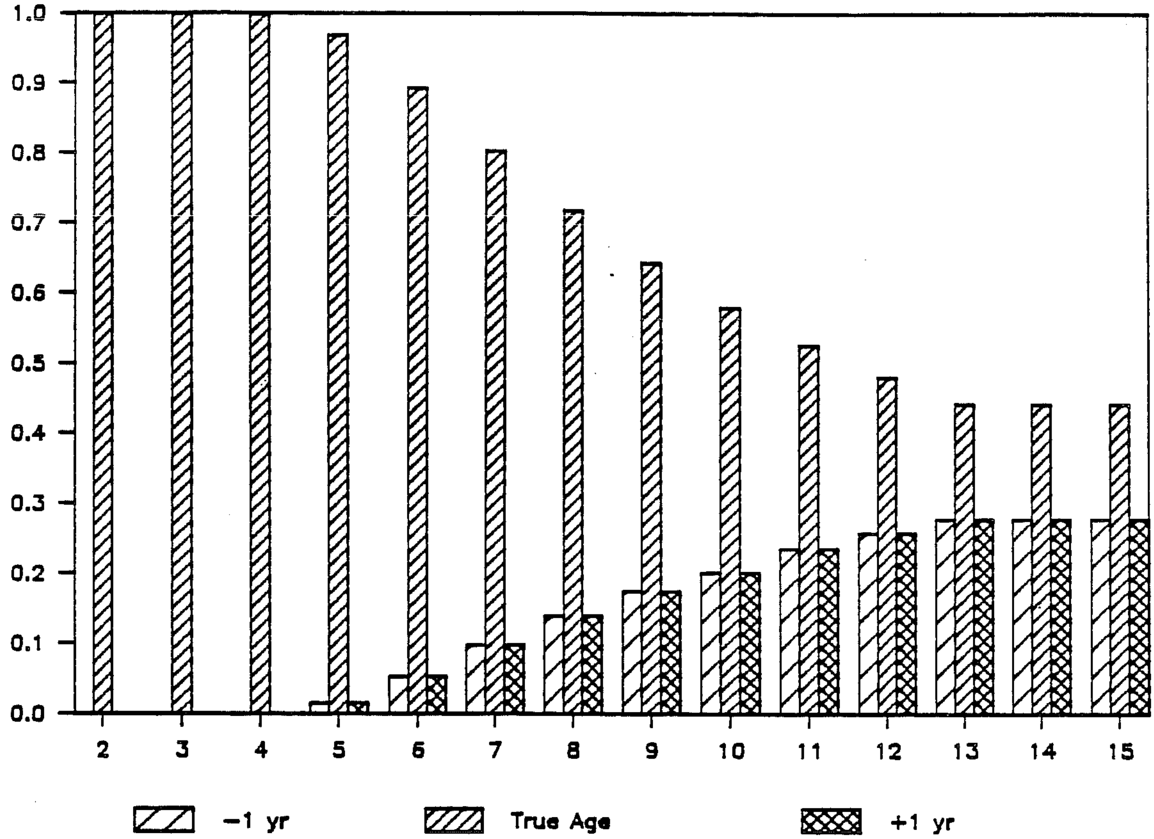

Figure 11. Probabilities by age for correctly ageing a fish, ageing the fish 1 year too young, or 1 year too old. Probabilities used to model ageing error as a multinomial distribution with three outcomes.

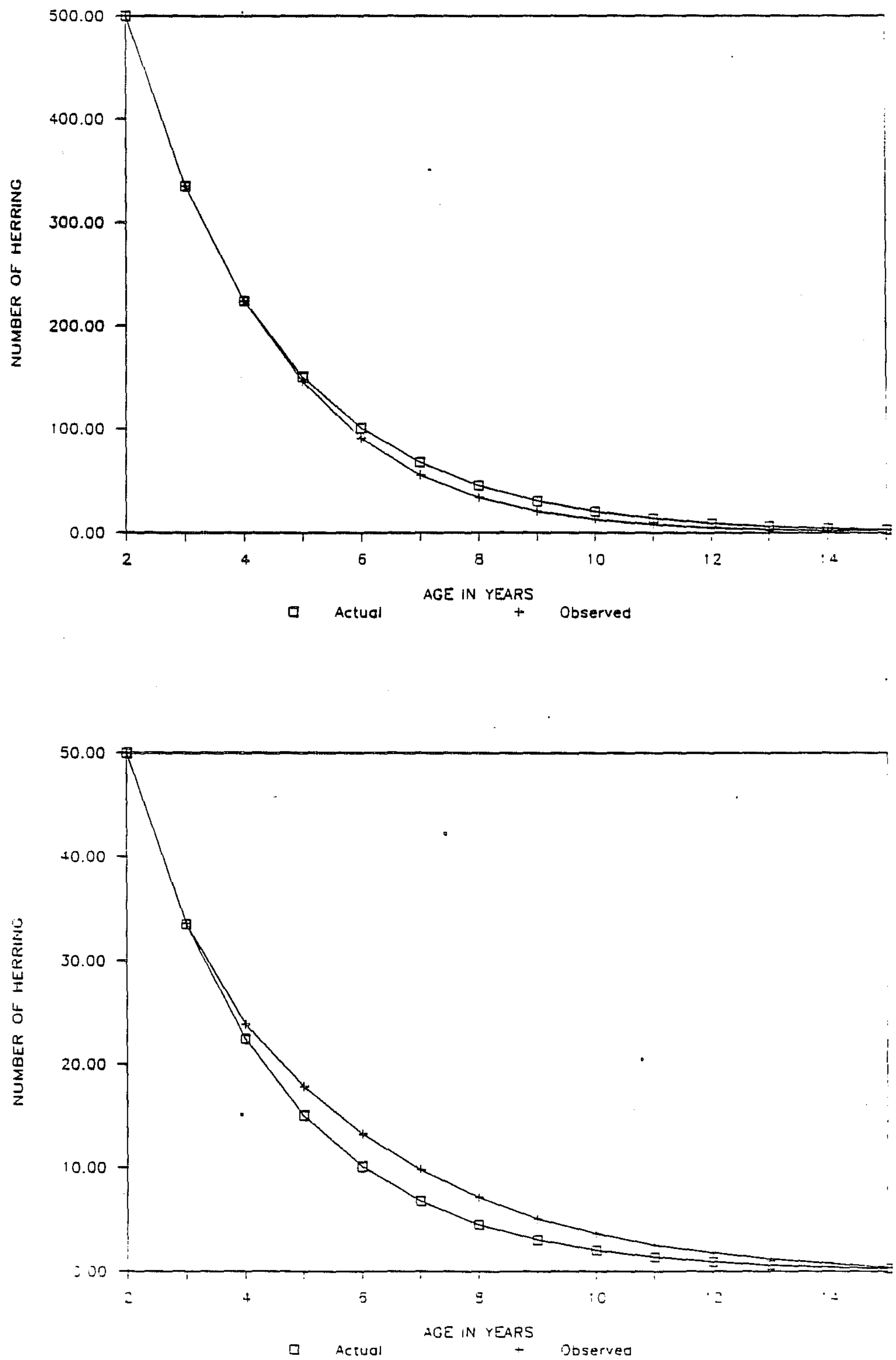

Figure 12. Abundance of herring by age for strong (top) and weak (bottom) cohorts. Observed lines as they differ from actual are the result of ageing imprecision.

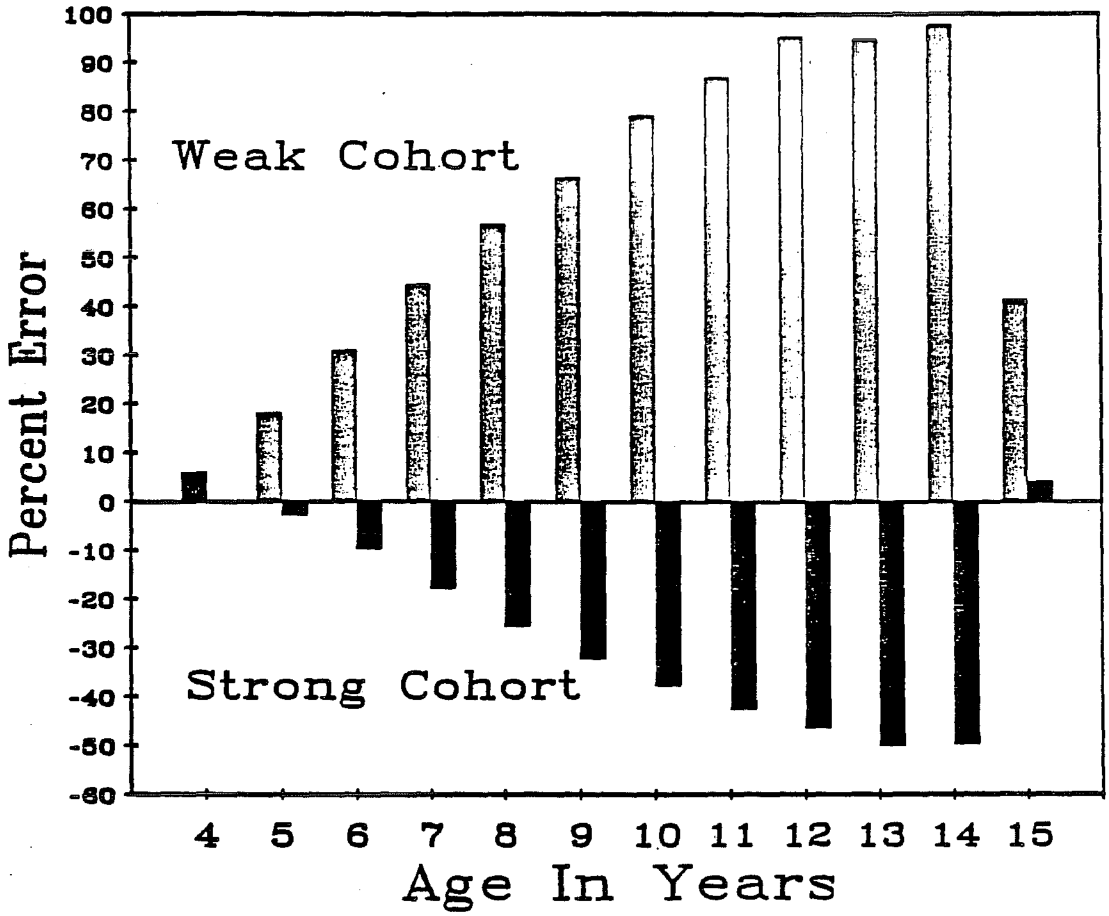

Figure 13. Average percent error of observed from actual for strong and weak cohorts.

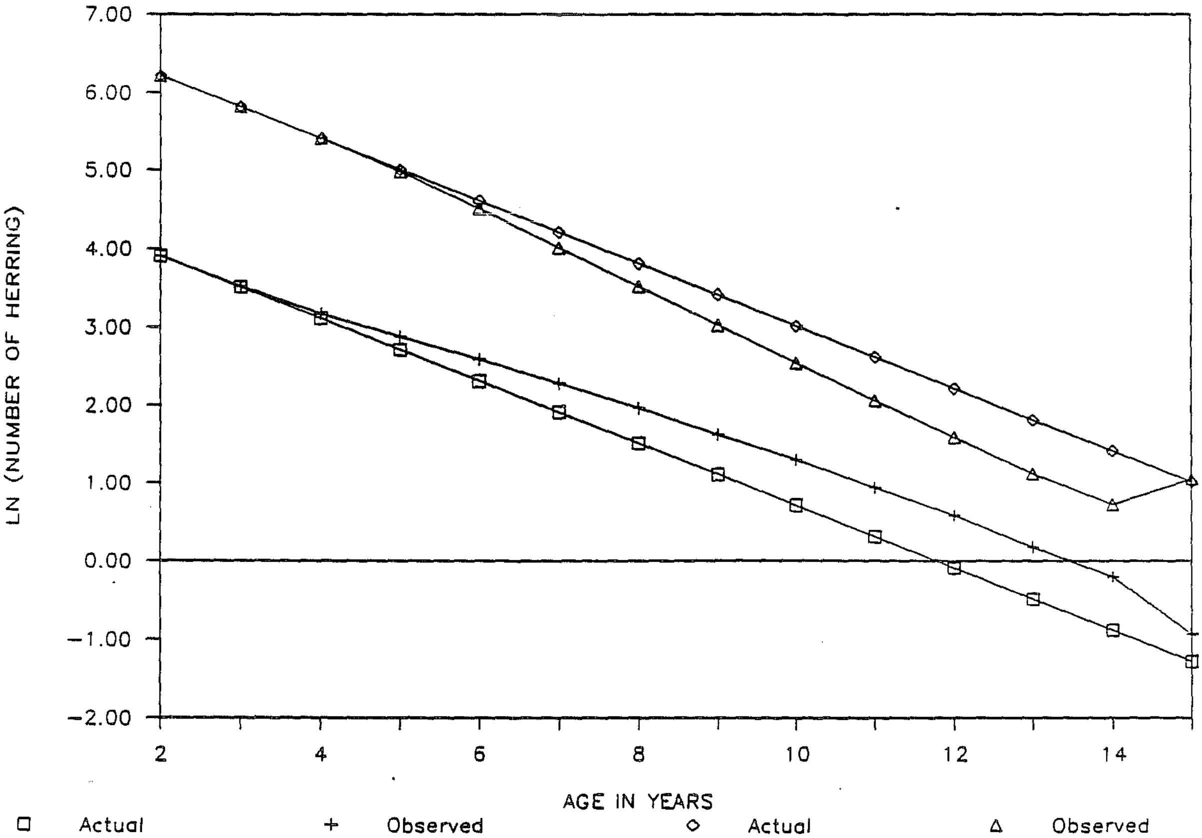

Figure 14. Abundance of herring by age for strong and weak cohorts. Observed lines as they differ from actual are the result of ageing imprecision.

Appendix A.1. Date and time spent ageing 240 herring scales for each reader and repetition.

Reader	Repetition	Date Read	Time Spent	Experience (Years)
1	1	6-15-88	2 h 44 min	3
	2	6-28-88	2 h 3 min	
2	1	6-16 & 6-17-88	2 h 41 min	5
	2	7-2-88	1 h 42 min	
3	1	6-20-88	2 h	1
	2	8-5-88	2 h	
4	1	6-22-88	52 min	1
	2	8-16 & 8-27-88	1 h 1 min	
5	1	7-13-88	1 h 30 min	5
	2	8-10-88	1 h 50 min	
6	1	7-29-88	1 h	10
	2	8-10-88	1 h	

Appendix A.2. Percent by age class of 240 herring scales as read by six readers for two repetitions. Only ageable scales were used.

Age (Years)	Percent by Age Class 1		2		3		4		5		6	
	1	2	1	2	1	2	1	2	1	2	1	2
2	0.4	0.9	0.5	0.5	0.5	0.0	0.4	0.4	0.4	0.4	0.4	0.4
3	6.2	6.0	5.9	5.6	6.0	6.1	6.0	6.1	6.2	6.7	6.6	6.1
4	8.4	9.1	8.8	8.9	11.5	9.1	7.7	8.7	8.4	8.5	9.2	9.2
5	17.7	16.8	15.7	16.4	19.8	17.7	18.5	17.3	15.9	15.7	17.9	17.5
6	20.4	20.3	20.6	20.7	20.3	18.6	19.7	20.8	16.3	19.3	20.1	21.5
7	14.6	15.9	17.6	16.4	15.2	17.3	16.7	14.3	15.9	13.9	18.3	16.7
8	12.8	12.5	15.7	14.6	12.4	12.1	14.2	16.0	13.2	15.2	10.9	11.8
9	8.0	6.9	6.9	7.0	5.5	7.8	7.7	6.5	12.8	10.8	7.4	8.3
10	5.3	5.2	3.9	4.2	4.1	5.2	4.7	6.1	8.4	5.8	5.7	4.8
11	2.2	2.6	2.5	3.3	3.7	3.0	3.4	3.0	1.8	2.7	2.6	2.6
12	3.1	3.4	1.5	1.4	0.5	2.2	0.0	0.0	0.4	0.4	0.4	0.4
13	0.0	0.0	0.0	0.5	0.0	0.9	0.9	0.9	0.4	0.4	0.0	0.0
14	0.9	0.4	0.5	0.5	0.5	0.0	0.0	0.0	0.0	0.0	0.4	0.4
Sample Size	226	232	204	213	217	231	233	231	227	223	229	228

Appendix A.3. Average percent error, average absolute difference and the minimum and maximum designated age for Pacific herring scales by age.

Mean Age	APE	AAD	Min.	Max.	SD	Estimated[a] SD
2	6.3	0.153	2	3	0.289	0.000
3	2.0	0.063	3	4	0.085	0.078
4	2.6	0.110	3	5	0.158	0.155
5	2.7	0.142	4	8	0.216	0.233
6	3.5	0.211	5	8	0.270	0.310
7	4.6	0.322	6	9	0.415	0.387
8	5.2	0.411	6	11	0.527	0.465
9	4.2	0.376	7	11	0.486	0.542
10	4.7	0.465	8	12	0.634	0.620
11	4.9	0.539	9	14	0.719	0.697
12	4.9	0.595	11	13	0.751	0.774
13	3.6	0.486	13	14	0.515	0.852
14						0.929
15						1.007

[a] Estimated from the linear relationship between age and the standard deviation in the adjacent column.