Bias in Using an Age–Length Key to Estimate Age-Frequency Distributions

S. J. WESTRHEIM AND W. E. RICKER

Department of Fisheries and the Environment, Fisheries and Marine Service, Pacific Biological Station, Nanaimo, B.C. V9R 5K6

WESTRHEIM, S. J., AND W. E. RICKER. 1978. Bias in using an age–length key to estimate age-frequency distributions. J. Fish. Res. Board Can. 35: 184–189.

Consider two representative samples of fish taken in different years from the same fish population, this being a population in which year-class strength varies. For the "parental" sample the length and age of the fish are determined and are used to construct an "age–length key," the fractions of the fish in each (short) length interval that are of each age. For the "filial" sample only the length is measured, and the parental age–length key is used to compute the corresponding age distribution. Trials show that the age–length key will reproduce the age-frequency distribution of the filial sample without systematic bias only if there is no overlap in length between successive ages. Where there is much overlap, the age–length key will compute from the filial length-frequency distribution approximately the parental age distribution. Additional bias arises if the rate of growth if a year-class is affected by its abundance, or if the survival rate in the population changes. The length of the fish present in any given part of a population's range can vary with environmental factors such as depth of the water; nevertheless, a sample taken in any part of that range can be used to compute age from the length distribution of a sample taken at the same time in any other part of the range, without systematic bias. But this of course is not likely to be true of samples taken from different populations of the species.

Key words: age–length key, bias, Pacific ocean perch, *Sebastes alutus*

WESTRHEIM, S. J., AND W. E. RICKER. 1978. Bias in using an age–length key to estimate age-frequency distributions. J. Fish. Res. Board Can. 35: 184–189.

Les auteurs ont prélevé deux échantillons de poissons, à des années différentes, dans une même population dont la répartition annuelle des âges varie. Ils ont déterminé l'âge et la taille des poissons de l'échantillon "parental" et établi à partir de là une table de l'âge en fonction de la longueur indiquant la proportion de poissons de chaque âge dans chaque court intervalle de longueurs. Pour établir la pyramide des âges de l'échantillon "filial", il a suffi de mesurer la taille des poissons et de se référer à la table précédente. Des essais ont montré que cette méthode ne produit pas de distorsion dans la répartition des âges de l'échantillon filial à condition que les longueurs correspondant à chaque âge ne se chevauchent pas. Si le chevauchement est considérable, la répartition des âges de l'échantillon "filial" sera virtuellement identique à celle de l'échantillon parental. Une distorsion supplémentaire survient quand le rythme de croissance d'une classe d'âges est modifié par la densité de la population ou que le taux de survie change. Dans l'aire de distribution d'une population, des facteurs du milieu, comme la profondeur de l'eau, peuvent modifier la taille des poissons; il est cependant possible, à partir d'un échantillon pris n'importe où dans cette aire, de déterminer, d'après la répartition des tailles, celle des âges d'un échantillon, prélevé simultanément ailleurs dans cette aire, sans qu'il y ait de distorsion. Cela ne vaut évidemment pas pour différentes populations de la même espèce.

Received May 10, 1977
Accepted October 3, 1977

Reçu le 10 mai 1977
Accepté le 3 octobre 1977

"AGEING" fish from scales, otoliths, or other structures is a tedious and time-consuming occupation. Many investigators have lightened their work by using an age–length key obtained from one or more samples in which both age and length were determined. For a series of rather narrow length intervals the age composition is computed as a fraction or percentage of the total fish of that length. These age compositions are then applied to the length-frequency distributions of samples collected at other times or in other areas, and age distributions are obtained by summation. We will call the age–length matrix used to construct the key the *parental* sample, and any length distribution to which it is applied will be the *filial* sample. It is understood that all samples are taken from the same interbreeding population in which, however, fish of different sizes may be concentrated in different parts of the range.

Printed in Canada (J4822)
Imprimé au Canada (J4822)

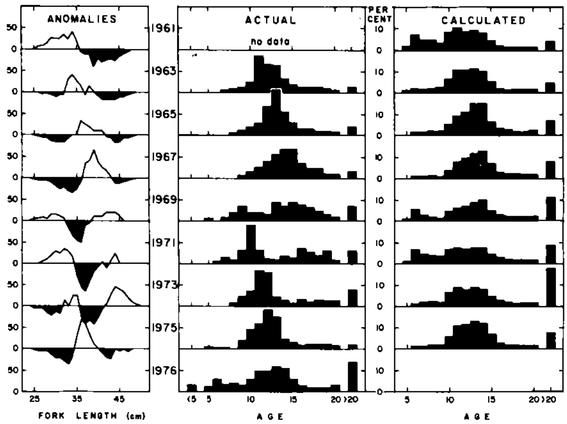

FIG. 1. Annual length-frequency anomolies, and actual and calculated age-frequencies for trawl-caught Pacific ocean perch landed in British Columbia from Queen Charlotte Sound (Goose Island Gully), 1961–75. For brevity, only odd-numbered years are shown.

It is probably generally recognized that two phenomena can affect the accuracy of an age distribution computed from an age–length key. The first is a change in survival rate between the parental and filial samples. If survival rate has decreased, then the key will compute too many fish of the older ages. However, after any change in survival rate in a multiaged species several years must elapse before its effect is fully felt among the older ages; hence, generally speaking, this has probably not been a major source of error. A second disturbing factor can be an inverse relation between year-class strength and growth rate. This obviously will change the representation of a particular age at a given length. However, in many populations, including that shown in Fig. 1, there is little or no relation between the abundance of a year-class and its rate of growth.

There is also a third source of bias which has had little general recognition, but whose effects can be extremely serious. Kimura (1977) alluded to it recently when he observed that "the age–length key will give biased results if applied to a population where the age composition differs from that of the population from which the age–length key was drawn."

In this article we consider bias that is a result of variable year-class strengths in a population, in combination with overlap in the length distribution of successive age-groups. It appears when an age–length key is prepared from one year's data and is then applied to length distributions obtained from samples of the same population in other years. This bias came to the attention of the senior author in the course of reviewing various unpublished reports dealing with assessment of stocks of Pacific ocean perch (*Sebastes alutus*), sablefish (*Anoplopoma fimbria*), and walleye pollock (*Theragra chalcogramma*) in the eastern North Pacific Ocean. Age distributions obtained using an age–length key turned out to be remarkably similar, in spite of large differences between the length-frequency distributions to which the keys were applied in successive years.

Example of Bias from Pacific Ocean Perch in Queen Charlotte Sound

The age composition of Pacific ocean perch landed by Canadian trawlers each year (Fig. 1) was obtained directly by reading all the otoliths collected. The length-

TABLE 1. Age–length matrix for a sample of 3017 Pacific ocean perch collected in Queen Charlotte Sound (Goose Island Gully) during September 1976.

Fork length (cm)	Age																		Total
	<5	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	>20	
<22	29	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	29
22	17	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	17
24	18	9	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	28
26	17	17	5	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	40
28	—	29	55	11	1	—	—	—	—	—	—	—	—	—	—	—	—	—	96
30	—	3	104	45	4	—	1	2	—	—	—	—	—	—	—	—	—	—	159
32	—	—	24	32	42	10	18	7	6	3	1	1	—	—	—	—	—	—	144
34	—	—	1	8	25	45	78	57	28	25	9	1	—	—	—	—	—	—	277
36	—	—	—	1	6	10	73	155	129	109	77	28	3	—	—	—	—	—	591
38	—	—	—	—	—	—	13	74	118	157	141	50	20	9	7	1	—	—	590
40	—	—	—	—	—	—	—	5	27	50	82	60	29	20	12	14	20	23	342
42	—	—	—	—	—	—	—	1	2	6	21	39	21	20	22	26	20	82	260
44	—	—	—	—	—	—	—	—	1	—	9	7	12	13	22	23	17	120	224
46	—	—	—	—	—	—	—	—	—	—	—	—	—	1	3	5	19	139	167
48	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	3	46	50
50	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	3	3
Total	81	58	190	98	78	65	183	301	311	350	340	186	85	63	66	70	79	413	3,017

TABLE 2. Age–length matrix of a population in which year-class strength varies in the ratio 2:1, and rate of increase in length decreases with age. (For simplicity, there is no growth between ages 8 and 9, or between ages 10 and 11.)

Age	4	5	6	7	8	9	10	11	
Strength	2	1	1	2	1	1	2	1	Total
				A. *Parental distribution*					
Length (cm)									
	6								6
40	61								61
	242								242
50	383								383
	242	2							244
60	61	23							84
	6	91							97
70		144	2						146
		91	17						108
80		23	68	3					94
		2	108	26	1				137
90			68	102	10	5			185
			17	162	38	19	3		239
100			2	102	61	30	11	1	207
				26	38	19	18	1	102
110				3	10	5	11	1	30
					1	0	3	0	4
Total	1001	376	282	424	159	78	46	3	2,369
%	42.2	15.9	11.9	17.9	6.7	3.3	1.9	0.1	
				B. *Age–length key* (%)					
	100								
40	100								
	100								
50	100								
	99	1							
60	73	27							
	6	94							
70		99	1						
		84	16						
80		24	72	3					
		1	79	19	1				
90			37	55	5	3			
			7	68	16	8	1		
100			1	49	29	15	5	1	
				25	37	19	18	1	
110				10	33	17	37	3	
					25	0	75	0	

frequency anomalies and the actual age compositions show the progression through the fishery of two groups of strong year-classes — 1951–54 and 1960–62.

In September 1976 a sample of 3017 fish was obtained by a research vessel from the same area. An age–length matrix from these fish (Table 1) was used to compute age frequencies from the lengths of earlier years. Age compositions based on this key, shown in the last column of Fig. 1, fail to reveal the progression of strong year-classes. On the other hand, a prominent mode at age 6 in the 1976 sample was reproduced in 6 of the 8 computed age distributions, although it did not occur at all in the actual commercial landings. Similarly a mode near age 13 tended to appear in the filial distributions whether or not it actually existed in the landings of the year in question.

Model of the Interaction of Variable Year-Class Strength with Size Overlap

The age and length distributions of Fig. 1 are subject to sampling variability, and some might regard them as less than completely convincing on that account. Consequently it is desirable to consider a model population in which all accidental sources of error are eliminated. The upper half of Table 2 shows a completely representative sample of a population in which both survival rate and rate of growth remain unchanged from

TABLE 3. A, age–length matrix like that of Table 2, except that the year-class strengths are shifted 1 yr to the right; B, age–length matrix computed from the total column of Table 3A and from the age–length key in Table 2B.

Age	4	5	6	7	8	9	10	11	
Strength	1	2	1	1	2	1	1	2	Total
				A. *Actual filial distribution*					
Length (cm)									
	3								3
40	30								30
	121								121
50	192								192
	121	4							125
60	30	46							76
	3	182							185
70		287	2						289
		182	17						199
80		46	68	1					115
		4	108	13	2				127
90			68	51	19	5			143
			17	81	77	19	1		195
100			2	51	121	30	6	1	211
				13	77	19	9	2	120
110				1	19	5	6	1	32
					2	0	1	0	3
Total	500	751	282	211	317	78	23	4	2,166
%	23.1	34.7	13.0	9.7	14.6	3.6	1.1	0.2	
				B. *Computed filial distribution*					
	3								3
40	30								30
	121								121
50	192								192
	124	1							125
60	55	21							76
	11	174							185
70		286	3						289
		167	32						199
80		28	83	3					114
		1	100	24	1				126
90			53	79	7	4			143
			14	133	31	16	2		196
100			2	103	61	32	11	2	211
				30	44	23	22	1	120
110				3	11	5	12	1	32
					1	0	2	0	3
Total	536	678	287	375	156	80	49	4	2,165
%	24.8	31.3	13.3	17.3	7.2	3.7	2.3	0.2	

year to year, the fish are always accurately aged, and there is no sampling error. In Table 3 the age–length distribution in the top half differs from that of Table 2 only in the size of the successive age-groups: the year-classes of Table 2 are moved 1 yr to the right, simulating the transition from one year to the next. In the lower half of Table 3 the age–length key of Table 2 is applied to the total length distribution in the last column of the frequency distribution above, and in the last line the computed filial frequencies are summed by age.

The age distribution recovered obviously differs greatly from the true age distribution of the sample immediately above it (Fig. 2) and it has considerable resemblance to that of the parental sample in Table 2. From this and other examples, as well as the illustration of Fig. 1, it is clear that:

1) when there is little overlap in size between ages, the computed age distribution will closely resemble that of the filial sample;

2) when there is much overlap in size between ages, the computed age distribution will closely resemble that of the parental sample.

Unbiased Situations

Still maintaining the condition of no secular change in age-specific mortality or growth rates, mention can

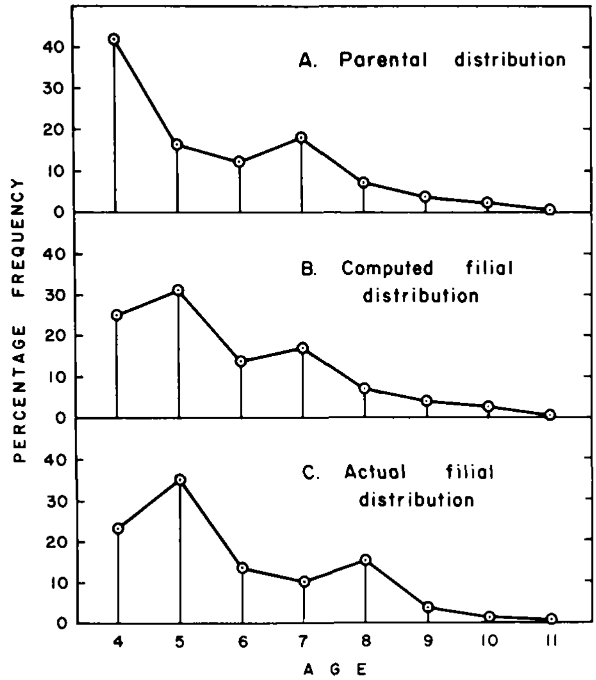

FIG. 2. Age-frequencies for the parental population and the actual and calculated filial populations (from Tables 2 and 3).

be made of three situations in which there is no systematic bias in using an age–length key. Two are rather trivial, but the third is of practical importance.

Firstly, if there is no overlap in size between successive ages in a population, then obviously an age–length key for one year will be applicable to all other years. A wrong age cannot be assigned.

Secondly, if year-classes are all of identical size at recruitment, and consequently their relative numbers at different ages do not change from year to year, there is really only one age–length matrix. The key that is obtained from any year's sample will be the same as that obtained from any other year's sample, apart from random variability.

Finally, an age–length key can be applied, without introducing systematic bias, to any sample taken *from the same population* at the same time, regardless of the relative numbers of successive length-groups present in the two samples. For example, different fishing gears usually have different selection characteristics, and fish of different sizes often occur predominantly at different depths; for either reason, one sample may differ considerably in length- and age-frequency distribution from another. Nevertheless, an age–length key computed from one such sample can safely be applied to the lengths from any other taken from the same population in the same year.

Because this is not intuitively obvious, an example is given in Table 4, which shows a sample taken from the population of Table 2 using a gear whose relative catching power decreases rapidly with decreasing fish length, as shown in the second column. By applying the age–length key of Table 2 to the totals column of Table 4, exactly the same age–length matrix is obtained as is shown in Table 4, except for small differences that arise from rounding all entries to the nearest unit.

In practice, random variability makes for considerable differences between age–length matrices even when taken from the same population in the same year and

TABLE 4. Age–length matrix derived from Table 2, by multiplying the entries in each row of 2 by the "efficiency" fraction shown in column two.

Length (cm)	Efficiency	Age 4	5	6	7	8	9	10	11	Total
	0.	0								0
40	0.	0								0
	0.05	12								12
50	0.1	38								38
	0.15	36								36
60	0.2	12	5							17
	0.3	2	27							29
70	0.4		58	1						59
	0.5		46	8						54
80	0.6		14	41	2					57
	0.7		1	76	18	1				96
90	0.8			54	82	8	4			148
	0.85			14	138	32	16	3		203
100	0.9			2	92	55	27	10	1	187
	0.95				25	36	18	17	1	97
110	1.				3	10	5	11	1	30
	1.					1	0	3	0	4
Total		100	151	196	360	143	70	44	3	1,067

by the same gear. This aspect of the use of age–length keys is considered by Kimura (1977), but for orientation a simple rule is that any entry in the original matrix can be treated as a Poisson variable having a standard error approximately equal to its own square root for numbers of moderate or large size, and somewhat larger for small numbers. For greater accuracy, the appropriate confidence limits are given in Appendix 2 of Ricker (1975).

Conclusions

Much overlap in size among age-groups, and moderate or large variations in year-class strength, are the rule among commercial fishes of the size range customarily captured. For such populations age–length keys computed from one year's data cannot be applied to samples taken in a different year. However, an age–length key *can* be applied to other samples taken at more or less the same time in the same year, regardless of any differences in the fraction of each length-group that is present in different samples, always provided that all the samples come from the same population.

KIMURA, D. K. 1977. Statistical assessment of the age–length key. J. Fish. Res. Board Can. 34: 317–324.

RICKER, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191: 1–382.

This article has been cited by:

1. S. Dale Hanson, Craig P. Stafford. 2017. Modeling Otolith Weight using Fish Age and Length: Applications to Age Determination. *Transactions of the American Fisheries Society* **146**:4, 778-790. [Crossref]
2. . References 283-304. [Crossref]
3. Sondre Aanes, Jon Helge Vølstad. 2015. Efficient statistical estimators and sampling strategies for estimating the age composition of fish. *Canadian Journal of Fisheries and Aquatic Sciences* **72**:6, 938-953. [Abstract] [Full Text] [PDF] [PDF Plus]
4. Douglas G. Chapman. Fisheries Research, Statistics in . [Crossref]
5. Hiroshi Yamaguchi, Yoko Goto, Noboru Hoshino, Kazushi Miyashita. 2014. Growth and age composition of northern shrimp Pandalus eous estimated by multiple length frequency analysis. *Fisheries Science* **80**:4, 665-678. [Crossref]
6. Lewis G. Coggins, Daniel C. Gwinn, Micheal S. Allen. 2013. Evaluation of Age–Length Key Sample Sizes Required to Estimate Fish Total Mortality and Growth. *Transactions of the American Fisheries Society* **142**:3, 832-840. [Crossref]
7. John Cotter, Benoit Mesnil, Peter Witthames, Matt Parker-Humphreys. 2009. Notes on nine biological indicators estimable from trawl surveys with an illustrative assessment for North Sea cod. *Aquatic Living Resources* **22**:2, 135-153. [Crossref]
8. A. G. Murta, C. Vendrell. 2009. Using the EM algorithm to age fish eggs. *ICES Journal of Marine Science* **66**:4, 607-616. [Crossref]
9. Charles L. Mesing, Richard L. Cailteux, P. Andrew Strickland, Eric A. Long, Mark W. Rogers. 2008. Stocking of Advanced-Fingerling Largemouth Bass to Supplement Year-Classes in Lake Talquin, Florida. *North American Journal of Fisheries Management* **28**:6, 1762-1774. [Crossref]
10. Debra A. Prince, Lyle W. Konigsberg. 2008. New Formulae for Estimating Age-at-Death in the Balkans Utilizing Lamendin's Dental Technique and Bayesian Analysis. *Journal of Forensic Sciences* **53**:3, 578-587. [Crossref]
11. JODY L. CALLIHAN, LYNN T. TAKATA, RYAN J. WOODLAND, DAVID H. SECOR. 2008. Cohort splitting in bluefish, Pomatomus saltatrix, in the US mid-Atlantic Bight. *Fisheries Oceanography* **17**:3, 191-205. [Crossref]
12. Faith A. Ochwada, James P. Scandol, Charles A. Gray. 2008. Predicting the age of fish using general and generalized linear models of biometric data: A case study of two estuarine finfish from New South Wales, Australia. *Fisheries Research* **90**:1-3, 187-197. [Crossref]
13. A J R Cotter, G M Pilling. 2007. Landings, logbooks and observer surveys: improving the protocols for sampling commercial fisheries. *Fish and Fisheries* **8**:2, 123-152. [Crossref]
14. Keith L. Hurley. 2007. Effects of a reduced subsampling protocol on calculation of growth and body condition indices. *Fisheries Research* **83**:2-3, 360-363. [Crossref]
15. Jessica H. Farley, Tim L.O. Davis, John S. Gunn, Naomi P. Clear, Ann L. Preece. 2007. Demographic patterns of southern bluefin tuna, Thunnus maccoyii, as inferred from direct age data. *Fisheries Research* **83**:2-3, 151-161. [Crossref]
16. Douglas G. Chapman. Fisheries Research, Statistics in . [Crossref]
17. K. Lekve, K. Enersen, S. E. Enersen, J. Gjosaeter, N. Chr. Stenseth. 2006. Interannual variability in abundance and length of young coastal cod in the subtidal zone. *Journal of Fish Biology* **68**:3, 734-746. [Crossref]

18. Nancy A. Nate, Mary T. Bremigan. 2005. Comparison of Mean Length at Age and Growth Parameters of Bluegills, Largemouth Bass, and Yellow Perch from Length-Stratified Subsamples and Samples in Michigan Lakes. *North American Journal of Fisheries Management* **25**:4, 1486-1492. [Crossref]

19. Daniel A. Isermann, Carey T. Knight. 2005. A Computer Program for Age–Length Keys Incorporating Age Assignment to Individual Fish. *North American Journal of Fisheries Management* **25**:3, 1153-1160. [Crossref]

20. David Milton, Markson Yarrao, Gary Fry, Charles Tenakanai. 2005. Response of barramundi, Lates calcarifer , populations in the Fly River, Papua New Guinea to mining, fishing and climate-related perturbation. *Marine and Freshwater Research* **56**:7, 969. [Crossref]

21. R. I. C. Chris Francis, Shelton J. Harley, Steven E. Campana, Peer Doering-Arjes. 2005. Use of otolith weight in length-mediated estimation of proportions at age. *Marine and Freshwater Research* **56**:5, 735. [Crossref]

22. V. S. Troynikov, S. G. Robertson. 2005. Estimating age composition using the Fredholm first-kind equation. *Marine and Freshwater Research* **56**:5, 745. [Crossref]

23. Phillip W. Bettoli, Leandro E. Miranda. 2001. Cautionary Note about Estimating Mean Length at Age with Subsampled Data. *North American Journal of Fisheries Management* **21**:2, 425-428. [Crossref]

24. Patrick M. Kocovsky, Robert F. Carline. 2001. Dynamics of the Unexploited Walleye Population of Pymatuning Sanctuary, Pennsylvania, 1997���1998. *North American Journal of Fisheries Management* **21**:1, 178. [Crossref]

25. Andrew J Paul, John R Post, George L Sterling, Carl Hunt. 2000. Density-dependent intercohort interactions and recruitment dynamics: models and a bull trout (Salvelinus confluentus) time series. *Canadian Journal of Fisheries and Aquatic Sciences* **57**:6, 1220-1231. [Abstract] [PDF] [PDF Plus]

26. Lyle W. Konigsberg, Susan R. Frankenberg. 1994. Paleodemography: "Not quite dead". *Evolutionary Anthropology: Issues, News, and Reviews* **3**:3, 92-105. [Crossref]

27. Lyle W. Konigsberg, Susan R. Frankenberg. 1992. Estimation of age structure in anthropological demography. *American Journal of Physical Anthropology* **89**:2, 235-256. [Crossref]

28. A. Tsangridis, N. Filippousis. 1991. Use of length-frequency data in the estimation of growth parameters of three Mediterranean fish species: bogue (Boops boops L.), picarel (Spicara smaris L.) and horse mackerel (Trachurus trachurus L.). *Fisheries Research* **12**:4, 283-297. [Crossref]

29. John M. Hoenig, Dennis M. Heisey, William D. Lawing, Dennis H. Schupp. 1987. An Indirect Rapid Methods Approach to Assessment. *Canadian Journal of Fisheries and Aquatic Sciences* **44**:S2, s324-s338. [Crossref]

30. Gary D. Grossman. 1979. Demographic characteristics of an intertidal bay goby (Lepidogobius lepidus). *Environmental Biology of Fishes* **4**:3, 207-218. [Crossref]