International Commission for the Northwest Atlantic Fisheries

Serial No. 5033
(D.c.9)

ICNAF Res.Doc. 77/VI/13

ANNUAL MEETING - JUNE 1977

Analysis of anomalies between readers and regressions of average length on age from the Ageing Workshop on Cod held at Vigo, Spain in October 1975

by

R. K. Misra and R. Wells
Fisheries and Environment Canada, Fisheries and Marine Service
Research and Resource Services
St. John's, Newfoundland, Canada A1C 1A1

Following the observation by the Assessments Subcommittee meeting at Woods Hole that differences in age compositions of some fish species as submitted by various countries to ICNAF existed, an ageing workshop was held at Vigo, Spain in October, 1975. Proceedings of the workshop were documented and a report presented at the annual meeting of ICNAF in June, 1976. Also see Summ. Doc. 76/VI/13 (revised July 1976) which was submitted to the 1976 annual meeting. A brief account of that part of their examination of cod otoliths which is pertinent to the present report is given as follows:

Two samples of cod otoliths were examined. These were designated 'Spanish" (sample 1) and 'Canadian' (sample 2). Sample 2 included photographs and projector slides of the otolith sections. Percentage agreement between pairs of readers ranged from 0 to 77% in sample 1 and from 0 to 96% in sample 2 over ICNAF Divisions considered separately. A modal age was derived for each specimen from the ages estimated by the various readers, except for those age determinations in which no clear mode was evident. Mode was considered as the best estimate of the true age of each specimen and anomalies (estimated age minus modal age) calculated. Mean anomaly for a reader indicated his bias with respect to the group as a whole. Based on anomalies, the agreement with the mode ranged from 12% to 94% for sample 1 and from 3% to 96% for sample 2 when Divisions were considered separately and to an average of about 65% for sample 1 and of about 60% for sample 2 when all Divisions were combined. From this preliminary analysis it seemed that the differences between pairs of readers were nearly as large as they could be and those between readers and model ages were severe. For otolith sample 2, age-length keys by reader were derived and graphical plots of length-at-age curves indicated that albeit the curves were generally similar, differences between readers may have occurred.

As knowledge of differences between readers is of vital importance in analytical assessment of fish stocks, further statistical analyses of anomalies and age-length keys were considered warranted. These, with the exception of the non-parametric test which was done later, were presented for discussion before the members of the ageing workshop held at St. John's, Newfoundland in February, 1977 and the general consensus then was that the report be sent for publication.

For fifteen readers who read sample 2, anomalies ranged from -5 to +7, sample sizes from 193 to 260 and mean anomalies from -0.008 for reader 14 to +1.196 for reader 16, who showed the greatest positive bias in ageing. Table 1 gives sample sizes (n), mean ($\bar{y}$) and standard errors (SE) of anomalies for each reader. Bartlett's test of homogeneity of fifteen variances (see, e.g., Snedcor and Cochran, 1967, p. 296) yielded $x^2 = 279.6$ on df (degrees of freedom) = 14, which is significant (probability level of significance, $p < 0.005$). ANOVA (analysis of variance) procedure was therefore not followed to test the equality of mean anomalies of readers. Instead, these mean values were compared in pairs following a weighted x^2 procedure described, e.g., in Armitage (1971, p. 197). Each mean anomaly, $\bar{y}$, of a pair was assigned a weight, w, which is the reciprocal of its variance and x^2 estimated as $\Sigma w\bar{y}^2 - (\Sigma w\bar{y})^2/\Sigma w$, on 1 df. In Table 2 the space for comparison of mean anomalies of a pair was marked '***' (without parentheses) if p for x^2 was ≤ 0.001, '**' if $0.001 < p \leq 0.01$, '*' if $0.01 < p \leq 0.05$ and 'ns' (not significant) if $p > 0.05$. Table 2 shows that 80 out of a possible 105 differences in mean anomalies were significant. At this point skewness of the distribution of anomalies for each reader was tested, by analyzing coefficient of skewness as explained, e.g., in Snedcor and Cochran (1967, p. 86). Coefficients of skewness were tested for significance by t-test (on df = ∞). Table 3 gives values of p. Only the distributions for readers 2 and 5 did not show any evidence of asymmetry. All other

distributions were 'highly' skew. Skewness was +ve for readers 1, 13, 14 16 and 17 and -ve for readers 3, 7, 8, 9, 11, 12 and 15. It was therefore suspected that x^2 tests for comparing mean anomalies, described earlier, may have been unduly sensitive to these more than 'moderate' departures from normality. Paired comparisons of anomalies of readers were therefore also done by a nonparametric test. Kruskel-Wallis test (see, e.g., Sokal and Rohlf, 1969, p. 388) was employed. This yielded a x^2 value, on 1 df, for each pair. Significant and insignificant differences were marked in the same manner as for the weighted x^2 test and these marks are shown in parentheses in Table 2. Again, a large number (viz., 79) of pairs differed significantly. It is interesting to observe that the agreement between the weighted x^2 and the nonparametric tests was good, as these yielded the same results for 98 (out of 105) pairs.

Age-length keys

From a management point of view, inconsistency in age-length keys by readers should perhaps be of greater concern than inconsistency in their anomalies. Age-length keys by reader, given in Table 5 of the June, 1976 report of ICNAF, referred earlier, were examined for consistency among readers. Graphic plots of average length-at-age by reader indicated that each distribution was remarkably linear in the range 2 to 10 years (there were no data for age 1). As there were only few observations, viz., 0.01% to 2.4% (exceptions: 7.4% for reader 16 and 3.4% for reader 17) beyond age 10 and as a straight line is a unique mathematical function in that it is easier to analyze it efficiently and to interpret it with minimal ambiguity, it was decided to analyze data in the range 2 to 10 years only. The analysis may be described as follows:

Regression of average length on age was analyzed following the ANOVA outline. For each reader MS (mean square) for deviations from the linear regression was compared with pooled 'within' reader MS (on df = 14586). All MS values were significant ($p < 0.01$ for reader 4 and < 0.001 for all other readers). MS for 'slope' of each linear regression was therefore compared with the MS for deviations from the linear regression. All slopes were significant ($p < 0.001$). As deviations from the linear regression were significant, presence of quadratic regression was also tested. Second degree terms were not significant for 6 (out of 15) sets of data and significant, but only at 5% probability level, for the other nine. As the number of age groups was small (10 readers had no data at age 2), search for the presence of still higher degree terms would not have been meaningful. Straight lines were therefore accepted as the best approximations of the relation of length and age. Table 4 gives the 'intercept' and 'slope' constants a and b, respectively, of the equation, length = a + b (age) of linear regression of length on age, for each reader.

Paired comparisons of the slope constants of readers were done. Significant and insignificant differences were marked in the same manner as in Table 2 for the earlier tests, and are shown in Table 5. Only 17 of 105 possible differences were significant. Furthermore, 15 of these 17 differences were those of readers 16 and 17 with other readers; the other 2 being those of reader 9 with readers 1 and 15. The analysis thus clearly indicated that (a) the agreements among all readers except readers 16 and 17 have been remarkably good and (b) readers 16 and 17 differ from most of the other readers.

Regression of average length on age for the age-length key derived from modal ages of otolith sample 2 yielded similar results, viz., the linear regression was significant ($p < 0.001$) and the second degree terms were not significant ($p > 0.05$). Again, readers 16 and 17 differed in their slope constants from the modal slope but other readers did not, with the exception of reader 15, who did.

It would thus seem that differences between readers may lead to more severe differences in age composition than in average lengths-at-age.

Summary

As knowledge of differences between readers is of vital importance in analytical assessment of fish stocks, further statistical analyses of anomalies and age-length keys were considered warranted. This report presents and discusses analyses of anomalies (estimated ages - modal ages) and regressions of average length on age for 15 readers of otolith sample 2 read at Vigo, Spain in October 1975. Weighted x^2 was done for paired comparisons of mean anomalies of 15 readers. Eighty, out of a possible 105, differences in mean anomalies were significant. As 13, out of 15, distributions of anomalies were 'highly' skew, it was decided to do paired comparisons by a non-parametric method as well. Kruskal-Wallis test was employed. The results were essentially the same, viz., 79 (out of 105) pairs differed in mean anomalies and 98 (out of 105) pairs yielded the same results in the two tests. Severe differences among readers in their mean anomalies were thus indicated.

Analysis of variance of regression of average lengths on age indicated that straight lines were the best approximations of the relation of length and age for each reader as well as for length and modal ages. When slope constants of 15 readers were compared with each other, in pairs, only 17 of the 105 possible differences were significant. Of these 15 were of readers 16 and 17 with other readers, and 2 were of reader 9 with readers 1 and 15. Readers 15, 16 and 17 differed in their slope constants from the modal slope. Thus, it would seem that differences between readers may lead to more severe differences in age composition than in average lengths-at-age.

References

Armitage, P. 1971. Statistical Methods in Medical Research. Blackwell Scientific Publications, Edinburgh, Scotland.

Snedecor, G.W. and W.G. Cochran. 1967. Statistical Methods, sixth edition. The Iowa State University Press, Ames, Iowa, U.S.A.

Sokal, R.R. and F. J. Rohlf. 1969. Biometry. W.H. Freeman and Company, San Francisco.

Table 1. Sample sizes (n), means ($\bar{y}$) and standard errors (SE) of anomalies of fifteen readers

Reader #	n	$\bar{y}$	SE	Reader #	n	$\bar{y}$	SE
1	225	0.173	0.056	11	254	-0.276	0.057
2	247	0.498	0.060	12	254	-0.039	0.044
3	250	-0.032	0.047	13	193	0.145	0.057
4	251	-0.438	0.053	14	244	-0.008	0.049
5	209	0.134	0.056	15	259	-0.208	0.057
7	254	-0.543	0.052	16	235	1.196	0.086
8	260	0.135	0.055	17	209	0.086	0.098
9	260	0.165	0.055				

Table 2. Significance of paired comparisons of mean anomalies of 15 readers. Markings without parentheses correspond to weighted x^2 test and those within parentheses to Kruskal-Wallis test.

Reader #	1	2	3	4	5	7	8	9	11	12	13	14	15	16
2	*** (***)													
3	** (*)	*** (***)												
4	*** (***)	*** (***)	*** (***)											
5	ns (ns)	*** (***)	* (*)	*** (***)										
7	*** (***)	*** (***)	*** (***)	ns (ns)	*** (***)									
8	ns (ns)	*** (***)	(*) (*)	*** (***)	ns (ns)	*** (***)								
9	ns (ns)	*** (***)	** (*)	*** (***)	ns (ns)	*** (***)	ns (ns)							
11	*** (***)	*** (***)	*** (***)	* (*)	*** (***)	*** (***)	*** (***)	*** (***)						
12	** (*)	*** (***)	ns (ns)	*** (***)	* (*)	*** (***)	* (*)	** (**)	*** (***)					
13	ns (ns)	(***) (***)	* (ns)	*** (***)	ns (ns)	*** (***)	ns (ns)	ns (ns)	*** (***)	* (ns)				
14	* (*)	(***) (***)	ns (ns)	*** (***)	ns (*)	*** (***)	ns (*)	* (**)	*** (***)	ns (ns)	* (ns)			
15	*** (***)	(***) (***)	* (**)	** (**)	*** (***)	*** (***)	*** (***)	*** (***)	ns (ns)	* (**)	*** (***)	** (**)		
16	*** (***)	(***) (***)	*** (***)	*** (***)	*** (***)	*** (***)	*** (***)	*** (***)	*** (***)	*** (***)	*** (***)	*** (***)	*** (***)	
17	ns (ns)	(***) (***)	ns (ns)	*** (***)	ns (ns)	*** (***)	ns (ns)	ns (*)	** (*)	ns (ns)	ns (ns)	ns (ns)	** (ns)	*** (***)

Table 3. Probability levels of significance (p) of tests of skewness of distributions of anomalies.

Skewness	Reader #
None: $p > 0.05$	2, 5
+ve: $p \leq 0.01$	14
+ve: $0.01 < p \leq 0.001$	1, 13, 16, 17
-ve: $0.01 < p \leq 0.001$	3, 4, 7, 8, 9, 11, 12 15

Table 4. Constants a and b of equations of linear regression of length on age, for 15 readers

Reader #	a	b	Reader #	a	b
1	-0.6096	0.2391	11	-0.5719	0.2444
2	-0.3656	0.2050	12	-0.3643	0.2074
3	-0.3368	0.2023	13	-0.6370	0.2454
4	-0.3936	0.2045	14	-0.7458	0.2698
5	-0.5539	0.2299	15	-0.5551	0.2240
7	-0.3375	0.1942	16	-0.4502	0.1997
8	-0.6313	0.2694	17	-0.5074	0.1980
9	-0.5190	0.2514			

Table 5. Significance of paired comparisons of slope constants of 15 readers.

Reader #	1	2	3	4	5	7	8	9	11	12	13	14	15	16
2	ns													
3	ns	ns												
4	ns	ns	ns											
5	ns	ns	ns	ns										
7	ns	ns	ns	ns	ns									
8	ns	ns	ns	ns	ns	ns								
9	*	ns	ns	ns	ns	ns	ns							
11	ns	ns	ns	ns	ns	ns	ns	ns						
12	ns	ns	ns	ns	ns	ns	ns	ns	ns					
13	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns				
14	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns			
15	ns	ns	ns	ns	ns	ns	ns	*	ns	ns	ns	ns		
16	ns	**	**	*	ns	*	***	**	*	**	ns	*	ns	
17	ns	ns	**	ns	ns	*	**	**	*	**	ns	ns	ns	ns