RESTRICTED

International Commission for the Northwest Atlantic Fisherie

Serial No. 3515
(E)

ICNAF Res.Doc. 75/36

ANNUAL MEETING - JUNE 1975

Preliminary report on a sampling study of Subdiv. 4Vn commercial herring landings for 1974

by

W. G. Doubleday
Fisheries & Marine Service
Biological Station
St. Andrews, N. B.

Abstract:

A sampling experiment carried out on herring caught in ICNAF division 4VN in November and December, 1974, and landed at the ports of Sydney, North Sydney, and Louisbourg, Nova Scotia, is analysed. The adequacy of length-weight and age-length keys is examined. Components of variance of catch composition within and between landings are estimated. Trends in catch composition with weight of landings, date of capture, and area of capture are considered. Two sampling techniques are compared for bias. A method of smoothing age-length keys is suggested.

Sampling Design:

The ICNAF Div. 4Vn herring fishery was chosen as site for the experiment because it is one of the best sampled Canadian herring fisheries on the east coast. As many catches as possible from the 4VN fishery landed at Sydney, North Sydney, and Louisbourg in November and December, 1974, were sampled by taking two five-gallon buckets of fish from separate locations in the load for length frequencies and retaining one fish per 1/2 cm length group per bucket as a stratified subsample for ageing. Eight special samples were taken to compare the use of five-gallon buckets with one-gallon buckets. Before unloading began, one five-gallon bucket and four one-gallon buckets of herring were taken from the top of the load. Four more one-gallon buckets of fish were taken, equally spaced through the load during unloading. A stratified sample of one fish per 1/2 cm length group was retained from the large bucket and none from the small buckets. All fish were measured for total length to the 1/2 cm below the actual fish length. Stratified samples were frozen and shipped to St. Andrews for ageing and determination of a weight-length relationship.

At St. Andrews, the fish in the stratified samples were measured to the nearest millimeter, weighed to the nearest 1/10 gram, and aged by otoliths.

Length-Weight Relationship:

Analysis of covariance was applied to the lengths and weights of fish in the stratified samples to determine whether an overall weight-length key was adequate to estimate sample weights for the length frequency samples. Two models were considered:

$$(W/1000) = a + b\ (L/100)^3 \quad \text{and}$$

$$\log_e\ (W/1000) = a + b\ \log_e\ (L/100)$$

Tables 1 and 2 show the analyses.

The inclusion of a constant term in the length cubed regressions is appropriate since a regression line is less steep than the corresponding relation in a bivariate population. Similarly, if weight were proportional to length cubed, then the slope of the theoretical logarithmic regression would be 3ρ when ρ is the correlation between log length and log weight for the stratified samples. For the overall regression, r = 0.98 so that a slope of 2.94 would be expected if the cube law were correct for the bivariate population. The observed slope of 2.877 is significantly less than this so that allometric growth is indicated.

The fitting of individual constant terms and of individual regressions allows departures of some landings from the overall relationship to be detected. For both models, reductions in the residual mean square due to individual constant terms and individual slopes were statistically significant. However, the total reduction was 25% of the overall residual mean square for the length cubed regression and 28% for the logarithmic regression so that biases in applying the overall key to individual samples represent no more than 1% of the total value.

Determining separate length-weight relationships for each landing inflates the variance of the estimated sample weight for length frequency samples considerably. In sampling designs employing length frequency samples of 100 fish and stratified samples of about 40 fish, the variance associated with estimating the mean weights is $\sigma^2/100$ (where σ^2 is the residual variance about the regression) using the overall regression, but this variance becomes $\sigma^2(1/100 + 1/40)$ using individual regressions. Thus, the overall regression is to be preferred unless the landings would be grouped into a few homogeneous groups.

The L^3 regression estimates sample weights to about ±4% (95% confidence), while the log regression is accurate to ±2%. There are slight signs of systematic over-estimation of weight for very small and very large fish using the log regression, but this should not be important unless a landing contains a high proportion of very large or very small fish.

In general, the usual logarithmic length-weight key was found to be satisfactory.

Age-length key:

The fourty-four two-bucket samples were used to construct an age length key (Table 3). This key was divided into two keys by the landing weights in order to examine differences between large (greater than 68 metric tons) and small landings. Differences were visible between the keys for large and small landings at the cross-over lengths where one age becomes more common than the previous age. Fig. 1 illustrates the transition in the proportion of three-year old's and four-year old's. A chi-square test showed that the differences at the transition of ages 2-3 and 3-4 were statistically significant (1%). However, the effects of this bias on estimated age composition is less than 10%. The presence of trends in age-length keys related to the size of landings suggests that the contributions to the key be weighted either by allocating a higher sampling rate to large landings at the design stage, or by giving samples from large landings more weight than samples from small landings when constructing the key. It appears that the length compositions of large and small landings differ as well, so that a proportional stratified sample (1 per every n fish in a length group in the length frequency sample) would be advisable if the weighting by landings leaves a significant bias. Further investigation is required to determine whether area of capture is affecting the key, since area of capture and weight of catch are related.

The age-length keys were examined to see whether smoothing over adjacent lengths was possible. Fig. 2 shows the logarithms of the numbers of age 3 & 4 fish at length in the stratified sample. Some age one fish are included with the two-year olds. The log numbers well approximate parabolas, being better behaved than the proportions of Fig. 1. Fig. 3 shows a parabola fitted to the log numbers of four-year olds at length in the overall key. The good fit of the parabola indicates that smoothing numbers at age is a promising topic for further investigation. Parabolas appear to fit well for all ages although close examination is required.

In view of the slow growth of herring in Nov. and Dec., the parabola of Fig. 3 is the theoretical distribution of log numbers of fish were schooling by mixing age groups of fixed length distribution (normally distributed) in varying proportions. This is contrary to the usual hypothesis of fixed age distribution at length and mixtures of lengths.

If smoothing by parabolas is feasible, fewer fish need be aged to produce a reliable age-length key.

Age Composition in Relation to Landing Weight, Area of Capture and Date of Capture:

Age composition of buckets based on the overall age-length key was plotted against weight of landing, date of capture, and area of capture. Age composition was relatively homogeneous within days and differences between areas were apparent. Slight trends in age composition with landing weights were visible with smaller landings containing fewer old and more young fish than larger landings. Further study, however, is needed to separate the joint contribution of day and area of capture. This preliminary analysis suggests that stratification by date and area of capture would substantially increase the precision of age composition estimates for the whole fishery.

Dispersion of Ages Given the Overall Age-Length Key:

The use of two buckets of fish from a landing enabled between and within landings dispersion matrices of age composition to be estimated. Tables 4 and 5 show the within landings dispersion and correlation matrices respectively. Tables 6 and 7 show the same quantities between landings. All variances are on a per bucket basis. The positive correlations between adjacent ages are largely due to the corresponding overlap of lengths at age. However, the negative correlations between young (2+3 year old) fish and old (6 & older) fish suggest that schools of different size composition are being fished. It is possible that stratification of landings by date and area of capture may considerably reduce the between landings dispersion.

Sampling Rates:

Landings were divided into four categories by weight: 0-20 metric tons, 20-75, 75-150, over 150. The corresponding sampling fractions (no. of landings sampled/ number of landings in the category) were 0.133, 0.278, 0.345, and 0.391 respectively. Thus, landings were not chosen with equal probability. In view of this, the age compositions of the samples cannot be weighted by landing weights of sampled landings and applied on a pro-rata basis to the total catch. Instead, a more complex weighting, taking into account the non-uniformity in choice of vessels is required. Stratification by date of capture and area of capture avoids this difficulty.

Size of Buckets:

Examination of the age compositions of large and small buckets showed that the small buckets, on the average, took 10% more fish of age two and three and 10% less fish of age six and greater than the large buckets. Thus, the small buckets yield a biased age distribution.

No trends were evident in the age composition of different parts of the landings. Thus, it does not seem to matter where a bucket of fish is taken.

Discussion:

Landing weights varied by a factor of forty. Thus, if landings were chosen at random, samples from the largest landings would be given a weighting factor 40 times that of samples from the smallest landings and would contribute 1600 times the variance of samples from the smallest landings to the estimated total catch composition. Considering the differences observed between samples from small and large landings, even if insignificant, suggest that either a larger proportion of large landings should be sampled (probability proportional to landing weight) or else (if possible) length frequencies from large and small landings for homogeneous groups of landings (strata) should be pooled before weighting takes place. Further study is needed to determine whether this pooling is possible for area-day strata.

Proportions in the age-length key are better estimated near the median length of each age class than mid-way between the median lengths of two age classes. The variance in estimations of a key could be reduced by sampling more heavily the lengths of transition from one age group to the next. This would require a sequential allocation of rates of sampling in the stratified samples based on processing the stratified samples as they are collected. This approach is not feasible for the fishery under discussion, but could possibly be applied elsewhere.

Possible biases in the age length key due to differences between landings can be reduced by proportional sampling within a length stratum and the weighting of contributions from different sizes of landings. This approach increases the number of fish to be aged for a given precision in estimation of the key.

Smoothing numbers at age and length in the age-length key by fitting parabolas to their logarithms appears promising and may permit separate keys for different strata of landings to be constructed reducing bias due to pooling over a heterogeneous set of landings.

The relative sizes of the within and between landings dispersions of age distributions suggest that one bucket of fish per landing is adequate with emphasis on increasing the number of landings sampled. However, there is little additional cost in taking an additional bucket and stratification, it appears, will reduce the between landings dispersion considerably. Therefore, samples of two buckets per landing are recommended.

In view of the differing sampling fractions for differing sizes of landings and of the regularity of catch composition for strata based on date and area of capture, it appears that samples should be weighted against total landings in their respective strata.

Acknowledgements:

I wish to thank Dr. W. T. Stobo for extensive contributions to this sampling study and Mr. T. K. Decker and Mr. C. F. Monaghan for collecting the samples.

Table 1. ANACOVA $Wgr/100 = a+b(Lmm/100)^3$

	Df	$(L^3)^2$	SS XP L^3W	W^2
Landings	43	21650.3	1661.70	131.861
Error	816	101040	7251.37	547.654
Total	859	122690	8913.07	679.515

Partition of E_{ww} and S_{ww} into Components for Regression and Deviations

	Regression		Deviations			
	Df	SS	Df	SS	MS	F
Error	1	520.411	815	27.243	0.03343	3.31 (*1%)
Sum	1	647.509	858	32.005		
Landings			43	4.762	0.1108	

Individual Regressions : SS_{Reg} = 523.456 , SS_{Dev} = 24.198 Df = 772

Overall Regression : SS_{Reg} = 647.5099 , SS_{Dev} = 32.01 Df = 851 R^2 = 0.953

Reduction on SS_{Dev} Individual Regressions from common slope 3.0446 Df = 43
F = 2.259 (* 0.1%)

Reduction on SS_{Dev} Common slope from overall regression 4.762 Df = 43
F = 3.316 (* 0.1%)

Slope : Overall regr 0.072647 , common slope 0.07177

Table 2. ANACOVA $\log_e(Wgr/100) = a+b \log_e(Lmm/100)$

	Df	SS & XP $(\log_e L)^2$	$(\log_e L)(\log_e W)$	$(\log_e W)^2$
Landings	43	3.3880	10.3927	32.9075
Error	816	14.8110	41.9657	123.516
Total	859	18.1990	52.3582	156.424

Partition of $E_{(\log w)^2}$ and $S_{(\log w)^2}$ into Components for Regression & Deviation

	Regr		Dev				
	Df	SS	Df	SS	MS	F	
Error	1	118.906	815	4.610	0.00566	4.85	(*1%)
Sum	1	150.632	858	5.792			
Landings			43	1.182	0.62748		

Individual Regressions : $SS_{Reg} = 119.320$, $SS_{Dev} = 4.196$ Df 772

Overall Regression : $SS_{Reg} = 150.632$, $SS_{Dev} = 5.792$ Df 851 $R^2 = 0.963$

Reduction on SS_{Dev} , Individual Regressions from common slope 0.41400 Df 43
F = 1.77139 (*1%)

Reduction on SS_{Dev} , Common slope from overall regression 1.182 Df 43
F = 4.85 (*1%)

Overall Regression $W_{gr} = 0.00001541 \; L_{mm}^{2.877}$, Common slope 2.8341

Table 3. Overall Age-Length Key

Length	1-2	3	4	5	6	7	8	9	10	11+	Total
15.5	1										1
16											0
16.5											0
17											0
17.5	1										1
18											0
18.5	3										3
19	5										5
19.5	2										2
20											0
20.5	9										9
21	7										7
21.5	14										14
22	25	1									26
22.5	26										26
23	43	2									45
23.5	37	1									38
24	39	1									40
24.5	25	11									36
25	16	22									38
25.5	9	36	1								46
26	1	35	1								37
26.5	1	57	2								60
27	0	50	16								66
27.5	0	37	30								67
28	1	20	58								79
28.5		7	57								64
29		4	88								92
29.5			66	3							69
30			98	4							102
30.5			66	6	2						74
31			57	14	2						73
31.5			44	21	6						71
32			14	39	9	4					66
32.5			3	22	23	10	2				60
33			1	12	15	13	2	1	1		45
33.5				1	18	17	8	1			45
34					6	19	20	5	2		52
34.5					1	15	11	12	5		44
35						10	16	17	9	2	54
35.5						2	13	12	13	4	44
36						3	7	9	11	9	39
36.5							4	16	5	11	36
37							1	5	16	19	41
37.5								2	9	15	26
38								3	3	11	17
38.5									1	8	9
39										5	5
39.5										4	4

Table 4. Within Landings Dispersion (per Bucket)

Age	1-2	3	4	5	6	7	8	9	10	11+
1-2	25.6	4.0	2.9	-2.3	-1.1	-0.8	-0.6	-0.3	-0.3	0.0
3	4.0	10.4	4.6	-1.0	-0.7	-0.9	-1.0	-0.6	-0.7	-0.3
4	2.9	4.6	38.8	1.7	-1.6	-2.8	-2.8	-2.6	-1.9	-2.0
5	-2.3	-1.0	1.7	1.6	0.5	0.0	-0.1	-0.1	-0.0	0.0
6	-1.1	-0.7	-1.6	0.5	0.7	0.5	0.1	0.0	0.0	0.1
7	-0.8	-0.9	-2.8	0.0	0.5	0.7	0.5	0.3	0.1	0.0
8	-0.6	-1.0	-2.8	-0.1	0.1	0.5	0.7	0.5	0.3	0.0
9	-0.3	-0.6	-2.6	-0.1	0.0	0.3	0.5	0.7	0.4	0.2
10	-0.3	-0.7	-1.9	-0.0	0.0	0.1	0.3	0.4	0.5	0.5
11+	0.0	-0.3	-2.0	0.0	0.1	0.0	0.0	0.2	0.5	1.2

Table 5. Within Landings Correlation (per Bucket)

Age	1-2	3	4	5	6	7	8	9	10	11+
1&2	1	0.2	0.0	-0.3	-0.2	-0.1	-0.1	-0.0	-0.0	0.0
3	0.2	1	0.2	-0.2	-0.2	-0.3	-0.3	-0.2	-0.2	-0.1
4	0.0	0.2	1	0.2	-0.3	-0.5	-0.5	-0.4	-0.4	-0.2
5	-0.3	-0.2	0.2	1	0.4	0.0	-0.1	-0.1	-0.0	0.0
6	-0.2	-0.2	-0.3	0.4	1	0.7	0.2	0.0	0.1	0.1
7	-0.1	-0.3	-0.5	0.0	0.7	1	0.7	0.4	0.3	0.0
8	-0.1	-0.3	-0.5	-0.1	0.2	0.7	1	0.7	0.5	0.0
9	-0.0	-0.2	-0.4	-0.1	0.0	0.4	0.7	1	0.6	0.2
10	-0.0	-0.2	-0.4	-0.0	0.1	0.3	0.5	0.6	1	0.6
11+	0.0	-0.1	-0.2	0.0	0.1	0.0	0.0	0.2	0.6	1

Table 6. Between Landings Dispersion (per bucket)

Age	1&2	3	4	5	6	7	8	9	10	11+
1&2	569.9	140.6	-157.4	-33.2	-18.0	-16.2	-19.9	-20.7	-17.2	-16.7
3	140.6	82.3	-10.9	-12.4	-8.7	-10.2	-11.1	-11.8	-9.7	-9.5
4	-157.4	-10.9	335.2	19.3	-6.2	-26.2	-32.6	-38.8	-31.1	-29.9
5	-33.2	-12.4	19.3	6.0	2.4	0.8	0.1	-0.3	-0.2	-0.4
6	-18.0	-8.7	-6.2	2.4	2.2	2.8	2.8	2.9	2.2	2.0
7	-16.2	-10.2	-26.2	0.8	2.8	5.5	6.4	7.1	5.6	5.1
8	-19.9	-11.1	-32.6	0.1	2.8	6.4	8.0	9.1	7.2	6.7
9	-20.7	-11.8	-38.8	-0.3	2.9	7.1	9.1	10.6	8.5	7.9
10	-17.2	-9.7	-31.1	-0.2	2.2	5.6	7.2	8.5	6.9	6.6
11+	-16.7	-9.5	-29.9	-0.4	2.0	5.1	6.7	7.9	6.6	6.5

Table 7. Between Landings Correlation

Age	1&2	3	4	5	6	7	8	9	10	11+
1&2	1	0.6	-0.3	-0.5	-0.5	-0.2	-0.2	-0.2	-0.2	-0.2
3	0.6	1	-0.0	-0.5	-0.6	-0.4	-0.4	-0.3	-0.4	-0.4
4	-0.3	-0.0	1	0.4	-0.2	-0.6	-0.6	-0.6	-0.6	-0.6
5	-0.5	-0.5	-0.4	1	0.6	0.1	0.0	-0.0	-0.0	-0.0
6	-0.5	-0.6	-0.2	0.6	1	0.7	0.6	0.5	0.5	0.5
7	-0.2	-0.4	-0.6	0.1	0.7	1	0.9	0.9	0.9	0.8
8	-0.2	-0.4	-0.6	0.0	0.6	0.9	1	0.9	0.9	0.9
9	-0.2	-0.3	-0.6	-0.0	0.5	0.9	0.9	1	0.9	0.9
10	-0.2	-0.4	-0.6	-0.0	0.5	0.9	0.9	0.9	1	0.9
11+	-0.2	-0.4	-0.6	-0.0	0.5	0.8	0.9	0.9	0.9	1

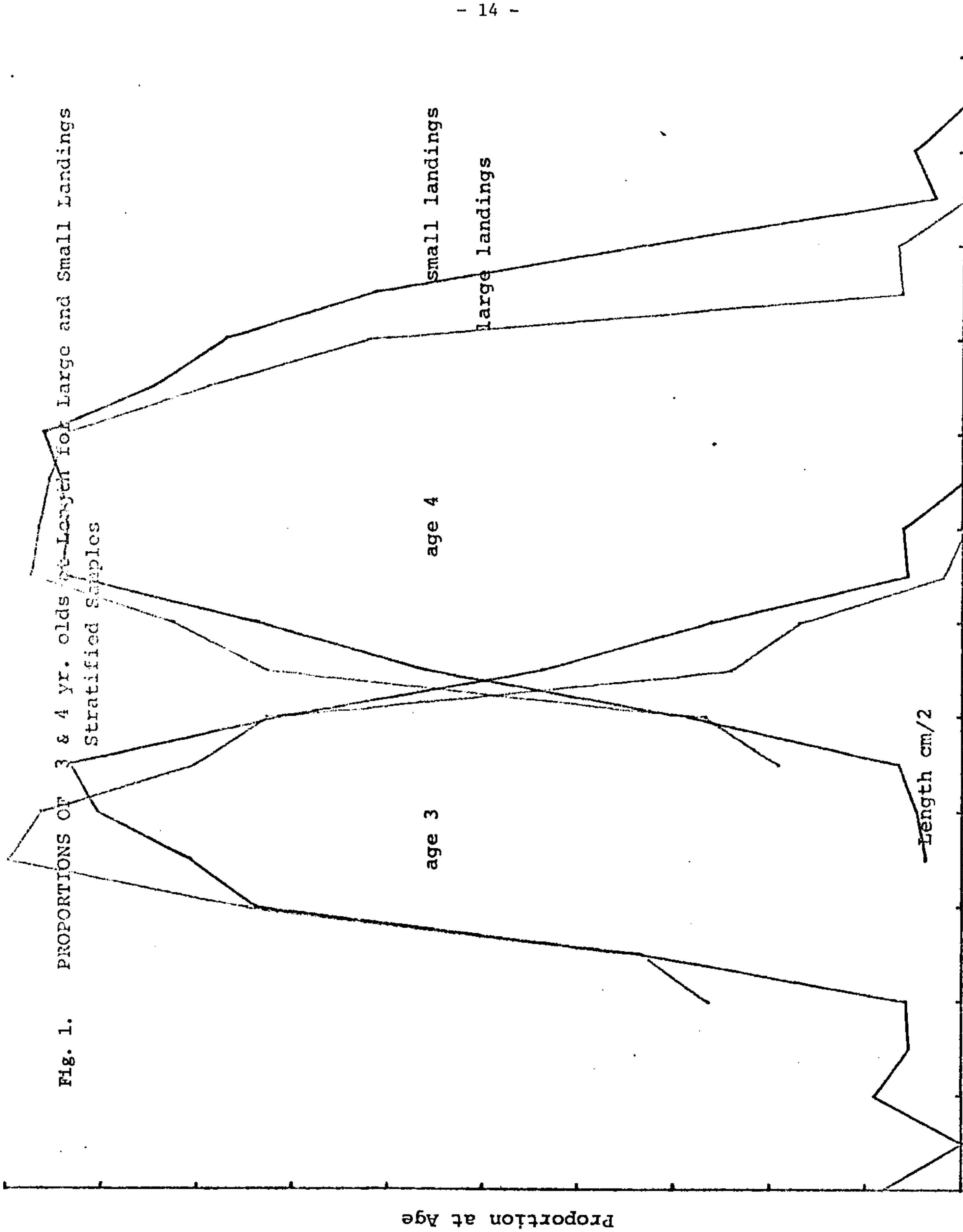


Fig. 1. PROPORTIONS OF 3 & 4 yr. olds at Length for Large and Small Landings
Stratified Samples

Fig. 2. Log Numbers at age 3 & 4 vs Length

Stratified Samples

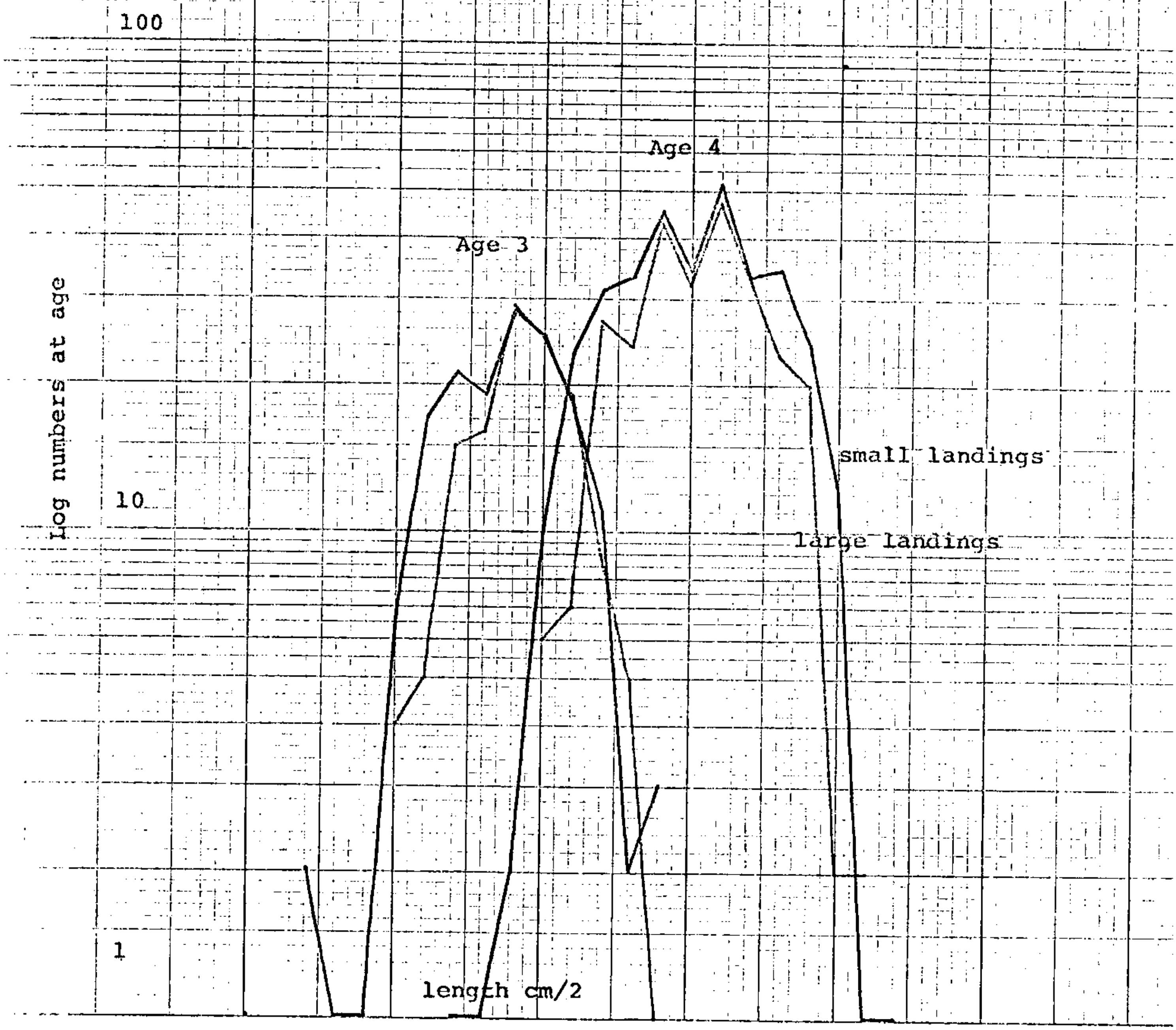
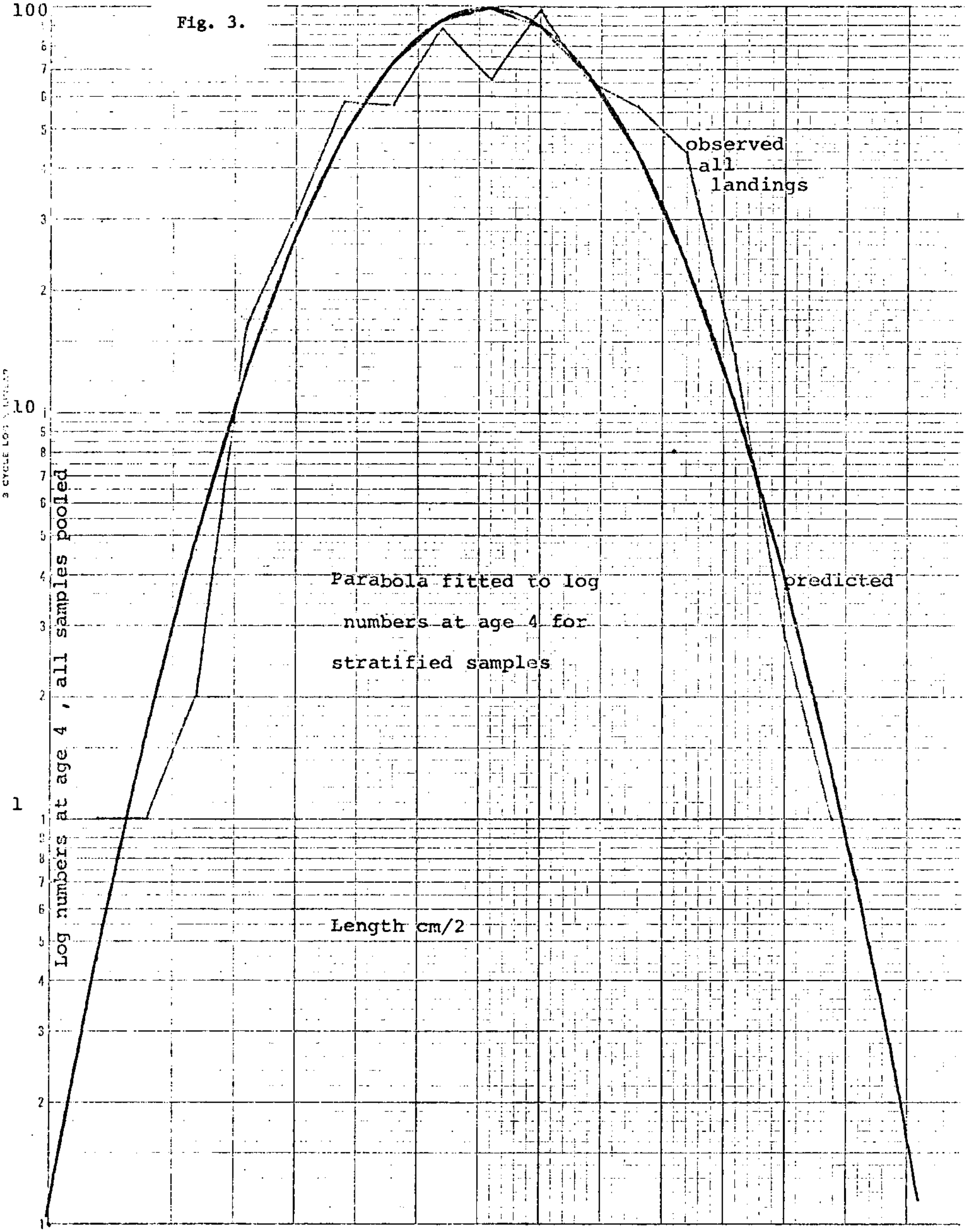


Fig. 3.


100
10
1
observed
all
landings
predicted
Parabola fitted to log
numbers at age 4 for
stratified samples
Length cm/2
Log numbers at age 4 , all samples pooled