International Commission for the Northwest Atlantic Fisheries

Serial No. 3373
(D.c.9)

ICNAF Res.Doc. 74/120

ANNUAL MEETING - JUNE 1974

Statistical Analysis for Sampling Mackerel in the ICNAF - Area

H. Borrmann
Institut für Hochseefischerei und Fischverarbeitung
Rostock-Marienehe

Introduction

The following analysis shall show the precision of the present sampling method used by the G.D.R., and shall give some recommendations how to sample mackerel to get precise age compositions.

Notations and Definitions

n = number of fish sampled for length

n_i = number of fish sampled for length that fall into the i^{th} length group ($i = 1, \ldots, L$)

n_j = number of fish sampled for length that fall into the j^{th} age group

a = total number of fish aged

a_i = number of fish aged in the length group i

a_{ij} = number of fish in the length group i, and age group j ($j = 1, \ldots, K$)

a_j = number of fish sampled in the age group j

p_i = proportion of the sample in the i^{th} length group ($i = 1,..,L$)

p_j = proportion of the sample in the j^{th} age group ($j = 1, ...,K$)

p_{ij} = proportion of age group j fish in the i^{th} length group

$q = 1 - p$

Var = variance

s = standard error ($s = \sqrt{Var}$)

C.V.= coefficients of variation

T_o = fixed total time that can be used in a month for sampling one age group

t_1 = time needed for aging

t_2 = time needed for measuring

D.S.= double sampling

S.R.S. = simple random sampling

VF = factor, by which the variance is reduced

Material and Methods

For the analysis mackerel samples were used which were taken from the catches during the fishery for mackerel in February and March 1972 in the Statistical subarea 6 A and in November and December 1972 in the Subdivision 5 Z e. The number of samples for age and length and the number of fish aged and measured in each month are listed on the bottom of table 1,2,3 and 4. All samples were random samples.

Two quarters were analysed, because they differed in their age compositions and sample sizes. In the sample from November and December the age groups 0 and 1 were also present.

For all calculations we used the computer programme AGECOM (ABRAMSON, 1971). This programme was adapted to our computer. The base for the calculations were the equations which were presented by KUTKUHN (1963) and MACKETT (1963).

The number of samples was not taken into consideration in this analysis.

1. Estimation of the variances of the age groups in the samples

   a) for simple random sampling:

$$\text{Var } p_i = \frac{1}{a} \cdot p_j \ (1 - p_j),$$

   b) for double sampling:

$$\text{Var } p_j = \sum_{i=1}^{L} \frac{p_i^2 (p_{ij} \cdot q_{ij})}{a_i} + \sum_{i=1}^{L} \frac{p_i (p_{ij} - p_j)^2}{n}.$$

   This equation is valid for any allocation of the sample for age.

The calculations were made for all mentioned months. Different age-length keys were used:

1. age samples were combined only from one month,
2. age samples were combined from the two months of the corresponding quarter.

In addition the age-length key (combined from the quarter) for March and December were changed proportionally so that we got following numbers of aged fish per length group:

	March	December
1.	52	up to length group 19 = 33, from length group 20=67
2.	20	up to length group 21 = 10, from length group 22=20

In the first option the total number of aged fish were not changed.

2. Estimation of the age sample sizes per age group needed to reduce standard erros

It was assumed that the length-frequency sample size remains constant and the sampling for age is proportional.

The needed number per age group is:

$$a_j = \frac{n \cdot \sum_{i=1}^{L} p_i \cdot p_{ij} \cdot q_{ij}}{n \cdot (1 - VF)^2 \cdot \mathrm{Var}\, p_j - \sum_{i=1}^{L} p_i \, (p_{ij} - p_i)^2},$$

where Var p_j is the corresponding result from the calculation in chapter 1 b.

The results for these calculations are presented for the options per month, where the age-length key was used from the combined age samples for the corresponding month.

3. Estimation of the sample sizes per age group needed for a minimum variance with fixed time for sampling and expected variances using these sample sizes.

a) Sample sizes

- for double sampling:

$$a_j = \frac{\sqrt{V_a}}{\sqrt{t_1}} \cdot \frac{T_o}{\sqrt{V_a \cdot t_1} + \sqrt{V_n \cdot t_2}}$$

$$n_j = \frac{\sqrt{V_n}}{\sqrt{t_2}} \cdot \frac{T_o}{\sqrt{V_a \cdot t_1} + \sqrt{V_n \cdot t_2}},$$

where $V_a = \sum_{i=1}^{L} p_i \cdot p_{ij} \cdot q_{ij}$ and

$$V_n = \sum_{i=1}^{L} p_i \, (p_{ij} - p_j)^2,$$

- for simple random sampling:

$$a_j = \frac{T_o}{t_1} .$$

b) Expected variances

- for double sampling:

$$\text{Var } p_j = \frac{\sqrt{V_a \cdot t_1} + \sqrt{V_n \cdot t_2}}{T_o},$$

- for simple random samling:

$$\text{Var } p_j = \frac{s^2 \cdot t_1}{T_o},$$

where

$$s^2 = p_j (1 - p_j).$$

c) Used times

Instead of the costs for sampling used by KUTKUHN (1963) and MACKETT (1963), we used times, for it was more usefull for us to determinate the needed times for aging and measuring. The used times are:

t_1 = 3 minutes

t_2 = 0,2 minutes

T_o = 260 minutes for February and March,
220 minutes for November and December.

To get an estimation for T_o it was assumed that we have a catch per month of 20 000 t in the 1st quarter and 10 000 t in the 2nd quarter, that 200 length measurements for each 1000 t are needed, and that 600 age determinations per month are sufficient for the age-length key. The sum of the needed times were divided by 10 (expected age groups) to get the time per age group. This is only a rough estimation, for it does not take into consideration the portions of the age groups in the total sample. But the computer programme does not allow an input of different T_o for each age group.

Results and Conclusions

Table 1, 2, 3, 4 show the age compositions, standard errors and coefficients of variation for simple random sampling and double sampling for the analysed months and different age-length keys.

It can be seen that

- for both methods - S.R.S. and D. S. - the coefficients of variations are below 10 % or about 10 % (February) for the important ($p_j > 0,10$) age groups, when we consider the unchanged age-length key for D. S.,
- the difference between S.R.S. and D.S. is small for the important age groups, when we have the sample sizes listed in the tables, the C.V. for D.S. are somewhat smaller in nearly all cases,
- for D.S. with changed age-length keys, when the total number of aged fish is not changed, the C.V. for the old age groups and age groups 0 and 1 for December become much smaller, but for the important age groups somewhat greater,
- for D.S., when only 20 fishes per length group are aged, the C.V. will become much greater for the important age groups (about 2 times), but for the old age groups and for the age groups 0 and 1 for December the C.V. stay below the C.V. for the unchanged age-length key,
- the age compositions differ for one month, when the age compositions calculated with the age-length key combined from one month and from two month of the corresponding quarter are compared (differences see Table 5). The differences for the two month of one quarter are opposite.

Table 10, 11, 12 and 14 show the sample sizes for D.S. and S.R. S. for a minimum variance with the estimated times for sampling and the expected standard errors using these sample sizes. The results show that

- for a minimum variance not so much age determinations are needed in the younger age groups as in the older ones,
- the expected standard errors are in nearly all cases somewhat greater for S.R.S. than for D.S.,
- for the calculated sample sizes the standard errors are greater than for the original sample sizes.

Table 6,7,8,9 show the age sample sizes per age group needed to reduce the standard errors of the original samples by several percentage. It can be seen that much more fish must be aged per age group, even if we want to reduce the standard error only by 10 %. But it must be taken into consideration that the standard errors of the original samples are already small.

From the results of the presented analysis it can be concluded that

- the precision of the age compositions resulted from sampling by the G. D.R. was good,

- for the used age sample sizes the C. V. for D.S. was not much smaller compared with S.R.S. (but the number of samples were not taken into consideration),

- double sampling is somewhat more efficient than simple random samling,

- it would be better for a small total number of aged fish in the age-length key, when the number sampled for age within a length group is proportional to the number measured to get at least a good precision for the important age groups,

- the age composition per month for D. S. is influenced, when we use an age-length key, where age samples of several months of the quarter are combined, and that it would be better to use an age-length key from the corresponding month.

References

ABRAMSON, N. J. — Computer programmes for Fish Stock Assessment
FAO Fish. Techn. Paper No. 101(1971)

KUTKUHN, J. H. — Estimating absolute age compostion of California salmon landings
Calif, Dept. Fish and Game, Fish.Bull. No. 120 (1963)

MACKETT, D.J. — A method of sampling the Pacific albacore (Thunnus germo) catch for relative age composition
Proc. World Sci. Meet. Biol. Tunas and Rel. Spec., FAO. Fish, Rypt. No 6 (1963), Vol. 3

Table 1

Age composition, standard errors and coefficients of variation for simple random sampling and double sampling
Mackerel, February 1972, statistical area 6 A

Age	S.R.S.			D.S. (Age samples from Feb.)			D.S. (Age samples from Feb.and Ma_{n_1})		
	p_i	s	C.V.	p_i	s	C.V.	p_i	s	C.V.
2	.057	.017753	.3115	.044	.008599[+)]	.1954	.050	.006410	.1282
3	.279	.026713	.0957	.275	.017627	.0641	.278	.012530	.0451
4	.237	.025314	.1068	.247	.025247	.1022	.275	.014637	.0532
5	.279	.026713	.0957	.286	.025865	.0904	.287	.014762	.0514
6	.081	.016272	.0201	.093	.017786	.0607	.080	.008872[+)]	.1109
7	.053	.013341	.2517	.047	.011698	.2489	.021	.004586	.2184
8	.007	.004988	.7126	.002	.001565	.7825	.006	.002276	.3793
9	.007	.004988	.7126	.004	.001701[+)]	.4252	.002	.000842	.4210
10							.001	.000548	.5480
11							.001	.000381	.3810
12							.0003	.000294	.0980
No. Age samples	3			3			10		
No. aged	283			283			931		
No. length samples				6			6		
No.measured				2474			2476		

[+)] standard error estimation invalid because subsamples were of size 1.

Table 2.

Age compositions; standard errors and coefficiens of variation for simple random sampling and double sampling
Mackerel, March 1972, Statistical area 6 A

Age	S.R.S. p_j	s	C.V.	D.S. (age samples from Mar) p_j	s	C.V.	D.S. (Age samples from Feb.a.Mar) p_j	s	C.V.
2	.068	.009890	.1453	.083	.009743	.1174	.083	.008052	.0970
3	.256	.017161	.0670	.277	.014518	.0524	.274	.011958	.0436
4	.265	.017360	.0654	.261	.016162	.0619	.250	.013245	.0530
5	.278	.017609	.0634	.267	.016205	.0607	.267	.013489	.0505
6	.085	.010957	.1291	.077	.009652+)	.1254	.079	.008259+)	.1045
7	.009	.003765	.4067	.008	.003454	.4317	.020	.004250	.2025
8	.015	.004846	.3231	.011	.003419	.3108	.011	.002978	.2707
9	.008	.003440	.4300	.005	.002350	.4700	.006	.002401	.4002
10	.006	.003080	.5133	.004	.002042	.5105	.004	.001918	.4795
11	.006	.003793	.6322	.004	.001990	.4975	.004	.001800	.4500
12	.003	.002181	.7270	.002	.001452	.7260	.002	.001394	.6970
No. age samples			7			7			10
No. aged			648			648			931
No. lengths samples						9			9
No.measured						3692			3692

+) standard error estimation invalid because subsamples were of size 1

Table 2 (continued)

Age	D.S. (age-length key changed[1])			D.S. (age-length key changed[2])		
	p_j	s	C.V.	p_j	s	C.V.
2	.082	.008750	.1067	.079	.012877	.1630
3	.270	.014147	.0524	.265	.021822	.0823
4	.252	.018555	.0736	.253	.030156	.1192
5	.269	.018485	.0687	.270	.030283	.1122
6	.082	.010526	.1284	.080	.016729	.2091
7	.018	.004132	.2296	.026	.009235	.3552
8	.010	.001960	.1960	.011	.003608	.3280
9	.007	.001356	.1937	.007	.002149	.3070
10	.004	.001000	.2500	.003	.001184	.3947
11	.004	.000940	.2350	.003	.001457	.4857
12	.002	.000616	.3080	.002	.000974	.4870

1) 52 aged fishes per length group

2) 20 aged fishes per length group

No. age samples		
No. aged	930	360
No. length samples	9	9
No. measured	3692	3692

Table 3

Age compositions, standard errors and coefficients of variation for simple random sampling and double sampling
Mackerel, November 1972, Subdivision 5 Ze

Age	S.R.S.			D.S.(age samples from Nov.)			D.S. (age samples from Nov. and Dec.)		
	p_j	s	C.V.	p_j	s	C.V.	p_j	s	C.V.
0	.056	.007613	.1359	.058	.003573[+)]	.0616	.052	.003106	.0597
1	.154	.011943	.0776	.147	.005692	.0387	.160	.004838	.0302
2	.123	.010874	.0885	.097	.008029	.0828	.086	.005244	.0610
3	.237	.014086	.0595	.216	.011757	.0544	.202	.007903	.0391
4	.193	.013075	.0678	.203	.012426	.0612	.223	.008693	.0390
5	.135	.011317	.0839	.156	.011411	.0731	.187	.008032	.0430
6	.082	.009102	.1100	.100	.008464	.0846	.070	.005239	.0748
7	.016	.004214	.2634	.018	.004334	.2408	.015	.002729	.1819
8	.003	.001897	.6323	.004	.002253	.5632	.004	.001297	.3242
9	.001	.001096	.9600	.002	.000467[+)]	.2335	.001	.000506	.5060
10							.002	.000508[+)]	.2540

No age samples	10	10	24
No aged	912	912	2003
No length samples		28	28
No measured		9319	9393

[+)] standard error estimation invalid because subsamples were of size 1

Table 4

Age compositions, standard errors and coefficients of variation for simple random sampling and double sampling

Mackerel, December 1972, Subdivision 5 Ze

Age	S.R.S.			D.S.(age samples from Dec.)			D.S. (age samples from Nov.and Dec.)		
	p_j	s	C.V.	p_j	s	C.V.	p_j	s	C.V.
0	.024	.004607	.1920	.057	.013617+)	.2389	.101	.006026+)	.0597
1	.179	.011599	.0648	.240	.014194	.0591	.197	.006922	.0351
2	.103	.009205	.0894	.094	.008905	.0947	.100	.005899	.0590
3	.218	.012441	.0571	.172	.009602	.0558	.182	.007231	.0397
4	.238	.012875	.0541	.203	.010522	.0518	.187	.007565	.0404
5	.171	.011387	.0666	.180	.008994	.0500	.157	.006827	.0435
6	.039	.005812	.1490	.034	.004591	.1350	.057	.004340	.0761
7	.017	.003951	.2324	.012	.002828	.2357	.014	.002276	.1626
8	.006	.002412	.4020	.004	.001637	.4092	.004	.001215	.3037
9	.004	.001826	.4565	.001	.000643	.6430	.002	.000602	.3010
10	.002	.001292	.6460	.001	.000508	.5080	.001	.000444	.4440
11	.001	.000914	.9140	.001	.000265+)	.2650	.001	.000266+)	.2660
13	.001	.000914	.9140	.0002	.000153+)	.7650	.0002	.000153+)	.7650
No. age samples			14			14			24
No. aged			1094			1094			2006
No. length samples			-			19			19
No. measured						9217			9217

+) standard error estimation invalid because subsamples were of size 1

Table 4 (continued)

Age	D.S. (age - length key changed [1]) p_j	s	C.V.	D.S. (age - length key changed [2]) P_j	s	C.V.
0	.102	.004195	.0411	.102	.006478	.0635
1	.184	.006098	.0331	.197	.007721	.0392
2	.098	.009666	.0986	.099	.017775	.1795
3	.188	.014059	.0748	.175	.025340	.1448
4	.192	.014487	.0754	.189	.026660	.1411
5	.159	.011005	.0692	.159	.020182	.1269
6	.057	.005570	.0977	.055	.009304	.1692
7	.014	.002204	.1574	.016	.004641	.2901
8	.004	.000642	.1605	.004	.000968	.2420
9	.002	.000342	.1710	.001	.000426	.4260
10	.001	.000189	.1890	.001	.000282	.2820
11	.001	.000266	.2660	.001	.000266	.2660
13	.0002	.000153	.7650	.0002	.000153	.7650
No. aged			1940			560
No. measured			9217			9217

1) up to length group 19 33 aged fishes, from length group 2o 67 aged fishes per length gropp

2) up to length group 21 10 aged fishes, from length group 22 20 aged fishes per length group

Table 5

The differences between age compositions calculated with different age - length keys

Age	February $P_j^{1)} - P_j^{2)}$	March $P_j^{1)} - P_j^{2)}$	November $P_j^{1)} - P_j^{2)}$	December $P_j^{1)} - P_j^{2)}$
0			+ 0.006	- 0.044
1			- 0.013	+ 0.043
2	- 0.006	0.000	+ 0.011	- 0.006
3	- 0.003	+ 0.003	+ 0.014	- 0.010
4	- 0.028	+ 0.011	- 0.020	+ 0.016
5	- 0.001	0.000	- 0.031	+ 0.023
6	+ 0.013	- 0.002	+ 0.030	- 0.023
7	+ 0.026	- 0.012	+ 0.003	- 0.002
8	- 0.002	0.000	0.000	0.000
9	+ 0.002	- 0.001	+ 0.001	- 0.001
10	- 0.001	0.000	- 0.002	0.000
11	- 0.001	0.000		0.000
12				
13				0.000

1) age samples for the age - length key only from the corresponding month

2) age samples for the age-length key from the months of the corresponding quarter

Table 6.

Estimated age sample sizes needed to reduce standard errors

Mackerel, February 1972, Statistical area 6 A

Age	s	percent reduction					
		0	10	20	30	40	50
2	.008599	16	337	452	644	1021	2020
3	.017627	79	358	487	712	1190	2750
4	.025247	67	312	397	524	723	1066
5	.025865	79	314	401	529	734	1091
6	.017786	23	273	348	460	636	942
7	.011698	15	312	400	532	744	1125
8	.001565	2	810	1123	1703	3080	9754
9	.001701	2	987	2118	+	+	+

Total length sample size 2474.

+ ... Standard error cannot be reduced by this percentage without increasing length sample

Table 7.

Estimated age sample sizes needed to reduce standard errors
Mackerel, March 1972, Statistical area 6 A

Age	s	percent reduction					
		0	10	20	30	40	50
2	.009743	44	683	891	1220	1795	2982
3	.014518	166	734	974	1369	2111	3896
4	.016162	172	791	1013	1346	1883	2841
5	.016205	180	804	1029	1366	1908	2872
6	.009652	55	854	1090	1441	2001	2979
7	.003454	6	829	1054	1384	1899	2773
8	.003419	10	1052	1346	1784	2486	3726
9	.002350	5	1089	1390	1838	2550	3795
10	.002042	4	1086	1395	1861	2622	4008
11	.001990	4	1107	1427	1915	2722	4229
12	.001452	2	1094	1398	1852	2576	3850

Total length sample size 3692.

Table 8.

Estimated age sample sizes needed to reduce standard errors
Mackerel, November 1972, Subdivision 5 Ze

Age	s	percent reduction					
		0	10	20	30	40	50
0	.003573	51	1013	1774	5260	+	+
1	.005692	140	1363	2155	4417	48870	+
2	.008029	112	1333	1708	2270	3175	4793
3	.011757	216	1167	1494	1986	2780	4198
4	.012426	176	1052	1342	1774	2461	3657
5	.011411	123	957	1223	1621	2255	3373
6	.008464	75	876	1135	1536	2213	3528
7	.004334	15	910	1161	1534	2127	3160
8	.002253	3	818	1039	1363	1869	2723
9	.000467	1	+	+	+	+	+

Total length sample size 9319.

+ ... Standard error cannot be reduced by this percentage without increasing length sample

Table 9.

Estimated age sample sizes needed to reduce standard errors
Mackerel, December 1972, Subdivision 5 Ze

Age	s	percent reduction					
		0	10	20	30	40	50
0	.013617	26	138	176	233	321	475
1	.014194	196	216	282	386	566	935
2	.008905	113	1034	1322	1753	2443	3661
3	.009602	236	1545	1982	2641	3708	5635
4	.010522	260	1450	1859	2474	3470	5262
5	.008994	187	1261	1659	2299	3455	6014
6	.004591	42	1337	1727	2327	3329	5235
7	.002828	19	1388	1788	2398	3405	5279
8	.001637	7	1440	1849	2468	3477	5316
9	.000643	4	3123	4217	6104	9972	21498
10	.000508	2	2786	3674	5110	7727	13637
11	.000265	1	+	+	+	+	+
13	.000153	1	+	+	+	+	+

Total length sample size 9217.

+ ... Standard error cannot be reduced by this percentage without increasing length sample

Table 10

Sample sizes needed for minimum variance with fixed time for sampling and expected standard errors using these sample sizes
Mackerel, February 1972, Statistical Area 6 A

Age	Sample sizes			expected standard errors	
	D.S. lengths	ages	S.R.S. ages	D.S.	S.R.S.
2	312	66	87	.018320	.024809
3	333	64	87	.038685	.048185
4	129	78	87	.047344	.045662
5	145	77	87	.049160	.048185
6	146	77	87	.031617	.029352
7	173	75	87	.022564	.024066
8	266	69	87	.004704	.008998
9	438	57	87	.004884	.008998

Table 11

Sample sizes needed for minimum variance with fixed time for sampling and expected standard errors using these sample sizes
Mackerel, March 72, Statistical Area 6 A

Aged	sample sizes D.S. lengths	ages	S.R.S. ages	expected standard errors D.S.	S.R.S.
2	220	72	87	.028059	.027024
3	253	70	87	.043581	.046890
4	136	78	87	.048035	.047431
5	131	78	87	.048442	.048113
6	114	79	87	.029417	.029937
7	77	82	87	.010160	0010288
8	112	79	87	.011518	.013241
9	101	80	87	.008006	.009399
10	130	78	87	.007039	.008413
11	142	77	87	.006958	.008413
12	106	80	87	.004972	.005958

Table 12

Sample sizes needed for minimum variance with fixed time for sampling and expected standard errors using these sample sizes
Mackerel, November 1972, Subdivision 5 Ze

Age	sample sizes D.S. lengths	ages	S.R.S. ages	expected standard errors D.S.	S.R.S.
0	489	41	73	.015054	.026831
1	406	46	73	.026460	.042095
2	138	64	73	.034477	.038327
3	146	64	73	.047655	.049646
4	130	65	73	.047319	.046084
5	145	64	73	.042022	.039888
6	219	59	73	.031537	.032081
7	137	64	73	.015474	.014852
8	94	67	73	.007369	.006686
9	1100	0	73	.001360	.003865

Table 13

Sample sizes needed for minimum variance with fixed time for sampling and expected standard errors using these sample sizes Mackerel, December 1972, Subdiv. 5 Ze

Age	sample sizes D.S. lengths	D.S. ages	S.R.S. ages	expected standard errors D.S.	S.R.S.
0	273	55	73	.022157	.017787
1	396	47	73	.032533	.044781
2	143	64	73	.033975	.035539
3	135	64	73	.044168	.048032
4	137	64	73	.046998	.049705
5	230	58	73	.039685	.043959
6	174	62	73	.020199	.022437
7	161	63	73	.012558	.015255
8	146	64	73	.007327	.009311
9	195	60	73	.004131	.007048
10	170	62	73	.003107	.004988
11	1100	0	73	.000769	.003529
12	82	68	73	.002153	0
13	1100	0	73	.000444	.003529