International Council for
the Exploration of the Sea

CM 2000/K:23
Theme Session: Incorporation of External Factors in Marine Resource Surveys (K)

Assessing Trawl-Survey Estimates of Frequency Distributions

Michael Pennington[1*], Liza-Mare Burmeister[2] and Vidar Hjellvik[1]

*Institute of Marine Research
PO Box 1870 Nordnes
N-58 17 Bergen, Norway

[2]Ministry of Fisheries and Marine Resources
of Namibia. NatMIRC
PO Box 912, Swakopmund, Namibia

Abstract

Marine trawl surveys catch a cluster of fish at each station and fish caught together tend to have more similar characteristics, such as length, age, stomach contents etc., than those in the entire population. When this is the case, the effective sample size of estimates of the frequency distribution of a population characteristic can be much smaller than the number of fish sampled during a survey. As examples, it is shown that the effective sample size for estimates of length-frequency distributions generated by trawl surveys conducted in the Barents Sea, off Namibia and off South Africa is on average approximately one fish per tow. It **is** concluded that many more fish than necessary are measured at each station and that one way to increase the effective sample size for these surveys and, hence, increase the precision of the length-frequency estimates, is to reduce tow duration and use the time saved to collect samples at more stations.

Key *words:* Trawl surveys; cluster sampling; intra-haul correlation: effective sample size

'Tel: +47-5523-6309; fax: +47-5523-8687; e-mail: michael@imr.no

Introduction

Survey-based assessments often appear to provide a more accurate prognosis of the status of a fish stock than catch-based assessments (Nakken, 1998; Pennington and Strømme, 1998; Korsbrekke, *et* al., 2000). An advantage that survey-based assessments have over those using commercial catch statistics is that the uncertainties associated with survey estimates can be studied and quantified, and based on such research, survey methodology, and ultimately stock assessments, can be improved (Godø, 1994). In contrast, it is generally difficult to determine either the accuracy or the precision of estimates based on commercial catch data, and it is not clear how to improve, at a reasonable cost, the collection of catch data so that these data would more accurately reflect the mortality caused by fishing (Christensen, 1996).

Marine bottom trawl surveys provide estimates of the abundance or relative abundance of demersal fish stocks and estimates of the frequency distribution of various population characteristics, such as length, age, stomach contents, etc. In this paper we examine the precision of survey-generated estimates of length-frequency distributions. The focus is on length, but the results are relevant for estimating the frequency distribution of other population characteristics.

It was found that, for the surveys examined, the precision of the estimated length-frequencies is rather low given the number of fish that were measured during the surveys. That is the 'effective sample size' is much smaller than the number of fish measured. The reason for the low effective sample sizes for survey estimates of length-frequency distributions is that fish that are caught together tend to be more similar than those in the general population. It is concluded that the only practical way to increase the precision of estimates of frequency distributions and to improve overall survey efficiency would be to reduce survey tow duration and use the time saved to collect samples at more stations.

Materials and methods

Survey length data

Bottom trawl survey length data for Northeast Artic cod (*Gadus mohua*) and Northeast Arctic haddock (*Melanogrammus aeglefinus*) are from the Institute of Marine Research (Norway) winter and summer surveys in the Barents Sea. The surveys are stratified systematic surveys and at each station the trawl is towed for 30 minutes (see Aglen, 1999; Mehl. 1999 for details).

The Namibian deepwater hake (*Meluccius paradoxus*) data were collected during bottom trawl surveys off Namibia conducted by the Ministry of Fisheries and Marine Resources of Namibia in conjunction with the Norwegian Agency for Foreign Aid (NORAD). For these surveys, tows of 30-minute duration are made at stations along transects perpendicular to the coast (see Anon., 1999).

The deepwater hake data for South Africa are from bottom trawl surveys off the west coast of South Africa. The surveys were conducted by the Marine and Coastal Management Centre, South Africa, using a stratified random design with 30-minute tows at each station (see Payne *et* al., 1985).

Assessing the precision of length frequency estimates

The sample of fish of a particular species measured during a survey is not a random sample of individual fish from the entire population but a sample of n clusters, one cluster from each station. Since fish caught together are usually more similar than those in the general population, a total of M fish collected in n clusters will contain less information about the population length distribution than M fish randomly sampled. One way to measure the information contained in a sample of length measurements is to estimate the number of fish that one would need to sample at random (the effective sample size) to obtain the same information on length contained in the cluster samples.

The effective sample size for cluster sampling can be defined and calculated as follows (Pennington and Vølstad, 1994; Folmer and Pennington, 2000). First estimate the population mean fish length and its variance based on the clusters of fish caught at n stations. Since both the lengths and the number of fish at a station are random variables, a ratio estimator is appropriate (Cochran, 1977). The ratio estimator, $\hat{R}$, of the mean length is given by

$$\hat{R} = \frac{\sum_{i=1}^{n} M_i \hat{\mu}_i}{\sum_{i=1}^{n} M_i} \tag{1}$$

where M_i is the number of fish caught (either actual or estimated) at station i and $\hat{\mu}_i$ denotes an estimate of the average length of fish at station i. For example, if the catch at a station is divided into strata and a random sample of fish are chosen in each stratum, then for that station $\hat{\mu}_i$ in eq. (1) would be the stratified estimate of mean length. The estimated variance of $\hat{R}$ is approximately given by

$$\mathrm{var}(\hat{R}) = \sum_{i=1}^{n} \frac{(M_i / \overline{M})^2 (\hat{\mu}_i - \hat{R})^2}{n(n-1)} \tag{2}$$

where $\overline{M} = \sum_{i=1}^{n} M_i / n$.

Next estimate the variance, σ_x^2, of the population length distribution. If m_i fish are randomly selected at each station (or if all fish are measured) then

$$\hat{\sigma}_x^2 = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m_i} (M_i / m_i)(x_{i,j} - \hat{R})^2}{M - 1} \tag{3}$$

is an estimator of σ_x^2, where $M = \sum_{i=1}^{n} M_i$ is the total number of fish caught during the survey and $x_{i,j}$ is the length of the j^{th} fish at station i. For other sampling schemes at a station, first estimate the number of fish caught during the survey in each of L length bins, then

$$\sigma_x^2 = \frac{\sum_{k=1}^{L} f_k (y_k - \hat{R})^2}{M - 1} \tag{4}$$

is an estimator of a,‘, where f_k is the frequency of fish in the k^{th} length bin and y'_k is the bin's midpoint.

Now if it were possible to sample m fish at random from the population then the variance of the sample mean would be equal to σ_x^2 / m . *The* effective sample, m_{eff}, is defined as the number of fish that would need to be sampled at random so that the sample mean would have the same precision as an estimate based on a sample of n clusters. An estimate of the effective sample size for a particular cluster sample can be derived by substituting the estimates from (2) and either (3) or (4) into the equation

$$\frac{\hat{\sigma}_x^2}{\hat{m}_{eff}} = \text{var}(\hat{R})\,. \tag{5}$$

Simulation techniques were used to examine the effect that reducing the total number of fish measured during a particular survey would have on the estimates of the mean length and the effective sample size. Length measurements consist of one or more subsamples from the fish caught at each station. The simulated estimates of the distributions of $\hat{m}_{eff}$ and $\hat{R}$ (given the actual fish measured during the survey) were generated by randomly selecting from every haul a maximum of k fish without replacement from each subsample. If fewer than k fish were in a subsample, then all were chosen. This was done 500 times for k = 10.30 and 100 and each run produced values of $\hat{m}_{eff}$ and $\hat{R}$.

To assess the precision of an estimated length distribution, bootstrapping (Efron. 1982) was used to generate 95% confidence intervals for the number of fish in each 5 cm length bin. For each of 500 runs, n stations (the number of tows made during the survey) were randomly sampled with replacement and the confidence interval for each 5 cm length bin was based on the resulting 500 estimates of the number of fish in that bin. Finally, bootstrapping was used to examine how much the length of the 95% confidence intervals would increase if a maximum of 10 fish were selected from each subsample.

Results

Estimates of the effective sample size and associated statistics for survey-based estimates of the length composition of cod in the Barents Sea are in Table 1. The results indicate that for cod the estimated effective sample size is small compared with the number of fish measured. For example during the 1995 winter survey, 175,006 cod were caught, 47,286 were measured and the effective sample size was 313 fish or 0.7% of the total number measured (Table 1). The average effective sample size for the winter surveys is 1.2 cod per tow and for the summer surveys, 1 .O cod per tow. The estimated effective sample sizes for the Northeast Arctic haddock survey data were, on average, approximately one fish per tow (Table 2).

The effective sample sizes for the survey estimates of the length distribution of deepwater hake off Namibia and off South Africa (Tables 3 and 4) followed the same pattern as for cod and haddock in the Barents Sea. In particular, the average effective sample size was 0.5 hake per tow for the Namibian surveys and 1.3 hake per tow for the South African surveys.

The simulated distributions of $\hat{m}_{eff}$ and $\hat{R}$, which demonstrate the effects of reducing the total number of measured fish on estimates of mean length, for the 1995 and 1999 winter surveys of cod in the Barents Sea are shown in Figures 1 and 2. For example, if a maximum of 30 fish were selected from each subsample at each station, then a total of 11,123 fish would have been measured during the 1995 survey compared with 47,286 fish that were actually measured. In 1995, $\hat{R}$ = 19.96 and the 95% confidence interval for $\hat{R}$ is (18.29, 21.63). As can be seen from Fig. 1, all 500 simulated estimates of the mean based on the reduced sample size are well within the 95% confidence limits for $\hat{R}$. When the number of fish measured was reduced to a maximum of 10 fish per subsample for a total sample of 2,597 fish, the simulated estimates were also well within the 95% confidence interval for $\hat{R}$ (Fig. 1). The results of the simulations for the winter survey in 1999 were similar to those for 1995 (Fig. 2).

Bootstrapped estimates of the 95% confidence intervals for the estimated number of fish in each 5 cm length bin for the 1995 and 1999 Barents Sea winter survey of cod are shown in Fig. 3. The inner brackets denote the confidence interval based on the total number of fish actually measured and the outer brackets denote the confidence intervals if a maximum of 10 fish were measured per subsample.

Discussion and conclusions

For all the surveys examined, the effective sample size for the survey length data was much smaller than the total number of fish measured. The average effective sample size was approximately one fish per tow and it seems to be typical that the effective sample size for estimating length distributions is relatively small for marine surveys. For example. the effective sample size for trawl surveys of haddock on Georges Bank was on average less than 0.5 fish per tow (Pennington and Vølstad, 1994) and for shrimp in a small area off West Greenland. about 3 shrimp per tow (Folmer and Pennington. 2000).

The reason that estimates of the length distributions were rather imprecise given the number of fish that were measured is that the sizes of the fish in a haul tend to be more similar than those in the entire population. An additional factor is that the density of fish in a survey region is usually quite variable. To see this, consider the equation for the expected value of $\mathrm{var}(\hat{R})$. If every fish is measured during a survey, then subject to some assumptions, the expected variance of $\hat{R}$ when n stations are sampled is given approximately by (Pennington and Vølstad, 1994)

$$\mathrm{Var}(\hat{R} = \frac{\sigma_x^2\{1+(\overline{M}-1+\sigma_m^2/\overline{M})\rho\}}{A4}$$

where $\overline{M}$ is the expected mean catch per tow, σ_m^2 is the tow-to-tow variance of catch, M (= $n\,\overline{M}$) is the expected total number of fish caught, σ_x^2 is the population variance of length, and ρ is the coefficient of intra-haul correlation (see Cochran, 1977, p. 209) for length. If $\rho = 0$, then $\mathrm{Var}(\hat{R}) = \sigma_x^2 / M$ and therefore the effective sample size is equal to M. However if $\rho > 0$ (i.e. fish of similar length tend to be caught together), then the terms in the parentheses can greatly increase the variance and thus drastically reduce the effective size. In particular, the term $\sigma_m^2 / \overline{M}$ is relatively large for marine surveys.

The precision of estimates of other population characteristics, such as the age distribution, can also be relatively low compared with the number of fish sampled if the particular attribute or measurement is more similar for fish caught together than for those in the general population. For example, the precision of estimates of mean stomach contents (Bogstad et al., 1995) or diet composition (Tirasin and Jørgensen, 1999) can be relatively low due to intra-haul correlation.

It appears, based on the bootstrapped estimates of precision and the sampling simulations, that reducing (or increasing) the number of fish measured at a station will not significantly affect the precision of the estimates of length distributions. It has also been observed that short tows are generally more efficient for estimating stock abundance than long tows (Gods *et* al., 1990; Pennington and Vølstad, 199 1; Gunderson, 1993; Kingsley et al., 2000). Therefore one way to increase the precision of marine survey estimates of frequency distributions and improve overall survey efficiency without increasing survey cost is to reduce tow duration and use the time saved to collect samples at more stations (Pennington and Vølstad, 1994).

Acknowledgement

We would like to thank Rob Leslie, Marine and Coastal Management Centre, South Africa, for providing us with the South African survey data.

Literature cited

Aglen, A., 1999. Report on demersal fish surveys in the Barents Sea and Svalbard area during summer/autumn 1996 and 1997. Fisket og Havet NR. 7- 1999.

Anonymous, 1999. Survey of the fish resources of Namibia. Preliminary Cruise Report No 1/99. NORAD-FAO/UNDP Project GLO92/013, 52 pp.

Bogstad, B., M. Pennington, and J. H. Vølstad, 1995. Cost-efficient survey designs for estimating food consumption by fish. Fish. Res. 23:37-46.

Christensen, V., 1996. Virtual population reality. Rev. Fish Biol. Fish. 6: 243-247.

Cochran, W. G., 1977. Sampling Techniques, 3rd ed. John Wiley and Sons, New York, NY, 428 pp.

Efron, B., 1982. The jackknife, the bootstrap, and other resampling plans. SIAM, CBMS-NSF Regional Conference Series in Applied Mathematics, No. 38, Philadelphia, PA.

Folmer O., and M. Pennington, 2000. A statistical evaluation of the design and precision of the shrimp survey off West Greenland. Fish. Res. (in press).

Godø, 0. R., 1994. Factors affecting the reliability of groundfish abundance estimates from bottom trawl surveys. In: Fernø, A., Olsen, S., (Eds). Marine Fish Behaviour in Capture and Abundance Estimation. Fishing News Books. Farnham. UK. 166-199.

Godø, O. R., M. Pennington, and J.H. Vølstad, 1990. Effect of tow duration on length composition of trawl catches. Fish. Res. 9:165-179.

Gunderson, D. R., 1993. Surveys of Fisheries Resources. John Wiley and Sons, New York, NY, 248 pp.

Kingsley, M.C.S., D.M. Carlsson, P. Kanneworft, and M. Pennington, 2000. Spatial structure of the resource of *Pandalus borealis* and its implications for trawl survey design. Fish. Res. (submitted).

Korsbrekke, K., S. Mehl, 0. Nakken, and M. Pennington, 2000. A survey-based assessment of the Northeast Arctic cod stock. ICES J. Mar. Sci. (submitted).

Mehl, S., 1999. Botnfiskundersøkingar i Barentshavet vinteren 1999. (Norwegian with English summary) Fisken og Havet NR. 13- 1999.

Nakken, O., 1998. Past, present and future exploitation and management of marine resources in the Barents Sea and adjacent areas. Fish. Res. 37:23-35.

Payne. A.I.L.. C.J. Augustyn, and R.W. Leslie, 1985. Biomass index and catch of Cape hake from random stratified sampling cruises in division 1.6 during 1984. Colln. Scient. Pap. Int. Corn. SE Atl. Fish. 12:99-123.

Pennington. M.. and T. Strømme, 1998. Surveys as a research tool for managing dynamic stocks. Fish. Res. 37:97-106.

Pennington, M., and J.H. Vølstad, 1991. Optimum size of sampling unit for estimating the density of marine populations. Biometrics 47:717-723.

Pennington, M., and J.H. Vølstad, 1994. Assessing the effect of intra-haul correlation and variable density on estimates of population characteristics from marine surveys. Biometrics 50:725-732.

Tirasin, E.M., and T. Jorgensen. 1999. An evaluation of the precision of diet description. Mar. Ecol. Prog. Ser. 182:243-252.

Table 1. Summary statistics for assessing the precision of the estimated length distributions of Northeast Arctic cod based on the winter (a) and summer (b) bottom trawl surveys in the Barents Sea. The estimated effective sample size is denoted by $\hat{m}_{eff}$, n is the number of stations at which cod were caught, M is the total number of cod caught, m is the number measured, $\hat{R}$ is the estimate of mean length and var($\hat{R}$) is its variance.

(a) Winter

	n	M	m	$\hat{R}$ (cm)	var($\hat{R}$)	$\hat{m}_{eff}$	$\hat{m}_{eff}/n$	$(\hat{m}_{eff}/m)\times100\%$
95	296	175006	47286	20.0	0.7	313	1.1	0.7
96	314	209114	44021	18.0	0.3	511	1.6	1.1
97	177	71418	25689	19.0	2.1	119	0.7	0.7
98	197	60746	32536	22.1	0.7	394	2.0	1.2
99	223	50192	21760	25.0	1.9	107	0.5	0.5
	Avg.	113295	34258			289	1.2	0.8%

(b) Summer

		M	m	$\hat{R}$ (cm)	var($\hat{R}$)	$\hat{m}_{eff}$	$\hat{m}_{eff}/n$	$(\hat{m}_{eff}/m)\times100\%$
95	329	66643	46161	31.2	1.4	252	0.8	0.6
96	341	15834	45286	24.4	0.6	478	.4	
97	266	72093	26947	23.1	0.8	266	.0	.0
98	218	72360	23461	25.1	.1	184	0.8	0.8
99	217	46593	23253	30.8	0.9	211	0.9	0.9
	Avg.	74705	33022			278	1.0	0.9%

Table 2. Summary statistics for assessing the precision of the estimated length distributions of Northeast Arctic haddock generated by the winter (a) and summer (b) bottom trawl surveys in the Barents Sea. The notation is the same as in Table 1.

(a) Winter

Year	n	M	m	$\hat{R}$ (cm)	var($\hat{R}$)	$\hat{m}_{eff}$	$\hat{m}_{eff}/n$	$(\hat{m}_{eff}/m)\times100\%$
95	199	66009	22938	25.0	1.0	168	0.8	0.7
96	235	54892	25525	32.0	2.9	69	0.3	0.3
97	140	37441	13273	22.0	0.8	185	0.8	.4
98	144	12704	9620	23.9	1.0	169	.2	1.8
99	182	41612	12152	13.4	0.4	188	.0	.6
	Avg.	42532	16702			155	0.8	1.2%

(b) Summer

Year	n	M	m	$\hat{R}$ (cm)	var($\hat{R}$)	$\hat{m}_{eff}$	$\hat{m}_{eff}/n$	$(\hat{m}_{eff}/m)\times100\%$
95	208	25771	15763	27.0	0.95	147	0.7	0.9
96	163	14139	7338	31	3.65	51	0.3	0.7
97	114	13560	4314	23.	.72	56	0.5	.3
98	89	7432	2699	21.3	0.34	170	1.9	6.3
99	140	11922	5489	20.1	0.36	197	1.4	3.6
	Avg.	14565	7536			124	1.0	2.6%

Table 3. Summary statistics for assessing the precision of the estimated length distribution of deepwater hake off Namibia based on bottom trawl surveys. The notation is the same as in Table 1.

Survey	n	M	m	$\hat{R}$ (cm)	var($\hat{R}$)	$\hat{m}_{eff}$	$\hat{m}_{eff}/n$	$(\hat{m}_{eff}/m)\times 100\%$
Sept. 90	37	6671	1837	24.6	1.8	35	1.0	1.9
Jan. 91	19	3887	1329	29.2	3.3	19	1.0	1.4
Oct. 92	53	22369	5090	30.1	2.1	30	0.6	0.6
Apr. 92	63	33107	5411	34.0	2.8	30	0.5	0.6
Oct. 93	88	36814	8480	30.3	2.6	41	0.5	0.5
Jan. 93	70	36247	8208	37.3	2.9	33	0.5	0.4
Apr. 93	84	25746	7023	32.7	4.2	35	0.4	0.5
Jan. 94	60	30134	7997	28.4	10.0	13	0.2	0.2
Apr. 94	103	72012	17694	35.3	1.4	63	0.6	0.4
Oct. 94	105	70817	17216	28.1	1.9	44	0.4	0.3
Apr. 95	79	47585	14661	26.0	7.0	20	0.3	0.2
Jan. 96	105	57540	27834	30.3	2.9	45	0.4	0.2
Sept. 96	105	78562	24975	28.7	1.6	50	0.5	0.2
Jan. 97	122	54995	27648	28.5	4.6	29	0.2	0.
Jan. 98	104	52573	15717	34.5	2.3	44	0.4	0.4
Jan. 99	104	68419	19305	28.4	4.	30	0.3	0.2
	Avg.	43592	13152			35	0.5	0.5%

Table 4. Summary statistics for assessing the precision of the estimated length distribution of deepwater hake off South Africa based on bottom trawl surveys. The notation is the same as in Table 1.

Survey	n	M	m	$\hat{R}$ (cm)	var($\hat{R}$)	$\hat{m}_{eff}$	$\hat{m}_{eff}/n$	$(\hat{m}_{eff}/m)\times 100\%$
Jan. 85	75	75883	13863	29.3	0.7	70	0.9	0.5
July 85	65	55704	10786	29.9	0.9	52	0.8	0.5
Jan. 86	86	82720	16216	28.2	0.3	132	1.5	0.8
Jan. 87	91	140685	18317	26.7	0.3	122	1.3	0.7
July 87	76	80476	14774	27.0	0.8	68	0.9	0.5
Feb. 88	88	91828	17187	24.3	0.5	98	1.1	0.6
Feb. 89	33	234796	10115	25.7	0.1	207	6.3	2.0
Jan. 90	75	150814	23093	26.0	1.6	43	0.6	0.2
Jan. 91	73	226234	17115	27.2	0.9	38	0.5	0.2
Feb. 92	83	174364	24334	25.2	0.8	58	0.7	0.2
Jan. 93	81	102395	24922	26.0	0.6	89	1.1	0.4
Jan. 94	63	139268	28621	27.1	0.7	68	1.1	0.2
Jan. 95	81	137225	37481	26.4	0.5	97	1.2	0.3
Jan. 96	77	167765	31538	25.2	1.1	46	0.6	0.1
Jan. 97	88	219218	33537	24.2	0.7	70	0.8	0.2
Jan. 98	76	251050	26328	25.9	0.3	116	1.5	0.4
Jan. 99	74	218438	26144	25.4	0.5	91	1.2	0.3
	Avg.	149933	22022			86	1.3	0.5%

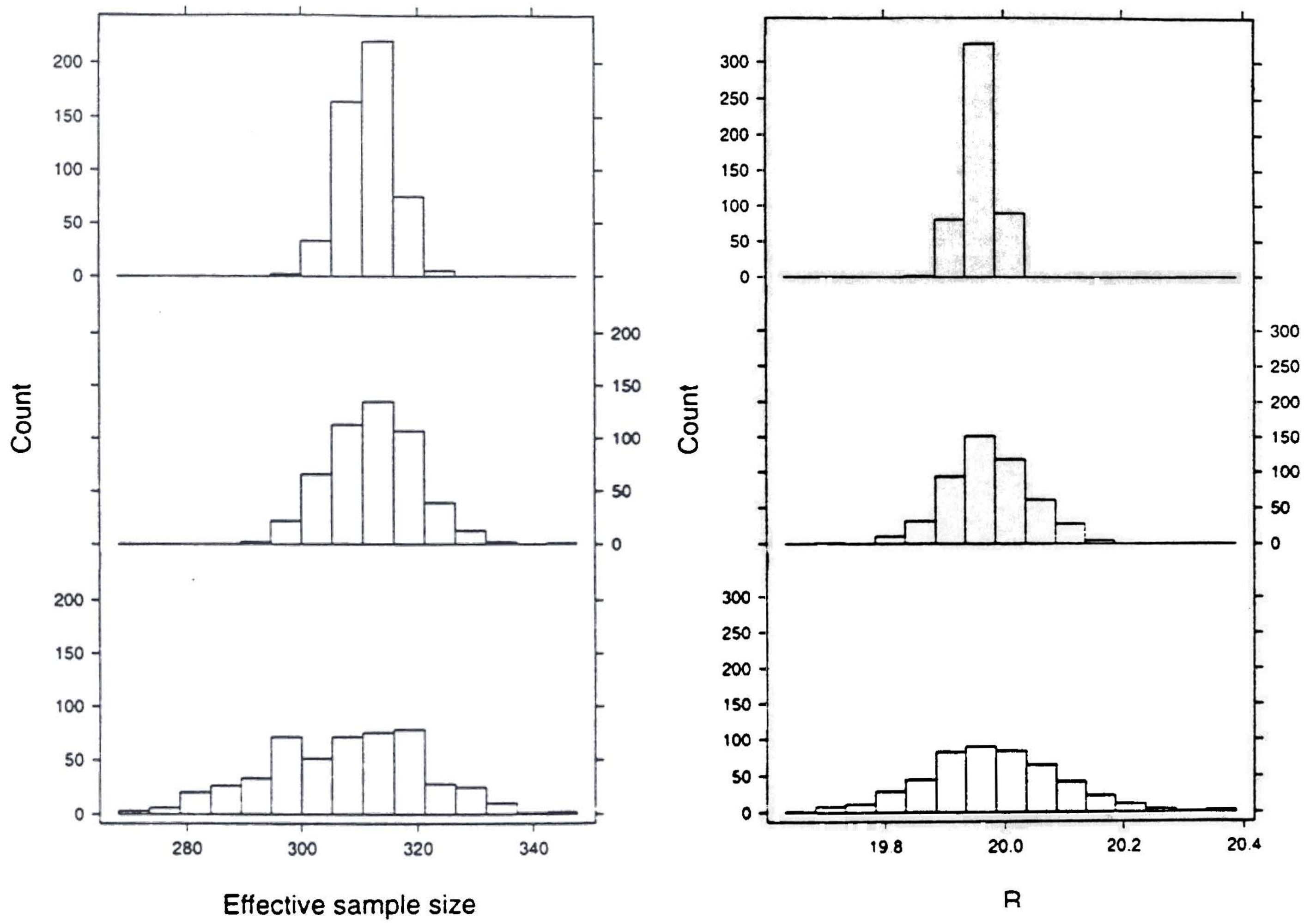

Fig. 1. Simulated estimates of the distribution of the effective sample size, $\hat{m}_{eff}$, and of the mean length, $\hat{R}$, when the total number of fish measured is reduced for the 1995 winter survey in the Barents Sea. The top panel is when a maximum of k = 100 fish are selected per subsample for a total of m = 30,403 fish in each run; the middle panel, k = 30, m = 11,123; and the bottom panel, k = 10, m = 3,911. The estimate of the population mean, R, based on the entire sample (m = 47,286) is 19.96 and its 95% confidence interval is (18.29, 21.63).

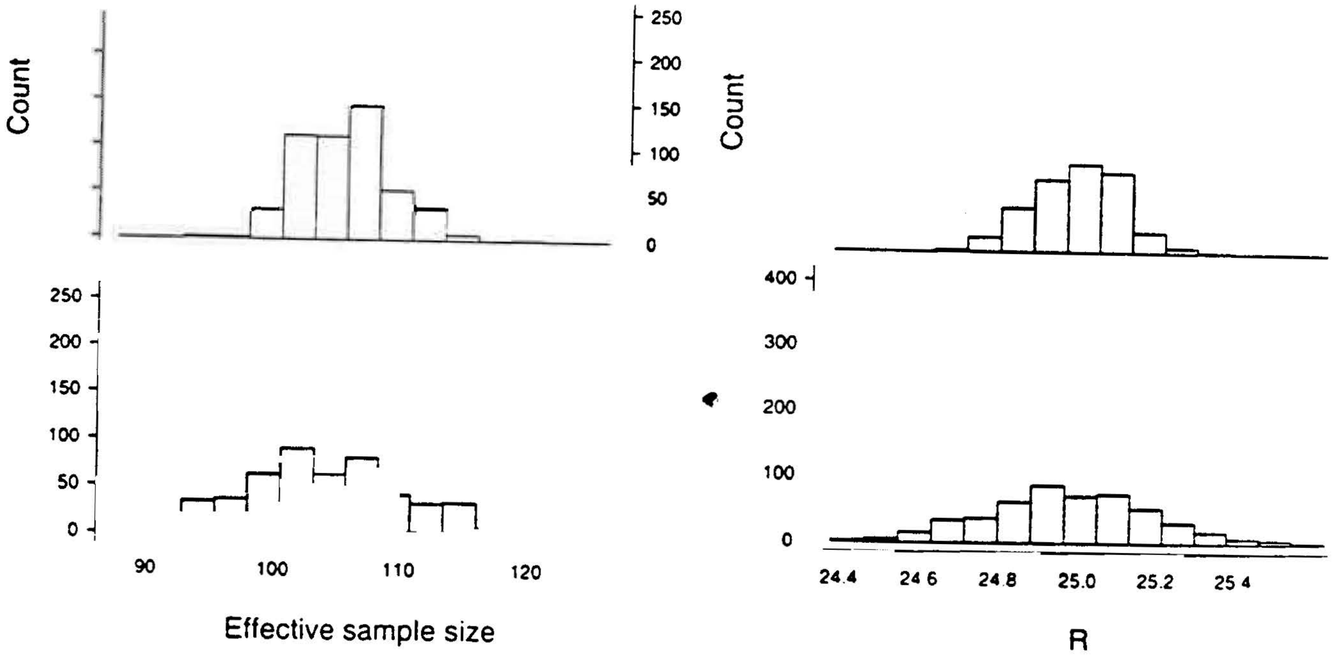

Fig. 2. Simulated estimates of the distribution of the effective sample size, $\hat{m}_{eff}$, and of the mean length, $\hat{R}$, when the total number of fish measured is reduced for the 1999 winter survey in the Barents Sea. The top panel is when a maximum of k = 100 fish are selected per subsample for a total of m = 17,615 fish in each run; the middle panel, k = 30, M = 7,240; and the bottom panel, k = 10, m = 2,597. The estimate of the population mean, R, based on the entire sample (m = 21,769) is 24.96 and its 95% confidence interval is (22.26, 27.66).

(a) (b)

Fig. 3. Bootstrapped estimates of the 95% confidence intervals for each 5 cm length bin for the length frequency distribution of cod in the Barents Sea in winter 1995 (a) and in winter 1999 (b). The inner brackets denote the confidence intervals if the estimates are based on all the cod measured during the surveys and the outer brackets, if 10 fish were measured per subsample.