RESTRICTED

International Commission for the Northwest Atlantic Fisheries

Serial No. 3853
(E)

ICNAF Res.Doc. 75/11

ANNUAL MEETING - JUNE 1975

Fish Market Sampling in Scotland[1]

by

Christine Johnston, M.D. Nicholson and J.A. Pope
Marine Laboratory, Aberdeen, Scotland

Abstract

This paper contains a brief description of the aims and procedures of fish market sampling in Scotland, with emphasis on sampling at Aberdeen Market. The authors present formulae for an estimate and its variance of the total landings of fish in any specified size range. The results are given for a special sampling study designed to provide estimates and variances of the total numbers of haddock in ten centimetre groupings for trawlers fishing on Faroe Plateau and landing in Aberdeen. These results are discussed in relation to the recommended arbitrary minimum level of 200 fish measured per 1,000 tons landed.

Introduction

The statistical information required for a full understanding of the population dynamics of exploited fish stocks includes reliable data on the amount of fish removed by fishermen together with the associated fishing effort and the size and age distribution of the individuals in the stock. In Scotland a complete record of all landings by commercial fishing vessels and the amount of fishing effort expended is maintained (Pope and Saville 1973). The size and age compositions of the landings themselves have to be estimated by sampling.

Regular sampling of landings is carried on at a number of ports in Scotland, being most intensive at Aberdeen. Here landings are sampled by members of the scientific staff from the nearby Marine Laboratory, currently for cod, haddock, whiting, saithe, hake, lemon sole, plaice and megrim. In this paper attention will be confined to haddock sampling on Aberdeen fish market. The sampling procedures used at Aberdeen are similar to those at other Scottish ports.

The objective of this sampling is to provide, for each of a number of defined areas, estimates of the absolute number of haddock of different lengths and ages landed in a month by commercial fishermen. From these, using information on fishing effort, indices of the abundance of the exploited components of the stock are determined. The areas used are groups of ICES statistical rectangles, chosen so as to reflect regions of different growth rate in the haddock populations exploited and also the pattern of Scottish fishing activity. A chart showing these areas is to be found in Statistical News Letters, No.61, Demersal Species 1972.

Samples are taken from landings by trawlers, Danish seiners, light trawlers and Nephrops trawlers, these being treated separately. A minimum sampling intensity at Aberdeen of four vessels per gear type, per area, per month is aimed at although, for various reasons, the actual number of boats sampled in any month may differ from this. In 1973 a total of 218 600 haddock were measured, distributed among the various fishing regions as shown in Table 1.

---

[1] Presented to the 1974 Annual Meeting of ICES as C.M. 1974/F:39

Table 1 Number of Haddock Measured and Total Weight of Fish Caught by Areas in 1973

Area	Number of fish measured				Total weight landed (m. tons)				
	Motor trawl	Light trawl	Seine net	Nephrops trawl	Motor trawl	Light trawl	Seine net	Nephrops trawl	Total
Northern	23 739	4 082	19 115	-	10 432	1 701	10 486	-	22 619
North Central	1 779	728	17 967	-	761	664	17 235	-	18 660
Central	539	-	12 360	-	508	-	5 237	-	5 745
Moray Firth	-	1 807	13 217	3 327	-	506	5 153	97	5 756
Eastern & Orkney	24 296	6 068	21 239	-	8 554	4 741	9 342	13	22 650
North Coast	12 265	-	-	-	3 631	-	640	-	4 271
Butt of Lewis	7 025	-	723	-	2 417	-	665	-	3 082
Inner Hebrides	-	1 133	2 696	919	-	313	1 047	78	1 438
Outer Hebrides	5 700	-	1 243	-	4 364	-	10	-	4 374
South Minch	-	400	3 431	1 198	1 024	38	805	119	1 986
Clyde	-	440	4 676	942	-	145	214	64	423
Faroe Bank	7 983	-	-	-	739	-	-	-	739
Faroe Plateau	14 067	-	-	-	3 252	-	-	-	3 252
Iceland	3 489	-	-	-	224	-	-	-	224

Sampling Procedures

Landings of haddock at Aberdeen fish market are stratified for marketing purposes into a number of size (length) categories termed, for example, small, medium, large, extra large. These are of a variable nature and their number and composition are to a large extent determined by the size distribution of fish in the catch. Thus, fish of say the medium category in one landing may have a significantly different size composition from fish of the medium category in another landing. Because of these differences in strata between vessels it is necessary to sample each stratum within each selected vessel. On discharge, fish are placed in boxes of identical size, the weight of fish in each box being nominally the same (1 cwt = 50.8 kg). Slight variations in weight from box to box do, of course, occur but this is ignored in the treatment of sample data.

The selection of a landing for sampling on any day is determined by the requirements of the area quota at the time and the position of the landing on the market. Discharging of catches usually begins about 3.30 am with sales (by auction) starting at one end of the market at 7.30 am and proceeding sequentially along the market. This usually requires one to two hours depending on the volume of fish exposed for sale on the day. Sampling commences sometime between 6.00 am and 6.30 am and usually lasts for about two hours. As landings are removed from the market almost immediately on being sold it is essential to choose vessels as far as possible from the starting point of sales to allow as much time as possible for measuring fish and taking otoliths.

These requirements preclude the selection of vessels for sampling by any random procedure but this is not likely to cause bias since the arrival of vessels and hence their position of unloading are completely unsystematic.

Estimation Procedures

Sampling is carried out in two stages, the first being the selection of vessels (primary units) and the second the selection of boxes (secondary units) within selected primary units. As already mentioned, sampling at the second stage is stratified with size categories forming the strata. Normally only one box is selected from each stratum of each chosen vessel although the results described below refer to a special study in which three boxes were chosen from each stratum.

The following notation will be used:

subscripts: i refers to vessel
j refers to stratum (category)
k refers to box

observations: y_{ijk} refers to the observed value for box k in category j of vessel i (e.g. y_{ijk} might be the total number of fish in the box)

Item	Population Value	Sample Value
Number of vessels	N	n
Number of size categories in each vessel	s	s
Number of boxes in category j of vessel i	M_{ij}	m_{ij}
Number of boxes in vessel i	$\sum_{j}^{s} M_{ij} = M_i$	$\sum_{i}^{s} m_{ij} = m_i$
Overall number of boxes	$\sum_{i}^{N} M_i = M$	$\sum_{i}^{n} m_i = m$
Mean number of fish in category j of vessel i	$\sum_{k} y_{ijk}/M_{ij} = \bar{Y}_{ij}$	$\sum_{k} y_{ijk}/m_{ij} = \bar{y}_{ij}$
Mean number of fish per box in vessel i	$\sum_{j}^{s} M_{ij}\bar{Y}_{ij}/M_i = \bar{Y}_i$	$\sum_{j}^{s} M_{ij}\bar{y}_{ij}/M_i = \bar{y}_i$
Overall mean number of fish per box	$\sum_{i}^{N} M_i\bar{Y}_i/M = \bar{\bar{Y}}$	$\sum_{i}^{n} M_i\bar{y}_i/\sum_{i}^{n} M_i = \bar{\bar{y}}$

As all boxes are assumed to hold the same weight of fish, W, the weight of fish landed by vessel i is $\sum_{j}^{s} WM_{ij} = WM_i$ and the total weight landed by all vessels in the population is $\sum_{i}^{N} WM_i = WM$. For those vessels in the sample the values of M_{ij} are known, but for vessels not in the sample only the values of M_i (and hence WM_i) are available.

The estimated total for y_{ijk} (i.e. the estimate of $M\bar{\bar{Y}}$) is obtained in two stages. First the totals for each sampled vessel are found and then, using these, the total for all vessels is obtained.

Since sampling at the second stage is stratified the estimate of the total for vessel i is

$$\sum_{j}^{s} M_{ij}\bar{y}_{ij} = M_i\bar{y}_i \qquad (1)$$

The estimated total for the population is then given by summing the values of (1) over all sampled boats and raising by the ratio of the number of boxes in the population to the number of boxes landed by the sampled boats, i.e. by

$$M\left(\sum_{i}^{n} M_i \bar{y}_i\right) \Big/ \sum_{i}^{n} M_i = M\bar{\bar{y}} \qquad (2)$$

Sampling error in this quantity can arise from two sources, one due to variability within sampled boats and the other due to variability between boats. It is shown in the Appendix that the variance of the estimate (2) is given by

$$\mathrm{var}(M\bar{\bar{y}}) = \frac{N}{n}\left\{ \sum_{ij}^{Ns} M_{ij}(M_{ij} - m_{ij}) \frac{S_{ij}^2}{m_{ij}} + \sum_{i}^{N} M_i^2 (\bar{y}_i - \bar{\bar{y}})^2 \right\} \qquad (3)$$

This quantity may be estimated by

$$s^2(M\bar{\bar{y}}) = \frac{N}{n}\left\{ \sum_{ij}^{ns} M_{ij}(M_{ij} - m_{ij}) \frac{s_{ij}^2}{m_{ij}} + \sum_{i}^{n} M_i^2 (\bar{y}_i - \bar{\bar{y}})^2 \right\} \qquad (4)$$

where S_{ij}^2 is the true within-stratum variance in category j of vessel i and s_{ij}^2 its estimate.

Sampling Study

In order to test the efficiency of the sampling system used at Aberdeen a special study was made in July 1974. Trawlers which had been fishing on Faroe Plateau during the month were taken as the population. A total of 54 landings from this area was made during the month, the total quantity of haddock landed being 18 091 cwt (= boxes). Five vessels were sampled, these having landed 178, 270, 614, 268 and 243 cwt, a total of 1 573 cwt or just under 10% of the total weight landed for the month. All landings were stratified into four size categories and three boxes of fish were measured in each category. This permitted the values of s_{ij}^2 (each based on 2 degrees of freedom) to be evaluated for every sampled category and hence the value of $s^2(M\bar{\bar{y}})$.

For the purpose of this study the y_{ijk} were taken successively to be the number of fish in a box in the following length groups, 30-39 cm, 40-49 cm, 50-59 cm and 60-69 cm. The estimates of the total number of fish in each of these size groups landed during the whole month of July, obtained by means of expression (2) above and their standard errors, as given by the square-root of expression (4), are shown in Table 2. (Estimates and standard errors are quoted to the nearest 100 fish.)

Table 2 Estimated Total Number of Fish Landed in Specified Length Groups and Standard Errors

Length Group (cm)	Estimated Total	Standard Error
30-39	1 155 900	39 100
40-49	222 200	6 200
50-59	123 400	6 000
60-69	64 600	3 300

These results show that, with the level of sampling used in this study, a standard error of 5% or less is achieved in the estimates. Since the contribution to $s^2(M\bar{\bar{y}})$ arising from between-vessel variation was in every case much higher than the contribution due to within-vessel variation, the percen-

tage standard error will increase only very slightly as the number of boxes sampled per category is reduced from 3 to 1, the usual number selected. The percentage standard error attaching to estimates of numbers in smaller length intervals may be expected to be somewhat higher on average, although this has not so far been investigated.

The total number of fish measured in this study was of the order of 3 900, equivalent to about 4 300 per 1 000 m tons landed. This is considerably more than the figure of 200 fish per 1 000 m tons put forward tentatively as a minimum requirement for commercial fish sampling. The results of this study clearly indicate that in order to achieve even the moderate accuracy of 5%, 200 fish measured per 1 000 m tons landed could be grossly inadequate. As the situation studied here relates to the sampling of landings made by a relatively homogeneous fleet of trawlers operating in a small area it is to be expected that this level of accuracy will not be achieved when combining samples from landings by a wider variety of vessels and gears fishing over larger areas. The number of fish required to achieve an acceptable standard error will vary not only according to the quantity being estimated but also according to the way sampling effort is distributed between and within boats.

The objective of Scottish haddock sampling is the production of estimates of the numbers and mean lengths of the different age groups in the landings. This introduces an additional stage into the estimation process, namely the employment of age-length keys. These keys, being based on samples, are subject to sampling error which will add a further component to the variance discussed in this paper. The sampling properties of age-length keys are currently being studied. Some results of this work are given by Nicholson and Armstrong (1974) and further results will be presented later.

Acknowledgement

The authors wish to express their thanks to Mr A.S. Jermyn and Mr D.C. Emslie who undertook all the sampling in the special study described here.

References

Nicholson M.D. & Armstrong D.W.	1974	A note on a possible source of bias in the estimation of mean length at age. ICES CM 1974 Demersal Fish (Nthn) Cttee, Doc. no. F:35.
Pope J.A. & Saville A.	1973	Scottish fisheries statistics. ICES CM 1973, Statistics Cttee, Doc. no. D:2.

APPENDIX. DERIVATION OF VARIANCE FORMULA

The estimate of the population total of y_{ijk} is $M\bar{\bar{y}}$. Evaluation of $\text{var}(M\bar{\bar{y}})$ is not straightforward as $\bar{\bar{y}}$ is a ratio estimate.

We first prove the following lemmas.

Lemma 1 $\bar{\bar{y}}$ is an almost unbiased estimate of $\bar{\bar{Y}}$.

Proof: $M_i\bar{y}_i$ is an unbiased estimate of the total for the i^{th} primary unit. Hence $NM_i\bar{y}_i$ is an unbiased estimate of the total for all N primary units, namely $M\bar{\bar{Y}}$. It follows that $N\sum_i^n M_i\bar{y}_i/n$ is an unbiased estimate of $M\bar{\bar{Y}}$.

Further, $\sum_1^n M_i/n$ is an unbiased estimate of $\sum_1^N M_i/N$ so $\sum_1^N M_i/\sum_1^n M_i$ is almost equal to N/n. Thus, $M(\sum_1^n M_i\bar{y}_i/\sum_1^n M_i)$ is almost equal to $N\Sigma M_i\bar{y}_i/n$ which is an unbiased estimate of $M\bar{\bar{Y}}$.

Lemma 2 $(\bar{\bar{y}} - \bar{\bar{Y}})$ is almost equal to $N\sum_1^n M_i(\bar{y}_i - \bar{\bar{Y}})/nM$.

Proof: $\bar{\bar{y}} - \bar{\bar{Y}} = (\sum_1^n M_i\bar{y}_i - \sum_1^n M_i\bar{\bar{Y}})/\sum_1^n M_i = \sum_1^n M_i(\bar{y}_i - \bar{\bar{Y}})/\sum_1^n M_i$.

The result follows on replacing $\sum_1^n M_i$ by Mn/N.

Theorem The variance of $M\bar{\bar{y}}$ is almost equal to

$$N\left\{ \sum_{ij}^{Ns} M_{ij}(M_{ij} - m_{ij})S_{ij}^2/m_{ij} + \sum_1^N M_i^2(\bar{Y}_i - \bar{\bar{Y}})^2 \right\} / nM^2$$

Proof:
$$\begin{aligned} \mathrm{var}(M\bar{\bar{y}}) &= M^2\mathrm{var}(\bar{\bar{y}}) \\ &= M^2E(\bar{\bar{y}} - \bar{\bar{Y}})^2 && \text{(by lemma 1)} \\ &= N^2E\{ \Sigma M_i(\bar{y}_i - \bar{\bar{Y}}) \}^2 / n^2 && \text{(by lemma 2)} \\ &= N^2E\{ M_i(\bar{y}_i - \bar{\bar{Y}}) \}^2 / n \end{aligned}$$

where E stands for expectation over all possible samples.

Now

$$M_i(\bar{y}_i - \bar{\bar{Y}}) = M_i(\bar{y}_i - \bar{Y}_i) + M_i(\bar{Y}_i - \bar{\bar{Y}})$$

and so

$$E\{ M_i(\bar{y}_i - \bar{\bar{Y}}) \}^2 = E\{ M_i^2(\bar{y}_i - \bar{Y}_i)^2 + M_i^2(\bar{Y}_i - \bar{\bar{Y}})^2 + 2M_i^2(\bar{y}_i - \bar{Y}_i)(\bar{Y}_i - \bar{\bar{Y}}) \}$$

The expression on the righthand side may be evaluated in two stages. First the expectation over all possible samples for i fixed is found. Only the first term on the right is involved and, since $\bar{y}_i$ is the mean of a stratified random sample, this term has expectation

$$\sum_j^s M_{ij}(M_{ij} - m_{ij})S_{ij}^2/m_{ij}$$

where $S_{ij}^2 = \Sigma(y_{ijk} - y_{ij})^2/(M_{ij} - 1)$

Following this by taking expectation over all values of i gives

$$E\{ M_i(\bar{y}_i - \bar{\bar{Y}}) \}^2 = \frac{1}{N}\left\{ \sum_{ij}^{Ns} M_{ij}(M_{ij} - m_{ij})S_{ij}^2/m_{ij} + \sum_i^N (\bar{Y}_i - \bar{\bar{Y}})^2 \right\}$$

since the expectation of the cross-product term is zero. The required result follows.