THIS PAPER NOT TO BE CITED WITHOUT PRIOR REFERENCE TO THE AUTHORS

International Council for the
Exploration of the Sea

CM 1977/D: 2
Statistics Committee

THÜNEN
Digitalization sponsored
by Thünen-Institut

BIAS IN A SAMPLING PROGRAMME

M J Holden and T Williams
MAFF, Fisheries Laboratory, Lowestoft, NR33 OHT, England

Most of the sampling programmes from which length composition data of demersal species are obtained for stock assessment at the Fisheries Laboratory, Lowestoft use port as a sampling stratum. The programmes are based on the hypothesis that positions of fishing are not correlated with either the day of landing or the position at which the fishing vessels berth to unload and that therefore the number of landings sampled from each statistical rectangle is proportional to the number of rectangles which are fished.

This paper deals with sampling at the port of Lowestoft where a specific number of samples of selected species is required weekly. The measuring team has free choice as to when these samples are measured but, as most teams aim to complete the week's programme in one day, the preferred sampling days are Mondays and Wednesdays when the number of boats landing is usually higher than on the other days of the week.

MATERIAL

Up to 10 trawlers land at Lowestoft each day from Monday to Friday. The mate of each trawler over 12.2 m (> 40') landing at Lowestoft is interviewed and the statistical rectangle in which the vessel fished is recorded. Occasionally, trawlers fish in several rectangles but in these instances, for the purpose of this analysis, only the statistical rectangle covering the main fishing ground has been used.

A sample of 425 landings, covering the months March, July-August and October-November 1976, was analysed. Of these 56 were sampled to obtain measurements of plaice, the only species considered in this paper.

RESULTS

Occurrence of bias: The data were analysed to test the hypothesis that the number of samples taken was proportional to the number of landings from each statistical rectangle. In order to satisfy the requirements for chi-squared analysis that no cell size should be less than 5, groups of statistical rectangles had to be used for both this and succeeding analyses. This in fact corresponds to the realities of sampling because the number of samples is so few that they cannot be directly proportional to the number of landings from each statistical rectangle. The results are shown in Table 1; the observations differ significantly from those expected on the basis of the hypothesis. Sampling is

positively biassed towards landings from north of 55°30'N but even within groups of rectangles within which, overall the observed number of samples is close to that expected, detailed inspection reveals bias; for example, as in group E (Figure 1) where 9 of the 10 samples were taken from vessels fishing one statistical rectangle from which only 35 of a total of 56 landings were made.

Origin of bias: The following two hypotheses were tested:

1 that for each group of statistical rectangles, the number of landings on any day of the week was proportional to the number of days in the sampling period.

2 that for each group of statistical rectangles, the number of landings at each position on the quay was proportional to the number of landings at that position in the sampling period.

The results of the first analysis, which are given in Table 2, show that the data do not depart significantly from those expected on the basis of the hypothesis. The results from the second analysis (Table 3) show that the data do depart significantly from those expected on the basis of the hypothesis. Four of the cells contribute 61% of the total chi-squared distribution; these cells refer to landings at position 1 from statistical rectangles north of 56°N and east of 4°E (more than expected) and at positions 5+ from all statistical rectangles north of 56°N (fewer than expected) and those south of 56°N and west of 4°E (more than expected).

Bias therefore originates in the landing positions on the market. It becomes fed into the sampling system because the vessels are unloaded sequentially along the market, starting with position 1. Landings from vessels in positions 1 to 4 are immediately available to the measuring team and time spent working between ships is minimised by measuring from ships that are close together. Furthermore, landings from vessels at position 5 onwards may be continuing after sales have started and it helps to maintain good working relations if the measuring team is off the market by the time that buyers arrive. Of the 56 samples of plaice, 49 were taken from vessels that landed at positions 1-4 and the remaining 7 were taken from positions 5 and 6.

DISCUSSION

The original sampling programme was set up by Gulland and formed the basis for his manual on sampling (Gulland, 1955). He considered origins of bias and it can be assumed that at that time this type of bias did not exist. In 1955 the fleet was relatively homogeneous but it can now be divided approximately into 2 sectors, one composed of large, powerful trawlers that fish mainly north of 56°N and east of 4°E and smaller, less powerful trawlers that fish south of 56°N. In 1955 position of landing was determined by time of arrival but it is now determined by time of sailing and it can only be assumed that, to obtain what are considered the prime selling positions of the first berths on the market, the larger trawlers sail earlier than the smaller ones. Crews are presumably prepared to do this both to

retain their jobs aboard the top earning vessels and to ensure that their vessel achieves one of the earliest selling berths.

That the bias remained undetected resulted from a lack of a continuous re- appraisal system. On the basis of this analysis samples would have been taken in direct proportion to the landings from each group of statistical rectangles if 2 ships had been sampled from positions 5+ for every one ship sampled at positions 1 to 4. Sampling on this basis has now been implemented.

ACKNOWLEDGEMENTS

The existence of bias within the system was originally detected in an analysis by Mr J Bridger.

REFERENCE

Gulland, J. A. 1955. Estimation of growth and mortality in commercial fish populations. Fishery Invest., Lond., Ser. II, 18 (9), 46 pp.

Table 1 Observed and expected number of samples from each of the groups of statistical rectangles shown in Figure 1

Group	A	B	C	D	E	F	Total
Observed	12	12	8	8	10	6	56
Expected	7.91	5.93	9.09	9.36	7.38	16.33	56

Chi-squared = 13.50 (continuity correction made)

$P < 0.02$ for 5 d.f.

Table 2 Observed and expected number of landings on each day of the week from 4 areas of the North Sea

			M	Tu	W	Th	F
North of 56°N	West of 4°E	Observed	10	10	10	6	8
		Expected	12.01	7.35	9.32	8.59	6.73
	East of 4°E	Observed	39	13	29	19	12
		Expected	30.57	18.71	23.72	21.87	17.13
South of 56°N	West of 4°E	Observed	46	33	38	39	33
		Expected	51.59	31.57	40.02	36.91	28.91
	East of 4°E	Observed	21	15	13	19	12
		Expected	21.84	13.36	16.94	15.62	12.24

Chi-squared = 9.88

P > 0.95 for 19 d.f.

Table 3 Observed and expected number of landings at each position on the market from 4 areas of the North Sea

			1	2	3	4	5+
North of 56°N	West of 4°E	Observed	10	11	6	11	6
		Expected	7.45	7.45	7.25	7.04	14.81
	East of 4°E	Observed	29	23	22	12	26
		Expected	18.97	18.97	18.45	17.92	37.69
South of 56°N	West of 4°E	Observed	20	25	30	32	82
		Expected	32.02	32.02	31.13	30.24	63.59
	East of 4°E	Observed	13	13	12	13	29
		Expected	13.55	13.55	13.18	12.80	26.92

Chi-squared = 29.88 (continuity correction made)

P < 0.05 for 19 d.f.

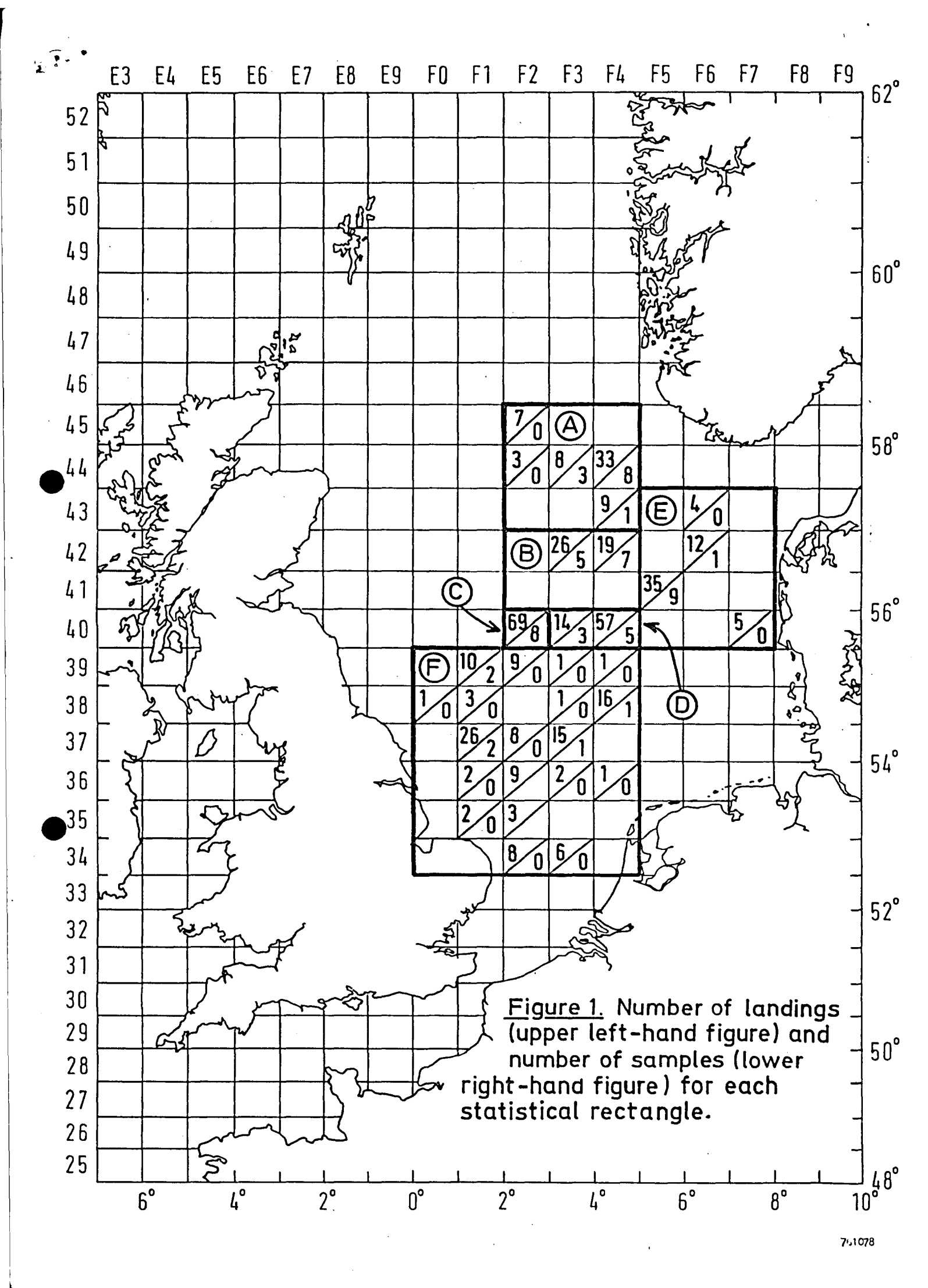

Figure 1. Number of landings (upper left-hand figure) and number of samples (lower right-hand figure) for each statistical rectangle.