International Commission for the Northwest Atlantic Fisheries

Serial No. 3922
(E)

ICNAF Res.Doc. 76/VI/100

ANNUAL MEETING - JUNE 1976

Report on the Special Sampling Project for Cod in Div. 2J & 3KL, Mackerel in Subareas 3-5 & Statistical Area 6, and Silver Hake in Subareas 4-5 & Statistical Area 6

by
W.G. Doubleday
Computing and Applied Statistics Directorate
Environment Canada

Editorial Note

This document, containing 5 pages of text, 38 pages of computer-printed tables and 92 pages of Maps, was issued in a limited edition at the 1976 Annual Meeting. Only the text of this document is being issued now for more general distribution. Copies of the Tables and Figures are on file at the Secretariat and a copy of these may be obtained upon request.

Introduction:

STACRES, at the 1974 annual meeting, recommended that individual length and age samples of Cod in Div. 2J & 3KL, Mackerel in SA 3 to 6, and Silver Hake in SA 4,5, and 6 for the years 1972 and 1973 be submitted by member countries and examined as a pilot study for data base improvement. Response to the recommendation was slow and too little data had been received prior to the 1975 annual meeting to justify analysis. At the 1975 annual meeting, the Statistics and Sampling subcommittee emphasized the importance of the study and extended its time scale to include 1974.

Member nations responded to the 1975 request with a large volume of samples in December 1975 and January 1976. The samples were keypunched at the ICNAF secretariat by Miss G. Langille under the supervision of Mr. S.A. Aikenhead.

Most of the samples were received by the author immediately prior to the April 1976 meeting of ICNAF, with some samples arriving as late as May 18.

Due to the shortage of time and limitations inherent in the data, this analysis does not claim to be comprehensive or definitive, but hopefully it will shed some light on the major problems.

Issues:

The following questions which can be asked of any catch sampling scheme isolate certain aspects of sampling methodology. In later sections, the ability of the special sampling project data to answer these questions is discussed.

1. Is the technique of selecting fish for a sample biased or otherwise unrepresentative?

2. Are the samples distributed evenly through the total catch of the fleet(s)?

3. Are sufficient samples collected to describe the total catch with adequate precision?

4. Is the stratification by location, time and gear adequate? Are major improvements possible?

5. Is the age reading technique adequate?

6. Are the assumptions underlying the use of age-length keys met?

7. Can improvements be made in data processing?

Limitations:

Keypunching errors were evident in several samples. Verification of all samples is recommended prior to further anlaysis or revision of estimated age compositions for these stocks. Editing the data for consistency would not detect all errors.

Difficulties arise when attempts are made to apply statistical methods in the analysis of the sampling study data. Firstly, there is no replication of sampling from a single haul or landing so that fine comparisons of samples from separate hauls or landings using a measure of pure sampling error in single hauls or landings are impossible. Secondly, the samples are spread unevenly through space and time for a given gear so that the effects of location and date are confounded and the "experimental design" of the data is highly unbalanced.

It is felt that a formal statistical analysis which could only distinguish between broad categories such as gears would not be useful to the committee and that on a finer scale differences due to date and location could not be distinguished. Consequently, quantitative analysis has been restricted to selected cluster analyses by principal components of length frequency samples.

It must be realized that, while data collected for administrative purposes can have real scientific worth, its analytic value is not comparable to that of data arising from a planned experiment.

Sampling Technique:

While sample sizes for mackerel are near 50, 100 and 200 fish, and silver hake samples are near 100 and 200 fish, cod samples vary from 15 to several hundred. Some countries use stratified subsampling of length groups while others do not. In one case, a cumulative stratified sampling scheme appears to be in effect with subsampling limited to the relatively empty rows of an age length key towards the end of a quarter.

Many of the samples were taken by research vessels. Contrasts between research and commercial samples are inconclusive in the few cases where samples of both types from the same gear, month, and division combination occur. Otter trawl samples of mackerel (research) were more variable than midwater trawl samples (commercial) and contained more 0 group and ages 1 & 2 fish.

The detection of biases in the selection of fish for a sample is impossible with the available data. Specially designed studies are necessary for this purpose. One general remark can be made, however. As the number of fish in a sample decrease choice in selection and consequent bias towards larger fish becomes increasingly difficult to avoid. Small samples of cod of the order of 20 fish could be quite biased.

Distribution of the Samples in Space and Time:

Tables 1 and 2 list the age and length samples respectively received by the author. The following samples were received by the Secretariat but not punched Canada (M) mackerel (no latitude + longitude), U.K. cod (by market categories without weighting factors), Denmark cod (one sample with no latitude or longitude).

A series of maps shows the locations of monthly samples by country. Samples taken at the same latitude and longitude in the same month are represented by a single letter. The small number of Canada (N) and U.S.A. samples is due to the repetition of identical coordinates. Some samples with impossible coordinates are plotted due to coding or punching errors. Samples of all types are combined.

There is a very evident clustering of samples by location, date and gear simultaneously. To some extent this follows the pattern of the various fisheries and it would be necessary to study catch data on the same scale as the sampling data to determine whether the sampling is adequately dispersed. What information there is in the samples suggests that samples are taken in bunches with possibly large unsampled gaps in the intervening catches.

The large variation in numbers of samples for a gear - month - division category renders the otherwise attractive use of multivariate analysis of variance ineffective.

Number of Samples:

Although detailed catch data were not at hand to calculate sampling rates by gears in segments of division and month, the overall sampling rate appears to be less than the ICNAF minimum of 100 fish per 1000 tons.

Calculations of sampling variance ordinarily assume that the samples are spread through the catch at random. There is some evidence that this is not true and that bunching of samples has occurred. This means that the sample variance is higher (if appropriate weights are used for different segments of the total landings) than would occur if the samples were more evenly dispersed. Biases in estimation can occur if variation in the sampling rate through the total catch are not allowed for by the use of weighting factors.

Adequacy of the Stratification:

Currently, reporting of sampling data is by gear by division with three monthly summaries of age-length and monthly length frequency samples. The value of stratifying by gear was considered sufficiently well established as not to merit further consideration so that attention was limited to space and time within gears. There was a tendency for older cod to be caught in Div. 2J than 3K and in 3K than in 3L. Within divisions, location was sometimes nominal and exerted no clear influence on age composition.

Selected length frequency samples for mackerel and silver hake were examined for clustering by plotting on the plane of the first two principal components of their respective covariance matrices. The method was applied to silver hake samples by sex by div. 5ZE in 1973. The first component explained 58.7% of the total variance and the second component a further 25.6%. The first component was mainly a contrast between length classes 26-28 cm and 30-34 cm, while the second contrasted 26 cm with 28-30 cm. Examination of the scatter plot showed that the first component was closely related to sex while the second component tended to oppose the two samples by sex from a single landing and thus related to the sex ratio. This is further evidence for reporting silver hake samples by sex.

The method was then applied to silver hake samples for 1972 from several divisions. In this case the first component, containing length groups 22-26 cm with 28-32 cm, explains 49.0% of the total variance. The second component, contrasting length groups 20-24 cm with 26-28 cm, explains a further 32.4% of the total variance. Samples from a single division appeared together with the exception of those from Div. 5ZW which formed two clusters. The first component appeared to be related to the date with a rapid transition near the beginning of September representing new recruits. Fine location explains some of the scatter within the division clusters but association in time appears to be generally more influential than association in space.

Finally, the method was applied to selected mackerel samples for 1973 for a number of divisions. The first component contrasted the 25-30 cm groups with the 32-37 cm groups and explained 44.8% of the total variance. The second component contrasted the 31-33 cm groups with the 35-38 cm groups. The first two components related well with month such that May-June samples appeared mainly in one area, August samples in another and September samples in a third etc. However, despite rapid transitions within a month, samples from each division tended to form seasonal clusters some of which (especially Div. 4X and Div. 5ZE) were clear and separate from other samples.

The principal component analyses are not conclusive, but do lead to some observations. Firstly, the length frequency samples are largely 2-dimensional with 70-80% of the total variance explained by the first two principal components. Secondly, the first two components in all cases essentially represented contrasts of adjacent blocks of length groups which suggests a close relationship with the growth and abundance of two or three dominant age classes. Thirdly samples from within the same division tended to form clusters corresponding to seasons of the year, often with rapid transitions corresponding to new recruitment. Time appears to be more important than space as a stratifying variable on a fine scale.

Due to the approach of the deadline for reporting, further cluster analyses were not attempted. The method appears to be promising in its ability to find structure in the somewhat scattered data and could be refined by analysing more homogeneous sets of samples by fixing the month and varying location or fixing a location and varying the month of sampling.

Ageing:

Large differences between age length keys for cod in Div. 2J between countries were observed. The differences were greatest for the oldest fish. Differences are also visible between countries for mackerel samples in the same month, although in this case different divisions are involved.

For older cod the age-length keys are very scattered as is usual. In view of the substantial disagreement concerning the interpretation of otoliths at ageing workshops (ICNAF Summ. Doc. 76/V1/13), this author suggests that as a working hypothesis it be assumed that the scatter is mainly due to uncertanties in ageing and not to inherent variability in length at age for older fish.

There are some inconsistencies in mackerel age length keys with, for example, two year old fish sometimes having a mode near 30 cm in January and near 26 cm in April of the same year.

Differences are evident in mean length at age of even two year old mackerel between countries in the same month. These differences may be related to location of capture but if there is only one stock involved, they may be due to differences in age reading technique.

Age-length Keys:

Age length keys from separate hauls or landings are ordinarily pooled without weighting under the assumption that for a given length and catches by a single gear over a limited time period the distribution of ages is essentially constant. Failure of this assumption is not critical if the size of a landing or haul is unrelated to its age composition, but otherwise can introduce biases into the estimates of age composition of the commercial catch.

In the case of 0 and 1 group mackerel in particular year classes can appear and disappear in samples separated by a few days and miles. These year classes tend to come and go as a whole with large changes in the relevant portion of the age-length keys. If sampling rates differ between catches with these groups represented and those without, serious biases in estimates of numbers caught in these age groups result.

For the silver hake and mackerel fisheries, the numbers of new recruits caught have a considerable influence on recommendations for regulation so that it is advisable to verify that sampling rates do not vary with catch composition. This can only be done by comparison of catch and sampling data on a small scale but does not require a specially designed experiment.

Data Processing:

The submission of individual samples to the Secretariat has definite advantages for data processing over processing by individual countries. Firstly, the combined volume of sampling data would justify the full time employment of skilled workers. Secondly, uniform, high standards of keypunching and editing could be developed and enforced. Thirdly, designated experts would be better able to apply the appropriate weighting to samples in calculating age compositions of catches with detailed sampling data, provided these were accompanied by catch in weight data on the same scale. Fourthly, studies on data base improvement would be facilitated.

Conclusions:

Within the limitations of the data it is possible to conclude that monthly age-length keys of silver hake and mackerel are desirable but that reporting of samples by 1^{o} or $30'$ squares is of less value and with the present bunching of samples would leave gaps in many cells. Catch statistics must be submitted on the same scale as sampling data. The fine spatial recording is unlikely to produce substantial increases in precision of estimation of age composition unless a national fleet has two or more components aimed at different age groups of the same stock.

Several suggestions have been made regarding possible areas of improvement of sampling methods. Of these, the dispersal of sampling evenly through the catch and the ageing of the fish appear to be the most serious problems.

Submission of individual samples to the Secretariat has advantages of scale for data processing and could be useful for designated experts in estimating catch compositions and in improving sampling methods.