

On estimating the age composition of the commercial catch of Northeast Arctic cod from a sample of clusters

Sondre Aanes and Michael Pennington

Aanes, S., and Pennington, M. 2003. On estimating the age composition of the commercial catch of Northeast Arctic cod from a sample of clusters. — ICES Journal of Marine Science, 60: 297–303.

Assessment of Northeast Arctic cod is based on estimates of the commercial catch in numbers at age. The age structure of the catch is estimated by sampling fish from commercial fishing trips. Although it is commonly assumed that a sample of individuals is a random sample from the population, fish sampled from the same trip (i.e. from a “cluster” of fish) tend to be more similar in age than those in the total catch. For Northeast Arctic cod, the intracluster correlation for age is positive, and therefore the effective sample size is much smaller than the number of fish aged. Given the number of fish aged, the precision of the estimated age distribution is rather low, and the number of fish aged from each trip could be reduced from approximately 85 to 20 without a significant loss in precision.

© 2003 International Council for the Exploration of the Sea. Published by Elsevier Science Ltd. All rights reserved.

Keywords: age distribution, cluster sampling, effective sample size, intracluster correlation.

Received 9 October 2002; accepted 18 December 2002.

S. Aanes, and M. Pennington: Institute of Marine Research, PO Box 1870 Nordnes, N-5817 Bergen, Norway; e-mail: michael.pennington@imr.no. Correspondence to S. Aanes: tel.: +47 55 238627; fax: +47 55 238687; e-mail: sondre.aanes@imr.no.

Introduction

The fishery for Northeast Arctic cod (*Gadus morhua*) is the largest commercial cod fishery in the world. The total catch of cod peaked in the 1950s with an average annual landing of 800 000 t (Nakken, 1994). The total catch has declined since the 1970s as a consequence of decreasing stock size and quota regulations, which were implemented in the 1970s. Recent landings have ranged from 730 000 t in 1996 to 400 000 t in 2000. Russia and Norway each land approximately half the total catch.

A critical input to fish stock assessments is an estimate of catch at age in numbers (Gudmundsson, 1994; Skagen and Hauge, 2002). This has led to a substantial effort to collect samples from commercial catches to determine the age distribution. However, the precision, the amount of information in the samples, and the sampling scheme for Northeast Arctic cod have not been examined, and therefore the uncertainty in the assessments is not quantified.

When determining the precision of the estimated age distribution of a fish population on the basis of a sample of age readings, it is usually assumed that the age readings are a random sample from the population (see, for example, Hoenig and Heisey, 1987; Richards *et al.*, 1992; Worthington *et al.*, 1995). Because it is generally impossible to sample a fish population randomly, samples of fish for age determination are taken from a number of clusters, e.g.

from individual trawls or fishing trips. The ages of fish caught together tend to be more similar than those in the entire population (i.e. intracluster correlation is positive), and, therefore, the resulting sample will often contain much less information on the age distribution than an equal number of fish sampled at random. If it is assumed that the sample of fish collected from clusters is a random sample of individuals, then the estimated age distribution will appear to be more precise than it actually is (Pennington and Vølstad, 1994; Pennington *et al.*, 2002).

In this article, we examine the precision of the estimates of the age distribution of the Norwegian commercial catch of Northeast Arctic cod obtained by applying two different estimators. A number of fish were aged from each of an assumed random sample of fishing trips; the sample, therefore, consists of a number of clusters, each from a larger cluster (i.e. the fish caught during a trip). In addition, because age determination of fish is time-consuming, we examined whether the number of fish aged from each trip could be reduced without significantly reducing the precision of the estimates.

Commercial age data

The weight of Northeast Arctic cod caught in various statistical regions of the Barents Sea and adjacent areas by

season and for various types of gear is available from a census, conducted by the Norwegian Directorate of Fisheries, of the commercial catch. This census does not collect information on the age or length composition of the catch. To estimate catch characteristics, such as the catch at age in numbers or weight, the Institute of Marine Research (IMR) collects fish from selected fishing trips. For practical reasons, not all the combinations of regions, seasons and gears are sampled. The sampling design for this survey does not take into account region when sampling the catches.

When a sample from a fishing trip is landed, the weight and the length of each fish are recorded along with the size of the catch taken during the trip. To determine the age of a fish, otoliths are removed and stored, and later placed under a microscope for the growth zones to be counted. Age determination is both time-consuming and difficult; currently, the average handling time for sampling a cod and determining its age is approximately 7–10 min. In 2000, 126 catches were sampled by IMR, and some 85 fish were subsampled from each catch for age determination.

In this work, we assume a random sample of catches (fishing trips) from the cod population in terms of catches (trips) in each of the four quarters of the year. We also assume that the fish were chosen randomly from each catch and that ages were determined without error.

Estimating the age composition

Given a random sample of n clusters (landings) and a random subsample of m_i fish from a total of M_i individual fish in cluster i , then the estimator based on the sampling design

$$\hat{\mu}_1 = \frac{\sum_{i=1}^n M_i \bar{x}_i}{\sum_{i=1}^n M_i} \quad (1)$$

is an approximately unbiased and consistent estimator of the mean age of the population if \bar{x}_i is the average age of the sample of m_i fish from cluster i , or the proportion at age in the population if \bar{x}_i is the estimated proportion of fish of specific age in cluster i (Skinner *et al.*, 1989). This is a weighted average of the \bar{x}_i s, where the cluster sizes are the weights. As both numerator and denominator are random variables this is a ratio-type estimator (Cochran, 1977), and an exact variance formula does not exist. The variance may be approximated using a Taylor expansion of Equation (1) or by resampling techniques, such as non-parametric bootstrapping (Efron, 1983).

An alternative to the design-based estimator is the unweighted average of the \bar{x}_i s:

$$\hat{\mu}_2 = \frac{\sum_{i=1}^n \bar{x}_i}{n} \quad (2)$$

An estimate of the variance of $\hat{\mu}_2$ is given by $\sum_{i=1}^n (\bar{x}_i - \hat{\mu}_2)^2 / (n(n-1))$. Bootstrapping techniques can also

be used to assess the precision of $\hat{\mu}_2$. In general, this unweighted estimator $\hat{\mu}_2$ may be biased and this bias may not decrease with increasing sample size, but if \bar{x}_i and M_i are uncorrelated, then $\hat{\mu}_2$ may be an acceptable estimator (Cochran, 1977).

To illustrate the factors that affect the relative precision of $\hat{\mu}_1$ and $\hat{\mu}_2$, consider the standard random effects model

$$x_{ij} = \mu + A_i + \varepsilon_{ij} \quad (3)$$

where x_{ij} is the measured quantity for individual j in cluster i , μ the overall mean, A_i the random effect for cluster i , and ε_{ij} is the residual error, and A_i and ε_{ij} are independent of cluster size. To simplify the algebra, consider the case where sampling is proportional to cluster size, i.e. $m_i \propto M_i$, or the entire cluster is sampled, i.e. $m_i = M_i$. The samples are then self-weighting and $\hat{\mu}_1$ can be rewritten as

$$\hat{\mu}_1 = \frac{\sum_{i=1}^n m_i \bar{x}_i}{\sum_{i=1}^n m_i} \quad (4)$$

Both $\hat{\mu}_1$ and $\hat{\mu}_2$ are model-unbiased, and it may be shown (Pennington and Vølstad, 1994), given model (3), that

$$\text{Var}(\hat{\mu}_1 | \mathbf{m}) = \frac{\sigma_x^2}{n\bar{m}} \left[1 + \left(\bar{m} - 1 + \frac{s_m^2}{\bar{m}} \right) \rho \right] \quad (5)$$

where \mathbf{m} is a vector of the n sample sizes $\sigma_x^2 = \sigma_A^2 + \sigma_\varepsilon^2$, where σ_A^2 and σ_ε^2 denote the variances of A_i and ε_{ij} , respectively; \bar{m} and s_m^2 are the mean and variance of the m_i s, respectively; and $\rho = \sigma_A^2 / (\sigma_A^2 + \sigma_\varepsilon^2)$ is the (model-based) intraclass correlation coefficient (Searle *et al.*, 1992). The conditional variance of $\hat{\mu}_2$ is given by

$$\text{Var}(\hat{\mu}_2 | \mathbf{m}) = \frac{\sigma_\varepsilon^2}{n^2} \sum_{i=1}^n \frac{1}{m_i} + \frac{\sigma_A^2}{n} \quad (6)$$

To compare the precision of the weighted average $\hat{\mu}_1$ and the unweighted average $\hat{\mu}_2$, Equation (5) can be expressed as

$$\text{Var}(\hat{\mu}_1 | \mathbf{m}) = \frac{\sigma_\varepsilon^2}{n\bar{m}} + \frac{\sigma_A^2}{n} \left(1 + \frac{s_m^2}{\bar{m}^2} \right) \quad (7)$$

If the intraclass correlation coefficient, ρ , is 0, then $\sigma_A^2 = 0$, and it follows from the Cauchy–Schwarz inequality that $\sigma_\varepsilon^2 / (n\bar{m}) \leq \sigma_\varepsilon^2 / n^2 \sum_{i=1}^n 1/m_i$ for all \mathbf{m} with equality only if the sample sizes, m_i , are equal. Therefore, if $\rho = 0$, then for any \mathbf{m} , the variance of $\hat{\mu}_1$ will be less than or equal to the variance of the unweighted estimator $\hat{\mu}_2$.

If $\rho > 0$, then whichever is smaller, $\text{Var}(\hat{\mu}_1)$ or $\text{Var}(\hat{\mu}_2)$, will depend on the sizes of the various components in Equations (6) and (7). In particular, it should be noted that the first term in Equation (6) can be considerably larger than the first term in Equation (7) when the samples consist of many small values and a few very large ones, which is often the case for scientific trawl surveys of fish stocks (Pennington *et al.*, 2002). Further, it should be stressed that $\hat{\mu}_2$ may be biased and, hence, its variance can be smaller

than that of $\hat{\mu}_1$, while its mean-squared error is larger. For more details on the possible drawbacks of model-based estimators, see Smith (1990).

To assess the information contained in the data, it is useful to consider the effective sample size, m_{eff} (Kish, 1965; Skinner *et al.*, 1989). The effective sample size is defined as the number of fish that would need to be sampled at random so that the estimates generated by simple random sampling would have had the same precision as the estimates obtained based on a more complex sampling scheme. To estimate the effective sample size, we first estimated the mean age of the catch, $\hat{\mu}_a$, its variance, and the variance of the age distribution of the catch, $\hat{\sigma}_a^2$. Then, the estimated effective sample size, \hat{m}_{eff} , is defined by

$$\frac{\hat{\sigma}_a^2}{\hat{m}_{\text{eff}}} = \text{var}(\hat{\mu}_a) \quad (8)$$

If the effective sample size is small, then this implies that the estimate of the entire age distribution is rather imprecise (Pennington and Vølstad, 1994; Pennington *et al.*, 2002).

As the usual formula for estimating the variance of a ratio estimator, which is based on a Taylor approximation, tends to overestimate the precision, especially for small sample sizes (Cochran, 1977; Pennington and Vølstad, 1994), non-parametric bootstrapping was used to estimate the variance (Efron, 1983). Each replicate was generated by first sampling the trips at random with replacement, then a number of individual fish ages from each selected trip were sampled with replacement. The effect of reducing the number of fish aged on the precision of the estimated age distribution was assessed by reducing the number aged from each trip in the resampling procedure. The intracluster correlation coefficients were estimated as in Searle *et al.* (1992).

Results

The estimate of the intracluster correlation coefficient was significantly greater than 0 in all quarters (Table 1). Because \bar{x}_i did not significantly correlate with M_i , for either average age or proportion at age, both the weighted estimator ($\hat{\mu}_1$, Equation 1), and the unweighted estimator ($\hat{\mu}_2$, Equation 2) were applied. Based on bootstrapped

estimates of the variance, it appeared that $\text{Var}(\hat{\mu}_1)$ is larger than $\text{Var}(\hat{\mu}_2)$ for all quarters except quarter 4 (Table 1).

The effective sample size, which is based on estimates of the average age and its variance (Equation 8), was much smaller than the number of fish aged in each region (Table 1). For example, if it were possible to sample fish at random from the total catch in quarter 2, then it would have been sufficient to determine the ages of slightly more than 200 fish (211 and 213, if applying $\hat{\mu}_1$ and $\hat{\mu}_2$ respectively), instead of the 2277 fish that were aged, to obtain the same precision for the estimates of mean age. To further illustrate the effect of a low effective sample size, the 95% confidence interval for the mean age for quarter 2 based on $\hat{\mu}_1$ is approximately 5.33 ± 0.04 , if the fish are assumed to be a random sample from the population, while the correct confidence interval is approximately 5.33 ± 0.13 .

The effect of reducing the number of fish aged from each catch during quarter 1 on the error coefficient of variance (ECV, the estimated standard error divided by the estimated mean) for cod aged between 3 and 11+ is shown in Figure 1. The weighted average appeared to be less precise than the unweighted average in quarter 1. Regardless of the estimator applied, the curves for ages 4 through 10 were rather flat in the range of 20–85 fish, and increased fairly rapidly when the number of fish sampled was <20. For fish aged 3, the ECV appeared to be highest, but age 3 cod are not fully recruited to the fishery. For ages 11+, the ECV was large even when 85 fish were sampled per catch (0.30 for $\hat{\mu}_1$ and 0.19 for $\hat{\mu}_2$). Therefore, it appeared that little is gained by sampling more than approximately 20 fish per catch. Results were the same for the other three quarters, except that the weighted average appeared to be more precise for quarter 4 (Table 1).

Figure 2 shows the estimated age distributions generated by the two estimators for each quarter and the stratified estimates for the entire year. The inner brackets denote the bootstrapped 95% confidence interval based on all fish aged, and the outer brackets the 95% intervals when the number of fish sampled within a catch was reduced from approximately 85 to 20. Again, it is apparent that reducing the number of cod aged from each catch would only slightly affect precision. An analysis of catch data from several other years, which included evaluating various

Table 1. Summary statistics for estimating the mean age of the Northeast Arctic cod catch in 2000 using estimators $\hat{\mu}_1$ and $\hat{\mu}_2$ (see text). The parameter n is the number of fishing trips sampled, m the total number of fish aged from the n trips, $\hat{\mu}_i$ and $\text{se}(\hat{\mu}_i)$ the estimated mean age and its standard error for $i = 1, 2$ respectively, $\hat{m}_{i,\text{eff}}$ the effective sample size for $i = 1, 2$, and $\hat{\rho}$ is the estimated intracluster correlation coefficient. The approximate 95% confidence intervals are in parentheses. The estimated standard errors and confidence intervals are based on 500 bootstrap replicates.

Quarter	n	m	$\hat{\mu}_1$	$\text{se}(\hat{\mu}_1)$	$\hat{\mu}_2$	$\text{se}(\hat{\mu}_2)$	$\hat{m}_{1,\text{eff}}$	$\hat{m}_{2,\text{eff}}$	$\hat{\rho}$
1	70	6000	6.75 (6.35,7.23)	0.23	7.25 (6.99,7.45)	0.11	59 (37,212)	223 (158,330)	0.26 (0.20,0.33)
2	26	2277	5.33 (5.20,5.46)	0.07	5.33 (5.21,5.46)	0.06	211 (93,393)	213 (135,333)	0.10 (0.06,0.15)
3	13	1077	5.23 (4.98,5.60)	0.17	5.27 (5.08,5.55)	0.12	32 (18,193)	56 (29,215)	0.20 (0.06,0.33)
4	17	1342	5.05 (4.89,5.18)	0.07	5.06 (4.83,5.29)	0.12	182 (81,428)	57 (29,206)	0.23 (0.07,0.38)

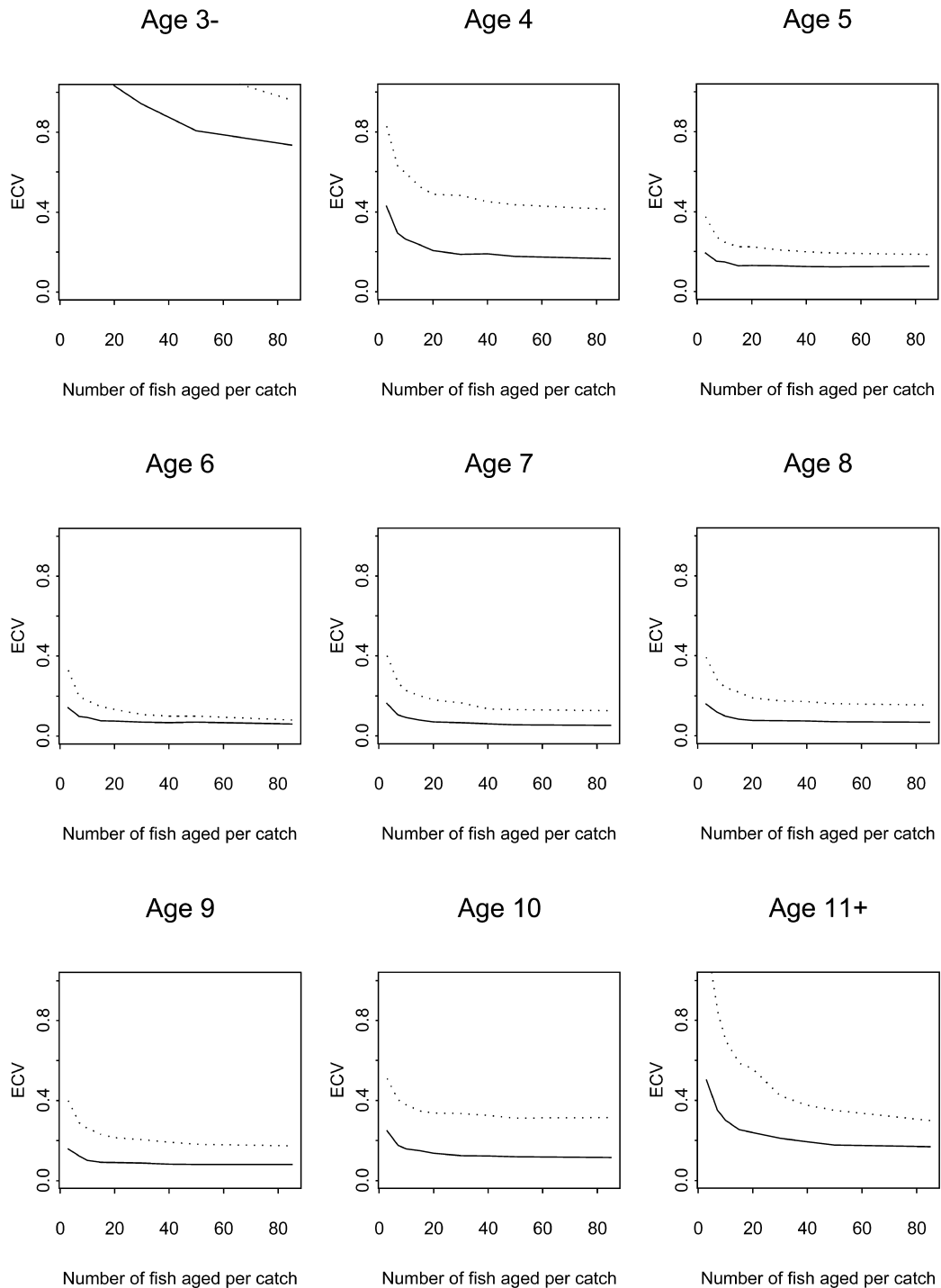


Figure 1. Error coefficient of variation (ECV) for the estimated proportion at age plotted against the number of cod aged per fishing trip during quarter 1 of 2000. The dotted line represents estimator $\hat{\mu}_1$ and the solid line, $\hat{\mu}_2$.

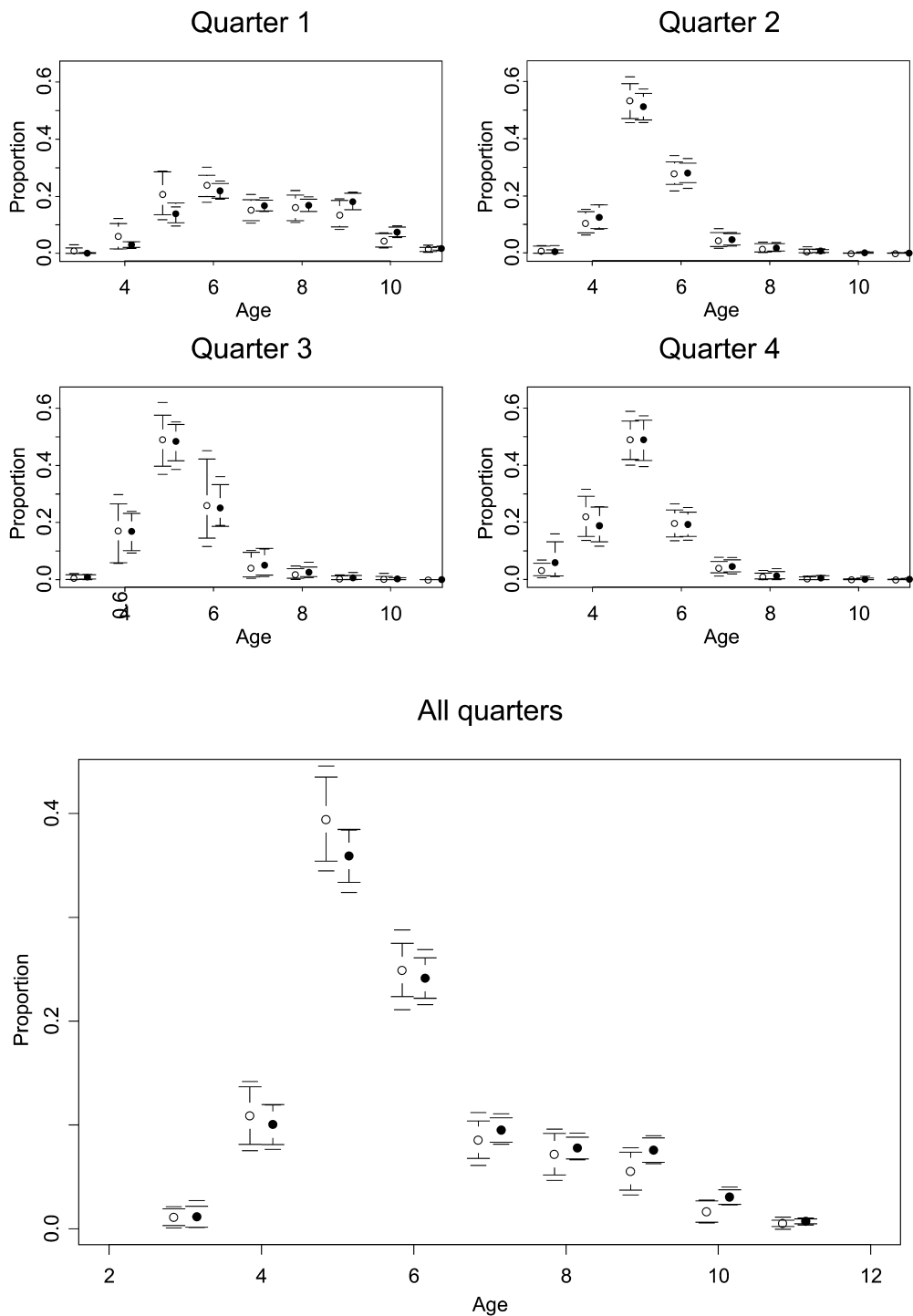


Figure 2. Estimated age distribution of the Norwegian commercial catch of cod in 2000 for each quarter and the whole year. The inner brackets denote the approximate 95% confidence interval based on all age readings, and the outer brackets the approximate 95% confidence intervals when the number of fish aged per catch was reduced from approximately 85 to 20. The open circles denote $\hat{\mu}_1$ and the filled circles $\hat{\mu}_2$.

post-stratification schemes, e.g. by gear, gave similar results: relatively small effective sample sizes and little loss of precision when the number of fish aged from each catch was reduced to 20.

The effect of varying the number of trips sampled and the number of fish sampled per trip when the intracluster correlation is positive is shown in Figure 3. The precision of the estimate was much more sensitive to varying the number of trips sampled than to varying the number of fish sampled per trip. In particular, if only the number of fish sampled per trip were to be increased, the variance tends towards an asymptote, which is equal to $(\sigma_x^2/n)\rho$.

Discussion

If it were feasible to select, at random, fish from the entire catch in each quarter, then a sample of about 500–600 cod would have provided an estimate of the mean age that was approximately as precise as the one based on 11 000 fish. However, even if an estimate of mean age is not needed for assessing a stock, a low effective sample size implies that the estimate of the entire age distribution is rather imprecise (Pennington and Vølstad, 1994; Pennington *et al.*, 2002). For example, even though the ages of some 11 000 cod caught during 126 trips were determined for 2000, the

estimate of the age distribution of the commercial catch was rather imprecise (Figure 2).

The low effective sample sizes were caused by the tendency for cod caught together to be more similar in age than those in the total catch. Similar results hold for scientific trawl surveys, which are conducted worldwide to assess the status of fish stocks. These surveys sample clusters of animals (a trawl catch), and because of intracluster correlation and variable density, the effective sample sizes for estimates of biological characteristics are small compared with the number of animals sampled (Pennington and Vølstad, 1994; Pennington *et al.*, 2002).

The method used to subsample a cluster, e.g. stratify by length and then select fish to age, will not significantly increase or decrease the effective sample size if there is intracluster correlation. For instance, even if all fish are measured during a bottom trawl survey, the effective sample sizes are still extremely small (Pennington and Vølstad, 1994).

Except in quarter 4, the unweighted estimator $\hat{\mu}_2$ appeared to be more precise for estimating the mean age and the proportion of catch at age than the weighted estimator $\hat{\mu}_1$ for this data set (Table 1, Figure 2). However, care should be taken while applying the unweighted estimator because it may be biased and inconsistent.

We assumed that the age of cod was determined without error. The strength of year classes varies considerably and,

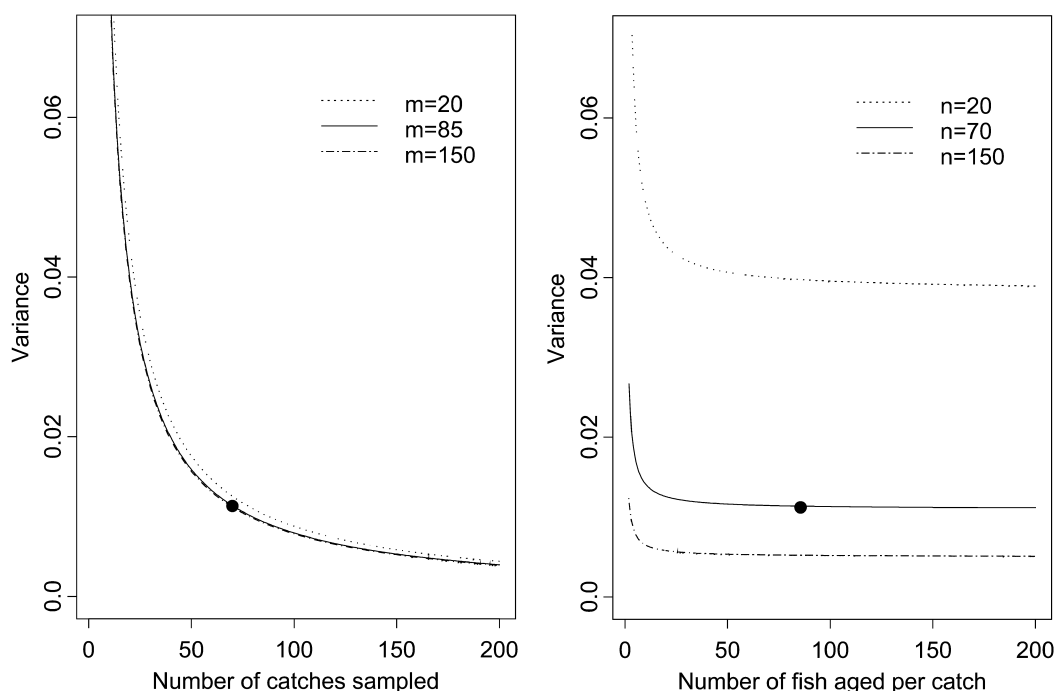


Figure 3. Variance of the estimate of mean age [Equation (5)] with equal cluster sizes and parameters as estimated for quarter 1 as a function of the number of trips, n , sampled (left panel) and the number of fish sampled per trip, m (right panel). The dots correspond to the sampling intensity in quarter 1 ($n = 70$, $m = 85$).

therefore, even low rates of misclassification error can cause severe overestimation of the proportion in the catch of a small year class and, to a lesser degree, underestimation of an abundant cohort (Richards *et al.*, 1992; Worthington *et al.*, 1995).

For routine surveys of the commercial catch of Northeast Arctic cod, one way to reduce possible classification errors would be to decrease the number of fish sampled from each catch and to use the time saved in being more rigorous in determining the age of each cod. As shown, such a reduction in the number of cod sampled within each cluster would only marginally decrease the precision of the estimates. The time saved by collecting and reading fewer otoliths is significant. For example, a reduction from about 85 to 20 cod from each catch sampled in 2000 would have reduced the number of fish sampled by Norway from 11 000 to 2500, which could decrease reading and handling costs by 150 workdays. Part of the time saved could be used to collect samples from more landings, which is the only practical way to improve the precision of the surveys. Motivated by these results, IMR will in future sample fewer fish from each catch and increase the number of landings sampled.

Acknowledgements

We thank Dr Kristin Helle (IMR) and Prof. Dag Tjøstheim (Department of Mathematics, University of Bergen) for their comments, Per Ågotnes (IMR) for an introduction to the practical side of collecting the catch data, and two anonymous reviewers for their constructive comments on the manuscript submitted. The Norwegian Research Council provided financial support for the work.

References

- Cochran, W. G. 1977. Sampling Techniques, 3rd edition. Wiley, New York.
- Efron, B. 1983. The Jackknife, the Bootstrap and Other Resampling Plans. 2nd edition. Society for Industrial and Applied Mathematics, Philadelphia.
- Gudmundsson, G. 1994. Time series analysis of catch at age observations. Applied Statistics – Journal of the Royal Statistical Society, Series C, 43: 117–126.
- Hoenig, J. M., and Heisey, D. M. 1987. Use of a log-linear model with the EM algorithm to correct estimates of stock composition and to convert length to age. Transactions of the American Fisheries Society, 116: 232–243.
- Kish, L. 1965. Survey Sampling. Wiley, New York.
- Nakken, O. 1994. Causes of trends and fluctuations in the Arcto-Norwegian cod stock. ICES Marine Science Symposia, 198: 212–228.
- Pennington, M., Burnmeister, L. M., and Hjellevik, V. 2002. Assessing the precision of frequency distributions estimated from trawl-survey samples. Fishery Bulletin, US, 100: 74–81.
- Pennington, M., and Vølstad, J. H. 1994. Assessing the effect of intra-haul correlation and variable density on estimates of population characteristics from marine surveys. Biometrics, 50: 725–732.
- Richards, L. J., Schnute, J. T., Kronlund, A. R., and Beamish, R. J. 1992. Statistical models for the analysis of ageing error. Canadian Journal of Fisheries and Aquatic Sciences, 49: 1801–1815.
- Searle, R. S., Casella, G., and McCulloch, C. E. 1992. Variance Components. Wiley, New York.
- Skagen, D., and Hauge, K. H. 2002. Recent development of methods for analytical fish stock assessment within ICES. ICES Marine Science Symposia, 215: 523–531.
- Skinner, C. J., Holt, D., and Smith, T. M. F. (Eds). 1989. Analysis of Complex Surveys. Wiley, New York.
- Smith, S. J. 1990. Use of statistical models for the estimation of abundance from groundfish survey data. Canadian Journal of Fisheries and Aquatic Sciences, 47: 894–903.
- Worthington, D. G., Fowler, A. J., and Doherty, P. J. 1995. Determining the most efficient method of age determination for estimating the age structure of a fish population. Canadian Journal of Fisheries and Aquatic Sciences, 52: 2320–2326.