ICES CM 2001

Theme session: P:01

Assessing the precision of the estimated age distribution of the commercial catch of Northeast Arctic cod

Sondre Aanes and Michael Pennington

The assessment of Northeast Arctic cod is based on estimates of the commercial catch numbers at age. The age structure of the catch is estimated by sampling individuals from commercial fishing trips. Though it is commonly assumed that the 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 have more similar ages than those in the total catch. For Northeast Arctic cod, the intra-cluster correlation for age is positive and thus the effective sample size is much smaller than the number of fish aged. It is shown, given the number of fish aged, that the precision of the estimated age distribution is rather low, and that the number of fish aged from each trip could be reduced from approximately 80 to 20 without a significant loss in precision.

Keywords: Sampling error, cluster sampling, intra-cluster correlation, variability in age determination of fish.

Sondre Aanes: Institute of Marine Research, P.O. Box 1870 Nordnes, N-5817 Bergen, Norway [tel: +47 55 23 86 27, fax: +47 55 23 86 87, e-mail:sondre.aanes@imr.no].
Michael Pennington: Institute of Marine Research, P.O. Box 1870 Nordnes, N-5817 Bergen, Norway [tel: +47 55 23 63 09, fax: +47 55 23 86 87, e-mail:michael.pennington@imr.no].

Introduction

When determining the precision of the estimated age distribution of a fish population based on a sample of age readings, it is usually assumed that the age readings are a random sample from the population (see, *e.g.*, Hoenig and Heisey, 1987; Richards *et al.*, 1992; Worthington *et al.*, 1995). Because it is generally impossible to randomly sample a fish population, samples of fish for aging are taken from a number of clusters, for example, from trawl hauls or fishing trips. The resulting sample of individuals will often contain much less information on the age distribution than an equal number of fish sampled at random, and 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 much more precise than it actually is (Pennington and Vølstad, 1994; Pennington *et al.*, 2001).

In this paper we examine the precision of estimates of the age distribution of the commercial catch of Northeast Arctic cod by Norway, which is based on a sample of fish from the catch. A number of fish are aged from each of an assumed random sample of fishing trips and therefore the sample consists of a number of clusters each of which is from a larger cluster (*i.e.* the fish caught during a trip). In addition, because aging fish is costly, we examine whether the number of fish aged from each trip could be reduced without significantly reducing the precision of the estimates.

Material and Methods

Data

The Norwegian fishery for Northeast Arctic cod is conducted throughout the year, and constitutes approximately half of the total catch. In order to get estimates of catch-characteristics, such as the age structure and catch at age in numbers or weight, samples of fish from selected fishing trips are collected and their ages, weights and lengths are recorded as well as the size of the catch. Staff from the Institute of Marine Research collects most of these data. For the years considered in this paper, 1997, 1998 and 1999, approximately 150-200 catches were sampled each year and from each catch, approximately 80 fish were subsampled for aging. The otoliths are removed from the sampled fish, and age is determined by counting the number of growth zones. It is important to note that this sampling scheme corresponds to a two-stage cluster sampling design (see, *e.g*. Skinner *et al*., 1989); the first stage is a sample of clusters of fish, and the second stage is a sample of individuals from each cluster. In this work we assume that we have a random sample of catches and a random subsample of otoliths from each catch. We also assume that the recorded catch sizes and total catches are known without error.

Statistical methods

Given a random sample of catches of size n and a random subsample of ages, y_{ij}, of size m_i from each catch, then an estimator of the mean age, μ_y, in the total catch is a ratio type estimator

$$\hat{\mu}_y = \frac{\sum_{i=1}^{n} \frac{M_i}{m_i} \sum_{j=1}^{m_i} y_{ij}}{\sum_{i=1}^{n} M_i} = \frac{\sum_{i=1}^{n} M_i \bar{y}_i}{\sum_{i=1}^{n} M_i}, \qquad (1)$$

where M_i is the number of fish in the i^{th} cluster (trip). Similarly, defining x_{ij} to be 1 if y_{ij} is age a, and 0 otherwise, an estimator for the proportion at age a in the total catch is

$$\hat{p}_a = \frac{\sum_{i=1}^{n} \frac{M_i}{m_i} \sum_{j=1}^{m_i} x_{ij}}{\sum_{i=1}^{n} M_i} = \frac{\sum_{i=1}^{n} M_i \hat{p}_{i,a}}{\sum_{i=1}^{n} M_i}. \qquad (2)$$

Estimators (1) and (2) are consistent and approximately unbiased for large samples (Cochran, 1977). Another option could be to use the simple unweighted average. However, the unweighted average may be a biased estimator such that the bias does not decrease as sample size increases. Because exact variance formulas for (1) and (2) do not exist (see, *e.g.* Thompson, 1992), bootstrapping was used to estimate the variances. Standard bootstrapping techniques (Efron, 1983) were used to estimate the variance. That is the empirical variance of a number of estimates (500) obtained from each replicate of the data. Each replicate was generated by first sampling the trips at random with replacement, then a number of individual otoliths from each selected trip were sampled with replacement. The effect of reducing the number of otoliths sampled within each trip was assessed by reducing the number of otoliths selected from each trip.

For evaluating the amount of information in the data, it is useful to consider the *intracluster correlation* coefficient, ρ, and the *effective sample size*, m_{eff} (Skinner *et al.*, 1989; Pennington and Vølstad, 1994; Pennington *et al.*, 2001). The intracluster correlation coefficient is defined as the correlation between individuals in the same cluster (*e.g.* Cochran, 1977), and the effective sample size is the number of individuals that would need to be sampled at random so that the estimates generated by simple random sampling would have the same precision as the estimates obtained based on the more complex sampling scheme. Specifically, m_{eff} is the number such that

$$\frac{\sigma_y^2}{m_{eff}} = Var(\hat{\mu}_y), \qquad (3)$$

where σ_y^2 is the variance of ages in the total catch. It should be stressed that if the effective sample size is low, this implies that the estimate of the entire distribution is rather imprecise (Pennington and Vølstad, 1994; Pennington *et al.*, 2001).

Furthermore, subject to some assumptions, it can be shown that the approximate variance of $\hat{\mu}_y$ can be written as

$$\frac{\sigma_y^2}{n\overline{M}}\left[1 + (\overline{M} - 1 + \sigma_M^2 / \overline{M})\rho\right], \qquad (4)$$

where $\overline{M}$ and σ_M^2 are the mean and variance, respectively, of the cluster sizes (Pennington and Vølstad, 1994). If $\rho > 0$, then the factor $(\overline{M} - 1 + \sigma_M^2 / \overline{M})\rho$ can greatly increase in the variance and corresponds to the phenomenon known as *overdispersion* that often arises when modeling polytomous responses (McCullagh and Nelder, 1989). If all clusters are of equal size, then $\sigma_M^2 = 0$ and (4) reduces to the formula for the variance when cluster sizes are equal (Cochran, 1977). Finally, if $\rho = 0$, then (4) reduces to the formula for the variance given a simple random sample of otoliths.

It should be noted that if there are errors in aging, then the above estimators would be biased. The age of a fish is determined by counting growth zones on the otolith and the aging process is subject to errors in determining these growth zones. The effects of classification errors on estimating the proportion at age will be to increase sampling variability and introduce bias (Tenenbein, 1970; Bross, 1954; Worthington *et al.*, 1995). To see this, let the true age structure be $P = (p_1, p_2, ..., p_r)'$, where r is the maximum possible age of the species and Q is an $r \times r$ -dimensional transition matrix specifying the probabilities that an individual of age j (column) is measured as age i (row). Then the distribution of ages that will be observed is equal to QP, which is a vector giving the observable proportions at age.

For example, Figure 1 demonstrates the bias caused by errors in age reading and shows that even a small error rate of 10% (5% to each side) may bias the estimates of the age distribution substantially. In general errors in age reading causes the observed age distribution to be much smoother than the true distribution. In particular, the highest proportions are underestimated while the smallest are overestimated. Increasing sample size will not decrease this bias, which can only be reduced if the fish are aged more accurately or Q is known.

Results

The effective sample sizes are much smaller than the number of fish aged (Table 1). For example, if it were possible to sample fish at random from the catch in region 1 in 1998, then it would have been sufficient to age 23 fish instead of 4338 to obtain the same precision for estimating the mean age. The low effective sizes are caused by positive intra-cluster correlation (Table 1). The lowest correlations are in region 4 for all years (ρ

ranges from 0.032 to 0.066) and the highest were in region 1 and 2 (ρ ranges from 0.398 to 0.745)

In Figures 2 to 4 are the estimated age distributions for each region and the stratified estimates for the entire area. The inner brackets denote the bootstrapped 95% confidence interval based on all the aged fish, and the outer brackets, the 95% intervals when the number of fish sampled within a catch is reduced from 80, on average, to 20.

The effect of reducing the number of fish sampled from each catch on the average for ages 4 to 9 of the error coefficient of variance, which is the standard error divided by the mean, is shown in Figures 5 to 7. The curves are rather flat in the range of 20 to 80 fish and increase fairly rapidly when the number fish sampled is less than 20.

Discussion

Because individuals within a cluster tend to be more alike than those in the entire catch, the variability of the estimates is more sensitive to the number of clusters sampled than the number of fish sampled within a cluster. Even though approximately 13,000 fish were aged each year, the low effective sample sizes imply that the estimates of the age distributions are rather imprecise (see Figures 2 to 4; Pennington and Vølstad, 1994; Pennington *et al.*, 2001).

It is common when analyzing fish-age data to assume the fish are a random sample of individuals from a population (see, *e.g.*, Hoenig and Heisey, 1987; Richards *et al.*, 1992; Worthington *et al.*, 1995). As shown, such an assumption would lead to a severe overestimation of the precision of the estimated age distribution.

To improve the precision of the estimated age distribution for the catch of Northeast Arctic cod, more catches from the fishery should be sampled, if possible, rather than sample more fish from each catch. Because the ultimate goal is to improve accuracy, that is increase precision and reduce bias, one should also focus on getting more accurate age readings. Though it appears that the variance component caused by errors in age readings is small as compared with the overall sampling variability, even relatively small reading errors may cause a significant bias, which does not decrease with increasing sample size. Therefore, considering the small increase in uncertainty if fewer fish are aged per trip, it is likely that the most efficient way to improve the accuracy of the estimates is to reduce aging errors. This could be accomplished by sampling fewer fish per trip, which would give the age readers more time to make more accurate age readings.

References

Bross, I. 1954. Misclassification in 2×2 tables. Biometrics 10: 478-486.

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

Efron, B. 1983. The Jackknife, the Bootstrap and Other Resampling Plans. 2nd ed. SIAM, Philadelphia, Pennsylvania.

Hoenig, J.M. and D.M. Heisey. 1987. Use of a log-linear model with the EM algorithm to correct estimates of stock-composition and to convert length to age. Trans. Am. Fish. Soc. 116: 232-243.

McCullagh, P. and J.A. Nelder. 1989. Generalized Linear Models. 2nd ed. Chapman & Hall, London.

Pennington, P. 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.

Pennington, P., L.M. Burmeister and V. Hjellvik. 2001. Assessing Trawl-Survey Estimates of Frequency Distributions. Fishery Bulletin. In Press.

Richards, L.J., J.T. Scnute, A.R. Kronlund, and R.J. Beamish. 1992. Statistical models for the analysis of ageing error. Can. J. Fish. Sci. 49: 1801-1815.

Skinner, C.J., D. Holt and T.M.F. Smith, eds. 1989. Analysis of Complex Surveys. Wiley. New York.

Snedecor, G.W. and W.G. Cochran. 1980. Statistical Methods. 7th ed. Iowa State University Press. U.S.A.

Tenenbein, A. 1970. A Double Sampling Scheme for Estimating from Misclassified Multinomial Data with Applications to Sampling Inspection. Technometrics 14: 187-202.

Thompson, S.K. 1992. Sampling. Wiley. New York.

Worthington, D.G., A.J. Fowler and P.J. Doherty. 1995. Determining the most efficient method of age determination for estimating the age structure of a fish population. Can. J. Fish. Sci. 52: 2320-2326.

Table 1. The estimated mean age, $\hat{\mu}_y$, of the Northeast Arctic cod catch and its estimated variance, $V\hat{a}r_B(\hat{\mu}_y)$, for each region and for the total area for 1997, 1998 and 1999, where M is the total number of fish, m the number of otoliths sampled for aging from the M fish, n the number of trips sampled, w the relative weight of the catch in each region, $\hat{m}_{eff}$ is the effective sample size and $\hat{\rho}$ is the estimate of intracluster correlation.

a);1997

Region	M	m	n	w	$\hat{\mu}_y$	$V\hat{a}r_B(\hat{\mu}_y)$	$\hat{m}_{eff}$	$\hat{\rho}$
1	59525	4858	62	0.261	5.997	0.013	114	0.283
2	77185	3928	53	0.280	6.049	0.108	26	0.745
3	67604	3126	42	0.263	7.304	0.009	87	0.190
4	5572	1222	17	0.196	7.556	0.008	212	0.049

b);1998

Region	M	m	n	w	$\hat{\mu}_y$	$V\hat{a}r_B(\hat{\mu}_y)$	$\hat{m}_{eff}$	$\hat{\rho}$
1	167783	4338	49	0.236	5.578	0.088	23	0.398
2	47752	2842	40	0.277	6.230	0.104	25	0.318
3	33965	2700	34	0.260	7.600	0.010	93	0.194
4	13729	2726	32	0.227	8.017	0.004	336	0.032

c);1999

Region	M	m	n	w	$\hat{\mu}_y$	$V\hat{a}r_B(\hat{\mu}_y)$	$\hat{m}_{eff}$	$\hat{\rho}$
1	56277	4801	58	0.275	5.170	0.011	147	0.247
2	32778	2998	38	0.350	5.900	0.077	41	0.445
3	22406	2206	28	0.223	7.231	0.086	28	0.272
4	3606	1388	21	0.152	8.222	0.017	149	0.066

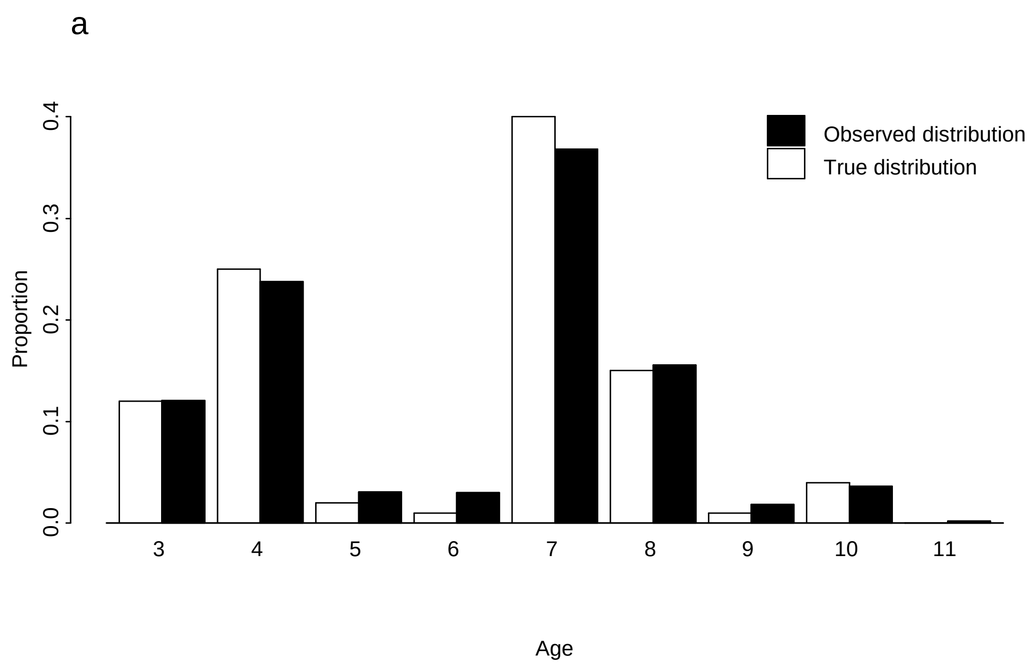

Figure 1. The true age distribution (white bars) and the observed distribution (black bars) if the probability of misaging an individual is approximately; a) 5%, and b) 15% to each side of the true age.

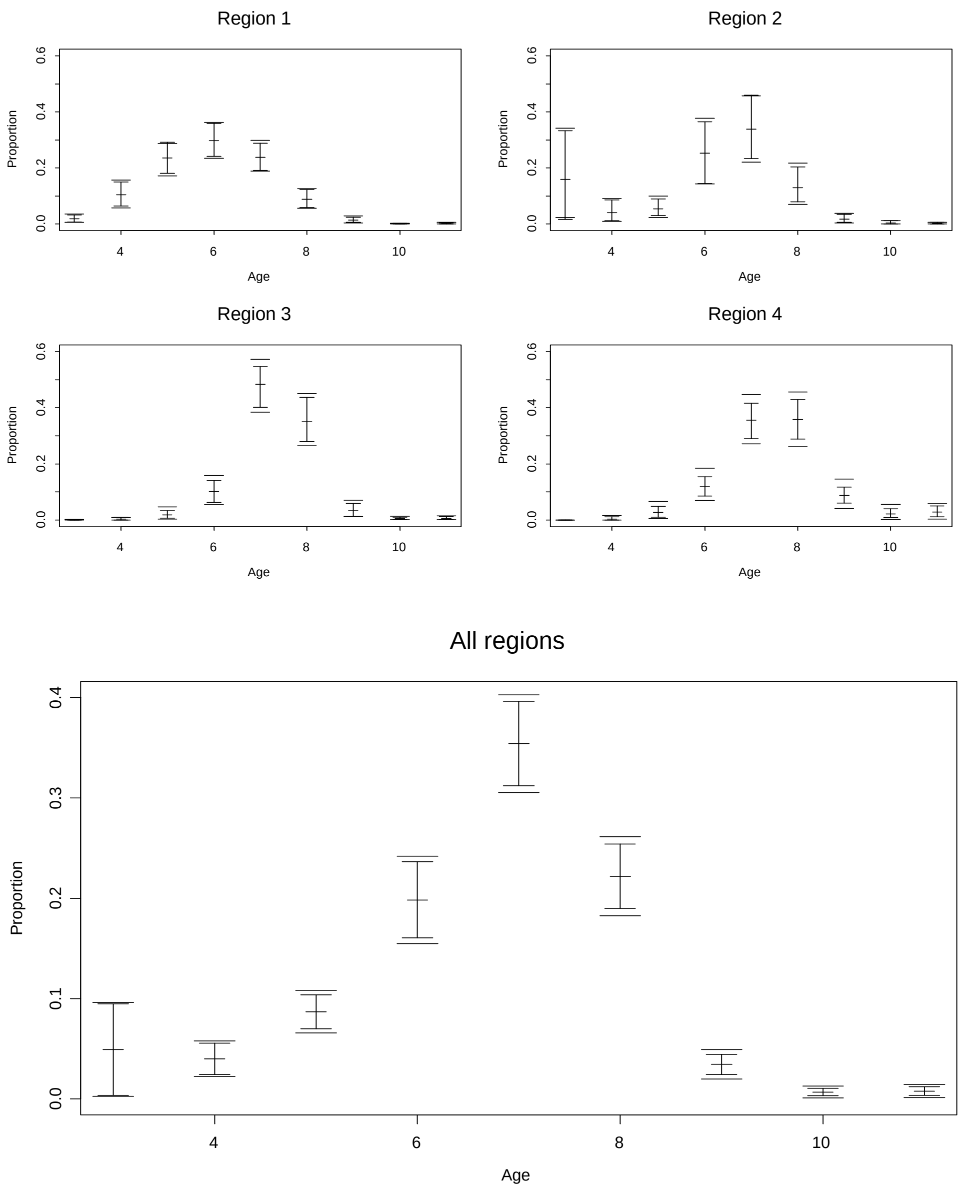

Figure 2. The estimated age distribution of the catch in 1997 for the 4 regions and the entire area. The inner brackets denote the 95% confidence interval based on all the age readings and the outer brackets, 95 % confidence intervals when the number of measured otoliths is reduced from approximately 80 to 20 per catch.

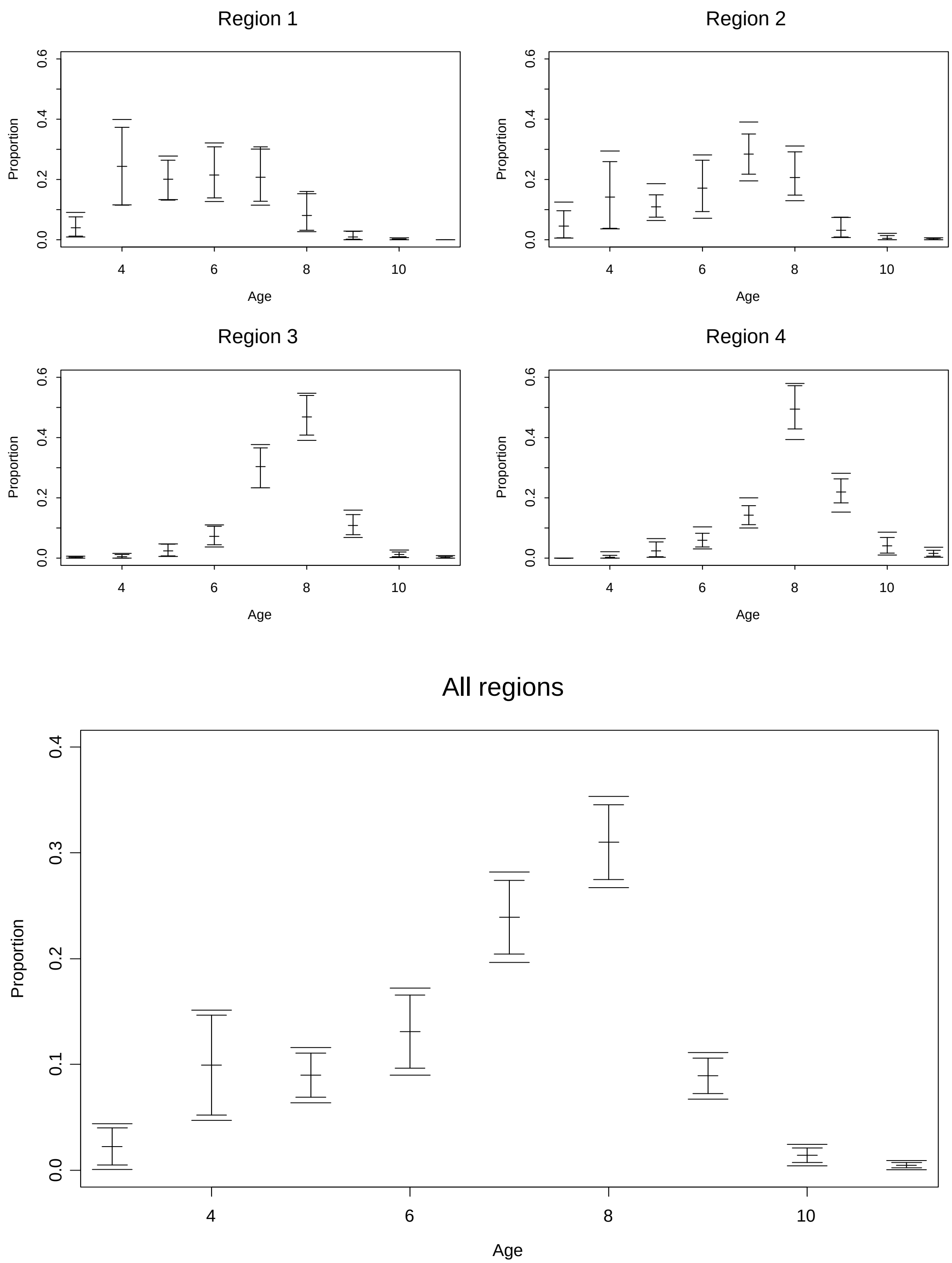

Figure 3. The estimated age distribution of the catch in 1998 for the 4 regions and the entire area. The inner brackets denote the 95% confidence interval based on all the age readings and the outer brackets, 95 % confidence intervals when the number of measured otoliths is reduced from approximately 80 to 20 per catch.

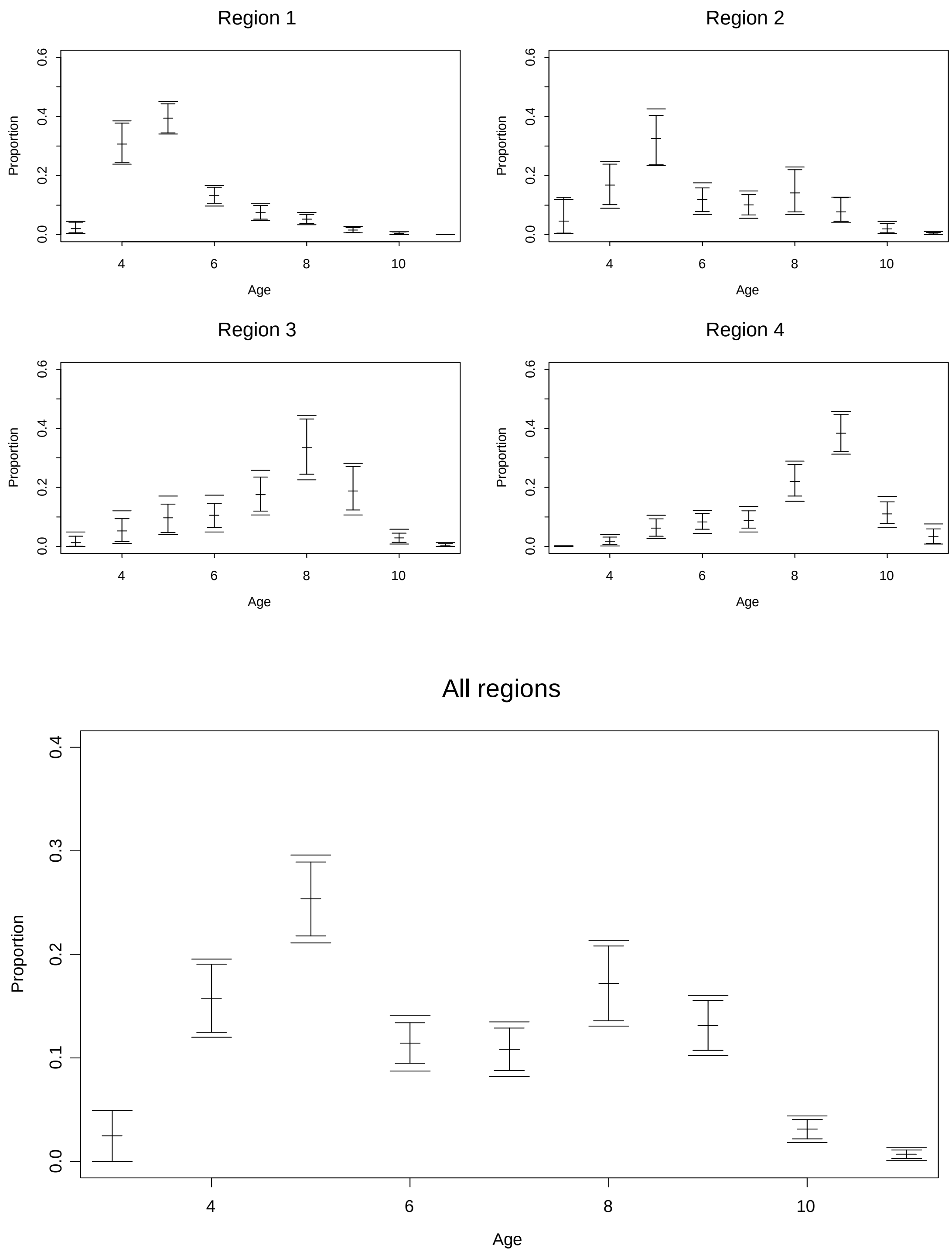

Figure 4. The estimated age distribution of the catch in 1999 for the 4 regions and the entire area. The inner brackets denote the 95% confidence interval based on all the age readings and the outer brackets, 95 % confidence intervals when the number of measured otoliths is reduced from approximately 80 to 20 per catch.

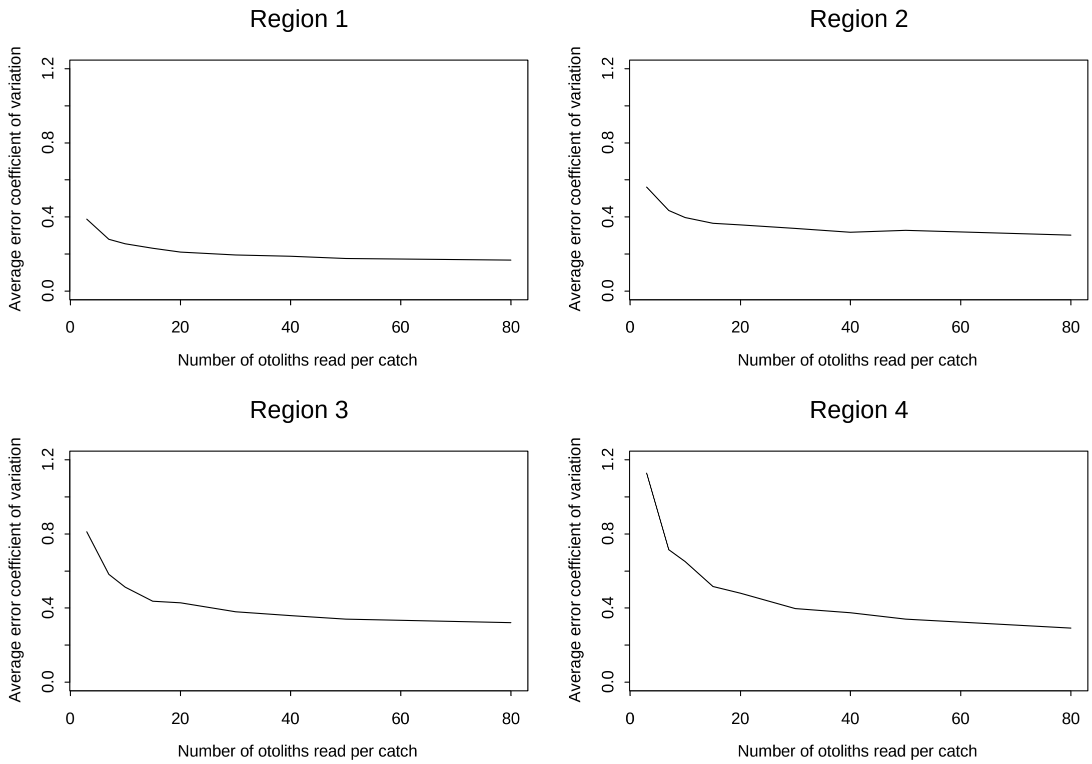

Figure 5. The average error coefficient of variation for ages 4-9 for the estimated proportion at age versus the number of age samples taken per trip in 1997.

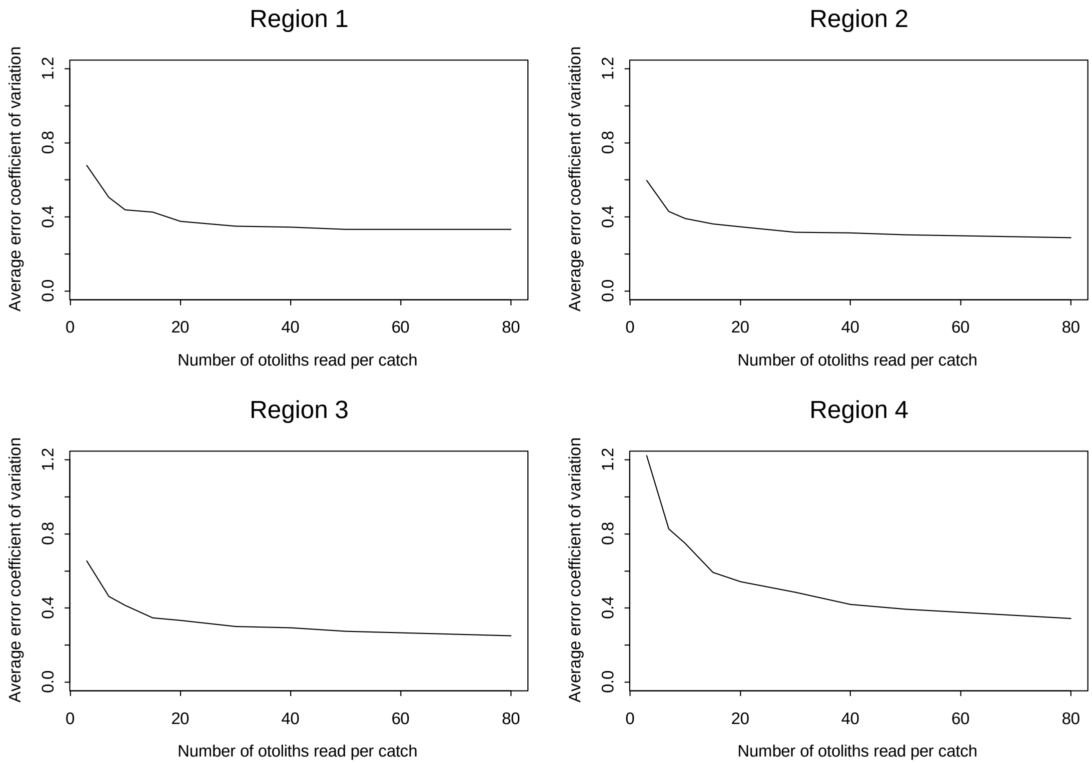

Figure 6. The average error coefficient of variation for ages 4-9 for the estimated proportion at age versus the number of age samples taken per trip in 1998.

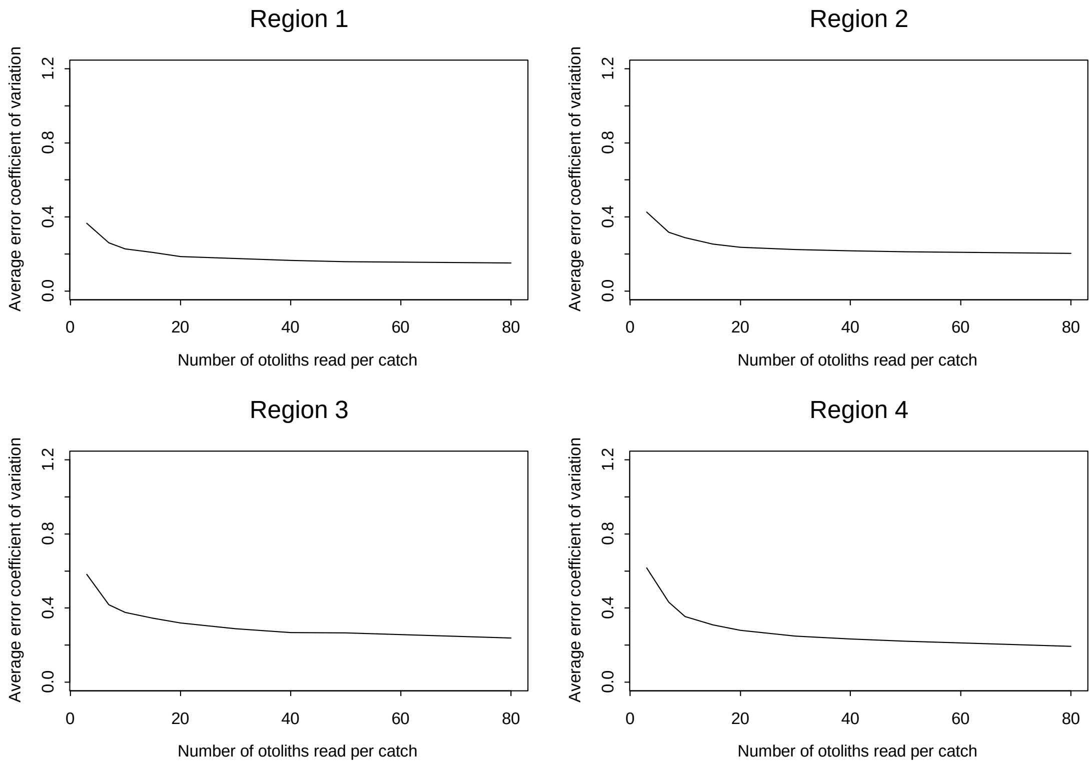

Figure 7. The average error coefficient of variation for ages 4-9 for the estimated proportion at age versus the number of age samples taken per trip in 1999.