ICES Journal of Marine Science, 59: 1072–1080. 2002
doi:10.1006/jmsc.2002.1258, available online at http://www.idealibrary.com on IDEAL®

A model for the uncertainty around the yearly trawl-acoustic estimate of biomass of Barents Sea capelin, *Mallotus villosus* (Müller)

Sigurd Tjelmeland

Tjelmeland, S. 2002. A model for the uncertainty around the yearly trawl-acoustic estimate of biomass of Barents Sea capelin, *Mallotus villosus* (Müller). – ICES Journal of Marine Science, 59: 1072–1080.

Management of Barents Sea capelin, *Mallotus villosus* (Müller), is based on the precautionary approach, and probabilistic short-term predictions are used directly as a tool. One important part of the overall uncertainty is the uncertainty in the yearly trawl-acoustic estimate made in September. Harvest-control rules are based on historical time-series for the spawning stock, the uncertainty around which therefore has a direct bearing on management of the species. Accordingly, one must quantify the uncertainty, not only of the September estimate the latest year, but also for the whole time-series. In this paper, a model for the uncertainty around the yearly trawl-acoustic estimate is developed. The uncertainty in the mean integrator value by standard one-by-two-degree rectangles is evaluated with a model for the distribution of the basic five-mile integrator values parameterized with data from many historical surveys. The uncertainty from the biological samples is quantified on the basis of the multinomial distribution. A large number of replicates of the historical time-series of September estimates are produced and stored on file for later use.

Keywords: Barents Sea, capelin, parametric bootstrap, stock assessment.

Manuscript received 26 July 2001; accepted 11 April 2002.

S. Tjelmeland: Institute of Marine Research, PO Box 1870, Nordnes, N-5817 Bergen, Norway; tel: +47 55 238400; fax: +47 55 238617; e-mail: sigurd@imr.no.

Introduction

Since 1972 there has been an annual trawl-acoustic survey of the Barents Sea stock of capelin, *Mallotus villosus* (Müller). Initially, the survey was conducted with only one vessel, but gradually more effort was put into the survey, until now, two Norwegian and 2–3 Russian vessels participate in a unified operation. One vessel is considered the coordinating vessel and the other vessels report data on a daily basis to it; the coordinating vessel has the mandate to instruct the other vessels to change planned survey tracks, if necessary. The backscattered sound from the transducers is integrated along the course in intervals of five nautical miles (the length unit used for survey data is customarily one nautical mile, which is used also here; one nautical mile is 1.852 km) and split on species during daily scrutiny of the echograms. Frequent trawling serves two purposes: to give information both for splitting the integrator values by species and for calculating the stock abundance in number by age and length. In the present study, data from 1992 and later are used, because the information stored (Norwegian vessels only) permits night and day values to be distinguished. Altogether, 17 000 five-mile values are used, of which 8000 contained capelin.

The survey has traditionally been treated as a measurement of absolute stock size, and the assessment and subsequent quota allocation has been based on it. The survey is semi-adaptive in the sense that the vessels adapt to what is perceived as the geographical border of the stock, and sometimes the survey tracks have been narrowed when the vessels have encountered large concentrations of capelin. Apart from this, the survey is conducted on a more-or-less fixed survey-track basis, and the distance between tracks has been fairly constant. The survey must be considered a systematic survey with a random starting point, because the distribution of fish can be considered random in relation to the distance between tracks. In a simulation study, Simmonds and Fryer (1996) found that such surveys are unbiased. However, they considered situations in which one transect across the whole distribution could be regarded

1054–3139/02/101072+09 $35.00/0

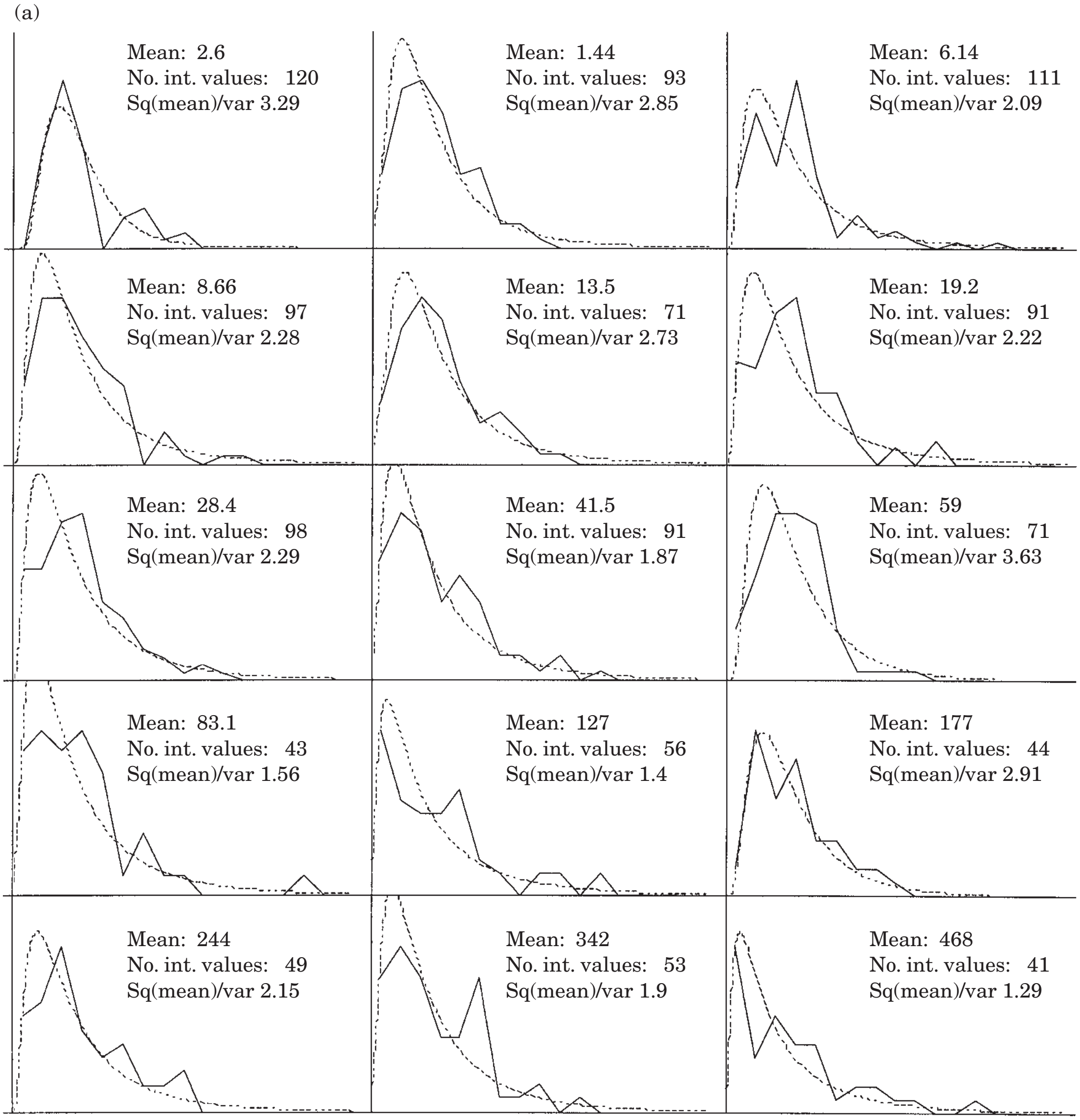

Figure 1(a).

as a single measurement, which is applicable to fish stocks with a pronounced elongated distribution. In the Barents Sea, the capelin distribution is essentially two-dimensional, a situation that is little studied theoretically.

Management of Barents Sea capelin is based on probabilistic forecasts of the spawning stock size at 1 April from the two-country trawl-acoustic survey during the previous September. A correct probability distribution is obtained only if the probability distribution of the September estimate correctly reflects the uncertainty in both acoustic and biological sampling. Moreover, because harvest-control rules should be based on analysis of historical data, the uncertainty connected to the harvest-control rule must be based on the probability distribution of the September surveys back to 1972.

The variable of interest is the number of capelin distributed not only by age but also by length, because the model used for estimating the maturing component is length-based.

The September trawl-acoustic estimate of Barents Sea capelin

The estimate is constructed in the following stages (Toresen *et al.*, 1998):

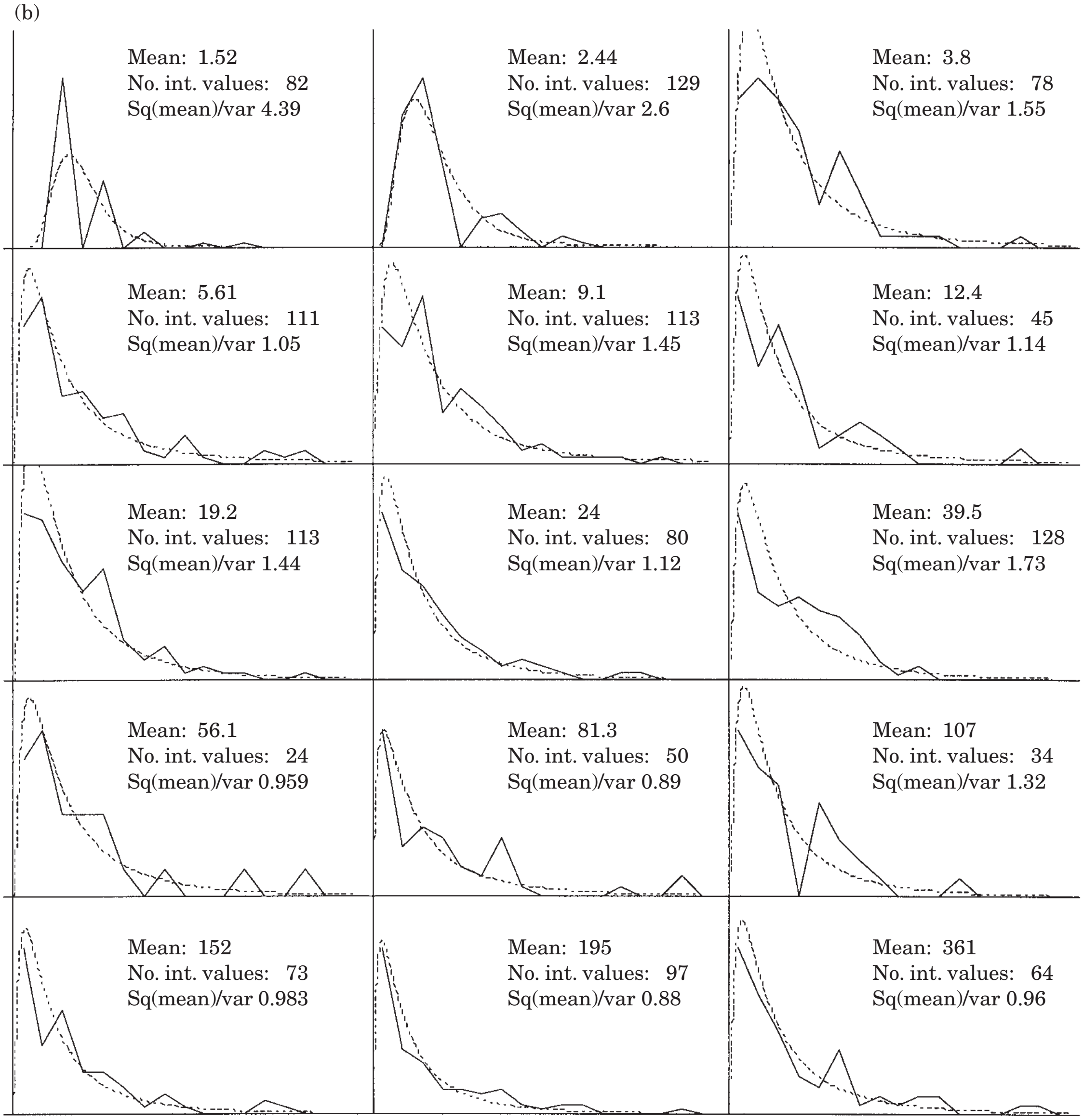

Figure 1(b).

Figure 1. Five-mile values partitioned in (a) 4-value and (b) 12-value intervals, grouped according to the mean value. The x-axis is the logarithm of the mean of integrator values in each bin, and the y-axis is the frequency, normalized so that the area under each curve is 1. The solid curve shows histogrammed data and the dashed curve the fitted lognormal distribution.

- The Barents Sea is divided into rectangles of 2 degrees east–west and 1 degree north–south, in each of which there are 1–4 vessel transects. In each transect, 6–10 five-mile integrator points are sampled, depending on latitude, and based on these samples, the arithmetic mean in each rectangle can be calculated.
- From biological sampling in or near each rectangle, the length distribution of capelin in that rectangle is calculated. Each sample may be subjectively weighted.
- Based on the mean integrator value and the length distribution, the total number of capelin in each length-class is calculated for each rectangle.
- The basic rectangles are aggregated over larger areas (typically 3–5), where each area is perceived to have more or less uniform age distributions. For each area, an age–length key is calculated and applied to the length distribution summed over rectangles to give an age–length estimate. The samples used in the age–length key may also be weighted subjectively.

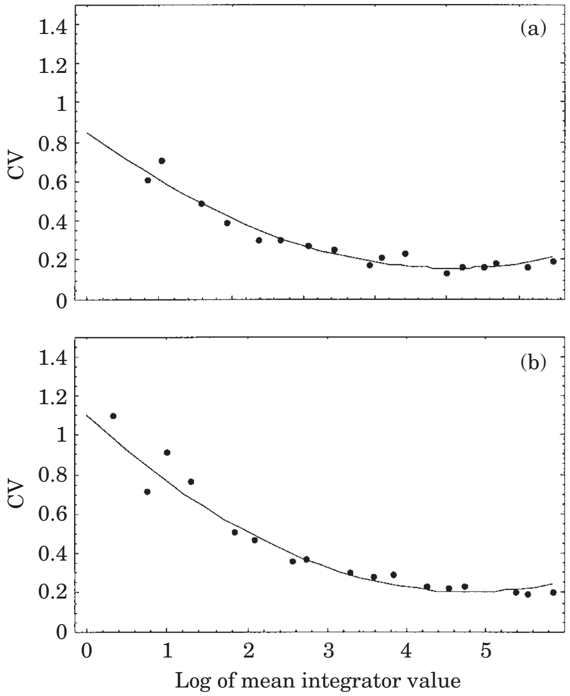

Figure 2. CV as a function of the mean of non-zero 5-mile integrator values when the tracks have been divided into intervals of (a) 4 and (b) 12 values. The curve is a polynomial fit of order 2.

In this paper, a method for calculating the uncertainty around the estimate is developed using a parametric bootstrap procedure. The estimate is made many times, each time applying different values for the mean integrator value in each rectangle, the length distribution in each rectangle, and the age–length key for each area.

Modelling the uncertainty associated with acoustic data

The model for the distribution of mean integrator values in each rectangle is based on a model for the distribution of individual five-mile values inside the rectangle, which is fitted to the data. During resampling, values are drawn from this model to calculate new mean integrator values.

For the model of mean integrator value by rectangle, an alternative method can be applied: to resample directly from the actual integrator values inside the rectangle. However, all integrator values used historically are not available, so one needs a generic model, in which the only input is the observed mean integrator value. One important factor in determining the precision of the estimate is the coverage of each rectangle, i.e. how many five-mile integrator values are there in it? This depends on the number of tracks through the rectangle and the number of values in each track. It is common practice to cover the rectangles in an east–west direction, and the number of values for each track depends on the east–west extent of the rectangle, which again depends on latitude. The number of tracks through each rectangle depends mainly on the number of vessels participating in the survey.

The aim is not to construct confidence intervals, which alone would be of little use to management. What is needed is the probability distribution of the full age–length estimate, which would be represented by replicates of the estimate, from which input values to the management procedure will be drawn at random with equal probability. No attempt to correct for bias is made. The bias correction should be done on the primary interest variable in the final assessment step. This would be the spawning stock, as forecast from the September estimate with subsequent use of a maturation model and a model for predation by cod (*Gadus morhua*; Gjøsæter *et al.*, 2002).

All analyses made in this paper are done by Mathematica. The work is also documented on www.assessment.imr.no, where too the Mathematica notebooks used can be inspected.

It has not been possible to avoid subjective elements altogether, but the subjective elements are based on operational elements that can be investigated further later.

A model for the distribution of five-mile integrator values

The model for the distribution of five-mile integrator values rests on the fundamental assumption that the distribution is dependent only upon the mean integrator value in the surrounding area and not on other factors such as geographical position or year. The model for the distribution of acoustic five-mile measurements treats zero values and non-zero values separately, analogous to the procedure used for analysis of groundfish trawl surveys (Stefánsson, 1996).

Each survey has been divided into intervals, where the number of five-mile values in each interval is 4, 5, . . ., 12 and the analysis is made for each of these partitions. In each interval, the zero values are grouped according to the logarithm of the mean value, and the non-zero values are grouped according to the logarithm of the mean of the non-zero values.

Figure 1 is based on data partitioned in 4 (Figure 1a) and 12 (Figure 1b) five-mile integrator value intervals. Gamma, Weibull, and lognormal distributions have been fitted to the data and all seem to fit reasonably, although the Weibull distribution seems to be slightly better than the other two. The lognormal distribution is somewhat narrow and skewed to the left when

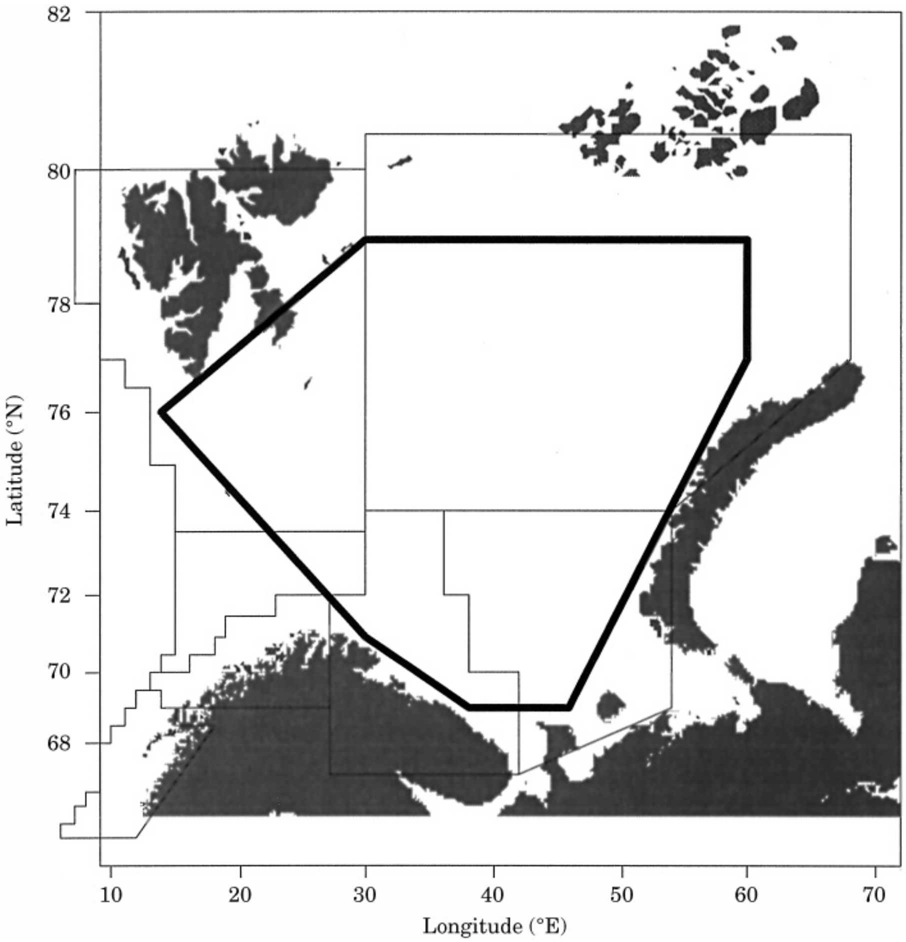

Figure 3. Area division for calculating number by age. Thin lines show the areas used for calculating number by age (Multspec areas). The heavy line shows the standardized distributional area for capelin used to convert vessel days to track density.

compared with the data and the two other distributions. However, the deviation is slight, and in view of its simplicity and ease of use, the lognormal distribution is used in the following analyses, and only the lognormal distribution is depicted graphically.

The expectation value of the distribution in each bin is taken as the mean of the individual data values on a log-scale. The lognormal distribution is then determined by one additional parameter, taken to be the coefficient of variation (CV) on a log-scale. Figure 2 shows the CV of the empirical data for interval sizes of 4 and 12 values. The CV decreases with mean integrator value, but flattens out at large mean integrator values and is remarkably well described by a polynomial of order 2.

When the mean integrator value in a rectangle is resampled, the partitioning appropriate for the latitude of that rectangle is first selected on the basis of its east–west extent, which depends on latitude. Next to be selected are the parameters of the polynomial fitted to the relationship between the CV and the mean integrator value for the given partition. On the basis of the mean integrator value of the rectangle, the CV of the lognormal distribution is calculated from the selected polynomial, and the empirical distribution of zeros for the given mean integrator value is selected. Finally, the number of zeros is drawn, and the number of non-zero values is resampled from the lognormal distribution.

Resampling the mean integrator value in each rectangle

The number of tracks through each rectangle depends on the number of vessel days utilized during a survey. Some time is lost unproductively through, for instance, bad weather and technical problems with the ships. As a crude estimate of the number of tracks in each rectangle, it is assumed that the survey area is adequately modelled by the dark curve shown in Figure 1. The number of vessel days is given by Gjøsæter *et al.* (1998) for

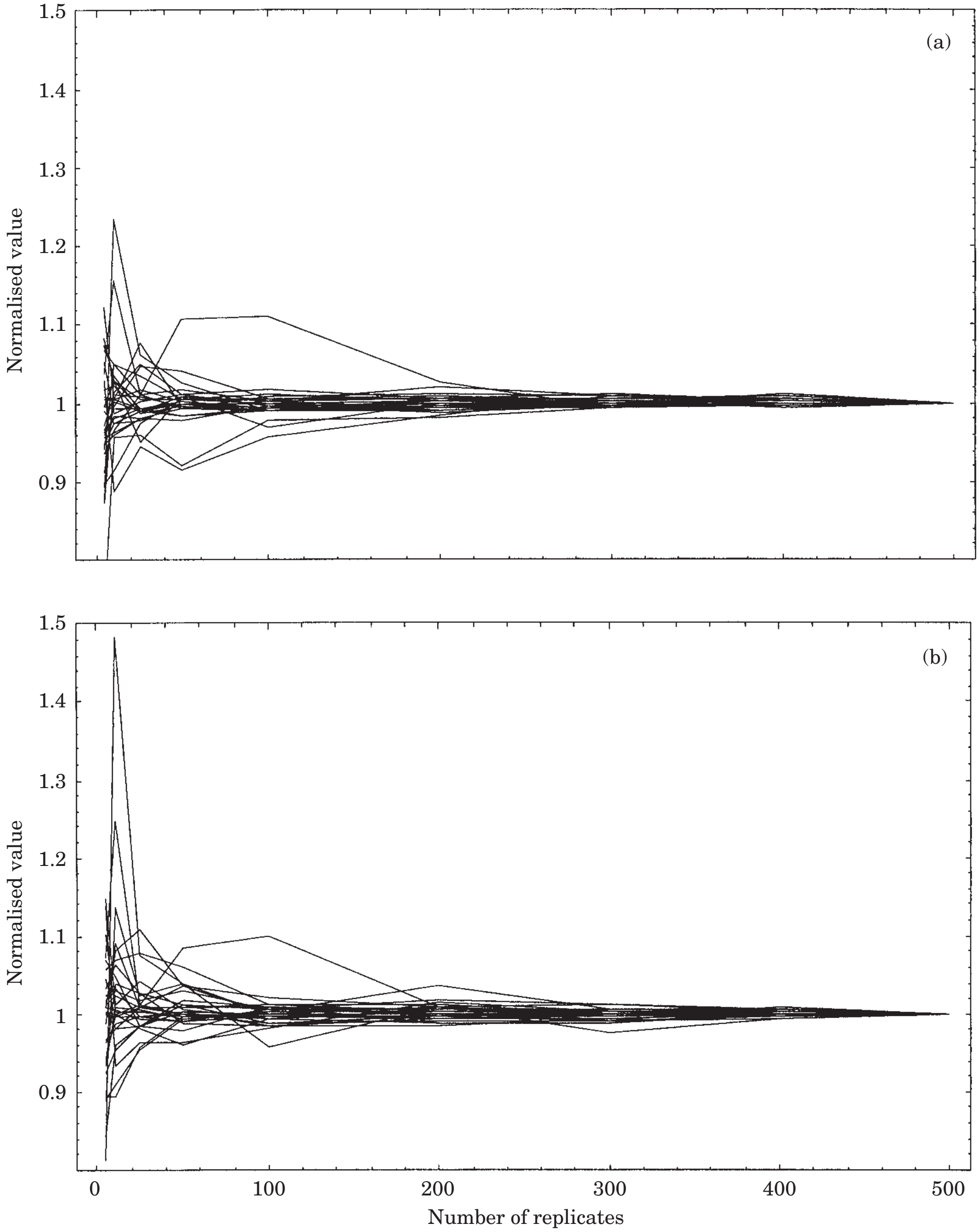

Figure 4. The CV (upper) and median (lower) of the number of 3-year-old capelin as a function of number of replicates for different years. For each year, the data have been normalized to the value obtained for 500 replicates.

the surveys of 1972–1997. Vessel days for the years 1998–2000 are taken from survey reports (H. Gjøsæter, pers. comm.). It is assumed that the effective speed of surveying is 7 knots and that two seadays of each vessel are lost through technical problems, etc. The number of transects through each rectangle is then calculated by spacing east–west transects evenly in a north–southwest direction. The effective coverage will usually be somewhat denser because the effort is not evenly distributed; rather more vessel days are spent in areas where capelin concentration is densest. This effect has not been accounted for in the present version of the model.

Zero values and five-mile integrator values are drawn for each rectangle on the basis of the historical mean integrator value as outlined above. Based on the drawn values, a new value of the mean integrator value in the rectangle is calculated. The drawn values are multiplied by $\mathrm{Exp}(-0.5\,\sigma^2)$ to correct for the bias in the expectation value of the lognormal distribution, where σ is the standard deviation.

Modelling the uncertainty connected with biological sampling

The biological samples are resampled assuming multinomial distributions in each length-class or age–length class. For the samples used to calculate the length distribution in each rectangle, the uncertainty may thus be underestimated, because only the uncertainty associated with each haul (i.e. that associated with age determination) is accounted for. The variability from haul to haul may be larger than that associated with single hauls (Pennington and Vølstad, 1994), owing to clustering effects such as schooling by size (Gjøsæter and Korsbrekke, 1990). However, this effect could be less severe for capelin than for demersal fish, because as a rule the capelin in September remain in layers in the upper 50 m by night, and single fish may lose their school identity. When the age–length estimate is calculated for larger areas, whole samples are resampled before the age–length distribution of the individual samples is calculated by drawing from the multinomial distribution, so accounting for the uncertainty associated with variability between hauls.

The original areas are not used. Instead, the number at age is calculated for a fixed area system (Tjelmeland and Bogstad, 1998) that permits analysis of changes in the geographical distribution from year to year. Figure 3 shows the area division used.

Additional considerations

A number of subjective elements enter into the estimate; in such cases, crude assumptions have been made that await further research.

Identification uncertainty

If the density of capelin is high, there is little difficulty in distinguishing the capelin traces from background noise during daily scrutiny of the echograms. If the capelin density is moderate or low, partitioning of the integrator value between capelin and other species may be difficult, especially at night. Somewhat speculatively, the five-mile integrator values drawn are adjusted with a random element that is decreased from a maximum of 0.7 by night and 0.2 by day at an integrator value of zero, to 0 at an integrator value of 500 by night and 150 by day. The new integrator value is drawn from a uniform distribution with its centre at the old value and a total spread equal to the centre value multiplied by the random element. The probability for day or night is set to 0.5. All values along a transect through a rectangle are scaled by the same factor, i.e. the whole transect is supposed to be covered either by day or by night.

Weighting samples that determine the length distribution in each rectangle

Weights are assigned to samples simply by assigning a sample several times to the rectangle (only integer weights are possible with the estimate of capelin in the Barents Sea). Usually, however, the samples are unweighted. In order to reflect the uncertainty associated with weighting of the samples, each is allocated a stochastic weight with a uniform distribution between 0.5 and 1.5.

Weighting samples that determine the age–length key in each area

During the regular September surveys, all samples in an area are lumped to construct that area's age–length key. A weight is associated with each sample to reflect the importance of that sample relative to other samples. This breaks down, however, when the areas are redefined as in the present paper. To reflect the increased uncertainty associated with such subjective weighting, the samples are resampled (block sampling). This is not an entirely satisfactory procedure, however, and in an improved version of the model, weights will be associated with the mean (resampled) integrator value in the rectangles to which the samples have been allotted.

Results

The number of replicates necessary is dependent on the application. For abundance estimation using acoustic data, Simmonds *et al.* (1991) recommend 100 replicates. Efron and Tibshirani (1993) state that very seldom are more than 200 replicates needed to estimate a standard error, and that 50 values will generally be adequate. However, much bigger values are needed to estimate confidence intervals.

In the present case, the number of replicates needed is dependent on the variable of interest. If the variable of interest is the expectation value of the number at age for the whole Barents Sea, fewer replicates are needed than if one wishes to investigate distributions by sex and/or area. For the present work, 500 replicates were chosen. Figure 4 shows how the CV and median value for 3-year-old capelin converge as the number of replicates is increased.

Figure 5 depicts the CV for capelin of age 2 and age 3 by year; the CVs are about 0.2 or lower, except for the years 1987–1991 for 3-year-olds and 1991 for 2-year-

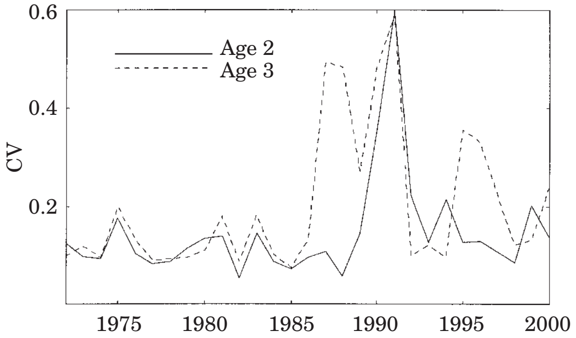

Figure 5. The CV for 500 replicates of the September estimate for capelin of age 2 and age 3 by year.

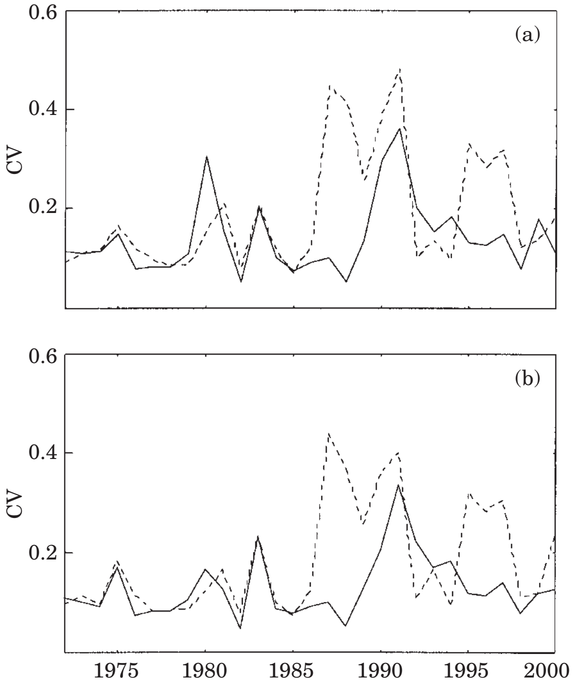

Figure 6. The CV for capelin of age 2 (solid curve) and age 3 (dashed curve) by year, with uncertainty only from acoustics, (a) applying identification uncertainty and (b) not applying it.

olds. Figure 6 shows the CV for ages 2 and 3 applying only uncertainty on the acoustics. There is not much difference whether the identification uncertainty is applied (Figure 6a) or not (Figure 6b), which means that this source of uncertainty is small compared with the uncertainty from spatial variation for the assumptions made. Figure 7 shows the CV for ages 2 and 3 applying only the uncertainty from the biological samples. With the exception of fairly large uncertainties in some years for capelin of age 3, the uncertainty is appreciably reduced from a situation when all uncertainty is applied (Figure 5), and lower than the uncertainty from acoustics. The acoustic uncertainty therefore seems to be more critical than the uncertainty from the biological sampling. Resampling the biological samples when the division on age is made by area seems to have an appreciable effect.

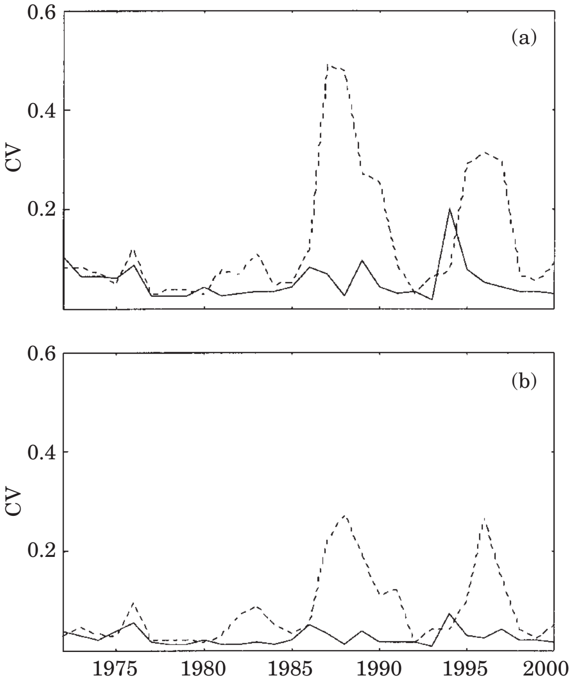

Figure 7. The CV for capelin of age 2 (solid curve) and age 3 (dashed curve) by year, with uncertainty only from biological sampling, (a) applying resampling from biological samples and (b) not applying it.

Discussion

Management of Barents Sea capelin is based on the precautionary approach in the sense that a lower threshold B_{lim} for the spawning stock is defined and the catch quota is constrained by requiring that B_{lim} can be exceeded with only a small probability (Gjøsæter *et al.*, 2002). To evaluate the probability of exceeding B_{lim} for different catch options, a large number of probabilistic projections need to be simulated. In such simulations, the initial stock structure and size can be taken from the present model, a different replicate for each trajectory. B_{lim} is based on historical data, at present being taken as the spawning stock in 1989, which was very small but still yielded very good recruitment. As the historical data are uncertain, so is B_{lim}, and the present model gives a necessary input for evaluating the uncertainty in B_{lim} (Gjøsæter *et al.*, 2002).

It is expected that, in the near future, a B_{target} will be established for Barents Sea capelin, and that it will be defined as the spawning stock biomass that in the long term will yield the biggest catches. For each possible time-series of historical estimates resampled from replicates calculated by the present model, a different relationship between spawning stock and recruitment will be calculated. As B_{target} will be based on long-term simulation runs using spawning stock–recruitment relationships estimated on the basis of historical data, the present model constitutes an important element in establishing the uncertainty in B_{target}.

In the present analysis, some subjective elements could not be avoided; identification uncertainty is clearly the most critical. It may in future be possible to give a quantified estimate of identification uncertainty by scrutinizing stored echograms. Also, the present model makes it feasible to investigate how large the identification uncertainty (and other subjective elements) must be in order to be significant in relation to the most crucial sources of uncertainty: biological sampling and the spatial variation of the individual five-mile integrator values.

Acknowledgements

Three anonymous referees provided valuable criticism and suggestions, which greatly improved the manuscript.

References

Efron, B., and Tibshirani, R. J. 1993. An Introduction to the Bootstrap. Chapman & Hall, New York. 436 pp.

Gjøsæter, H., Bogstad, B., and Tjelmeland, S. 2002. Assessment methodology for Barents Sea capelin, *Mallotus villosus* (Müller). ICES Journal of Marine Science, 59: 1086–1095.

Gjøsæter, H., Dommasnes, A., and Røttingen, B. 1998. The Barents sea capelin stock 1992–1997. A synthesis of results from acoustic surveys. Sarsia, 83: 497–510.

Gjøsæter, H., and Korsbrekke, K. 1990. Schooling-by-size in the Barents Sea capelin stock. ICES CM 1999/D: 28, 9 pp.

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.

Simmonds, E. J., and Fryer, R. J. 1996. Which are better, random or systematic acoustic surveys? A simulation study using North Sea herring as an example. ICES Journal of Marine Science, 53: 39–50.

Simmonds, A. J., Williamson, N. J., Gerlotto, G., and Aglen, A. 1991. Survey design and analysis procedures: a comprehensive review of good practice. ICES CM 1991/B: 54, 132 pp.

Stefánsson, G. 1996. Analysis of groundfish survey abundance data: combining the GLM and delta approaches. ICES Journal of Marine Science, 53: 577–588.

Tjelmeland, S., and Bogstad, B. 1998. MULTSPEC – a review of a multispecies modelling project for the Barents Sea. Fisheries Research, 37: 127–142.

Toresen, R., Gjøsæter, H., and de Barros, P. 1998. The acoustic method as used in the abundance estimation of capelin (*Mallotus villosus* Müller) and herring (*Clupea harengus*, Linné) in the Barents Sea. Fisheries Research, 34: 27–34.