# How Should The Sardine Length Frequencies Be Weighted? Exploring Alternative Methods

**C.L. de Moor* and D.S. Butterworth***

**Abstract**

The effect of alternative methods of weighting trawls when estimating the proportions-at-length in November surveys is examined. The current method used, in which the length frequencies from the individual trawls are weighted by the acoustic weighting associated by that trawl as a proportion of the total acoustic weighting of all trawls, is compared to methods in which the length frequencies of individual trawls are weighted equally or those with low sample number are down weighted. In addition, the effect of using only night time trawls in this weighting is investigated. We recommend that a method of weighting the length frequencies of trawls equally, except that those with sample sizes less than 40, which should be down-weighted proportional to their size, replace the current method used to calculate the survey length frequencies.

## Introduction

Length frequencies are calculated for the annual sardine November surveys using samples from trawls undertaken during the spawner biomass acoustic surveys and using a method of calculation that was developed by Roel and probably also Hampton (Coetzee pers. comm.). These length frequencies are used in estimating the acoustic biomass and the proportion-at-age at the time of the survey. de Moor and Butterworth (2009) developed an alternative method consisting of logical steps in which the weighting factor applied to each trawl was clear and easily modified. The weighting factor applied in de Moor and Butterworth (2009) mimicked that of the current method and the same proportions-at-length were obtained from both methods.

In this document, the method developed by de Moor and Butterworth (2009) is compared with alternative forms of trawl weighting factors. All results presented use data up to Port Alfred only, as recommended by de Moor and Butterworth (2009).

## Method

The method currently used to calculate the length frequency of the spawner biomass survey is detailed in Appendix A. In essence, the length frequencies from the individual trawls are weighted by the acoustic weighting associated by that trawl as a proportion of the total of the acoustic weighting of all trawls. A new method, termed "alternative method", has been developed with the following in mind (Appendix B, de Moor and Butterworth 2009):

* MARAM (Marine Resource Assessment and Management Group), Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch, 7701, South Africa. Email: carryn.demoor@uct.ac.za , doug.butterworth@uct.ac.za.

i) the logic behind calculating the weightings (equations A.2 to A.4) in the current method is unclear;

ii) the individual trawls and their proportions-at-length are weighted using the acoustic density in the manner assumed to underlie the current method;

iii) a method with logical steps was required, in which the weighting factor applied to each trawl was clear and easily modified given changed assumptions.

In order to test the effect of weighting the length frequencies of the individual trawls based on the number of fish in the sub-sample, further methods are also tested:

i) an "equal weighting" method in which all trawls are weighted equally. In this case, the weighting factor becomes $w_{sj} = \frac{1}{n_s}$ where $n_s$ is the number of trawls in the stratum and therefore equation (B.7) changes to:

$$p_{sl} = \sum_{j=1}^{ns} \left( p_{sjl}^{trawl} \cdot \frac{1}{n_s} \right)$$

ii) a "down weighted" method in which small trawl sub-samples are downweighted. Although scientists aim to sample 100 fish from each trawl (Coetzee pers comm.), the actual number of fish sub-sampled can vary considerably (Figure 1). As 100 fish is currently used as a 'cut-off' when sampling, the weighting factor remains $w_{sj} = \frac{1}{n_s}$ if $T_{s,j} > 100$, but is $w_{sj} = \frac{T_{s,j}}{100 * n_s}$ if $T_{s,j} \leq 100$, i.e. equation (B.7) changes to:

$$p_{sl} = \begin{cases} \sum_{j=1}^{ns} \left[ p_{sjl}^{trawl} \cdot \frac{1}{n_s} \right] & T_{sj} > 100 \\ \sum_{j=1}^{ns} \left[ p_{sjl}^{trawl} \cdot \frac{T_{sj}}{n_s \times 100} \right] & T_{sj} \leq 100 \end{cases}$$

iii) further "down weighted" methods with lower cut-off values; and

iv) the "equal weighting" method, but only using trawls that occurred during night time (taken as between 5.30pm and 5am). In this case, the weighting factor becomes $w_{sj} = \frac{1}{n_s^{night}}$ where $n_s^{night}$ is the number of night time trawls in the stratum and therefore equation (B.7) changes to:

$$p_{sl} = \sum_{j=1}^{n\_s\hat{}night} \left( p_{sjl}^{trawl} \cdot \frac{1}{n_s^{night}} \right)$$

v) the "down weighted" method using a cut-off value of 40, but only using trawls that occurred during night time. In this case, the weighting factor becomes $w_{sj} = \frac{1}{n_s^{night}}$ if $T_{s,j} > 40$, but is $w_{sj} = \frac{T_{s,j}}{40 * n_s^{night}}$ if $T_{s,j} \leq 40$, i.e. equation (B.7) changes to:

$$p_{sl} = \begin{cases} \sum_{j=1}^{n\_s^{\wedge}night} \left[ p_{sjl}^{trawl} \cdot \frac{1}{n_s^{night}} \right] & T_{sj} > 40 \\ \sum_{j=1}^{n\_s^{\wedge}night} \left[ p_{sjl}^{trawl} \cdot \frac{T_{sj}}{n_s^{night} \times 40} \right] & T_{sj} \leq 40 \end{cases}$$

The standard deviation and CVs of the estimated proportions-at-age (see Appendix C) and proportions-at-length are calculated from 1000 bootstrapped iterations (see Appendix D). This is done only for years in which survey age-length keys were available (1993, 1994, 1996, 2001, 2002, 2003, 2004, 2006), since the age-length-key using commercial data up to Port Alfred only was not easily obtainable. A few assumptions were made regarding the ALKs when a non-zero proportion was calculated for a length class, while the length class in the ALK had no age sample[1][2][3][4][5][6]. These assumptions are only made for the comparative purpose of this document and will not be assumed in future, i.e. in forthcoming assessments.

The latest sardine assessment model was not fit to the proportion-at-age data due to concerns in the ageing data (Cunningham and Butterworth 2007). However, the model predicted proportions-at-age (unchanged from that estimated by the model in Cunningham and Butterworth 2007) were compared to the proportions-at-age obtained using the alternative, equal weighting and down weighted methods. As a rough indicator of the contribution to a log likelihood, the following sum was calculated for each method:

$$\sum_{y} \sum_{a=1}^{5} \left\{ n_y \, p_{y,a,Nov}^{S} \left( \ln p_{y,a,Nov}^{S} - \ln\left(p_{y,a}^{S}\right) \right)^2 \right\}, \qquad y = 1993, 1994, 1996, 2001, 2002, 2003, 2004, 2006 \quad (1)$$

**Results and Discussion**

*Proportions-at-length*

The average (taken over the 8 years) estimated and bootstrapped median proportions-at-length obtained using the alternative, equal weighting and down weighted methods are compared in Figure 2a,b. The average bootstrapped standard deviations for each proportion-at-length are compared in Figure 3a, while Figure 4 shows these bootstrapped standard deviations for each individual year. The median proportions-at-length are very similar for all the methods. Figures 3a and 4 indicate that the alternative and equal weighting methods may have more cases of large standard deviations than the down weighted method (with alternative cut-off values). The bootstrapped standard deviation obtained using the down weighted method was compared to that

---

[1] All sardine in length classes less than that for which a 1-yer-old has been recorded will be assumed to be 1-year-olds (eg Lc 6-9cm in 1993, Lc 6-8cm in 1994, Lc 6-8.5cm in 1996 etc.)

[2] All sardine in length classes greater than that for which a 5+-year-old has been recorded will be assumed to be 5+-year-olds (eg Lcs 22-23cm in 1994, 21.5-23cm in 1996, 21-23cm in 2001 etc.)

[3] Lcs 9cm and 9.5cm in 2002 are assumed to be 1 year olds.

[4] Lc 18.5cm in 2003 are assumed to be 3 and 4 year olds in equal proportion.

[5] Lc 10.5cm in 2004 are assumed to be 1 year olds.

[6] Lc 21.5cm and 22cm in 2004 are assumed to be 5+ year olds

obtained using the alternative and equal weighting methods in Tables 1 and 2, respectively. Although the total difference decreases (indicating a smaller, more desirable, standard deviation) with decreasing cut-off value, a sample value of less than 40 could see small samples containing outliers have an unduly influential impact. In addition, there seems to be a clear difference between a large number of small (<40) sample sizes and those >100 (Figure 1). For this reason a cut-off number of 40 is preferred. The individual differences in the bootstrapped standard deviations between the alternative and equal weighting methods and the down weighted method with a cut-off value of 40 are shown in Figures 5 and 6, respectively. (Note here that a negative value indicates the standard deviation for the down weighted method is smaller.)

The average (taken over the 8 years) bootstrapped median proportions-at-length obtained using the equal and down weighted methods (with a cut-off value of 40) and using night trawls only in weighting, are shown to be very similar to the down weighted method (with a cut-off value of 40) using all trawls in Figure 2c. The corresponding average bootstrapped standard deviations for each proportion-at-length are compared in Figure 3b. These distributions indicate that the standard deviation is, on average, higher if only the night trawls are used when weighting the trawls.

*Proportions-at-age*

The average (taken over the 8 years) bootstrapped median proportions-at-age obtained using the alternative, equal weighting and down weighted methods are compared in Figure 7a. These bootstrapped medians are shown together with ±2SDs in Figure 8. It is clear that although the median proportions-at-age of the three methods are similar, the standard deviations around these estimates of proportions-at-age are lower for the equal and down weighted methods, compared to the alternative method (Figures 8 and 9a).

The average (taken over the 8 years) bootstrapped median proportions-at-age obtained using the equal and down weighted methods (with a cut-off value of 40) and using night trawls only in weighting, are shown to be very similar to the down weighted method (with a cut-off value of 40) using all trawls in Figure 7b. These bootstrapped medians are shown together with ±2SDs in Figures 8e, f. The standard deviations around these estimates of proportions-at-age are lower on average for the down weighted method (with a cut-off value of 40) using all trawls in weighting, compared to the equal and down weighted method (with a cut-off value of 40) that used only night trawls in weighting (Figure 9b).

The sum obtained from equation (1) was 8282 for the alternative method, 6293 for the equal weighting method, 6953 for the down weighted method with a cut-off value of 100 and 6602 for the down weighted method with a cut-off value of 40. When using night time trawls only, the sum obtained from equation (1) was 7186 for the equal weighting method and 7146 for the down weighted method with a cut-off value of 40. Although not the exact term used for the log-likelihood in an assessment due to the exclusion of a constant of proportionality parameter, this indicates that the assessment model predicted November survey proportions-at-age that are closer to those estimated using the equal and down weighted methods (based on weighting using all trawls) than those estimated using the alternative method or using only night trawls in the weighting. The model predicted proportions-at-age are compared to those obtained from the three methods using all trawls in weighting in Figure 10.

The bootstrapped standard deviation for the proportions-at-age obtained using the equal weighting method and the sum obtained from equation (1) were lower than those obtained using the down weighted method. However, as noted in the methods section, there were some assumptions made regarding the ALKs used to calculate the proportions-at-age. The comparisons of proportions-at-age are useful and informative, though the comparisons of proportions-at-length are more accurate for choosing between methods used to calculate length frequencies.

**Conclusions and Recommendations**

The equal and down weighted methods have been shown to perform better than the alternative method, having a smaller standard deviation around the estimated proportions-at-age and length. In addition, the estimated proportions-at-age obtained using the equal and down weighted methods were closer to those predicted by the latest full sardine assessment model than those estimated using the alternative method. In addition, the down weighted method using all trawls in weighting has been shown to perform better than the equal weighting and down weighted methods using night trawls only in the weighting, since it has a smaller average standard deviation around the estimated proportions-at-age and length and the estimated proportions-at-age are closer to those predicted by the latest full sardine assessment model. .

Given that the alternative method was developed using the same weighting as that in the current method used to generate survey length frequencies, and that the proportions-at-length in the length frequencies were the same for the alternative and current methods (de Moor and Butterworth 2009), we recommend:

<u>* that the down weighted method, with a cut-off value of 40 replace the current method used to calculate the survey length frequencies (Figure 11).</u>

The idea of down weighting small samples seems more intuitive since very low sample sizes should not receive as high a weight as others. There may be a concern that the low sample sizes may correspond to trawls of schools of larger fish, which are more able to avoid the trawl. In down weighting these low sample size trawls, one may therefore underestimate the number of large fish. However, the down weighting suggested herein is "weak", with a low cut off value (40) suggested.

This could also alter the length frequency and hence the associated acoustic target strength for each trawl of each survey, with the knock-on effect of a change in the overall estimated survey biomass. As a next step in this process we therefore recommend:

<u>* that the sardine November survey biomass be recalculated for a few years (eg a year of high, intermediate and low abundance) using the down weighted method with a cut-off value of 40.</u>

The effect on the survey biomass using the down weighted method in contrast to the current method should be compared for a few years first before a recommendation might be made to revise the target strength used for all years.

**Acknowledgements**

We thank Janet Coetzee for providing the current method used to weight the length frequencies and for comments on earlier parts of this work. Deon Durholtz is thanked for providing the ALKs and discussions on their use.

**References**

de Moor, CL and Butterworth, DS 2009. How Should the Sardine Length Frequencies be Weighted? MCM Document MCM/2009/SWG-PEL/12. 12pp.

Cunningham, CL and Butterworth, DS 2007. Base Case Assessment of the South African Sardine Resource. MCM Document MCM/2007/SEP/SWG-PEL/06. 30pp.

***Table 1***. *The annual and total difference, summed over all length classes, between the bootstrapped standard deviation of proportions-at-length for the alternative method, and for the down weighted method with alternative cut-off values.*

| | Cut-Off Value | | | | |
|---|---|---|---|---|---|
| Year | 100 | 60 | 40 | 30 | 20 |
| 1993 | -0.016 | -0.020 | -0.021 | -0.022 | -0.023 |
| 1994 | -0.079 | -0.082 | -0.083 | -0.087 | -0.091 |
| 1996 | -0.109 | -0.111 | -0.112 | -0.111 | -0.110 |
| 2001 | -0.069 | -0.069 | -0.068 | -0.067 | -0.069 |
| 2002 | -0.066 | -0.079 | -0.085 | -0.088 | -0.090 |
| 2003 | -0.130 | -0.110 | -0.114 | -0.116 | -0.117 |
| 2004 | -0.224 | -0.225 | -0.230 | -0.232 | -0.231 |
| 2006 | -0.088 | -0.107 | -0.125 | -0.140 | -0.158 |
| Total | -0.754 | -0.803 | -0.837 | -0.862 | -0.889 |

***Table 2.*** *The annual and total difference, summed over all length classes, between the bootstrapped standard deviation of proportions-at-length for the equal weighting method, and for the down weighted method with alternative cut-off values.*

| | Cut-Off Value | | | | |
|---|---|---|---|---|---|
| Year | 100 | 60 | 40 | 30 | 20 |
| 1993 | 0.002 | -0.001 | -0.003 | -0.004 | -0.004 |
| 1994 | 0.019 | 0.017 | 0.015 | 0.012 | 0.007 |
| 1996 | -0.072 | -0.074 | -0.075 | -0.075 | -0.073 |
| 2001 | -0.049 | -0.049 | -0.048 | -0.047 | -0.049 |
| 2002 | -0.023 | -0.037 | -0.043 | -0.045 | -0.048 |
| 2003 | -0.010 | -0.017 | -0.021 | -0.023 | -0.024 |
| 2004 | -0.061 | -0.062 | -0.067 | -0.069 | -0.068 |
| 2006 | 0.041 | 0.022 | 0.004 | -0.011 | -0.029 |
| Total | -0.153 | -0.202 | -0.236 | -0.261 | -0.288 |

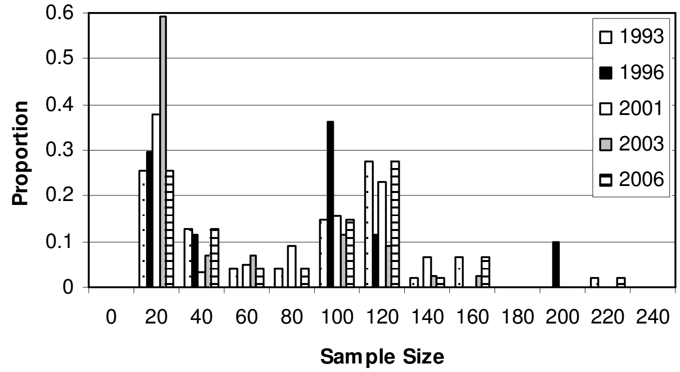



***Figure 1.*** *The histogram of the number of fish in each trawl sub-sample for selected years.*

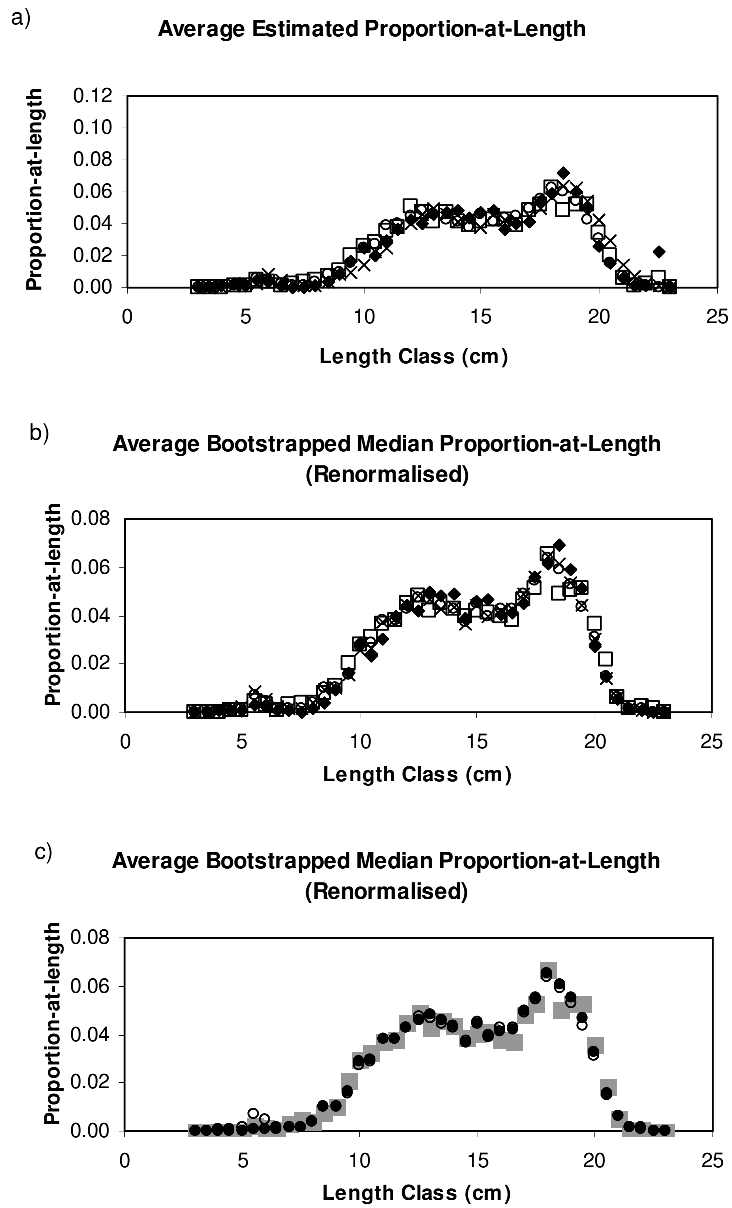



***Figure 2.*** *The average (over 8 years) a) estimated proportion-at-length and b) the renormalized bootstrapped median proportion-at-lengths for the alternative method (solid diamonds), the equal weighting method (open squares) and the down weighted method (cut-off 100 – crosses; cut-off 40 – open circles). The average (over 8 years) renormalized bootstrapped median proportion-at-lengths for the down weighted method (cut-off 40 – open circles), and the equal weighting method (grey squares) and the down weighted method (cut-off 40 – closed circles) using night trawls only in the weighting, are compared in plot c).*

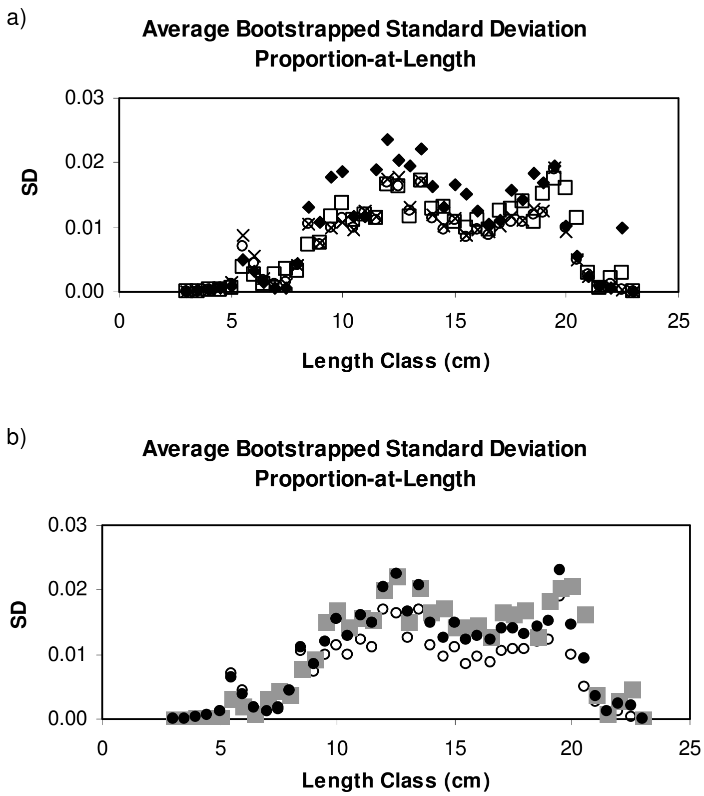



***Figure 3.*** *a) The average (over 8 years) bootstrapped standard deviation around the proportions-at-length for the alternative method (solid diamonds), the equal weighting method (open squares) and the down weighted method (cut-of 100 – crosses, cut-off 40 – open circles). b) The average (over 8 years) bootstrapped standard deviation around the proportions-at-length for the down weighted method (cut-off 40 – open circles), and the equal weighting method (grey squares) and the down weighted method (cut-off 40 – closed circles) using night trawls only in the weighting.*

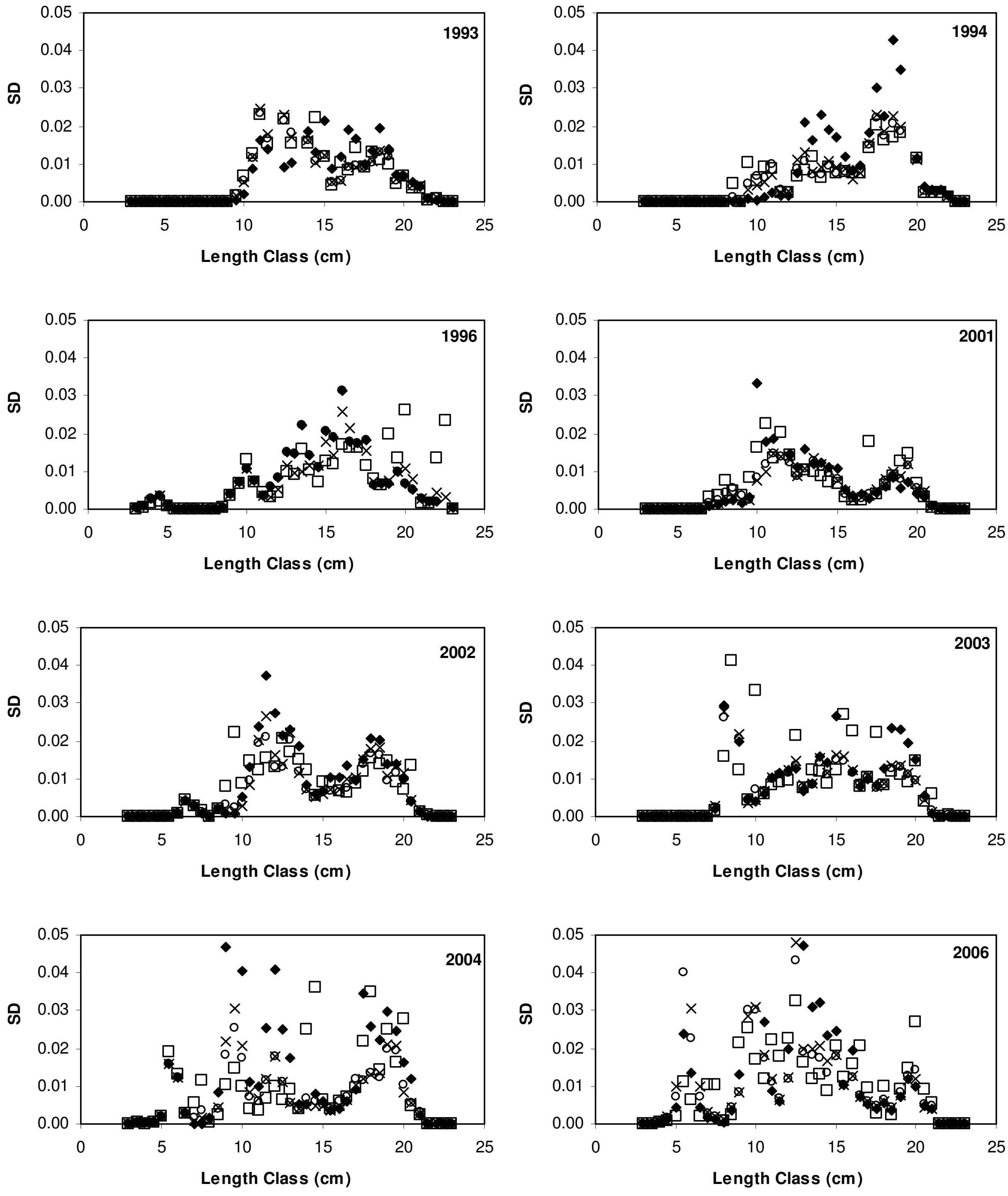



***Figure 4.*** *The annual bootstrapped median proportions-at-length by year for the alternative method (solid diamonds), the equal weighting method (open squares) and the down weighted method (cut-of 100 – crosses, cut-off 40 – open circles).*

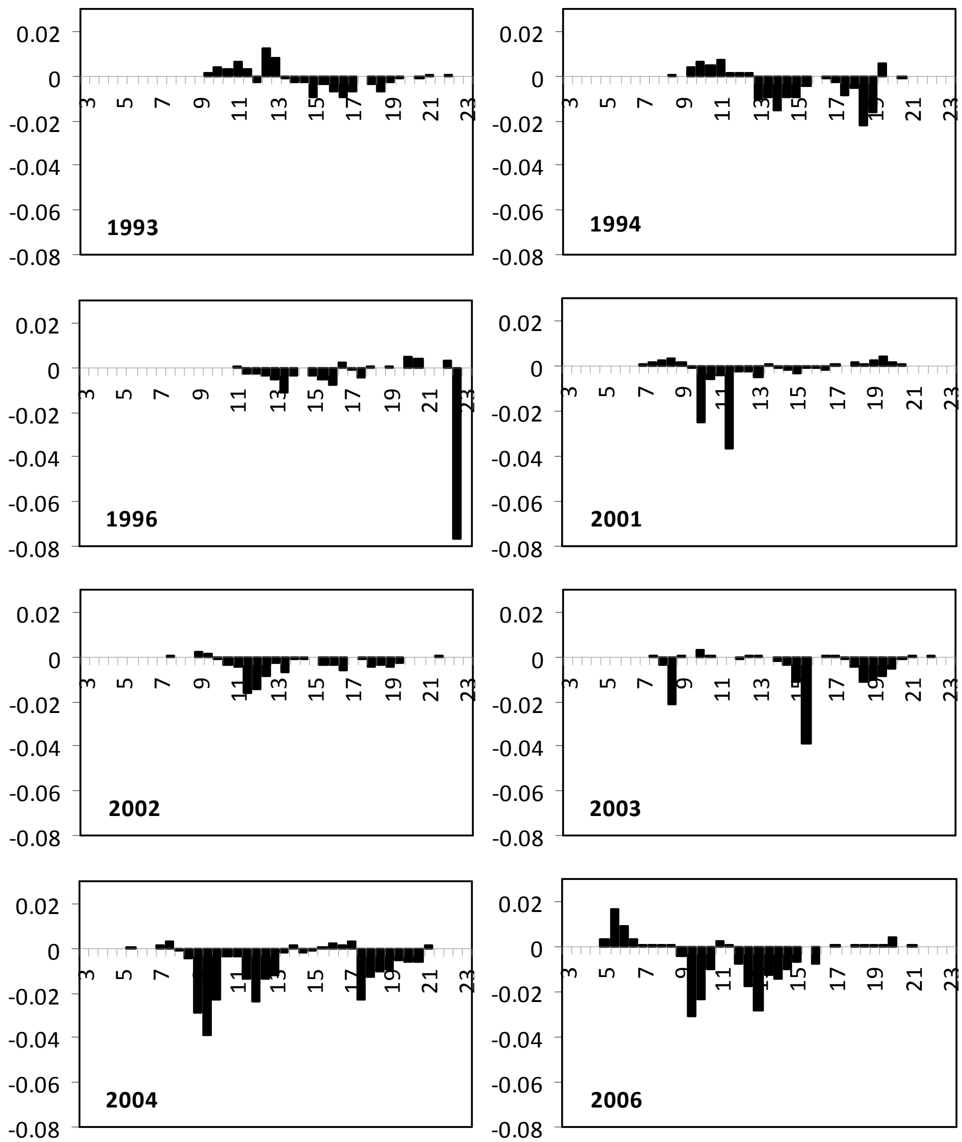



***Figure 5.*** *The difference each year between the bootstrapped standard deviation of proportions-at-length for the down weighted method with a cut-off value of 40 and the alternative method.*

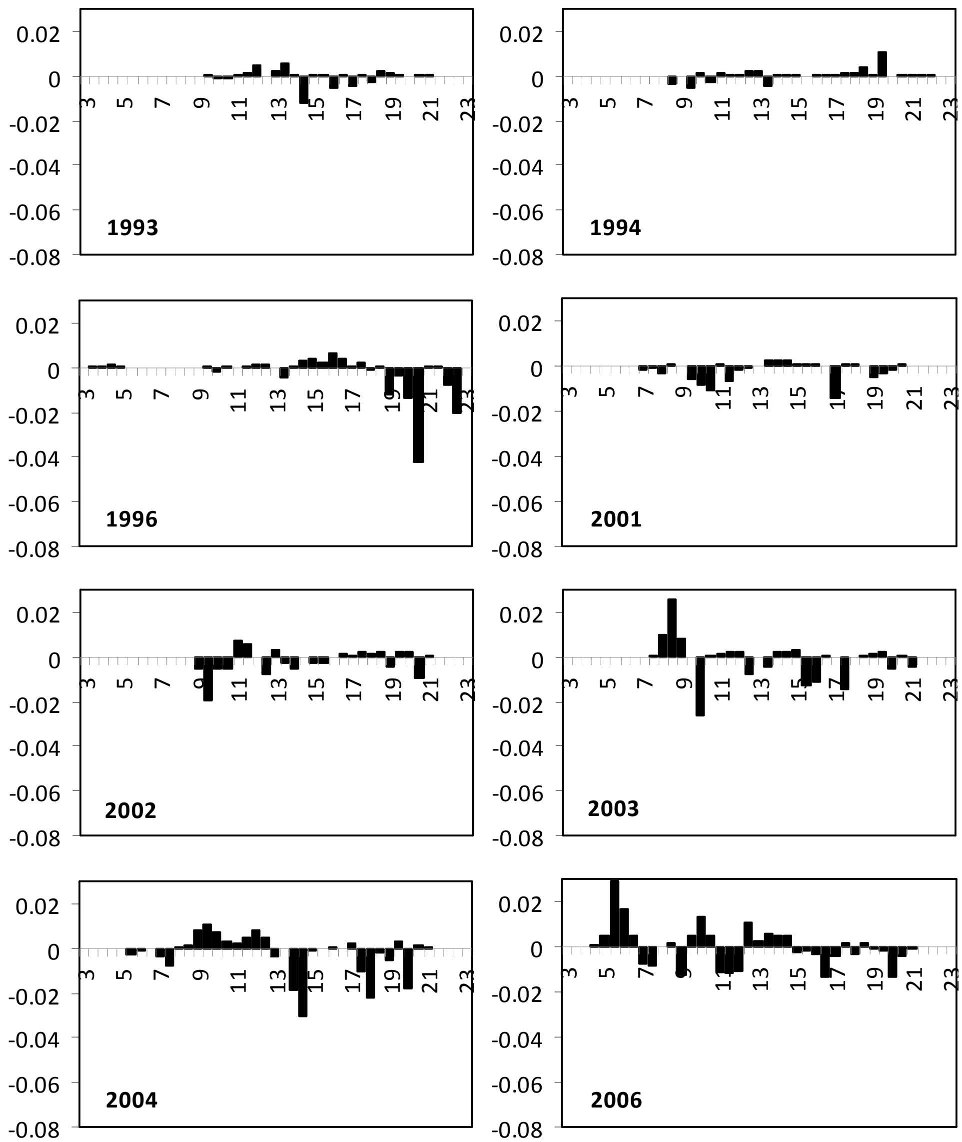



***Figure 6.*** *The difference each year between the bootstrapped standard deviation of proportions-at-length for the down weighted method with a cut-off value of 40 and the equal weighting method.*

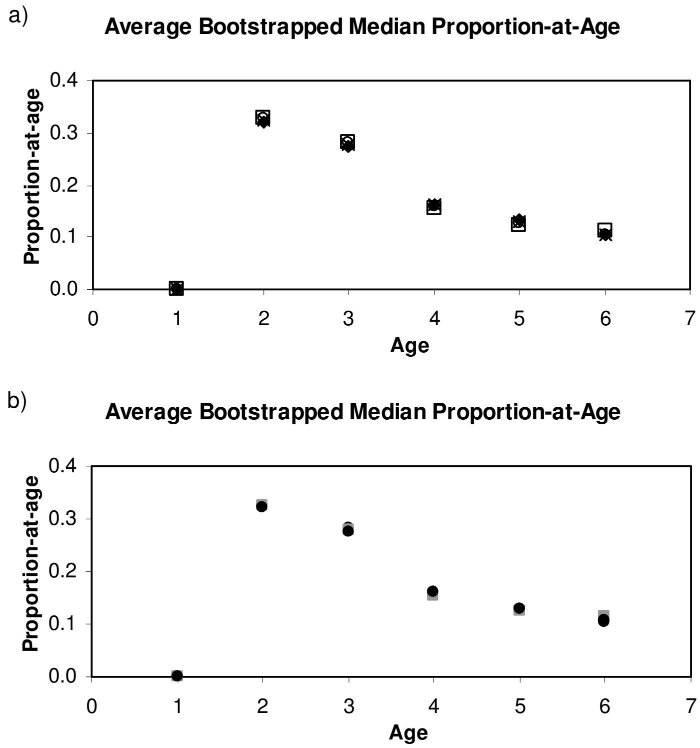



***Figure 7.*** *a) The average (over all years) bootstrapped median proportions-at-age for the alternative method (solid diamonds), the equal weighting method (open squares) and the down weighted method (cut-of 100 – crosses, cut-off 40 – open circles). b) The average (over 8 years) bootstrapped median proportions-at-age for the down weighted method (cut-off 40 – open circles), and the equal weighting method (grey squares) and the down weighted method (cut-off 40 – closed circles) using night trawls only in the weighting.*

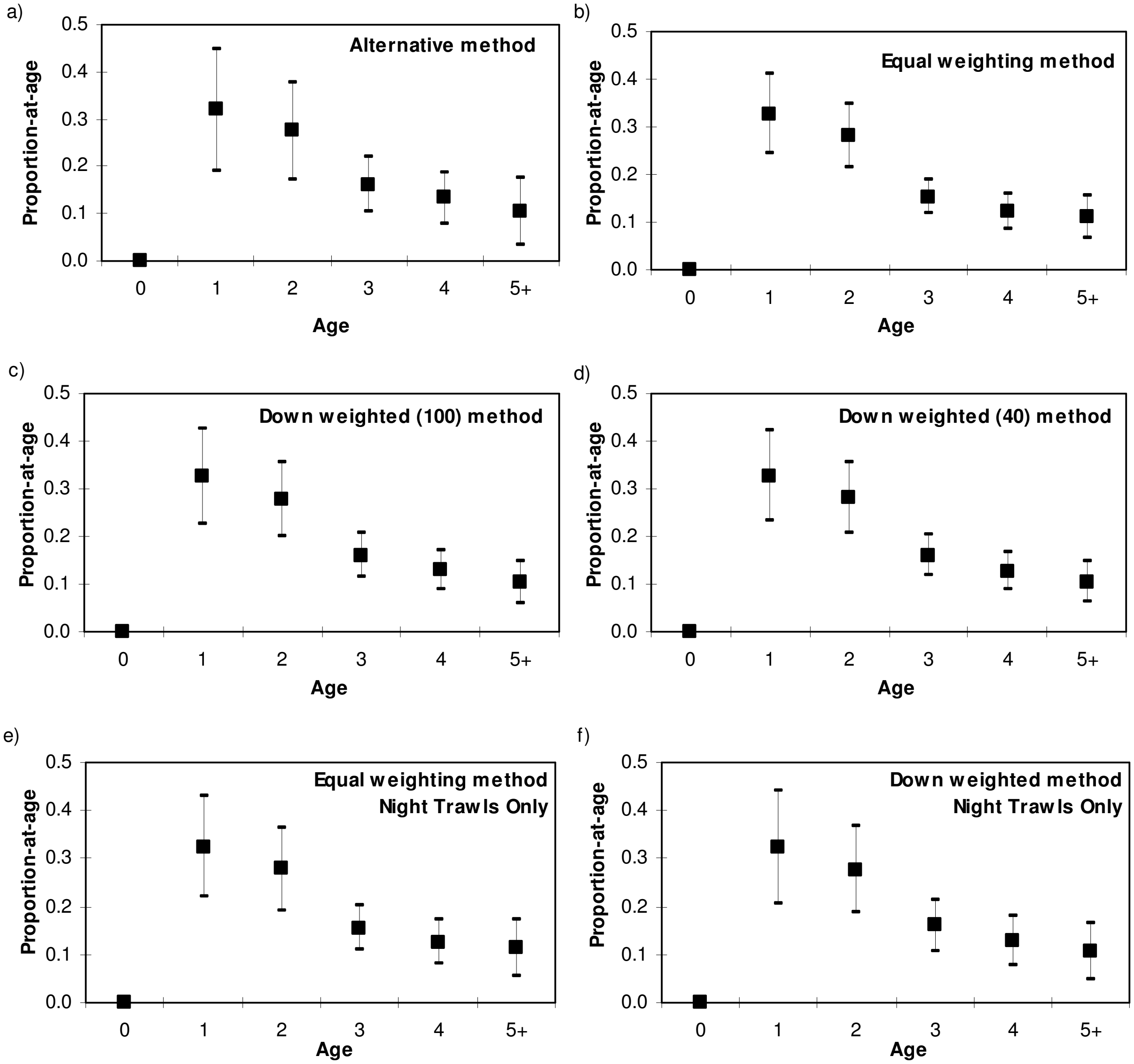
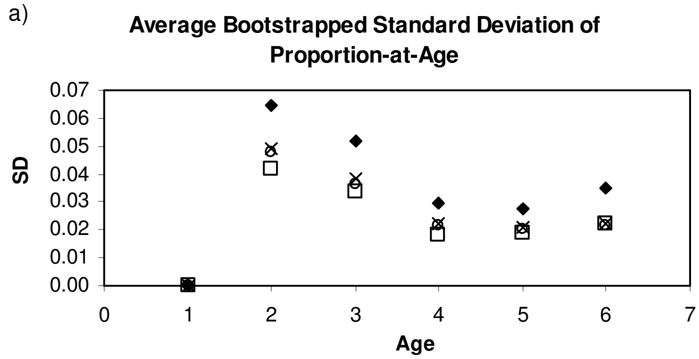



***Figure 8.*** *The average (over all years) bootstrapped median proportions-at-age using a) the alternative method, b) the equal weighting method, c) the down weighted method with a cut-off value of 100, d) the down weighted method with a cut-off value of 40, e) the equal weighting method using night trawls only and f) the down weighted method with a cut-off of 40 using night trawls only, shown with ±2 SDs.*

a)




b)

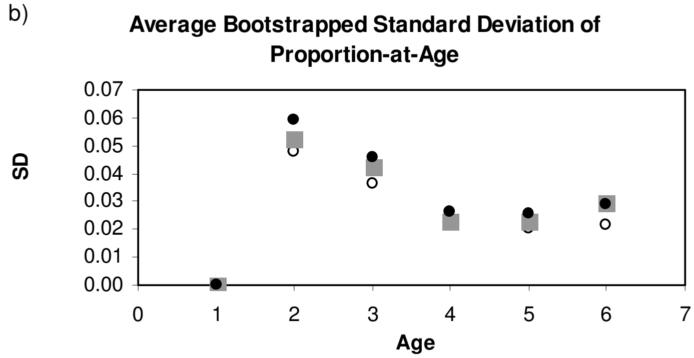



***Figure 9.*** *a) The average (over all years) bootstrapped standard deviation around the proportions-at-age for the alternative method (solid diamonds), the equal weighting method (open squares) and the down weighted method (cut-of 100 – crosses, cut-off 40 – open circles). b) The average (over 8 years) bootstrapped standard deviation around the proportions-at-age for the down weighted method (cut-off 40 – open circles), and the equal weighting method (grey squares) and the down weighted method (cut-off 40 – closed circles) using night trawls only in the weighting.*

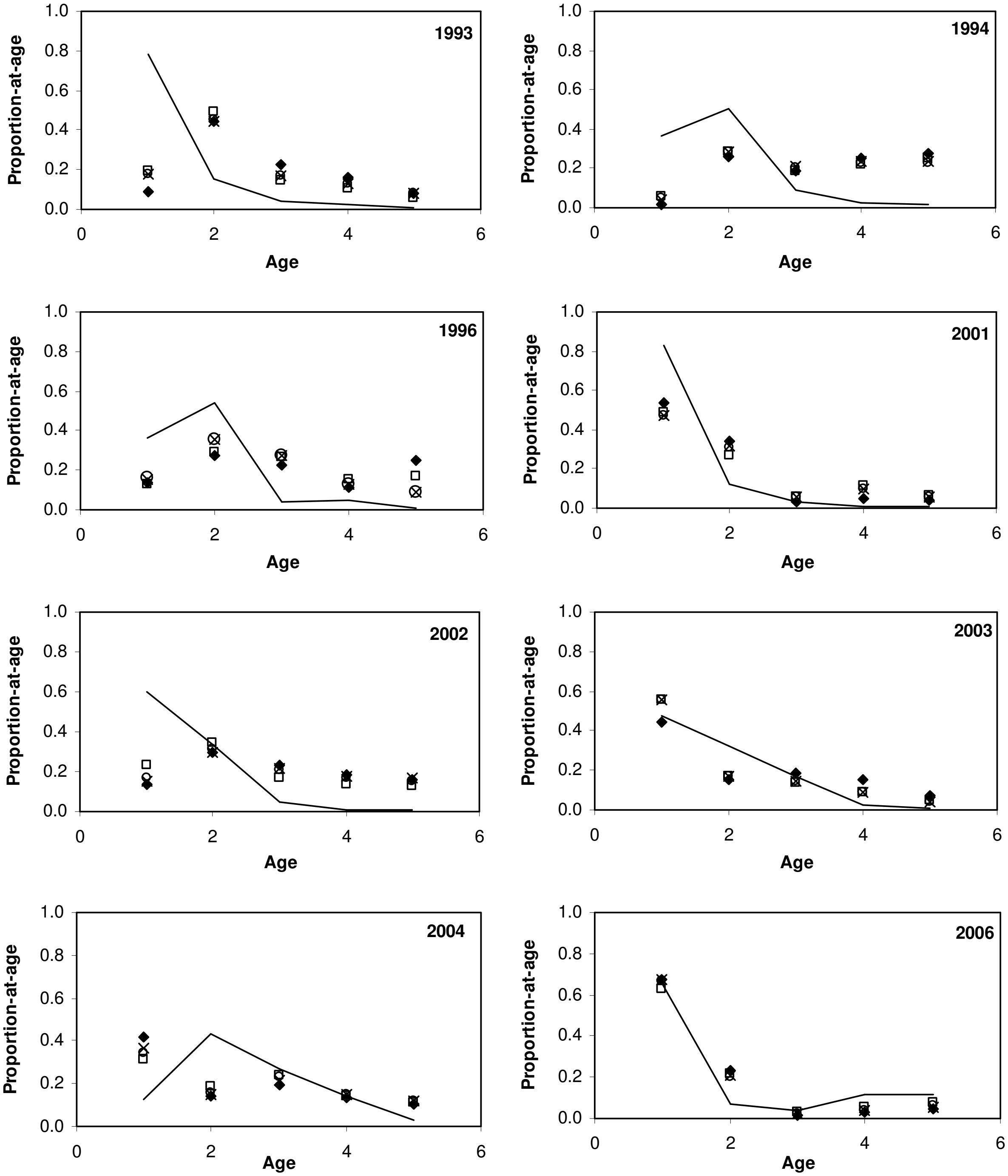



***Figure 10.*** *The proportions-at-age estimated using the alternative method (solid diamonds), the equal weighting method (open squares) and the down weighted method (cut-of 100 – crosses, cut-off 40 – open circles), compared to the model predicted proportions-at-age (line, Cunningham and Butterworth (2007))*

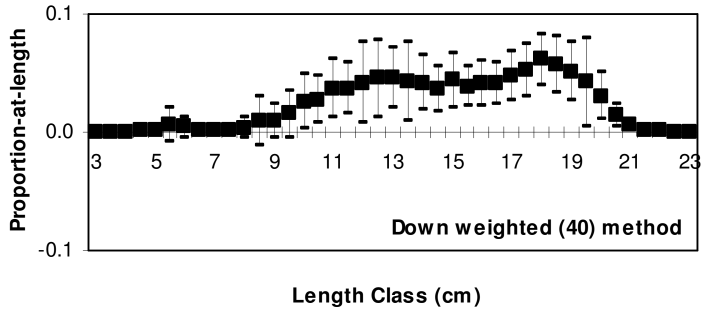



***Figure 11.*** *The average bootstrapped median proportion-at-lengths using the down weighted method with a cut-off value of 40, shown with ± 2 SDs.*

**Appendix A: The Current Method Used to Calculate the Weighted Length Frequencies**

The acoustic weighting for each trawl sample *j* in stratum *s* is given by:

$$Z_{sj} = \sum_i L_{sji} \cdot \rho_{sji} \qquad \text{(A.1)}$$

where

$L_{sji}$ is the mean interval length (nmi) for trawl sample *j* and interval (ESDU) *i* in stratum *s*, and

$\rho_{sji}$ is the mean acoustic interval density (g.m$^{-2}$) for trawl sample *j* and interval (ESDU) *i* in stratum *s*.

To weigh individual trawls, one needs to convert the acoustic weighting factor into a factor in terms of numbers. The trawl weighting factor is given by:

$$Q_{sj} = \frac{Z_{sj}}{W_{sj}} \qquad \text{(A.2)}$$

where

$W_{sj}$ is the total mass (kg) (also termed length frequency mass) of all the fish in the sub-sample from trawl *j* in stratum *s*.

The weighted length frequency in stratum *s* is the vector $\underline{\mathbf{T}}_s$, which has elements

$$T_{sl} = \sum_j T_{sjl} \cdot Q_{sj} \qquad \text{(A.3)}$$

where

$T_{sjl}$ is the number of fish of length *l* in the sub-sample from trawl *j* in stratum *s*.

The total number of fish of length *l* estimated by the survey to be in stratum *s* is then given by:

$$N_{sl} = \frac{T_{sl} \cdot B_s}{\sum_j Z_{sj}} \qquad \text{(A.4)}$$

where

$B_s$ is the biomass (kg) in stratum *s*.

The total number of fish of length *l* estimated by the survey is then given by:

$$N_l^{tot} = \sum_s N_{sl} \qquad \text{(A.5)}$$

**Appendix B: The Alternative Method to Calculate the Weighted Length Frequencies**

The acoustic weighting for each trawl sample *j* in stratum *s* is given by:

$$Z_{sj} = \sum_i L_{sji} \cdot \rho_{sji} \qquad \text{(B.1)}$$

where

$L_{sji}$ is the mean interval length (nmi) for trawl sample *j* and interval (ESDU) *i* in stratum *s*, and

$\rho_{sji}$ is the mean acoustic interval density (g.m$^{-2}$) for trawl sample *j* and interval (ESDU) *i* in stratum *s*.

From the total number of fish of all species sampled from a trawl, a sub-sample of a species of interest is drawn for the purposes of estimating length distribution. The total number of fish in this sub-sample from trawl *j* in stratum *s* is given by:

$$T_{sj} = \sum_l T_{sjl} \qquad \text{(B.2)}$$

where

$T_{sjl}$ is the number of fish of length *l* in the sub-sample from trawl *j* in stratum *s*.

The mean mass (kg) of an individual fish in the sub-sample from trawl *j* in stratum *s* is given by:

$$\overline{W}_{sj} = \frac{W_{sj}}{T_{sj}} \qquad \text{(B.3)}$$

where

$W_{sj}$ is the total mass (kg) of all the fish in the sub-sample from trawl *j* in stratum *s*.

The number of fish estimated by the survey to be in stratum *s* that are associated with trawl *j* is then given by:

$$N_{sj} = \frac{k_s \cdot Z_{sj}}{\overline{W}_{sj}} \qquad \text{(B.4)}$$

where $k_s = \dfrac{B_s}{\sum_j Z_{sj}}$ is the conversion factor (in kgs) which relates the estimate of biomass for the stratum, $B_s$ (NB: in kgs) to the sum of the acoustic weights.

The total number of fish estimated by the survey to be in stratum *s* is then calculated as follows:

$$N_s = \sum_j N_{sj} = k_s \sum_j \frac{Z_{sj}}{\overline{W}_{sj}} \qquad \text{(B.5)}$$

The proportion of fish of length *l* in the sub-sample from trawl *j* in stratum *s* is then given by:

$$p_{sjl}^{trawl} = \frac{T_{sjl}}{\sum_{l} T_{sjl}} \qquad \text{(B.6)}$$

and the proportion of all the fish of length *l* in stratum *s* is then estimated as:

$$p_{sl} = \sum_{j} p_{sjl}^{trawl} \cdot \frac{N_{sj}}{N_s} \qquad \text{(B.7)}$$

In this context, therefore, the proportion of fish of length *l* in stratum *s* is estimated from the proportion of fish of length *l* in the sub-samples taken from each trawl *j* in the stratum, weighted by the proportion of the total number of fish estimated to be in that stratum which is associated with trawl *j*, i.e. $w_{sj} = \frac{N_{sj}}{N_s}$.

The estimated number of fish of length *l* in stratum *s* is therefore

$$N_{sl} = p_{sl} \cdot N_s = \left( \sum_{j} p_{sjl}^{trawl} \cdot \frac{N_{sj}}{N_s} \right) \cdot N_s = \sum_{j} p_{sjl}^{trawl} N_{sj} \qquad \text{(B.8)}$$

and the estimated total number of fish of length *l* in the complete area surveyed is given by:

$$N_l^{tot} = \sum_{s} N_{sl} \qquad \text{(B.9)}$$

**Appendix C: The Calculation of the Proportion-at-Age in the Survey, given the Length Frequency**

The estimated total number of fish of age *a* in length class *l* in the complete area surveyed is given by:

$$N_{la} = \frac{n_{la}}{\sum_a n_{la}} N_l^{tot} \tag{C.1}$$

where

$n_{la}$ is the number of otoliths in length class *l* allocated to age *a*.

The estimated proportion by age of the total number in the complete area surveyed is then given by

$$p_a = \frac{\sum_l N_{la}}{\sum_a \sum_l N_{la}}. \tag{C.2}$$

**Appendix D: Bootstrapping**

For $b = 1,\ldots,1000$ resample with replacement the sets of data ($Z_{sj}^b, W_{sj}^b, T_{sjl}^b \; _{l=4,\ldots,23}$) as used in equation (C.1), (C.2) and (C.3) above[7].

As before, the total number of fish in this sub-sample from trawl *j* in stratum *s* is given by:

$$T_{sj}^b = \sum_l T_{sjl}^b \qquad \text{(D.1)}$$

The mean mass (kg) of an individual fish in the sub-sample from trawl *j* in stratum *s* is given by:

$$\overline{W}_{sj}^b = \frac{W_{sj}^b}{T_{sj}^b} \qquad \text{(D.2)}$$

The number of fish estimated by the survey to be in stratum *s* that are associated with trawl *j* is then given by:

$$N_{sj}^b = \frac{k_s^b \cdot Z_{sj}^b}{\overline{W}_{sj}^b} \qquad \text{(D.3)}$$

where $k_s^b = \dfrac{B_s}{\sum_j Z_{sj}^b}$ is the conversion factor (in kgs) which relates the estimate of biomass for the stratum, $B_s$ (NB: in kgs) to the sum of the acoustic weights.

The total number of fish estimated by the survey to be in stratum *s* is then calculated as follows:

$$N_s^b = \sum_j N_{sj}^b = k_s^b \sum_j \frac{Z_{sj}^b}{\overline{W}_{sj}^b} \qquad \text{(D.4)}$$

The proportion of fish of length *l* in the sub-sample from trawl *j* in stratum *s* is then given by:

$$p_{sjl}^{trawl,b} = \frac{T_{sjl}^b}{\sum_l T_{sjl}^b} \qquad \text{(D.5)}$$

and the proportion of all the fish of length *l* in stratum *s* is then estimated as:

$$p_{sl}^b = \sum_j p_{sjl}^{trawl,b} \cdot \frac{N_{sj}^b}{N_s^b} \qquad \text{(D.6)}$$

The estimated number of fish of length *l* in stratum *s* is therefore

$$N_{sl}^b = p_{sl}^b \cdot N_s^b \qquad \text{(D.7)}$$

and the estimated total number of fish of length *l* in the complete area surveyed is given by:

$$N_l^{tot,b} = \sum_s N_{sl}^b \qquad \text{(D.8)}$$

The estimated total number of fish of age *a* in length class *l* in the complete area surveyed is given by:

---

[7] Thus if there were originally *n* trawls in stratum *s*, then *j*=1…*n* sets of data are resampled with replacement from the trawls of stratum *s*.

$$N_{la}^{b} = \frac{n_{la}}{\sum_{a} n_{la}} N_{l}^{tot,b} \qquad \text{(D.9)}$$

and the estimated proportion by age of the total number in the complete area surveyed is then given by

$$p_{a}^{b} = \frac{\sum_{l} N_{la}^{b}}{\sum_{a}\sum_{l} N_{la}^{b}} . \qquad \text{(D.10)}$$