Northwest Atlantic Fisheries Organization

Serial No. N5411 **NAFO SCR Doc. 07/59**

SCIENTIFIC COUNCIL MEETING – JUNE 2007

A nonparametric bootstrap of the 2006 XSA assessment for Greenland Halibut (*Reinhardtius hippoglossoides*) in NAFO Subarea 2 + Divisions 3KLMNO using Fisheries Libraries in R (FLR)

David C. M. Miller and Peter A. Shelton

Department of Fisheries and Oceans, Northwest Atlantic Fisheries Centre,
P.O. Box 5667, St. John's, NL, A1C 5X1, Canada,

Abstract

Non-parametric bootstrap is applied to XSA (Extended Survivors Analysis) estimates from the 2006 NAFO assessment of the 2+3KLMNO Greenland halibut stock. Analyses are carried out within the FLR (Fisheries Libraries in R) environment. We found that XSA estimates using FLR are identical to those using the standard Lowestoft Assessment Suite, but that the standard errors and scaled weights of indices are different. This may indicate that further work is required with regard to the XSA code within FLR. The means and the medians from the bootstrap distributions are close to the XSA point estimates for ages up to age 7. For ages 8 to 10, the XSA point estimates are lower than the bootstrap means and medians, but this is reversed for ages 11 to 14. The degree of shrinkage has a major effect on the point estimates, whereas the bootstrap estimates are relatively robust. With less shrinkage, the point estimates for ages 8-10 decrease whereas those for ages 11 to 14 increase. This sensitivity to shrinkage is of concern and needs to be investigated further. The internal standard errors from the XSA greatly underestimate CV for the older ages (>10) relative to the bootstrap estimates, most likely due to shrinkage applied in the formulation of the model. Bootstrap estimates indicate that XSA bias is low up to and including age 7, but ages 8-10 have large positive bias and ages 11-14 have large negative bias. This distinct pattern of bias by age needs to be examined more closely to determine causes. Trying to correct such large biases using the bootstrap bias correction percentile technique would not seem to be appropriate. Efforts should rather be focused on improving the efficiency of the XSA estimator with respect to the application to Greenland halibut.

Keywords: Greenland halibut, XSA, Bootstrap, Bias, Bias correction, Rebuilding Plan, Projections

Introduction

Greenland halibut in NAFO Subarea 2 + Divisions 3KLMNO is one of the last significant commercial groundfish stocks on the Grand Banks and it is important that the NAFO Fisheries Commission consider the risk associated with alternative management strategies. One of the major sources of uncertainty in the provision of scientific advice is the uncertainty associated with the estimates of survivors in year *t*, based on data up to year *t*-1, used to provide TAC advice for year *t*+1. For a number of years the annual assessment has been based on Extended Survivors Analysis (XSA; Shepherd, 1999) fitted within the Lowestoft assessment suite (Darby and Flatman, 1994). In the 2003 assessment Darby *et al.* (2003) applied a nonparametric bootstrap procedure to generate 1 000 fits of the model to the survey data in order to estimate uncertainty in the XSA parameter estimates and computed time series. The residuals of log catchability from the fitted model were sampled independently for each age, with replacement, and new survey indices were computed for each year and age. The XSA model was re-fitted to the new survey indices and new population and exploitation parameters estimated. Bias corrected percentiles were generated for each parameter and the population time series. The nonparametric bootstrap seems a reasonable approach for quantifying the uncertainty, particularly in the case of XSA which is a non-statistical approach to model fitting. It is also useful for examining the bias in estimates and, where necessary, applying bias-correction.

Although nonparametric bootstrap estimates were used in the 2003 and 2004 assessments, this approach was dropped for 2005 and 2006 in favor of internal standard errors computed within XSA by the Lowestoft assessment suite (e.g. Healey and Mahé, 2006). Since XSA does not formally involve a statistical fitting procedure, it is not immediately clear to most fisheries scientists how the internal standard errors are computed and whether or not they represent a valid estimate of uncertainty in the estimates. Further, no covariance in the estimates is provided. We support a return to the nonparametric bootstrap approach to provide bias corrected estimates of survivors as well as confidence intervals for the Greenland halibut stock assessment. We also support the use of bootstrap samples for stochastic projections of the stock under alternative management options once bias correction problems are resolved.

A new stock assessment environment, Fisheries Libraries in R, or FLR (http://flr-project.org), has been developed in order to provide an integrated suite of software that allows data exploration, conditioning of models (the estimation of parameters consistent with the data and hypotheses about how these were generated), implementation of stock assessments (e.g. methods for stock assessments and forecasts) and the testing of management strategies and economic impact assessments, all within a common environment. FLR allows XSA to be carried out with enhanced features useful in management strategy evaluation (MSE). However, at the present time no bootstrap component has been written for XSA within FLR. In this paper we provide the R-code and results for a bootstrap application to the 2006 Greenland halibut assessment data. We compare our results with those obtained in the 2006 assessment. To support the use of the nonparametric bootstrap procedure, we provide a description of how the method is applied and the interpretation of results. This was lacking in the previous application to the Greenland halibut stock. Our approach also differs somewhat from Darby *et al.* (2003) in that they resampled residuals within age whereas we examine the effect of different methods of resampling residuals: across ages and years under the assumption that the residuals are independent and identically distributed (i.i.d.), standardizing residuals and then resampling across ages and years, and resampling within age and year.

Methods

Extended Survivors Analysis

The inputs for the XSA-bootstrap analysis were identical to those used in the 2006 assessment by Healey and Mahé, (2006). Catch numbers at age obtained from Canadian age-length keys for 1975-2005 were used. Catch weights-at-age are computed as weighted means of the values from national sampling (in Healey and Mahé, 2006; Tables 2 and 3 respectively).

Data from three surveys were available as age disaggregated indices of population abundance (mean numbers per tow, MNPT):

1. EU 3M - a European Union summer survey in Div. 3M from 1995-2005, ages 1-12 (González Troncoso *et al.*, 2006).
2. Can 2J+3K autumn survey, true Campelen data from 1996-2005, ages 1 to 14 (Healey *et al.*, 2006).
3. Can 3LNO spring survey, true Campelen data from 1996-2005, ages 1 to 8 (Healey *et al.*, 2006).

The XSA model was fitted using the FLXSA component of FLR. The XSA settings (FLXSA control group) are presented in Table 1. Survivors on 1 January 2006 (i.e. in the terminal year), 5+ (exploitable biomass), 10+ (as a proxy for spawner biomass), recruitment (numbers at age 1) and the mean fishing mortality (F) for ages 5-10 were calculated from the XSA results.

XSA Residuals

The residuals in XSA are calculated as follows:

For each index I, at age a, and year y:

$$r_{i,a,y} = \log(N_{i,a,y}) - \log(N_{XSA,a,y}) \qquad (1)$$

Where:

$r_{i,a,y}$ = the residual for index *i* in year *y* at age *a*,

$N_{i,a,y}$ = observed numbers at age *a* at the beginning of year *y* for index *I*,

$N_{XSA,a,y}$ = XSA predicted numbers at age *a* at the beginning of year *y*.

Also:

$$N_{i,a,y} = \frac{I_{i,a,y}}{\hat{q}_{i,a}} \qquad (2)$$

and the XSA predicted index (the XSA predicted numbers at age in terms of MNPT for index *i*) is:

$$\hat{I}_{i,y,a} = N_{XSA,a,y}\hat{q}_{i,a}$$

Therefore:

$$N_{XSA,a,y} = \frac{\hat{I}_{i,a,y}}{\hat{q}_{i,a}} \qquad (3)$$

Where;

$I_{i,a,y}$ = observed MNPT for index *i* in year *y* at age *a*,

$\hat{I}_{i,a,y}$ = XSA predicted MNPT for index *i* in year *y* at age *a*,

$\hat{q}_{i,a}$ = the XSA predicted catchability for index *i* at age a, common across all years.

It follows from (1) that:

$$\begin{aligned} r_{i,a,y} &= \log\left(\frac{I_{i,a,y}}{\hat{q}_{i,a}}\right) - \log\left(\frac{\hat{I}_{i,a,y}}{\hat{q}_{i,a}}\right) \\ &= \log\left(\frac{I_{i,a,y}\hat{q}_{i,a}}{\hat{I}_{i,a,y}\hat{q}_{i,a}}\right) \qquad (4) \\ &= \log\left(\frac{I_{i,a,y}}{\hat{I}_{i,a,y}}\right) \end{aligned}$$

Hence residuals are independent of the catchability at age. This is important given that the three indices have notably different catchabilities at age (Healey and Mahé, 2006).

The distribution of residuals for each of the three indices and the distribution of residuals by age are shown in Figs. 1 and 2, respectively. It appears that the residuals are very similarly distributed across age and index and therefore, for the purposes of the nonparametric bootstrap, they could be considered to be independent and identically distributed (i.i.d.). Standardizing the residuals is likely to result in distributions that more completely satisfy the i.i.d. assumptions by making standard deviations between age and index groups comparable.

Non-parametric bootstrapping of the XSA model

A nonparametric bootstrapping procedure was written in R (see Appendix for the R code) to generate 1000 fits ('bootstrap replicates') of the XSA model to the survey data. This was done in order to estimate the uncertainty in the XSA parameter estimates. The nonparametric bootstrap only accounts for observation error. Patterson *et al.* (2000) noted that variance in other parameters such as maturity, natural mortality and catch-at-age will also contribute to the uncertainty. We have not accounted for these additional sources of uncertainty in the current analysis. Although not accounting for all uncertainty, the short-term projections based on nonparametric bootstrap projections under different management options do provide reasonable relative measures of the risks associated with harvesting practices and should lead to better informed management decisions (Darby *et al*. 2004). They also allow for the construction of risk profiles by creating cumulative frequency distributions of the estimated fisheries management quantities of interest (Gavaris, 1999).

The nonparametric bootstrap was implemented as follows. By taking the exponent of both sides of equation (4) we get:

$$\exp(r_{i,a,y}) = \frac{I_{i,a,y}}{\hat{I}_{i,a,y}}$$

Therefore:

$$I_{i,a,y} = \exp(r_{i,a,y})\hat{I}_{i,a,y} \tag{5}$$

The residual $r_{i,a,y}$ can be replaced by r_{bs}, a residual randomly resampled, with replacement, from a given set of residuals. The residuals were resampled in three ways:

1. Across all indices and ages, assuming residuals are i.i.d.
2. By standardizing residuals and then resampling across all ages and indices.
3. By only resampling residuals within each age and index.

Sampling residuals randomly across all indices in the bootstrap is the most straightforward approach and appears *a priori* to be consistent with the i.i.d. assumption. Resampling within each age and index often results in only a limited number of residuals being available for resampling and may give inaccurate estimates of uncertainty. However, because XSA weights indices in the estimation, residuals for heavier weighted indices should be smaller than those from indices with a lesser weighting. Therefore, residuals will not be i.i.d. across indices. Also, difficulty in accurately aging older fish (Treble and Dwyer, 2006), and differences in the ability of surveys to sample fish of different ages, means that residuals may not be i.i.d. across ages either. Standardizing residuals, by dividing by the standard deviation for residuals from each age for each index, should make residuals comply with the assumptions of i.i.d.

The residuals were standardized as follows:

$$r'_{i,a,y} = \frac{\log(I_{i,a,y}) - \log(\hat{I}_{i,a,y})}{\hat{\sigma}_{i,a}} \tag{6}$$

Where: $r'_{i,a,y}$ = The standardized residuals for index *i*, age *a* and year *y*,

$\hat{\sigma}_{i,a}$ = The standard deviation of residuals for index *i* and age *a*.

The new pseudo-index values I' for each index type, age and year can be computed by adding these randomly selected residuals to each XSA predicted index value $\hat{I}_{i,a,y}$

$$I'_{i,a,y} = \exp(r_{bs})\hat{I}_{i,a,y} \tag{7}$$

Or, in the case of the standardized residuals:

$$I'_{i,a,y} = \exp(r'_{bs} \times \hat{\sigma}_{i,a})\hat{I}_{i,a,y} \tag{8}$$

Where: r'_{bs} is a randomly resampled residual.

The XSA model can then be re-fitted to the new survey pseudo-indices and the population size and fishing mortality parameters re-estimated.

Measures of uncertainty

Bootstrap standard error and percentile distributions of parameter estimates can be calculated for parameters from the set of new bootstrap XSAs and bias correction can be considered. Efron (1981) noted that a bootstrap sample size of 512 was more than sufficient for routine error analyses.

For a given parameter θ, let B be a set of n bootstrap replicate estimates θ^{BS}. The standard error of a sample of bootstrap estimates is equivalent to the standard deviation of the sample (Efron and Tibshirani, 1993):

$$SE_B = \sqrt{\frac{\sum_{b=1}^{n}\left[\theta_b^{BS} - \vec{\theta}^{BS}\right]^2}{n-1}} \tag{9}$$

Where:
$\vec{\theta}^{BS}$ = the mean of the bootstrap replicates, θ^{BS}.

The bootstrap co-efficient of variation (CV_B) is calculated as follows:

$$CV_B = \frac{SE_B}{\vec{\theta}^{BS}} \tag{10}$$

The bootstrap co-efficient of variation (CV_B) was calculated for the survivors at age in the terminal year and compared to the internal standard error from the original XSA, used as an estimate for CV in the stochastic projection done by Healey and Mahé (2006).

Detecting bias using bootstrap replicates

Ideally, estimators will be unbiased. However, estimators of fisheries management quantities from stock assessment models may be biased as a result of the non-linearity in the models (Gavaris, 1999). When the bias is large this may indicate serious problems with the model formulation. Bootstrap techniques (Efron and Tibshirani, 1993) provide a method for examining bias in estimates. For a given parameter θ, let B be a set of n bootstrap replicate estimates θ^{BS}. Then the bias of B, $bias_B$, can be calculated as follows:

$$bias_B = \frac{\sum_{b=1}^{n} \theta_b^{BS} - \hat{\theta}}{n} \qquad (11)$$

In other words, the bias of *B* is the mean of the differences between each bootstrap estimate, θ^{BS}, and $\hat{\theta}$, the best estimate of *θ*. Given that bootstrap distributions are skewed, due to fitting indices on a logarithmic scale in the XSA, one might think it best to compare the best estimate to the median rather than mean of the distribution. However computationally, comparing the best estimate with each bootstrap replicate individually and taking the mean of that difference as the bias (equation 10), a reasonable approach, is equivalent to comparing the mean of the distribution with the best estimate.

Bootstrap bias correction

When bias is detected in an estimated parameter, bias-correction is an issue for consideration. Efron and Tibshirani (1993) suggest that bias correction is not necessary if the bias is less than 25% of the bootstrap sample standard error, SE_B. When the bias is large, it is not clear to us whether or not bias correction is appropriate since a large bias may indicate serious problems with the model formulation and bias correction would just cover up the problem. In such a case it may be more beneficial to reassess the original model for possible sources of error rather than correcting the substantial bias. However, in cases where bias is large but not too severe, it is assumed that bias correction will provide more accurate parameter estimates. We explored bias-correction methods for XSA point estimates and distributions of bootstrap replicates.

The idea behind using the bootstrap replicates to correct for bias is that for any given parameter, on average, the mean bootstrapped XSA will over- or underestimate the original XSA by the same amount that the original XSA over- or under-estimates reality. The formulation of the XSA model may result in the distribution of bootstrap replicates of a given parameter *θ* being centered either to the low or high side of $\hat{\theta}$, the best estimate of *θ* (i.e. $\hat{\theta}$ is the estimate of *θ* derived by XSA from the original index values). The bootstrap replicates can therefore be used to calculate the bias of the XSA point estimates (equation 10), which can then be used to bias-correct them. The bias corrected estimate of *θ*, $\bar{\theta}$, is calculated as follows:

$$\bar{\theta} = \hat{\theta} - bias_B \qquad (12)$$

The ratio of bias to standard error ($\mathrm{bias_B/SE_B}$) can then be calculated for all the estimated parameters and compared with the 25% criterion of Efron and Tibshirani (1993).

If a bias in the point estimate of a parameter is detected, then it implies that the bootstrap distribution is also biased. Given that the bootstrap mean over- or underestimates the best estimate by the same amount that the best estimate over- or underestimates reality, the bootstrap mean is twice as biased as the best estimate. Gavaris (1999) found that the most suitable method for producing non-biased distributions from bootstrap replicates was the bootstrap bias corrected percentile technique. First the percentile method is used to form an empirical cumulative frequency distribution:

$$\hat{\Omega}(x) = \mathrm{P}\{\hat{\eta} \le x\} = \frac{\#\{\hat{\eta}^b \le x\}}{B} \qquad (13)$$

Where:

$\hat{\eta}$ = Quantity being estimated (e.g. number-at-age),

$\hat{\eta}^b$ = Bootstrap replicate of $\hat{\eta}$,

$\#\left\{\hat{\eta}^{b} \leq x\right\}$ = the number of bootstrap replicates less than or equal to *x*,
B = Total number of bootstrap replicates.

This can be considered as a set of paired values ($\alpha, \hat{\eta}^{b}$) where $\hat{\eta}^{b}$ are the ordered bootstrap values and *α* are the respective probability levels equal to 1/*B*, 2/ *B*, 3/ *B* … *B* / *B*.

Using this cumulative frequency distribution, each bootstrap replicate, $\hat{\eta}^{b}$, is bias corrected, $\vec{\eta}_{BC}^{b}$, to produce a bias corrected bootstrap distribution:

$$\vec{\eta}_{BC}^{b} = \hat{\Omega}^{-1}\left(\Phi\left(2z_{0} + z_{\alpha}\right)\right) \tag{14}$$

Where:
Φ, $\hat{\Omega}$ = Cumulative frequency distribution of the bootstrap distribution and a standard normal variate, respectively (i.e. the probability of a given critical value).

Φ^{-1}, $\hat{\Omega}^{-1}$ = Inverse distribution functions (i.e. critical value corresponding to the specified probability level).

$z_{\alpha} = \Phi^{-1}\left(\alpha\right)$ where *α* is the probability level of a given bootstrap value, $\hat{\eta}^{b}$.
$z_{0} = \Phi^{-1}\left(\hat{\Omega}\left(\hat{\eta}\right)\right)$

This last factor determines the bias. If median($\hat{\eta}^{b}$) = $\hat{\eta}$ then $\hat{\Omega}\left(\hat{\eta}\right)$=0.5. Therefore z_0 = 0 and $\vec{\eta}_{BC}^{b}$ = $\hat{\eta}^{b}$. i.e. no bias correction will occur.

Results and Discussion

We found that XSA estimates using FLR are identical to those using the Lowestoft Assessment Suite, but that the standard errors and scaled weights of indices are different. This needs to be investigated further, but we assume that the Lowestoft suite estimates are correct and that the FLR code requires further attention. Our results are based on comparing bootstrap XSA estimates within FLR with those from the NAFO 2006 assessment made using the Lowestoft Suite. We do not use the FLR version of XSA estimates of standard errors.

In Table 2 we compare the XSA survivors at age for the start of 2006 with those obtained from the nonparametric bootstrap analysis. The means and the medians from the bootstrap distributions are close to the XSA point estimates for ages up to age 7. For ages 8 to 10 the XSA point estimates are lower than the bootstrap means and medians, but this is reversed for ages 11 to 14. For younger ages the CVs from the two methods are similar, but the bootstrap CV indicates substantially wider confidence intervals than those estimated by the XSA (based on internal SE) for ages 11 to 14.

We investigated whether these differences were related to shrinkage in some way by varying the permitted standard error of *F* when shrinking to mean *F* (Fse). In Table 3 we compare estimates with Fse = 0.5 (same as Table 2) and 1.0 (less shrinkage effect) based on 1000 runs. It is clear that changing the degree of shrinkage has a major effect on the point estimates, whereas the bootstrap estimates are relatively robust. With less shrinkage (higher Fse), the point estimates for ages 8-10 decrease whereas those for ages 11 to 14 increase. This sensitivity to shrinkage is of concern and needs to be investigated further.

Table 3 also provides the bias corrected estimates of survivors based on the bootstrap results. A negative bias-corrected estimate for age 9 at the lower level of shrinkage (higher value of Fse) is clearly not feasible. Table 4 provides results for the same shrinkage comparison, but in this case for the 2006 estimates of 5+ biomass, 10+ biomass and average *F*. When aggregated across ages, impact of shrinkage is less evident because of the reversal of effects on ages 8-10 vs. 11-14. Given the increase in the older, heavier ages with greater shrinkage, there is however a considerable increase in the 10+ biomass.

The difference between the bootstrap estimates and the XSA point estimates of exploitable (5+) biomass and 10+ biomass are illustrated in Figs. 3 and 4. For 5+ biomass the XSA point estimates are close to the median of the bootstrap estimates. For the 10+ biomass the XSA point estimates are also close to the median of the bootstrap distribution for the early 2000s, but by 2006 come to lie closer to the 25th percentile. The confidence intervals for the 5+ biomass are relatively narrow compared to the 10+ biomass. The point estimates from XSA for numbers at age 1 (recruitment) are close to the median of the bootstrap distribution (Fig. 5). The confidence intervals are relatively wide, particularly for the period in the early 2000s. The XSA point estimates for average *F* fall above the 95th percentile of the bootstrap estimates in 2005 (Fig. 6). The confidence intervals are relatively wide for the last two years.

In addition to the estimates of survivors and the resulting estimates of 5+ and 10+ biomass, it is interesting to compare the bootstrap estimates of fishing mortality reference points with those applied in the last assessment. In Fig. 7 the distribution of 200 bootstrap estimates of $F_{0.1}$ and $F_{\max}$ are shown together with the point estimates. In this case it is only variability in partial recruitment (PR), averaged from the *F*s for the last three years in each bootstrap sample, that is influencing the calculations. It would appear that both the $F_{0.1}$ and $F_{\max}$ point estimates used in the XSA are relatively high compared with the bootstrap distributions.

The CVs (measures of uncertainty) for survivors at age in 2006 from the bootstrap results for each of the different resampling methods are compared with the internal standard errors used by Healey and Mahé (2006) in Fig. 8. For almost all ages, resampling all residuals, without standardizing, leads to the highest CV. The patterns of CVs by age are very similar when residuals are resampled within ages and indices, or standardized residuals are resampled across ages and indices. CVs are lowest for the highly selected ages (5-10) and highest for the early ages and older fish (including the plus group). The internal standard errors from the XSA greatly underestimate CV for the older ages (>10) relative to the bootstrap estimates, most likely due to shrinkage used in the formulation of the model.

Survivors-at-age estimates are the primary output from the XSA, and other statistics such as 5+ and 10+ biomass are derived statistics. As can be seen in Table 3, bootstrap estimates of survivors at age overestimate the best estimate for younger ages and underestimate it for older ages. Statistics calculated across ages, such as 5+ biomass, could be misleading because errors could cancel when the age groups are amalgamated. In other words, the statistics represented in Table 4 and Figures 3, 4 and 6 combine biases from the different survivor at age estimates and negative and positive biases for different ages may cancel. Bias should be examined for the primary outputs of the XSA - the survivors at age. If bias correction is necessary, it should be done on survivors-at-age estimates before further statistics based on these are calculated.

The bias to standard error ratios for the bootstrap estimates of survivors at age are shown in Fig. 9.. It can be seen that there is little difference in either variation or bias between the different methods of resampling residuals. Bias is low up to and including age 7, but ages 8-10 have large positive bias and ages 11-14 have large negative bias. Ages 5-8 are heavily selected by the fishery. With the exception of age 8, these estimates have low bias. It would be a concern if there was a large bias for ages that make up the bulk of the population, as this could result in retrospective problems. This distinct pattern of bias by age needs to be examined more closely to determine exactly what the causes are.

The results of the bootstrap distribution bias correcting methods are illustrated in Fig. 10 for survivors on 1 January 2006 at three ages: age 4, age 9 and age 12. In relation to the standard error of the bootstrap distributions, the bias of the estimate for age 4 survivors was very low (0.15), while there was a large positive bias for age 9 survivors (1.40) and a large negative bias for age 12 survivors (-1.14). The bootstrap bias corrected percentile technique does not allow for the bias corrected bootstrap distribution to lie outside the range of the observed bootstrap distribution. This in itself appears to be reasonable, but given a large bias (and in particular a large positive bias), this can lead to distributions being highly constrained by either the lower or upper bounds of the bootstrap distribution (i.e. to pile up at the bound). This can clearly be seen in the case of age 9 survivors, where the range of the bias corrected distribution is extremely small as a result of most bias corrected bootstrap replicates being corrected to the minimum value of the distribution. Trying to correct such large biases using the bootstrap bias correction percentile technique would not seem to be appropriate and efforts should rather be focused on improving the efficiency of the estimator so that less bias-correction is required.

Both the EU 3M and the Canadian spring 3LMNO surveys appear to have cohort effects in their residuals in the 2006 XSA, while the Canadian Fall 2J3K survey has mainly year effects in the residuals (Fig. 11). Cohort effects in residuals indicate some misspecification in the XSA model. The bootstrap resampling of residuals mixes up the residuals from cohorts that have been consistently under-estimated or over-estimated with the residuals from other cohorts. It is therefore to be expected that a large bias will be apparent when comparing the XSA estimates of survivors from a cohort that has predominantly negative or positive residuals. The 1995 cohort (age 12 survivors on 1 January 2006) had predominantly positive residuals in both the EU and Canadian Spring surveys, leading to a large negative bias. Conversely, the 1998 cohort (age 9 survivors on 1 January 2006) had predominantly negative residuals, leading to a large positive bias. Low biases tend to result from cohorts that no strong patterns in the residuals (age 4 survivors on 1 January 2006). Large biases resulting from cohort patterns in the residuals point to a misspecification within the XSA model design as applied to this stock.

It was found that estimates of bias and variance were relatively robust to the way residuals were resampled in the bootstrap. We consider the approach of standardizing the residuals and them resampling from the complete set of residuals to be the most logical. It is assumed that by standardizing residuals they should conform to the i.i.d. assumptions and bootstrap methods should provide good estimates of error associated with XSA estimates. We think that the nonparametric bootstrap provides a more reasonable estimation of the CV of XSA estimates than the internal standard error from the XSA, especially at older ages where it appears the internal SEs from the XSA are affected by the shrinkage applied in the model. With regards to estimates for older ages, the nonparametric bootstrap is relatively robust shrinkage. There are notable concerns with the bias in the XSA estimates of survivors at age for the older (>7) ages. These biases are communicated into biased estimates of 5+, 10+ biomass, and consequently F. $F_{0.1}$ and $F_{\max}$ reference points also appear to be biased.

Bias-correction of the bootstrap distribution using the bootstrap bias correction percentile technique poses some technical difficulties which were not able to overcome and which may be further indicative of problems in the XSA formulation for this stock. Bias could possibly be linked to cohort effects in the residuals of indices. Also, while the XSA does not converge without some shrinkage, if standard errors permitted around the estimate of F under shrinkage (Fse in the control group) is allowed to be larger, (i.e. less shrinkage) the bias is somewhat reduced. We would like to follow up on the possible causes of the bias through simulation testing on a known, simulated stock.

Acknowledgements

We thank Stratis Gavaris for correspondence on the appropriate methods for bootstrap bias corrected percentile methods. Brian Healey and Jean-Claude Mahé generously provided advice with respect to XSA as well as the output from the 2006 assessment. Carmen Fernandez provided advice on how to standardize residuals and then how to generate new indices using these standardized residuals. This research is funded by the International Governance program on straddling stocks by the Department of Fisheries and Oceans, Canada.

References

DARBY, C. D., W. R. BOWERING, and J.-C. MAHÉ. 2003. An Assessment of Stock Status of the Greenland Halibut Resource in NAFO Subarea 2 and Divisions 3KLMNO Based on Extended Survivors Analysis with Short and Medium-term projections of Future Stock Development. *NAFO SCR Doc*., No. 03/64, Ser. No. N4883.

DARBY, C. D., and S. FLATMAN. 1994. Virtual Population Analysis: Version 3.1 (Windows/Dos) user guide. *Info. Tech. Ser., MAFF Direct. Fish. Res., Lowestoft*, (1): 85 p.

DARBY, C. D., B.P. HEALEY, J.-C. MAHÉ and W. R. BOWERING. 2004. Greenland Halibut (*Reinhardtius hippoglossoides*) in Subarea 2 and Divisions 3KLMNO: an Assessment of Stock Status Based on Extended Survivors Analysis, ADAPT, and ASPIC Analyses, with Stochastic projections of Potential Stock Dynamics. *NAFO SCR Doc*., No. 04/55, Ser. No. N5008.

EFRON, B. 1981. Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika*, **68**(3): 589-599.

EFRON, B. and R.J. TIBSHIRANI. 1993. An Introduction To The Bootstrap. Chapman & Hall. New York. 436pp.

GAVARIS, S. 1999. Dealing With Bias in Estimating Uncertainty and Risk. *Proceedings, 5th NMFS NSAW.* NOAA Tech. Memo. NMFS-F/SPO-40.

GONZÁLEZ TRONCOSO, D., J. M. CASAS, and F. SABORIDO-REY. 2006. Results from bottom trawl survey on Flemish Cap of July-August 2005. *NAFO SCR Doc*., No. 16, Ser. No. N5231.

HEALEY, B.P. and J.-C. MAHÉ. 2006. An Assessment of Greenland Halibut (*Reinhardtius hippoglossoides*) in NAFO Subarea 2 and Divisions 3KLMNO. *NAFO SCR Doc.,* No. 06/51, Ser. No. N5281.

HEALEY, B. P., K. S. DWYER, and D. MADDOCK PARSONS. 2006. Greenland halibut (*Reinhardtius hippoglossoides*) in NAFO Subarea 2 and Divisions 3LKMNO: stock trends based on annual Canadian research vessel survey results during 1978-2005. *NAFO SCR Doc.*, No. 06/34, Ser. No. N5257.

PATTERSON, K. R., R. M. COOK, C. D. DARBY, S. GAVARIS, B. MESNIL, A. E. PUNT, V. R. RESTREPO, D. W. SKAGEN, G. STEFANSSON, and M. SMITH. 2000. Validating three methods for making probability statements in fisheries forecasts. *ICES C.M. Doc*., No. 2000 V:06.

SHEPHERD, J. G. 1999. Extended survivors analysis: An improved method for the analysis of catch-at-age data and abundance indices *ICES J. Mar. Sci*., 56(5): 584-591.

TREBLE, M. A. and DWYER, K. S. 2006. Report of the Greenland Halibut (*Reinhardtius hippoglossoides*) Age Determination Workshop. *NAFO SCR Doc*., No. 06/49, Ser. No. N5276.

Table 1. XSA control group used to mimic the 2006 assessment of Greenland halibut.

Argument	Description	Value for G. Hal. (2006 assess.)
tol	Convergence tolerance	1.00E-09
maxit	Maximum number of iterations	60
min.nse	Minimum value of SE permitted in estimate of N	0.5
fse	SE of F when shrinking to mean F	0.5
rage	Oldest age for which two parameter model is used fro catchability at age	0
qage	Age after which catchability is no longer estimated	11
shk.n	Shrinkage to mean N	TRUE
shk.f	Shrinkage to mean F	TRUE
shk.yrs	No. years for shrinkage of F for terminal year	5
shk.ages	No. ages for shrinkage of F for terminal age	3
window	Time window to consider in model fitting	100
tsrange	No. years to use in time series weighting	20
tspower	Power to be used in the time series weighting	0
vpa	VPA (T) or cohort analysis (F)	FALSE

Table 2. Survivors-at-age on 1 January 2006 as estimated by the XSA and 1000 bootstrap replicates of the XSA (standardized residuals resampled).

	XSA		Bootstrap Replicates			Bootstrap Replicate Percentiles					Bootstrap Range		
Age	Estimate	Int. SE	Mean	SE	CV	5%	25%	Median	75%	95%	Min	Max	Range
2	28548	0.32	30518	9047	0.30	17737	24222	29458	35888	46327	11799	70998	59199
3	44456	0.24	45708	10092	0.22	31387	38353	44933	51901	63559	24350	85124	60775
4	52583	0.20	53665	9325	0.17	39269	47212	52956	59559	69674	31114	87472	56358
5	43964	0.16	43254	5644	0.13	34344	39360	42966	47048	52648	25954	65196	39242
6	32586	0.14	31331	3971	0.13	25122	28646	30980	33799	38433	20476	44185	23709
7	19952	0.13	19408	2880	0.15	14851	17343	19329	21342	24196	11439	28850	17411
8	3570	0.16	4843	989	0.20	3489	4120	4703	5377	6680	2814	9803	6989
9	1656	0.20	2448	483	0.20	1804	2118	2377	2689	3324	1393	4400	3007
10	1244	0.20	1648	363	0.22	1163	1387	1595	1858	2329	857	3229	2372
11	959	0.19	853	198	0.23	570	698	833	981	1218	411	1768	1356
12	604	0.19	445	131	0.29	281	348	426	521	695	211	1248	1037
13	223	0.20	188	61	0.32	110	143	178	220	301	77	474	397
14	260	0.21	220	78	0.36	126	165	204	257	368	82	626	544

Table 3. XSA estimates at two levels of shrinkage, Fse=0.5 as per the 2006 assessment and Fse=1 to implement less shrinkage. A no-shrinkage run was also attempted but the XSA does not converge and the estimated cannot be compared with runs in which shrinkage is applied. Standardized residuals were resampled for 1000 bootstrap estimates.

Fse = 0.5 (original XSA)

Survivors-at-age in 2006	XSA Estimate	Bootstrap Mean	Bootstrap Median	Bootstrap SE	Bootstrap CV	Bias	Bias/SE	Bias Corrected Estimate
2	28548	30518	29458	9047	0.30	1970	0.22	26578
3	44456	45708	44933	10092	0.22	1252	0.12	43204
4	52583	53665	52956	9325	0.17	1082	0.12	51501
5	43964	43254	42966	5644	0.13	-710	-0.13	44674
6	32586	31331	30980	3971	0.13	-1255	-0.32	33841
7	19952	19408	19329	2880	0.15	-544	-0.19	20496
8	3570	4843	4703	989	0.20	1274	1.29	2296
9	1656	2448	2377	483	0.20	791	1.64	865
10	1244	1648	1595	363	0.22	404	1.11	840
11	959	853	833	198	0.23	-106	-0.53	1065
12	604	445	426	131	0.29	-159	-1.22	763
13	223	188	178	61	0.32	-36	-0.58	259
14	260	220	204	78	0.36	-40	-0.51	300

Fse = 1.0 (less shrinkage)

Survivors-at-age in 2006	XSA Estimate	Bootstrap Mean	Bootstrap Median	Bootstrap SE	Bootstrap CV	Bias	Bias/SE	Bias Corrected Estimate
2	27635	30948	29746	10440	0.34	2399	0.23	25236
3	43882	46427	45360	10323	0.22	1971	0.19	41911
4	51990	53602	53014	9916	0.18	1019	0.10	50971
5	44589	44516	43561	7264	0.16	552	0.08	44037
6	34030	32778	32570	4720	0.14	192	0.04	33838
7	20505	20272	20141	3564	0.18	320	0.09	20185
8	1225	4470	4196	1648	0.37	900	0.55	325
9	36	2341	2130	987	0.42	685	0.69	-649
10	457	1611	1467	694	0.43	367	0.53	90
11	1501	975	919	371	0.38	16	0.04	1485
12	1275	553	522	233	0.42	-52	-0.22	1327
13	455	205	193	97	0.47	-19	-0.19	474
14	497	239	216	119	0.50	-21	-0.17	518

Table 4. Comparison of point estimates from the XSA under two levels of shrinkage with the nonparametric bootstrap estimates.

Fse = 0.5 (Original XSA) n = 200

2006	XSA Estimate	Bootstrap Median	Bootstrap Mean	Bootstrap SE	Bootstrap CV	Bias	Bias/SE	Bias Corrected Estimate
5+Biomass	68415	69245	69516	5240	0.08	1101	0.21	67314
10+Biomass	9773	9225	9436	1602	0.17	-337	-0.21	10110
F	0.63	0.55	0.55	0.04	0.07	-0.08	-2.06	0.71

Fse = 1 (Less shrinkage) n = 200

2006	XSA Estimate	Bootstrap Median	Bootstrap Mean	Bootstrap SE	Bootstrap CV	Bias	Bias/SE	Bias Corrected Estimate
5+Biomass	70786	71549	71691	7421	0.10	905	0.12	69880
10+Biomass	12527	10081	10237	2561	0.25	-2290	-0.89	14817
F	0.74	0.58	0.58	0.08	0.13	-0.16	-2.05	0.89

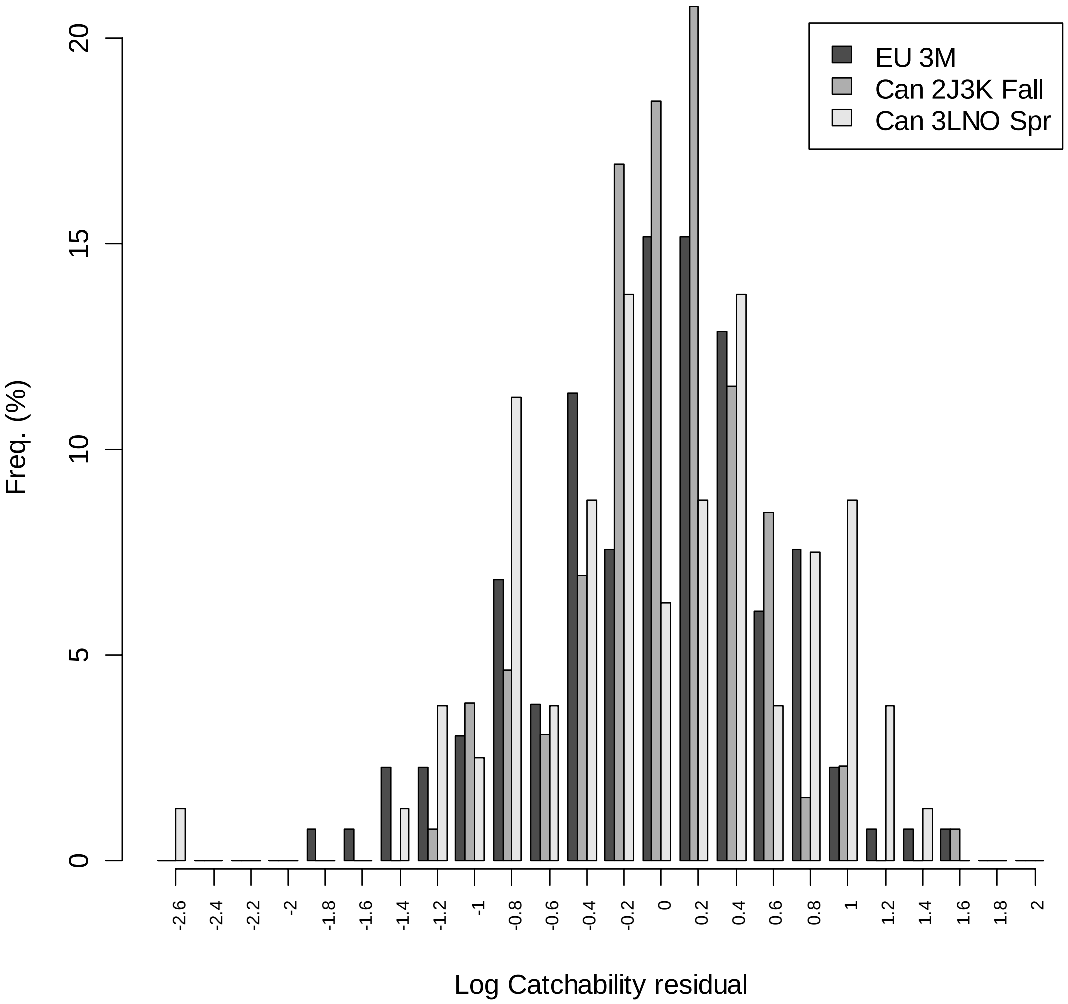


Fig. 1 Frequency distribution for the log catchability residuals for each tuning index from XSA applied to the 2006 data for Greenland halibut.

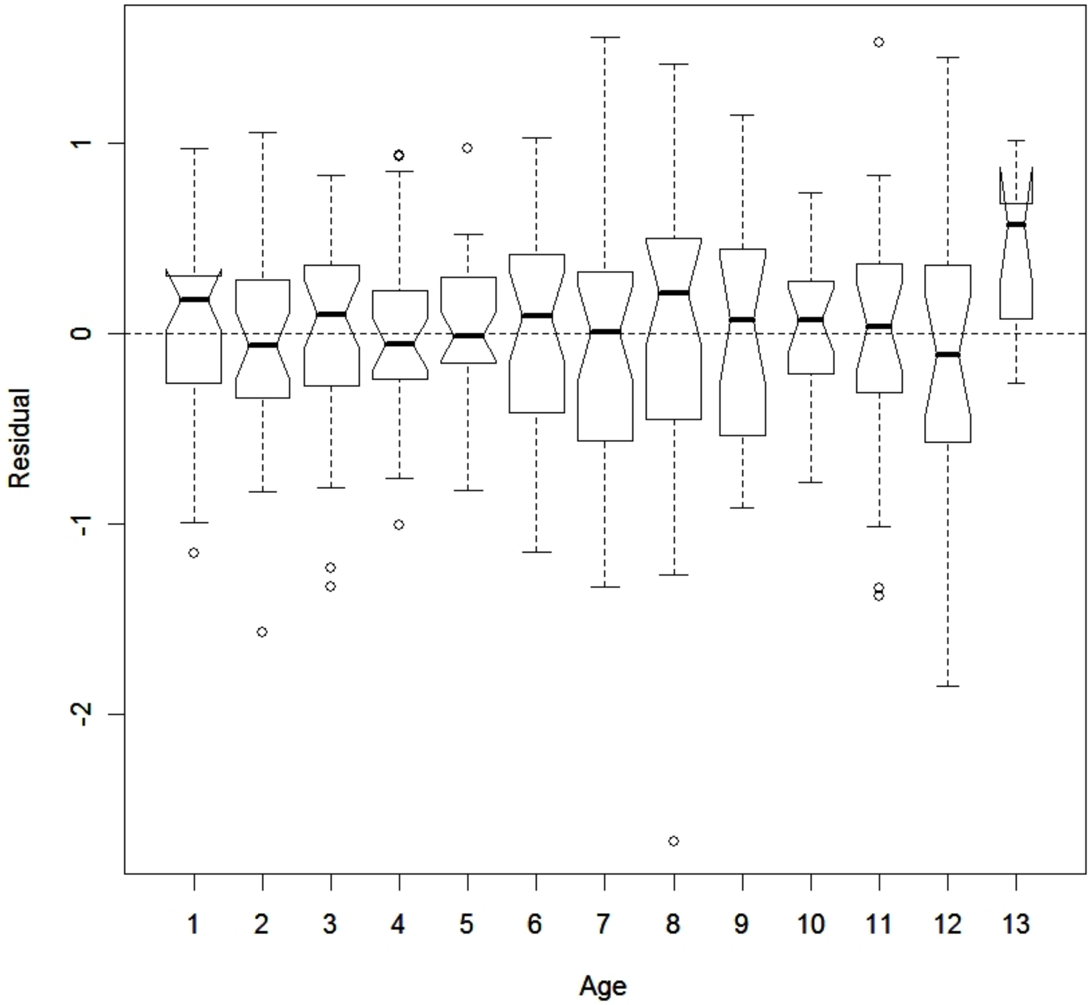


Fig. 2. Box-and-whisker plot of the distributions of the log catchability residuals for each age from XSA applied to the 2006 data for Greenland halibut. The box indicates the inter-quartile range and the whiskers extend to the furthest point within 1.5*the box range. Distributions do not differ significantly if the notches in the side of the boxes overlap. Outliers are marked. The width of the boxes is relative to the number of observations.

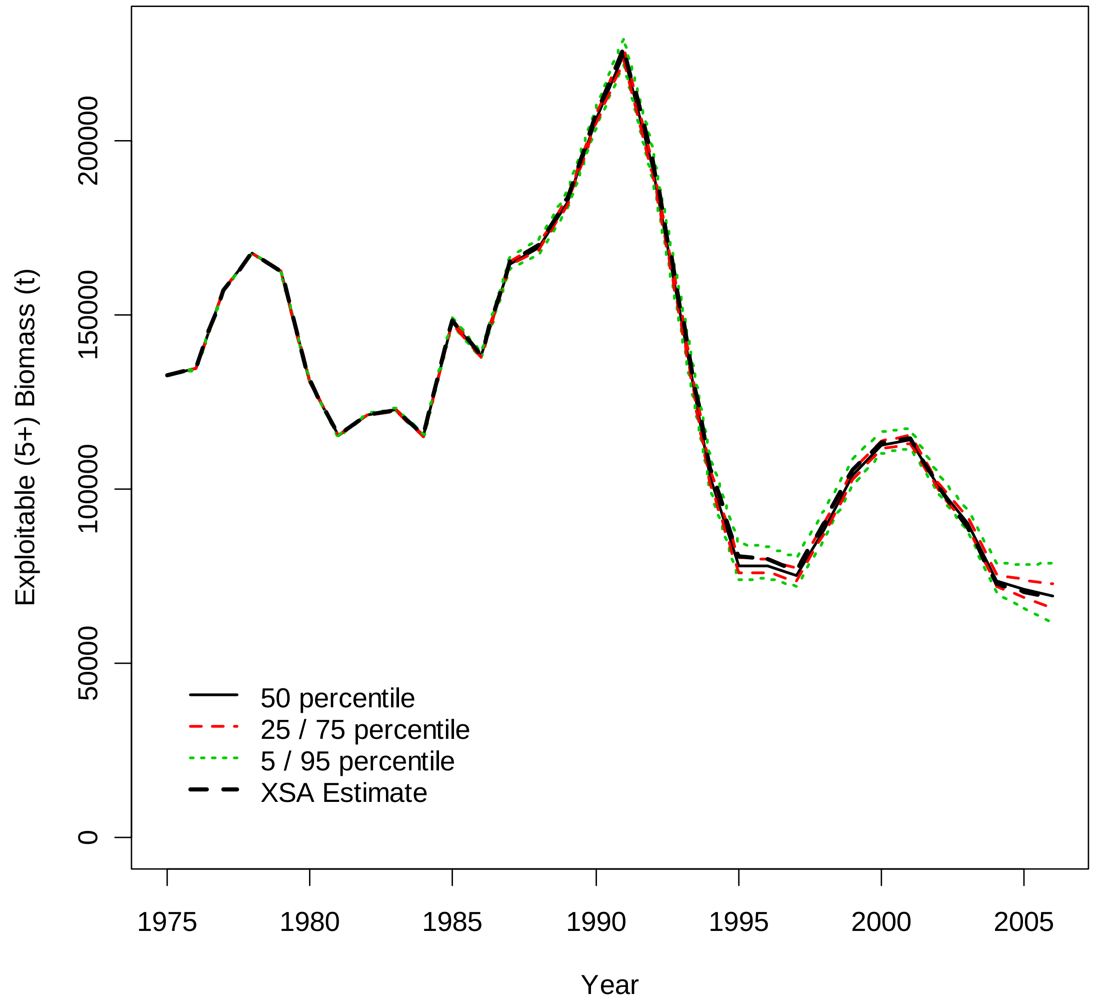


Fig. 3. Comparison of the bootstrap percentiles with the 2006 XSA point estimates for exploitable (5+) biomass.

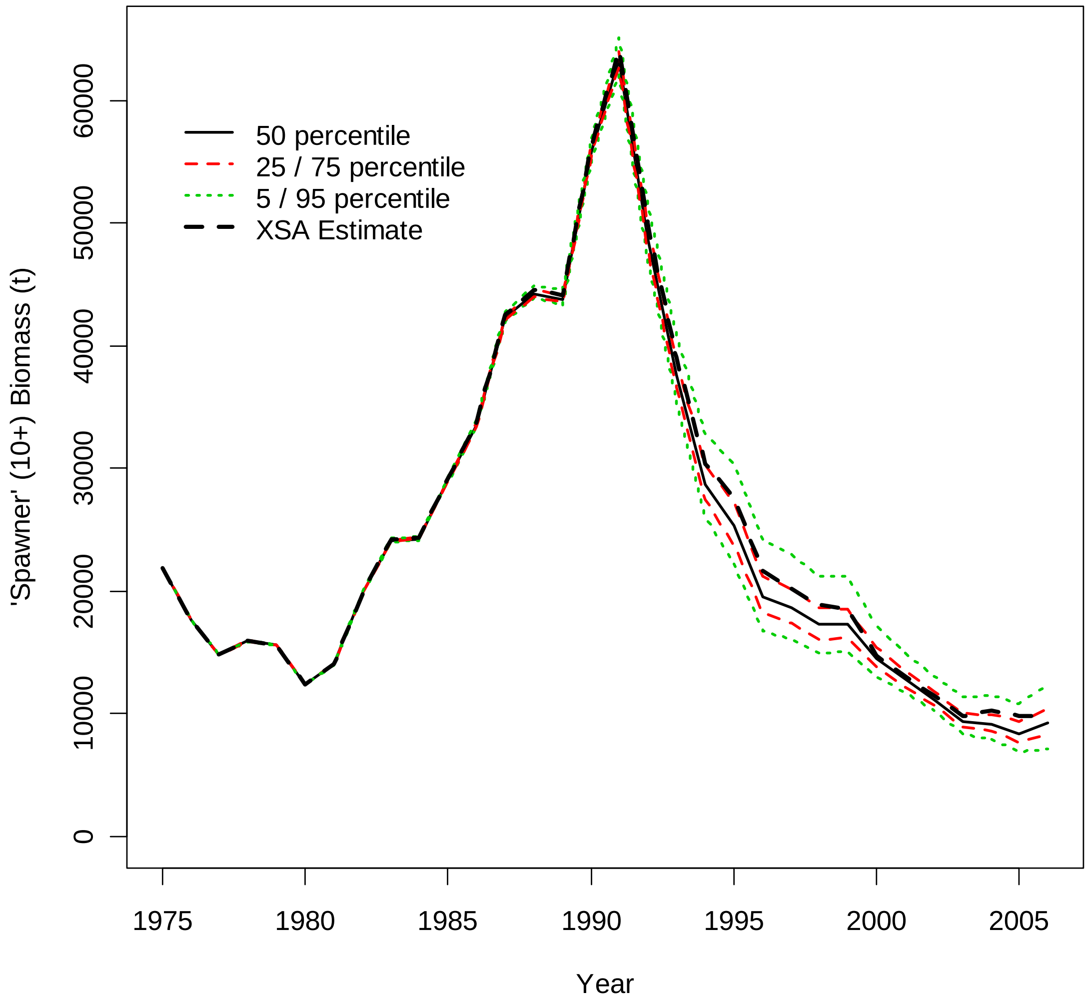


Fig. 4. Comparison of the bootstrap percentiles with the 2006 XSA point estimates for 10+ biomass.

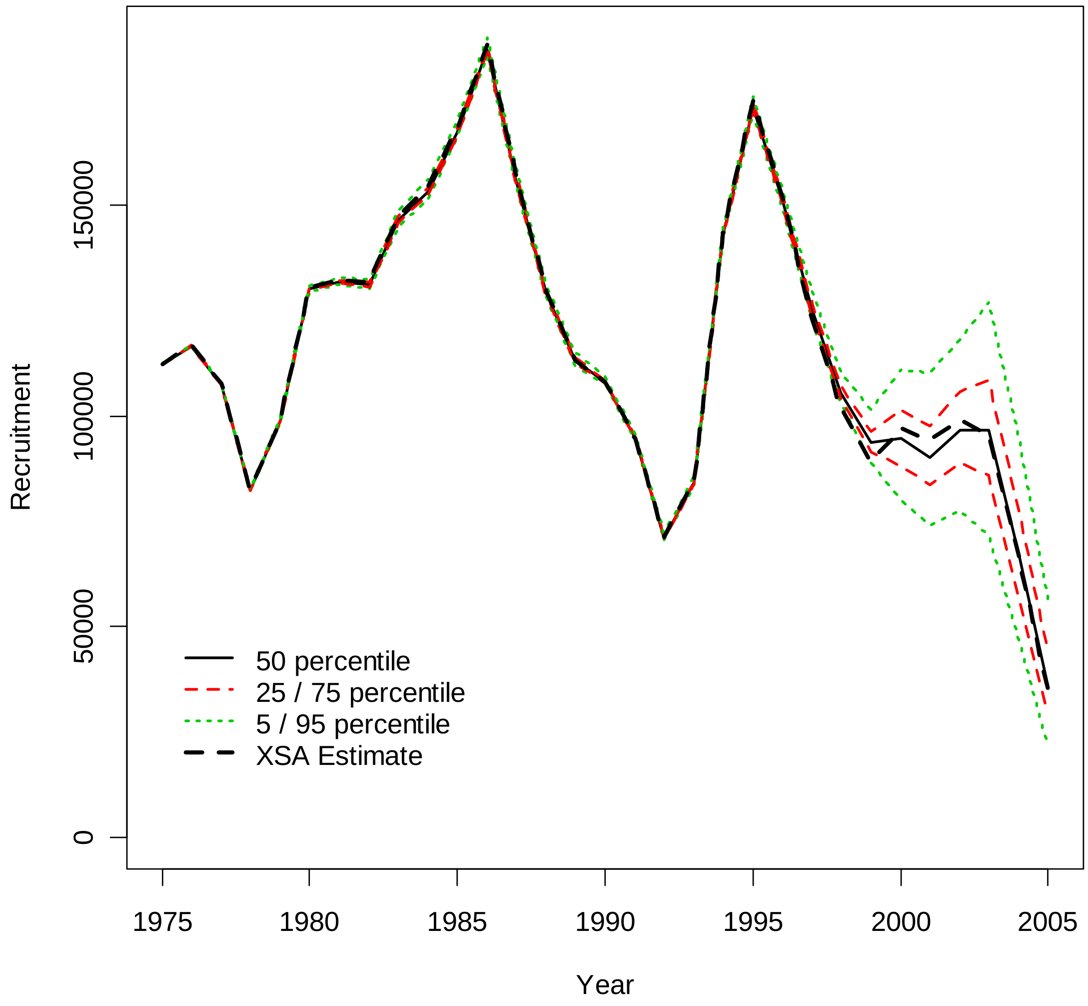


Fig. 5. Comparison of the bootstrap percentiles with the 2006 XSA point estimates for recruitment (age 1).

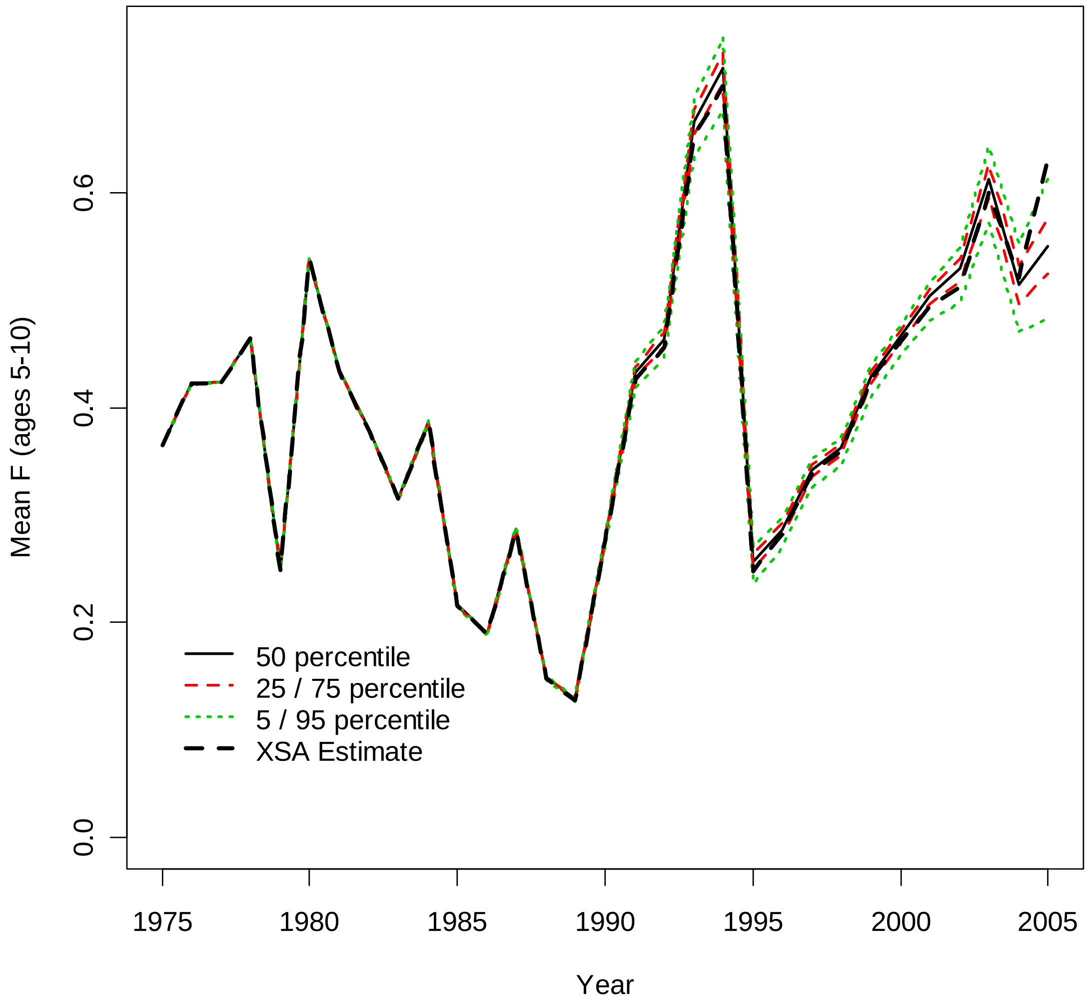


Fig. 6. Comparison of the bootstrap percentiles with the 2006 XSA point estimates for mean *F* (ages 5 to 10).

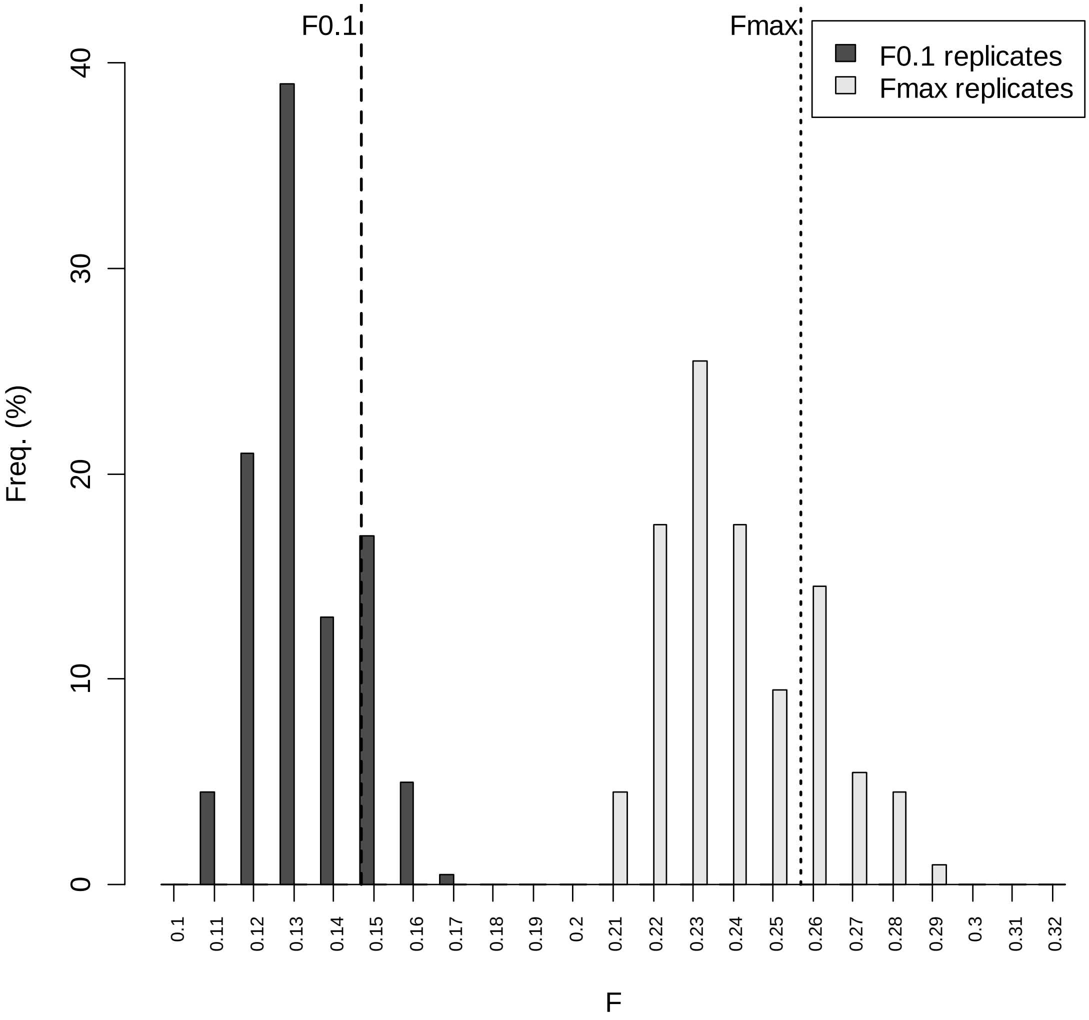


Fig. 7. Bootstrap frequency distribution of estimates of $F_{0.1}$ and $F_{\max}$ from 200 replicates. The vertical broken lines show the location of the 2006 XSA point estimates.

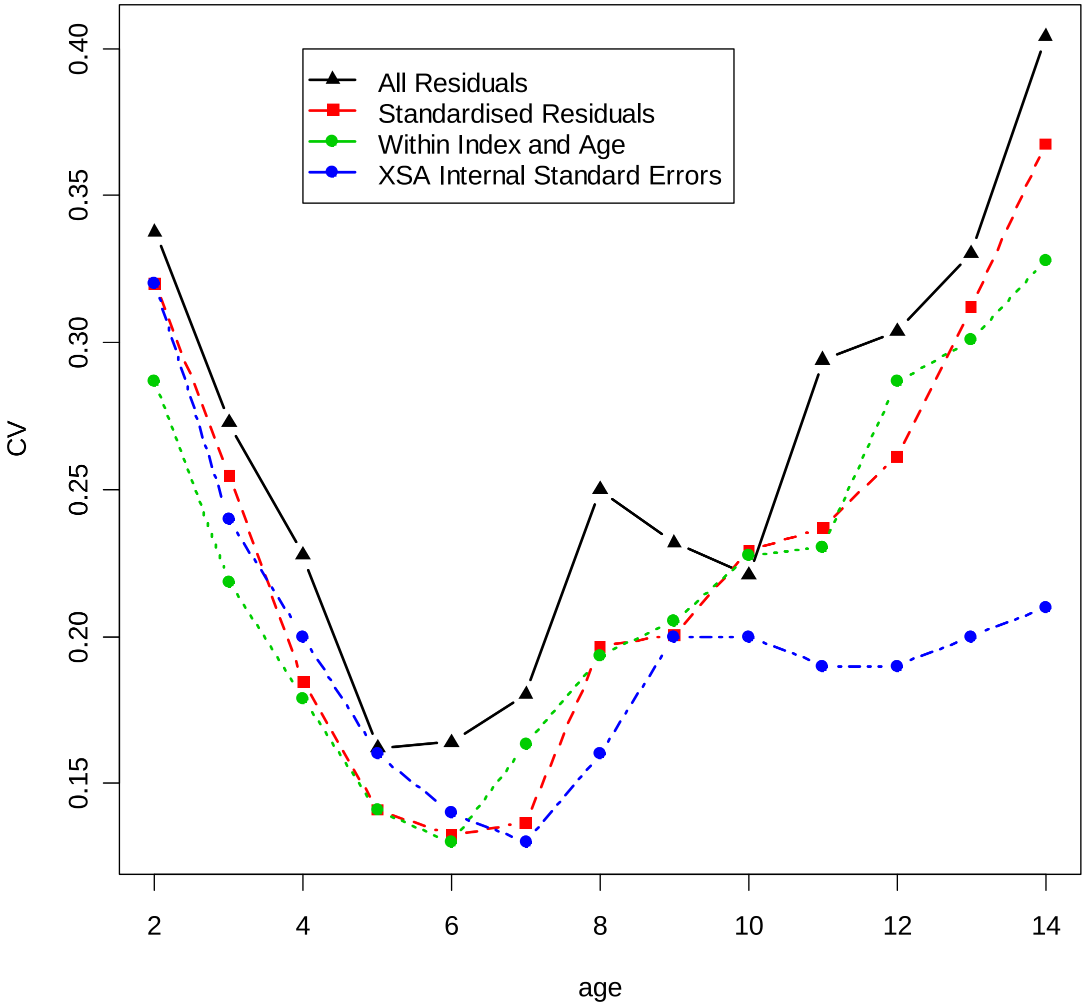


Fig. 8. Comparison of bootstrap coefficients of variation and internal standard errors (used in the stochastic projections of Healey and Mahé 2006)) for the survivors-at-age in 2006 (the terminal year).

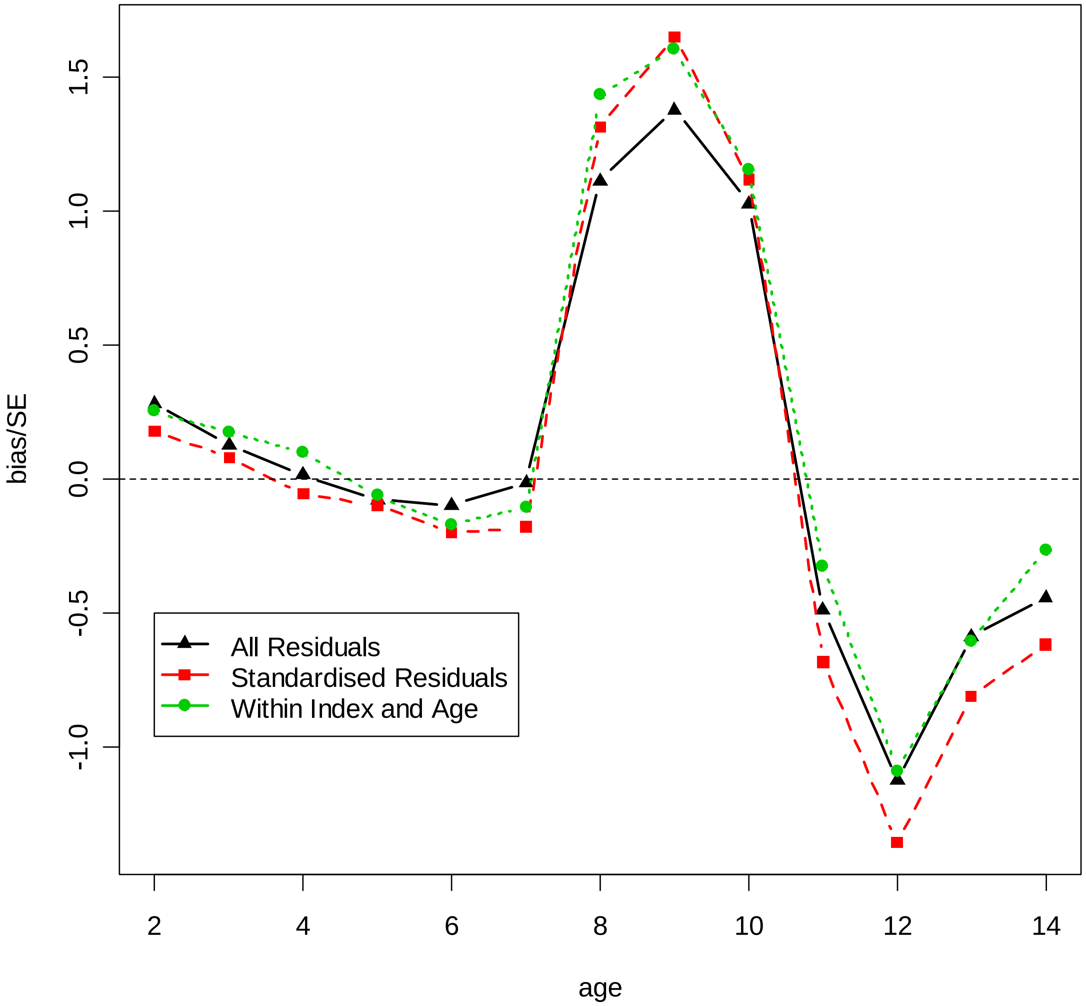
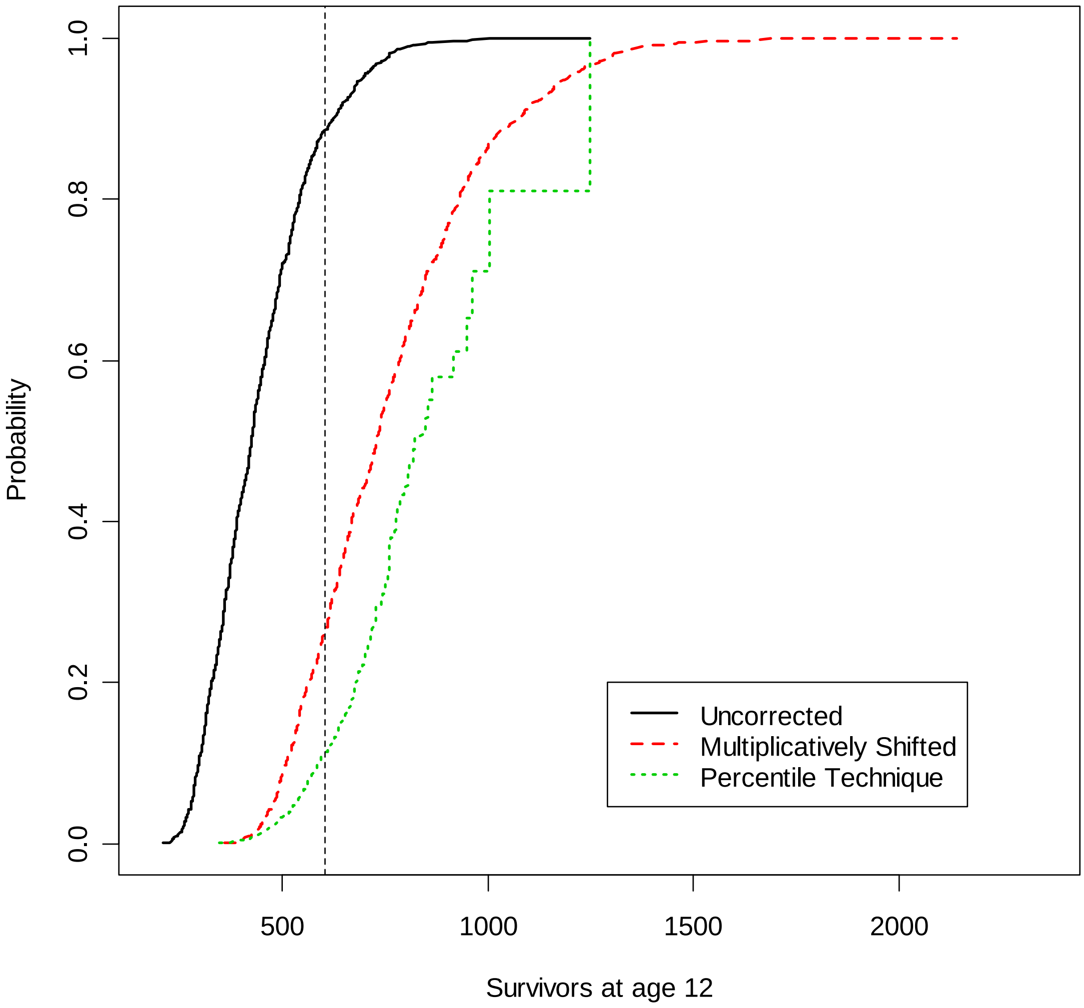
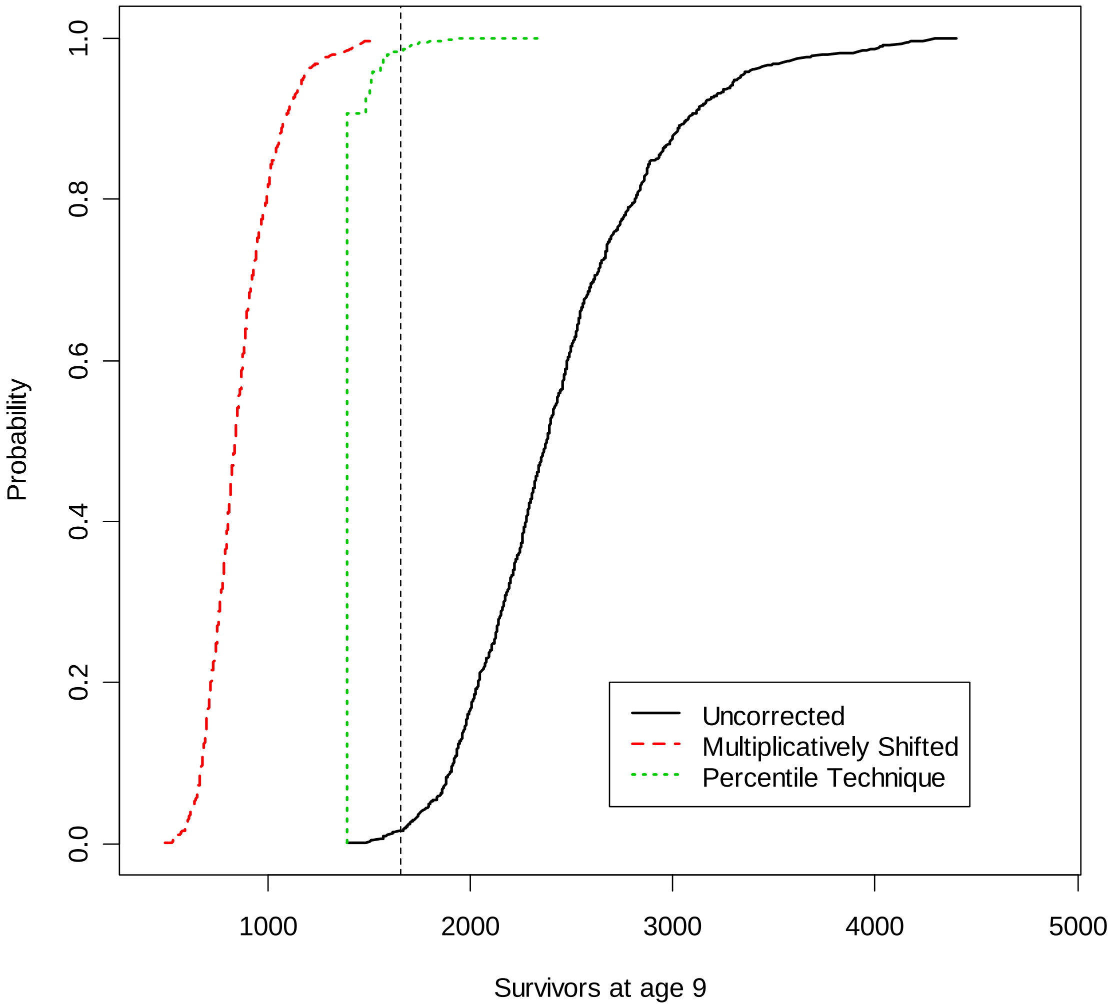


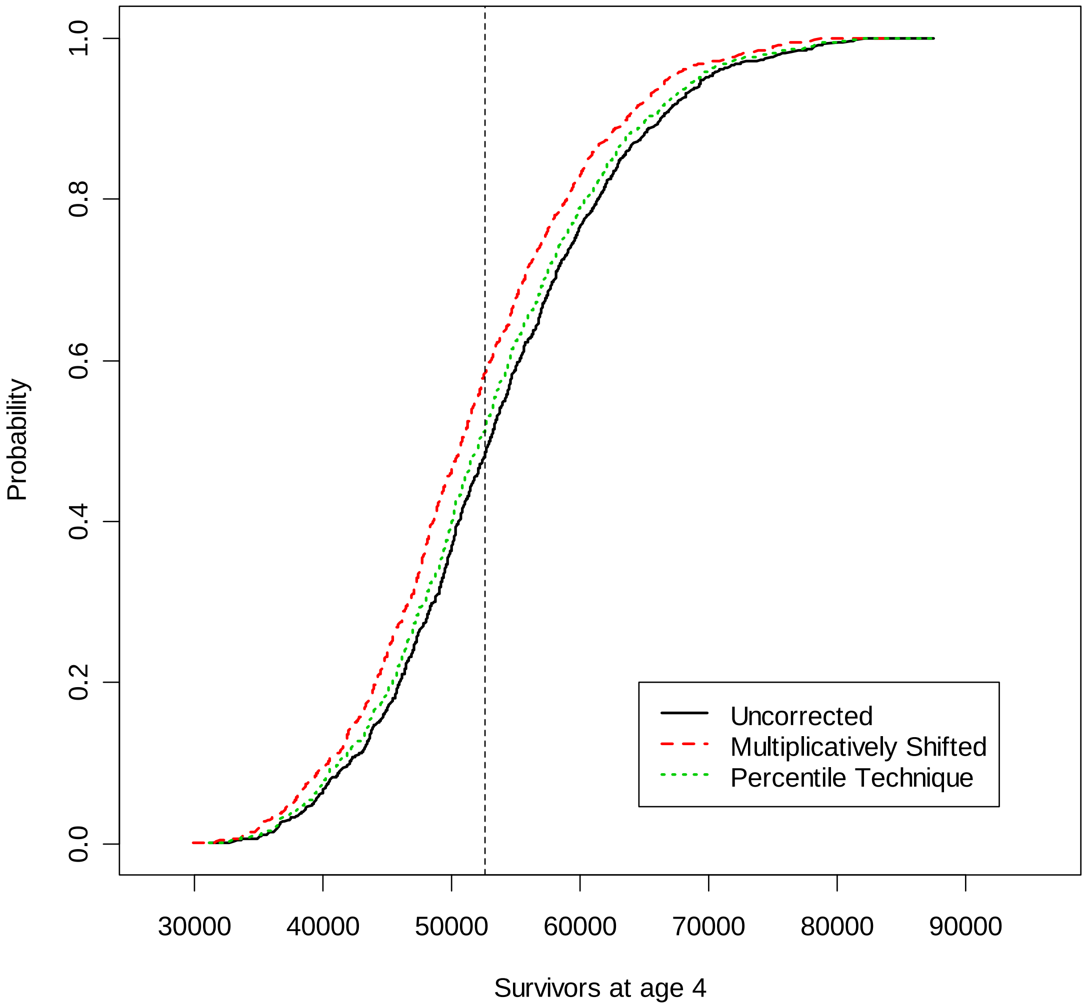
Fig. 9. Bias to standard error ratios for the bootstrap estimates of survivors at age using three different methods of resampling residuals.



Age 12 survivors (XSA estimate = 604)

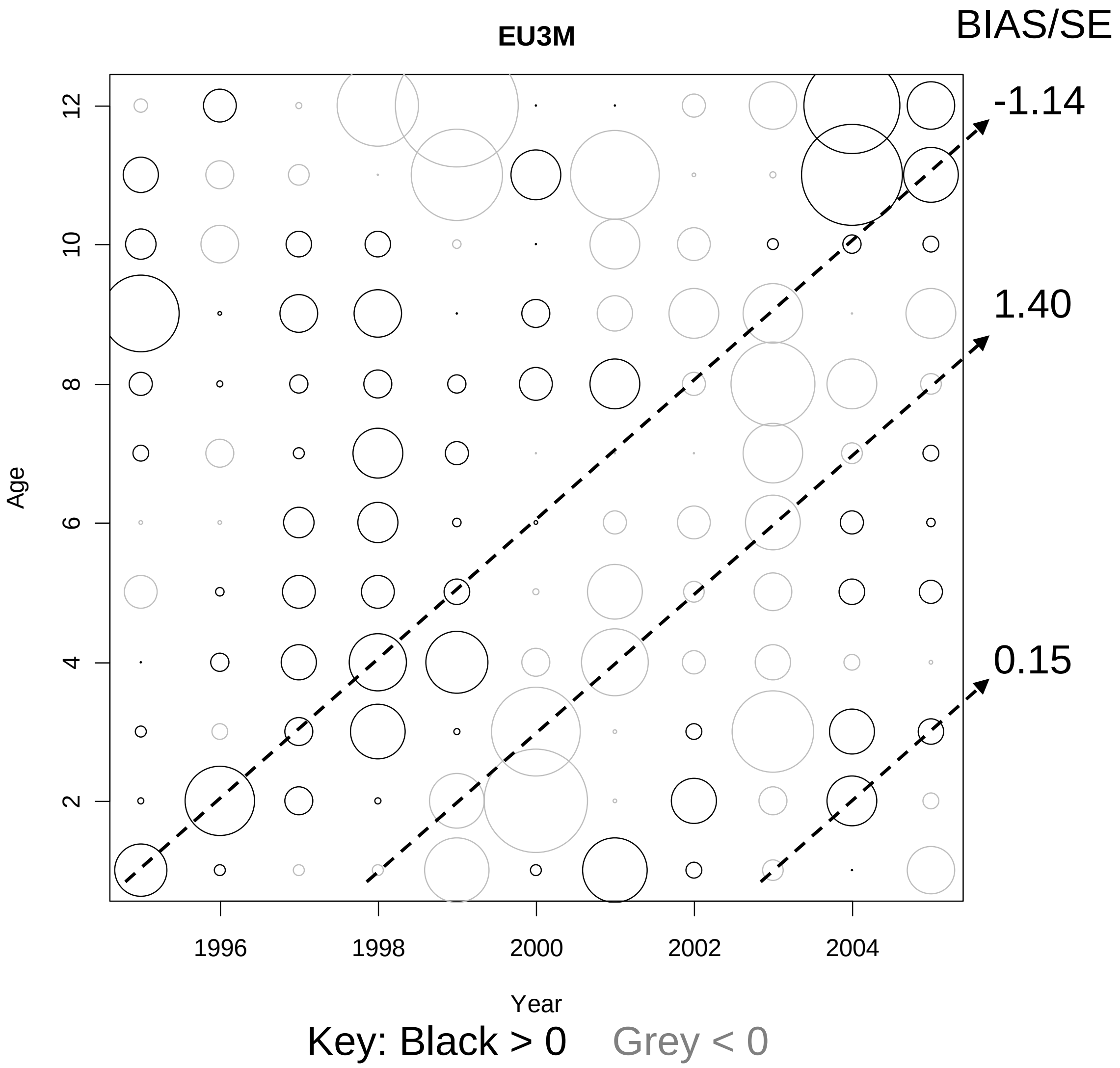


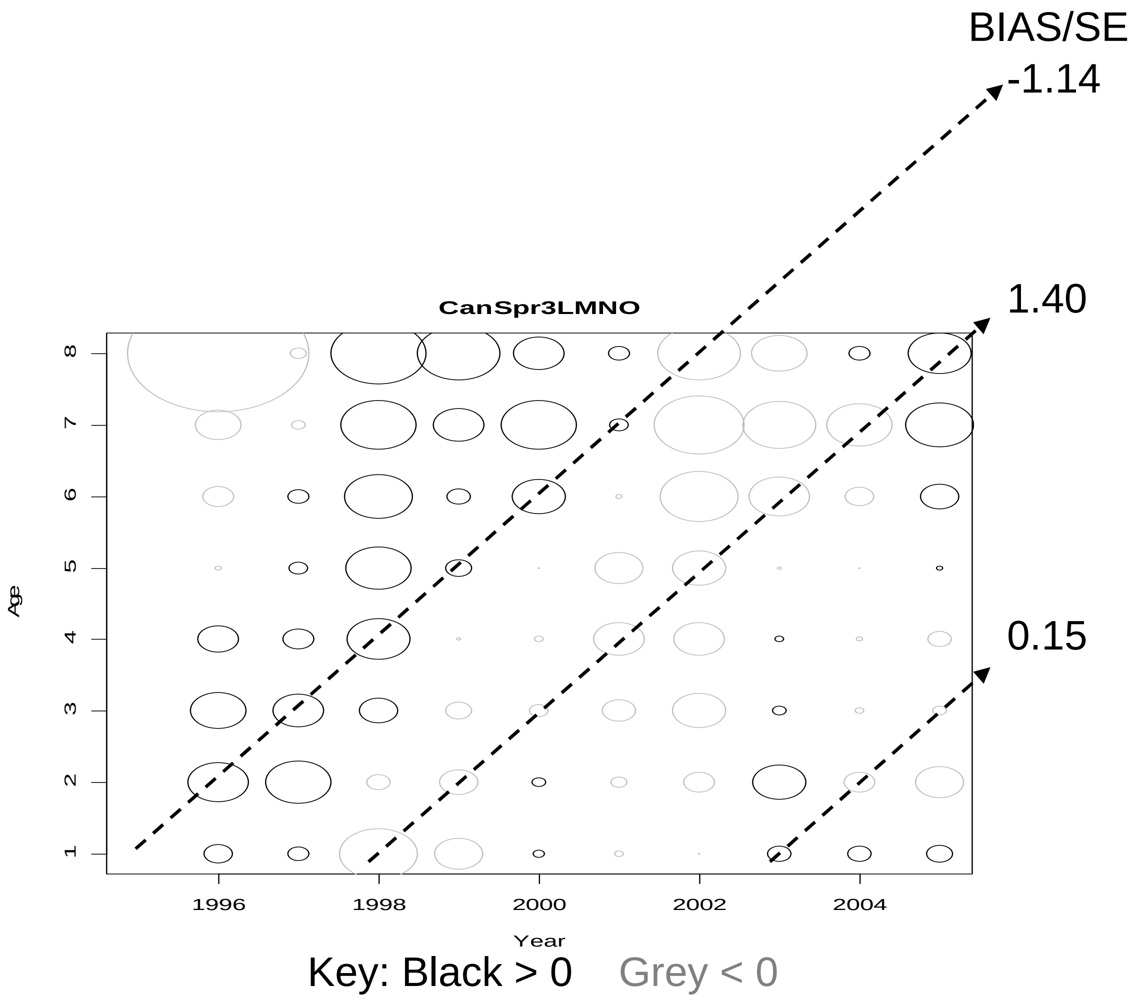
Age 9 survivors (XSA estimate = 1656)



Age 4 survivors (XSA estimate = 52583)

Fig. 10. Biased (uncorrected) and bias-corrected bootstrap probability distributions of survivors at ages 4, 9 and 12 in the terminal year (1 Januray 2006). Standardized residuals were resampled. Vertical lines represent the estimates of survivors-at-age from the original XSA. Multiplicatively Shifted distributions involve simply shifting the entire distribution, multiplicatively, such that the mean of the distribution lines up with the bias-corrected value of the best estimate.


EU3M
BIAS/SE
-1.14
1.40
0.15
Age
12
10
8
6
4
2
1996
1998
2000
2002
2004
Year
Key: Black > 0 Grey < 0


BIAS/SE
-1.14
1.40
0.15
CanSpr3LMNO
Age
1
2
3
4
5
6
7
8
1996
1998
2000
2002
2004
Year
Key: Black > 0 Grey < 0

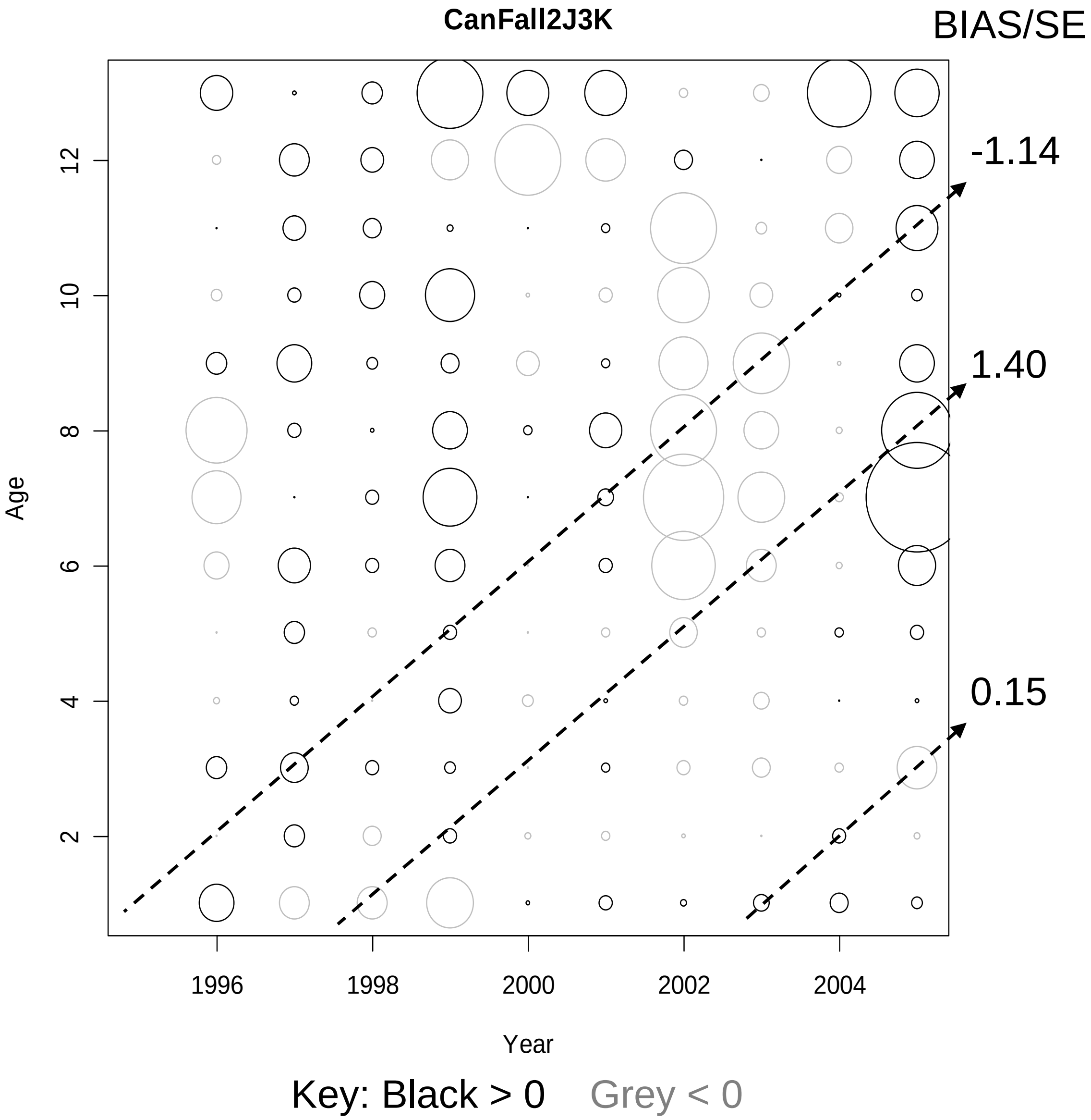


Fig. 11. Bubble plots of the XSA residuals for the EU 3M, Can Fall 2J3K and Can Spr 3LMNO indices used in the Greenland Halibut XSA, indicating how cohort effects lead to bias in the bootstrap distributions. Bias/SE values show the ratio of bias to standard error for estimates of survivors at age at the start of 2006 for ages 4, 9 and 12. Dashed arrows lie over the respective cohorts.

APPENDIX 1: R code to run a bootstrap replicate of the XSA for Greenland Halibut

Input:

GH.stock = An FLStock object with the relevenat Greenland Halibut information (catch etc.)
GH.index = An FLIndices object with the three research surveys of the Greenland Halibut stock
("EU 3M","Can 2J3K Fall","Can 3LNO Spr")
GH.xsa = An FLXSA object with the results of the XSA carried out using GH.stock and GH.index above (the control group for the XSA is given below and is the same one used for the bootstrap replicate)

 # Control file for XSAs

```
maxIt <- 60   # Set the max number of iterations
GH.control<-FLXSA.control( tol = 1e-09, maxit = maxIt, min.nse = 0.5, fse = 0.5,
         rage = 0, qage = 11, shk.n = TRUE, shk.f = TRUE, shk.yrs = 5,
         shk.ages = 3, window = 100, tsrange = 20, tspower = 0, vpa = FALSE)
```

 # Get values for ranges

 # Stock object

```
maxYear <- GH.stock@range[[5]]; minYear <- GH.stock@range[[4]]; numYears <- maxYear-minYear+1
minAge <- GH.stock@range[[1]]; maxAge <- GH.stock@range[[2]]; numAges <- maxAge-minAge+1
```

 # Indices

```
EUnumYears <- GH.index2[[1]]@range[[5]]-GH.index2[[1]]@range[[4]]+1
EUmaxAge <- GH.index2[[1]]@range[2]
EUnumCases <- EUmaxAge*EUnumYears

CanFallnumYears <- GH.index2[[2]]@range[[5]]-GH.index2[[2]]@range[[4]]+1
CanFallmaxAge <- GH.index2[[2]]@range[2]-1 # Because can't get residuals for age 14 - in plusgroup
CanFallnumCases <- CanFallmaxAge*CanFallnumYears

CanSprnumYears <- GH.index2[[3]]@range[[5]]-GH.index2[[3]]@range[[4]]+1
CanSprmaxAge <- GH.index2[[3]]@range[2]
CanSprnumCases <- CanSprmaxAge*CanSprnumYears

AllresidsnumCases <- EUnumCases+CanFallnumCases+CanSprnumCases
```

 # Create vectors with all the log catchability residuals for each index

```
EUresid <- array(NA,dim=c(EUnumCases))
```

```
CanFallresid <- array(NA,dim=c(CanFallnumCases))
CanSprresid <- array(NA,dim=c(CanSprnumCases))
Allresids <- array(NA,dim=c(AllresidsnumCases))

EUresid <- as.numeric(GH.xsa@index.res[[1]])
CanFallresid <- as.numeric(GH.xsa@index.res[[2]])
CanSprresid <- as.numeric(GH.xsa@index.res[[3]])
Allresids <- c(EUresid, CanFallresid, CanSprresid)

# To standardize residuals
EUresidSE <- apply(EUresid,1,sd)
CanFallresidSE <- apply(CanFallresid,1,sd)
CanSprresidSE <- apply(CanSprresid,1,sd)

for (eur in 1:12) { stdEUresid[eur,] <- EUresid[eur,]/EUresidSE[eur] }
for (eur in 1:13) { stdCanFallresid[eur,] <- CanFallresid[eur,]/CanFallresidSE[eur] }
for (eur in 1:8) { stdCanSprresid[eur,] <- CanSprresid[eur,]/CanSprresidSE[eur] }

stdAllresids <- c(stdEUresid,stdCanFallresid,stdCanSprresid)

# exponents of log(q') to get q' i.e. vector of catchability residuals
expEUresid <- exp(EUresid)
expCanFallresid <- exp(CanFallresid)
expCanSprresid <- exp(CanSprresid)
expAllresids <- exp(Allresids)

# Calculate index model PREDICTED values (= XSA stock numbers at age and year multiplied by index catchability at age (avg.
over years))
EUpredIndex <- array(NA,, dim=c(EUmaxAge,EUnumYears))
for (r in 1:EUmaxAge) { for (cc in 1:EUnumYears) {
      EUpredIndex[r,cc] <- GH.xsa@stock.n[r,numYears-EUnumYears+cc,,,]*GH.xsa@q.hat[[1]][r,,,,]
      }}

CanFallpredIndex <- array(NA,, dim=c(CanFallmaxAge,CanFallnumYears))
for (r in 1:CanFallmaxAge) { for (cc in 1:CanFallnumYears) {
      CanFallpredIndex[r,cc] <- GH.xsa@stock.n[r,numYears-CanFallnumYears+cc,,,]*GH.xsa@q.hat[[2]][r,,,,]
      }}

CanSprpredIndex <- array(NA,, dim=c(CanSprmaxAge,CanSprnumYears))
for (r in 1:CanSprmaxAge) { for (cc in 1:CanSprnumYears) {
      CanSprpredIndex[r,cc] <- GH.xsa@stock.n[r,numYears-CanSprnumYears+cc,,,]*GH.xsa@q.hat[[3]][r,,,,]
```

```
      }}

# Create base index file (values rescaled to the beginning of the year and start and end of fishing set at 0)
# This is done because index predicted values calculated above are for survivors at the START of each year, therefore index must reflect this.
GH.index00 <- GH.index

for (ind in 1:3){
    for (j in 6:7) { GH.index00[[ind]]@range[j] <- 0 }
    if (ind == 1 || ind ==3){
      GH.index00[[ind]]@catch.n <- GH.xsa@index[[ind]]
      GH.index00[[ind]]@index <- GH.xsa@index[[ind]]
      }
    if (ind == 2) {
      GH.index00[[ind]]@catch.n[1:13] <- GH.xsa@index[[ind]]
      GH.index00[[ind]]@index[1:13] <- GH.xsa@index[[ind]]
      }
    }

# Replace negative values in the index file with NAs
GH.index00 <- replaceNegative(GH.index00)   # A function from the FLR website

# Running a bootstrapped replicate of the XSA

# Create index file for run i from base index file  (and stock file from GH.stock)
GH.indexMC <- GH.index00
GH.stockMC <- GH.stock

# Model predicted index is multiplied by the a ramndomly sampled, exponentiated residual
# residSamp is a dummy variable to choose Between resampling methods

if (residSamp==1) {        # Resample accross all residuals
for (r in 1:EUmaxAge) { for (cc in 1:EUnumYears) { GH.indexMC[[1]]@index[r,cc,,,] <-
   EUpredIndex[r,cc]*expAllresids[round(runif(1,min=0.5,max=AllresidsnumCases+0.5),0)] }}
for (r in 1:CanFallmaxAge) { for (cc in 1:CanFallnumYears) { GH.indexMC[[2]]@index[r,cc,,,] <-
   CanFallpredIndex[r,cc]*expAllresids[round(runif(1,min=0.5,max=AllresidsnumCases+0.5),0)] }}
for (r in 1:CanSprmaxAge) { for (cc in 1:CanSprnumYears) { GH.indexMC[[3]]@index[r,cc,,,] <-
    CanSprpredIndex[r,cc]*expAllresids[round(runif(1,min=0.5,max=AllresidsnumCases+0.5),0)] }}
```

```
    }

if (residSamp==2) {   # Resample accross all standardised residuals
for (r in 1:EUmaxAge) { for (cc in 1:EUnumYears) { GH.indexMC[[1]]@index[r,cc,,,] <-
   EUpredIndex[r,cc]*exp(stdAllresids[round(runif(1,min=0.5,max=length(stdAllresids)+0.5),0)]
   *EUresidSE[r])   }}

for (r in 1:CanFallmaxAge) { for (cc in 1:CanFallnumYears) { GH.indexMC[[2]]@index[r,cc,,,] <-
   CanFallpredIndex[r,cc]*exp(stdAllresids[round(runif(1,min=0.5,max=length(stdAllresids)+0.5),0)]
   *CanFallresidSE[r]) }}

for (r in 1:CanSprmaxAge) { for (cc in 1:CanSprnumYears) { GH.indexMC[[3]]@index[r,cc,,,] <-
   CanSprpredIndex[r,cc]*exp(stdAllresids[round(runif(1,min=0.5,max=length(stdAllresids)+0.5),0)]
   *CanSprresidSE[r]) }}
    }

if (residSamp==3) {   # Resample within indices and ages
for (r in 1:EUmaxAge) { for (cc in 1:EUnumYears) { GH.indexMC[[1]]@index[r,cc,,,] <-
   EUpredIndex[r,cc]*expEUresid[r,round(runif(1,min=0.5,max=ncol(EUresid)+0.5),0)] }}
for (r in 1:CanFallmaxAge) { for (cc in 1:CanFallnumYears) { GH.indexMC[[2]]@index[r,cc,,,] <-
    CanFallpredIndex[r,cc]*expCanFallresid[r,round(runif(1,min=0.5,max=ncol(CanFallresid)+0.5),0)] }}
for (r in 1:CanSprmaxAge) { for (cc in 1:CanSprnumYears) { GH.indexMC[[3]]@index[r,cc,,,] <-
    CanSprpredIndex[r,cc]*expCanSprresid[r,round(runif(1,min=0.5,max=ncol(CanSprresid)+0.5),0)] }}
    }

for (ind in 1:3) { GH.indexMC[[ind]]@catch.n <- GH.indexMC[[ind]]@index }

# Run XSA
GH.xsaBS <- FLXSA(GH.stockMC,GH.indexMC,GH.control)

# If the XSA does not converge in the maximum # of iterations, resample residuals and try again until it does
numIts <- GH.xsaBS@control@maxit

if (numIts > (maxIt-1)) {
  while (numIts > (maxIt-1)) {
for (r in 1:EUmaxAge) { for (cc in 1:EUnumYears) { GH.indexMC[[1]]@index[r,cc,,,] <-
EUpredIndex[r,cc]/expAllresids[round(runif(1,min=0.5,max=AllresidsnumCases+0.5),0)] }}

for (r in 1:CanFallmaxAge) { for (cc in 1:CanFallnumYears) { GH.indexMC[[2]]@index[r,cc,,,] <-
CanFallpredIndex[r,cc]/expAllresids[round(runif(1,min=0.5,max=AllresidsnumCases+0.5),0)] }}
```

```
for (r in 1:CanSprmaxAge) { for (cc in 1:CanSprnumYears) { GH.indexMC[[3]]@index[r,cc,,,] <-
CanSprpredIndex[r,cc]/expAllresids[round(runif(1,min=0.5,max=AllresidsnumCases+0.5),0)] }}

for (ind in 1:3) { GH.indexMC[[ind]]@catch.n <- GH.indexMC[[ind]]@index }

GH.xsaBS <- FLXSA(GH.stockMC,GH.indexMC,GH.control, paste("Run#: ",i, sep = ""))
numIts <- GH.xsaBS@control@maxit
    }
}
```

APPENDIX 2: R code methods to run the bootstrap bias corrected percentile method on a set of bootstrap replicates

```
####################################################################################
# "invPnorm"
## Inverse cumulative distribution function of a standard normal variate
# Author: David C. M. Miller (millerdcm@dfo-mpo.gc.ca)
# Date: 25 March 2007

setGeneric("invPnorm", function(Obj){
      standardGeneric("invPnorm")
      }
)
setMethod("invPnorm", signature(Obj="numeric"),
  function(Obj){
    pot <- 0
    x <- -4.001
      while ((Obj - pot) > 0.00001) {
        x <- x+0.001
        pot = pnorm(x)
      }
    return(round(x,7))
      }
)

####################################################################################
# "empCumFreq"
##  Empirical cumulative frequency distribution
# Author: David C. M. Miller (millerdcm@dfo-mpo.gc.ca)
# Date: 26 March 2007

setGeneric("empCumFreq", function(Obj, Value, ...){
      standardGeneric("empCumFreq")
      }
)
setMethod("empCumFreq", signature(Obj="matrix"),
  function(Obj, Value){
```

```
    numObs <- length(Obj)
    sortedProb <- matrix(data=NA,2,numObs)
    sortedProb[1,] <- sort(Obj)
    sortedProb[2,] <- seq((1/numObs),1,(1/numObs))

  diff <- array(NA,dim=c(1,numObs))

  for (colm in 1:numObs) {
    diff[1,colm] <- abs(sortedProb[1,colm]-Value)
  }

  minDiff <- as.numeric(which.min(diff[1,]))

  return(sortedProb[2,minDiff])
      }
)
```

##

 # "invEmpCumFreq"

Inverse empirical cummulative frequency distribution

 # Author: David C. M. Miller (millerdcm@dfo-mpo.gc.ca)

 # Date: 26 March 2007

```
setGeneric("invEmpCumFreq", function(Obj, Value, ...){
      standardGeneric("invEmpCumFreq")
      }
)
setMethod("invEmpCumFreq", signature(Obj="matrix"),
  function(Obj, Value){
    numObs <- length(Obj)
    sortedProb <- matrix(data=NA,2,numObs)
    sortedProb[2,] <- sort(Obj)
    sortedProb[1,] <- seq((1/numObs),1,(1/numObs))

  diff <- array(NA,dim=c(1,numObs))

  for (colm in 1:numObs) {
    diff[1,colm] <- abs(sortedProb[1,colm]-Value)
  }
```

```
  minDiff <- as.numeric(which.min(diff[1,]))

  return(sortedProb[2,minDiff])
      }
)
```

```
################################################################################
# "biasCorPercent"
## Bootstrap bias corrected percentile technique
# Author: David C. M. Miller (millerdcm@dfo-mpo.gc.ca)
# Date: 26 March 2007
# Input matrix of bootstrap samples and vector (1 row/column) of observed values for bias correction of each column of bootstrap samples
```

```
setGeneric("biasCorPercent", function(Obj, Obs, ...){
      standardGeneric("biasCorPercent")
      }
)
setMethod("biasCorPercent", signature(Obj="matrix"),#, Obs="matrix"),
  function(Obj, Obs){
    numYr <- length(Obj[1,])
    noBoots <- length(Obj[,1])

    biasCorValues <- matrix(data=NA,noBoots,numYr,dimnames=Obj@dimnames)

    for (colm in 1:numYr) {
      wrkCol <- matrix(sort(Obj[,colm]),noBoots,1)
      #obs <- Obs[1,colm]
      sortedCol <- matrix(data=NA,2,noBoots)
      sortedCol[2,] <- wrkCol
      sortedCol[1,] <- seq((1/noBoots),1,(1/noBoots))
      Z0 <- invPnorm(empCumFreq(wrkCol,Obs[colm]))#obs))
      for (btstrp in 1:noBoots){
        Za <- invPnorm(sortedCol[1,btstrp])
        biasCorValues[btstrp,colm] <- invEmpCumFreq(wrkCol,pnorm(2*Z0+Za))
        }
    }

    return(biasCorValues)
```

```
                }
)

################################################################################
```