VIMS Articles   Virginia Institute of Marine Science

2015

Properties of age compositions and mortality estimates derived from cohort slicing of length data

LE Ailloud
Virginia Institute of Marine Science

MW Smith
Virginia Institute of Marine Science

AY Then
Virginia Institute of Marine Science

KL Omori
Virginia Institute of Marine Science

GM Ralph
Virginia Institute of Marine Science

See next page for additional authors

Authors

LE Ailloud, MW Smith, AY Then, KL Omori, GM Ralph, and JM Hoenig

Original Article

Properties of age compositions and mortality estimates derived from cohort slicing of length data

Lisa E. Ailloud, Matthew W. Smith, Amy Y. Then, Kristen L. Omori, Gina M. Ralph, and John M. Hoenig*

Virginia Institute of Marine Science, College of William & Mary, P.O. Box 1346, Gloucester Point, VA 23062, USA

Corresponding author: tel: +1 804 684 7125; e-mail: hoenig@vims.edu

Cohort slicing can be used to obtain catch-at-age data from length frequency distributions when directly measured age data are unavailable. The procedure systematically underestimates the relative abundance of the youngest age groups and overestimates abundance at older ages. Cohort-sliced catch-at-age data can be used to estimate total mortality rate (Z) using a regression estimator or the Chapman–Robson estimator for right truncated data. However, the effect of cohort slicing on accuracy and precision of resulting Z estimates remains to be determined. We used Monte Carlo simulation to estimate the per cent bias and per cent root mean square error of the unweighted regression, weighted regression, and Chapman–Robson mortality estimators applied to cohort-sliced data. Incompletely recruited age groups were truncated from the cohort-sliced catch-at-age data using previously established recommendations and a variety of plus groups was used to combine older age groups. The sensitivity of the results to a range of plausible biological combinations of Z, growth parameters, recruitment variability, and length-at-age error was tested. Our simulation shows that cohort slicing can work well in some cases and poorly in others. Overall, plus group selection was more important in high K scenarios than it was in low K scenarios. Surprisingly, defining the plus group to start at a high age worked well in some cases, although length and age are poorly correlated for old ages. No one estimator was uniformly superior; we therefore provide recommendations concerning the appropriate estimator and plus group to use, depending on the parameters characterizing the stock. We further recommend that simulations be performed to determine exactly which plus group would be most appropriate given the scenario at hand.

Keywords: age composition, age distribution, age slicing, catch-at-age, cohort slicing, mortality.

Introduction

While there has been a recent shift in stock assessment methods towards using catch-at-length-based models, much of modern stock assessment remains based on catch-at-age models, which estimate population sizes and derive exploitation history by summing catches over time on a cohort-by-cohort basis. Size-structured models like MULTIFAN-CL (Fournier *et al.*, 1998) and Stock Synthesis (Methot, 2005) are often more informative than the simpler catch-at-age models, but these highly complex integrated assessment methods also tend to require more data, leaving simpler models like virtual population analysis (VPA) still used for data-poor species.
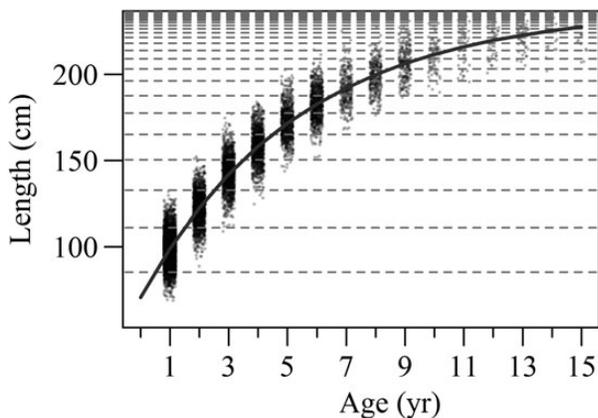
The catch-at-age approach is predicated on having reliable data on the age composition of the catch in each year. Age data can often be obtained from hard parts (e.g. otoliths, vertebrae, spines), but such techniques are labour-intensive and time-consuming, and

not applicable to many invertebrates. This information is therefore not always available to stock assessment scientists who have to extract age composition from the available fisheries catch-at-length data. The most common approaches used when no age estimates are available are Pauly and David's (1981) ELEFAN, and Fournier *et al.*'s (1990) MULTIFAN. When limited direct observations on age are available, an inverse age–length key (Hoenig and Heisey, 1987; Kimura and Chikuni, 1987) or a combined forward and inverse key (Hoenig *et al.*, 1994) might be used. While these tools reduce the need for direct ageing studies, they still require some age–length data to be collected, which is not always practical.

An alternative is to estimate the age composition from the length frequency distribution of the catch using cohort slicing (also known as age slicing). This requires a growth equation to be available but does not require information on variability in size at age. With this method, a length interval or "bin" is specified for each age

group and the number at each age is estimated as the number of observations in the corresponding length bin. The bin definitions are determined from a von Bertalanffy (or other) growth equation, following the assumption that ages are clearly separated by length bounds. The oldest age groups are lumped together in a catch-all "plus" group because, as fish grow, the relationship between body size and age weakens to the point that the oldest nominal ages are largely mixtures of ages (Figure 1). This method is currently being used in the assessment of many highly migratory species, including swordfish, *Xiphias gladius,* yellowfin tuna, *Thunnus albacares,* bigeye tuna, *Thunnus obesus,* Atlantic bluefin tuna, *Thunnus thynnus*, and North Atlantic albacore, *Thunnus alalunga* (International Commission for the Conservation of Atlantic Tunas, 2010, 2011, 2012a, b, 2014) as well as a number of demersal fisheries, including the witch flounder, *Glyptocephalus cynoglossus* (International Council for the Exploration of the Sea, 2012), European hake, *Merluccius merluccius*, red mullet, *Mullus barbatus*, red shrimp, *Aristeus antennatus*, and deep-water pink shrimp, *Parapenaeus longirostris* (General Fisheries Commission for the Mediterranean, 2012).

Cohort slicing is predicated on the assumption that there is no overlap in length among cohorts. Strictly speaking, this assumption is never met—size distributions for the oldest age groups always overlap. While the properties of cohort slicing have not yet been evaluated comprehensively, a few studies have explored the implications of its assumptions for the estimation of age composition. Mohn (1994) and Restrepo (1995) were the first to point out that cohort slicing tends to underestimate recruitment variability. When the cohorts are of equal abundance, the younger cohort contributes as much to the estimate of the older cohort as the older cohort contributes to the estimate of the younger cohort. Hence, the errors of misclassification cancel out. But, when the cohorts are of unequal size, the more abundant cohort contributes more to perceived size of the weaker cohort than the weaker cohort contributes to the more abundant one. Consequently, the abundance of weaker year classes tends to be overestimated and the abundance of stronger year classes tends to be underestimated. Furthermore, Kell and Kell (2011) compared cohort slicing with a more

sophisticated statistical method, mixture analysis, and suggested that cohort slicing underestimated the contribution of younger fish in an analysis of data on swordfish. Similar observations were made by Goodyear (1987) in the assessments of red drum, *Sciaenops ocellatus*, where he noted that young fish, being fast-growing in nature, were consistently misclassified as being older, while older fish, which are typically slow growing, were consistently misclassified to younger age classes, leading to underestimates of the first couple of age groups and overestimates of the remaining age groups (Figure 2).

In a study by Rodríguez-Marín *et al.* (2001), cohort slicing, length frequency analysis by MULTIFAN, and age–length key (from dorsal spine readings) were compared to determine the relative performance of each method in estimating catch-at-age for juvenile Atlantic bluefin tuna. They found no significant difference in the catch-at-age estimated using the three methods, but noticed that both cohort slicing and MULTIFAN underestimated the abundance of the strong 1994 cohort compared with estimates made using the age–length key. Similar observations were made by Turner and Terceiro (1994) who compared catch-at-age estimated by MULTIFAN with cohort slicing in juvenile Western bluefin tuna. In addition, they found that differences between the two methods widened in older age groups.

Cohort-sliced catch-at-age models have been criticized for being markedly inferior to catch-at-length models (Polacheck and Preece, 2001; Butterworth and Rademeyer, 2013). While it is true that, in data-rich situations, catch-at-length models can resolve problems associated with cohort slicing, under more difficult assessment conditions (i.e. when less data are available for the stock), it is not always evident that catch-at-length models perform better than cohort-sliced catch-at-age models (Kurota *et al.*, 2001; Kolody *et al.*, 2004). With increasing demand for more stock assessments to be carried out, but limited resources available to obtain the necessary data and perform sophisticated statistical analyses, cohort slicing may be a valuable tool for obtaining preliminary results. Furthermore, there is value in having a simpler model with which to compare other, more sophisticated models; those simpler models allow stock assessment scientists to explore what individual components of the dataset may indicate and may be very helpful when transitioning to more sophisticated models. Cohort slicing is still being used for a
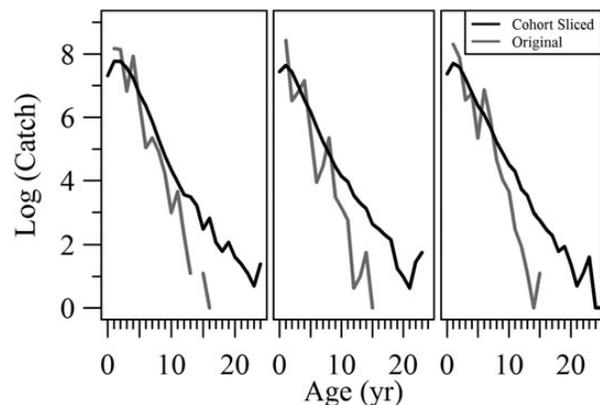

Figure 1. Simulated swordfish lengths plotted against age. The lengths were generated by adding a random error to the predicated lengths from the von Bertalanffy growth curve currently used in stock assessment (shown as the solid line). The dashed lines indicate the length bin separations used in cohort slicing. The parameters used in the simulation are: $K = 0.185$ year^{-1}, $L_\infty = 238$ cm LJF, $t_0 = -1.404$ year, $Z = 0.38$, $\sigma_r = 0.3$, and $\sigma_l = 4\% L_\infty$.

Figure 2. Differences between the actual sample composition and the cohort-sliced composition in three realizations of a simulation scenario. The parameter values used in the simulation are: $K = 0.185$ year^{-1}, $L_\infty = 238$ cm LJF, $t_0 = -1.404$ year, $Z = 0.6$, $\sigma_r = 0.7$, and $\sigma_l = 12\% L_\infty$.

number of species and as such it must be emphasized that practitioners know its properties and limitations, so they may take these into account when interpreting the stock assessment model results. For example, VPA is still being used for some stocks assessed by the International Commission for the Conservation of Atlantic Tunas (ICCAT) and will continue to be used for a number of years as transitions are made to statistical catch-at-age models; cohort slicing is used to obtain the required catches-at-age. There is a need to determine if perceived changes in stock dynamics are due to changes in abundance or to changes in methodology.

Having recognized the qualitative properties of the derived age composition, we chose to summarize the implications of age misclassification in subsequent use of the catch-at-age data. We used a Monte Carlo simulation to examine how ageing errors induced by cohort slicing affect estimates of total mortality rate derived from catch curve analysis, in part because catch curve analysis is used as one component of the assessment of Mediterranean swordfish. Of interest to our study was to: (i) determine which conditions produce large errors (i.e. evaluate the implications of having different definitions of the plus group, different magnitudes of recruitment variability and growth variability, etc.), and (ii) provide recommendation as to which method of mortality estimation is most suited for age-sliced data (i.e. given the existing biases, which estimation technique produces estimates with the smallest root mean square error).

Methods

We quantified errors in estimating age composition by repeatedly generating datasets with known properties and then analysing each dataset using cohort slicing. We used a factorial design to examine the effects of recruitment variability, individual growth variability, the von Bertalanffy growth coefficient, K, the total mortality rate, Z, and the first age in the plus group on catch curve estimates of total mortality rate. We simulated knife-edge selectivity by age (not length); we assumed that below a certain age, fish suffer no fishing mortality at all. The parameters used to generate the datasets were loosely patterned after the biology and population dynamics of *X. gladius* in the Mediterranean Sea, a species assessed by cohort slicing. We also bracketed these parameter values to see how robust the conclusions are under a variety of conditions (parameter values are listed in Table 1).

For each combination of factors, 10 000 populations were simulated and analysed according to the following procedure.

Table 1. Parameter values used in the simulation and case study of Mediterranean swordfish.

Parameter	Simulation	Mediterranean swordfish
K	0.185 or 0.4 year^{-1}	0.185 year^{-1}
L_∞	238 cm LJF	238 cm LJF
t_0	−1.404 years	−1.404 years
σ_l	4, 8, or 12% L_∞	4% L_∞
σ_r	0.3, 0.7, or 1.1	0.3
Z	0.3, 0.6, 1.0	0.38
Plus group	5, 7, 9, 11, 13, or 15 years	5, 7, 9, 11, 13, or 15 years

Recruitment variability parameter (σ_r) values were chosen from Myers *et al.* (1995) to cover the range that might be expected for *X. gladius*. LJF, lower jaw–fork length. Entries for Plus group refer to the youngest age in the plus group. For the case study, parameter values were patterned after the latest stock assessment for Mediterranean swordfish (International Commission for the Conservation of Atlantic Tunas, 2011), which uses the von Bertalanffy parameters estimated by Tserpes and Tsimenides (1995).

Generate an age composition with 41 ages (arbitrary number made sufficiently large to ensure the complete elimination of the population after 41 years) by generating 41 initial cohort sizes as lognormal random variables and projecting the abundance of each cohort forward to a specified age in the range 0–40 according to a constant mortality rate, Z. Thus, at one point in time, the age composition is given by,

$$N_t = N_{0t}e^{-Zt}, \quad t = 0, 1, \ldots, 40, \quad (1)$$

where N_t is the number of animals in cohort t at age t, N_{0t} the initial size of the cohort that is age t in the sample, $N_{0t} \sim \ln(\mu, \sigma_r^2)$, and $\mu = \ln(10\,000)$. Thus, the expected initial size of a cohort t is $E(N_{0t}) = \exp[\mu + (\sigma_r^2/2)] > 10000$.

Assign a length to each animal in the population by adding a normally distributed random error to the expected length of the animal as specified by the von Bertalanffy growth equation. Thus,

$$L_i = L_\infty(1 - e^{(-K(t_i-t_0))}) + \varepsilon_i, \quad (2)$$

where L_i is the length of the ith animal, t_i the age of the ith animal, the parameters K, L_∞, and t_0 are from the von Bertalanffy equation, and $\varepsilon_i \sim N(0, \sigma_l^2)$.

Establish length bins for cohort slicing and count the number of occurrences in each bin. The number of animals of age t, N_t, is estimated to be the number of animals where $LB_t \leq L < UB_t$ where the lower bound LB_t is the predicted length from the von Bertalanffy equation at age $t - \frac{1}{2}$ and UB_t is the predicted length-at-age $t + \frac{1}{2}$.

Estimate the total mortality rate from the estimated age composition using the Chapman–Robson estimator for truncated age distributions (Robson and Chapman, 1961), the unweighted regression estimator (see Ricker, 1975), and the weighted regression estimator (Maceina and Bettoli, 1998; Smith *et al.*, 2012).

The Chapman–Robson estimator (Chapman and Robson, 1960) was used in the form published by Robson and Chapman (1961) for truncated age distributions,

$$\hat{Z} = -\log\left(\frac{T}{n - m + T}\right), \quad (3)$$

where $\hat{Z}$ is the estimated total mortality rate, n the sample size, m the frequency of fish in the plus group (fish of age > k), and T is defined as,

$$T = N_1 + 2N_2 + 3N_3 + \cdots + kN_k + m(k + 1), \quad (4)$$

where N_i corresponds to the number of fish at age i, starting with the age of maximum catch plus 1 year [following recommendations by Smith *et al.* (2012)] and ending with the last age group before the plus group, N_k.

The unweighted regression estimator as defined in Ricker (1975) was used to estimate Z by fitting a linear regression to the log-transformed catch-at-age data and calculating the negative of the slope obtained by ordinary least squares. The first age group used was the age of maximum catch, following recommendations of Smith *et al.* (2012), and the oldest ages were truncated following different plus group definitions. The weighted regression estimator (Maceina and Bettoli, 1998; Smith *et al.*, 2012) was used as an alternative to simple truncation for dealing with the low and often sporadic catches of older age groups. Weights were calculated following the method of Maceina and Bettoli (1998) who recommend using

the predicted log catch-at-age values from an unweighted regression model as weights in a subsequently fitted model. This weighting scheme, although *ad hoc*, successfully shifts weight away from the older age groups which could otherwise be highly influential in determining model fit.

The ability of each technique to recover the true value of Z was assessed by calculating per cent bias (%BIAS) and per cent root mean square error (%RMSE) for each scenario according to the equations

$$\%BIAS(\hat{Z}) = 100\frac{E(\hat{Z}) - Z}{Z}, \qquad (5)$$

and

$$\%RMSE(\hat{Z}) = \frac{100\sqrt{E(\hat{Z} - Z)^2}}{Z}, \qquad (6)$$

where $E()$ denotes expectation, which is approximated by averaging over simulation results.
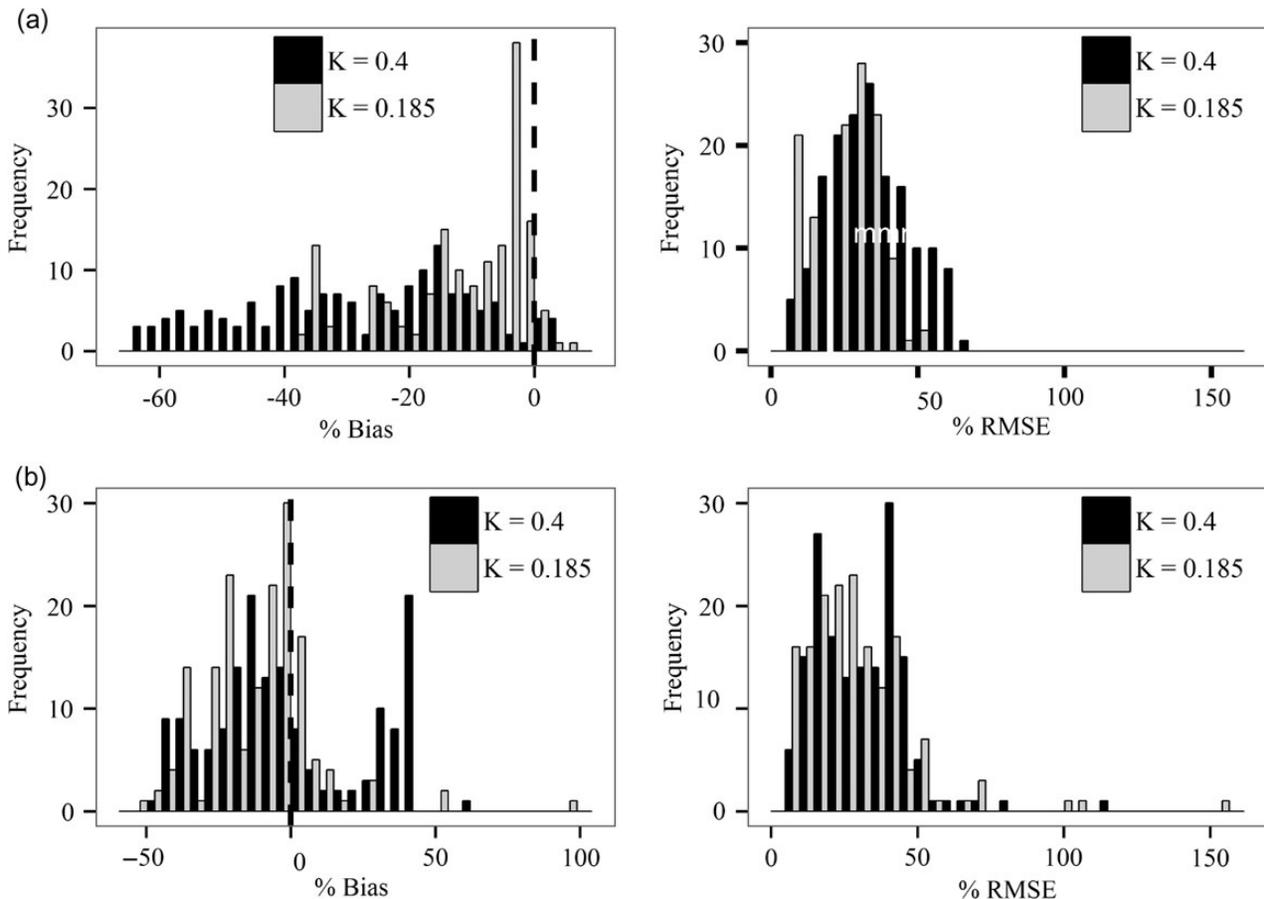
Results

Cohort slicing was found to systematically underestimate the abundance of the youngest age groups, while overestimating the contribution of the oldest age groups (Figure 2). Estimates of total

mortality rate were negatively biased in most scenarios (Figure 3). The weighted regression yielded nearly identical results to the unweighted regression but, in rare instances, performed slightly better than the unweighted regression. We therefore focus on the weighted regression for simplicity.

Simulations with low K ($K = 0.185$)

For the Chapman–Robson estimator, biases were mostly negative, ranging from −38 to +5%, with most biases occurring between 0 and −15% (Figure 3). Small positive per cent bias was observed at young plus groups for scenarios combining low Z, low to moderate length-at-age error, and medium to high recruitment variability (Figure 4). When length-at-age error was low, per cent biases were near zero throughout the range of plus groups, recruitment error, and Z. As length-at-age error increased, the first age in the plus group became more important, with per cent bias becoming increasingly negative with decreasing first age in the plus group. The largest negative per cent bias occurred at high Z and high length-at-age error, where bias reached −38%.

For the Chapman–Robson estimator, recruitment variability was an important factor in determining per cent RMSE, which was not the case for per cent bias (Figure 4). As a rule, when per cent bias was low, per cent RMSE was very dependent on the level of recruitment variability and when per cent bias was high, per cent RMSE was fairly insensitive to the level of recruitment

Figure 3. Histograms showing the range and distribution of per cent bias and per cent RMSE results across scenarios resulting from (a) the Chapman–Robson estimator and (b) the weighted regression estimator. Note the different scales used in the bias panels.
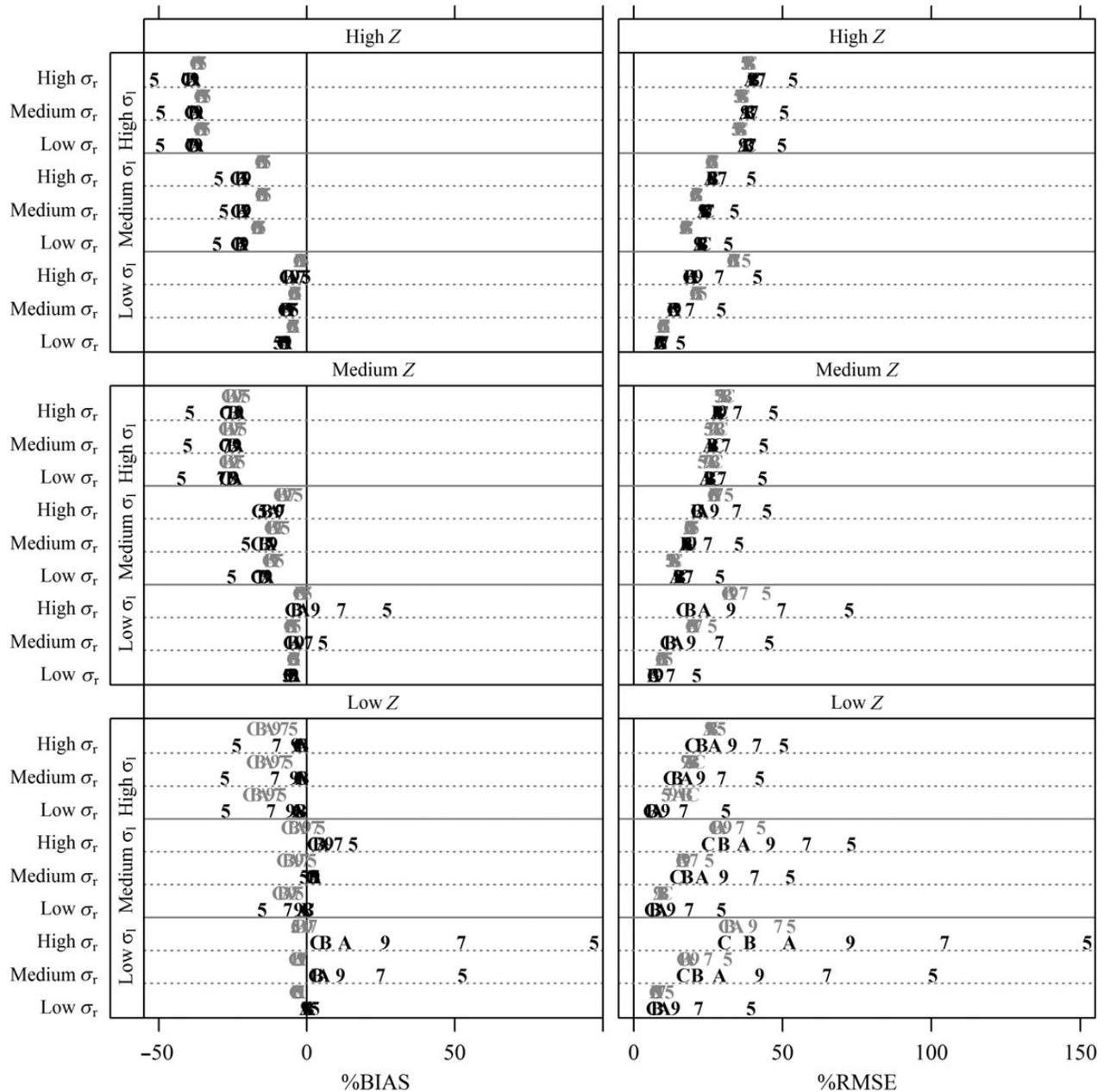
Figure 4. Simulation results from the low growth rate scenario ($K = 0.185$ year^{-1}). The left panel indicates per cent bias in the estimate of Z, the right panel indicates per cent RMSE in the estimate of Z. The black vertical lines present in the left and right panels indicate 0% bias and 0% RMSE, respectively. The numbers and letters represent the first age in the plus group: 5, 5 years; 7, 7 years; 9, 9 years; A, 11 years; B, 13 years; and C, 15 years. The grey numbers and letters are results from the Chapman–Robson estimator and the black numbers and letters are results from the weighted regression estimator. Solid horizontal grey lines divide results from scenarios with high, medium, and low length-at-age error. Dotted horizontal grey lines divide results from scenarios with high, medium, and low recruitment error. The parameters used for the last line of results (low Z, low σ_l and low σ_r) are patterned after Mediterranean swordfish.

variability (Figure 4). Overall, per cent RMSE could be controlled by selecting for an older plus group.

For the weighted regression estimator, bias reached 97% at high recruitment error, low Z, and low length-at-age error when the youngest plus group (age 5+) was being used (Figure 4), but, overall, most biases were around −30 to +5% (Figure 3). At low Z, per cent bias remained close to zero in the older plus groups, regardless of recruitment error and length-at-age error (Figure 4). At

high Z and low length-at-age error, per cent bias was close to zero regardless of recruitment error and plus group selection. Overall, selecting an older plus group helped keep per cent bias low, which is the opposite of what was observed in almost all cases for the Chapman–Robson estimator (Figure 4).

Per cent RMSE resulting from the weighted regression estimator followed the same patterns as per cent bias in that selecting for a higher plus group consistently brought per cent RMSE closer to

zero. Like per cent bias, per cent RMSE reached a peak (of 152%) when Z was low, length-at-age error was low, recruitment error was high, and the youngest plus group was being used (Figure 4). Higher Z and higher length-at-age error caused an increase in per cent RMSE and employing older plus groups helped keep per cent RMSE low. As was seen in the Chapman–Robson results, when either Z or length-at-age error was low, per cent RMSE from the weighted regression was very sensitive to recruitment error and plus group selection; as recruitment variability increased, so did per cent RMSE. The differences were most notable for younger plus groups.

The Chapman–Robson and regression estimators both yielded largely negative biases (Figure 3). Smallest biases ($\approx 0\%$) were observed when all variables were low, and highest biases ($\approx -40\%$) were reached when Z and length-at-age error were high, except the 97% bias observed at high recruitment error, low Z, young plus groups, and low length-at-age error when the regression estimator was used (Figure 4). Under both methods, increasing recruitment error increased per cent RMSE drastically at lower plus groups, but per cent bias was only affected when the regression estimator was used (Figure 4). The main difference between the two methods was the influence of plus group selection on per cent bias and per cent RMSE (Figure 4). With the weighted regression, selecting for a younger plus group inflated the per cent bias and per cent RMSE, whereas the Chapman–Robson estimator was much less sensitive to plus group selection and performed better, or almost as well, for young plus groups as for old plus groups. Overall, the range of per cent bias and per cent RMSE observed was narrower when the Chapman–Robson estimator was used compared with the regression estimator (Figure 3).

Simulations with high K ($K = 0.4$)

Increasing K from 0.185 to 0.4 increased the spread of per cent bias and per cent RMSE observed throughout the range of scenarios (Figure 3). Under the Chapman–Robson estimator, biases ranged from -64 to $+4\%$, with RMSE reaching up to 65%. Under the regression estimator, biases ranged from -45 to $+61\%$ with RMSE reaching 117% (Figures 3 and 5). Aside from the exceptional case where the recruitment variability was highest and Z and length-at-age error lowest, all scenarios displayed low per cent bias at a low plus group (5+ to 9+), regardless of which mortality estimator was used. The same pattern was observed for per cent RMSE when the Chapman–Robson estimator was used, but patterns in per cent RMSE were not as clear when the weighted regression was being used (Figure 5).

As was the case in the low K scenario, the Chapman–Robson estimator performed better in terms of per cent bias with younger plus groups than with older plus groups, but under the higher K scenarios ($K = 0.4$), the Chapman–Robson estimates were much more sensitive to plus group selection, performing worse with increasingly older plus groups (Figure 5). Per cent RMSE for the Chapman–Robson estimates was generally lowest, or close to lowest, when the first age in the plus group was young, the exceptions being when high recruitment variability was paired with low Z and length-at-age error (Figure 5).

When weighted regression was used, high positive biases around 40% appeared for all low Z scenarios when the youngest age in the plus group was high (9+ and up). High positive per cent bias also occurred for some low Z scenarios when the youngest age in the plus group was <9 (Figure 5). Per cent bias was generally lowest when young plus groups were used (5+ to 9+) and highly negative

for cases with high Z and medium to high length-at-age error (Figure 5). Per cent RMSE was lowest at medium Z, when length-at-age error and recruitment variability were low. In all scenarios where length-at-age error was low, per cent RMSE was highly sensitive to plus group definition, with the youngest plus group (5+) almost consistently performing worse than the other plus groups. On the contrary, when Z was low and length-at-age error was high, weighted regression performed better when the youngest plus group was being used. The highest per cent RMSE was caused by high recruitment error, when both Z and age-at-length error were low (Figure 5).

The bias of the Chapman–Robson estimator always becomes increasingly negative with an increase in the first age of the plus group. In contrast, the weighted regression estimator displays a more complex behaviour (Figure 5). In general, choosing a young plus group minimized per cent bias for both estimators (Figure 5); however, there are some exceptions for the regression estimator (Figure 5, high recruitment variability with low length-at-age error and low to medium Z). For both estimators, proper plus group selection was key to reducing both per cent bias and per cent RMSE. In a high K situation, the preferred estimator, in terms of both minimum per cent bias and minimum per cent RMSE, varied across scenarios (Figure 5).

Discussion

Our simulations show that cohort slicing can work well in some cases and poorly in others. The impact of certain parameters on the per cent bias and per cent RMSE of the mortality estimate is more predictable than others. Length-at-age error, recruitment variability, and mortality rate influenced the per cent bias and per cent RMSE in a similar way across estimators; as a rule, an increase in these variables resulted in an increase in per cent bias and per cent RMSE. The influence of K on per cent bias and per cent RMSE was not always straightforward, but, overall, plus group selection was more important in high K scenarios than it was in low K scenarios. A surprising result that came out of the study was that sometimes a high plus group provides good results for mortality estimation (although length and age are poorly correlated for old ages).

No one estimator was uniformly superior. With high K, the weighted regression performed as well or better than the Chapman–Robson estimator for medium and high Z scenarios, while the Chapman–Robson estimator performed better for low Z scenarios. The opposite was true with low K: the Chapman–Robson estimator performed better than the weighted regression for medium and high Z scenarios, while weighted regression performed better for low Z scenarios. Our recommendations concerning the appropriate estimator and plus group to use are outlined in Table 2. These points are important guidelines for reducing bias induced by cohort slicing; however, the variables taken into consideration in this study do not account for all possible sources of uncertainty likely to affect the resulting Z estimate. As with any other assessment tool, additional sources of uncertainty, such as the quality of the length data and variations in the population dynamics of the stock, are important factors to consider when providing recommendations in the stock assessment process as these are likely to exacerbate the biases observed in this simulation. That being said, the parameters included in our simulation can act as proxies for understanding the effect additional sources of uncertainty would have on the results. If the concern is measurement or recording error of body lengths, and if this error is assumed to be
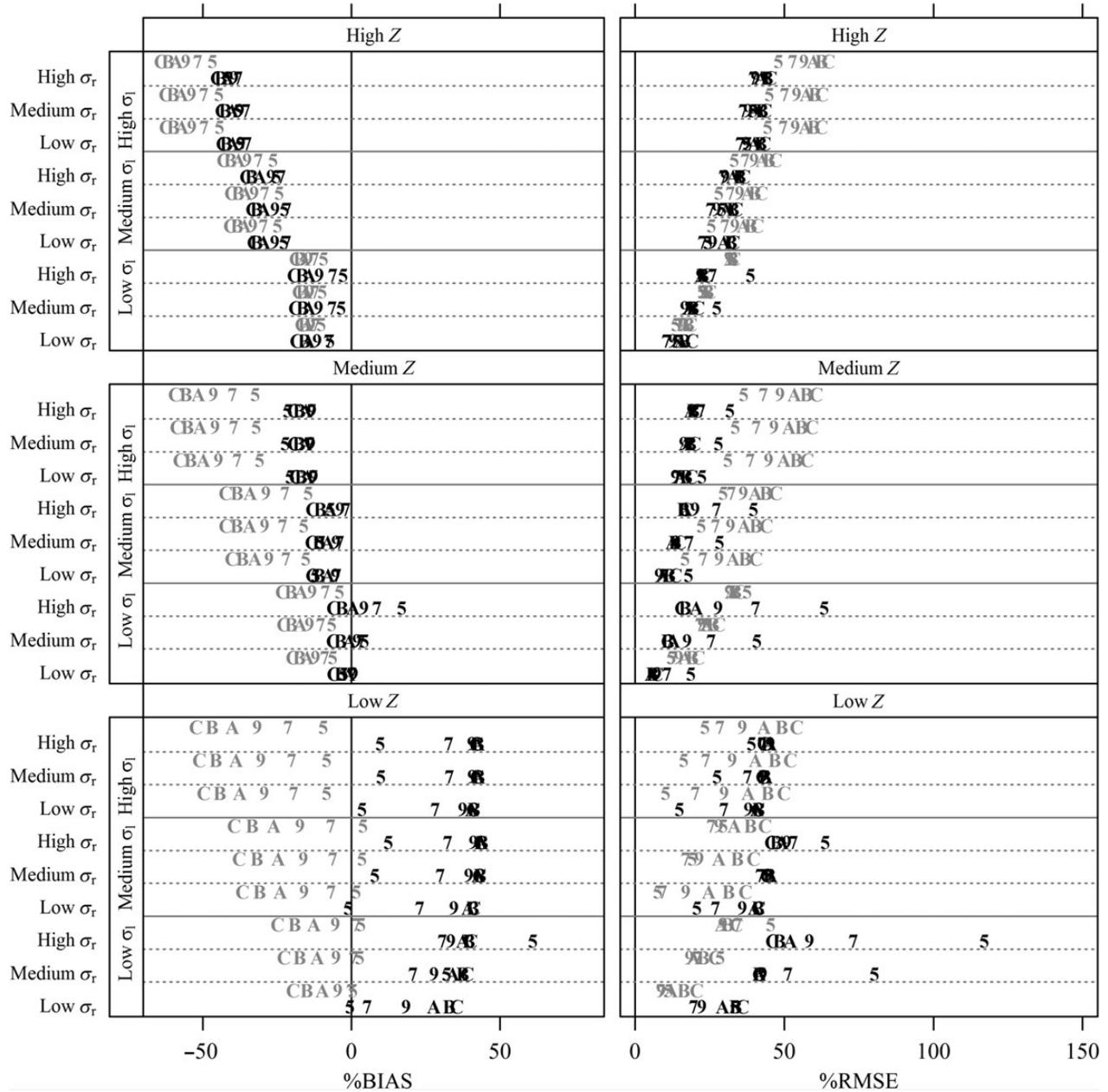
Figure 5. Simulation results from the high growth rate scenario ($K = 0.4$ year^{-1}). The left panel indicates per cent bias in the estimate of *Z*, and the right panel indicates per cent RMSE in the estimate of *Z*. The black vertical lines present in the left and right panels indicate 0% bias and 0% RMSE, respectively. The numbers and letters represent the first age in the plus group: 5, 5 years; 7, 7 years; 9, 9 years; A, 11 years; B, 13 years; and C, 15 years. The grey numbers and letters are results from the Chapman–Robson estimator and the black numbers and letters are results from the weighted regression estimator. Solid horizontal grey lines divide results from scenarios with high, medium, and low length-at-age error. Dotted horizontal grey lines divide results from scenarios with high, medium, and low recruitment error.

random with mean 0, then this source of error would be expected to act like the length-at-age error included in our simulations. Similarly, if the concern is that cohort-specific schooling behaviour causes cohorts to be missed at random, then variation in recruitment strength would have the same effect (i.e. strong and weak cohorts in the catch-at-age sample).

We consider swordfish in the Mediterranean Sea as a case study. Appropriate plus group selection was discussed in the latest stock assessment (International Commission for the Conservation of Atlantic Tunas, 2011). In 2007, a plus group of 10 was used for cohort slicing, but, in the following assessment, the plus group was reduced to five, with the working group stating that there was not enough information available on the length distribution of older ages to justify splitting the catch-at-size data into ages >5 (International Commission for the Conservation of Atlantic Tunas, 2011). From Figure 1, patterned after swordfish, it likewise seems unlikely that we would be able to estimate age composition accurately for ages higher than 5. However, if our goal is to

Table 2. Summary table of recommendations on which estimator and plus group to use for estimating Z depending on the magnitude of K, Z, and σ_I assumed for the population under study.

Low K	High σ_I	Medium σ_I	Low σ_I
High Z	CR[a]	CR[a]	CR[a]
Medium Z	CR low plus group	CR[a]	CR high plus group
Low Z	WR highest plus group	WR highest plus group	WR highest plus group
High K	**High σ_I**	**Medium σ_I**	**Low σ_I**
High Z	WR intermediate to low plus group	WR intermediate to low plus group	WR intermediate to low plus group
Medium Z	WR intermediate plus group	WR intermediate plus group	WR intermediate plus group
Low Z	CR lowest plus group	CR intermediate to low plus group	CR intermediate to low plus group

CR stands for Chapman–Robson and WR for weighted regression methods. Recruitment variability, σ_r, was not included in this summary table because increasing σ_r was either not influential or exacerbated per cent bias and per cent RMSE in nearly all scenarios but did not change conclusions concerning the best method and plus group to use in the estimation of Z.
[a]Results are not sensitive to the choice of plus group.

perform a catch curve analysis, our results show that the derived age composition provides information on the mortality rate even if the first age in the plus group is high (see the last line in Figure 4). In fact, if we look at our results in the context of swordfish, which is characterized by the parameters outined in Table 1, we notice that plus group selection can matter. If the weighted regression method is being used to estimate Z, then switching from a plus group starting at 10 to a plus group starting at 5 would inflate the RMSE from 7 to 40%, or even higher if recruitment variability for the species is higher than the case simulated here (Figure 4, last three lines pertaining to weighted regression estimator). However, if the Chapman–Robson method is being used, the practitioner has more flexibility in defining the plus group since the estimator, under this set of parameter values, shows little sensitivity to plus group selection (Figure 4). Overall, the best estimate of Z is obtained by using the weighted regression method with a plus group starting at 15 (Figure 4). Thus, proper care should be given in selecting an appropriate estimator and plus group given the variables at hand. The ICCAT working group performed both catch curve analysis and catch-at-age analysis (extended survivors analysis; International Commission for the Conservation of Atlantic Tunas, 2011). Our results are pertinent to the former. In a larger context, the choice of plus group is also influenced by the need to be able to estimate spawning biomass; this is made more difficult if the plus group contains immature fish. Also, it is advantageous to define a plus group that is homogeneous with respect to exploitation rate. It is thus fortunate that the simulations indicate there is some flexibility in the choice of plus group definition.

Though it is clear that cohort slicing will give biased estimates of age composition, it is not known how these biases will propagate through the stock assessment process. There is a need to explore the implications of such bias for management recommendations. Kell and Kell (2011) pointed out that negatively biased estimates of Z at younger ages from catch-at-age obtained from cohort slicing resulted in an overestimation of biological reference points. As our simulation results suggest, when Z is high, it is likely to be underestimated regardless of the estimation technique used. This is of particular concern for stocks with mortality rates approaching the target identified by management bodies, because if the bias causes Z to be below the identified target, it could give a false indication that the stock is healthy and could lead to total allowable catches being set too high. If mortality is consistently being underestimated year after year, this could lead to long-term detrimental effects for the stock in question, as was experienced with North Atlantic groundfish stocks (Steele *et al.*, 1992). Of less concern for

conservation would be if the opposite situation were to happen: Z is low so using the weighted regression estimator leads to high positive biases in Z estimates. This situation is less likely to be detrimental to the stock as it would lead scientists to take a more precautionary approach than is necessary; but the result may be detrimental to the fishery. It may cause confusion in the stock assessment process as other variables may indicate that the stock is in fact stable or rebuilding. The confusion could also create conflict among stakeholders as inappropriate management decisions would be reached for the stock and for neighbouring stocks if the stock was part of a large-scale mixed stocks management complex (Tuckey *et al.*, 2007). Porch (2000) furthermore called attention to the fact that since signs of changes in year-class strength or mortality rates are made harder to discern through cohort slicing, it may lead scientists into thinking a population is in fact stable, which is likely to result in inappropriate management recommendations for the stock (International Commission for the Conservation of Atlantic Tunas, 2010; Kell and Kell, 2011).

Several important pelagic species, among them swordfish, are managed based on the results from an age-structured model conducted on age compositions derived from cohort slicing. It becomes important to examine how errors propagate from cohort slicing to catch-at-age analysis to revision of biological reference points and quota setting, and how these errors are affected by the choice of plus group. Such an evaluation can be accomplished by simulating an "actual" population, drawing a set of samples to determine "perceived" stock status, calculating catch quotas based on the assessment results, and the harvest control rule, then feeding the harvest, recruitment, and other parameters back into the simulation of the actual population to arrive at the next population state. This process could, additionally, involve updating the biological reference points and harvest control rule based on the latest information. The process is alternated between population updates and population assessment to determine long-term behaviour of the assessment and management procedure; the whole simulation is repeated many times to characterize the variability arising from random events (recruitment, sampling, etc.). Especially when the perception of the stock dynamics does not match the actual stock dynamics closely, it becomes necessary to simulate many possible states of nature to ensure that the conclusions are generally applicable rather than dependent on the particular stock dynamics chosen for the simulation. It may also be necessary to evaluate alternative harvest control rules if the status quo does not appear to perform well, e.g. simulate control rules that reduce the target exploitation rate.

Cohort slicing is a crude method for estimating age composition and its performance can be good or bad, depending on circumstances. It remains useful in specific areas, especially stock assessments in data-limited fisheries. It can also provide a test of the reasonableness of results from other methods. However, because its performance varies considerably from one situation to the next, care must be taken in the use and interpretation of the results from cohort slicing.

Acknowledgements

References

Butterworth, D. S., and Rademeyer, R. A. 2013. A comparison of initial statistical catch-at-age and catch-at-length assessments of eastern Atlantic Bluefin tuna. ICCAT Collective Volume of Scientific Papers, 69: 710–741.

Chapman, D. G., and Robson, D. S. 1960. The analysis of a catch curve. Biometrics, 16: 354–368.

Fournier, D. A., Hampton, J., and Sibert, J. R. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. Canadian Journal of Fisheries and Aquatic Sciences, 55: 2105–2116.

Fournier, D. A., Sibert, J. R., Majkowski, J., and Hampton, J. 1990. MULTIFAN: a likelihood-based method for estimating growth parameters and age composition from Multiple Length Frequency Datasets illustrated using data from southern bluefin tuna (*Thunnus maccoyi*). Canadian Journal of Fisheries and Aquatic Sciences, 47: 301–317.

General Fisheries Commission for the Mediterranean (GFCM). 2012. Report of the working group on stock assessment of demersal species. Split, Croatia, 5–9 November 2012.

Goodyear, C. P. 1987. Status of the red drum stocks of the Gulf of Mexico. USDOC, NMFS, SEFC, Miami Laboratory Contribution CRD 86/87-34. 113 pp.

Hoenig, J. M., and Heisey, D. M. 1987. Use of a loglinear model with the EM algorithm for correcting estimates of stock composition and converting length to age. Transactions of the American Fisheries Society, 116: 232–243.

Hoenig, J. M., Heisey, D. M., and Hanumara, R. C. 1994. A new approach to age–length keys: using last year's and this year's data to estimate age composition. ICCAT Collective Volume of Scientific Papers, 42: 124–128.

International Commission for the Conservation of Atlantic Tunas (ICCAT). 2010. Report of the 2010 bigeye tuna stock assessment session (Pasaia, Gipuzkoa, Spain, 5–9 July 2010). ICCAT Secretariat, Madrid.

International Commission for the Conservation of Atlantic Tunas (ICCAT). 2011. Report of the 2010 ICCAT Mediterranean swordfish stock assessment meeting (Madrid, Spain, 28 June to 2 July 2010). ICCAT Collective Volume of Scientific Papers, 66: 1405–1470.

International Commission for the Conservation of Atlantic Tunas (ICCAT). 2012a. Report of the 2012 Atlantic bluefin tuna stock assessment session (Madrid, Spain, 4–11 September 2012). ICCAT Collective Volume of Scientific Papers, 33: 652–880.

International Commission for the Conservation of Atlantic Tunas (ICCAT). 2012b. Report of the 2011 ICCAT Yellowfin tuna stock assessment session (San Sebastián, Spain, 5–12 September 2011). ICCAT Collective Volume of Scientific Papers, 68: 655–817.

International Commission for the Conservation of Atlantic Tunas (ICCAT). 2014. Report of the 2013 ICCAT North and South Atlantic Albacore stock assessment meeting (Sukarrieta, Spain, 17–24 June 2013). ICCAT Collective Volume of Scientific Papers, in press.

International Council for the Exploration of the Sea (ICES). 2012. Report of the working group on assessment of new MoU species. Denmark, 5–9 March 2012. ICES CM 2012/ACOM: 20.

Kell, L. T., and Kell, A. 2011. A comparison of age slicing and statistical age estimation for Mediterranean Swordfish (*Xiphias gladius*). ICCAT Collective Volume of Scientific Papers, 66: 1522–1534.

Kimura, D., and Chikuni, S. 1987. Mixtures of empirical distributions: an iterative application of the age–length key. Biometrics, 43: 23–35.

Kolody, D. S., Jumppanen, P. C., Ricard, D. G., Hartog, J. R., Preece, A. L., and Polacheck, T. 2004. SESAME: a simulation-estimation stock assessment model evaluation project focused on large pelagic species. CSIRO Marine Laboratories Report 241. CSIRO Publishing, Collingwood.

Kurota, H., Tsuji, S., Takahashi, N., Hiramatsu, K., and Itoh, I. 2001. Exploration of cohort analysis based on catch at length data for southern bluefin tuna. CCSBT-SC/0108/32, 40 pp.

Maceina, M. J., and Bettoli, P. W. 1998. Variation in largemouth bass recruitment in four mainstream impoundments on the Tennessee River. North American Journal of Fisheries Management, 18: 998–1003.

Methot, R. D. 2005. Technical description of the Stock Synthesis II assessment program. NOAA Technical Memorandum SEDAR 16-AW-04.

Mohn, R. 1994. A comparison of three methods to convert catch at length data into catch-at-age. ICCAT Collective Volume of Scientific Papers, 42: 110–119.

Myers, R. A., Bridson, J., and Barrowman, N. J. 1995. Summary of worldwide spawner and recruitment data. Canadian Technical Report of Fisheries and Aquatic Sciences 2024, 331 pp.

Pauly, D., and David, N. 1981. ELEFAN I, a BASIC, program for the objective extraction of growth parameters from length-frequency data. Reports on Marine Research, 28: 205–211.

Polacheck, T., and Preece, A. 2001. An Integrated Statistical Time Series Assessment of the Southern Bluefin Tuna Stock based on Catch at Age Data. CCSBT-SC/0108/20.

Porch, C. E. 2000. Status of the red drum stocks of the Gulf of Mexico. Southeast Fisheries Science Center. Sustainable Fisheries Division Contributions: SFD-99/00-85. 62 pp.

Restrepo, V. R. 1995. Application of cohort slicing and tuned VPA to simulated data that includes variability in length at age. ICCAT Collective Volume of Scientific Papers, 44: 67–71.

Ricker, W. E. 1975. Computation and Interpretation of Biological Statistics of Fish Populations. Bulletin of the Fisheries Research Board of Canada 191.

Robson, D. S., and Chapman, D. G. 1961. Catch curves and mortality rates. Transactions of the American Fisheries Society, 90: 181–189.

Rodríguez-Marín, E., Rodríguez-Cabello, C., Ortíz De Zárate, V., and Cort, J. L. 2001. Comparison of three methods to estimate age composition of juvenile east Atlantic bluefin tuna (*Thunnus thynnus*). ICCAT Collective Volume of Scientific Papers, 52: 1215–1225.

Smith, M. W., Then, A. Y., Wor, C., Ralph, G., Pollock, K. H., and Hoenig, J. M. 2012. Recommendations for catch-curve analysis. North American Journal of Fisheries Management, 32: 956–967.

Steele, D. H., Andersen, R., and Green, J. M. 1992. The managed commercial annihilation of northern cod. Newfoundland Studies, 8: 34–68.

Tserpes, G., and Tsimenides, N. 1995. Determination of age and growth of swordfish, *Xiphias gladius* L., 1758, in the eastern Mediterranean using anal-fin spines. Fishery Bulletin US, 93: 594–602.

Tuckey, T., Yochum, N., Hoenig, J., Lucy, J., and Cimino, J. 2007. Evaluating localized vs. large-scale management: the example of tautog in Virginia. Fisheries, 32: 21–28.

Turner, S. C., and Terceiro, M. 1994. Estimation of west Atlantic bluefin tuna, *Thunnus thynnus*, age composition with length composition analysis. ICCAT Collective Volume of Scientific Papers, 42: 173–180.

Handling editor: Howard Browman