Taylor & Francis
Taylor & Francis Group

North American Journal of Fisheries Management

ISSN: 0275-5947 (Print) 1548-8675 (Online) Journal homepage: http://afs.tandfonline.com/loi/ujfm20

Recommendations for Catch-Curve Analysis

Matthew W. Smith , Amy Y. Then , Catarina Wor , Gina Ralph , Kenneth H. Pollock & John M. Hoenig

To cite this article: Matthew W. Smith , Amy Y. Then , Catarina Wor , Gina Ralph , Kenneth H. Pollock & John M. Hoenig (2012) Recommendations for Catch-Curve Analysis, North American Journal of Fisheries Management, 32:5, 956-967, DOI: 10.1080/02755947.2012.711270

To link to this article: http://dx.doi.org/10.1080/02755947.2012.711270

Published online: 25 Sep 2012.

North American Journal of Fisheries Management 32:956–967, 2012

ISSN: 0275-5947 print / 1548-8675 online
DOI: 10.1080/02755947.2012.711270

ARTICLE

Recommendations for Catch-Curve Analysis

Matthew W. Smith,* Amy Y. Then, Catarina Wor, and Gina Ralph
Virginia Institute of Marine Science, College of William and Mary, Post Office Box 1346, Gloucester Point, Virginia 23062, USA

Kenneth H. Pollock
Department of Biology, North Carolina State University, Post Office Box 7617, Raleigh, North Carolina 27695-7617, USA

John M. Hoenig
Virginia Institute of Marine Science, College of William and Mary, Post Office Box 1346, Gloucester Point, Virginia 23062, USA

Abstract
Three common cross-sectional catch-curve methods for estimating total mortality rate (Z) are the Chapman–Robson, regression, and Heincke estimators. There are five unresolved methodological issues: (1) which is the best estimator, (2) how one should determine the first age-group to use in the analysis, (3) how the variance estimators perform; and, for regression estimators, (4) how the observations should be weighted, including (5) whether and how the oldest ages should be truncated. We used analytical methods and Monte Carlo simulation to evaluate the three catch-curve methods, including unweighted and weighted versions of the regression estimator. We evaluated four criteria for specifying the first age-class used. Regression estimators were evaluated with four different methods of right data truncation. Heincke's method performed poorly and is generally not recommended. The two-tailed χ^2 test and one-tailed z-test for full selectivity described by Chapman and Robson did not perform as well as simpler criteria and are not recommended. Estimates with the lowest mean square error were generally provided by (1) the Chapman–Robson estimator with the age of full recruitment being the age of maximum catch plus 1 year and (2) the weighted regression estimator with the age of full recruitment being the age of maximum catch and with no right truncation. Differences in performance between the two methods were small (<6% of Z). The Chapman–Robson estimator of the variance of $\hat{Z}$ had large negative bias when not corrected for overdispersion; once corrected, it performed as well as or better than all other variance estimators evaluated. The regression variance estimator is generally precise and slightly negatively biased. We recommend that the traditional Chapman–Robson approach be corrected for overdispersion and used routinely to estimate Z. Weighted linear regression may work slightly better but is completely ad hoc. Unweighted linear regression should no longer be used for analyzing catch-curve data.

There is a rich literature on using age-frequency data, commonly referred to as catch-curve data, to estimate the instantaneous total mortality rate (Z; Ricker 1975; Seber 1982; Dunn et al. 2002; Tuckey et al. 2007; Thorson and Prager 2011). This suite of methods has been extensively studied, with the result that the Chapman and Robson (1960) and the regression-based methods (see Ricker 1975 or Seber 1982) are generally preferred for the analysis of catch-curve data (Ricker 1975; Dunn et al. 2002). Use of the early method of Heincke (1913) has persisted (e.g., Jensen 1996; Weber et al. 2011); however, much of the use of Heincke's method is unnecessary because the more accurate and precise regression and Chapman–Robson estimators could be used.

The assumptions usually made for catch-curve analysis are that (1) there are no errors in the estimation of age composition, (2) recruitment is constant or at least varies without trend over

*Corresponding author: mws212@vims.edu
Received November 30, 2011; accepted July 4, 2012
Published online September 25, 2012

time, (3) Z is constant over time and across ages, and (4) above some determined age, all animals are equally available and vulnerable to the fishery and the sampling process. Graphical depiction of catch-curve data can reveal violations of the assumptions. Tuckey et al. (2007) discussed the situation in which mortality changes over time or across ages. Jensen (1984) discussed the problem of heterogeneous mortality rates among subpopulations. In addition, graphical analysis of catch-curve data often reveals an ascending limb and a descending limb: the ascending limb is believed to represent catches of age-groups that are not fully recruited to the fishing gear, and the descending limb depicts the catches of fully recruited fish. Standard catch-curve analysis only utilizes the descending limb to estimate Z; deciding what age marks the beginning of the descending limb (age of full recruitment) is necessary for the implementation of these estimators.

Alternative approaches to catch-curve analysis have been developed that utilize maximum likelihood to simultaneously estimate gear selectivity and Z, which eliminates the need to choose heuristically an age of full recruitment (Thorson and Prager 2011). This type of approach is attractive because it utilizes information in the ascending limb that is lost in standard catch-curve analysis and has been demonstrated to improve estimation accuracy when compared with standard catch-curve methods. Despite these advantages, it is unlikely that this method will either completely supplant the use of standard catch-curve analysis or be used in the absence of standard catch-curve analysis, as multiple estimators are useful for examining the robustness of results. Therefore, how best to deal with nonuniform selectivity when standard catch-curve methods are applied has yet to be determined.

There is little guidance in the literature on choice of the age of full recruitment for catch-curve analysis. Most studies use the age of peak abundance (hereafter, "Peak" criterion) as the fully recruited age. This decision appears to be logical if the assumptions for catch-curve analysis are met, if nonsampling error is low, and if the gear selectivity curve is steep (e.g., the first age-group is 50% selected by the fishery and the sampling gear, while the second and subsequent age-groups are 100% selected). Failure of these assumptions can result in the sample peak age of abundance occurring at a younger or older age than the true age of full recruitment. For example, if one were to look at a time series of cross-sectional catch curves, a very strong year-class may be identified as the age of peak abundance for several consecutive years despite constant gear selectivity and Z. When the selectivity curve is gradual (e.g., age 1 is 50% selected, age 2 is 80% selected, age 3 is 90% selected, and ages 4 and above are 100% selected), one may by chance observe a peak abundance at an age younger than the true age of full recruitment. That is, the age-group with peak abundance may not correspond to the true age of full recruitment. Pauly (1984) suggested using the age-group 1 year beyond the age of peak abundance (hereafter, "Peak Plus" criterion) as the age of full recruitment, thus potentially sacrificing one usable data point for protection against bias due to incomplete recruitment. Ricker (1975) suggested that one may wish to use the second or even third age-group beyond the peak. However, starting at an older age implies a reduction in information and thus a reduction in precision. This is particularly problematic if fish are subjected to high Z, and thus few age-groups are available.

Chapman and Robson (1960) proposed a one-tailed z-test for identifying the age of full recruitment. Robson and Chapman (1961) modified the z-test into a χ^2 test. However, the procedure for implementing the χ^2 test and the associated statistical properties are not clear (see Appendix). The z-test utilizes the Heincke and Chapman–Robson survival estimators to identify discrepancies in the catch frequency of the youngest age-group relative to the older age-groups. The z-test will reject a candidate age for the first fully recruited age-class only if the catch is lower than expected and will accept the candidate age if the catch at that age exceeds the expected catch at age or is deficient but within the accepted sampling error. We have seen no guidance on the choice of α; therefore, the choice seems arbitrary. To our knowledge, the z-test has not been used in real application and its performance has not been evaluated.

Dunn et al. (2002) used simulation to study the behavior of the Chapman–Robson and regression mortality rate estimators over a range of Z and errors. This thorough and well-executed study did not, however, deal with the issue of selectivity and, hence, the choice of age at full recruitment. Dunn et al. (2002) did not evaluate weighted regression methods as suggested by Maceina and Bettoli (1998). Furthermore, the performance of the associated variance estimators for the estimated mortality rates has not been examined. Through simulation, we investigated the effect of the Peak, Peak Plus, χ^2 test, and z-test criteria for selecting age of full recruitment on the performance of the Chapman–Robson and regression-based (weighted and unweighted) Z estimators. We also looked at variance estimators, including the use of a correction for the Chapman–Robson method. Performance of the Heincke estimator was evaluated using only the Peak selection criterion. The estimators were evaluated over combinations of Z, sample size, and nonsampling error. Two different patterns of gear selectivity were also explored to evaluate effects of selectivity on estimator performance. This study provides guidelines for selecting the most appropriate age of full recruitment, method of estimating Z, and method for estimating the variance of the estimated Z.

THE ESTIMATORS

Chapman and Robson (1960) developed a method for estimating the annual survival rate (S) based on the assumption that the duration of life follows a geometric distribution (this implies annual reproduction). They also presented an estimator for Z, noting that it was biased but only slightly so at large sample sizes. The Chapman–Robson estimator ($CR[\hat{Z}]$), written in the form presented by Hoenig et al. (1983), is

$$CR(\hat{Z}) = \log_e\left(\frac{1+\bar{T}-T_C-\frac{1}{N}}{\bar{T}-T_C}\right) - \frac{(N-1)(N-2)}{N[N(\bar{T}-T_C)+1][N+N(\bar{T}-T_C)-1]}, \quad (1)$$

where $\bar{T}$ is the mean age of fish in the sample that are greater than or equal to age T_C; T_C is the age of full recruitment; and N is the sample size of fish greater than or equal to age T_C. The first term transforms the Chapman–Robson estimate of S into an estimate of Z; the second term reduces the bias that is induced by the transform.

An approximate (large-sample) variance estimator for $CR(\hat{Z})$ is

$$V\hat{A}R[CR(\hat{Z})] = \frac{\left[1-e^{-CR(\hat{Z})}\right]^2}{Ne^{-CR(\hat{Z})}}. \quad (2)$$

The method developed by Heincke (1913) estimates S as

$$\hat{S} = \frac{N-N_0}{N}, \quad (3)$$

where N_0 is the number of fish in the first fully recruited age (i.e., age T_C) and N is the number of fully recruited fish in the sample (i.e., age T_C and older). This method assumes that the catch of the youngest fully recruited age-group N_0 is binomially distributed with parameters (N, $1-S$). Heincke's method does not directly use any of the information in the other age-classes. The variance of the Heincke estimator follows directly from the binomial distribution such that

$$V\hat{A}R(\hat{S}) = \frac{\hat{S}(1-\hat{S})}{N}. \quad (4)$$

The estimated total mortality rate ($\hat{Z}$) is found by taking $-\log_e(\hat{S})$, and a large-sample variance estimator is derived using the delta method as

$$V\hat{A}R(\hat{Z}) = \frac{1-\hat{S}}{N\hat{S}}. \quad (5)$$

All of the unweighted regression-based estimators estimate Z as the negative of the slope obtained using ordinary least squares to regress $\log_e(N_X)$ on X, where N_X is the catch at age X. This model assumes that $VAR[\log_e(N_X)]$ is independent of age.

Several papers have noted that regression mortality estimators are prone to negative bias from the inclusion of infrequently caught older age-groups. Right truncation of the data set is commonly applied to reduce bias in regression estimators, as suggested by Chapman and Robson (1960) and evaluated by Dunn et al. (2002). Four methods of right truncation are considered here to investigate the relative impact on the Z estimators in conjunction with the age of full recruitment selection criteria (Figure 1). Regular regression (RG) uses all age-groups with nonzero catch. Regression 1 (R1) uses all age-groups up to but not including the first age at which catch is 0 or 1. Regression 1 prime (R1′) uses all age-groups with nonzero catches up to and including the age after which all catches are less than 2. Regression 1 double prime (R1″) uses all age-groups that have catches greater than 1. The RG and R1 methods were evaluated by Dunn et al. (2002); the R1′ and R1″ methods are new.

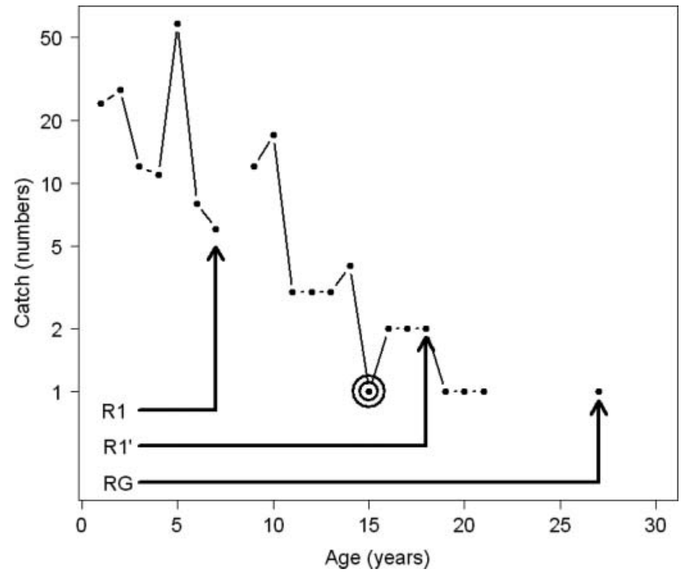

FIGURE 1. Plot of catch at age (on a log scale) versus age, showing truncation points resulting from application of the four proposed right truncation methods. The regular regression (RG) method uses all ages for which catch is nonzero (i.e., all indicated ages up to and including age 27). Regression 1 (R1) uses all ages up to but not including the first age at which catch is 0 or 1 (i.e., ages 1–7). Regression 1 prime (R1′) uses all ages with nonzero catches up to and including the age after which all catches are less than 2 (i.e., ages 1–18 except age 8 [zero catch]). Regression 1 double prime (R1″) uses all ages that have catches greater than 1 (i.e., ages 1–18 except age 8 [zero catch] and age 15 [shown by the bullseye]).

Regardless of right truncation method, the variance of the regression estimator is estimated by using the usual formula for the estimated variance of the slope for ordinary least-squares regression.

The mortality rate can also be estimated from the slope of a weighted linear regression, as suggested by Maceina and Bettoli (1998). An unweighted regression line was fitted with a candidate age of full recruitment and a right truncation method as previously described. The resulting model was used to predict log-transformed catch at age ($\log_e[\hat{N}_x]$), which was used as the weight in a subsequent fit using the same age of full recruitment and right truncation criteria. Frequently, $\log_e(\hat{N}_x)$ of older age-classes was negative; when this occurred, the weight of the age-group in question was set equal to 0. The remaining weights were scaled to sum to 1. We also explored the use of iteratively reweighted least squares, but this did not perform well and is not discussed further.

AGE OF FULL RECRUITMENT SELECTION CRITERIA

Four criteria for selecting the age of full recruitment (Peak, Peak Plus, χ^2 test, and z-test) were investigated in this study. The Peak method identifies the first fully recruited age as that with the greatest abundance. The Peak Plus method identifies the first fully recruited age as the age-group that is 1 year older than the Peak age. The χ^2 test is discussed in the Appendix. The z-test represents a statistical approach to identifying the first fully recruited age. The z-test method tests the null hypothesis that the youngest age-group's catch is consistent with what is expected given that the complete age-frequency data follow a geometric distribution. This is done by comparing survival estimates obtained using two unbiased estimators of S (Heincke's method and the Chapman–Robson estimator) and determining whether the difference between these two estimators can be attributed to sampling error alone. The z-test, which Chapman and Robson (1960) stated is valid as long as sample size exceeds 100, has test statistic

$$Z_{test} = \frac{\left(\frac{N-N_0}{N} - \frac{T}{N+T-1}\right)}{\sqrt{\frac{T(T-1)(N-1)}{N(N+T-1)^2(N+T-2)}}}, \tag{6}$$

where $T = \sum_{i=1}^{N} T_i$ and the T_i are the recoded ages of the sample such that the age of the youngest age-group is set equal to 0, the age of the next older age-group is set equal to 1, and so on. Under the null hypothesis that the first age-group is fully recruited, the random variable Z_{test} has a mean of 0 and a variance of 1 and is asymptotically distributed as a standard normal. A one-tailed z-test can be conducted to determine whether the observed deficit in the youngest age-group is significant at some predetermined type I error rate α. If the youngest age-group is rejected, the remaining age-groups are recoded as described above and the test is repeated. This process continues until the test fails to reject an age-group.

SIMULATION METHODS

We explored the performance of the mortality rate estimators and the age of full recruitment selection criteria over a range of true Z-values, sample sizes, selectivity effects, and total errors. For the z-test, we used α values ranging from 0.01 to 0.10. For the first set of simulations, we set the natural mortality rate (M) equal to 0.2 year^{-1} and the full fishing mortality rate (F) at values ranging from 0 to 0.8 year^{-1} in increments of 0.1 such that Z ranged from 0.2 to 1.0 year^{-1}. This range encompasses most scenarios in which catch-curve analysis would be used. Few commercially exploited populations have Z below 0.2 year^{-1} (Pauly 1980; Hoenig 1983), and age-frequency data are sparse at mortality rates greater than 1.0 year^{-1}, making the utility of catch-curve analysis questionable outside the simulated range. We used two partial recruitment (PR) vectors to modify the fishing mortality at age. The steep PR vector multiplied F by 0.5 for the first age-group, and all other ages had full recruitment to the fishery. The gradual PR vector multipliers for the first three age-groups were 0.50, 0.85, and 0.90, respectively, and all other age-groups had full recruitment to the fishery. Thorson and Prager (2011) found that modeling M as a function of age did little to improve estimation accuracy in the presence of additional and more substantial sources of nonsampling error. As a result, no attempt to model M was made in this study. However, in a second set of simulations, we specified an M-value equal to 0.5 year^{-1} to see if the conclusions held across a wider range of conditions.

Nonsampling errors encompass all sources of stochastic variation that influence a population age-frequency distribution. Here, we assumed that nonsampling variability encompasses deviations in recruitment, deviations in annual total mortality rate, and aging errors; we represented the error (ε) in frequency at age as lognormally distributed ($\varepsilon \sim \text{LN}[0, \sigma^2]$), uncorrelated, and without trend, where σ was fixed at either 0.3 or 0.6 to allow for different magnitudes of error.

Total mortality estimators, their associated variance estimators, and the age of full recruitment selection criteria were evaluated using simulated catch-at-age data. Five-thousand unique population age-frequency distributions that included nonsampling error as described above were generated for each simulation scenario. The true age proportions (FR) for fish at age i in simulation j was generated as

$$FR_{i,j} = \begin{cases} \dfrac{\varepsilon_1}{\varepsilon_1 + \sum_{k=2}^{30} \psi_k \varepsilon_k}, & i = 1,\ j = 1, \ldots, 5{,}000 \\[2ex] \dfrac{\varepsilon_i \psi_i}{\varepsilon_1 + \sum_{k=2}^{30} \psi_k \varepsilon_k}, & 2 \le i \le 30,\ j = 1, \ldots, 5{,}000 \end{cases}, \tag{7}$$

where ψ_i was the cumulative survival rate up to age i, given by

$$\psi_i = \exp\left[-\sum_{k=1}^{i-1} (PR_k F - M) \right], \tag{8}$$

and PR_k was the partial recruitment value for age k. The observed catch of simulated population j was then obtained by drawing a random sample of size N from a multinomial distribution with probability of selecting age i set equal to $FR_{i,j}SEL_i$, where SEL_i is the sampling selectivity for age i. In this paper, SEL_i equals PR_i for all ages i.

By chance, samples could be generated that were not suitable for analysis with certain combinations of mortality estimator, age of full recruitment criterion, and in the case of the regression estimators, right truncation method. The combination of these elements that resulted in the most severe age truncation (both left and right) was the R1 regression estimator with the Peak Plus age of full recruitment criterion. Use of this combination on a simulated sample can result in only two age-groups being available, which precludes estimation of the slope with its variance. Samples were rejected when they contained fewer than three age-groups when the R1 truncation and Peak Plus selection criteria were used. When a rejection occurred, a new population frequency and sample were generated as described above, and the new sample was tested. Only samples that could be analyzed by all combinations of estimator, age of full recruitment selection criterion, and right truncation method were used.

Performance of the Z estimators was measured with percent bias (%BIAS) and percent root mean square error (%RMSE). For each estimator, we calculated %BIAS of $\hat{Z}$ as

$$\%BIAS(\hat{Z}) = 100[E(\hat{Z}) - Z]/Z, \tag{9}$$

and %RMSE as

$$\%RMSE(\hat{Z}) = \frac{100\sqrt{E(\hat{Z} - Z)^2}}{Z}, \tag{10}$$

where $E()$ denotes expectation, which is approximated by averaging over simulation results. Additionally, the %BIAS and %RMSE of the $S\hat{E}(\hat{Z})$ were calculated for each set of estimators as

$$\%BIAS[S\hat{E}(\hat{Z})] = 100\{E[S\hat{E}(\hat{Z})] - SE(\hat{Z})\}/SE(\hat{Z}), \tag{11}$$

and

$$\%RMSE[S\hat{E}(\hat{Z})] = \frac{100\sqrt{E\left[S\hat{E}(\hat{Z}) - SE(\hat{Z})\right]^2}}{SE(\hat{Z})}, \tag{12}$$

where $S\hat{E}(\hat{Z})$ is the estimated standard error of the estimated Z; and $SE(\hat{Z})$ is the true SE, which is calculated as the SD of the 5,000 estimates of Z obtained through simulation.

As seen in equations (1) and (2), the $CR(\hat{Z})$ and $V\hat{A}R[CR(\hat{Z})]$ estimators are functionally dependent. This can lead to overdispersion, which exists when the variability in the data exceeds the nominal amount expected by the model. The statistical literature has long recognized this issue (Wedderburn 1974), and commonly applied procedures that address overdispersion have been developed (see Burnham and Anderson 2002). A variance inflation factor ($\hat{c}$) was calculated as the usual chi-square goodness-of-fit statistic divided by the corresponding df, where the number of df was set equal to the number of age-groups used in a given analysis minus 1. Estimated SEs of the Chapman–Robson Z estimator were multiplied by $\sqrt{\hat{c}}$.

RESULTS

Patterns of performance were generally highly consistent across scenarios of sample size, nonsampling error, and selectivity/PR patterns. Consequently, we show only representative results where patterns were consistent. Because the choice of significance level in the z-test resulted in only minor differences in performance of the estimators, we only show results for an α of 0.05. Results for all scenarios are presented by Smith and Hoenig (2012).

Heincke's Estimator

The %RMSE of Heincke's method exceeded those for all other methods tested and was often in excess of 100% for low values of Z (Figure 2). Under ideal conditions (i.e., large sample size and low error), Heincke's estimator never achieved a %RMSE below 30% and it exceeded the next-worst estimator's %RMSE by at least 10% of Z (Figure 2). Heincke's estimator can have a large bias (>20% in absolute value) for any combination of nonsampling error, sample size, and selectivity/PR pattern examined (Figure 2). The SE estimator for Heincke's method performed poorly in terms of %BIAS and %RMSE (Figure 2). No further results are presented for Heincke's estimator because of its poor performance.

Regression Estimators

Weighted regressions.—The RG right truncation criterion combined with the Peak or z-test age of full recruitment criterion were generally the best-performing weighted regression estimators in terms of minimum %RMSE and %BIAS (Figure 3A, C). The RG Peak method was preferred in the majority of cases over the RG z-test based on %RMSE and, when not preferred, was almost as good as the RG z-test. Root mean square error increased with Z (Figure 3A). The %BIAS of the RG Peak weighted regression ranged from −17% to 0% depending on the tested combination of selectivity curve, Z-value, sample size, and error. Best results were obtained when selectivity was steep, Z-value was intermediate (0.5–0.9), sample size was large (N = 600), and error was small (σ = 0.3). The %BIAS of RG Peak was relatively insensitive to changes in sample size and error but increased substantially at high Z-values when the selectivity/PR curve was gradual (Smith and Hoenig 2012).

Weighted versus unweighted regression.—Weighted regression with the RG right truncation criterion had a lower %RMSE than all other unweighted regressions except occasionally at high values of Z (Figure 3B; see also Smith and Hoenig 2012). For all regression methods, the RMSE increased with Z.

In nearly all of the cases simulated, the best two weighted regressions had lower %BIAS than the unweighted regressions (Figure 3D). Only the R1 unweighted regressions had less bias than the two best weighted regressions, but the differences in %BIAS were negligible (Figure 3D).

Chapman–Robson Estimator

Under most of the conditions simulated, the Chapman–Robson estimator with the Peak Plus criterion had the lowest %RMSE in comparison with the other Chapman–Robson estimators (Figure 4).

For the Chapman–Robson estimator with the Peak Plus criterion, %BIAS was typically negative and ranged from –2.6% to

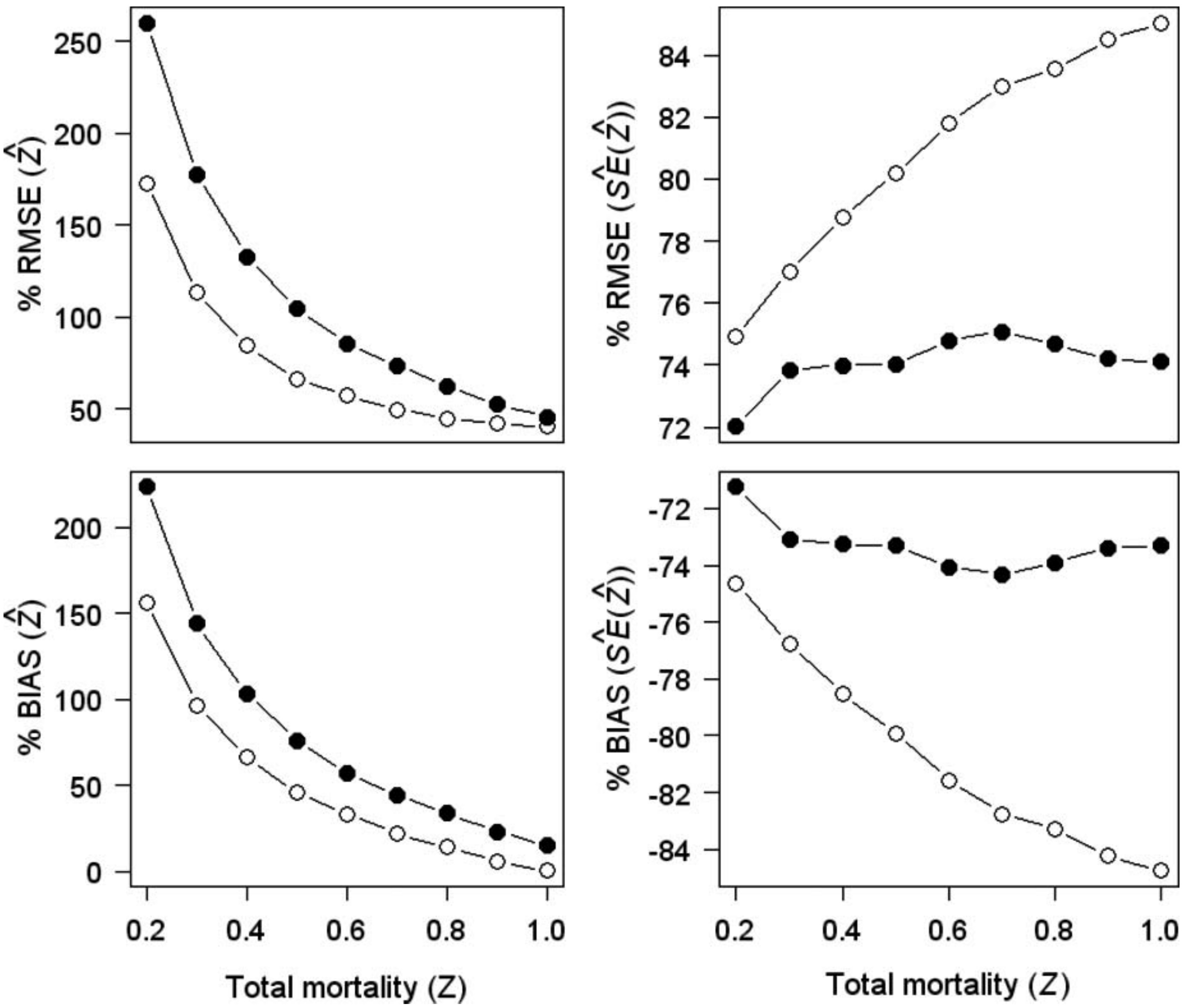

FIGURE 2. Percent bias (%BIAS) and percent root mean square error (%RMSE) of Heincke's total mortality rate (Z) estimator (left panels) and the SE estimator (right panels). Results plotted with black-shaded circles represent performance when instantaneous natural mortality rate (M) is 0.2 year^{-1}, sample size (N) is 200, total error (σ) is 0.6, and selectivity/partial recruitment is steep. Results plotted with open circles depict performance when M is 0.2 year^{-1}, N is 600, σ is 0.3, and selectivity/partial recruitment is steep.

–14.7% when selectivity was steep, sample size was small, and error was large (Figure 5). When selectivity was steep, sample size was large, and error was small, %BIAS ranged from 0.21% to –3.9%. In all cases, %BIAS increased as Z increased. Increasing M had little effect on the bias of the Chapman–Robson estimator; however, bias increased by up to 200% when the selectivity curve was gradual. The %BIAS obtained under gradual selectivity/PR never exceeded −22% and was substantially lower when sample size was large and Z was low (Smith and Hoenig 2012).

Comparison of the Weighted Regression and Chapman–Robson Estimators

In terms of %RMSE, the weighted regression with the Peak criterion generally outperformed the Chapman–Robson estimator with the Peak Plus criterion (Figure 6). The Chapman–Robson estimator's performance was best when sample size was small ($N = 200$); however, it was not always uniformly better under this condition (Figure 6). When sample size was large ($N = 600$) or when error was large ($\sigma = 0.6$), the weighted regression method generally outperformed the Chapman–Robson estimator.

In terms of bias, neither method was clearly superior. The %BIAS was always negative but rarely exceeded 15% of Z except occasionally at high values of Z (Figure 5).

Performance of Variance Estimators

The uncorrected Chapman–Robson variance estimator performed poorly because it underestimated the variance and thus had a large %RMSE. Once corrected for overdispersion, the Chapman–Robson Peak Plus SE estimator performed as well as or better than all other estimators evaluated, as measured by minimum RMSE (Figure 7). The corrected Chapman–Robson SE estimator was slightly positively biased except when nonsampling error (σ) was low and Z was high (Figure 7). Performance of the corrected Chapman–Robson Peak Plus SE

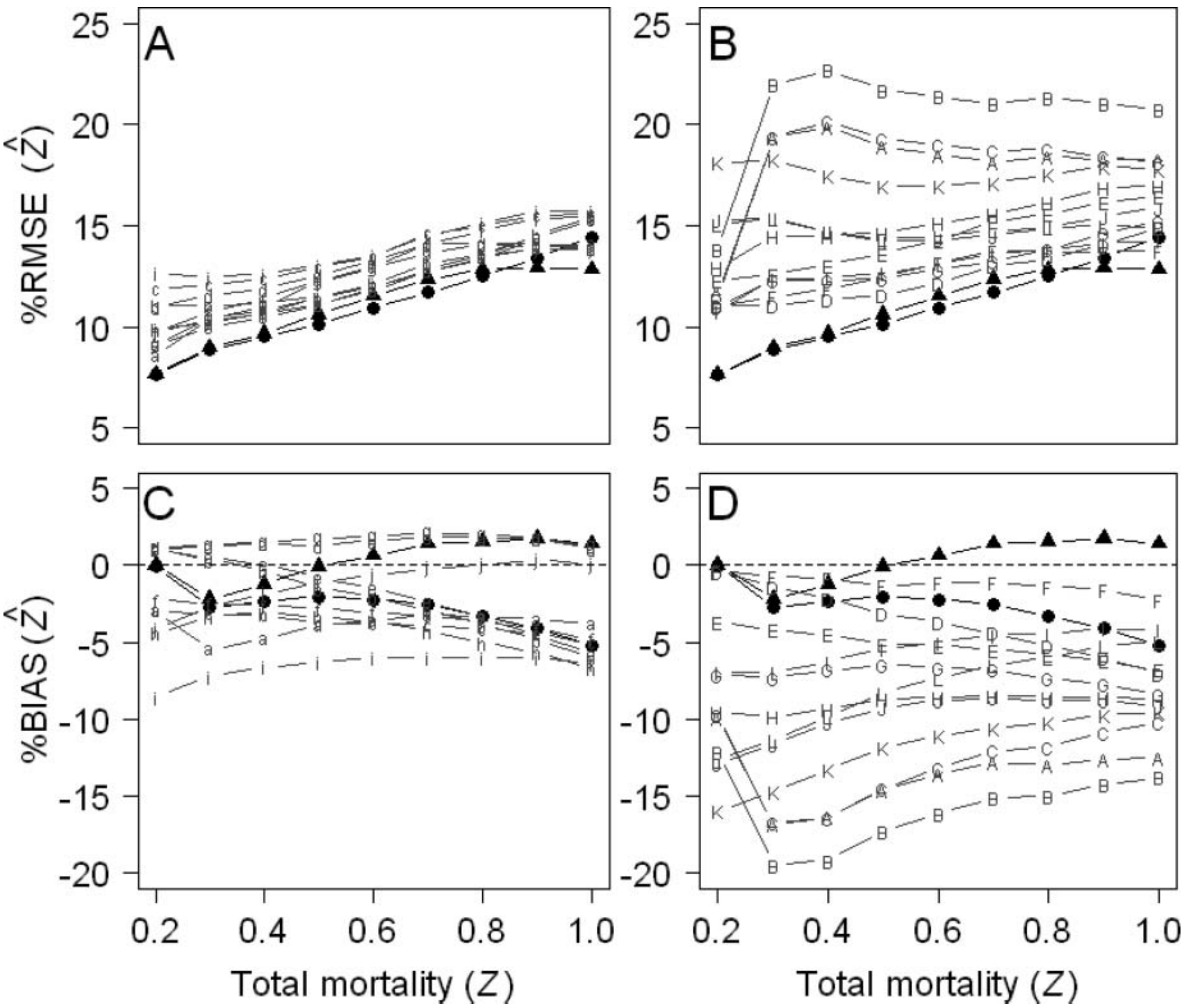

FIGURE 3. Percent bias (%BIAS; lower panels) and percent root mean square error (%RMSE; upper panels) of various total mortality rate (Z) estimators as a function of Z. Panels **(A)** and **(C)** show the preferred weighted regression methods (black circles = RG Peak; black triangles = RG z-test; see the main text for definition of terms) in comparison with all other simulated weighted regressions (a = RG Peak Plus; b = R1 Peak; c = R1 Peak Plus; d = R1 z-test; e = R1′ Peak; f = R1′ Peak Plus; g = R1′ z-test; h = R1′′ Peak; i = R1′′ Peak Plus; j = R1′′ z-test). Simulation conditions were an instantaneous natural mortality rate (M) equal to 0.2 year^{-1}, steep selectivity/partial recruitment curves, sample size (N) equal to 600, and total error (σ) equal to 0.3. Panels **(B)** and **(D)** show the preferred weighted regressions (as indicated above) in comparison with all simulated unweighted regressions (A = RG Peak; B = RG Peak Plus; C = RG z-test; D = R1 Peak; E = R1 Peak Plus; F = R1 z-test; G = R1′ Peak; H = R1′ Peak Plus; I = R1′ z-test; J = R1′′ Peak; K = R1′′ Peak Plus; L = R1′′ z-test). Simulation conditions were the same as in panels A and C.

estimator was insensitive to changes in simulated sample size within a given level of nonsampling error (σ; Smith and Hoenig 2012).

For the weighted regression estimators, use of the RG Peak criterion generally provided the best performance (Figure 7). The weighted regression SE estimators were almost always negatively biased except occasionally at very low Z (Figure 7).

DISCUSSION

Heincke's Method

Heincke's method performs poorly in the presence of nonsampling error and is not recommended when age-structured data are available. This poor performance is tied to the fact that Heincke's method is extremely sensitive to recruitment variability. When the youngest fully recruited age-group is highly abundant, Heincke's estimator will estimate a low S. Likewise, when the youngest fully recruited age-group is less abundant, Heincke's estimator will estimate a high S. As a result, estimates of S obtained using Heincke's estimator can be misleading. This does not mean that Heincke's method is without merit; it remains a valuable tool in data-poor situations due to its ability to produce Z estimates when only a few age-groups can be distinguished. However, Jensen's (1996) suggestion that Heincke's approach be widely adopted is not tenable.

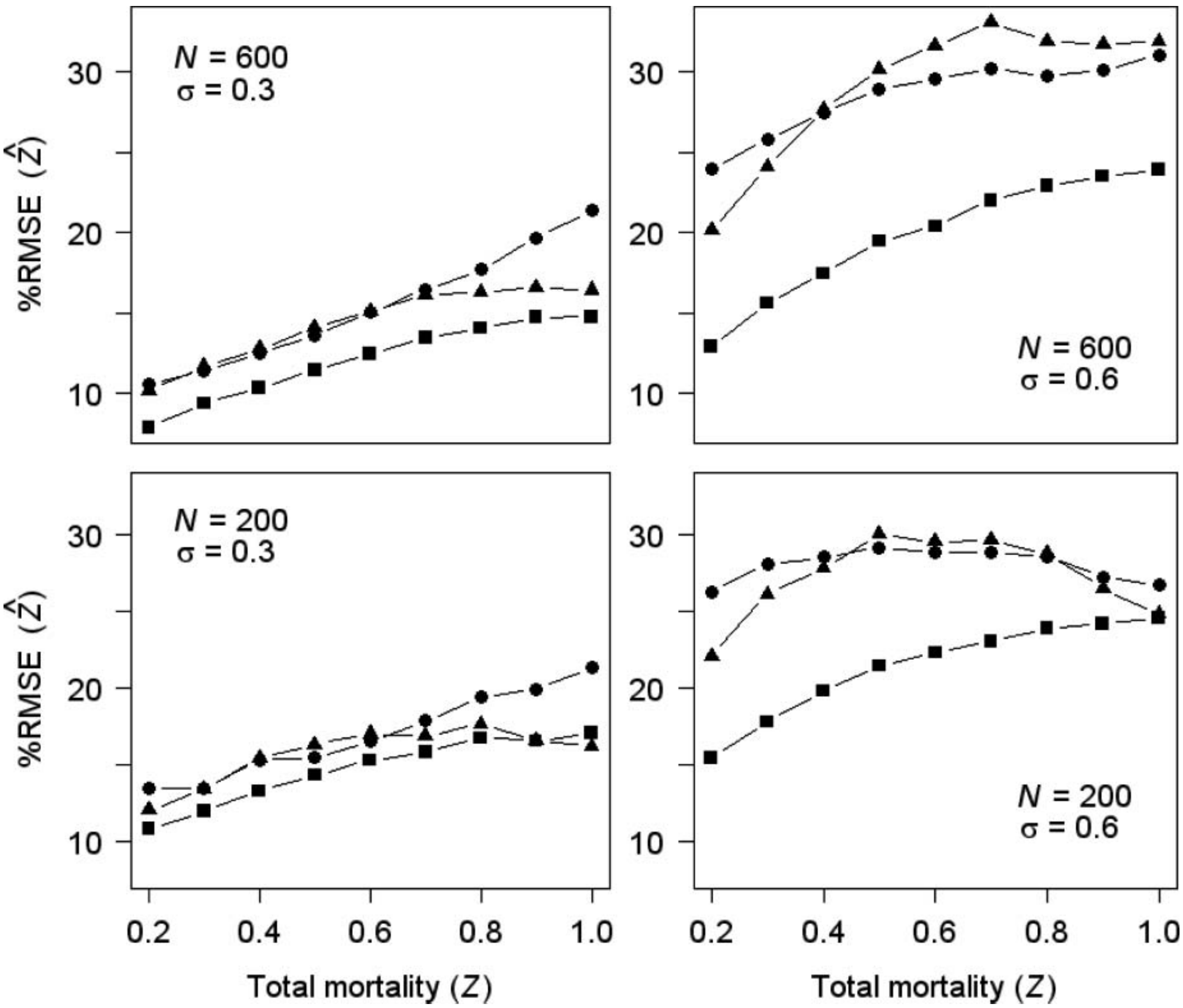

FIGURE 4. Percent root mean square error (%RMSE) of the Chapman–Robson total mortality rate (Z) estimator with three methods of selecting the first fully recruited age: Peak (black circles), Peak Plus (black squares), and z-test (black triangles; see the main text for definition of terms). Conditions were an instantaneous natural mortality rate (M) of 0.2 year^{-1} and a steep selectivity/partial recruitment pattern.

The Best Estimators

We evaluated many estimators of total mortality rate. Two stand out as superior to the others: the Chapman–Robson estimator with the Peak Plus criterion, and the weighted regression with the Peak criterion and no right truncation. Unfortunately, neither estimator is uniformly superior. At a low sample size ($N = 200$ in our simulations), the Chapman–Robson estimator generally has lower mean square error than the weighted regression estimator, whereas at a large sample size ($N = 600$) the opposite is observed (Figure 6). Overall, the weighted regression estimator outperformed the Chapman–Robson estimator in a majority of the scenarios but usually not by much (Figure 6). Hence, the two approaches are competitive and either approach could be used. We recommend the Chapman–Robson approach with the variance estimate corrected for overdispersion because this approach is based on well-established quasilikelihood theory as opposed to the ad hoc nature of the regression weighting. There is no guarantee that the regression results would hold up under different scenarios of error structure.

The Chapman–Robson Peak Plus estimator also proved to be relatively insensitive to changes in sample size for a given level of error (Figure 4). This has great logistical value in that the regression methods (unweighted or weighted) require a 600-fish sample to achieve a relative performance similar to that of the Chapman–Robson Peak Plus estimator with a sample size of 200. When the main source of uncertainty is sampling error, the mean square error will decrease as the sample size increases. However, in catch-curve analysis, nonsampling error associated with recruitment variability is present and will not go away as the sample size increases. Apparently, the variance of the Chapman–Robson estimator approaches an asymptotic minimum more quickly than that of the weighted regression estimator; hence, for the sample sizes simulated,

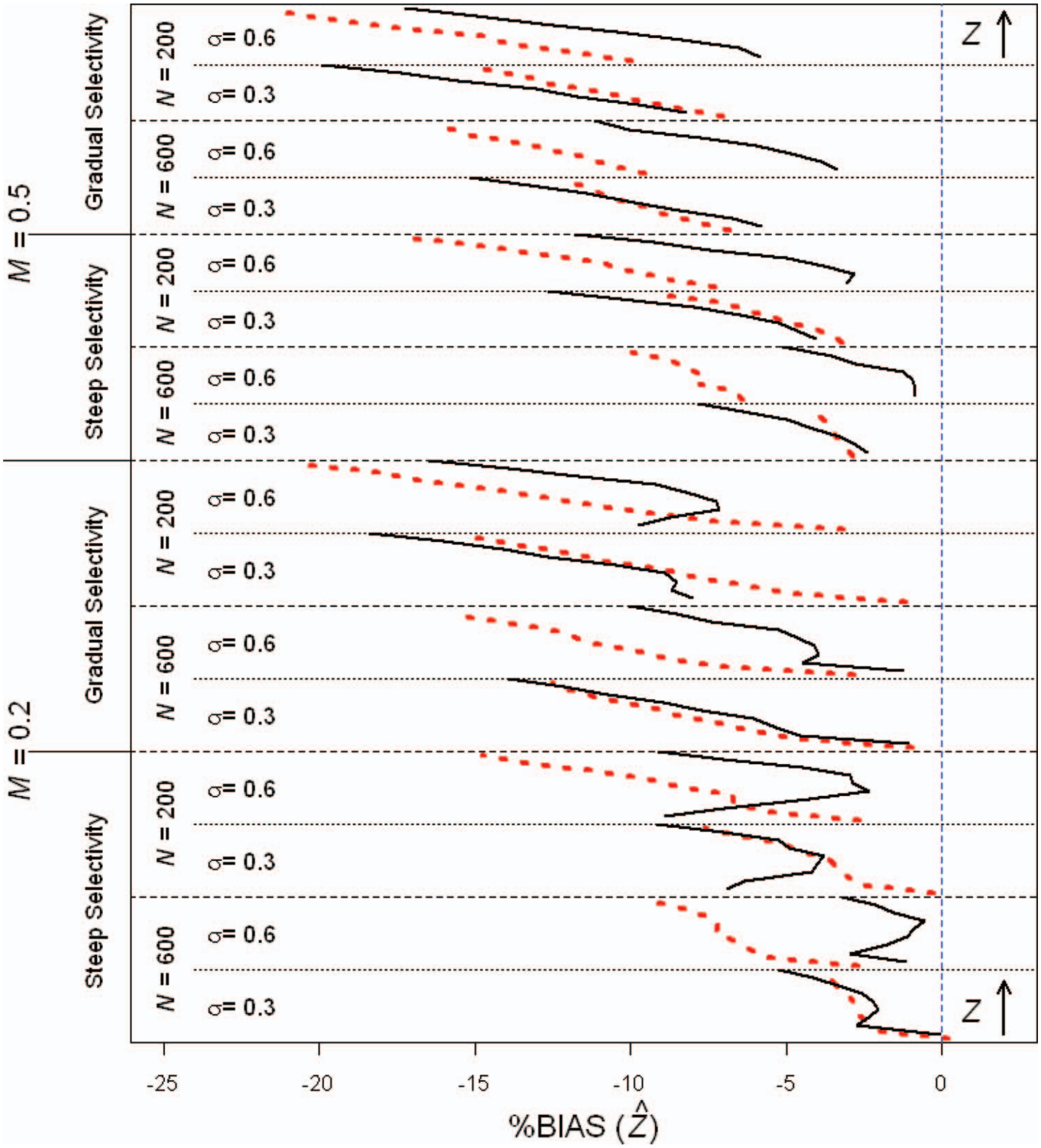

FIGURE 5. Percent bias (%BIAS) of the Chapman–Robson total mortality rate (*Z*) estimator with the Peak Plus age of full recruitment (dotted line) and the weighted regression estimator with the Peak age of full recruitment (solid line) as a function of natural mortality rate (*M*), selectivity/partial recruitment pattern, sample size (*N*), nonsampling error (σ), and *Z*. In each strip, *Z* increases from bottom to top. Results of all 128 scenarios are shown. [Figure available in color online.]

the Chapman–Robson estimator seems to be less sensitive to sample size than the regression estimator.

Unweighted regression methods are widely used for the analysis of catch-curve data. This study confirms the findings of Dunn et al. (2002) that unweighted regression methods are typically negatively biased and that right truncation of the data set significantly reduces bias. However, since the variance of the slope estimate depends on the range of the independent variable (age), discarding data increases the variance and thus the mean square error. This study demonstrates that weighted regression using no right truncation can achieve reductions in bias that are equal to or greater than those achieved by unweighted regression using any of the four right truncation methods tested. These bias reductions do not come at the cost of precision due to the fact that older age-classes are only occasionally discarded when a weight is set to 0. Weighted regression

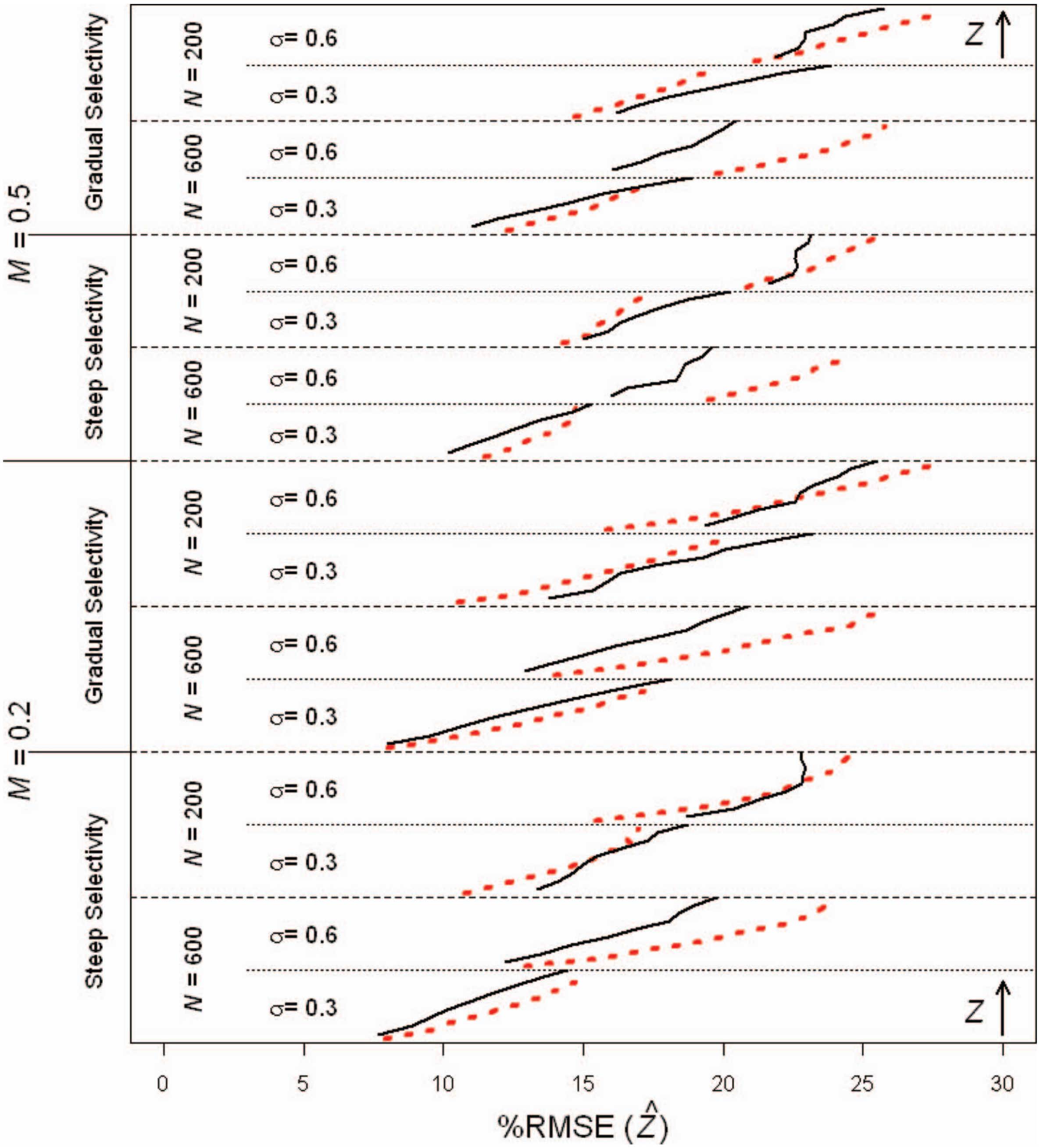

FIGURE 6. Percent root mean square error (%RMSE) of the Chapman–Robson total mortality rate (*Z*) estimator with the Peak Plus age of full recruitment (dotted line) and the weighted regression estimator with the Peak age of full recruitment (solid line) as a function of natural mortality rate (*M*), selectivity/partial recruitment pattern, sample size (*N*), nonsampling error (σ), and *Z*. In each strip, *Z* increases from bottom to top. Results of all 128 scenarios are shown. [Figure available in color online.]

routinely achieves lower mean square error in comparison with unweighted regression. Consequently, the continued use of unweighted regression cannot be justified at present.

Recommendations

1. For catch-curve analysis, we recommend the use of either the Chapman–Robson mortality estimator with the first age-group used being 1 year older than the age of peak abundance (i.e., Peak Plus criterion) or the weighted regression estimator with the first age-group used being the age of peak abundance (i.e., Peak criterion). We prefer the Chapman–Robson method because it performed well in our simulations and in those of Dunn et al. (2002). We acknowledge that the weighted regression estimator generally outperformed the Chapman–Robson estimator in our simulations. However, differences between the methods were slight, at least under

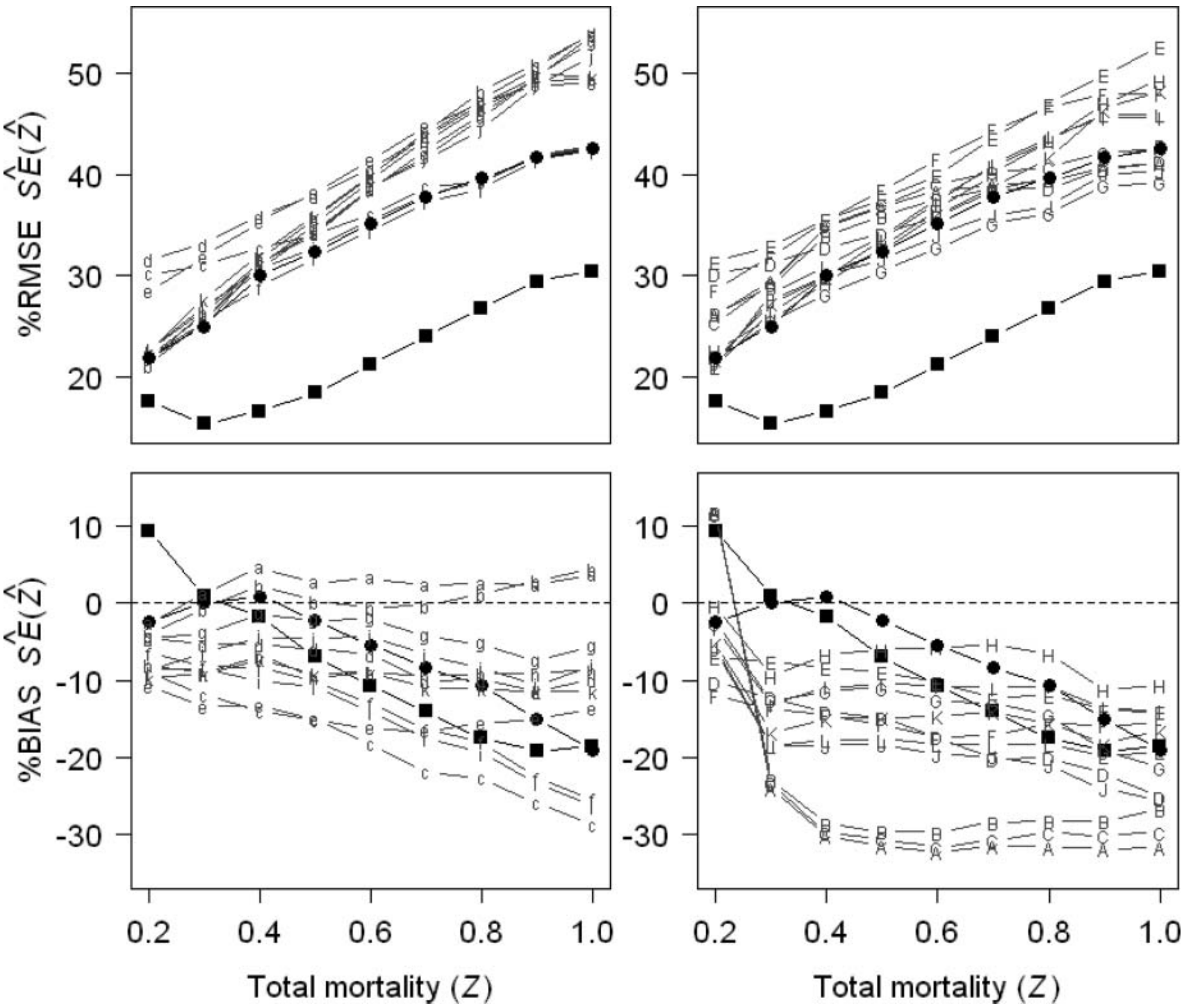

FIGURE 7. Performance of the Chapman–Robson Peak Plus variance estimator with correction for overdispersion (black squares) and the RG Peak weighted regression variance estimator (black circles) in terms of percent root mean square error (%RMSE; upper panels) and percent bias (%BIAS; lower panels). Estimators are plotted with all other weighted (left panels) and unweighted (right panels) regression estimators. Plotting symbols (a–j and A–L) are defined in Figure 3. Simulation conditions were an instantaneous natural mortality rate (M) of 0.2 year^{-1}, steep selectivity/partial recruitment curves, a sample size (N) of 600, and total error (σ) of 0.3.

the conditions we simulated, and the weighting procedure was purely ad hoc.

2. The Chapman–Robson variance estimator should be corrected for overdispersion to avoid a large negative bias and to reduce the mean square error. The corrected estimator performed as well as or better than all of the other variance estimators tested.
3. Unweighted linear regression should not be used for catch-curve analysis because it is inferior to other methods.
4. Heincke's estimator performs poorly and should be avoided. However, there are situations in which only one or two age-groups can be distinguished, and Heincke's method may be useful in those situations.

ACKNOWLEDGMENTS

We thank Romuald Lipcius and Norman Hall for helpful comments and encouragement during the development and drafting of the manuscript. This is Virginia Institute of Marine Science Contribution Number 3248.

REFERENCES

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd edition. Springer-Verlag, New York.

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

Dunn, A., R. I. C. C. Francis, and I. J. Doonan. 2002. Comparison of the Chapman–Robson and regression estimators of Z from catch-curve data when non-sampling stochastic error is present. Fisheries Research 59:149–159.

Heincke, F. 1913. Investigations on the plaice—general report: 1. plaice fishery and protective measures, preliminary brief summary of the most important points of the report. Rapports et Procès-Verbaux des Réunions, Conseil International pour l'Exploration de la Mer 16.

Hoenig, J. M. 1983. Empirical use of longevity data to estimate mortality rates. U.S. National Marine Fisheries Service Fishery Bulletin 81:898–903.

Hoenig, J. M., W. D. Lawing, and N. A. Hoenig. 1983. Using mean age, mean length and median length data to estimate the total mortality rate. International Council for the Exploration of the Sea, CM 1983/D:23, Copenhagen.

Jensen, A. L. 1984. Non-linear catch curves resulting from variation in mortality among subpopulations. ICES Journal of Marine Science 41:121–124.

Jensen, A. L. 1996. Ratio estimation of mortality using catch curves. Fisheries Research 27:61–67.

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

Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. ICES Journal of Marine Science 39:175–192.

Pauly, D. 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. International Center for Living Aquatic Resources Management, ICLARM Studies and Reviews 8, Manila.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Fisheries Research Board of Canada Bulletin 191.

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

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. Macmillan, New York.

Smith, M. W., and J. M. Hoenig. 2012. Simulated performance of catch curve methods for estimating total mortality rate. Virginia Institute of Marine Science, Data Report 60, Gloucester Point, Virginia. Available: www.vims.edu/GreyLit/VIMS/dr060.pdf. (August 2012).

Thorson, J. T., and M. H. Prager. 2011. Better catch curves: incorporating age-specific natural mortality and logistic selectivity. Transactions of the American Fisheries Society 140:356–366.

Tuckey, T., N. Yochum, J. Hoenig, J. Lucy, and J. Cimino. 2007. Evaluating localized vs. large-scale management: the example of tautog in Virginia. Fisheries 32:22–28.

Weber, M. J., M. J. Hennen, and M. L. Brown. 2011. Simulated population responses of common carp to commercial exploitation. North American Journal of Fisheries Management 31:269–279.

Wedderburn, R. W. M. 1974. Quasi-likelihood functions, generalized linear models, and the Gauss–Newton method. Biometrika 61:439–447.

APPENDIX: CHI-SQUARE TEST FOR INCOMPLETE RECRUITMENT

Chapman and Robson (1960) proposed a one-tailed z-test for assessing whether the first age-group is “deficient” (their term)—that is, incompletely recruited. Later, Robson and Chapman (1961) noted that the square of a standard normal variate (Z) is distributed as a χ^2 variable with 1 df. They illustrated the use of this testing procedure by incrementally testing candidate ages, beginning with the youngest, until the test failed to reject an age as being deficient in numbers. The z-test and χ^2 test were suggested to be equivalent, but this point needs clarification.

As was noted by Seber (1982), a χ^2 test is normally a two-tailed test, and Seber suggested that this is the case for Robson and Chapman’s (1961) test. This would not be a reasonable approach since an age-group could be rejected as incompletely recruited because it appears to be too abundant. Indeed, in our simulations, application of the two-tailed χ^2 test sometimes results in all ages being rejected until there are only two ages left. However, Robson and Chapman (1961) appear to have presented a one-tailed test with a two-stage procedure for testing the null hypothesis (H_o = the first age-group is fully recruited) against the alternative hypothesis (H_a = the first age-group is deficient). The first stage of their test is to determine whether the Heincke estimate is lower or higher than the Chapman–Robson estimate. If the Heincke estimate is lower, there is no evidence that the first age-group is deficient; if the Heincke estimate is higher, proceed to stage 2. Stage 2 is to conduct the χ^2 test at some level of α. With this procedure, the H_o will be rejected with probability $\alpha/2$ when the H_o is true. This follows from the fact that the H_o is rejected only if the Heincke estimate is higher than the Chapman–Robson estimate (50% chance under the H_o) followed by the χ^2 test giving a significant result ($[100 \times \alpha]$% chance under the H_o). Thus, the chance of rejecting the H_o when the H_o is true is 0.5α. To achieve an overall α level of 0.05, for example, one would need to conduct the χ^2 test with α equal to 0.10. If this is done, then the z-test with an α equal to 0.05 will be equivalent to the two-stage procedure for the χ^2 test with 2 times the nominal α. Of course, the choice of α is arbitrary in the absence of consideration of the costs of type I and type II errors. Furthermore, when multiple tests are done, the relationship between the nominal (individual) α and the overall α is complex.