2017

Extensions and Applications of Mean Length Mortality Estimators for Assessment of Data-Limited Fisheries

Quang C. Huynh
College of William and Mary - Virginia Institute of Marine Science, qhuynh@vims.edu

Extensions and applications of mean length mortality estimators for assessment of data-limited fisheries

A Dissertation

Presented to

The Faculty of the School of Marine Science

The College of William and Mary in Virginia

In Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

by

Quang C. Huynh

January 2018

APPROVAL PAGE

This dissertation is submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

Quang C. Huynh

Approved by the Committee, December, 2017

John M. Hoenig, Ph.D.
Committee Chair/Advisor

Mark J. Brush, Ph.D.

John E. Graves, Ph.D.

Ross J. Iaci, Ph.D.
Department of Mathematics

John F. Walter III, Ph.D.
Southeast Fisheries Science Center
Miami, Florida, USA

Table of Contents

Acknowledgements

I would like to thank first and foremost my advisor, Dr. John Hoenig, for his guidance and mentorship over the past five years. You taught me to be methodical and thorough in my work, and you pushed me to expand my boundaries. I know this experience will be invaluable down the line. Thanks to my committee, Drs. Mark Brush, John Graves, Ross Iaci, and John Walter, for their support first as instructors, whether formal or informal, as I was beginning my studies and subsequently as committee members for help and feedback on my research and writing. A special thanks is also needed for Todd Gedamke, whose dissertation helped me start mine and impromptu phone calls helped me see the big picture and stay on track.

This dissertation would not be possible without the funding support from NOAA and Virginia Sea Grant through Cooperative Agreements NA14OAR4170297 and NA15OAR4170184 (Population and Ecosystem Dynamics Fellowship), as well as the VIMS Office of Academic Studies. I am particularly appreciative of Academics Studies for the funding to attend training workshops and courses, which were invaluable for the quantitative skills training and networking opportunities. Many thanks to Linda, Jen, and Cathy at Academic Studies for helping me navigate through the administrative side of things.

This dissertation would also not be possible without the work of the anglers, port samplers, data managers who collect and maintain the data that were used in my analyses.

I would also like to acknowledge the many scientists and collaborators at the Southeast Fisheries Science Center (SEFSC), Pacific Islands Fisheries Science Center (PIFSC), and the International Council for the Exploration of the Sea (ICES) with whom I have had the privilege to work.

Special thanks to the past and present members of the Hoenig lab for the adventures, commiseration, support, and ideas: Matt Smith, Amy Then, Lisa Ailloud, Kristen Omori, Liese Carleton, Lydia Goins, Dan Gonzales, and Phil Sadler.

Thanks to all involved in the VIMS community, especially fellow students and colleagues, who made this journey a fun one. To my climbing partners who took me to new heights. To the Sail and Paddle Club and the crew of Kingfisher for the sailing fun. Finally, to my friends and family for their love and support from afar!

List of Tables

List of Figures

Abstract

For data-limited fisheries, length-based mortality estimators are attractive as alternatives to age-structured models due to the simpler data requirements and ease of use of the former. This dissertation develops new extensions of mean length-based mortality estimators and applies them to federally-managed stocks in the southeastern U.S. and U.S. Caribbean.

Chapter 1 presents a review of length-based methods from the literature. Common themes regarding the methodology, assumptions, and diagnostics in these length-based methods are discussed. In Chapter 2, a simulation study evaluates the performance of the length-converted catch curve (LCCC), Beverton-Holt equation (BHE), and Length Based-Spawner Potential Ratio (LB-SPR) over a range of scenarios. Although the LCCC and BHE are older methods than LB-SPR, the former outperformed LB-SPR in many scenarios in the simulation. Overall, it was found that the three length-based mortality estimators are less likely to perform well for low *M/K* stocks (*M/K* is the ratio of the natural mortality rate and the von Bertalanffy growth parameter; this ratio describes different life history strategies of exploited fish and invertebrate populations), while various decision rules for truncating the length data for the LCCC and BHE were less influential.

In Chapter 3, a multi-stock model is developed for the non-equilibrium mean length-based mortality estimator and then applied to the deepwater snapper complex in Puerto Rico. The multispecies estimator evaluates synchrony in changes to the mean length of multiple species in a complex. Synchrony in mortality can reduce the number of estimated parameters and borrows information from more informative species to lesser sampled species in the model. In Chapter 4, a new method is developed to estimate mortality from both mean lengths and catch rates (MLCR), which is an extension of the mean length-only (ML) model. To do so, the corresponding behavior for the catch rate following step-wise changes in mortality is derived. Application of both models to Puerto Rico mutton snapper shows that the MLCR model can provide more information to support a more complex mortality history with the two data types compared to the ML model.

In Chapter 5, a suite of mean length-based mortality estimators is applied to six stocks (four in the Gulf of Mexico and two in the U.S. Atlantic) recently assessed with age-structured models. There was general agreement in historical mortality trends between the age-structured models and the mean length-based methods, although there were some discrepancies which are discussed. All models also agreed on the overfishing status in the terminal year of the assessment of the six stocks considered here when the mortality rates were compared relative to reference points.

This dissertation develops new length-based assessment methods which consider multiple sources of data. The review guides prospective users on potential choices for assessment with length-based methods. Issues and diagnostics associated with the methods are also discussed in the review and highlighted in the example applications.

Author's Note

The chapters that comprise this dissertation were written in manuscript format for a scientific publication. Thus, the formatting of each chapter follows the guidelines of the publication to which the manuscript was or will be submitted. At the time of writing, citations for individual chapters are as follows:

Chapter 1:

Huynh, Q. C. et al.* In prep. Size-based mortality estimators and diagnostic procedures for assessment of data-limited fisheries. Intended for submission to Fish and Fisheries.

Chapter 2:

Huynh, Q. C., Beckensteiner, J., Carleton, L. M., Marcek, B. J., Nepal KC, V., Peterson, C. P., Wood, M. A., Hoenig, J. M. In review. Comparative performance of three length-based mortality estimators. Submitted to Marine and Coastal Fisheries.

Chapter 3:

Huynh, Q. C., Gedamke, T., Hoenig, J. M., Porch, C. 2017. Multispecies Extensions to a Nonequilbrium Length-based Mortality Estimator. Marine and Coastal Fisheries 9:68-78.

Chapter 4:

Huynh, Q. C., Gedamke, T., Porch, C. E., Hoenig, J. M., Walter, J. F., Bryan, M., Brodziak, J. 2017. Estimating Total Mortality Rates from Mean Lengths and Catch Rates in Nonequilibrium Situations. Transactions of the American Fisheries Society 146:803-815.

Chapter 5:

Huynh, Q. C., et al.* In prep. How well do length-based mortality estimators and age-structured models agree for stock status? A comparison with six southeastern United States stocks. Intended for submission to ICES Journal of Marine Science.

* Co-authors to be determined.

Extensions and applications of mean length mortality estimators for assessment of data-limited fisheries

Chapter 1: Sized-based mortality estimators and diagnostic procedures for assessment of data-limited fisheries

1.1. Abstract

We review the suite of methods which use size composition data for mortality estimation and stock status determination for data-limited fisheries. Methods that are currently available can be grouped into two categories: mean length-based methods, in which the mortality rate is estimated from the observed mean length of the catch, and composition-based methods, in which the mortality rate is estimated based on the shape of the size distribution. The simplest methods assume equilibrium conditions. Advances in methodology provide the opportunity to assess goodness of fit and allow for mortality estimation in nonequilibrium situations from multiple years of data and additional data types such as fishing effort and indices of abundance. We discuss six issues that are often encountered with size-based mortality estimators, including assumptions regarding growth, selectivity, recruitment, and mortality (with respect to time and age). We describe the effects of violations of these assumptions, summarize simulation studies which have evaluated the performance of these methods, and propose diagnostic procedures for identifying violations. From this discussion, we provide general guidelines and solutions to address violations in model assumptions.

1.2. Introduction

Since the mid-20th century, an array of size-based methods has been developed for the assessment of exploited marine resources. For most of the world's fisheries, data-limited is the norm rather than the exception (Newman, Berkson, & Suatoni, 2015). Such situations often arise when there are limitations in time or money for data collection or study (Bentley, 2015; Chrysafi & Kuparinen, 2016). With limited resources, priority is often given to assessments of higher-valued stocks. Data-limited, size-based methods are attractive in cases where other data types are not available for traditional stock assessment models, such as statistical catch-at-age models. These methods generally estimate mortality rates for a stock from size composition and life history parameters. Survival is informed by the relative abundance of small and large animals in the catch despite the diversity of modeling approaches. Thus, the depletion of large animals implies a population that experiences a high mortality rate because few animals survive to large sizes.

There are two general families of size-based methods: mean-length based and composition-based methods. In the former, information from the length composition is summarized by calculating the mean length using lengths larger than a specified size of full vulnerability L_c (Beverton & Holt, 1956; Gedamke & Hoenig, 2006). From a single calculation of the mean length, the mortality rate is estimated from a moment estimator under equilibrium assumptions. To allow for the equilibrium assumption to be relaxed, a likelihood framework can be used to estimate a series of historical mortality rates from multiple years of mean lengths. Recent composition-based methods use the shape of the size composition to estimate the mortality rate and selectivity parameters which produce

the best fit to the data in a likelihood-based framework (Hordyk, Ono, Prince, & Walters, 2016; Hordyk, Ono, Valencia, Loneragan, & Prince, 2015; Kokkalis, Thygesen, Nielsen, & Andersen, 2015; Rudd & Thorson, in press). An older composition-based method, the length-converted catch curve, estimates mortality from a linear regression fitted to a subset of the length composition (Pauly, 1983). Both equilibrium and nonequilibrium size composition methods have been developed.

Historically, length-based methods, such as ELEFAN (Pauly, 1987; Taylor & Mildenberger, 2017) and MULTIFAN (Fournier, *et al*., 1990), were developed as estimators of growth parameters and empirical stock status indicators (Cope & Punt, 2009; Geromont & Butterworth, 2015; Jardim, Azevedo, & Brites, 2015; Punt, Campbell, & Smith, 2001). In this paper, we only consider analytical methods that estimate mortality rates for data-limited applications where growth is known but the catch is not necessarily known.

Data-limited, size-based methods require knowledge of life history parameters and may make simplifying assumptions regarding the population dynamics underlying the data. Selectivity may be estimated or fixed in the analysis (for example, only fully selected lengths are analyzed for some methods, which require the user to determine the length of full selectivity beforehand). Constant-rate assumptions are often made regarding recruitment and mortality and some methods assume deterministic growth. In situations when only a single sample of lengths is available, these assumptions may be needed for mortality estimation to remain tractable. However, when multiple years of data are used in a single analysis, some equilibrium assumptions can be relaxed in the model. Length-based methods are methodologically rich by allowing for the inclusion of

additional data types. For example, an index of recruitment, in conjunction with length data for several years, would allow for mortality estimation without assumptions of equilibrium recruitment or constant mortality (Gedamke, Hoenig, DuPaul, & Musick, 2008).

In this paper, we review the current methods available for data-limited, size-based assessments. We first describe various methods and their modeling approaches for estimating mortality rates. Next, we list the assumptions and discuss the effect of violations of assumptions on mortality estimation. We summarize the results of simulation studies which have evaluated the robustness of these methods to these simplifying assumptions and describe diagnostic procedures that can be used to evaluate the extent to which mortality estimation has been biased. Finally, we describe how some assumptions may be relaxed.

1.3. Description of the size-based mortality estimators

In this section, a brief description of the size-based methods, with their data requirements and assumptions, is provided (Table 1.1). The data requirements can be generally split into three categories in terms of availability: (1) a single length composition (from a single year or pooled from multiple years), (2) multiple years of length composition, and (3) multiple years of length composition as well as auxiliary data. In all three cases, some growth and life history parameters are assumed to be known. The required life history parameters and assumptions regarding growth (variability in size-at-age), recruitment, selectivity, and mortality (with respect to time and age) vary depending on the method.

1.3.1. Mean length-based mortality estimators

The mean length-based mortality estimators include the Beverton-Holt equation (BHE; Beverton & Holt, 1956) and the Gedamke-Hoenig (GH) nonequilibrium Beverton-Holt extension (Gedamke & Hoenig, 2006). Assuming constant recruitment and constant selectivity (knife-edge above length L_c) over time, the mean length is function of the mortality rate experienced in the population. The BHE is an equilibrium moment estimator in which the sample mean length is equated with its expected value as a function of Z. After solving for Z, the estimator is expressed as

$$Z = \frac{K(L_\infty - \bar{L})}{\bar{L} - L_c}, \tag{1.1}$$

where L_∞ and K are the asymptotic length and the growth coefficient, respectively, from the von Bertalanffy equation, L_c is the fully selected length, and $\bar{L}$ is the mean length of animals larger than L_c. For the BHE, a single mean length is calculated from either a single year or from multiple years when constant Z with time and age is to be assumed.

The GH estimator is an expansion of the BHE for estimating total mortality rates in historical time stanzas from multiple years of mean length. When mortality changes, the equilibrium BHE will underestimate the magnitude of the change in mortality (Hilborn & Walters, 1992). To correct for this, the nonequilibrium estimator divides the time series of mean lengths into time stanzas where total mortality is constant within time stanzas. The transitory behavior of the mean length from one mortality rate to another is then modeled so that the predicted mean length is a function of prior mortality rates and the time elapsed since those mortality rates were experienced. The estimated parameters

are the mortality rates in each stanza and the change points between time stanzas (in calendar time) that maximize the log-likelihood (log L), which is proportional to the negative of the weighted sum of squares of the deviations between observed and predicted mean lengths,

$$\log L \propto -\sum_{y} m_y \left(\bar{L}_y - \mu_y\right)^2 , \tag{1.2}$$

where y indexes year, $\bar{L}_y$ is the observed mean length, μ_y is the predicted mean length as a function of estimated parameters, and m_y is the sample size used to calculate $\bar{L}_y$. Typically, the model is fitted successively with differing numbers of change points and model selection procedures (e.g., Akaike Information Criterion, AIC) are used to select the model with the appropriate number of change points (Figure 1.1).

GH forms a general framework for mortality estimation using mean lengths with other data types, including indices of recruitment (Gedamke, *et al*., 2008) and indices of abundance (Huynh, Gedamke, Porch, *et al*., 2017). The method has also been extended to analyze stock complexes with synchronous changes in fishing mortality among several stocks (Huynh, Gedamke, Hoenig, & Porch, 2017). In all of these, the mortality rates are stanza-specific. Along with the BHE, fishing mortality F can be obtained by subtracting natural mortality M from the estimate of total mortality ($F = Z - M$), although this is not necessary to use the methods.

Another extension of GH has been developed to estimate year-specific mortality (GH with effort; Then, Hoenig, & Huynh, in press). In this model, fishing effort is used as an index of mortality. Total mortality in year y is parameterized as

$$Z_y = F_y + M = qf_y + M , \tag{1.3}$$

where the fishing mortality F in year y is the product of the fishing effort f and the catchability coefficient q, and total mortality is the sum of the natural mortality rate M and fishing mortality (the estimated parameters are q and M). Estimates of q and M are often highly negatively correlated, but M may be fixed to an assumed value. In an application to stocks of Norway lobster *Nephrops norvegicus*, the effective effort was calculated as the ratio of the commercial catch and catch per unit effort, and the estimated M was often close to the assumed value (Then, *et al*., in press).

All mean length-based methods assume no growth variability, constant recruitment, knife-edge selectivity above length L_c, and perfectly-known growth parameters.

1.3.2. Composition-based mortality estimators

Four composition-based methods are considered in this paper: the length-converted catch curve (LCCC), a linear regression based method, and the LB-SPR, S6, and LIME models, which are likelihood-based.

The LCCC is a length-based analogue of a cross-sectional, age-based catch curve (Pauly, 1983). With an age-based catch curve, the estimate of total mortality Z is the magnitude of the slope in a linear regression of the logarithm of catch-at-age versus age of fully selected animals (Ricker, 1975). With the length-converted catch curve, a length frequency distribution is instead used where the lengths are deterministically converted to relative ages. Assuming no growth variability, constant recruitment, and von Bertalanffy growth, an estimate of Z is obtained from a linear regression of the logarithm of catch-at-relative-age (C_i) from fully selected length bins versus relative-age (t_i),

$$\log(C_i) = a + \left(1 - \frac{Z}{K}\right) t_i , \tag{1.4}$$

where t_i is the relative age and is defined as $t_i = \log(1 - L_i / L_\infty)$ with L_i as the length at the midpoint of the *i*-th length bin, *a* is the intercept of the linear regression, and $1 - Z/K$ is the slope of the regression line (Figure 1.2).

Small size classes are removed from the analysis due to incomplete selection. Two additional considerations for the LCCC are also needed. First, length bins whose midpoints are larger than L_∞ are removed because relative ages cannot be calculated for these length bins. Secondly, additional size classes can also be removed if there is perceived observation error or dome shaped selectivity, or if overlapping ages in the large size bins break down the linear relationship in Equation 1.4 (Sparre & Venema, 1998). There are several parameterizations of the LCCC, all of which are equivalent (Pauly, 1983; Punt, Huang, and Maunder, 2013).

The Length-Based Spawner Potential Ratio (LB-SPR) method is an equilibrium mortality estimator which uses the *M/K* ratio, L_∞, the coefficient of variation (CV) in length-at-age (also referred to as the CV of L_∞), and a single length composition to estimate *F/M.* Assuming constant recruitment and logistic selectivity, the shape of the length composition for an unexploited population is determined from *M/K* (Hordyk, Ono, Sainsbury, Loneragan, & Prince, 2015), while the observed length composition is a function of *Z/K* (Pauly, 1984). Based on the extent of truncation in the observed length frequency relative to an unexploited length composition (Figure 1.3), the fishing mortality relative to natural mortality, *F/M*, is estimable because

$$\frac{F}{M} = \frac{\frac{Z}{K} - \frac{M}{K}}{\frac{M}{K}}. \quad (1.5)$$

Estimates of selectivity (logistic function of length) and *F/M* are those which maximize the log-likelihood which assumes a multinomial distribution for the observed length composition,

$$\log L \propto \sum_i N p_i \log(\tilde{p}_i), \quad (1.6)$$

where N is the sample size of the length composition and p_i and $\tilde{p}_i$ are the observed and predicted proportion, respectively, in length bin i.

Individual estimates of *M* and *K* are unnecessary because the ratio of mortality and growth determines the size structure of the virgin population. This is advantageous when the estimates of *M* and *K* are highly uncertain. By using the ratio, less information is needed to use LB-SPR. Meta-analysis has shown that there is invariance in *M/K* based on life history and reproductive strategies (Prince, Hordyk, Valencia, Loneragan, & Sainsbury, 2015), although this may not necessarily hold true in all taxa (Nadon & Ault, 2016). From only *M/K*, the magnitude of *F* is not estimable, but *F/M* can also be used to obtain F/F_{MSY} (section 1.5.1). The spawner potential ratio is used as the biological reference point.

There are two versions of LB-SPR, in which either an age-structured model or length-structured model is used. In the age-structured model, the predicted age composition is converted to a length composition via an age-length transition matrix to fit to the observed lengths (Hordyk, Ono, Valencia, *et al*., 2015). Selectivity is age-based and the variability in length-at-age in the model is always normally distributed. In the

length-structured version of LB-SPR, length-based selectivity is implemented, which more quickly fishes out faster-growing individuals in a cohort than slower-growing individuals (see Figure 1 of Punt, *et al*., 2013). The length-structured LB-SPR typically estimates a higher mortality rate from the same length composition than the age-structured LB-SPR which ignores this phenomenon (Hordyk, *et al*., 2016). However, with length-based selectivity, growth in the underlying population is distorted based on the selectivity function and the magnitude of fishing mortality. The extent to which growth is distorted due to length-based selectivity, and guidance on whether to use age or length-based selectivity, has not been extensively studied (Sampson, 2014).

The Single Species, Size-Structured, Steady State (S6) model uses a weight-based theory of population dynamics for mortality estimation (Andersen & Beyer, 2015). In equilibrium, the energy budget available to an individual animal is a power function of weight. Energy is devoted to activity, growth, and reproduction. As animals grow towards the asymptotic weight (W_∞), the growth rate decreases and investment in reproduction increases. The expected size distribution of a population is therefore a function of growth, natural mortality and fishing mortality. The functions describing the metabolic processes are typically power functions while maturity and fishing mortality are weight-based logistic functions. The weight at 50% maturity, used to model energy devoted to reproduction, is assumed to be $0.25W_\infty$. Most parameters for the power functions in the model are typically assumed to be invariant among fish stocks. The exception is the physiological constant of mortality, which describes mortality and growth in an unexploited population and is analogous and proportional to *M/K*. Based on simulation, S6 was most sensitive to the physiological constant. Thus, this parameter can

be specified in the model by the user by providing a value of *M/K*. Assuming constant mortality with time and constant recruitment, the S6 model estimates fishing mortality *F*, selectivity (as a weight-based logistic function), and W_∞ from a single weight composition sample. Deterministic growth appears to be assumed in S6 because all individuals grow to W_∞. The population dynamics model and likelihood are presented in Table 1 of Kokkalis *et al*. (2015).

The Length-based Integrated Mixed Effects (LIME) model is a general framework for estimating fishing mortality without assumptions of constant mortality with time and constant recruitment. It assumes that the data are informative that year-specific recruitment can be estimated as a random effect variable and year-specific fishing mortality can be estimated as a random walk variable. To do so, multiple years of length composition data are needed. The estimation model is an age-structured model in which the predicted age compositions are converted to length compositions to fit to the observed lengths. The primary intent of the LIME model is to use length composition data for mortality estimation, but the model optionally allows for the use of an index of abundance to better estimate mortality. Using both data types together provides more information on estimating recruitment than using either alone (section 1.4.5). The equations for the population dynamics model are presented in Tables 2-4 of Rudd & Thorson (in press). The log-likelihood of the length composition, assuming a multinomial distribution, is

$$\log L \propto \sum_y \sum_i N_y p_{i,y} \log\left(\tilde{p}_{i,y}\right), \tag{1.7}$$

where the variables are the same as Equation 1.6 with an additional subscript for year *y*. Alternatively, a Dirichlet-multinomial distribution can be used to account for

overdispersion in the composition data (Thorson, Johnson, Methot, & Taylor, 2017). The LIME model allows for multiple years of length and index data to be analyzed. Growth, natural mortality, and the CV in length at age are needed life history parameters for this model.

1.4. Violations of model assumptions: implications, diagnostics, and solutions

In any application, the appropriateness of a stock assessment model needs to be evaluated. In this section, we discuss how violations of assumptions regarding growth, selectivity, recruitment, and mortality may occur when using the size-based methods and their effects on mortality estimation (Table 1.2). We propose diagnostics to determine whether violations have occurred and develop solutions to address such violations. Six issues and questions relevant to application of size-based mortality estimators are discussed:

1. Life history parameters are not known well, and the mortality estimators assume perfect knowledge. How sensitive are the models to misspecification of life history parameters?
2. Deterministic growth is assumed in the mean length-based mortality estimators and LCCC, while other methods assume variability in size-at-age is known. How critical are these assumptions to estimating mortality?
3. There is some fishing mortality at lengths below L_c, but the mean length-based methods and LCCC assume knife-edge selectivity (no fishing mortality below L_c and constant mortality above L_c). What is the best choice of L_c?

4. There may be dome-shaped selectivity, but size-based methods typically assume flat-top selectivity (full selection of large animals). How can dome-shaped selectivity be detected? What options are available when there is dome selectivity?
5. There is a suspected large year class in the population, but all size-based methods assume constant recruitment. How can a pulse in recruitment be diagnosed? What can be done for mortality estimation?
6. There is a trend in mortality over time. How does this affect models that assume constant mortality?

Overall, size-based mortality estimators allow for modifications and extensions to models to relax assumptions. The ability to diagnose and address violations of assumptions is generally improved with the availability of a time series of auxiliary data.

1.4.1. Uncertainty in life history parameters

The requisite life history information for size-based mortality estimators includes growth parameters and, in some cases, natural mortality (Table 1.1). Size-based methods model growth with the von Bertalanffy function (with parameters L_∞ and K) and typically use a single value of M over all ages. Although it is generally accepted that M varies with size, it may be reasonable to use a single value of M (constant with size) in the mean length methods and the LCCC since these methods only evaluate size classes that are fully selected by the gear. For the composition-based mortality estimators, size-

specific *M* can be modeled as a monotonically decreasing power function (Hordyk *et al*., 2016, Kokkalis *et al*., 2015).

Based on simulation, mortality estimates from size-based methods are more sensitive to L_∞ than *K* (Gedamke & Hoenig, 2006; Hordyk, Ono, Valencia, *et al*., 2015; Huynh *et al*., in review; Rudd & Thorson, in press). With an overestimate of L_∞, there will be fewer animals near L_∞ than when a correct estimate is used. Thus, mortality estimates will be positively biased with an overestimate of L_∞ and negatively biased with an underestimate. The direction of the bias is the same with over/underestimates of *K*, but the magnitude of the bias appears to be smaller (Gedamke & Hoenig, 2006; Rudd & Thorson, in press). For the LB-SPR and LIME models, overestimates of *M/K* and *M* will result in a negative bias in *F/M* and *F* estimates, respectively (Hordyk, Ono, Valencia, *et al*., 2015; Rudd & Thorson, in press).

Life history parameters may not be known well or may be unavailable for data-limited stocks. In the case where estimates are unavailable for the stock of interest, information may be borrowed from other stocks of the same species or empirical estimates from meta-analytic relationships can be used. For example, Then, Hoenig, Hall, & Hewitt (2015) developed an empirical estimator for natural mortality based on maximum observed age and Nadon & Ault (2016) modeled empirical relationships among the maximum observed length (L_{max}), growth parameters, length at maturity, and natural mortality for several tropical reef fish taxa. Meta-analyses of life history parameters are beyond the scope of this review but have been extensively covered in the literature (Hoenig *et al*., 2016).

Uncertainty in mortality estimation can be quantified in two ways in a given application. First, stochastic mortality estimates can be obtained via Monte Carlo sampling from probability distributions of life history parameters (Kokkalis *et al*., 2016; Nadon, 2017; Prince, Victor, *et al*., 2015). Second, a Bayesian implementation of the size-based estimators can be used to derive posterior distributions of mortality based on the priors for the life history parameters, although such models currently remain in preliminary development (Brodziak, *et al*., 2012; Harford, Bryan, & Babcock, 2015). High uncertainty is incorporated with a large variance in the priors, and the variance of the posterior distribution is subsequently evaluated. Secondly, sensitivity analyses can more simply determine the range in mortality estimates based on the plausible range of parameter values.

Monte Carlo sampling and Bayesian modeling provide a distribution of mortality estimates from which confidence intervals and credibility intervals, respectively, can be obtained. These are useful if one is unsure of the direction of the bias in the life history parameter. On the other hand, sensitivity analysis is more useful when one suspects an overestimate or underestimate in the life history parameter.

If a value of a life history parameter cannot be established with reasonable certainty, some information regarding mortality remains available. If the natural mortality rate cannot be established, then the mean length-based methods or the LCCC can be used to estimate total mortality, and approaches that use only the total mortality rate for management advice can be explored (Die & Caddy, 1997). On the other hand, if K is unknown or uncertain, then relative changes in mortality are estimable. For example, let Δ be the relative change in total mortality, defined as

$$\Delta = \frac{Z_2 - Z_1}{Z_1} \tag{1.8}$$

where Z_1 and Z_2 are total mortality rates from two different time periods (assuming equilibrium in each time period). If the BHE is used to estimate mortality, then Z_1 and Z_2 are defined as

$$Z_1 = \frac{K(L_\infty - \bar{L}_1)}{\bar{L}_1 - L_c} \tag{1.9a}$$

and

$$Z_2 = \frac{K(L_\infty - \bar{L}_2)}{\bar{L}_2 - L_c}, \tag{1.9b}$$

where $\bar{L}_1$ and $\bar{L}_2$ are the mean lengths in the two different time periods. After substitution of Equations 1.9a and 1.9b into Equation 1.8, it is shown that Δ is independent of *K*.

1.4.2. Deterministic versus stochastic size-at-age

The mean length-based methods and the LCCC assume no variability in length-at-age. The implications of the growth assumption have been tested by simulation. The BHE and LCCC performed better with higher growth variability, i.e., the methods are robust to the failure of the assumption of deterministic growth (Then, Hoenig, Gedamke, & Ault, 2015; Huynh *et al*. in review). It would also be expected that Gedamke-Hoenig, as an extension of the BHE, would also be robust to violation of this assumption.

Alternatively, the LB-SPR and LIME models parameterize some growth variability in the estimation model. The coefficient of variation (CV) in length at age is provided to the model by the user. An estimate could be obtained from an age-growth

study, but otherwise, a CV of 0.1 has been assumed in applications if the value is unknown (Prince, Victor, Kloulchad, & Hordyk, 2015). Both LIME and LB-SPR are relatively insensitive to misspecification of the CV in growth compared to other life history parameters (Hordyk, Ono, Valencia, *et al*., 2015; Rudd & Thorson, in press).

In S6, no conversion between the age and size compositions is needed because the population is modeled natively in terms of weight units. For all other methods considered here, issues concerning variability in size-at-age appear to be minor compared to others.

1.4.3. Knife-edge selectivity versus gradual selectivity

The mean length-based methods and the LCCC assume knife-edge selectivity with size and age (this correspondence arises from the assumption of deterministic growth). To meet this assumption for the mean length-based methods, a value of L_c must be selected by the user and lengths smaller than L_c are removed from the analysis. For the LCCC, an analogous step is also needed where the length composition data is truncated to remove length classes that are believed to be incompletely selected by the fishing gear.

Knife-edge selectivity is typically an approximation to logistic selectivity. Length classes smaller than the first fully selected length experience a reduced mortality rate, but the associated methods (BHE, Gedamke-Hoenig, and LCCC) only estimate the apical mortality rate. The choice for L_c for the mean length-based methods is usually be the first fully selected length. Another suggestion for L_c is to use the presumed the length of 50% selectivity. Under this interpretation, size classes smaller than the first fully selected length are also impacted by fishing mortality and are to be included in the mortality estimation procedure. However, this choice tends to underestimate mortality due to the

inclusion of incompletely selected size classes (Figure 1.4). Sparre & Venema (1998) suggested an additional variant of the BHE where the L_c is the length at 50% selectivity, but all size classes are used to calculate the mean length. This variant has not been evaluated for use. Although it is likely to be biased high, the variant may be useful if historical records indicate the mean length of the entire catch but the histogram is not available to determine L_c.

Some caution is needed with the assumption of the mode as the fully selected length. In an unexploited stock with constant recruitment, the abundance at length is a balance of growth towards L_∞ and natural mortality; this process can be summarized in the *M/K* life history trait (Beverton 1992; Prince, Hordyk, *et al*. 2015). A low *M/K* population is expected to have a mode near L_∞ rather than the length of full selectivity due to the "pile-up" effect where many ages comprise a single length class as length increases (Figure 1.3). The mode does not correspond to the fully selected length unless exploitation is high. On the other extreme, a high *M/K* population features a monotonically decreasing abundance at length and the ascending limb in the observed catch frequency would reflect selectivity. The mode is also less sensitive to increases in fishing mortality in high *M/K* situations compared to low *M/K* situations.

When exploitation is high, i.e., high *Z/K* scenario, the mode appears to be a viable choice for L_c regardless of *M/K*. The mode should remain near the length of full selectivity for high *M/K* stocks regardless of fishing mortality. However, the fully selected length can be smaller than the mode in low *M/K* stocks (Figure 1.3). Huynh *et al*. (in review) evaluated the performance of the BHE and LCCC with difference values of L_c and left truncation lengths, respectively. Candidate truncation lengths included the mode

and the half-modal abundance lengths (i.e., the first length at which the abundance is at least 50% of that at the mode). They considered two values of M/K = 0.5 and 2. There did not appear to be significant differences in performance with the different truncation lengths in the high M/K scenarios. While the half-modal abundance length was the preferred over the modal length as the truncation length in the low M/K scenarios, there still was high bias and low precision in mortality estimates. Presumably, the half-modal length was closer to the fully selected length than the modal length, but still did not precisely correspond to the fully selected length. A large value of L_c would be more likely to guarantee that partially recruited sizes are removed. However, this step also removes potentially useful data and thus increases the variance of mortality estimates.

As alternatives to methods which use knife-edge selectivity, the composition-based methods LB-SPR, S6, and LIME estimate logistic selectivity. However, simulations suggest that LB-SPR also performs poorly for low M/K stocks (Hordyk, Ono, Valencia, *et al*., 2015). Assuming equilibrium, it will generally be difficult to identify the fully selected length in low M/K stocks from the length frequency data alone, which results in potentially arbitrary choices for L_c and biased or imprecise mortality estimates.

In nonequilibrium situations, selectivity in low M/K stocks may be estimable if data from years in which fishing mortality was high are available. The shape of the length composition in years with high F would be informative for selectivity. This situation would avoid the problem of selectivity estimation from the ascending limb of the length frequency distribution. The LIME model would be an appropriate model in this scenario.

1.4.4. Flat-top versus dome-shaped selectivity

Dome-selectivity is conflated with high mortality and the two are not simultaneously estimable because both situations reduce the abundance of large animals. Thus, the length-based methods in general will overestimate mortality when there is dome-shaped selectivity, as demonstrated through simulation (Then, Hoenig, Gedamke, *et al*., 2015; Huynh *et al*., in review; Hordyk, Ono, Valencia, *et al*., 2015). However, the influence of dome selectivity is diminished when mortality is very high because large animals are fished out of the population regardless of the selectivity (Huynh *et al*., in review).

A potential diagnostic for dome selectivity is to compute estimates of Z from the BHE for an increasing series of L_c values (Figure 1.4, Then *et al*., in press). Under logistic selectivity, mortality estimates should plateau with increasing values of L_c because the selectivity assumptions are being met. This plateau would not occur when there is dome selectivity because, with increasing L_c, the cryptic abundance of large animals comprises a larger proportion of the assumed catch. Thus, the estimates of Z continue to increase as L_c increases. This behavior is also consistent with an overestimate of L_∞, but this diagnostic could be used to identify if either is (or both are) occurring in the analysis.

Several strategies can combat this overestimation. First, one could incorporate external selectivity estimates in the mortality estimators. Experimental field studies can be used to obtain empirical selectivity estimates (Cadrin, DeCelles, & Reid, 2016). Second, the LCCC could be used where length bins over which dome selectivity is occurring are removed. This would allow for the estimation of the apical total mortality rate from the remaining length bins (Figure 1.2).

A variant of the BHE has been developed to allow for mortality estimation from mean length when there is dome selectivity (Ehrhardt & Ault, 1992). Similar to the simplification of logistic selectivity with knife-edge selection, dome selectivity is simplified assuming knife-edge truncation at length L_λ on the right (where $L_c < L_\lambda < L_\infty$), which is fixed by the user. Simulations showed that the behavior of the Ehrhardt-Ault variant, in cases where there is length truncation, is unpredictable and complex (Then, Gedamke, Hoenig, *et al*., 2015). Owing to variability in size-at-age in the population, mortality estimates remained biased even if L_λ was correctly specified and were often less precise than those from the BHE. Overall, routine use of the Ehrhardt-Ault estimator is not recommended.

1.4.5. Large year class in recruitment

Recruitment is estimable in the LIME model, while recruitment is constant in all other models. Large deviations in cohort strength relative to neighboring years could be identified through evaluation of length compositions and indices of abundances. In the length compositions, a large cohort would be represented by a mode in the size composition which progresses through the size structure over time while large cohorts would increase the index based on the increased abundance in the population (Figure 1.5). If only a single length composition is available, a large cohort may be identified if the distribution is multimodal, although a time series provides more support.

We propose that the mean length and index of abundance be used as a diagnostic for detecting mortality and recruitment changes (Figure 1.5). If there is a change in mortality, then the mean length and index are affected in the same direction, that is, the

increased mortality reduces both the mean length and index by decreasing the abundance of large animals. Likewise, decreased mortality increases both. On the other hand, a temporary increase in recruitment initially produces opposing changes where the mean length decreases while the index increases. Over time, the effect is dampened as the cohorts age. These patterns could help determine whether the observed lengths and indices are predominantly controlled by recruitment or mortality dynamics.

While stochastic recruitment may be perceived as observation noise in a model, trends in recruitment due to autocorrelation from environmental conditions or the stock-recruitment relationship over a considerable time period will bias mortality estimates from length data. For example, decreased recruitment would increase the mean length due to fewer or missing animals in small size classes. Mortality estimates would be biased high.

For both mean length-based and composition-based methods, either an index of recruitment or an index of abundance can be utilized to account for variable recruitment. In the former, the index of recruitment is used in a model in lieu of the constant recruitment assumption. A variant of Gedamke-Hoenig with an index of recruitment was developed for the assessment of barndoor skate *Dipturus laevis* in Georges Bank and the Gulf of Maine (Gedamke, *et al*., 2008). In this model, it is presumed that the index of recruitment is known perfectly, but this approach can be taken if the estimation error is more adversely affected by assuming constant recruitment than an error-free index. The index can be used because only the relative strength among cohorts, not the absolute magnitude, is relevant in the model. Composition-based methods could also be modified to use an index of recruitment, but this has not been done. The availability of a

recruitment index is usually limited to a small subset of data-limited stocks, for example, if less valuable or bycatch species are caught in fishery-independent surveys. If survey catch is identified by length, then indices can be developed for length classes which correspond to recruits.

With an index of abundance, mortality can be estimated in an integrated framework which uses both time series of length and index data in the likelihood (Huynh, Gedamke, Porch, *et al*., 2017; Rudd & Thorson, in press). The LIME model separates cohort strength from observation noise. Thus, recruitment is estimated as a random effect. A large cohort is identifiable from a mode in the length composition which progresses in size over multiple years. The Gedamke-Hoenig approach has been generalized to estimate mortality from both mean lengths and abundance indices in the likelihood function (Huynh, Gedamke, Porch, *et al*., 2017). Simulations suggest that the integrated model generally performs better than the mean length only model when recruitment is stochastic. Overfitting would be less likely to occur compared to when only mean lengths are used in the base Gedamke-Hoenig estimator because changes in mortality are estimated only if there are synchronous changes in both data types. Changes in mean length and index of abundance due to recruitment dynamics would result in negatively correlated residuals. The explanatory power of this extension of Gedamke-Hoenig could be more limited compared to LIME because the constant recruitment assumption remains in the former.

1.4.6. Trend in mortality

Equilibrium methods assume constant mortality over time. When multiple years of data are available, all data may be collated into one length composition or independent analyses performed for each year of the data, but equilibrium methods will underestimate the true rate of change in mortality if mortality is not constant. On the other hand, models that incorporate multiple years of data (Gedamke-Hoenig and its various extensions and LIME) allow for time-varying mortality estimates within a single analysis.

There are three approaches for time-varying mortality estimation. First, the time stanza approach of the Gedamke-Hoenig estimator because there is limited information for estimating mortality in a fixed effects-only approach without over-parameterizing the model (Figure 1.1). Second, year-specific mortality can be estimated as random effects through a random walk (Brodziak, *et al*., 2012; Rudd & Thorson, in press). The estimated mortality rate in a given year is allowed to vary but constrained to be close to the previous year's estimate. This would allow for trend-based mortality estimates over time instead of stepwise changes at certain points in time. The LIME model estimates fishing mortality with the random walk approach. However, the variance of the random walk needs to be evaluated to avoid overfitting. Third, effort data could be used to provide year-specific mortality estimates, as is done in the Gedamke-Hoenig with effort model. Composition-based methods can also be modified accordingly.

1.5. Considerations for applications

Size-based mortality estimators have been popular analytical methods in sub-tropical and tropical regions, including the U.S. Caribbean (Puerto Rico and U.S. Virgin Islands), Hawaii, Palau, and Florida (Table 1.3). The methods have also been evaluated

for European stocks (ICES, 2015; ICES, 2016). For management purposes, it is often practical to estimate fishing mortality relative to biological reference points and to provide some measure of uncertainty. Software packages allow for streamlining the workflow associated with using the analytical methods.

1.5.1. Biological reference points

Estimates of mortality can be used in the context of biological reference points for management advice. From the model, F or F/M is compared with F_{proxy} or F_{proxy}/M, respectively, the proxy fishing mortality rate for maximum sustainable yield, to determine stock status, i.e. overfishing occurs if $F/F_{proxy} > 1$. Possible reference point proxies include F_{max}, the fishing mortality that maximizes yield-per-recruit (YPR; Beverton & Holt, 1957); $F_{0.1}$, the fishing mortality where the slope of the YPR curve is 10% of that at $F = 0$; $F_{X\%}$, the fishing mortality that reduces the spawning potential ratio (SPR; Goodyear, 1993) to $X\,\%$ of that at $F=0$, or F_{XM}, the fishing mortality at $X\,\%$ of the assumed M (Zhou, Yin, Thorson, Smith, & Fuller, 2012). Maturity, natural mortality, and length-weight parameters are usually needed for reference point calculation in addition to the life history parameters for mortality estimation. If only F/M is estimated (e.g., in LB-SPR), then F/F_{proxy} can be obtained without necessarily knowing the magnitude of F, since

$$\frac{F}{F_{proxy}} = \frac{F}{M}\left(\frac{F_{proxy}}{M}\right)^{-1}. \tag{1.10}$$

The F_{proxy}/M ratio can be obtained from SPR, or it can be defined as a scalar proportion of M.

Applications of the length-based methods have used reference points from YPR and SPR for F_{proxy}. The mean length-based methods were developed without specific reference points, but M and SPR proxies have been used with the Ehrhardt-Ault variant of the BHE (Ault, Smith, & Bohnsack, 2005; Nadon, 2017; Nadon, Ault, Williams, Smith, & DiNardo, 2015). Proxies from both YPR and SPR have been used with Gedamke-Hoenig (ICES, 2015). The likelihood-based composition methods have been developed with particular reference points. LB-SPR reports the SPR associated with the estimated F/M, S6 uses F_{max} as the proxy, while LIME reports both SPR and YPR-based reference points.

1.5.2. Software packages

An array of computer software is available for using length-based mortality estimators either in the form an R package or a stand-alone GUI program (Table 1.3). Typically, R packages provide supporting functions for standardizing the workflow associated with the methods. In addition to parameter estimation, functions are needed for data processing, plotting figures, and performing statistical diagnostics. The R packages are in active development with version control through Github. Most software packages also contain extensive supporting documentation for potential users.

The packages implement the standard application of the described model. Stock-specific applications can be accommodated by varying a numeric parameter. For example, fecundity is specified as a power function of length in the calculation of spawning potential ratio in the LB-SPR model. For species with determinate biological fecundity, such as elasmobranchs, the exponent of the power function can be set to zero

whereas larger exponents can be used for teleosts. Modifications that require a different functional form, e.g., the Gompertz or Richards growth functions in lieu of the von Bertalanffy function, would require a new model derivation (in the case of the mean length-based methods) or case-specific computer code that implement these modifications numerically.

Routines for performing uncertainty evaluation in software packages are currently limited. Some software, such as the R packages for Gedamke-Hoenig and LB-SPR report the asymptotic standard error of model estimates. However, these intervals may not be appropriate because the intervals are conditional on correct model specification, i.e., these intervals do not consider whether the equilibrium assumptions are met as part of the uncertainty evaluation. The fishmethods R package bootstraps the length data to obtain confidence intervals in the estimated Z from the BHE. Generally, the R computing environment is flexible to allow for users to write supplemental code as necessary for Monte Carlo sampling of parameters and data bootstrapping. This can then be used with the core estimation functions in the R packages.

1.6. Conclusions

There is a spectrum of data requirements and flexibility in model structure among data-limited assessment methods. This is desirable because data-limited fisheries also vary in data quality and quantity (Bentley, 2015). Consider two fisheries that are classified as data-limited, but one only has opportunistic sampling while another has a long-standing, designed sampling program. Equilibrium methods could be options in the former scenario, but the assessment capabilities will be much greater in the latter scenario

due to presumed higher quality of data. With multiple years of data, diagnostics are available to determine whether certain assumptions are appropriate and methods which relax some assumptions are available for assessment. Methods that incorporate multiple years of data provide the most flexibility in mortality estimation, but simple, equilibrium methods remain valuable diagnostic tools to evaluate mismatch (if any) between assumptions, data, and life history parameters.

We recommend the following practices associated with size-based mortality estimators:

1. Use sensitivity analyses and Monte Carlo sampling to ascertain uncertainty in mortality estimates due to uncertainty in life history parameters.
2. Length-based mortality estimators are generally least sensitive to assumptions regarding variability in length at age.
3. For high *M/K* stocks, the mode of the length frequency distribution is generally the length of full selectivity for the LCCC and L_c for the mean length-based mortality estimators. For low *M/K* stocks, lengths smaller than the modal length may plausibly be the length of full selectivity.
4. To address dome-shaped selectivity, external selectivity estimates can be used to parameterize selectivity in the mortality estimator. The LCCC can also be used to estimate mortality over a subset of fully selected lengths.
5. Use multiple years of length data and auxiliary data where available. Potential pulses in recruitment can be evaluated and the constant mortality over time assumption can be relaxed.

1.7. References

Andersen, K. H., & Beyer, J. E. (2015). Size structure, not metabolic scaling rules, determines fisheries reference points. *Fish and Fisheries, 16*, 1-22.

Ault, J. S., Smith, S. G., & Bohnsack, J. A. (2005). Evaluation of average length as an estimator of exploitation status for the Florida coral-reef fish community. *ICES Journal of Marine Science, 62*, 417-423.

Babcock, E. A., Coleman, R., Karnauskas, M., & Gibson, J. (2013). Length-based indicators of fishery and ecosystem status: Glover's Reef Marine Reserve, Belize. *Fisheries Research, 147*, 434-445.

Bentley, N. (2015). Data and time poverty in fisheries estimation: potential approaches and solutions. *ICES Journal of Marine Science, 72*, 186-193.

Beverton, R. J. H. (1992). Patterns of reproductive strategy parameters in some marine teleost fishes. *Journal of Fish Biology, 41 (Supplement B)*, 137-160.

Beverton, R. J. H., & Holt, S. J. (1956). A review of methods for estimating mortality rates in fish populations, with special reference to sources of bias in catch sampling. *Rapports et Procès-verbaux des Reunions, Conséil International Pour L'Exploration de la Mer, 140*, 67-83.

Beverton, R. J. H., & Holt, S. J. (1957). *On the dynamics of exploited fish populations*. U.K. Ministry of Agriculture, Fisheries, Food, and Fishery Investigations Series II, Vol. XIX.

Brodziak, J., Gedamke, T., Porch, C., Walter, J., Courtney, D., O'Malley, J, & Richards, B. (2012). A Workshop on Methods to estimate Total and Natural Mortality Rates Using Mean Length Observations and Life History Parameters. U.S. Department of Commerce, NOAA Technical Memorandum, NOAA-TM-NMFS-PIFSC-32, 26 pp + Appendix.

Cadrin, S. X., DeCelles, G. R, & Reid, D. (2016). Informing fishery assessment and management with field observations of selectivity and efficiency. *Fisheries Research, 184*, 9-17.

Chrysafi, A., & Kuparinen, A. (2016). Assessing abundance of populations with limited data: Lessons learned from data-poor fisheries stock assessment. *Environmental Reviews, 24*, 25-38.

Cope, J. M., & Punt, A. E. (2009). Length-Based Reference Points for Data-Limited Situations: Applications and Restrictions. *Marine and Coastal Fisheries, 1*, 169-186.

Die, D. J., & Caddy, J. F. (1997). Sustainable yield indicators from biomass: are there appropriate reference points for use in tropical fisheries? *Fisheries Research, 32*, 69-79.

Ehrhardt, N. M., & Ault, J. S. (1992). Analysis of Two Length-Based Mortality Models Applied to Bounded Catch Length Frequencies. *Transactions of the American Fisheries Society, 121*, 115-122.

Gedamke, T., & Hoenig, J. M. (2006). Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of Goosefish. *Transactions of the American Fisheries Society, 135*, 476–487.

Gedamke, T., Hoenig, J. M., DuPaul, W. D., & Musick, J. A. (2008). Total mortality rates of the barndoor skate, *Dipturus laevis*, from the Gulf of Maine and Georges Bank, United States, 1963-2005. *Fisheries Research, 89*, 17-25.

Geromont, H. F., & Butterworth, D. S. (2015). Generic management procedures for data-poor fisheries: forecasting with few data. *ICES Journal of Marine Science, 72*, 251-261.

Goodyear, C. P. (1993). Spawning stock biomass per recruit in fisheries management: foundation and current use. In S.J. Smith, J.J. Hunt, & D. Rivard (Eds). *Risk evaluation and biological reference points for fisheries management* (pp. 67-81). Canadian Special Publication of Fisheries and Aquatic Science 120.

Gayanilo, Jr., F. C., Sparre, P., & Pauly, D. (2005). *FAO-ICLARM Stock Assessment Tools II (FiSAT II). Revised version. User's guide*. FAO Computerized Information Series (Fisheries). No. 8, Revised version. FAO, Rome. 168 pages.

Harford, W., Bryan, M, & Babcock, E. A. (2015). Probabilistic assessment of fishery status using data-limited methods. SEDAR46-DW-03. SEDAR, North Charleston, SC. 5 pp.

Hilborn, R., & Walters, C. J. (1992). *Quantitative fisheries stock assessment: Choice, dynamics and uncertainty*. Chapman & Hall, New York. 570 pp.

Hoenig, J. M., Then, A. Y.-H., Babcock, E. A., Hall, N. G., Hewitt, D. A., & Hesp, S. A. (2016). The logic of comparative life history studies for estimating key parameters, with a focus on natural mortality rate. *ICES Journal of Marine Science, 73*, 2453-2467.

Hordyk, A., Ono, K., Sainsbury, K., Loneragan, N., & Prince, J. (2015). Some explorations of the life history ratios to describe length composition, spawning-per-recruit, and the spawning potential ratio. *ICES Journal of Marine Science, 72*, 204-216.

Hordyk, A., Ono, K., Valencia, S., Loneragan, N., & Prince, J. (2015). A novel length-based empirical estimation method of spawning potential ratio (SPR), and tests of its

performance, for small-scale, data-poor fisheries. *ICES Journal of Marine Science, 72*, 217-231.

Hordyk, A. R., Ono, K., Prince, J. D., & Walters, C. J. (2016). A simple length-structured model based on life history ratios and incorporating size-dependent selectivity: application to spawning potential ratios for data-poor stocks. *Canadian Journal of Fisheries and Aquatic Science, 73*, 1787-1799.

Huynh, Q. C. (2016). Estimating total mortality rates and calculating overfishing limits from length observations for six U.S. Caribbean stocks. SEDAR46-RW-01. SEDAR, North Charleston, SC. 19 pp.

Huynh, Q. C., Beckensteiner, J., Carleton, L. M., Marcek, B. J., Nepal KC, V., Peterson, C. P., Wood, M. A., & Hoenig, J. M. In review. Comparative performance of three length-based mortality estimators. *Marine and Coastal Fisheries*.

Huynh, Q. C., Gedamke, T., Hoenig, J. M., & Porch, C. (2017). Multispecies extension to a non-equilibrium length-based mortality estimator. *Marine and Coastal Fisheries, 9*, 68-78.

Huynh, Q. C., Gedamke, T., Porch, C. E., Hoenig, J. M., Walter, J. F., Bryan, M., & Brodziak, J. (2017). Estimating total mortality rates of mutton snapper from mean lengths and aggregate catch rates in a non-equilibrium situation. *Transactions of the American Fisheries Society, 146*, 803-815.

ICES (2015). Report of the Fifth Workshop on the Development of Quantitative Assessment Methodologies based on Life-history Traits, Exploitation Characteristics and other Relevant Parameters for Data-limited Stocks (WKLIFE V), 5–9 October 2015, Lisbon, Portugal. ICES CM 2015/ACOM:56. 157 pp.

ICES (2016). Report of the Workshop to consider MSY proxies for stocks in ICES category 3 and 4 stocks in Western Waters (WKProxy), 3–6 November 2015, ICES Headquarters, Copenhagen. ICES CM 2015/ACOM:61. 183 pp.

Jardim, E., Azevedo, M., & Brites, N. M. (2015). Harvest control rules for data limited stocks using length-based reference points and survey biomass indices. *Fisheries Research, 171*, 12-19.

Kokkalis, A., Thygesen, U. H., Nielsen, A., & Andersen, K. H. (2015). Limits to the reliability of size-based fishing status estimation for data-poor stocks. *Fisheries Research, 171*, 4-11.

Kokkalis, A., Eikeset, A. M., Thygesen, U. H., Steingrund, P., & Andersen, K. H. (2016). Estimating uncertainty of data limited stock assessments. *ICES Journal of Marine Science, 74*, 69-77.

Mildenberger, T. K., Taylor, M. H., & Wolff, M. (2017). TropFishR: an R package for fisheries analysis with length-frequency data. *Methods in Ecology and Evolution*. doi:10.1111/2041-210X.12791

Nadon, M. O. (2017). Stock assessment of the coral reef fishes of Hawaii, 2016. U.S. Department of Commerce, NOAA Technical Memorandum, NOAA-TM-NMFS-PIFSC-60, 212 p.

Nadon, M. O., & Ault, J. S. (2016). A stepwise stochastic simulation approach to estimate life history parameters for data-poor fisheries. *Canadian Journal of Fisheries and Aquatic Science, 73*, 1874-1884.

Nadon, M. O., Ault, J. S., Williams, I. D., Smith, S. G., & DiNardo, G. T. (2015). Length-Based Assessment of Coral Reef Fish Populations in the Main and Northwestern Hawaiian Islands. *PLoS ONE, 10(8)*, e0133960. doi:10.1371/journal.pone.0133960

Newman, D., Berkson, J., & Suatoni, L. (2015). Current methods for setting catch limits for data-limited fish stocks in the United States. *Fisheries Research, 164*, 86-93.

Pauly, D. (1983). Length-converted catch curves: A powerful tool for fisheries research in the tropics (Part I). *Fishbyte, 1(2)*, 9-13.

Pauly, D. (1984). *Fish population dynamics in tropical waters: a manual for use with programmable calculators*. International Center for Living Aquatic Resources Management, Manila.

Pauly, D. (1987). A review of the ELEFAN system for analysis of length-frequency data in fish and aquatic invertebrates. In D. Pauly, & G. R. Morgan (Eds). *Length-based methods in fisheries research.* (pp. 7-34). International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.

Prince, J., Hordyk, A., Valencia, S. R., Loneragn, N., & Sainsbury, K. (2015). Revisiting the concept of Beverton–Holt life-history invariants with the aim of informing data-poor fisheries assessment. *ICES Journal of Marine Science, 72*, 194-203.

Prince, J., Victor, S., Kloulchad, V., & Hordyk, A. (2015). Length based SPR assessment of eleven Indo-Pacific coral reef fish populations in Palau. *Fisheries Research, 171*, 42-58.

Punt, A. E., Campbell, R. A. & Smith, A. D. M. (2001). Evaluating empirical indicators and reference points for fisheries management: application to the broadbill swordfish fishery off eastern Australia. *Marine and Freshwater Research, 52*, 819-832.

Punt, A. E., Huang, T., & Maunder, M. N. (2013). Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusk species. *ICES Journal of Marine Science, 70*, 16-33.

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

Rudd, M., & Thorson, J. T. (In press). Accounting for variable recruitment and fishing mortality in length-based stock assessments for data-limited fisheries. *Canadian Journal of Fisheries and Aquatic Science.*

Sampson, D. B. (2014). Fishery selection and its relevance to stock assessment and fishery management. *Fisheries Research, 158*, 5-14.

SEDAR (2007). SEDAR 14: Caribbean Mutton Snapper stock assessment report. SEDAR, North Charleston, South Carolina.

SEDAR (2011a). SEDAR 26: U.S. Caribbean Redtail Parrotfish stock assessment report. SEDAR, North Charleston, South Carolina.

SEDAR (2011b). SEDAR 26: U.S. Caribbean Queen Snapper stock assessment report. SEDAR, North Charleston, South Carolina.

SEDAR (2011c). SEDAR 26: U.S. Caribbean Silk Snapper stock assessment report. SEDAR, North Charleston, South Carolina.

SEDAR (2013a). SEDAR 30: U.S. Caribbean Queen Triggerfish stock assessment report. SEDAR, North Charleston, South Carolina.

SEDAR (2013b). SEDAR 30: U.S. Caribbean Blue Tang stock assessment report. SEDAR, North Charleston, South Carolina.

SEDAR (2014). SEDAR 35: U.S. Caribbean Red Hind stock assessment report. SEDAR, North Charleston, South Carolina.

Sparre, P., & Venema, S. C. (1998). *Introduction to tropical fish stock assessment. Part 1. Manual*. FAO Fisheries Technical Paper 306.1 Rev. 2. Rome, FAO.

Taylor, M. H., & Mildenberger, T. K. (2017). Extending electronic length frequency analysis in R. *Fisheries Management and Ecology, 24*, 330-338.

Then, A. Y., Hoenig, J. M., Gedamke, T., & Ault, J. S. (2015). Comparison of Two Length-Based Estimators of Total Mortality: a Simulation Approach. *Transactions of the American Fisheries Society, 144*, 1206-1219.

Then, A. Y., Hoenig, J. M., Hall, N. G., Hewitt, D. A. (2015). Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. *ICES Journal of Marine Science, 72*, 82-92.

Then, A. Y., Hoenig, J. M., & Huynh, Q. C. (In press). Estimating fishing and natural mortality rates, and catchability coefficient, from a series of observations on mean length and fishing effort. *ICES Journal of Marine Science.*

Zhou, S., Yin, S., Thorson, J. T., Smith, A. D. M., & Fuller, M. (2012). Linking fishing mortality reference points to life history traits: an empirical study. *Canadian Journal of Fisheries and Aquatic Science, 69*, 1292-1301.

1.8. Tables

Table 1.1. Summary of the data-limited, size-based mortality estimators. Descriptions of the methods are provided in Section 1.3.

Method	Data	Required life history parameters	Estimates	Other Assumptions
Mean length-based				
BHE	Single mean length above L_c	L_∞, K	Constant Z with time and age	Constant recruitment; deterministic size-at-age; knife-edge selectivity at L_c
Gedamke-Hoenig (GH, including extensions with indices)	Multiple years of mean length (above L_c); extensions require either index of recruitment[1] or index of abundance[2]	L_∞, K, b[3]	Period-specific Z, constant with age; Time bounds of the periods	Constant recruitment (without index of recruitment); deterministic size-at-age; knife-edge selectivity at L_c
GH with effort	Multiple years of mean length (above L_c) and fishing effort	L_∞, K, M[4]	Year-specific estimates of F[5], constant with age	Constant recruitment; deterministic size-at-age; knife-edge selectivity at L_c
Composition-based				
LCCC	Single, truncated length composition of fully selected animals	L_∞, K	Constant Z with time and age	Constant recruitment; deterministic size-at-age; included lengths are fully selected
LB-SPR	Single length composition	M/K, L_∞, CV of length-at-age	Constant F/M with time, varies with length by estimated selectivity	Constant recruitment; logistic selectivity
S6	Single weight composition	M/K	Constant F, varies with weight by estimated selectivity; W_∞	Constant recruitment; logistic selectivity
LIME	Multiple years of length composition; index of abundance is optional[6]	L_∞, K, M, CV of length-at-age	Random effects estimation of year-specific recruitment and F; F varies with length by estimated selectivity	Logistic selectivity

Symbols: L_∞ = von Bertalanffy asymptotic length, K = von Bertalanffy growth coefficient, b = length-weight allometric exponent, W_∞ = asymptotic weight, Z = total mortality, M = natural mortality, F = fishing mortality

[1] Index of recruitment is numbers-based.

[2] Index of abundance can be either weight or numbers-based.

[3] b is only required when a weight-based index of abundance is used.

[4] M can be fixed or estimated in the model.

[5] Year-specific F is the product of the effort and the estimated catchability coefficient q.

[6] Index of abundance is weight-based.

Table 1.2. How assumptions of size-based mortality estimators are addressed.

Assumptions (Models)	**Violation**	**Diagnostic**	**Impact of violation**	**Options to address violation of assumption**
Life history parameters perfectly known (all)	Parameters not known well	For possible overestimate in L_∞, trend in Z estimates with increasing L_c in BHE	Potential bias of mortality estimates (e.g., overestimate mortality with L_∞ overestimate)	Estimate mortality trends rather than magnitude; sensitivity analysis of alternative values of parameters
Deterministic size-at-age (BHE, GH, LCCC)	Variability in size-at-age	N/A	Mean length methods and LCCC are robust to this assumption	Proceed with mean length methods or LCCC; use LB-SPR or LIME which model variability in size-at-age
Length of knife-edge selectivity is known (BHE, GH, & LCCC)	Logistic selection (fishing mortality at lengths < L_c is ignored)	N/A	Underestimate Z if incompletely selected lengths are not truncated	Use modal length for L_c for high M/K stocks, smaller values for low M/K; use models (LB-SPR, S6, LIME) which estimate logistic selectivity
Flat-top selectivity, either knife-edge or logistic (all)	Dome-shaped selectivity	Increasing trend in Z estimates from BHE with increasing L_c	Overestimate mortality	Use LCCC and omit length bins affected by dome selectivity; incorporate selectivity estimates into various methods
Constant recruitment (all except LIME)	Trend in recruitment, large year-class strength	Opposite trends in mean length and index over time; strong cohorts in length compositions over time	Increase in recruitment overestimates mortality, and vice versa	Use GH with index of recruitment; use LIME to estimate recruitment deviations (needs informative data)
Constant mortality (all except GH and LIME)	Trend in mortality over time	Mean length changes over time	Lag, underestimate magnitude of change in mortality in equilibrium estimators	Use models (GH, LIME) which incorporate multiple years of data

Table 1.3. Applications of and software packages for the size-based mortality estimators.

Method	Applications	Software Packages
Mean length-based		
Beverton-Holt (including variants)	Ault, *et al*. (2005); Babcock, *et al.* (2013); Nadon, *et al*. (2015); more cited in Then, Hoenig, Gedamke, *et al*. (2015)	fishmethods[1], TropFishR[1,2]
Gedamke-Hoenig (GH, including extensions with indices)	SEDAR (2007); SEDAR (2011a); SEDAR (2011b); SEDAR (2011c); SEDAR (2013a); SEDAR (2013b); SEDAR (2014); ICES (2015); Huynh (2016)	fishmethods, SEINE[3], MLZ[4]
GH with effort	ICES (2015); Then, *et al.* (In press)	MLZ
Composition-based		
LCCC	Many, some cited in Huynh *et al*. (in review)	FiSAT[5], TropFishR
LB-SPR	Prince, Victor, *et al*. (2015); ICES (2015)	LBSPR[1]
S6	Kokkalis, *et al*. (2016); ICES (2015)	s6model[6]
LIME	--	LIME[7]

[1] R package available on CRAN
[2] Mildenberger, Taylor, & Wolff (2017)
[3] Stand-alone program (without extensions for index of recruitment or abundance) available at: http://www.nft.nefsc.noaa.gov
[4] R package available on Github at: http://www.github.com/quang-huynh/MLZ
[5] Stand-alone program available at: http://www.fao.org/fishery/topic/16072/en (Gayanilo, Jr, *et al*., 2005)
[6] R package available on Github at: http://www.github.com/alko989/s6model
[7] R package available on Github at: http://www.github.com/merrillrudd/LIME

1.9. Figures

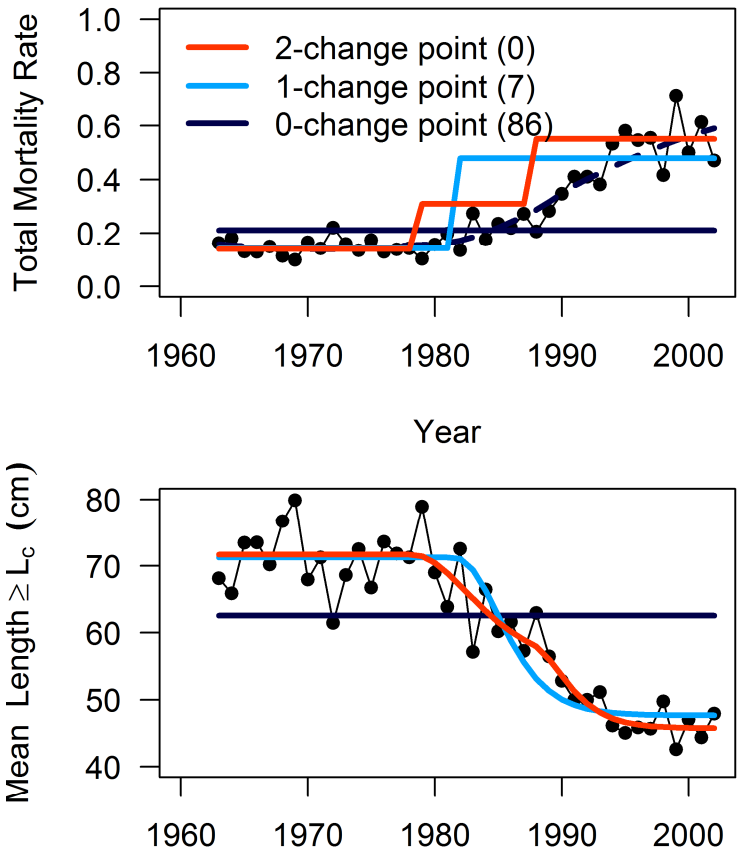

Figure 1.1. Application of the mean length-based mortality estimators for the Northern management (New England) stock of goosefish (*Lophius americanus*). *Top figure*: estimates of instantaneous total mortality Z (year^{-1}) from successive fits of Gedamke-Hoenig with differing number of change points (colored lines) and independent year-specific estimates from the BHE (points with dotted loess regression line). Parentheses in legend indicate ΔAIC values for different change points with the Gedamke-Hoenig models. Compared to Gedamke-Hoenig, the BHE will underestimates the magnitude of change in the mortality rate until a new equilibrium mean length is reached. *Bottom figure*: observed mean lengths (points) and predicted mean lengths from the Gedamke-Hoenig models (colored lines). Gedamke-Hoenig allows for evaluation of goodness of fit to the mean length data. With the 1-change point model, the model the mean length is underestimated during 1987-1993 and generally overestimated from 1994-2001. This trend in residuals is removed with a 2-change point model, which is supported with AIC. Data and life history parameters were obtained from Gedamke & Hoenig (2006).

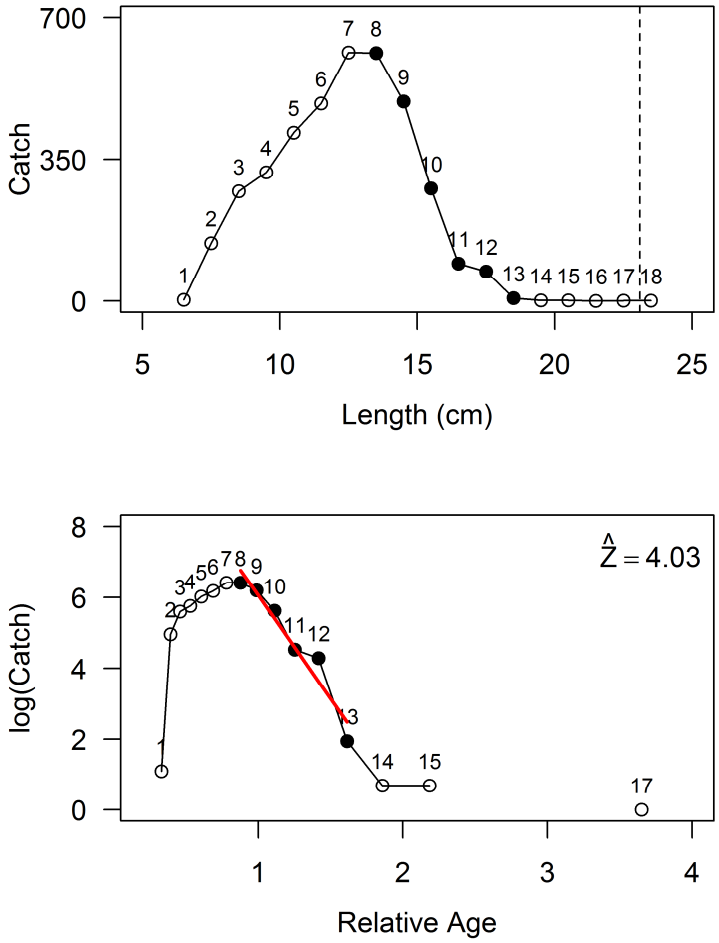

Figure 1.2. Application of the LCCC to yellow-striped goatfish (*Upeneus vittatus*) in Manila Bay, Philippines. *Top figure*: the observed length composition (with 1-cm length bins), the vertical dotted line marks L_∞. *Bottom figure*: conversion of lengths to relative ages and linear regression (red line) on relative ages to estimate total mortality Z (Equation 1.4). Numbers above points index length bins; relative ages could not be calculated because there was zero catch in length bin #16 and the length bin #18 was larger than L_∞. Solid points indicate the length bins used in the LCCC. Open points indicate truncated length bins because they are assumed to be incompletely selected (bins 1-7). Length bins close to L_∞ (bins 14-18) were also truncated because the log-linear relationship between relative age and the catch breaks down at lengths near L_∞ due to the effect of (1) dome selectivity, (2) outlier observations relative to other length bins, or (3) significant overlap of multiple ages. Data and life history parameters were obtained from Sparre & Venema (1998) through the TropFishR software package (Mildenberger, Taylor, & Wolff, 2017).

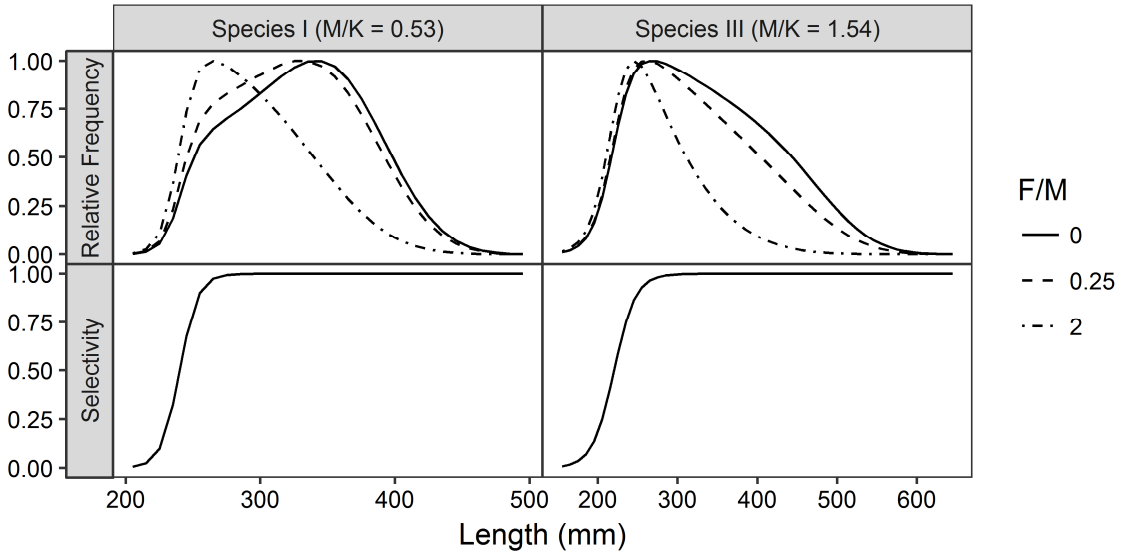

Figure 1.3. Length distributions in different *F/M* scenarios for two species which vary in *M/K*. With increasing *F/M*, the shape of the length distribution changes and there is truncation with reduced abundance in large size classes. In the low *M/K* species (Species I), the modal length is much larger than the first fully selected length when *F/M* is low. As *F/M* increases, the mode moves towards the left. In the high *M/K* species (Species III), the mode of the distribution appears to be more stable and the ascending limb of the length distribution reflects selectivity (regardless of *F/M*). In low *M/K* scenarios, more caution is needed when using the mode as the L_c for the mean length-based methods, a length smaller than the mode will be more appropriate. In high *M/K* scenarios, the mode can be used more reliably as the L_c. The age-structured LB-SPR was the operating model, with life history parameters for Species I and Species III obtained from Hordyk, Ono, Valencia, *et al*. (2015).

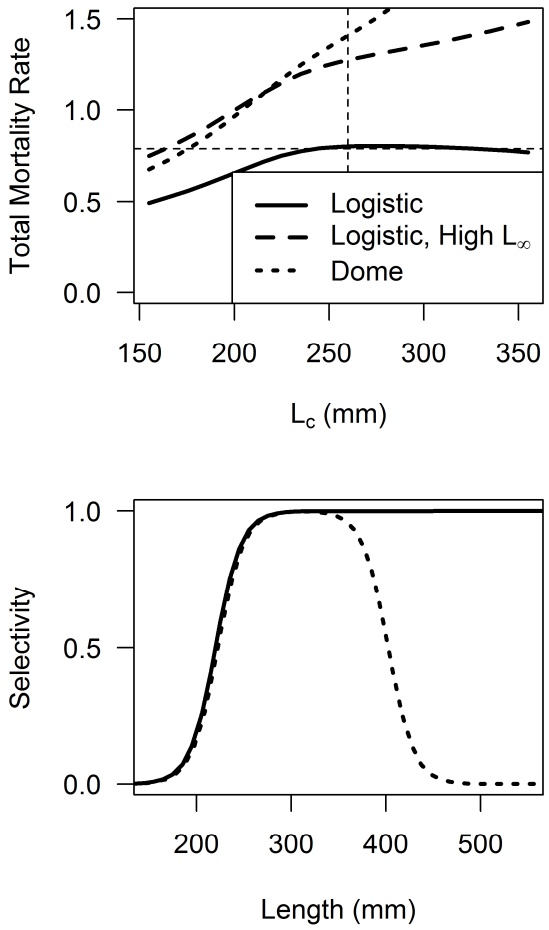

Figure 1.4. Diagnostic of dome selectivity from the BHE. *Top figure*: Estimates of total mortality Z from the BHE based on increasing values of L_c. Horizontal, dashed line indicates the true $Z = 0.79$ ($F/M = 0.25$) and vertical, dashed line indicates the length of 95% selectivity. Three scenarios are evaluated: L_∞ is known perfectly with logistic selectivity in the length composition (Logistic), a 20% overestimate of L_∞ is used in the BHE with logistic selectivity (Logistic, High L_∞), and L_∞ is known perfectly with dome selectivity (Dome). *Bottom figure*: Logistic (solid line) and dome (dashed line) selectivity. In the Logistic scenario, Z estimates from the BHE plateau when lengths that are near or above the length of full selectivity are chosen as the L_c. The increasing trend in estimates of Z in the High L_∞ and Dome scenarios could be used to evaluate whether there is either dome selectivity or an overestimate of L_∞ is occurring, although these two causes may not be differentiable. Length compositions were generated from the age-based LB-SPR model with the Species III life history with $M/K = 1.54$ (Hordyk, Ono, Valencia, *et al*., 2015).

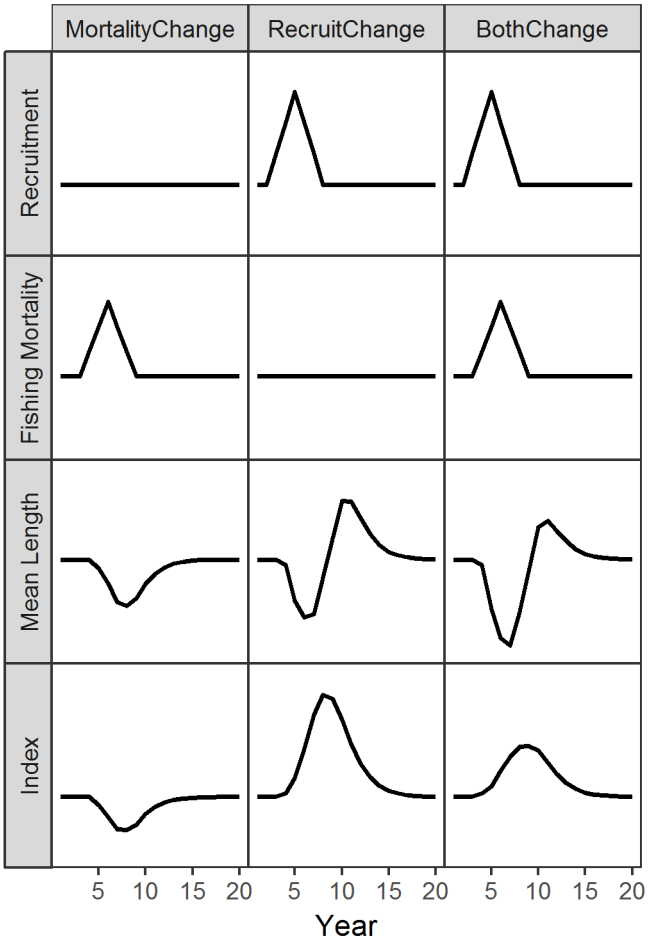

Figure 1.5. Diagnostic of a recruitment trend over time based on the response in mean length and index of abundance when there is a change in mortality (MortalityChange), recruitment (RecruitChange), or both (BothChange). With a change in mortality, both the mean length and index change in the same direction. With a change in recruitment, the mean length and index change in different directions. The trend in the mean length and index is suggested as a diagnostic for evaluating whether the changes in mortality versus recruitment can be identified. LIME was the operating model for data generation, and life history parameters from the Medium life history type from Rudd & Thorson (in press) were used.

Chapter 2: Comparative performance of three length-based mortality estimators

2.1. Abstract

Length-based methods provide alternatives for estimating the instantaneous total mortality rate (Z) in exploited marine populations when data are not available for age-based methods. We compared the performance of three equilibrium length-based methods: the length-converted catch curve (LCCC), the Beverton-Holt equation (BHE), and Length-Based Spawning Potential Ratio (LB-SPR) method. The LCCC and BHE are two historically common procedures that use length as a proxy for age. From a truncated length-frequency distribution of fully selected animals, the LCCC estimates Z with a regression of the logarithm of catch-at-length by the midpoint of the length bins, while the BHE estimates Z as a function of the mean length. The LB-SPR method is a likelihood-based population dynamics model, which, unlike the LCCC and BHE, does not require data truncation. Using Monte Carlo simulations across a range of scenarios with varying mortality and life history characteristics, our study showed that neither the LCCC nor the BHE was uniformly superior in terms of bias or root mean square error across simulations, but these estimators performed better than LB-SPR which had the largest bias in most cases. Generally, if M/K (the ratio of natural mortality M to von Bertalanffy K) is low, then the BHE is most preferred, although there is likely to be high bias and low precision. If M/K is high, then the LCCC and BHE performed better and similarly to each other. Differences in performance among commonly-used truncation

methods for the LCCC and BHE were small. LB-SPR did not perform as well as the classical methods, but may still be of interest because it provides estimates of a logistic selectivity curve. The *M/K* ratio provided the most contrast in the performance of the three methods, suggesting that it be considered for predicting the likely performance of length-based mortality estimators.

2.2. Introduction

Length-based methods for assessing exploited marine populations are of significant interest largely because of their applicability to the study of data-limited stocks for which age-based methods may not be available or suitable (Punt et al. 2013). Hard tissues, such as scales and otoliths, may lack distinct growth marks, for example, in tropical fish species. Such species are often assessed using length-based methods (Pauly 1984c), because length measurements are collected both easily and non-lethally.

Historically, the most common methods used to estimate the total instantaneous mortality rate, Z (year^{-1}), from length composition data are the length-converted catch curve (LCCC; Pauly 1983, 1984a, 1984b) and the Beverton-Holt equation (BHE; Beverton and Holt 1956, 1957). These methods are based on a linear regression and a moment estimator, respectively. Improvements in computational power over time have allowed for the development and use of nonlinear models that use derivative-based optimization methods. Recently, Hordyk et al. (2015b) developed the Length-Based Spawner Potential Ratio (LB-SPR) method to estimate mortality using a nonlinear model.

An analogue of the age-based catch curve (Ricker 1975), the LCCC uses the natural logarithm of catch (C_j) in the *j*-th length interval of a length-frequency distribution (LFD) regressed on the relative age (t'_j) at the midpoint of the length bin (ℓ_j). Only fully-selected lengths are considered in the analysis. Under the assumption of deterministic growth following a von Bertalanffy function (with parameters constant across time and cohorts), the relative age at the *j*-th length bin is defined as:

$$t'_j = -\log\left(1 - \frac{\ell_j}{L_\infty}\right), \tag{2.1}$$

where L_∞ is the asymptotic maximum length from the von Bertalanffy growth function. The regression is of the form

$$\log(C_j) = a + bt'_j + \varepsilon_j, \tag{2.2}$$

where *a* and *b* are the intercept and slope, respectively, of the linear regression and ε_j is the normally-distributed residual error. Total mortality ($\hat{Z}$) is estimated using the estimated slope of the linear regression ($\hat{b}$) and the von Bertalanffy growth rate parameter (*K*):

$$\hat{Z} = K(1 - \hat{b}), \tag{2.3}$$

where the circumflex (^) denotes an estimate. The slope is positive if $Z/K < 1$ and negative if $Z/K > 1$. The derivation for the LCCC is provided in Appendix A.

Similar to the age-based catch curve, the LCCC assumes a steady state population, with constant total mortality (over age and time) and constant recruitment (Pauly 1984a). Additionally, all selected fish assumed to be equally vulnerable to the sampling gear, and the sample size is large enough to effectively represent the average

population structure over the time period considered (Pauly 1983, 1984a, 1984b). Length-converted catch curves have been criticized for overestimating *Z* when individual growth varies seasonally (Isaac 1990; Sparre 1990). However, this bias has been overcome by modified LCCCs that accommodate seasonally varying growth (Pauly 1990). Simulations have also shown that individual growth variability creates a negative bias, while reduced size-selectivity for smaller sizes produces a positive bias for the LCCC (Isaac 1990). Analogous to the age-based catch curve, a bend in the regression line could be an indication of a change in mortality with time or with age (Pauly 1984c; Tuckey et al. 2006). However, this method does not generalize easily to account for non-equilibrium conditions.

Beverton and Holt (1956, 1957) derived the total mortality rate as a function of the observed mean length:

$$\hat{Z} = \frac{K(L_\infty - \bar{L})}{\bar{L} - L_c}, \tag{2.4}$$

where $\hat{z}$, *K*, and L_∞ are as in Equations 2.1 and 2.3, L_c is the critical length above which all animals are fully selected by the fishery, and $\bar{L}$ is the mean length of animals larger than L_c. Gedamke and Hoenig (2006) provide a recent derivation of the BHE. Similar to the LCCC, the BHE also assumes steady-state conditions, deterministic von Bertalanffy-type growth, constant mortality rate of all fully recruited fish, and continuous and constant recruitment to the fishery.

A criticism of the BHE is that it tends to overestimate total mortality when the largest size classes in the population are truncated from the sample (Then et al. 2015). On the other hand, Laurec and Mesnil (1987) and Then et al. (2015) observed that the BHE is

generally robust to individual variability in growth. The BHE has also been criticized as overly simplistic because of its stringent equilibrium assumptions (Hilborn and Walters 1992). Gedamke and Hoenig (2006) generalized the BHE to allow for the estimation of total mortality from a time series of mean lengths under non-equilibrium conditions in a maximum likelihood framework. Further extension of this non-equilibrium model has allowed for variable recruitment by incorporating a year-specific index of recruits into the model (Gedamke et al. 2008).

The LB-SPR mortality estimator is an equilibrium age-structured model which converts the predicted age distribution of the catch to a length distribution. Unlike the LCCC and BHE, variability in growth is explicitly modeled with a coefficient of variation (CV) of length-at-age generally assumed to be 0.1 (Hordyk et al. 2015b). Ages are converted to lengths via an age-length transition matrix in which the probability of length-at-age sums to one for a given age. Logistic selectivity parameters are estimated concurrently with mortality, allowing for the use of the entire LFD in the likelihood function. The method also assumes constant recruitment. Unlike the LCCC and BHE, the LB-SPR method explicitly pairs the mortality estimator with the biological reference points obtained from spawner potential ratio analyses for management. The same can be done for the LCCC and BHE, although this was not the focus of the current study.

The LCCC has been implemented in ELEFAN II (Pauly 1987; Isaac 1990). Currently, the LCCC can be applied using the FAO-ICLARM Stock Assessment Tools (FiSAT) software package (Gayanilo et al. 2005). Recently, it has been used to estimate mortality in reef fishes in North Carolina, United States (Rudershausen et al. 2008), albacore tuna in the Mediterranean (Anonymous 2012), Japanese threadfin bream in the

Indian Ocean (Kalhoro et al. 2014), red king crab in the Barents Sea (Windsland 2015), wahoo in the Southwest Pacific Ocean (Zischke and Griffiths 2015), blood cockle in Malaysia (Mirzaei et al. 2015), and red lionfish in the Gulf of Mexico (Rodriguez-Cortes et al. 2015). Recent applications of the BHE and LB-SPR are cited in Then et al. (2015) and presented in Prince et al. (2015b), respectively. The LB-SPR method can be implemented using the LBSPR R package available on CRAN (Hordyk 2017).

The three length-based methods have been studied individually, but they have not been directly compared. Importantly, the methods differ in handling selectivity. While LB-SPR estimates selectivity as a logistic function, the LCCC and BHE assume knife-edge selectivity (i.e., full selectivity of animals greater than a certain length) and thus, only animals larger than a certain size are included in the analysis. Previous simulations (e.g., Isaac 1990) have not examined the effect of different decision rules for truncating the data on the performance of the estimators. This study compares the performance of the length-based methods in estimating total mortality by applying these methods to populations with known parameters. First, we examined the performance of each of the estimators individually and relative to each other using a common simulation framework. Second, we examined the choice of decision rules in selecting the truncation points for the LCCC and choosing the L_c parameter for the BHE. Third, we examined the robustness of each method to violations in the assumptions of growth and recruitment variability across several life histories and exploitation scenarios. Finally, we conducted sensitivity analyses of the three mortality estimators to total sample size and length bin width.

2.3. Methods

2.3.1. Simulation design

Length samples were generated using a factorial design for von Bertalanffy *K*, fishing mortality *F*, growth variability, recruitment variability, and selectivity (Table 2.1), accumulating to a total of 108 combinations. In this study, the different values of *K* and *F* are presented as ratios with respect to natural mortality, *M* (i.e., *M/K* and *F/M*, respectively), with $M = 0.2$ year^{-1} for all scenarios. Ratios used because the relative values provide a better description of the life history and magnitude of exploitation, respectively, than the absolute values. On a per-recruit basis, the *M/K* ratio describes the balance between growth and mortality which affects the shape of the LFD of a population in an unexploited state (Hordyk et al. 2015a), while *Z/K* describes the shape of the LFD of an exploited population. The *F/M* ratio can provide an indication of the relative impact of fishing pressure because a scalar multiple of *M* is often used as a proxy for fishing at maximum sustainable yield (e.g., $F_{MSY} = 0.75M$; Zhou et al. 2012).

The simulation used an age-structured model for the population. The model was run for 25 years to burn in deviates in growth trajectories and recruitment strength among cohorts. Fishing was assumed to occur throughout the 25 years, but the length distribution of the catch was only obtained at the end of the 25 years. Growth was assumed to vary among cohorts (Whitten et al. 2013). The mean length ($L_{y,a}$) in year *y* at age *a* was

$$L_{y,a} = \begin{cases} L_0 \exp(\nu_c) & a = 0 \\ L_{y-1,a-1} + (L_{y-1,a-1} - L_\infty)\{\exp(-K) - 1\}\exp(\nu_c) & a = 1,2,...,A \end{cases}, \qquad (2.5)$$

where L_0 was the expected length at age-0, $c = y - a$ indexed the age-0 recruitment that gave rise to the cohort of age *a* in year *y*, and $\nu_c \sim N\left(-0.5\sigma_C^2, \sigma_C^2\right)$ was the cohort-specific deviation in growth increments. The von Bertalanffy asymptotic length L_∞ was set to 500 (arbitrary) units, L_0 to 75 units, and σ_C to 0.15 across all factorials. Two values of K = 0.4 and 0.1 year^{-1}, corresponding to *M/K* ratios of 0.5 and 2.0, respectively, were used in the factorial design.

Variability in length-at-age ($\sigma_{y,a,L}$) assumed a constant CV to the mean length-at-age,

$$\sigma_{y,a,L} = CV \times L_{y,a}, \tag{2.6}$$

with three CVs of 0.03, 0.06, and 0.09 in the factorial design. These values were based on the evaluation of size-at-age data of several species by Then et al. (2015).

Lognormally-distributed recruitment (R_y) was simulated as a first-order autoregressive process with autocorrelation coefficient (ρ) and residual deviations (δ_y) (Thorson et al. 2014):

$$\log(R_y) = \begin{cases} \delta_y & y = 1 \\ \rho \log(R_{y-1}) + \sqrt{1-\rho^2}\,\delta_y & y = 2,...,Y \end{cases}, \tag{2.7}$$

where $\delta_y \sim N\left(-0.5\sigma_R^2 \times \frac{1-\rho}{\sqrt{1-\rho^2}}, \sigma_R^2\right)$ (Thorson et al. 2016) and *Y* was the terminal year of the age-structured model. Two levels of residual standard deviation for recruitment σ_R = 0.6 and 1.0 were included in the factorial design, with ρ = 0.45. The values of these parameters were guided by the meta-analysis of Thorson et al. (2014). Mean

recruitment was set to 1.0 because the population was stationary over time and magnitude of recruitment was not relevant for estimating mortality.

Three length-based selectivity patterns were used in the factorial design: a logistic function with either broad ascending limb ("Gradual"), a logistic function with a steep ascending limb ("Steep"), and a dome-shaped logistic-normal function ("Dome") (Figure 2.1). The logistic function, parameterized by the lengths of 50% (L_{50}) and 95% (L_{95}) selectivity, defines selectivity at length *L* as

$$sel_L = \left(1+\exp\left[-\log_e(19)\frac{L-L_{50}}{L_{95}-L_{50}}\right]\right)^{-1}. \tag{2.8}$$

The dome-shaped selectivity function was defined as a piecewise-defined function,

$$sel_L = \begin{cases} \left(1+\exp\left[-\log_e(19)\frac{L-L_{50}}{L_{95}-L_{50}}\right]\right)^{-1} & L < \mu_d \\ g(L;\mu_d,\sigma_d)/\max\left[g(L;\mu_d,\sigma_d)\right] & L \geq \mu_d \end{cases}, \tag{2.9}$$

where selectivity at length *L* is a logistic function for the ascending limb and the right-half of a normal probability density function $g(L)$ with mean μ_d and standard deviation σ_d for the descending limb, with the latter standardized to a value of 1 at μ_d. The Steep and Gradual selectivity functions evaluated the effect of logistic selectivity on data truncation with the LCCC and BHE while the Dome selectivity function tested the effect of violating the assumption of constant total mortality of fully selected lengths in all three mortality estimators.

The population abundance ($N_{y,a}$) was defined by

$$N_{y,a} = \begin{cases} R_y & a = 0 \\ s_{y-1,a-1}N_{y-1,a-1} & a = 1,2,.,A \end{cases}, \tag{2.10}$$

where $s_{y,a}$ is the survival and A is the maximum age in the model. A maximum age of 23 years, the age when 1% of a cohort survives given the natural mortality rate, was used in the simulation for computational convenience.

To calculate survival due to length-based selectivity, a population length-age matrix ($N_{y,a,\ell}$) was created for the beginning of each year *y*, where

$$N_{y,a,\ell} = N_{y,a} P(\ell | y,a). \tag{2.11}$$

With a normal distribution for variability in length-at-age, the length-at-age probability vector $P(\ell | y,a)$ is

$$P(\ell | y,a) = \begin{cases} \phi\left(\ell'_{j+1}\right) & j = 1 \\ \phi\left(\ell'_{j+1}\right) - \phi\left(\ell'_{j}\right) & j = 2,...,J-1, \\ 1 - \phi\left(\ell'_{j}\right) & j = J \end{cases} \tag{2.12}$$

where ℓ' is the length at the lower boundary of the length bin with midpoint ℓ, $j = 1,2,...,J$ indexes the length bins and $\phi(\cdot)$ is the cumulative density function of a normal distribution with mean $L_{y,a}$ and standard deviation $\sigma_{y,a,L}$. The length bin width in the population model was 5 units (larger bins were subsequently used for mortality estimation).

The abundance of survivors ($N^{s}_{y,a,\ell}$) at the end of year *y* was calculated as

$$N^{s}_{y,a,\ell} = N_{y,a,\ell} \exp\left[-(sel_{\ell} F + M)\right]. \tag{2.13}$$

This study used three values of apical fishing mortality F = 0.05, 0.2, and 1.0 year^{-1}, corresponding *F/M* ratios of 0.25, 1.0, and 5.0, respectively, with mortality occurring after growth. Survival, which is dependent on age and year due to cohort-specific growth, was calculated as

$$s_{y,a} = \frac{\sum_{\ell} N^{s}_{y,a,\ell}}{N_{y,a}}. \tag{2.14}$$

To approximate continuous recruitment assumed in the mortality estimators, quarterly time steps were used in the simulation. Recruitment occurred quarterly with all cohorts within a year having the same growth trajectory and recruitment strength. All rate parameters were adjusted accordingly from year^{-1} to season^{-1} with growth updated after every season.

In the terminal time step of the simulation, the length-age catch matrix ($C_{a,\ell}$) was created using the Baranov catch equation,

$$C_{a,\ell} = \frac{sel_{\ell} F}{sel_{\ell} F + M} N_{y,a,\ell} \left\{1 - \exp\left[-(sel_{\ell} F + M)\right]\right\}, \tag{2.15}$$

and the catch-at-length vector (C_{ℓ}) was obtained by summing over ages,

$$C_{\ell} = \sum_{a} C_{a,\ell}. \tag{2.16}$$

A data set was obtained by sampling 2,000 individuals from the terminal catch-at-length vector using a multinomial distribution. The sample size of 2,000 was chosen to evaluate the robustness of the estimators to the variables in the factorial design when there is little observation error. For each data set, 2,000 length observations were obtained and a length frequency histogram was generated by dividing the data set into length bins with a bin width of 10 units (2% of L_{∞}). For each factorial combination, 1,000 stochastic data sets were generated.

2.3.2. Mortality estimation

To use the LCCC, a subset of length bins from the LFD corresponding to fully selected lengths must be chosen for the linear regression. The LFD typically features an ascending limb, representing some lengths that may not be fully selected to the fishing year, followed by a descending limb of numbers-at-length (Figure 2.2). The first usable length bin may be defined as the peak of the LFD (hereafter, "Peak") (Wetherall et al. 1987), although Pauly (1983) suggested that the first size class to be included in the LCCC be the size class immediately to the right of the most frequent size class ("Peak-plus").

Because the LCCC assumes deterministic growth, length bins greater than L_∞ must be excluded from the analysis. Furthermore, length bins close to L_∞ may be assigned unreasonably large relative ages. High observation error in length bins with few observations may affect the slope of the regression line (Isaac 1990; Punt et al. 2013). To combat this, Pauly (1983) recommended that animals within 5 - 30% of L_∞ or length bins with fewer than five individuals be excluded from the analysis. Such approaches sacrifice data in an attempt to avoid bias due to decreased selectivity by the fishing gear and overestimation of relative age of large individuals. This approach can be problematic, however, when the sample size is low or when only few size classes are available.

For the BHE, the length data are usually binned to examine the length frequency distribution and identify the critical length, L_c. The mean length is then calculated from the subset of animals larger than L_c. Wetherall et al. (1987) suggested that L_c be defined as the length corresponding to the peak of the LFD. Alternatively, Peak-plus truncation can be applied to select a value for L_c. Consequently, length observations from the

ascending limb of the LFD are removed from the mean length calculation for the BHE based on the choice of L_c.

To reduce bias associated with the BHE, Laurec and Mesnil (1987) recommended summarizing length data in fine detail and grouping length frequencies in narrow size bins. Animals within 30% of L_∞ were excluded from their analyses. However, simulation analysis demonstrated that the BHE performed well when all lengths greater than L_c, including those larger than L_∞, were retained (Then et al. 2015).

In this study, three candidate length bins were selected for left truncation: the first length bin on the ascending limb of the LFD corresponding to at least half of the frequency of that at the peak ("Half-peak abundance"), the length bin of the peak ("Peak"), and the first length bin after the peak ("Peak-plus") (Figure 2.2). If there are few length bins larger than the peak, then a portion of the ascending limb of the LFD may consist of fully selected animals (Figure 7 of Hordyk et al. 2015a). While arbitrary, the Half-peak abundance decision rule can be used to select a length on the ascending limb relative to the Peak across a variety of shapes in the length distribution (Figures 2.2, 2.3). This decision rule has also been used in several applications of methods evaluating length data (for example, ICES 2014).

Similarly, there were three candidate length bins for right truncation: the largest length bin containing at least 5 individuals (5+), the length bin at 90% of L_∞ (90% L_∞), and the length bin at L_∞ (100% L_∞). The length bin at L_∞ was chosen if the 5+ right truncation rule selected a length bin with a midpoint larger than L_∞. For the 90% L_∞ and

100% L_∞ decision rules, if the bin contained no observations, then we selected the next smallest bin containing any observations as the truncation point.

Nine methods, labeled L1-L9, were tested with the LCCC using the combinations of left and right truncation (Table 2.2). For the BHE, the lower boundary of the three candidate length bins for left truncation (Half-peak abundance, Peak, Peak-plus) was identified as the L_c, comprising methods B1-B3. No truncation was necessary to use the LB-SPR method.

For each data set, total mortality was estimated with the data truncation methods described for LCCC, BHE, and LB-SPR. The values of von Bertalanffy parameters L_∞ and K used in the mortality estimators were sampled from a bivariate normal distribution around the true values with a CV of 0.1 and a correlation of -0.9. This step is designed to simulate the scenario in which only length data are available and growth information is obtained externally (e.g., via a literature search).

2.3.3. Performance analysis

To quantify the performance of the decision rules for the estimators in terms of bias and precision, the relative percent bias (%*Bias*) and relative percent root mean square error (%*RMSE*) for each decision rule in each factorial combination were calculated respectively as:

$$\%Bias = \frac{\bar{\hat{Z}} - Z}{Z} \times 100 \tag{17}$$

and

$$\%RMSE = \frac{1}{Z}\sqrt{\frac{\sum_i (\hat{Z}_i - Z)^2}{n}} \times 100, \quad (18)$$

where $\bar{\hat{Z}}$ is the mean of the estimated total mortality rates from n out of 1,000 data sets which produced a feasible estimate, $Z = F + M$ is the true underlying mortality rate for the factorial combination in the simulation, and $\hat{Z}_i$ is the estimated mortality rate from each data set $i = 1, 2, \ldots, n$ within each factorial combination. Unfeasible estimates occurred using the LCCC if only one length bin was selected using the respective decision rule, in which case the linear regression was not possible, or if the slope of the regression line in Equation 2.3 was greater than 1, which resulted in a negative estimate of Z. With the BHE, a negative total mortality rate was estimated if the mean length was larger than L_∞.

The %*Bias* and %*RMSE* were calculated for each method in all 108 factorials. From this set, the median %*Bias* and median %*RMSE* for each method were calculated among factorials with common *M/K* and *F/M* ratios. The median %*Bias* and median %*RMSE* were further stratified across levels of the other factorial variables (growth variability, recruitment variability, and selectivity) within each group of *M/K* and *F/M*. The best decision rules can be identified as those with the lowest absolute values of the median %*Bias* and the median %*RMSE*.

2.3.4. Sensitivity analyses

Sensitivity analyses were performed with respect to sample size, length bin, and assumed growth parameters. For the sample size analysis, 200 and 500 length

observations were selected without replacement as subsets of the original data sets, with bin size of 10 units in the LFD. This analysis allowed us to test the effect of observation (sampling) error on mortality estimation. For the bin width analysis, mortality was re-estimated by re-binning the data with bin widths of 25 and 50 units (5% and 10% of L_∞, respectively). Sample sizes of 2,000 were used to analyze the effect of bin width separately from observation error. For these two sensitivity analyses, mortality was estimated again using the same decision rules for data truncation and the %*Bias* and %*RMSE* were calculated for each decision rule in each factorial combination. Finally, the variability in individual estimates of mortality was also evaluated when assumed growth parameters were stochastically sampled. All simulations and analyses were performed in R version 3.3 (R Core Team 2017).

2.4. Results

Our factorial design generated several functionally distinct LFDs based on *M/K* and *F/M* (Figure 2.3). Compared to the Gradual and Dome selectivity functions, the Steep selectivity function produced a shorter ascending limb of the LFD, which truncated the length structure of the sample. In contrast, the Dome selectivity function only showed a discernable difference in the descending limb when *F/M* = 0.25 or 1.

Based on a sample size of 2,000, performance of the methods varied the most by *M/K* and *F/M* scenarios, with best performance of the methods when *M/K* = 2 in conjunction with *F/M* = 0.25 or 1 (Figure 2.4A). The methods have the least bias in these scenarios, with the magnitude of the median *%Bias* generally less than 20% and the *%RMSE* less than 50%. The ranges of the *%Bias* and *%RMSE* among factorials were also

relatively small in these scenarios. Most methods performed similarly, although LB-SPR did not perform as well as the LCCC (L1-L9) and BHE methods (B1-B3).

Performance was worst when $M/K = 0.5$ in conjunction with $F/M = 0.25$ or 1 (Figure 2.4A). While there were some factorial combinations where the methods produced low *%Bias* and *%RMSE*, the range in *%Bias* and *%RMSE* of all methods was large (with the performance metrics as high as 300-400%) indicating high variability in performance. In all cases, the bias was positive. The median *%Bias* and median *%RMSE* were usually larger than 100%. The best performing methods were B1 (BHE with Half-peak abundance as the L_c), closely followed by L1 and L3 (both L1 and L3 methods use the LCCC with Half-peak abundance for left truncation).

When $F/M = 5$, all methods improved in terms of bias for $M/K = 0.5$ (the magnitude of *%Bias* generally less than 20%), but worsened for $M/K = 2$ (the magnitude of *%Bias* increasing up to 40%), relative to lower F/M (Figure 2.4A). Overall, the sign of the bias trends from positive to negative with increasing F/M, with the trend most noticeable for $M/K = 0.5$. No single method appeared to perform the best when $F/M = 5$. While LB-SPR had the lowest bias with $M/K = 2$, it also had the highest mean square error. In other M/K and F/M scenarios, LB-SPR did not appear to perform as well as the LCCC and BHE.

2.4.1. Performance across factorial variables

In this and the next section, we present the results for B1 when $M/K = 0.5$ and L5 when $M/K = 2$ in the main text. Method B1 performed the best when $M/K = 0.5$ (and $F/M = 0.25$ or 1). Method L5 was chosen arbitrarily because there was no clear best method

when *M/K* = 2. The performance across factorial variables and sensitivity analyses for individual factorial combinations for all decision rules are described in the main text, with supporting figures and tables in the Supplementary Materials.

Within *M/K* and *F/M* combinations, the performance metrics were further stratified by growth variability type, magnitude of growth variability, recruitment variability, and selectivity. Observed trends in performance remained similar to those described in the previous section (Tables S1-S6).

Bias and precision generally improved with increasing growth variability when *F/M* = 0.25 or 1 (Figures 2.5, S1 –S13). When *F/M* = 5, the differences in bias and precision among different growth variabilities were small to negligible. Larger bias and mean square error was associated with high variability than with low variability in recruitment (Figures 2.6, S14-S26), although there were negligible differences when *F/M* = 5. All three methods were much more positively biased with Dome selectivity than with the logistic selectivities (Gradual and Steep) when *F/M* = 0.25 or 1 (Figures 2.7, S27-39). However, the effect of Dome selectivity is minimal at *F/M* = 5. There were no major differences in performance common to all methods between the Steep and Gradual selectivity functions.

2.4.2. Sensitivity analyses

At the sample size of 200, most methods had larger bias and less precision compared to when a sample size of 2,000 was used, but the magnitude of the difference between sample sizes was not particularly large (Figures 2.4B, S40-S52). Methods L4 and L7 were notable in that their median *%Bias* was lower, but median *%RMSE* was

higher when the sample size was 200 instead of 2,000. However, the general trends remained unchanged.

Length bin width generally did not affect the %*Bias* and %*RMSE* of the mortality estimators (Figures 2.8, S53-S65). Methods L7, L8, and L9 showed improvement in some scenarios when $M/K = 0.5$ in conjunction with $F/M = 0.25$ where larger length bins performed better, but these scenarios still appeared to be outliers. The magnitude of the performance metrics remained large (*%Bias* > 100 %) for these scenarios.

We examined the correlation of total mortality estimates with the assumed values of L_∞ and K (Figure 2.9). In general, higher estimates of mortality are obtained with a larger value of L_∞. However, underestimates of Z did not often occur with low M/K scenario and $F/M = 0.25$ or 1. On the other hand, when $F/M = 5$ (for both M/K scenarios), overestimates of Z did not often occur.

2.5. Discussion

2.5.1. Performance of mortality estimators

Our simulations suggest that the M/K ratio strongly affects the performance of the three length-based methods, with poor performance at low M/K for all three methods. This finding is consistent with previous simulations on LB-SPR (Hordyk et al. 2015b). When M/K is low, the peak of the LFD may not correspond to the true length of full selectivity (Figure 2.3). The best decision rules for both the LCCC and BHE used half-peak abundance length as the left truncation point (methods L1 and L3 for the LCCC and B1 for the BHE), although there was still a large bias associated with them. When M/K was high, there was no clearly superior method in both bias and precision.

The performance of all three length-based methods worsened in situations with extreme shapes in the LFD (i.e., low *M/K* with low *F/M* or high *M/K* with high *F/M*; Figure 2.3). In spite of this, if a stock is exploited over a broad range of sizes, then a qualitative assessment of the mortality rate is still possible based on life history and the shape of the LFD. High mortality can be inferred with truncation of the size structure due to low survival of animals to large size classes. Populations with a low *M/K* ratio and low *F/M* will have a protracted ascending limb due to the 'pile-up' effect where there are many large animals in the LFD due to low mortality. The use of length as a proxy for age by assuming deterministic growth in the LCCC and BHE did not appear to work well in such scenarios, because a much more substantial portion of the length distribution consists of lengths larger than L_∞ due to variability in growth (Figure 2.3). However, contrary to what might be expected, the LB-SPR method, which explicitly models variability in growth and selectivity (removing the need to truncate the data to meet model assumptions), did not perform more reliably than the LCCC and BHE in these situations.

In our study, all three length-based methods were robust to high growth variability. This result is surprising for the LCCC and BHE because both methods assume no variability in growth. Previous simulations with the LCCC showed that the estimator performed better with less growth variability (Isaac 1990), although Then et al. (2015) found that the BHE performed better with higher growth variability if the selectivity was dome-shaped. On the other hand, it was not surprising that LB-SPR performed worse when the CV of growth in the population was lower than that assumed in the estimation model. However, this assumption is not as critical in LB-SPR compared to the other two

methods because the CV of growth variability can be adjusted in the former when external information from a growth study is available.

The estimators were robust to the magnitude of recruitment variability as long as the recruitments were random. Trends in recruitment are more likely to be a problem because they would be conflated with mortality.

Dome selectivity had a noticeable effect on the bias only when $F/M = 0.25$ or 1, due to the high abundance of large individuals present in the population but missing in the catch (Figure 2.3). The length-based methods all assume logistic selectivity because dome selectivity is conflated with high mortality. To estimate a mortality rate, selectivity external selectivity estimates would be needed. For example, Ehrhardt and Ault (1992) developed a modified version of the BHE to estimate mortality when there is an upper length truncation in the LFD of the catch (Ehrhardt and Ault 1992) with that length of upper truncation estimated externally and then provided to the equation. Simulations have found the behavior of the Ehrhardt-Ault estimator to be complex (Then et al. 2015). Contrary to the methods tested in our study, the performance of their estimator often worsened with higher growth variability. In some cases, lower bias but higher variance was observed with using the Ehrhardt-Ault compared to the original equation, although the best input length for upper truncation for minimum bias was often larger the true length of upper truncation. Then et al. (2015) did not recommend the Ehrhardt-Ault estimator for routine use.

At high F/M, low survival to large size classes minimizes the effect of dome selectivity. Then et al. (2015) found a positive bias associated with the truncation of large animals in the length distribution when using the BHE for all mortality scenarios, but

they assumed knife-edge selection of small animals in their simulations. Our simulations also examined the effect of left truncation for the LCCC and the BHE when selectivity was not knife-edged. In theory, Steep selectivity more closely corresponds to the knife-edge selectivity assumption, compared to Gradual selectivity. However, all three methods (including LB-SPR) were robust to different logistic selectivity functions as indicated by the small differences in performance among truncation methods.

2.5.2. Sensitivity analyses

We examined the estimators in the ideal situation with large sample sizes and little observation error. The sensitivity analysis indicated that the estimators were generally robust to smaller sample sizes. We assumed that the generated data set was a random sample of animals from the vulnerable population. In reality, data are generally collected in clusters from samples of fishing trips or from schools of animals, often with similar lengths and ages within trips or schools. Cluster sampling reduces the effective sample size of the observed LFD and increases the uncertainty surrounding estimates of mortality (Chih 2011). With knowledge of the sampling program used to collect the data, the effective sample size can be estimated via bootstrapping methods (Stewart and Hamel 2014) or design-based formulas (Thorson 2014). Stewart and Hamel (2014) suggested that the number of sampled trips may be an appropriate proxy for the effective sample size. This suggestion may be applicable in a data-limited context if the ratio of within- to among-trip variance is low due to the cluster effect. The effective sample size would be important to determine if an appropriate range of size classes have been sampled, because

the sampling would affect the shape of the LFD used to apply the length truncation methods for the LCCC and BHE and estimate selectivity in LB-SPR.

The performance of the LCCC and BHE did not appear to diminish with larger length bins, and in some cases (with low *F/M*), performance was better. While sensitivity of length bins can be examined in individual applications of the equilibrium mortality estimators, low sample sizes may preclude the use of small length bins to describe the length composition of the catch in data-limited situations.

Our study design assumed that information on growth was stochastic, arising from a bivariate distribution with a highly negative correlation often associated with estimating parameters of the von Bertalanffy growth equation (Gallucci and Quinn 1979). Sensitivity analyses can be used to determine the influence of growth parameters on mortality estimation. An overestimate of L_∞ may create a positive bias for the estimate of total mortality, because fewer large animals are observed than are expected. We examined the correlation of total mortality estimates with the assumed values of L_∞ and K (Figure 2.9). In light of the systematic biases of the estimators among different *M/K* scenarios, overestimates of mortality may be more likely and underestimates less likely when *M/K* is low. Similarly, overestimates of mortality are unlikely when *M/K* and *F/M* are high. Such information could be used to assess the direction and magnitude of estimation error based on mis-specified growth in future applications of the mortality estimators.

Mortality estimates with length-based methods are dependent on the values of growth parameters. The quality of external growth estimates is affected by, among other things, the choice of the growth model (Gwinn et al. 2010) and the representativeness of

the size-at-age data to the population when sampling gears with different selectivity patterns are used (Wilson et al. 2015), and should be assessed in future applications of length-based methods. If size-at-age data are available, integrated modeling approaches for estimating growth simultaneously with mortality and selectivity also exist (Taylor et al. 2005).

2.5.3. Life history considerations

Overall, our study found that life history expressed in the *M/K* ratio was a good predictor of the performance of length-based mortality estimators. Meta-analyses of life history traits have shown a negative correlation between *M/K* and relative length of maturity (the ratio of the length at maturity to L_∞) in teleost families (Prince et al. 2015a). Better performance in high *M/K* scenarios was observed compared to low *M/K*. This result supports those in previous studies of length-based methods (e.g., Hordyk et al. 2015b). Two features unique to low *M/K* populations may result in their poor performance. First, the protracted ascending limb of the length composition in the population conflates selectivity with abundance-at-length. It may be difficult to select appropriate truncation lengths or estimate selectivity. Second, there is a high abundance of large animals from the 'pile-up' effect. In a low *M/K* population, a large age range is encompassed in a small spectrum of lengths and the length structure is a poor proxy for the age structure of the population. We recommend caution when using length-based methods in low *M/K* situations as the results are likely to be positively biased even in equilibrium situations. From a management standpoint, this behavior can dictate data collection priorities for alternative data types for assessment of low *M/K* stocks. In a data-

limited context, meta-analyses can be used to identify taxa with low *M/K* ratios (Prince et al. 2015a). Nevertheless, if length-based methods are to be used, our study suggests that classical methods (LCCC and BHE) remain viable options for mortality estimation.

2.6. Conclusion

Our study examined the performance of three length-based mortality estimators. When *M/K* is low (*M/K* = 0.5 in our simulation), we recommend using the BHE using half-peak abundance as the L_c, although the method is still likely to be positively biased and imprecise. When *M/K* is high (*M/K* = 2 in our simulation), both the LCCC and BHE performed well and were robust to variation in commonly-used truncation rules. For optimal performance, the length-based estimators require some *a priori* judgment of the life history and expected fishing pressure on the stock of interest. We recommend caution in using length-based methods for populations with low *M/K*. Overall, this study demonstrated that relative to LB-SPR, both the LCCC and BHE produced less biased and more precise estimates of total mortality. While LB-SPR did not perform as well compared to the other two methods when estimating mortality, the method has an advantage of providing estimates of selectivity, if desired. The LCCC and BHE methods performed comparably and no firm recommendation is made for choosing between these two methods.

2.7. References

Anonymous. 2012. Report of the 2011 ICCAT South Atlantic and Mediterranean Albacore Stock Assessment Sessions (Madrid, Spain – July 25 to 29, 2011). Collect. Vol. Sci. Pap. ICCAT 68:387-491.

Beverton, R. J. H. and S. J. Holt. 1956. A review of methods for estimating mortality rates in fish populations, with special reference to sources of bias in catch sampling. Rapports et Procès-verbaux des Reunions, Conséil International Pour L'Exploration de la Mer 140:67-83.

Beverton, R. J. H. and S. J. Holt. 1957. On the dynamics of exploited fish populations. U.K. Ministry of Agriculture, Fisheries, Food, and Fishery Investigations Series II, Vol. XIX.

Chih, C-P. 2011. The design effects of cluster sampling on the estimation of mean lengths and total mortality of reef fish. Fisheries Research 109:295-302.

Ehrhardt, N.M. and J.S. Ault. 1992. Analysis of Two Length-Based Mortality Models Applied to Bounded Catch Length Frequencies. Transactions of the American Fisheries Society 121:115-122.

Gallucci, V. F. and T. J. Quinn II. 1979. Reparameterizing, Fitting, and Testing a Simple Growth Model. Transactions of the American Fisheries Society 108:14-25.

Gayanilo, Jr., F. C., P. Sparre, and D. Pauly. 2005. FAO-ICLARM Stock Assessment Tools II (FiSAT II). Revised version. User's guide. FAO Computerized Information Series (Fisheries). No. 8, Revised version. FAO, Rome. 168 pages.

Gedamke, T., and J. M. Hoenig. 2006. Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of Goosefish. Transactions of the American Fisheries Society 135:476–487.

Gedamke, T., J. M. Hoenig, W. D. DuPaul, and J. A. Musick. 2008. Total mortality rates of the barndoor skate, *Dipturus laevis*, from the Gulf of Maine and Georges Bank, United States, 1963-2005. Fisheries Research 89:17-25.

Gwinn, D. C., M. S. Allen, and M. W. Rogers. 2010. Evaluation of procedures to reduce bias in fish growth parameter estimates resulting from size-selective sampling. Fisheries Research 105:75-79.

Hilborn, R. and C. J. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman & Hall, New York. 570 pages.

Hordyk, A., K. Ono, K. Sainsbury, N. Loneragan, and J. Prince. 2015a. Some explorations of the life history ratios to describe length composition, spawning-per-recruit, and the spawning potential ratio. ICES Journal of Marine Science 72:204-216.

Hordyk, A., K. Ono, S. Valencia, N. Loneragan, and J. Prince. 2015b. A novel length-based empirical estimation method of spawning potential ratio (SPR), and tests of its performance, for small-scale, data-poor fisheries. ICES Journal of Marine Science 72:217-231.

Hordyk, A. 2017. LBSPR: Length-Based Spawning Potential Ratio. R package version 0.1.1. Available at: https://CRAN.R-project.org/package=LBSPR

ICES. 2014. Report of the Workshop on the Development of Quantitative Assessment Methodologies based on LIFE-history traits, exploitation characteristics, and other relevant parameters for data-limited stocks (WKLIFE IV), 27–31 October 2014, Lisbon, Portugal. ICES CM 2014/ACOM:54. 223 pp.

Isaac, V. J. 1990. The Accuracy of Some Length-Based Methods for Fish Population Studies. International Center for Living Aquatic Resources Management, Manila.

Kalhoro, M. A., Q. Liu, K. H. Memon, M. S. Chang, and K. Zhang. 2014. Population dynamics of Japanese threadfin bream *Nemipterus japonicus* from Pakistani waters. Acta Oceanologica Sinica 33:49-57.

Laurec, A. and B. Mesnil. 1987. Analytical investigations of errors in mortality rates estimated from length distribution of catches. Pages 239-281 *in* D. Pauly and G.R. Morgan, editors. Length-based methods in fisheries research. International Center for Living Aquatic Resources Management, Conference Proceedings 13, Manila and Kuwait Institute for Scientific Research, Safat, Kuwait.

Mirzaei, M. R., Z. Yasin, and A. T. S. Hwai. 2015. Length-weight relationship, growth and mortality of *Anadara granosa* in Penang Island, Malaysia: an approach using length-frequency data sets. Journal of the Marine Biological Association of the United Kingdom 95:381-390.

Pauly, D. 1983. Length-converted catch curves: A powerful tool for fisheries research in the tropics (Part I). Fishbyte 1(2):9-13.

Pauly, D. 1984a. Length-converted catch curves: A powerful tool for fisheries research in the tropics (Part II). Fishbyte 2(1):17-19.

Pauly, D. 1984b. Length-converted catch curves: A powerful tool for fisheries research in the tropics (III: Conclusion). Fishbyte 2(3):9-10.

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

Pauly, D. 1987. A review of the ELEFAN system for analysis of length frequency data in fish and aquatic invertebrates. Pages 7-34 *in* D. Pauly and G.R. Morgan, editors. Length-based methods in fisheries research. International Center for Living Aquatic Resources Management, Conference Proceedings 13, Manila and Kuwait Institute for Scientific Research, Safat, Kuwait.

Pauly, D. 1990. Length-Converted catch curves and the seasonal growth of fishes. Fishbyte 8(3): 24-29.

Prince, J., A. Hordyk, S. R. Valencia, N. Loneragan, and K. Sainsbury. 2015a. Revisiting the concept of Beverton–Holt life-history invariants with the aim of informing data-poor fisheries assessment. ICES Journal of Marine Science 72:194-203.

Prince, J., S. Victor, V. Kloulchad, and A. Hordyk. 2015b. Length based SPR assessment of eleven Indo-Pacific coral reef fish populations in Palau. Fisheries Research 171:42-58.

Punt, A. E., T. Huang, and M. N. Maunder. 2013. Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusk species. ICES Journal of Marine Science 70:16-33.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

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

Rodriguez-Cortes, K. D., A. Aguilar-Perera, and J. L. Bonilla-Gomez. 2015. Growth and mortality of red lionfish *Pterois volitans* (Actinopterygii: Scorpaeniformes: Scorpaenidae) in the Parque Nacional Arrecife Alacranes, Southern Gulf of Mexico as determined by size-frequency analysis. Acta Ichthyologica et Piscatoria 45:175-179.

Rudershausen, P.J., E. H. Williams, J. A. Buckel, J. C. Potts, and C. S. Manooch III. 2008. Comparison of Reef Fish Catch per Unit Effort and Total Mortality between the 1970s and 2005-2006 in Onslow Bay, North Carolina. Transactions of the American Fisheries Society 137:1389-1405.

Sparre, P. 1990. Can we use traditional length-based fish stock assessment when growth is seasonal? Fishbyte 8(3): 29-32.

Stewart, I. J. and O. S. Hamel. 2014. Bootstrapping of sample sizes for length- or age-composition data used in stock assessments. Canadian Journal of Fisheries and Aquatic Sciences 71:581-588.

Taylor, N. G., C. J. Walters, and S. J. D. Martell. 2005. A new likelihood for simultaneously estimating von Bertalanffy growth parameters, gear selectivity, and natural and fishing mortality. Canadian Journal of Fisheries and Aquatic Science 62:215-223.

Then, A.Y., J. M. Hoenig, T. Gedamke, and J. S. Ault. 2015. Comparison of Two Length-Based Estimators of Total Mortality: a Simulation Approach. Transactions of the American Fisheries Society 144:1206-1219.

Thorson, J. T. 2014. Standardizing compositional data for stock assessment. ICES Journal of Marine Science 71:1117-1128.

Thorson, J. T., O. P. Jensen, and E. F. Zipkin. 2014. How variable is recruitment for exploited marine fishes? A hierarchical model for testing life history theory. Canadian Journal of Fisheries and Aquatic Sciences 71:973-983.

Thorson, J. T., O. P. Jensen, and E. F. Zipkin. 2016. Corrigendum: How variable is recruitment for exploited marine fishes? A hierarchical model for testing life history theory. Canadian Journal of Fisheries and Aquatic Sciences 73:1014.

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 32:21-28.

Wetherall, J. A., J. J. Polovina, and S. Ralston. 1987. Estimating growth and mortality in steady-state fish stocks from length frequency data. Pages 53-74 *in* D. Pauly and G.R. Morgan, editors. Length-based methods in fisheries research. International Center for Living Aquatic Resources Management, Conference Proceedings 13, Manila and Kuwait Institute for Scientific Research, Safat, Kuwait.

Whitten, A. R., N. L. Klaer, G. N. Tuck, and R. W. Day. 2013. Accounting for cohort-specific variable growth in fisheries stock assessments: A case study from south-eastern Australia. Fisheries Research 142:27-36.

Wilson, K. L., B. G. Matthias, A. B. Barbour, R. N. M. Ahrens, T. Tuten, and M. S. Allen. 2015. Combining Samples from Multiple Gears Helps to Avoid Fishy Growth Curves. North American Journal of Fisheries Management 35:1121-1131.

Windsland, K. 2015. Total and natural mortality of red king crab (*Paralithodes camtschaticus*) in Norwegian waters: catch-curve analysis and indirect estimation methods. ICES Journal of Marine Science 72:642-650.

Zischke, M. T. and S. P. Griffiths. 2015. Per-recruit stock assessment of wahoo (*Acanthocybium solandri*) in the southwest Pacific Ocean. Fishery Bulletin 113:407-418.

Zhou, S., S. Yin, J. T. Thorson, A. D. M. Smith, and M. Fuller. 2012. Linking fishing mortality reference points to life history traits: an empirical study. Canadian Journal of Fisheries and Aquatic Science 69:1292-1301.

2.8. Tables

Table 2.1. Parameter values used for data generation in the simulation study. Parameters with multiple values were included in factorial design. Parameters L_{50} and L_{95} are the lengths of 50% and 95% selectivity, respectively, using a logistic function. Parameters μ_d and σ_d are the mean and standard deviation of the normal probability density function, respectively, with values standardized to 1 at length μ_d for dome-shaped selectivity.

Parameter	Symbol	Values
Ratio of natural mortality rate *M* and von Bertalanffy *K*	*M/K*	Low: 0.5 (*K* = 0.4) High: 2.0 (*K* = 0.1)
Ratio of fishing mortality *F* and natural mortality rates	*F/M*	Low: 0.25 (*F* = 0.05) Medium: 1.0 (*F* = 0.2) High: 5.0 (*F* = 1.0)
Coefficient of variation of length-at-age ($L_{y,a}$)	*CV*	Low: 0.03 Medium: 0.06 High: 0.09
Recruitment residual standard deviation	σ_R	Low: 0.6 High: 1.0
Selectivity-at-length	sel_L	Gradual: $L_{50} = 175,\ L_{95} = 200$ Steep: $L_{50} = 175,\ L_{95} = 275$ Dome: $L_{50} = 175,\ L_{95} = 275,\ \mu_d = 325,\ \sigma_d = 65$
Recruitment autocorrelation coefficient	ρ	0.45
von Bertalanffy asymptotic length	L_∞	500
Expected length at age-0	L_0	75
Cohort growth standard deviation	σ_C	0.15
Maximum age	*A*	23

Table 2.2. Truncation methods of the length data for estimating total mortality Z with the length-converted catch curve (LCCC) and Beverton-Holt equation (BHE). No truncation is associated with the LB method (LB-SPR).

Method	Estimator	Left truncation	Right truncation
L1	LCCC	Half-peak abundance	5+
L2	LCCC	Half-peak abundance	90% L_∞
L3	LCCC	Half-peak abundance	100% L_∞
L4	LCCC	Peak	5+
L5	LCCC	Peak	90% L_∞
L6	LCCC	Peak	100% L_∞
L7	LCCC	Peak-plus	5+
L8	LCCC	Peak-plus	90% L_∞
L9	LCCC	Peak-plus	100% L_∞
B1	BHE	Half-peak abundance	N/A
B2	BHE	Peak	N/A
B3	BHE	Peak-plus	N/A
LB	LB-SPR	N/A	N/A

2.9. Figures

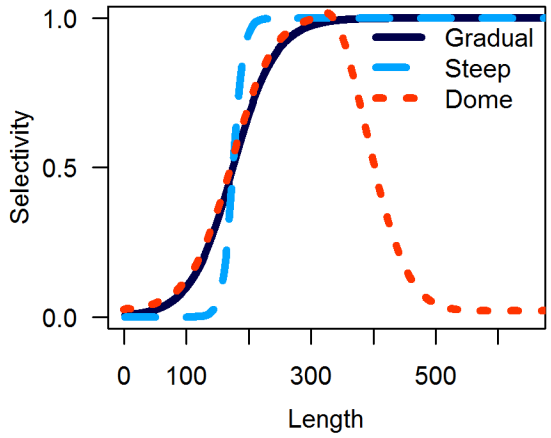

Figure 2.1. Length-based selectivity functions used in the simulation.

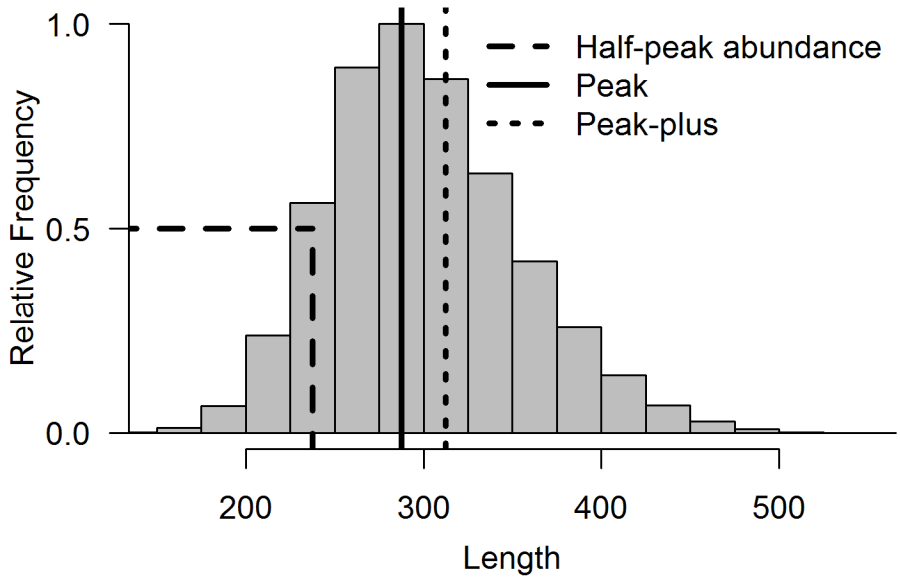

Figure 2.2. Histogram of a length frequency distribution with the left-handed decision rules (Half-peak abundance, Peak, and Peak-plus) used to select the length bin of left truncation for the LCCC and value of L_c for the BHE in the simulation study.

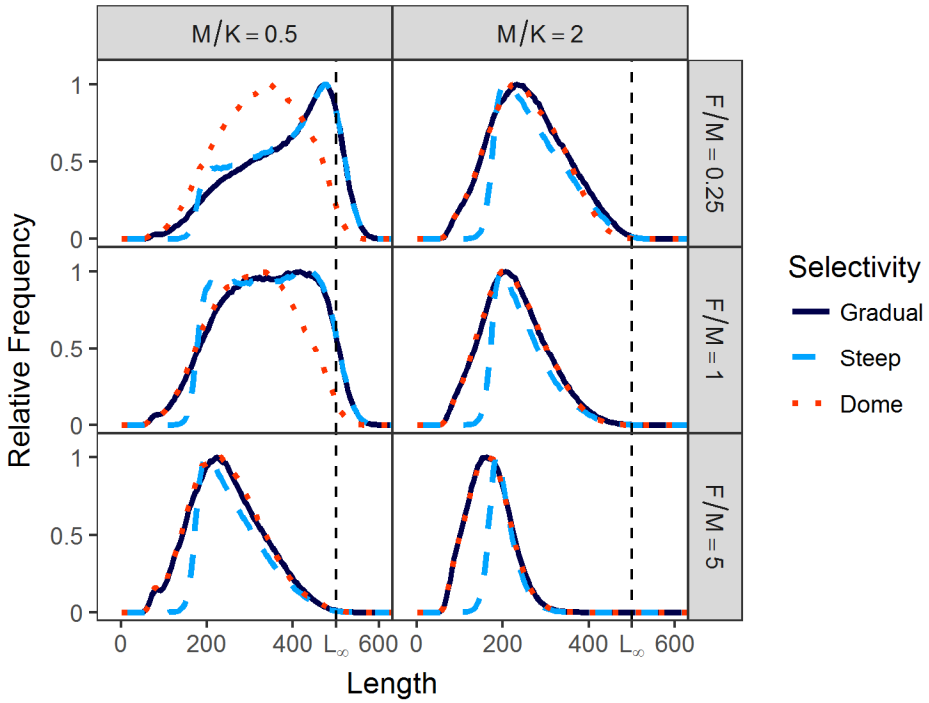

Figure 2.3. Expected length frequency distributions obtained from the sum of 1,000 data sets from the simulation stratified by the factorial design for *M/K*, *F/M*, and selectivity. Selectivity functions correspond to those in Figure 2.1. In all panels, medium growth variability and low recruitment variability was assumed in the sample. Dashed vertical lines indicate $L_{\infty} = 500$ (Table 2.1).

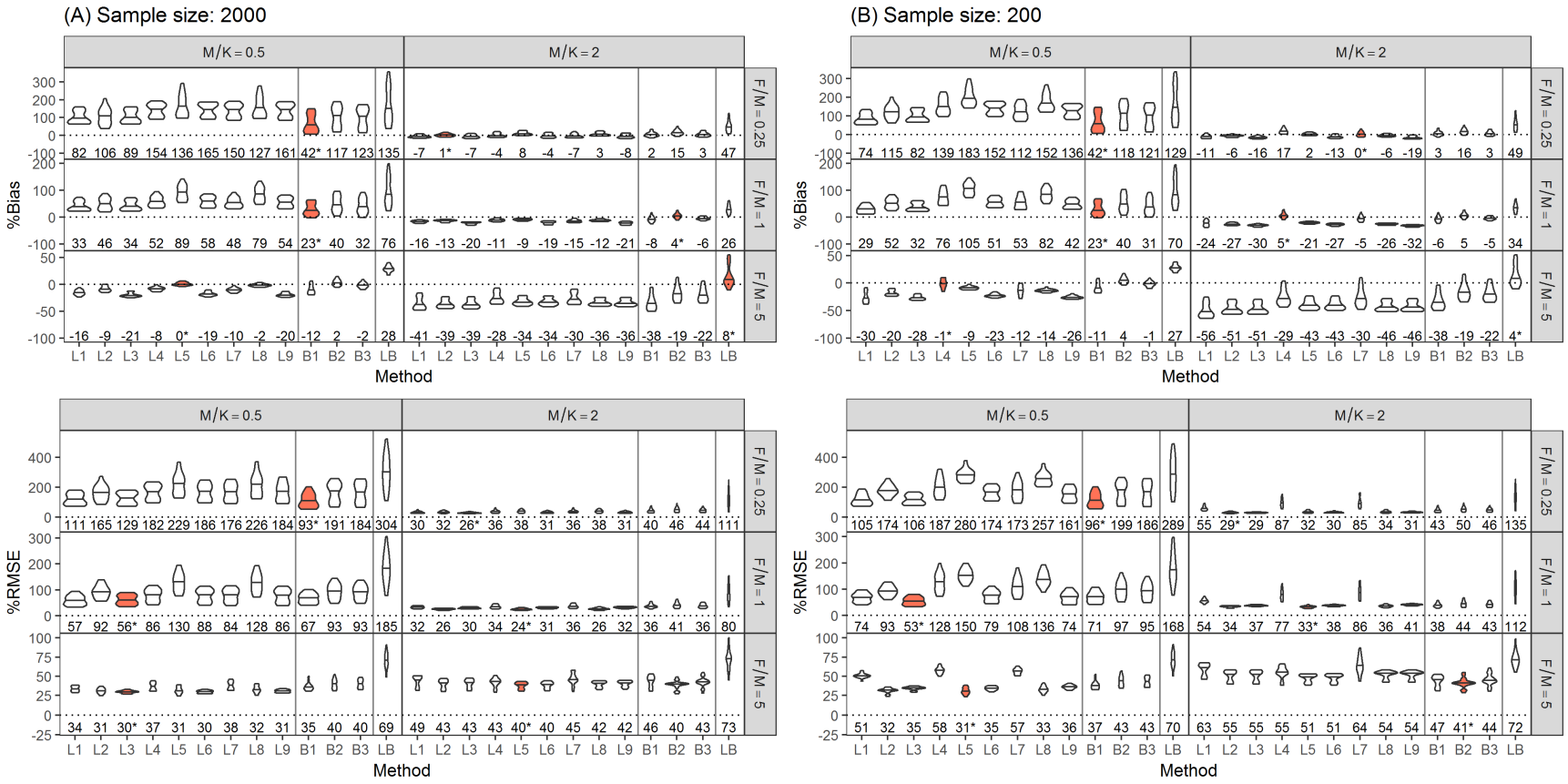

Figure 2.4. %*Bias* (top grids) and %*RMSE* (bottom grids) from the simulation study when the data set sample size is 2000 (A) and 200 (B). For each method, factorial combinations are stratified by *M/K* and *F/M*. Numbers and horizontal lines in the violin plots indicate median %*Bias* and %*RMSE*, with the numbers rounded to the nearest whole number for clarity. The shape of the violin plots shows the distribution of values. Asterisks and shaded violin plots indicate the method with the lowest median value in each grid cell (not subject to rounding error). Rows in each grid have separate scales on the y-axis to show the shape of the violin plots.

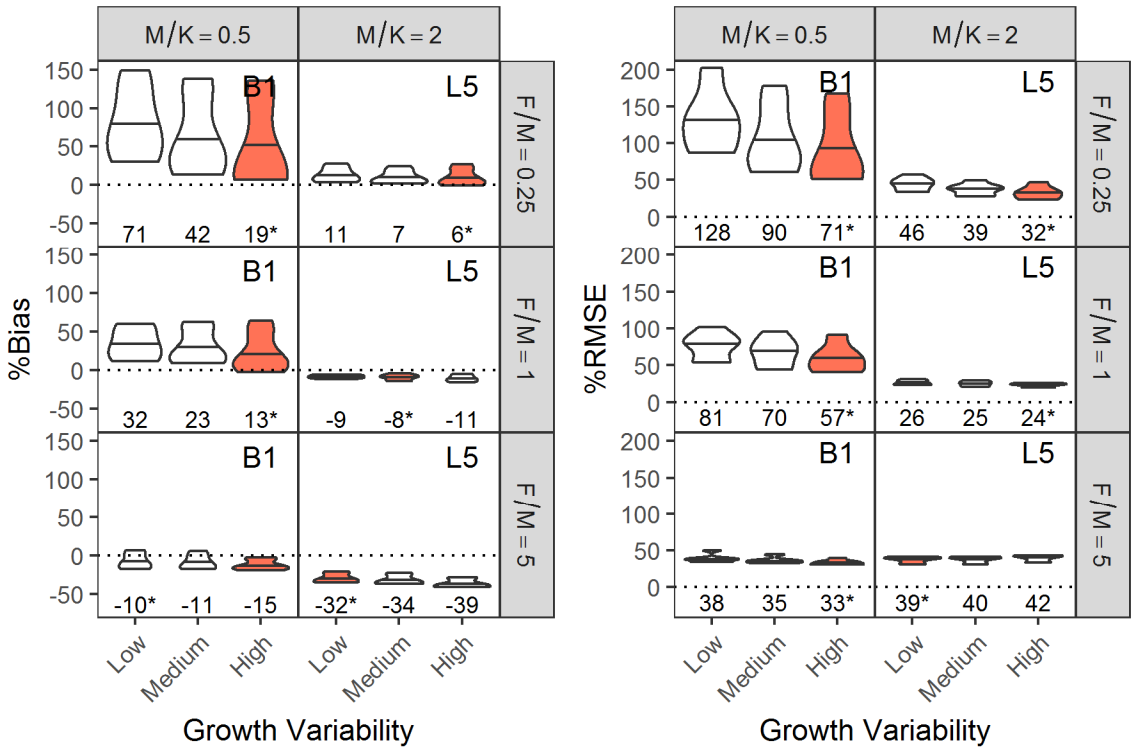

Figure 2.5. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability. Only methods B1 and L5 are shown (corner text in the corners indicate the method shown). Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and the shape of violin plot shows the distribution of values. Asterisks and shaded violin plots indicate the method with the lowest median value in each grid cell (not subject to rounding error).

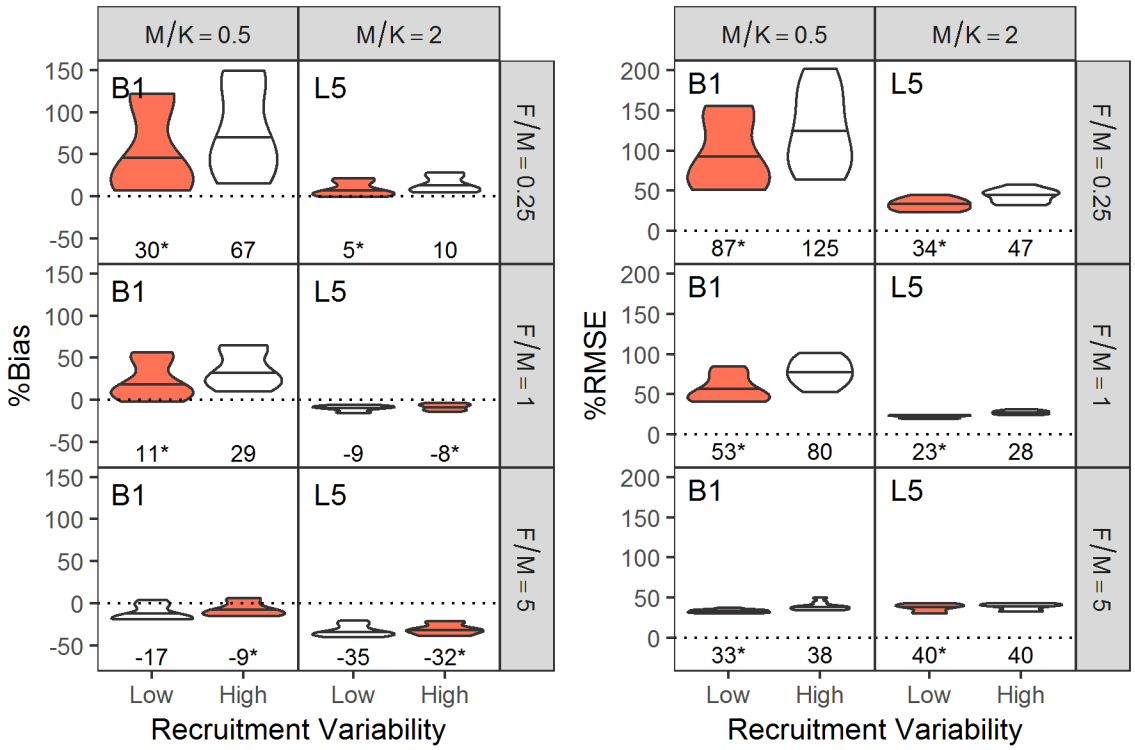

Figure 2.6. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability. Only methods B1 and L5 are shown (corner text in the corners indicate the method shown). Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and the shape of violin plot shows the distribution of values. Asterisks and shaded violin plots indicate the method with the lowest median value in each grid cell (not subject to rounding error).

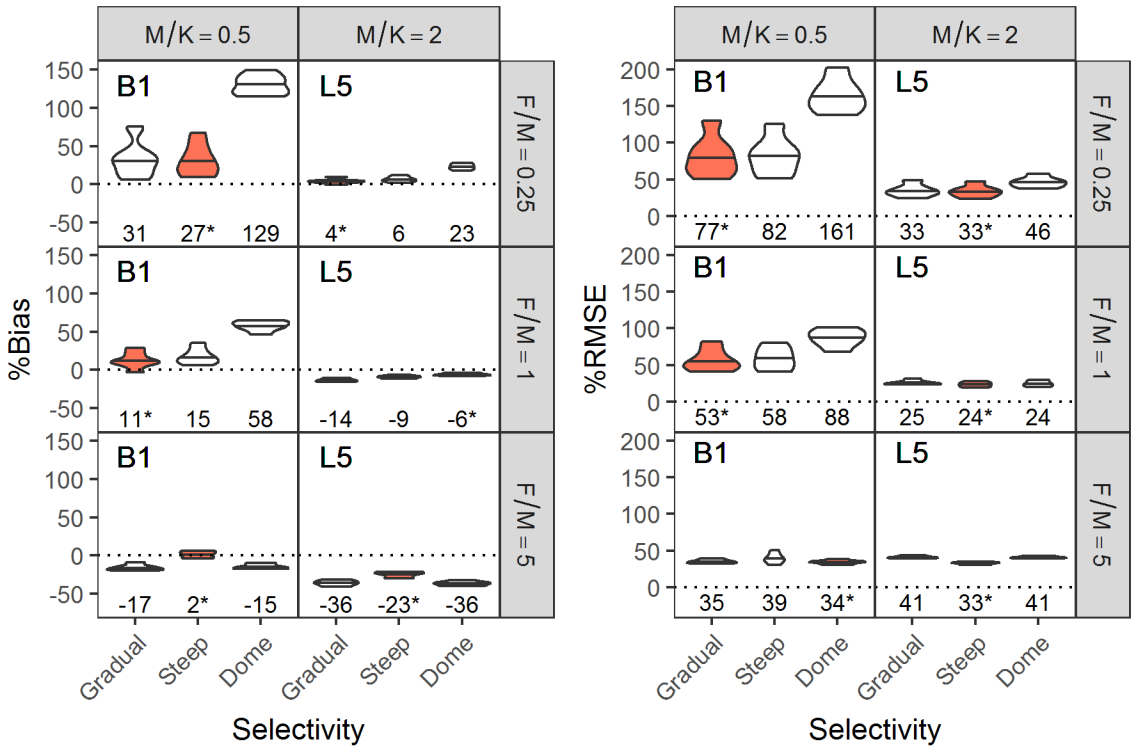

Figure 2.7. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity. Only methods B1 and L5 are shown (corner text in the corners indicate the method shown). Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and the shape of violin plot shows the distribution of values. Asterisks and shaded violin plots indicate the method with the lowest median value in each grid cell (not subject to rounding error).

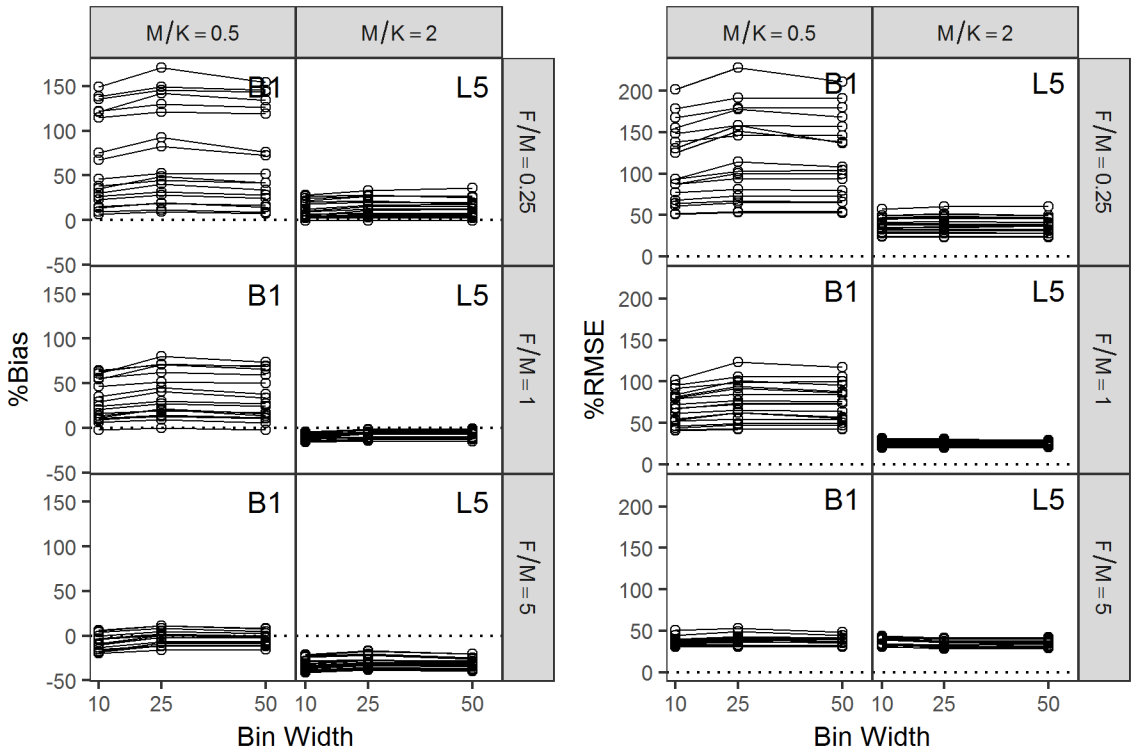

Figure 2.8. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for the length-based estimators. Only methods B1 and L5 are shown (corner text in the corners indicate the method shown). Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

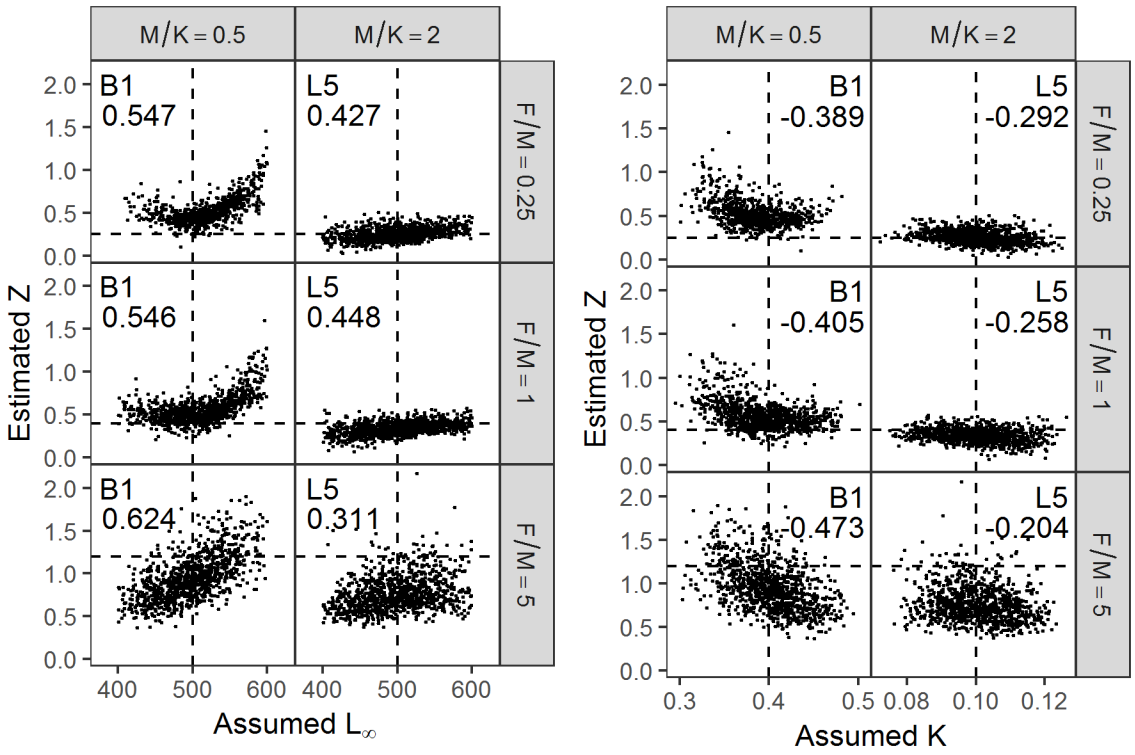

Figure 2.9. Individual estimates of total mortality *Z* based on the assumed values of von Bertalanffy parameters L_∞ (left grid) and *K* (right grid) in the estimation model (parameters are sampled from a bivariate normal distribution around the true values with a correlation of -0.9). Only methods B1 and L5 are shown (corner text in the corners indicate the method shown). Plotted estimates are from the simulations with medium growth variability, low recruitment variability, and the gradual selectivity function stratified in separate cells by *M/K* and *F/M*. Dotted lines indicate the true value of mortality and growth parameter in the respective cell.

Chapter 3: Multispecies Extensions to a Nonequilibrium Length-Based Mortality Estimator

3.1. Abstract

Recent advances in methodology allow the history of the total mortality rate experienced by a population to be estimated from periodic (e.g., annual) observations on the mean length of the population. This approach is generalized to allow data on several species that are caught together to be analyzed simultaneously based on the theory that changes in fishing effort are likely to affect several species; thus, the estimation of times when the mortality rate changes for one species borrows strength from data on other, concurrently caught species. Information theory can be used to select among models describing the degree of synchrony (if any) in mortality changes for a suite of species. This approach is illustrated using data on Puerto Rican handline fishery catches of three snapper species: Silk Snapper *Lutjanus vivanus*, Blackfin Snapper *L. buccanella*, and Vermilion Snapper *Rhomboplites aurorubens*. We identified the best model as the one that provided for simultaneous decreases in mortality rate around the year 1997 and for separate, species-specific magnitudes of change in total mortality. The simultaneous estimation of parameters for multiple species can provide for more credibility in the inferred mortality trends than is possible with independent estimation for each species.

3.2. Introduction

The mean length of animals in a population is a dynamic statistic that reflects the recent and past mortality rates experienced by the population (Baranov 1918, cited by Ricker 1975; Beverton and Holt 1956, 1957). Under stringent equilibrium conditions, the total instantaneous mortality rate (Z; year^{-1}) can be estimated by the well-known Beverton–Holt equation (Beverton and Holt 1956, 1957),

$$Z = \frac{K(L_\infty - \overline{L})}{\overline{L} - L_c}, \tag{3.1}$$

where K and L_∞ are parameters of the von Bertalanffy growth curve (K = Brody growth coefficient; L_∞ = asymptotic length), L_c is the smallest length of animals that are fully vulnerable to the fishing gear, and $\overline{L}$ is the mean size of animals that are larger than L_c. Gedamke and Hoenig (2006) generalized this approach by analyzing an annual series of mean length measurements to estimate (1) the years in which the mortality rate changed and (2) the magnitudes of mortality in the separate time periods identified. To do this, Gedamke and Hoenig (2006) derived the transitional behavior of the mean length statistic as the population moved from one equilibrium state toward another equilibrium state after a change in mortality.

The present work was motivated by the consideration that most types of fishing gear catch a variety of fish species. Consequently, if fishing effort changes over time, one might reasonably expect that an assemblage or "complex" of species that are caught together would show synchronous changes in mortality rate. For example, if handline fishing effort doubles from one year to the next, one might suspect that all fish species

that are caught by the handline gear would experience an increase in mortality during the same year. Data from all species in the complex could then be used to estimate the common years of change, resulting in greater efficiency in estimation. Note that all species in the complex do not necessarily experience the same time schedule of mortality changes. For example, one species may have a restricted range centered on the main fishing grounds where the increase in fishing occurred, and another species might have a broader distribution occurring on both the main fishing ground and a secondary fishing ground where effort has declined; those two species might not show synchronous changes in mortality. In addition, the magnitude of changes in mortality will not necessarily be the same for all species, as species with low catchability may exhibit a smaller change in mortality than species with high catchability.

In the present work, we sought to model changes in the mean length of several species simultaneously so as to estimate period-specific mortality rates and years of change. We develop four progressively restrictive and nested models for the estimation of mortality when multiple species are considered: (1) trends in total mortality (both the timing and magnitude of the change in mortality) are independent for each species; (2) the year in which the change in total mortality occurs (hereafter, "change point") is common to all species, but the magnitude of the change is independent; (3) the change point is common to all species, and the magnitude of the change in fishing mortality is the product of a time and species effect; and (4) both the change point and the magnitude of the change in fishing mortality are common to all species. The latter three approaches are attractive because they borrow strength from disparate data with common underlying trends in mortality that otherwise might not be detected by analyzing each species

separately. Additionally, these approaches allow us to distinguish the dynamics that are common to all species (i.e., the change points) from those that are not (i.e., the magnitude of the change in mortality). We describe our application of the models to three species that occur together in Puerto Rican handline catches: Silk Snapper *Lutjanus vivanus*, Blackfin Snapper *L. buccanella*, and Vermilion Snapper *Rhomboplites aurorubens*.

3.3. Methods

3.3.1. Model Development and Model Fitting

We consider a complex of N species that tend to be caught together and for which we suspect the patterns of fishing effort are similar over time. We hold the following assumptions for each species considered:

1. Individual growth follows the von Bertalanffy growth function, with the parameters K and L_∞ known and constant over time.
2. There is no individual variability in growth.
3. Recruitment is constant and continuous over time; or if recruitment fluctuates, it does so randomly (i.e., with no time trend).
4. Total mortality rate Z is constant with age after the age of recruitment (t_c) that corresponds to L_c.

These assumptions are discussed by Gedamke and Hoenig (2006) and Then et al. (2015a).

Consider first the simplest case in which a population starts at equilibrium with a total instantaneous mortality rate of $Z_{1,n}$ and then experiences a second mortality rate $Z_{2,n}$, where the first subscript indexes the time period of the respective mortality rate

and the second subscript indexes species $n = 1, 2, \ldots, N$. Generalizing the results of Gedamke and Hoenig (2006) by adding a subscript for species, the mean length d years after the change point is computed as

$$\mu_{(n)}(d) = L_{\infty(n)} - \frac{Z_{1,n}Z_{2,n}(L_{\infty(n)} - L_{c(n)})\{Z_{1,n} + K_n + (Z_{2,n} - Z_{1,n})\exp[-(Z_{2,n} + K_n)d]\}}{(Z_{1,n} + K_n)(Z_{2,n} + K_n)\{Z_{1,n} + (Z_{2,n} - Z_{1,n})\exp(Z_{2,n}d)\}}, \quad (3.2)$$

where $\mu_{(n)}(d)$ is the predicted mean length of animals of species n as a function of time d years after the change point; the remaining symbols are as in equation (3.1), with subscripts added for time period and species as needed.

Gedamke and Hoenig (2006) computed the mean length at any time when there was an arbitrary number of change points over time. Cardinale et al. (2009) analyzed length frequency data from an annual trawl survey conducted for over 100 years, and they estimated nine separate total mortality rates (and eight change points). The general equation for computing the predicted mean length in each year given a history of mortality rates is provided by Gedamke and Hoenig (2006; their Appendix 2).

From a time series of mean length observations, the mortality rates and change points can be estimated as described by Gedamke and Hoenig (2006). By virtue of the central limit theorem, we model the observed mean length $\overline{L}_{t,n}$ of species n in year $t = 1, 2, \ldots, T$ as a normally distributed random variable with mean $\mu_{t,n}$ and variance $\sigma_n^2 / m_{t,n}$, where $m_{t,n}$ is the sample size of observed lengths above L_c in year t for species n. A

common variance σ_n^2 over time is assumed so that the precision of the mean length in year *t* depends on the sample size $m_{t,n}$ in that year for species *n*. Annual means are modeled as in equation (3.2) by rewriting the equation in terms of calendar years instead of the number of years since a change in mortality occurred or—more generally—as in Gedamke and Hoenig (2006; their Appendix 2), assuming more than one change point in the time series.

The log-likelihood ($\log\lambda_n$) of observing mean length $\overline{L}_{t,n}$ for species *n* over *T* years is proportional to

$$\log\lambda_n \propto -T\log(\sigma_n) - \frac{1}{2\sigma_n^2}\sum_{t=1}^{T} m_{t,n}(\overline{L}_{t,n} - \mu_{t,n})^2 \,. \tag{3.3}$$

Maximum likelihood estimates of the mortality rates and change points for species *n* are found numerically by estimating the values of the parameters that maximize the log-likelihood. The maximum likelihood estimate for the residual variance $\hat{\sigma}_n^{\,2}$ is solved analytically,

$$\hat{\sigma}_n^{\,2} = \frac{1}{T}\sum_{t=1}^{T} m_{t,n}(\overline{L}_{t,n} - \hat{\mu}_{t,n})^2 \,, \tag{3.4}$$

where $\hat{\mu}_{t,n}$ is the maximum likelihood estimate of the mean length.

3.3.2. Modifications for Multispecies Estimation

The total log-likelihood ($\log\Lambda$) for all *N* species is simply the sum of the individual species' log-likelihoods,

$$\log\Lambda = \sum_{n=1}^{N} \log\lambda_n \,. \tag{3.5}$$

Assuming that there are I change points, let each species $n = 1, 2,\ldots, N$ have its own vector of parameters $\theta_n = \{Z_{i,n}, D_{i,n}\}$, where $Z_{i,n}$ is the vector of period-specific total mortality rates through time (with $i = 1, 2,\ldots, I + 1$); and $D_{i,n}$ is the vector of change points in calendar years from mortality rate $Z_{i,n}$ to $Z_{i+1,n}$ (with $i = 1, 2,\ldots, I$). We can then envision a suite of progressively more restrictive models for the patterns of mortality across species: a single-species model (SSM) and three multiple-species models (MSM1, MSM2, and MSM3), as described below.

Single-species model.—In the single-species scenario, what happens to one species is not reflected in what happens to other species. For example, fishers may target certain species at certain times so that changes in total mortality for one species are independent of the changes experienced by another species. In this case, the mortality rates ($Z_{i,n}$) and change points ($D_{i,n}$) are all estimated parameters. Maximizing log Λ (equation 3.5) is equivalent to applying the mean length estimator to each species independently because there are no parameters in common among species.

Multispecies model 1: common years of change.—In MSM1, changes in fishing effort simultaneously affect all species in the complex being considered; however, the magnitudes of the changes in Z are independent. Thus, all species have a common set of change points when the mortality rate changed (i.e., $D_{i,n} = D_{i,n'}$ for all periods i and for all pairs of species n and n'). This feature is included in the following two models, and we drop the second subscript for the change points in the subsequent text.

Multispecies model 2: common years of change with species-specific proportional changes in fishing mortality.—In MSM2, all species experience synchronous changes in mortality (i.e., common change points as in MSM1), but in addition we employ separability to model changes in fishing mortality with temporal and species components. The total instantaneous mortality rate $Z_{i,n}$ is broken down into its components,

$$Z_{i,n} = F_{i,n} + M_n, \tag{3.6}$$

where $F_{i,n}$ is the fishing mortality rate for species *n* during period *i*; and M_n is the time-invariant and age-invariant natural mortality rate for species *n*. A change in fishing effort in the next time period would cause the fishing mortality for a reference species to change by a factor δ. Not all species can be expected to experience the same proportional change in fishing mortality (e.g., due to different catchability coefficients), so we incorporate species-specific effects (ε_n). Subsequent mortality rates for species *n* are

$$Z_{i+1,n} = \begin{cases} \delta_{i,1} F_{i,n} + M_n & n = 1 \\ \delta_{i,1} \varepsilon_n F_{i,n} + M_n & n = 2,\ldots,N \end{cases}, \tag{3.7}$$

where $\delta_{i,1}$ is the proportional change in fishing mortality for a reference species $n = 1$ at time D_i; and ε_n is the multiplicative species effect relative to the reference species for all other species $n = 2,\ldots, N$.

In MSM2, the following parameters are estimated: the first total mortality rate $Z_{1,n}$ for all *N* species; the residual variances σ_n^2 for each species; the common change points $D_{i,n}$; the proportional changes $\delta_{i,1}$ in fishing mortality for the reference species; and the species-specific effects ε_n for all other species in the complex. Successive total mortality rates $Z_{i,n}$ ($i > 1$) are derived by propagating equation (3.7). The values of M_n

are obtained externally to the analysis of the mean length data—for example, via the methods described by Then et al. (2015b) or Hamel (2015). If there is only one change point, MSM2 is equivalent to MSM1.

Multispecies model 3: common years of change and common proportional changes in fishing mortality.—The MSM3 assumes that changes in mortality are concurrent among species in the complex, with the same proportional changes in fishing mortality for all species. For all species in the complex, we modify equation (3.7) such that

$$Z_{i+1,n} = \delta_i F_{i,n} + M_n, \tag{3.8}$$

where δ_i is the proportional change in fishing mortality at time D_i. By dropping the species subscript, the proportional change δ_i is common to all species.

In MSM3, estimated parameters include the first total mortality rate $Z_{1,n}$ for each species, the residual variances σ_n^2 for each species, the common change points D_i, and the corresponding δ_i; the values of M_n are again obtained externally. Successive total mortality rates $Z_{i,n}$ ($i > 1$) are derived by propagating equation (3.8).

3.3.3. Model Complexity and Model Selection

Knowledge of fishing practices for species in the complex (e.g., spatial extent; depth strata fished) can be used to guide the choice of model for estimating total mortality rates. If regulations on fishing effort are implemented in different years or if there are changes in species targeting, this information may be considered when selecting

the appropriate model. In the absence of such knowledge or if an empirical approach is desirable, information theory (Akaike's information criterion corrected for small sample sizes [AIC_c]) can be used to select the best model (Burnham and Anderson 2002). Accounting for both parsimony and goodness of fit, the model with the lowest AIC_c value (AIC_c difference [ΔAIC_c] = 0) can be chosen as the most plausible model, whereas the support for other candidate models decreases with higher AIC_c values. Model selection is conditional on the values of natural mortality (*M*) that are specified in MSM2 and MSM3. Uncertainty in *M* may be addressed with a sensitivity analysis using alternative values.

By expanding the mean length analysis to incorporate multiple species and common parameters between species, the number of estimated parameters is reduced. Table 3.1 presents the formulas for the number of estimated parameters given *I* changes in mortality and *N* species. For example, with three species and one change in mortality, a total of 12, 10, 10, and 8 parameters would be estimated by the SSM, MSM1, MSM2, and MSM3, respectively. Incorporating an additional change point would increase the number of estimated parameters by six in the SSM but only by four in MSM1. Adding another species in the analysis would increase the number of estimated parameters by four in the SSM but only by three in MSM1.

3.3.4. Application to Deepwater Snappers in the Puerto Rican Handline Fishery

We demonstrate the application of the multispecies mean length mortality estimators by using data for the Silk Snapper, Blackfin Snapper, and Vermilion Snapper in the Puerto Rican handline fishery. The distributions of these three species are

overlapping in terms of habitat and depth strata (Sylvester 1974; Boardman and Weiler 1980; Claro et al. 2001). Catches in the deepwater snapper complex have historically been dominated by Silk Snapper, followed by Vermilion Snapper and then Blackfin Snapper (Claro et al. 2001; SEDAR 2011). Currently, the three species are managed together as species complex (Snapper Unit 1) by the Caribbean Fishery Management Council (USOFR 2005). The present analysis is intended to be a demonstration of our methods and not an assessment of the three stocks.

Length data for the three deepwater snapper species from 1983 to 2013 were obtained from the commercial handline fishery through the Trip Interview Program (TIP) of the National Marine Fisheries Service's Southeast Fisheries Science Center. Fishing occurred at depths of up to 519 m (280 fathoms), with most animals caught above 370 m (200 fathoms). The K and L_∞ parameters of the von Bertalanffy growth function for each species were obtained from the literature (Table 3.2). For Silk Snapper and Vermilion Snapper, there was considerable variability in the reported growth parameters. Ultimately, we used estimates from the 2011 assessment for Silk Snapper (SEDAR 2011) and estimates from Caribbean stocks for Blackfin Snapper and Vermilion Snapper (Table 3.2). Analyses of sensitivity to the misspecification of von Bertalanffy parameters have shown that the mean length mortality estimator is most sensitive to the overestimation or underestimation of L_∞ (Gedamke and Hoenig 2006).

The L_c parameter was determined for each species by examining the respective length frequency data spanning the entire time period (Figure 3.1). In general, the length near the peak of the histogram was chosen. For Silk Snapper, the data showed an

increase in the peak over time from 260 mm to 310 mm; in this case, we selected the peak of the most recent time period (i.e., 310 mm) as the L_c (Table 3.3).

The annual mean length of animals above L_c (i.e., $\bar{L}$) and the respective sample sizes were then calculated; 7,497 records for Silk Snapper, 1,902 records for Blackfin Snapper, and 3,836 records for Vermilion Snapper were available after we removed records of lengths smaller than L_c. Sample sizes generally increased through time, although there were several intermittent years without length data for a given species (Figure 3.2).

Using the SSM, MSM1, and MSM3, mortality rates were estimated by assuming that there was one change point in the time series. For Blackfin Snapper, the total mortality rate in the SSM was also estimated by assuming zero change points (equilibrium conditions), as there was equal support for zero and one change point. Because the data only suggested one change in mortality over time (Figure 3.2), we did not implement MSM2 due to redundancy (see Section 3.3.2). For MSM3, estimates of M were obtained via the method of Then et al. (2015b) by using von Bertalanffy parameters rather than maximum ages because the latter were not available (Table 3.3).

A sensitivity analysis of MSM3 to the prescribed M was performed. A factorial design was implemented to specify M for each of the three species at 60, 80, 100, 120, and 140% of the base values in Table 3.3, resulting in a total of 125 factorial combinations for all levels in all three species. The percent deviation (%DEV) in the estimated total mortality rate from each factorial combination was calculated as

$$\%DEV = 100 \times \frac{Z_{sens} - Z_{base}}{Z_{base}}, \tag{3.9}$$

where Z_{sens} is the estimated total mortality rate in the respective sensitivity run; and Z_{base} is the total mortality rate that was estimated by using the base values of *M*.

3.4. Results

3.4.1. Application to Snappers in the Puerto Rican Handline Fishery

Model fits to the mean length data are shown in Figure 3.2. Total mortality rates were first estimated independently for each of the three snapper species under the assumption of one change in mortality (i.e., two mortality rates estimated per species) with the SSM. For all three species, the data suggested a decrease in total mortality during the observed time series (Table 3.4). However, the change points varied widely: the year 1996.8 for Silk Snapper, 1985.9 for Blackfin Snapper, and 1997.7 for Vermilion Snapper. Change points are estimated in continuous time with the decimal representing tenths of a year.

For Blackfin Snapper, the equilibrium *Z* in the data was estimated as 0.46 when no change point was specified; this value was intermediate to the two mortality rates that were estimated with one change point. Using AIC_c, there was almost equal support for the equilibrium model and the one change-point model, but the former produced a slight trend in the residual fit. Thus, we proceeded with the analysis for Blackfin Snapper under an assumption of one change point.

Next, total mortality rates were estimated by assuming a common change point using MSM1. The estimated mortality rates for Silk Snapper and Vermilion Snapper were virtually unchanged, and the estimated common change point in year 1997.5 did not

noticeably depart from the change points generated by the SSM for Silk Snapper and Vermilion Snapper. For Blackfin Snapper, the first mortality rate Z_1 was estimated to be lower and the change point occurred much later in MSM1 than in the SSM. The estimated second mortality rate Z_2 for Blackfin Snapper was unchanged between the two models.

Multispecies model 3 estimated an earlier common change point (i.e., 1994.8) than was estimated by MSM1. The corresponding common value of δ was estimated to be 0.52. A decrease in mortality over time was still inferred, but the fit to the data was distinctly different from that of the SSM and MSM1 (Figure 3.2). For Silk Snapper and Blackfin Snapper, values of Z_2 from MSM3 were virtually unchanged compared to those from the SSM and MSM1, whereas values of Z_1 varied (Table 3.4). For Vermilion Snapper, Z_1 was lower, and a smaller reduction in mortality during 1994 was inferred.

Multispecies model 1 was the best-fitting model among the three, closely followed by the SSM (Table 3.5). There was little support for estimating common changes in fishing mortality, as the ΔAIC_c value for MSM3 was more than 10 units. Between MSM1 and SSM, there was not much support for estimating additional parameters (i.e., species-specific change points).

As an indicator of model certainty and the benefit of using multiple species to infer trends in mortality, we examined the asymptotic SE of the change point in the models (Table 3.4). For MSM1, the asymptotic SE of the change point was 0.84, compared to 2.26 for MSM3 and a mean SE of 1.77 from the three species-specific

change points in the SSM. The likelihood profile for the change point was also used to examine the contribution of the data from each species to the goodness of fit in MSM1 (Figure 3.3). The likelihood profile indicated that the data for Vermilion Snapper contributed the most information for estimating the change point. Together, these diagnostics suggested that the mean length data for the three species were best modeled by a common change point and independent trends in mortality.

3.4.2. Sensitivity Analysis of Natural Mortality Specification

The *%DEV* values of the estimated total mortality rates for the three snapper species were all less than 15%, indicating that the estimates in MSM3 were not considerably affected by the specification of *M* (Table 3.6). There was no consistent trend in the estimated *Z* across the range of *M* values (Figure 3.4). For Silk Snapper and Blackfin Snapper, there was little to no trend in the estimation of Z_2 in the sensitivity analysis, whereas more variability and a slight trend were observed in Z_1. For Vermilion Snapper, both Z_1 and Z_2 were generally more variable with some trend given M, although the trends were in opposite directions for the two total mortality rates (Figure 3.4). In comparison with the SSM and MSM1, the ΔAIC_c values obtained from the sensitivity analysis for MSM3 were all greater than 6.2 and therefore did not affect model selection.

3.5. Discussion

There is interest in developing methods for multispecies assessments, especially for ensembles of data-rich and data-poor stocks or species (Punt et al. 2011). Here, we

have demonstrated the application of the mean length mortality estimator in a multispecies context.

Mean length models that incorporate common trends in mortality can use data for well-sampled species to inform trends in species with fewer observations. This can be accomplished because sample sizes are incorporated into the likelihood function to give more weight to the trends in species with more observations. In our example, the length data for Vermilion Snapper were valuable for estimating the change point. The sample sizes for Vermilion Snapper were relatively large and consistent for the duration of the time series, whereas the observations of Blackfin Snapper were historically sparser and the sample sizes for Silk Snapper were low before the year 2000.

When a common change point was estimated for all three snapper species, the trends in mortality for Blackfin Snapper and, to a lesser extent for Silk Snapper, seem to have "borrowed strength" (Punt et al. 2011) from those for Vermilion Snapper. This behavior can be diagnosed with the likelihood profile for the change point in MSM1. As a result, there was higher confidence in the timing of the common change point than in the timing of the individual change points estimated by the SSM. More credibility can be given for the change point in Blackfin Snapper to occur synchronously with those of the other two species in MSM1. The SSM did not considerably improve the goodness of fit relative to MSM1 because the estimated mortality rates for Silk Snapper and Vermilion Snapper and the most recent mortality rate for Blackfin Snapper were very similar between the two models. This was reflected in the ΔAIC_c, which did not indicate more support for the estimation of separate change points.

The data did not provide support for modeling synchronous changes in fishing mortality with MSM3, and the most noticeable change in model fit was observed with the Vermilion Snapper data. From the SSM and MSM1, total mortality was inferred to be much higher in Vermilion Snapper than in Silk Snapper and Blackfin Snapper, which implies that the changes in total mortality occurred with different proportional changes in fishing mortality in Vermilion Snapper relative to the other two species since the *M* values for all three species were assumed to be very similar. Estimating a common proportional change in fishing mortality altered the goodness of fit to the data. There was not enough information (i.e., more than one change point) in the data to explicitly model species-specific effects with MSM2. Thus, MSM1 provided the empirically best fit to the data in our application.

The sensitivity analysis of MSM3 in our snapper example showed that the range in *Z*-estimates was smaller than that specified for *M*. This behavior occurs because the model fit is determined by Z, and any overestimation of M is expected to be compensated for by an underestimation of fishing mortality and vice versa. Since values of *M* are obtained externally to the models and can vary widely depending on the empirical method used to derive those values (Hamel 2015), we recommend sensitivity analyses for applications of MSM2 or MSM3 and any possible effects on model selection.

For Vermilion Snapper, there was some uncertainty in the rather large magnitude of the estimated mortality rates. Examination of the length frequency histogram showed a truncated length structure, which could also arise from the overestimation of L_∞ specified in the model or from size-selective fishing that did not catch large animals. Identifying

the cause of the truncation will require more information on the life history of the stock. As long as they are time invariant, these factors should not affect the fact that the mortality rate decreased (Gedamke and Hoenig 2006). However, knowledge of the absolute magnitude of mortality is still desirable for management purposes. This case study highlights the importance of choosing the appropriate life history parameters for mean length mortality estimators and for data-limited assessment methods in general. In our study, the growth function for Vermilion Snapper was obtained from a different stock. A high-quality growth study of the Vermilion Snapper stock in Puerto Rican waters would have conferred greater certainty of the causes underlying the large total mortality estimates in our analysis.

The advantage of using the mean length mortality estimator on a multispecies basis is to identify concomitant trends in mortality. The work we have presented can be further modified to account for particular exploitation patterns. For example, for a given species complex, certain change points can be synchronous while others can remain independent if there is information available to guide that model specification. A switch in the fishery's target species or a regulation to reduce total effort in the fishery may produce a synchronous change in mortality, whereas the opening of a new fishery may elicit an independent change in mortality. The general framework presented here for expanding the mean length mortality estimator to multiple species by using models of varying complexity and synchrony in mortality trends allows for improved inference for one species within the context of companion species in a complex, especially if the companion species are better sampled and studied.

3.5.1. Selection of the Minimum Length of Vulnerability to the Fishery

The Beverton–Holt nonequilibrium method presented here assume time-invariant, knife-edge selection in the observed length data. However, selectivity is more likely to be logistic, with partially selected small animals in the length frequency distribution. To conform to the assumption of knife-edge selection in the model, a value of L_c is chosen for use in calculating the $\overline{L}$ of the truncated distribution. In a stock that is fully exploited or that has a high *M/K* ratio, the length distribution above L_c is monotonically descending (Figure 5.13 of Pauly 1984; Figure 7 of Hordyk et al. 2015). Thus, the choice of modal length for L_c would be appropriate, with lengths smaller than the modal length assumed to be incompletely selected. However, a stock that is lightly exploited and that has a low *M/K* ratio will have a modal length near L_∞, with fully selected animals comprising a significant portion of the ascending limb to the left of the mode. In this scenario, the modal length would not be suitable if the stock is exploited over a large size range; an L_c value smaller than the modal length would be more appropriate. Based on the modal length relative to L∞ over the time series, the modal length was used as the L_c, assuming that the length distributions for all three deepwater snapper species reflected fully exploited stocks (Figure 3.1).

For Silk Snapper, an increase in the modal length was detected from the annual length frequency distributions in the mid-2000s, whereas the modal lengths for Blackfin Snapper and Vermilion Snapper were more stable over time (Figure 3.5). This shift in selectivity may have resulted from the temporary 16-in (406.4 mm) minimum size limit implemented in 2004–2006 (SEDAR 2011). The large modal lengths for Silk Snapper in 2005 and 2006 reflect this management regulation. A smaller shift in selectivity appeared

to have persisted after the regulation was repealed. The mortality analysis presented here parameterized L_c to reflect the most recent selectivity pattern for Silk Snapper. The size limit regulation did not appear to affect the mortality estimation procedure, as no change in mortality was inferred to have occurred when the regulation was in effect.

Nonetheless, to examine the implications of a change in modal length for Silk Snapper, we analyzed the sensitivity of the mortality estimates to the chosen value of L_c by using the SSM under the assumption of one change point (Figure 3.6). The early mortality rate Z_1 fluctuated widely because the high values of L_c compared to the modal lengths early in the time series led to the truncation of large proportions of the length data. On the other hand, the recent mortality rate Z_2 was relatively stable when values near the recent modal length were used for L_c. For a situation in which selectivity has changed for a stock, it may be preferable to configure the mean length mortality estimator to reflect the most recent conditions so as to estimate current exploitation while allowing for greater uncertainty when inferring past trends in mortality.

3.5.2. Other Assumptions and Considerations

Improvements in fishing technology may also increase the spatial extent of exploitation of fish stocks. With serial depletion of coastal stocks, fishing effort generally moves further offshore over time as inshore areas become depleted. Changes in mean length could be a result of changes in targeting rather than changes in mortality. The ability to detect serial depletion requires high-resolution spatiotemporal data (Walters 2003; Cardinale et al. 2011). However, data on the location of catch were generally

unavailable in the TIP database (SEDAR 2009). If a fishery has expanded spatially over time, one could restrict the mean length analysis to areas fished within certain depth strata or spatial strata if such data are available. If the stock actually consists of several distinct units with little interchange, then separate analyses would allow for estimation of the mortality experienced within each stratum. However, if individual fish inhabit different depths as they grow, then the observed size range in the strata would also reflect movement with age in addition to mortality. In such a scenario, effective management of the entire resource would have to rely on expert judgment to determine the extent of the spatial expansion and appropriate management measures. Analyses of additional data types, such as CPUE, can provide further insight on serial depletion (Cardinale et al. 2011).

Since the models we have examined assume that recruitment is constant, large pulses of recruitment can confound estimates of mortality. Although changes in mean length alone cannot differentiate between a recruitment pulse and a change in mortality, the length frequency distribution may provide more information. A large recruitment cohort will progress through the length structure of the catch over time; this cohort would cause the mean length to temporarily decrease. On the other hand, if there is a poor recruitment year-class, a gap in the size distribution may be observed and will also progress over time. It is important to ensure that changes in mortality are not inferred from the model when large or small year-classes are concurrently observed. Variability in recruitment may elicit small trends in mean length (Gedamke and Hoenig 2006), but model selection using AIC_c provides the balance between detection of long-term changes

in mortality and spurious changes in mortality due to overfitting; the latter may be confounded with recruitment.

3.6. References

Baranov, F. I. 1918. On the question of the biological basis of fisheries. Nauchn. Issled. Ikhtiologicheskii Inst. Izv. 1:81-128.

Beverton, R. J. H., and S. J. Holt. 1956. A review of methods for estimating mortality rates in fish populations, with special reference to sources of bias in catch sampling. Rapports et Procès-verbaux des Reunions, Conséil International Pour L'Exploration de la Mer 140:67-83.

Beverton, R. J. H., and S. J. Holt. 1957. On the dynamics of exploited fish populations. U.K. Ministry of Agriculture, Fisheries, Food, and Fishery Investigations Series II, Vol. XIX.

Boardman, C., and D. Weiler. 1980. Aspects of the Life History of Three Deepwater Snappers around Puerto Rico. Proceedings of the Gulf and Caribbean Fisheries Institute 32:158-172.

Burnham, K. P., and D. R. Anderson. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd edition. Springer-Verlag, New York, New York.

Cardinale, M., J. Hagberg, H. Svedang, V. Bartolino, T. Gedamke, J. Hjelm, P. Borjesson, and F. Noren. 2009. Fishing through time: population dynamics of plaice (*Pleuronectes platessa*) in the Kattegat-Skagerrak over a century. Population Ecology 52:251-262.

Cardinale, M., Nugroho, D., and P. Jonson. 2011. Serial depletion of fishing grounds in an unregulated, open access fishery. Fisheries Research 108:106-111.

Claro, R., J. A. Basire, K. C. Lindeman, and J. P. Garcia-Arteaga. 2001. Cuban Fisheries: Historical Trends and Current Status. Pages 194-219 *in* R. Claro, K. C. Lindeman, and L. R. Parenti, editors. Ecology of the Marine Fishes of Cuba. Smithsonian Institution Press, Washington, D.C.

Espinosa, L., and E. Pozo. 1982. Edad y crecimiento del sesí (*Lutjanus buccanella* [Cuvier, 1828]) en la plataforma suroriental de Cuba. Revista Cubana de Investigaciones Pesqueras 7:80-100.

Gedamke, T. and J. M. Hoenig. 2006. Estimating Mortality from Mean Length Data in Non-equilibrium Situations, with Application to the Assessment of Goosefish. Transactions of the American Fisheries Society 135:476-487.

Hamel, O. S. 2015. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. ICES Journal of Marine Science 72:62-69.

Hordyk, A., K. Ono, K. Sainsbury, N. Loneragan, and J. Prince. 2015. Some explorations of the life history ratios to describe length composition, spawning-per-recruit, and the spawning potential ratio. ICES Journal of Marine Science 72:204-216.

Manickchand-Heileman, S. C., and D. A. T. Phillip. 1999. Contribution to the biology of the vermilion snapper, *Rhomboplites aurorubens*, in Trinidad and Tobago, West Indies. Environmental Biology of Fishes 55:413-421.

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

Punt, A. E., D. C. Smith, and A. D. M. Smith. 2011. Among-stock comparisons for improving stock assessments of data-poor stocks: the "Robin Hood" approach. ICES Journal of Marine Science 68:972-981.

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

SEDAR (SouthEast Data, Assessment, and Review). 2009. Caribbean Fisheries Data Evaluation: SEDAR Procedures Workshop 3. January 26 – 29, 2009. SEDAR, North Charleston, South Carolina.

SEDAR (SouthEast Data, Assessment, and Review). 2011. SEDAR 26 Stock Assessment Report: U.S. Caribbean Silk Snapper. SEDAR, North Charleston, South Carolina.

Sylvester, J. R. 1974. A preliminary study of the length composition, distribution and relative abundance of three species of deepwater snappers from the Virgin Islands. Journal of Fish Biology 6:43-49.

Then, A. Y., J. M. Hoenig, T. Gedamke, and J.S. Ault. 2015a. Comparison of Two Length-Based Estimators of Total Mortality: a Simulation Approach. Transactions of the American Fisheries Society 144:1206-1219.

Then, A. Y., J. M. Hoenig, N. G. Hall, and D. A. Hewitt. 2015b. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES Journal of Marine Science 72:82-92.

USOFR (U.S. Office of the Federal Register). 2005. Fisheries of the Caribbean, Gulf of Mexico, and South Atlantic; Comprehensive Amendment to the Fishery Management Plans of the U.S. Caribbean. Federal Register 70:208(28 October 2005):62073–62084.

Walters, C. 2003. Folly and fantasy in the analysis of spatial catch rate data. Canadian Journal of Fisheries and Aquatic Sciences 60:1433-1436.

3.7. Tables

Table 3.1. Number of estimated parameters for the single-species model (SSM) and multispecies models (MSM1, MSM2, and MSM3), where N is the number of species, I is the number of change points (years during which a change in total mortality occurred), and $I + 1$ is the number of estimated mortality rates. Values include the estimated residual variance for each species.

	SSM	**MSM1**	**MSM2**	**MSM3**
General formula	$2N(I+1)$	$N+I+N(I+1)$	$3N+2I-1$	$2(N+I)$
One change point	$4N$	$3N+1$	$3N+1$	$2N+2$
Additional parameters with:				
Additional species	4	3	3	2
Additional change point	$2N$	N+1	2	2

Table 3.2. Von Bertalanffy growth parameters (K = Brody growth coefficient; L_∞ = asymptotic length) for the three deepwater snapper species.

Species	***K***	***L*∞ (mm)**	**Source**
Silk Snapper	0.10	794	SEDAR (2011)
Blackfin Snapper	0.10	635	Espinosa and Pozo (1982)
Vermilion Snapper	0.13	532	Manickchand-Heileman and Phillip (1999)

Table 3.3. Estimates of the length at full fishery selectivity (L_c), which was used to calculate mean lengths and natural mortality rates (M) for the three deepwater snapper species.

Species	L_c (mm)	M
Silk Snapper	310	0.18
Blackfin Snapper	250	0.19
Vermilion Snapper	220	0.25

Table 3.4. Estimates (SE in parentheses) of the total instantaneous mortality rate (Z) and change points (years during which a change in total mortality occurred; Z_1 = total mortality before the change point; Z_2 = total mortality after the change point) from application of the single-species model (SSM) and multispecies models 1 and 3 (MSM1 and MSM3) for the three deepwater snapper species. The proportional change in fishing mortality (i.e., δ) for MSM3 was estimated as 0.52 (SE = 0.08).

Parameter	SSM	MSM1	MSM3
Silk Snapper			
Z_1	0.63 (0.08)	0.62 (0.07)	0.76 (0.09)
Z_2	0.49 (0.02)	0.49 (0.02)	0.49 (0.02)
Change point	1996.8 (2.93)	1997.5 (0.81)	1994.8 (2.26)
Blackfin Snapper			
Z_1	0.77 (0.28)	0.50 (0.05)	0.60 (0.06)
Z_2	0.43 (0.03)	0.43 (0.03)	0.41 (0.03)
Change point	1985.9 (1.66)	1997.5 (0.81)	1994.8 (2.26)
Vermilion Snapper			
Z_1	1.89 (0.27)	1.90 (0.27)	1.39 (0.16)
Z_2	0.60 (0.06)	0.61 (0.06)	0.85 (0.09)
Change point	1997.7 (0.74)	1997.5 (0.81)	1994.8 (2.26)

Table 3.5. Difference in Akaike's information criterion corrected for small sample sizes (ΔAIC_c) from application of the single-species model (SSM) and multispecies models 1 and 3 (MSM1 and MSM3) to the three deepwater snapper species.

Model	ΔAIC_c	Parameters
Multispecies Model 1	0.0	10
Single Species Model	2.2	12
Multispecies Model 3	11.9	8

Table 3.6. Range of percent deviation (*%DEV*; min = minimum; max = maximum) in estimates of total mortality (Z_1 = total mortality before the change point; Z_2 = total mortality after the change point) from the sensitivity analysis of natural mortality specification in Multispecies model 3 as applied to the three deepwater snapper species.

	Silk Snapper		**Blackfin Snapper**		**Vermilion Snapper**	
	%DEV		***%DEV***		***%DEV***	
Parameter	**Min.**	**Max.**	**Min.**	**Max.**	**Min.**	**Max.**
Z_1	-10.9	13.8	-6.9	7.3	-6.3	10.7
Z_2	-1.9	1.6	-2.2	2.7	-13.3	10.9

3.8. Figures

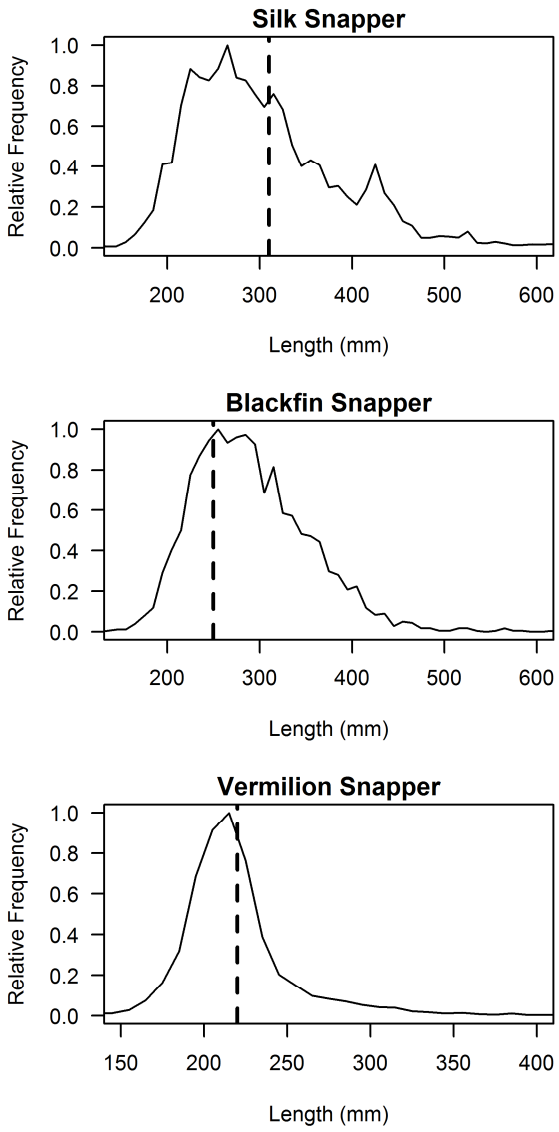

Figure 3.1. Length frequency histograms for the three deepwater snapper species captured in the Puerto Rican handline fishery from 1983 to 2013. Dashed vertical lines indicate the length of full selectivity (L_c), above which the annual mean lengths were calculated for the multispecies models.

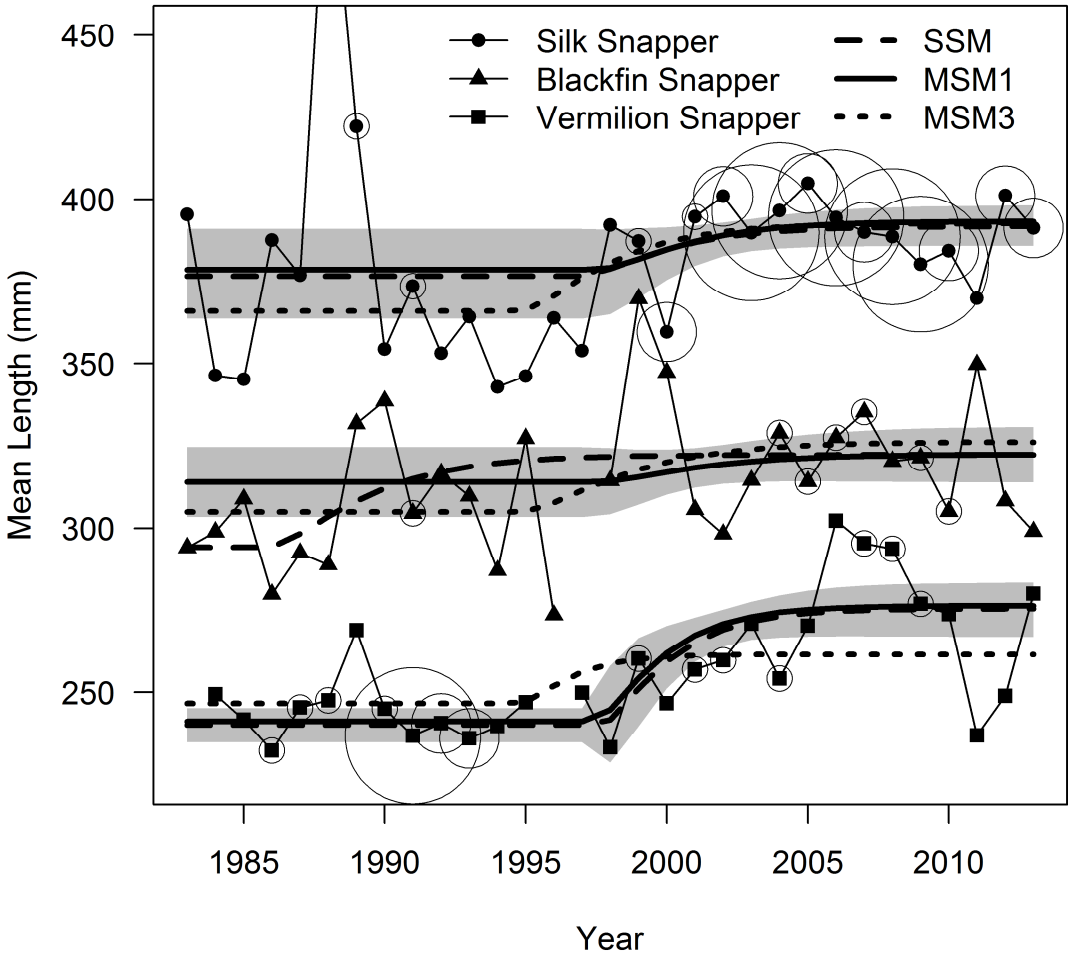

Figure 3.2. Observed (points and thin lines) and predicted mean lengths (bold lines) from the Single Species Model (SSM), Multispecies Model 1 (MSM1), and Multispecies Model 3 (MSM3) for Silk Snapper, Blackfin Snapper, and Vermilion Snapper. The grey shaded region indicates the 95% confidence interval of the predicted mean length from Multispecies Model 1 using the derived asymptotic SEs. Concentric circles indicate the annual sample size of observed lengths (small circles = 100-249, medium circles = 250-499, large circles = 500 or more). No circles were drawn for sample sizes less than 100. The observed mean length in 1988 for Silk Snapper (514 mm from 29 samples) is not shown but was used in the analysis.

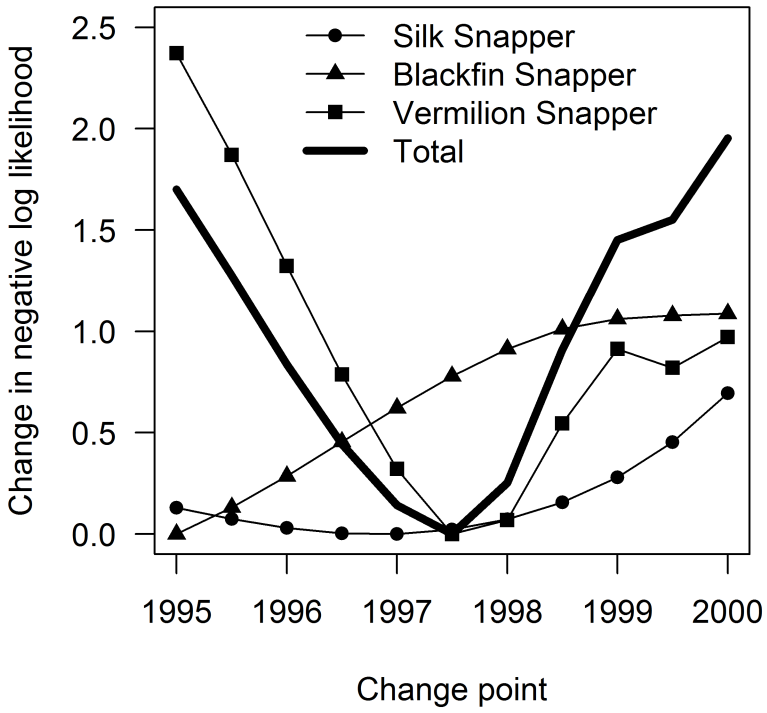

Figure 3.3. Likelihood profile for the change point (year during which the change in mortality occurred) from Multispecies Model 1 in the application to the Puerto Rican deepwater snapper complex.

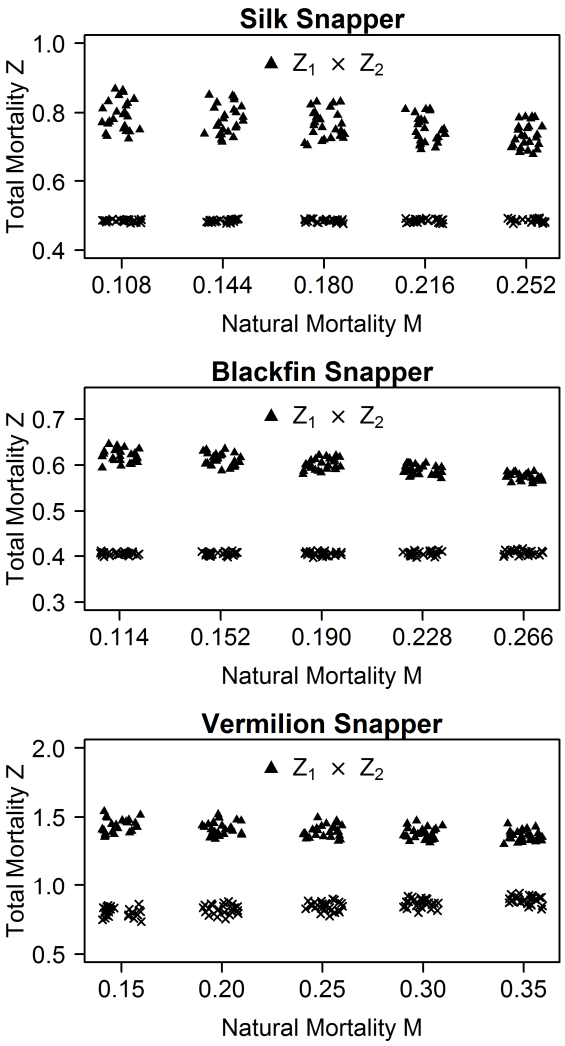

Figure 3.4. Estimates of the total mortality rate (Z) for the three deepwater snapper species based on the sensitivity analysis of Multispecies Model 3 to different specified values of natural mortality (M). The x-axis is jittered to enhance visibility of the Z-values obtained in each "bin" of M.

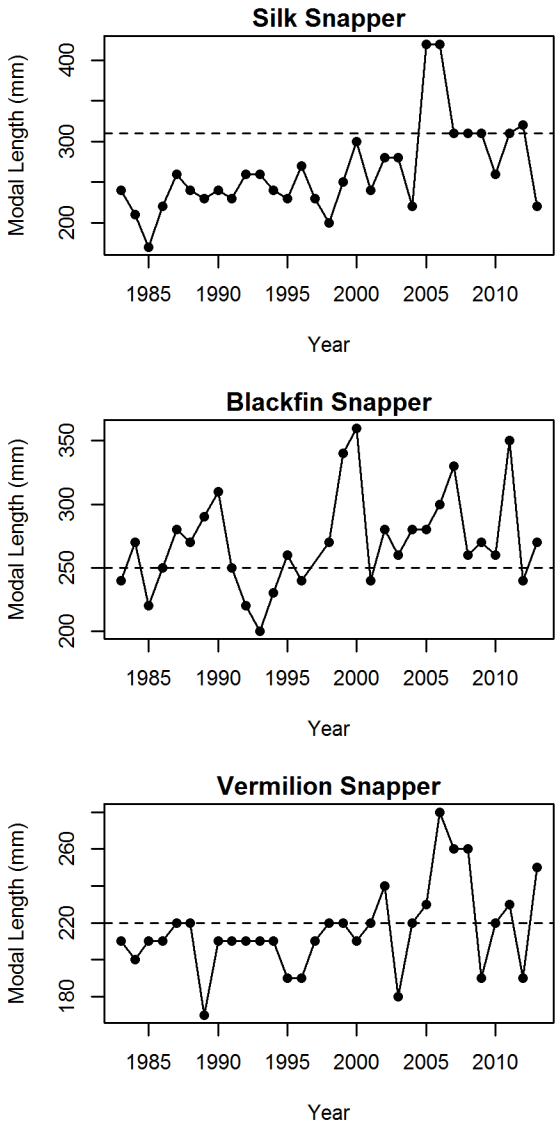

Figure 3.5. Modal lengths from annual length frequency distributions for the three Puerto Rican deepwater snapper species. The dashed horizontal line in each panel shows the length of full fishery selectivity (L_c), which was used for mortality estimation.

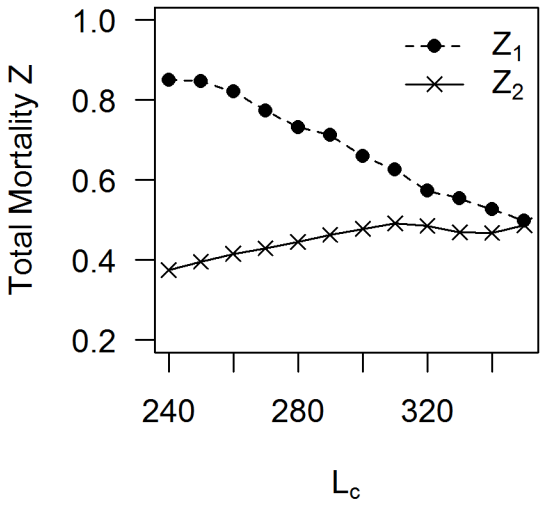

Figure 3.6. Sensitivity analysis of the estimated total mortality rates (Z_1 = total mortality before the change point; Z_2 = total mortality after the change point) in Silk Snapper when different lengths at full fishery selectivity (L_c) are used.

Chapter 4: Estimating Total Mortality Rates from Mean Lengths and Catch Rates in Nonequilibrium Situations

4.1. Abstract

A series of estimates of the total mortality rate (Z) can be obtained by using the Beverton–Holt nonequilibrium-based approach of Gedamke and Hoenig (2006) on observations of population mean length over time (ML model). In contrast, only relative mortality rates (not absolute values) can be obtained from a time series of catch rates. We derived the transitional behavior of the catch rate following a change in total mortality in the population. From this derivation, we developed a new method to estimate Z that utilizes both mean lengths and catch rates (MLCR model). Both the ML model and the MLCR model assume constant recruitment in the population. We used a simulation study to test performance when recruitment is variable. Simulations over various scenarios of Z and recruitment variability showed that there may be correlated residuals in the mean lengths and catch rates arising from fluctuations in recruitment. However, the root mean square errors of the Z estimates and the change point (i.e., the year when mortality changed) were smaller in the MLCR model than in the ML model, indicating that the MLCR model can better account for variable recruitment. Both methods were then applied to Mutton Snapper *Lutjanus analis* in Puerto Rico to illustrate their potential application to assess data-limited stocks. The ML model estimated an increase in Z, but the MLCR model also estimated a subsequent reduction in Z when the catch rate data were considered.

4.2. Introduction

Beverton and Holt (1956, 1957) developed an approach to estimating the instantaneous total mortality rate (Z; year^{-1}) from the mean length in the population and information on growth rates and selectivity. The estimator is based on the assumptions that the population is in equilibrium and that recruitment is constant. Despite these stringent assumptions, the estimator has found widespread usage because of its minimal data requirements. Gedamke and Hoenig (2006) derived the transitional behavior of the mean length statistic following stepwise changes in mortality rate over time and developed an estimator for period-specific mortality rates. The required information for applying this extension of the Beverton and Holt equation to nonequilibrium situations includes the von Bertalanffy growth parameters (L_∞ and K), the length of first capture (L_c; the smallest size at which animals are fully vulnerable to the fishery and sampling gear), and a time series of mean length ($\bar{L}$) of animals above the L_c. The methodology and applications to Goosefish *Lophius americanus* and to Barndoor Skate *Dipturus laevis* were described by Gedamke and Hoenig (2006) and Gedamke et al. (2008), respectively.

This approach can be generalized to integrate additional data types when they are available. For example, Gedamke et al. (2008) relaxed the assumption of constant recruitment by incorporating an index of recruitment in the model, and Then et al. (in press) incorporated information on fishing effort. Here, we develop a model to incorporate a time series of catch rates (indices of abundance) into the mean length estimator to better detect changes in mortality and to better estimate Z compared to using the mean length-only estimator.

In the absence of information on recruitment, the mortality estimators using only mean lengths (ML model) and mean lengths plus catch rates (MLCR model) assume constant recruitment over time. Exploited fish populations can exhibit high interannual variability in recruitment, and while Gedamke and Hoenig (2006) demonstrated the effect of a weak or failed year-class on the mean length statistic over time, the effects of variable recruitment on estimating mortality have not been examined. Here, we evaluate via simulation the effect of variable recruitment on mortality estimation in the two models, which assume constant recruitment. We focus on recruitment because it is often identified as the largest source of variability in fish populations (Thorson et al. 2014).

In this study, we derive the transitional behavior of catch rates expressed in terms of either abundance (number per unit effort [NPUE]) or biomass (weight per unit effort [WPUE]) to develop the MLCR model. Next, we evaluate the effect of variable recruitment on the performance of the ML model and the MLCR model in estimating the parameters of interest: Z and the years when Z changed. In addition to standard performance metrics of bias and root mean square error, we examine correlations in paired residuals of mean length and catch rate and runs in the sign (positive or negative) of residuals. Finally, we apply both the ML model and the MLCR model to estimate historical values of Z for Mutton Snapper *Lutjanus analis* in Puerto Rico.

4.3. Methods

4.3.1. Relationship between the Catch Rate and Mortality Rate

Consider what information about Z can be obtained from catch rates in the simplest scenario when we assume equilibrium conditions, including constant recruitment

(R) and a constant catchability coefficient (q). We assume that the abundance-based catch rate, NPUE, is a function of the probability of an individual being captured by 1 unit of effort (q) and the abundance N:

$$NPUE = qN \ . \tag{4.1}$$

Under equilibrium, abundance will be related to total mortality as follows (Ricker 1975):

$$N = \int_{t_c}^{\infty} R \cdot \exp[-Z(t - t_c)]dt = \frac{R}{Z}, \tag{4.2}$$

where $R \cdot \exp[-Z(t - t_c)]$ is the abundance at age t; and t_c is the age at which animals are fully selected by the fishing gear, corresponding to length L_c via the von Bertalanffy growth equation. With integration, parameter t_c drops out of the equation and is not used here. We substitute equation (4.2) into equation (4.1),

$$NPUE = \frac{qR}{Z} = \frac{\tilde{q}}{Z}, \tag{4.3}$$

where $\tilde{q}$ is a scaling parameter that is the product of q and R. Two equilibrium catch rates, $NPUE_1$ and $NPUE_2$, corresponding to time periods with mortality rates Z_1 and Z_2, can provide estimates of the relative change in Z if $\tilde{q}$ is assumed to be constant. Thus,

$$NPUE_1 = \frac{\tilde{q}}{Z_1} \tag{4.4a}$$

and

$$NPUE_2 = \frac{\tilde{q}}{Z_2} \tag{4.4b}$$

imply that the ratio of NPUE is an estimate of the ratio of mortality rates,

$$\frac{NPUE_2}{NPUE_1} = \frac{Z_1}{Z_2}. \tag{4.5}$$

In real-world situations, obtaining reliable estimates of both *q* and *R* is extremely rare; therefore, it is unlikely that absolute values of *Z* can be estimated from catch rates alone. However, there is information on relative mortality rates, and we can incorporate this information into the length-based nonequilibrium mortality estimator of Gedamke and Hoenig (2006).

In a nonequilibrium framework, overall abundance and the corresponding NPUE will not respond instantaneously to changes in *Z*. Equation (4.3) will only reflect the new mortality rate when enough time has passed for the new equilibrium age structure to be established. Assume that *d* years have elapsed since a change in mortality from Z_1 to Z_2. The nonequilibrium NPUE will be equal to

$$NPUE(Z_1, Z_2, d) = \widetilde{q} \cdot \widetilde{N}(Z_1, Z_2, d), \tag{4.6}$$

where the NPUE and $\widetilde{N}$ are now functions of the mortality rates and the time elapsed since the change in mortality. Using the derivations from Gedamke and Hoenig (2006), the relative abundance $\widetilde{N}(Z_1, Z_2, d)$, after dividing out recruitment *R*, has two components and can be expressed as

$$\widetilde{N}(Z_1, Z_2, d) = \int_{t_c}^{t_c+d} \exp\left[-Z_2\left(t - t_c\right)\right]dt + \int_{t_c+d}^{\infty} \exp\left[-Z_2 d\right]\exp\left\{-Z_1\left[t - \left(t_c + d\right)\right]\right\}dt. \tag{4.7}$$

In equation (4.7), the first integral represents fish recruited after the change in mortality; these animals have only experienced mortality rate Z_2 and are of ages t_c to $t_c + d$. The second integral represents fish that were recruited before the change in mortality; these fish have experienced both the old (Z_1) and the new (Z_2) mortality rates and are of ages $t_c + d$ and older. The implications of equation (4.7) can be envisioned by considering

the response of NPUE when Z hypothetically doubles (Figure 4.1). The NPUE drops rapidly after the change in mortality rate and then approaches the new asymptotic value, which is half of the starting value.

As shown by Gedamke and Hoenig (2006, their Appendix 1), after integration and simplification, equation (4.7) becomes

$$\widetilde{N}(Z_1, Z_2, d) = \frac{1-\exp(-Z_2 d)}{Z_2} + \frac{\exp(-Z_2 d)}{Z_1}. \tag{4.8}$$

Equations (4.7) and (4.8) can then be modified to incorporate any number of changes in mortality (see Gedamke and Hoenig 2006: their equation A.2.2). The corresponding derivation for the behavior of WPUE is provided in Appendix B.

4.3.2. Integrating Mean Lengths and Catch Rates in a Model

Using the transitional behavior of the mean length and catch rate, we construct a likelihood-based model to estimate Z and change points (the calendar years when the mortality rate changed) from a time series of mean lengths and catch rates. The assumptions of the MLCR model include those in the ML estimator as described by Gedamke and Hoenig (2006), but additionally it is assumed that the NPUE and WPUE are proportional to population abundance and biomass, respectively, by a scaling coefficient $\tilde{q}$. Given k changes in mortality, maximum likelihood estimation is used to estimate the vector of $k + 1$ total mortality rates (denoted by $\underline{Z} = \{Z_1, Z_2, ..., Z_{k+1}\}$) and the vector of k change points (denoted by $\underline{D} = \{D_1, D_2, ..., D_k\}$) that best predict the observed data. We construct the joint log-likelihood function, $\ln \Lambda(\underline{Z}, \underline{D})$, of the MLCR model to be proportional to

$$\ln \Lambda(\underline{Z},\underline{D}) \propto \ln \lambda_{\bar{L}}(\underline{Z},\underline{D}) + \ln \lambda_I(\underline{Z},\underline{D}), \tag{4.9}$$

where $\ln \lambda_{\bar{L}}(\underline{Z},\underline{D})$ and $\ln \lambda_I(\underline{Z},\underline{D})$ are the log-likelihoods of the mean lengths and catch rates (either in NPUE or WPUE), respectively. Assuming a normal distribution for the annual observed mean lengths (Gedamke and Hoenig 2006), the log-likelihood of the mean lengths is proportional to

$$\ln \lambda_{\bar{L}}(\underline{Z},\underline{D}) \propto -n_L \ln \sigma_L - \frac{1}{2\sigma_L^2} \sum_{y=1}^{n_L} m_y (\bar{L}_y - \hat{\bar{L}}_y)^2 , \tag{4.10}$$

where y indexes year; n_L is the number of years with mean length observations; $\bar{L}_y$ and $\hat{\bar{L}}_y$ are the observed and predicted mean lengths, respectively, of animals larger than L_c in year y; m_y is the sample size of observed lengths above L_c in year y; and σ_L^2 is the variance of lengths. The log-likelihood of the catch rates, assuming either a normal or lognormal distribution, is proportional to

$$\ln \lambda_I(\underline{Z},\underline{D}) \propto -n_I \ln \sigma_I - \frac{1}{2\sigma_I^2} \sum_{y=1}^{n_I} (I_y - \hat{I}_y)^2 \tag{4.11a}$$

or

$$\ln \lambda_I(\underline{Z},\underline{D}) \propto -n_I \ln \sigma_I - \frac{1}{2\sigma_I^2} \sum_{y=1}^{n_I} (\ln I_y - \ln \hat{I}_y)^2 , \tag{4.11b}$$

respectively, where n_I is the number of years with catch rate observations; I_y and $\hat{I}_y$ are the observed and predicted catch rates, respectively, in year y; and σ_I^2 is the catch rate variance in either normal (equation 4.11a) or log-transformed (equation 4.11b) space.

Equation (4.9) can be maximized to produce the asymptotically most efficient (maximum likelihood) estimates with SEs and confidence intervals generated for the estimated mortality rates ($\hat{\underline{Z}}$) and change points ($\hat{\underline{D}}$), where the circumflex (^) denotes an estimate. A grid search over change points is recommended to identify and avoid local extrema in the log-likelihood function. The required life history information for the MLCR model includes the von Bertalanffy parameters L_∞ and K when NPUE is modeled, whereas the allometric exponent b from the length–weight relationship is also required when WPUE is modeled (Appendix B). Compared to the ML model with the same number of change points, two additional parameters are estimated for the MLCR model: the catch rate scaling coefficient ($\tilde{q}$) and the catch rate SD (σ_I). Different numbers of change points can be specified, with model selection procedures used to identify the best-fitting model (Burnham and Anderson 2002). In this study, we use values of Akaike's information criterion with correction for small sample sizes (AIC_c) and identify the best-fitting model as the one having the smallest value (i.e., Akaike difference [ΔAIC_c] = 0), with less support for models with larger AIC_c values.

4.3.3. Simulation Study of the Mortality Estimators

Effect of variable recruitment on the mean length and catch rate.—To illustrate the dynamics of the mean length and catch rate relative to changes in recruitment, consider an age-structured population in which recruitment varies stochastically with a constant mean and variance, while Z is constant (Figure 4.2). During periods of poor recruitment relative to the mean, the mean length increases and the catch rate decreases as fewer small animals recruit to the fishery. This pattern produces positive and

negative residuals in the mean lengths and catch indices, respectively, from values predicted under constant recruitment. Similarly, the patterns reverse in periods of relatively good recruitment. We observe two general patterns in the residuals of mean lengths and catch rates. First, variable recruitment produces opposing effects in the residuals of the two data types, creating a negative correlation between the paired residuals. Second, the trends in the residuals can persist as the relative strength of a cohort progresses through the age structure of the population over time.

Simulation design.—To examine the implications of using models that assume constant recruitment, we implemented a simulation study with variable recruitment in the population while meeting the other assumptions. The goals of the simulation were to (1) compare the performance of the MLCR mortality estimator relative to that of the ML estimator, (2) compare the performance of both models with variable recruitment in the population, and (3) provide guidance on interpreting the behavior of both models under variable recruitment.

In the simulation model, an age-structured population was constructed,

$$N_{t,y} = \begin{cases} R_y & t = t_c \\ N_{t-1,y-1}\exp(-Z_{y-1}) & t = t_c + 1,\ldots,t_{\max} \end{cases}, \tag{4.12}$$

where $N_{t,y}$ is the abundance at age t in year y; R_y is the recruitment of animals of age t_c in year y; Z_{y-1} is the instantaneous total mortality rate in year $y-1$; and t_{max} is the maximum age. Recruitment followed a lognormal distribution,

$$R_y = \exp(\varepsilon_R - 0.5\sigma_R^2), \tag{4.13}$$

where $\varepsilon_R \sim N(0,\sigma_R^2)$ are normally distributed deviations in log space. The expected median recruitment was equal to 1.0 since the magnitude was not relevant in the

simulation. To simulate the continuous recruitment assumed in the models, monthly cohorts were created with monthly time steps. Recruitment within each calendar year was held constant.

The population was projected for 20 years. A factorial design for the simulation was created across four values of interannual recruitment variability (σ_R) and four mortality scenarios (A–D), with a stepwise change in mortality at the beginning of year 11 (Table 4.1). The mean length ($\bar{L}_y$) and abundance-based catch rate ($NPUE_y$) in year y were calculated as

$$\bar{L}_y = \frac{\sum_{t=t_c}^{t_{max}} L_t N_{t,y}}{\sum_{t=t_c}^{t_{max}} N_{t,y}} + \varepsilon_L, \tag{4.14}$$

$$NPUE_y = \tilde{q} \sum_{t=t_c}^{t_{max}} N_{t,y} + \varepsilon_I, \tag{4.15}$$

where $\varepsilon_L \sim N(0,\sigma_L^2)$ and $\varepsilon_I \sim N(0,\sigma_I^2)$ are normally distributed deviations in mean length and NPUE, respectively; and L_t is the length of an animal at age t following a von Bertalanffy growth function and is calculated as $L_t = L_\infty \{1 - \exp[-K(t - t_0)]\}$. The mean length and catch rate were observed at the beginning of each year. For each factorial combination, 10,000 stochastic time series of mean lengths and catch rates were generated. The values of the life history parameters, scaling coefficient, and SD parameters for the simulation are defined in Table 4.1, with growth parameters partly based on a Mutton Snapper stock.

The two mortality rates and single change point were then estimated by using only mean lengths (ML model) or by using both mean lengths and catch rates (MLCR

model). In each factorial combination, the percent bias (%Bias) and percent root mean square error (%RMSE) for these parameters were calculated for both models as

$$\%Bias = \frac{\overline{\hat{X}} - X}{X} \times 100, \tag{4.16}$$

$$\%RMSE = \frac{1}{X}\sqrt{\frac{\sum_i (\hat{X}_i - X)^2}{n}} \times 100, \tag{4.17}$$

where X is the true value of the parameter of interest, $\hat{X}_i$ is the estimate in the i-th simulation, and $\overline{\hat{X}}$ is the mean of the estimates from simulations i = 1, 2,…, 10,000.

Pearson's product-moment correlation of paired mean length and NPUE residuals in the MLCR model was calculated in each factorial combination. Two sets of residuals were examined: (1) the difference between the simulated value and the value expected under constant recruitment with the true mortality rate (true residual); and (2) the difference between the simulated value and the value predicted in the application of the MLCR model (fitted-MLCR residual).

To analyze the trends in residuals, we calculated the mean of the longest run of positive and negative residuals of the mean lengths and catch rates from the 10,000 time series in each factorial combination. For both data types, we calculated the true residual and the fitted-MLCR residual. For the mean lengths, we also calculated a third type of residual from the difference between observed values and the values predicted by the ML model (fitted-ML residual).

4.3.4. Application to the Mutton Snapper Pot Fishery in Puerto Rico

The Mutton Snapper is one of the more important commercially caught fishes in Puerto Rico, yet it is a data-limited stock with unknown stock status (SEDAR 2007). Annual mean lengths (1983–2006) of Mutton Snapper larger than 30 cm (i.e., the assumed length of full vulnerability, L_c) were calculated. Standardized WPUEs (1990–2006) were obtained from Cummings (2007) and were used to index biomass trends (SEDAR 2007). Mortality rates were estimated using the ML and MLCR models implemented in AD Model Builder (Fournier et al. 2012). Life history values for the analyses were obtained from Burton (2002): L_∞ was 86.9 cm, K was 0.16 year^{-1}, and b was 3.05.

Sensitivity of the MLCR model to growth parameters was evaluated by refitting the model with alternative values of L_∞, K, and b. The parameters were sampled 100 times from a multivariate normal distribution with the means from the base analysis. The covariance matrix was created, assuming coefficients of variation (CVs) of 0.15, 0.04, and 0.0098, respectively, based on the estimated SEs reported by Burton (2002). The von Bertalanffy parameters L_∞ and K were sampled assuming a correlation of –0.90, while both were independent of b.

4.4. Results

4.4.1. Simulation Study of the Mortality Estimators

The bias in estimates of Z and the change point was generally small for both the ML model and the MLCR model in all factorial combinations. The %Bias metric was less than 10% in almost all cases, although it increased with increasing recruitment

variability (Figure 4.3). For most factorial combinations in the simulation, both the ML model and the MLCR model produced a positive bias in the mortality estimates. This result was consistent with the findings in previous simulation studies of the Beverton–Holt equation (Then et al. 2015), although the positive bias was less likely to occur when recruitment variability was highest (σ_R = 1.0).

Compared to the ML model, the MLCR model generally produced less-biased estimates of the higher mortality rate (i.e., Z_1 when mortality decreased or Z_2 when mortality increased). The %Bias in the estimate of the change point and the difference in bias between the ML model and MLCR model were small (all < 2.0%). The estimates from the MLCR model also had a %RMSE that was equal to or lower than those from the ML model for all parameters in all factorial combinations, indicating higher precision when using the MLCR model (Figure 4.4).

When the MLCR model was used to estimate Z, the correlation between paired residuals of mean length and NPUE was negative (Figure 4.5). The true residuals were uncorrelated when there was no recruitment variability (σ_R = 0.0). However, the fitted-MLCR residuals showed a slight negative correlation coefficient of approximately –0.05 even when there was no recruitment variability. With increasing recruitment variability, the correlation coefficient of fitted-MLCR residuals ranged from –0.2 to –0.6, with the most extreme correlation value observed when recruitment variability was highest. In all cases, the fitted-MLCR residuals had stronger correlations than the true residuals.

The magnitude of the largest run of positive or negative residuals in the mean lengths and catch rates increased as recruitment variability increased (Figure 4.6). When

there was no recruitment variability (σ_R = 0.0), the mean largest run in the true residual for mean lengths and NPUE was approximately 4.6. The mean largest run in the true residual increased to as much as 6.0 for mean length and as high as 8.9 for NPUE with high recruitment variability. In all recruitment and mortality scenarios, residuals from the both the ML model and the MLCR model showed shorter runs than the true residuals. For the fitted-MLCR residuals, the mean largest run was as high as 5.2 in mean length and 6.2 in NPUE with high recruitment variability. For the mean length, the fitted-ML residual had shorter runs than the fitted-MLCR residual of the corresponding scenarios.

4.4.2. Application to the Mutton Snapper Pot Fishery in Puerto Rico

Time-specific *Z* was estimated from mean lengths by using the ML model (Figure 4.7). With one change point in the mortality rate over time, mortality was estimated to have increased from $\hat{Z}_1$ = 0.51 year^{-1} to $\hat{Z}_2$ = 1.00 year^{-1}, with the change point $\hat{D}_1$ at 1992.7 (change points are estimated in continuous time, with the decimal representing tenths of a year). With two change points, mortality increased from $\hat{Z}_1$ = 0.51 year^{-1} to $\hat{Z}_2$ = 1.25 year^{-1} and subsequently decreased to $\hat{Z}_3$ = 0.79 year^{-1}, with change points of $\hat{D}_1$ = 1993.3 and $\hat{D}_2$ = 1998.9, respectively. There was strong support for the one-change model over the two-change model, as the AIC_c value increased by 2.3 units for the latter model with the additional change point (Table 4.2).

Next, the MLCR model was used to estimate *Z* from mean lengths and WPUEs (Figure 4.8). Assuming one change in mortality, *Z* was estimated to have increased from $\hat{Z}_1$= 0.51 year^{-1} to $\hat{Z}_2$ = 0.81 year^{-1} in 1987 (Table 4.3). An examination of the predicted

and observed WPUE values showed a clear pattern to the residuals, suggesting that the one-change model did not fit the data well (Figure 4.8). The model was reformulated to include a second change in mortality. In the two-change model, mortality was estimated to have increased from $\hat{Z}_1 = 0.51$ year^{-1} to $\hat{Z}_2 = 1.19$ year^{-1} with $\hat{D}_1 = 1987.3$, followed by a reduction to $\hat{Z}_3 = 0.61$ year^{-1} with $\hat{D}_2 = 1997$. The two-change model was the better fit to the mean lengths and catch rates, with the AICc value reduced by 22.8 units despite the need to estimate an additional mortality rate and change point.

The sensitivity analysis of growth parameters was performed for the MLCR model with two change points. Model estimates were all highly correlated with L_∞ (positively) and K (negatively). The magnitude of the correlation between L_∞ and model estimates was greater than 0.85 in all cases, whereas there was little correlation (magnitude all less than 0.05) between b and the model estimates. The CV of the estimate of the most recent mortality rate (Z_3) was 0.27 (Figure 4.9).

4.5. Discussion

4.5.1. Simulation Study of the Mortality Estimators

In the simulation, the %RMSE for Z_1 was similar between the ML model and the MLCR model. The data (mean lengths and catch rates) available to estimate Z_1 were in equilibrium in our simulation, and the model did not need to account for the transitory behavior of the data that would occur after a change in mortality. This is apparent because the %RMSE was larger for estimating the mortality rate that followed the change point (i.e., Z_2). For the change point, the %RMSE was larger when the magnitude of the change in mortality was small (mortality scenarios B and C). In such situations, the

changes in mean lengths and catch rates were relatively small and more difficult to detect. As a result, having two data types in the MLCR model noticeably reduced the %RMSE in estimating Z_2 and the change point.

Generally, when using either mean length or catch rate alone, there is no ability to distinguish between a change in mortality and a change or variability in recruitment. However, the two data types together can be used to estimate stepwise changes in mortality when recruitment is variable. The MLCR model essentially splits the difference in the information between the two data types. Both ML and MLCR fit the model to produce shorter-than-expected residual runs, but in doing so the MLCR model produced correlations that were more negative than expected in paired residuals (Figures 4.5, 4.6). The runs of mean length in the ML model were shorter than those in the MLCR model, but this resulted in less precision in mortality estimation for the former model. From our simulations, the estimates of mortality and inference on the mortality history using both data types were better (by reducing the %RMSE) despite correlations in residuals and residual patterns, both of which should be expected given the variability of recruitment in fish populations.

The values of σ_R used in the simulation encompassed the range of recruitment variability likely to be encountered in marine fish stocks (Thorson et al. 2014). Our simulation generated data by only including two sources of error: recruitment variability and observation error. Increased or decreased random observation error in the data will decrease or increase, respectively, correlations and runs in residuals. Other nonrandom sources of error, such as the extent to which the catch rate is representative of stock abundance and the length composition data of the population size structure, also need to

be evaluated for each application of the MLCR model. These factors make it unlikely to be able to estimate recruitment variability based on residual behavior in the data-limited model.

The behavior of residuals arises from the constant recruitment assumption in the model. Recruitment could be explicitly modeled in a mortality estimation procedure, but this would require a different derivation of the mean length and catch rate than the one presented in the current study.

The log-likelihood function (equation 4.9) assumes that observation errors for length and catch rate are uncorrelated because the length and catch rate data are sampled independently from each other. If a stock exhibits characteristics that would cause correlated observation errors—for example, if schooling behavior occurs in certain size-classes, resulting in concurrent high catch rates, and the data are sampled as paired observations (e.g., within individual fishing trips or gear hauls)—then correlations in observation error can occur. The log-likelihood can be modified on a case-by-case basis to account for such situations.

4.5.2. Application to the Mutton Snapper Pot Fishery in Puerto Rico

In application of the MLCR model to the Mutton Snapper data specifying two change points, there was some disagreement in the signals from the mean length and WPUE data, as was apparent from examination of the residuals (Figure 4.8). For the mean lengths, there was a run of five positive residuals in 1992–1996 and four negative residuals in 2002–2006, which were within the range of the mean largest runs observed in the simulation. On the other hand, there did not appear to be a pattern in the residuals of

the catch rates. The correlation of the 17 paired residuals was –0.07 but was not statistically significant (t = –0.27, df = 15, P = 0.79). These diagnostics could suggest relatively low recruitment variability in Mutton Snapper. Low sample sizes of lengths during some years may also contribute to large residuals, but this aspect was not evaluated in the simulation.

Recruitment variability has been found to be higher when oceanographic conditions are less stable (Myers and Pepin 1994; Myers 2001). Thus, the tropical distribution of Mutton Snapper suggests that recruitment variability of the stock is likely to be lower than those of higher-latitude species. Stock assessments of several lutjanid species based on age-structured models have also suggested low values of recruitment variability, with σ_R no greater than 0.3 (SEDAR 2003, 2015, 2016), although a formal meta-analysis has not been performed for lutjanids. In the context of the ML and MLCR models and the simulation study, recruitment refers to the cohort entering the fishery at age t_c, at which time cohort strength can be dampened by density-dependent processes. In contrast, recruits can also be defined at age 0 when individuals reach the settlement phase immediately after the larval stage of the life cycle. High interannual variability in these post-settlement recruits of tropical coral reef fish species has been observed (Shulman 1985; Rankin and Sponaugle 2014).

According to the transitional behavior of the mean length and catch rate presented by Gedamke and Hoenig (2006) and in the present study, respectively, the mean length data suggested an initial change in mortality during 1992–1993 (Figure 4.7; Table 4.3), whereas the stability of WPUE during 1990–1997 suggested that a mortality change

occurred prior to 1990 (Figure 4.8). When both types of data were used in the MLCR model, the WPUE data apparently received more weight given the current log-likelihood equation, resulting in the early estimates of the first change point in the two-change-point model.

Additional modifications are also possible to balance the contribution of the length and catch rate log-likelihoods in the MLCR model. Weight coefficients can be assigned to each log-likelihood component in equation (4.9); as more weight is given to the length data, the results will approach those in the ML model at the expense of model fit to the catch rates. When there is high interannual variability in the precision of the catch rate, annual estimates of the catch rate SD (often obtained via standardization techniques; Maunder and Punt 2004) can serve as input into the log-likelihood to weight each annual value accordingly.

Results from the ML model and MLCR model illustrate that different models can produce different interpretations about the pattern of mortality experienced by the Mutton Snapper stock. All four models presented here (ML and MLCR models, each with one or two change points) predicted the same initial increase in Z around 1989–1993 because only length data were available prior to 1990. There appeared to be an increase in the mean length data in the early 2000s, which would suggest a second change in Z (Figures 4.7, 4.8). However, this trend alone did not provide sufficiently strong evidence of a reduction in mortality using the ML model until the concurrent increase in the catch rate was also considered in the MLCR model.

The predicted catch rates in the MLCR model specifying two changes in mortality tracked the observed values very well, but the high variability of the mean length data

still resulted in some uncertainty surrounding the estimates of Z since 1990 (i.e., $\hat{Z}_2$ and $\hat{Z}_3$; Table 4.3). This is expected when we consider that the clear trend in the catch rates will only provide information on the relative change in mortality; information from the mean lengths is still required to estimate the absolute mortality rates.

The four models agreed on the initial Z before 1988, but the mortality rates estimated after that year were extremely variable. Variability of the mean length data in this time period was partly attributable to low sample sizes in some years. Since the likelihood function for the mean lengths weights the time series by annual sample size, years with few length observations may produce large outliers and large residuals, as was seen in this application. Additionally, the WPUE time series does not provide information for estimating the ratio of the first change in mortality since data prior to 1990 were not available.

All four models indicated that since 1998, mortality has either (1) decreased if mortality had been very high (ML and MLCR models with two change points) or (2) held constant at a more moderate value (ML and MLCR models with one change point). Reliable estimates of Z over time will require a more intensive, standardized fishery sampling program or alternatively a standardized survey index of Mutton Snapper relative abundance through time. Utilizing the best-fit model that considers all available data (i.e., MLCR with two changes in mortality) implies that the Z in the terminal year of the time series was markedly smaller than values from the other models (Tables 4.2, 4.3).

Estimates of mortality are also conditional on the values of life history parameters used in the model. In the MLCR model, mortality estimation is partly based on the magnitude of mean lengths relative to L_∞. Larger values of L_∞ imply a larger mortality

rate, which is why we observed a strong correlation between the two in the sensitivity analysis. Larger uncertainty in L_∞ would result in a proportionate increase in the uncertainty of mortality estimates. This behavior is consistent with what has been observed in the ML model (Gedamke and Hoenig 2006). In contrast, there is generally less uncertainty in the allometric exponent b relative to von Bertalanffy parameters.

The estimates of *Z* from these models could be used to obtain fishing mortality if an external estimate of natural mortality is available. Biological reference points from spawning potential ratio or yield-per-recruit analyses can then be used to evaluate stock status. We did not do so here because the Mutton Snapper application was an illustration of the methods and not intended to be an assessment of the stock.

4.6. Conclusions

In this study, we derived the transitional behavior of the catch rate following a change in *Z*. Since catch rates can provide additional information on mortality trends over time, we developed a mortality estimator that uses both mean length and catch rate data. Simulations showed that when the assumption of constant recruitment was violated, patterns in the residuals were generated. Despite this, the %Bias values for the Z-estimates and change point were relatively low when σ_R was less than 1.0, and the %RMSE was reduced in all situations with the inclusion of catch rates compared to when mean lengths were used alone. Thus, residual patterns arising from non-constant recruitment are unlikely to substantially bias the estimates of *Z*. The application to Mutton Snapper highlights the value of considering the catch rates together with the

mean lengths; an additional change in mortality was estimated when both mean length and catch rate data were used in one model.

4.7. References

Beverton, R. J. H., and S. J. Holt. 1956. A review of methods for estimating mortality rates in fish populations, with special reference to sources of bias in catch sampling. Rapports et Procès-verbaux des Réunions, Conseil International pour l'Exploration de la Mer 140:67-83.

Beverton, R. J. H., and S. J. Holt. 1957. On the dynamics of exploited fish populations. Fishery Investigations Series II, Marine Fisheries, Great Britain Ministry of Agriculture, Fisheries and Food 19.

Burnham, K. P., and D. R. Anderson. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd edition. Springer-Verlag, New York, New York.

Burton, M. L. 2002. Age, growth and mortality of mutton snapper, *Lutjanus analis*, from the east coast of Florida, with a brief discussion of management implications. Fisheries Research 59:31-41.

Cummings, N. J. 2007. Updated Commercial Catch per Unit of Effort Indices for Mutton snapper line and pot fisheries in Puerto Rico, 1983-2006. SEDAR14-AW01-Addendum 1. SEDAR, North Charleston, SC.

Fournier, D. A., H. J. Skaug, J. Ancheta, J. Ianelli, A. Magnusson, M. N. Maunder, A. Nielsen, and J. Sibert. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods and Software 27: 233-249.

Gedamke, T., and J. M. Hoenig. 2006. Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of goosefish. Transactions of the American Fisheries Society 135: 476-487.

Gedamke, T., J. M. Hoenig, W. DuPaul, and J. A. Musick. 2008. Total Mortality Rates of the Barndoor Skate, *Dipturus laevis,* from the Gulf of Maine and Georges Bank, United States, 1963-2005. Fisheries Research 89:17-25.

Maunder, M. N., and A. E. Punt. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research 70:141-159.

Myers, R. A. 2001. Stock and recruitment: generalizations about maximum reproductive rate, density dependence, and variability using meta-analytic approaches. ICES Journal of Marine Science 58:937-951.

Myers, R. A., and P. Pepin. 1994. Recruitment variability and oceanographic stability. Fisheries Oceanography 3:246-255.

Rankin, T. L. and S. Sponaugle. 2014. Characteristics of Settling Coral Reef Fish Are Related to Recruitment Timing and Success. PLoS ONE 9(9):e108871.

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

SEDAR (SouthEast Data, Assessment, and Review). 2003. Complete Stock Assessment Report of Yellowtail Snapper in the Southeastern United States – SEDAR 3 Assessment Report 1. SEDAR, North Charleston, SC.

SEDAR (SouthEast Data, Assessment, and Review). 2007. SEDAR 14 – Caribbean Mutton snapper Stock Assessment Report. SEDAR, North Charleston, SC.

SEDAR (SouthEast Data, Assessment, and Review). 2015. Stock Assessment of Red Snapper in the Gulf of Mexico 1872-2013 with Provisional 2014 Landings. SEDAR, North Charleston, SC.

SEDAR (SouthEast Data, Assessment, and Review). 2016. SEDAR 45 – Gulf of Mexico Vermilion Snapper Stock Assessment Report. SEDAR, North Charleston, SC.

Shulman, M. J. 1985. Variability in recruitment of coral reef fishes. Journal of Experimental Marine Biology and Ecology 89:205-219.

Then, A. Y., J.M. Hoenig, and Q.C. Huynh. In review. Estimating fishing and natural mortality rates, and catchability coefficient, from a series of observations on mean length and fishing effort. Submitted to ICES Journal of Marine Science.

Then, A. Y., J. M. Hoenig, T. Gedamke, and J. S. Ault. 2015. Comparison of Two Length-Based Estimators of Total Mortality: a Simulation Approach. Transactions of the American Fisheries Society 144:1206-1219.

Thorson, J. T., O. P. Jensen, and E. F. Zipkin. 2014. How variable is recruitment for exploited marine fishes? A hierarchical model for testing life history theory. Canadian Journal of Fisheries and Aquatic Sciences 71:973-983.

4.8. Tables

Table 4.1. Factorial design and values of parameters used for the simulation study (Z = total mortality rate).

Variable	Symbol	Values
SD of recruitment	σ_R	0.0, 0.25, 0.50, 1.0
Mortality scenario	A	$Z_1 = 0.4$, $Z_2 = 0.8$
	B	$Z_1 = 0.4$, $Z_2 = 0.6$
	C	$Z_1 = 0.8$, $Z_2 = 0.6$
	D	$Z_1 = 0.8$, $Z_2 = 0.4$
Change point	D	Year 11
Age of full recruitment	t_c	3
Maximum age	t_{max}	18
Catch rate scaling coefficient	$\tilde{q}$	1
von Bertalanffy asymptotic length	L_∞	80
von Bertalanffy growth parameter	K	0.15
von Bertalanffy location parameter	t_0	-1
Observation error SD of mean lengths	σ_L	1
Observation error SD of NPUE	σ_I	$0.25\tilde{q}$

Table 4.2. Estimates of total mortality (Z) and change points (D) for Mutton Snapper from the mean length-only model (ΔAIC_c = difference in Akaike's information criterion with correction for small sample sizes). Coefficients of variation (CVs) for the parameter estimates are shown in parentheses; in CV calculations for the change points, the number of years elapsed since the first year of the model (i.e., 1983) was used in the denominator.

Parameter	One change point ($\Delta AIC_c = 0.0$)	Two change points ($\Delta AIC_c = 2.3$)
Z_1	0.51 (0.06)	0.51 (0.06)
D_1	1992.70 (0.14)	1993.30 (0.23)
Z_2	1.00 (0.14)	1.25 (0.02)
D_2	-	1998.90 (0.06)
Z_3	-	0.79 (0.03)

Table 4.3. Estimates of total mortality (Z) and change points (D) for Mutton Snapper from the mean length–catch rate model (ΔAIC_c = difference in Akaike's information criterion with correction for small sample sizes). Coefficients of variation (CVs) for the parameter estimates are shown in parentheses; in CV calculations for the change points, the number of years elapsed since the first year of the model (i.e., 1983) was used in the denominator.

Parameter	**One change point (ΔAIC_c = 22.8)**	**Two change point (ΔAIC_c = 0.0)**
Z_1	0.51 (0.08)	0.51 (0.08)
D_1	1987.00 (0.24)	1987.30 (0.20)
Z_2	0.81 (0.12)	1.19 (0.27)
D_2	-	1997.20 (0.04)
Z_3	-	0.61 (0.16)

4.9. Figures

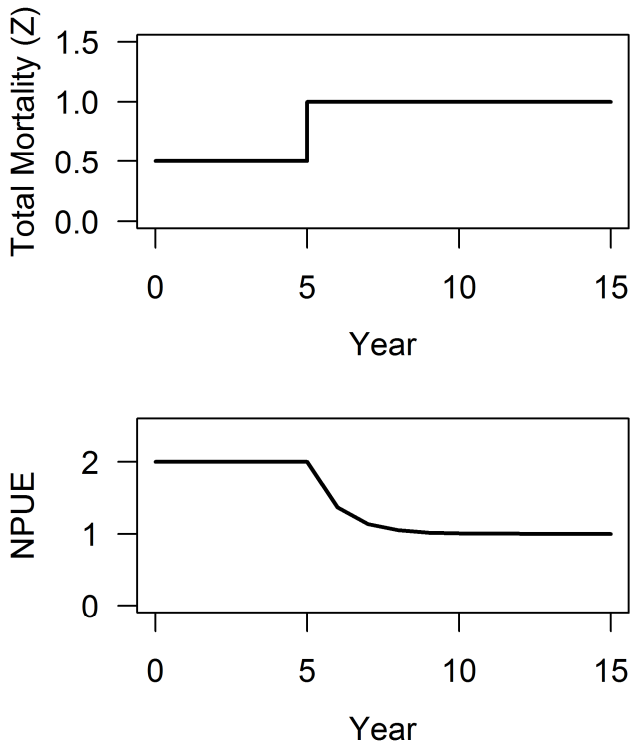

Figure 4.1. Response of a number-per-unit-effort (*NPUE*) index of abundance (lower panel) to a 100% increase in total mortality (Z) from 0.5 year^{-1} to 1.0 year^{-1} (upper panel). The new asymptotic value of the catch rate will be half of the original equilibrium catch rate. The values of the catch rate are scaled by $\widetilde{q}$ which is the product of the catchability coefficient q and recruitment R.

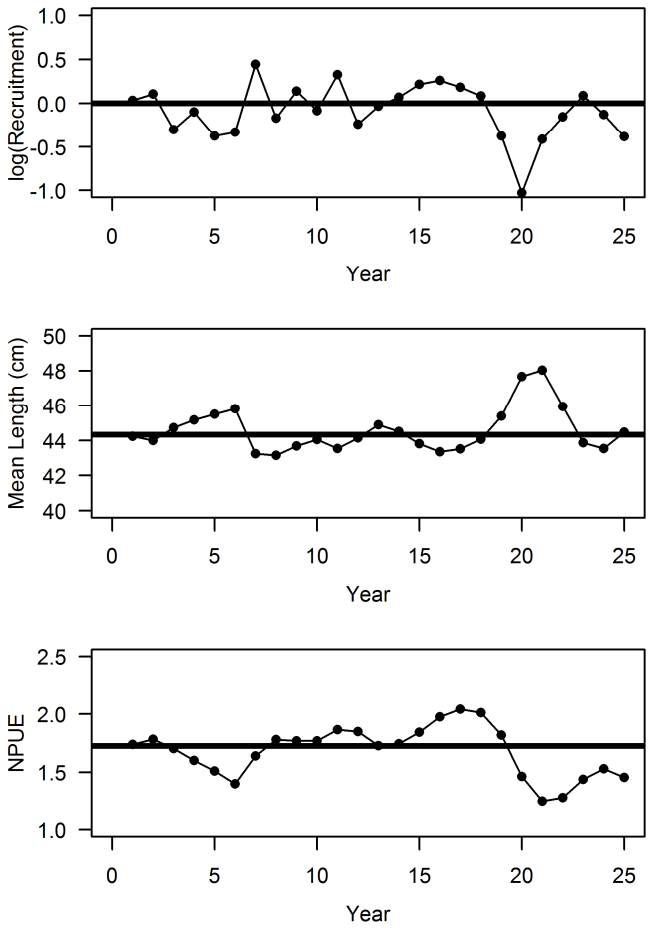

Figure 4.2. Hypothetical time series of stochastic recruitment that is lognormally distributed around a stationary mean (top panel) and the corresponding response of mean length (middle panel) and catch rate (bottom panel). Mortality is held constant over time. Solid horizontal lines indicate values predicted under constant recruitment. Life history values from Table 4.1 were used.

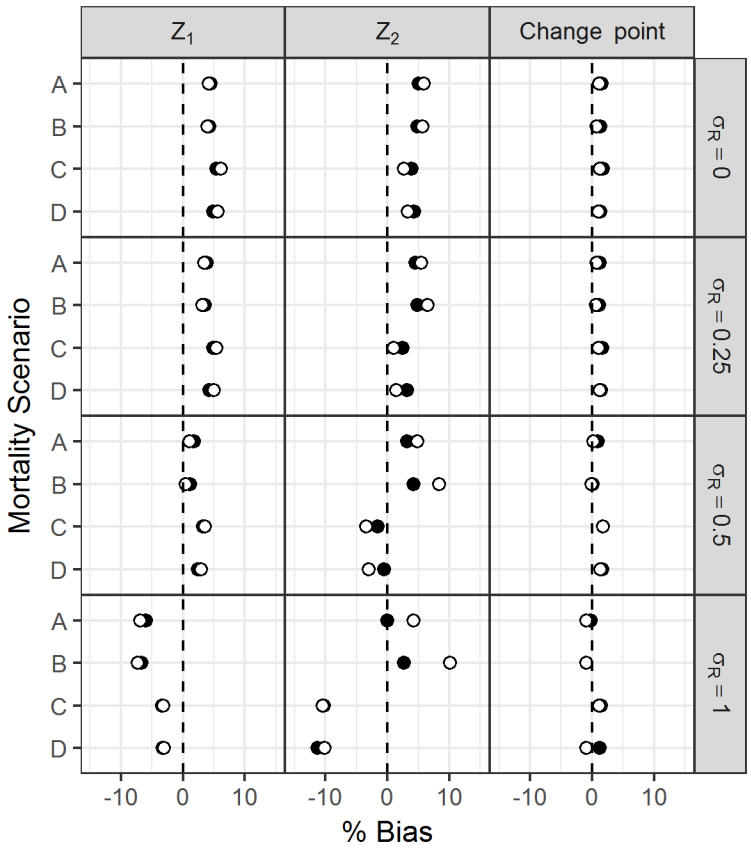

Figure 4.3. Percent Bias (*%Bias*) of estimated total mortality rates Z_1 and Z_2, and the change point based on mean lengths only (ML; open circles) or based on mean lengths and catch rates (MLCR; filled circles) from the simulation. The four mortality scenarios (A-D) and four values of recruitment variability (σ_R) from the simulation are described in Table 4.1. Dashed vertical lines indicate %Bias = 0. In some cases, open circles directly overlap filled circles.

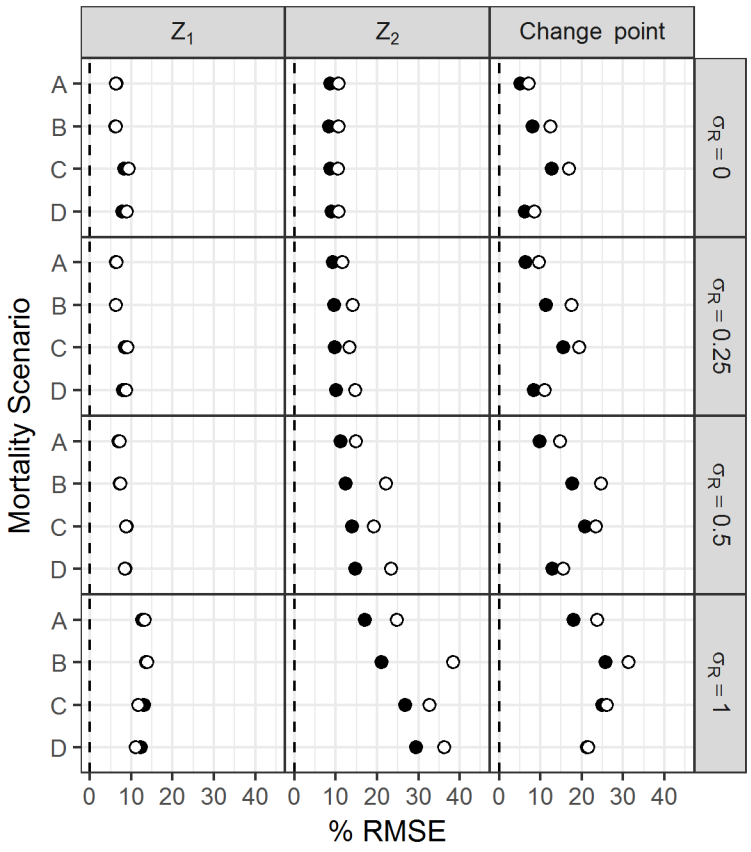

Figure 4.4. Percent root mean square error (*%RMSE*) of estimated total mortality rates Z_1 and Z_2, and the change point based on mean lengths only (ML; open circles) or based on mean lengths and catch rates (MLCR; filled circles) from the simulation. The four mortality scenarios (A-D) and four values of recruitment variability (σ_R) from the simulation are described in Table 4.1. Dashed vertical lines indicate %RMSE = 0. In some cases, open circles directly overlap filled circles.

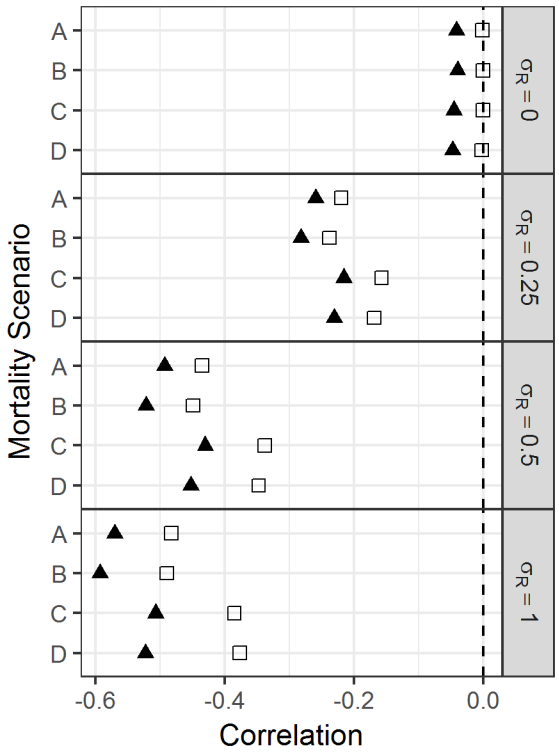

Figure 4.5. Pearson's product-moment correlation coefficients between paired residuals of mean length and catch rate (open squares = true residuals; filled triangles = fitted-MLCR residuals [i.e., mean length-catch rate model]) from the simulation. The four mortality scenarios (A-D) and four levels of recruitment variability (σ_R) are described in Table 4.1. The dashed vertical line indicates a correlation coefficient of zero.

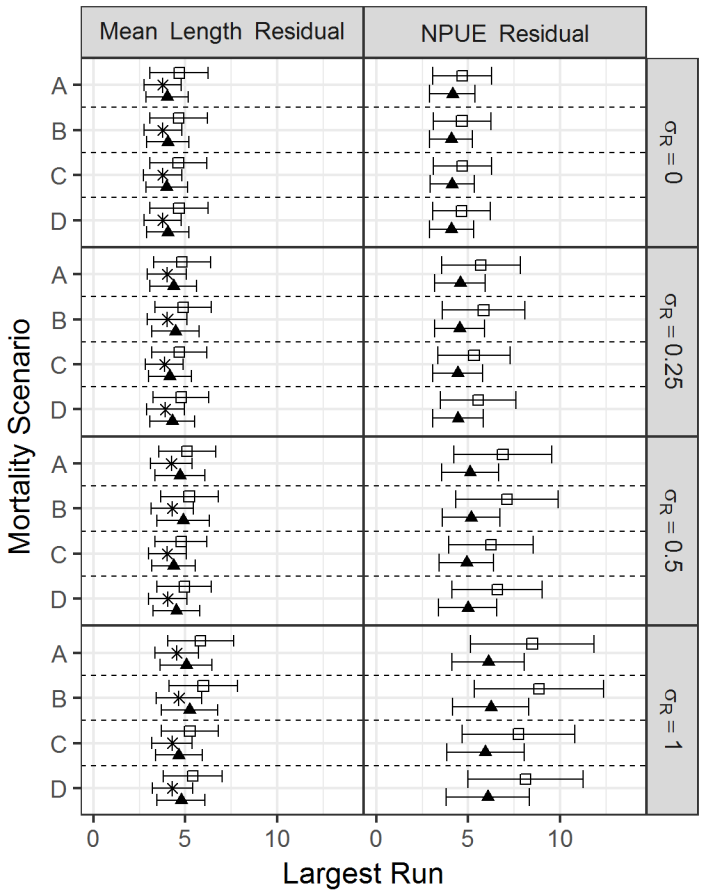

Figure 4.6. The mean (+/- SD; n = 10,000) of the largest run for sequences of positive and negative residuals of mean lengths and catch rate (number-per-unit-effort [NPUE]; open squares = true residuals; asterisks = fitted-ML; filled triangles = fitted-MLCR residuals [i.e., mean length-catch rate model]) in a 20-year time series. The four mortality scenarios (A-D) and four levels of recruitment variability (σ_R) are described in Table 4.1.

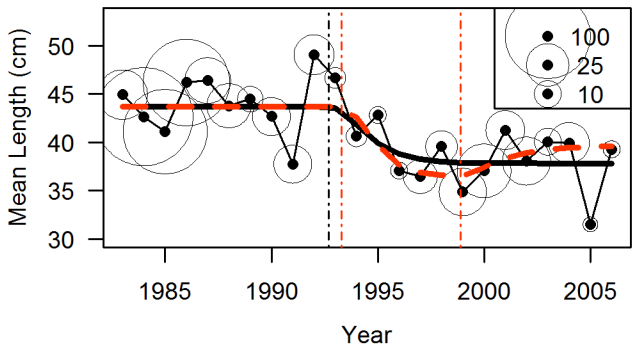

Figure 4.7. Observed (points) and predicted mean lengths assuming one change in mortality (solid black line) or two changes in mortality (dashed red line) for Mutton Snapper based on the mean length-only model. Dot-dashed vertical lines indicate the estimated change points for the respective model (black = one change; red = two changes). Concentric circles around mean lengths indicate the annual sample size of observations used in the likelihood function (with legend provided); the area of the circle is proportional to the sample size.

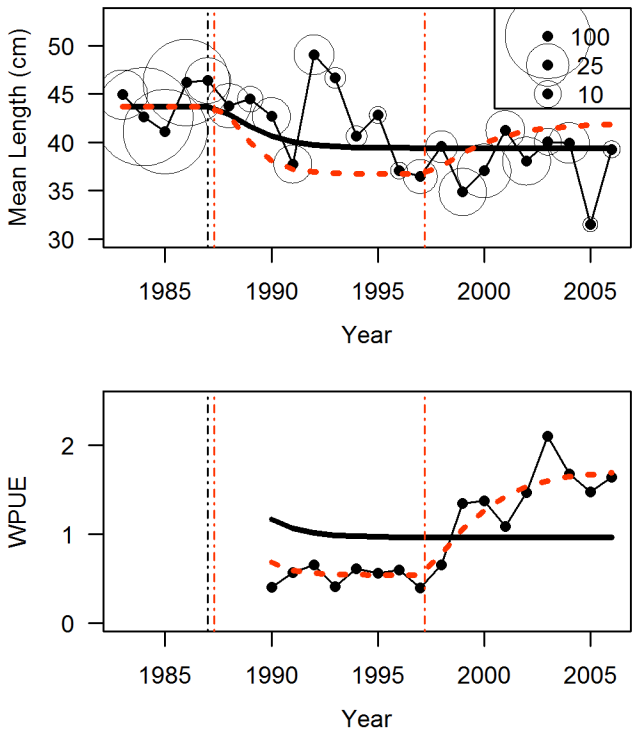

Figure 4.8. Observed (points) and predicted mean lengths (upper panel) and weight per unit effort (WPUE; bottom panel) assuming one change in mortality (solid black line) or two changes in mortality (dashed red line) for Mutton Snapper based on the mean length-catch rate model. Dot-dashed vertical lines indicate the estimated change points for the respective model (black = one change; red = two changes). Concentric circles around mean lengths indicate the annual sample size of observations used in the likelihood function (with legend provided); the area of the circle is proportional to the sample size.

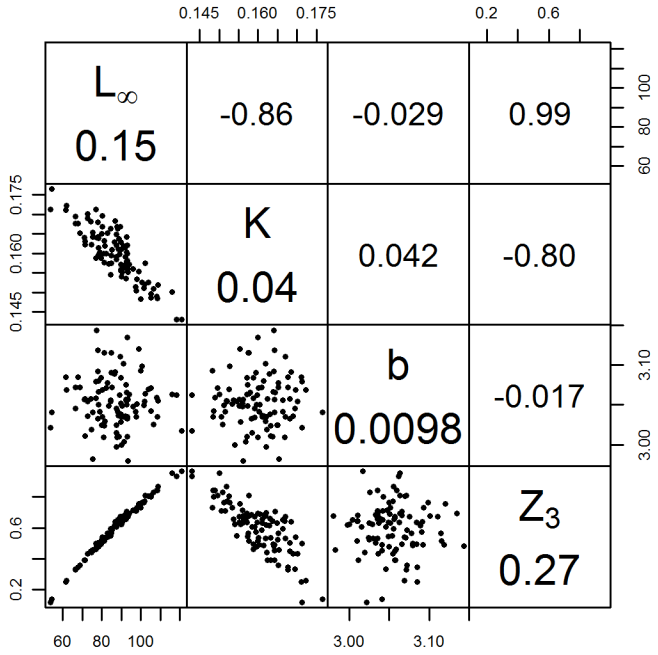

Figure 4.9. Scatterplots (lower triangle), correlation coefficients (upper triangle), and coefficients of variation (diagonal) of life history parameters sampled from a multivariate normal distribution (L_∞ , K, and b; symbols defined in Table 4.1 and the mean values for the sensitivity values are defined by Burton [2002]) and the resulting mortality estimate (Z_3) in the terminal year of the time series for Mutton Snapper based on the mean length-catch rate model with two change points.

Chapter 5: How well do length-based mortality estimators and age-structured models agree for stock status? A comparison with six southeastern United States stocks

5.1. Abstract

Several methods have recently been developed to estimate historical mortality rates for data-limited stocks from a time series of mean lengths, with extensions for auxiliary data, including indices of abundance and fishing effort. In this study, we used three methodologically-related mean length-based methods to estimate mortality for six stocks in the southeastern United States, four in the Gulf of Mexico (greater amberjack, Spanish mackerel, cobia, and king mackerel) and two in the Atlantic (cobia and king mackerel). The analysis with the mean length-based methods used the same length compositions from the recreational fleet as those in the most recent benchmark age-structured assessments of these stocks, which allowed for comparisons using the same subset of data. Generally, there was agreement in the historical trends in mortality among the three mean length-based models and the age-structured assessments. The Gulf of Mexico Spanish mackerel stock produced the most divergent results among the models, but diagnostic steps were taken to evaluate goodness of fit of the mean length-based models. Reduction in shrimp bycatch mortality, corroborated by the results of the age-structured assessment, is hypothesized to have increased recruitment to the recreational gear, which affected the observed trends in the mean length and index. For the six stocks, all models agreed on the overfishing status in the terminal year of the assessment, and there was

high agreement in the number of years with overfishing within different historical periods. The stock status in the terminal year did not differ based upon the choice of models considered in this study. Based on our case study, applications of length-based methods in data-limited situations are likely to be consistent with what might be obtained from an age-structured model.

5.2. Introduction

Simpler, alternative stock assessment methods for the assessment of exploited stocks are generally desirable when a more comprehensive stock assessment model may not be viable (Chrysafi and Kuparinen, 2016). Simple methods are generally used in "data-limited" situations, where the data available for an assessment may be restricting, for example, due to lack of availability or a short time series (Bentley, 2015). Tractable assessment methods typically make simplifying assumptions regarding the population. On the other hand, a more comprehensive stock assessment model, such as an age-structured model (ASM), is typically used in "data-rich" scenarios where multiple sources of data are available (Dichmont *et al*., 2016). In both data-limited and data-rich scenarios, analytical methods are used to estimate historical trends in fishing mortality and/or biomass. The most current estimates of these two quantities relative to reference points can then be used to provide short-term management advice.

In data-limited situations, length-based methods are attractive due to their ease of use and the general availability of length information. In conjunction with growth parameters, simple methods typically estimate mortality from a single size composition or mean length (Hordyk *et al*., 2015; Hordyk *et al*., 2016; Kokkalis *et al*., 2015; Beverton

and Holt, 1956; Ehrhardt and Ault, 1992). Recently, three related mean length-based methods have been developed to analyze time series of length data, with extensions that include indices of abundance or fishing effort (Gedamke and Hoenig, 2006; Huynh *et al*., 2017; Then *et al*., in press).

The mean length (ML) mortality estimator of Gedamke and Hoenig (2006) was developed to estimate a series of historical total mortality rates (Z) based on a nonequilibrium formulation of the Beverton and Holt (1956) mean length mortality estimator. From annual observations of mean length of animals larger than L_c, the first fully selected length, the time series is partitioned into stanzas of constant mortality. The mortality rate and the duration of each stanza is then estimated. Mortality is modeled as a step-wise change from one stanza to another, and the mean length is modeled as a continuous feature of time to reflect how mean length changes gradually in response. The model is systematically fitted by varying the number of stanzas and a model selection procedure (e.g., AIC; Akaike Information Criterion) is used to select the best model.

A second approach uses a formulation of an index of abundance which was combined with the mean length model, termed the mean length-catch rate model (MLCR; Huynh *et al*., 2017). In this model, both the mean length and the index are predicted to decrease gradually after a step-wise increase in mortality and, similarly, to increase after a decrease in mortality. This allows for an evaluation of the consistency between the length and index data for mortality estimation using this framework. The systematic fitting procedure used in the ML model is also used here to select the best model.

A third approach estimates year-specific mortality rates from mean lengths by using estimates of effort as an index of mortality (MLeffort; Then *et al*., in press). In this

model, fishing mortality F is proportional to fishing effort via the estimated catchability coefficient q. Total mortality Z in year y of the model is

$$Z_y = qf_y + M \,, \tag{5.1}$$

where f is the effort, and natural mortality M can be estimated or fixed in the model. This formulation precludes the need to estimate mortality in time stanzas. A technical description of these methods is provided in Appendix C.

To evaluate how simpler, data-limited methods may perform relative to age-structured models, the former can be applied to data sets from stocks for which there are age-structured assessments (Dick and MacCall, 2011; Kokkalis *et al*., 2016). Synchrony in the results among models, i.e. whether or not the historical stock trends are in agreement, can be a form of endorsement for the data-limited methods. While there is no guarantee that the age-structured model is correct nor that it produces precise and accurate estimates, benchmark assessments often undergo a peer-review process (Dichmont *et al*., 2016) and the results of the age-structured models usually represent our best knowledge of the system. If similar results are obtained among models, then the use of simpler models is inconsequential. The comparison of stock status and whether it would differ based on the choice of model can be accomplished by examining fishing mortality and biomass estimates relative to reference points (F/F_{MSY} and B/B_{MSY}, respectively; Kokkalis *et al*., 2016).

In this study, we use the three multi-year, mean length-based methods to estimate historical mortality trends in six stocks in the southeastern United States that are of interest because they have been assessed using age-structured models. The stocks are Gulf of Mexico (GOM) greater amberjack *Seriola dumerili*, GOM Spanish mackerel

Scomberomorus maculatus, GOM cobia *Rachycentron canadum*, Atlantic (ATL) cobia, GOM king mackerel *S. cavalla*, and ATL king mackerel. The Beaufort Assessment Model (BAM; Williams and Shertzer, 2015) was used for ATL cobia, while Stock Synthesis (SS; Methot and Wetzel, 2013) was used for all others. For these stocks, length composition data were used in the age-structured assessments which were accepted as the basis for management advice. The length data from these assessments were obtained for the mean length-based methods, which allowed for comparison of historical mortality rates among models which had a common subset of data.

5.3. Methods

5.3.1. Stocks of interest

Greater amberjack is managed under the Reef Fish Fishery Management Plan, and Spanish mackerel, cobia, and king mackerel are managed under the Coastal Migratory Pelagic Fishery Management Plan of the Gulf of Mexico Fishery Management Council and South Atlantic Fishery Management Council. Each of the four species are divided into separate Gulf of Mexico (GOM) and Atlantic (ATL) stocks. Over time, these stocks have been managed with seasonal closures, bag limits, minimum size limits, and catch limits. Size limits, i.e. minimum retention sizes, have generally increased over time for the recreational fleet (Table 5.1).

Benchmark assessments for these stocks occurred in 2013 - 2014 (SEDAR, 2013a, 2013b, 2013c, 2014a, 2014b, 2014c). Data inputs for the age-structured models have typically included landings, discards, indices of abundance, length composition, and length-at-age observations from commercial and recreational sectors. Fishery-

independent indices and length compositions from surveys were also included in the assessment, although the time series is shorter than for fishery-dependent data.

5.3.2. Mortality estimation

For the mean length-based estimators, length compositions, catch, and indices and abundance were obtained directly from the assessments (Table 5.2). Only the length compositions of retained catch were used. In the southeastern U.S., the largest targeted fishing effort has historically come from the recreational fleet (Siegfried *et al*., 2016). The indices from the recreational fleet have generally had the lowest root mean square error (RMSE) in the age-structured assessments, indicating that the fleet were most informative for inference on stock trends (Sagarese *et al*., 2016). Thus, for the length-based methods, the analyses based on the data from the recreational fleets are presented here. Data from fishery-independent sources were not used due to the shorter length of the time series.

To use the ML, MLCR, and MLeffort models, an estimate of L_c is needed and is determined based on the data. It is assumed that all animals larger than length L_c are fully selected. Here, the mode of the length composition compiled for all years was chosen to be the L_c (Figure 5.1). There was generally no trend in the modal length for most years for the 6 stocks. Von Bertalanffy growth parameters L_∞ and K were then obtained from the assessments (Table 5.3). With this information, the ML model could be used to estimate total mortality rates. Models were fitted assuming zero, one, or two change points in mortality.

With the index of abundance, the MLCR model was then fitted using the same successive strategy as for the ML model. For both models, AIC was used to select the best fitting model (i.e., the model with the lowest AIC score). To avoid overfitting, models with more parameters were accepted only if the reduction in AIC was greater than 2 units. From total mortality estimates, fishing mortality *F* was obtained by subtracting the value of natural mortality assumed in the assessments (Table 5.3).

For the MLeffort model, the effort time series was obtained by taking the ratio of the landings (thousands of fish) and index of abundance (catch-per-unit-effort, number of fish per angler hour). Values of *M* were fixed in the model to estimate *q*, which was then used to obtain *F*. The equilibrium effort prior to the first year of the model was assumed to equal to the effort in the first year. Since the model requires a full time series of effort, the initial year of the model was set to the first year with available indices of abundance. Landings estimates prior to the year when composition and index data were based on historical reconstruction (Siegfried *et al*., 2016).

Model performance was evaluated by analysis of residuals of the observed and predicted values of both mean lengths and indices.

5.3.3. Comparison among models

For comparison with the mean length-based models, annual estimates of the summary *F* from the age-structured assessments were obtained from assessment reports (SEDAR, 2013a, 2013b, 2013c, 2014a, 2014b, 2014c). Only estimates since the first year of length composition data are considered here (Table 5.2).

Two sets of relative mortality rates were calculated to facilitate comparison among the different models. First, the absolute magnitude of the estimates was scaled through transformation to Z-scores by subtracting the mean and dividing by the standard deviation of the respective time series (scaled *F*). For the ASM and MLeffort time series, an additional step was taken with the scaled *F* to smoothen the estimates with a loess regression line. The scaled *F* allowed for better comparisons of the trends over time among models.

Second, the estimates were divided by biological reference points (F/F_{MSY}, relative *F*) calculated from their respective time series. The ratio of F/F_{MSY} is relevant to management for classification of overfishing status. Here, $F_{30\%}$, the fishing mortality rate that reduces the spawning potential ratio to 0.3, was generally used as the proxy for F_{MSY} (Restrepo and Powers, 1999). The exception was in the case of ATL Cobia for which the Beaufort Assessment Model was used as the ASM. Instead of using a proxy, F_{MSY} (the fishing mortality that maximized equilibrium yield) was directly estimated (SEDAR, 2013c).

Estimates of F_{MSY} or their proxies from the age-structured assessments were obtained from the assessment reports. For the mean length-based mortality estimators, $F_{30\%}$ was calculated separately with the life history information in Table 5.3 (Appendix D). Spawning potential ratio calculations differ between the age-structured models and the mean length-based models based on different assumptions regarding selectivity and maturity.

To evaluate the synchrony of relative *F*, the proportion of years in which overfishing is estimated to occur was calculated for 4 time periods: (1) pre-1995

(approximately the first half of the time series for all 6 stocks), (2) post-1995 (approximately the second half of the time series), (3) the most recent 5 years, and (4) the terminal year of the time series.

All analyses were performed in R using the MLZ package, which is publicly available on Github (http://ww.github.com/quang-huynh/MLZ).

5.4. Results

5.4.1. Trends in fishing mortality

While the mean length-based methods estimate Z, we assume, as many age-structured models do, that M is constant over time. Thus, the trends that we examine are due to changes in F. For most stocks analyzed here, both the age-structured assessment models and the mean length mortality estimators indicated high mortality in the 1980-1990s followed by a reduction in mortality since then (Figure 5.2). This pattern is common to many southeastern U.S. stocks (Siegfried *et al*., 2016).

For GOM greater amberjack, there was high synchrony in the mortality estimates over time. Both the ASM and MLeffort models showed an increase in F from 1981 – 1993 followed by a gradual decrease from 1993 – 2012. Both models exhibited very similar descents in mortality. The ML and MLCR models showed two changes in mortality, an initial increase to an extended plateau in mortality during the 1990s corresponding to the time period surrounding the peak in the ASM and MLeffort models, followed by a reduction in the 2000s.

For GOM Spanish mackerel, the ASM, ML, and MLCR models all showed a general reduction in mortality over time, although the trends and timing differ (the

MLeffort model did not converge). Compared to the ML and MLCR models, the ASM showed a much larger reduction in *F* from the beginning of the time series to 2011 (the terminal year of the data). The ML model indicated two changes in mortality, with a decrease in mortality during the early 1990s from the initial mortality rate. This was caused by the large increase in mean length from 1990-1995 (Figure 5.3). Afterwards, a modest increase to an intermediate mortality rate until the present time was estimated. The trends in the index, however, did not support two changes in mortality. Thus, only one change in mortality, a modest decrease, was inferred in the MLCR model.

For GOM cobia, all four models indicated a reduction in mortality since the 1990s. The ASM showed an initial ramp in mortality followed by a gradual decrease after 1990. The MLeffort model showed a large decrease prior to 1986-1990 (effort data was not available prior to 1986). After 1990, the gradual decrease in mortality mimicked that in the ASM. The ML and MLCR models both estimated two changes in mortality, with a temporary decrease in mortality in the late-1990s followed by a modest increase to a mortality rate that is less than the initial estimated mortality rate. This pattern was inferred based on the synchronous increase and decrease in the mean length and index in the late-1990s.

For ATL cobia, differing trends in mortality were inferred among the four models. The loess smoother indicated a recent increase in mortality in the ASM, although there was high variability in annual estimates (Figure 5.4). While there were trends in the mean length over time, the ML model indicated zero changes in mortality based on AIC. On the other hand, the MLCR model indicated a decrease in mortality, largely based on the

increase in the index after 1995. The MLeffort model showed a gradual decrease in mortality over time.

For GOM king mackerel, different models produced differing trends in mortality. According to the ASM and MLeffort models, mortality had generally decreased, with a pronounced drop in the late 2000s, although the fit to the mean length data in the MLeffort model was generally poor (Figure 5.4). With the ML and MLCR models, three mortality time stanzas were estimated, with a temporary increase in mortality in the early 2000s followed by a decrease to a mortality rate that is slightly larger than the initial mortality rate.

For ATL king mackerel, all models showed an increase followed by a decrease in mortality over the examined time period 1979-2012. In the ASM, the maximum F occurred around 1995 and fishing mortality continued to decrease in the most recent years. The MLeffort model showed an earlier peak in mortality around 1985, followed by a decrease in mortality until 2005. The mortality rate has been steady since then. The ML and MLCR captured the general trend in mortality estimated in the other two models with two stepwise changes in mortality, with an increase around 1985. In the late 1990s, another mortality rate was estimated which was lower than the initial mortality rate.

5.4.2. Stock status

To compare the models with respect to stock status, mortality trends were compared relative to the F_{MSY} proxies (relative F). For all six stocks, the four models agreed in the overfishing status in the terminal year of the time series, i.e., overfishing is

occurring ($F/F_{MSY} > 1$) for GOM greater amberjack and not for the other five stocks (Figures 5.4, 5.5).

For GOM greater amberjack, all models showed that overfishing was occurring in 2012, the terminal year of the time series. The magnitude of F/F_{MSY} over time were very similar among the four models, with a very large relative *F* in the late 1980s and 1990s. A reduction in relative *F* followed, but overfishing was still occurring in 2012. The four models generally agreed on the extent of overfishing within the four time periods. A lower proportion of years with overfishing was inferred in the most recent 5 years for the MLeffort model compared to the other three models, but this appeared to be a result of the high inter-annual variability in relative *F*.

For GOM Spanish mackerel, the ASM showed more contrast in fishing history, with overfishing occurring in eight out of 14 years (57%) in the pre-1995 period. The ML and MLCR models showed that overfishing had not occurred in the stock history. All three models agreed that overfishing had not occurred post-1995.

For GOM cobia, the relative *F* in the MLeffort model was lower over time than in the other three models. Pre-1995, an increase and decrease in relative *F* corresponded to overfishing in one out of nine years (11%) in the MLeffort model, but seven out of 16 years (44%) in the ASM. During the same time period, the ML and MLCR estimated a plateau mortality rate which indicated overfishing in all included years. Post-1995, overfishing has generally not occurred in all four models (the ML and MLCR models estimated a mortality reduction shortly after 1995).

For ATL cobia, overfishing has not occurred based on the relative *F* of all four models.

For GOM king mackerel, the ASM showed that overfishing was occurring over much of the pre-1995 period, contrary to the other three models which showed no overfishing in the same time period. In the early part of the post-1995 period, the ASM, ML, and MLCR showed that overfishing was occurring (20-40% of years post-1995) until mortality was reduced by 2000. The relative *F* in the MLeffort model was lower than those in the other three models over time and did not indicate overfishing in the stock history.

For ATL king mackerel, the ASM and MLeffort models indicated that overfishing occurred in 29% (five out of 17 years) and 27% (four out of 15 years), respectively, of pre-1995 years. Those years generally did not overlap (Figure 5.4), with overfishing estimated pre-1990 with MLeffort and post-1990 with the ASM. Post-1995, there were fewer years with overfishing in the MLeffort model than in the ASM. While the trends in mortality with ML and MLCR followed those with the ASM and MLeffort, the former set of models did not indicate overfishing in the stock history.

5.4.3. Residual analysis

For the mean length-based models, residuals can be analyzed to determine goodness of fit (Appendix E). The model selection procedure with the ML model generally selected the model which minimized any residual trends except in the case of ATL cobia (Figure E.1). In the MLCR model, an extensive trend of positive and negative residuals of the mean lengths and index, respectively, was observed over time for GOM Spanish mackerel (Figure E.2). Similarly, negatively correlated residuals were also present for ATL king mackerel in the most recent years of the analysis. In the MLeffort,

there were trends in residuals over the course of the entire time series for both GOM and ATL king mackerel (Figure E.3).

5.5. Discussion

Based on the estimated trends in mortality over time and stock status for the terminal year of the analyses, there was strong agreement in the stock perception among the mean length-based models and the age-structured models for the six case studies presented here. Similar inferences could be obtained from the mean length-based models despite using only a subset of the data that were included in the ASM.

5.5.1. Life history parameters

The mean length-based models and their corresponding proxies require simpler life history assumptions than the ASM. In the former, growth is assumed to be deterministic and parameters need to be provided prior to the analysis, though simulations have suggested robustness of the mean length-based models to this assumption (Then *et al*., 2015; Huynh *et al*., in review). With age-structured models, growth incorporates variability in size at age and parameters may be estimable within the model (Francis, 2016).

The life history parameters for the mean length-based models were extensively evaluated for the assessments. Growth and maturity were typically estimated from large historical datasets of otoliths and gonad samples, respectively. The observed maximum age was used to estimate natural mortality (Hoenig, 1983). During the assessments, the observed maximum age was evaluated to ascertain whether it was an appropriate

indicator of longevity, especially when there were among-stock differences for the same species. In data-limited situations, uncertainty in mortality estimates can be evaluated with Monte Carlo sampling of life history parameters from parametric distributions (e.g., Huynh *et al*., 2017) or a sensitivity analysis (Gedamke and Hoenig, 2006).

In many ASM, including the six presented here, natural mortality was parameterized to decline with age with a Lorenzen function rather than be constant with age. While age-varying natural mortality would violate the assumption that Z is constant with age, the extent of the violation would be minimal in a Lorenzen-type parameterization of M because older animals experience similar natural mortality rates. Differences in natural mortality are assumed to be largest among youngest ages but length bins corresponding to these ages are typically not considered in the mean length-based methods if those lengths are not fully selected.

5.5.2. Selectivity and retention behavior

In terms of selectivity, age-structured models allow for modeling of complex fishing behavior, albeit at the cost of estimating many, often confounding, parameters. Multiple fishing fleets with disparate selectivity patterns and fishing behaviors are typically modeled separately and there may be enough information to model logistic and dome-shaped selectivity functions. Discard and retention length composition allow for estimation of the vulnerability and retention functions, the product of which would be the effective selectivity of the gear for retained catch. Finally, changes in size regulations can be modeled with time-varying features of the ASM (Methot and Wetzel, 2013). For the mean length models, knife-edge selectivity is assumed at length L_c. Thus, the analysis

uses a subset of the length composition data so that only animals assumed to be fully selected are included in the calculation of the mean length.

Application of the data-limited models should consider if changes in mean length occurred due a change in retention behavior as opposed to a change in mortality. We chose values of L_c that were larger than any implemented minimum retention size for the stocks in this study. In this way, all lengths larger than L_c would have the same presumed selectivity to minimize the effect of the management regulations. On the other hand, to the extent that there has been variable fishing over time on fish smaller than L_c, the assumption of constant recruitment is violated by confounded fishing mortality. Changes in bag limits could alter discard and retention behavior; for example, the implementation of a bag limit may increase discarding of smaller animals in favor of larger ones. To account for this, one would need to evaluate whether there were significant changes in the length distribution of retained catch once those regulations were implemented.

The age-structured assessments estimated dome-shaped selectivity for the recreational fleet for three of the six stocks, these being GOM greater amberjack and both GOM and ATL stocks of king mackerel. This contrasts with the knife-edge selectivity assumption made with the mean length-based models. If the selectivity of the fleets were dome-shaped, then it is presumed that mortality would be overestimated by the length-based models. For determination of stock status with the mean length-based models, the F_{MSY} proxies were also calculated assuming logistic selectivity for these three stocks. The estimated stock status relative to overfishing in the terminal year among the four methods did not change based on presumed selectivity estimated by the ASM. The time series of

F/F_{MSY} from the mean length models were either similar in magnitude or more optimistic than those from the age-structured assessments (Figure 5.4).

5.5.3. Trends in recruitment to the recreational fishery

The assumption of constant recruitment to length L_c was likely violated for GOM Spanish mackerel due to the dynamics of the shrimp fleet which had bycatch of smaller animals. In the age-structured assessment, the shrimp fleet was the highest source of fishing mortality historically (with 100% discard mortality assumed) until the late-1990s, when fishing effort subsequently decreased (SEDAR 2014b; Figure 5.6). This reduction would increase survival and recruitment to size L_c (39 cm in this study). Such an effect could have caused the decrease in the observed mean length from the recreational fleet (Figure 5.5).

For the MLeffort model, non-convergence for GOM Spanish mackerel was caused by the data conflict where the recreational effort was estimated to have decreased (Figure 5.6), yet the mean length also decreased (an increase would have been expected based on the trend in effort). Concurrently, the gradual increase in the index of abundance with the decrease in mean length since mid-1990s would support the hypothesis of increased recruitment to the recreational fishery (Huynh *et al*., 2017). A simpler mortality history, i.e., with fewer change points, was inferred with the MLCR model compared to the ML model to avoid overfitting spurious trends in the mean length due to hypothesized changes in recruitment. The observed trends in the paired residuals of mean length and the index in the MLCR model were also consistent with hypothesized increased recruitment.

While trends in mortality are affected by factors external to the recreational fleet, the analysis of residuals in the MLCR model and non-convergence of the MLeffort model allowed us to diagnose issues in the application of the mean length-based models for GOM Spanish mackerel without external information. With the ASM, we can corroborate that bycatch mortality may have been the primary driver of the historical stock dynamics. In isolation, the length composition from the recreational fleet may not provide sufficient information on the stock history. This is evident in the contrasting trends in mortality in the ML model and ASM since the mid-1990s (Figure 5.2). Overall, the general presence of large animals in the length composition relative to L_∞ would indicate that the GOM Spanish mackerel stock is in generally good shape (Figure 5.1). For GOM and ATL king mackerel, bycatch from the shrimp fleet was a smaller source of mortality relative to the recreational fleet. The impact of bycatch mortality would not be as noticeable for these stocks.

For ATL king mackerel, large residuals in the mean lengths and index were observed in the most recent years of the MLCR model. The increasing mean length increasing and decreasing index since 2007 would be consistent with decreasing recruitment (of animals of length L_c). The ASM for ATL king mackerel estimated a decreasing trend in recruitment of age-0 animals since 2003. Here, the qualitative information about recruitment trends from the MLCR model are also supported by the recruitment estimates from the ASM after accounting for the time lag from age 0 to the age of full selection to the recreational fishery.

5.5.4. Uncertainty in catch and effort

In data-limited situations, as well as in any assessment, the quality of the data and their representativeness to the underlying population dynamics should be evaluated. For example, there were generally large coefficients of variation associated with discard estimates (Siegfried *et al*., 2016) and consequently, discard data were not used with the mean length-based methods. In data-limited situations, discard data may not be available, but it would also be important to consider the magnitude of discard mortality in a management context. As another example, expert judgment is needed to decide if the catch per unit effort (CPUE) can serve as index of abundance. Spanish mackerel and cobia are reported to be opportunistically caught by the recreational fleet, resulting in high percentages of zero catch (Bryan and Saul, 2012). This may degrade the quality of the CPUE as an index of abundance.

One must obtain length compositions from multiple years for the mean length models used in this study. Data from several fleets could be combined if the fleets are believed to behave similarly. Otherwise, mortality estimates can be confounded by the contrasting fishing effort and selectivity of the different fleets. In this study, the recreational data were obtained from MRFSS (Marine Recreational Fisheries Statistics Survey) and MRIP (Marine Recreational Information Program), which is a designed-based sampling program for the charter and private boat fleet, or SRHS (Southeast Region Headboat Survey), which strives to be a census of all headboats in the region (Table 5.4). We followed the decision of the assessment team in regards to combining or separating the data from these two programs. Uncertainty in the composition data could be evaluated by comparing annual lengths from the different gear sectors.

The MLeffort model provides year-specific mortality rates, but the fit to the mean lengths varies from good in the case of GOM greater amberjack to poor, as in the case of GOM king mackerel (Figure 5.3). For mixed fisheries, nominal effort such as days fished may not be an indicator of targeted effort due to switches in targeting. As effort in the recreational fisheries examined here are not allocated on a species-specific basis, indices from these fleets should be obtained from a subset of fishing trips that are believed to have targeted the stock of interest based on catch of associated species (Stephens and MacCall, 2004). Unfortunately, this was often not sati employed for the six stocks analyzed here due to poor model performance during and often was not used to standardize the indices of abundance. Coupled with relatively high uncertainty in recreational effort, these factors likely contributed to poor performance of the MLeffort model for GOM and ATL king mackerel. Methodological advancements of the MLeffort model can smoothen the estimates of effort prior to using the model or treat effort as a state-space variable, although the latter may result in overfitting.

On the other hand, the ML and MLCR models produced similar fishing mortality estimates except in the case of GOM Spanish mackerel and ATL cobia (Figures 5.2, 5.4). For the other four stocks, the index supports the mortality estimates based on mean lengths, which indicates that the length and index data are in agreement. For GOM Spanish mackerel, bycatch mortality may be affecting the stock dynamics as discussed earlier. For ATL cobia, there appears to be little contrast in the fishing mortality in the stock history because overfishing has not occurred in the history of the stock based on the four models. The lack of contrast in mortality may result in more variability in the estimates among models.

5.6. Conclusion

The results from this study show that the mean length-based methods can provide very similar results, i.e., mortality trends and stock status over the course of the entire time series and time periods within, to those from the age-structured assessments. Life history parameters evaluated for the benchmark assessments were used in the analyses with the length-based methods, and a subset of the data from the fleet believed to drive stock dynamics most strongly was used with these methods. Simple models can still be applied to nonequilibrium situations, and the historical increase and decrease in mortality during the 1980s to the 2010s common to many southeastern U.S. stocks was captured in the analyses. All methods were in agreement with regards to classifying current status relative to overfishing.

In data-limited situations, the mean length-based methods can be used to explore historical changes in mortality over time, with results likely to be consistent with what might be obtained from an age-structured model. The ML and MLCR models provide a series of historical mortality rates, although the changes in mortality over time will be coarser than in models with year-specific mortality rates. This is due to the stepwise, time stanza structure of the ML and MLCR models. The MLeffort model can provide year-specific mortality rates, and inter-annual variability can be smoothed to describe the trend over time.

As a rule, age-structured models should generally not be replaced by simpler methods. Age-structured models provide more modeling options to accommodate multiple drivers of fishing mortality and productivity, as well as more diagnostic tools to

evaluate the quality of the assessment. Nevertheless, in data-rich scenarios, the mean length-based methods can be used as a diagnostic to evaluate and explain how the mean length has changed over time (through fishing mortality or other causes). When there are conflicting results, diagnostic procedures can provide additional insight on the causes of model or data conflict. Models which incorporate multiple data types are advantageous, because the agreement (or lack of) between data types can be evaluated to determine whether the chosen model is appropriate for the stock of interest. As a large majority of stocks worldwide do not and will not likely have fully age-structured assessments in the future, studies such as this are useful in to results of applying mean length-based mortality estimators.

5.7. References

Bentley, N. 2015. Data and time poverty in fisheries estimation: potential approaches and solutions. ICES Journal of Marine Science, 72: 186-193.

Beverton, R. J. H., and Holt, S. J. 1956. A review of methods for estimating mortality rates in fish populations, with special reference to sources of bias in catch sampling. Rapports et Procès-verbaux des Reunions, Conséil International Pour L'Exploration de la Mer, 140: 67-83.

Bryan, M., and Saul, S. 2012. Recreational indices for cobia and Spanish mackerel in the Gulf of Mexico. SEDAR28-DW-22. SEDAR, North Charleston, South Carolina. 44 pages.

Chrysafi, A., and Kuparinen, A. 2016. Assessing abundance of populations with limited data: Lessons learned from data-poor fisheries stock assessment. Environmental Reviews, 24: 25-38.

Dichmont, C. M, Deng, R. A., Punt, A. E., Brodziak, J., Chang, Y.-J., Cope, J. M., Ianelli, J. N., *et al*. 2016. A review of stock assessment packages in the United States. Fisheries Research, 183: 447-460.

Dick, E. J., and MacCall, A. D. 2011. Depletion-Based Stock Reduction Analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. Fisheries Research, 110: 331-341.

Ehrhardt, N. M., Ault, J. S. 1992. Analysis of Two Length-Based Mortality Models Applied to Bounded Catch Length Frequencies. Transactions of the American Fisheries Society, 121: 115-122.

Francis, R. I. C. C. 2016. Growth in age-structured stock assessment models. Fisheries Research, 180: 77-86.

Gedamke, T., and Hoenig, J. M. 2006. Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of Goosefish. Transactions of the American Fisheries Society, 135: 476–487.

Hoenig, J. M. 1983. Empirical use of longevity data to estimate mortality rates. Fishery Bulletin, 82: 898-903.

Hordyk, A., Ono, K., Valencia, S., Loneragan, N., and Prince, J. 2015. A novel length-based empirical estimation method of spawning potential ratio (SPR), and tests of its performance, for small-scale, data-poor fisheries. ICES Journal of Marine Science, 72: 217-231.

Hordyk, A. R., Ono, K., Prince, J. D., and Walters, C. J. 2016. A simple length-structured model based on life history ratios and incorporating size-dependent selectivity: application to spawning potential ratios for data-poor stocks. Canadian Journal of Fisheries and Aquatic Science, 73: 1787-1799.

Huynh, Q. C., Beckensteiner, J., Carleton, L. M., Marcek, B. J., Nepal KC, V., Peterson, C. P., Wood, M. A., & Hoenig, J. M. In review. Comparative performance of three length-based mortality estimators. Marine and Coastal Fisheries.

Huynh, Q. C., Gedamke, T., Porch, C. E., Hoenig, J. M., Walter, J. F., Bryan, M., and Brodziak, J. 2017. Estimating total mortality rates of mutton snapper from mean lengths and aggregate catch rates in a non-equilibrium situation. Transactions of the American Fisheries Society, 146: 803-815.

Kokkalis, A., Thygesen, U. H., Nielsen, A., and Andersen, K. H. 2015. Limits to the reliability of size-based fishing status estimation for data-poor stocks. Fisheries Research, 171: 4-11.

Kokkalis, A., Eikeset, A. M., Thygesen, U. H., Steingrund, P., and Andersen, K. H. 2016. Estimating uncertainty of data limited stock assessments. ICES Journal of Marine Science, 74: 69-77.

Linton, B. 2012. Methods for estimating shrimp bycatch of Gulf of Mexico Spanish mackerel and cobia. SEDAR28-DW-06. SEDAR, North Charleston, South Carolina. 14 pages.

Lombardi, L. 2014. Growth models for king mackerel from the south Atlantic and Gulf of Mexico. SEDAR38-AW-01. 62 pages.

Methot, Jr., R. D., and Wetzel, C. R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research, 142: 86-99.

Murie, D. J., and Parkyn, D. C. 2008. Age, Growth and Sex Maturity of Greater Amberjack (*Seriola dumerili*) in the Gulf of Mexico. SEDAR33-RD-13. SEDAR, North Charleston, South Carolina. 41 pages.

Restrepo, V. R., and Powers, J. E. 1999. Precautionary control rules in US fisheries management: specification and performance. ICES Journal of Marine Science, 45: 846-852.

Sagarese, S. R., Walter III, J. F., Bryan, M. D, and Carruthers, T. R. 2016. Evaluating Methods for Setting Catch Limits for Gag Grouper: Data-Rich versus Data-Limited. In: T. J. Quinn II, J. L. Armstrong, M. R. Baker, J. D. Heifetz, and D. Witherell (eds.), Assessing and Managing Data-Limited Fish Stocks. Alaska Sea Grant, University of Alaska Fairbanks.

SEDAR. 2013a. SEDAR 28: Gulf of Mexico Cobia stock assessment report. SEDAR, North Charleston, South Carolina.

SEDAR. 2013b. SEDAR 28: Gulf of Mexico Spanish Mackerel stock assessment report. SEDAR, North Charleston, South Carolina. 712 pages.

SEDAR. 2013c. SEDAR 28: South Atlantic Cobia stock assessment report. SEDAR, North Charleston, South Carolina.

SEDAR. 2014a. SEDAR 33: Gulf of Mexico Greater Amberjack stock assessment report. SEDAR, North Charleston, South Carolina. 490 pages.

SEDAR. 2014b. SEDAR 38: Gulf of Mexico King Mackerel stock assessment report. SEDAR, North Charleston, South Carolina.

SEDAR. 2014c. SEDAR 38: South Atlantic King Mackerel stock assessment report. SEDAR, North Charleston, South Carolina.

Siegfried, K. I., Williams, E. H., Shertzer, K. W., and Coggins, L. G. 2016. Improving stock assessments through data prioritization. Canadian Journal of Fisheries and Aquatic Science, 73: 1703-1711.

Stephens, A., and MacCall, A. 2004. A multispecies approach to subsetting logbook data for purposes of estimating CPUE. Fisheries Research, 70: 299-310.

Then, A. Y., Hoenig, J. M., Gedamke, T., and Ault, J. S. 2015. Comparison of Two Length-Based Estimators of Total Mortality: a Simulation Approach. Transactions of the American Fisheries Society, 144: 1206-1219.

Then, A. Y., Hoenig, J. M., and Huynh, Q. C. In press. Estimating fishing and natural mortality rates, and catchability coefficient, from a series of observations on mean length and fishing effort. ICES Journal of Marine Science.

Williams, E. H., and Shertzer, K. W. 2015. Technical documentation of the Beaufort Assessment Model (BAM). NOAA Technical Memorandum NMFS-SEFSC-671, U.S. Department of Commerce. 43 pages.

5.8. Tables

Table 5.1. Summary of size regulations from the recreational fishery (in terms of fork length). Only years preceding the year of the assessment are considered.

Stock	Minimum Legal Size Limit	Years
GOM greater amberjack	28-in (71.1 cm)	1990-2007
	30-in (76.2 cm)	2008-2012
GOM Spanish mackerel	12-in (30.5 cm)	1993-2011
GOM & ATL cobia	33-in (83.8 cm)	1985-2011
GOM & ATL king mackerel	12-in (30.5 cm)	1990-1991
	20-in (50.8 cm)	1992-1999
	24-in (61.0 cm)	2000-2012

Table 5.2. Summary of assessment models and the length composition and index of abundance for the length-based mortality estimators. The Recreational fleet combines the data from both the Charter/Private and the Headboat fleets.

Stock	Assessment Model	Fleet for length analyses	Length time series	Index time series
GOM greater amberjack	SS	Charter/Private	1981-2012	1986-2012
GOM Spanish mackerel	SS	Recreational	1981-2011	1981-2011
GOM cobia	SS	Recreational	1979-2011	1986-2011
ATL cobia	BAM	Recreational	1982-2011	1985-2011
GOM king mackerel	SS	Headboat	1985-2012	1986-2012
ATL king mackerel	SS	Headboat	1978-2012	1980-2012

Table 5.3. Life history parameters used in the analyses for the length-based mortality estimators. Parameters are defined in Table D.1.

Stock	**L_∞ (cm)**	**K (yr^{-1})**	**a_0 (yr)**	**L_c (cm)**	**L_{mat} (cm)**	**α**	**β**	**a_{max} (yr)**	**M (yr^{-1})**	**Source**
GOM greater amberjack	143.6	0.18	-0.95	77.5	90	7.0e-5	2.63	15	0.28	SEDAR, 2014a; Murie and Parkyn, 2008
GOM Spanish mackerel	56.0	0.61	-0.50	39	31	1.5e-5	2.86	11	0.38	SEDAR, 2013b
GOM cobia	128.1	0.42	-0.53	88	70	9.6e-6	3.03	11	0.38	SEDAR, 2013a
ATL cobia	132.4	0.27	-0.47	95	70	2.0e-9	3.28	16	0.26	SEDAR, 2013b
GOM king mackerel	128.9	0.12	-4.08	80	58	7.3e-6	3.01	24	0.17	SEDAR, 2014b; Lombardi, 2014
ATL king mackerel	121.1	0.15	-3.73	80	58	7.3e-6	3.01	26	0.16	SEDAR, 2014c; Lombardi, 2014

5.9. Figures

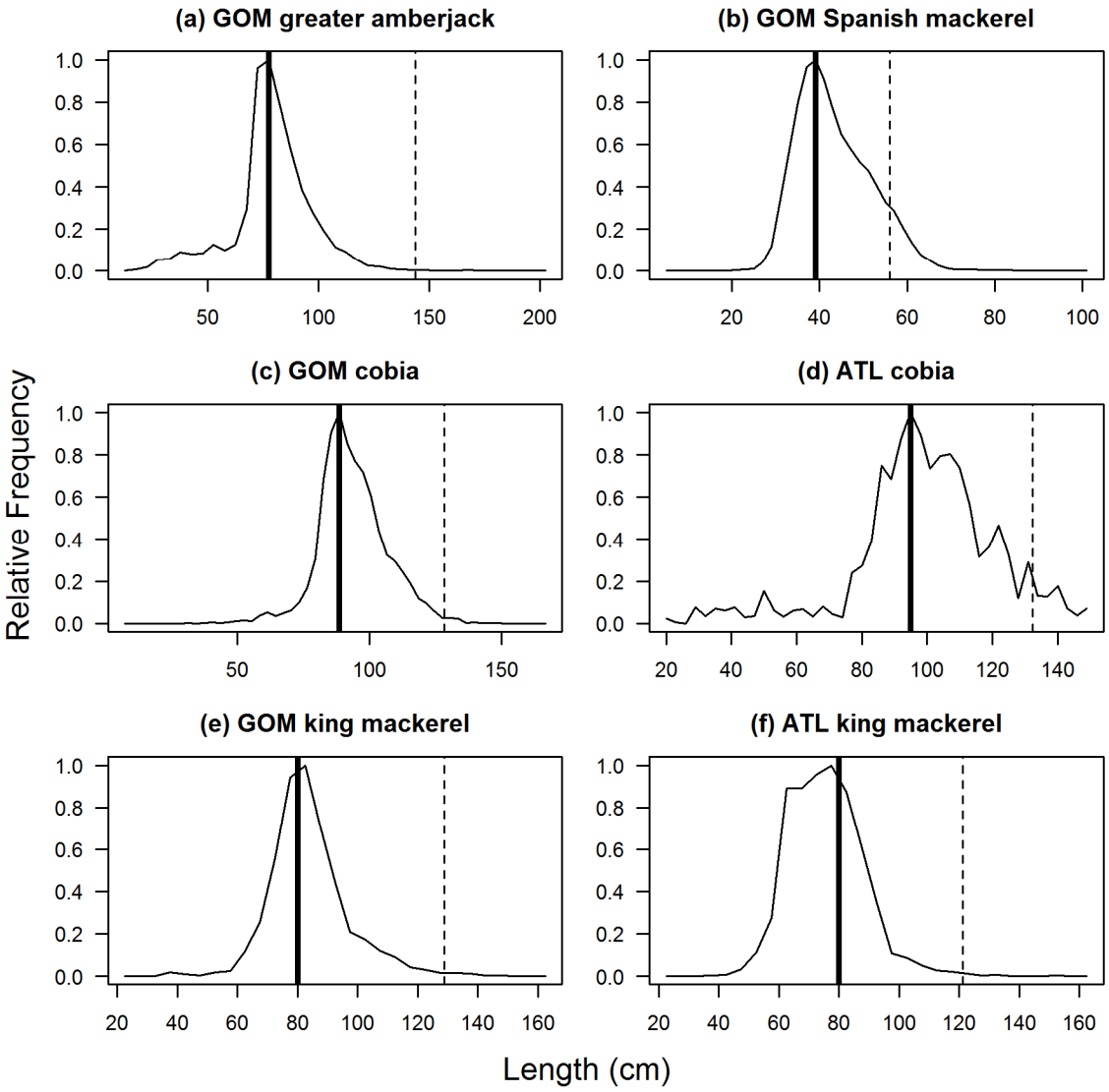

Figure 5.1. Summary length compositions summed across all available years of data for the six stocks for the mean length mortality estimators. Solid vertical line indicates L_c and dashed vertical line indicates L_∞.

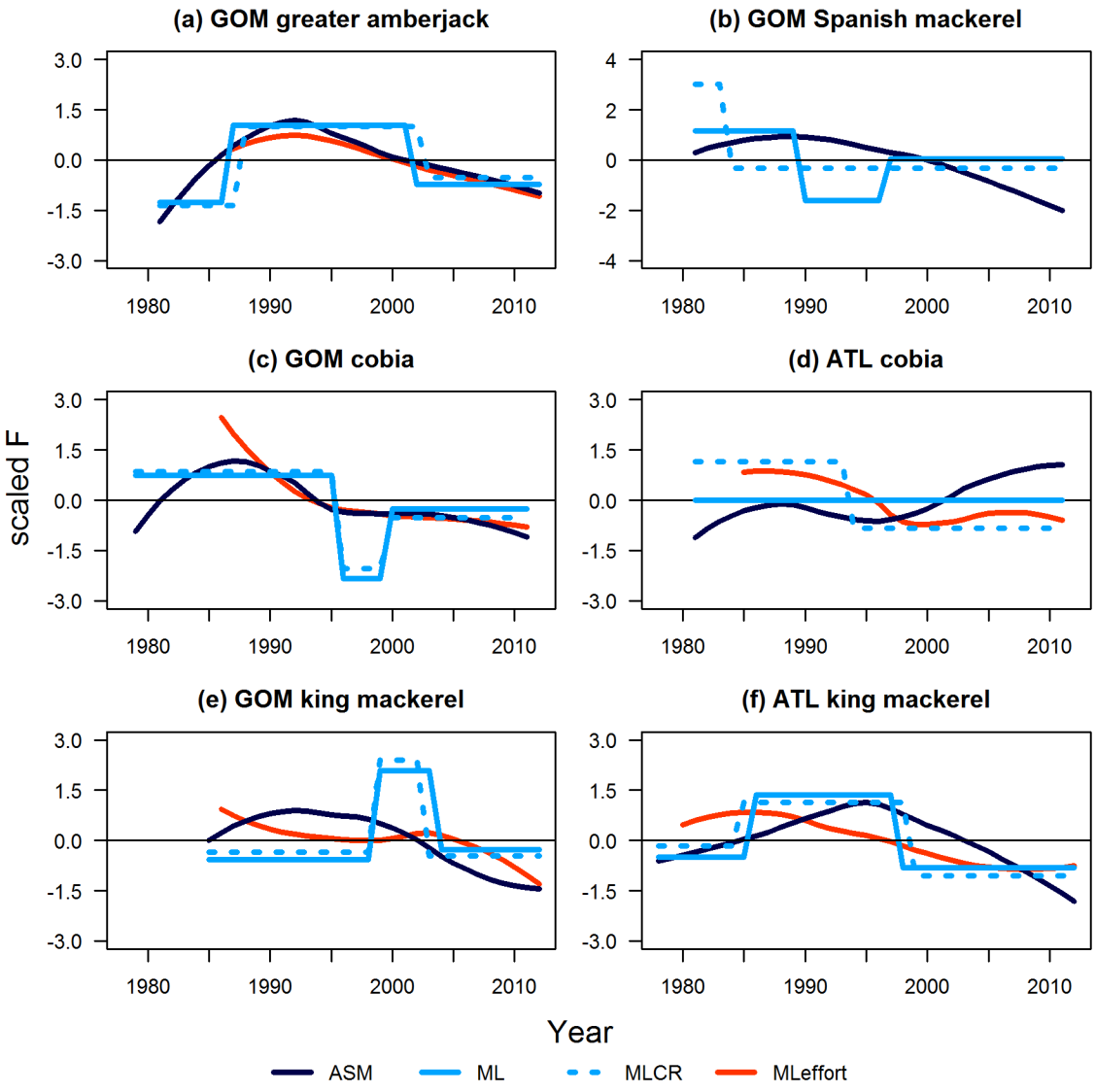

Figure 5.2. Estimates of scaled *F* from the four models (ASM = age-structured model, ML = mean length, MLCR = mean length with catch rate, MLeffort = mean length with effort). Annual estimates were converted to Z-scores and, for ASM and MLeffort, smoothed over time with a loess regression line. The MLeffort model did not converge for GOM Spanish mackerel.

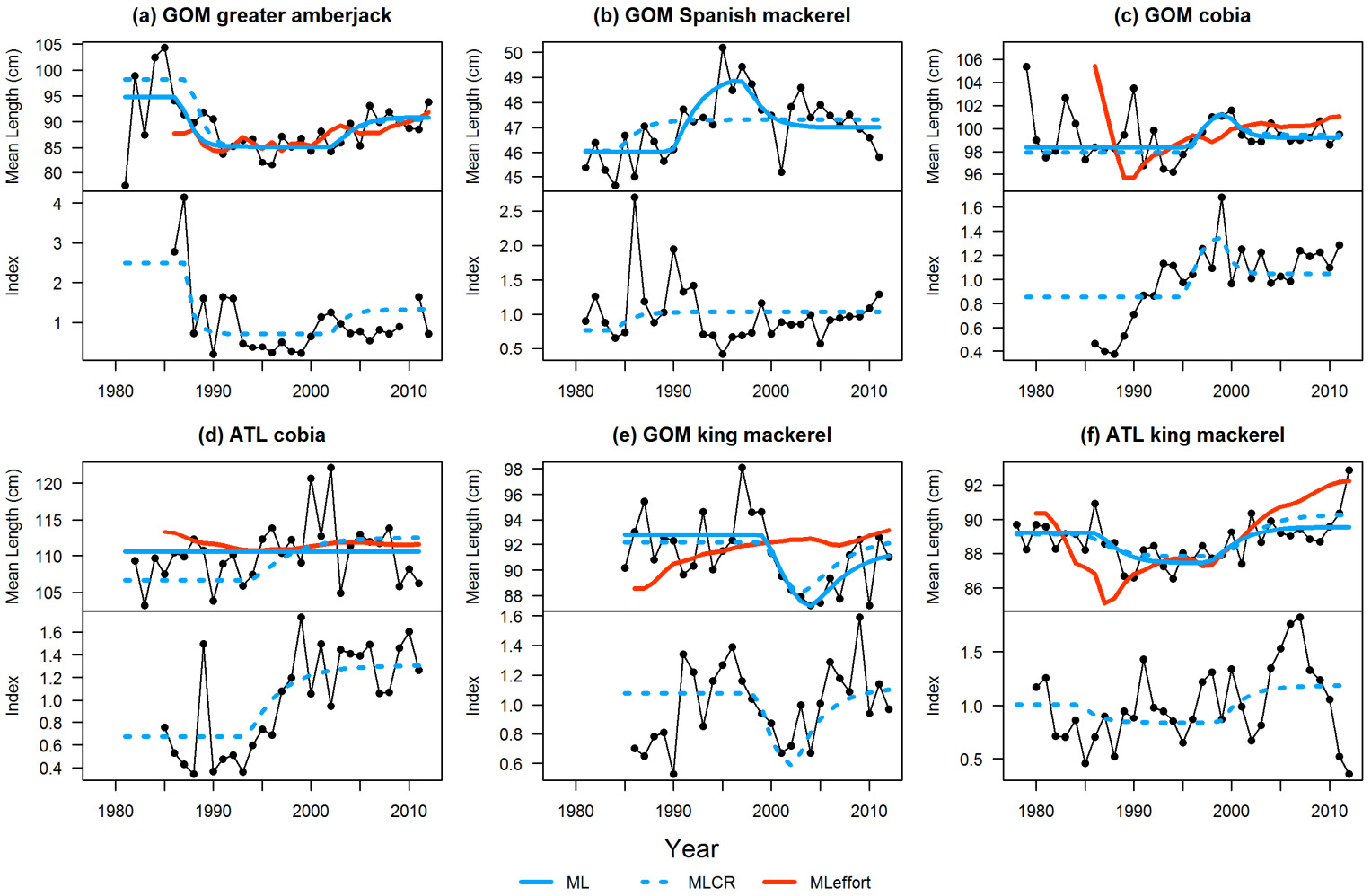

Figure 5.3. Observed (connected points) and predicted mean lengths (colored lines) from the three length-based mortality estimators (ML = mean length, MLCR = mean length with catch rate, MLeffort = mean length with effort) and observed and predicted index for the MLCR model.

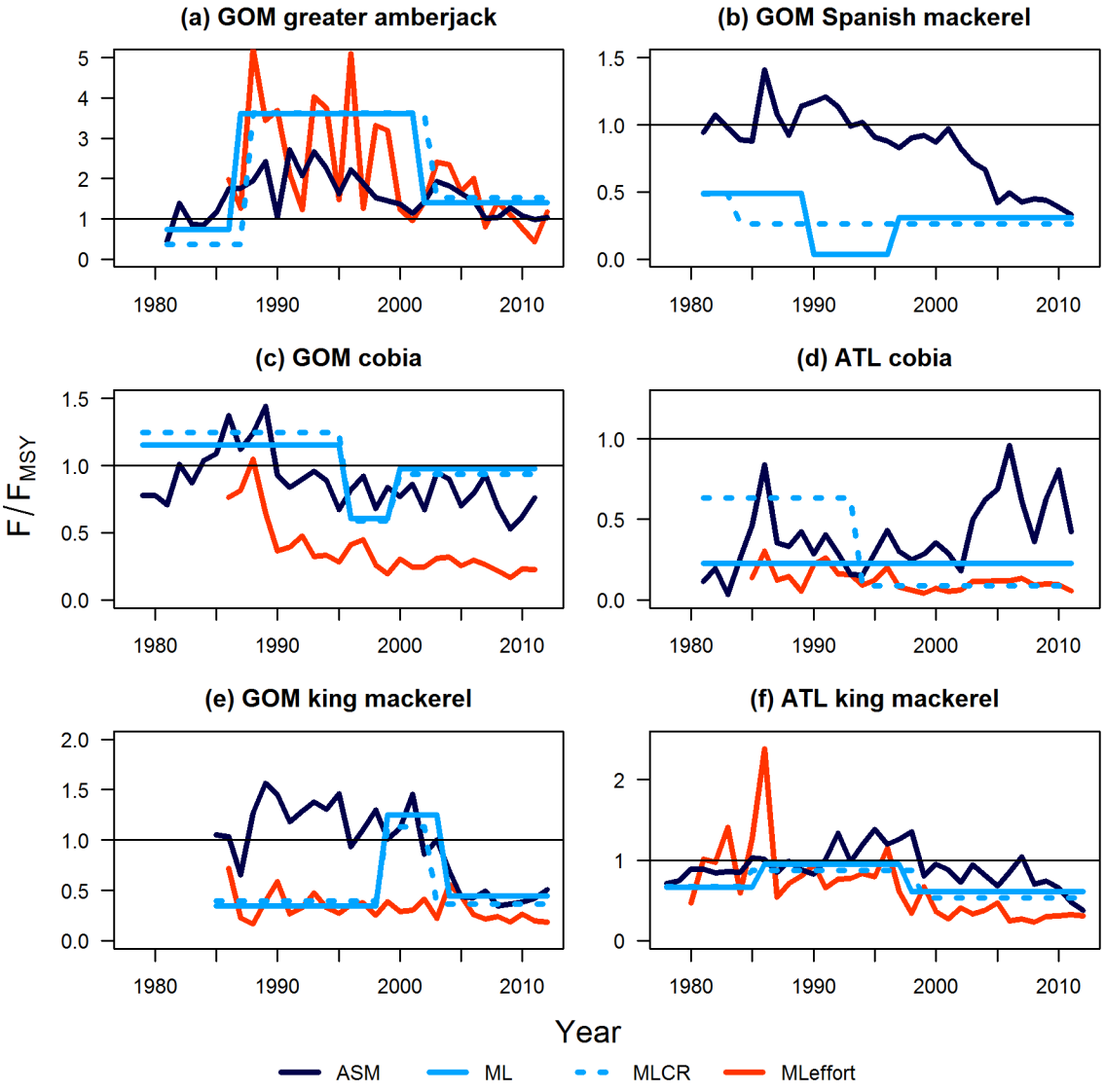

Figure 5.4. Annual estimates of F/F_{MSY} (relative F) from the four models (ASM = age-structured model, ML = mean length, MLCR = mean length with catch rate, MLeffort = mean length with effort). The ASM was the Beaufort Assessment Model for ATL Cobia and Stock Synthesis for all other stocks. F_{MSY} was estimated in the ASM for ATL Cobia while for all other methods, the F_{MSY} proxy is $F_{30\%}$.

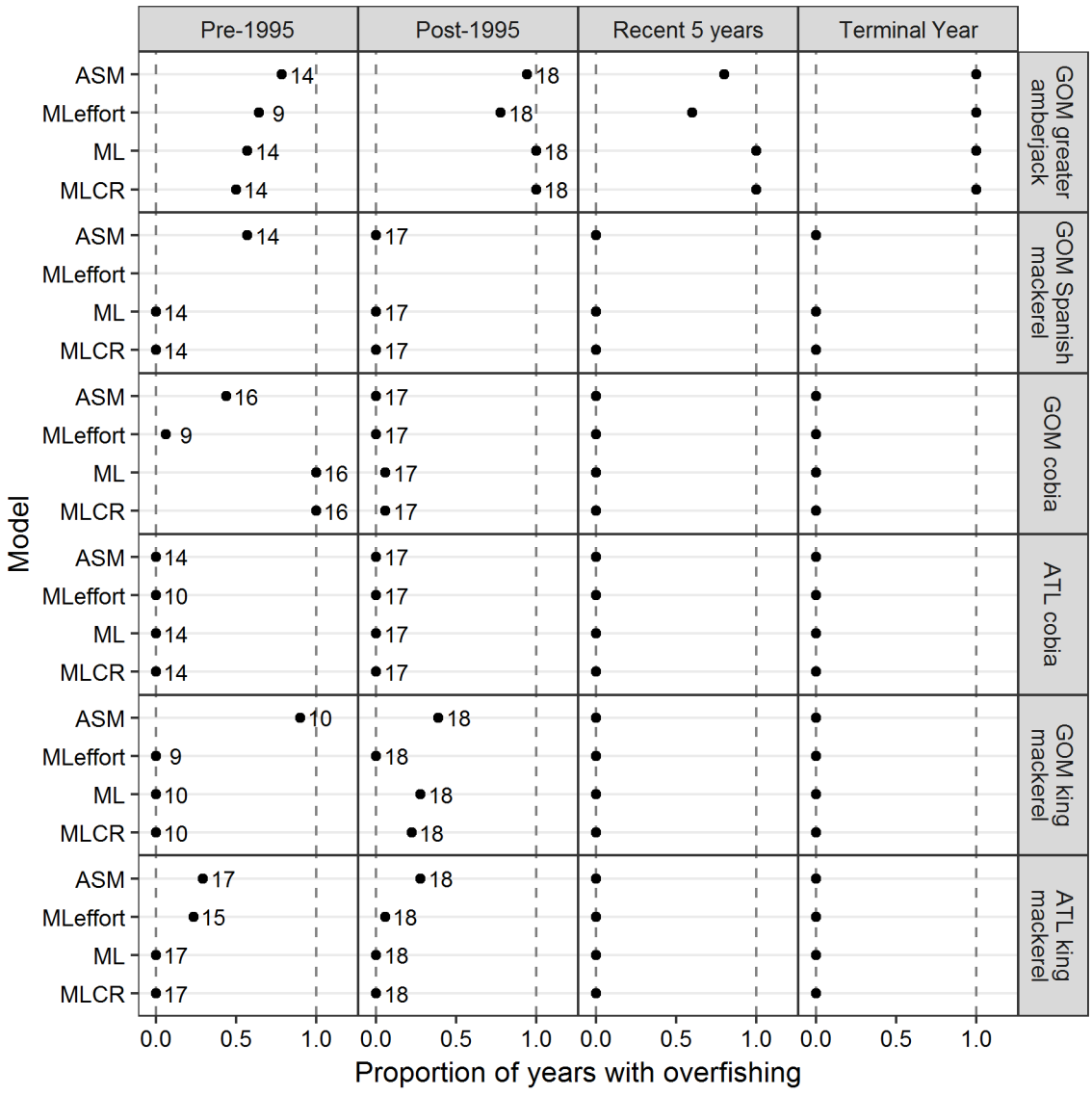

Figure 5.5. The proportion of years with overfishing as estimated with the four models within the respective time periods for the 6 stocks. The MLeffort model did not converge for GOM Spanish mackerel. For Pre-1995 and Post-1995, numbers indicate the number of years in the assessment for the respective time period.

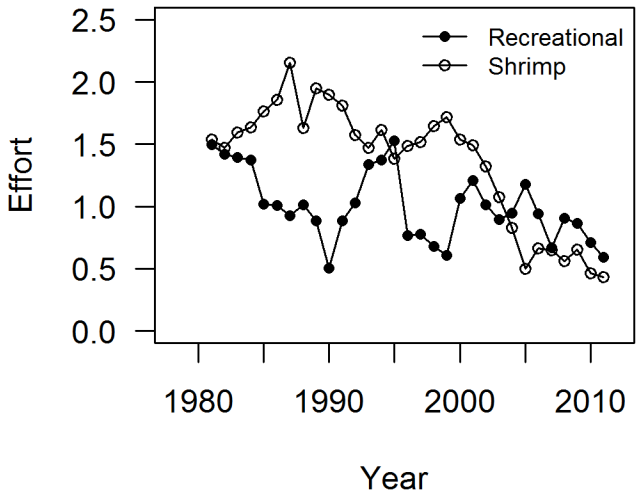

Figure 5.6. Estimates of relative effort for GOM Spanish mackerel from the recreational fleet, obtained as the ratio of the recreational catch and index of abundance, and the shrimp bycatch fleet, estimated as described in Linton (2012). Estimates are scaled so that the time series mean is one.

Chapter 6: Conclusions

6.1. Next steps

The central challenge of any stock assessment is the ability to determine whether the applied methods are appropriate for providing management advice. This dissertation provides insight for stock assessors on the challenges and opportunities with using mean length-based mortality estimators for stock assessment. Data-limited stocks are not limited by geography or management system. Developing countries may lack capacity for formal stock assessments. In developed regions, such as the United States and Europe, the desire to provide scientific advice results in a demand for assessments of all managed stocks (Berkson and Thorson 2015). Limitations in time and expertise create challenging conditions for the assessment and management of marine resources. The best scientific advice for management is traditionally thought to be that obtained from assessments with age-structured models. However, those assessments should be considered as exceptions to the norm. In U.S. federally managed waters, 70% (354 out of 504) of stocks do not have age-structured assessments (Newman et al. 2015). The International Council for the Exploration of the Sea (ICES) manages over 200 stocks in the Northeast Atlantic, and more than half of those stocks do not have assessment advice. (Jardim et al. 2015).

Alternative methods, such as those presented in this dissertation, will be important tools for addressing these limitations. The features and challenges of the mean length-based mortality estimators were addressed through discussion, simulation, and case

studies. A brief description of the many size-based mortality estimators and the available computer software used to implement them was provided in the review chapter (Chapter 1). Case studies throughout this dissertation provided examples of how data are processed and analyzed with mean length mortality estimators (Chapters 3-5). This dissertation will serve as an introduction and guide to prospective users of these methods.

For stock assessment, the mean length-based methods can be used in three ways. First, they can be used to infer historical trends in mortality, as shown in Chapters 3 and 4. The mean length-based mortality estimators are methodologically rich and can accommodate multiple sources of data when available. In a single-species context, a model was developed to estimate mortality from both mean length and catch rate data (Chapter 4). The analysis for Puerto Rico mutton snapper in illustrated a common situation where assessments produce conflicting results. From the mean length data alone, the mortality rate was estimated to be relatively high in the year 2002, the terminal year of the data. When the catch rates were considered with the mean length data, a smaller mortality rate is estimated because a reduction in mortality occurred just prior to 2002. Although the model results were conflicting, the trend in the mean length did not necessarily contradict that in the catch rate. High variability in the mean lengths precluded acceptance of a more complex model with the recent reduction in mortality. The reduction in mortality was more evident in the catch rates which were less noisy. Thus, the trends length and catch rate data are consistent with each other when considered together.

In a multispecies context, multiple stocks that are caught together by the same fishing gear may show synchrony in changes in fishing mortality over time. In a multiple

likelihood framework, it is possible to perform multispecies analyses by specifying parameters in the model that are common to all species (Chapter 4). The benefit of these two extensions is to allow for corroboration of estimated mortality trends among different time series of data.

Second, the mean length-based methods can be used to ascertain stock status. After estimating total mortality, fishing mortality can be derived with an estimate of natural mortality. With a proxy for fishing mortality at maximum sustainable yield, one can make a determination of whether overfishing is occurring (Chapter 5). Statistical diagnostics provide a mechanism for determining whether the method is internally consistent (without external validation from other data) and deciding whether to accept or reject an analysis. For stocks such as Gulf of Mexico greater amberjack and U.S. Atlantic king mackerel, the three length-based models agree with each other in estimated mortality trends and there are good diagnostics in terms of residuals. The trends agree with those from the accepted assessment model. For Gulf of Mexico Spanish mackerel, model diagnostics are poorer, with non-convergence and trends in residuals among the different mean length-based models.

Finally, the output from these methods, and size-based methods in general, can be used to provide advice for the management of these stocks (Hordyk et al. 2015). Advice in the form of a harvest control rule can be tested via management strategy evaluation (MSE), a closed-loop simulation in which a population model is projected into the future based on successive implementations of the control rule from generated data and assessments (Walters and Martell 2004; Punt et al. 2016). Here, potential control rules could be based on the ratio of mortality estimates relative to reference points (ICES

2017), as demonstrated in Chapter 5. In the MSE, management goals, e.g., reducing fishing mortality to be sufficiently precautionary and rebuilding the stock biomass to be in good condition, are defined and operationalized. These simulations can evaluate whether the management goals are met with candidate harvest control rules that use mean length mortality estimators.

Overall, mean length-based methods are attractive options as alternative assessment methods for providing estimates of historical mortality estimates and current status for the many unassessed stocks (ICES 2015; ICES 2016), and work is underway to develop them as tools for providing management advice (ICES 2017).

6.2. References

Berkson, J., and J. T. Thorson. 2015. The determination of data-poor catch limits in the United States: is there a better way? ICES Journal of Marine Science 72:237-242.

Hordyk, A. R., N. R. Loneragan, and J. D. Prince. 2015. An evaluation of an iterative harvest strategy for data-poor fisheries using the length-based spawning potential ratio methodology. Fisheries Research 171:20-32.

ICES. 2015. Report of the Fifth Workshop on the Development of Quantitative Assessment Methodologies based on Life-history Traits, Exploitation Characteristics and other Relevant Parameters for Data-limited Stocks (WKLIFE V), 5–9 October 2015, Lisbon, Portugal. ICES CM 2015/ACOM:56. 157 pp.

ICES. 2016. Report of the Workshop to consider MSY proxies for stocks in ICES category 3 and 4 stocks in Western Waters (WKProxy), 3–6 November 2015, ICES Headquarters, Copenhagen. ICES CM 2015/ACOM:61. 183 pp.

ICES. 2017. Report of the Workshop on the Development of the ICES approach to providing MSY advice for category 3 and 4 stocks (WKMSYCat34), 6–10 March 2017, Copenhagen, Denmark. ICES CM 2017/ ACOM:47. 53 pp.

Jardim, E., M. Azevedo, and N. M. Brites. 2015. Harvest control rules for data limited stocks using length-based reference points and survey biomass indices. Fisheries Research 171:12-19.

Newman, D., J. Berkson, and L. Suatoni. 2015. Current methods for setting catch limits for data-limited fish stocks in the United States. Fisheries Research 164:86-93.

Punt, A. E., D. S. Butterworth, C. L. de Moor, J. A. A. de Oliveira, and M. Haddon. 2016. Management strategy evaluation: best practices. Fish and Fisheries 17:303-334.

Walters C. J. and S. J. D. Martell. 2004. Fisheries Ecology and Management. Princeton University Press, Princeton, N.J.

Appendix A: Derivation of the length-converted catch curve (Chapter 2)

The age-based catch curve is of the form:

$$\log(C_t) = a - Zt\,, \tag{A.1}$$

where C_t is the catch at age t, Z is total mortality, and a is a constant.

In a length frequency distribution, length bins of larger animals contain more age groups than bins with smaller ones due to the decreasing growth rate of older individuals. Thus, abundance at size in an equilibrium population is a function of individual growth rate and mortality over time (Ricker 1975; van Sickle 1977; Pauly 1983). Assuming the length bins are narrow, the length-based catch curve is of the form:

$$\log\left(C_i \frac{dL_{t_i}}{dt}\right) = a - Zt_i\,, \tag{A.2}$$

where C_i is the catch in the i-th length bin, t_i is the ages at the midpoint of the i-th length bin in the length frequency distribution (assuming deterministic growth), $L_{t_i} = L_\infty\{1 - \exp[-K(t_i - t_0)]\}$ is the von Bertalanffy growth equation for length at age t_i, and $\frac{dL_{t_i}}{dt}$ is the instantaneous growth rate evaluated at the corresponding midpoint of the i-th length bin. The following substitutions are made:

$$\log\left(\frac{dL_{t_i}}{dt}\right) = \log(KL_\infty) - K(t_i - t_0) \tag{A.3}$$

$$t' = K(t_i - t_0)\,, \tag{A.4}$$

where t' is the relative age defined as a variable transformation. After substitution and simplification, Equation A.2 reduces to:

$$\log(C_{t'}) = \tilde{a} + \left(1 - \frac{Z}{K}\right)t', \tag{A.5}$$

where $\tilde{a}$ is a nuisance parameter of all constant terms. Equation A.5 is a linear equation of the form:

$$\log(C_{t'}) = \tilde{a} + bt', \tag{A.6}$$

where $\tilde{a}$ and b are the intercept and slope, respectively. Using Equations A.5 and A.6, total mortality Z is solved:

$$Z = K(1-b). \tag{A.7}$$

From a length frequency distribution, the midpoint of the length bins can be converted to relative ages t' also defined by the von Bertalanffy growth equation:

$$t' = -\log\left(1 - \frac{L_{t_i}}{L_\infty}\right), \tag{A.8}$$

with the logarithm of the catch in that length bin used in a linear regression to estimate the slope of Equation A.6 and thus total mortality in Equation A.7.

Appendix A: References

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

Van Sickle, J. 1977. Mortality Rates from Size Distributions: The Application of a Conservation Law. Oecologia (Berlin) 27:311-318.

Pauly, D. 1983. Length-converted catch curves: A powerful tool for fisheries research in the tropics (Part I). Fishbyte 1(2):9-13.

Appendix B: Derivation of the Transitional Behavior of Weight per Unit Effort (Chapter 4)

The catch rate in weight per unit effort (*WPUE*) is:

$$WPUE = qB, \tag{B.1}$$

where *q* is the catchability coefficient and *B* is the biomass. In equilibrium, the biomass is modeled as

$$B = \int_{t_c}^{\infty} N_t W_t dt, \tag{B.2}$$

where N_t is the abundance at age *t* and is calculated as $N_t = R \cdot \exp\left[-Z\left(t - t_c\right)\right]$; t_c is the age of full selectivity; and W_t is the abundance at age *t* and is calculated as $W_t = W_\infty \left\{1 - \exp\left[-K\left(t - t_0\right)\right]\right\}^b$. The weight at age is a composite of the allometric weight-length equation, $W_t = aL_t^{\ b}$, with the von Bertalanffy equation for length at age, $L_t = L_\infty \left\{1 - \exp\left[-K\left(t - t_0\right)\right]\right\}$, where $W_t = aL_\infty^{\ b}$. After substitution, equation (B.2) becomes

$$B = RW_\infty \int_{t_c}^{\infty} \exp[-Z(t - t_c)]\{1 - \exp[-K(t - t_0)]\}^b dt, \tag{B.3}$$

where *R* is the recruitment at age t_c; *Z* is the instantaneous total mortality rate; and W_∞, *K*, and t_0 are growth parameters. To evaluate equation B.3, the substitution $u = 1 - \exp[-K(t - t_0)]$ is made, which implies

$$t = t_0 - \frac{1}{K}\ln(1 - u) \tag{B.4}$$

and

$$dt = \frac{1}{K(1-u)} du. \tag{B.5}$$

After substitution, Equation B.3 simplifies to

$$B = \frac{RW_\infty}{K}\left(1-\frac{L_c}{L_\infty}\right)^{-Z/K}\left[Beta\left(1;b+1,\frac{Z}{K}\right)-Beta\left(\frac{L_c}{L_\infty};b+1,\frac{Z}{K}\right)\right], \qquad \text{(B.6)}$$

where L_c is the length at age t_c; and $Beta(x;\alpha,\beta)=\int_0^x u^{\alpha-1}(1-u)^{\beta-1}du$ is the incomplete beta function to be evaluated numerically.

Assuming one change in mortality, the biomass is a function of the time elapsed, d years, since the change in mortality from Z_1 to Z_2,

$$B(Z_1,Z_2,d) = RW_\infty\left[\begin{array}{l}\int_{t_c}^{t_c+d}\exp[-Z_2(t-t_c)]\{1-\exp[-K(t-t_0)]\}^b dt + \\ \int_{t_c+d}^{\infty}\exp(-Z_2 d)\exp\{-Z_1[t-(t_c+d)]\}\{1-\exp[-K(t-t_0)]\}^b dt\end{array}\right]$$

$$= \frac{RW_\infty}{K}\left[\begin{array}{l}\left(1-\frac{L_c}{L_\infty}\right)^{\frac{-Z_2}{K}}\left[Beta\left(\lambda;b+1,\frac{Z_2}{K}\right)-Beta\left(\frac{L_c}{L_\infty};b+1,\frac{Z_2}{K}\right)\right]+ \\ \exp[-(Z_2-Z_1)d]\left(1-\frac{L_c}{L_\infty}\right)^{\frac{-Z_1}{K}}\left[Beta\left(1;b+1,\frac{Z_1}{K}\right)-Beta\left(\lambda;b+1,\frac{Z_1}{K}\right)\right]\end{array}\right]$$

, (B.7)

where $\lambda = 1-(1-L_c/L_\infty)\exp(-Kd)$. The first term represents the biomass of animals recruited after the change in mortality, and the second term represents the biomass of those recruited before the change.

Assume there have been k changes in mortality. Let $\underline{\theta} = \{Z_1, Z_2, ..., Z_{k+1}, d_1, d_2, ..., d_k\}$, where $Z_1, Z_2, ..., Z_{k+1}$ is the vector of total mortality rates that the population that has sequentially experienced over time and d_i is the elapsed duration of mortality rate Z_{i+1}. The general solution for the biomass is:

$$B(\vec{\theta}) = \frac{RW_\infty}{K}\sum_{i=1}^{k+1}\left[a_i w_i\left(1-\frac{L_c}{L_\infty}\right)^{\frac{-Z_{k+2-i}}{K}}\left[Beta\left(\gamma_i; b+1, \frac{Z_{k+2-i}}{K}\right) - Beta\left(\lambda_i; b+1, \frac{Z_{k+2-i}}{K}\right)\right]\right]$$

, (B.8)

where

$$a_i = \begin{cases} 1 & i = 1 \\ \exp\left(-\sum_{j=1}^{i-1} Z_{k+2-j} d_{k+1-j}\right) & i = 2,\ldots,k+1 \end{cases}$$

$$w_i = \begin{cases} 1 & i = 1 \\ \exp\left(Z_{k+2-i}\sum_{j=1}^{i-1} d_{k+1-j}\right) & i = 2,\ldots,k+1 \end{cases}$$

$$\gamma_i = \begin{cases} 1-\left(1-\frac{L_c}{L_\infty}\right)\exp\left(-K\sum_{j=1}^{i} d_{k+1-j}\right) & i = 1,\ldots,k \\ 1 & i = k+1 \end{cases}$$

$$\lambda_i = \begin{cases} \frac{L_c}{L_\infty} & i = 1 \\ 1-\left(1-\frac{L_c}{L_\infty}\right)\exp\left(-K\sum_{j=1}^{i-1} d_{k+2-j}\right) & i = 2,\ldots,k+1 \end{cases}.$$

The corresponding *WPUE* after *k* changes in mortality is

$$WPUE(\underline{\theta}) = q \cdot B(\underline{\theta}) = \widetilde{q} \cdot \widetilde{B}(\underline{\theta}), \quad \text{(B.9)}$$

where $\widetilde{q} = \frac{qRW_\infty}{K}$ is a scaling parameter for the *WPUE* and $\widetilde{B}(\underline{\theta}) = \frac{K}{RW_\infty} B(\underline{\theta})$ is the biomass excluding *K*, *R*, and W_∞. In this way, trends in biomass can be modeled without W_∞. Compared to the model of *NPUE*, the only additional information required to model *WPUE* is the allometric growth exponent *b*.

Appendix C: Technical description of the mean length mortality estimators (Chapter 5)

The ML and MLCR models estimate mortality rates and change points in mortality based on the transitional behavior of the mean length and index following a change in mortality (Gedamke and Hoenig, 2006; Huynh *et al*., 2017). Assume there are k+1 time stanzas (k changes in mortality in the time series). The predicted mean length ($\overline{L}_y$), abundance-based index ($NPUE_y$) and weight-based index ($WPUE_y$) in year y are calculated as

$$\overline{L}_y = L_\infty \frac{\sum_{i=1}^{k+1} \frac{a_{i,y} v_{i,y}}{Z_{k+2-i}} - \left(1 - \frac{L_c}{L_\infty}\right) \sum_{i=1}^{k+1} \frac{r_{i,y} s_{i,y}}{Z_{k+2-i} + K}}{\sum_{i=1}^{k+1} \frac{a_{i,y} v_{i,y}}{Z_{k+2-i}}}, \tag{C.1}$$

$$NPUE_y = \widetilde{q} \sum_{i=1}^{k+1} \frac{a_{i,y} v_{i,y}}{Z_{k+2-i}}, \tag{C.2}$$

and

$$WPUE_y = \widetilde{q} \sum_{i=1}^{k+1} \left\{ a_{i,y} w_{i,y} \left(1 - \frac{L_c}{L_\infty}\right)^{-Z_{k+2-i}/K} \left[\begin{array}{l} Beta\left(\gamma_{k+2-i,y}; b+1, \frac{Z_{k+2-i}}{K}\right) \\ - Beta\left(\lambda_{k+2-i,y}; b+1, \frac{Z_{k+2-i}}{K}\right) \end{array} \right] \right\} \tag{C.3}$$

where

$$a_{i,y} = \begin{cases} 1 & i = 1 \\ \exp\left(- \sum_{j=1}^{i-1} Z_{k+2-j} d_{k+1-j,y} \right) & i = 2,..., k+1 \end{cases} \tag{C.4}$$

$$v_{i,y} = \begin{cases} 1 - \exp\left(Z_{k+2-i} d_{k+1-i,y}\right) & i = 1,..., k \\ 1 & i = k+1 \end{cases} \tag{C.5}$$

$$r_{i,y} = \begin{cases} 1 & i = 1 \\ \exp\left(-\sum_{j=1}^{i-1} \left[Z_{k+2-j} + K \right] d_{k+1-j,y} \right) & i = 2,...,k+1 \end{cases} \tag{C.6}$$

$$s_{i,y} = \begin{cases} 1 - \exp\left(\left[Z_{k+2-i} + K \right] d_{k+1-i,y} \right) & i = 1,...,k \\ 1 & i = k+1 \end{cases} \tag{C.7}$$

$$w_{i,y} = \begin{cases} 1 & i = 1 \\ \exp\left(Z_{k+2-i} \sum_{j=1}^{i-1} d_{k+1-j,y} \right) & i = 2,...,k+1 \end{cases} \tag{C.8}$$

$$Beta(x;\alpha,\beta) = \int_0^x u^{\alpha-1}(1-u)^{\beta-1} du \tag{C.9}$$

$$\gamma_{k+2-i,y} = \begin{cases} 0 & i = 1,...,k \;\&\; y \le D_{k+1-i} \\ 1 - \left(1 - \frac{L_c}{L_\infty} \right) \exp\left(-K \sum_{j=1}^{i} d_{k+1-j,y} \right) & i = 1,...,k \;\&\; y > D_{k+1-i} \\ 1 & i = k+1 \end{cases} \tag{C.10}$$

$$\lambda_{k+2-i,y} = \begin{cases} 0 & i = 1 \;\&\; y \le D_{k+1-i} \\ L_c / L_\infty & i = 1 \;\&\; y > D_{k+1-i} \\ 0 & 1 < i \le k \;\&\; y \le D_{k+1-i} \\ L_c / L_\infty & 1 < i \le k \;\&\; D_{k+1-i} < y \le D_{k+2-i} \\ 1 - \left(1 - \frac{L_c}{L_\infty} \right) \exp\left(-K \sum_{j=1}^{i-1} d_{k+1-j,y} \right) & 1 < i \le k \;\&\; y > D_{k+2-i} \\ L_c / L_\infty & i = k+1 \;\&\; y \le D_{k+2-i} \\ 1 - \left(1 - \frac{L_c}{L_\infty} \right) \exp\left(-K \sum_{j=1}^{i-1} d_{k+1-j,y} \right) & i = k+1 \;\&\; y > D_{k+2-i} \end{cases} \tag{C.11}$$

and

$$d_{i,y} = \begin{cases} 0 & y \le D_i \\ y - D_i & D_i < y \le D_{i+1} \\ D_{i+1} - D_i & y > D_{i+1} \end{cases}. \tag{C.12}$$

All additional variables are defined in Table C.1.

Table C.1. Definitions of variables for the ML and MLCR models.

Variable	Definition
i	Index for time stanza ($i = 1, \ldots, k + 1$)
j	Index for time stanzas experienced prior to time stanza i ($j = 1, \ldots, i - 1$)
y	Calendar year
Z	Instantaneous total mortality rate ($year^{-1}$)
D	Change point for mortality (calendar year)
L_∞	Von Bertalanffy asymptotic length
K	Von Bertalanffy growth parameter
$\tilde{q}$	Scaling parameter for index
b	Length-weight exponent

Appendix C: References

Gedamke, T., and Hoenig, J. M. 2006. Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of Goosefish. Transactions of the American Fisheries Society, 135: 476–487.

Huynh, Q. C., Gedamke, T., Porch, C. E., Hoenig, J. M., Walter, J. F., Bryan, M., and Brodziak, J. 2017. Estimating total mortality rates of mutton snapper from mean lengths and aggregate catch rates in a non-equilibrium situation. Transactions of the American Fisheries Society, 146: 803-815.

Appendix D: Spawning potential ratio for the mean length estimators (Chapter 5)

The spawning potential ratio (*SPR*) is calculated as,

$$SPR = \frac{SSBPR(F = F_{SPR\%})}{SSBPR(F = 0)}, \tag{D.1}$$

which is the ratio of the spawning stock biomass per recruit (*SSBPR*) at $F = F_{SPR\%}$ compared to that at $F = 0$. The spawning stock biomass per recruit is

$$SSBPR(F) = \sum_{a=0}^{a_{\max}} N_a w_a m_a \tag{D.2}$$

where the abundance at age a (N_a) is

$$N_a = \begin{cases} 1 & a = 0 \\ N_{a-1}\exp(-Z_{a-1}) & a = 1,...,a_{\max}-1, \\ \dfrac{N_{a-1}\exp(-Z_{a-1})}{1-\exp(-Z_a)} & a = a_{\max} \end{cases} \tag{D.3}$$

the weight at age (w_a) is

$$w_a = \left(\alpha L_\infty{}^{\beta}\right)\{1-\exp[-K(a-a_0)]\}^{\beta}, \tag{D.4}$$

the maturity at age (m_a) is

$$m_a = \begin{cases} 0 & a = 0,...,a_{mat}-1 \\ 1 & a = a_{mat},...,a_{\max} \end{cases}, \tag{D.5}$$

and total mortality at age (Z_a) is

$$Z_a = \begin{cases} M & a = 0,...,a_c - 1 \\ F + M & a = a_c,...,a_{\max} \end{cases}. \tag{D.6}$$

From L_c, the fully selected length, the corresponding age a_c is obtained from the inverse of the von Bertalanffy function,

$$a_c = a_0 - \frac{1}{K}\log\left(1 - \frac{L_c}{L_\infty}\right). \quad \text{(D.7)}$$

From L_{mat}, the length of knife-edge maturity, the corresponding age a_{mat} is also obtained from the inverse of the von Bertalanffy function,

$$a_{mat} = a_0 - \frac{1}{K}\log\left(1 - \frac{L_{mat}}{L_\infty}\right). \quad \text{(D.8)}$$

To obtain the $F_{30\%}$ reference point, Equation D.1 is solved for $F_{SPR\%}$ such that $SPR = 0.3$. All variables are defined in Table D.1.

Table D.1. Definition of variables for spawning potential ratio calculation.

Variable	Definition
F	Instantaneous fishing mortality rate (year^{-1})
M	Instantaneous natural mortality rate (year^{-1})
Z	Instantaneous total mortality rate (year^{-1})
α	Length-weight allometric constant
β	Length-weight allometric exponent
L_∞	Von Bertalanffy asymptotic length
K	Von Bertalanffy growth parameter
a_0	Von Bertalanffy theoretical age at length zero
$a_{\max}$	Maximum age (plus-group)
L_{mat}	Length at maturity
a_{mat}	Age at maturity
L_c	Fully selected length (knife-edge selectivity)
a_c	Fully selected age

Appendix E: Residuals in the application of the mean length-based mortality estimators (Chapter 5)

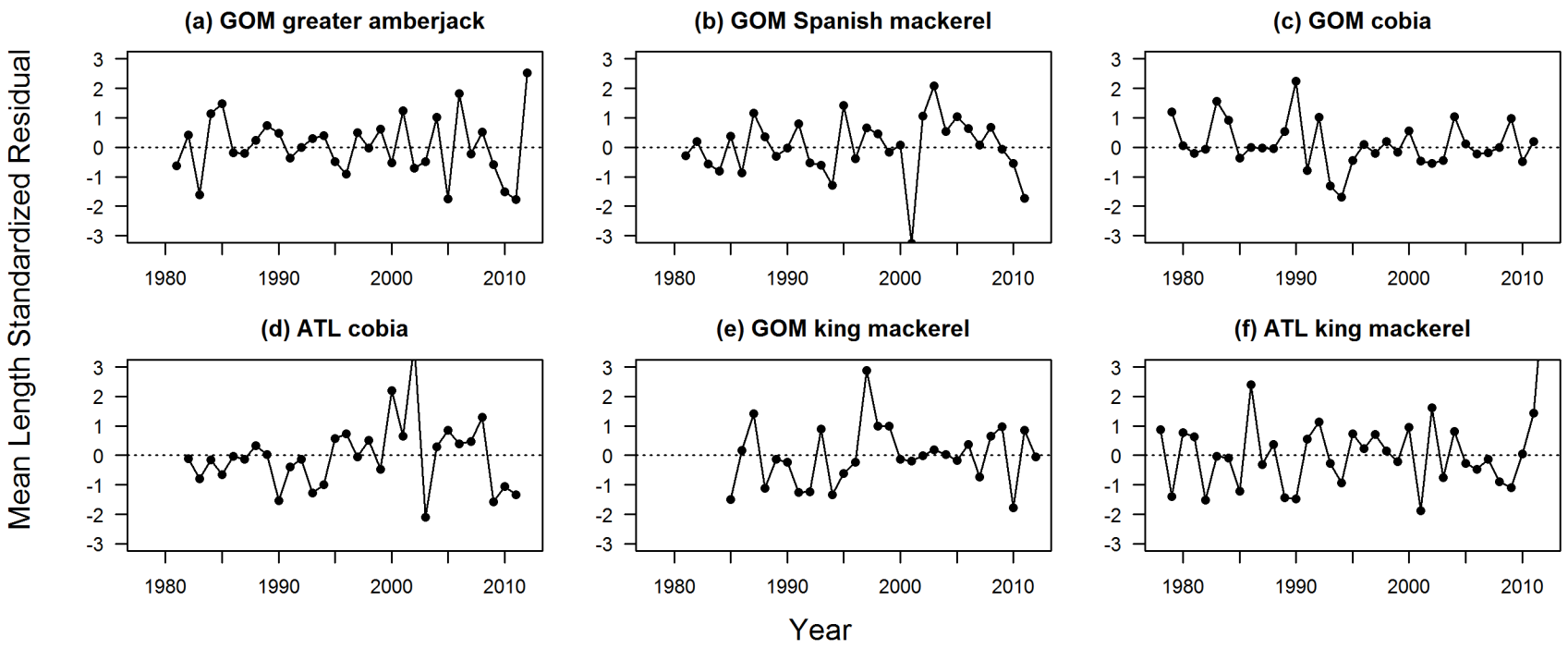

Figure E.1. Standardized residuals of mean length from the ML model. Residuals were calculated by subtracting the predicted value from the observed value and then dividing the difference by the estimated standard deviation.

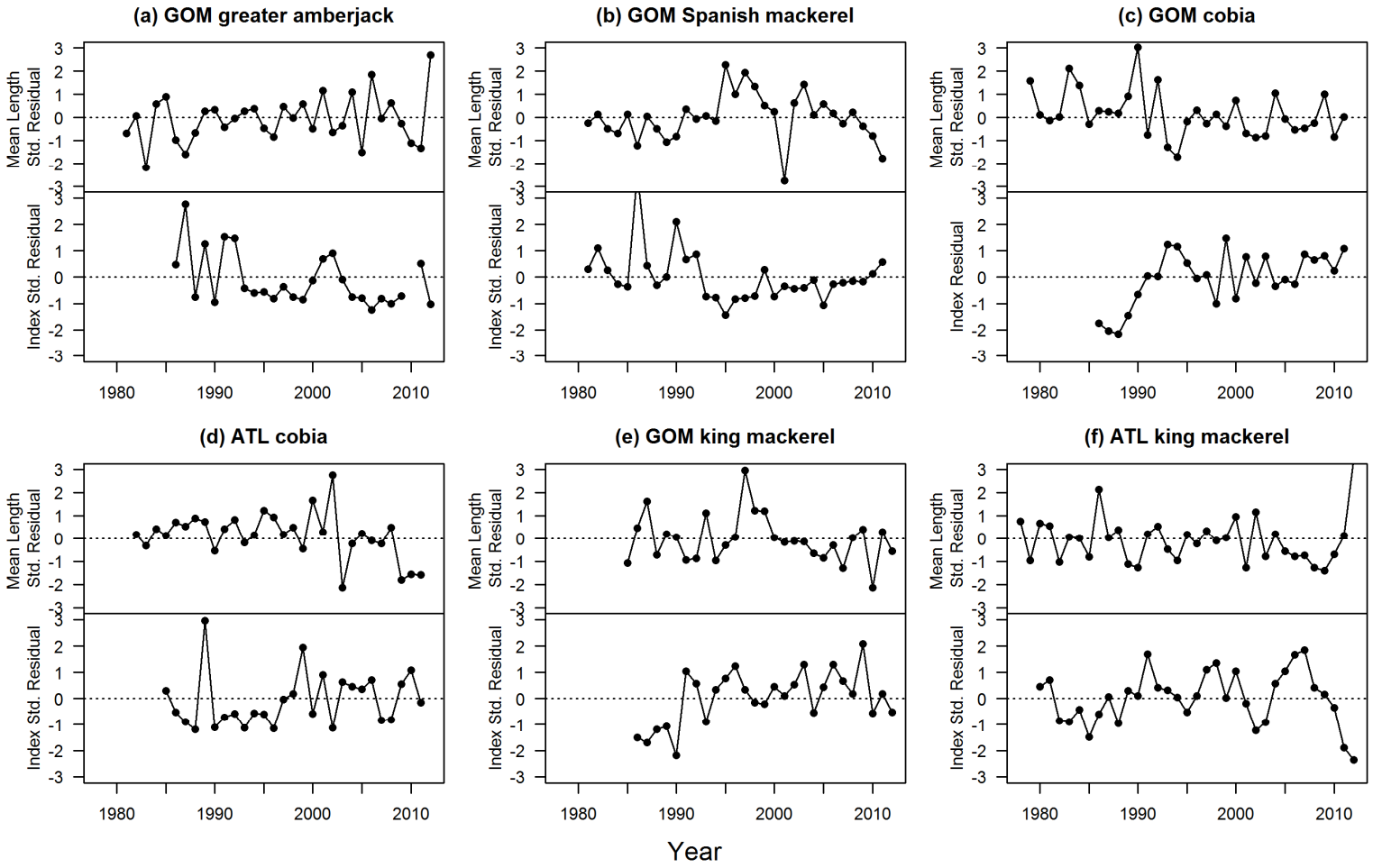

Figure E.2. Standardized residuals of mean length and index from the MLCR model. Residuals were calculated by subtracting the predicted value from the observed value and then dividing the difference by the estimated standard deviation.

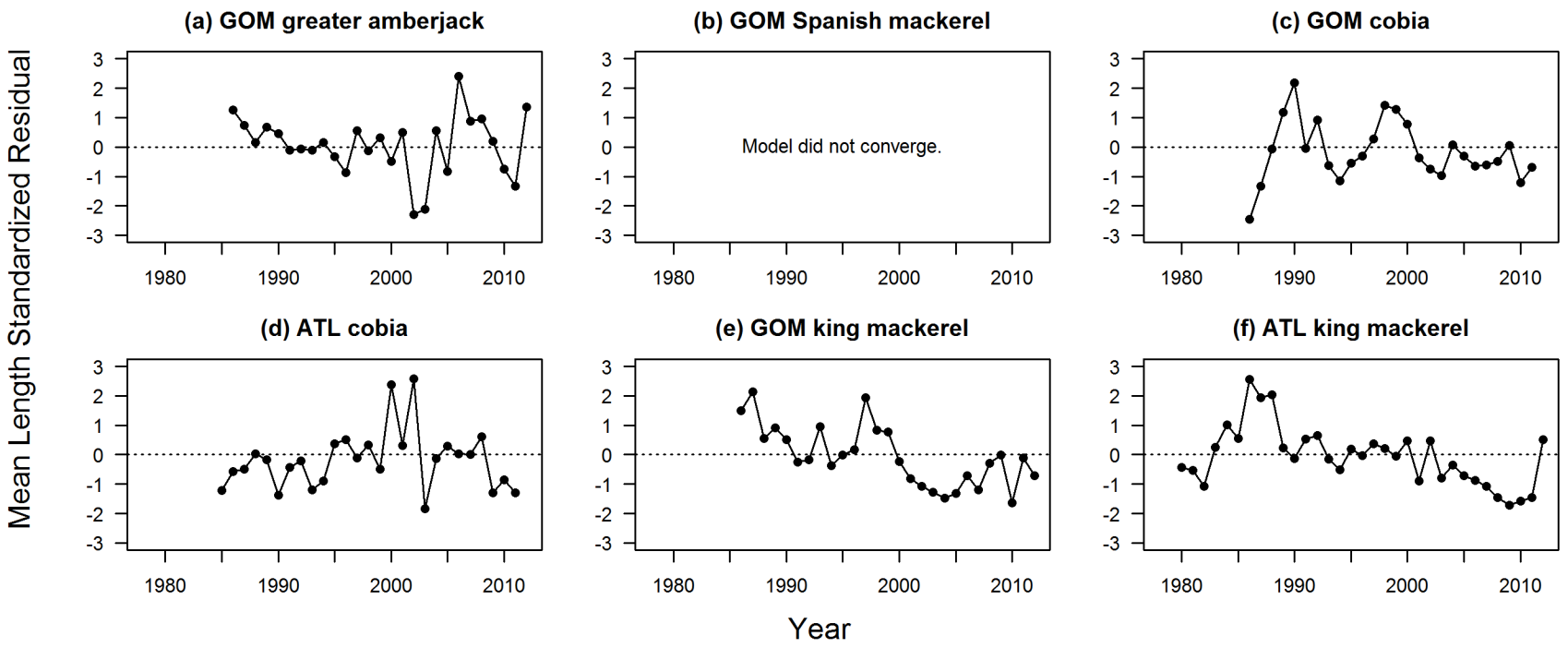

Figure E.3. Standardized residuals of mean length from the MLeffort model. Residuals were calculated by subtracting the predicted value from the observed value and then dividing the difference by the estimated standard deviation.

Supplement to Chapter 2

Additional figures and tables from the simulation study are presented. The results from all methods used in the simulation are reported. All results are reported by stratifying the simulation scenarios by factorial combinations of *M/K* and *F/M*. First, the median *%Bias* and median *%RMSE* are reported across factorial combinations for growth variability, recruitment variability, and selectivity function (Tables S1-S6; Figures S1-S39). Then, the results of sensitivity analyses of sample size (Figures S40-S52) and bin width (Figures S53-S65) are presented.

Table S1. Median %*Bias* of the different methods from factorial combinations stratified by *M/K*, *F/M*, and growth variability. Bold-with-asterisk values indicate the method with the lowest median %*Bias* within each stratum.

F/M	Growth Variability	L1	L2	L3	L4	L5	L6	L7	L8	L9	B1	B2	B3	LB
		M/K = 0.5												
0.25														
	Low	104.1	136.5	115.8	168.9	177	178.5	170.2	162	177.3	**71.2***	141.9	140.9	146.3
	Medium	82.2	84.8	83.2	117.6	119.7	115.9	114	118.2	112.2	**42***	70.8	66.5	139.2
	High	72.8	63.4	72.8	98.2	113.6	97.8	93.7	110.2	93.3	**19.3***	31.6	26.2	73
1														
	Low	37.7	67.3	44.3	65.3	109.7	70.5	60.5	102.6	66.3	**32.2***	61.4	51.2	83.9
	Medium	33.2	45.6	33.5	49.7	79	47.9	46.7	74.5	44.9	**22.6***	37.2	30.3	76
	High	28.1	28.8	27.2	39.6	64.3	37.7	37	58.5	35.1	**13.1***	21.7	16.3	55
5														
	Low	-18.2	-8.9	-20	-9.2	**2.6***	-15.1	-12.7	-0.8	-17.3	-9.8	4.4	-4	31.4
	Medium	-15.1	-8	-20.2	-7.3	1.1	-18	-9.1	-0.9	-19.5	-11.4	4.9	**0.3***	30.8
	High	-15.2	-11	-21.4	-8.2	-1.1	-20.3	-9.2	-2.4	-21.5	-15.3	**0.8***	-1.1	26.6
		M/K = 2												
0.25														
	Low	-6.2	4	-3.4	-4.7	10.9	**-2***	-10.4	3.6	-5.8	11.5	24.8	4.3	59.2
	Medium	-6.6	1.7	-7.9	-2.9	7.1	-6.5	-5.9	3.2	-8.4	**0.7***	13.8	4.3	47
	High	-8.2	**-1.2***	-14	-3.9	5.7	-12.9	-5.7	3.2	-14.3	-3.2	8.4	2.5	31.6
1														
	Low	-16.6	-12.2	-18.4	-12.5	-8.5	-17.8	-17.9	-12.6	-20.8	**1.4***	10.6	-6.5	34.7
	Medium	-16.1	-12.2	-21	-9.9	-8.4	-21	-12.5	-11	-22.8	-7.8	**3.2***	-4.7	30.2
	High	-15.9	-13.5	-27.2	-11.4	-11.2	-27.5	-13	-12.8	-28.7	-15.5	**-1.7***	-5.4	19.2
5														
	Low	-38.1	-37.1	-37.2	-25.7	-31.5	-31.6	-28.7	-34.5	-34.6	-32	**-10.3***	-16.5	11.5
	Medium	-40.8	-39.5	-39.8	-28.2	-34.1	-34.5	-29	-35.6	-35.9	-39	-19.9	-21.7	**7.5***
	High	-45.8	-44	-44.4	-33.7	-38.8	-39.4	-33.4	-39.5	-40	-47.3	-30.2	-30.1	**1.1***

Table S2. Median %*RMSE* of the different methods from factorial combinations stratified by *M/K*, *F/M*, and growth variability. Bold-with-asterisk values indicate the method with the lowest median %*RMSE* within each stratum.

F/M	Growth Variability	L1	L2	L3	L4	L5	L6	L7	L8	L9	B1	B2	B3	LB
			M/K = 0.5											
0.25														
	Low	141.1	211.5	157.9	220.8	275.2	232.1	233.4	272	247.2	**127.8***	223	244.8	319.8
	Medium	100.5	146.9	101	133.8	178.6	130	132.2	171.2	127.9	**90.4***	124.9	123	309.4
	High	84.3	118.3	84.1	105	152.2	104.3	101.4	145	100.6	**70.6***	85.8	82.9	177.7
1														
	Low	68.6	115	**74***	95.6	165.6	101.3	98.2	165.3	102.7	80.8	117	109.8	198.4
	Medium	57.1	94.3	**54.7***	70.5	118.9	64.4	69.5	116	63	69.6	86.6	83.9	184.6
	High	46	73.4	**43.5***	53.1	94.4	48.6	51.7	90.4	47	57.1	68.1	66	142.5
5														
	Low	35.5	32.6	29.6	36.9	34.2	**29.3***	38.5	36	30.7	38	42.1	41.7	74.6
	Medium	33.9	30.9	**29.7***	38.6	32.1	29.8	39.9	33.5	30.7	35.4	42	41.7	71.8
	High	32.2	30.1	30.5	34.7	**27.9***	30.2	35.7	28.9	31	33.2	36.9	37.7	65.5
			M/K = 2											
0.25														
	Low	31.5	36	**29.8***	40.3	45.9	35.2	40.3	46.2	35	49.1	60.6	49.3	144.1
	Medium	29.9	33.2	**27.1***	37.2	38.8	30.6	37.2	39.8	31.2	39.7	48.9	46	126.6
	High	28.5	31.2	**25.6***	34	32.3	26.3	34.2	32.8	27	34.2	41.9	40.9	81.9
1														
	Low	35.1	**25.2***	28.6	38.3	26.1	29.5	40.3	28.3	31.2	41.6	46.3	40.6	95.7
	Medium	34	26.1	31	34.3	**25.1***	31.5	35.7	26.7	32.5	37.3	39.5	38.2	90.8
	High	32.1	25.8	32.5	32	**23.8***	32.9	32.9	24.8	33.8	34.2	32.7	33.6	66.5
5														
	Low	49.1	42.3	42.4	47.1	39.4	**39.7***	55.2	41.8	42.1	45.5	43	48.3	79.3
	Medium	48	43.3	43.6	42.5	40	40.3	45.2	41.3	41.7	46	**38.4***	41.4	73.4
	High	50.3	46.4	46.9	42.4	42.3	43	43.4	42.9	43.7	51.1	**39.9***	41.1	59

Table S3. Median %*Bias* of the different methods from factorial combinations stratified by *M/K*, *F/M*, and recruitment variability. Bold-with-asterisk values indicate the method with the lowest median %*Bias* within each stratum.

F/M	Recruitment Variability	L1	L2	L3	L4	L5	L6	L7	L8	L9	B1	B2	B3	LB
		M/K = 0.5												
0.25														
	Low	75.1	105.6	85.5	152.4	135	162.1	144.7	124.6	158.5	**30.2***	108	110.3	111.1
	High	103.7	132.1	114.7	169.3	175	169.1	168	158.5	163.6	**67.1***	136.8	141.8	145.3
1														
	Low	25.8	36.3	29.4	52	85.4	57.2	47.1	75.5	53.3	**11.4***	36.3	28.9	64
	High	36.6	67	43	63.4	103.4	68	59.2	95.6	64.2	**29.2***	59	50.2	81.6
5														
	Low	-18.4	-11.5	-21.8	-11.1	-0.6	-19.2	-12.8	-2.1	-20.4	-17.3	**-0.5***	-4.2	25.8
	High	-15.6	-8.8	-20.7	-7.2	2.6	-18.3	-9.7	**-0.9***	-19.8	-9.4	5.6	-1.4	31.7
		M/K = 2												
0.25														
	Low	-8.4	**0.1***	-9.7	-6.8	4.7	-5.7	-9.2	**0.1***	-9	-1.7	8.2	-2.3	30.2
	High	-4.8	3.4	-6.2	**-2.6***	9.8	-4.3	-5.6	4.6	-6.4	11.2	20.1	8.1	56.6
1														
	Low	-17.6	-13.1	-20.1	-12.4	-8.9	-18.2	-17.7	-12.8	-21.1	-10.6	**0.9***	-6	22.2
	High	-14.3	-11.6	-22	-8.5	-7.9	-22	-13.7	-11.2	-24.2	-5.6	10.6	**0.3***	40.5
5														
	Low	-43.8	-40.2	-40.4	-32.8	-35.4	-35.7	-35.1	-37.4	-37.6	-40.1	-22.9	-28.5	**4.8***
	High	-39	-37.4	-37.5	-26.6	-31.9	-32	-27.7	-34.8	-35	-32.8	-13.9	-17.6	**9.6***

Table S4. Median %*RMSE* of the different methods from factorial combinations stratified by *M/K*, *F/M*, and recruitment variability. Bold-with-asterisk values indicate the method with the lowest median %*RMSE* within each stratum.

F/M	Recruitment Variability	L1	L2	L3	L4	L5	L6	L7	L8	L9	B1	B2	B3	LB
		M/K = 0.5												
0.25														
	Low	108.6	156.2	125.8	177	228.2	184.8	168.1	225.2	182.4	**87***	186.8	182.8	257.8
	High	140.7	181.7	157.1	199.7	254.8	188.8	195.4	251.1	184.8	**125.4***	222	217	322
1														
	Low	45.7	83.6	**53.3***	84.9	122.1	85.2	79.1	119.3	83.7	53.4	90.2	92.2	158.1
	High	**67.5***	105.3	73.3	101.7	142.9	93.8	103.7	140.6	91.4	80	116.7	115.6	190.8
5														
	Low	30.5	28.7	29.1	31.5	29.4	28.1	33.4	30.6	**29***	32.9	34.3	35.2	65.7
	High	37.7	34	**30.9***	41.6	36.6	31.3	43.3	38	32.5	38.3	44.8	45.4	78.6
		M/K = 2												
0.25														
	Low	26.3	28.4	25.1	29	34	27.5	30.3	34.8	**28***	31.8	39.3	34.7	85.4
	High	34.5	38.3	30.4	39.8	47	34.6	40.2	47	**34.4***	51.2	56.8	52.1	139
1														
	Low	28.7	**22.9***	27.6	30	**22.9***	28.5	31.1	24.6	30.1	33	34	33.1	60.6
	High	36.1	**27.5***	31.6	40.7	27.7	32.5	42.4	29.2	33.6	41.8	50	44.7	105.6
5														
	Low	49	42.7	43	42.3	**39.5***	39.9	45	41.4	41.7	47	**39.5***	41.6	68.9
	High	49.1	42.4	42.5	44.5	40.4	40.8	48.5	42.7	43.2	45.3	**40.3***	44.1	76.3

Table S5. Median %*Bias* of the different methods from factorial combinations stratified by *M/K*, *F/M*, and selectivity. Bold-with-asterisk values indicate the method with the lowest median %*Bias* within each stratum.

F/M	Selectivity	L1	L2	L3	L4	L5	L6	L7	L8	L9	B1	B2	B3	LB
		M/K = 0.5												
0.25														
	Gradual	76.3	70.8	78	114.6	113.1	113.5	111.6	110.3	110.3	**31***	61.6	57.9	112.9
	Steep	74.7	80.4	79.2	113.4	120.8	112	109.5	118.2	107.9	**26.7***	60.2	57.6	82.2
	Dome	146.2	149.3	148.6	175	227.8	170.9	168.1	221.7	167.2	**128.8***	165.3	151.1	278.5
1														
	Gradual	26.6	31.6	29.2	46.4	73.5	45.8	43.4	68.7	42.9	**10.7***	29.4	23.4	64.1
	Steep	27.7	35.5	30.7	45.8	72.1	44.5	43.1	66.7	41.6	**14.9***	30.4	24	48.7
	Dome	60.5	69.3	64.9	77.7	119.3	75.7	74.7	113.9	72.5	**58.3***	83.5	73.4	153.9
5														
	Gradual	-18.4	-12.2	-21.9	-10.3	-1.9	-18.7	-12.2	-3.8	-20.1	-16.6	**-1.2***	-4.3	27.6
	Steep	-10.4	-1.6	-14.9	-5.9	3	-13.9	-8.9	**0.1***	-15.4	1.8	8	2.4	31.2
	Dome	-17	-11	-22.1	-9.2	**-0.4***	-19.7	-10.7	-1.9	-21.2	-15.4	0.8	-2.8	27.9
		M/K = 2												
0.25														
	Gradual	-12.1	-6	-12.6	-7	3.8	-9.5	-11	**-0.5***	-11.3	-2.8	7.7	-2.1	34
	Steep	-7.3	1.2	-7.9	-5.1	6.1	-6.5	-8.8	2.6	-8.4	6.8	11.5	**0.7***	28.1
	Dome	**3.3***	11.1	6.6	15	22.9	12.8	11	19.9	9.9	9.4	29.5	18.3	79.6
1														
	Gradual	-21	-18.5	-26	-15.1	-13.6	-24.4	-19.4	-16.2	-26.3	-14.2	**-1.5***	-9.3	23
	Steep	-13.4	-10.6	-21	-10.3	-8.9	-21	-13.2	-12	-22.8	**0.1***	7.5	-2.9	21.5
	Dome	-16.3	-12.3	-18.6	-8.2	-6.3	-15.9	-10.6	-8.2	-17.1	-10.4	**5.4***	-2.8	38.2
5														
	Gradual	-44.1	-40.8	-41	-32.8	-35.5	-35.8	-33.7	-37.4	-37.6	-41.9	-23.6	-27.6	**5.4***
	Steep	-20.1	-25.8	-26.2	-11.2	-23.3	-23.7	-12.9	-25.4	-25.6	-18.7	**-2.6***	-3.6	39.8
	Dome	-44.3	-41.1	-41.4	-33.6	-36.1	-36.3	-34.1	-37.9	-38.1	-42.3	-24.8	-28.9	**3.5***

Table S6. Median %*RMSE* of the different methods from factorial combinations stratified by *M/K*, *F/M*, and selectivity. Bold-with-asterisk values indicate the method rule with the lowest median %*RMSE* within each stratum.

F/M	Selectivity	L1	L2	L3	L4	L5	L6	L7	L8	L9	B1	B2	B3	LB
		M/K = 0.5												
0.25														
	Gradual	92.7	128.6	94.5	130	166.5	126.8	129.1	156.8	125.3	**77.4***	117.4	117.2	262.5
	Steep	94.5	136.5	94.5	129.1	170.6	125.8	127.5	163.1	123.5	**82***	115.8	115.4	223.2
	Dome	165.6	197.5	165.9	195.9	266.6	190.5	192.7	263.2	185.8	**161.3***	213.8	201.9	420.3
1														
	Gradual	44.5	80.6	49.2	64.5	110.1	61.4	63.5	107.2	60.1	**52.9***	76.4	74.9	160.5
	Steep	49.5	83.3	50.5	63.3	108.5	59.7	62.2	104.5	58.5	**58.2***	80.1	77.1	135.8
	Dome	80	106.4	80.9	99.1	148.9	92.6	97.5	146.7	90.5	**87.9***	121.9	112.6	248
5														
	Gradual	34	31.7	30.4	35.9	30.8	**30.2***	36.9	32.5	31.4	34.6	37.7	38.9	73.1
	Steep	34.5	31	**28.8***	38	31.5	29.3	39.1	32.7	30.2	38.6	42.4	42	66.8
	Dome	33.1	31	**29.7***	35.4	30.1	29.6	36.6	31.6	30.6	34.5	37.4	38.5	72.3
		M/K = 2												
0.25														
	Gradual	29.4	30.2	**25.8***	31.6	33.4	27.2	32.4	34.7	27.8	36.6	41.9	38.9	95.5
	Steep	28.6	29.8	**25.6***	32	33.1	27.4	32.8	33.7	27.9	40.2	44.7	40.1	75.4
	Dome	32.3	35.8	**29.9***	45.2	45.7	37	44.7	45.1	36.1	40.7	57.2	50.9	162.8
1														
	Gradual	33.1	27.9	33.1	33.8	**25***	33.3	35.5	26.7	34.5	35.3	38.2	36.5	82.9
	Steep	32	23.7	30.1	35.2	**23.6***	31.1	37.3	25.4	32.5	36.5	40.6	37.4	62
	Dome	31.9	25.1	28.3	34.8	**24.3***	28.4	36.6	25.5	29.3	36.1	41.3	39.1	105
5														
	Gradual	49.5	43.4	43.7	43.9	40.7	41	47	42.7	43.3	47.9	**40.4***	43.4	73.1
	Steep	34.7	32.7	33.3	35.3	**33.1***	33.7	40.7	34.9	36	33.8	32.8	37.4	79.1
	Dome	49.1	44	44.3	44.6	40.6	41	47.6	42.7	43.2	47.7	**40.2***	43.4	72.3

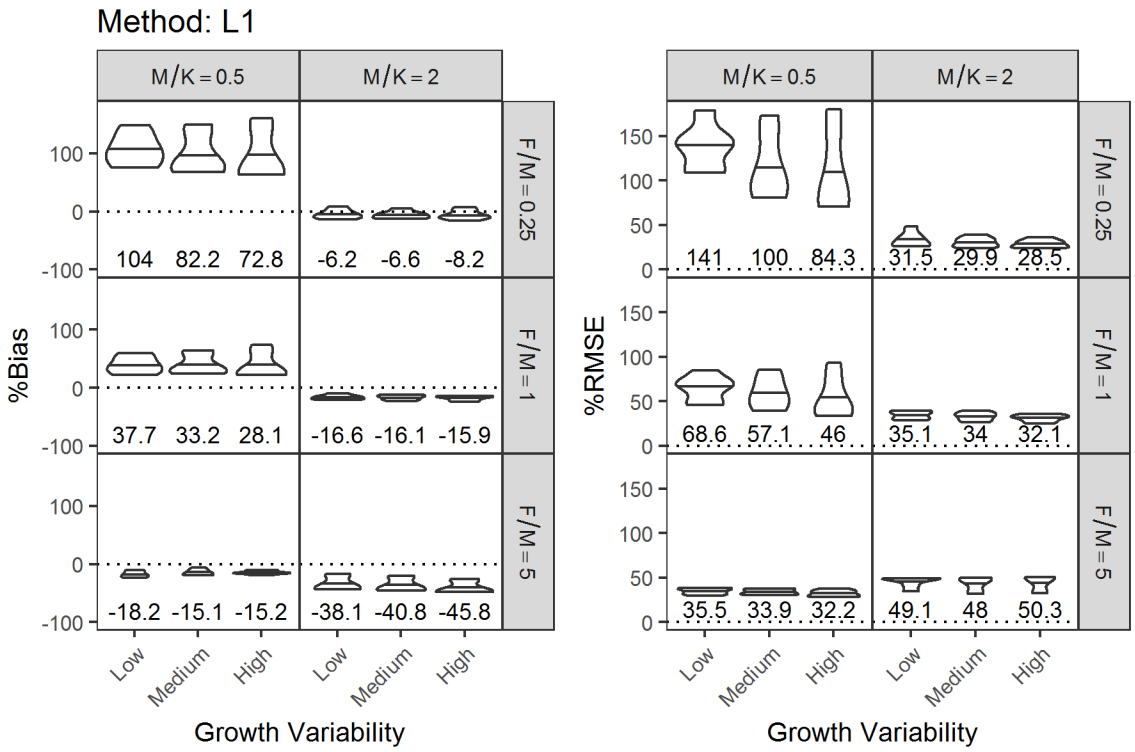

Figure S1. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability for method L1. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

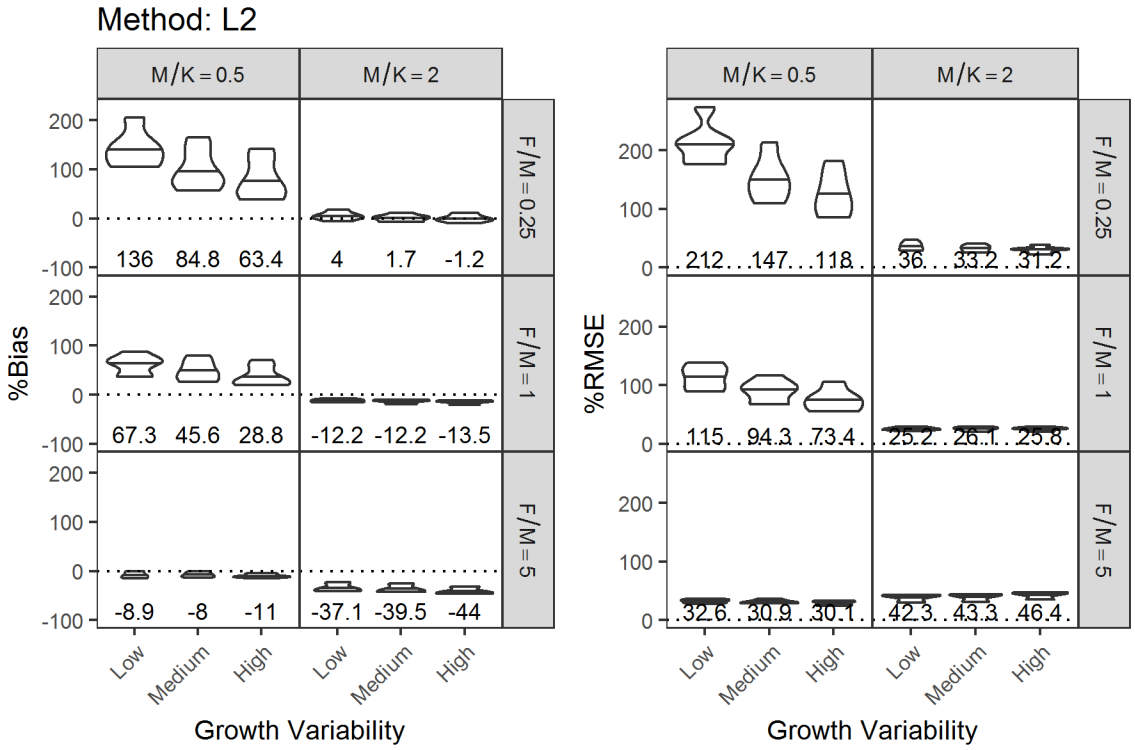

Figure S2. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability for method L2. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

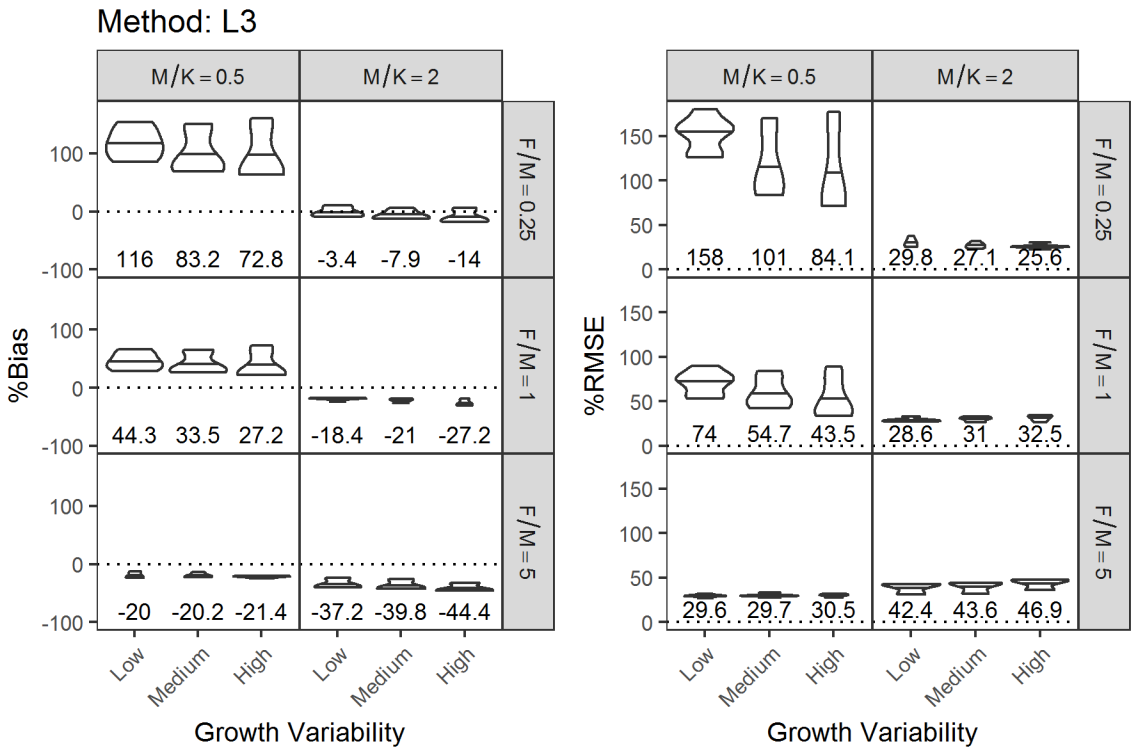

Figure S3. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability for method L3. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

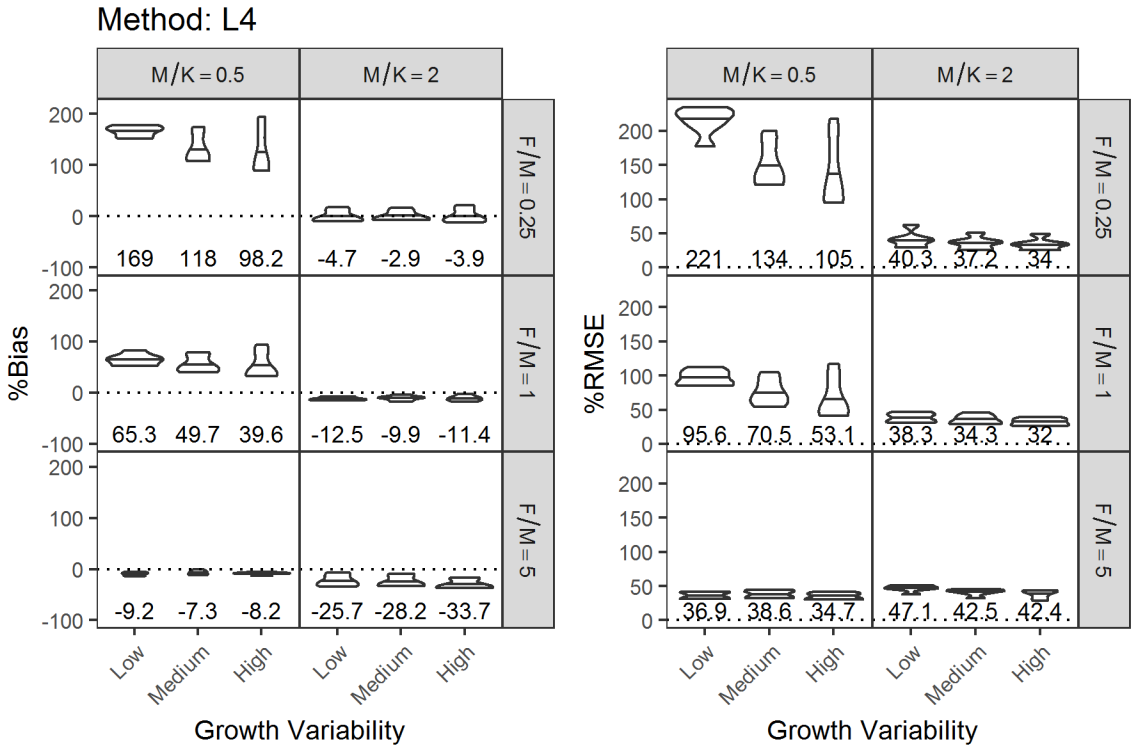

Figure S4. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability for method L4. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

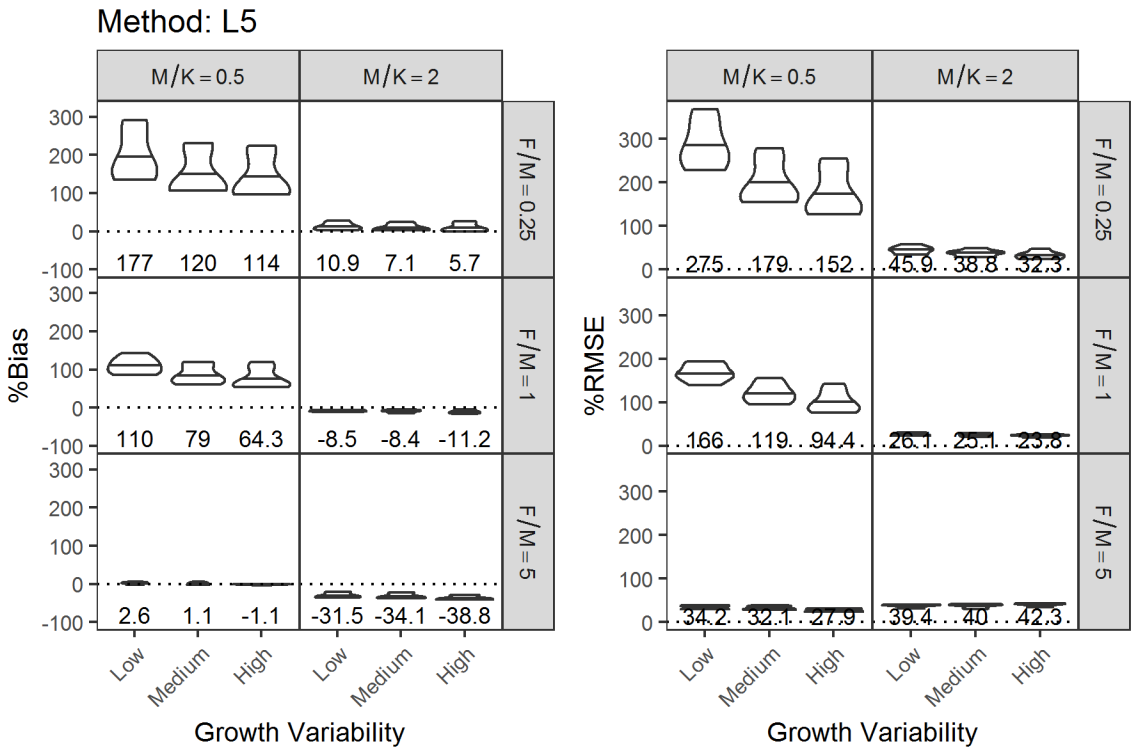

Figure S5. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability for method L5. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

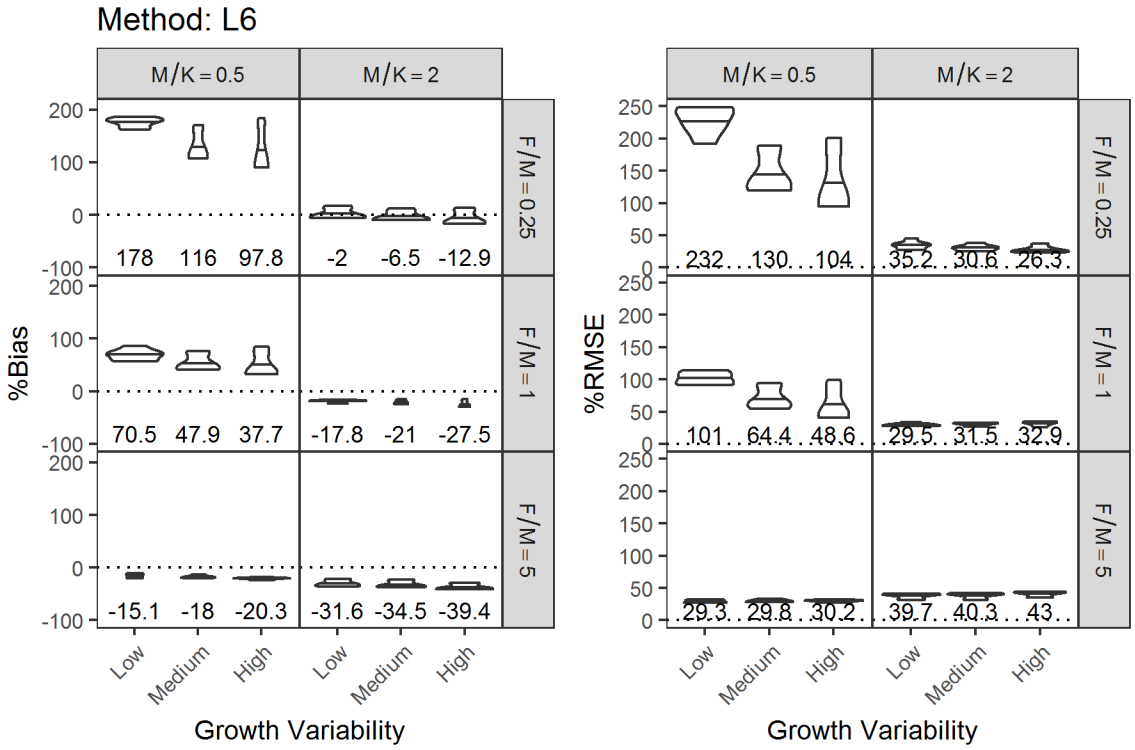

Figure S6. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability for method L6. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

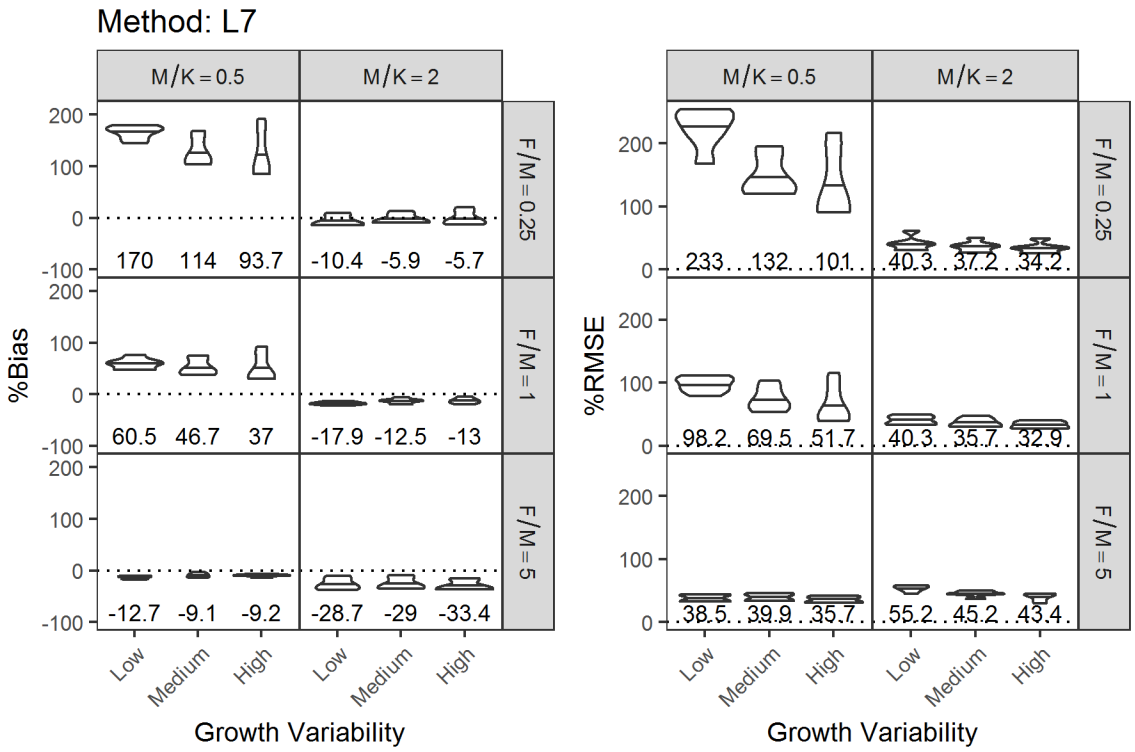

Figure S7. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability for method L7. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

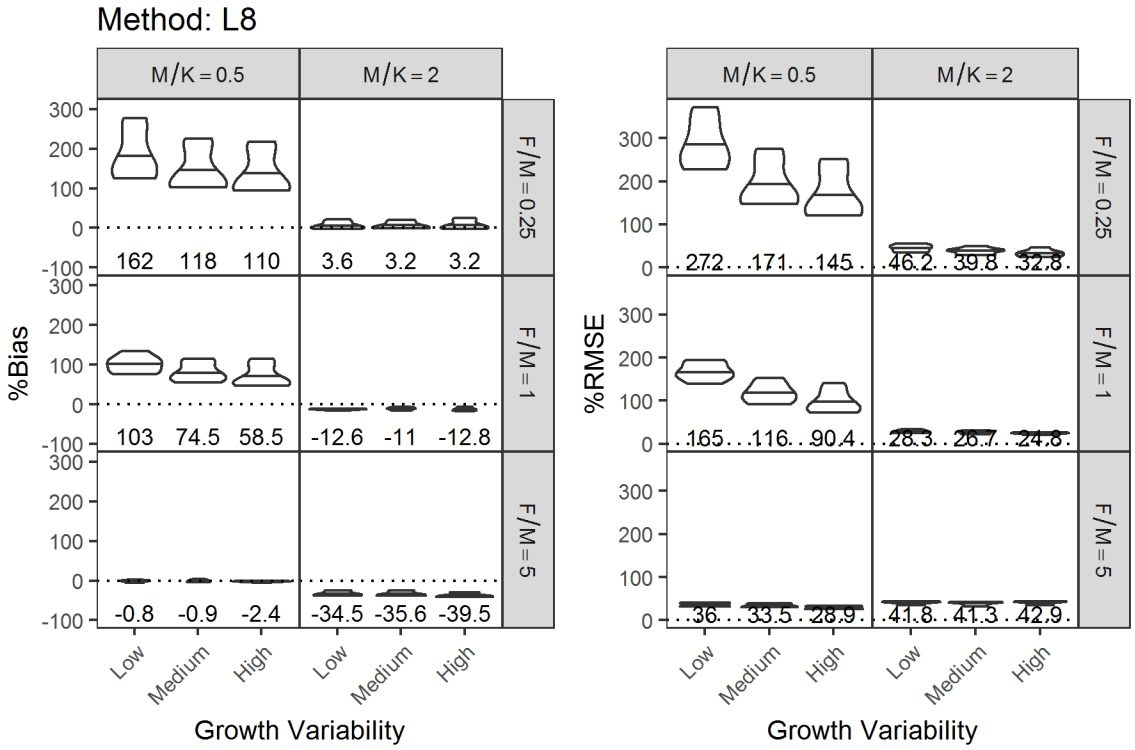

Figure S8. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability for method L8. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

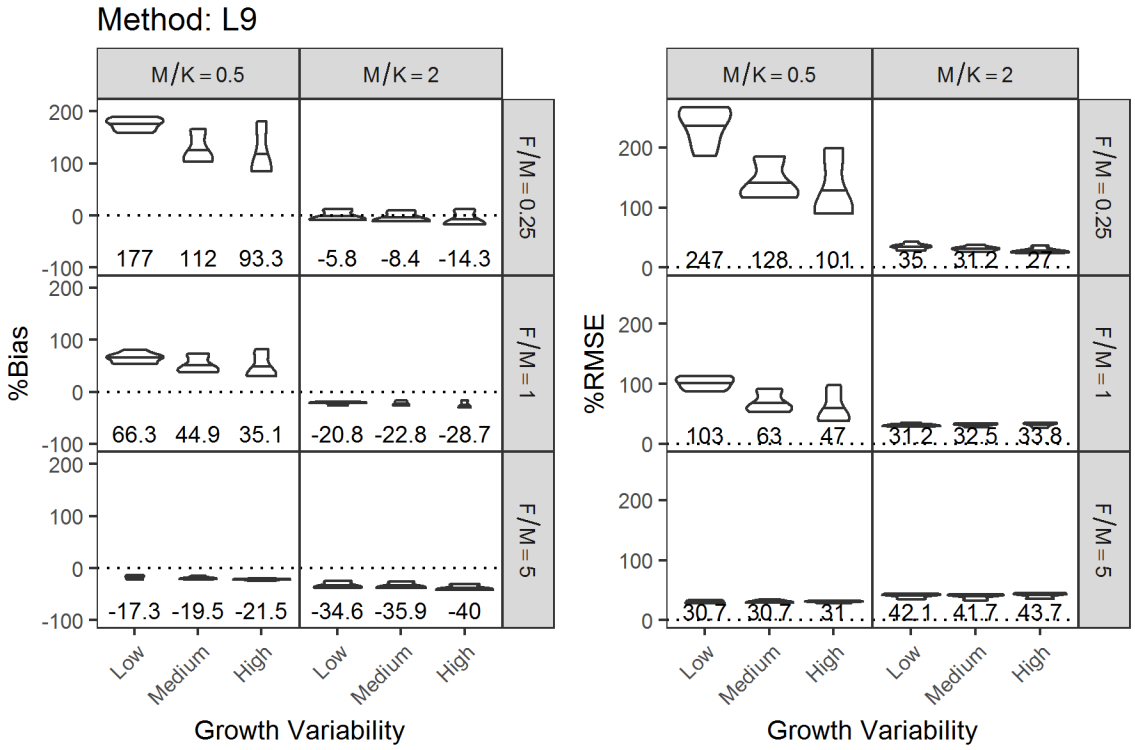

Figure S9. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability for method L9. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

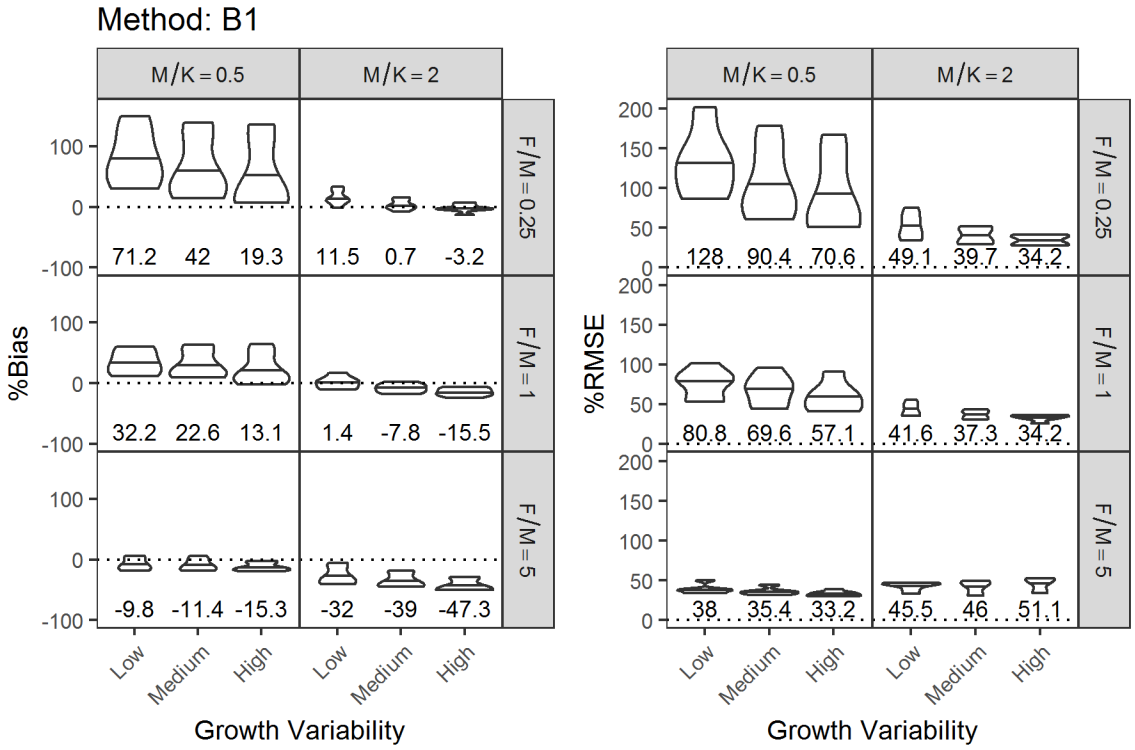

Figure S10. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability for method B1. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

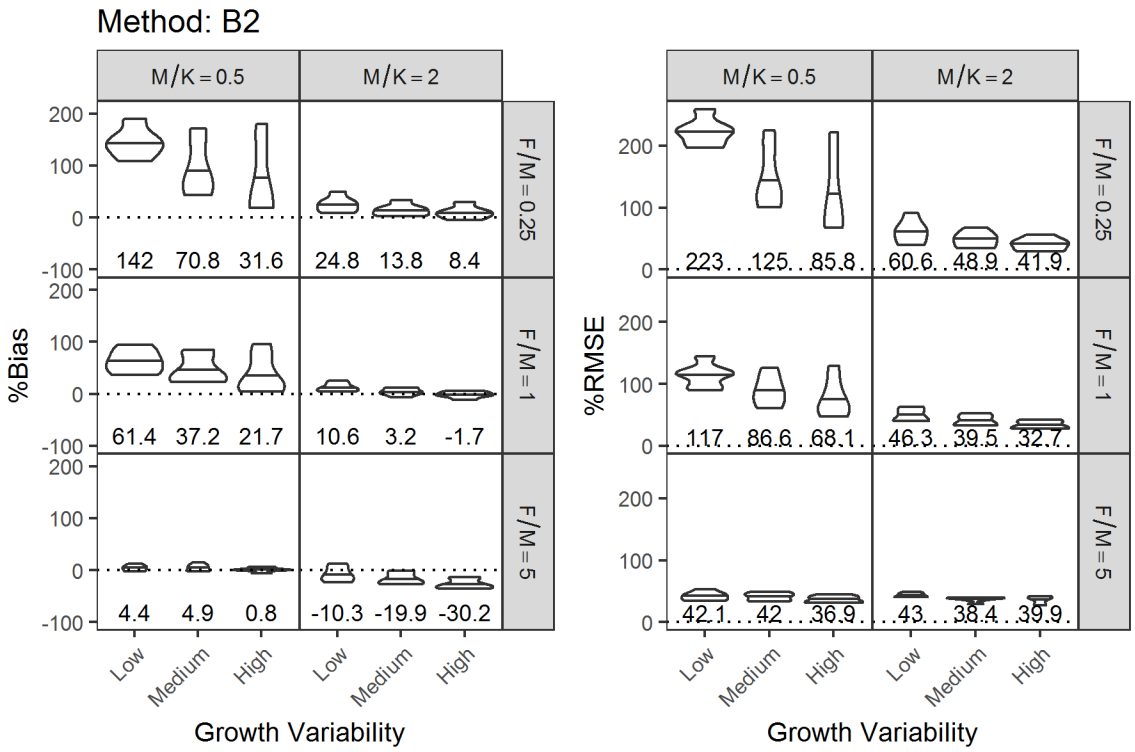

Figure S11. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability for method B2. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

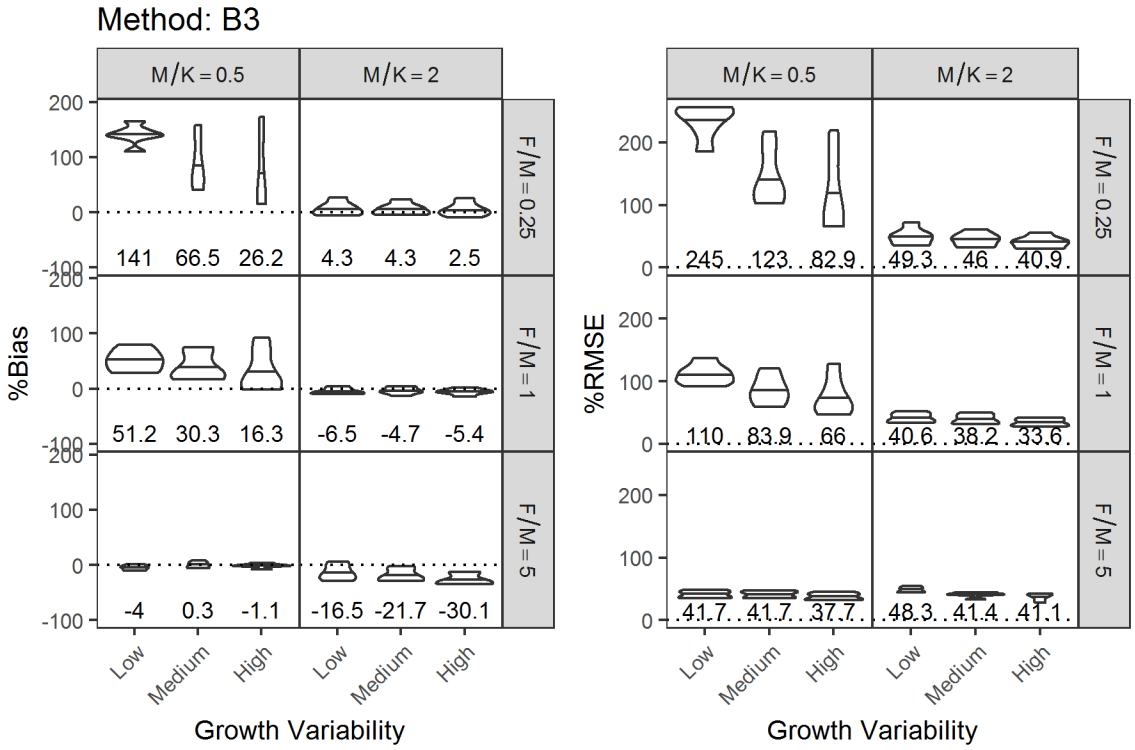

Figure S12. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability for method B3. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

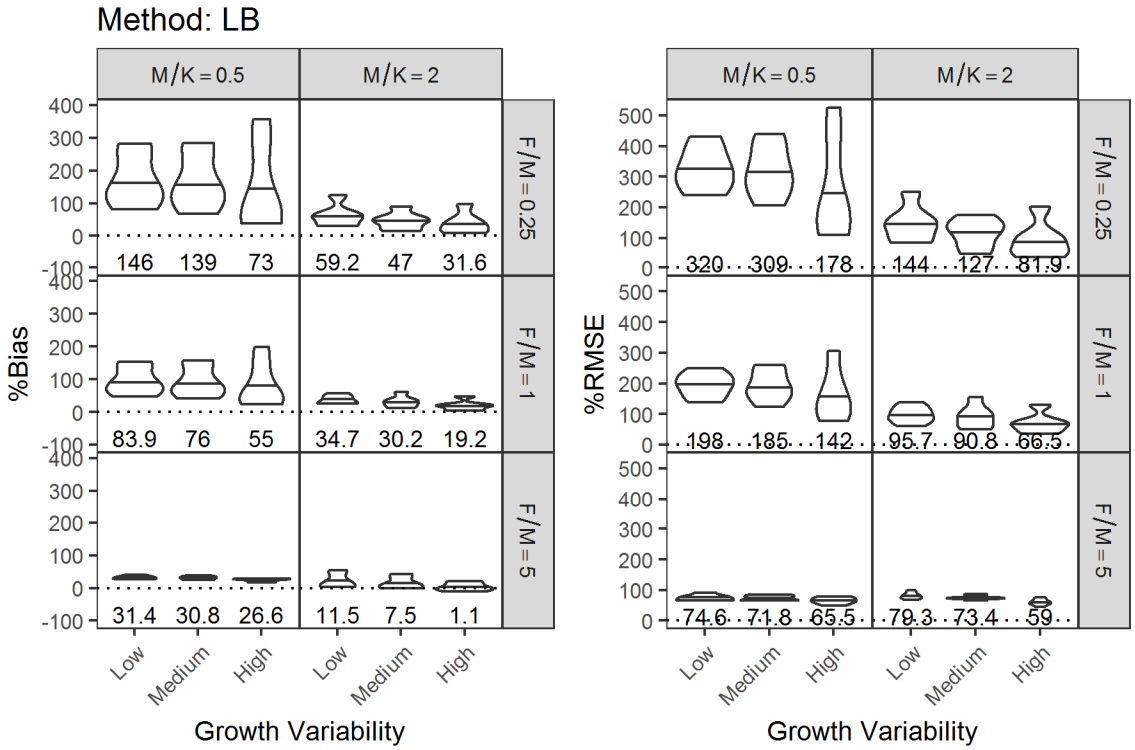

Figure S13. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and growth variability for method LB. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

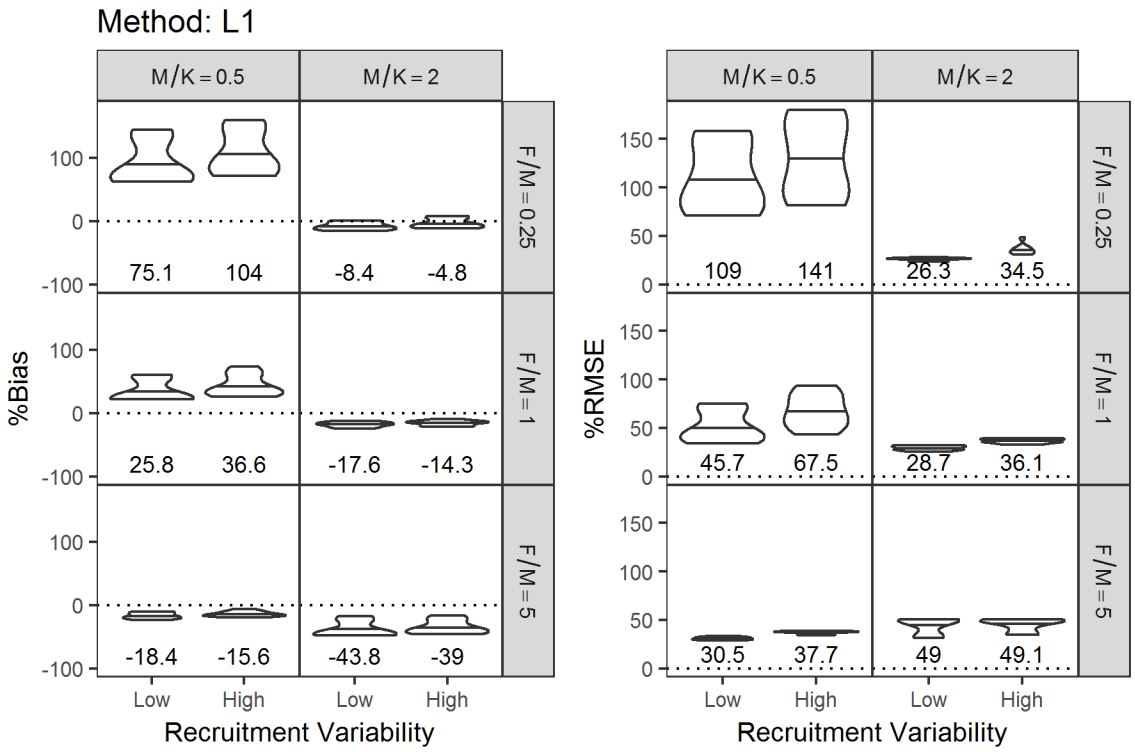

Figure S14. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability for method L1. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

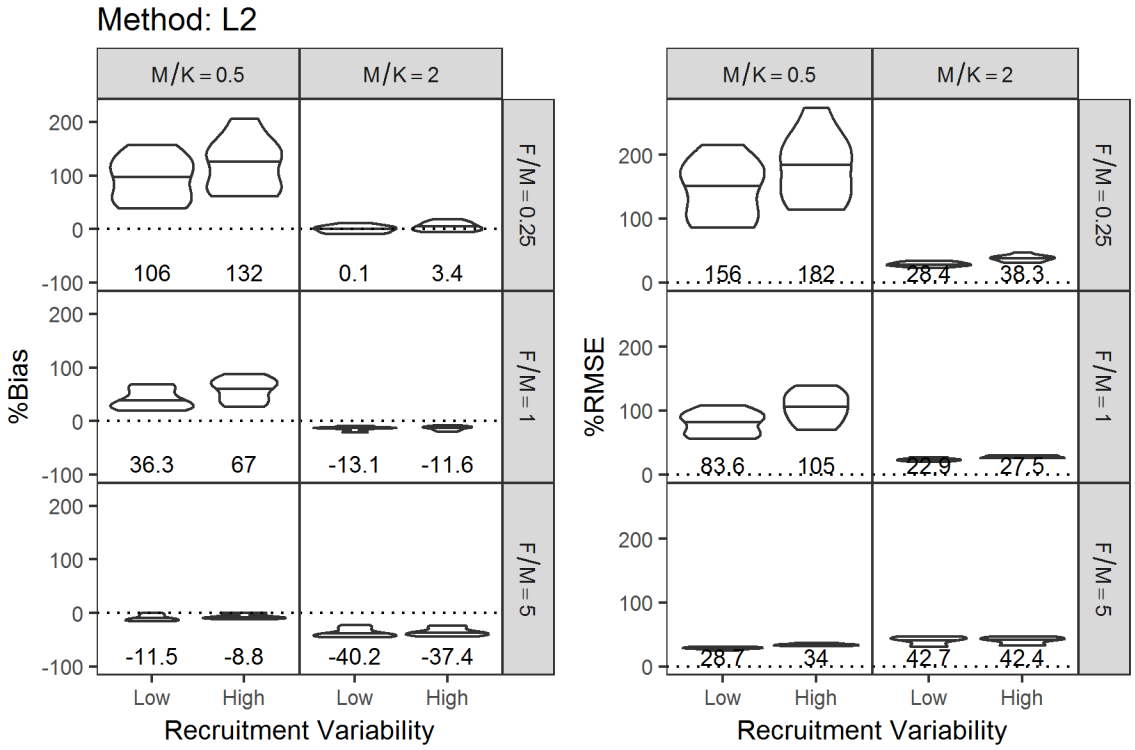

Figure S15. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability for method L2. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

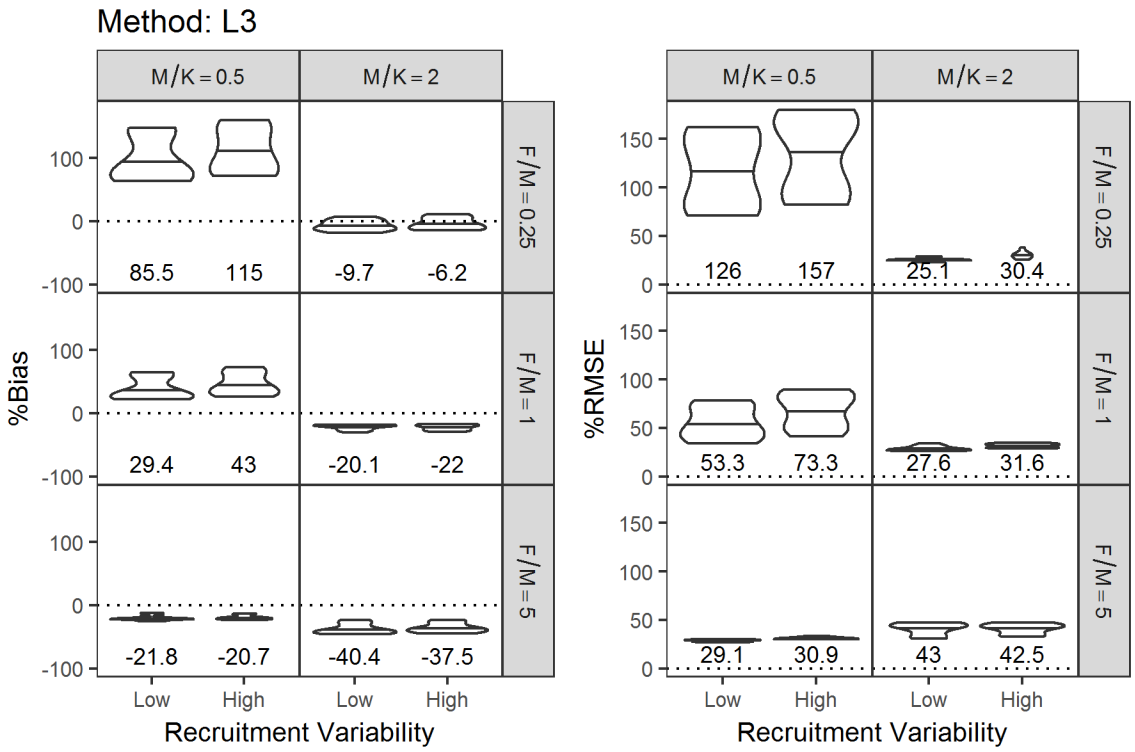

Figure S16. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability for method L3. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

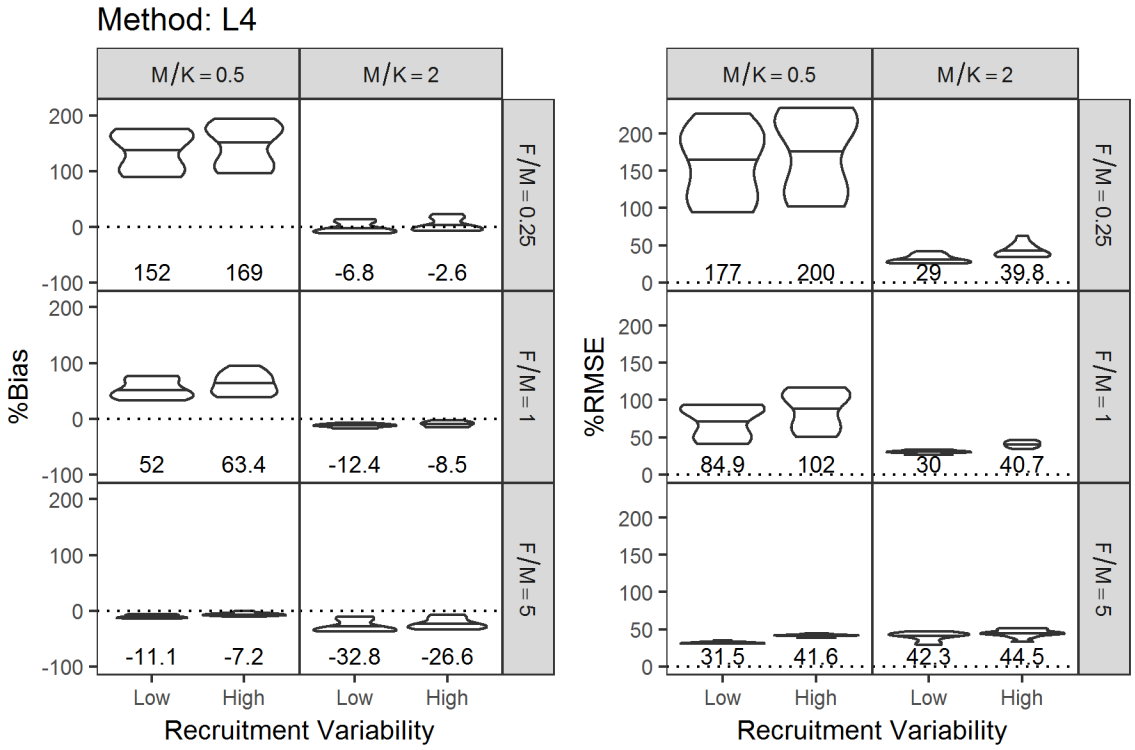

Figure S17. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability for method L4. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

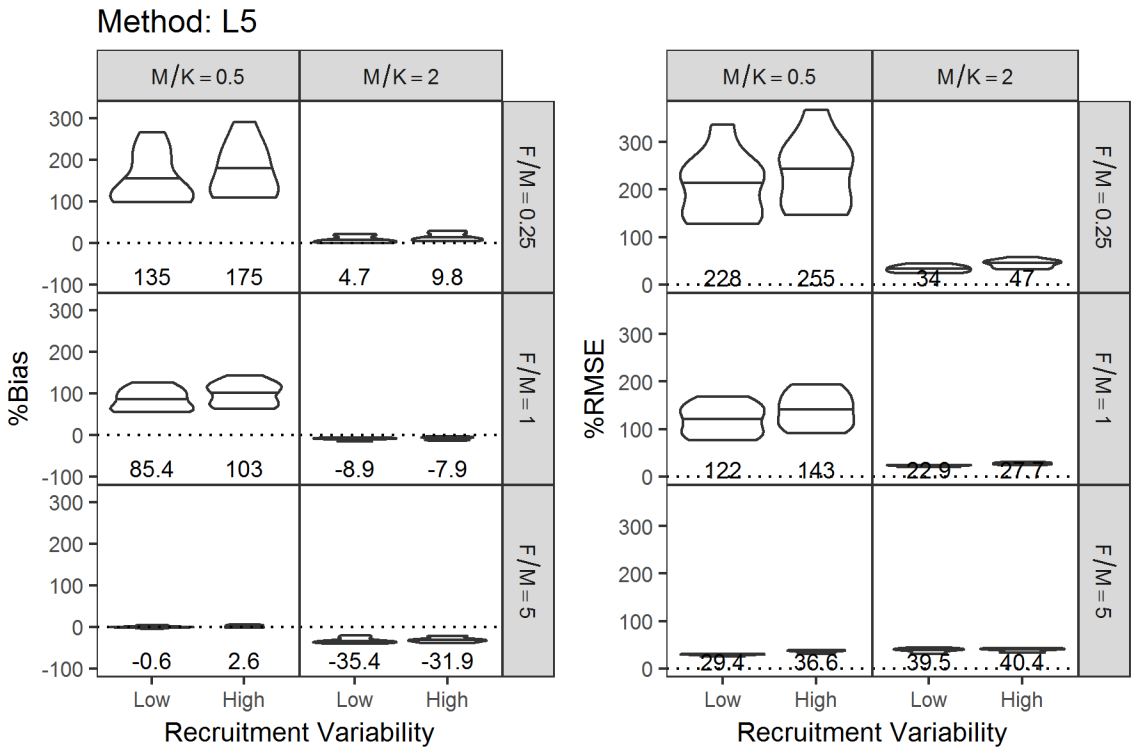

Figure S18. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability for method L5. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

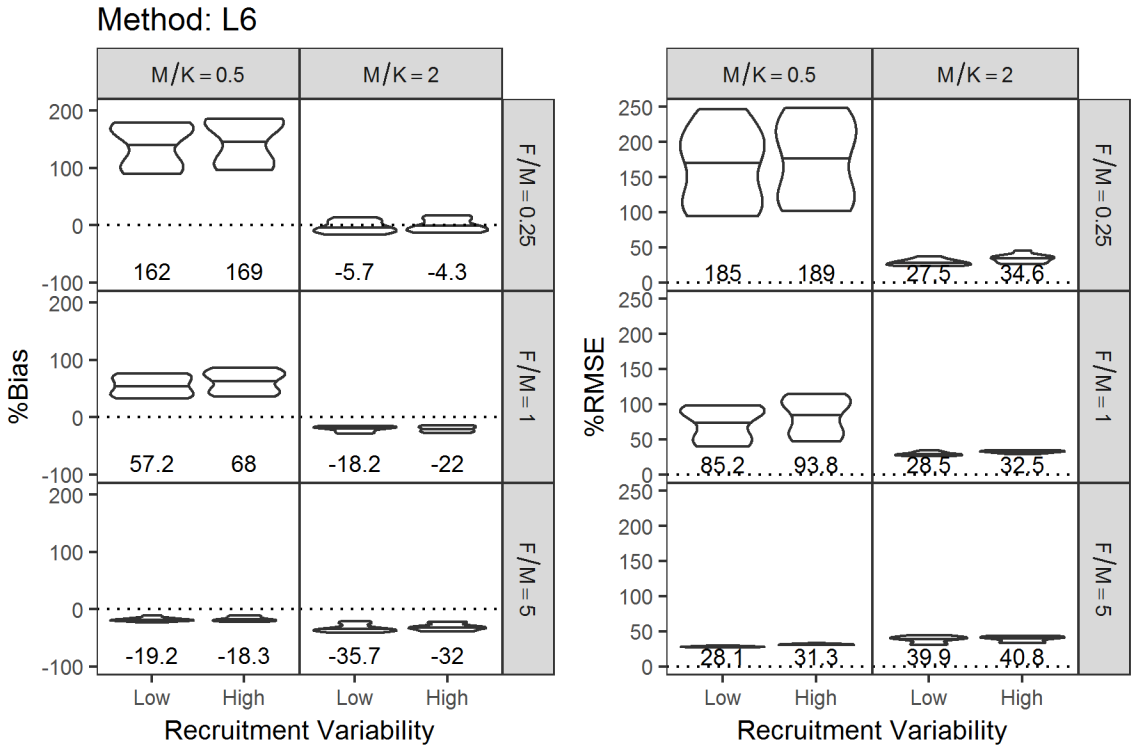

Figure S19. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability for method L6. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

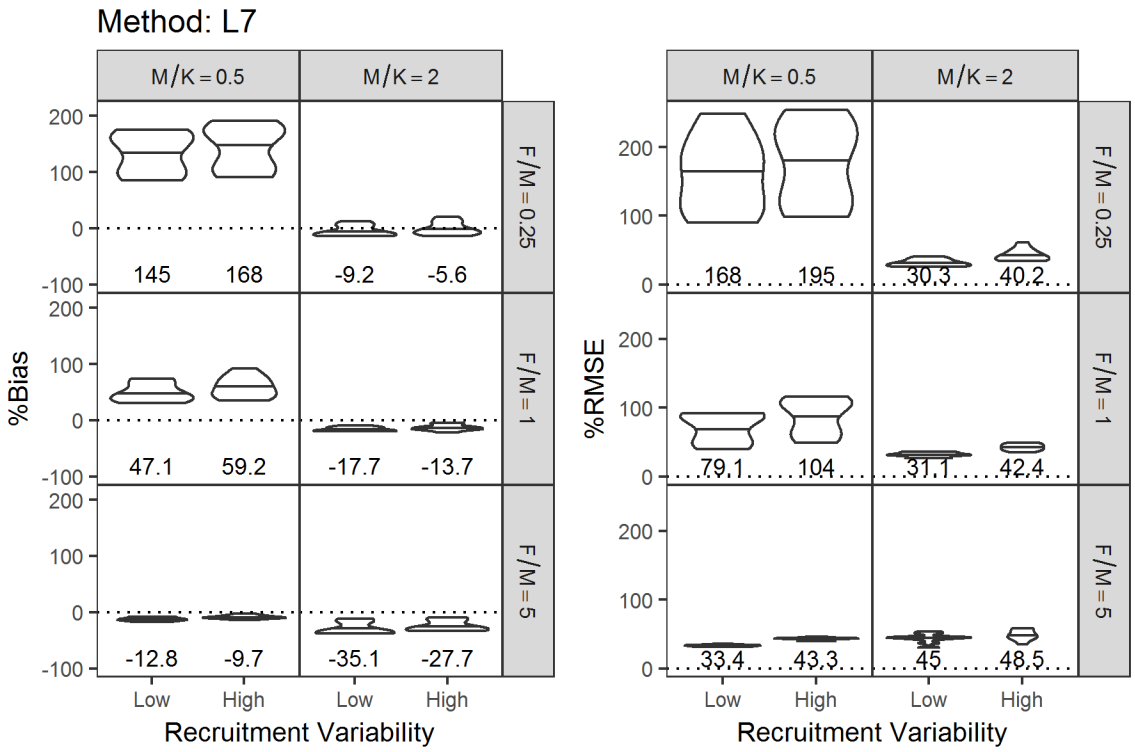

Figure S20. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability for method L7. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

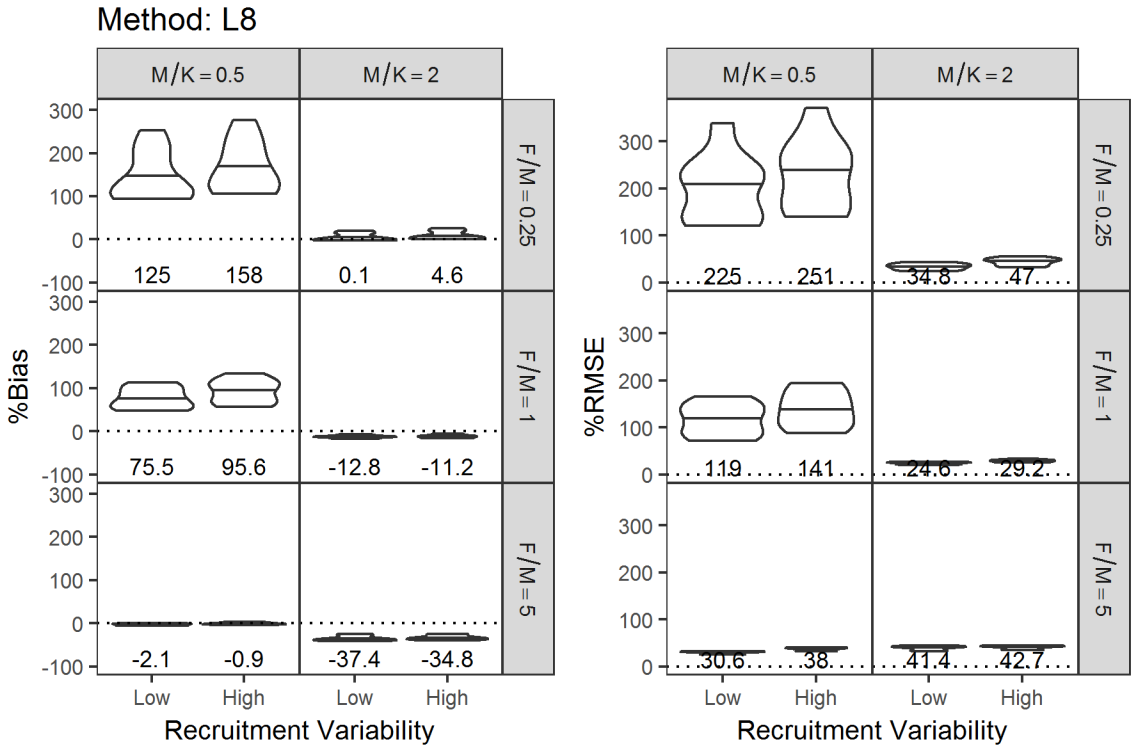

Figure S21. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability for method L8. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

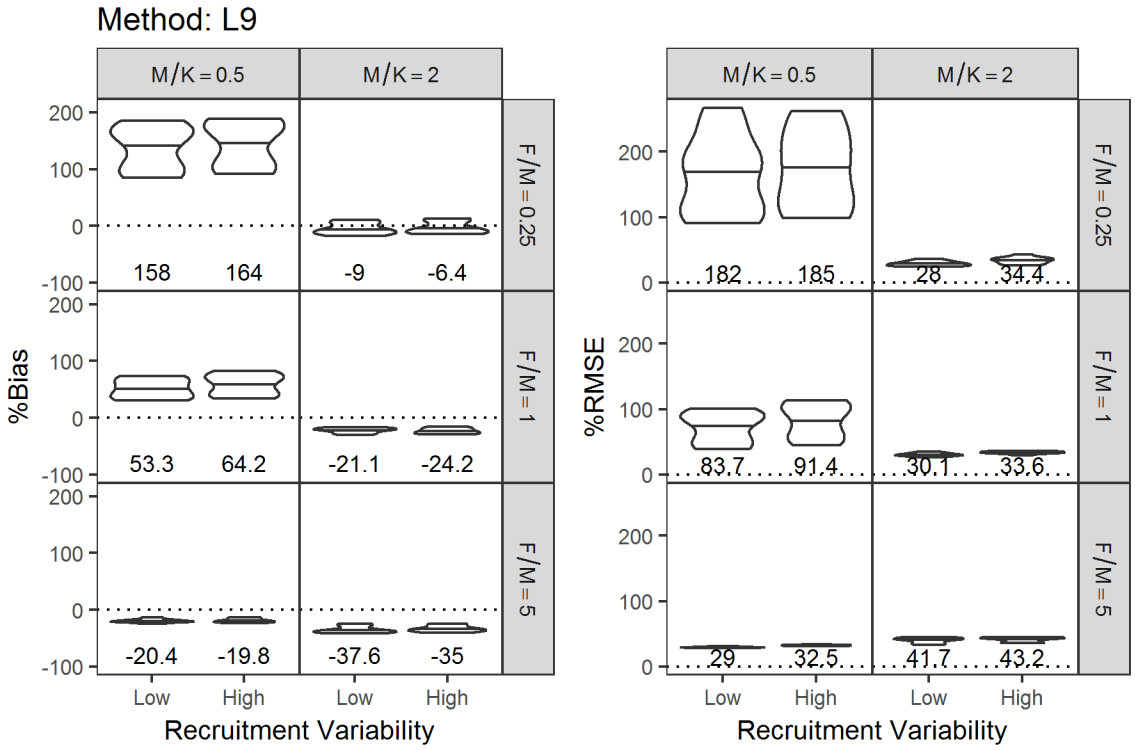

Figure S22. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability for method L9. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

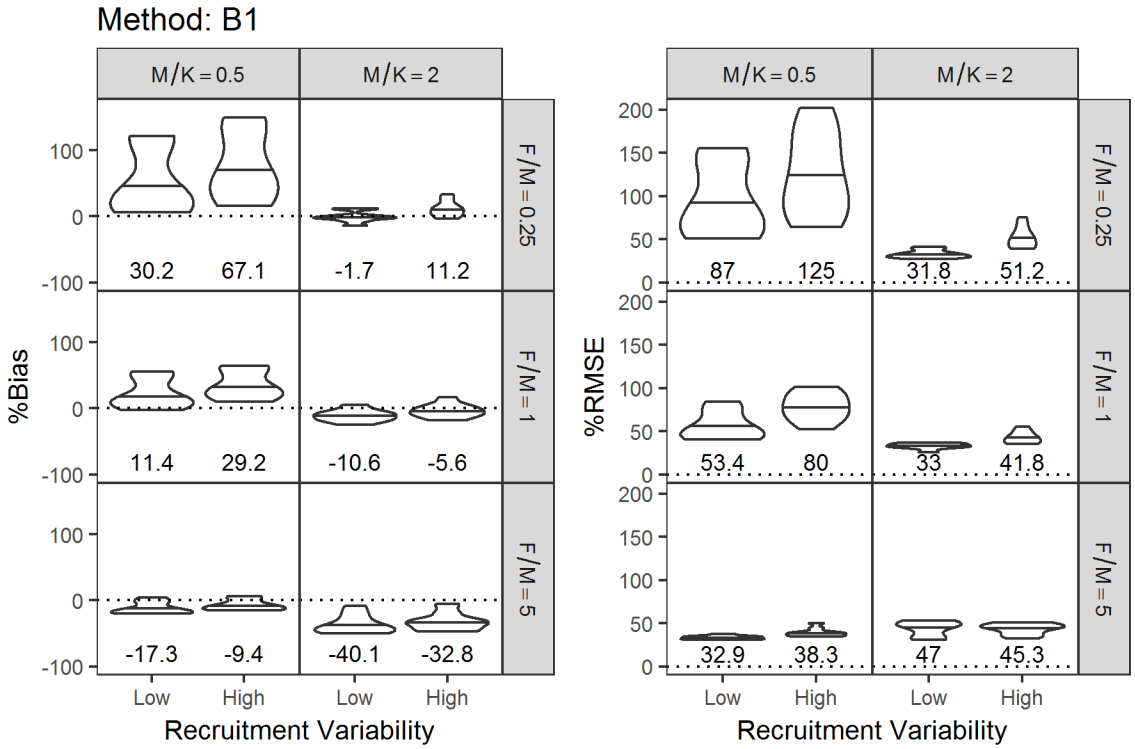

Figure S23. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability for method B1. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

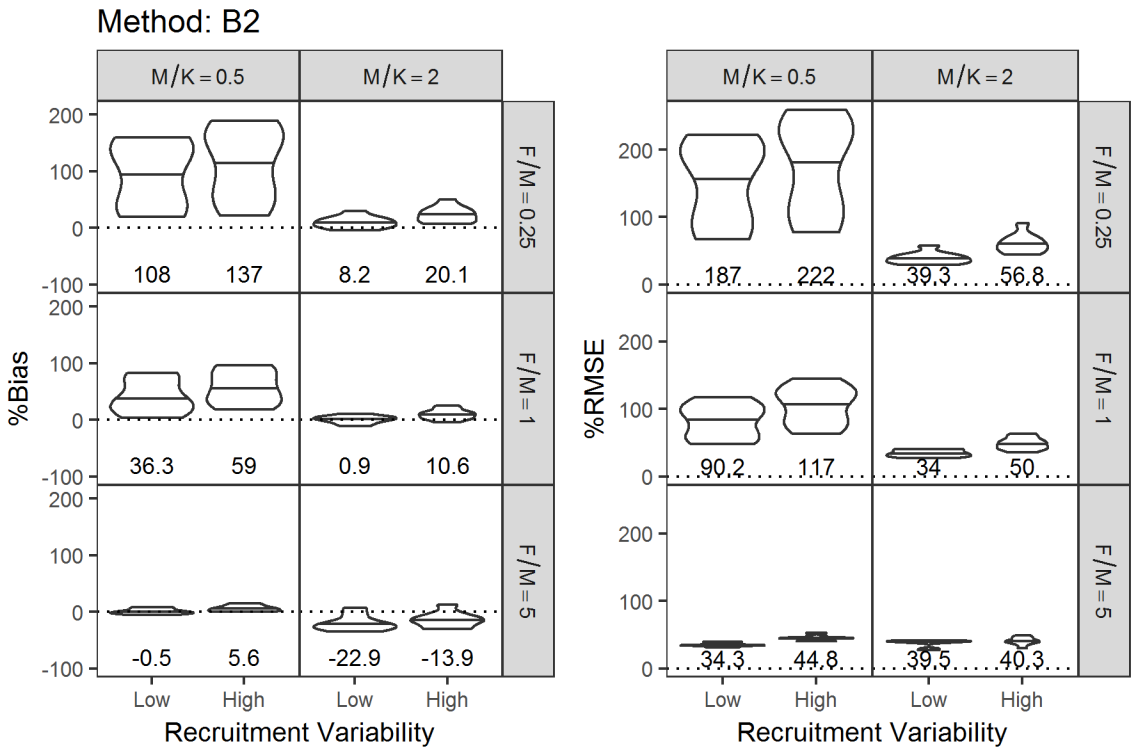

Figure S24. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability for method B2. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

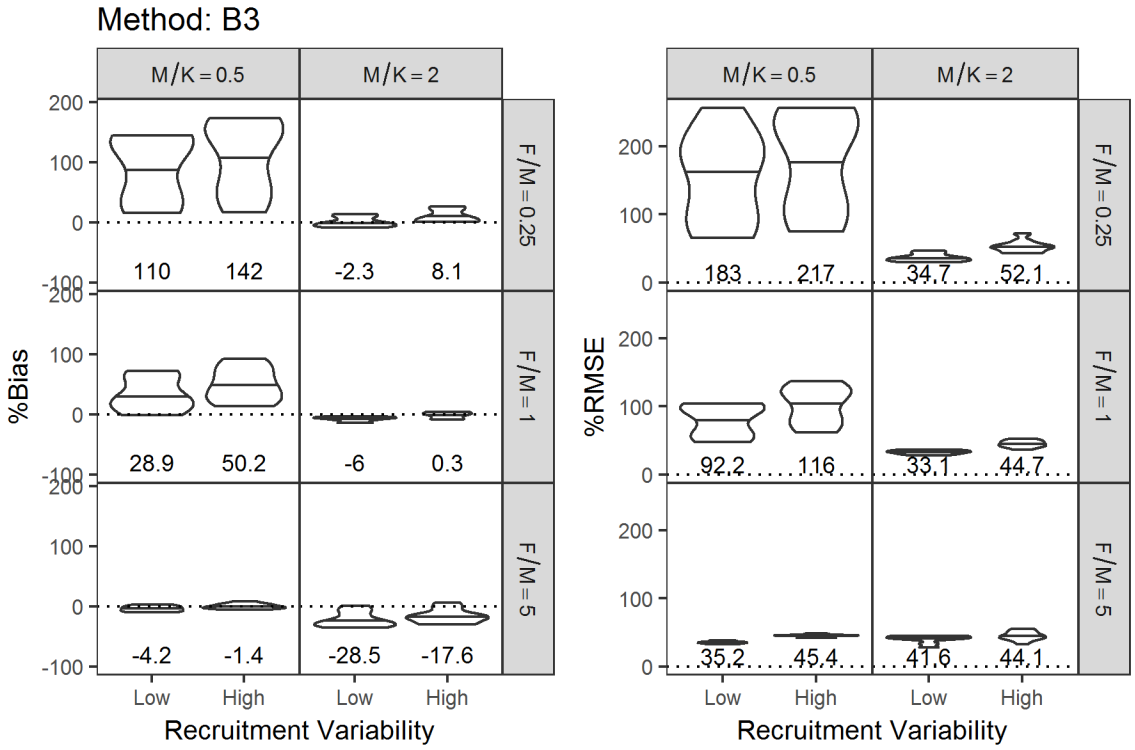

Figure S25. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability for method B3. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

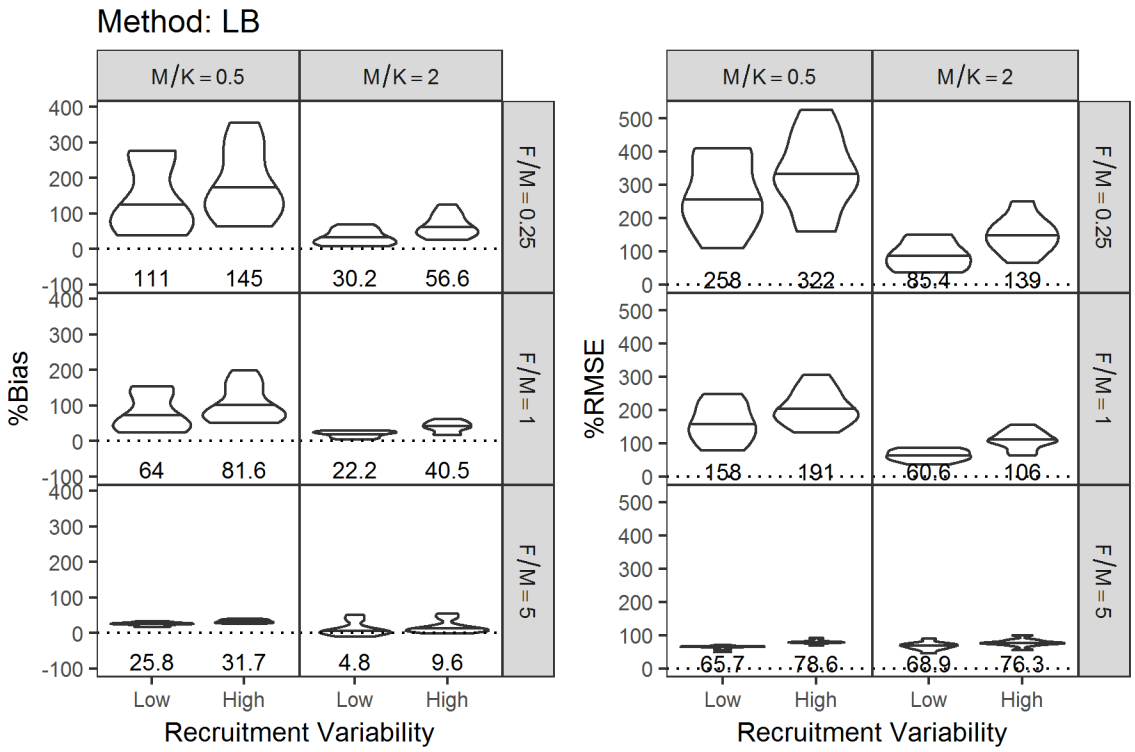

Figure S26. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and recruitment variability for method LB. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

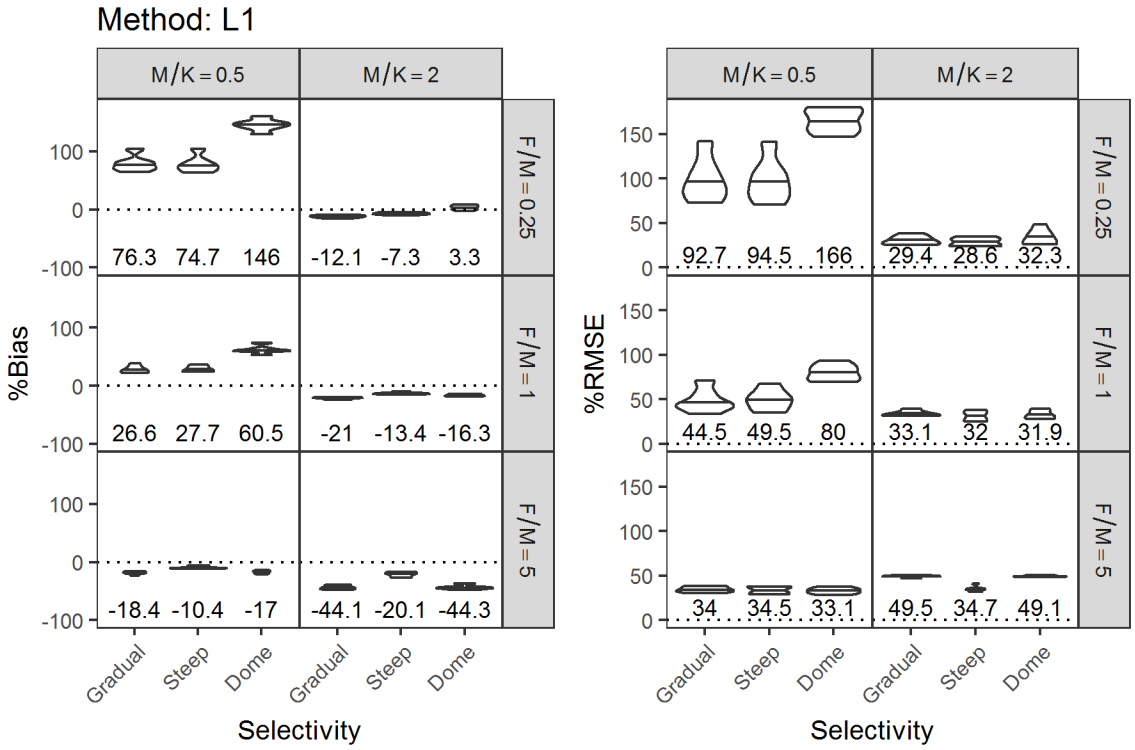

Figure S27. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity for method L1. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

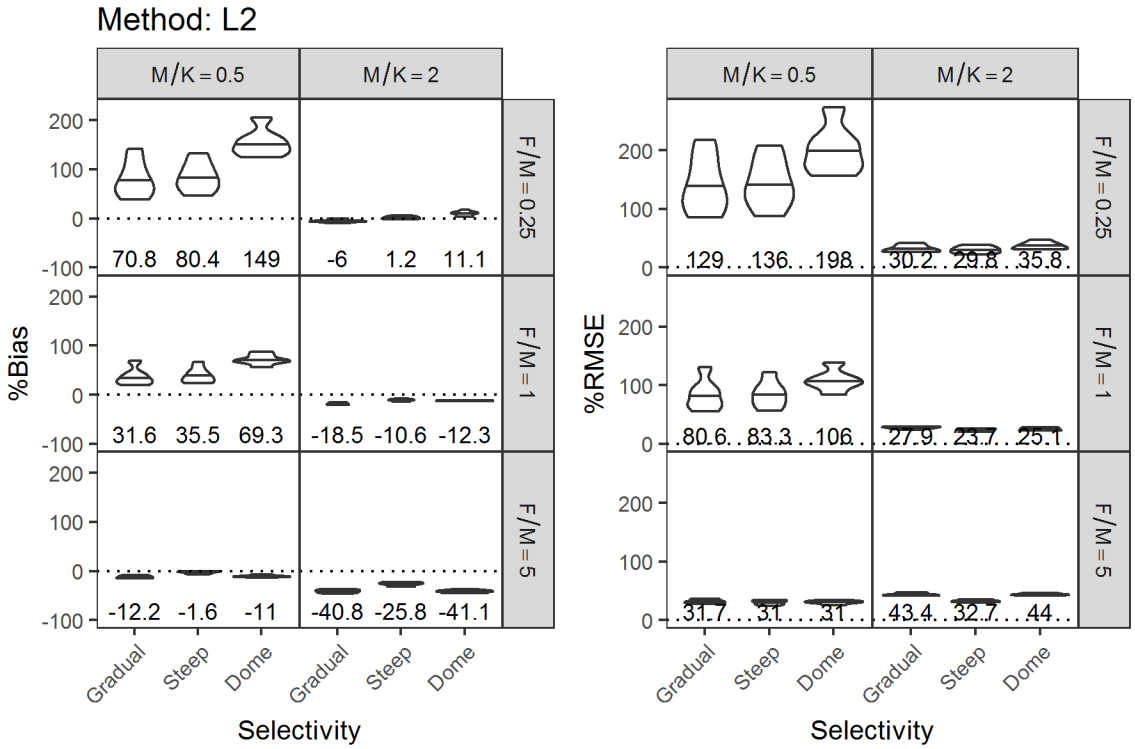

Figure S28. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity for method L2. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

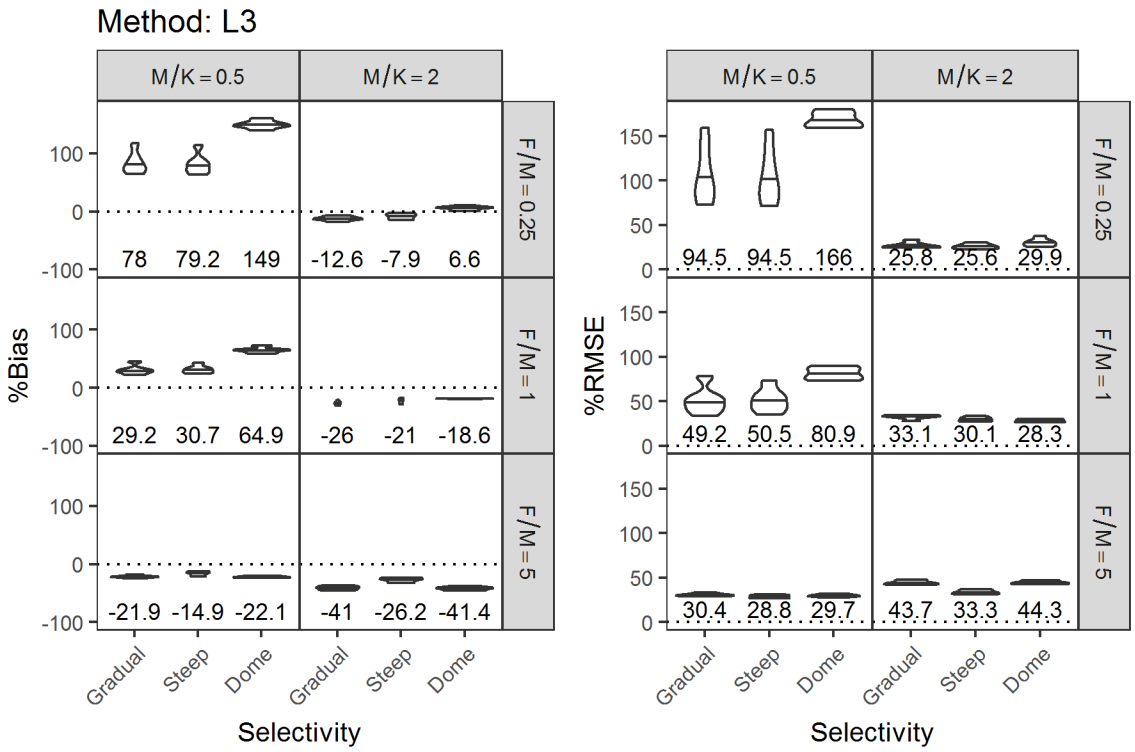

Figure S29. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity for method L3. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

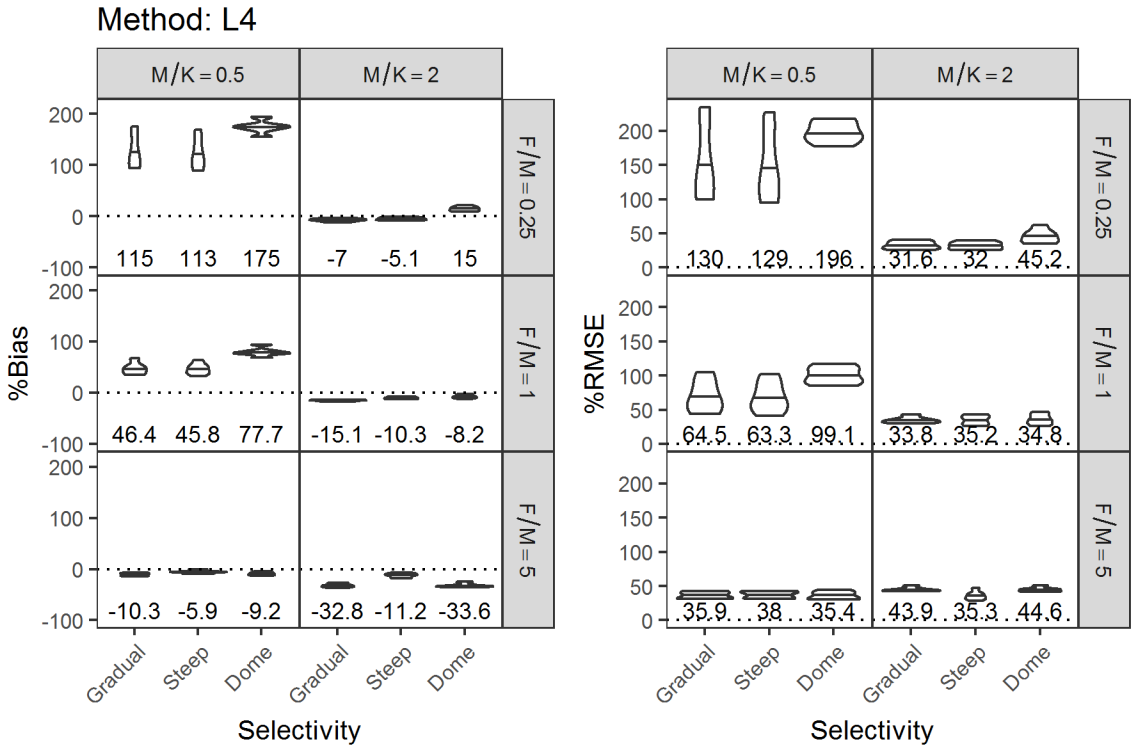

Figure S30. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity for method L4. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

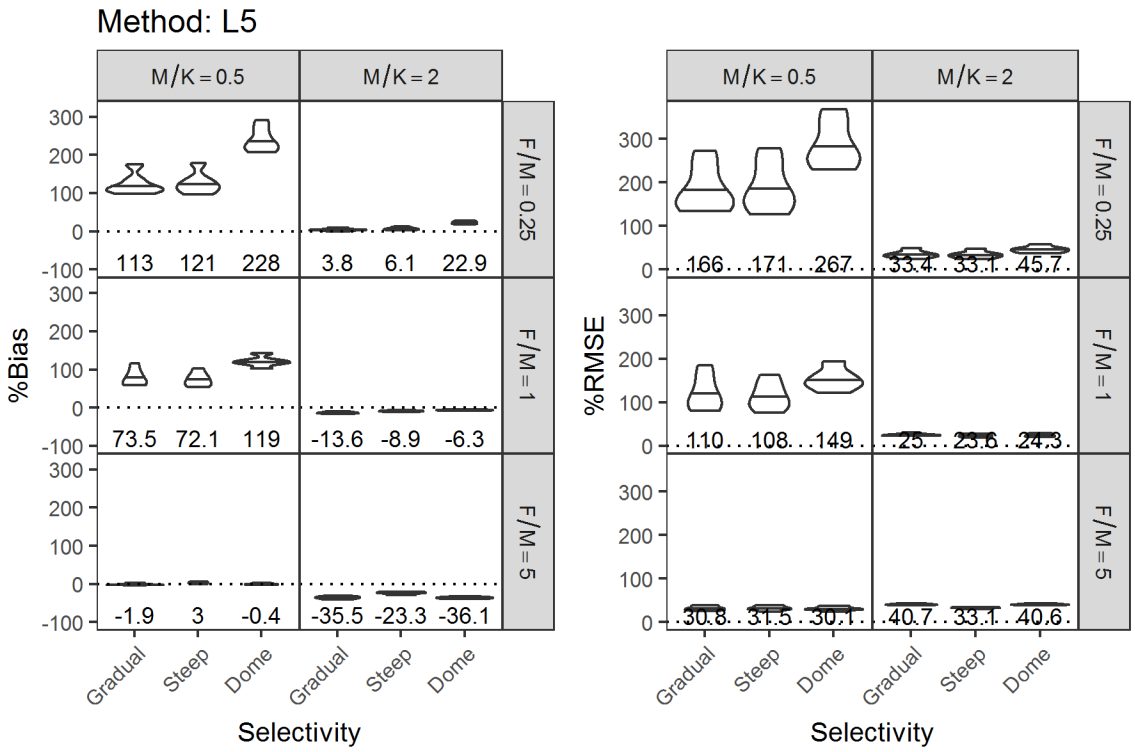

Figure S31. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity for method L5. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

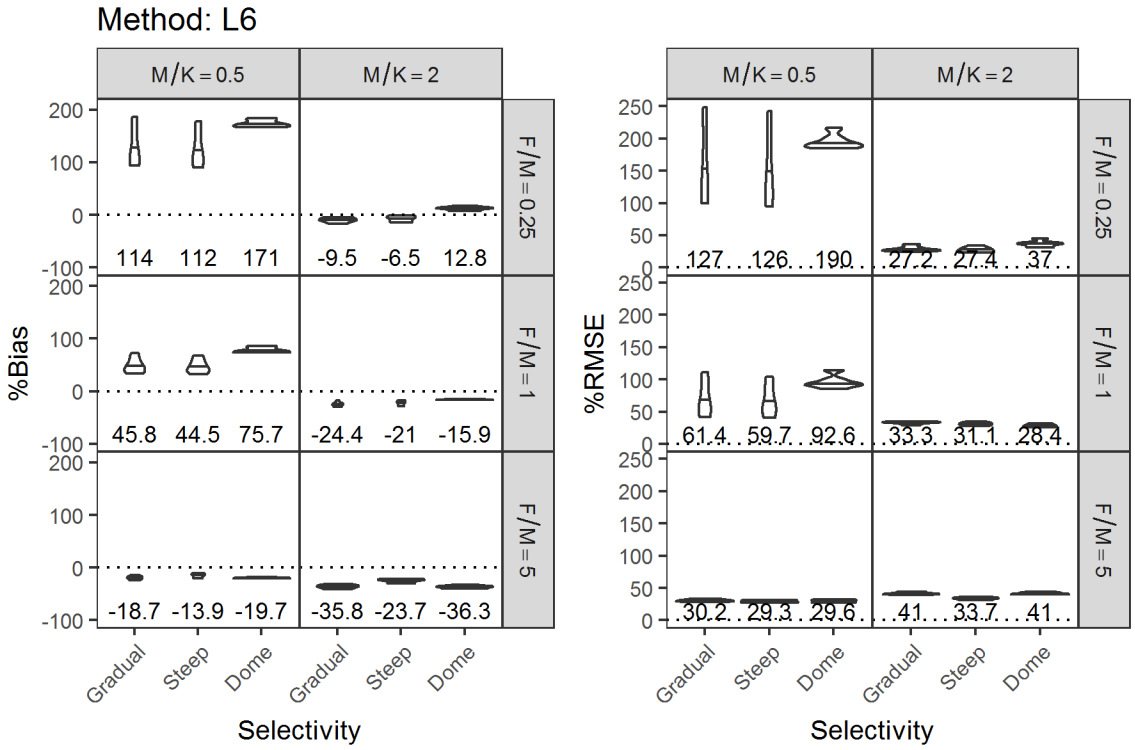

Figure S32. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity for method L6. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

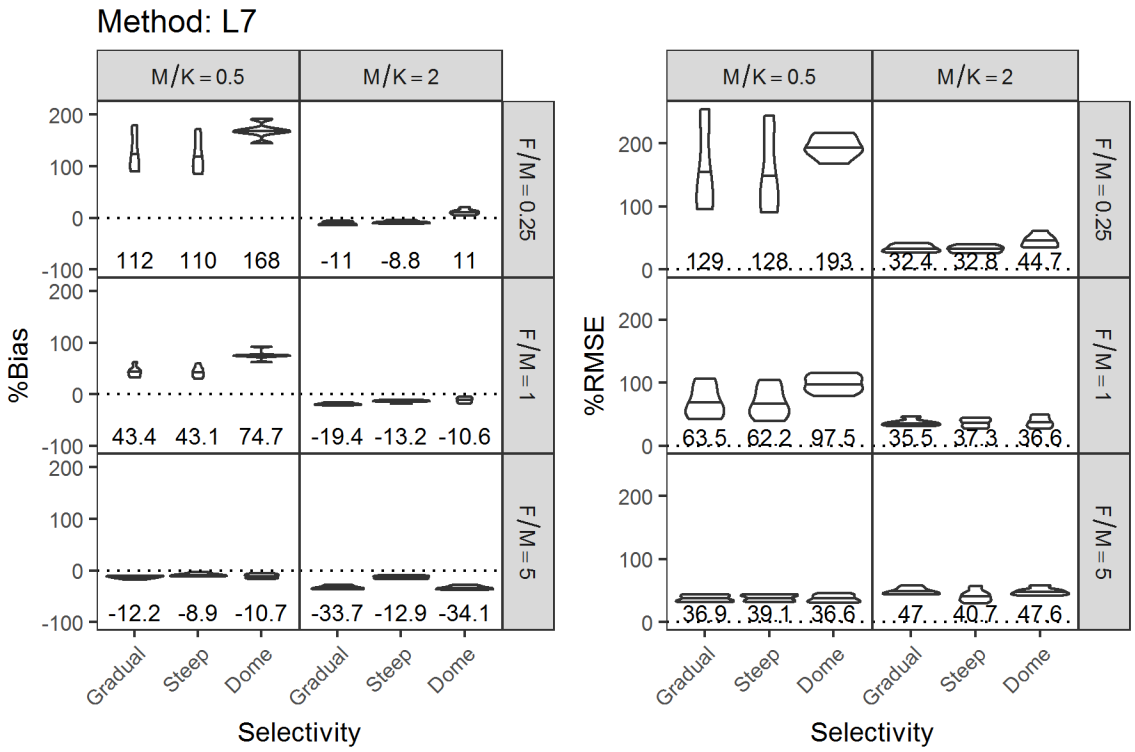

Figure S33. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity for method L7. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

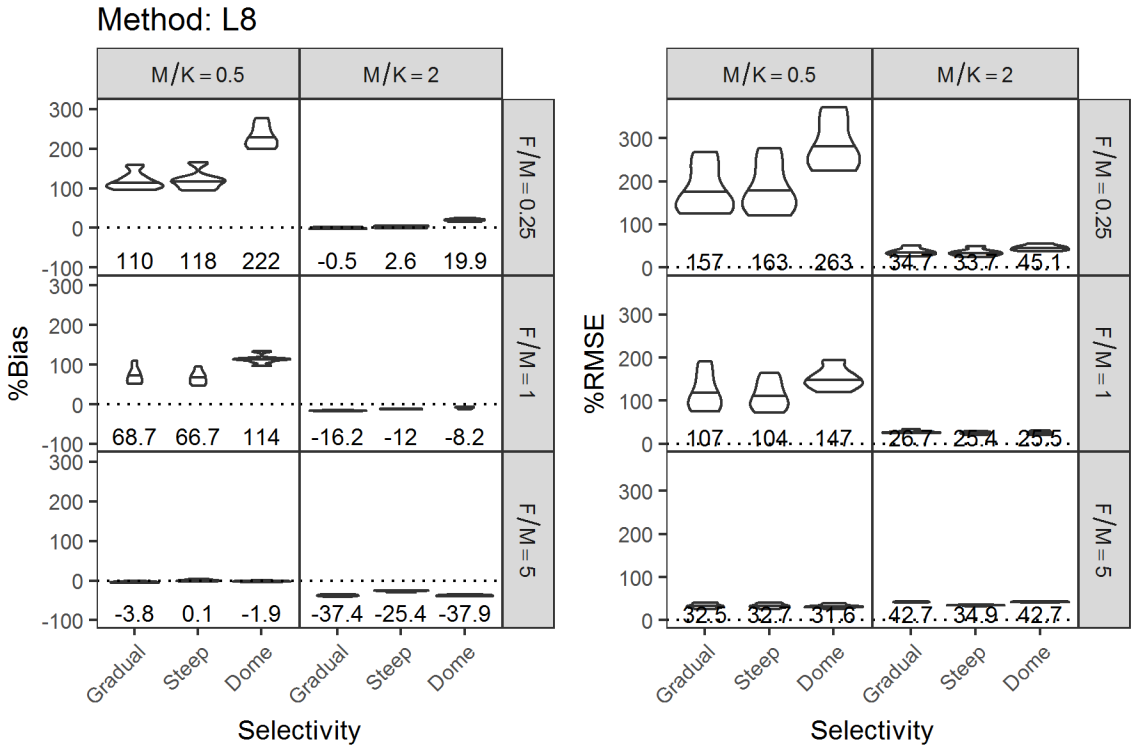

Figure S34. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity for method L8. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

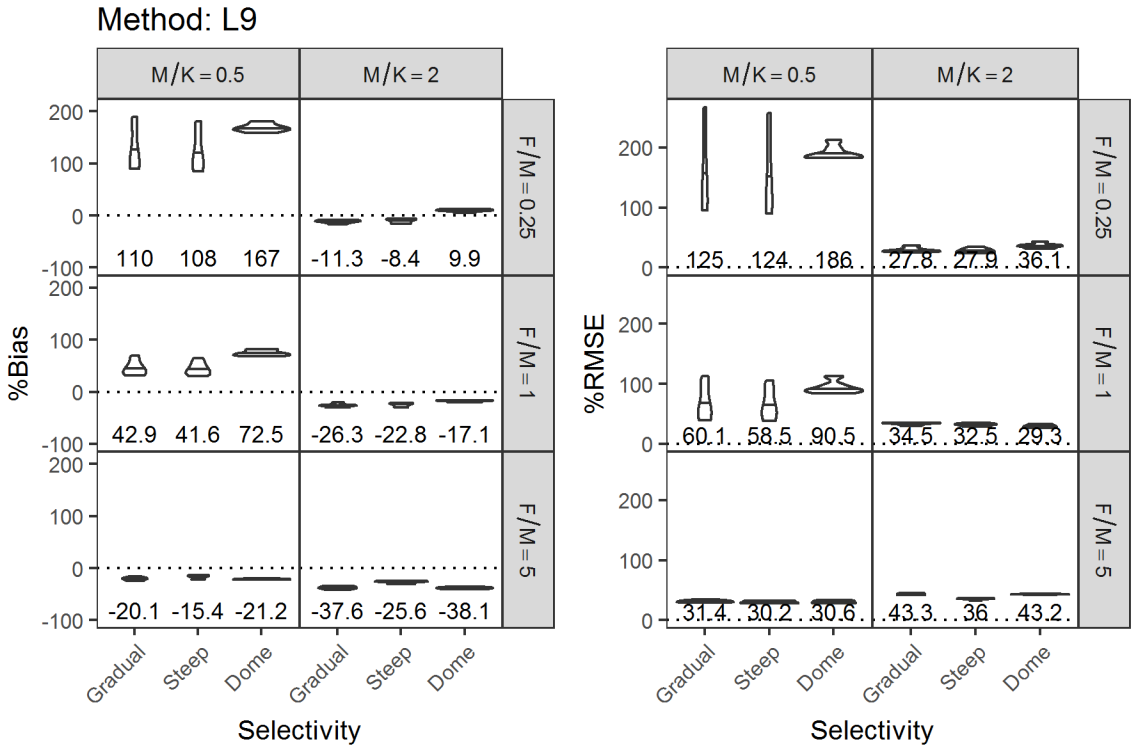

Figure S35. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity for method L9. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

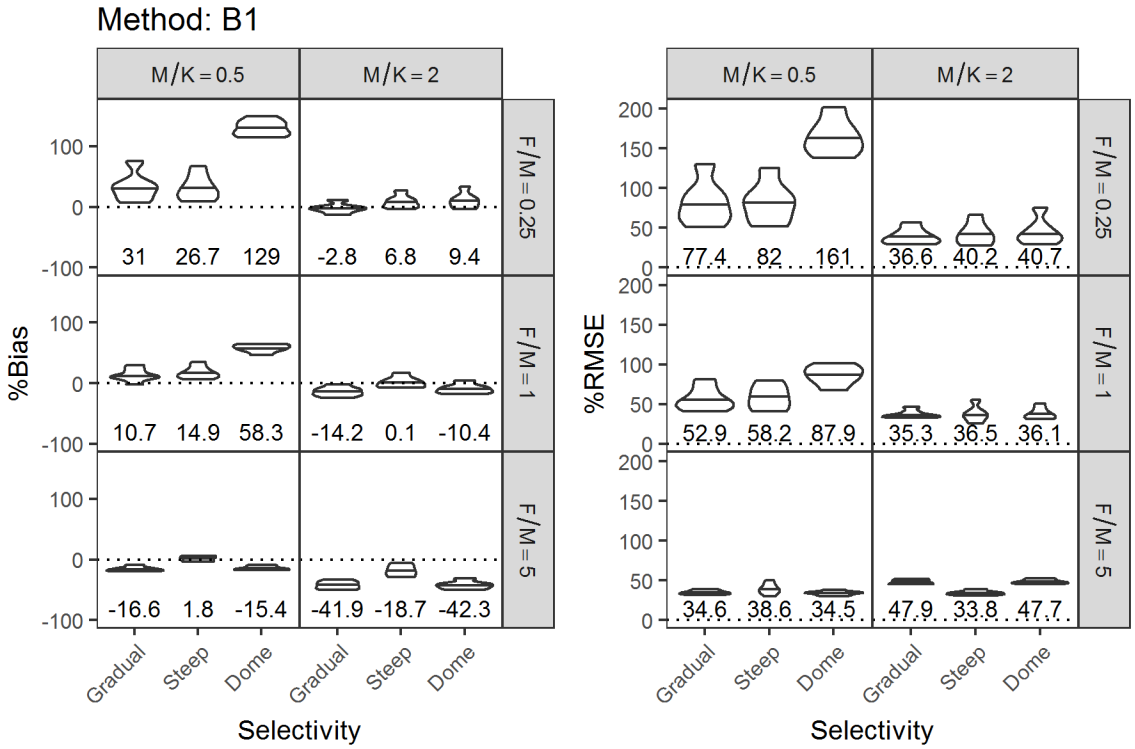

Figure S36. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity for method B1. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

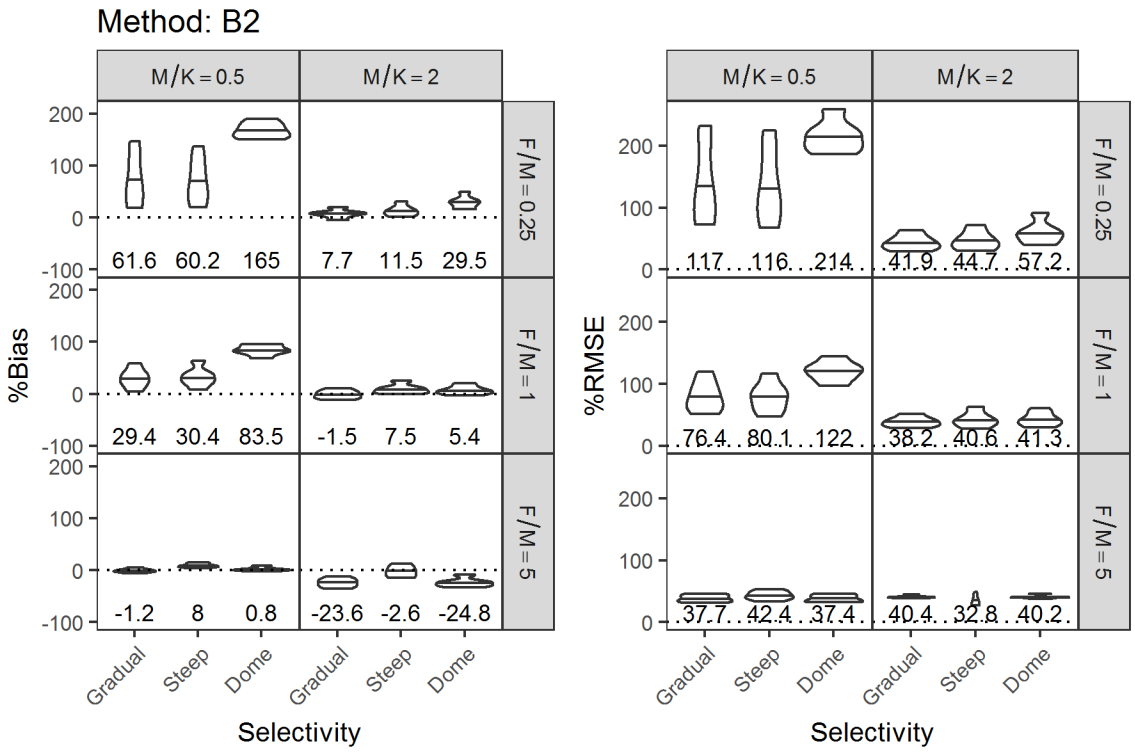

Figure S37. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity for method B2. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

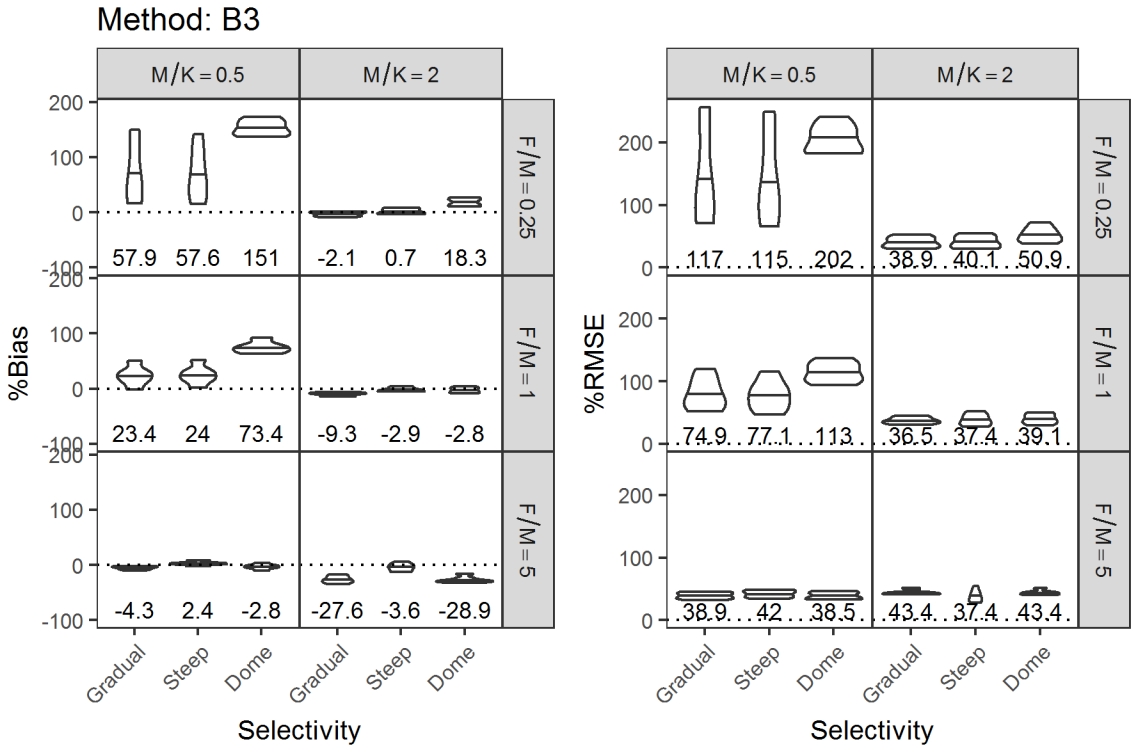

Figure S38. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity for method B3. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

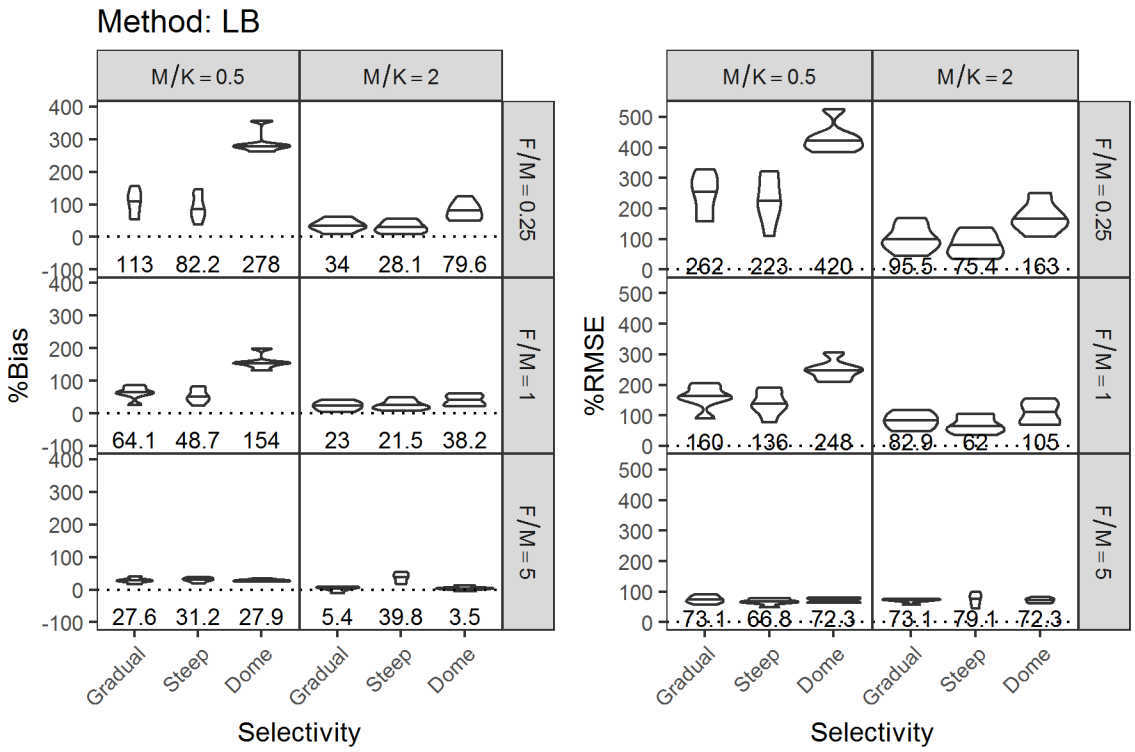

Figure S39. %*Bias* (left grid) and %*RMSE* (right grid) stratified by *M/K*, *F/M*, and selectivity for method LB. Numbers and horizontal lines in the violin plot indicate median %*Bias* and %*RMSE* and shape of violin plot show distribution of values.

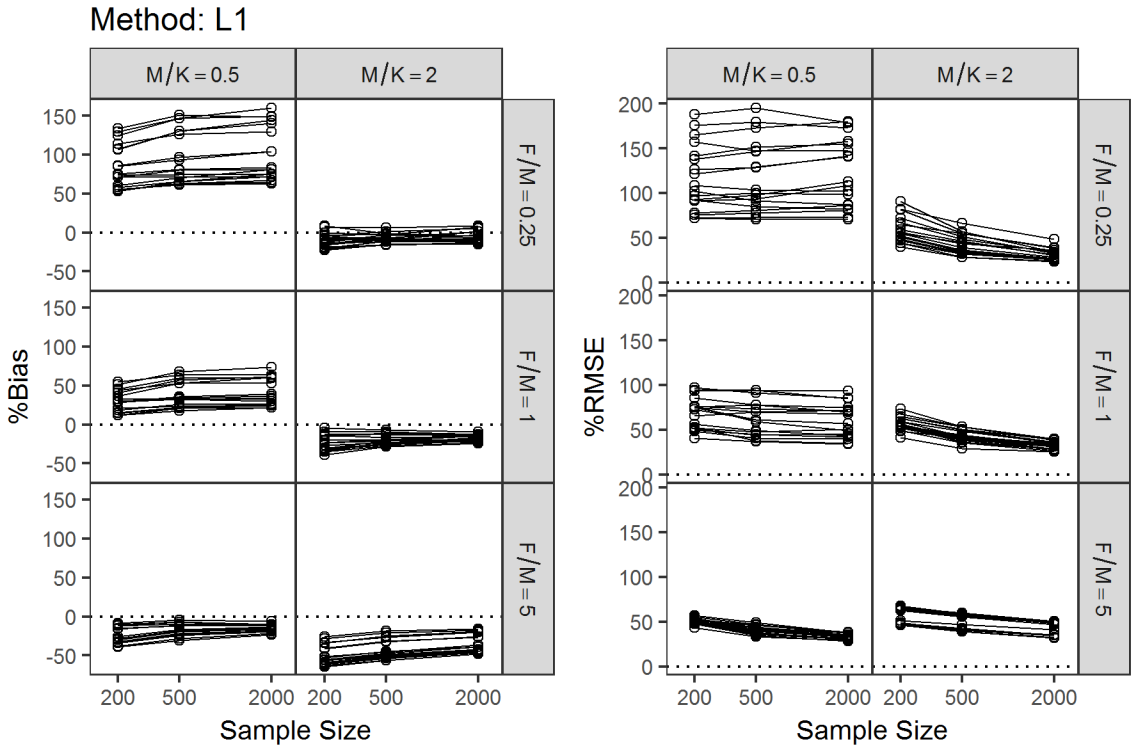

Figure S40. The effect of sample size on %*Bias* (left grid) and %*RMSE* (right grid) for method L1. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

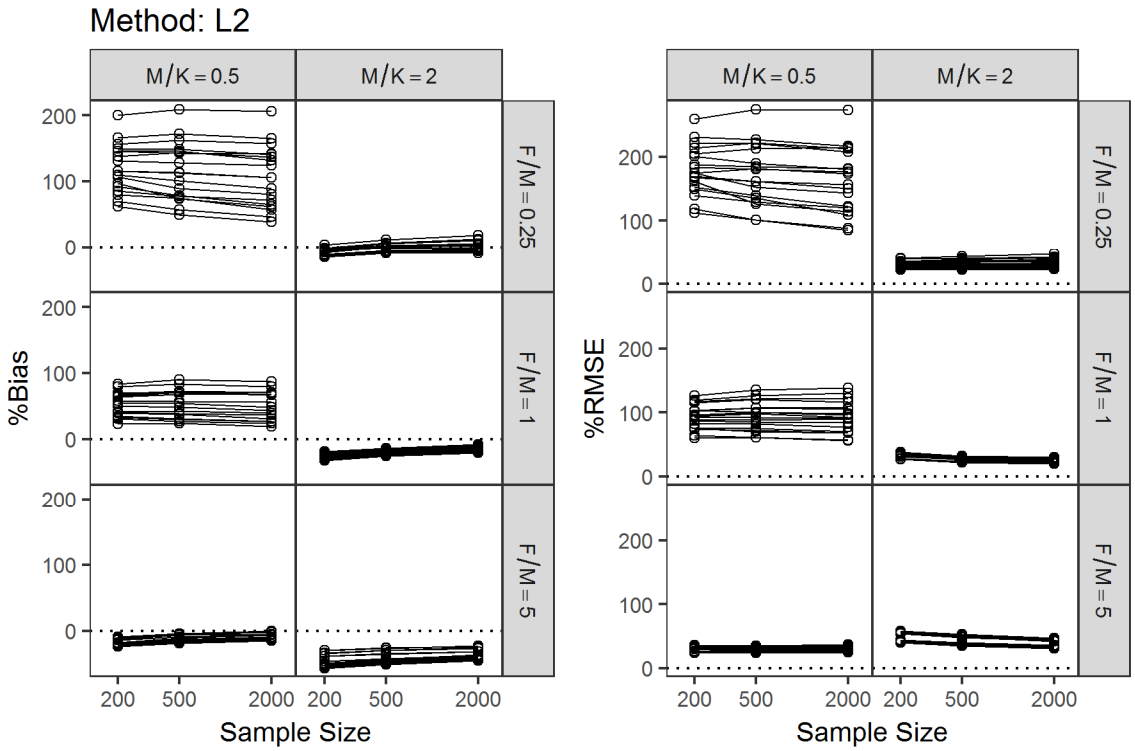

Figure S41. The effect of sample size on %*Bias* (left grid) and %*RMSE* (right grid) for method L2. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

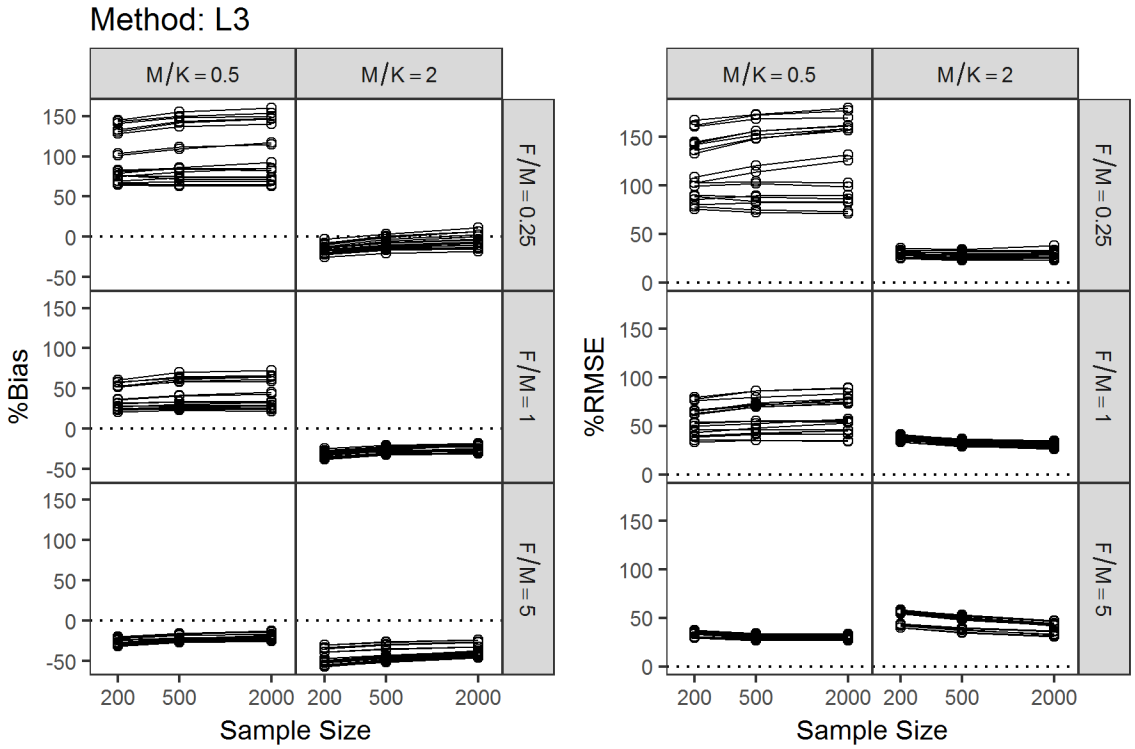

Figure S42. The effect of sample size on %*Bias* (left grid) and %*RMSE* (right grid) for method L3. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

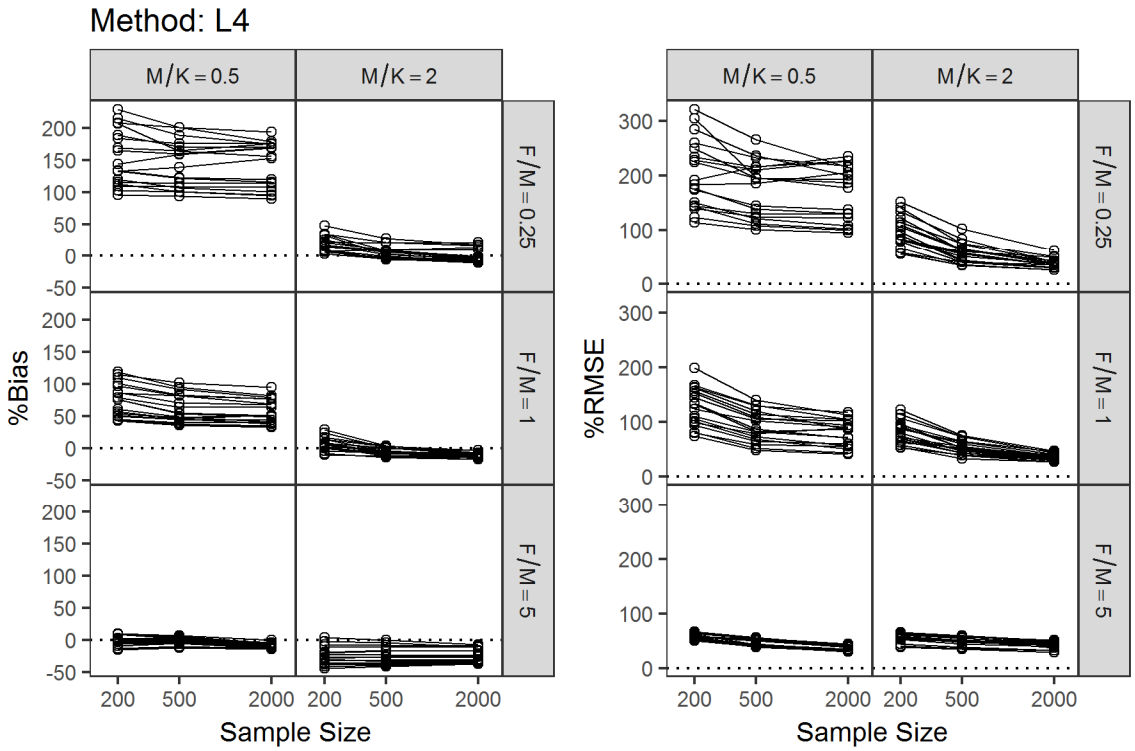

Figure S43. The effect of sample size on %*Bias* (left grid) and %*RMSE* (right grid) for method L4. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

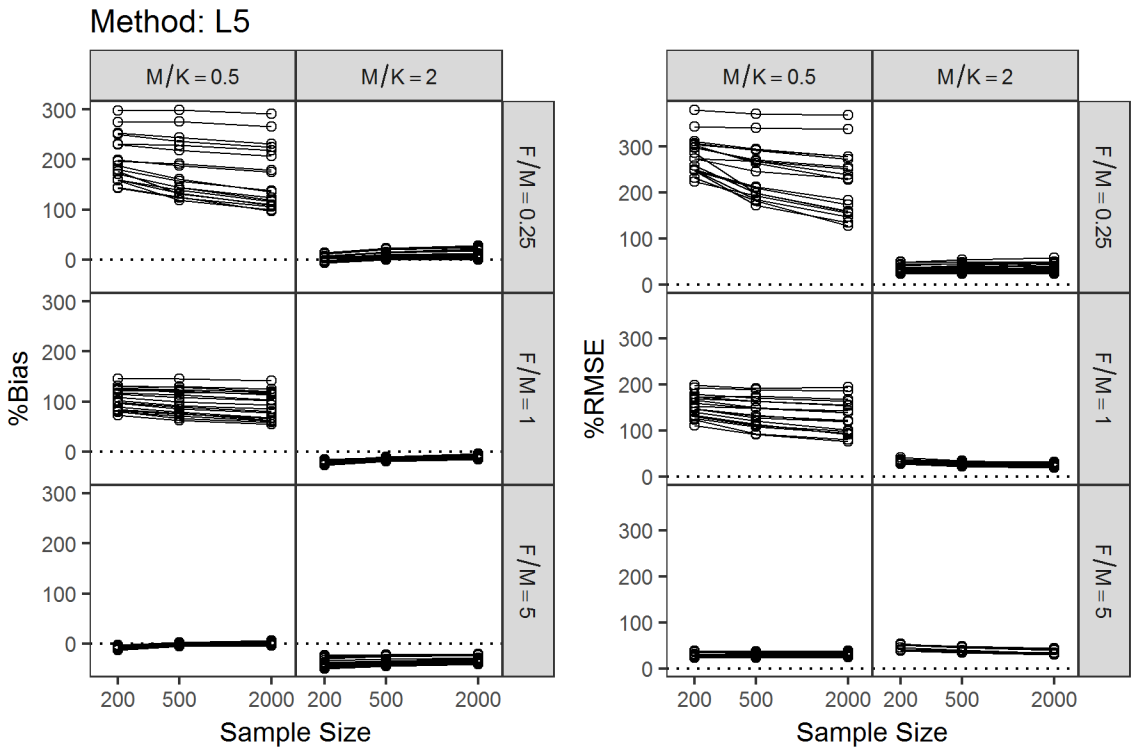

Figure S44. The effect of sample size on %*Bias* (left grid) and %*RMSE* (right grid) for method L5. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

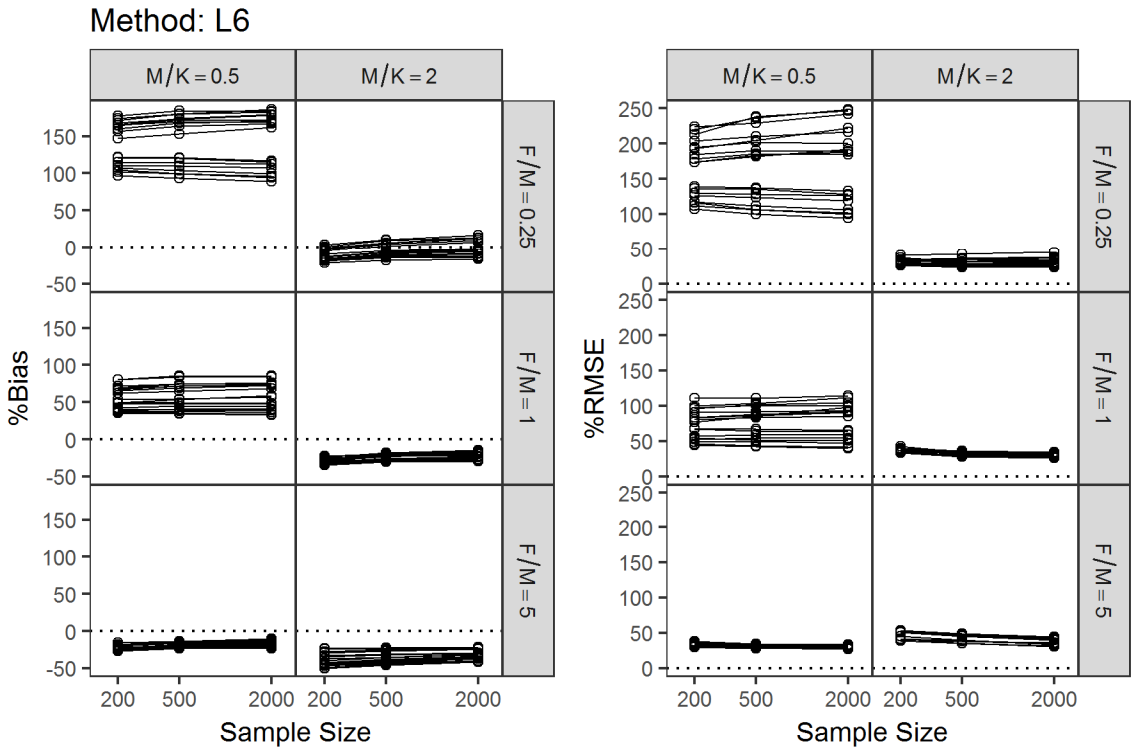

Figure S45. The effect of sample size on %*Bias* (left grid) and %*RMSE* (right grid) for method L6. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

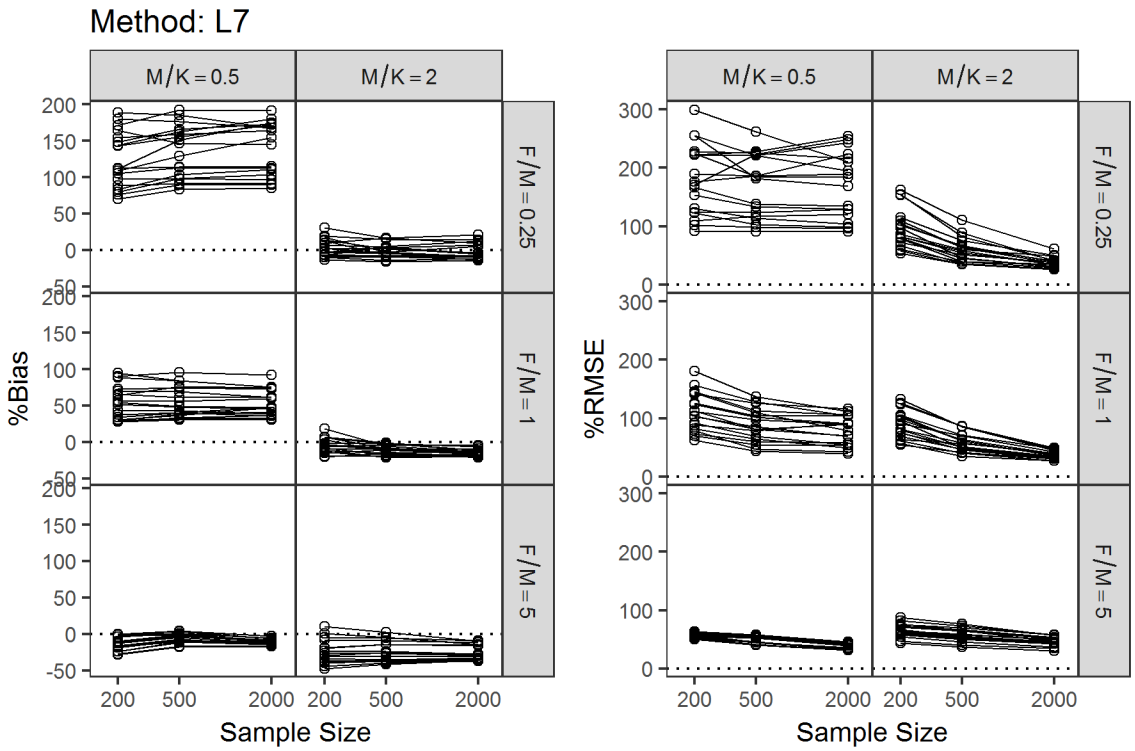

Figure S46. The effect of sample size on %*Bias* (left grid) and %*RMSE* (right grid) for method L7. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

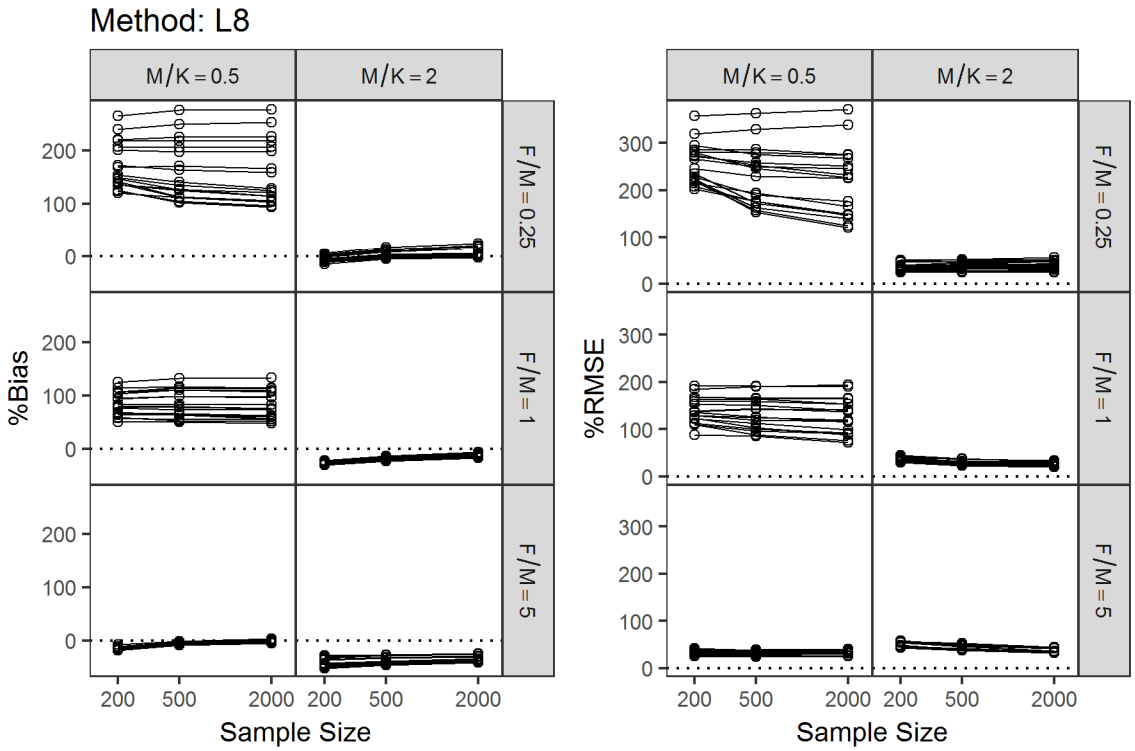

Figure S47. The effect of sample size on %*Bias* (left grid) and %*RMSE* (right grid) for method L8. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

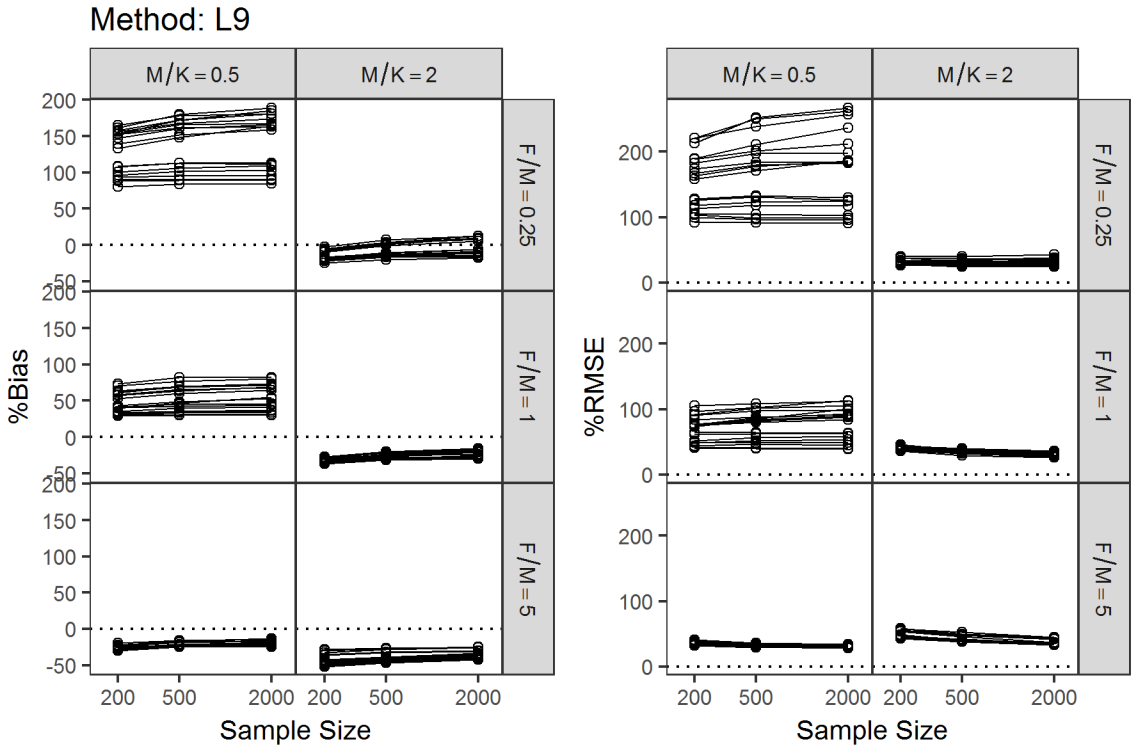

Figure S48. The effect of sample size on %*Bias* (left grid) and %*RMSE* (right grid) for method L9. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

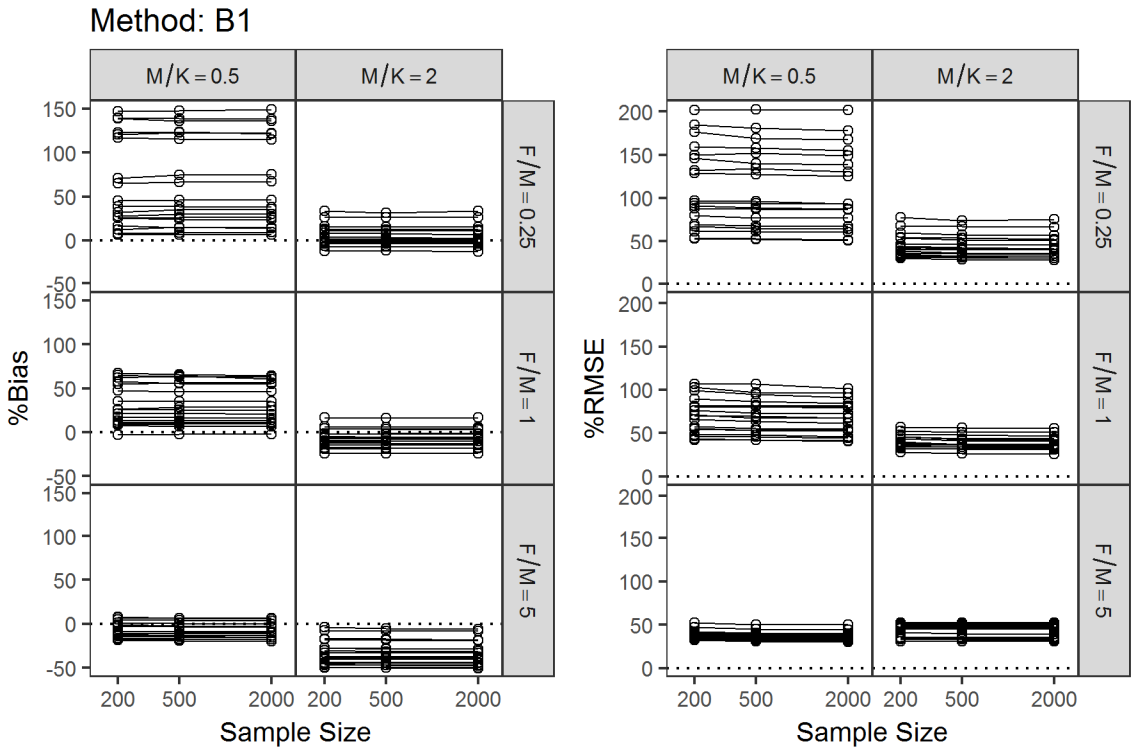

Figure S49. The effect of sample size on %*Bias* (left grid) and %*RMSE* (right grid) for method B1. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

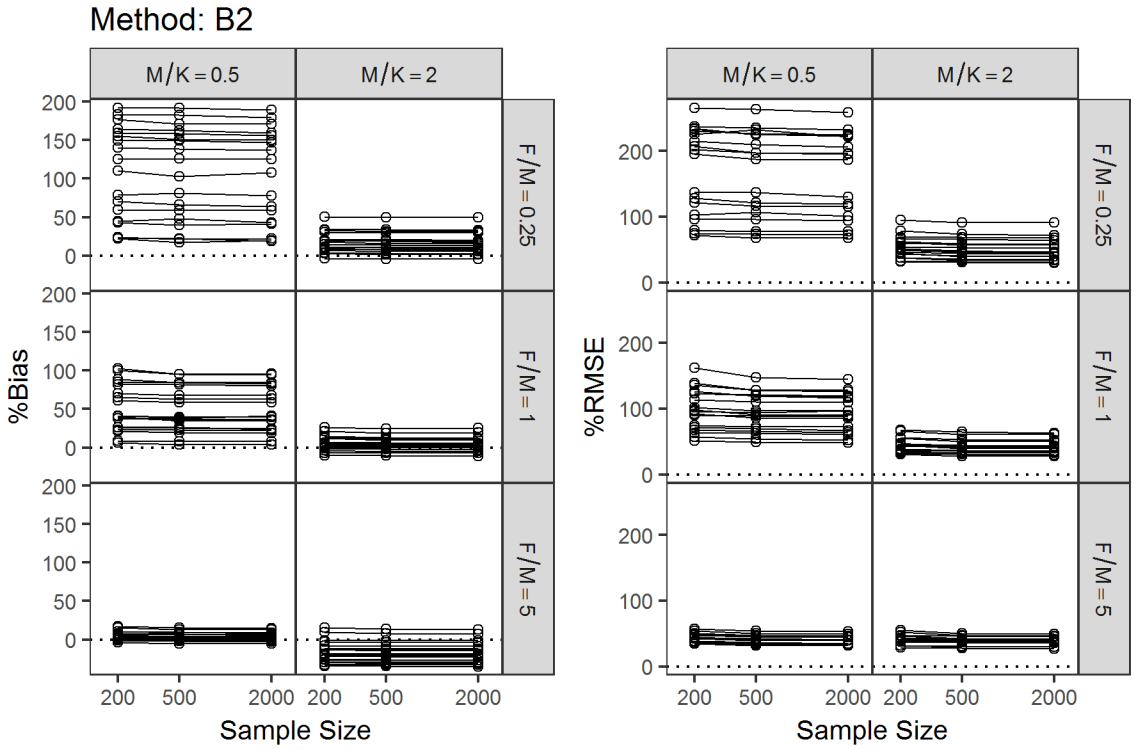

Figure S50. The effect of sample size on %*Bias* (left grid) and %*RMSE* (right grid) for method B2. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

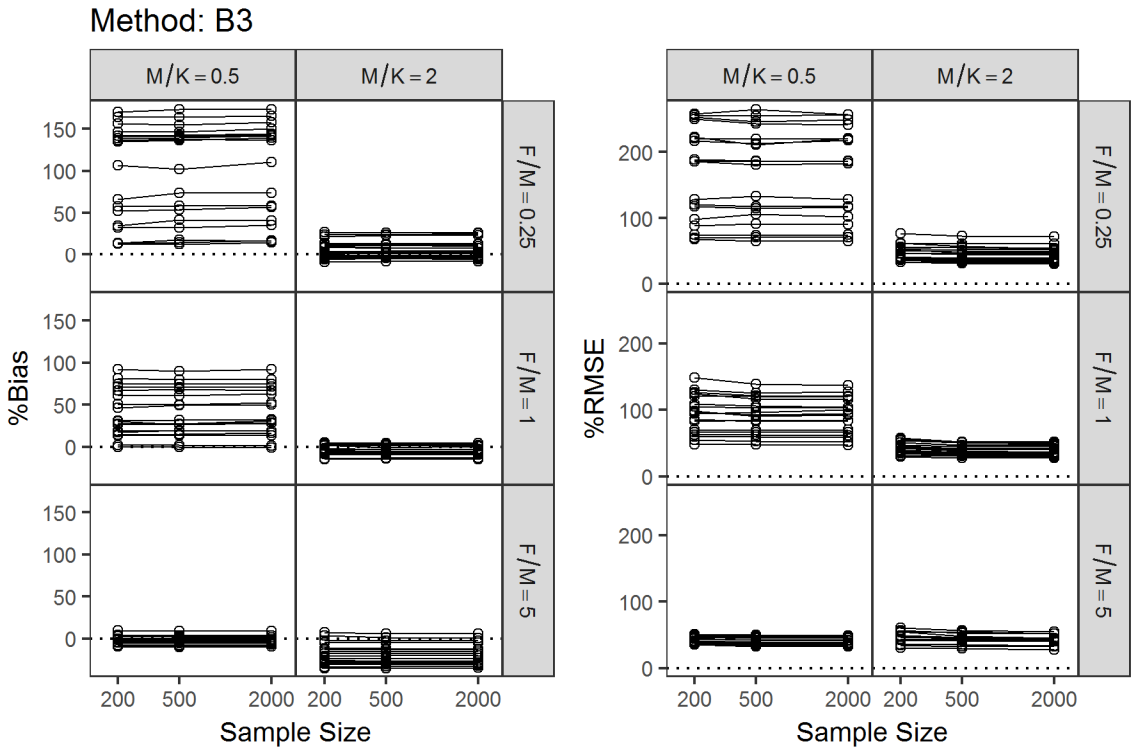

Figure S51. The effect of sample size on %*Bias* (left grid) and %*RMSE* (right grid) for method B3. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

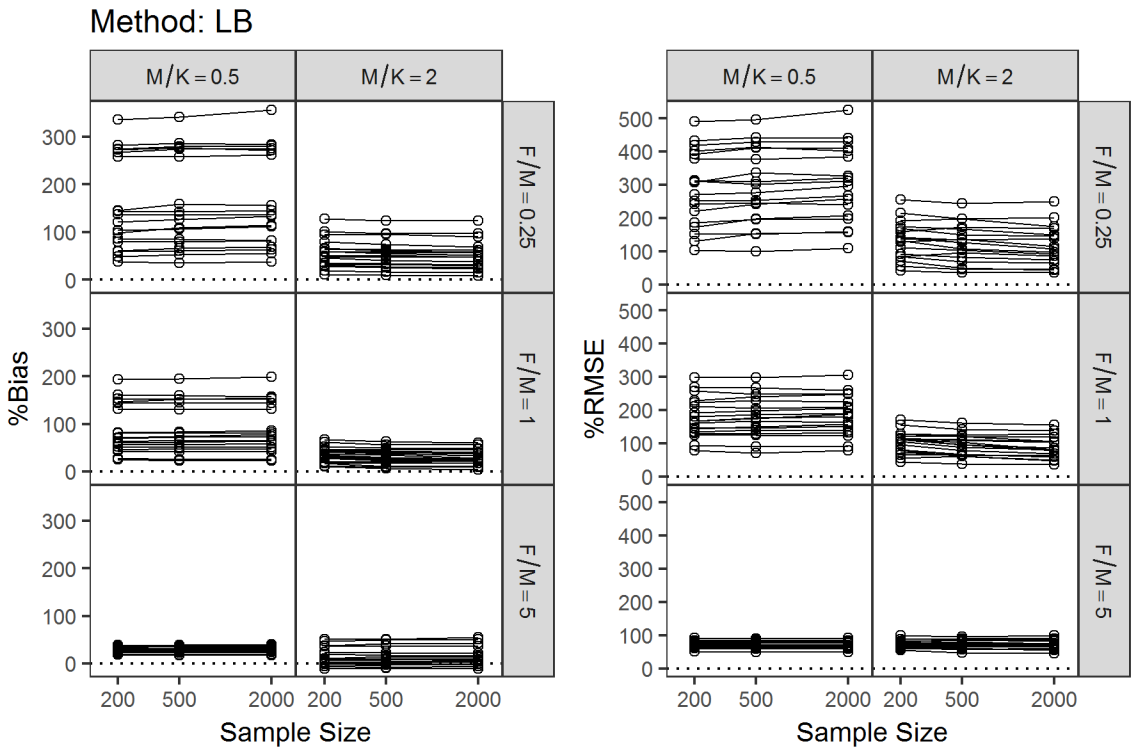

Figure S52. The effect of sample size on %*Bias* (left grid) and %*RMSE* (right grid) for method LB. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

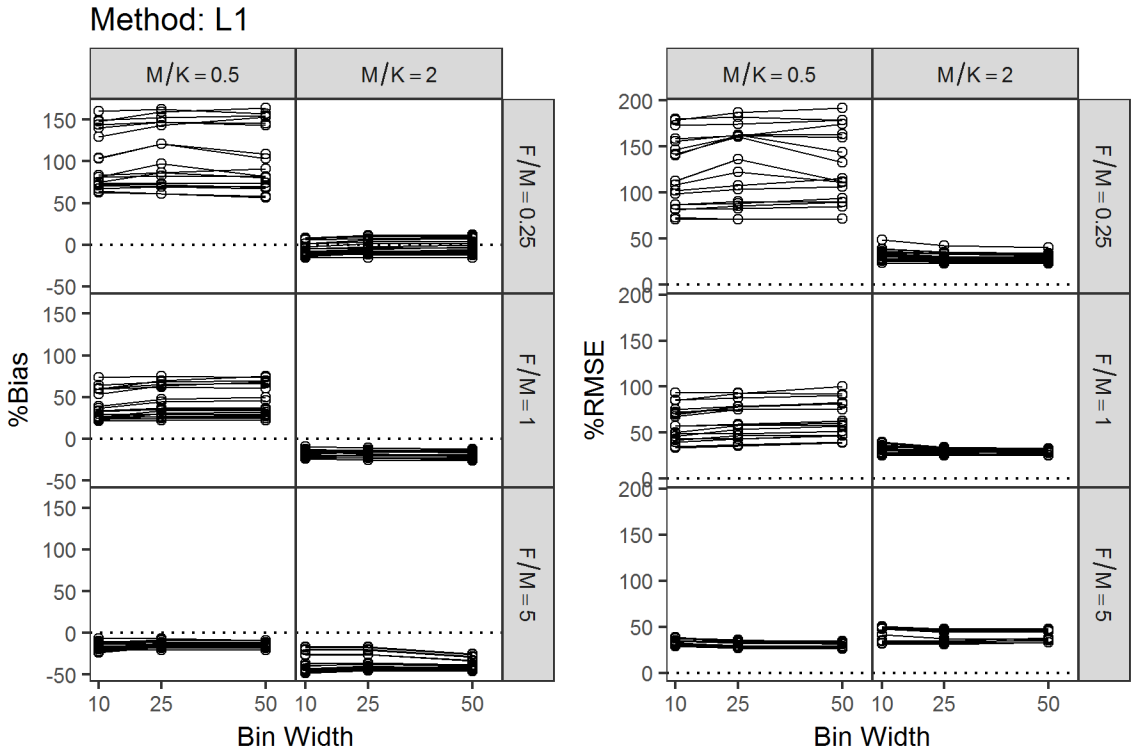

Figure S53. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for method L1. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

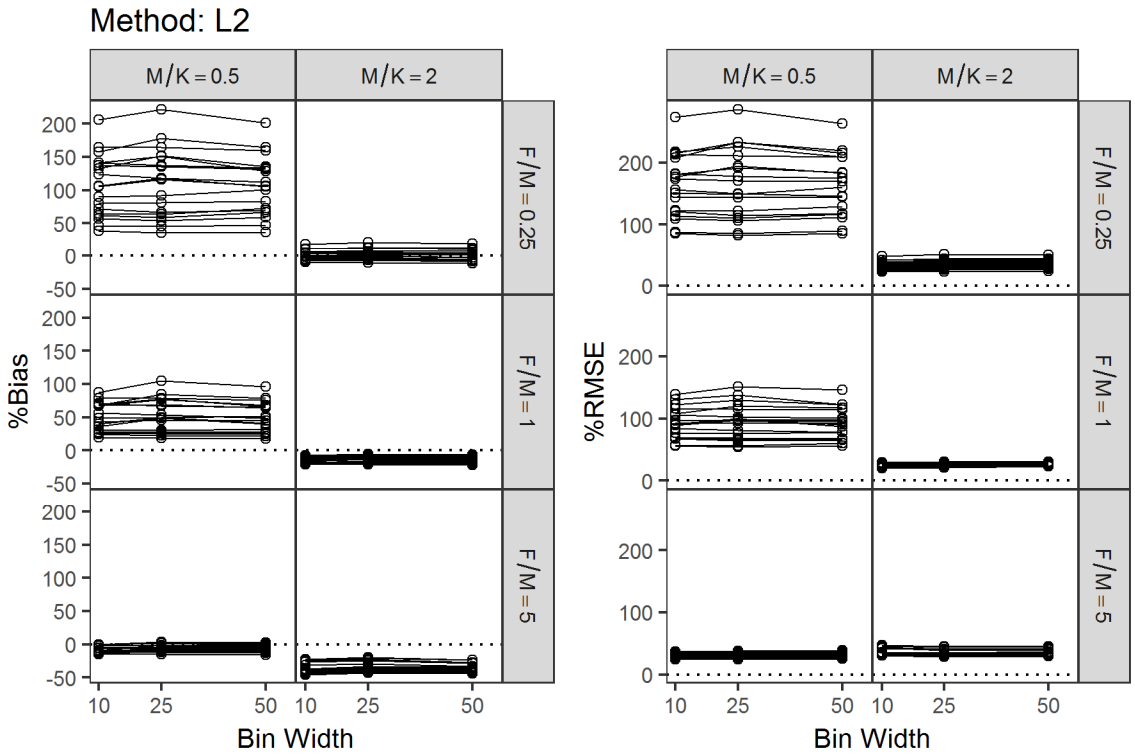

Figure S54. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for method L2. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

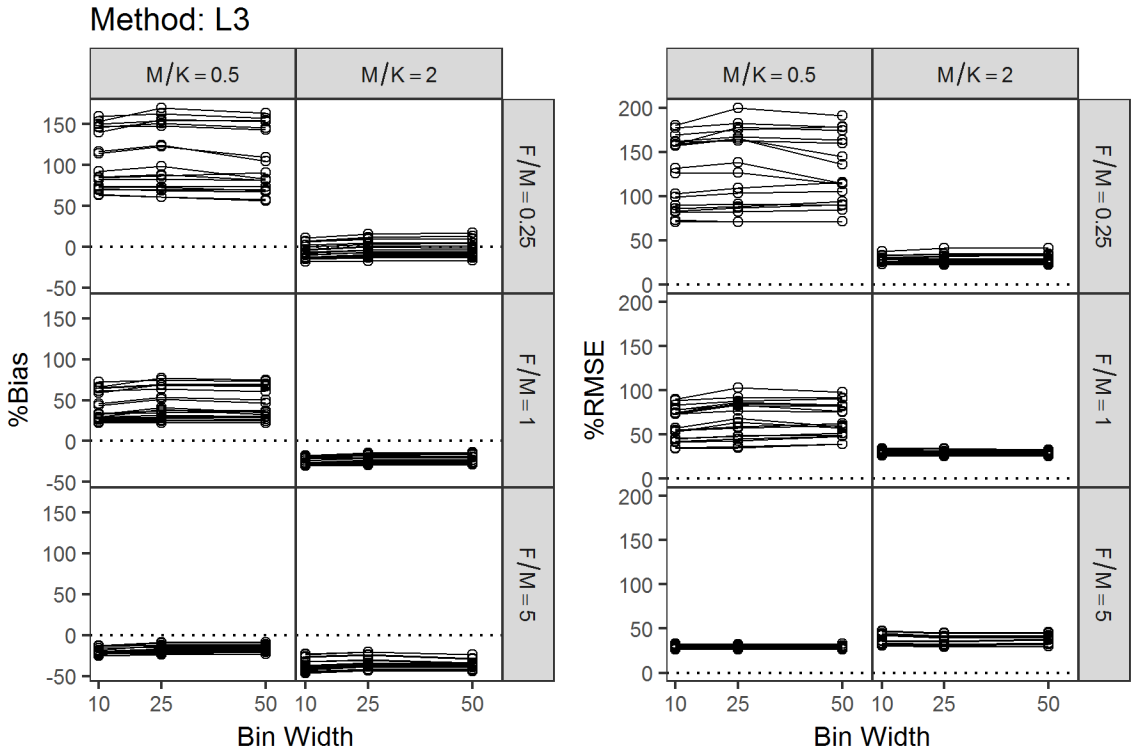

Figure S55. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for method L3. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

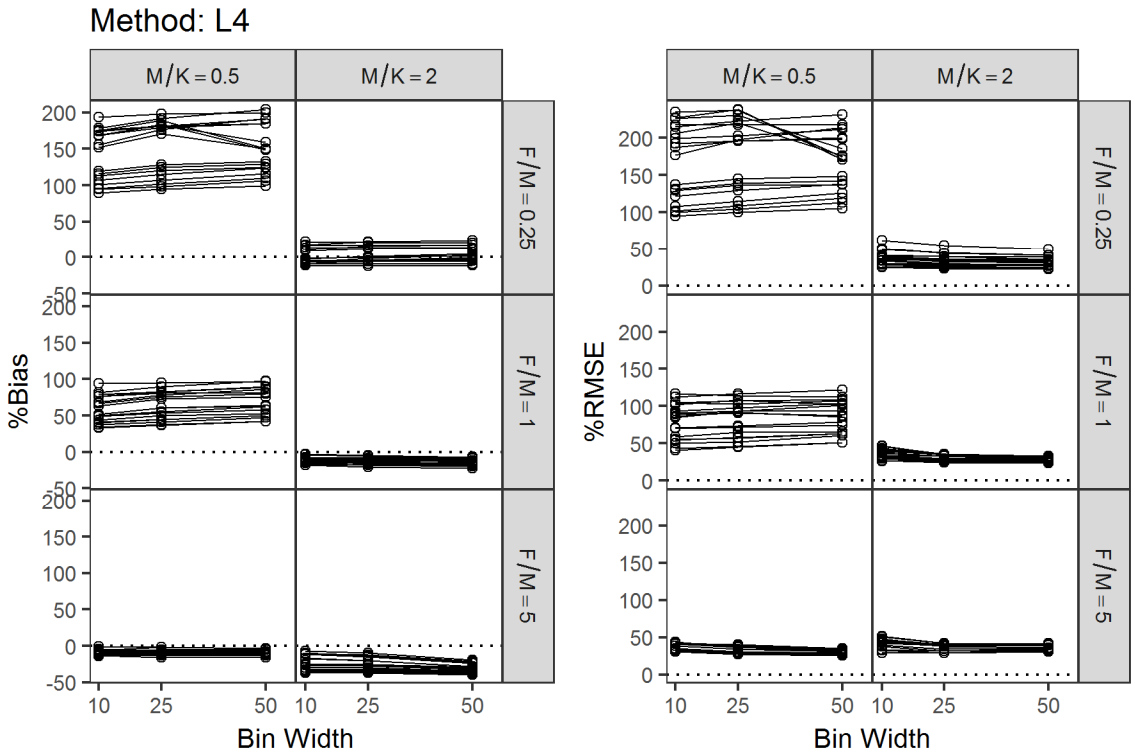

Figure S56. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for method L4. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

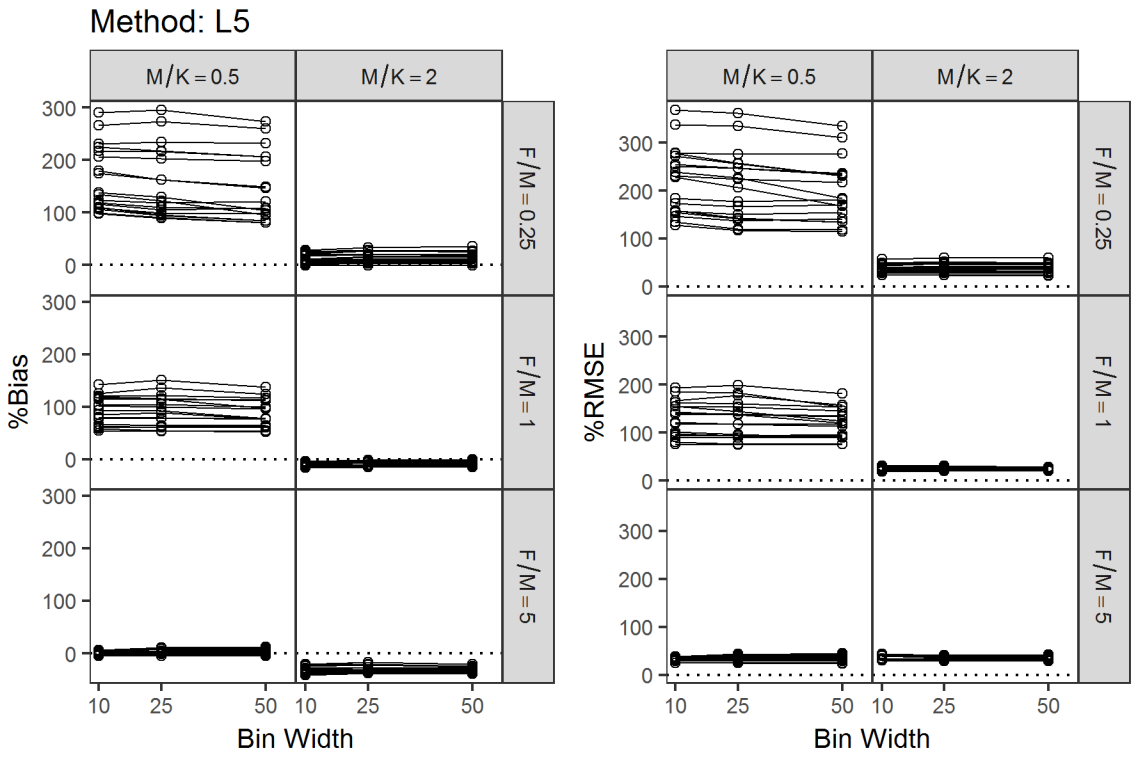

Figure S57. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for method L5. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

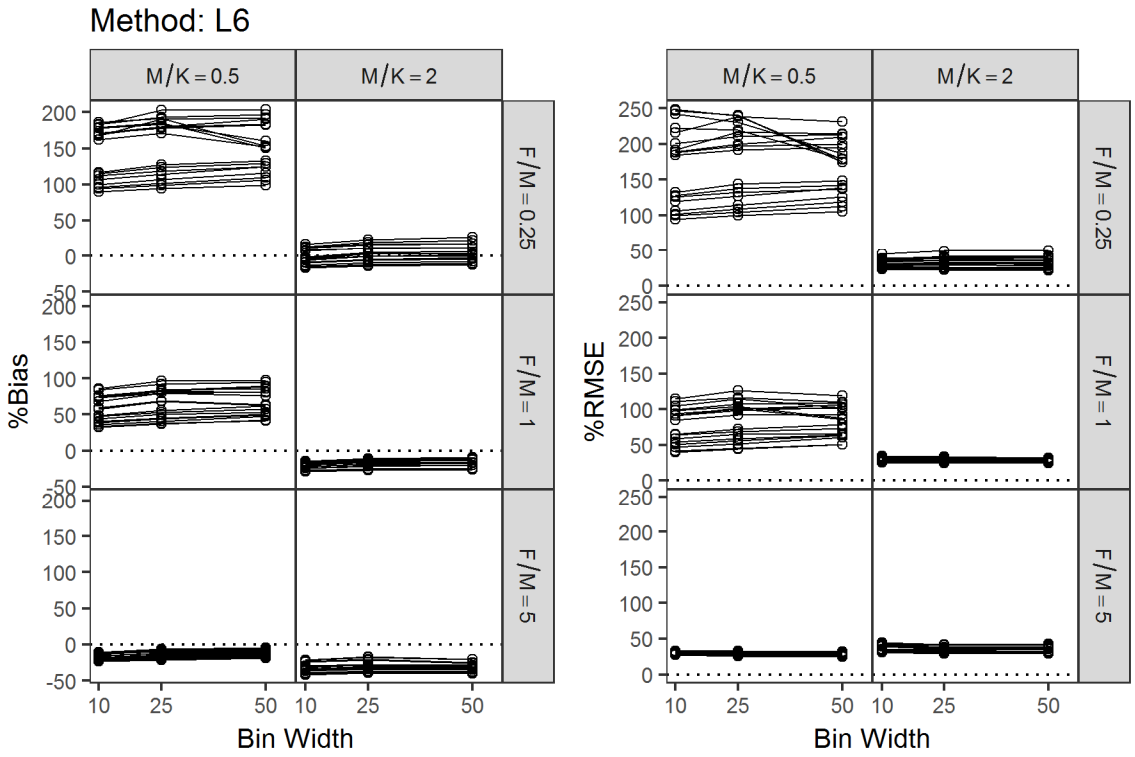

Figure S58. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for method L6. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

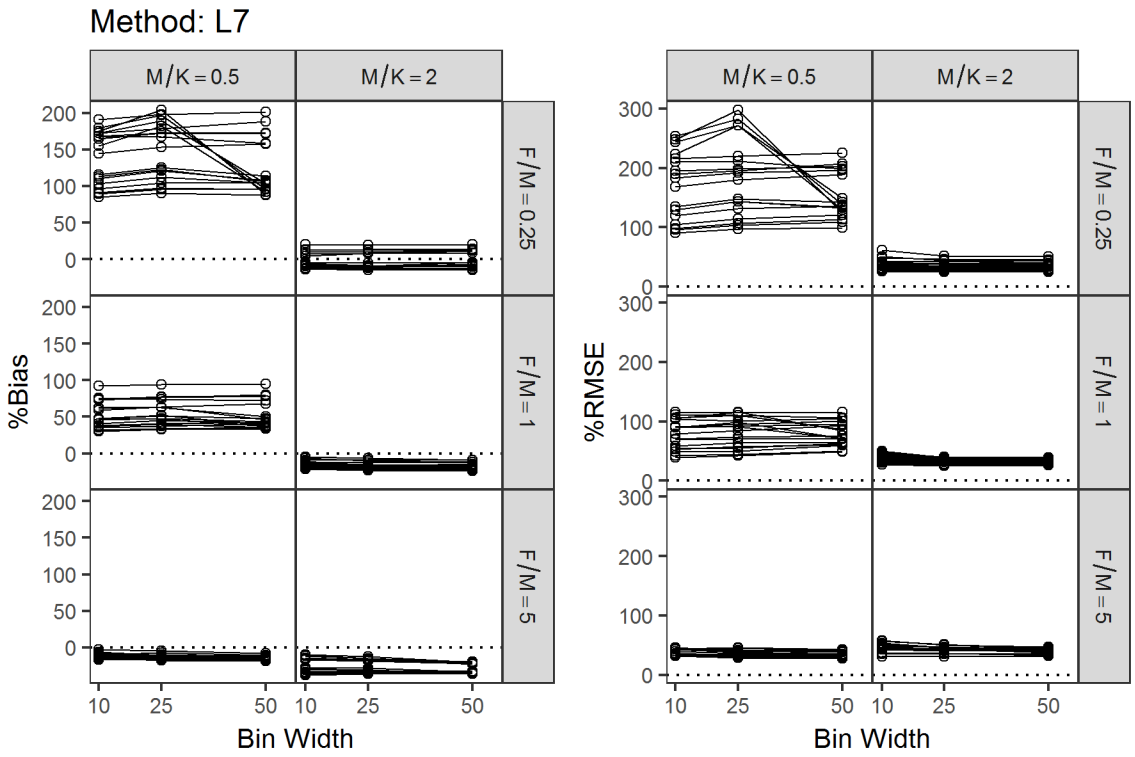

Figure S59. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for method L7. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

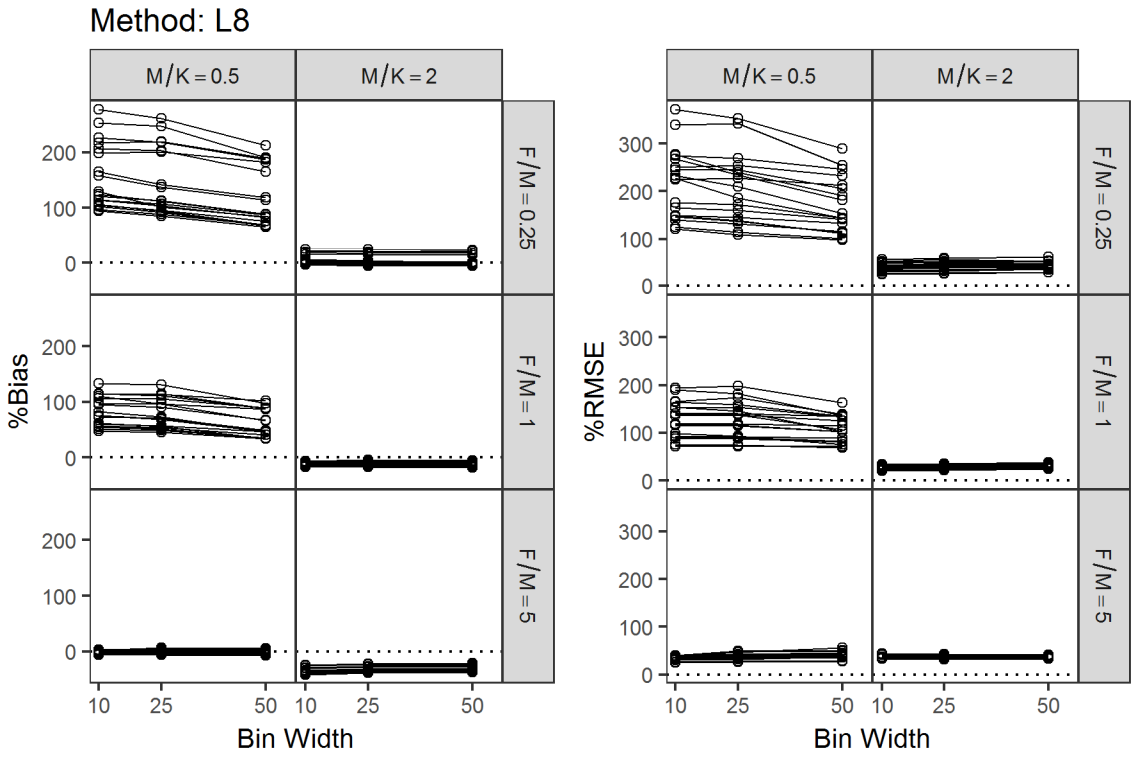

Figure S60. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for method L8. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

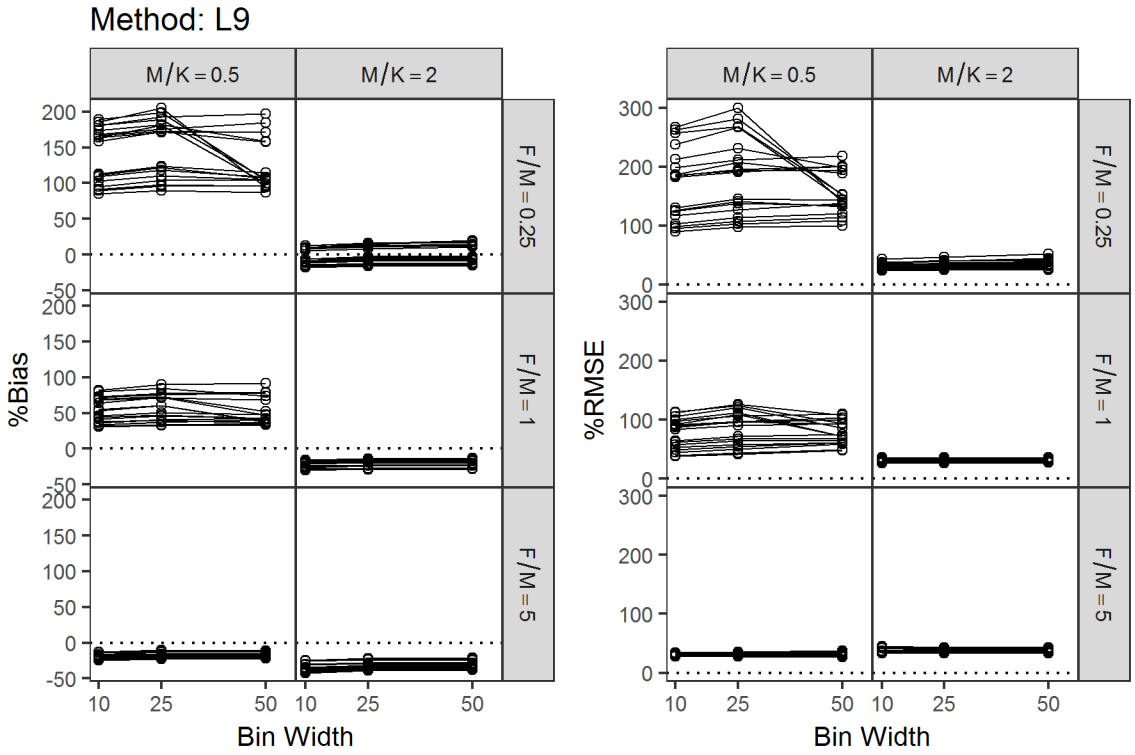

Figure S61. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for method L9. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

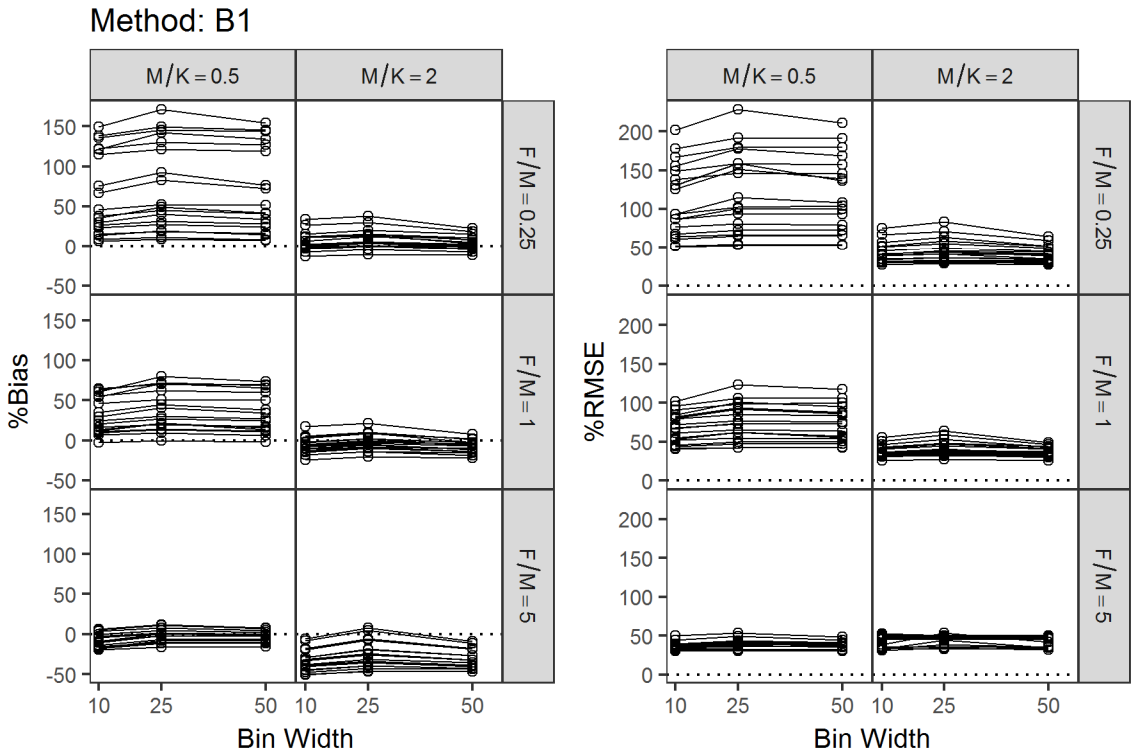

Figure S62. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for method B1. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

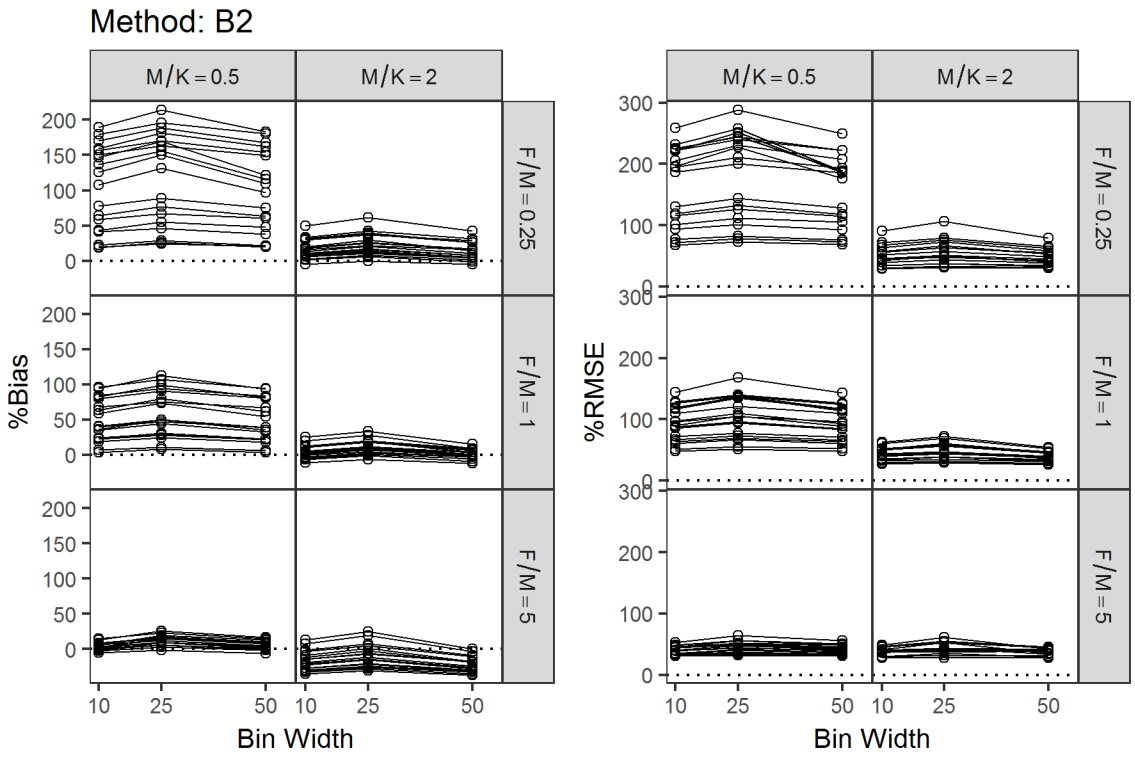

Figure S63. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for method B2. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

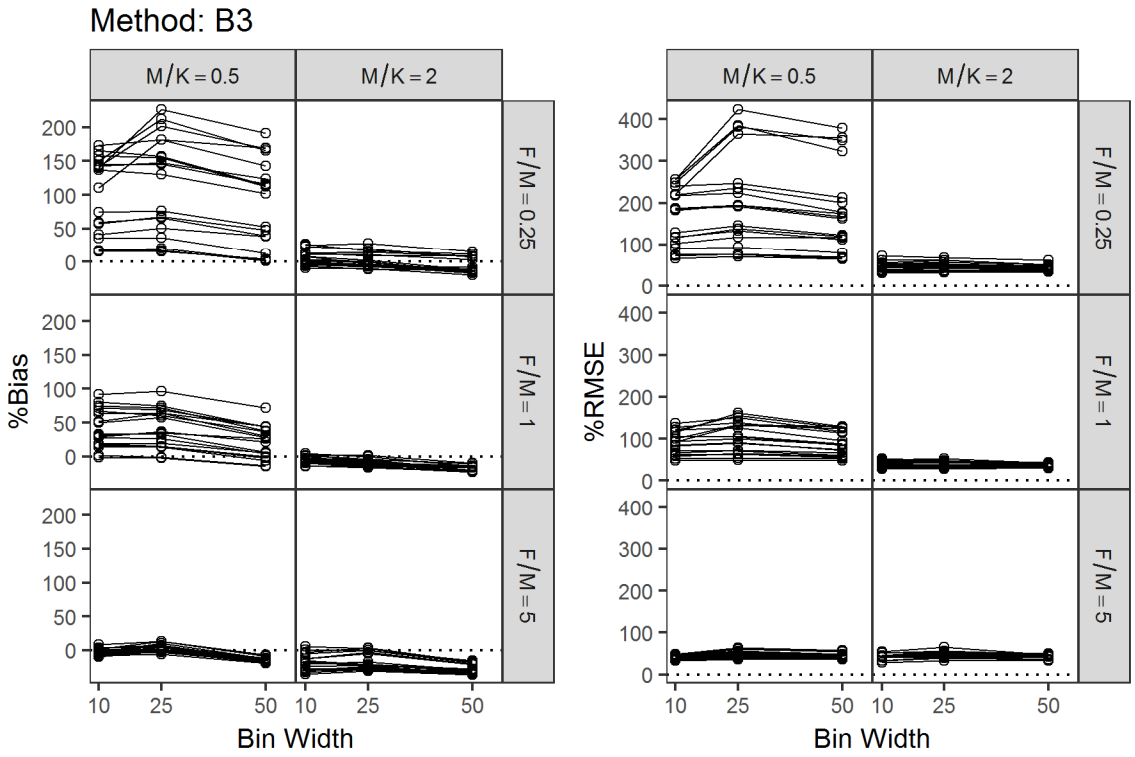

Figure S64. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for method B3. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

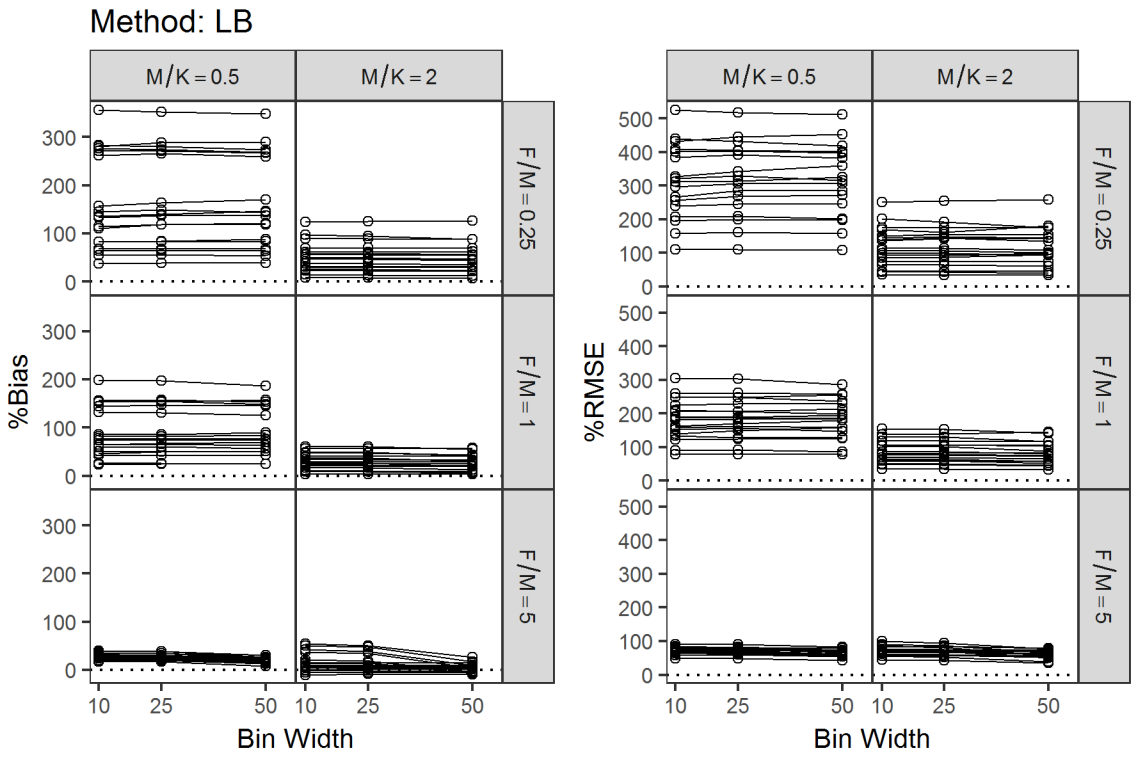

Figure S65. The effect of bin width on %*Bias* (left grid) and %*RMSE* (right grid) for method LB. Each line represents individual factorial combinations stratified in separate cells by *M/K* and *F/M*.

Vita

Quang C. Huynh

Born in Pulau Bidong, Malaysia on August 5, 1988. Graduated from Brooklyn Technical High School in Brooklyn, New York in 2006. Graduated from Swarthmore College in with a Bachelor of Arts in Biology with minors in Music and Environmental Studies in 2010. Worked as a research assistant at the Hudson River Foundation in New York, New York from 2010 – 2013. Enrolled in the Master of Science program at the Virginia Institute of Marine Science, College of William & Mary in 2013 and bypassed to the Doctor of Philosophy program in 2015. Will work as a Postdoctoral Research Fellow at the Institute of Oceans and Fisheries of the University of British Columbia in Vancouver, Canada.